

Engineering Antisense Oligonucleotides for Site-directed RNA Editing with Endogenous ADAR

Dissertation

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

vorgelegt von
Tobias Merkle
aus Laichingen

Tübingen
2020

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

| | |
|-----------------------------------|-------------------------------|
| Tag der mündlichen Qualifikation: | 09.07.2020 |
| Dekan: | Prof. Dr. Wolfgang Rosenstiel |
| 1. Berichterstatter: | Prof. Dr. Thorsten Stafforst |
| 2. Berichterstatter: | Prof. Dr. Dirk Schwarzer |
| 3. Berichterstatter: | Prof. Dr. Claudia Höbartner |

Danksagung/ Acknowledgements

Mein erster Dank geht an meinen Doktorvater Prof. Dr. Thorsten Stafforst. Vielen herzlichen Dank nicht nur für die Überlassung vieler spannender Projekte, sondern auch für die ständige Bereitschaft zum wissenschaftlichen Austausch und für viele gute Ratschläge. Außerdem bin ich sehr dankbar für den offenen Umgang und die Gelegenheit auch eigene Ideen entwickeln und verfolgen zu dürfen. Vielen Dank für die sehr gute und erfolgreiche Zusammenarbeit! Auch bei Prof. Dr. Dirk Schwarzer möchte ich mich herzlich bedanken für das Interesse an meiner Arbeit und die freundliche Übernahme der Zweitbetreuung.

Ein großer Dank geht auch an alle meine Kollegen im Team Stafforst für die tolle Atmosphäre, auch außerhalb des Labors. Special thanks go also to my upstairs office mates Karthika Devi Selvasravanan and Ngadhnjim "Gimi" Latifi for the good time in the office. Vielen herzlichen Dank auch an Gimi Latifi für das Korrekturlesen meiner Arbeit. Außerdem möchte ich mich bei Dr. Paul Vogel und Alfred Hanswillemenke für die freundschaftliche Zusammenarbeit und den produktiven wissenschaftlichen Austausch bedanken.

I also want to thank Dr. Qin Li and Prof. Dr. Jin Billy Li for the great collaboration and the valuable contribution to our publication. Ein besonderer Dank gilt auch den zahlreichen Studenten, die ich betreuen durfte und einen wertvollen Beitrag für unsere Manuskripte geleistet haben. Besonders Sarah Merz, Andreas Blaha, Clemens Lochmann, Carolin Schlitz und Laura Pfeiffer möchte ich für ihre große Motivation und ihren unermüdlichen Einsatz danken.

Außerdem danke ich meinen Freunden aus Tübingen, Konstanz und meiner Heimat, der Schwäbischen Alb. Danke für die vielen guten Stunden und auch den Rückhalt in schwierigen Zeiten! Zum Schluss geht ein ganz besonderer Dank an meine Familie, besonders an meine Eltern, die mich in jeder Lebenssituation gefördert und bedingungslos unterstützt haben. Vielen Dank für alles!

Table of contents

| | |
|---|------|
| List of abbreviations..... | I |
| Zusammenfassung | V |
| Abstract..... | VII |
| List of publications and personal contribution | VIII |
| 1 Introduction | 1 |
| 1.1 Oligonucleotide therapeutics and chemical modifications | 1 |
| 1.1.1 Antisense oligonucleotides | 1 |
| 1.1.2 Small interfering RNAs | 4 |
| 1.2 RNA editing and ADARs | 6 |
| 1.2.1 ADARs in general..... | 7 |
| 1.2.2 Localization of ADARs | 8 |
| 1.2.3 Recoding ADAR substrates..... | 9 |
| 1.2.4 Non-coding ADAR substrates..... | 10 |
| 1.2.5 Structure and enzymatic mechanism of ADAR..... | 11 |
| 1.3 Site-directed RNA editing..... | 15 |
| 1.3.1 The SNAP-ADAR approach | 15 |
| 1.3.2 The λ N-ADAR approach | 16 |
| 1.3.3 The MS2-ADAR approach..... | 17 |
| 1.3.4 The CRISPR-Cas13-ADAR approach – REPAIR and RESCUE..... | 18 |
| 1.3.5 The bump-hole approach | 19 |
| 1.3.6 The CIRT5-ADAR approach..... | 20 |
| 1.3.7 Recruiting wild-type ADAR..... | 20 |
| 1.4 Diseases for targeting with RNA editing | 23 |
| 1.4.1 Mucopolysaccharidosis/ Hurler/ Scheie syndrome | 23 |

| | | |
|--------|---|----|
| 1.4.2 | Alpha-1-Antritrypsin deficiency | 24 |
| 1.4.3 | Rett syndrome..... | 25 |
| 2 | Aim of the study..... | 27 |
| 3 | Results and discussion | 28 |
| 3.1 | Architecture and production of ASOs for site-directed RNA editing | 28 |
| 3.2 | RNA editing with ectopic expression of ADAR | 29 |
| 3.3 | Harnessing endogenous ADAR for 3'UTR editing – RESTORE v1 | 33 |
| 3.3.1 | Which endogenous ADAR is recruited? | 35 |
| 3.3.2 | Potency and duration of RESTORE v1 ASO treatment..... | 35 |
| 3.3.3 | Cell scope of RESTORE v1..... | 36 |
| 3.4 | Harnessing endogenous ADAR for ORF editing – RESTORE v1 | 37 |
| 3.4.1 | Optimizing design and modification pattern | 37 |
| 3.4.2 | Global off-target analysis | 41 |
| 3.4.3 | SERPINA editing and AAT ELISA | 42 |
| 3.5 | RESTORE v2 | 44 |
| 3.5.1 | Optimizing length and symmetry | 44 |
| 3.5.2 | Optimizing chemical modifications..... | 48 |
| 3.5.3 | Cell scope of RESTORE v2..... | 54 |
| 3.3.1 | Which endogenous ADAR is recruited? – RESTORE v2..... | 55 |
| 3.5.4 | Optimizing chemical modification pattern for nuclease resistance | 56 |
| 3.5.5 | Gymnotic uptake..... | 60 |
| 3.5.6 | STAT1 Y701 editing | 60 |
| 3.5.7 | MECP2 W104X editing | 61 |
| 3.5.8 | SERPINA1 PiZ editing..... | 62 |
| 3.5.9 | IDUA editing and restoration of enzymatic activity..... | 65 |
| 3.5.10 | Recruiting murine ADARs..... | 68 |

| | | |
|---|-------------------|----|
| 4 | Conclusions | 70 |
| 5 | References | 71 |
| 6 | Appendix | 85 |

List of abbreviations

| | |
|------------|---|
| 2'-F | 2'-fluoro |
| 2'-OMe | 2'-O-methyl |
| AATD | α 1-Antitrypsin deficiency |
| AAV | adeno-associated virus |
| ACTB | β -actin |
| ADAR | adenosine deaminase acting on RNA |
| AGS | Aicardi-Goutières syndrome |
| ASGPR | asialoglycoprotein receptor |
| AZIN1 | antizyme inhibitor |
| Bdf2 | Bromodomain-containing factor 2 |
| BG | O6-benzylguanine |
| C. elegans | Caenorhabditis elegans |
| CDA | cytidine deaminase |
| cEt | constrained ethyl |
| CFTR | cystic fibrosis transmembrane conductance regulator |
| CIRTS | CRISPR-Cas-inspired RNA targeting system |
| CNS | central nervous system |
| COPD | chronic obstructive pulmonary disease |
| CRISPR | clustered regularly interspaced short palindromic repeats |
| DGCR8 | DiGeorge syndrome chromosomal region 8 |
| DMD | Duchenne muscular dystrophy |
| DMSO | dimethyl sulfoxide |
| DNA | deoxyribonucleic acid |
| ds | double-stranded |
| dsRBD | double-stranded RNA binding domain |
| EGFP | enhanced green fluorescent protein |
| EMA | European Medicines Agency |
| ER | endoplasmic reticulum |
| ERT | Enzyme replacement therapy |
| FDA | Food and Drug Administration |

List of abbreviations

| | |
|--------------------------|--|
| FLNA | filamin- α |
| FLNB | filamin- β |
| GABA | γ -aminobutyric acid |
| GABRA3 | GABA _A receptor subunit α 3 |
| GAGs | glycosaminoglycans |
| GalNAc | N-acetylgalactosamine |
| GAPDH | glyceraldehyde 3-phosphate dehydrogenase |
| GFP | green fluorescent protein |
| GRIA2 (GluR2, GluR-B) | glutamate ionotropic receptor AMPA type subunit 2 |
| gRNA | guide RNA |
| hATTR | hereditary transthyretin-mediated amyloidosis |
| HCV | hepatitis C virus |
| HEK | human embryonic kidney |
| HSCT | hematopoietic stem cell transplantation |
| HTR2C | serotonin receptor 2C |
| HUVEC | human umbilical vein endothelial cells |
| IDUA | α -L-Iduronidase |
| ISGs | interferon-stimulated genes |
| KPNA1 | karyopherin subunit α 1 (importin subunit α 1) |
| KRAS | Kirsten rat sarcoma |
| K _v 1.1/KCNA1 | Potassium voltage-gated channel subfamily A member 1 |
| LDL | low-density lipoprotein |
| LNA | locked nucleic acid |
| LNP | lipid nanoparticle |
| MAVS | mitochondrial antiviral signaling adaptor protein |
| MDA5 | melanoma differentiation-associated protein 5 |
| miRNA | micro RNA |
| mMECP2 | murine Methyl CpG Binding Protein |
| MOE | 2'-O-methoxyethyl |
| NES | nuclear export signal |
| NLS | nuclear localization signal |

List of abbreviations

| | |
|-----------|---|
| NPOM | 6-nitropiperonyloxymethyl |
| nt | nucleotide |
| ORF | open reading frame |
| OTC | ornithine transcarbamylase |
| PAGE | polyacrylamide gel electrophoresis |
| PCSK9 | proprotein convertase subtilisin/kexin type 9 |
| PINK1 | PTEN-induced putative kinase 1 |
| PKR | protein kinase R |
| PMO | phosphorodiamidate morpholino |
| PNK | polynucleotide kinase |
| PS | phosphorothioate |
| PTMs | post-translational modifications |
| REPAIR | RNA Editing for Programmable A to I Replacement |
| RESCUE | RNA Editing for Specific C to U Exchange |
| RESTORE | recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing |
| RIG-1 | retinoic acid-inducible gene 1 |
| RISC | RNA-induced silencing complex |
| RNA | ribonucleic acid |
| SERPINA1 | serine proteinase inhibitor A1 |
| siRNA | small interfering RNA |
| ss | single-stranded |
| SSO | splice switching oligonucleotide |
| STAT1 | signal transducer and activator of transcription 1 |
| TadA | tRNA adenosine deaminase |
| TBP | TAR binding protein |
| TRN1 | transportin-1 |
| UTR | untranslated region |
| XPO1/CRM1 | exportin-1 |
| XPO5 | exportin-5 |

Zusammenfassung

Mit mehreren klinischen Zulassungen ist das Gebiet der Oligonukleotid-Therapeutika in den letzten Jahren erwachsen geworden. Gleichzeitig hat die Entdeckung des CRISPR-Cas-Systems die Manipulation genetischer Informationen in Zellen und Organismen revolutioniert. Die therapeutische Anwendung der DNA-Editierung mit CRISPR leidet jedoch unter den ungelösten Sicherheitsproblemen aufgrund unvorhersehbarer potenzieller Nebenwirkungen durch ungewollte DNA-Schädigungen. In jüngerer Zeit wurden verschiedene Ansätze zur ortsspezifischen RNA-Editierung entwickelt, um dem Risiko einer dauerhaften DNA-Schädigung zu entgehen, indem stattdessen RNA als Eingriffspunkt gewählt wurde. Alle Ansätze für die ortsspezifische RNA-Editierung erfordern jedoch zusätzlich zur Expression oder Anwendung eines RNA-Moleküls die ektopische Expression eines Proteins und leiden unter teilweise deutlicher RNA-Editierungen außerhalb des Zielbereichs, sogenannter Off-Target Editierungen. In dieser Arbeit wurde versucht, die Vorteile der ortsgerichteten RNA-Editierung mit den Fortschritten der Oligonukleotid-Therapeutika zu kombinieren. Aus diesem Grund wurden Antisense-Oligonukleotide entwickelt, um das endogene RNA-Editing-Enzym ADAR für die ortsspezifische RNA-Editierung zu nutzen. Dieser Ansatz wird als RESTORE bezeichnet (Rekrutierung von endogenem ADAR für spezifische Transkripte für Oligonukleotid-vermitteltes RNA-Editing). Verschiedene chemische Modifikationen führten zu einer präzisen und effizienten Editierung im 3'UTR und ORF mit einem überlegenen Off-Target-Editierungsprofil im Vergleich zu allen anderen bestehenden RNA-Editierungsansätzen. Die Anwendbarkeit von RESTORE konnte in einer breiten Palette von humanen Zelllinien und mit noch besseren Editierungsausbeuten von bis zu 80% im ORF von humanen Primärzellen gezeigt werden. Darüber hinaus konnten pathogene Mutationen von schweren genetischen Störungen wie dem Rett-Syndrom, dem Alpha-1-Antitrypsin-Mangel und dem Hurler-Syndrom editiert werden. Um das therapeutische Potenzial von RESTORE zu demonstrieren, wurde die IDUA W402X-Mutation in primären Fibroblasten von zwei Hurler-Syndrom Patienten editiert. Wichtig ist, dass nicht nur auf RNA-Ebene der Wildtyp-Phänotyp teilweise wiederhergestellt werden konnte, sondern auch eine bis zu 6-fach höhere Enzymaktivität als beim viel milderen Scheie-Syndrom erreicht werden konnte. Um diesen vielversprechenden Ansatz

Zusammenfassung

auf *in-vivo* Anwendungen zu übertragen, wurden die Antisense-Oligonukleotide mit chemischen Modifikationen weiter stabilisiert, und es konnte eine verbesserte Stabilität in Serum und Cerebrospinalflüssigkeit erreicht werden. Darüber hinaus ermöglichte dies die freie „gymnotische“ Aufnahme der Antisense-Oligonukleotide in Primärzellen ohne weitere Unterstützung. Zusammen mit der erfolgreichen Rekrutierung von Maus-ADARs ebnet dies den Weg für *in vivo* Anwendungen und die Entwicklung von RESTORE als neue Klasse von Oligonukleotid-Therapeutika.

Abstract

With several clinical approvals, the field of oligonucleotide therapeutics has come of age in the last years. Simultaneously, the discovery of the CRISPR-Cas system has revolutionized manipulation of genetic information in cells and organisms. However, therapeutic application of DNA editing with CRISPR suffers from the unresolved safety issues due to unpredictable potential off-target effects. More recently, several approaches have evolved to escape the risk of permanent DNA damage by targeting RNA instead. Yet, all approaches for site-directed RNA editing require the ectopic expression of a protein in addition to the expression or application of an RNA molecule and suffer from partially severe off-target RNA editing. In this thesis, it was sought to combine the advantages of site-directed RNA editing with the advances of oligonucleotide therapeutics. Therefore, antisense oligonucleotides to harness the endogenous RNA editing enzyme ADAR for site-directed RNA editing were designed, an approach we refer to as RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing). Various chemical modifications resulted in precise and efficient editing in the 3'UTR and ORF with a superior off-target editing profile compared to all other existing RNA editing approaches. The applicability of RESTORE could be demonstrated in a wide panel of human cell lines and with even better editing yields of up to 80% in the ORF of human primary cells. Furthermore, pathogenic mutations found in severe genetic disorders as Rett syndrome, alpha-1-antitrypsin deficiency and Hurler syndrome could be edited. To demonstrate the therapeutic potential of RESTORE, the IDUA W402X mutation was edited in primary fibroblasts donated from two Hurler syndrome patients. Importantly, the wild-type phenotype could be partially restored and an enzyme activity of up to 6-fold higher than that of the much milder Scheie syndrome could be reached. Finally, to transfer this promising approach to *in vivo* applications, the antisense oligonucleotides were further improved with chemical modifications, enhancing stability in serum and cerebrospinal fluid. Moreover, this made unassisted gymnotic uptake of the antisense oligonucleotides into primary cells possible. Together with the successful recruitment of murine ADARs, this paves the way to *in vivo* applications and the development of RESTORE as a new class of oligonucleotide therapeutics.

List of publications and personal contribution

Manuscript 1 (published):

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

Personal contribution: I cloned the gRNA constructs, performed and evaluated all experiments for the recruitment of various editases by Cas13-gRNAs (**Supplementary Figure 11** and **Supplementary Note 2**).

Manuscript 2 (published):

T. Merkle, S. Merz, P. Reautschnig, A. Blaha, Q. Li, P. Vogel, J. Wettengel, J. B. Li, T. Stafforst, Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nature Biotechnology* **37**, 133-138 (2019)

Personal contribution: I established the in vitro transcription and ligation procedure for the ASOs. Also, I designed all ASOs used in this study together with Thorsten Stafforst. I designed, conducted and analyzed all experiments depicted in **Figure 2 A-C, E, F, H, Figure 3 B, D, Supplementary Figure 6, 7, 10, 14, 15**. Furthermore, I designed the experiments and analyzed the results of all and conducted some experiments in **Figure 2 G, 3 E** and co-supervised the intern Andreas Blaha conducting and analyzing the rest of the experiments in these figures. I conducted the wet lab experiments for **Figure 3 C, Supplementary figure 11, 12, 13** and took part in the analysis of the data together with Qin Li, Paul Vogel, Jin Billy Li and Thorsten Stafforst. I designed the experiments and analyzed the results of the experiments in **Figure 1 C, Supplementary Figure 2, 3, 5, 8, 9** and co-supervised the Bachelor student Sarah Merz conducting and analyzing these experiments. In collaboration with Thorsten Stafforst and Paul Vogel, I prepared the figures in **Figure 1 A, B, Figure 3 A, Supplementary Figure 15**. I contributed to writing the manuscript.

Manuscript 3 (accepted):

T. Merkle and T. Stafforst, New frontiers for site-directed RNA editing – harnessing endogenous ADARs. *Methods in Molecular Biology* (accepted)

Personal contribution: I optimized the protocols for in vitro transcription, ligation and transfection of the ASOs. I wrote the material and methods part and prepared the figures.

Manuscript 4 (in preparation):

T. Merkle, C. Schlitz, L. Pfeiffer, C. Lochmann, T. Stafforst, Improved antisense oligonucleotides for efficient and precise RNA editing with endogenous ADARs

Personal contribution: I designed all ASOs used in this study together with Thorsten Stafforst. I designed, conducted and analyzed all experiments depicted in **Figure 1 C, D, F, Figure 2 C, D, Figure 3 A, SI Figure 1 A-C** and **SI Figure 4**. I designed the experiments and analyzed the results of all and conducted some experiments (THP-1 and partly HeLa) in **Figure 1 E, SI Figure 3** and partly **Figure 2A** and co-supervised the intern Laura Pfeiffer conducting and analyzing the rest of the experiments in these figures. I designed the experiments and analyzed the results of all experiments in **Figure 2 A, B** and **SI Figure 2** and co-supervised the Bachelor student Clemens Lochmann conducting and analyzing the experiments in these figures. I designed the experiments and analyzed the results of all experiments in **Figure 3 B-H, SI Figure 5, 6, 7, 8, 9, 10** and co-supervised the Bachelor student Carolin Schlitz conducting and analyzing the experiments in these figures. I contributed to writing the manuscript and prepared the Figures together with Thorsten Stafforst.

1 Introduction

1.1 Oligonucleotide therapeutics and chemical modifications

The discovery of the antisense effect of a synthetic DNA oligonucleotide by Zamecnik and Stephenson in 1978 marks the beginning of the therapeutic modality of antisense oligonucleotides (ASOs) and probably oligonucleotide therapeutics in general^{1, 2}. However, it took another 20 years until the first ASO drug fomivirsen was approved as a therapy for cytomegalovirus retinitis^{3, 4}. Major challenges in developing oligonucleotide drugs has been and still remain cellular delivery, stability in body fluids, and poor pharmacokinetics⁵⁻⁷. These challenges originate mainly from the polyanionic character of oligonucleotides that supports rapid renal clearance and impairs migration over the lipid bilayer of cellular membranes^{5, 8}. On the other hand, the high abundance of nucleases, especially RNases, in body fluids leads to almost immediate degradation of unprotected nucleic acids⁷. To overcome these hurdles, different strategies have been developed in the last 40 years, most significantly the introduction of diverse chemical modifications and conjugations^{9, 10}.

1.1.1 Antisense oligonucleotides

An ASO in general is a synthetic single-stranded (ss) oligonucleotide that is at least in part reverse complementary or “antisense” to a target RNA. Depending on identity, target and mode of action, there are different classes of ASOs. One major class is RNase H-dependent ASOs that bind to a specific mRNA and can recruit RNase H that subsequently cleaves the target effecting in a knock-down¹¹. Other classes are ASOs that mediate splice switching, translational arrest or miRNA binding (**Figure 1**)¹¹⁻¹⁴.

Since RNase H recognizes DNA-RNA hybrids, the first generation of RNase H ASOs like fomivirsen consisted of short (~21nt) ssDNA. In order to support nuclease-stability, the ASO backbone was completely modified with phosphorothioates instead of phosphates^{4, 15}.

The synthesis of the phosphorothioate (PS) modification, substituting a non-bridging oxygen in the backbone phosphate by sulfur, was first reported by Fritz Eckstein in 1966¹⁶. The incorporation of sulfur creates a stereocenter at the phosphorus. Consequently, a

Introduction

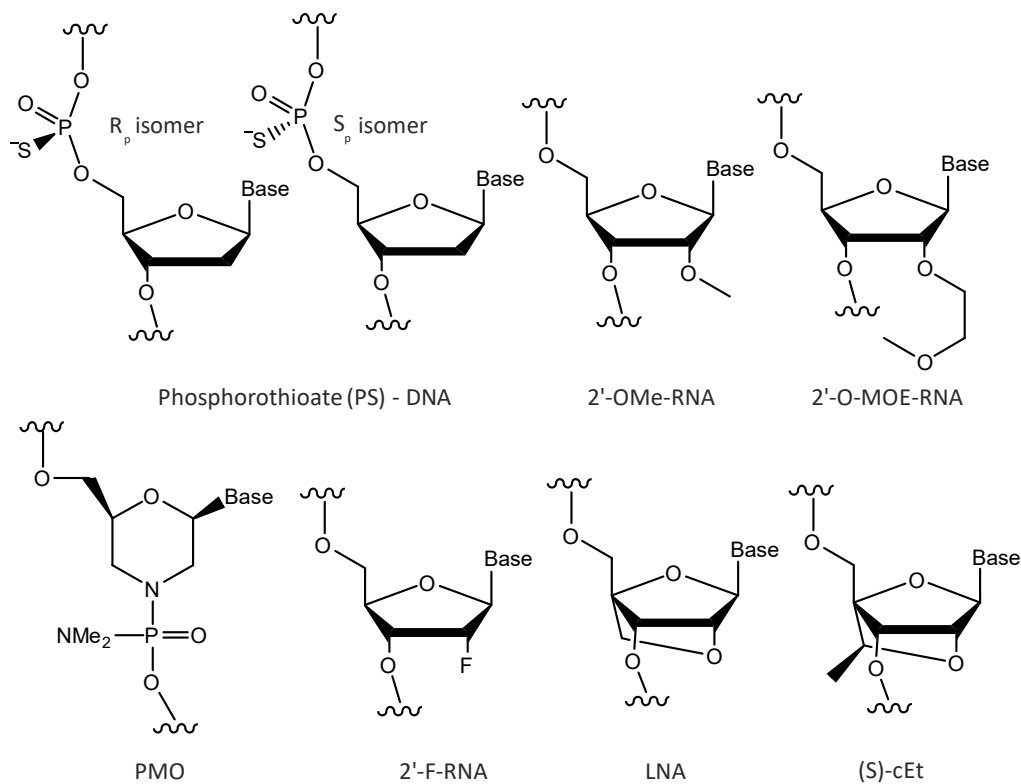
mixture of a great number of diastereomers results from standard oligonucleotide synthesis¹⁵. Interestingly, the S_p diastereomer revealed to have a higher nuclease-resistance than the R_p diastereomer. On the other hand, the binding affinity of the S_p diastereomer is lower compared to its counterpart. Furthermore, RNase H cleavage was more efficient with an all- R_p -ASO^{17, 18}. In a random PS diastereomer mixture, the overall binding affinity is reduced in comparison to a phosphate backbone modification. Due to the enhanced plasma protein binding, PS modifications improve the pharmacokinetic profile of ASOs significantly¹⁹. Moreover, PS improve cellular uptake of ASOs by surface protein binding^{15, 20, 21}. These two properties make PS the most widely used modification in ASOs despite PS-related toxicity found in the context of binding to immune receptors^{22, 23}.

Although PS backbone modifications increase the overall nuclease stability, PS-DNA ASOs are rapidly degraded *in vivo*. Thus, in second generation ASOs more nuclease-stable 2'-substituted nucleotides were incorporated, flanking the central 8-10 nt DNA gap required for RNase H cleavage, also referred to as gapmers. The oldest and also naturally occurring 2'-ribose modifications is the 2'-OMe, which increases nuclease-resistance and improves binding affinity to the target²⁴⁻²⁶. A fully 2'-OMe and PS modified ASO was also used for the eventually FDA-rejected exon skipping oligo drisapersen against Duchenne muscular dystrophy (DMD)²⁷. Newer ASOs in clinical trials use, in addition to PS, the 2'-O-methoxyethyl (MOE) modifications. Examples for this are the FDA approved RNase H ASO mipomersen (Kynamro) against familial hypercholesterolemia and the splice-switching oligonucleotide (SSO) nusinersen (Spinraza) against spinal muscular atrophy^{28, 29}. Due to the larger 2'-O-methoxyethyl substituent that increases steric hindrance at the nuclease cleaving site, MOE exhibits even higher nuclease resistance than the 2'-OMe group³⁰. Furthermore, the RNA-like C3'-endo sugar conformation is even more favored than in 2'-OMe which results in higher binding affinity of MOE compared to 2'-OMe modified ASOs^{30, 31}.

Another special modification, the phosphorodiamidate morpholino (PMO) nucleic acid analog, was used in the SSO eteplirsen (Exondys51) to treat DMD³². Although PMOs are uncharged, the cellular uptake remains challenging and therefore only moderate activity in the muscle cells of a DMD mouse model without a delivery vehicle was obtained³³.

However, despite the low efficacy in clinical studies, eteplirsen was approved by the FDA but not by the EMA^{34, 35, 36}.

A



B

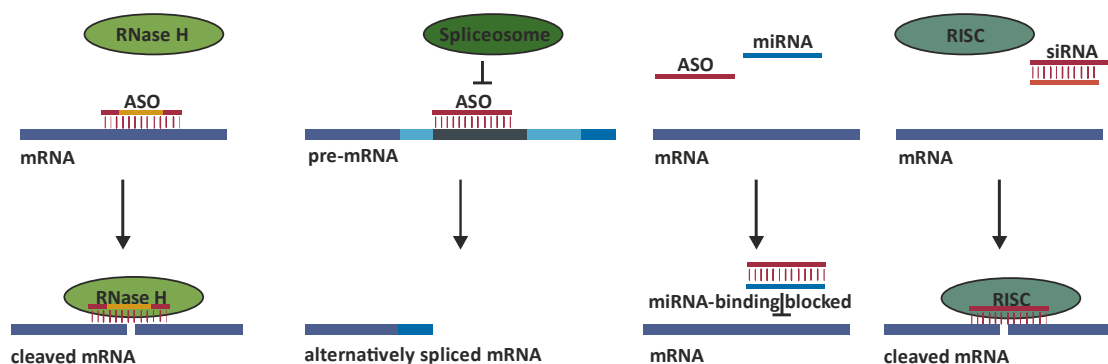


Figure 1: Oligonucleotide therapeutics - chemical modifications and mode of action. A) depicts chemical modifications in nucleic acids that are commonly used in ASO or siRNA therapeutics. In B) different mode of actions of oligonucleotide therapeutics are illustrated. While gapmer ASOs and siRNAs mediate mRNA degradation via RNase H or RISC enzyme activities, ASOs can also switch splicing or inhibit miRNA activity by binding to pre-mRNA or miRNA.

In the next generation of ASOs that are under investigation, locked nucleic acid (LNA) and constrained ethyl (cEt) building blocks are applied. In both of these modifications the 2' oxygen is linked with the 4'-carbon and therefore "locking" the ribose in the 3'-endo

Introduction

sugar pucker and increasing base stacking. As a result, the binding affinity towards RNA is significantly increased with a melting temperature elevation ranging from 4°C to 8°C per LNA. Furthermore, both modifications exhibit high nuclease-resistance³⁷⁻⁴¹. The high affinity allows the use of shorter ASOs with gapmers of 12-15nt in length and the potency was reported to be significantly higher. Additionally, the “gymnotic” uptake, meaning administration of the naked ASO without any ligand or vehicle is improved^{37, 42, 43}. The high affinity of LNA ASOs put also miRNAs as an attractive target in focus. One of the most advanced ASO miravirsin, targets miRNA-122 as a therapy against hepatitis C virus (HCV). The 15nt LNA/DNA hybrid ASO containing 8 LNA bases is currently in late phase 2 clinical trials^{13, 44}.

1.1.2 Small interfering RNAs

Another major class of oligonucleotide therapeutics are small interfering RNAs (siRNAs). For their discovery about potent RNA interference (RNAi) with double-stranded (ds)RNA in *C.elegans*, Mello and Fire were awarded with the Nobel prize⁴⁵. Only three years after this discovery, efficient transcript knock-down with a 21nt siRNA duplex was demonstrated in mammalian cells⁴⁶. And in 2018, finally with patisiran (Onpattro) the first siRNA drug was approved^{47, 48}. In patisiran only the pyrimidines in the guide strand and some of the pyrimidines in the passenger strand were 2'-OMe modified to increase nuclease stability. Also, for the overhangs required for RNAi, two 5' inverse dT nucleotides were included to further increase stability⁴⁷.

Patisiran revealed to be more beneficial in clinical trials than the MOE modified RNase H ASO inotersen (Tegsedi) that was also approved for treatment of hereditary transthyretin-mediated amyloidosis (hATTR). Not only was the knock-down duration of the liver-derived transthyretin higher, resulting in infusion of patisiran every 3 weeks in comparison to a weekly application of inotersen. But more strikingly, patisiran could significantly improve neuropathy in patients while inotersen could only slow down neuropathic disease progression compared to the placebo control^{49, 50}.

In RNAi, the RNA-induced silencing complex (RISC) separates the two strands of the siRNA and the phosphorylated guide strand is loaded into the RISC. Subsequently, mRNA that is antisense to the guide strand is recruited and cleaved. However, the RISC is very sensitive

to chemical modifications^{51, 52}. Exchanging the 2'-OH group with a 2'-F revealed to be the most suitable modification for siRNAs⁵¹. 2'-F modifications can decrease innate immune stimulation, increase nuclease stability while maintaining or even improving siRNA activity⁵³⁻⁵⁷. With the small size and the high electronegativity it has similar properties as the OH group⁵⁸. However, instead of forming hydrogen bonds with water in the minor groove, the 2'-F modification exhibits enhanced affinity due to increased base-stacking, Watson-Crick base pairing and lower hydration^{51, 59}.

Since fully chemically modified siRNAs performed best *in vivo* and neither 2'-F nor 2'-OMe modifications alone are optimal, a mixture of 2'-OMe and 2'-F modifications is widely used⁶⁰⁻⁶³. Additional PS modifications at the termini of both strands for enhanced nuclease stability are often used⁹. Nuclease stability is of high importance when it comes to delivery of siRNAs. While single-stranded phosphorothioate modified ASOs are taken up by several tissues after systemic administration, predominantly by liver and kidney, this is not the case for siRNAs^{5, 64}. Early attempts to deliver siRNAs utilized lipid nanoparticles (LNPs) as in the example of patisiran⁴⁸. Although delivery to the liver was possible and only partly modified siRNAs could be used, there are several disadvantages using LNPs. Their large size not only reduced the pharmacokinetic properties but also diffusion in tissues is very limited. Thus, delivery beyond the liver remains challenging with LNPs⁵. In other approaches hydrophobic conjugates were analyzed and especially cholesterol conjugates with fully stabilized siRNAs gave promising results *in vivo* in brain, liver, spleen and placenta^{63, 65, 66}. Intracerebroventricular injection of stabilized, PS-containing divalent siRNAs showed also promising distribution and silencing, similar or even better than intrathecal injection PS ASOs, probably due to their high PS content⁶⁷.

The possibly most successful strategy to deliver siRNA and ASOs into the liver is the use of triantennary N-acetylgalactosamine (GalNAc) conjugates that bind to the asialoglycoprotein receptor (ASGPR)^{68, 69}. The reason for this efficient uptake lies in the ASGPR that is highly expressed in hepatocytes and has an extraordinarily high recycling rate of about 15 min. Unfortunately, no other similar receptor with that high expression and recycling rate has been found to date and therefore effective cellular uptake of oligonucleotides in extrahepatic tissues still remains challenging⁵.

Introduction

Inclisiran is a typical example of a triantennary GalNAc passenger strand conjugated, fully 2'-F/2'-OMe modified siRNA with terminal PS backbone modifications⁷⁰. With the cholesterol lowering proprotein convertase subtilisin/kexin type 9 (PCSK9) mRNA targeting inclisiran, the first potential blockbuster RNA therapeutic is in late phase clinical trials^{70, 71}. Clinical data revealed efficient PCSK9 knock-down and low-density lipoprotein (LDL) cholesterol reduction with minimal adverse effects and superior duration that made injection intervals of 3-6 months possible^{71, 72}. The approval of inclisiran might be a breakthrough milestone for the class of oligonucleotide therapeutics and clearly demonstrating the potency and transferability of this platform technology.

1.2 RNA editing and ADARs

The successful approval and promising candidates in clinical trials of oligonucleotide therapeutics emphasize the strength of chemically modified oligonucleotides for therapeutic application. The unique advantage to easily transfer the pharmacokinetic properties of distinct chemical modifications of oligonucleotides to other targets by just altering the sequence makes this class of therapeutics especially attractive⁹. Therefore, it would be highly desirable to expand the scope of oligonucleotide therapeutics beyond knock-down or splice-switching.

With the CRISPR-Cas9 technology, a relatively easy way to alter genetic information in a programmable oligonucleotide-dependent manner was discovered^{73, 74}. Undoubtedly, it became very rapidly a versatile, cheap and easy laboratory tool⁷⁵. However, transferring it to the clinics revealed several hurdles and challenges⁷⁶. The most serious obstacle is probably undesired insertions, deletions and other off-target effects associated with the risk of introducing mutations leading eventually to cancer. To avoid DNA double-strand breaks and therefore reduce the risk of off-target effects, CRISPR base editors utilizing fusion proteins of deaminases and Cas9 protein that is not able to introduce double-strand breaks were invented^{77, 78}. Although reducing the risk of off-target effects, deamination by the base editors was not limited to only a single base, but deamination in neighboring bases was observed⁷⁹.

An alternative to base editing on the genome level, represents base editing on the transcriptome level. One major advantage of RNA editing over DNA editing is the reversibility. Potential off-target effects are always temporary due to the nature of mRNA stability. Thus, RNA editing also offers the possibility of only temporary manipulation of genetic information. Another significant advantage of RNA editing is that it is a natural process mediated by an endogenous machinery. Hijacking this machinery with synthetic oligonucleotides for a therapeutic purpose comparable to RNase H dependent ASOs or siRNAs would circumvent ectopic expression of synthetic CRISPR fusion proteins and expand the toolbox of the promising area of oligonucleotide therapeutics⁸⁰.

1.2.1 ADARs in general

A-to-I editing is the most abundant form of natural RNA editing and is mediated by ADARs (Adenosine deaminase acting on RNA). ADARs are members of an enzyme family that catalyze the deamination of adenosine to inosine in double-stranded RNA (**Figure 2 A**)^{81, 82}. The first evidence of ADAR date back to 1987, when its ability to unwind double-stranded RNA in embryos of *Xenopus laevis* was first described^{83, 84}. Shortly after that, the unwinding activity was found to base in fact on the A-to-I conversion and the resulting weaker I-U base pairing^{85, 86}. However, ADARs are not only found in *Xenopus*, but very early also in human and murine cells^{86, 87} and occur in multicellular animals but not in yeast or plants⁸⁸. There are three ADAR genes in vertebrates⁸⁹. While all three ADARs share a common C-terminal deaminase domain, only ADAR1 and ADAR2 revealed to be catalytically active^{90, 91}. Another common functional domain is the double-stranded RNA binding domain (dsRBD). While ADAR1 contains three, ADAR2 and ADAR3 share only two dsRBDs (**Figure 2 D**). For ADAR1 two isoforms are known. The short constitutively expressed 110 kDa ADAR1 p110 and the longer 150 kDa ADAR1 p150 which is expressed from an alternative, interferon inducible promoter^{92, 93}. In contrast to the other ADARs, the N-terminal part of ADAR1 p110 comprises one Z-DNA binding domain (Z β), whereas the long isoform ADAR1 p150 has two Z-DNA binding domains (Z α and Z β), respectively⁹⁴.

Introduction

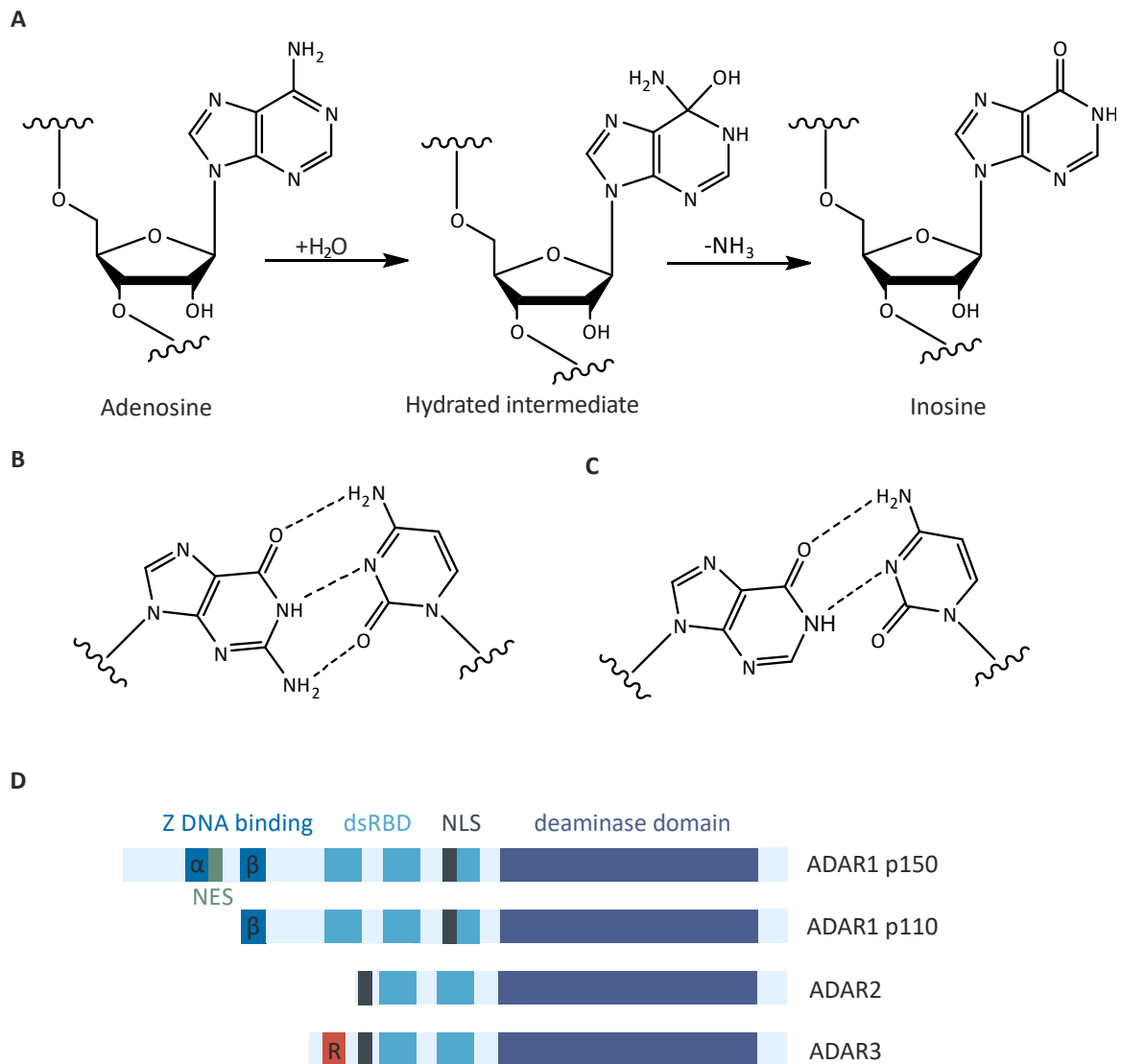


Figure 2: ADAR deamination and ADAR isoforms. A) Deamination reaction catalyzed by ADAR. An adenosine is converted to an inosine via a hydrated intermediate by nucleophilic attack of a water molecule and loss of an ammonia molecule. The resulting inosine is also able to base-pair to cytosine **C)**. However, in contrast to guanosine **B)** only two instead of 3 Watson-Crick hydrogen bonds can be formed. In **D)** the human ADAR variants with their different domains are depicted. All ADAR isoforms comprise of at least two dsRBDs, an NLS and a deaminase domain, although ADAR3 is catalytically inactive. ADAR1 p150 has two additional Z DNA binding domains with a NES in the Z α domain. The shorter ADAR1 isoform lacks the Z α domain and therefore also the NES.

1.2.2 Localization of ADARs

The nuclear localization signal (NLS) in the third dsRBD of ADAR1 results in an accumulation of ADAR1 p110 in the nucleus and nucleolus^{95,96}. On the other hand, ADAR1 p150 is found to be predominantly cytoplasmic due to a strong nuclear export signal (NES) in its N-terminal Z α domain^{96,97}. However, both ADAR1 isoforms shuttle between the cytoplasm and nucleus⁹⁷. While the nuclear export of ADAR1 p150 is mediated by nuclear

export factor exportin-1 (XPO1) binding to the NES in concert with RAN-GTP⁹⁸, exportin-5 (XPO5) together with RAN-GTP regulate the export of ADAR1 p110 by dsRNA binding to the dsRBDs⁹⁷. The import of ADAR1 into the nucleus is dependent on transportin-1 (TRN1) binding to dsRBD3 in the absence of dsRNA⁹⁹. Like ADAR1 p110, ADAR2 localizes to the nucleus, predominantly to the nucleolus⁹⁵. This is determined by karyopherin subunit α 1 (KPNA1, also known as importin subunit α 1) and KPNA3 by binding to the N-terminal NLS of ADAR2¹⁰⁰. Both ADAR1 and ADAR2 are ubiquitously expressed in different human tissues, whereas ADAR3 is reported as brain-specific.^{91, 101, 102}

1.2.3 Recoding ADAR substrates

While guanosine can form three hydrogen bonds to cytidine, inosine can form two hydrogen bonds to cytidine (**Figure 2 C, D**) and is read by the translational machinery as G. Therefore, ADARs are capable of introducing formal A-to-G mutations on RNA level¹⁰³. The first A-to-I editing in mammals was discovered at specific sites in the GRIA2 (GluR2/ GluR-B) subunit of the AMPA glutamate receptor (GluR)¹⁰³⁻¹⁰⁵. The probably most prominent editing sites in that context are the Q/R and R/G site, named after their respective amino acid alteration. Q (CAG) to R (CGG, here CIG) editing results in loss of Ca²⁺-ion permeability of GluR2 containing channels (**Figure 3**)^{103, 104}. R to G recoding was not only found in GluR2 but also in GluR3 and GluR4 subunits of AMPA glutamate receptor and alters receptor desensitization. Similar to GluR2, GluR5 and GluR6 subunits of kainate glutamate receptors contain a Q/R recoding site that also affects Ca²⁺-ion permeability^{103, 105}.

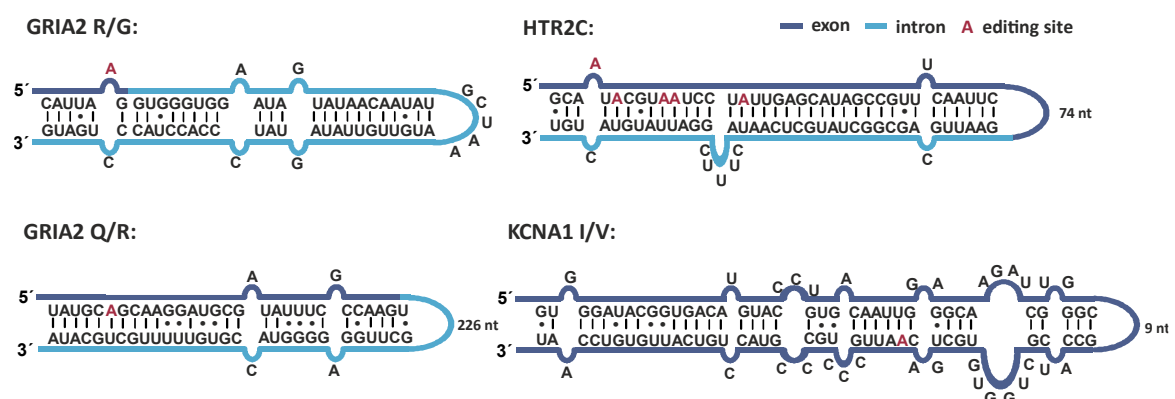


Figure 3: Recoding ADAR substrates. Exemplary recoding editing sites (in red) and their surrounding predicted secondary structure are depicted. Frequently, the dsRNA substrate for ADAR is formed by an exon (dark blue) and an intron (light blue) in recoding sites. Editing in those sites usually results in altered protein properties.

Introduction

Furthermore, several other editing sites in coding regions have been found such as the I/M site in the GABA_A receptor subunit $\alpha 3$ (GABRA3)^{106,107}, the I/V site in the voltage gated K⁺ channel (K_v1.1/KCNA1)¹⁰⁸ and multiple sites in the serotonin receptor 2C (HTR2C)¹⁰⁹. Interestingly, in contrast to the R/G and the Q/R site in GluR2, where the double-stranded substrate for ADAR is formed by a duplex of intronic and exonic RNA, the I/V site of the K_v1.1 and the I/M site in GABRA3 reside in double-stranded exons^{104, 107, 108}. All of the above recoding editing sites lead to physiological alterations related to neurotransmission^{103, 105-109}. While ADAR2 predominantly edits coding sites in the brain, where it is most highly expressed¹¹⁰, ADAR1 is the major enzyme for editing non-coding sites¹⁰². Remarkably, in ADAR2 null mutant mice only the lack of Q/R site editing in the GluR2 subunit of AMPA receptors is responsible for epileptic seizures and shortened lifespan of only several weeks. These seizures and finally death are a result of Ca²⁺ influx in neurons¹¹¹. All other recoding sites edited by ADAR2 are believed to be at least partly compensated by ADAR1 editing¹¹². Some recoding sites revealed to be cancer associated like the Q/R site in the filamin- α (FLNA)¹¹³, the M/V site in filamin- β (FLNB)¹¹⁴ or the S/G site in antizyme inhibitor 1 (AZIN1)^{114, 115}.

1.2.4 Non-coding ADAR substrates

Editing in coding sites represents only a minor fraction of A-to-I editing events. Beyond that, editing was found to occur in primary miRNA transcripts¹¹⁶⁻¹¹⁸. On the one hand, RNA editing of miRNAs can lead to altered specificity, especially if it occurs in the seed region, for example in miR-376a¹¹⁷. On the other hand, miRNA maturation by the Drosha-DGCR8 complex can be inhibited by altering the RNA hairpin recognition of miRNA precursors due to RNA editing¹¹⁸.

The vast majority with 97.7% of all identified editing sites as revealed by RNA-seq of the human transcriptome, occur in repetitive sites like Alu repeats and in noncoding regions for example introns and 3'UTRs^{101, 102, 119-126}. One functional consequence of editing in introns can be the generation of splice donor (GU) and acceptor (AG) sites¹²¹. An example for that is the exonization of an Alu-exon in human nuclear prelamin A by the editing dependent generation of an AG acceptor splice site¹²⁷. Another prominent example is the creation of an acceptor splice site by ADAR2 editing in an intronic sequence of its own pre-

mRNA¹²⁸. As a result, a non-functional ADAR2 protein is expressed. Therefore, ADAR2 expression is negatively regulated by its own editing activity¹²⁹.

However, editing of endogenous non-coding dsRNA by ADAR1 is crucial to suppress interferon activation. Deficiency of ADAR1 in mice leads to death at the embryonic stage due to interferon upregulation, defective hematopoiesis and apoptosis¹³⁰⁻¹³⁴. Interestingly, the mutation of mitochondrial antiviral signaling adaptor protein (MAVS) in Adar1-null mice can rescue the mice from embryonic lethality¹³⁵. The same is true for the melanoma differentiation-associated protein 5 (MDA5)¹³⁴. Since both MAVS and MDA5 are important in the interferon activation pathway, inosines in dsRNA originating from ADAR1 are essential to inhibit antiviral inflammation mediated by cytosolic innate immune system¹³⁴⁻¹³⁶. Indeed, it has been shown that hyper-edited dsRNA with multiple I-U wobble base pairs results in suppression of interferon-stimulated genes (ISGs) in cell culture by inhibition of MDA5 and retinoic acid-inducible gene 1 (RIG-1)^{136, 137}.

Mutations of ADAR1 were identified to be a cause of the autosomal-recessive inflammatory disease Aicardi-Goutières syndrome (AGS)¹³⁸. The increased interferon expression and aberrant innate immune response in AGS patients is a result of decreased editing of endogenous Alu repeats¹³⁵. This consequently drives “non-self” recognition of this endogenous dsRNA by MDA5 and protein kinase R (PKR)^{136, 139}. The role of ADAR1 is not only to prevent MDA5 and PKR activation by endogenous dsRNA but it also avoids translational shutdown of endogenous RNAs during interferon response¹³⁹.

1.2.5 Structure and enzymatic mechanism of ADAR

Although there is no crystal structure of any full-length ADAR, the structure and function of the deaminase domain (dd) of ADAR2 (**Figure 4**) has been well characterized^{140, 141}. Furthermore, the solution structures of both of ADAR2's dsRBDs give some insights into the binding of ADARs to their substrate¹⁴². Unfortunately, only little is known about ADAR1.

In the catalytic center of ADAR2 deaminase domain there is a zinc ion, coordinated by amino acid residues H394, C451 and C516 (**Figure 4 C**). Additionally, a water molecule is coordinated by the zinc ion and residue E396. That water molecule is thought to replace the ammonia in the adenosine deamination¹⁴⁰. Analogous active sites can be found in

Introduction

other deaminases like cytidine deaminase (CDA) and TadA (tRNA-specific adenosine deaminase)^{143, 144}. In close proximity of the coordinated zinc ion there is an inositol hexakisphosphate surrounded by several arginine and lysine residues. Unexpectedly, inositol hexakisphosphate revealed to be a crucial cofactor for the catalytic activity of ADAR2¹⁴⁰. Structural studies utilizing a dsRNA substrate with 8-azanebnularine mimicking the hydrated intermediated of the deamination showed that the hADAR2 deaminase domain binds approximately 20 nt of dsRNA by contacts to both strands (**Figure 4 A,B**)¹⁴¹. This is in good accordance with previous reports using ribonuclease footprinting analyses, where ADAR2 deaminase domain protected a 23nt RNA strand with 18nt to the 5' end of the editing site and 5 nt to the 3' end¹⁴⁵. The protein contacts the RNA mainly through the phosphodiester-ribose backbone close to the target adenosine¹⁴¹. Importantly, the earlier postulated¹⁴⁰ base-flipping mechanism could be proved¹⁴¹.

The target A is flipped out of the dsRNA by a loop in the ADAR2 deaminase domain. The E488 residue in this loop occupies the free space in and forms hydrogen bonds with the base opposing the target A (**Figure 4 C-E**). This results in a conformational change of the RNA from its usual A-form. This makes the adenosine accessible to the active site where it can interact with several amino acids and the deamination takes place¹⁴¹. Replacement of the E488 by a Q results in a hyperactive variant of ADAR that is especially superior in the context of a G as 5' nearest neighbor of the target regarding its editing efficiency¹⁴⁶

In fact, the editing efficiency of the target A depends strongly on the sequence context. The identity of the base opposite of the target A is one critical factor. While an A-C mismatch yields the highest editing, A-U base pairing is also well tolerated in many substrates¹⁴⁷. Purines on the other hand, especially G, have been found to strongly impair the editing¹⁴⁷. This can be explained by a clash of the larger purines with E488 of ADAR2¹⁴¹.

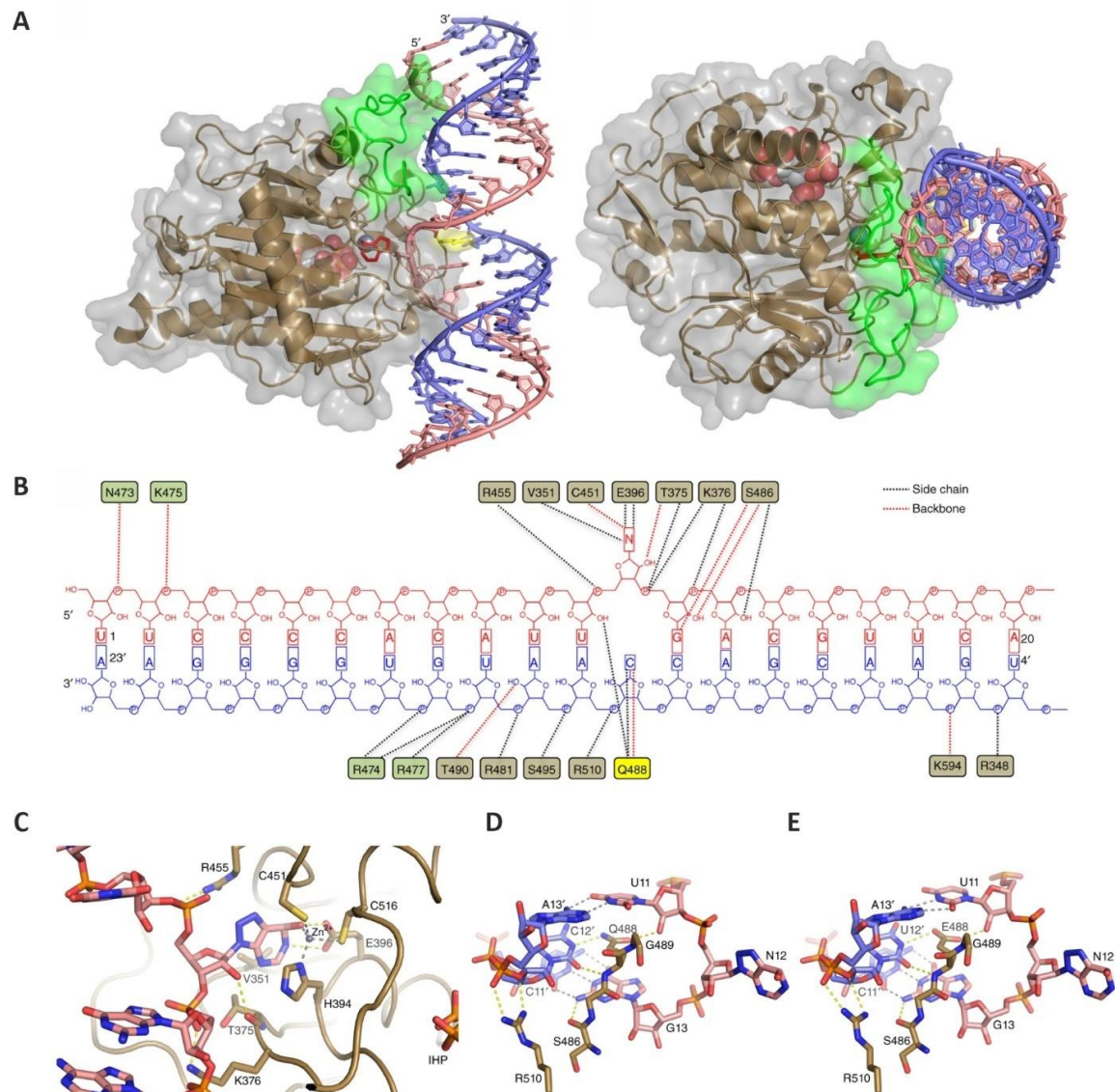


Figure 4: Structure of ADAR2 deaminase domain. **A)** hADAR2 E488Q deaminase domain bound to the natural ADAR substrate Bdf2-C. The flipped-out base N is highlighted in red, the zinc-ion in grey and the Q488 in yellow. **B)** the BDF-2C dsRNA substrate is depicted with the respective contacts of the protein. **C)** The catalytic center with the coordinated zinc ion and the flipped-out target A. **D)** contacts of the target A opposed, orphan nucleotide in the Bdf2-C substrate with the hADAR2 E488 deaminase domain. **E)** The same site as in **D)** but with the Bdf2-U substrate in contact with the hADAR2 wt. Figures taken with permission from¹⁴¹.

The nearest neighbors, that is the bases 5' and 3' of the target A have also decisive influence on editing efficiency. For the base 5' of the target A the base preference is equal for ADAR1 and ADAR2 and is $U>A>C>G$ ¹⁴⁸. This is in contrast to the 3' preference, where G is the most preferred base for both enzymes. However, apart from that, the 3' preference differs between the two enzymes. For ADAR1 it is: $G>C\approx A>U$ while the preference for ADAR2 is $G>C>U\approx A$ ¹⁴⁸. The full length ADARs have slightly different 3' preferences than the ADAR deaminase domains alone¹⁴⁸. The 5' neighbor preference of

Introduction

ADAR can be explained by a clash of a 2-amino group (provided by a G-C pair) in the minor groove that is expected to clash with the G489 residue of ADAR2¹⁴¹. On the other hand, the 3' preference of a G can be explained by a hydrogen bond between the 2-amino group of the G and S486. This hydrogen bond to the minor groove can only be formed by a G and not the other three common bases¹⁴¹. Furthermore, a loop in the ADAR2 deaminase domain (aa454-477) could be identified that binds the RNA close to the editing site. Interestingly, this loop region is well conserved in the ADAR2 enzyme family but differs in the ADAR1 enzyme family. Therefore, it is suggested to be at least in part responsible for the difference in substrate specificity of ADAR1 and ADAR2¹⁴¹.

Although ADAR deaminase domains are sufficient to recognize and edit some targets, the dsRBDs significantly increase editing yields or are necessary in most targets^{148, 149}. They consist of a typical α - β - β - β - α protein conformation with two α -helices being packed along three antiparallel β -sheets (**Figure 5**)^{142, 149}. Similar to other dsRNA binding proteins a conserved lysine-rich sequence in the N-terminal part of α 2 interacts non-sequence specific with the phosphate backbone in the major groove^{142, 150}.

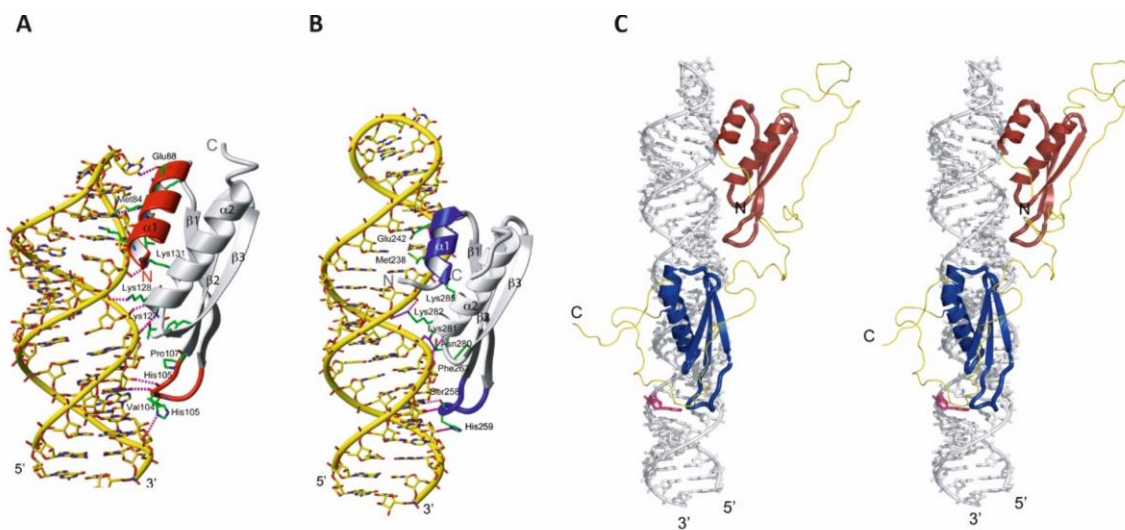


Figure 5: Structure of the dsRBDs of ADAR2. **A)** NMR-structure of the ADAR2 dsRBD1 bound to the GRIA2 R/G upper stem loop and **B)** the dsRBD2 bound to the GRIA2 R/G lower stem loop. **C)** depicts a combined model of both dsRBDs of ADAR2 bound to the GRIA2 R/G site. Figures taken with permission from¹⁴².

An NMR solution structure of the single dsRBDs of ADAR2 with parts of the GluR2's R/G stem loop structure revealed that 12-14 base pairs spanning two minor and one major groove are bound by ADAR2's dsRBDs. Surprisingly, it was found that the dsRBDs interact in a sequence specific manner with both minor grooves¹⁴². This might be another reason

for the different substrate specificity of ADAR2 and ADAR1, apart from their different deaminase domains. However, despite considerable progress in the field it remains elusive how the full length ADAR2 with both dsRBDs and the deaminase domain together binds and edits its target RNA. Although the structure of the deaminase domain and the dsRBD2 of ADAR2 binding to the GluR2 R/G site suggest to bind overlapping regions which seems to be impossible, some models of concurrent binding of both subunits exist¹⁵¹. Furthermore, it was reported that ADARs not only edit RNA but also DNA in DNA/RNA hybrids¹⁵². The physiological relevance of this however remains subject of research. There is also evidence that ADARs require dimer formation to be active¹⁵³⁻¹⁵⁶. But the dimer formation and RNA binding are still being investigated. Due to the lack of structural data, even less is known about ADAR1.

1.3 Site-directed RNA editing

The basic concept of site-directed RNA editing is to recruit either ADAR or an ADAR-dd fusion protein to a specific target site on a transcript within a cell. *In vitro* studies could confirm that inosine is identified as guanosine by the translational machinery with few exceptions and apart from the UUI codon, protein truncation seems to affect only codons with multiple inosines¹⁵⁷. Thus, a formal A-to-G mutation can be introduced at the desired target site on RNA level. Consequently, site-directed RNA editing opens the possibility of recoding 12 amino acids, change splice sites, start, and stop codons¹⁵⁸.

1.3.1 The SNAP-ADAR approach

The first approach to exploit the deaminase activity of ADARs for site-directed RNA editing utilizing an engineered fusion protein, the so called SNAP-ADAR was first published in 2012¹⁵⁹. This protein comprises a SNAP-tag and a deaminase domain of ADAR1 or ADAR2¹⁵⁹. Instead of the dsRBDs, SNAP-ADAR proteins use covalently bound guide RNAs (gRNAs) that form a dsRNA with the target site for substrate recognition. The covalent binding is mediated by a benzylguanine (BG) modification at the 5' end of the gRNA that reacts with the SNAP-tag¹⁵⁹.

The positioning of pyrimidines, especially C opposite of the target A in the center of these 14-20 nt gRNAs yielded highest editing¹⁵⁸. Chemical modification with 2'-OMe and

Introduction

phosphorothioates of the gRNAs revealed to be well accepted at most positions, except for the central triplet opposite the target A and improved editing yields, while simultaneously reducing bystander off-target editing in cell culture^{160,161}. With the *in vitro* editing of Factor V Leiden mutation, a first example for using the SNAP-ADAR approach in a biomedical context was demonstrated¹⁶¹. First *in vivo* experiments with light-inducible BG-gRNAs, using the NPOM photoprotection group, mRNA encoded SNAP-ADAR and EGFP-reporter constructs, were demonstrated in *Platynereis dumerilii*¹⁶². In another application of these light-inducible gRNAs, the alteration of cellular localization of different proteins could be achieved by SNAP-ADAR mediated RNA editing¹⁶³. The use of engineered fusion proteins allowed to boost the editing efficiency, by utilizing the hyperactive E/Q variants of ADAR1 and ADAR2¹⁶⁴. Therefore, also the editing of difficult codons like the GAN codons was possible¹⁶⁴.

Furthermore, a codon scope of the hyperactive SNAP-ADAR1 and SNAP-ADAR2 constructs similar to the scope of wt ADARs was found^{148,164}. With the editing of KRAS and STAT1 potential application of the technique were established. Also, multiple endogenous targets could be edited simultaneously with high yields up to 90%. It was demonstrated that editing in the 3'UTR of endogenous transcripts resulted in higher editing than in the ORF or the 5'UTR. Especially the SNAP-ADAR fusion proteins with the wt deaminase domains achieved very few global off-target editing events¹⁶⁴. However, although the target editing efficiency of the hyperactive mutants of SNAP-ADARs is very high and advantageous, especially for difficult codons, their global off-target editing in cells that stably overexpress those proteins is very high and might lead to unpredictable side effects¹⁶⁴.

1.3.2 The λ N-ADAR approach

Another approach, published shortly after the SNAP-ADAR approach, makes also use of an artificial fusion protein. Here the deaminase domain of ADAR2 was fused to a λ N peptide. This phage derived peptide is able to bind an RNA BoxB-hairpin with high affinity. Therefore, gRNAs antisense to the target mRNA containing a 17 nt BoxB hairpin in different positions relative to the target A were designed to recruit the λ N-ADAR2-dd fusion protein. Same as for the SNAP-ADAR approach the base opposite the target was a C¹⁶⁵.

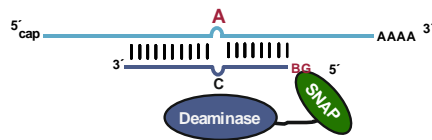
With that first approach, the premature stop codon of cystic fibrosis causing mutation cystic fibrosis transmembrane conductance regulator (CFTR) W496X could be edited quantitatively *in vitro* and up to 20% in *Xenopus* oocytes. Furthermore, functional chloride currents across the membrane could be restored in *Xenopus* oocytes by injection of mRNA of deaminase fusion protein, target mRNA and gRNA¹⁶⁵. Also, the fluorescence of a W58X GFP reporter could be restored in human cells with editing yields of 20% by ectopic expression of target, gRNA, and fusion protein from plasmids. However, moderate bystander off-target editing was detected¹⁶⁵. In a further study, fusion of up to 4 λ N peptides to the ADAR2-dd and including 2 BoxB hairpins increased editing yields of a GFP W58X reporter construct in HEK cells. Additionally, employing the hyperactive ADAR2-dd E488Q mutant in the fusion protein further improved the editing yields up to 70%¹⁶⁶. However, especially the ADAR2-dd E488Q fusion protein revealed to produce severe off-target editing¹⁶⁶. Sinnamon et al. successfully expressed the improved λ N-ADAR2-dd E488Q fusion protein together with six copies of the two BoxB containing gRNAs from an adeno-associated virus (AAV) in murine hippocampal neurons *ex vivo*. Targeting the endogenous Methyl CpG Binding Protein (MECP2) R106Q Rett syndrome related mutation, they could achieve editing yields of up to 72% in a CAA codon context of the target A¹⁶⁷. Moreover, the wild-type phenotype of MeCP2 heterochromatin enrichment could be verified in the edited cells by fluorescence microscopy of the GFP-tagged MECP2 R106Q¹⁶⁷. While the SNAP-ADAR approach yields high editing and the chemical modifications of the gRNA can suppress bystander off-target editing, the main advantage of the λ N-ADAR approach is that all components are genetically encodable and can be delivered to cells by AAVs.

1.3.3 The MS2-ADAR approach

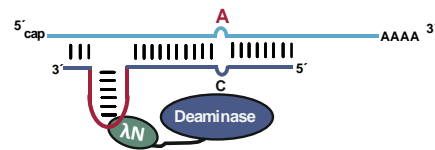
A newer approach utilizes the MS2 protein fused to the deaminase domain of ADAR1 together with a gRNA comprising a 21nt antisense part connected to 6 MS2 RNA loops on the 5' end¹⁶⁸. Same as for all the other approaches a mismatched C opposite the target A was inserted. However, only 5% editing on a EGFP premature UAG stop codon could be achieved overexpressing all components from plasmids in HEK cells¹⁶⁸.

Introduction

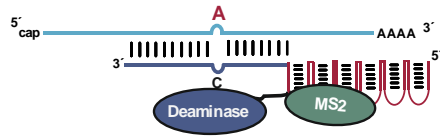
SNAP-ADAR:



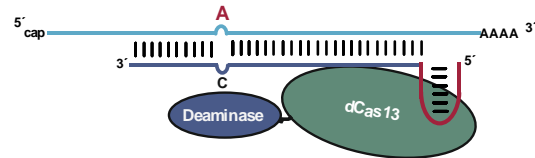
λ N-ADAR :



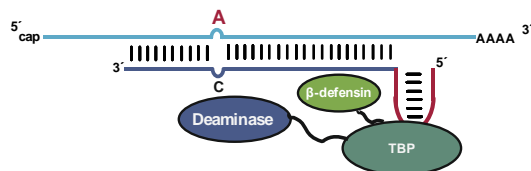
MS2-ADAR:



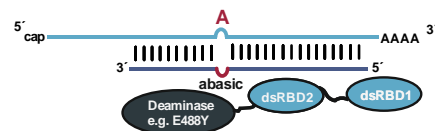
dCas13-ADAR:



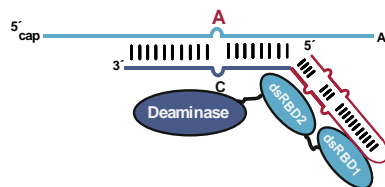
CIRTS-ADAR:



bump-hole approach:



wild-type-ADAR:



wild-type-ADAR:

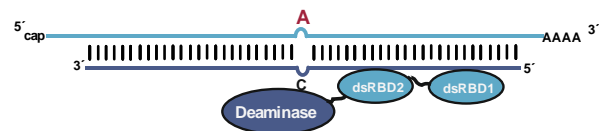


Figure 6: Site-directed RNA editing strategies. Different approaches for site-directed RNA editing utilizing either artificial deaminase fusion proteins (green) for targeting with the respective guide RNAs (antisense part in dark blue and fusion protein specific recruiting domain in red) or wild-type ADAR.

1.3.4 The CRISPR-Cas13-ADAR approach – REPAIR and RESCUE

The CRISPR-Cas13 variant was shown earlier to target mRNA instead of DNA¹⁶⁹. In this approach named REPAIR (RNA Editing for Programmable A to I Replacement), the catalytically inactive Cas13 protein was fused with the hyperactive E/Q variant of ADAR1-dd or ADAR2-dd respectively¹⁷⁰. Although the hyperactive deaminase domains of both ADARs were tested, ADAR2 seemed to yield better editing. A typically 50 nt antisense part with an A-C mismatch at the target site with an additional 36 nt direct repeat sequence to recruit the dCas13 serves as guide RNA. With all components ectopically expressed from

plasmids, 34 disease relevant targets in a UAG codon context could be edited, with efficiencies ranging from almost zero up to 28%. Interestingly, the codon scope deviated from that of wild-type ADAR and the variation of editing yields was not as pronounced as for the wild-type ADARs^{148, 170}. However, this approach suffered from very high off-target editing. For this reason, Cox et al. mutated the ADAR deaminase domain to find a version referred to as REPAIRv2 with an additional T375G mutation in the E488Q ADAR2-dd. REPAIRv2 did not only show significantly lower off-target editing, but also decreased on-target editing. In the endogenous PPIB target less than half of the editing compared to REPAIRv1 was achieved¹⁷⁰. Unlike the fusion proteins in all the other approaches, the size of the dCas13-ADAR-dd protein was too large to fit the packaging limit of AAVs. Thus, the authors performed a C-terminal truncation of the protein that allowed packaging into AAVs without the loss of activity¹⁷⁰. However, they did not demonstrate editing with the components expressed from an AAV.

Abudayyeh and Gootenberg et al. developed very recently a system called RESCUE (RNA Editing for Specific C to U Exchange) expanding the scope of RNA editing from A-to-I to C-to-U editing¹⁷¹. RESCUE utilizes the same dCas13 platform as REPAIR, however, an ADAR2-dd was evolved to deaminate Cs. A shortened guide RNA with a 30 nt antisense part and a C or U opposite of the target C revealed to be optimal. Editing of several phosphorylation sites of the endogenous β -catenin transcript yielded 5-28% editing. As a functional consequence, the Wnt/ β -catenin signaling was induced and increased growth of HEK293FT and HUVEC cells¹⁷¹. However, not only caused this version of RESCUE severe C-to-U off-target editing, but also A-to-I off-target editing was discovered. The authors exploited this A-to-I off-target editing for site-directed multiplexed A-to-I and C-to-U editing. But again, to reduce the overall off-target editing they evolved the ADAR2-dd even more and identified S375A as the mutant with highest specificity¹⁷¹.

1.3.5 The bump-hole approach

The bump-hole approach is another attempt based on the structural features of ADAR2 to yield site-directed A-to-I editing with minimal off-target sites¹⁷². Here, residue 488 of the full length ADAR2 was substituted by either phenylalanine, tyrosine or tryptophan to introduce a steric clash with the nucleobase opposite of the target A. Thereby, the editing efficiency in regular dsRNA and hence global off-target editing by ectopic expression of

Introduction

the enzyme is significantly reduced. In order to maintain the editing capacity at the target, an abasic site opposite of the target A on the 39 nt guide RNA was introduced. Similar to the shorter SNAP-ADAR gRNAs, the gRNAs of this approach are fully 2'-OMe modified except for a few nucleotides opposite of the target A and two linkages at the 5' and 3' ends are phosphorothioates¹⁷². Transfecting both, a plasmid with the mutated enzyme and the chemically modified gRNA in HEK293T cells yielded editing of up to 55% in the 3'UTR of endogenous RAB7A transcripts, comparable to the wild-type hADAR2 with a C opposite the target A instead of an abasic site. For the 3'UTR of endogenous β -actin transcripts, all three mutants combined with abasic site gRNAs were significantly worse than the corresponding wild-type control. However, the editing of six endogenous off-target sites could either be completely abolished or drastically reduced by using the bump-hole approach instead of the wild-type ADAR¹⁷².

1.3.6 The CIRTS-ADAR approach

The CRISPR-Cas-inspired RNA targeting system, or short CIRTS approach, is very similar to the CRISPR-Cas13-ADAR system¹⁷³. However, the main objective here was to engineer a fully genetically encodable platform with a fusion protein that consists only of human proteins, to prevent potential adverse effects, originating from bacterial proteins like CRISPR or the λ N peptide. For the RNA editing platform the authors fused either the ADAR2-dd or its corresponding hyperactive variant with the RNA hairpin binding TAR binding protein (TBP) and the ssRNA binding protein β -defensin 3. The recruitment to the target site was mediated by a 31 nt TAR hairpin scaffold binding to TBP on the 5' end of an antisense guide sequence that binds the target mRNA. The β -defensin part served as protection of the free gRNA from degradation. This approach reached approximately 15% editing with the wild-type ADAR2-dd and almost 50% editing with the hyperactive E488Q version in a premature UAG stop codon of a dual luciferase reporter¹⁷³.

1.3.7 Recruiting wild-type ADAR

The probably most attractive approach for site-directed RNA editing in a therapeutic setting, is probably the recruitment of endogenous wild-type ADAR, since this does not require the delivery of an artificial fusion protein. The first approaches to harness the deaminase function of wild-type ADARs to introduce site-specific A-to-I conversions were reported in 1995 by Woolf et al ¹⁷⁴. After injection of a 52 nt completely antisense guide

RNA hybridized to a dystrophin mRNA with a premature UAG stop codon followed by a luciferase reporter into *Xenopus* oocytes, they could partly restore the luciferase activity with endogenous ADAR only¹⁷⁴. Additionally, they were able to show that a shorter 34nt version was also active in cell extracts, although to a lower extent than the 52 nt gRNA, whereas a 25 nt gRNA did not reveal any activity. Fully 2'-OMe modification of the 34nt version completely abolished editing activity. Even an end-blocked version with only the last five nucleotides on both ends modified with 2'-OMe and phosphorothioates showed very weak activity compared to the unmodified gRNA¹⁷⁴.

In 2016 Wettengel et al. and independently from that Fukuda et al. reported a plasmid-borne approach, where both the wild-type ADAR2 enzyme and a guide RNA were expressed from plasmids in human cells^{175, 176}. The so called R/G gRNA comprises a 16-29 nt part antisense to the target transcript with an A-C mismatch opposite of the target A, and a 45nt hairpin inspired by the natural ADAR2 R/G editing site in the GluR-B transcript¹⁷⁵. Under optimized conditions ADAR2 with such an R/G gRNA yielded almost quantitative deamination on a UAG codon in a reporter transcript *in vitro*. However, although the antisense part of the gRNA was very short, some bystander off-target editing could be detected¹⁷⁵. Utilizing HEK293T cells with genomically integrated ADAR2, up to 65% editing could be achieved in a UAG premature stop codon of an EGFP reporter by co-transfection of separate plasmids carrying the reporter and the gRNA sequence under control of a U6 promoter¹⁷⁵. Ectopic expression of ADAR2 and R/G gRNAs in HEK293T cells yielded much lower editing on UAG targets of endogenous transcripts in the range of 0-38%¹⁷⁵. As an example of therapeutic use, a premature UAG codon in the PINK1 transcript was targeted in HeLa cells, overexpressing all components from plasmids. Mutations of PINK1 are related to Parkinson's disease¹⁷⁷. However, although editing yields of only 10% could be detected, in 85% of the cells expressing all components a PINK/Parkin mediated mitophagy phenotype could be rescued¹⁷⁵.

In another study editing with these R/G guide RNAs could be expanded to ADAR1 p110 and ADAR1 p150¹⁷⁸. Although the sequence of the R/G hairpin part of the guide RNA could be varied to avoid auto-editing within the gRNA, ADAR2 performed always better than the other two enzymes¹⁷⁸.

Introduction

Shortly after we published manuscript 2, the first study demonstrating *in vivo* RNA editing with R/G guide RNA and the MS2 approach was presented¹⁷⁹. For the R/G gRNAs, a screen with a 20-100 nt antisense part, 0-2 R/G hairpins and additional hairpins, as editing enhancer elements was performed. With a 100 nt antisense gRNA more than 10% editing on an endogenous RAB7A target could be achieved in HEK293T cells without the overexpression of ADAR2. However, with additional ADAR2 the editing yield of around 30% was not exceeded and a longer antisense part was not beneficial. In a comparison of the Cas13-approach, the R/G gRNA approach and the MS2 approach on an endogenous UAG codon in HEK293T cells, overexpression of ADAR2 together with R/G gRNAs revealed good editing with very low off-target editing. The MS2-ADAR2-dd with an additional NES resulted in similar editing with few off-target sites. Utilizing the ADAR1-dd or the hyperactive mutants led to severe off-target editing. In this comparison the REPAIR approach performed worst with low editing and high off-target editing¹⁷⁹. For *in vivo* proof of concept, Katrekar et al. chose two different mouse models, the *spf^{ash}* mouse model of ornithine transcarbamylase (OTC) deficiency and *mdx* mouse model of Duchenne muscular dystrophy (DMD)¹⁷⁹. For delivery two copies of gRNAs under the control of U6 promoters and one copy of full length ADAR2 or its hyperactive E488Q mutant were packed in an AAV8 vector. A conversion of the challenging UAA ochre stop codon to UII of up to 1% with ADAR2 and R/G gRNAs in the *mdx* mouse could be achieved. Furthermore, 1-2.5% protein restoration could be detected in the treated muscles of the *mdx* mice. Application of the MS2 approach with the hyperactive ADAR1 E1008Q variant to the same mouse model yielded up to 2.4% conversion to UII. Even more pronounced was the effect in the *spf^{ash}* mouse model, where the hyperactive ADAR2 with gRNA yielded 5-34% editing¹⁷⁹.

Very recently and several months after we published our RESTORE approach, another study demonstrated the recruitment of endogenous ADARs in cell lines and primary cells using gRNAs that are completely antisense, except for the target A-C mismatch¹⁸⁰. However, with 111-151 nt the used gRNAs are the longest among all systems. Despite their length, for most endogenous targets only low to moderate editing yields could be achieved. Lentiviral transduction of a 151 nt gRNA construct resulted in less than 10% editing in HEK293T cells on an endogenous target. For the restoration of the human IDUA

W402X mutation in Hurler syndrome patient derived fibroblasts up to 30% editing was detected with a synthetic 111 nt oligonucleotide. The three nucleotides at both ends of the oligonucleotide were equipped with 2'-OMe modifications and phosphorothioate linkages. In an enzymatic assay Qu et al. were able to restore the IDUA enzyme activity to the level of the activity of the less severe Scheie disease phenotype¹⁸⁰.

1.4 Diseases for targeting with RNA editing

About 58% of the more than 54000 disease-associated genetic variations in humans represent point mutations⁷⁹. Interestingly, G-to-A mutations are significantly overrepresented compared to the other possible changes⁷⁹. Thus, site-directed A-to-I RNA editing appears to be a promising tool for the direct targeting and reversal of a huge number of disease-related mutations. Furthermore, modulation of post-translational modifications (PTMs) such as phosphorylation, ubiquitination, acetylation and glycosylation to reverse disease phenotypes by RNA editing could be envisioned. This opportunity arises by the possibility of selectively targeting and thereby changing amino acids important for PTMs like tyrosines, threonines, some serines, lysines and some arginines. In the following sections some disease related targets that were studied for the application of site-directed RNA editing as a therapeutic agent during this thesis are discussed.

1.4.1 Mucopolysaccharidosis/ Hurler/ Scheie syndrome

Hurler syndrome is the most severe type of the class of lysosomal storage disease Mucopolysaccharidosis type I (MPS-I). It is named after Gertrud Hurler who described the symptoms of the disease including mental retardation, hearing loss, corneal clouding and predominantly multiple skeletal abnormalities in 1919¹⁸¹. Depending on the severity of the symptoms, it is distinguished between three types of MPS-I. Beside the severe Hurler syndrome, and the much milder Scheie syndrome, an intermediate form, referred to as Hurler-Scheie syndrome exists¹⁸¹⁻¹⁸³. Without treatment, Hurler patients have a decreased lifespan of less than 10 years resulting from brain damage or cardiorespiratory complications¹⁰¹. The underlying cause of MPS-I has been identified as the dysfunction of α -L-iduronidase (IDUA), an enzyme responsible for the degradation glycosaminoglycans

Introduction

(GAGs)¹⁸⁴. This leads to enrichment of GAGs in the lysosome resulting eventually in cell death in several organs and tissues. A high number of over 100 mutations in the IDUA gene are known¹⁸². However, the most common mutation in the autosomal recessively inherited disorder is the W402X mutation in the IDUA gene. Patients with two alleles of such a nonsense mutation always suffer from the severe Hurler syndrome¹⁸². The model case of the significantly milder Scheie syndrome is compound heterozygous with one allele bearing the W402X mutation and a G→A mutation in intron 5, which creates an alternative splice site resulting in a premature termination codon^{185, 186}. Despite these mutations, Scheie syndrome patients are not affected of impaired intelligence and have a normal lifespan¹⁸⁵. However, in addition to cardiovascular problems they also suffer from cloudy corneas and decreased joint mobility¹⁸⁵.

Unfortunately, to date there are only very limited possibilities for causal treatment of MPS-I. For patients with severe Hurler syndrome hematopoietic stem cell transplantation (HSCT) below the age of 2.5 years is the treatment of choice¹⁸⁷. Enzyme replacement therapy (ERT) is the only other existing alternative to HSCT for treatment¹⁸⁷. However, while HSCT can significantly improve the lifespan of Hurler patients, a major limitation is the availability of suitable donors. ERT on the other hand is not suitable for treatment of the CNS phenotype of Hurler syndrome, due to the impermeability of the blood-brain barrier^{188, 189}. Newer therapeutic approaches focus on gene therapy and genome editing¹⁹⁰⁻¹⁹². Since the most common Hurler syndrome mutation is the W402X (UGG→UAG), RNA editing seems to be a valuable tool for treatment considering that UAG is the most preferred codon by ADARs and restoration of small enzyme activity, as found in the much milder Scheie syndrome, could lead to a significant improvement.

1.4.2 Alpha-1-Antitrypsin deficiency

α1-Antitrypsin deficiency (AATD) is a genetic disease caused by mutations in the SERPINA1 gene coding for α1-antitrypsin, a serine proteinase inhibitor (SERPIN) family protein¹⁹³. Although a great number of mutations were found in the SERPINA1 gene, the two most common ones in AATD patients are the E342K or PiZ and the E264V or PiS mutation¹⁹⁴⁻¹⁹⁶. With at least 3.4 million individuals worldwide carrying two of the deficiency alleles (PiZZ, PiSZ or PiSS), AATD is suggested to be one of the most frequent genetic diseases¹⁹⁷.

AAT is primarily expressed in hepatocytes and secreted into the blood circulation. The main function of AAT was identified to be the inhibition of neutrophil elastase^{198, 199}. This is of distinct importance in the lung, where uncontrolled neutrophil elastase activity can lead to severe tissue damage²⁰⁰.

The underlying reason of AATD is a loop-sheet polymer formation of the PiS and more pronounced of the PiZ mutant AAT protein^{201, 202}. These polymers subsequently accumulate in the endoplasmic reticulum (ER) of hepatocytes²⁰³. As a result, blood AAT levels are decreased in AATD patients. Homozygous PiS mutant individuals have only about 60% of AAT levels in the blood compared to the wild-type PiM variant, while PiZ homozygotes reach blood AAT levels of only 10-15%²⁰². In particular in combination with smoking the risk of COPD (chronic obstructive pulmonary disease) and emphysema is drastically increased^{195, 204}. On the other hand, PiZ homozygous patients can also develop severe liver disease caused by AAT polymer inclusions²⁰⁵. This liver disease can manifest as juvenile hepatitis, liver cirrhosis and hepatocellular carcinoma²⁰².

To date, the only causal treatment for AATD, besides liver transplantation, is enzyme replacement therapy using pooled human AAT²⁰⁶. Furthermore, gene therapy approaches are under investigation²⁰⁷. To address the liver disease, antisense oligos for knock-down of polymer-causing PiZ mutant AAT have been developed and successfully tested in a mouse model²⁰⁸. Site-directed RNA editing is especially attractive as a therapeutic approach for AATD, since the PiZ mutant could be reversed and therefore both the lung and the liver phenotype could be treated simultaneously.

1.4.3 Rett syndrome

The Rett syndrome is an X-chromosome linked progressive brain disorder with a typical onset of symptoms at 6-18 months after birth. Besides the typical hand moving pattern, symptoms include mental retardation, autism, ataxia and dementia^{209, 210}. Due to the X-linked character and the fact that most of the mutations are de novo mutations of paternal origin, this disease is almost exclusively found in females and only very rare cases of affected males have been identified^{211, 212}. However, the underlying cause, a mutated methyl-CpG-binding protein 2 (MECP2) gene was only found in 1999²¹³. The role of MeCP2 is complex, but it plays a key role in gene expression, acting both as transcriptional

Introduction

activator and repressor²¹⁴. Mutations in MECP2 linked to Rett syndrome result in morphological alterations in the brain such as smaller neurons and dendritic changes²¹⁵.

At present, Rett syndrome cannot be cured but importantly, since it is not a neurodegenerative disease, reversal is possible and could be shown in a mouse model²¹⁶. RNA editing seems to be the ideal way to treat Rett syndrome, since alternative methods like gene therapy can result in abnormally high production of MeCP2 protein that can also lead to adverse effects as reported for MECP2 duplication²¹⁷. Therefore, first steps for treating Rett syndrome with RNA editing have been made with the λ N-ADAR approach in mouse neurons¹⁶⁷.

2 Aim of the study

Recruiting endogenous ADAR instead of using ectopic expression of fusion proteins appears to be a very promising alternative to existing site-directed RNA editing strategies. Especially in the context of therapeutic applications, treatment with only an ADAR-recruiting ASO could not only be comparable to other established and approved oligonucleotide therapeutics but expand the scope of these. Therefore, major challenges of protein delivery required for other site-directed RNA editing approaches, CRISPR-based methods or gene therapy could be circumvented. More importantly, since RNA is targeted, the inherent risk of irreversible DNA damage existing for CRISPR and related DNA editing technologies is avoided. However, the major challenge is to find an ASO design that not only permits the recruitment of endogenous ADAR, but also fulfills other properties of an oligonucleotide therapeutic such as serum stability, high potency and good cellular delivery. Therefore, the aim of this study was to engineer chemically modified ASOs that enable harnessing of endogenous ADARs for site-directed RNA editing. As a first goal of this study, a method for producing such ASOs by *in vitro* transcription and ligation had to be established. Starting from the design of previously established gRNA from the plasmid-borne approach, the ASOs were optimized for their activity with different human ADARs, varying length and chemical modifications. In a next step, their ability to harness endogenous ADARs in different human cell lines and primary cells was explored. Additionally, the interferon dependency and differences of 3'UTR and ORF editing within endogenous targets was analyzed. Besides the determination of potency and duration of the RNA editing after ASO treatment, the off-target profile of this approach was analyzed. Also, first attempts for repairing disease-relevant mutations were evaluated. After characterization of this first generation of ASOs, further potential for optimization was anticipated. The aim for these second generation ASOs was not only to achieve higher efficacy, but also significantly decrease length by adjusting symmetry and modification patterns. To assess the therapeutic application of this approach, not only several pathogenic mutations were analyzed but also restoration of enzymatic activity in patient fibroblasts was evaluated. Finally, the use of ASOs for *in vivo* studies was envisioned and serum stability, delivery possibilities beyond transfection agents, and murine ADAR recruitment was assessed.

3 Results and discussion

3.1 Architecture and production of ASOs for site-directed RNA editing

Since the primary goal was to establish an oligonucleotide therapeutic that utilizes endogenous ADAR for site-directed RNA editing, the first challenge was the design, modification pattern and production of such an oligonucleotide. Those oligonucleotides are referred to as RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing) ASOs. One mandatory part of such a RESTORE ASO, is a part that is “anti-sense” to the target mRNA. Previously published data from the SNAP-ADAR approach indicated that dense modification with 2'-OMe and PS was well accepted by the ADAR deaminase domain and improved editing yields while decreasing bystander off-target editing¹⁶¹. Thus, the modification pattern from this approach served as an initial blueprint.

Besides this specificity domain, another crucial component that recruits ADAR enzymes to the target was necessary (**Figure 7**). In contrast to other site-directed RNA editing approaches with artificial fusion proteins^{161, 165, 170, 218} that form high affinity or even covalent interactions with the antisense guide RNA, only the dsRNA binding properties of wild-type ADAR could be employed. Therefore, the gRNA design from a plasmid-borne approach utilizing the natural GRIA2 R/G motif for ADAR attraction, developed in our group, was used as a starting point¹⁷⁵. Despite the ability of those gRNAs to recruit ADAR1 and ADAR2 for site-directed RNA editing, editing levels on endogenous targets were very low and overexpression of ADAR was indispensable^{175, 178}. However, the combination of a chemically modified specificity domain and the ADAR-recruiting domain resulted in a length of more than 70 nucleotides. This makes chemical synthesis of such molecules difficult and expensive. Therefore, a ligation strategy was developed where a chemically synthesized RNA was ligated to an *in vitro* transcribed RNA (**Figure 7**). This allowed the combination of any short chemically modified specificity domain and longer *in vitro* transcribed ADAR-recruiting domains. The advantage of the *in vitro* transcription using T7 RNA polymerase was that it is a relatively cheap and fast way to produce several different RNAs. As templates synthetic DNA oligonucleotides were used and 15% DMSO was added

to the overnight reaction to increase efficiency and avoid byproducts²¹⁹. After urea-PAGE (polyacrylamide gel electrophoresis) purification the RNA could be used for the ligation.

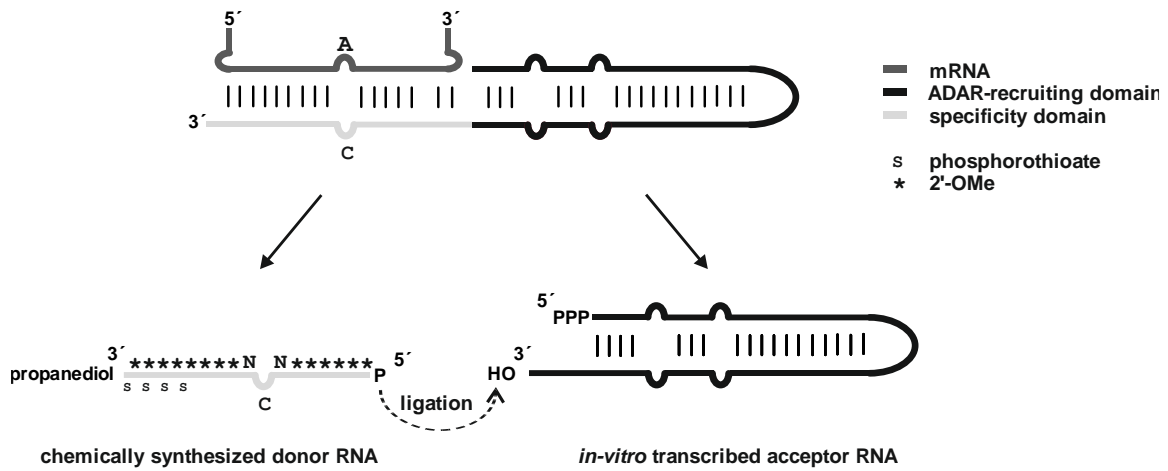


Figure 7: Manuscript 3, Figure 2: Ligation scheme of a RESTORE ASO. The chemically synthesized specificity domain is ligated by T4 RNA ligase 1 to an in-vitro-transcribed acceptor RNA to form a RESTORE ASO for RNA editing. For ligation the 5' end of the donor RNA must be phosphorylated and the 3' end of the acceptor RNA must have a free hydroxyl group.

The ligation was performed with T4 RNA ligase. This ssRNA ligase requires a 5' phosphorylated donor RNA and a free hydroxyl group at the 3' terminus of the acceptor RNA strand. Therefore, the chemically synthesized RNA was enzymatically phosphorylated with T4 polynucleotide kinase (PNK). Initially, four unmodified nucleotides at the 5' end of the modified RNA were used to ensure phosphorylation. However, 2'-OMe modified nucleotides at the 5' terminus revealed to be equally well accepted by the T4 PNK and the T4 RNA ligase. Thus, for all further phosphorylation and ligation reactions 2'-OMe modified 5' termini were used for the donor strand. To avoid byproducts like self-ligation of the donor strand, the chemically modified donor contained a modification that blocks the 3' terminus, usually a propanediol or C6-aminolinker. After ligation, the oligos were purified on a urea PAGE to separate it from unligated RNA (**Figure 3**, manuscript 3) and finally eluted from the gel.

3.2 RNA editing with ectopic expression of ADAR

In a first screen an 18 nt specificity domain with four 3' terminal PS linkages and fully 2'-OMe modifications except for three unmodified nucleotides in the central triplet was used. The central triplet is the nucleotide opposite the target A and the two neighboring

Results and discussion

nucleotides. Except for the C opposite of the target A (hereafter called central C) that revealed to be optimal for ADAR^{141, 158}, all other 17 nucleotides were reverse complementary to the target mRNA. As targets As in a UAG sequence context in the 3'UTR in the housekeeping genes of GAPDH and ACTB were chosen. This had the advantage that the mRNAs were expressed on an endogenous level in contrast to ectopically expressed reporters and all human cell lines regardless of their plasmid transfection efficiency could be analyzed. Furthermore, editing in the 3'UTR was found earlier to be more efficient compared to the ORF or the 5'UTR in the SNAP-ADAR approach (manuscript 1). For the ADAR-recruiting domain, different variations of the original R/G motif¹⁷⁵ were analyzed in combination with the respective specificity domains (**Figure 8 B**). The ASO version 4 had the same length but three A-U base pairs within the R/G-motif where auto-editing was detected earlier¹⁷⁸ were replaced by G-C pairs compared to the ASO v1 with the initial R/G motif. In ASO v9.4 the ADAR-recruiting domain of ASO v4 was prolonged by five additional base pairs. It was anticipated that the prolonged dsRNA offers more space for dsRBDs to bind and therefore improve the editing especially for ADAR1 which in contrast to ADAR2 bears three dsRBDs.

All these three versions for both ACTB and GAPDH 3'UTR targets were tested in previously established Flp-In T-REx 293 cells with inducible expression of ADAR1 p110, ADAR1 p150 or ADAR2. Most experiments in this chapter utilizing ectopic expression of ADAR were performed by the Bachelor student Sarah Merz under my co-supervision (**Figure 8 C, SI Figure 2, 3, 5, 8, 9**). After 48h of ADAR induction with doxycycline the cells were reverse transfected with 5 pmol per well of a 96 well-plate ASO using 0.75µL Lipofectamine 2000. 24h post-transfection the cells were harvested, their RNA was isolated and reverse transcribed. The resulting cDNA was amplified by Taq PCR, gel purified and subsequently sent for Sanger sequencing. Then the sequencing trace was analyzed for the editing yield at the respective sites. This procedure served as a standard for all the following editing experiments.

Surprisingly, all three ASO versions with only minor differences yielded very high editing in Flp-ADAR1 p150 cells with 75%-85% (**Figure 8 C**). In contrast to that, in both ADAR1 p110 and ADAR2 cells the editing yields were significantly lower. The main reason for the significantly higher editing yields for ADAR1 p150 might be the predominant cytoplasmic

localization of the latter. The timespan for the editing reaction might be just longer in the cytoplasm due to the rapid export of mRNA from the nucleus where ADAR1 p110 and ADAR2 are mainly localized. Also, it might be possible that more ASO is present in the cytoplasm than in the nucleus.

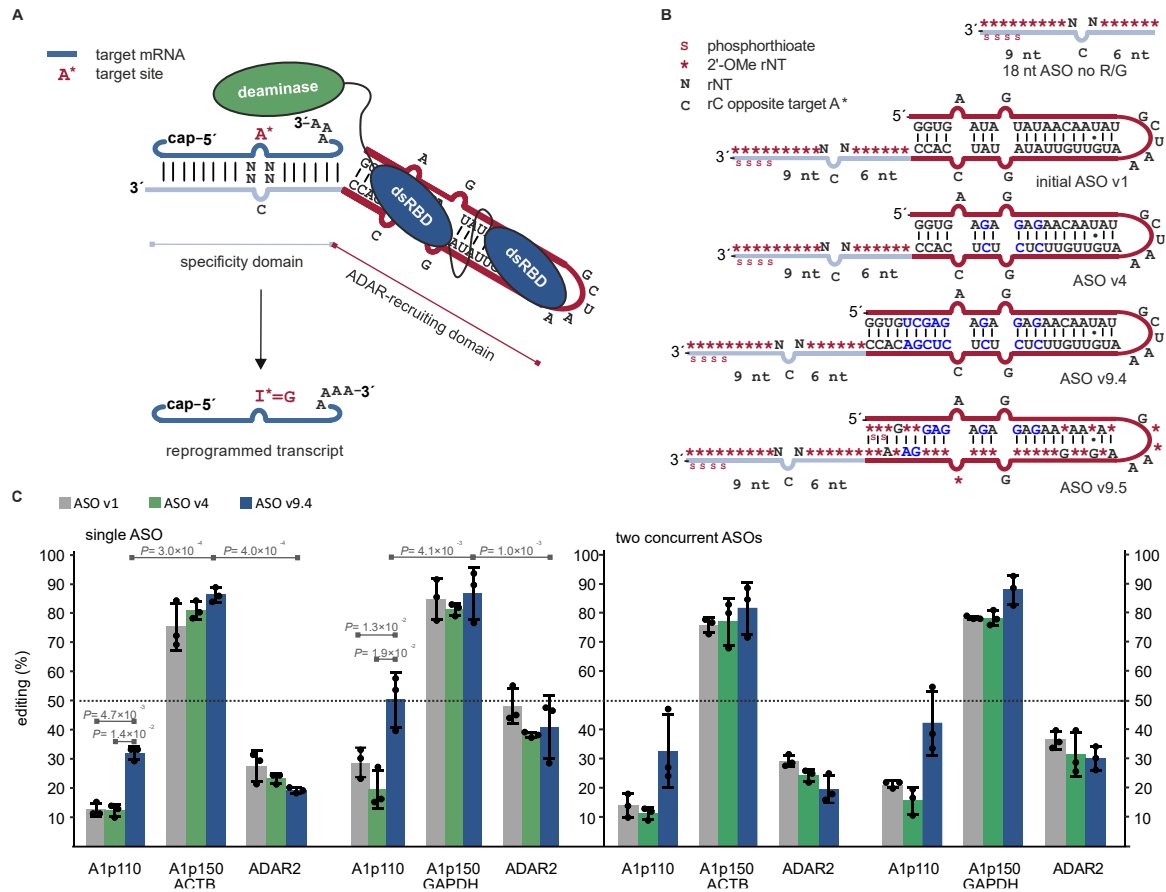


Figure 8: Manuscript 2, Figure 1: Design of ADAR-directing ASOs and characterization in engineered ADAR-expressing cell lines (293 Flp-In T-REx). **A)** Principle of RESTORE: ASOs comprise a programmable specificity domain that determines target mRNA binding and an invariant ADAR-recruiting domain to steer endogenous ADAR to the ASO:mRNA hybrid. Site-directed RNA editing at the mRNA is controlled by the chemically modified ASO and results in a specific adenosine-to-inosine change (functionally equivalent to an adenosine-to-guanosine change). dsRBD, double-stranded RNA-binding domain; A* or I*, adenosine or inosine base at target site, respectively. **B)** Sequences and chemical modifications of ASOs (see also Supplementary Table 1). rNT, natural ribonucleotide; rC, cytidine. **C)** Comparative editing of two endogenous transcripts (ACTB, GAPDH) by transfection of the respective chemically modified ASOs into the indicated ADAR-expressing cell line. Either a single ASO (against GAPDH or ACTB) or two ASOs (against GAPDH and ACTB) were transfected. Data in **C)** are shown as the mean \pm s.d., $N = 3$ independent experiments; significance (P) was calculated with a two-tailed paired t -test. A1p110 represents the ADAR1 p110 isoform, A1p150 the ADAR1 p150 isoform; n.d., no editing was detectable. All targets are given in Supplementary Note 1.

Remarkably, using the same constructs transiently expressed from a plasmid, the highest editing could be achieved with version1 in ADAR2 expressing cells (**Supplementary Figure 1**, manuscript 2). My colleague Philipp Reautschnig performed the experiment in

Results and discussion

Supplementary Figure 1, manuscript 2. Since the gRNAs are transcribed in the nucleus editing in the nucleus, where ADAR2 is mainly localized might be more efficient.

However, while ASO v1 and ASO v4 exhibited similar editing, with ASO v9.4 editing yields could be approximately doubled for ADAR1 p110 but was slightly inferior in ADAR2 expressing cells compared to ASO v1. Between the versions the trend was the same for the plasmid approach as for the ASO approach. Version 9.4 achieved the highest editing for ADAR1 p110 and the lowest editing for ADAR2, while the opposite was true for version 1. An explanation for this could be that the third dsRBD of ADAR1 can also bind to ASO 9.4 due to the longer dsRNA and therefore the editing yield is increased for ADAR1 p110 but not for ADAR2.

The overall tendencies were the same for ACTB and GAPDH, however, the editing yield for ADAR1 p110 and ADAR2 were higher for the GAPDH transcript than for the ACTB transcript. Of note, simultaneous transfection of 2.5 pmol/96 well of each ASOs for ACTB and GAPDH resulted in comparable results as transfection of only 5 pmol/96 well ASO against only one target (**Figure 8 C**). This proves that simultaneous editing of two targets is also possible. It can be envisioned that editing of even more than two different targets at the same time is possible since this was already successful for the SNAP-ADAR approach (manuscript 1).

Transfection of the 18nt chemically modified specificity alone did not result in editing for ADAR-expressing cell line (**Supplementary Figure 2**, manuscript 2). In order to investigate the effect of the chemical modifications incorporated in the specificity domain all three ASO versions were fully in vitro transcribed without any modified nucleotides. While the overall tendencies and preferences for the different ADARs remained the same, editing yields were generally lower (**Supplementary Figure 3**, manuscript 2). Especially in the ADAR1 p150 cells the editing efficiency dropped more than 20% compared to the modified versions. Therefore, the modified versions proved to be superior for editing probably due to the increased stability against nucleases in their 3' terminus. The lack of modifications in the genetically encoded gRNAs might be also an explanation for their lower editing yields compared to the chemically stabilized ASOs.

Next, the potential for alterations in coding regions was evaluated. Therefore, two UAG sites in the ORF of GAPDH were chosen that did not alter protein information upon editing. In comparison to the 3'UTR the editing yield dropped drastically for all ADARs. From 85% to approximately 50% in ADAR1 p150 cells and from 50% to less than 15% in ADAR1 p110 for both ORF sites (**Supplementary Figure 8**, manuscript 2). The drop was similar for ADAR2 expressing cells.

Due to the beneficial effect of modifications in editing further modifications were introduced into the ASO v9.4. This ASO v9.5 comprised of additional two terminal 2'-OMe nucleotides with PS linkages on the 5' end and all pyrimidine nucleotides were also 2'-OMe modified. The potency of ASO v9.5 was somewhat higher than the potency of ASO 9.4, as revealed in a dilution series from 20 pmol to 1 nmol ASO/96well for the 3'UTR UAG target in GAPDH. ASO 9.5 showed even at 1 nmol 30% editing in ADAR1 p150 cells while only 10% editing could be detected for ASO 9.4 (**Supplementary Figure 5**, manuscript 2). For ADAR1 p110 the same trend with lower editing yields was found. Therefore, even at significantly lower doses, moderate to good editing could be achieved. Furthermore, this highlights another strength of the site-directed RNA editing approach in comparison to DNA editing since at least for the ADAR1 p150 cells, editing is tunable from 10%-90% depending only on the dose of ASO.

3.3 Harnessing endogenous ADAR for 3'UTR editing – RESTORE v1

Although very high editing yields could be obtained with the newly constructed ASOs outperforming the plasmid-borne approach and offering a competitive alternative to other site-directed RNA editing approaches, engineered hyperactive deaminase fusions in the SNAP-ADAR or λ N system outcompete this strategy with higher editing yields. However, the unique advantage of these ASOs is their ability to recognize wild-type ADAR. Consequently, recruiting the endogenous ADARs instead of overexpressed ADAR would greatly simplify the approach. And in contrast to all other approaches ectopic expression of a protein would not be required any more.

Thus, encouraged by the high editing yields for ADAR1 p150 cells, the different ASO versions were tested in cells without ectopically expressed ADARs. Again, the 3'UTR UAG

Results and discussion

GAPDH and ACTB sites were targeted. For this, 50,000 HeLa cells were directly reverse transfected in a 96-well format with 5 pmol ASO/well and 0.5 μ L Lipofectamine 2000/well. While ASO v1 and v4 resulted in low but clearly detectable editing, v9.4 was superior and yielded approximately 40% editing (**Figure 9 A**).

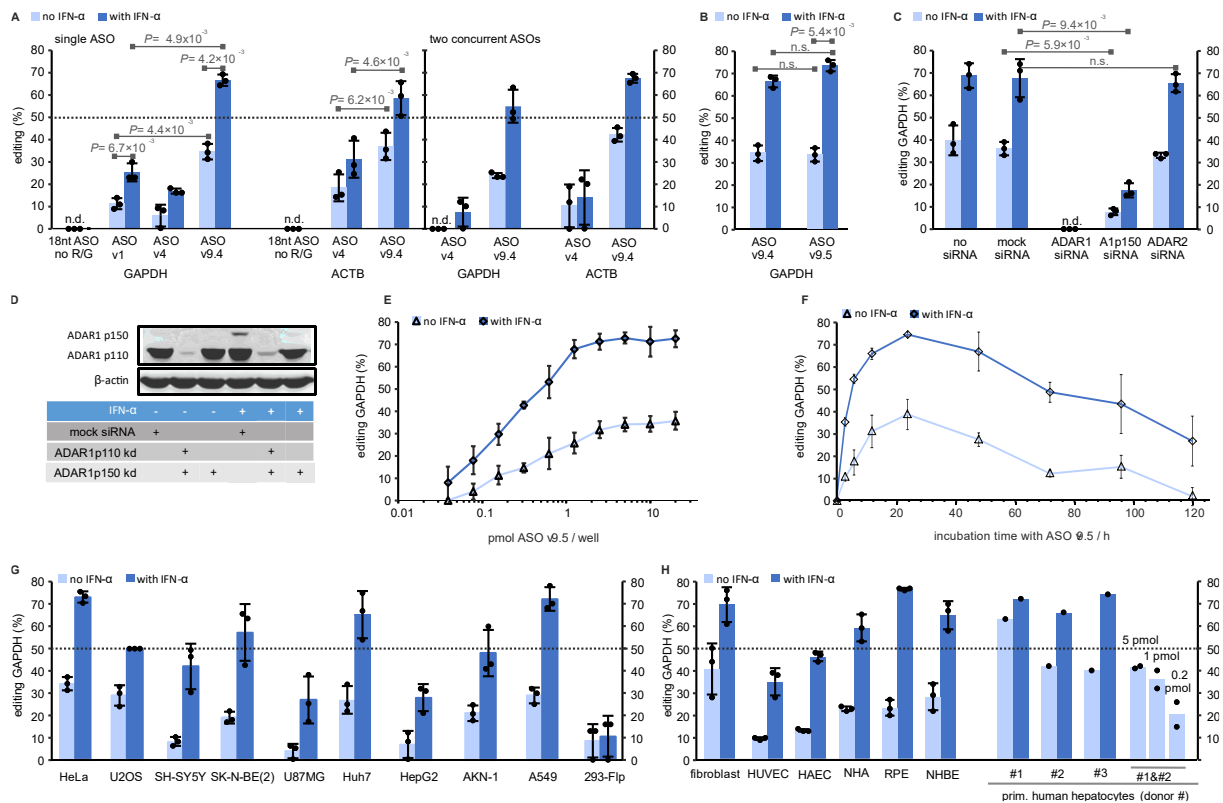


Figure 9: Manuscript 2, Figure 2: Applying RESTORE to edit endogenous transcripts (GAPDH and ACTB, each with a targeted 5' UAG triplet in the 3' UTR) in various cell lines by transfection with ASOs, performed in presence or absence of IFN- α , as indicated. A) Comparing ASO designs for the recruitment of endogenous ADAR in HeLa cells. Either a single ASO (against GAPDH or ACTB) or both ASOs (against GAPDH and ACTB) were transfected. "no R/G" indicates an ASO lacking the ADAR-recruiting domain. B) Comparative editing of ASO v9.4 and v9.5 on GAPDH. C) Effect of isoform-specific ADAR knockdown on the GAPDH editing yield in HeLa cells. D) the knockdown efficiency was verified by western blot in technical duplicate. The western blot is composed of two images with different exposure times. The full blots are given in Supplementary Fig. 4. e, Determination of the half-maximal effective dose (ED50) of ASO v9.5 for editing GAPDH in HeLa cells. F) Time course of GAPDH editing yields in HeLa cells. G) GAPDH editing yields with ASO v9.5 in various standard (cancer) cell lines. H) GAPDH editing yields with ASO v9.5 in various primary human cells. HUVEC, human umbilical vein endothelial cells; HAEC, human aortic endothelial cells; NHA, normal human astrocytes; RPE, human retinal pigment epithelium; NHBE, normal human bronchial epithelium. Data in A-H) are shown as the mean \pm s.d., N = 3 independent experiments; experiments in hepatocytes are single determinations for each donor (donors 1–3) as indicated. Significance (P) was calculated with a two-tailed paired t-test; n.s., P > 0.05; A1p150, ADAR1 p150; n.d., no editing was detectable.

Due to the very efficient editing of ADAR1 p150, HeLa cells were treated with Interferon (IFN)- α to stimulate induction of endogenous ADAR1 p150. In fact, editing yields for all versions and targets increased, and nearly doubled after IFN treatment. Editing yields of up to 70% with v9.4 were possible. Again, co-transfection of 2.5 pmol ACTB and GAPDH

ASOs resulted in similar editing levels for both targets as transfection of 5 pmol of ASO for a single target (**Figure 9 A**). Comparison of the densely modified v9.5 revealed similar editing levels as for v9.4 (**Figure 9 B**) This indicates that not only the ADAR deaminase domain but also the dsRBDs tolerate chemical modifications in the ASO without loss of editing efficiency.

Of note, neither the 18nt specificity domain nor the ADAR-recruiting domain or any non-targeting ASO resulted in detectable editing in Flp-ADAR1 p150 or HeLa cells (**Figure 9 A**, **Supplementary Figure 2, 6, 7**, manuscript 2) proving clearly the necessity of all components.

3.3.1 Which endogenous ADAR is recruited?

The high editing levels in the ADAR1 p150 overexpressing cells and the IFN induction were strong indicators for the recruitment of endogenous ADAR1 p150 in HeLa cells. In order to verify this hypothesis, western blot analysis and siRNA knock-down of the different ADARs in HeLa cells was performed. My colleague Philipp Reautschnig performed the knock-down and western blot analysis (**Figure 9 D**). While ADAR1 p110 gave a strong signal in western blot analysis, the signal for the longer isoform ADAR1 p150 was only faintly visible but increased clearly after IFN induction (**Figure 9 D**). ADAR2 was not detectable at all. For siRNA knock-down experiments, HeLa cells were transfected with siRNA 48 hours before ASO v9.5 transfection. Neither the mock siRNA nor the ADAR2 siRNA transfected cells exhibited any difference in editing yields compared to the no siRNA control (**Figure 9 C**). However, the treatment with ADAR1 siRNA and thus simultaneous knock-down of both ADAR1 isoforms completely abolished editing. Knock-down of only ADAR1 p150 decreased editing yields significantly to less than 10% without IFN treatment. This clearly confirms that ADAR1 p150 was responsible for the majority of site-directed RNA editing with ASO 9.5. Even though ADAR1 p110 was expressed in much higher quantity than ADAR1 p150 it seemed to play only a minor role for editing.

3.3.2 Potency and duration of RESTORE v1 ASO treatment

Similarly, as for the ADAR expressing cells, also a dilution series of ASO v9.5 was applied to HeLa cells to determine the potency. A half maximum editing yield of 0.4 pmol ASO per well of a 96-well plate was achieved and with additional IFN- α treatment 0.2 pmol ASO

Results and discussion

per well was reached (**Figure 9 E**). Again, a clear dose-dependency of the editing yield was observed and opens the possibility of tuning exactly the desired editing yield by adjusting the ASO dose.

So far all editing experiments were stopped and analyzed after 24 hours. In order to investigate the development of editing yield over time at different time points after ASO v9.5 transfection (5 pmol/96 well) cells were harvested and analyzed. In the first hours after transfection the editing yield increased rapidly until reaching a maximum at 24 hours (**Figure 9 F**). After 24 hours editing yields slowly decreased until five days after transfection hardly any editing was detectable. However, when continuously treated with IFN- α even after five days 27% editing was detected on the GAPDH 3'UTR UAG target.

The gradual decrease of editing yield over time can be explained by the dilution of ASO within the cells due to the rapid growth of HeLa cells. Another factor may be the intracellular degradation of ASOs. Nevertheless, editing activity could be detected over several days after application which might be sufficient for disease treatment. Considering that a potential dilution effect will not take place in non-dividing cells and that other modification patterns might increase nuclease stability of the ASOs, a prolonged activity *in vivo* can be anticipated. Besides, for some therapeutic applications like cancer treatment it might be of great benefit that editing is completely reversible after some days.

3.3.3 Cell scope of RESTORE v1

Next, the transferability of the approach to other cells of different origin was addressed. Hence, a panel of standard human cell lines was tested with ASO v9.5. Most of the cell lines (except for HeLa, Huh7 and 293-Flp) were screened for RNA editing by the intern Andreas Blaha under my co-supervision (**Figure 9 G**). Notably, in all cell lines editing could be detected, however, editing yields varied largely (**Figure 9 G**). Editing in empty HEK293-Flp-In T-REx cells was at the detection limit which was exploited in the previous chapter for selective overexpression of different ADAR version. Similar to HeLa cells, INF- α treatment considerably improved editing yields in all cells. Whereas some cells like lung cancer derived A549, liver tumor derived Huh7 and osteosarcoma derived U2OS yielded editing in a similar range as HeLa, some other cell lines like HepG2, U87MG and SH-SY5Y

required IFN- α treatment to exhibit more than 10% editing. However, transfection conditions were optimized for every cell line and it is possible that low editing yields rather correlate with poor transfection efficiencies than with low ADAR levels.

To further assess the cell scope, I examined some commercially available human primary cells. Since some of them are more sensitive than commonly used cell lines, the transfection protocol was changed. Forward transfection with 100 000 cell/well in a 24-well plate seeded 24 hours in advance was found to be a good condition. Furthermore, instead of Lipofectamine 2000, 1.5 μ L of the less toxic Lipofectamine RNAiMAX was used. Again, a two-to threefold increase in editing yield was detected after IFN- α treatment. The lowest editing yields were found in the human aortic endothelial cells (HAEC) with 13% and the human umbilical vein endothelial cells (HUVEC) with 10% (**Figure 9 H**). Editing yields between 23-28% for normal human astrocytes (NHA), retina pigment epithelium (RPE) cells and normal human bronchial epithelium (NHBE) cells were achieved. And with up to 77% after INF- α treatment editing levels in primary cells were comparable to the cell lines with best editing yields. Notably, the human fibroblasts and all donors of the primary human hepatocytes (PHH) reached higher editing yields than HeLa cells. In summary, this approach for recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing or short RESTORE is applicable in a wide range of cell lines and primary cells. The high editing yields in primary cells especially in primary hepatocytes seem to be very promising, since delivery of oligonucleotide therapeutics to the liver is well established⁵. Additionally, liver would be the target organ of diseases like AATD¹⁹⁵.

3.4 Harnessing endogenous ADAR for ORF editing – RESTORE v1

3.4.1 Optimizing design and modification pattern

However, all the editing experiments harnessing endogenous ADAR so far were performed in a UAG codon in the 3'UTR of endogenous transcripts. Although this was the first time that endogenous human ADAR was recruited and good editing yields could be achieved, all attempts to accomplish editing in the ORF with the current ASO designs failed. Therefore, further versions were screened under the assumption that RNA editing

Results and discussion

was in competition with translation. Consequently, a higher affinity to the target site might increase editing. To easily screen a number of versions, all versions were completely in vitro transcribed as described in manuscript 3 (**Table 1**). To obtain higher editing yields, initially the 3'UTR UAG GAPDH target in ADAR1 p150 overexpressing cells was used. V10 to V18 yielded at best similar editing yields as the in vitro transcribed version 9.4. While the in vitro transcribed version 9.4 yielded 74% editing in the 3'UTR of GAPDH, only 16% was achieved in the ORF of ADAR1 p150 overexpressing cells. Interestingly, version 17, a 43 nt specificity domain without ADAR-recruiting domain yielded 18% editing.

Systematic extension of the specificity domain at the 5' end of v9.4 (v20- v24) could indeed increase editing yields up to 32% with version 24, bearing a 12 nt longer specificity domain (**Figure 10 B**). However, the increase was not linear. With up to 6 nt extension (v21), the editing yields were even lower compared to the original v9.4. Systematic extension of the 3' end of the in vitro transcribed version 9.4 up to 22 nt (v26-v32) did not result in any editing.

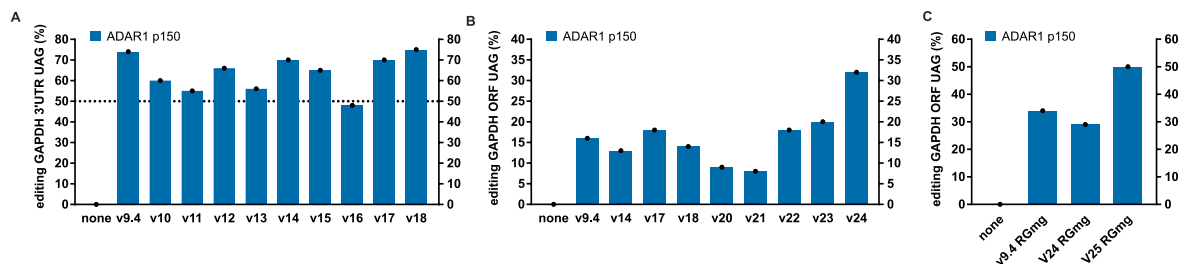


Figure 10: Screening for ASO versions that can edit ORF targets. A) different in vitro transcribed ASO versions targeting a GAPDH 3'UTR UAG codon were screened in ADAR1 p150 expressing cells. **B)** In the same cells different versions based on the results of A) were screened with a UAG GAPDH ORF target. **C)** The best versions were ligated to an 18 nt modified 3' end and analyzed as in B). Sequences and modifications can be found in Table 1. The experimental procedure was the same as described in manuscript 2 for the respective cells. Data in A)-C) reflect N=1 experiment.

Table 1: ASO sequences used in Figure 10. The ADAR-recruiting domain is depicted in green letters, constant part of the specificity domain in red letters, the variable part in blue letters and mismatched bases in the specificity domain are depicted in black. The central C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, *=Phosphorothioate linkage.

| GAPDH 3'UTR UAG ASO sequences (5' to 3') : | |
|--|---|
| v9.4 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} AGGGGUC C ACAUGGCAAC) |
| v10 | (GGGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCC} UUCA AGGGGUC C ACAUGGCAAC) |
| v11 | (GGAGAACAUAUGC ^{UAAUGUUGUUCUG} CCUCUUA AGGGGUC C ACAUGGCAAC) |
| v12 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} AGGGGUC C ACAUGGCAAC UGUGAGG) |
| v13 | (GGGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCCU} CCCUCUUA AGGGGUC C ACAUGGCAAC) |
| v14 | (GGAGAACAUAUGC ^{UAAUGUUGUUCUG} CCUCCCCUUA AGGGGUC C ACAUGGCAAC) |
| v15 | (GGCUAGGCCCUCCUCUUA AGGGGUC C ACAUGGCAAC) |
| v16 | (GGCUAGGCCCUCCUGUUA AGGGGUC C ACAUGGCAAC) |
| v17 | (GGCUCCUAGGCCCUCCUCUUA AGGGGUC C ACAUGGCAAC) |
| v18 | (GGCUCCUAGGCGCCUGCCUCUUA AGGGGUC C ACAUGGCAAC) |
| GAPDH ORF1 UAG ASO sequences (5' to 3') : | |
| v9.4 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} GGGGUC C AAGCAGUUGG) |
| v14 | (GGAGAACAUAUGC ^{UAAUGUUGUUCUG} AUGACCUUGGCCA GGGGUC C AAGCAGUUGG) |
| v17 | (GG ^{AGUUGUCAUGGAUGACCUUGGCCA} GGGGUC C AAGCAGUUGG) |
| v18 | (GG ^{AGUUGUCAUGGGUGAGCUGGCCA} GGGGUC C AAGCAGUUGG) |
| v20 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} GCCA GGGGUC C AAGCAGUUGG) |
| v21 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} UGGCCA GGGGUC C AAGCAGUUGG) |
| v22 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} CUUGGCCA GGGGUC C AAGCAGUUGG) |
| v23 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} ACCUUGGCCA GGGGUC C AAGCAGUUGG) |
| v24 | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} UGACCUUGGCCA GGGGUC C AAGCAGUUGG) |
| v9.4 RGgm | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC}) [GGG GUG](C C A)[AGCAG*U*U*G*G] Propanediol |
| v24 RGgm | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} UGACCUUGGCCA) [GGGGUG](C C A)[AGCAG*U*U*G*G] Propanediol |
| v25 RGgm | (GGUGUCGAGAAGAGGAGAACAUAUGC ^{UAAUGUUGUUCUCGUCUCCUCGACACC} UUGUCAUGGAUGACCUUGGCCA) [GGGGUG](C C A)[AGCAG*U*U*G*G] Propanediol |

Next, version 9.4 with a modified specificity domain was compared to v24 (+12 nt 5' extension) and v25(+22 nt 5' extension) and both contained the same modification pattern at the 18 nt at the 3'termini of the ASOs. ASO v25 performed with 50% editing yield in ADAR1 p150 overexpressing cells the best (Figure 10 C). However, in HeLa cells

Results and discussion

even after IFN- α treatment none of the versions could achieve editing. Thus, 3 LNA nucleotides were included in the 18nt modified part of the specificity domain to further increase the affinity of the ASO. Whereas neither v9.4 nor v24 yielded any editing with three LNA nucleotides, after IFN- α treatment v25 yielded 14% editing (**Supplementary Figure 10**, manuscript 2).

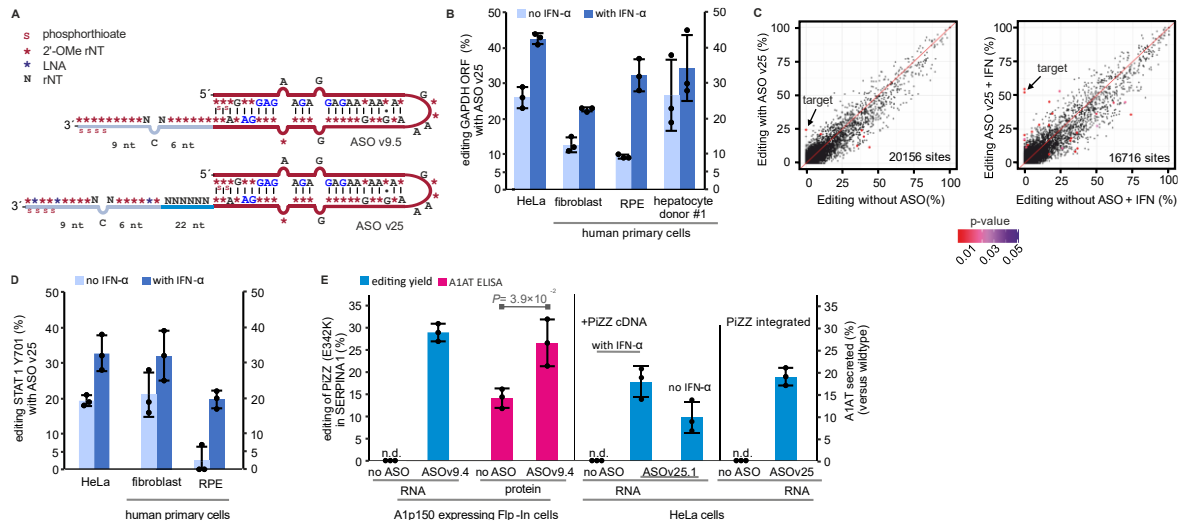


Figure 11: Manuscript 2, Figure 3: Applying RESTORE for ORF editing with ASO v25, off-target analysis, and editing of disease-relevant sites. A) ASO design v25. **B)** Editing of 5' UAG site no. 1 in the ORF of GAPDH with ASO v25 in HeLa and human primary cells. **C)** Analysis of off-target editing in the poly(A)-transcriptome when recruiting endogenous ADAR from HeLa cells to 5' UAG site no. 1 in the ORF of GAPDH with ASO v25, in absence (left) or presence (right) of IFN- α . Scatter plots show differential editing at $\sim 18,000$ sites per experiment comparing editing levels in cells treated with ASO v25 compared to empty transfected cells. Experiments were done in two independent replicates. The on-target editing is indicated by an arrow. Significantly differently edited sites ($P < 0.01$, Fisher's exact test, two-sided, $N > 50$) are highlighted in red. **D)** Editing of the Tyr701 site (5' UAU codon) of STAT1 in HeLa and primary cells. **E)** Editing of the PiZZ mutation causing α 1-antitrypsin deficiency (E342K in SERPINA1, 5' CAA codon) either in ADAR1 p150-expressing 293 Flp-In T-Rex cells with v9.4 ASO or in HeLa cells with v25 ASO (3-nt gap) or v25.1 (2-nt gap). The SERPINA1 E342K cDNA was either cotransfected or genetically integrated into HeLa cells. α 1-Antitrypsin (A1AT) secretion was normalized to the secretion when transfecting wild-type SERPINA1. Data in **B**, **D**, **E** are shown as the mean \pm s.d., $N = 3$ independent experiments; significance (P) in **e** was calculated with a two-tailed paired t-test. n.d., no editing was detectable.

Of note, when longer specificity domains with LNA nucleotides were used, before the reverse transcription step the isolated RNA was heated to 95°C with a fully 2'OMe modified oligonucleotide reverse complementary to the specificity domain, referred to as sense strand. This was necessary because the ASO binding seemed to impair the reverse transcription. Further modification of the specificity domain as for v9.5 (**Figure 11 A**) improved editing yield to 26% and 43% with IFN- α treatment in the ORF1 UAG target of GAPDH in HeLa cells (**Figure 11 B**). This version was referred to as the final v25 (**Figure 11 A**). Probably the combination of higher target affinity due to the 22 nt longer specificity

domain, the LNA nucleotides and extended dsRNA serving as substrate for improved ADAR binding were the main factors for the success of v25.

With this version 13% (23% with IFN- α) editing in the ORF UAG target was obtained in primary fibroblasts, 9% (32% with IFN- α) in RPE and 27% (34%) in PHH (**Figure 11 B**). As another endogenous target the signal transducer and activator of transcription 1 (STAT1) Y701 phosphorylation site was chosen. By editing of a UAU codon a Y701C mutation can be inserted removing the important phosphorylation site²²⁰. In HeLa and primary fibroblasts around 20% (30% with IFN- α) editing was achieved with ASO v25. In RPE only 7% editing and 20% editing with IFN- α treatment could be accomplished (**Figure 11 C**).

3.4.2 Global off-target analysis

One major challenge in site-directed RNA is managing off-target editing. All current approaches using overexpressed deaminase fusion proteins, suffer from a huge number of editing sites beyond the target site, especially when hyperactive deaminase domains are employed⁸⁰. To analyze the off-target editing of this approach deep RNA sequencing was performed with ASO v25 for the GAPDH UAG ORF1 target in HeLa. The computational analysis (**Figure 11 C, Supplementary Figure 11-13**, manuscript 2) was performed by our collaborator Qin Li (Billy Li lab, Stanford). The analysis revealed three out of 20,156 sites to be significantly differently edited compared to the control without ASO treatment (**Figure 11 C**). This off-target sites were located in noncoding areas such as introns and 3'UTRs. In the IFN- α treated sample 14 significantly differently edited sites were found. Again, all 14 sites were located in noncoding regions. Of all 14 off-target sites, 11 were known editing sites²²¹. All off-target sites were ASO dependent and high sequence similarity to the GAPDH target site was found (**Supplementary Figure 11** manuscript 2). In five of the 14 off-target sites editing levels were reduced compared to the control. This reduction could be explained by ASO binding and blocking of the natural editing sites (**Supplementary Figure 12**, manuscript 2). The overall editing homeostasis and ADAR expression was not influenced by ASO treatment compared to the respective controls (**Supplementary Figure 13**, manuscript 2). However, upon IFN- α treatment ADAR expression increased as observed earlier in the western blot analysis (**Figure 9, D**). Consequently, 116 sites exhibited significantly more editing in the IFN- α treated control compared to the none-treated control (**Supplementary Figure 13 B**, manuscript 2).

Results and discussion

With 25% target editing in the absence of IFN- α and 52% editing after IFN- α treatment the results were in good accordance to the Sanger sequencing results (**Figure 11 B, C**).

In summary, only very few ASO-dependent off-target editing sites in non-coding regions were found and no significant alteration of the natural editing homeostasis could be detected. Compared to the SNAP-ADAR, in the λ N-ADAR and the Cas13-ADAR approach much less off-target editing was detected especially considering that using the hyperactive deaminase domain all approaches had more than 1000 off-target sites⁸⁰. Even when deaminases with better off-target profiles were chosen this RESTORE approach is superior concerning off-target editing. Therefore, the very low off-target editing and thus precision is another major advantage of RESTORE besides the unique advantage that no separate protein expression is necessary.

3.4.3 SERPINA editing and AAT ELISA

In order to demonstrate a therapeutic application of RESTORE, the SERPINA1 E342K mutant, also known as PiZ mutant, was edited. In contrast to the UAG and UAU codons edited before, a CAA codon had to be edited to reverse the E342K mutation in SERPINA1. This codon is more challenging since it is one of the less preferred substrates of ADAR¹⁴⁸. Therefore, initially ADAR1 p150 overexpressing cells were used. Most experiments with SERPINA editing in the ADAR1 p150 expressing cells and the ELISA analysis were performed by the intern Andreas Blaha under my supervision, after I have optimized the experimental conditions (**Figure 11 E**). 24h after doxycycline induction cells were forward transfected with plasmids bearing either the cDNA sequence of wild-type or PiZ mutant SERPINA1. Subsequently, 24 hours later the cells were reverse transfected with ASO 9.4 and after another 24 hours the cell medium was collected and analyzed with an ELISA detecting AAT and cells were harvested and RNA editing was analyzed. The detected AAT protein in the ELISA corresponded to the AAT secreted from the cells. Since the cause of AATD is that AAT with the PiZ mutation aggregates inside liver cells and is not secreted any more the amount of secreted AAT is an important measure for the severity of AATD¹⁹⁵.

In the ADAR1 p150 expressing cells 29% editing could be achieved. Simultaneously, the level of secreted AAT relative to wild-type SERPINA1 transfected cells increased from 14% to 27% with ASO v9.4 treatment (**Figure 11 E**). However, also some bystander off-target

editing at the A proximal in the CAA codon was detected (**Supplementary Figure 14 A**, manuscript 2). To harness the endogenous ADAR, the procedure was transferred to HeLa cells but ASO v25 was transfected. Unfortunately, the editing yields of 10% and 18% with IFN- α was not sufficient for significant changes in AAT secretion levels detected by ELISA. In a further experiment the cDNA of the SERPINA1 PiZ mutant was stably integrated in HeLa cells using the piggyBac system²²². By this, it was anticipated to yield more equally distributed and lower expression across cell population closer to an endogenous target than artificial overexpression. Nevertheless, with 19% the editing yield of the IFN- α treated and ASO 25 transfected condition was comparable to the detected yield for the cells after plasmid overexpression.

However, surprisingly the bystander off-target editing of the proximal A was lower in the stably integrated cells (**Supplementary Figure 14 B**, manuscript 2). With the intention to avoid this bystander editing, the nucleotide opposite of the proximal A in the 3 nt gap containing unmodified nucleotides was also modified with a 2'-OMe nucleotide. 2'-OMe modifications were found previously to suppress bystander off-target editing in the SNAP-ADAR approach¹⁶¹. Especially for codons with a proximal A, modification of the opposite nucleotide with a 2'-OMe was reported to decrease bystander off-target editing (**Figure 2 A**, manuscript 1). Indeed, the bystander off-target editing could be suppressed from 20% to less than 5% with this additional modification in ASO 25.1 with only minimal loss in the on-target editing yield (**Supplementary Figure 15**, manuscript 2).

In summary, it could be demonstrated that the RESTORE approach can be applied to a more difficult codon and bystander off-target editing can be suppressed by further chemical modifications of the ASO. Importantly, alterations on RNA level by RESTORE ASOs were transferred and could be detected on protein level. However, although the protein secretion of AAT could be increased by RNA editing, higher editing rates, and more critical, endogenous ADAR needs to be recruited to apply RESTORE in a therapeutic context.

Results and discussion

3.5 RESTORE v2

3.5.1 Optimizing length and symmetry

Despite the unique advantage of recruiting endogenous ADAR compared to all other techniques for site-directed RNA editing at the time of publication, the RESTORE approach still had three significant disadvantages preventing RESTORE from therapeutic applicability.

First of all, only additional treatment with IFN- α resulted in good editing yields in most cells and especially in the more difficult CAA codon even with IFN- α editing yields were rather low. Despite the clinical approval of IFN- α decades ago²²³, it would be desirable to reach high editing levels without IFN- α treatment. This might be achieved by harnessing the more strongly and ubiquitously expressed ADAR1 p110 isoform.

Secondly, although a lot of stabilizing modifications like 2'-OMe and PS were already included in the ASO design, especially the 2-3nt gap opposite of the target A and the extension of the specificity domain in ASO v25 were not modified. However, stability against nucleases is one key element for a successful oligonucleotide therapeutic⁹.

And thirdly, with 95 nucleotides the size of ASO 25 is rather large compared to other oligonucleotide therapeutics which makes delivery more challenging^{5,9}. Furthermore, manufacturing of such ASOs is more complex and costly.

Therefore, the ASO design was further optimized, inspired by the in vitro transcribed ASO version 17 that comprised of only a 43-nucleotide long specificity domain but yielded 18% editing in the ORF1 UAG codon of GAPDH in ADAR1 p150 overexpressing cells (**Figure 10**). It was hypothesized that instead of a 55 nt long ADAR recruiting domain, extension of the specificity domain could have similar effects for ADAR recruiting. Hence, ASOs of different lengths without ADAR-recruiting domains were designed for the GAPDH UAG ORF1 target and in vitro transcribed. At first only the end 5' of the central C was varied (**Figure 12**). The in vitro transcribed ASOs were analyzed in Flp-ADAR1 p150 cells. With 34 nt on the 5' end already 49% editing could be detected (**Figure 12 A**) which was better than what was found for the best version (v24, 32%) with ADAR recruiting domain in the earlier screen (**Figure 10**). Extension up to 45 nt on the 5' end could increase the editing yield up to 66%. However, further extension slightly decreased the editing yield. The extension was

probably helpful due to higher affinity binding of the ASOs and more importantly the longer dsRNA offered more space for the dsRBDs to bind. Therefore, more than 45 nt on the 5' end did not increase editing yields any further because ADAR does not require more than 45 nt for the dsRBDs to bind.

To investigate the influence of the length 3' of the central C in the ASO, the 5' end was kept constant at a length of 49 nt and the 3' end was varied from 8 up to 30 nt. These in vitro transcribed ASOs were analyzed the same way as the ASOs for the 5' screen. The overall trend of increased editing yields with longer 3' extensions ranging from 52% with only 8 nt (49-1-8) to 92% with 30 nt (49-1-30) was observed (**Figure 12 B**) with only two exceptions. However, since experiments were not repeated it was unclear if this was a significant effect.

In summary, extension at both ends of an ASO with just a specificity domain resulted in increased editing yields. Interestingly, it did not make a big difference if the 5' end or the 3' end was the longer (**Figure 12 C**). This led to the hypothesis that a symmetrical design where the central C was in the center of the ASO might be optimal. Consequently, ASO versions were produced by in vitro transcription that had similar or equal extensions at the 3' or 5' termini. Notably, when shortening an ASO with a 44 nt 5' and 25 nt 3' end (44-1-25) to a 34-1-25 oligo the editing yield in ADAR1- p150 expressing cells was still higher than 90%. An in total 11 nt longer oligo (49-1-11) yielded only 49% editing in comparison (**Figure 12 D**). Therefore, symmetry and length were very important. Shortening a completely symmetrical ASO from 40-1-40 stepwise to a 27-1-27 ASO only minor losses in editing were detected. However, 23 nucleotides and less on each side led to a dramatic decrease and transfection of a 15-1-15 ASO did not result in any detectable editing.

Results and discussion

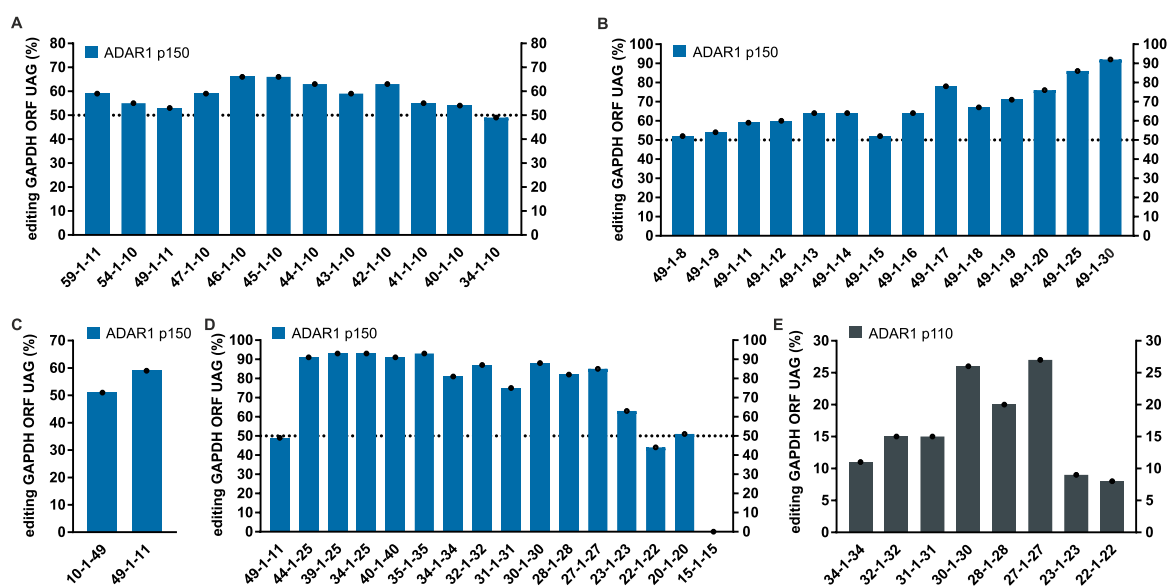


Figure 12: Screen of *in vitro* transcribed ASOs of different length and symmetry. The names reflect the numbers of nucleotides in the ASO e.g. 59-1-11 means 59 nt are 5' of the C opposite the target A and 11 nt are 3' of the C respectively. In **A**) ASO length was varied on the 5' terminus and analyzed in ADAR1 p150 expressing cells whereas in **B**) the 3' terminal length was varied with a constant 5' end. In **C**) a reversed symmetry was analyzed in ADAR1 p150 expressing cells and in **D**) the efficacy of asymmetrical and symmetrical ASOs of different length were analyzed. In **E**) symmetrical ASOs of different lengths were analyzed in ADAR1 p110 expressing cells. Sequences and modifications can be found in **Table 2**. The experimental procedure was the same as described in manuscript 2 for the respective cells with the exception that 10 pmol instead of 5 pmol ASO were transfected. Data in **A**)-**E**) reflect N=1 experiment.

Table 2: Sequences of *in vitro* transcribed ASOs used in Figure 12. The constant part of the specificity domain is depicted in red letters, the variable part in blue letters and mismatched bases in the specificity domain are depicted in black. The C opposite of the target A is highlighted in bold. (N)=RNA base

| GAPDH ORF1 UAG ASO sequences (5' to 3') | |
|---|--|
| 59-1-11 | (GG UGGUCAUGAGUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGU) |
| 54-1-10 | (GG CAUGAGUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 49-1-11 | (GG GUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGU) |
| 47-1-10 | (GG CCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 46-1-10 | (GG CUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 45-1-10 | (GG UCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 44-1-10 | (GG UCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 43-1-10 | (GG CCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 42-1-10 | (GG CACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 41-1-10 | (GG ACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 40-1-10 | (GG CGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 34-1-10 | (GG CAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 49-1-8 | (GG GUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUU) |
| 49-1-9 | (GG GUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUG) |
| 49-1-11 | (GG GUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 49-1-12 | (GG GUCCUCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGU) |

Results and discussion

| | |
|---------|---|
| 49-1-13 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUG) |
| 49-1-14 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGG) |
| 49-1-15 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGU) |
| 49-1-15 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUG) |
| 49-1-16 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGC) |
| 49-1-17 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCA) |
| 49-1-18 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAG) |
| 49-1-19 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGG) |
| 49-1-20 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGA) |
| 49-1-25 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAU) |
| 49-1-30 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUG) |
| 10-1-49 | (GG CCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUGAUGAUCUUGAGGCGUUGU) |
| 49-1-11 | (GG GUCCUUCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGG) |
| 44-1-25 | (GG UCCACGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAU) |
| 39-1-25 | (GG GAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAU) |
| 34-1-25 | (GG AUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAU) |
| 40-1-40 | (GG CGAUACCAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUGAUGAUCUUGA) |
| 35-1-35 | (GG CAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUGAUGAU) |
| 34-1-34 | (GG CAAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUGAUGA) |
| 32-1-32 | (GG AAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUGAU) |
| 31-1-31 | (GG AGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUGA) |
| 30-1-30 | (GG GUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGCUG) |
| 28-1-28 | (GG UGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUGC) |
| 27-1-27 | (GG GUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUUG) |
| 23-1-23 | (GG UGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGC) |
| 22-1-22 | (GG GGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGG) |
| 20-1-20 | (GG AUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGA) |
| 15-1-15 | (GG CUUGGCCAGGGGUGC C AAGCAGUUGGUGGUG) |

Some of the shorter symmetrical designs were also transfected in ADAR1 p110 expressing cells (**Figure 12 E**). The editing yields were much lower in comparison to the ADAR1 p150 expressing cells. With 20%-27% editing symmetrical versions with 27-30 nucleotides on each side of the central C performed best.

Results and discussion

3.5.2 Optimizing chemical modifications

Next, the influence of chemical modifications in a symmetrical ASO was analyzed. A 59 (29-1-29, referred to as v117) nucleotides ASO was designed because this was in the range of the shortest oligos with highest editing yields in ADAR1 p150 and in ADAR1 p110 expressing cells (**Figure 12 D**). Additionally, this length allowed affordable synthesis of the ASO as a whole. This version 117 was modified with different patterns of 2'-OMe (v117.1-v117.11) or LNA (v117.12-v117.15) modifications (**Table 3**). In ADAR1 p150 expressing cells v117.1 that contained only three terminal 2'-OMe modifications on every side yielded with 89% the highest editing (**Figure 13 A**). All other patterns with 2'-OMe and LNA showed decreased editing. Generally, an increasing number of 2'-OMe modifications decreased the editing yield independently of the modification pattern. One reason might be the altered dsRNA helix properties with 2'-OMe modifications that impede dsRBD binding. LNA modifications unexpectedly also decreased the editing yields. Probably the enhanced binding affinity was not beneficial due to the already high affinity resulting from the long ASO. It might be possible that elevated binding affinity negatively correlated with editing.

However, except for the 2'-OMe end-blocked v117.1 all other versions were inferior to the v25.1 in ADAR1 p150 expressing cells. In ADAR1 p110 expressing Flp cells v25.1 was still clearly superior even compared to v117.1 (**Figure 13 B**). In HeLa cells editing was either not detectable or at the detection limit in the GAPDH UAG ORF1 target.

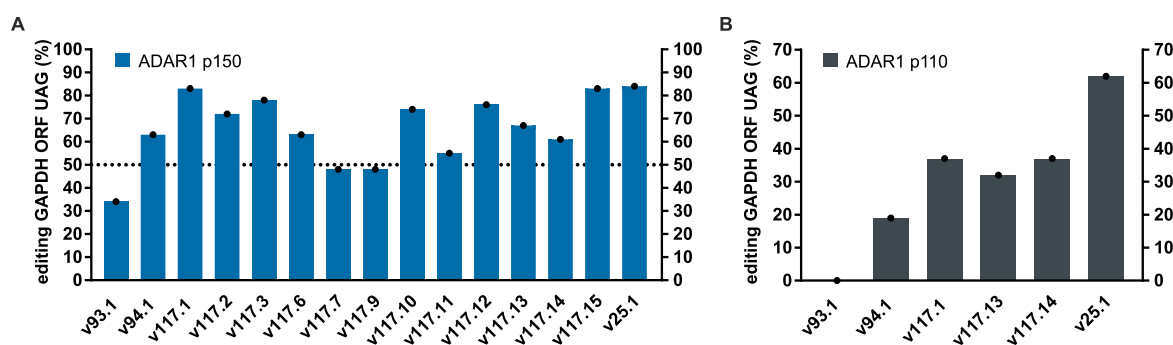


Figure 13: Screen of single-stranded ASOs with different patterns of 2'-OMe and LNA modifications. In **A**) ASOs were analyzed in ADAR1 p150 expressing cells whereas in **B**) selected ASOs were analyzed in ADAR1 p110 expressing cells. Sequences and modifications can be found in **Table 3**. The experimental procedure was the same as described in manuscript 2 for the respective cells. Data in **A**) and **B**) reflect N=1 experiment.

Table 3: Sequences of single-stranded ASOs with different patterns of 2'-OMe and LNA modifications used in Figure 13 . The C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, {N}=LNA base.

| GAPDH ORF1 UAG ASO sequences (5' to 3') : | |
|---|--|
| v93.1 | [AUG](ACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCA)[GGA] |
| v94.1 | [CAU](GGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGG)[CAU] |
| v117.1 | [UUG](UCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCAUU)[GCU] |
| v117.3 | [UUG](UCA)[UGGAU](GACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAG)[GAGGC](AUU)[GCU] |
| v117.6 | [UUG](UCA)[UGGAU](GACCUUGGCC)[AGGGGUG](C C A)[AGCAGUU](GGUGGUGCAG)[GAGGC](AUU)[GCU] |
| v117.7 | [UUG](UCA)[UGGAU](GA)[CCUUG](GCC)[AGGGGUG](C C A)[AGCAGUU](GGU)[GGUGC](AG)[GAGGC](AUU)[GCU] |
| v117.9 | [UUG] [UC](A) [U](GGA)[U](GA)[CCUU](G G)[CC](AGGGG)[U](GC C AAG)[C](AG)[UU](GG)[U](GG)[U](G)[C](AGGAGG)[C](A)[UU][GCU] |
| v117.10 | [UUG](UCAUGGAUGACCUUGGCCAGGGGUGC C A)[AGCAGUU](GGU)[GGUGC](AG)[GAGGC](AUU)[GCU] |
| v117.11 | [UUG](U)[C](A)[U](G)[G](A)[U](G)[A](C)[C](U)[U](G)[G](C)[C](A)G[G](U)[GC C AA)[G](C)[A](G)UG[U](G)[G](U)[G](C)[A](G)[G](A)G[C](A)U[GCU] |
| v117.12 | {T}[UG](UCAUGGAUGACCUUGGCCAG){G}(GGUGC C AAGCA){G}(UUGGUGGUGCAGGAGGCAUU)[GC]{T} |
| v117.13 | {T}[U]{G}(UCAUGGAUGACCUUGGCCAG){G}(GGUGC C AAGCA){G}(UUGGUGGUGCAGGAGGCAUU){G}{C}{T} |
| v117.14 | {T}[UG](UCAUGGAUGACCUUGGCC){A}(G){G}(GGUGC C AAGCA){G}(U){T} GGUGGUGCAGGAGGCAUU)[GC]{T} |
| v117.15 | {T}[U]{G}(UCAUGGAUGACCUUGGCC){A}(G){G}(GGUGC C AAGCA){G}(U){T}(GGUGGUGCAGGAGGCAUU){G}{C}{T} |
| v25.1 | [G*G*U](G)[UC](GAGAAGAGGAGAA)[C](AA)[U](A)[U](G)[CU](AAA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GA)[C](A)[C C](UUGUCAUGGAUGACCUUGGCCA)[G]{G}[GGUG](C C A)[AGCA]{G*}[U*U*]{G*}[G] AminoC6 |

Subsequently, the influence of PS-modifications was analyzed in the same setting. The three nucleotide long terminal end-block of version 117.1 was maintained and stepwise the number of PS-linkages was increased starting from the termini. Interestingly, all versions tested in ADAR1 p150 overexpressing cells reached more than 90% editing (**Figure 14 A**). In ADAR1 p110 expressing cells the difference between the versions was more pronounced. In contrast to the 2'-OMe modifications, the editing yield increased with increasing numbers of PS linkages (**Figure 14 B**). Five PS linkages on each terminus (V117.16) resulted already in editing of more than 60%, comparable to v25.1 and remarkably less than v117.1 without PS linkages. Further increase of PS linkage number on each terminus up to 20 (V117.18) or 25 (V117.19) even improved the editing yields (**Figure 14 B**). However, replacing all phosphate with phosphorothioate linkages decreased editing yields similar to the level of v117.16. Placing 20 PS linkages in the middle of the ASO around the C opposite of the target A (v117.21), only 50% instead of 69% editing with 10 PS linkages on each end of the ASO (v117.17) was detected (**Figure 14 B**).

Results and discussion

The beneficial effect of PS linkages might be due to enhanced protein binding that might not only support localization to the nucleus and therefore increase editing with the nuclear ADAR1 p110. But also binding of ADAR to the ASO might be promoted by the PS linkages. Additionally, the PS linkages also stabilize the RNA against nucleases. However, because of the mixture of diastereomers and a number of non-modified linkages, efficient protection is only expected against exonucleases. PS linkages in the middle of the ASO seemed to reduce the editing yields (v117.17 and v117.21). This can be explained by the many contacts of the ADAR deaminase domain with phosphates on the strand opposite of the target site (**Figure 4**). Exchanging these phosphates by PS- linkages probably results in unfavorable binding of the ADAR deaminase domain.

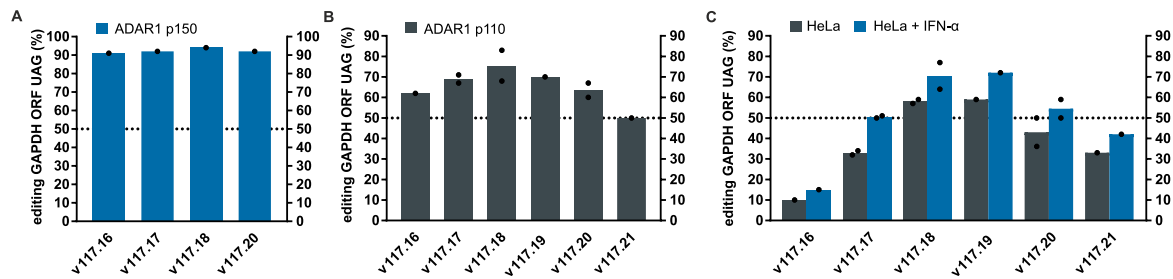


Figure 14: Manuscript 4, SI Figure1: Screen of single-stranded ASOs with different amount and patterns of phosphorothioate linkages. In **A**) ASOs were analyzed in ADAR1 p150 expressing cells in **B**) in ADAR1 p110 expressing cells. **C**) ASOs were analyzed in HeLa cells with and without IFN-α treatment. The experimental procedure was the same as described in manuscript 2 for the respective cells. Sequences and modifications can be found in **Table 4**. Data in **A**)-**C**) reflect N=1 or 2 independent experiments as indicated by the dots.

Table 4: Manuscript 4, SI Table 1: Sequences of single-stranded ASOs with different patterns of phosphorothioate linkages used in Figure 14. The C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, * = phosphorothioate linkage.

| GAPDH ORF1 UAG ASO sequences (5' to 3') : | |
|---|---|
| v117.16 | [U*U*G*](U*C*AUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGGCA*U*U*)[G*C*U] |
| v117.17 | [U*U*G*](U*C*A*U*G*G*A*U*G*G*A*UGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGG*A*G*G*C*A*U*U*)[G*C*U] |
| v117.18 | [U*U*G*](U*C*A*U*G*G*A*U*G*G*A*U*G*A*C*U*U*G*G*C*G*G*GUGC C AAGCAGUUG*G*U*G*G*U*G*C*A*G*G*A*G*G*C*A*U*U*)[G*C*U] |
| v117.19 | [U*U*G*](U*C*A*U*G*G*A*U*G*G*A*U*G*A*C*U*U*G*G*C*A*G*G*GUGC C AAGC*A*G*U*U*G*G*U*G*G*U*G*C*A*G*G*A*G*G*C*A*U*U*)[G*C*U] |
| v117.20 | [U*U*G*](U*C*A*U*G*G*A*U*G*G*A*U*G*A*C*U*U*G*G*C*A*G*G*G*U*G*C* C* A*A*G*C*A*G*U*U*G*G*U*G*G*U*G*C*A*G*G*A*G*G*C*A*U*U*)[G*C*U] |
| v117.21 | [UUG](UCAUGGAUGACCUUGGC*C*A*G*G*G*U*G*G* C* A*A*G*C*A*G*U*U*G*GUGGUGCAGGAGGCAUU)[GCU] |

Most importantly, transfecting the PS-modified versions in HeLa cells resulted in editing yields of up to 59% (v117.19) and 72% (v117.19) after IFN- α treatment (**Figure 14 C**). Generally, the trend observed in HeLa was similar to the trend in ADAR1 p110 expressing cells but more pronounced.

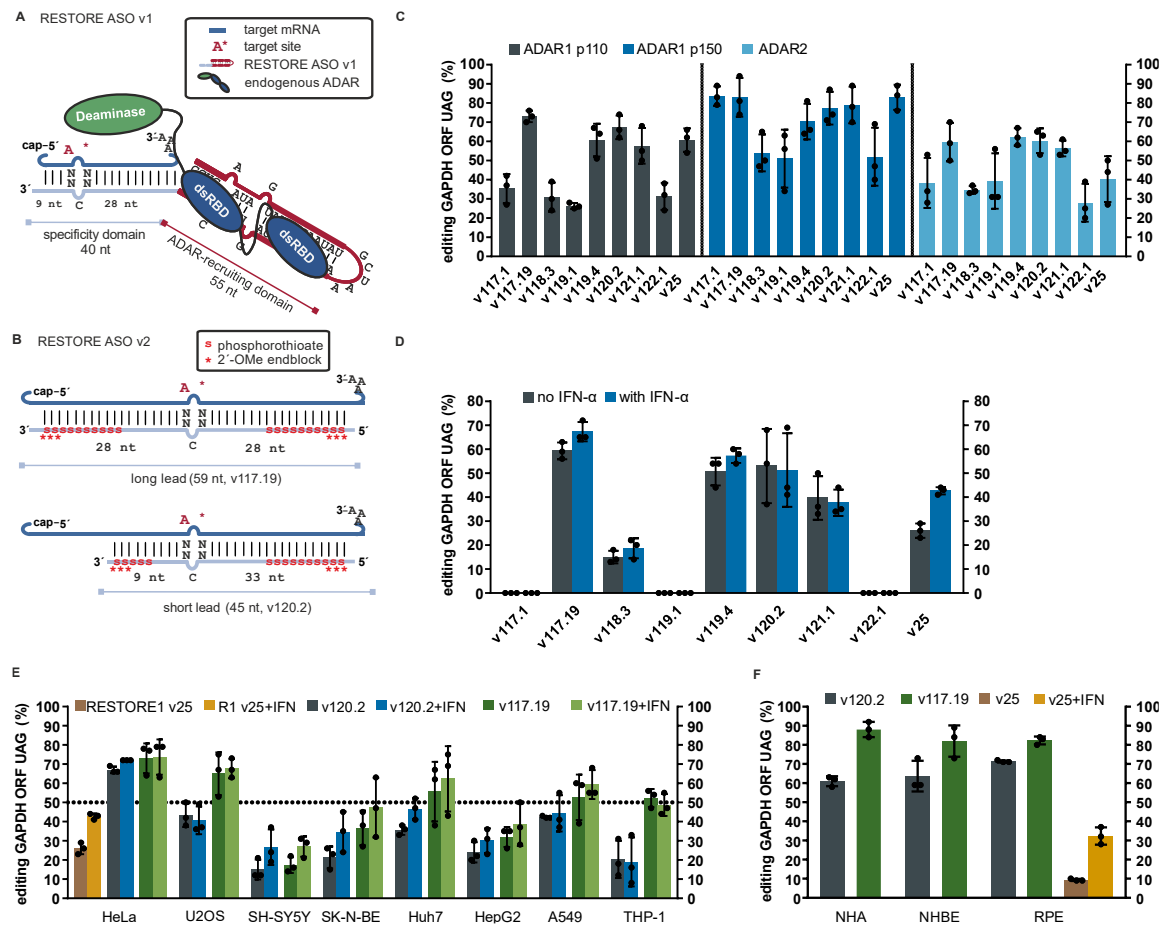


Figure 15: Manuscript 4, Figure 1: ASO screening. A) Scheme of the old RESTORE ASO v1, comprising of specificity and ADAR-recruiting domain. **B)** RESTORE v2: Schematic view of the two new lead designs, symmetric and long (v117), and asymmetric and short (v120). **C)** Effect of shorting symmetric and asymmetric ASOs for the recruitment of stably overexpressed ADARs (ADAR1p110, ADAR1p150 or ADAR2) have been overexpressed from transgenic 293 Flp-IN-T-REx cell lines, as described before). **D)** Activity of ASOs to recruit endogenous ADAR in HeLa cells, with vs. without IFN- α treatment. **E)** Cell line screen of the two lead designs (long and short). **F)** Activity of ASOs in primary human cells. The complete sequence and modification pattern can be found in **Table 5**. NHA = normal human astrocytes, NHBE = normal human bronchial epithelium, RPE = retinal pigment epithelium; Data are shown as the mean \pm s.d., N = 3 independent experiments.

decreased editing yields in ADAR overexpressing cells. For v122.1 no editing could be detected in HeLa cells. The only 5 nt longer v121.1 achieved with 36% already more editing than v25. And with 54% both the 59 nt symmetrical v117.19 and the asymmetrical 45nt v120.2 showed more than double the editing yield of the 95 nt v25 in the GAPDH UAG ORF1 target. Of note, the increase of editing after IFN- α treatment was not as pronounced as for v25 (**Figure 15 D**).

The reason for the superior editing of the symmetrical v117.19 might be that from both sides enough space for ADAR binding is offered which increases the probability of binding and therefore editing. The drastic drop of editing when this symmetrical version is shortened might be explained by insufficient space for ADAR binding on both sides of the ASO. This would be also in good accordance with the shortening of the asymmetrical version. While the 59 nt v119.4 did not bring any significant improvement over the much shorter 45 nt v120.2 even the 40 nt v121.1 yielded good editing with endogenous ADAR. This v121.1 offered still a 29 nt long RNA strand 5' to the central C which should be sufficient for two dsRBDs and the deaminase domain to bind. On the other hand, the 47 nt symmetrical v118.3 offered only 23 nt on each side of the C opposite of the target A which probably is not enough for the binding of both dsRBDs. However, the most important factor for efficient recruitment of ADAR1 p110 and ADAR2 as well as endogenous ADAR levels was the incorporation of PS linkages as discussed above. Shortening of the ASOs also resulted in a reduction of PS-linkages which might also be a factor for decreased editing in the shorter versions.

As a conclusion of all the screening data to this point, two lead ASOs were identified: The symmetrical 59 nt v117.19 and the shorter asymmetrical 45 nt v120.2 (**Figure 15 B**).

Results and discussion

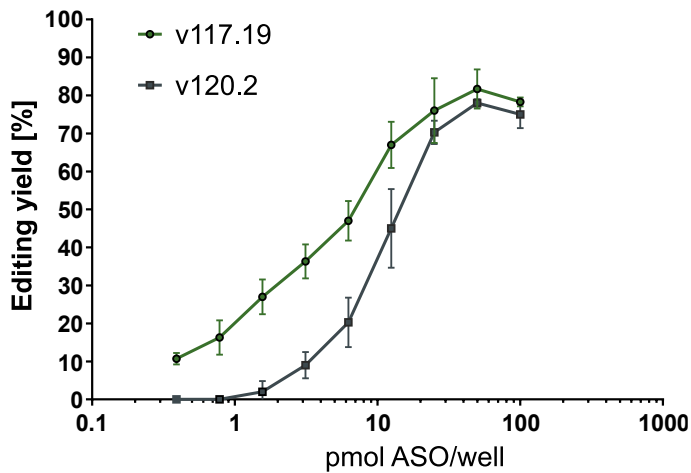


Figure 16: Potency determination of the two lead ASOs in HeLa cells. ASOs with GAPDH ORF1 UAG target were transfected with the respective amount of ASO per 24-well. The experimental procedure was the same as described in manuscript 4 for immortalized cell lines except that ASO amount was varied. Data are shown as the mean \pm s.d., $N = 3$ independent experiments.

For better reproducibility and lower toxicities due to the transfection agent, all further experiments except for the ADAR knock-down were performed in 24-well plates and, if not stated otherwise, 25 pmol ASO was transfected with 1.5 μ L Lipofecatmine RNAiMAX 24 hours after seeding. To assess the potency of the two new lead ASOs, different amounts from 100 pmol down to 390 fmol were transfected. The experiments for the potency analysis in HeLa cells in this chapter was performed by the Bachelor student Clemens Lochmann under my co-supervision (**Figure 16**). In this dilution series v117.19 had a clearly higher potency than v120.2 that did not show editing below 1.56 pmol while v117.19 showed even at 390 fmol more than 10% editing (**Figure 16**). However, both versions reached a maximum of almost 80% editing with more than 25 pmol ASO.

3.5.3 Cell scope of RESTORE v2

In a next step, as before for RESTORE v1, the cell scope of the two new ASOs was evaluated. However, this time instead of a 3'UTR UAG codon an ORF UAG codon in GAPDH was used as a target. A large part of the cell line screen (except for THP-1) was performed by the intern Laura Pfeiffer under my co-supervision (**Figure 15 E**). Notably, the editing yields of the two new lead versions were higher than the yields for the v9.5 in the 3'UTR in absence of IFN- α (**Figure 15 E** and **Figure 9 G**). With v117.19 in 5 out of the 8 analyzed cell lines more than 50% editing could be achieved. The effect of IFN- α was very low and mostly not significant at all. Except for HeLa, in all the other cell lines v120.2 was inferior to v117.19. However, in all analyzed cell lines editing could be detected with both lead

ASOs even in absence of IFN- α . Furthermore, both ASOs were analyzed in three different primary cells. As observed earlier for v9.5, editing yields in the primary cells were even higher than for the other cell lines (**Figure 15 F**). In NHA, NHBE and RPE v120.2 with editing yields between ~60% and ~70% was inferior to v117.19 that yielded more than 80% editing in the GAPDH ORF1 UAG codon in all three primary cell types. Remarkably, in absence of IFN- α less than 10% editing could be achieved with v25 in the GAPDH ORF of RPE cells. In comparison, v120.2, that is not only half the size of v25, ~70% editing was achieved. With such high editing yields the efficacy of RESTORE v2 is not only clearly higher than that of RESTORE v1 but it is also superior to all other reported approaches for site-directed RNA editing except for the SNAP-ADAR and λ N-ADAR approaches using hyperactive deaminase mutants. However, those mutants are reported to have the most severe off-target editing and need ectopic expression of an artificial protein⁸⁰.

3.3.1 Which endogenous ADAR is recruited? – RESTORE v2

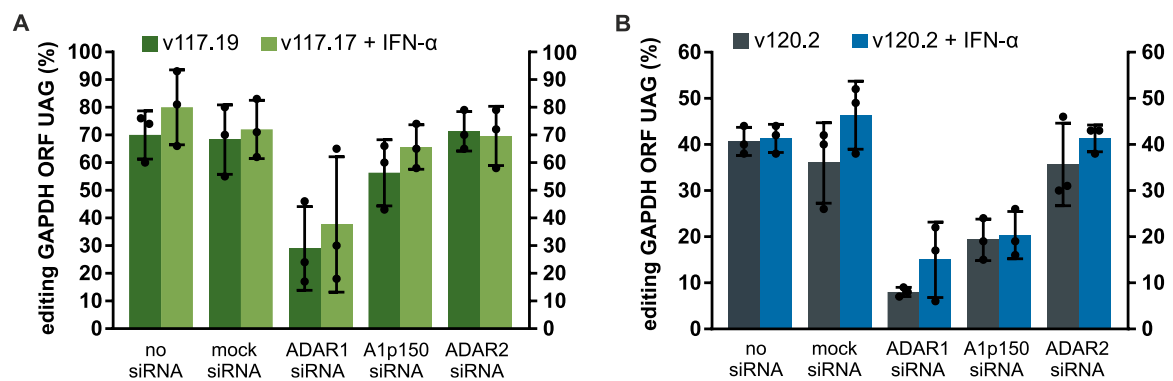


Figure 17: Effect on GAPDH ORF editing after siRNA knock-down of different ADAR isoforms in HeLa. A) editing was performed with lead v117.19 after siRNA knock-down of different ADAR isoforms or **B)** wit lead v120.2. The experimental procedure was the same as described in manuscript 2. Data are shown as the mean \pm s.d., N = 3 independent experiments.

Applying the same procedure as before, editing with the two lead ASOs was analyzed after knock-down of the different ADARs in HeLa. The editing experiments after ADAR knock-down were partly performed by the intern Laura Pfeiffer under my co-supervision (**Figure 17**). As before, the transfection of a mock siRNA or ADAR2 siRNA prior ASO transfection did not change the editing yields. Interestingly, for v117.19 ADAR1 p150 knock-down did not result in a significant decrease of editing yields (**Figure 17 A**). Only the knock-down of both ADAR1 isoforms resulted in substantial decrease but not in a complete loss of editing. This indicates that ADAR1 p110 was the enzyme primarily recruited for editing with ASO v117.19. Data from ADAR1 p110 overexpressing Flp cells showing high editing

Results and discussion

yields with this ASO support this hypothesis (**Figure 15 C**). The high standard deviation and the faint ADAR1 p110 signal in the western blot analysis after knock-down explain the residual editing even after ADAR1 knock-down.

Interestingly, v120.2 showed reduced editing after ADAR1 p150 siRNA treatment (**Figure 17 B**). Again, after ADAR1 knock-down the editing decreased clearly compared to the respective controls but was only slightly lower than after ADAR1 p150 siRNA treatment. This indicates that in contrast to v117.19, ADAR1 p150 is the predominantly recruited enzyme for v120.2. However, the low increase in editing after IFN- α induction and the high editing yield in ADAR1 p110 overexpressing cells contradicts that hypothesis. On the other hand, it is possible that the IFN- α effect was low because ASO v120.2 already stimulated ADAR1 p150 production. It might be also possible that ASO v120.2 due to its different length and PS content localizes more in the cytoplasm, whereas ASO v117.19 localizes more in the nucleus where ADAR1 p110 is more efficiently recruited. However, to clarify this finding the experiment should be repeated with higher knock-down efficiencies and further experiments investigating the immune stimulatory effects and localization of the different ASOs should be performed.

3.5.4 Optimizing chemical modification pattern for nuclease resistance

So far, the focus was on optimizing length and efficacy of ASOs for RNA editing. However, stability against nucleases is vital for successful ASO application *in vivo* and in the clinic. Analysis of the stability in 10% FBS at 37°C of the two lead ASOs v117.19 and v120.2 revealed that the ASOs were almost immediately completely degraded (**Figure 19 A**). Thus, further modifications that stabilize the ASO without loss of editing activity were necessary. The stability assays and the modification screen in HeLa cells in this chapter was performed by the Bachelor student Clemens Lochmann under my co-supervision (**Figure 18, Figure 19 A**). The design was inspired by the 2'-F and 2'-OMe modifications of siRNAs⁹ and ASOs, that recruit the ADAR2-like dsRBDs containing protein ILF3^{224, 225}. Initially, only the 5' region of v120.2, which was assumed to be bound by ADARs dsRBDs, was modified. However, a mixmer with two 2'-F modifications followed by an RNA nucleotide (v120.13, **Table 6**) decreased editing remarkably (**Figure 18**). Even worse was the performance of the modification pattern with alternating 2'-F and 2'-OMe nucleotides (v120.14), the 2'F/2'OMe mixmer v120.15) and the 2'F/DNA mixmer (120.16). Since these densely

modified ASOs performed poorly, it was tried to reach stability with minimal modifications. Therefore, only the pyrimidines were 2' modified, as this was reported to exhibit higher stability against nucleases⁵⁴. However, because the three nucleotides opposite of the target A were found to be sensitive to 2'-OMe modifications¹⁶¹, the pyrimidines in this region were incorporated as DNA nucleotides that were found to be well accepted by SNAP-ADAR (unpublished data). Indeed, v120.17, with 2'-F modification of the pyrimidines and DNA modification of the pyrimidines in the central triplet, yielded more than 60% editing. Using 2'-OMe instead of 2'F modifications (v120.18) decreased editing yield and with MOE (v120.19) no editing could be detected (**Figure 18**). Introducing additional LNAs to v120.17 did not increase editing yield (v120.20-v120.22).

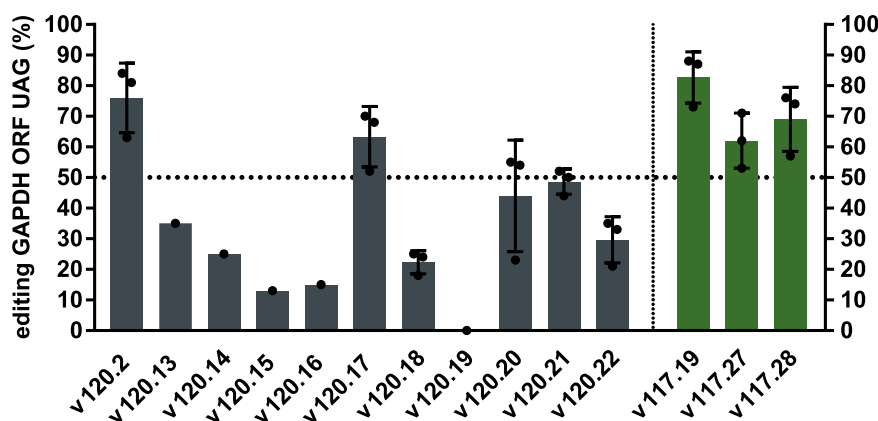


Figure 18: Manuscript 4, SI Figure 2: ASO screen for activity of stabilizing modifications in HeLa. The complete sequence and modification pattern can be found in **Table 6**. Data are shown as the mean \pm s.d, where applicable. $N=1-3$ independent experiments as indicated by data points.

Results and discussion

Table 6: Manuscript 4, SI Table 3 : Sequences of single-stranded ASOs with different lengths and modifications patterns for stabilization used in Figure 18. The C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, <N>=MOE, fN=2'-F RNA base, N=DNA base, {N}=LNA base * = phosphorothioate linkage.

| GAPDH ORF1 UAG ASO sequences (5' to 3') : | |
|---|--|
| v120.2 | [C*A*A*](A*G*U*U*G*U*C*A*U*G*G*A*U*G*A*C*U*U*G*G*C*A*G*G*G*GUGC C AAGC*A*G*U*)[U*G*G] |
| v120.13 | [C*A*A*](A*)fG*fU*(U*)fG*fU*(C*)fA*fU*(G*)fG*fA*(U*)fG*fA*(C*)fC*fU*(U*)fG*fG*(C*)fC*fA*(G*G*G*GUGC C AAGC*A*G*U*)[U*G*G] |
| v120.14 | [C*]fA*[A*]fA*[G*]fU*[U*]fG*[U*]fC*[A*]fU*[G*]fG*[A*]fU*[G*]fA*[C*]fC*[U*]fU*[G*]fG*[C*]fC*[A*](G*G*G*GU GC C AAGC*A*G*U*)[U*G*G] |
| v120.15 | [C*]fA*fA*[A*]fG*fU*[U*]fG*fU*[C*]fA*fU*[G*]fG*fA*[U*]fG*fA*[C*]fC*fU*[U*]fG*fG*[C*]fC*fA*(G*G*G*GUGC C AAGC*A*G*U*)[U*G*G] |
| v120.16 | <u>C</u> *fA*fA* <u>A</u> *fG*fU* <u>T</u> *fG*fU* <u>C</u> *fA*fU* <u>G</u> *fG*fA* <u>T</u> *fA* <u>C</u> *fC*fU* <u>T</u> *fG*fG* <u>C</u> *fC*fA*(G*G*G*GUGC C AAGC*A*G*U*)[U*G*G] |
| v120.17 | [C*A*A*](A*G*)fU*fU*(G*)fU*fC*(A*)fU*(G*G*A*)fU*(G*A*)fC*fC*fU*fU*(G*G*)fC*fC*(A*G*G*G*G)fU(G) <u>C</u> <u>C</u> (AAG)fC*(A*G*)fU*[U*G*G] |
| v120.18 | C6-disulfide [C*A*A*](A*G*)[U*U*](G*)[U*C*](A*)[U*](G*G*A*)[U*](G*A*)[C*C*U*U*](G*G*)[C*C*](AGG*G*G)[U](G) <u>C</u> <u>C</u> (AAG)[C*](A*G*)[U*U*G*G] |
| v120.19 | C6-disulfide <C*A*A*>(A*G*)<U*U*>(G*)<U*C*>(A*)<U*>(G*G*A*)<U*>(G*A*)<C*C*U*U*>(G*G*)<C*C*>(AGG*G*G)<U>(G) <u>C</u> <u>C</u> (AAG)<C*>(A*G*)<U*U*G*G> |
| v120.20 | C6-disulfide [C*A*A*](A*G*)fU*fU*(G*)fU*fC*(A*)fU*(G*G*A*)fU*(G*A*)fC*fC*fU*fU*(G*G*)fC*fC*(AGG*G*G)fU(G) <u>C</u> <u>C</u> (AAG)fC*(A*G*)fU*[U*G*G] |
| v120.21 | C6-disulfide [C*A*A*](A*G*)fU*fU*(G*)fU*fC*(A*)fU*(G*G*A*)fU*(G*A*)fC*fC*fU*fU*(G*G*)fC*fC*(AG)(G*)fU(G) <u>C</u> <u>C</u> (AAG)fC*(A*)[G*]fU*[U*]{G*}[G] |
| v120.22 | C6-disulfide {C*A*A*}(A*G*)fU*fU*(G*)fU*fC*(A*)fU*(G*G*A*)fU*(G*A*)fC*fC*fU*fU*(G*G*)fC*fC*(AGG*G*G)fU(G) <u>C</u> <u>C</u> (AAG)fC*(A*G*)fU*{T*G*G} |
| v117.19 | [U*U*G*](U*C*A*U*G*G*A*U*G*A*C*U*U*G*G*C*A*G*G*G*GUGC C AAGC*A*G*U*U*G*G*U*G*G*U*G*C*A*G*G*A*G*G*C*A*U*U*)[G*C*U] |
| v117.27 | C6-disulfide [U*U*G*U*C*](A*)[U*](G*G*A*)[U*](G*A*)[C*C*U*U*](G*G*)[C*C*](A*G*G*G*G)[U](G) <u>C</u> <u>C</u> (AAG)[C*](A*G*)[U*U*](G*G*)[U*](G*G*)[U*](G*)[C*](A*G*G*A*G*G*)[C*](A*)[U*U*G*C*U] |
| v117.28 | C6-disulfide [U*U*G*]fU*fC*(A*)fU*(G*G*A*)fU*(G*A*)fC*fC*fU*fU*(G*G*)fC*fC*(A*G*G*G*G)fU(G) <u>C</u> <u>C</u> (AAG)fC*(A*G*)fU*fU*(G*G*)fU*(G*G*)fU*(G*)fC*(A*G*G*A*G*G*)fC*(A*)fU*fU*[G*C*U] |

Analyzing the stability revealed that even after 7 days of incubation in 10% FBS a strong signal of full length ASO v120.17 could be detected (Figure 19 A). In 100% CSF after 7 days only a faint signal was detected. However, compared to v120.2 the stability was enhanced from almost immediate degradation to up to 7 days stability. The same modification

pattern with 2'-OMe (v117.27) and 2'-F (v117.28) was also applied to ASO v117.19. Interestingly, the 2'-OMe pyrimidine pattern (v117.27) performed with 62% only slightly worse than the 2'-F pyrimidine pattern (69%, v117.28). The stability analysis with the 2'-F pyrimidine and DNA modified v117.28 revealed the same stability in 10% FBS and 100% CSF as the shorter v120.17 (**Figure 19 A**). These two stabilized lead ASOs were also tested in NHA, NHBE and RPE primary cells. Except for the low editing of ~30% in RPE with the stabilized v117.28 the general tendencies in the different versions and cells was the same (**Figure 19 C**). Both stabilized versions were approximately 20% less effective than their non-stabilized counterpart.

In summary, the 2'-F modification of all pyrimidines combined with the DNA nucleotides in the central triplet of both lead ASOs increased stability with only moderate losses in editing yield. The 2'-F modification mimics closely the 2'-OH group in RNA and was therefore the most preferred modification. While 2'-OMe modifications were well tolerated by ADAR's deaminase domain, the dsRBDs seemed to be very sensitive to chemical modifications especially with 2'-OMe and MOE except for PS linkages that actually improved editing yields. Probably the bulky 2' OMe and the even bulkier 2'-MOE severely interfered with dsRBD binding.

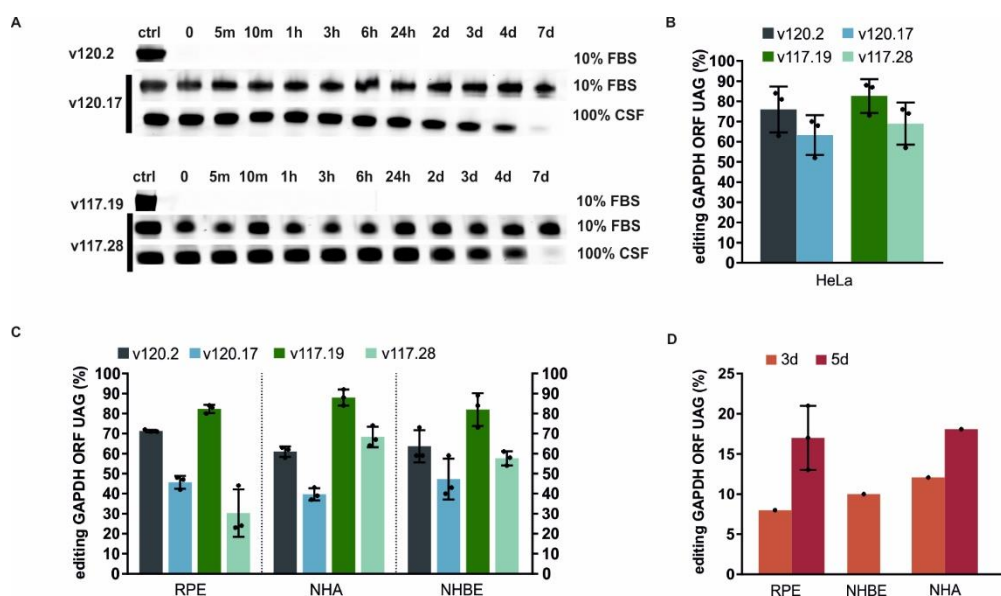


Figure 19: Manuscript 4, Figure 2. Further optimization of the lead ASOs. A) The inclusion of additional backbone modifications at all pyrimidine bases (2'-F/DNA) achieved effective stabilization of both lead ASOs in FBS and CSF. ASOs targeting the ORF of GAPDH. **B)** The stabilized ASOs are highly active in HeLa cells, and **C)** in primary cells. **D)** Chemical stabilization further enabled gymnotic uptake in primary cells with ASO v117.28. The sequences and modification patterns of all ASOs are given in **Table 6**. Data are shown as the mean \pm s.d, where applicable. N =1-3 independent experiments as indicated by data points.

Results and discussion

3.5.5 Gymnotic uptake

The now stabilized versions opened the opportunity to ASO delivery beyond transfection. Typically, RNase H-dependent ASOs are delivered by gymnotic uptake in cell culture and *in vivo* without any additional transfection agent⁵. Yet, these ASOs are usually 15-20 nucleotides long⁹, whereas our lead ASOs were 45-59 nt in length. However, a recent study demonstrated the unassisted delivery of a divalent siRNA scaffold into the CNS⁶⁷. They identified the amount of PS linkages as a crucial factor. Due to the high PS content, it was anticipated that gymnotic uptake could be an option for delivery of the RESTORE ASOs. Indeed, the incubation of three days in medium with 5 μ M ASO v117.28 yielded up to ~10% editing in human primary cells (**Figure 19 D**). Extending the incubation time to 5 days, in both RPE and NHA editing levels of more than 15% could be achieved. On the other hand, for ASO v120.17 no editing was detected. The reason for this could be either that the PS content of v120.17 is too low for gymnotic uptake or simply that the amount of intracellular ASO was not enough due to the lower potency of v120.17 compared to v117.28.

Although the editing yields were low, in principle, gymnotic uptake was feasible for v117.28. With longer incubation times editing yields might increase. Moreover, for tissue specific delivery the conjugation to a triantennary GalNAc or encapsulation in LNPs could be valuable options.

3.5.6 STAT1 Y701 editing

Next, the STAT1 Y701 UAU target was edited again to abolish the Y701 phosphorylation site. Both lead ASOs v120.2 and v117.19 were more effective in HeLa than ASO v25 (**Figure 21 A**). However, in contrast to the GAPDH UAG site, v120.2 for STAT1 was with 28% editing only slightly better than v25, that yielded 19% editing. On the other hand, ASO v117.19 reached 66% and its stabilized counterpart 59%. Consistent with previous results from GAPDH, Huh7 cells yielded slightly lower editing than HeLa. With 65% in NHA and 78% in RPE STAT1 editing with v117.19 was very efficient in primary cells (**Figure 21**). Especially editing in RPE demonstrated with 78% the superior efficacy of ASO v117.19 compared to the former lead ASO v25, where editing in absence of IFN- α was barely over the detection limit.

In the 2'-F and DNA stabilized version of v117.19 (v117.28) all three nucleotides were DNA. Of note, not only the stability against nucleases was clearly increased (**SI Figure 3**, manuscript 4), but also all bystander off-target sites were abolished by the additional modifications (**SI Figure 4**, manuscript 4). The stability assays were performed by the intern Laura Pfeiffer under my co-supervision (**SI Figure 3**, manuscript 4). This not only proves that our lead ASO design gives high editing yields in another codon context, but also that the modifications for enhanced stability are transferable to other sequences and additionally suppress bystander off-target editing.

3.5.7 MECP2 W104X editing

In a further step the applicability of RESTORE v2 ASOs for disease relevant targets was assessed. As mentioned above, RNA editing has several advantages for the treatment of Rett syndrome. As a proof of principle and because there is an existing mouse model for this modification the mMECP2 W104X UAG codon was chosen as a target. The MECP2 editing experiments in this chapter (**Figure 21 B**, **SI Figure 5,6,7**, manuscript 4) were performed by the Bachelor student Carolin Schlitz under my co-supervision. The cDNA was either directly transfected on a plasmid 24 hours prior ASO transfection or genomically integrated into HeLa cells using the piggyBac system. The editing yield with ASO v120.2 was with 31% in the plasmid overexpressing cells considerably lower than the 52% editing that was achieved in the cells with genomically integrated mMECP2 W104X (**Figure 21 B**). The stabilized version 120.17 performed with 25% in the plasmid overexpressing cell and 42% in the cells with integrated cDNA slightly worse than v120.2, similar to what was found before for the GAPDH target. Again, the influence of IFN- α treatment on editing was low. Therefore, the ability to edit the mMECP2 W104X with endogenous ADAR could be demonstrated and good yields were achieved even with the less potent v120.2 and its stabilized counterpart v120.17. With both versions also the fluorescence of the W104X-EGFP reporter could be rescued in HeLa cells (**SI Figure 6**). Sinnamon et al could show the efficacy of RNA editing with the λ N approach in primary neurons on protein level with similar editing yields as achieved with v120.2 in HeLa cells¹⁶⁷. Therefore, similar results could be expected with the RESTORE ASOs harnessing endogenous ADAR and the groundwork for *in vivo* experiments to assess the therapeutic application of RESTORE ASOs was laid.

Results and discussion

3.5.8 SERPINA1 PiZ editing

Next, the potential of the new lead ASOs for the SERPINA PiZ mutation was evaluated in HeLa cells with SERPINA1 cDNA transiently expressed from plasmids. The SERPINA editing experiments in this chapter) and the ELISA analysis were performed by the Bachelor student Carolin Schlitz under my co-supervision (Figure 21 E, F and SI Figure 9, manuscript 4).

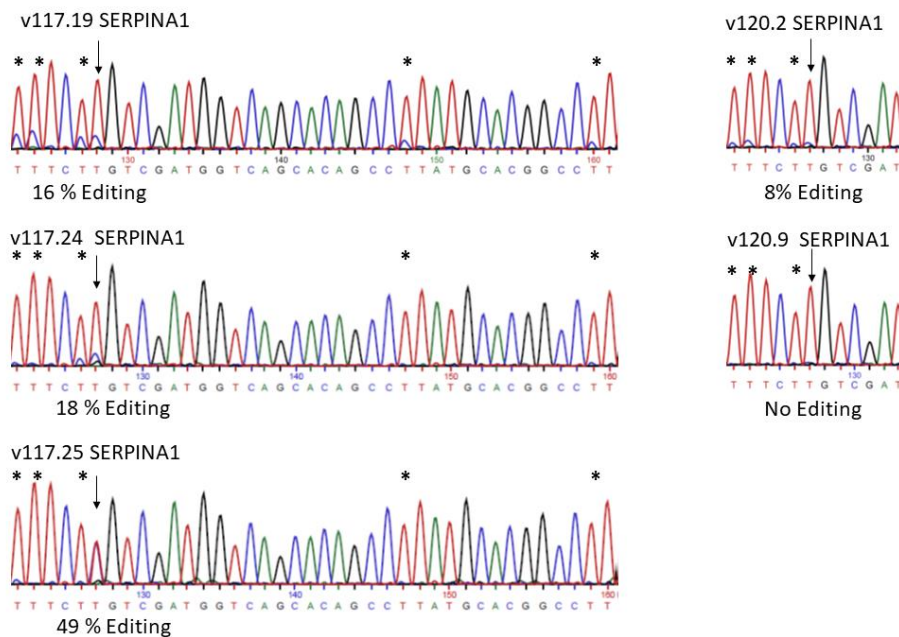


Figure 20: Manuscript 4, SI Figure 9: SERPINA editing and bystander off-target analysis. Editing experiments were performed in HeLa transiently transfected with SERPINA cDNA expressing plasmid. Different ASOs were analyzed to suppress bystander off-target editing with high on-target editing. Arrows indicate target site and asterisks indicate bystander off-target sites. The exact sequences and modifications pattern of all ASOs are given in Table 7.

With ASO 120.2 only 8% editing could be observed and in addition to the proximal off-target site in the CAA codon two other bystander off-target sites were detected. Incorporating 2'-OMe modifications opposite of the bystander off-target sites to suppress undesired bystander editing not only suppressed off-target but also on-target editing (Figure 20). Consequently, further experiments were conducted only with the longer lead ASO v117.19 and its derivatives. The editing yield of 16% was remarkably higher than with ASO v120.2, however, two additional bystander off-target sites occurred (Figure 20). Again, the nucleotides opposite of the off-target sites were 2'-OMe modified for off-target suppression. Additionally, in the proximity of the off-target site close to the target site further 2'-OMe modifications were inserted. With this v117.24 indeed bystander off-

target editing could be suppressed except for the one proximal to the target site. Simultaneously, the on-target editing did not decrease. However, the editing yield of v117.24 was still poor compared to the editing yields achieved in the GAPDH UAG codon or the STAT1 UAU codon with the lead ASO v117.19. Moreover, ASO v117.24 was most likely not resistant against nucleases. Thus, in the central triplet opposite of the target A, DNA nucleotides were inserted, and all other RNA pyrimidine nucleotides were replaced by the respective 2'-F modified nucleotides.

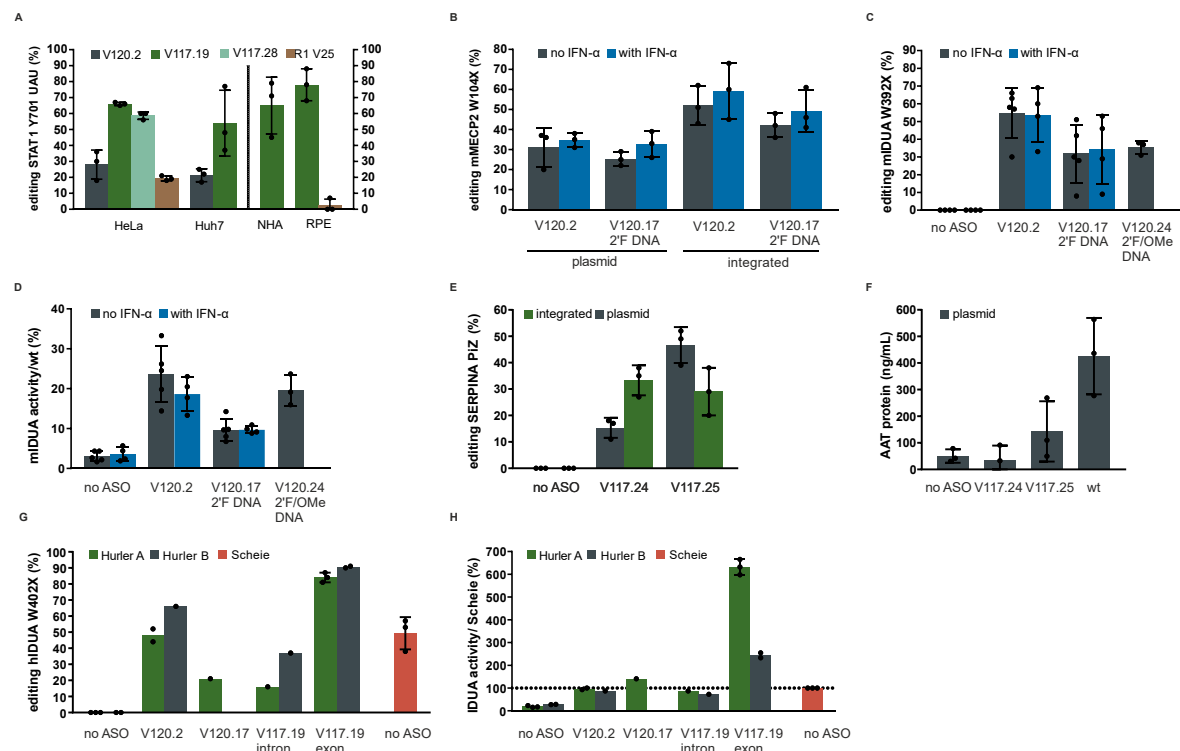


Figure 21: Manuscript 4, Figure 3: Application of RESTORE v2 ASO. A) Editing of Y701 in endogenous STAT1. B) Editing of murine MeCP2 W104X in transgenic HeLa cells (PiggyBac) or under cDNA transfection (plasmid). C) Editing of murine IDUA W392X in HeLa cells or under cDNA transfection. D) Restoration of IDUA enzyme activity after editing. E) Editing of human Serpina1 E342K in transgenic HeLa cells or under cDNA transfection. F) Restoration of α -1-antitrypsin secretion after editing. G) Editing of endogenous human IDUA W402X in two different patient fibroblasts (A, B). Long ASOs are either targeting the pre-mRNA (intron) or the mature mRNA (exon). H) Restoration of IDUA enzyme activity after editing, normalized to IDUA activity of the residual activity from a patient suffering from the more benign Scheie phenotype. The exact sequences and modifications pattern of all ASOs are given in **Table 7. Data are shown as the mean \pm s.d., where applicable. N =1-5 independent experiments as indicated by data points.**

Results and discussion

Table 7: Manuscript 4, SI Table 4: Sequences of ASOs for disease relevant targets used in Figure 21. The C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, fN=2'-F RNA base, N=DNA base, {N}=LNA base * = phosphorothioate linkage.

| STAT1 Y701C UAU ASO sequences (5' to 3'): | |
|---|--|
| v120.2 | [A*A*C*](U*U*C*A*G*A*C*A*C*A*G*A*A*U*C*A*A*C*U*C*A*G*U*C*U*U*GAUA C AUCC*A*G*U*)[U*C*C] |
| v117.19 | [C*A*G*](A*C*A*C*A*G*A*A*U*C*A*A*C*U*C*A*G*U*C*U*U*GAUA C AUCC*A*G*U*U*C*C*U*U*U*A*G*G*G*C*A*U*C*A*A*G*U*)[U*C*C] |
| v117.28 | [C*A*G*](A*)fC*(A*)fC*(A*G*A*A*)fU*fC*(A*A*)fC*fU*fC*(A*G*)fU*fC*fU*fU*(GA)fU <u>C</u> <u>C</u> <u>A</u> fUfCfC*(A*G*)fU*fU*fC*fC*fU*fU*fU*(A*G*G*G*)fC*fC*(A*)fU*fC*(A*A*G*)fU*[U*C*C] |
| v25 | [G*G*U](G)[UC](GAGAAGAGGAGAA)[C](AA)[U](A)[U](G)[CU](A AA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GACACCCA GACACAGAAAUCAACUCAGU)[C]{ T }[UGAU](A C A) [UCCA]{ G* }[U*U*]{ C* }[C] Aminolinker |
| mMECP2 W104X UAG ASO sequences (5' to 3'): | |
| v120.2 | [U*C*G*](G*C*C*A*G*A*C*U*U*C*C*U*U*U*G*U*U*U*A*G*C*U*U*U*C*G*UGUC C AACCU*U*C*)[A*G*G] |
| v120.17 | [U*C*G*](G*)fC*fC*(A*G*A*)fC*fU*fU*fC*fC*fU*fU*fU*(G*)fU*fU*fU*(A*A*G*)fC*fU*fU*fU*fC*(G*)fU(G)fU <u>C</u> <u>C</u> (AA)fCfC*fU*fU*fC*[A*G*G] |
| miDUA W392X UAG ASO sequences (5' to 3'): | |
| v120.2 | [G*U*C*](C*A*A*C*A*C*A*G*C*C*C*A*G*C*C*U*U*U*G*A*G*A*C*C*U*C*UGCC C AGAG*U*U*G*)[U*U*C] |
| v120.17 | [G*U*C*]{fC*(A*A*)fC*(A*)fC*(A*G*)fC*fC*fC*fC*(A*G*)fC*fC*fU*fU*fU*(G*A*G*A*)fC*fC*fU*fC*fU(G)fC <u>C</u> <u>C</u> (AGAG*)fU*fU*(G*)[U*U*C] |
| v120.24 | [G*U*C*C*](A*A*)fC*(A*)[C*](A*G*)fC*[C*]{fC*[C*](A*G*)fC*[C*]{fU*[U*]{fU*(G*A*G*A*)[C*]{fC*[U*]{fC*[U](G)fC <u>C</u> <u>C</u> <u>C</u> (AGAG*)[U*]{fU*(G*)[U*U*C] |
| hSERPINA1 E342K CAA ASO sequences (5' to 3'): | |
| v120.2 | [A*A*A*](A*A*C*A*U*G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*U*UUCU C GUCG*A*U*G*)[G*U*C] |
| v120.9 | [A*A*A*](A*A*C*A*U*G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*)[U*U](UC)[U] (C GUCG*A*U*G*)[G*U*C] |
| v117.19 | [C*A*U*](G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*U*UUCU C GUCG*A*U*G*G*U*C*A*G*C*A*G*C*U*U*A*U*G*C*A*)[C*G*G] |
| v117.24 | [C*A*U*](G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*)[U*UUCU] (C GUCG*A*U*G*G*U*C*A*G*C*A*G*C*)[U*U*](A*U*G*C*A*)[C*G*G] |
| v117.25 | [C*A*U*](G*G*)fC*fC*fC*fC*(A*G*)fC*(A*G*)fC*fU*fU*fC*(A*G*)fU*fC*fC*fC*fU*fUfUf <u>C</u> <u>C</u> <u>f</u> fUfC(G*A*)fU*(G*G*)fU*fC*(A*G*)fC*(A*)fC*(A*G*)fC*fC*fU*fU*(A*)fU*(G*)fC*A*)[C*G*G] |
| hiDUA W402X UAG ASO sequences (5' to 3'): | |
| v120.2 | [G*U*C*](C*A*G*G*A*G*G*U*C*C*G*G*C*U*G*C*G*A*A*C*U*U*C*GGCC C AGAG*C*U*G*)[C*U*C] |
| v120.17 | [G*U*C*]{fC*(A*G*G*A*)fC*(G*G*)fU*fC*fC*fC*(G*G*)fC*fC*fU*(G*)fC*(G*A*)fC*(A*)fC*fU*fU*fC*(GG)fC <u>C</u> <u>C</u> (AGAG*)fC*fU*(G*)[C*U*C] |
| v117.19 | [G*G*A*](C*G*G*U*C*C*G*G*C*U*U*G*C*G*A*A*C*U*U*C*GGCC C AGAG*C*U*G*C*U*C*U*C*A*U*C*U*G*C*G*G*G*C*G*G*) [G*G*G] |
| intron | |
| v117.19 | [G*G*A*](C*G*G*U*C*C*G*G*C*U*U*G*C*G*A*A*C*U*U*C*GGCC C AGAG*C*U*G*C*U*C*U*C*A*U*C*U*G*C*G*G*G*C*G*)[A*G*C] |
| exon | |

Although this modification pattern was expected to increase stability against nucleases, no improvement in editing yields was assumed. To also address this issue the dG opposite of the C in the CAA codon was replaced by a dI. Based on the structural insights from the ADAR2 deaminase domain a G-C base pair 5' of the target A hinders the enzyme by a clash of the 2-amino group of the G with residue G489 of ADAR2. Replacing the G by an I lacking this 2-amino group, editing efficiency of the ADAR2 deaminase domain could be increased¹⁴¹. Hence, it was anticipated that the inserted dI would be beneficial for editing in the CAA codon. Indeed, with 47% editing and no detected off-target editing v117.25 was clearly superior to v117.24 that achieved only 15% (**Figure 20** and **Figure 21 E**).

Subsequently, the collected medium from the ASO treated cells was analyzed for AAT content via ELISA. While no difference in protein amount was observed between the no-ASO control and ASO v117.24 treatment, for ASO v117.25 more protein could be detected in the medium (**Figure 21 F**). This indicates that more wild-type protein could be restored and was secreted after ASO v117.25 treatment. However, the standard deviation was very high and for determination of accurate numbers, the settings for the ELISA such as amount and incubation time of the medium in the antibody coated wells should be considered.

Nevertheless, a successful strategy to edit this difficult CAA codon with good editing yields and no bystander editing could be developed. Furthermore, a potential application to treat a genetic disease with RESTORE v2 ASOs was demonstrated. RESTORE seems to be especially well suited, since both the liver and lung phenotype of AATD could be restored simultaneously, a therapy option that currently is only given by liver transplantation. Furthermore, AAT is almost exclusively secreted from the liver which is a well accessible organ for oligonucleotide therapeutics.

3.5.9 IDUA editing and restoration of enzymatic activity

After the successful editing of the AATD causing PiZ mutant, another disease-related target was chosen. The UAG codon to reverse the IDUA W402X mutation, the most common mutation in the severe Hurler syndrome was targeted¹⁸². The IDUA editing and enzymatic assay experiments in this chapter (**Figure 21 C, D, G, H, Figure 22** and **SI Figure 10**) were performed by the Bachelor student Carolin Schlitz under my co-supervision. Initially the respective murine IDUA W392X was cloned. Subsequently, the different lead

Results and discussion

ASOs were assessed in HeLa cells with the transiently expressed mIDUA W392X mutant. The murine IDUA was chosen because *in vivo* experiments with a mouse model for Hurler syndrome with that mutation were anticipated if positive results could be gained from cell culture experiments²²⁶. With ASO 120.2 an editing yield of 45% was achieved in HeLa (**Figure 21 C**). The editing yield of the stabilized v120.2 was 15% lower, consistent with earlier observations. Likewise, IFN- α treatment was not beneficial for editing. In order to investigate the functional rescue of the IDUA activity, an earlier published enzymatic assay was established²²⁷. It is based on the cleavage of the 4-methylumbelliferyl- α -L-iduronide (4MU-iduronide) by the IDUA enzyme resulting in the fluorescent 4MU. Therefore, with a simple fluorescence readout the enzymatic activity could be determined using a calibration curve of non-conjugated 4MU. For better comparability, the activity was referenced to total protein concentration of cell lysates measured by Bradford or Bicinchoninic acid (BCA) assay. In lysates of HeLa cells treated with ASO v120.2 after transient transfection of mIDUA cDNA 24% enzyme activity relative to the wild-type activity was observed (**Figure 21 D**). In contrast, only 10% enzyme activity was reached with the stabilized v120.17. It was anticipated that the high 2'-F content resulted very high affinity of the ASO and hence partially blocked the protein translation. To verify this hypothesis IDUA specific ASOs and control ASOs were transfected in transiently overexpressing IDUA wild-type cells. If a translation blockage took place it should affect also the activity of cell expressing the wild-type enzyme. And indeed, while ASO v120.2 treated cells showed the same enzyme activity as cells without ASO treatment, v120.17 treated cells had slightly decreased enzyme activity of only 86% (**Figure 22**). Thus, another version with every other 2'-F modification exchanged to a 2'-OMe modification was designed and analyzed. However, this version 120.24 had only marginally higher enzyme activity on wild-type IDUA transfected HeLa cells than v120.17 (**Figure 22**). And although approximately the same editing yield as for v120.17 was observed, the restored enzyme activity was with 20% twice as high as for v120.17 (**Figure 21 D**). Therefore, despite some minor translation blockage was found for v120.17, it probably does not explain the higher enzyme activity restoration with v120.24. It can only be speculated at this point if other off-target effects due to the higher affinity of v120.17 or due to the higher 2'-F content might give an explanation to the different enzyme activity restoration.

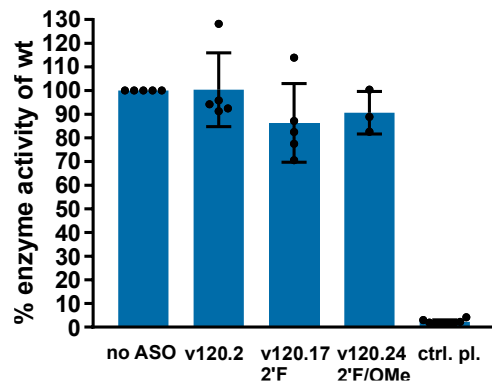


Figure 22: Manuscript 4, SI Figure 8: Effect of ASOs on translation. In HeLa cells transiently expressing *mIDUA* wt cDNA different ASOs were transfected and the enzymatic activity of the *IDUA* protein was analyzed. The exact sequences and modifications pattern of all ASOs are given in **Table 7**. Data are shown as the mean \pm s.d, $N=3-5$ independent experiments as indicated by data points.

Encouraged by the high editing yields and good restoration of enzymatic activity in HeLa, ASOs were transfected in primary fibroblasts donated from two patients (Hurler A and Hurler B) with Hurler syndrome to restore the *IDUA* W402X mutation. Primary fibroblasts from patients with the much milder Scheie syndrome served as control. ASO v120.2 yielded 39% editing in Hurler A and 66% editing in Hurler B fibroblasts, whereas with the stabilized v120.17 only 21% editing could be observed in Hurler A cells (**Figure 21 G**). Of note, due to nonsense-mediated decay (NMD) resulting from the premature stop codon, editing yields are generally overestimated in these cells. However, the restored enzymatic activities of both versions were with approximately on the same level as for the Scheie syndrome fibroblasts (**Figure 21 H**). Next, the more potent v117.19 was analyzed. Due to the close proximity of the editing site to a splice site two different version were designed. One version spanning the intron-exon junction of the pre-mRNA (referred to as v117.19 intron) and one version that binds to the spliced mRNA (referred to as v117.19 exon). Interestingly, with 84%-91% editing, the exon version performed much better than the intron version that yielded only 16% editing in Hurler A and 44% editing in Hurler B fibroblasts (**Figure 21 G**). This strongly indicates that most of the RNA editing process happens after splicing, which surprisingly contradicts what Ou et al. found in the same cells¹⁸⁰. However, their ASOs were with 111 nt much longer and only the three terminal nucleotides on each side were modified. The superior editing yields of the v117.19 exon variant could also translate to considerably enhanced enzymatic activity of more than 6-fold increase in Hurler A fibroblasts compared to the Scheie fibroblasts (**Figure 21 H**). Although using a 111 nt long ASO in their study Qu et al report less than 30% editing with

Results and discussion

a restored enzyme activity of approximately 1.5 fold higher than that of the Scheie fibroblasts¹⁸⁰.

Scheie syndrome has a much milder disease progression without symptoms of mental retardation and importantly, patients survive until adulthood. On the other hand patients with Hurler syndrome suffer from severe mental and physical disabilities and die without treatment before the age of 10¹⁸⁹. Therefore, reaching 6-fold higher enzymatic activities after treatment with RESTORE ASOs would be of enormous benefit for patients.

3.5.10 Recruiting murine ADARs

Finally, after the stability and the feasibility for different disease targets was demonstrated, the possibility to recruit endogenous murine ADARs was assessed before the therapeutic potential of RESTORE can be evaluated *in vivo*. Due to the lack of UAG codons in the ORF of murine GAPDH, that do not alter genetic information upon editing, a UAG ORF and 3'UTR target site in murine eukaryotic elongation factor 2 (mEef2) was chosen as an alternative.

Editing of these two sites was performed in MEF cells and Hepa 1.6. In both cell lines the highest editing of 57%-66% was observed in the 3'UTR with v120.2 (**Figure 23**). Editing in the ORF was lower with 30%-44% in MEF cells and v117.19 performed slightly better than v120.2 consistent with earlier observations. In Hepa 1.6 ORF editing was even lower with 14%-30% but with the same trend of better performance of v117.19. Treatment with mIFN- α did in most cases not considerably increase editing yields.

Overall, the editing yields especially in the ORF were poor relative to editing in HeLa or human primary cells. On the other hand, in other human cells lines like HepG2 also poor editing yields were observed. Probably transfection efficiency might be the reason for the poor editing in some cells like HepG2, Hepa 1.6 or MEF.

4 Conclusions

With our RESTORE v1 ASOs we could demonstrate the possibility of harnessing human endogenous ADAR in a wide variety of human cells lines and primary cells. To achieve this, not only the length and symmetry, but also the chemical modifications played a major role. Especially the latter proved to be a clear advantage compared to the plasmid-borne approach. In contrast to all other approaches for site-directed RNA editing, no ectopic expression of a protein was necessary. And due to the use of endogenous ADAR and chemically modified ASOs, the global off-target profile was superior to that of other approaches⁸⁰. The disadvantages of high IFN- α dependency, low ORF editing yields and large size of ASOs were overcome with RESTORE v2. Despite the remarkable size reduction to 62% (v117) and 47% (v120) of the length of ASO v25, choosing the right modification pattern and especially including phosphorothioate linkages could in many cases more than double the editing yields compared to RESTORE v1 ASOs. These improvements allowed the editing of pathogenic mutations in the MECP2, SERPINA and IDUA transcripts. And the functional restoration of IDUA enzymatic activity in fibroblasts derived from Hurler patients could be demonstrated. Further modification of RESTORE v2 ASOs with 2'-F RNA and DNA bases lead to remarkably higher half-lives in serum and CSF enabling gymnotic uptake. This opened, together with the successful recruitment of murine ADAR, the possibility of *in vivo* experiments and thus built the foundation for the therapeutic application of RESTORE. It can be envisioned that even higher stability and efficacy could be achieved by optimizing the modification pattern or introducing mismatches according to the structural constraints of ADAR1. Binding to the different symmetries of ASOs in the context of potential ADAR dimerization requires further experimental illumination and might result in even further advanced ASO designs. One major challenge will be the efficient delivery to target tissues and it remains elusive if gymnotic uptake will be sufficient *in vivo*. Thus, mechanisms for targeted delivery like LNPs or triantennary GalNAc conjugation could be other promising solutions. In summary, with RESTORE, we could contribute to expand the toolbox of oligonucleotide therapeutics by the class of RNA editing. This class can benefit from the advances in the oligonucleotide therapeutics field and might be applied in the future for the treatment of many currently untreatable diseases.

5 References

1. Zamecnik, P.C. & Stephenson, M.L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc Natl Acad Sci U S A* **75**, 280-284 (1978).
2. Stephenson, M.L. & Zamecnik, P.C. Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proc Natl Acad Sci U S A* **75**, 285-288 (1978).
3. Roehr, B. Fomivirsen approved for CMV retinitis. *J Int Assoc Physicians AIDS Care* **4**, 14-16 (1998).
4. Perry, C.M. & Balfour, J.A. Fomivirsen. *Drugs* **57**, 375-380; discussion 381 (1999).
5. Dowdy, S.F. Overcoming cellular barriers for RNA therapeutics. *Nature Biotechnology* **35**, 222 (2017).
6. Crooke, S.T., Witztum, J.L., Bennett, C.F. & Baker, B.F. RNA-Targeted Therapeutics. *Cell Metab* **27**, 714-739 (2018).
7. Lieberman, J. Tapping the RNA world for therapeutics. *Nature Structural & Molecular Biology* **25**, 357-364 (2018).
8. Crooke, S.T., Wang, S., Vickers, T.A., Shen, W. & Liang, X.-h. Cellular uptake and trafficking of antisense oligonucleotides. *Nature Biotechnology* **35**, 230 (2017).
9. Khvorova, A. & Watts, J.K. The chemical evolution of oligonucleotide therapies of clinical utility. *Nature Biotechnology* **35**, 238 (2017).
10. Wan, W.B. & Seth, P.P. The Medicinal Chemistry of Therapeutic Oligonucleotides. *Journal of Medicinal Chemistry* **59**, 9645-9667 (2016).
11. Yamamoto, T., Nakatani, M., Narukawa, K. & Obika, S. Antisense drug discovery and development. *Future Med Chem* **3**, 339-365 (2011).
12. Disterer, P. et al. Development of therapeutic splice-switching oligonucleotides. *Hum Gene Ther* **25**, 587-598 (2014).
13. Lindow, M. & Kauppinen, S. Discovering the first microRNA-targeted drug. *The Journal of Cell Biology* **199**, 407-412 (2012).
14. Rigo, F., Seth, P.P. & Bennett, C.F. Antisense oligonucleotide-based therapies for diseases caused by pre-mRNA processing defects. *Adv Exp Med Biol* **825**, 303-352 (2014).
15. Eckstein, F. Phosphorothioates, essential components of therapeutic oligonucleotides. *Nucleic acid therapeutics* **24**, 374-387 (2014).
16. Eckstein, F. Nucleoside Phosphorothioates. *Journal of the American Chemical Society* **88**, 4292-4294 (1966).
17. Boczkowska, M., Guga, P. & Stec, W.J. Stereodefined phosphorothioate analogues of DNA: relative thermodynamic stability of the model PS-DNA/DNA and PS-DNA/RNA complexes. *Biochemistry* **41**, 12483-12487 (2002).

References

18. Koziolkiewicz, M., Krakowiak, A., Kwinkowski, M., Boczkowska, M. & Stec, W.J. Stereodifferentiation--the effect of P chirality of oligo(nucleoside phosphorothioates) on the activity of bacterial RNase H. *Nucleic Acids Res* **23**, 5000-5005 (1995).
19. Watanabe, T.A., Geary, R.S. & Levin, A.A. Plasma protein binding of an antisense oligonucleotide targeting human ICAM-1 (ISIS 2302). *Oligonucleotides* **16**, 169-180 (2006).
20. Miller, C.M. et al. Stabilin-1 and Stabilin-2 are specific receptors for the cellular internalization of phosphorothioate-modified antisense oligonucleotides (ASOs) in the liver. *Nucleic Acids Research* **44**, 2782-2794 (2016).
21. Koller, E. et al. Mechanisms of single-stranded phosphorothioate modified antisense oligonucleotide accumulation in hepatocytes. *Nucleic Acids Research* **39**, 4795-4807 (2011).
22. Yanai, H. et al. Suppression of immune responses by nonimmunogenic oligodeoxynucleotides with high affinity for high-mobility group box proteins (HMGBs). *Proc Natl Acad Sci U S A* **108**, 11542-11547 (2011).
23. Shen, W. et al. Chemical modification of PS-ASO therapeutics reduces cellular protein-binding and improves the therapeutic index. *Nature Biotechnology* (2019).
24. Bobst, A.M., Rottman, F. & Cerutti, P.A. Effect of the methylation of the 2'-hydroxyl groups in polyadenylic acid on its structure in weakly acidic and neutral solutions and on its capability to form ordered complexes with polyuridylic acid. *J Mol Biol* **46**, 221-234 (1969).
25. Monia, B.P. et al. Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *The Journal of biological chemistry* **268**, 14514-14522 (1993).
26. Cummins, L.L. et al. Characterization of fully 2'-modified oligoribonucleotide hetero- and homoduplex hybridization and nuclease sensitivity. *Nucleic Acids Res* **23**, 2019-2024 (1995).
27. Flanigan, K.M. et al. Pharmacokinetics and safety of single doses of drisapersen in non-ambulant subjects with Duchenne muscular dystrophy: results of a double-blind randomized clinical trial. *Neuromuscul Disord* **24**, 16-24 (2014).
28. Crooke, S.T. & Geary, R.S. Clinical pharmacological properties of mipomersen (Kynamro), a second generation antisense inhibitor of apolipoprotein B. *Br J Clin Pharmacol* **76**, 269-276 (2013).
29. Ottesen, E.W. ISS-N1 makes the first FDA-approved drug for spinal muscular atrophy. *Translational neuroscience* **8**, 1-6 (2017).
30. Teplova, M. et al. Crystal structure and improved antisense properties of 2'-O-(2-methoxyethyl)-RNA. *Nature Structural Biology* **6**, 535-539 (1999).
31. Martin, P. Ein neuer Zugang zu 2'-O-Alkylribonucleosiden und Eigenschaften deren Oligonucleotide. *Helvetica Chimica Acta* **78**, 486-504 (1995).

32. Cirak, S. et al. Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet* **378**, 595-605 (2011).
33. Wu, B. et al. Effective rescue of dystrophin improves cardiac function in dystrophin-deficient mice by a modified morpholino oligomer. *Proceedings of the National Academy of Sciences* **105**, 14814-14819 (2008).
34. Lim, K.R., Maruyama, R. & Yokota, T. Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des Devel Ther* **11**, 533-545 (2017).
35. FDA Approves Eteplirsen for Duchenne Muscular Dystrophy: The Next Chapter in the Eteplirsen Saga. *Nucleic acid therapeutics* **27**, 1-3 (2017).
36. A Sequel to the Eteplirsen Saga: Eteplirsen Is Approved in the United States but Was Not Approved in Europe. *Nucleic acid therapeutics* **29**, 13-15 (2019).
37. Seth, P.P. et al. Short antisense oligonucleotides with novel 2'-4' conformationally restricted nucleoside analogues show improved potency without increased toxicity in animals. *J Med Chem* **52**, 10-13 (2009).
38. Koshkin, A.A. et al. LNA (Locked Nucleic Acid): An RNA Mimic Forming Exceedingly Stable LNA:LNA Duplexes. *Journal of the American Chemical Society* **120**, 13252-13253 (1998).
39. Koshkin, A.A. et al. LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **54**, 3607-3630 (1998).
40. Obika, S. et al. Synthesis of 2'-O,4'-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar pucker. *Tetrahedron Letters* **38**, 8735-8738 (1997).
41. Obika, S. et al. Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2'-O,4'-C-methylenerybonucleosides. *Tetrahedron Letters* **39**, 5401-5404 (1998).
42. Straarup, E.M. et al. Short locked nucleic acid antisense oligonucleotides potently reduce apolipoprotein B mRNA and serum cholesterol in mice and non-human primates. *Nucleic Acids Res* **38**, 7100-7111 (2010).
43. Stein, C.A. et al. Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic acids research* **38**, e3-e3 (2010).
44. Bonneau, E., Neveu, B., Kostantin, E., Tsongalis, G.J. & De Guire, V. How close are miRNAs from clinical practice? A perspective on the diagnostic and therapeutic market. *Ejifcc* **30**, 114-127 (2019).
45. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811 (1998).
46. Elbashir, S.M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498 (2001).
47. Hoy, S.M. Patisiran: First Global Approval. *Drugs* **78**, 1625-1631 (2018).

References

48. Adams, D. et al. Patisiran, an RNAi Therapeutic, for Hereditary Transthyretin Amyloidosis. *New England Journal of Medicine* **379**, 11-21 (2018).
49. Benson, M.D. et al. Inotersen Treatment for Patients with Hereditary Transthyretin Amyloidosis. *New England Journal of Medicine* **379**, 22-31 (2018).
50. Kristen, A.V. et al. Patisiran, an RNAi therapeutic for the treatment of hereditary transthyretin-mediated amyloidosis. *Neurodegenerative Disease Management* **9**, 5-23 (2019).
51. Manoharan, M. et al. Unique gene-silencing and structural properties of 2'-fluoro-modified siRNAs. *Angew Chem Int Ed Engl* **50**, 2284-2288 (2011).
52. Matranga, C., Tomari, Y., Shin, C., Bartel, D.P. & Zamore, P.D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**, 607-620 (2005).
53. Braasch, D.A. et al. RNA interference in mammalian cells by chemically-modified RNA. *Biochemistry* **42**, 7967-7975 (2003).
54. Chiu, Y.L. & Rana, T.M. siRNA function in RNAi: a chemical modification analysis. *Rna* **9**, 1034-1048 (2003).
55. Harborth, J. et al. Sequence, chemical, and structural variation of small interfering RNAs and short hairpin RNAs and the effect on mammalian gene silencing. *Antisense Nucleic Acid Drug Dev* **13**, 83-105 (2003).
56. Sioud, M. Single-stranded small interfering RNA are more immunostimulatory than their double-stranded counterparts: a central role for 2'-hydroxyl uridines in immune responses. *Eur J Immunol* **36**, 1222-1230 (2006).
57. Cekaite, L., Furset, G., Hovig, E. & Sioud, M. Gene expression analysis in blood cells in response to unmodified and 2'-modified siRNAs reveals TLR-dependent and independent effects. *J Mol Biol* **365**, 90-108 (2007).
58. Hagmann, W.K. The Many Roles for Fluorine in Medicinal Chemistry. *Journal of Medicinal Chemistry* **51**, 4359-4369 (2008).
59. Pallan, P.S. et al. Unexpected origins of the enhanced pairing affinity of 2'-fluoro-modified RNA. *Nucleic Acids Res* **39**, 3482-3495 (2011).
60. Hassler, M.R. et al. Comparison of partially and fully chemically-modified siRNA in conjugate-mediated delivery in vivo. *Nucleic Acids Res* **46**, 2185-2196 (2018).
61. Allerson, C.R. et al. Fully 2'-Modified Oligonucleotide Duplexes with Improved in Vitro Potency and Stability Compared to Unmodified Small Interfering RNA. *Journal of Medicinal Chemistry* **48**, 901-904 (2005).
62. Deleavey, G.F. et al. Synergistic effects between analogs of DNA and RNA improve the potency of siRNA-mediated gene silencing. *Nucleic Acids Res* **38**, 4547-4557 (2010).
63. Turanov, A.A. et al. RNAi modulation of placental sFLT1 for the treatment of preeclampsia. *Nature Biotechnology* **36**, 1164 (2018).

64. Geary, R.S., Norris, D., Yu, R. & Bennett, C.F. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Advanced Drug Delivery Reviews* **87**, 46-51 (2015).
65. Biscans, A. et al. Hydrophobicity of Lipid-Conjugated siRNAs Predicts Productive Loading to Small Extracellular Vesicles. *Molecular Therapy* **26**, 1520-1528 (2018).
66. Alterman, J.F. et al. Hydrophobically Modified siRNAs Silence Huntingtin mRNA in Primary Neurons and Mouse Brain. *Mol Ther Nucleic Acids* **4**, e266 (2015).
67. Alterman, J.F. et al. A divalent siRNA chemical scaffold for potent and sustained modulation of gene expression throughout the central nervous system. *Nature Biotechnology* **37**, 884-894 (2019).
68. 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).
69. Huang, Y. Preclinical and Clinical Advances of GalNAc-Decorated Nucleic Acid Therapeutics. *Molecular Therapy - Nucleic Acids* **6**, 116-132 (2017).
70. Khvorova, A. Oligonucleotide Therapeutics — A New Class of Cholesterol-Lowering Drugs. **376**, 4-7 (2017).
71. Fitzgerald, K. et al. A Highly Durable RNAi Therapeutic Inhibitor of PCSK9. *The New England journal of medicine* **376**, 41-51 (2017).
72. German, C.A. & Shapiro, M.D. Small Interfering RNA Therapeutic Inclisiran: A New Approach to Targeting PCSK9. *BioDrugs* (2019).
73. Jinek, M. et al. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
74. Cong, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819-823 (2013).
75. Barrangou, R. & Doudna, J.A. Applications of CRISPR technologies in research and beyond. *Nature Biotechnology* **34**, 933-941 (2016).
76. Dai, W.-J. et al. CRISPR-Cas9 for in vivo Gene Therapy: Promise and Hurdles. *Molecular Therapy - Nucleic Acids* **5**, e349 (2016).
77. 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).
78. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424 (2016).
79. Rees, H.A. & Liu, D.R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature Reviews Genetics* (2018).
80. Vogel, P. & Stafforst, T. Critical review on engineering deaminases for site-directed RNA editing. *Current Opinion in Biotechnology* **55**, 74-80 (2019).
81. Bass, B.L. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* **71**, 817-846 (2002).

References

82. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* **79**, 321-349 (2010).
83. Bass, B.L. & Weintraub, H. A developmentally regulated activity that unwinds RNA duplexes. *Cell* **48**, 607-613 (1987).
84. Rebagliati, M.R. & Melton, D.A. Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell* **48**, 599-605 (1987).
85. Bass, B.L. & Weintraub, H. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **55**, 1089-1098 (1988).
86. Wagner, R.W., Smith, J.E., Cooperman, B.S. & Nishikura, K. A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and *Xenopus* eggs. *Proc Natl Acad Sci U S A* **86**, 2647-2651 (1989).
87. Wagner, R.W. & Nishikura, K. Cell cycle expression of RNA duplex unwindase activity in mammalian cells. *Molecular and Cellular Biology* **8**, 770-777 (1988).
88. Jin, Y., Zhang, W. & Li, Q. Origins and evolution of ADAR-mediated RNA editing. *IUBMB Life* **61**, 572-578 (2009).
89. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat Rev Mol Cell Biol* **17**, 83-96 (2016).
90. 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).
91. Melcher, T. et al. RED2, a brain-specific member of the RNA-specific adenosine deaminase family. *The Journal of biological chemistry* **271**, 31795-31798 (1996).
92. 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. *Proceedings of the National Academy of Sciences* **96**, 4621-4626 (1999).
93. 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. *Molecular and Cellular Biology* **15**, 5376-5388 (1995).
94. Herbert, A. et al. A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc Natl Acad Sci U S A* **94**, 8421-8426 (1997).
95. Desterro, J.M. et al. Dynamic association of RNA-editing enzymes with the nucleolus. *J Cell Sci* **116**, 1805-1818 (2003).
96. Strehlow, 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).

97. 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. *Molecular and Cellular Biology* **29**, 1487-1497 (2009).
98. Poulsen, H., Nilsson, J., Damgaard, C.K., Egebjerg, J. & Kjems, J. CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol Cell Biol* **21**, 7862-7871 (2001).
99. Barraud, P., Banerjee, S., Mohamed, W.I., Jantsch, M.F. & Allain, F.H. A bimodular nuclear localization signal assembled via an extended double-stranded RNA-binding domain acts as an RNA-sensing signal for transportin 1. *Proc Natl Acad Sci U S A* **111**, E1852-1861 (2014).
100. Maas, S. & Gommans, W.M. Identification of a selective nuclear import signal in adenosine deaminases acting on RNA. *Nucleic acids research* **37**, 5822-5829 (2009).
101. Picardi, E. et al. Profiling RNA editing in human tissues: towards the inosinome Atlas. *Scientific Reports* **5**, 14941 (2015).
102. Tan, M.H. et al. Dynamic landscape and regulation of RNA editing in mammals. *Nature* **550**, 249-254 (2017).
103. 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).
104. 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).
105. Lomeli, H. et al. Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science* **266**, 1709-1713 (1994).
106. Daniel, C., Wahlstedt, H., Ohlson, J., Bjork, P. & Ohman, M. Adenosine-to-inosine RNA editing affects trafficking of the gamma-aminobutyric acid type A (GABA(A)) receptor. *The Journal of biological chemistry* **286**, 2031-2040 (2011).
107. Ohlson, J., Pedersen, J.S., Haussler, D. & Ohman, M. Editing modifies the GABAA receptor subunit $\alpha 3$. *Rna* **13**, 698-703 (2007).
108. Bhalla, T., Rosenthal, J.J.C., Holmgren, M. & Reenan, R. Control of human potassium channel inactivation by editing of a small mRNA hairpin. *Nature Structural & Molecular Biology* **11**, 950-956 (2004).
109. Burns, C.M. et al. Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* **387**, 303-308 (1997).
110. Melcher, T. et al. A mammalian RNA editing enzyme. *Nature* **379**, 460-464 (1996).
111. 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).

References

112. Chalk, A.M., Taylor, S., Heraud-Farlow, J.E. & Walkley, C.R. The majority of A-to-I RNA editing is not required for mammalian homeostasis. *Genome Biology* **20**, 268 (2019).
113. Levanon, E.Y. et al. Evolutionarily conserved human targets of adenosine to inosine RNA editing. *Nucleic acids research* **33**, 1162-1168 (2005).
114. Chan, T.H. et al. A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma. *Gut* **63**, 832-843 (2014).
115. Chen, L. et al. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. *Nat Med* **19**, 209-216 (2013).
116. Kawahara, Y. et al. Frequency and fate of microRNA editing in human brain. *Nucleic Acids Research* **36**, 5270-5280 (2008).
117. Kawahara, Y. et al. Redirection of Silencing Targets by Adenosine-to-Inosine Editing of miRNAs. *Science* **315**, 1137-1140 (2007).
118. Yang, W. et al. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol* **13**, 13-21 (2006).
119. Ramaswami, G. et al. Accurate identification of human Alu and non-Alu RNA editing sites. *Nature Methods* **9**, 579-581 (2012).
120. Ramaswami, G. et al. Identifying RNA editing sites using RNA sequencing data alone. *Nature Methods* **10**, 128 (2013).
121. Athanasiadis, A., Rich, A. & Maas, S. Widespread A-to-I RNA Editing of Alu-Containing mRNAs in the Human Transcriptome. *PLOS Biology* **2**, e391 (2004).
122. 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).
123. Levanon, E.Y. et al. Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nature Biotechnology* **22**, 1001-1005 (2004).
124. Li, J.B. et al. Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. *Science* **324**, 1210-1213 (2009).
125. Peng, Z. et al. Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat Biotechnol* **30**, 253-260 (2012).
126. Porath, H.T., Carmi, S. & Levanon, E.Y. A genome-wide map of hyper-edited RNA reveals numerous new sites. *Nature communications* **5**, 4726 (2014).
127. Lev-Maor, G. et al. RNA-editing-mediated exon evolution. *Genome Biol* **8**, R29 (2007).
128. Rueter, S.M., Dawson, T.R. & Emeson, R.B. Regulation of alternative splicing by RNA editing. *Nature* **399**, 75-80 (1999).
129. 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).
130. Hartner, J.C., Walkley, C.R., Lu, J. & Orkin, S.H. ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling. *Nat Immunol* **10**, 109-115 (2009).

131. Wang, Q. et al. Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *The Journal of biological chemistry* **279**, 4952-4961 (2004).
132. Hartner, J.C. et al. Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *The Journal of biological chemistry* **279**, 4894-4902 (2004).
133. XuFeng, R. et al. ADAR1 is required for hematopoietic progenitor cell survival via RNA editing. *Proc Natl Acad Sci U S A* **106**, 17763-17768 (2009).
134. Liddicoat, B.J. et al. RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* **349**, 1115-1120 (2015).
135. Mannion, N.M. et al. The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. *Cell Rep* **9**, 1482-1494 (2014).
136. Ahmad, S. et al. Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation. *Cell* **172**, 797-810.e713 (2018).
137. Vitali, P. & Scadden, A.D. Double-stranded RNAs containing multiple IU pairs are sufficient to suppress interferon induction and apoptosis. *Nat Struct Mol Biol* **17**, 1043-1050 (2010).
138. 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).
139. Chung, H. et al. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. *Cell* **172**, 811-824.e814 (2018).
140. Macbeth, M.R. et al. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**, 1534-1539 (2005).
141. 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).
142. 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).
143. Betts, L., Xiang, S., Short, S.A., Wolfenden, R. & Carter, C.W., Jr. Cytidine deaminase. The 2.3 Å crystal structure of an enzyme: transition-state analog complex. *J Mol Biol* **235**, 635-656 (1994).
144. Kuratani, M. et al. Crystal Structure of tRNA Adenosine Deaminase (TadA) from *Aquifex aeolicus*. *Journal of Biological Chemistry* **280**, 16002-16008 (2005).
145. Phelps, K.J. et al. Recognition of duplex RNA by the deaminase domain of the RNA editing enzyme ADAR2. *Nucleic Acids Research* **43**, 1123-1132 (2015).
146. Kuttan, A. & Bass, B.L. Mechanistic insights into editing-site specificity of ADARs. *Proc Natl Acad Sci U S A* **109**, E3295-3304 (2012).
147. Wong, S.K., Sato, S. & Lazinski, D.W. Substrate recognition by ADAR1 and ADAR2. *Rna* **7**, 846-858 (2001).

References

148. Eggington, J.M., Greene, T. & Bass, B.L. Predicting sites of ADAR editing in double-stranded RNA. *Nature communications* **2**, 319 (2011).
149. Stefl, R., Xu, M., Skrisovska, L., Emeson, R.B. & Allain, F.H. Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure* **14**, 345-355 (2006).
150. Wang, X., Vukovic, L., Koh, H.R., Schulten, K. & Myong, S. Dynamic profiling of double-stranded RNA binding proteins. *Nucleic Acids Research* **43**, 7566-7576 (2015).
151. Thomas, J.M. & Beal, P.A. How do ADARs bind RNA? New protein-RNA structures illuminate substrate recognition by the RNA editing ADARs. *BioEssays : news and reviews in molecular, cellular and developmental biology* (2017).
152. Zheng, Y., Lorenzo, C. & Beal, P.A. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res* (2017).
153. Jaikaran, D.C., Collins, C.H. & MacMillan, A.M. Adenosine to inosine editing by ADAR2 requires formation of a ternary complex on the GluR-B R/G site. *The Journal of biological chemistry* **277**, 37624-37629 (2002).
154. Valente, L. & Nishikura, K. RNA binding-independent dimerization of adenosine deaminases acting on RNA and dominant negative effects of nonfunctional subunits on dimer functions. *The Journal of biological chemistry* **282**, 16054-16061 (2007).
155. Cho, D.S. et al. Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *The Journal of biological chemistry* **278**, 17093-17102 (2003).
156. Poulsen, H. et al. Dimerization of ADAR2 is mediated by the double-stranded RNA binding domain. *Rna* **12**, 1350-1360 (2006).
157. Licht, K. et al. Inosine induces context-dependent recoding and translational stalling. *Nucleic Acids Research* **47**, 3-14 (2018).
158. 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).
159. Stafforst, T. & Schneider, M.F. An RNA-deaminase conjugate selectively repairs point mutations. *Angew Chem Int Ed Engl* **51**, 11166-11169 (2012).
160. Vogel, P. & Stafforst, T. Site-directed RNA editing with antagomir deaminases--a tool to study protein and RNA function. *ChemMedChem* **9**, 2021-2025 (2014).
161. 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 Engl* **53**, 6267-6271 (2014).
162. 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. *Journal of the American Chemical Society* **137**, 15875-15881 (2015).

163. Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching Protein Localization by Site-Directed RNA Editing under Control of Light. *ACS Synthetic Biology* **6**, 1642-1649 (2017).
164. Vogel, P. et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat Methods* **15**, 535-538 (2018).
165. 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. *Proceedings of the National Academy of Sciences* **110**, 18285-18290 (2013).
166. Montiel-Gonzalez, M.F., Vallecillo-Viejo, I.C. & Rosenthal, J.J. An efficient system for selectively altering genetic information within mRNAs. *Nucleic Acids Res* **44**, e157 (2016).
167. Sinnamon, J.R. et al. Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proceedings of the National Academy of Sciences* (2017).
168. 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 therapy* **24**, 779-786 (2017).
169. Abudayyeh, O.O. et al. RNA targeting with CRISPR–Cas13. *Nature* **550**, 280 (2017).
170. Cox, D.B.T. et al. RNA editing with CRISPR-Cas13. *Science* (2017).
171. Abudayyeh, O.O. et al. A cytosine deaminase for programmable single-base RNA editing. *Science*, eaax7063 (2019).
172. Monteleone, L.R. et al. A Bump-Hole Approach for Directed RNA Editing. *Cell Chem Biol* **26**, 269-277.e265 (2019).
173. Rauch, S. et al. Programmable RNA-Guided RNA Effector Proteins Built from Human Parts. *Cell* **178**, 122-134.e112 (2019).
174. Woolf, T.M., Chase, J.M. & Stinchcomb, D.T. Toward the therapeutic editing of mutated RNA sequences. *Proceedings of the National Academy of Sciences* **92**, 8298-8302 (1995).
175. 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 Research* (2016).
176. Fukuda, M. et al. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Scientific Reports* **7**, 41478 (2017).
177. Valente, E.M. et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science* **304**, 1158-1160 (2004).
178. 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).
179. Katrekar, D. et al. In vivo RNA editing of point mutations via RNA-guided adenosine deaminases. *Nature Methods* (2019).

References

180. Qu, L. et al. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nature Biotechnology* (2019).
181. Hurler, G. Über einen Typ multipler Abartungen, vorwiegend am Skelettsystem. *Zeitschrift für Kinderheilkunde* **24**, 220-234 (1920).
182. Terlato, N.J. & Cox, G.F. Can mucopolysaccharidosis type I disease severity be predicted based on a patient's genotype? A comprehensive review of the literature. *Genet Med* **5**, 286-294 (2003).
183. Scheie, H.G., Hambrick, G.W., Jr. & Barness, L.A. A newly recognized forme fruste of Hurler's disease (gargoylism). *Am J Ophthalmol* **53**, 753-769 (1962).
184. Scott, H.S. et al. Molecular genetics of mucopolysaccharidosis type I: Diagnostic, clinical, and biological implications. **6**, 288-302 (1995).
185. Moskowitz, S.M., Tieu, P.T. & Neufeld, E.F. Mutation in Scheie syndrome (MPS IS): A G→A transition creates new splice site in intron 5 of one IDUA allele. *Human Mutation* **2**, 141-144 (1993).
186. Giugliani, R. et al. Mucopolysaccharidosis I, II, and VI: brief review and guidelines for treatment. *Genetics and Molecular Biology* **33**, 589-604 (2010).
187. de Ru, M.H. et al. Enzyme replacement therapy and/or hematopoietic stem cell transplantation at diagnosis in patients with mucopolysaccharidosis type I: results of a European consensus procedure. *Orphanet J Rare Dis* **6**, 55 (2011).
188. Eisengart, J.B. et al. Long-term outcomes of systemic therapies for Hurler syndrome: an international multicenter comparison. *Genet Med* **20**, 1423-1429 (2018).
189. Moore, D., Connock, M.J., Wraith, E. & Lavery, C. The prevalence of and survival in Mucopolysaccharidosis I: Hurler, Hurler-Scheie and Scheie syndromes in the UK. *Orphanet J Rare Dis* **3**, 24 (2008).
190. Ou, L. et al. ZFN-Mediated In Vivo Genome Editing Corrects Murine Hurler Syndrome. *Molecular Therapy* **27**, 178-187 (2019).
191. Ponder, K.P. & Haskins, M.E. Gene therapy for mucopolysaccharidosis. *Expert opinion on biological therapy* **7**, 1333-1345 (2007).
192. Schuh, R.S. et al. In vivo genome editing of mucopolysaccharidosis I mice using the CRISPR/Cas9 system. *J Control Release* **288**, 23-33 (2018).
193. Janciauskiene, S.M. et al. The discovery of alpha1-antitrypsin and its role in health and disease. *Respir Med* **105**, 1129-1139 (2011).
194. Nukiwa, T. et al. Identification of a second mutation in the protein-coding sequence of the Z type alpha 1-antitrypsin gene. *The Journal of biological chemistry* **261**, 15989-15994 (1986).
195. Hazari, Y.M. et al. Alpha-1-antitrypsin deficiency: Genetic variations, clinical manifestations and therapeutic interventions. *Mutat Res* **773**, 14-25 (2017).
196. Yoshida, A., Ewing, C., Wessels, M., Lieberman, J. & Gaidulis, L. Molecular abnormality of PI S variant of human alpha1-antitrypsin. *American journal of human genetics* **29**, 233-239 (1977).

197. de Serres, F.J. Worldwide racial and ethnic distribution of alpha1-antitrypsin deficiency: summary of an analysis of published genetic epidemiologic surveys. *Chest* **122**, 1818-1829 (2002).
198. Llewellyn-Jones, C.G., Lomas, D.A., Carrell, R.W. & Stockley, R.A. The effect of the Z mutation on the ability of alpha 1-antitrypsin to prevent neutrophil mediated tissue damage. *Biochim Biophys Acta* **1227**, 155-160 (1994).
199. Gooptu, B. & Lomas, D.A. Conformational pathology of the serpins: themes, variations, and therapeutic strategies. *Annu Rev Biochem* **78**, 147-176 (2009).
200. Ogushi, F., Fells, G.A., Hubbard, R.C., Straus, S.D. & Crystal, R.G. Z-type alpha 1-antitrypsin is less competent than M1-type alpha 1-antitrypsin as an inhibitor of neutrophil elastase. *J Clin Invest* **80**, 1366-1374 (1987).
201. Lomas, D.A., Evans, D.L., Finch, J.T. & Carrell, R.W. The mechanism of Z alpha 1-antitrypsin accumulation in the liver. *Nature* **357**, 605-607 (1992).
202. Lomas, D.A. & Mahadeva, R. α 1-Antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy. *The Journal of Clinical Investigation* **110**, 1585-1590 (2002).
203. Carrell, R.W. & Lomas, D.A. Alpha1-antitrypsin deficiency--a model for conformational diseases. *The New England journal of medicine* **346**, 45-53 (2002).
204. Janus, E., Phillips, N. & Carrell, R. SMOKING, LUNG FUNCTION, AND α 1-ANTITRYPSIN DEFICIENCY. *The Lancet* **325**, 152-154 (1985).
205. Sveger, T. Liver disease in alpha1-antitrypsin deficiency detected by screening of 200,000 infants. *The New England journal of medicine* **294**, 1316-1321 (1976).
206. Wewers, M.D. et al. Replacement therapy for alpha 1-antitrypsin deficiency associated with emphysema. *The New England journal of medicine* **316**, 1055-1062 (1987).
207. Stecenko, A.A. & Brigham, K.L. Gene therapy progress and prospects: alpha-1 antitrypsin. *Gene therapy* **10**, 95-99 (2003).
208. Guo, S. et al. Antisense oligonucleotide treatment ameliorates alpha-1 antitrypsin-related liver disease in mice. *The Journal of Clinical Investigation* **124**, 251-261 (2014).
209. Rett, A. [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. *Wien Med Wochenschr* **116**, 723-726 (1966).
210. Hagberg, B., Aicardi, J., Dias, K. & Ramos, O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* **14**, 471-479 (1983).
211. Topcu, M. et al. Somatic mosaicism for a MECP2 mutation associated with classic Rett syndrome in a boy. *Eur J Hum Genet* **10**, 77-81 (2002).
212. Trappe, R. et al. MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am J Hum Genet* **68**, 1093-1101 (2001).
213. Amir, R.E. et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics* **23**, 185-188 (1999).

References

214. Lyst, M.J. & Bird, A. Rett syndrome: a complex disorder with simple roots. *Nature Reviews Genetics* **16**, 261-275 (2015).
215. Armstrong, D., Dunn, J.K., Antalffy, B. & Trivedi, R. Selective dendritic alterations in the cortex of Rett syndrome. *J Neuropathol Exp Neurol* **54**, 195-201 (1995).
216. Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. Reversal of Neurological Defects in a Mouse Model of Rett Syndrome. *Science* **315**, 1143-1147 (2007).
217. Van Esch, H. et al. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* **77**, 442-453 (2005).
218. 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. *Angewandte Chemie International Edition* **53**, 6267-6271 (2014).
219. Helmling, C. et al. Rapid NMR screening of RNA secondary structure and binding. *Journal of Biomolecular NMR* **63**, 67-76 (2015).
220. O'Shea, J.J. et al. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu Rev Med* **66**, 311-328 (2015).
221. Ramaswami, G. & Li, J.B. RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res* **42**, D109-113 (2014).
222. Woodard, L.E. & Wilson, M.H. piggyBac-ing models and new therapeutic strategies. *Trends Biotechnol* **33**, 525-533 (2015).
223. Antonelli, G., Scagnolari, C., Moschella, F. & Proietti, E. Twenty-five years of type I interferon-based treatment: a critical analysis of its therapeutic use. *Cytokine Growth Factor Rev* **26**, 121-131 (2015).
224. Jayachandran, U., Cook, A.G. & Grey, H. Nuclear factor 90 uses an ADAR2-like binding mode to recognize specific bases in dsRNA. *Nucleic Acids Research* **44**, 1924-1936 (2015).
225. Rigo, F. et al. Synthetic oligonucleotides recruit ILF2/3 to RNA transcripts to modulate splicing. *Nat Chem Biol* **8**, 555-561 (2012).
226. Wang, D. et al. Characterization of an MPS I-H knock-in mouse that carries a nonsense mutation analogous to the human IDUA-W402X mutation. *Molecular Genetics and Metabolism* **99**, 62-71 (2010).
227. Ou, L., Herzog, T.L., Wilmot, C.M. & Whitley, C.B. Standardization of α -L-iduronidase enzyme assay with Michaelis–Menten kinetics. *Molecular Genetics and Metabolism* **111**, 113-115 (2014).

6 Appendix

Manuscript 1 (Appendix p.1) + SI (Appendix p.10)

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

Manuscript 2 (Appendix p.63) + SI (Appendix p.74):

T. Merkle, S. Merz, P. Reautschnig, A. Blaha, Q. Li, P. Vogel, J. Wettengel, J. B. Li, T. Stafforst, Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nature Biotechnology* **37**, 133-138 (2019)

Manuscript 3 (Appendix p.102)

T. Merkle and T. Stafforst, New frontiers for site-directed RNA editing – harnessing endogenous ADARs. *Methods in Molecular Biology* (accepted)

Manuscript 4 (Appendix p.118) + SI (Appendix p.128)

T. Merkle, C. Schlitz, L. Pfeiffer, C. Lochmann, T. Stafforst, Improved antisense oligonucleotides for efficient and precise RNA editing with endogenous ADARs (in preparation)

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

Paul Vogel¹, Matin Moschref¹, Qin Li², Tobias Merkle¹, Karthika D. Selvasarayanan¹, 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 *SA1/2*. 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 *SA1/2* 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

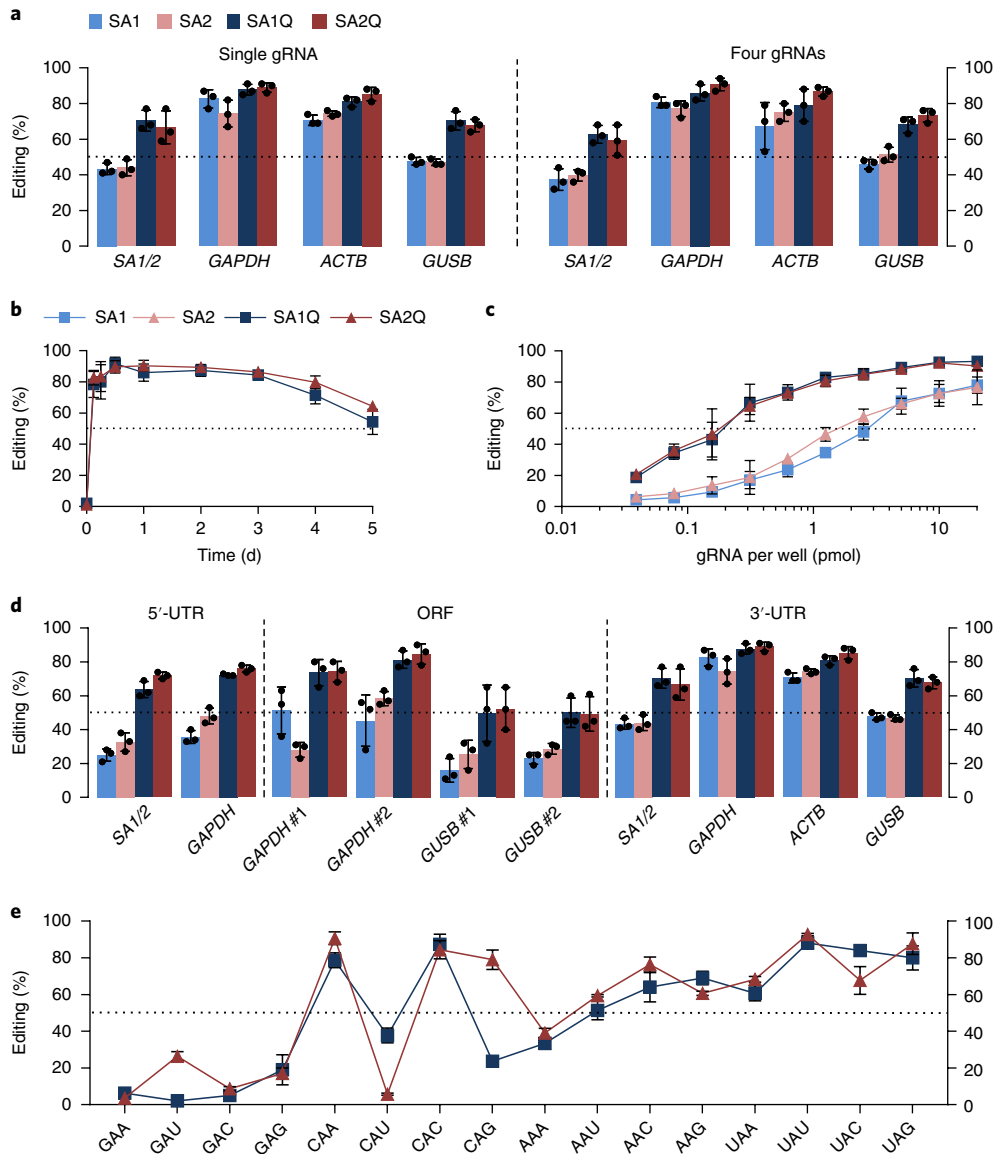


Fig. 1 | Editing performance of four SNAP-ADARs. **a**, Engineered 293 cell lines expressing one of four editase enzymes (see key) were transfected with either a single gRNA or four gRNAs targeting 5'-UAG triplets in the indicated endogenous transcripts. **b,c**, Time (**b**) and dose (**c**) dependency 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).

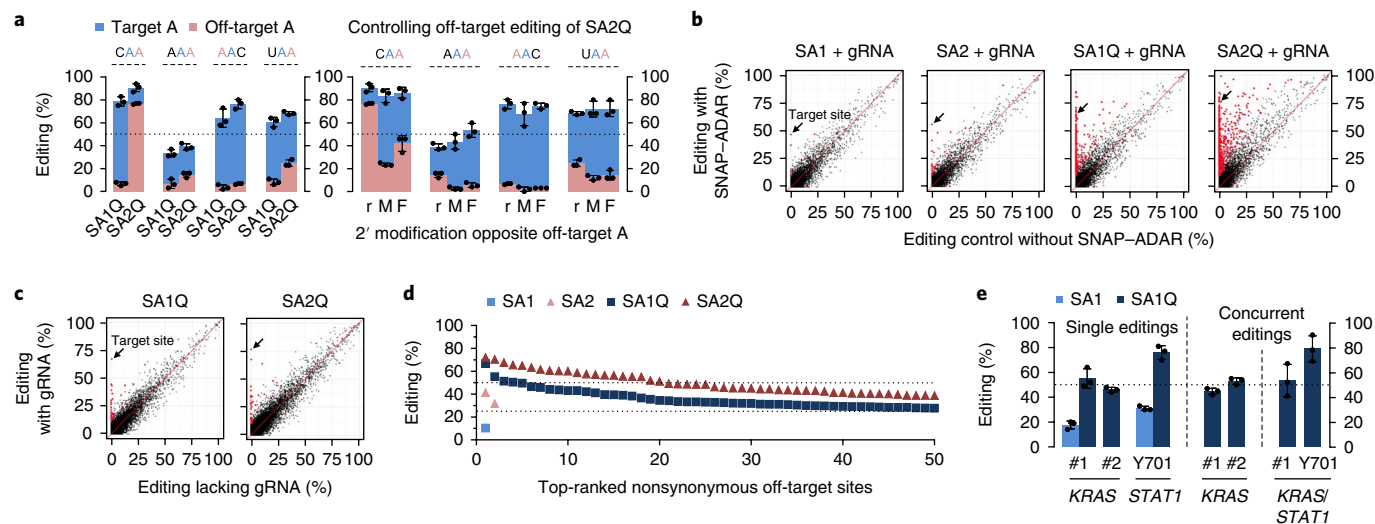


Fig. 2 | Editing specificity and application. **a**, Off-target editing of adjacent adenines (A) in A-rich triplets. r, M, and F refer to the chemical modification opposite the off-target A; r, natural ribonucleotide; M, 2'-methoxy; F, 2'-fluoro. **b, c**, Scatter plots of differential editing at ~50,000 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'-UAA 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).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

► Experimental design

1. Sample size

Describe how sample size was determined.

All experiments for evaluating editing yields of endogenous targets via Sanger sequencing were done in triplicate (independent experiments) to validate reproducibility and to provide appropriate standard deviations. NGS analysis was performed with two independent replicates per sample; the required sequencing depth was determined in a pilot experiment and saturated with 50 Mio 100 bp paired-end reads at 25 000 detected transcripts. This sequencing depth was also similar to other very recent papers on global off-target effects of site-directed RNA editing (Cox et al. Science 2017, Rosenthal et al. RNA Biol. 2018)

2. Data exclusions

Describe any data exclusions.

There was no data exclusion; details for the NGS analysis pipeline and filter settings are given in the online methods in full detail

3. Replication

Describe whether the experimental findings were reliably reproduced.

all experiments could reliably be reproduced

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

no randomization was performed, samples were treated according to the same protocols side-by-side with the respective controls

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

no blinding was performed, editing experiments were allocated to several experimentators

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

As outlined in full detail in the online methods all software tools used for NGS are publicly available: Mapping of RNA-seq and reads: BWA was used to align the reads to a combination of the reference genome sequences (hg19) and exonic sequences surrounding known splicing junctions from known gene models, obtained through the UCSC Genome Browser for Gencode, RefSeq, Ensembl, and UCSC Genes. Unique and non-duplicate reads were subjected to local realignment and base score recalibration using the IndelRealigner and TableRecalibration from the Genome Analysis Toolkit (GATK). 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. We first removed all known human 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/>). Finally, variants were annotated using ANNOVAR based on 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 database and were subsequently referred to as 'known' sites if also found in RADAR, or 'novel' sites if not found.

The manuscript and Supplementary Information were written with Microsoft Word 2016, Sanger editing data was analyzed (mean, SD) and plotted with GraphPad Prism 7.04 and Excel 2016. Figures were prepared with CorelDraw 2017.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). [Nature Methods guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

no unique material was used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

The protocol is provided in the manuscript: primary antibodies have been used against the SNAP-tag (1:1000, P9310S, New England Biolabs, USA) and β -actin (1:40000, A5441, Sigma Aldrich, USA). Both antibodies are well established commercial antibodies. The SNAP-tag antibody was validated by the fact that total protein from parental empty cells (not expressing a SNAP-tagged protein) do not stain in the immunoblot. After integration of the SNAP-tagged protein, the total protein showed a clear doxycycline-inducible protein band of the expected size. The secondary antibodies were HRP-conjugated anti-rabbit (1:10000, 111-035-003, Jackson Immuno Research Laboratories, USA) and anti-mouse (1:10000, 115-035-003, Jackson Immuno Research Laboratories, USA). Both are well-known and validated commercial secondary antibodies.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

We generated cell lines derived from the parental Flip-In T-REx cell line (Catalog no. R78007, Thermo Fisher scientific)

b. Describe the method of cell line authentication used.

Cell line authentication was confirmed by antibiotic selection before and after recombination

c. Report whether the cell lines were tested for mycoplasma contamination.

all cell lines were tested negative for mycoplasma contamination

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

no commonly misidentified cell lines were used

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

no animals were used

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

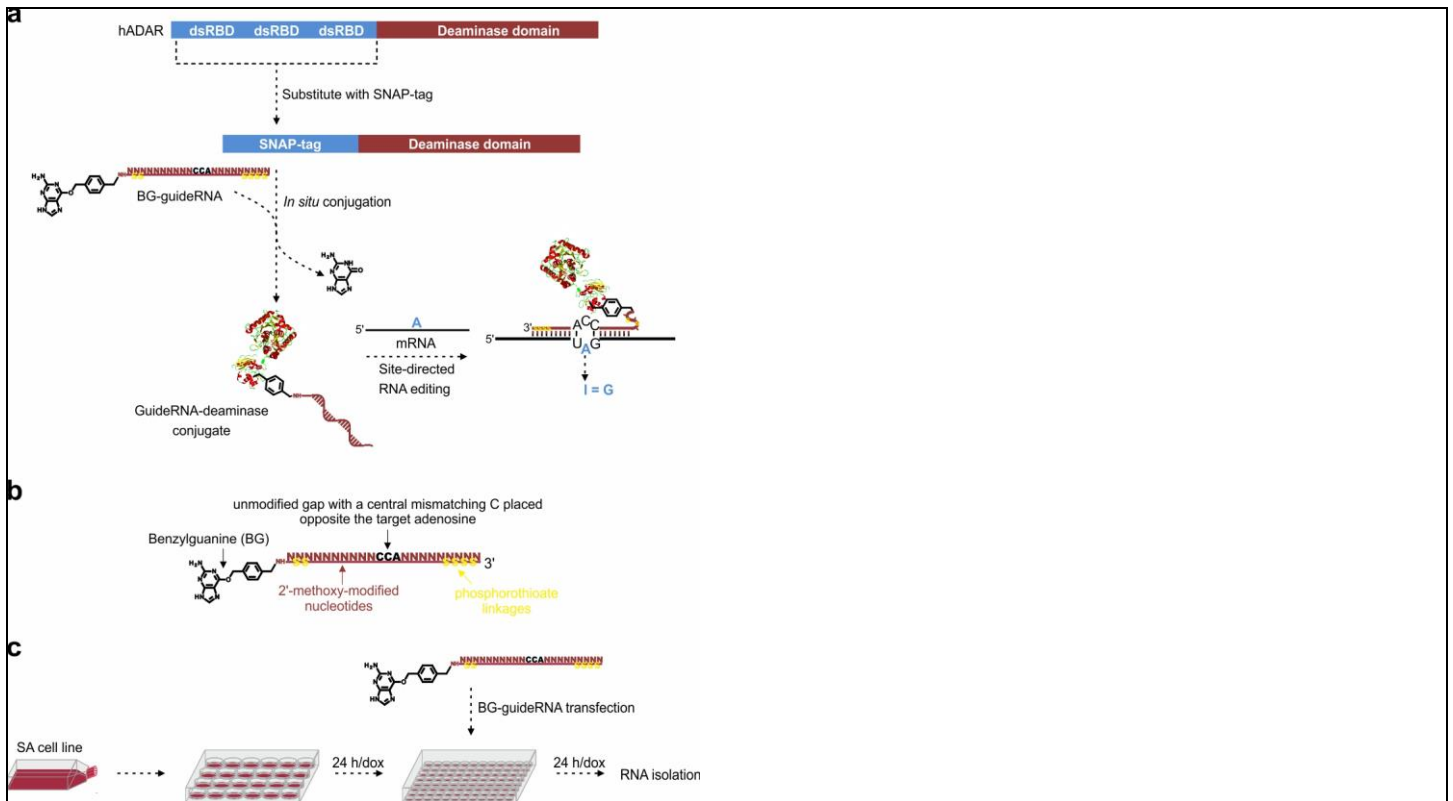
the study did not involve human research participants

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*}

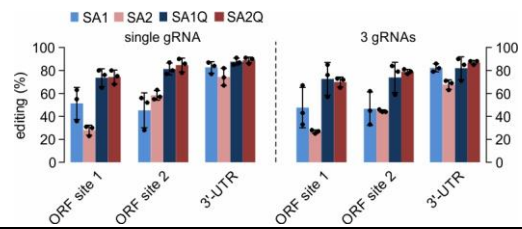
¹Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. ²Department of Genetics, Stanford University, Stanford, CA, USA.
*e-mail: thorsten.stafforst@uni-tuebingen.de



Supplementary Figure 1

Site-directed RNA editing by SNAP-tagged ADARs driven by short, chemically modified guide RNAs.

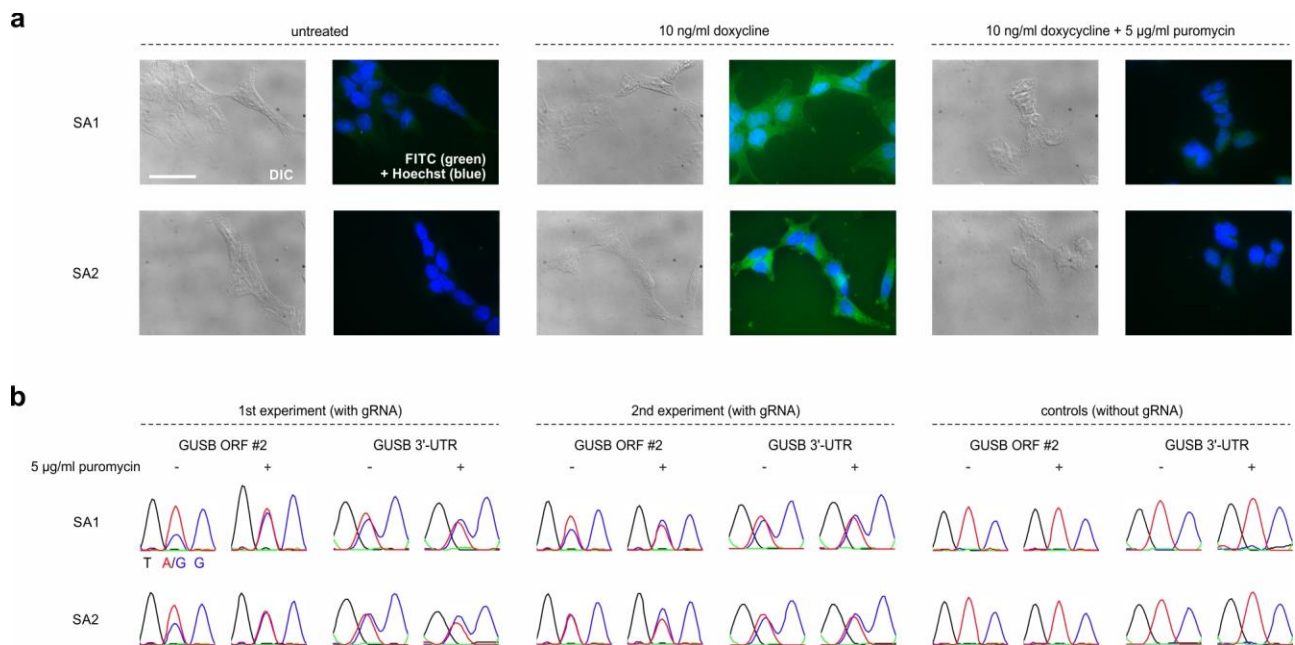
a) General concept: The double-stranded RNA-binding domains (dsRBDs) of hADAR have been substituted with the SNAP-tag. The latter is able to form a covalent bond to a guideRNA that is modified with benzylguanine (BG). When bound to the SNAP-ADAR, the guideRNA steers the attached SNAP-ADAR protein to 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.

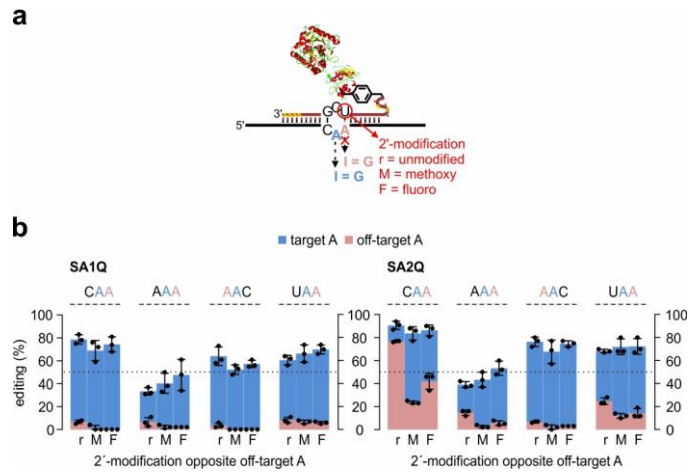
The respective SA-expressing cells were transfected with either a single gRNA or 3 gRNAs against distinct sites on the endogenous *GAPDH* transcript. Data are shown as the mean \pm SD, N=3 independent experiments, black dots represent individual data points.



Supplementary Figure 3

Effect of global translation inhibition (puromycin) on RNA editing.

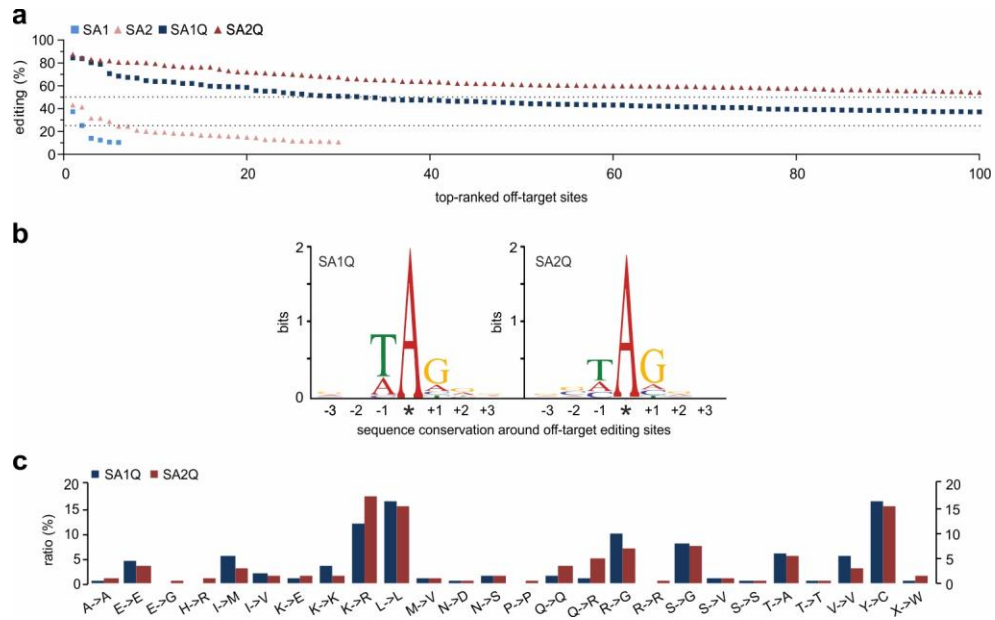
a) SA1 and SA2 cells were seeded on poly-D-lysine-coated glass slides. After one day, doxycycline (end concentration = 10 ng/ml) was added to induce SA enzyme expression. To inhibit translation, cells were additionally supplemented with 5 µg/ml puromycin, respectively. After 12h, cells were stained with 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.

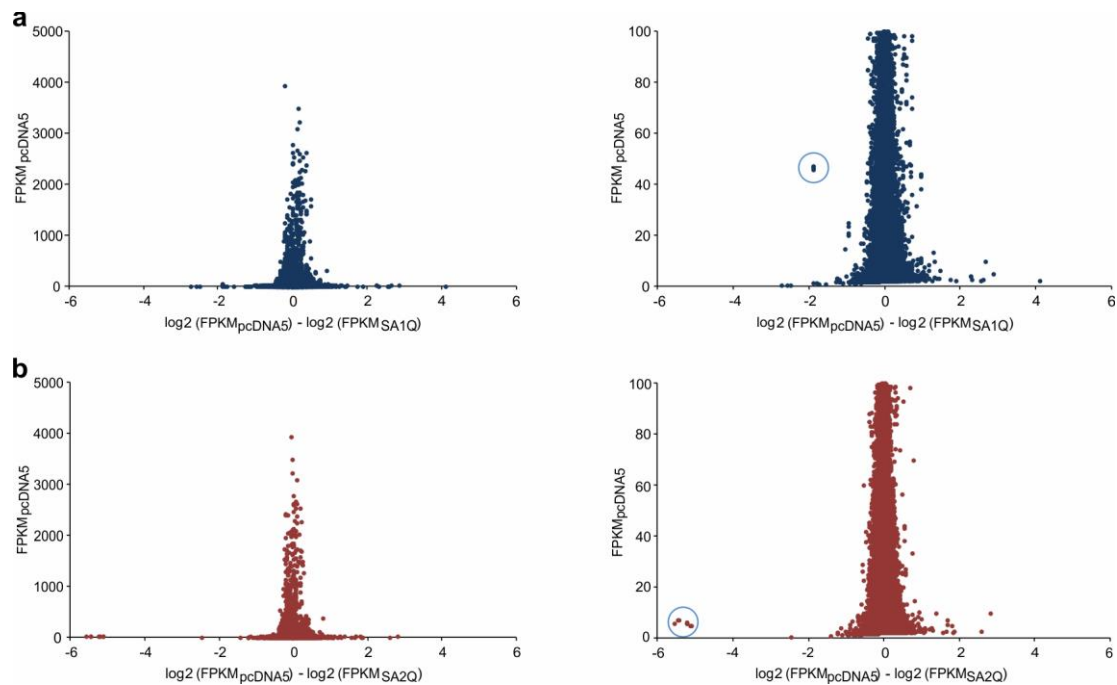
a) General strategy. To avoid unintended editing of an adjacent adenosine at the target site, the opposing base in the guideRNA can be modified by 2'-methoxylation (M) or 2'-fluorination (F). This is exemplary shown for the triplet CAA. **b)** In the study, off-target editing of an adjacent adenosine 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.

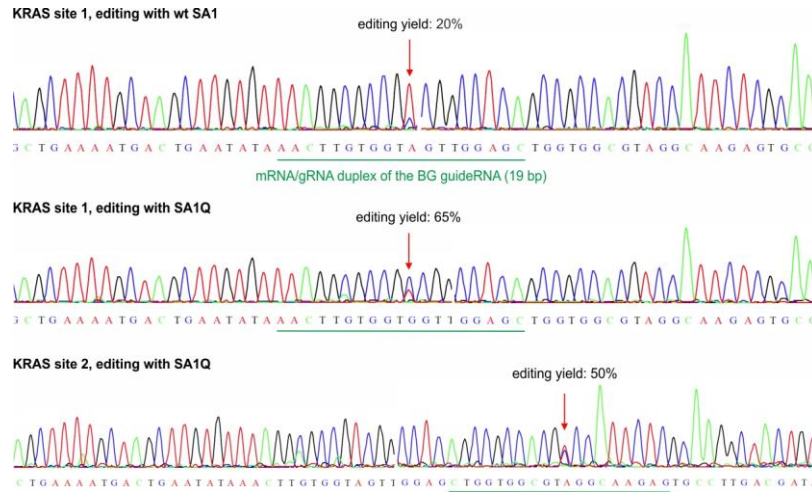
a) Overall off-target sites are ranked by editing yields. **b)** Logo represents the sequence conservation around all significant off-target sites for SA1Q and SA2Q. **c)** Analysis of the codon changes for all off-target editings that were found in SA1Q and SA2Q cells. The ratio was calculated in relation to the total number of editing events happened in the coding region of the transcripts.



Supplementary Figure 6

Gene expression analysis.

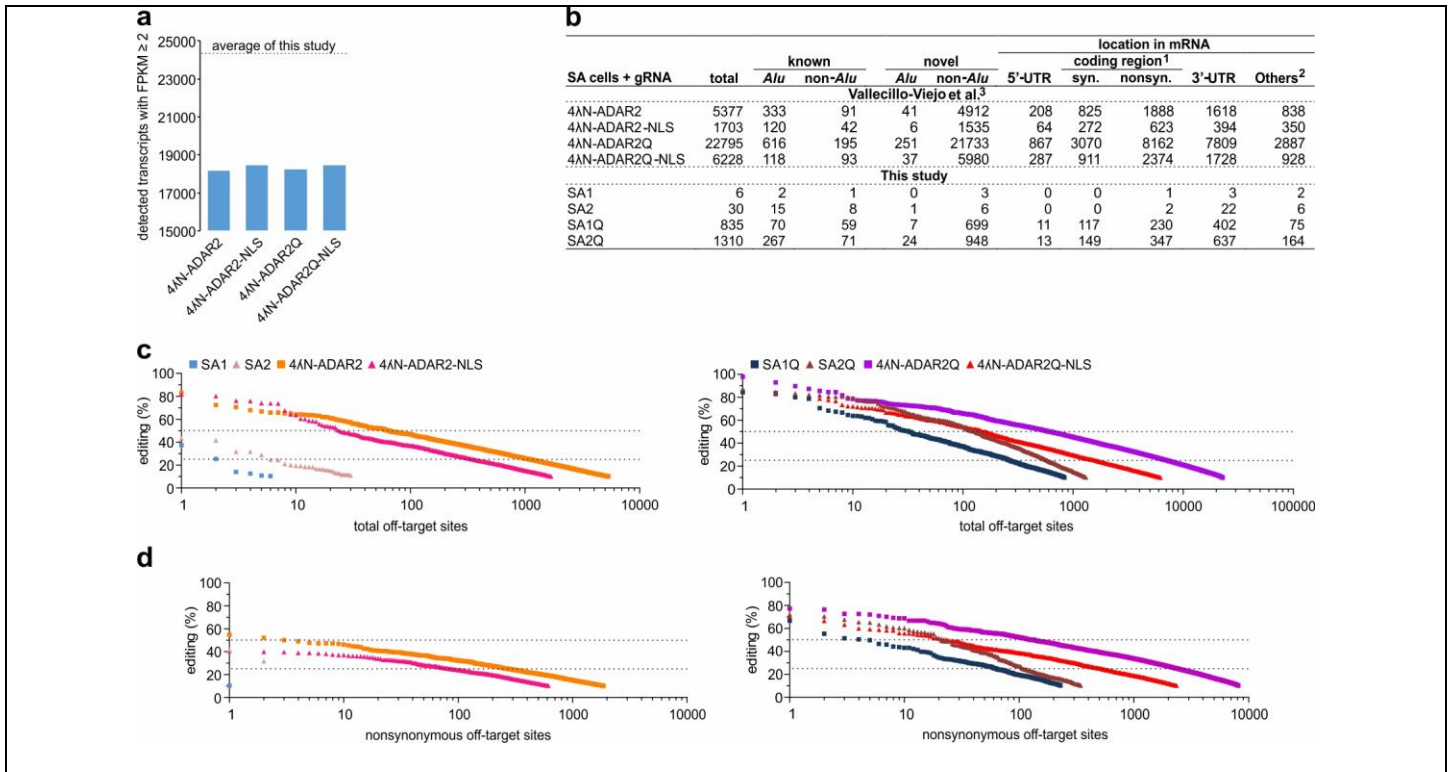
FPKM values of approximately 25,000 expressed transcripts are compared between cells containing the empty pcDNA5 vector with SA1Q cells + gRNA (a) or SA2Q cells + gRNA (b). Plotted is the log₂ fold change in expression against the FPKM of the respective transcript in the control cell line (pcDNA5). The 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 log₂-fold changes > 1. Log₂-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.

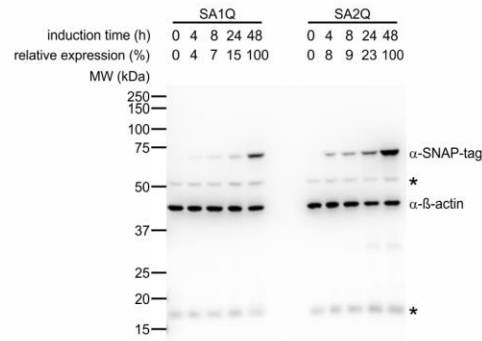
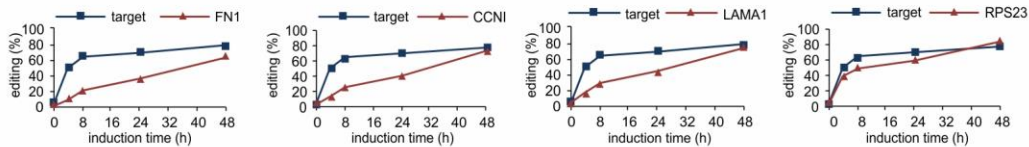
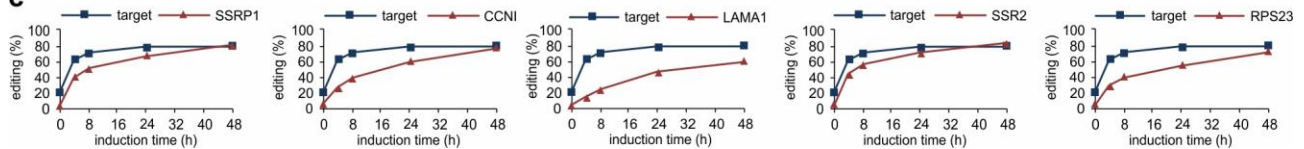
The applied BG-gRNAs form 19 bp duplex structures with the target transcript. No off-target editing was detected within these mRNA/gRNA duplexes in *KRAS* mRNA. For further discussion, see also Supplementary Note 1. N=3 independent experiments were performed with similar results.



Supplementary Figure 8

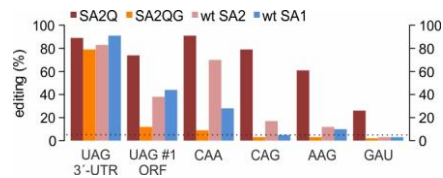
Reanalysis of NGS data produced by Vallecillo-Viejo et al.¹⁵ according to our pipeline (Methods).

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. ³Editings 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).

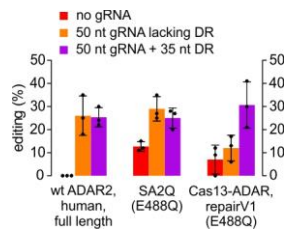
a) The expression of SA1Q or SA2Q was varied and quantified by western blot analysis (shown in relative expression, asterisks (*) indicate unspecific protein bands). We assessed the effect of reduced SA expression levels on editing the on-target (GAPDH, ORF site #2) versus several high-ranked off-targets 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).

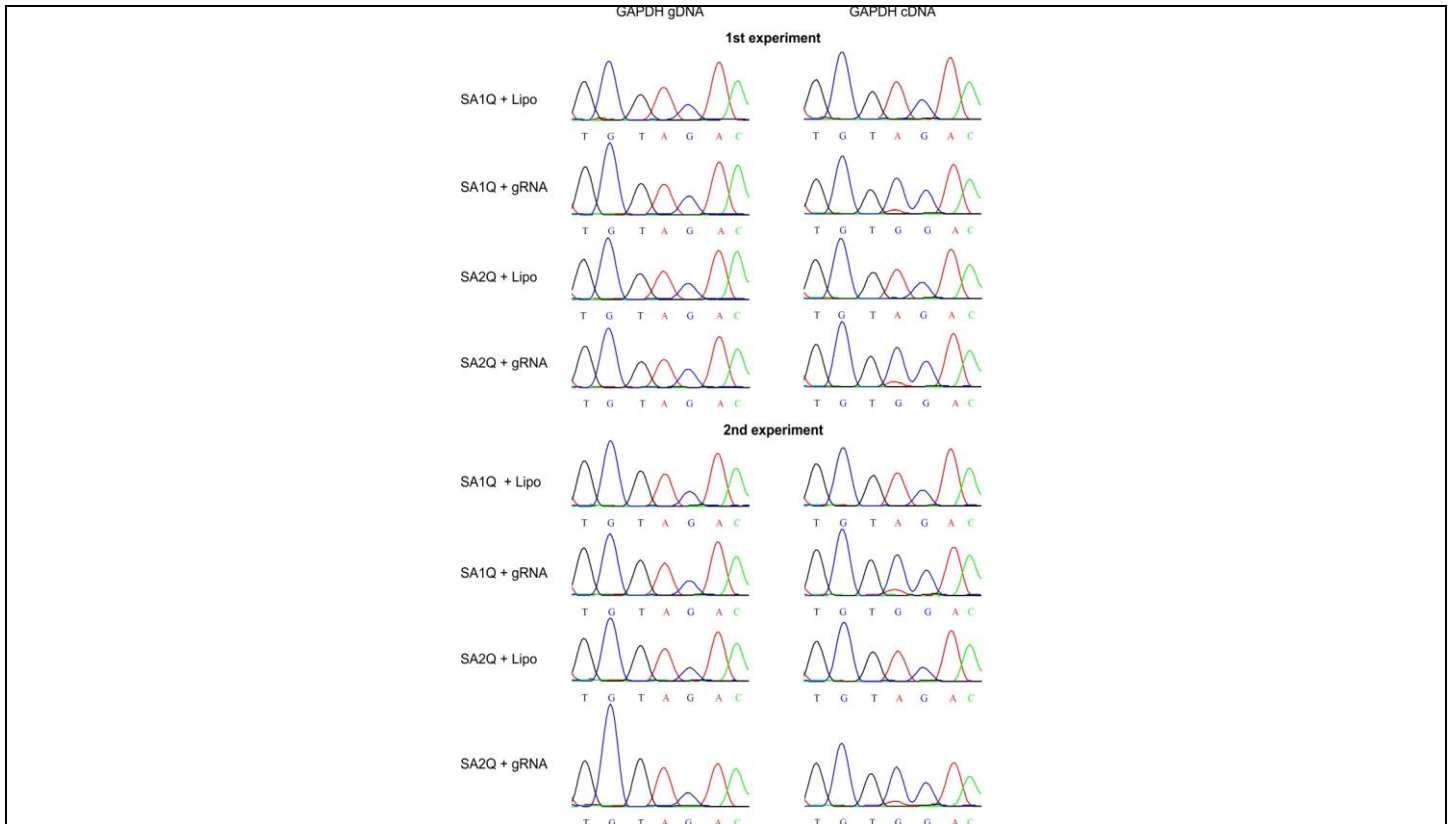
The respective double mutant (SA2QG) was genomically integrated into Flip-In cells analog as described for the other four enzymes (SA1, SA2, SA1Q, SA2Q). We then studied the editing of 6 different sites in the GAPDH transcript entirely analog as described. Interestingly, SA2QG was only active in the 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).

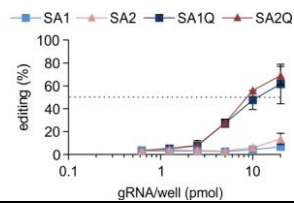
Overexpressed Cas13-guideRNAs can also recruit human ADAR2 or SA2Q to elicit editing in co-transfected reporter transcripts to yields similar as Cas13b-ADAR repairV1 does. For further details and discussion, see Supplementary Note 2. Data are shown as the mean±SD, N=3 independent experiments, black dots represent individual data points. DR = Cas13 directing domain



Supplementary Figure 13

Editing control gDNA versus cDNA .

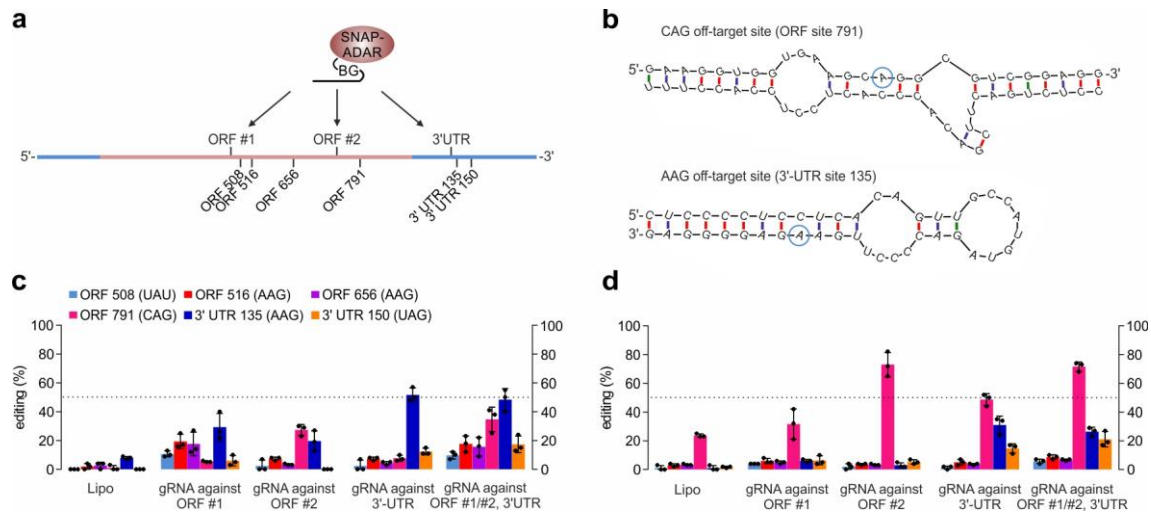
To ensure that editing occurred only at the transcript and not the genomic DNA (gDNA), sequencing traces of gDNA and cDNA derived from mRNA were compared after site-directed editing in the 3'-UTR of GAPDH. Only the cDNA traces showed an A-to-G change for SA1Q and SA2Q, indicating that both enzymes target 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±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'- AG GAGGUGAU AG CAU UUG -3' |
| AAGAB | 5'- CC AGG U UGAU AG CAU 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 |

| | | |
|--|---|--|
| | <p>2) clearly proven, covalent RNA targeting</p> <p>3) very short guideRNA/mRNA duplex, unlikely to interfere with endogenous ADARs or translation</p> <p>4) simple co-transfection of guideRNAs enables concurrent editing</p> | <p>overexpression, in particular with hyperactive ADAR deaminases, low editing yields with wildtype or less active ADAR domains like version2)</p> |
|--|---|--|

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.

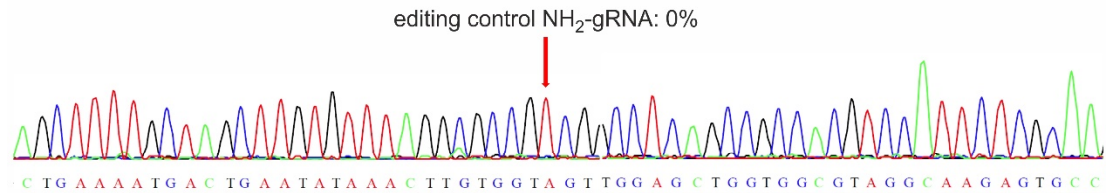
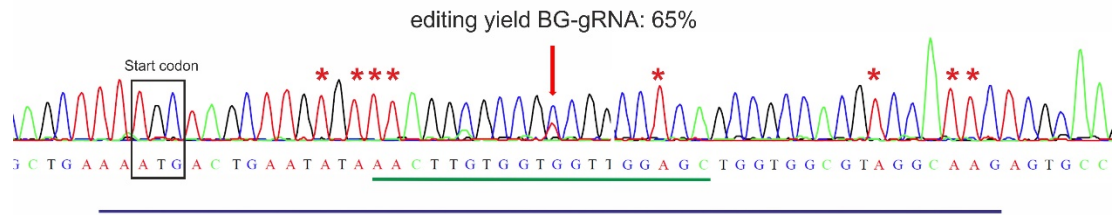
| target | gRNA sequence | applied gRNA amount ^{a)} |
|---|--|-----------------------------------|
| editing of various endogenous transcripts | | |
| 5'-UTR SNAP-ADAR | 5'-UsCsAUUAAACG CCA GAGUCsCsGsGsA-3' | 5 pmol |
| 5'-UTR GAPDH isoform 2 | 5'-UsCsUGAAUAAU CCA GGAAAsAsGsCsA-3' | 5 pmol |
| ORF #1 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 UACCsUsCsCsC-3' | 5 pmol |
| 3'-UTR GUSB | 5'-UsAsUUUCCUG CCA GAAUsGsAsUsG-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 AAGGsGsUsCsA-3' | 5 pmol |
| 5'-CAC | 5'-AsAsCGCCAGGG GCG CUAAGsCsAsGsU-3' | 5 pmol |
| 5'-CAG | 5'-UsAsCGCAUGGA CCG UGUUCsAsUsGsA-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 UACCAAsGsAsGsU-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)

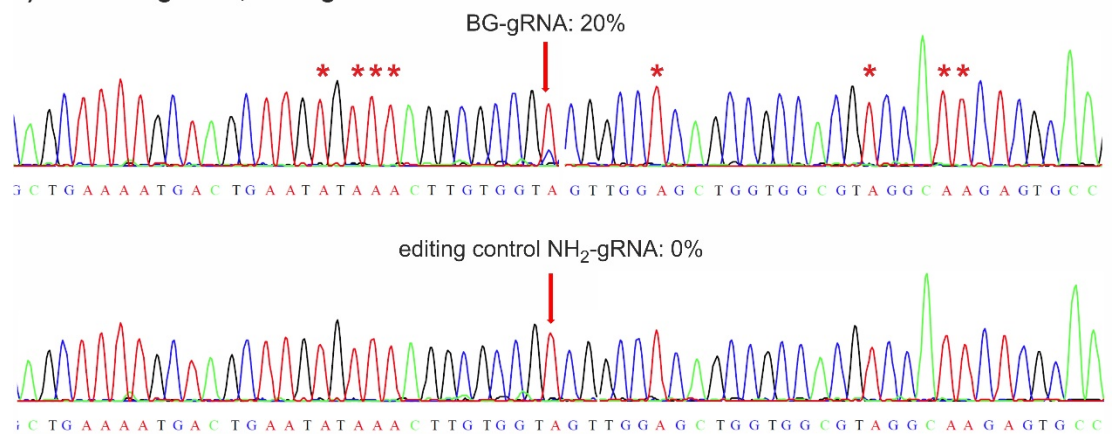
Editing of two sites in endogenous KRAS as previously reported by Cox et al. with Cas13b-ADAR

a) KRAS Target 1/A, editing with SA1Q

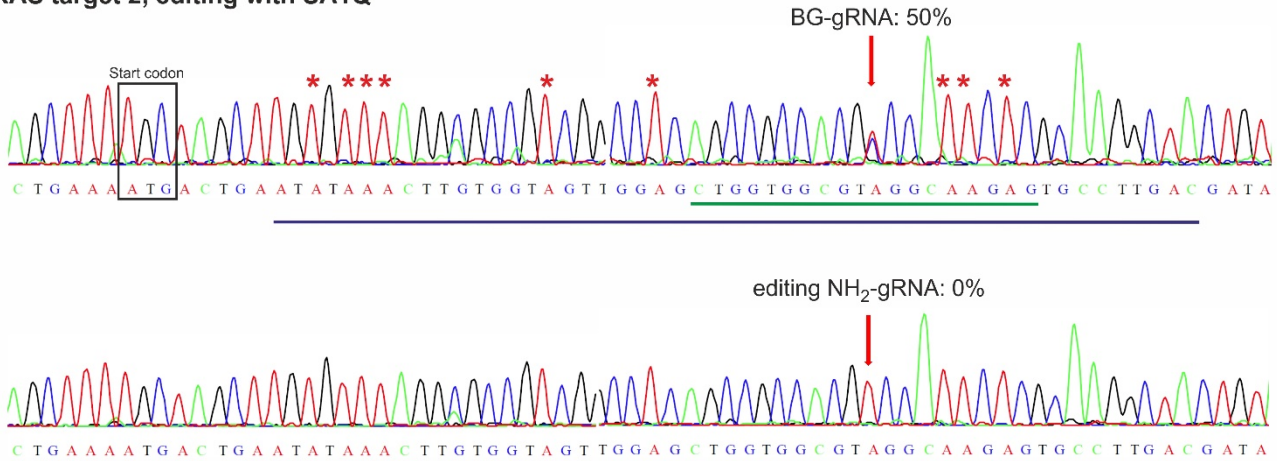
mRNA/gRNA duplex of the BG guideRNA (19 bp)
mRNA/gRNA duplex of the Cas13 guideRNA (50 bp)
* off-target sites for Cas13-ADAR



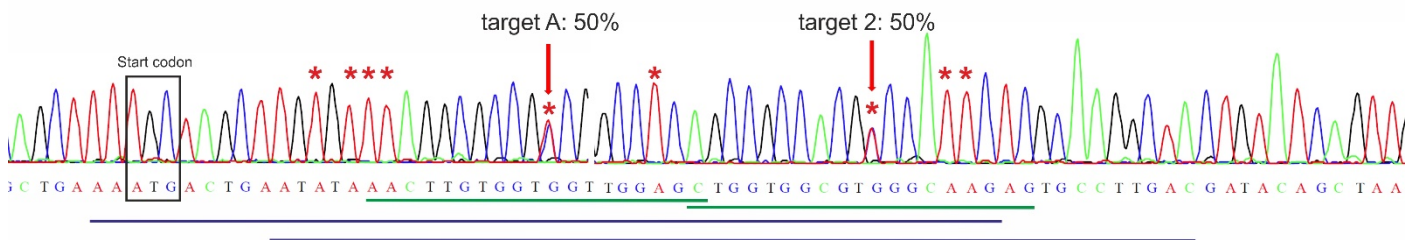
b) KRAS Target 1/A, editing with wt SA1



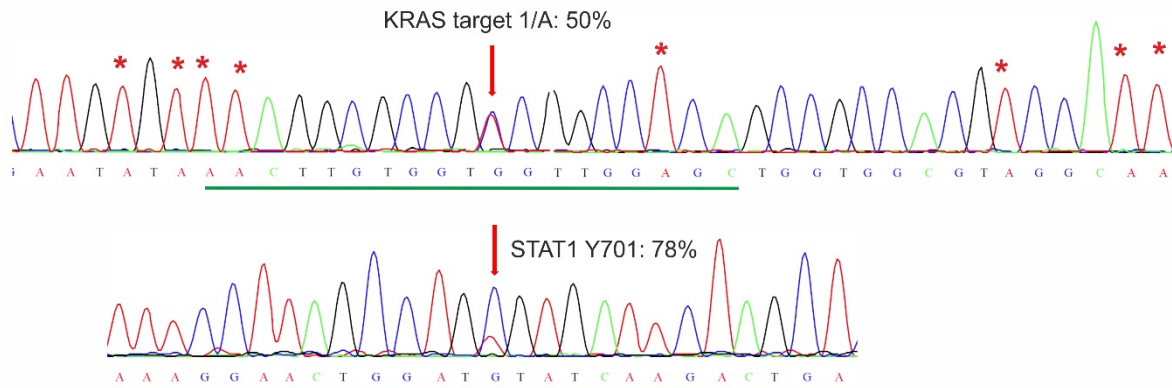
c) KRAS target 2, editing with SA1Q



d) KRAS concurrent editing target1/A and target 2 with SA1Q / BG-guideRNAs



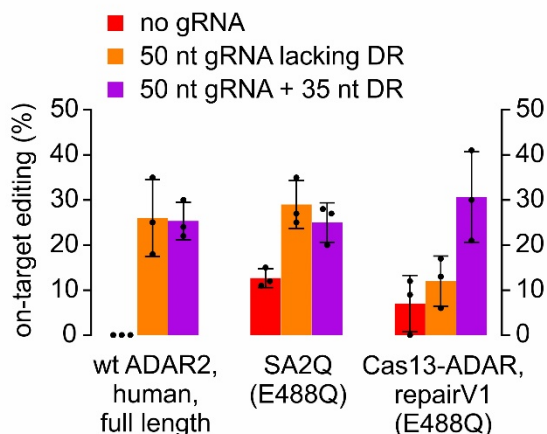
e) Concurrent editing KRAS target 1/A and STAT1 Tyr701 with SA1Q



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.

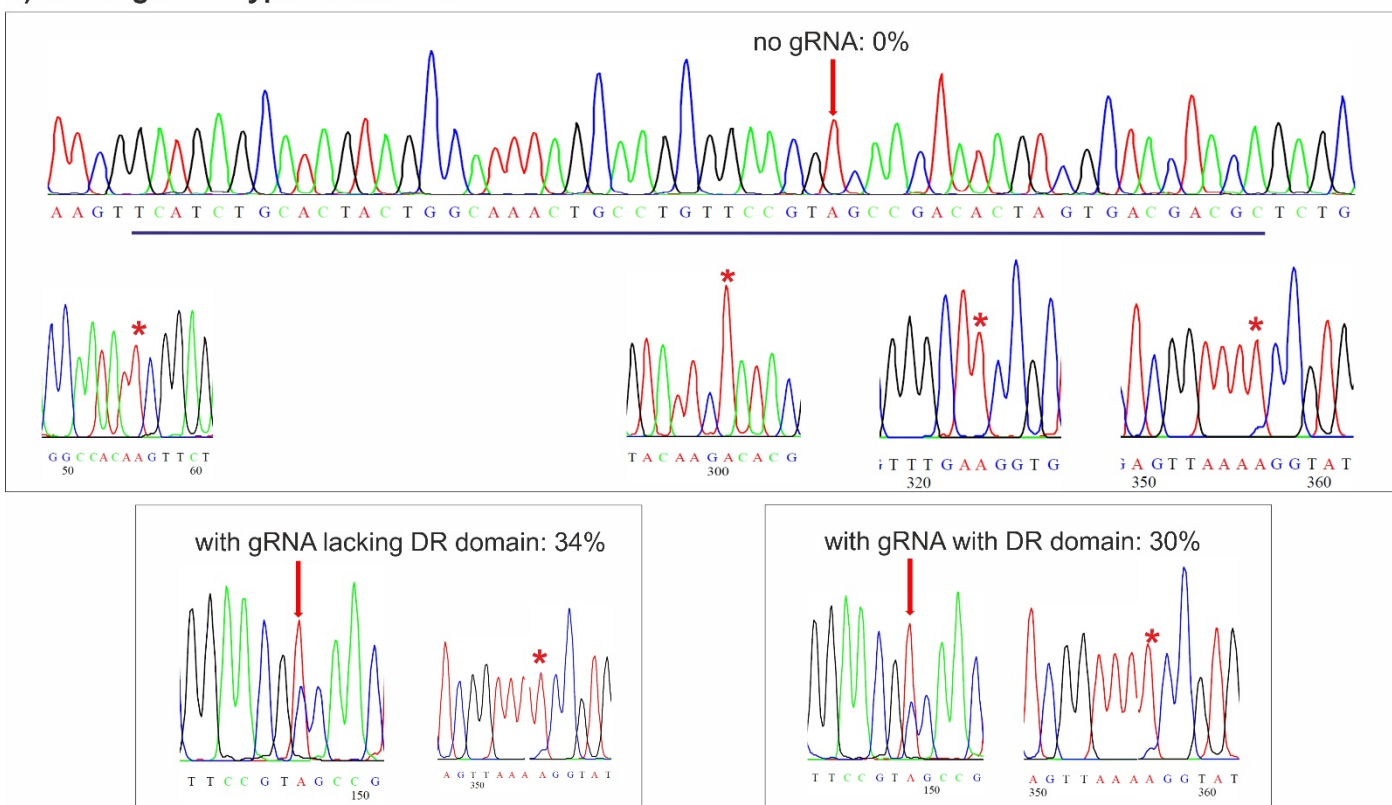
Editing of overexpressed GFP reporter W58X with co-overexpressed Cas13-guideRNAs and co-overexpressed, different ADAR fusions (SA2Q, human full length ADAR2, and Cas13-ADAR version1) following exactly the protocol given by Cox et al.

a) Overview of n=3 experiments

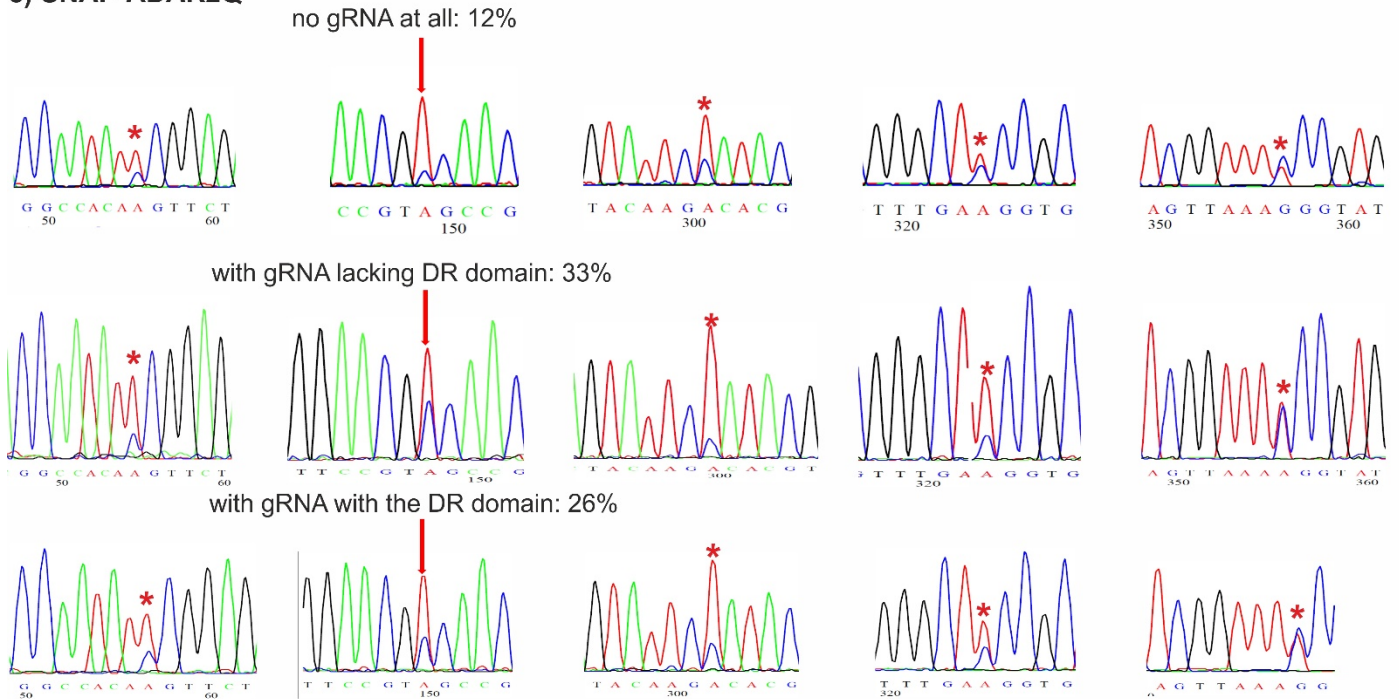


Selected Sanger sequencing traces (selected was always the trace with the highest on-target yield out of 3 experiments)

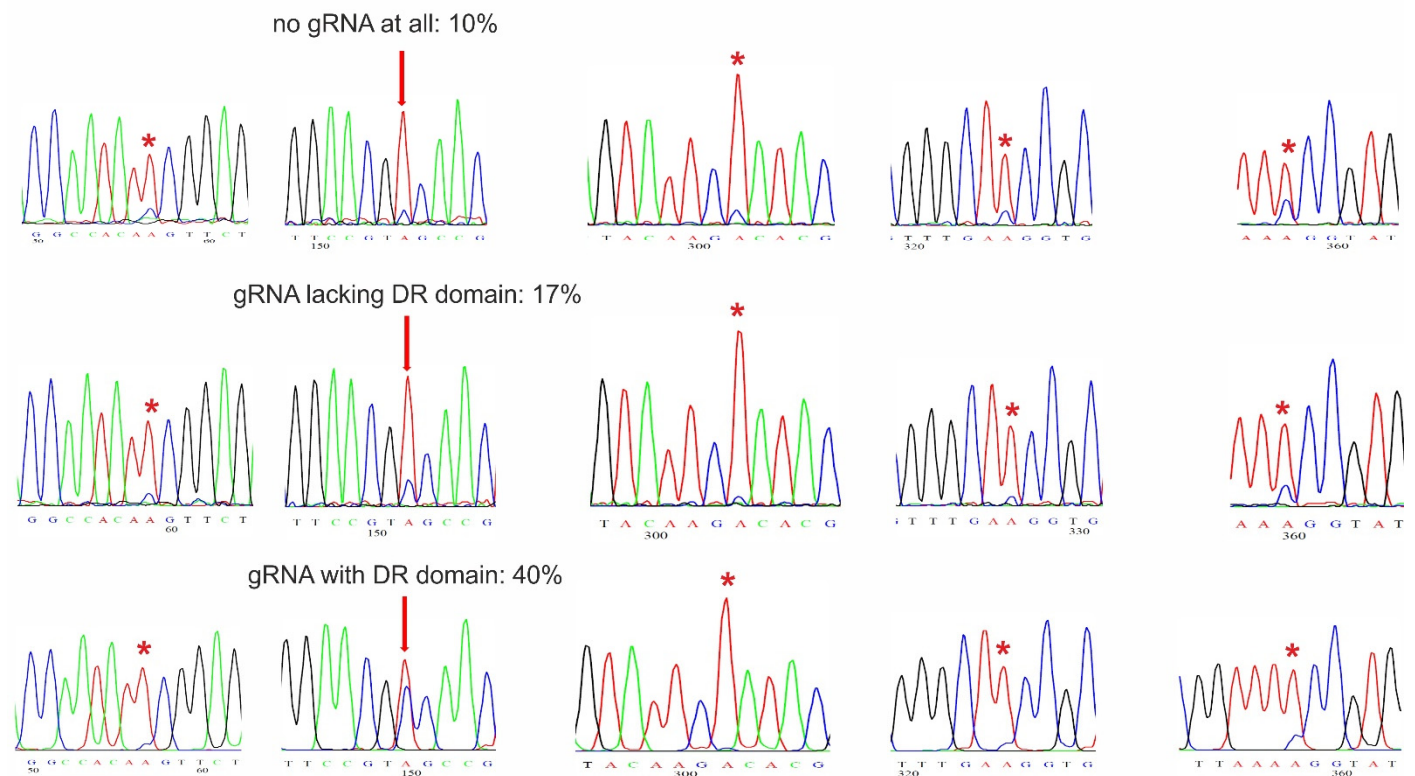
b) full length wildtype human ADAR2



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).

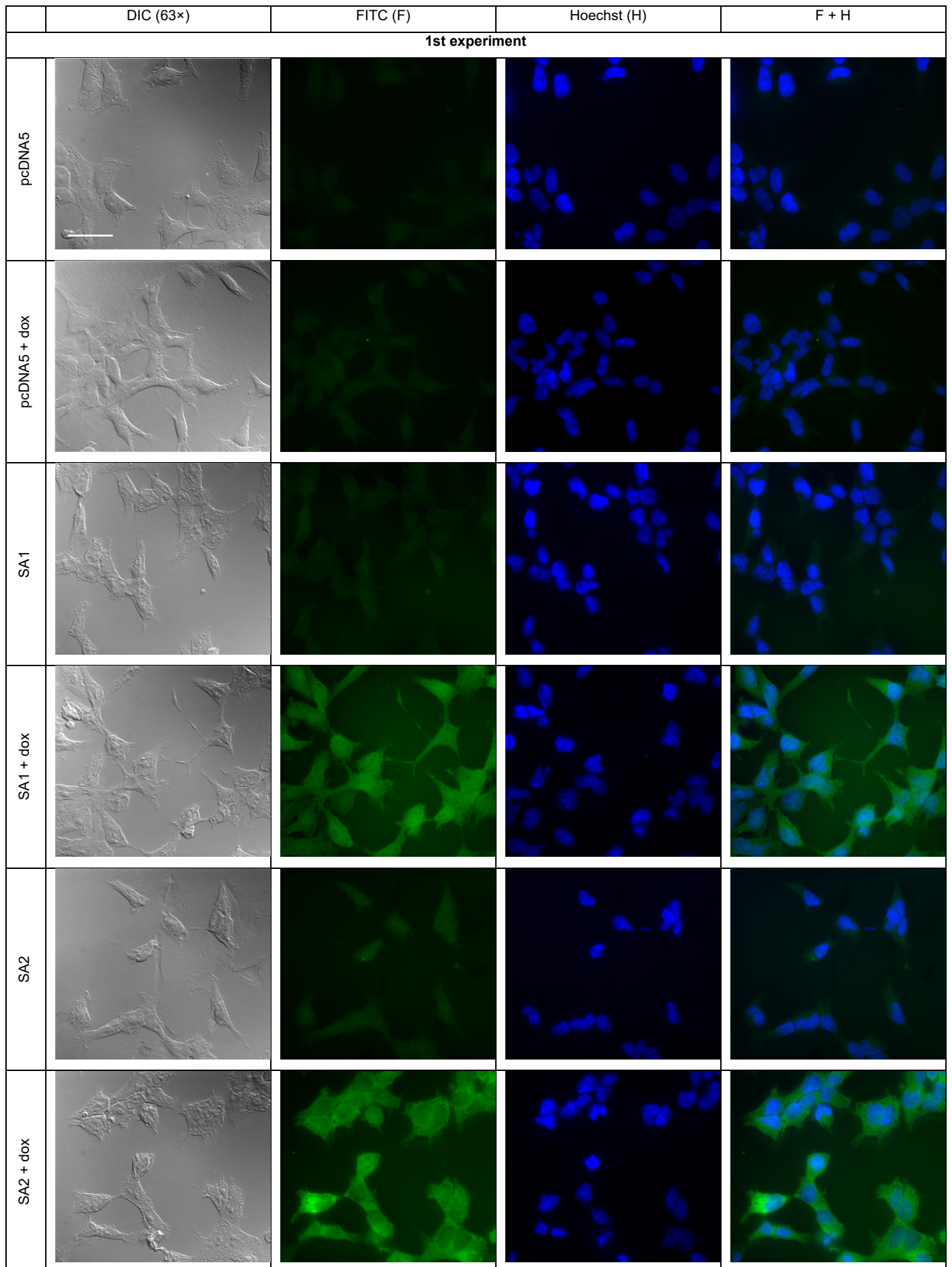
For this, we designed a Cas13 guideRNA according to Cox et al. containing a 50 nt part antisense to our GFP reporter (W58amber), putting the targeted A into mismatch with C. Mismatch position was 34. We constructed guideRNAs with the 3'-terminal DR hairpin for Cas-targeting but also lacking the DR motif (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):GCGTCACTAGTGTCGGCCACGGAACAGGCAGTTTGCCAGTAGTGCAGATGAgttgtggaaggtccagtttgaggggctattacaac. 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.

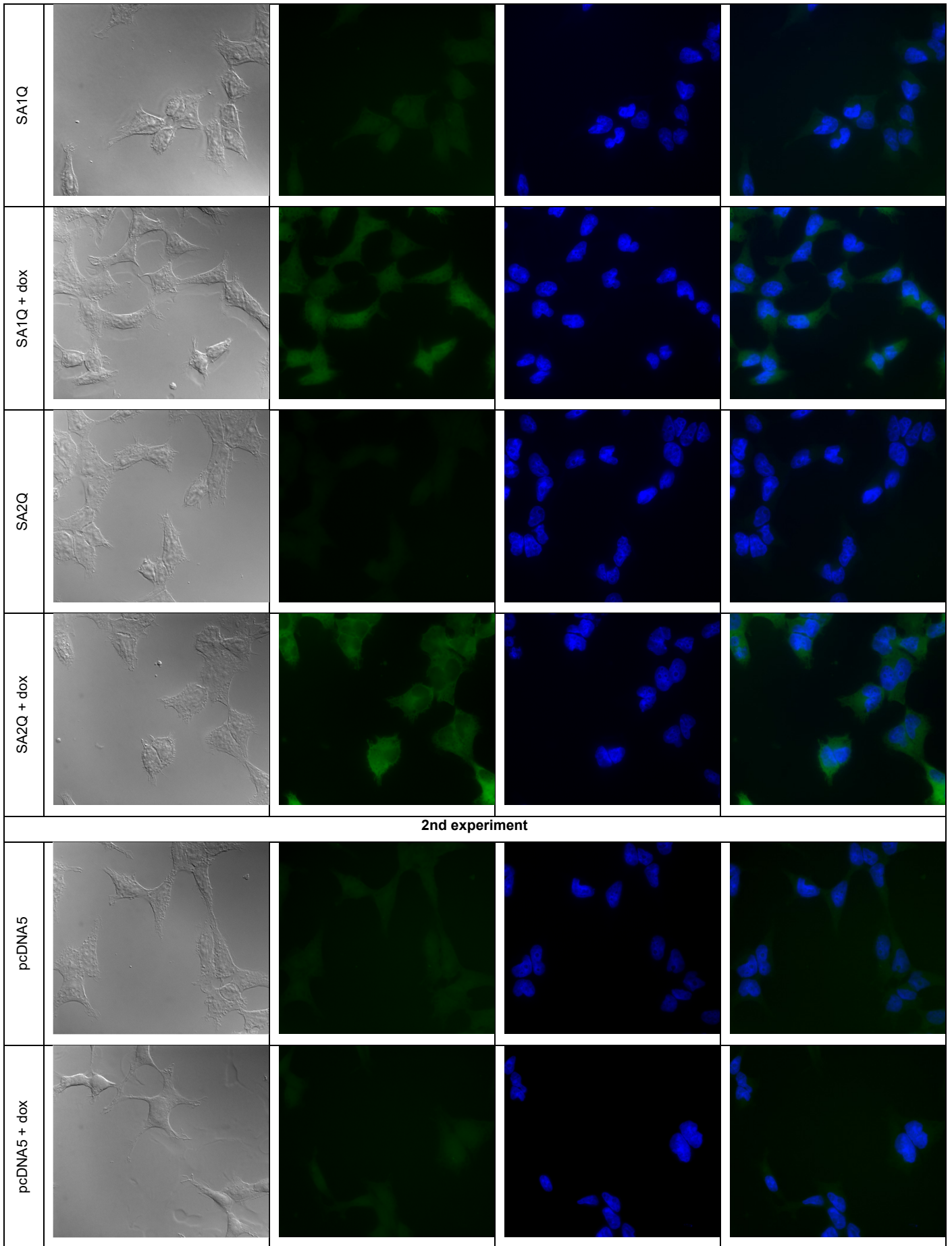
a) shows that the Cas13-guideRNA can also recruit human ADAR2 or SNAP-ADAR2Q to elicit editing yields similar to Cas13-ADAR. The average editing levels (25-30%) are very similar to those described by Cox et al. for various similar overexpression / reporter experiments in their Figures 2-4 (15-30%). As 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.

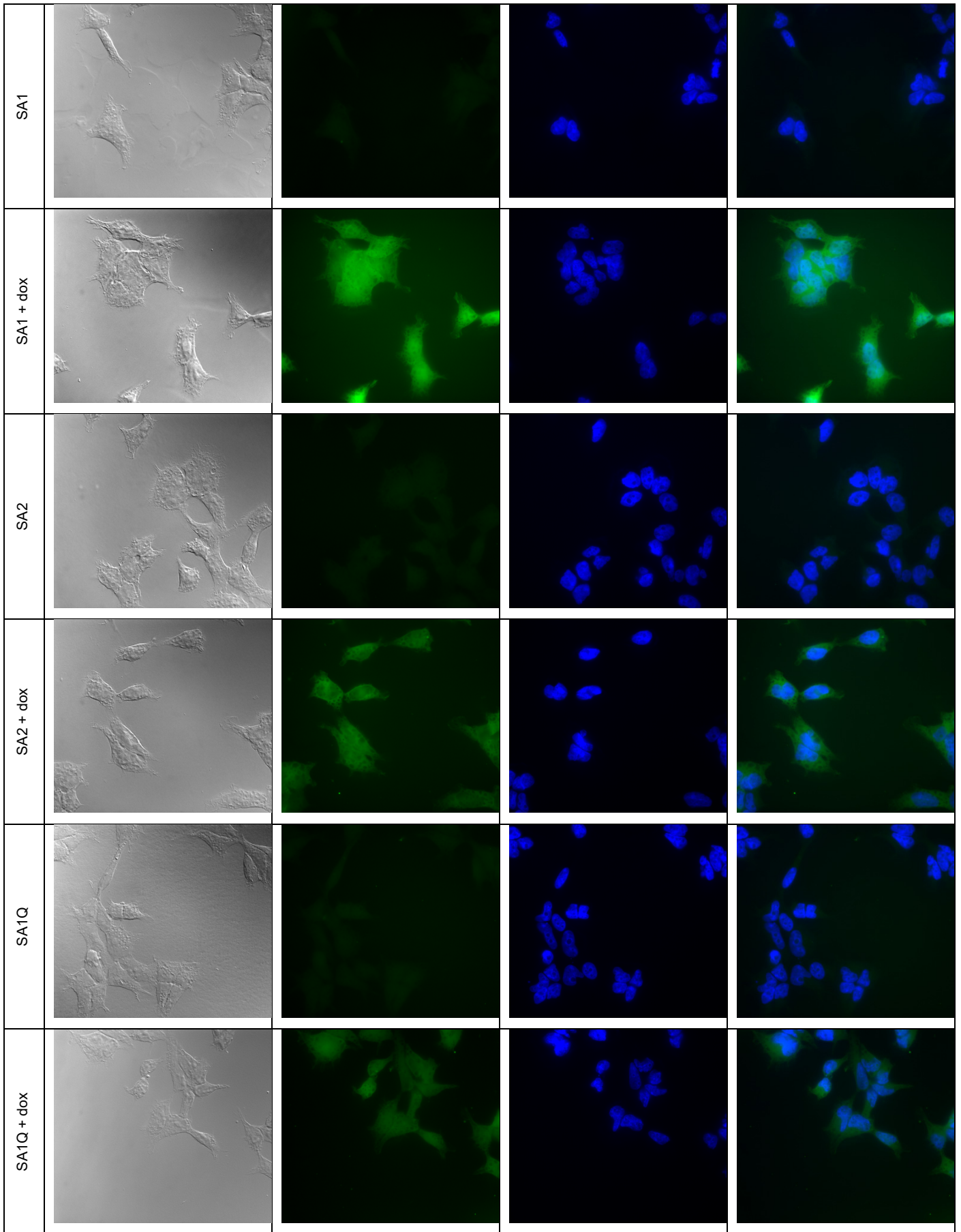
b-d) show selected Sanger sequencing traces (always the trace with the highest on-target editing yield was chosen) to give an idea of off-target editing. While ADAR2 (**b**) gives decent on-target editing (25%) there was only very little off-target editing seen and on-target editing was fully dependent 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.

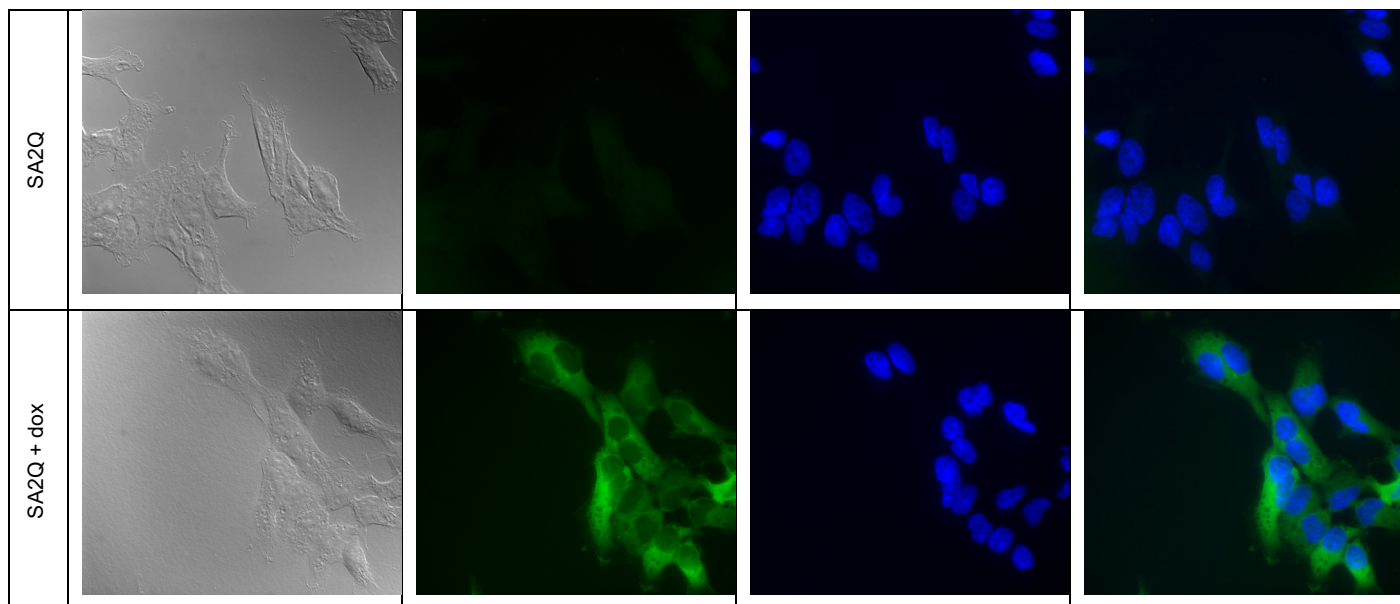
Together, panels **a-d**) suggest that the conditions (overexpression & reporters) under which Cas13-ADAR has mostly been characterized today are not sufficient to support the general claims made by Cox et al.

Determination of intracellular SNAP-ADAR localization by fluorescence microscopy





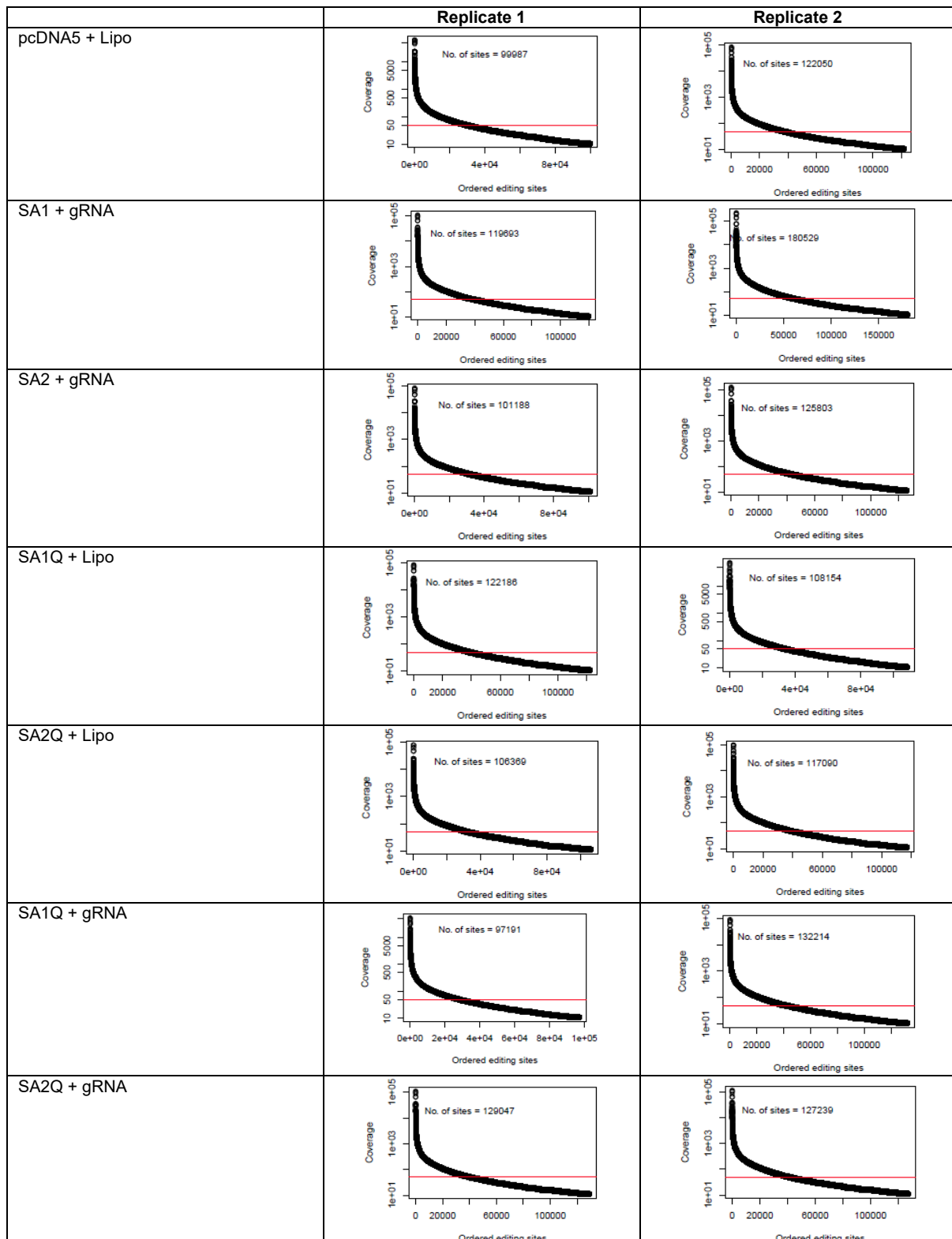




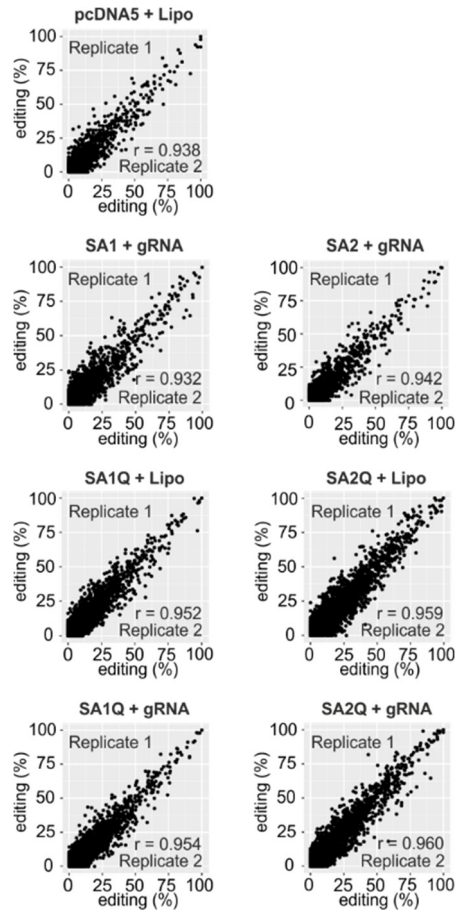
Supplementary Note 3. Protein expression was induced by doxycycline (dox) for 24 h. Cells were incubated with BG-FITC to stain SNAP-ADARs (green) and with Hoechst 33342 to stain nuclei (blue). Microscopy was performed with a Zeiss CellObserverZ1 under 630x total magnification. The scale bar represents 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)

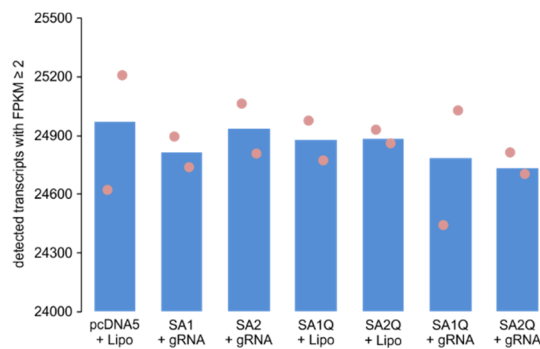
Additional NGS quality data



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     GCTGCTTTTAACTCTGGTAAAAGTGGATTATTGTTGCCATCAATGACCCCTCATTGACCTC
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    AACTACATGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATCCATGGCACCGTC
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    AAGGCTGAGAACGGGAAGCTTGTCAATGGAAATCCCATCACCATCTCCAGGAGCGA
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    GTCTTCACCACCATGGAGAAGGCTGGGGCTCATTGTCAGGGGGGAGCCAAAGGGTCATC
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    GACAACAGCCTCAAGATCATCAGCAATGCCTCCTGCACCACCAACTGCTTAGCACCCCTG
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    GCCAAGGTCAATCCATGACAACCTTTGGTATCGTGGAAGACTCATGACCACAGTCCATGCC
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    ATCACTGCCACCCAGAAGACTGTGGATGGCCCCCTCCGGGAAACTGTGGCGTGATGGCCCGC
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    GGGGCTCTCCAGAACATCATCCCTGCCTCTACTTGGCGCTGCCAAGGCTGTGGGCAAGGTC
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    TCAGTGGTGGACCTGACCTGCCGTCTAGAAAAACCTGCCAAATATGATGACATCAAGAAG
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    GTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGCATCCTGGGCTACACTGAGCACCAG
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    ATTGCCCTCAACGACCACCTTGTCAAGCTCATTTCCTGGTATGACAACGAATTTGGCTAC
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.

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
GGTCAACGCCTACTTTCCACGCTGAGGCCATCGAGGAGTTCCTGTGCCAGCCCTGC
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
TGGTGAAGTTCGGAGAGTTCATCAGCTACAGCCACCTGGCCGCCCTGGCCGGCAATCCCG
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
AAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTCTGCAG
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
GAGAAGGCACAATCCCTGTGGAATCCAGTGACATGTGCTACGTGGGATGGCATTCGGC
440 G E 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
GCCTGCAAGGGGCACTGTTGACCCACTTCTGCAGCCATTTATCTCAAATCTGTACAT
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 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 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 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).

1 10 20 30 40 50 60
1 GGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAAGACACCGGGACCGATCCAGCCTCC
1
61 70 80 90 100 110 120
20 GGACTCTAGCGTTTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGGACAAAGACT
M D K D
121 130 140 150 160 170 180
40 GCGAAATGAAGCGCACCACCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCG
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
60 AACAGGGCTGCACCGTATCATCTTCTGGGCAAGGAACATCTGCCGCCGACGCCGTGG
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
80 AAGTGCCTGCCCCAGCCGCGTGTGGGCGGACCAGAGCCACTGATGCAGGCCACCGCCT
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
100 GGCTCAACGCCTACTTTTACCAGCCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
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
120 ACCACCCAGTGTTCACGAGGAGAGCTTTACCCGCCAGGTGCTGTGGAAGTGTGAAAG
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
140 TGGTGAAGTTCGAGAGGTCATCAGCTACAGCCACCTGGCCGCTGGCCGCAATCCCG
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
160 CCGCCACCGCCGCGTGAACCCGCTGAGCGGAAATCCCGTGCCCATCTGATCCCT
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
180 GCCACCGGGTGGTGCAGGCGCACCTGGACGTGGGGGCTACGAGGGCGGGCTCGCCGTGA
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
200 AAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGTCTCGCAG
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
220 GCGGAGGCGCCAGGGTCTGGCGGGCAGTAAGAAGCTTGCCAAGGCCCGGGCTGCGC
G G G A P G S G G G S K K L A K A R A A
721 730 740 750 760 770 780
240 AGTCTGCCCTGGCCGCAATTTTTAACTTGCCTGGATCAGACGCCATCTCGCCAGCCTA
Q S A L A A I F N L H L D Q T P S R Q P
781 790 800 810 820 830 840
260 TTCCAGTGAGGGTCTTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCAGCC
I P S E G L Q L H L P Q V L A D A V S R
841 850 860 870 880 890 900
280 TGGTCTGGGTAAGTTTGGTGACCTGACCGACAACCTTCTCCTCCCTCACGCTCGCAGAA
L V L G K F G D L T D N F S S P H A R R
901 910 920 930 940 950 960
300 AAGTGCTGGTGGAGTCGTATGACAACAGGCACAGATGTTAAAGATGCCAAGGTGATAA
K V L A G V M T T G T D V K D A K V I
961 970 980 990 1000 1010 1020
320 GTGTTTCTACAGGAACAAATGTATTAATGGTGAATACATGAGTATCGTGGCCTGTCAT
S V S T G T K C I N G E Y M S D R G L A
1021 1030 1040 1050 1060 1070 1080
340 TAAATGACTGCCATGCAGAAATAATATCTCGGAGATCCTTGCTCAGATTTCTTTATACAC
L N D C H A E I I S R R S L L R F L Y T
1081 1090 1100 1110 1120 1130 1140
360 AACTTGAGCTTTACTTAAATAACAAGATGATCAAAAAAGATCCATCTTCAGAAATCAG
Q L E L Y L N N K D D Q K R S I F Q K S
1141 1150 1160 1170 1180 1190 1200
380 AGCGAGGGGGTTTAGGCTGAAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTC
E R G G F R L K E N V Q F H L Y I S T S
1201 1210 1220 1230 1240 1250 1260
400 CCTGTGGAGATGCCAGAATCTTCTCACCACATGAGCCAATCCTGGAAGAACCAGCAGATA
P C G D A R I F S P H E P I L E E P A D
1261 1270 1280 1290 1300 1310 1320
420 GACACCCAAATCGTAAAGCAAGAGGACAGCTACGGACCAAAATAGAGTCTGGTGAGGGGA
R H P N R K A R G Q L R T K I E S G E G
1321 1330 1340 1350 1360 1370 1380
440 CGATTCCAGTGCCTCCAAATGCGAGCATCCAACGTGGGACGGGTGCTGCAAGGGGAGC
T I P V R S N A S I Q T W D G V L Q G E
1381 1390 1400 1410 1420 1430 1440
460 GGCTGCTCACCATGCTCTGCAGTACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGG
R L L T M S C S D K I A R W N V V G I Q
1441 1450 1460 1470 1480 1490 1500
480 GATCCCTGCTCAGCATTTTCTGGAGCCCATTTACTTCTCGAGCATCATCCTGGGCAGCC
G S L L S I F V E P I Y F S S I I L G S
1501 1510 1520 1530 1540 1550 1560
500 TTTACCACGGGGACCACCTTTCCAGGGCCATGTACCAGCGGATCTCCAACATAGAGGACC
L Y H G D H L S R A M Y Q R I S N I E D
1561 1570 1580 1590 1600 1610 1620
520 TGCCACCTCTACACCCTCAACAAGCCTTTGCTCAGTGGCATCAGCAATGCAGAAGCAC
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 AGGTCATCAACGCCACGACTGGGAAGGATGAGCTGGGCCGCGCTCCCGCTGTGTAAGC
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 ACGCGTTGTACTGTGCGTGGATGCGTGTGCACGGCAAGGTTCCCTCCCCTTACTACGCT
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 CCGCCAAGGCGCGTCTGTTCACAGCCTTCATCAAGCGGGGCTGGGGGCTGGGTGGAGA
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 AGCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCTAGATGACTGCCTGTTCCG TAGC
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
AACAGGGCTGCACCGTATCATCTTCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGG
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
GGCTCAACGCCTACTTTCCACGAGCTGAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGC
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
TAACCCAGTGACAGGGGCCAGTCTCAGAAACTATGCTCCTCCTCAAGGTCCCCAG
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
GCCACTACCTGTCTTCGAGAATCCCAAACAAGGAAAGCTCCGCACCAAGGTGGAGAACG
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
GCCTGCAAGGGGCACTGTTGACCCACTTCTGCAGCCATTTATCTCAAATCTGTACAT
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 AGCTCTGCTCCTCCGTTACCGCAGGGATCTACTGAGACTCTCCTATGGTGAGGCAAGA
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 GAGGAAAAGAACTTTTATCTCTGCCAGTAT
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 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 CCGCCACCGCCCGCTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGATCCCT
 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 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 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 TGGTCTGGGTAAGTTTGGTGACCTGACCGACAACCTTCTCCTCCCTCAGCTCGCAGAA
 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 GTGTTTCTACAGGAACAAATGTATTAATGGTGAATACATGAGTGTGCGTGGCCTGCAT
 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 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 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 GGCTGCTCACCATGTCCTGCAGTACAAGATTGCACGCTGGAACGTGGTGGGCATCCAGG
 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   CCGCCAAGGCGCTCTGTTCACAGCCTTCATCAAGCGGGGCTGGGGGCTGGGTGGAGA
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   AGCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCTAGATGACTGCCTGTTCCG TAGC
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-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 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 AACAGGGCTGCACCGTATCATCTTCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGG
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 CCGCCACCGCCCGCTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCATCTGATCCCCCT
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 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 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 TTCCAGTGAGGGTCTTCAGCTGCATTTACCGCAGGTTTTAGCTGACGCTGTCTCAGGCC
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 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 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 CGATTCCAGTGCCTCCAAATGCGAGCATCCAACGTGGGACGGGGTGTGCAAGGGGAGC
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   ACGCGTTGTACTGTGCTGGATGCGTGTGCACGGCAAGGTTCCCTCCCACTTACTACGCT
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   AGCCACCGAGCAGGACCAGTTCTCACTCACGCCCTCTAGATGACTGCCTGTTCCG TAGC
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
ATTATTCTCTGATTGGTTCGTATTGGGCGCCTGGTCACCAGGGCTGCTTTTAACTCTGGT

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
CTTGTCAATGAAATCCCATCACCATCTTCCAGGAGCGAGATCCCTCCAAAATCAAG
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
TGGGGCGATGCTGGCGCTGAGTACGTCGTTGGAGTCCACTGGCGTCTTACCACCATGGAG
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
ATCAGCAATGCCTCTGTCACCACCAACTGCTTAGCACCCCTGGCCAAGGTCATCCATGAC
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
ACTGTGGATGGCCCTCCGGGAAACTGTGGCGTATGGCCGGGGGCTCTCCAGAACATC
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
AAGCTCACTGGCATGGCCTTCCGTGTCCCACTGCCAACGTGTGAGTGGTGGACCTGACC
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
GAGGGCCCCCTCAAGGGCATCCTGGGCTACACTGAGCACCAGGTGGTCTCCTCTGACTTC
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
GGGAGCCGCACCTTGTATGTACCATCAATAAAGTACCCTGTGCTCAACCAGTTAAAAAA
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     TCACCAGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTTGCCATCAATGACCCCTTCA
        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    ATGGCCGCGGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGG
        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    GCAAGGTCACTCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCGTGTCCTCCACTG
        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    CCAACGTGTCAGTGGTGGACCTGACCTGCCGCTCTAGAAAAACCTGCCAAATATGATGACA
        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    TCAAGAAGGTGGTGAAGCAGGCGTTCGGAGGGCCCCCTCAAGGGCATCCTGGGCTACACTG
        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    GGGCTGGCATTGCCCTCAACGACCACTTTGTCAAGCTCATTTCTGGTATGACAACGAAT
        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    GACCCCTTGAAGAGGGGAGGGCCCTAGGGAGCCGCACCTTGTTCATGTACCATCAATAAAG
1381  1390    1400    1410    1420
460    TACCCTGTGCTCAACCAGTTAAAAAAAAAAAAAAAAAAAAA

```

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 GAGCTGCGTGTGGCTCCCGAGGAGCACCCCGTGTGCTGCTGACCGAGGCCCCCTGAACCCC
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 AAGGCCAACCGCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCGCCATG
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 ATGGACTCCGGTGACGGGTCACCCACACTGTGCCCATCTACGAGGGTATGCCCTCCCC
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 GAGCGGTTCCGCTGCCCTGAGGCACTTCCAGCCTTCCTTCCTGGGCATGGAGTCCCTGT
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 CTGTACGCCAACACAGTGTGCTGTGCGCGCACCACCATGTACCCTGGCATGCGGACAGG
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 TTTTTTTTTTTTTTTTGGCTTGACTCAGGATTTAAAACTGGAACGGTGAAGGTGACAGC
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 CATCCTAAAAGCCACCCACTTCTCTCTAAGGAGAATGGCCAGTCTCTCCCAAGTCCA
501
1570 1580 1590 1600 1610 1620
1561 CACAGGGAGGTGATAGCATTGCTTTCGTGTAATTTATGTAATGCAAAATTTTTTAATC
521
1630 1640 1650 1660 1670 1680

```

1621   TTCGCCTTAATACTTTTTTATTTTGTGTTTTATTTTGAATGATGAGCCTTCGTGCCCCCCT
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 GTCCCTCAACCAAGATGGCGCGGATGGCTTCAGGCGCATCACGACACCCGGCGCGTCACGCG
1
70 80 90 100 110 120
61 ACCCGCCCTACGGGCACCTCCCGCGCTTTTCTTAGCGCCGAGACGGTGGCCGAGCGGGG
20
130 140 150 160 170 180
121 GACCGGGAAGCATGGCCCCGGGGTTCGGCGGTTGCCTGGGCGCGCTCGGGCCGTTGTTGT
40 M A R G S A V A W A A L G P L L
190 200 210 220 230 240
181 GGGGCTGCGCGCTGGGGCTGAGGGCGGGATGCTGTACCCCGAGGAGCCCGTGGCGGG
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 GGGGCTTCGAGGAGCAGTGGTACCAGCGCGCGCTGTGGGAGTCAGGCCACCGTGGACA
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 GCTGGGTGTGGTACGAACGGGAGTGATCCTGCCGGAGCGATGGACCCAGGACCTGGCGCA
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 AGTTGGAAGTGCCTTTTTGGATGCAGAAAACAAAGTCGTGGCGAATGGGACTGGGACCC
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 CTGCCTATCTGTATTTCATTTGGAGGTGCAGCTGACTGCACAGACGCTACTGGGCGCTGTGT
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 CTGACTTCTACACTCCCTGTGGGATCCGCACTGTGGCTGTCAACAGAGCCAGTTCC
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 TCATCAATGGGAACCTTTCTATTTCCACGGTGTCAACAAGCATGAGGATGGGGACATCC
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 GAGGGAAGGGCTTCGACTGGCCGCTGCTGGTGAAGGACTCAACCTGCTTCGCTGGCTTG
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 ACCGCTATGGGATTTGTGGTCATCGATGAGTGTCCCGCGTGGGCCTGGCGCTGCCGAGT
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 ACAAGAACCACCCCGGCTGTGATGTGGTCTGTGGCCAACGAGCCTGCGTCCCACC 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 GGCCTGTGACCTTTGTGAGCAACTTAACATATGCAGCAGACAAGGGGCTCCGTATGTGG
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 TTCAGAGCGAGTATGGAGCAGAAACGATTGCAGGTTTCACCAGGATCCACCTCTGATGT
580 I Q S E Y G A E T I A G F H Q D P P L M
      1810      1820      1830      1840      1850      1860
1801 TCACTGAAGAGTACCAGAAAAGTCTGCTAGAGCAGTACCATCTGGGTCTGGATCAAAAAC
600 F T E E Y Q K S L L E Q Y H L G L D Q K
      1870      1880      1890      1900      1910      1920
1861 GCAGAAAATACGTGGTTGGAGAGCTCATTGGAATTTTGCCGATTCATGACTGAACAGT
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 CACCGACGAGAGTCTGGGGAATAAAAAGGGGATCTTCACTCGGCAGAGACAACCAAAAA
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 ACTCAGTAGCCAAGTCACAATGTTTGGAAAACAGCCTGTTTACTTGAGCAAGACTGATAC
680 H S V A K S Q C L E N S L F T *
      2110      2120      2130      2140      2150      2160
2101 CACCTGCGTGTCCCTTCCTCCCCGAGTCAGGGCGACTTCCACAGCAGCAGAACAAGTGCC
700
      2170      2180      2190      2200      2210      2220
2161 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 AATGAGGGACCAGTACATGAGGACTGGGGAG
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    GGCTTTCTTTGTGTTTGGCATAAATAACTAAATCATTGGAAGATATTCACCATTAT
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    TAAGGCATACTAGTACAAGTGGTAATTTTGTACATTACACTAAATTATTAGCATTGTGTT
261

```

Sequence of KRAS mRNA (NM_004985.4) with chosen editing sites (yellow).

10 20 30 40 50 60
1 GCTGAGCGCGGAGCCGCCGGTGTATTGGTGGGGGCGGAAGGGGGCCGGGCGCCAGCGCTG
1
70 80 90 100 110 120
61 CCTTTTCTCCTGCCGGTAGTTTCGCTTTCTGCGCAGAGTCTGCGGAGGGGCTCGGCTG
21
130 140 150 160 170 180
121 CACCGGGGGATCGCGCCTGGCAGACCCAGACCGAGCAGAGGCGACCCAGCGCTCGG
41
190 200 210 220 230 240
181 GAGAGGCTGCACCGCCGCCCGCCCTAGCCCTTCCGGATCCTGCGCGCAGAAAAGTTT
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 GGCCCTTGTGGGGCACAAGGTGGCAGGATGTCTCAGTGGTACGAACTTCAGCAGCTTGA
121 M S Q W Y E L Q Q L D
421 CTCAAAATTCCTGGAGCAGGTTCCACCAGCTTTATGATGACAGTTTCCCATGGAATCAG
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 GAGCACAGTGTGTAGACAAAACAGAAAGAGCTTGACAGTAAAGTCAGAAATGTGAAGGA
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 TCAGAAAACAAGAAGCCTGTTACTCAAGAAGATGTATTTAATGCTTGACAATAAGAGAAA
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 GATTAATGATGAAGTAGTGGAGTGAAGCGGAGACAGCAGAGCGCTGTATTGGGGGCC
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 TGACCCTATCACAAAAACAAGTGTATGGGACCGCACCTTCAGTCTTTCCAGCA
421 D P I T K N K Q V L W D R T F S L F Q Q
1321 GCTCATTACAGAGCTCGTTTGTGGTGAAGACAGCCCTGCATGCCAACGCACCCCTCAGAG
441 L I Q S S F V V E R Q P C M P T H P Q R
1381 GCCGCTGGTCTTGAAGACAGGGTCCAGTTCAGTGTGAAGTTGAGACTGTGGTGAATTT
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 AAATACAGTAAAAGGATTTAGGAAGTTCACATTTTGGGCACGCACACAAAAGTATGAA
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 CTCCTTAGTTTTGAAACCAATTGTGCCAGCCTGGTTTTGGTAATTGACCTCGAGACGAC
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 CTCTCTGCCCGTTGTGGTGTACTCCAACGTCAGCCAGCTCCCGAGCGGTTGGCCTCCAT
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 ATGTGCACGATGGGCTCAGCTTTTCCAGAAGTGTGAGTTGGCAGTTTTCTTCTGTACCAA
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 CAGCCCCGATGGTCTCATTCCCGTGGACGAGGTTTTGTAAAGGAAAATATAAATGATAAAAA
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 TCTCTGGAATGATGGGTGCATCATGGGCTTCATCAGCAAGGAGCGAGAGCGTGCCTGT
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 GAAGGACCAGCAGCCGGGACCTTCTGCTGCGGTTAGTGAGAGCTCCCGGGAAGGGGC
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 GGAACCTGATGGCCCTAAAGGAAGTGGATATCAAGACTGAGTTGATTTCTGTGTCTGA
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).

Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides

Tobias Merkle¹, Sarah Merz¹, Philipp Reautschnig¹, Andreas Blaha¹, Qin Li², Paul Vogel¹,
Jacqueline Wettengel¹, Jin Billy Li² and Thorsten Stafforst^{1*}

Site-directed RNA editing might provide a safer or more effective alternative to genome editing in certain clinical scenarios. Until now, RNA editing has relied on overexpression of exogenous RNA editing enzymes or of endogenous human ADAR (adenosine deaminase acting on RNA) enzymes. Here we describe the engineering of chemically optimized antisense oligonucleotides that recruit endogenous human ADARs to edit endogenous transcripts in a simple and programmable way, an approach we call RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing). We observed almost no off-target editing, and natural editing homeostasis was not perturbed. We successfully applied RESTORE to a panel of standard human cell lines and human primary cells and demonstrated repair of the clinically relevant PiZZ mutation, which causes α 1-antitrypsin deficiency, and editing of phosphotyrosine 701 in STAT1, the activity switch of the signaling factor. RESTORE requires only the administration of an oligonucleotide, circumvents ectopic expression of proteins, and represents an attractive approach for drug development.

Adenosine-to-inosine editing in RNA diversifies the transcriptome by recoding of amino acid codons, Start codons and Stop codons, and by alteration of splicing, among other mechanisms¹. Steering such enzymes to specific sites at selected transcripts, a strategy called site-directed RNA editing^{2,3}, holds great promise for the treatment of disease and as a tool to study protein and RNA function. Unlike DNA editing, RNA editing manipulates genetic information in a reversible and tunable manner. These properties may enable manipulations that are either lethal or quickly compensated when done at the genome level⁴. Furthermore, RNA editing could be safer because potential adverse effects should be reversible and dose-dependent.

We and others have recently published several RNA editing strategies based on expression of exogenous engineered deaminases^{5–7}. However, in a therapeutic setting, harnessing of the widely expressed endogenous ADARs, including ADAR1 and ADAR2, would be preferable⁸ as it would replace ectopic expression of an engineered protein with administration of an oligonucleotide drug. Recently, we engineered a plasmid-borne guide RNA (gRNA) that recruits human ADAR2 to elicit programmable, site-specific RNA editing⁹. Such gRNAs comprise two parts: an invariant ADAR-recruiting domain and a programmable specificity domain (Fig. 1a). The ADAR-recruiting domain forms an imperfect 20-bp hairpin (Fig. 1a) and was adapted from a well-known ADAR2 target site in the GluR2 mRNA, and thus was called the R/G motif. The specificity domain is a programmable, short (~18 nt), single-stranded

sequence reverse complementary to the target mRNA (Fig. 1a). We optimized the gRNA for ADAR2 recruitment and demonstrated its expression from a U6 promoter⁹. However, sufficient editing of endogenous transcripts such as *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) or *ACTB* (β -actin) always required co-overexpression of ADAR2, whereas expression of the gRNA alone failed to achieve editing⁹. In the present study, we have engineered antisense oligonucleotides (ASOs) that recruit endogenous ADAR to edit endogenous transcripts in cancer cell lines and in primary human cells.

We applied a plasmid-borne approach⁹ to screen for better ADAR-recruiting domains (Supplementary Fig. 1). While testing 15 different designs, we found sequence variant 9.4 (with an additional 5 bp at the 5' site of the R/G motif). Although less effective with ADAR2, variant 9.4 almost doubled editing with the ADAR1 isoform p110. Using ADAR1 for RNA editing could be beneficial as its expression is particularly widespread¹⁰.

To further enhance editing efficiency, we tested chemically stabilized ASOs¹¹ (RESTORE) instead of plasmid-borne⁹ gRNA expression. In the first round, we tested three ASO designs (v1, v4, v9.4). The ASOs comprised an ADAR-recruiting domain composed entirely of natural ribonucleotides and a specificity domain that was chemically modified (2'-O-methylations, phosphorothioate, Fig. 1b), containing a modification gap opposite the editing site, much like what was described before¹².

Using ASOs targeting a 5' UAG site in the 3' untranslated region (3' UTR) of either *ACTB* or *GAPDH*, we assessed the ADAR preferences of the ASOs. We lipofected them into engineered Flp-In T-REx 293 cells expressing a specific ADAR isoform (ADAR2, ADAR1 p110 or ADAR1 p150)^{9,13}. We found the highest editing efficiency (75%–85%) in ADAR1 p150-expressing cells (Fig. 1c). Editing yields were lower for ADAR1 isoform p110 (12%–50%), but showed a strong (two- to threefold) benefit of ASO v9.4 compared to ASO v1. Editing with ADAR2 was similar to editing with ADAR1 p110, however, ASO v9.4 was inferior to ASO v1. Chemical modification of the ASO was required to obtain high editing yields (Supplementary Figs. 2 and 3). Also, the presence of the ADAR-recruiting domain was essential (Supplementary Fig. 2). Finally, we tested the concurrent editing of both transcripts by cotransfection of two ASOs (Fig. 1c, right). Notably, the editing yields stayed virtually unchanged, demonstrating that site-directed RNA editing can be carried out at several transcripts simultaneously.

In HeLa cells, targeting 5' UAG codons in the 3' UTRs of *GAPDH* and *ACTB*, ASO v1 and v4 gave some editing (Fig. 2a). However, the ASO v9.4 gave clearly higher editing of both transcripts (~40%). A control ASO lacking the ADAR-recruiting domain did not elicit

¹Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. ²Department of Genetics, Stanford University, Stanford, CA, USA.

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

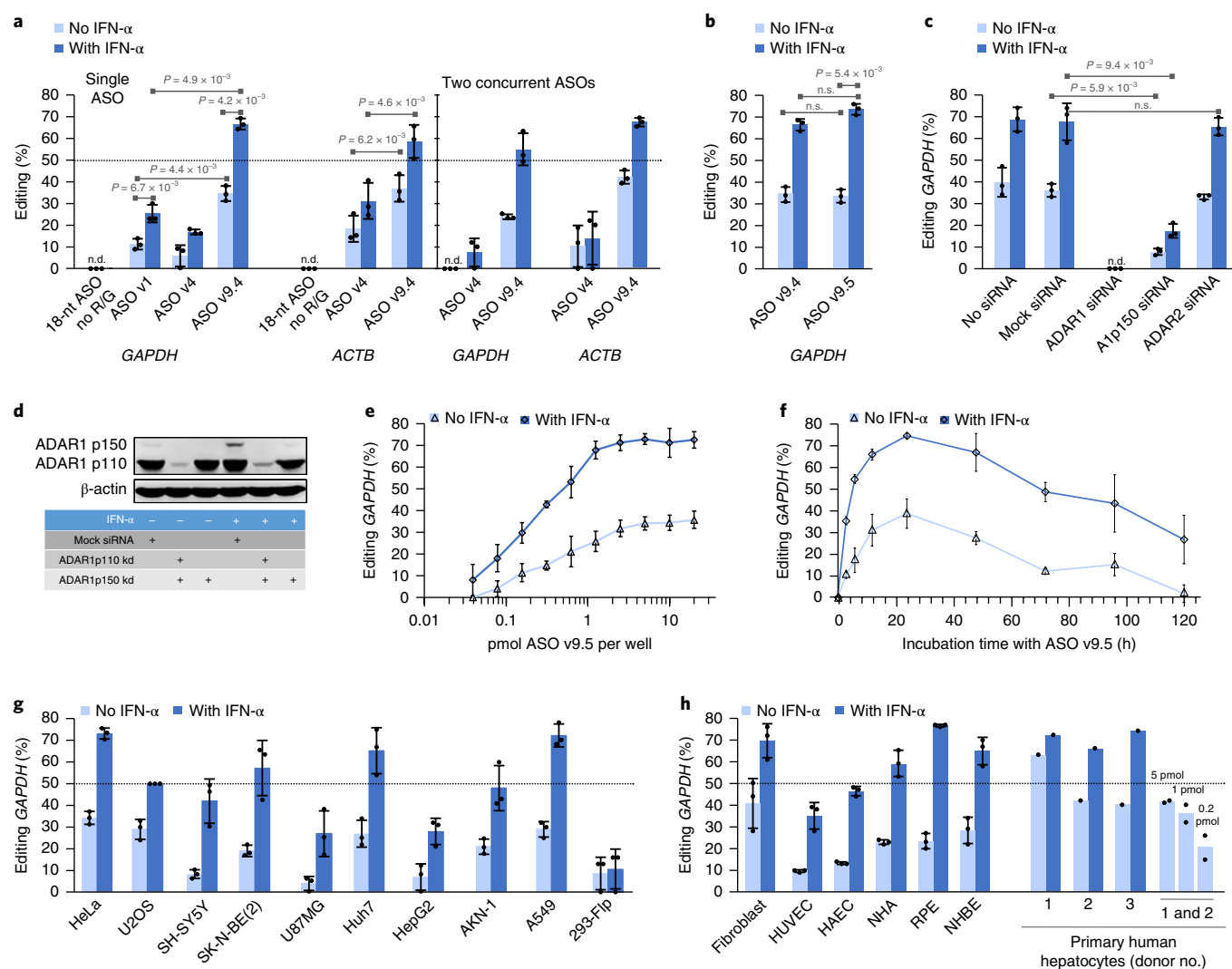


Fig. 2 | Applying RESTORE to edit endogenous transcripts (*GAPDH* and *ACTB*, each with a targeted 5' UAG triplet in the 3' UTR) in various cell lines by transfection with ASOs, performed in presence or absence of IFN- α , as indicated. **a, Comparing ASO designs for the recruitment of endogenous ADAR in HeLa cells. Either a single ASO (against *GAPDH* or *ACTB*) or both ASOs (against *GAPDH* and *ACTB*) were transfected. “no R/G” indicates an ASO lacking the ADAR-recruiting domain. **b**, Comparative editing of ASO v9.4 and v9.5 on *GAPDH*. **c**, Effect of isoform-specific ADAR knockdown on the *GAPDH* editing yield in HeLa cells. **d**, The knockdown efficiency was verified by western blot in technical duplicate. The western blot is composed of two images with different exposure times. The full blots are given in Supplementary Fig. 4. **e**, Determination of the half-maximal effective dose (ED_{50}) of ASO v9.5 for editing *GAPDH* in HeLa cells. **f**, Time course of *GAPDH* editing yields in HeLa cells. **g**, *GAPDH* editing yields with ASO v9.5 in various standard (cancer) cell lines. **h**, *GAPDH* editing yields with ASO v9.5 in various primary human cells. HUVEC, human umbilical vein endothelial cells; HAEC, human aortic endothelial cells; NHA, normal human astrocytes; RPE, human retinal pigment epithelium; NHBE, normal human bronchial epithelium. Data in **a–h** are shown as the mean \pm s.d., $N = 3$ independent experiments; experiments in hepatocytes are single determinations for each donor (donors 1–3) as indicated. Significance (P) was calculated with a two-tailed paired t -test; n.s., $P > 0.05$; A1p150, ADAR1 p150; n.d., no editing was detectable.**

expressed p150 isoform contributed more. This is in good agreement with the observed positive effect of IFN- α (Fig. 2) and the better performance of the ASO in ADAR1 p150-expressing 293 Flp-In T-REx cells (Fig. 1c). It remains unclear why the weakly expressed p150 isoform is more effective than the more strongly expressed p110 isoform. Reasons could be the different intracellular localization, different regulation, or the additional N-terminal part of the p150 isoform—for example, the Z-DNA binding α domain¹.

We found a sigmoidal dose dependency for ASO v9.5-mediated RNA editing, reaching half-maximum editing yield at 0.2 pmol ASO per well of a 96-well plate with IFN- α and 0.4 pmol/well without IFN- α (Fig. 2e and Supplementary Fig. 5). The maximum editing yield was obtained at ≥ 2 pmol/well, a dose similar to that used for siRNA duplexes in RNA interference¹⁵. As additional controls, we tested

the effect of a nontargeting ASO v9.5 and of an ADAR-recruiting domain v9.5 lacking any specificity domain on the on-target editing of *GAPDH* with ASO v9.5 (Supplementary Figs. 6 and 7). The on-target yield was not affected by the cotransfected components, indicating that the endogenous editing capacity is not limiting. We further assayed the time profile of the editing yield over 5 d in rapidly dividing HeLa cells (10% FBS, 5 pmol/well ASO). The maximum editing yield was observed 12–48 h after transfection and dropped slowly (Fig. 2f).

To assess the application scope of RESTORE, we applied ASO v9.5 in a panel of ten immortalized human cell lines (Fig. 2g). Editing yield was cell line dependent, with yields ranging from 4% to 34% (average 18.5%). Yields were two- to threefold higher after IFN- α treatment, ranging from 11% to 73% (average 46.8%). As ADAR

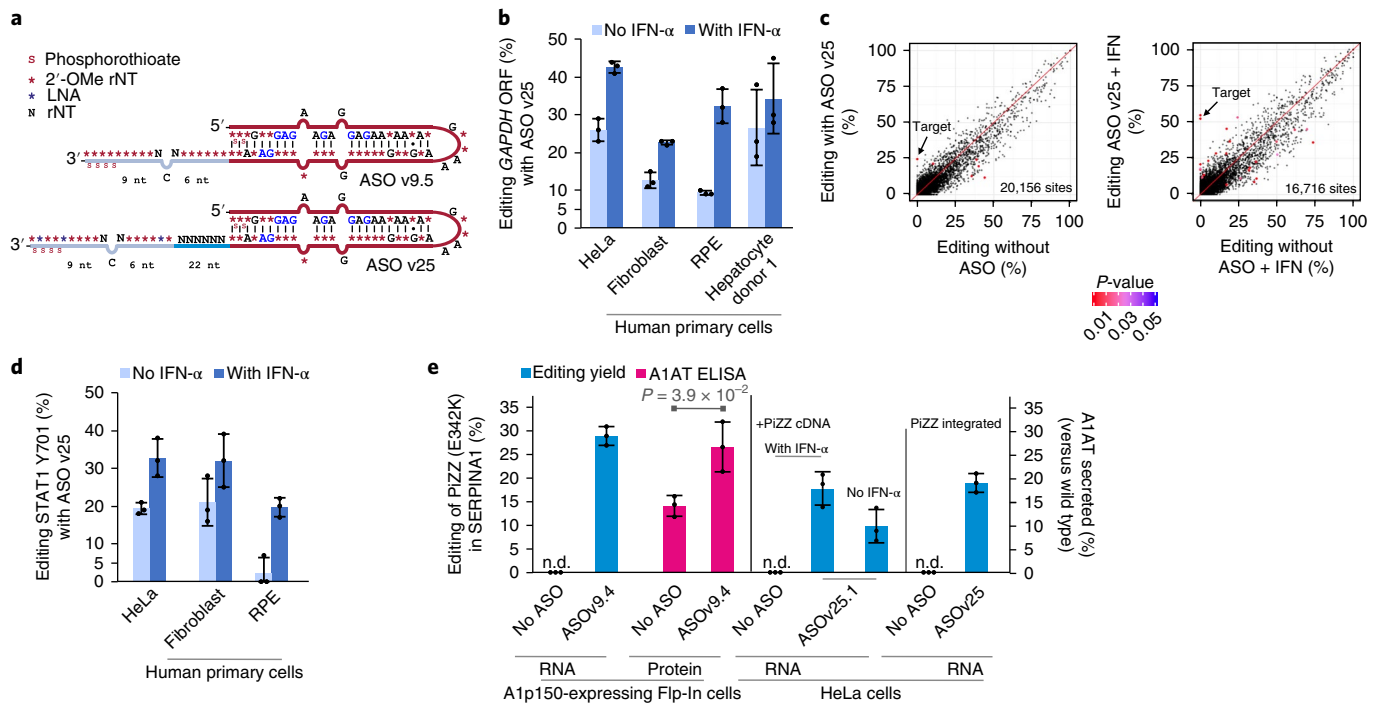


Fig. 3 | Applying RESTORE for ORF editing with ASO v25, off-target analysis, and editing of disease-relevant sites. a, ASO design v25. **b**, Editing of 5' UAG site no. 1 in the ORF of *GAPDH* with ASO v25 in HeLa and human primary cells. **c**, Analysis of off-target editing in the poly(A)⁺ transcriptome when recruiting endogenous ADAR from HeLa cells to 5' UAG site no. 1 in the ORF of *GAPDH* with ASO v25, in absence (left) or presence (right) of IFN- α . Scatter plots show differential editing at ~18,000 sites per experiment comparing editing levels in cells treated with ASO v25 compared to empty transfected cells. Experiments were done in two independent replicates. The on-target editing is indicated by an arrow. Significantly differently edited sites ($P < 0.01$, Fisher's exact test, two-sided, $N > 50$) are highlighted in red. **d**, Editing of the Tyr701 site (5' UAU codon) of STAT1 in HeLa and primary cells. **e**, Editing of the PIZZ mutation causing α 1-antitrypsin deficiency (E342K in SERPINA1, 5' CAA codon) either in ADAR1 p150-expressing 293 Flp-In T-REx cells with v9.4 ASO or in HeLa cells with v25 ASO (3-nt gap) or v25.1 (2-nt gap). The *SERPINA1* E342K cDNA was either cotransfected or genetically integrated into HeLa cells. α 1-Antitrypsin (A1AT) secretion was normalized to the secretion when transfecting wild-type *SERPINA1*. Data in **b**, **d**, **e** are shown as the mean \pm s.d., $N = 3$ independent experiments; significance (P) in **e** was calculated with a two-tailed paired t -test. n.d., no editing was detectable.

expression can differ between cancer and normal cells¹⁶, we further tested a panel of seven primary cells from different tissues, including patient fibroblasts¹⁷ and commercially acquired astrocytes, hepatocytes, retinal pigment epithelium cells (RPE), bronchial epithelial cells, and endothelial cells from arterial and venous vessels (Fig. 2h). We found higher editing levels in primary cells than in immortalized cells, obtaining editing levels of 10%–63% (average 31.5%). Notably, in all hepatocyte samples and in the fibroblasts, the editing levels were higher than in HeLa cells. Again, editing yields increased after IFN- α treatment (35%–77%, average 62.6%). We transfected a series of ASO dilutions (0.2–25 pmol ASO v9.5 per well of a 24-well plate, no IFN- α treatment) into hepatocytes of donors 1 and 2 and found a clear dose dependency (Fig. 2h).

We then tested the editing of a 5' UAG triplet in the open reading frame (ORF) of *GAPDH* in ADAR-expressing 293 Flp-In T-REx cells with an ASO v9.4. The editing in the ORF followed the same trend as in the 3' UTR (ADAR1 p150 > ADAR1 p110 \approx ADAR2), but with generally lower yields (11%–55%, Supplementary Fig. 8). Editing required the presence of the ADAR-recruiting domain (Supplementary Fig. 9). ASO v9.4 did not achieve editing in the ORF of *GAPDH* with endogenous ADAR in HeLa or A549 cells.

Thus, we further optimized the ASO design. We assumed that editing in the ORF might be kinetically limited by translation, as we had observed before^{18,19}. To improve on-target binding kinetics, we increased the length of specificity domain and included locked nucleic acid (LNA)^{20,21} modifications. We tested stepwise elongation of the specificity domain and found elongation at the 5' site to improve performance. Finally, we identified ASO v25,

which contains a 40-nt specificity domain partly modified with 2'-O-methylation, phosphorothioate, and three LNA bases (Fig. 3a). After transfection into HeLa cells, ASO v25 achieved editing yields of $26 \pm 3\%$ (without IFN- α) and $42.7 \pm 1.5\%$ (with IFN- α ; Fig. 3b). The chemical modification of the otherwise unchanged ADAR-recruiting domain was important. Without chemical modifications, v25 gave no editing in the absence and only moderate editing ($14 \pm 4\%$) in the presence of IFN- α (Supplementary Fig. 10). We then tested ASO v25 in several primary cells for the editing of the 5' UAG site in the ORF of *GAPDH*. Editing levels of $12.7 \pm 2.1\%$ (fibroblast), $9.3 \pm 0.6\%$ (RPE), and $27 \pm 10\%$ (hepatocyte) were obtained. As before, IFN- α increased the editing levels, to $22.7 \pm 0.6\%$ (fibroblast), $32.3 \pm 4.5\%$ (RPE) and $34 \pm 9\%$ (hepatocyte).

Off-target editing is a major problem of recent editing strategies. Also ADAR-directing ASOs could potentially elicit off-target editing or perturb the natural editing homeostasis. We conducted deep RNA sequencing (50 Mio 2×100 nt paired end reads per experiment) for the editing of the *GAPDH* ORF with ASO v25 in HeLa cells with and without IFN- α . The editing was precise, producing little off-target editing and keeping the natural editing homeostasis intact. In absence of IFN- α , only 3 out of 20,156 sites were significantly differently edited ($P < 0.01$, Fisher's exact test) compared to the control lacking ASO transfection (Fig. 3c, left, and Supplementary Datasets 1–3). All off-target sites were known sites, in noncoding regions (introns, 3' UTR). With IFN- α , the on-target yield increased from 25% to 52% (Fig. 3c, right), and 14 significantly differently edited off-target sites were detected, all in noncoding regions (yields 17–55%). Most sites (11 of 14) were known.

Both the 3 novel sites and the 11 known sites represented ASO-dependent off-targets effects as supported by sequence alignment with the ASO (Supplementary Fig. 11). Notably, 5 of 14 off-target sites showed attenuated editing in presence of the ASO. Sequence analysis suggests that this was due to a steric blockade of those specific natural editing sites by the ASO (Supplementary Fig. 12) and was not due to a general sequestering of ADAR by the ADAR-recruiting domain of the ASO. For comparison, the effect of IFN- α on ADAR1 expression and on the editing homeostasis was clearly visible (Supplementary Fig. 13), whereas no effects on ADAR1 expression (Supplementary Fig. 13) and global editing homeostasis (Fig. 3c) were detectable for the transfection of the ASO under both conditions (with or without IFN- α).

To illustrate the therapeutic potential of RESTORE, we give two examples. First, we targeted the functionally important phosphotyrosine 701 in endogenous signal transducer and activator of transcription 1 (STAT1)²². With an ASO v25 we achieved editing yields of $21.0 \pm 6.2\%$ in primary fibroblasts and up to 7% in RPE cells without IFN- α (Fig. 3d). With IFN- α , the yields increased to $32 \pm 7\%$ (fibroblasts) and $19.7 \pm 2.5\%$ (RPE). Overall, editing of the endogenous STAT1 transcript was possible in moderate yields in primary cell lines and HeLa cells. Second, we edited the PiZZ mutation (E342K) in SERPINA1 (serpin family A member 1), the most common cause of α 1-antitrypsin (A1AT) deficiency²³. Loss of functional antitrypsin due to the PiZZ allele causes severe damage to the lungs and the liver. Initially, we edited the E342K mutation (5' CAA triplet) by overexpression of the mutated SERPINA1 cDNA in ADAR1 p150-expressing 293 Flp-In T-REx cells. When applying an ASO v9.4, we achieved an editing yield of $29 \pm 2\%$ (Fig. 3e). The secretion of A1AT was measured by ELISA and was normalized to the secretion by cells transfected with wild-type SERPINA1 cDNA. The secretion level was elevated from $14 \pm 1.8\%$ before to $27 \pm 4.3\%$ after repair. The 5' CAA triplet contains an additional editable adenosine in closest proximity to the targeted A. We indeed found off-target editing at this proximal site (Supplementary Fig. 14); however, this was strongly reduced by further chemical modification of the ASO (Supplementary Fig. 15), as described before in the SNAP-ADAR system¹⁹. To test the repair of the PiZZ mutation with endogenous ADAR, we created a HeLa cell line stably expressing mutated SERPINA1 cDNA using the piggyBac²⁴ system or by plasmid-borne overexpression of the cDNA. With an ASO v25, we obtain editing yields of $19 \pm 2\%$ (piggyBac, with IFN- α), $18 \pm 4\%$ (plasmid-borne, with IFN- α) and $10 \pm 4\%$ (plasmid-borne, without IFN- α).

Several strategies for site-directed adenosine-to-inosine RNA editing have been described so far, including SNAP-ADAR⁵, λ N-ADAR⁶ and Cas13b-ADAR⁷. However, they all have severe limitations with respect to therapy. First, all systems require the codelivery of an artificial deaminase together with a gRNA in appropriate stoichiometry. Second, they all suffer from massive off-target editing (tens of thousands of sites) due to the overexpression of ADAR fusions^{7,19,25}, an unsolved problem²⁶. By contrast, our RESTORE approach simplifies the delivery and only a few off-target editing events were observed in our experiments. Our ASOs recruit endogenous ADARs to edit endogenous transcripts in good to moderate yields in many primary human cells. The editing yields are in the range of or even better than those achieved with the recently published Cas13b-ADAR strategy⁷ in HEK293 cells. The codon scope of RESTORE is probably limited by the codon preferences of natural ADARs²⁷, but we have already demonstrated here the editing of three different codons. The codon scope can be extended when using engineered hyperactive deaminases²⁶; however, this is hampered by massive off-target editing²⁶. In contrast, our data suggest that RESTORE allows editing with minimal off-target effects and without perturbing the natural editing homeostasis, unlike the other strategies.

ASOs have been developed as drugs to interact with RNase H, RNA interference, and splicing¹¹. RESTORE now adds the reprogramming of genetic information at specific sites by interaction with ADARs. We demonstrated the editing of two disease-relevant transcripts, SERPINA1 and STAT1, with v25 ASOs. Notably, the delivery of therapeutically effective, chemically stabilized siRNAs and ASOs into human liver has been achieved recently^{28,29}. We found primary hepatocytes comparably suitable for the RESTORE approach, and good editing has already been achieved in absence of IFN- α . Hepatocytes would also be the target for many inherited genetic diseases, including α 1-antitrypsin deficiency.

In the past, optimization of ASO sequence and chemistry was crucial to creating drugs that are effective in the clinic^{11,27,29}. We found here that our ASOs accept dense chemical modification and outcompete plasmid-borne gRNAs to recruit endogenous ADARs. There is still a large sequence and chemistry space to further improve the pharmacological properties of ADAR-directing ASOs—for example, to make the ASO shorter, to recruit ADARs more efficiently, and to expand the approach to other ADAR isoforms. This last might allow good editing without IFN- α -driven induction of ADAR1 p150 in the future. However, we expect IFN- α treatment to be more suitable in a therapeutic setting than ectopic expression of ADARs^{9,30}, as the latter could be difficult to deliver and control, whereas IFN- α is an approved drug³¹. Together, this work sets the stage for the development of a new drug system to reprogram the transcriptome using only antisense oligonucleotides.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability, and associated accession codes are available at <https://doi.org/10.1038/s41587-019-0013-6>.

Received: 1 June 2018; Accepted: 11 December 2018;

Published online: 28 January 2019

References

- Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **17**, 83–96 (2016).
- Vogel, P. & Stafforst, T. Site-directed RNA editing with antagomir deaminases—a tool to study protein and RNA function. *ChemMedChem*. **9**, 2021–2025 (2014).
- Gagnidze, K., Rayon-Estrada, V., Harroch, S., Bulloch, K. & Papavasiliou, F. N. A new chapter in genetic medicine: RNA editing and its role in disease pathogenesis. *Trends Mol. Med.* **24**, 294–303 (2018).
- Rossi, A. et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* **524**, 230–233 (2015).
- Stafforst, T. & Schneider, M. F. An RNA-deaminase conjugate selectively repairs point mutations. *Angew. Chem. Int. Ed.* **51**, 11166–11169 (2012).
- 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).
- Cox, D. B. T. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
- 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).
- 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).
- Picardi, E. et al. Profiling RNA editing in human tissues: towards the inosinome atlas. *Sci. Rep.* **5**, 14941 (2015).
- 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).
- Vogel, P., Schneider, M. F., Wettengel, J. & Stafforst, T. Improving site-directed RNA editing in vitro and in cell culture by chemical modification of the gRNA. *Angew. Chem. Int. Ed. Engl.* **53**, 6267–6271 (2014).
- Heep, M., Mach, P., Reautschnig, P., Wettengel, J. & Stafforst, T. Applying human ADAR1p110 and ADAR1p150 for site-directed RNA editing-G/C substitution stabilizes gRNAs against editing. *Genes (Basel)* **8**, 34 (2017).

14. Patterson, J. B., Thomis, D. C., Hans, S. L. & Samuel, C. E. Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology* **210**, 508–511 (1995).
15. Kim, D.-H. et al. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.* **23**, 222–226 (2005).
16. Xu, X., Wang, Y. & Liang, H. The role of A-to-I RNA editing in cancer development. *Curr. Opin. Genet. Dev.* **48**, 51–56 (2018).
17. Valente, E. M. et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science* **304**, 1158–1160 (2004).
18. Vogel, P., Hanswillemenke, A. & Stafforst, T. Switching protein localization by site-directed RNA editing under control of light. *ACS Synth. Biol.* **6**, 1642–1649 (2017).
19. Vogel, P. et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat. Methods* **15**, 535–538 (2018).
20. Singh, S. K., Koshkin, A. A., Wengel, J. & Nielsen, P. LNA (locked nucleic acids): synthesis and high-affinity nucleic acid recognition. *Chem. Commun. (Camb)*. 455–456 (1998).
21. Obika, S. et al. Synthesis of 2'-O,4'-C-methylneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar pucker. *Tetrahedr. Lett.* **38**, 8735–8738 (1997).
22. O'Shea, J. J. et al. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu. Rev. Med.* **66**, 311–328 (2015).
23. Lomas, D. A. & Mahadeva, R. α 1-antitrypsin polymerization and the serpinopathies: pathobiology and prospects for therapy. *J. Clin. Invest.* **110**, 1585–1590 (2002).
24. Woodard, L. E. & Wilson, M. H. piggyBac-ing models and new therapeutic strategies. *Trends Biotechnol.* **33**, 525–533 (2015).
25. Vallecillo-Viejo, I. C. et al. Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme. *RNA Biol.* **15**, 104–114 (2018).
26. Vogel, P. & Stafforst, T. Critical review on engineering deaminases for site-directed RNA editing. *Curr. Opin. Biotechnol.* **55**, 74–80 (2018).
27. Eggington, J. M., Greene, T. & Bass, B. L. Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* **2**, 319 (2011).
28. Prakash, T. P. et al. Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. *Nucleic Acids Res.* **42**, 8796–8807 (2014).
29. Fitzgerald, K. et al. A highly durable RNAi therapeutic inhibitor of PCSK9. *N. Engl. J. Med.* **376**, 41–51 (2017).
30. Fukuda, M. et al. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Sci. Rep.* **7**, 41478 (2017).
31. Antonelli, G., Scagnolari, C., Moschella, F. & Proietti, E. Twenty-five years of type I interferon-based treatment: a critical analysis of its therapeutic use. *Cytokine Growth Factor Rev.* **26**, 121–131 (2015).

Acknowledgements

We gratefully acknowledge the donation of primary fibroblasts from E. M. Valente (Università degli Studi di Salerno, Fisciano, Italy), of the hepatocytes cell line AKN-1 from A. Nüssler (BG Klinik, Tübingen, Germany), and of the U2OS Flp-In cell line from E. Schiebel (Universität Heidelberg, Germany). We gratefully acknowledge support from the Deutsche Forschungsgemeinschaft to T.S. (STA 1053/3-2; STA 1053/7-1). This work is supported by the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63) with an intramural innovation grant for J.W. This work is supported by National Institutes of Health grants R01GM102484 and R01GM124215 to J.B.L.

Author contributions

T.M., S.M., A.B., P.R., J.W., P.V. and T.S. conceived, performed and analyzed the experiments. Q.L. and J.B.L. analyzed and all authors interpreted next-generation sequencing data. All authors contributed to writing the manuscript.

Competing interests

T.S., J.W. and P.V. hold a patent on site-directed RNA editing (PCT/DE2016/000309). T.S., J.W., P.R. and T.M. are inventors of a filed patent based on the work published here.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41587-019-0013-6>.

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

Correspondence and requests for materials should be addressed to T.S.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Antisense oligonucleotides. Unmodified RNA oligonucleotides were produced by *in vitro* transcription from linear synthetic DNA templates (purchased from Sigma-Aldrich, Germany) with T7 RNA polymerase (Thermo Scientific, USA) at 37°C overnight. The resulting RNA was precipitated in ethanol and purified via urea (7 M) polyacrylamide (15%) gel electrophoresis (PAGE), extracted into water, precipitated with ethanol and resuspended and stored in nuclease-free water. All chemically modified RNA oligonucleotides purchased from Biospring (Germany), Eurogentec (Belgium) or Dharmacon (USA). Long sequences were assembled from two pieces by ligation. Sequences and modification patterns of all ASO are given in Supplementary Table 1.

Analysis of RNA editing. Total RNA was extracted from the cells with the RNeasy MinElute Kit (Qiagen, Germany). After DNase I (NEB, USA) treatment and reverse transcription with M-MuLV reverse transcriptase (NEB, USA), a subsequent PCR with Taq DNA polymerase (NEB) was performed. The resulting DNA was purified on an agarose gel and analyzed by Sanger sequencing (Eurofins Genomics, Germany). Adenosine-to-inosine editing yields were quantified by measuring the height of the guanosine and adenosine peaks at the respective site and dividing the guanosine peak height by the sum of the guanosine and adenosine peak heights. If the reverse primer was used for sequencing, cytidine and thymidine peaks were treated accordingly.

Cloning and editing with the plasmid-borne approach. Firefly luciferase was expressed under control of a CMV promoter from a pShuttle-CMV plasmid (see Supplementary Note 1). The W417X amber mutation was introduced via overlap PCR. Sequences of the cloned products were verified by Sanger sequencing. The R/G gRNAs were expressed under control of the U6 promoter from a modified pSilencer backbone as described before⁹. Sequences of the cloned products were verified by Sanger sequencing. Sequences of all applied R/G gRNAs are given in Supplementary Table 1. Flp-In 293 T-REX cells (R78007, Thermo Fisher Scientific) containing the respective genomically integrated ADAR version were generated previously^{9,13}. Cells were cultured in DMEM plus 10% FBS plus 100 µg/mL hygromycin B plus 15 µg/mL blasticidin S. For editing, 2.5×10^5 cells/well (ADAR1p110, ADAR1p150) or 3×10^5 cells/well (ADAR2) were seeded into poly-D-lysine-coated 24-well plates in 500 µL DMEM plus 10% FBS plus 10 ng/mL doxycycline. Twenty-four hours later, transfection was performed with the luciferase reporter plasmid (300 ng) and the R/G gRNA (1,300 ng) using a ratio of Lipofectamine 2000 to plasmid of 3:1. The medium was changed every 24 h until harvest. RNA was isolated and sequenced 72 h after transfection, as described above. Results are reported in Supplementary Fig. 1.

Editing procedure with ASOs in ADAR-expressing 293 cells. Forty-eight hours before ASO transfection, 2×10^5 of the respective ADAR-Flp-In 293 T-REX cells per well were seeded in 24-well plates in DMEM plus 10% FBS containing 10 ng/mL doxycycline for induction of ADAR gene expression. After 48 h cells were detached and reverse-transfected in 96-well plates. For this, the respective ASO (5 pmol/well unless stated otherwise) and Lipofectamine 2000 (0.75 µL/well) were each diluted with OptiMEM to a volume of 10 µL in separate tubes. After 5 min, the two solutions were mixed and 100 µL cell suspension (5×10^4 cells) in DMEM plus 10% FBS plus 10 ng/mL doxycycline was added to the transfection mixture inside 96-well plates. Twenty-four hours later, cells were harvested for RNA isolation and sequencing as described above. Results are reported in Fig. 1c and Supplementary Figs. 2, 3, 6 and 8a.

Editing procedure with ASO in HeLa cells. HeLa cells (cat. no. ATCC CCL-2) were cultured in DMEM plus 10% FBS plus P/S (100 U/mL penicillin and 100 µg/mL streptomycin). 5×10^4 cells in 100 µL DMEM plus 10% FBS (plus 600 units IFN- α , Merck, cat. no. IF007, lot number 2937858) were added to a transfection mix of 0.5 µL Lipofectamine 2000 and 5 pmol gRNA/well in a 96-well format. For concurrent editing with two different ASOs, 2.5 pmol of each respective ASO were cotransfected. After 24 h cells were harvested for RNA isolation and sequencing. Results are reported in Fig. 2a–f and Supplementary Fig. 7.

siRNA knockdown of ADAR isoforms and western blot. HeLa cells were reverse transfected in 12-well format with 2.5 pmol siRNA against ADAR1 (both isoforms, Dharmacon, SMARTpool: ON-TARGETplus ADAR1 (103) siRNA, L-008630-00-0005), ADAR1p150 (Ambion (Life Technologies), sense strand: 5'-GCCUCGCGGGCGCAAUGAAAt; antisense strand: 5'-UUCAUUGCGCCGCGAGGCat), ADAR2 (Dharmacon, SMARTpool: ON-TARGETplus ADAR2 (104) siRNA, L-009263-01-0005) or mock (Dharmacon, siGENOME Non-Targeting siRNA Pool #2, D-001206-14-05). For this, 200 µL of transfection mix, containing 2.5 µL of the respective siRNA (1 nM) and 3 µL HiPerFect (Qiagen, Germany) and OptiMEM, were distributed evenly in each well before adding 800 µL DMEM plus 10% FBS containing 1.2×10^5 HeLa cells. Medium was changed every 24 h. For RNA editing experiments, cells were detached 48 h after siRNA transfection and were reverse-transfected with the respective ASO as described above. For western blotting, cells were harvested and lysed in urea lysis buffer (8 M urea, 100 mM NaH₂PO₄, 10 mM Tris, pH 8.0) 72 h

after reverse transfection of the siRNA. Shear force was applied using a 23-gauge syringe, and the cell debris was removed by centrifugation at 30,000 g for 15 min at 4°C. Then a Bradford assay was used to normalize total protein amounts, and appropriate amounts of protein lysate in 1× Laemmli buffer were loaded for SDS-PAGE (4% stacking, 12% separating gel). Proteins were transferred on a PVDF membrane using a tank-blotting system at 30 V overnight. The membrane was blocked in 5% nonfat dry milk TBST plus 50 µg/mL avidin for 2 h at room temperature, and was afterwards incubated with the primary antibodies (5% nonfat dry milk TBST plus 1:1,000 anti-ADAR1, Santa Cruz, sc-73408 or anti-ADAR2, Santa Cruz, sc-73409, plus 1:40,000 anti-beta-actin, Sigma Aldrich, A5441) at 4°C overnight. The secondary antibodies (5% nonfat dry milk TBST plus 1:10,000 anti-mouse-HRP plus 1:50,000 Precision Protein StrepTactin-HRP Conjugate, Bio-Rad, cat. no. 1610381) were incubated for 1.5 h at room temperature. After each antibody incubation, the membrane was washed three times for 5 min with TBST. Detection was performed using 1 mL of Clarity Western ECL Substrate (Bio-Rad) and a Fusion SL Vilber Lourmat (Vilber). For antibodies, see also Supplementary Table 2. Results are reported in Fig. 2c,d and Supplementary Fig. 4.

Potency determination. For potency determination, HeLa cells were transfected as described above with varying ASO amounts (39 fmol–20 pmol per well of a 96-well plate). Results are reported in Fig. 2e.

Time course. For time course experiments, HeLa cells were transfected as described above. Prior transfection cells were treated with IFN- α for 24 h (where indicated). Cells were harvested for RNA isolation at the respective time points indicated. For time points later than 24 h after transfection, cells were detached after 24 h and transferred into 24-well plates to avoid overgrowth. Medium (containing IFN- α where indicated) was changed every 24 h. Results are reported in Fig. 2f.

Screening of immortalized cell lines. ASO transfection was not systematically optimized. All cells were cultured in DMEM plus 10% FBS plus P/S. 5×10^4 cells of the respective cell line per well of a 96-well plate (HeLa cells (cat. no. ATCC CCL-2), U2OS-Flp-In T-REX³³ (kind donation from Elmar Schiebel), SK-N-BE(2) (cat. no. ATCC CRL-2271), U87MG (cat. no. ATCC HTB-14), Huh7 (CLS GmbH, Heidelberg, cat. no. 300156), HepG2 (DSMZ, Braunschweig, Germany, cat. no. ACC180), AKN-1 (kind donation from the Nüssler laboratory³³), empty HEK-Flp-In T-REX (R78007, Thermo Fisher scientific, stably transfected with empty pcDNA5 vector) and A549 (European Collection of Authenticated Cell Cultures ECACC 86012804)) were reverse transfected with the respective ASO (5 pmol per well of a 96-well plate) as described above for HeLa cells without further optimization. Only SH-SY5Y (cat. no. ATCC CRL-2266) cells were reverse transfected differently, in a 24-well format: to 100 µL transfection mix consisting of 2.5 µL Lipofectamine 2000 and 25 pmol ASO in OptiMEM, 5×10^5 cells in 500 µL medium (plus 3,000 U IFN- α) were added. Results are reported in Fig. 2g.

Screening of human primary cell lines. ASO transfection was not systematically optimized. All primary cells were purchased from Lonza except for the primary fibroblasts, which were a kind gift from the Valente laboratory¹⁷. Primary fibroblasts were cultured in DMEM plus 20% FBS. The other cell lines were cultured in their respective commercial media as indicated: human umbilical vein endothelial cells (HUVEC, Lonza cat. no. CC-2517) and human aortic endothelial cells (HAEC, Lonza cat. no. CC-2535) in medium 200PRF (Thermo Fisher Scientific cat. no. M200PRF500) with Low Serum Growth Supplement (Thermo Fisher Scientific cat. no. S00310), normal human astrocytes (NHA, Lonza cat. no. CC-2565) in ABM Basal Medium (Lonza cat. no. CC-3187) with AGM SingleQuot Kit Supplementary & Growth Factors (Lonza cat. no. CC-4123), human retinal pigment epithelial cells (H-RPE, Lonza cat. no. 194987) in EpiLife Medium (Thermo Fisher Scientific cat. no. MEPI500CA) with Human Corneal Growth Supplement (Thermo Fisher Scientific cat. no. S0095), and normal human bronchial epithelial cells (NHBE, Lonza cat. no. CC-2540) in Airway Epithelial Cell Basal Medium (LGC Standard cat. no. ATCC-PCS-300-030) with the Bronchial Epithelial Cell Growth Kit (LGC Standard cat. no. ATCC-PCS-300-040). Primary human hepatocytes (PHH, Lonza cat. no. HUCPI) were thawed in Cryo HH thawing medium (Lonza cat. no. MCHT50), seeded in Hepatocyte Plating Medium with Supplement (Lonza cat. no. MP100) and, 6 h after seeding, cultured in Hepatocyte Maintenance Medium with Supplement (Lonza cat. no. MM250). 3.5×10^5 HUVEC and HAEC, 1×10^5 NHA, H-RPE and NHBE and 4.5×10^5 PHH cells were seeded 24 h before ASO transfection in 24-well format. For PHH, rat collagen I-coated 24-well plates (GreinerBioOne) were used. Shortly before forward transfection, medium was changed; 3,000 U IFN- α in 500 µL medium per well was included if indicated. For each 24-well, 1.5 µL Lipofectamine RNAiMAX (Thermo Fisher Scientific) and 25 pmol ASO were diluted separately in a total volume of 50 µL OptiMEM, respectively. After 5 min incubation the two solutions were combined, and after another 20 min incubation, the 100 µL transfection mix was evenly distributed in one well. After 24 h cells were harvested for RNA isolation and sequencing. Results are reported in Fig. 2h.

ORF editing. If not indicated, ORF editing experiments were performed the same as editing experiments in the 3' UTR for the respective cell lines as described

above. For PHH, 7.5 μ L RNAiMAX per well were used. Before reverse transcription of RNA from cells treated with design v25 ASOs, total RNA was incubated with an RNA strand reverse complementary to the respective ASO and heated to 95 °C for 3 min. Results are reported in Fig. 3b and Supplementary Figs. 8–10.

Next-generation RNA sequencing experiment. The RNA editing experiment was done by transfection of 5 pmol ASO against 5'-UAG ORF site #1 in the ORF of *GAPDH* into HeLa cells as described above. For samples with IFN- α , HeLa cells were treated with IFN- α 24 h before reverse transfection as described above. Overall, four settings were carried out, each with an independent duplicate. Those settings include (1) empty lipofection, (2) empty lipofection plus IFN- α , (3) ASO transfection, and (4) ASO transfection plus IFN- α . RNA was isolated with the RNeasy MinElute Kit, treated with DNase I, incubated with an RNA strand reverse complementary to the respective ASO and heated to 95 °C for 3 min and purified again with the RNeasy MinElute Kit. Purified RNA was delivered to CeGaT (Germany) for poly(A)⁺ mRNA sequencing. The library was prepared from 200 ng RNA with the TruSeq Stranded mRNA Library Prep Kit (Illumina, USA) and sequenced with a NovaSeq 6000 (50M reads, 2 \times 100 bp paired end, Illumina, USA). Results are reported in Fig. 3c and Supplementary Figs. 11–13.

Mapping of RNA-seq and reads. We adopted a previously published pipeline to accurately align RNA-seq reads onto the genome^{34,35}. We used BWA (version 0.7.10)³⁶ to align the reads to a combination of the reference genome sequences and exonic sequences surrounding known splicing junctions from known gene models. Each of the paired-end reads was mapped separately using the commands “bwa aln fastqfile” and “bwa samse -n4”. We then chose the length of the splicing junction to be slightly shorter than the RNA-seq reads to prevent redundant alignment (i.e., 95 bp for reads of 100 bp length). The reference genomes used were hg19 and the gene models were obtained through the UCSC Genome Browser for Gencode, RefSeq, Ensembl, and UCSC Genes. We considered only uniquely mapped reads with mapping quality $q > 10$ and used Picard³⁷ to remove clonal reads (PCR duplicates) mapped to the same location. Of these identical reads, only the read with the highest mapping quality was kept for downstream analysis. Unique and nonduplicate reads were subjected to local realignment and base score recalibration using the IndelRealigner and TableRecalibration from the Genome Analysis Toolkit (GATK, version 3.6)³⁷. The above steps were applied separately to each of the RNA-seq samples.

Identification of editing sites from RNA-seq data. We used the UnifiedGenotyper from GATK³⁷ to call variants from the mapped RNA-seq reads. In contrast to the usual practice of variant calling, we identified the variants with relatively loose criteria by using the UnifiedGenotyper tool with options stand_call_conf 0, stand_emit_conf 0, and output mode EMIT_VARIANTS_ONLY. Variants from nonrepetitive and repetitive non-Alu regions were required to be supported by at least three reads containing mismatches between the reference genome sequences and RNA-seq. Supporting of one mismatch read was required for variants in Alu regions. This set of variant candidates was subject to several filtering steps to increase the accuracy of editing site calling. We first removed all known human SNPs present in dbSNP build 137 (except SNPs of molecular type “cDNA”; database version 135; <http://www.ncbi.nlm.nih.gov/SNP/>), the 1000 Genomes Project, and the University of Washington Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>). To remove false-positive RNA-seq variant calls due to technical artifacts, further filters were applied as previously described^{34,35}. In brief, we required a variant call quality $Q > 20$ (refs. ^{34,35}), discarded variants if they occurred in the first 6 bases of a read³⁶, removed variants in simple repeats³⁸, removed intronic variants that were within 4 bp of splice junctions, and discarded³⁷ variants in homopolymers. Moreover, we removed reads mapped to highly similar regions of the transcriptome by BLAT³⁹. Finally, variants were annotated using ANNOVAR (version 11122013)⁴⁰ based on gene models from Gencode, RefSeq, Ensembl and UCSC.

Assignment of known versus novel sites. The resulting sets of sites identified from RNA-seq data were compared with all sites available in the RADAR database⁴¹ and were subsequently referred to as ‘known’ sites if found in RADAR or ‘novel’ sites if not found.

Identification of significantly differently edited sites. We quantified editing levels of edited sites with ≥ 50 reads coverage (combined coverage of both replicates)

and performed Fisher’s exact tests followed by Benjamini–Hochberg’s multiple test correction (adjusted $P < 0.01$) to identify significantly differently edited sites across the samples (absolute editing difference $> 10\%$). Additional next-generation sequencing quality data are given in the Supplementary Information.

SERPINA1 editing and A1AT-ELISA. To obtain *SERPINA1* cDNA for cloning, total RNA was isolated from HepG2 cells and reverse transcribed. The E342K mutation was inserted into the cDNA by PCR and both *SERPINA1* wild-type and the E342K mutant were each cloned on a pCDNA3.1 vector under control of the CMV promoter using HindIII and ApaI restriction. For genomic integration of *SERPINA1* using the piggyBac transposon system, the wild-type and mutant cDNA was cloned on a PB-CA vector using the same restriction sites as above. 1×10^6 HeLa cells were seeded in a six-well plate 24 h before transfection. 1 μ g of the piggyBac transposase vector (Transposagen Biopharmaceuticals) and 2.5 μ g of the *SERPINA1* PB-CA vector were cotransfected using 10.5 μ L FuGENE6 (Promega) according to the manufacturer’s protocol. After 24 h, cells were selected for 2 weeks in DMEM plus 10% FBS containing 10 μ g/mL puromycin. For editing, stably transfected or plasmid transfected (300 ng plasmid/0.9 μ L FuGENE6 for HeLa and 100 ng plasmid/0.3 μ L Lipofectamine 2000 for Flp-ADAR1p150 cells) cells were reverse transfected with the respective ASO as described above. After 24 h, cell culture supernatant was collected for the A1AT ELISA and cells were harvested for RNA isolation and sequencing. The A1AT ELISA was performed with a commercial kit (cat. no. ab108799, Abcam) according to the manufacturer’s protocol. Samples from three biological replicates were measured in technical duplicates. The A1AT protein amount was calculated from a standard curve using linear regression. ASO v25 refers to the ASO with the common 3-nt gap around the editing site; ASO v25.1 refers to an ASO of the same sequence but with an additional chemical modification (2' O-methyl) close to the editing site (2-nt gap; see also Supplementary Fig. 15 and Supplementary Table 1). Results are reported in Fig. 3e and Supplementary Figs. 14 and 15.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

This manuscript provides Supplementary Information on primary data and further controls (Supplementary Figs. 1–15), and it contains a table of ASOs (Supplementary Table 1), a list of target sequences (Supplementary Note 1), and spreadsheets with significantly differently edited sites (Supplementary Datasets 1–3). The original next-generation sequencing data have been deposited in the NCBI GEO database under accession code GSE121573. Code is available at <http://lilab.stanford.edu/SNPiR/>.

References

- Malecki, M. J. et al. Leukemia-associated mutations within the NOTCH1 heterodimerization domain fall into at least two distinct mechanistic classes. *Mol. Cell. Biol.* **26**, 4642–4651 (2006).
- Nüssler, A. K. et al. Isolation and characterization of a human hepatic epithelial-like cell line (AKN-1) from a normal liver. *In Vitro Cell. Dev. Biol. Anim.* **35**, 190–197 (1999).
- Ramaswami, G. et al. Accurate identification of human Alu and non-Alu RNA editing sites. *Nat. Methods* **9**, 579–581 (2012).
- Ramaswami, G. et al. Identifying RNA editing sites using RNA sequencing data alone. *Nat. Methods* **10**, 128–132 (2013).
- Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
- McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
- Ramaswami, G. & Li, J. B. RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res.* **42**, D109–D113 (2014).

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

No custom-made code was required to collect data. The Western blot pictures were taken with the FusionCapt Advance SL4 (16.09b) software installed on a Fusion SL Vilber Lourmat (Vilber) western blot analyzer. No further image processing was done with respect to brightness or contrast. Next-generation sequencing of Poly(A)+ mRNA was done by CeGaT (Germany). The library was prepared with the TruSeq Stranded mRNA Library Prep Kit (Illumina, USA), and sequenced with the NovaSeq 6000 (50 M reads, 2 × 100 bp paired end, Illumina, USA).

Data analysis

As outlined in full detail in the online methods all software tools used for NGS are publically available: Mapping of RNA-seq and reads: BWA (version 0.7.10) was used to align the reads to a combination of the reference genome sequences (hg19) and exonic sequences surrounding known splicing junctions from known gene models, obtained through the UCSC Genome Browser for Gencode, RefSeq, Ensembl, and UCSC Genes. Unique and non-duplicate reads were subjected to local realignment and base score recalibration using the IndelRealigner and TableRecalibration from the Genome Analysis Toolkit (GATK, version 3.6). 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. We first removed all known human SNPs present in dbSNP, build 137 (except SNPs of molecular type "cDNA"; database version 135; <http://www.ncbi.nlm.nih.gov/SNP/>), the 1000 Genomes Project, and the University of Washington Exome Sequencing Project (<http://evs.gs.washington.edu/EVS/>). Finally, variants were annotated using ANNOVAR (version 11122013) based on 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 database and were subsequently referred to as 'known' sites if also found in RADAR, or 'novel' sites if not found. Code can also be downloaded at: <http://lilab.stanford.edu/SNPiR/>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

This manuscript provides additional Supplementary Information on primary data and further controls (Supplementary Figures 1-15), it contains a Table of all oligonucleotides used (Supplementary Table 1) and a list of all target sequences (Supplementary Note 1). Furthermore it contains 3 excel spread sheets with the NGS off-target analysis. The accession code of the NGS raw data is not yet applicable but will be available before publication.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|-----------------|--|
| Sample size | Experiments for evaluating editing yields of endogenous targets via Sanger sequencing were mostly done in triplicate (independent biological experiments) in rare cases in duplicate to validate reproducibility and to provide appropriate standard deviations. Single data points are always given. NGS analysis was performed with two independent replicates per sample; the required sequencing depth was determined in a previous study (Vogel et al. Nature Methods 2018) and saturated with 50 Mio 100 bp paired-end reads at 25 000 detected transcripts. This sequencing depths was also similar to other very recent papers on global off-target effects of site-directed RNA editing (Cox et al. Science 2017, Rosenthal et al. RNA Biol. 2018) |
| Data exclusions | no data was excluded |
| Replication | all experiments could be reproduced, as shown in the manuscript, the number of replications and nature of replicates is always given in the figure caption |
| Randomization | no randomization was performed, samples were treated according to the same protocols side-by-side with the respective controls |
| Blinding | no blinding was performed, editing experiments were allocated to several experimentators |

Reporting for specific materials, systems and methods

Materials & experimental systems

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Unique biological materials |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |

Methods

- | | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials

All commercial primary cells were obtained from Lonza. Three cell lines, which are not standard cell lines, were donations from other labs, primary fibroblasts (Dr. Enza Maria Valente, Università degli Studi di Salerno, Fisciano, Italy), U2OS Flp In cells (Dr. Elmar Schiebel, ZMBH, Heidelberg, Germany), and AKN-1 cells (Dr. Andreas Nüssler, UKT, Tübingen, Germany), as indicated below. Please ask the respective lab to obtain the respective cell line from us or them.

Antibodies

Antibodies used

ADAR1 antibody (α -ADAR1, mouse monoclonal IgG, Santa Cruz cat. no.: sc-73408, clone no. 15.8.6, lot no. C2514, used in 1:1000 dilution) against amino acids 440-826 corresponding to the middle region of ADAR1 of human origin, ADAR2 antibody (α -ADAR2, mouse monoclonal IgG, Santa Cruz cat. no.: sc-73409, clone no. 1.3.1, lot no. G1613, used in 1:1000 dilution) against N-terminal region corresponding to amino acids 2-179 of ADAR2 of human origin, Clone AC-15 (α -ACTB, mouse monoclonal IgG, Sigma Aldrich cat. No.: A5441) against Actin N-terminal peptide, Ac-Asp-Asp-Asp-Ile-Ala-Ala-Leu-Val-Ile-Asp-Asn-Gly-Ser-Gly-Lys.

Validation

ADAR1 antibody:
Validated in our lab via siRNA KO and Western Blot (in several cell lines), and by overexpression / Western blot
PMID: # 28669490
PMID: # 28278381
PMID: # 27573237
PMID: # 27907896
ADAR2 antibody:
Validated in our lab via overexpression and Western Blot
PMID: # 26601943
PMID: # 24345557
PMID: # 27907896
Clone AC-15 antibody:
PMID: # 15809369
PMID: # 15048076
PMID: # 21217779

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

HeLa: ATCC (Cat.No.: ATCC CCL-2), U2OS-Flp-In T-Rex: kind donation from Prof. Elmar Schiebel, SK-N-BE(2) ATCC: (Cat.No.: ATCC CRL-2271), SK-N-BE(2): ATCC (Cat.No.: ATCC CRL-2271), U87MG: ATCC (Cat.No.: ATCC HTB-14), Huh7:CLS (CLS GmbH, Heidelberg, Cat.No.: 300156), HepG2 DSMZ (DSMZ, Braunschweig, Germany Cat.No.: ACC180), AKN-1: kind donation from the Nüssler lab, A549: ECACC (European Collection of Authenticated Cell Cultures ECACC 86012804), SH-SY5Y: ATCC (Cat.No.: ATCC CRL-2266), HEK-Flp-In T-Rex-A1p110 (R78007, Thermo Fisher scientific, stably transfected with ADAR1 p110 vector in our lab), HEK-Flp-In T-Rex-A1p150 (R78007, Thermo Fisher scientific, stably transfected with ADAR1p150 vector in our lab), HEK-Flp-In T-Rex-ADAR2 (R78007, Thermo Fisher scientific, stably transfected with ADAR2 vector in our lab), empty HEK-Flp-In T-Rex (R78007, Thermo Fisher scientific, stably transfected with empty pcDNA5 vector), primary fibroblasts: kind gift from the Valente lab. Human Umbilical Vein Endothelial Cells: Lonza (HUVEC, Cat.No.:CC-2517), Human Aortic Endothelial Cells: Lonza (HAEC, Cat.No.:CC-2535), Normal Human Astrocytes: Lonza (NHA, Cat.No.: CC-2565), Human Retinal Pigment Epithelial Cells: Lonza (H-RPE, Cat.No.: 194987), Normal Human Bronchial Epithelial Cells: Lonza (NHBE, Cat.No.: CC-2540) and Primary Human Hepatocytes: Lonza (PHH, Cat.No.: HUCPI)

Authentication

HUVEC, HAEC, NHA, RPE, NHBE, and primary hepatocytes were authenticated by the supplier. Commercial standard cell lines like HeLa, SK-N-BE, U87MG, Huh7, SH-SY5Y, empty HEK-293-Flp-In T-Rex were authenticated by the respective suppliers. Cell lines were not additionally authenticated by us.

Mycoplasma contamination

HUVEC, HAEC, NHA, RPE, and NHBE were certified as mycoplasma-free by the supplier. Flp-In T-REx 293 ADAR1p110, Flp-In T-REx 293 ADAR1p150, Flp-In T-REx 293 ADAR2, HeLa, SK-N-BE(2), Huh7, A549 have been tested as mycoplasma-free in house.

Commonly misidentified lines
(See [ICLAC](#) register)

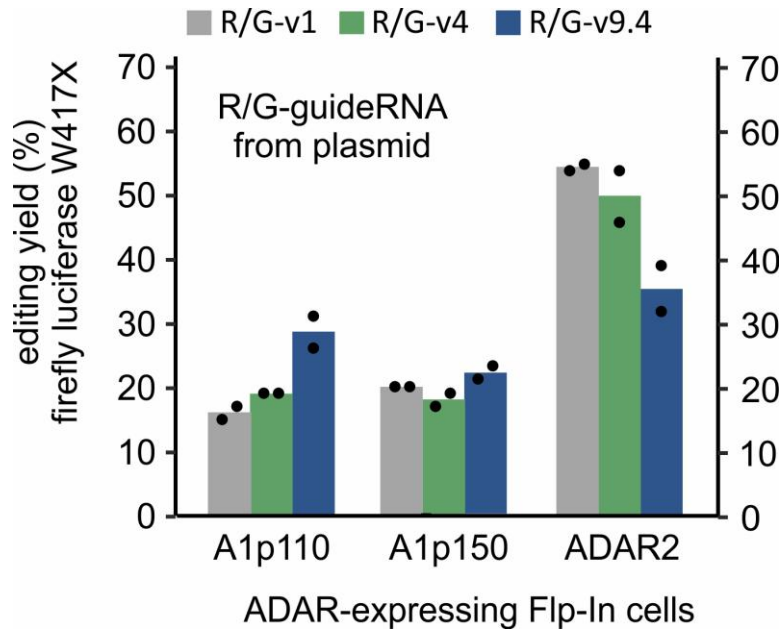
none were used

In the format provided by the authors and unedited.

Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides

Tobias Merkle¹, Sarah Merz¹, Philipp Reautschnig¹, Andreas Blaha ¹, Qin Li², Paul Vogel ¹,
Jacqueline Wettengel¹, Jin Billy Li ² and Thorsten Stafforst ^{1*}

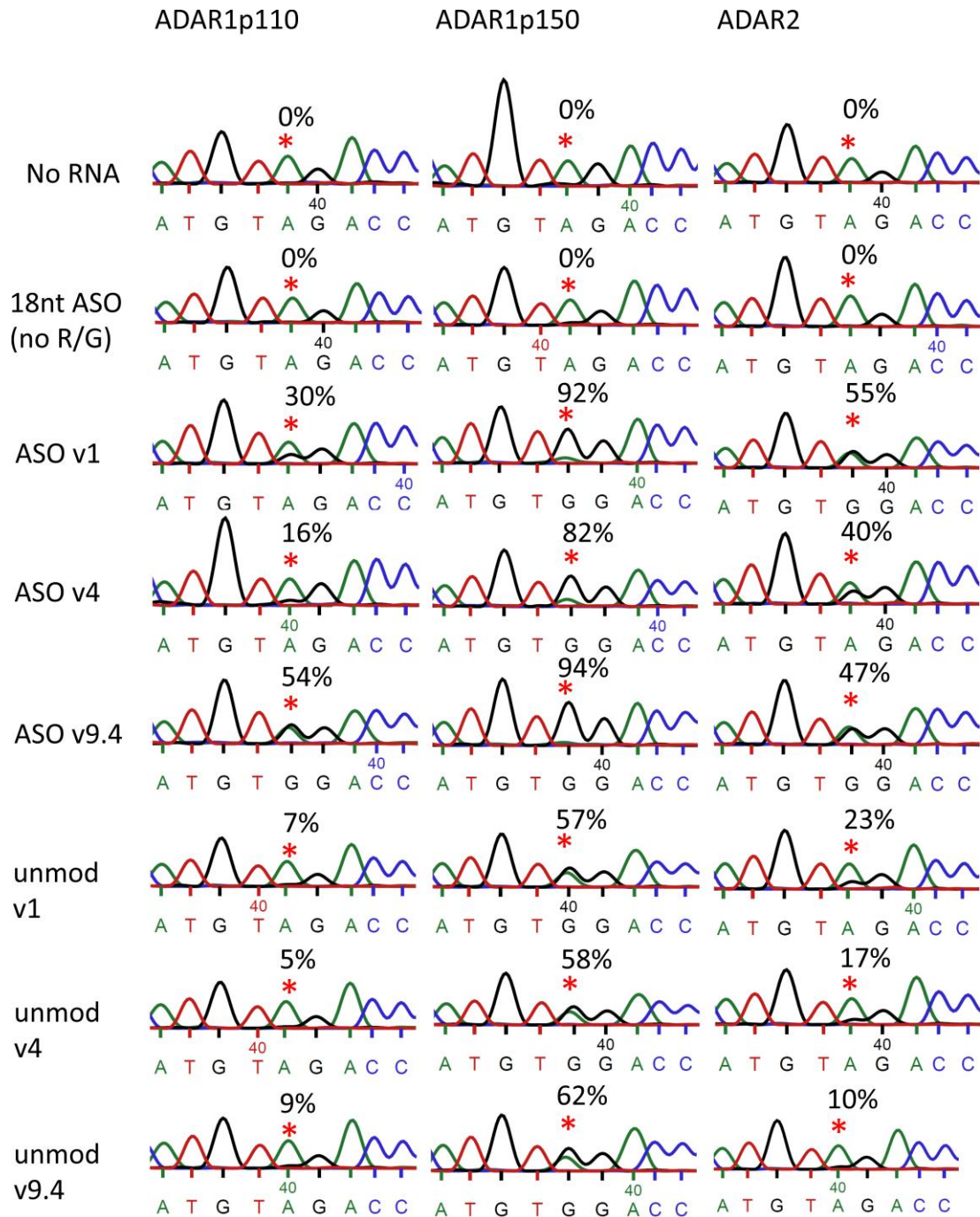
¹Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. ²Department of Genetics, Stanford University, Stanford, CA, USA.
*e-mail: thorsten.stafforst@uni-tuebingen.de



Supplementary Figure 1

Screening to improve the ADAR-recruiting domain

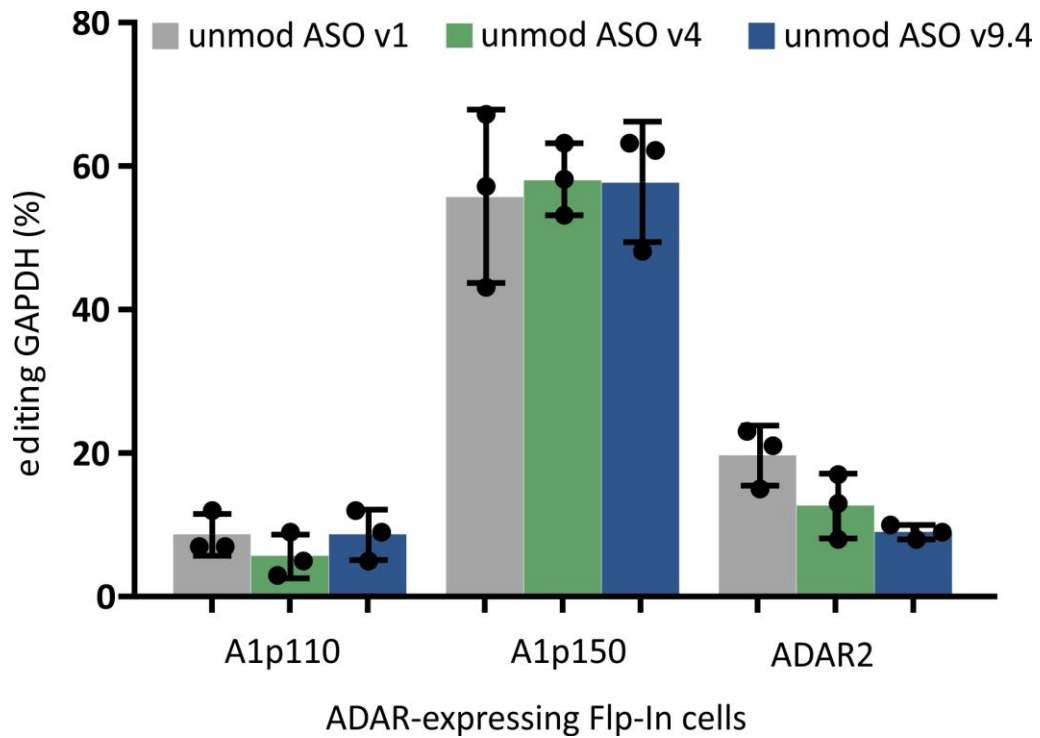
A plasmid borne screening assay was applied to screen for improved ADAR-recruiting domains. For this, plasmids expressing the respective ASO as a chemically unmodified guideRNA from a U6 promotor were prepared. The guide RNA plasmids were co-overexpressed together with a reporter construct (firefly luciferase) in 293 Flp-In T-REx cells expressing a specific ADAR isoform (A1p110 = ADAR1p110; A1p150 = ADAR1p150). Editing yields were determined by Sanger sequencing. Data are shown as the mean \pm SD, N=2 independent experiments



Supplementary Figure 2

Sequencing traces for editing of a 5' UAG site in the 3' UTR of GAPDH in 293 Flp-In T-REx ADAR cells

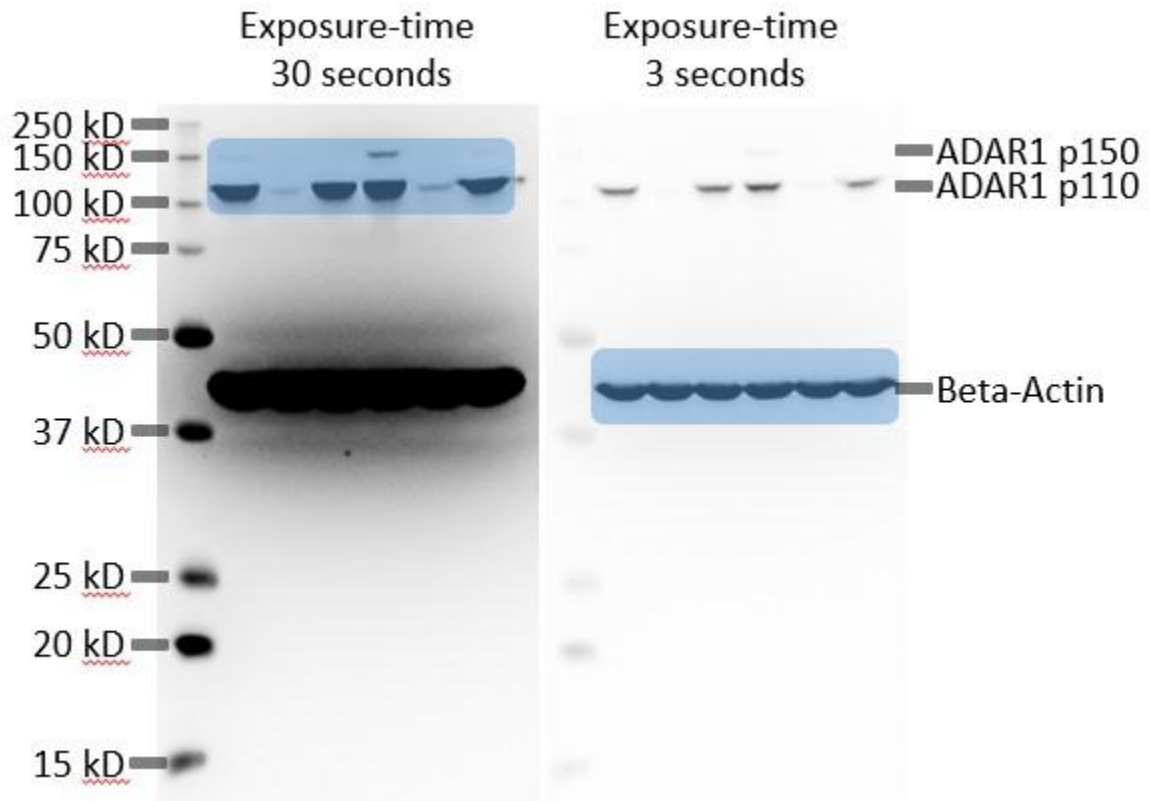
Exemplary editing traces for the editings shown in Figure 1C in the manuscript, but including additional controls ("No RNA" = empty transfection; "18nt ASO no R/G" = ASO lacking the ADAR-recruiting domain; "unmod" means chemically unmodified, *in-vitro* transcribed ASOs of the indicated design v1, v4 or v9.4. Red asterisks indicate the editing sites.



Supplementary Figure 3

Editing yields for targeting a 5' UAG codon in the 3' UTR of *GAPDH* with chemically unmodified, in vitro transcribed ASOs in 293 Flp-In T-REx ADAR cells

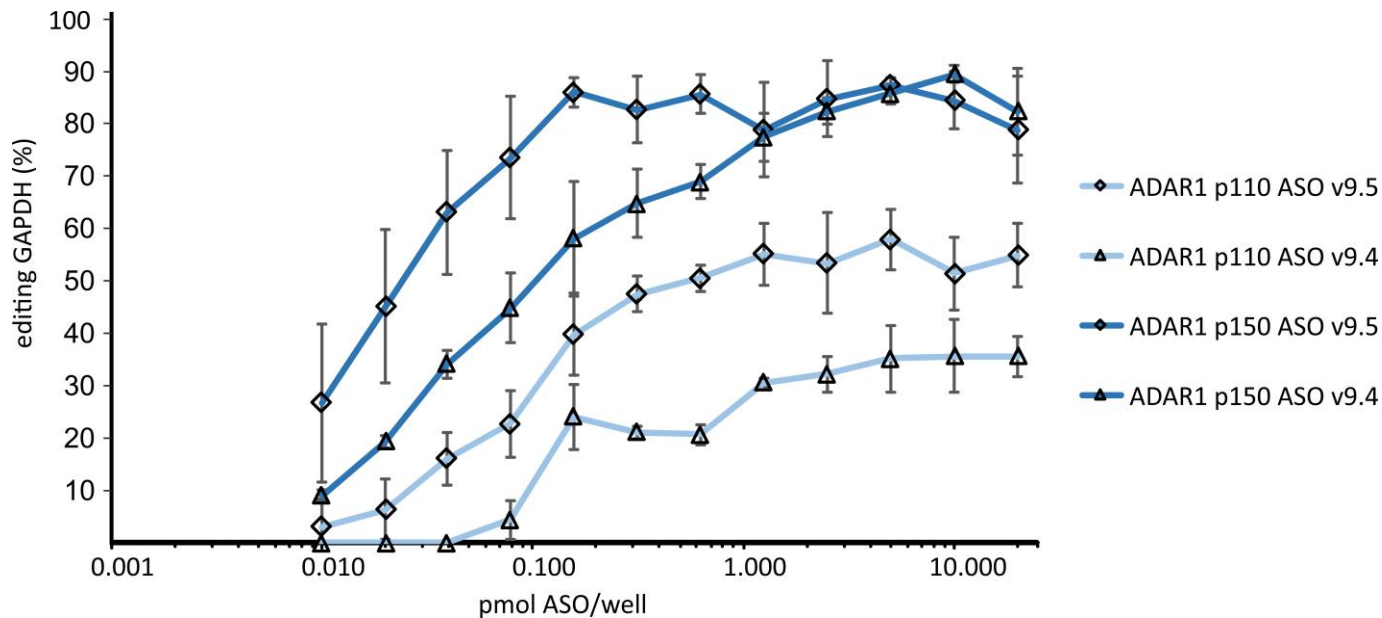
Unmodified *in-vitro* transcribed ASOs v1, v4 and v9.4 (5 pmol / 96well) were transfected into the respective ADAR-expressing Flp-In cell line. Data are shown as the mean \pm SD, N=3 independent experiments. A1p110 = ADAR1p110; A1p150 = ADAR1p150



Supplementary Figure 4

Western blot analysis of ADAR knockdown

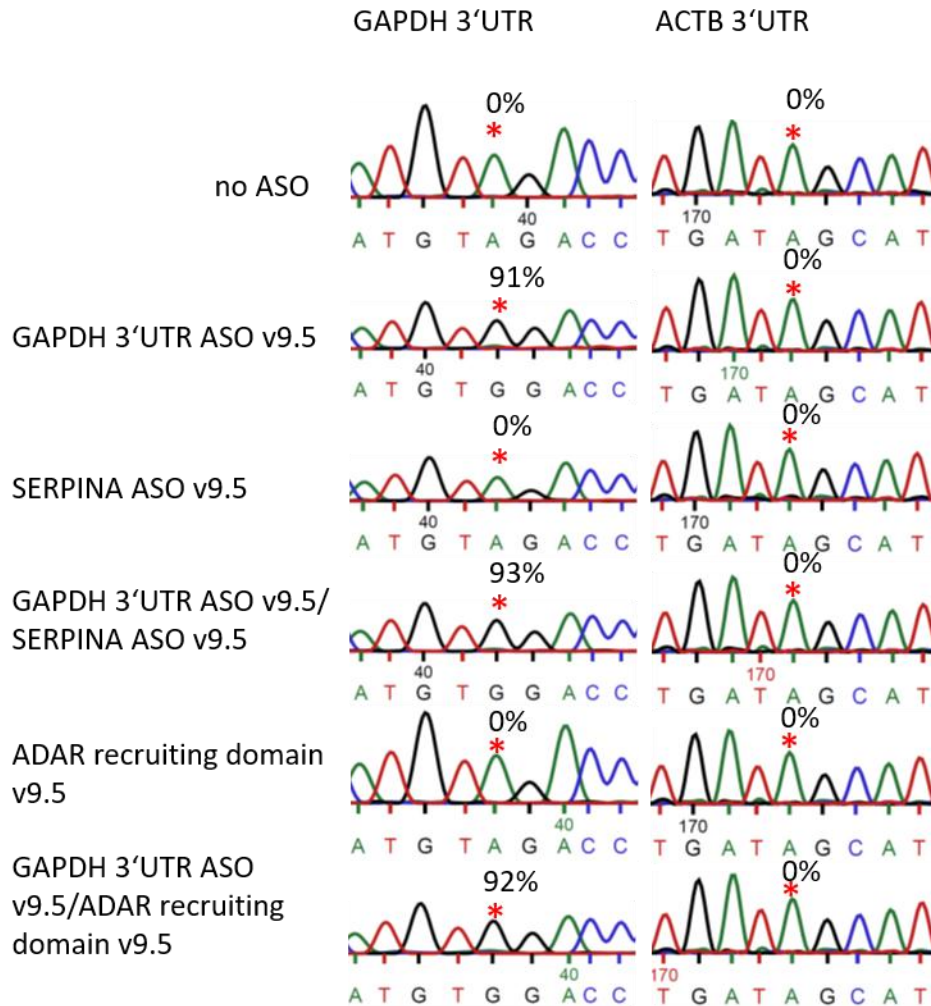
The western blot shown in Figure 2D in the manuscript was merged from images generated with two different exposure times. The part showing the ADAR bands comes from a 30 second exposure. The part showing β -actin from a 3 second exposure. The pictures were captured by the FusionCapt Advance SL4 (16.09b) software installed on the Fusion SL Vilber Lourmat (Vilber) western blot analyzer. No further image processing with respect to contrast or brightness was done. The western blot was done in technical duplicate.



Supplementary Figure 5

Determination of the effective dose (ED_{50}) of the respective ASO for editing *GAPDH* in the respective 293 Flp-In T-REx ADAR cells

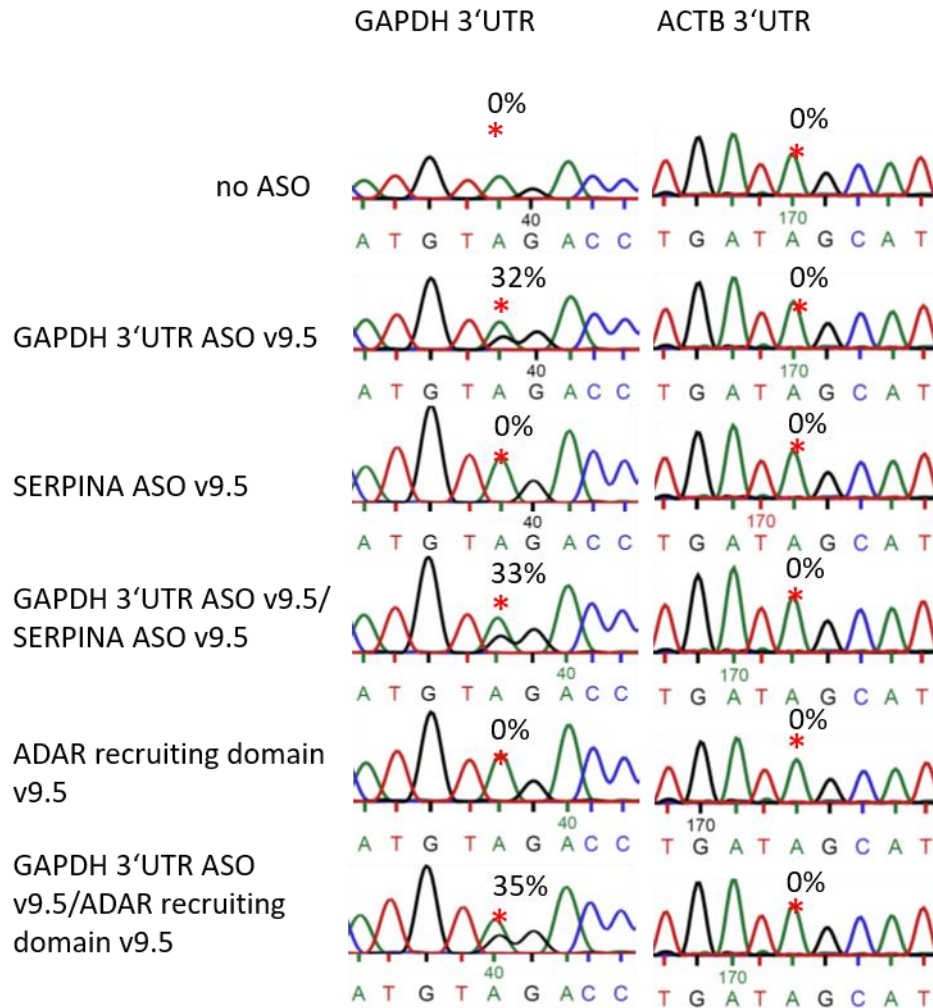
Shown is an experiment completely analog to that shown in the manuscript in Figure 2E, but in the indicated ADAR-expressing 293 Flp-In T-REx cell. Data are shown as the mean \pm SD, N=3 independent experiments.



Supplementary Figure 6

Effect of cotransfection of a nontargeting ASO v9.5 or the chemically stabilized ADAR-recruiting domain v9.5 alone on the *GAPDH* 3'-UTR editing with ASO v9.5 in ADAR1p150-expressing 293 Flp-In T-REx cells

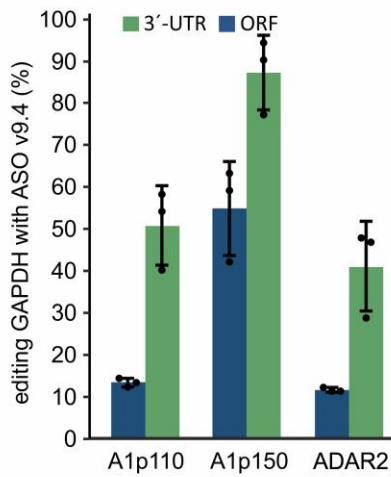
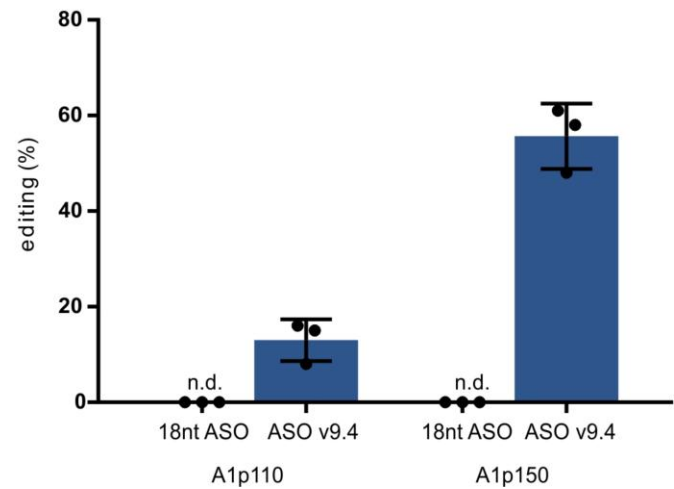
This is an additional control experiment. The on-target is the 5'-UAG codon in the 3'-UTR of *GAPDH*. A surveyed potential off-target is the 5'-UAG site in the 3'-UTR of *ACTB*. *SERPINA* ASO v9.5 acts as a non-targeting control, as the target (*SERPINA1*) is not expressed in this cell line. Another control is the ADAR-recruiting domain v9.5. This is the isolated, chemically stabilized ADAR-recruiting domain lacking any specificity domain. An ASO v9.5 against the on-target was co-transfected with either the non-targeting control or the control lacking a specificity domain. On-target editing requires the presences of the matching ASO. The surveyed potential off-target (*ACTB*) was not edited to detectable level under any condition. The on-target yield was not perturbed by the presence of the non-targeting ASO or the ADAR-recruiting domain alone, suggesting that only the combination of matching specificity and ADAR-recruiting domain enables site-directed RNA editing. It further suggests that the natural editing capacity is not limiting the editing reaction. (5 pmol ASO/96 well have been used)



Supplementary Figure 7

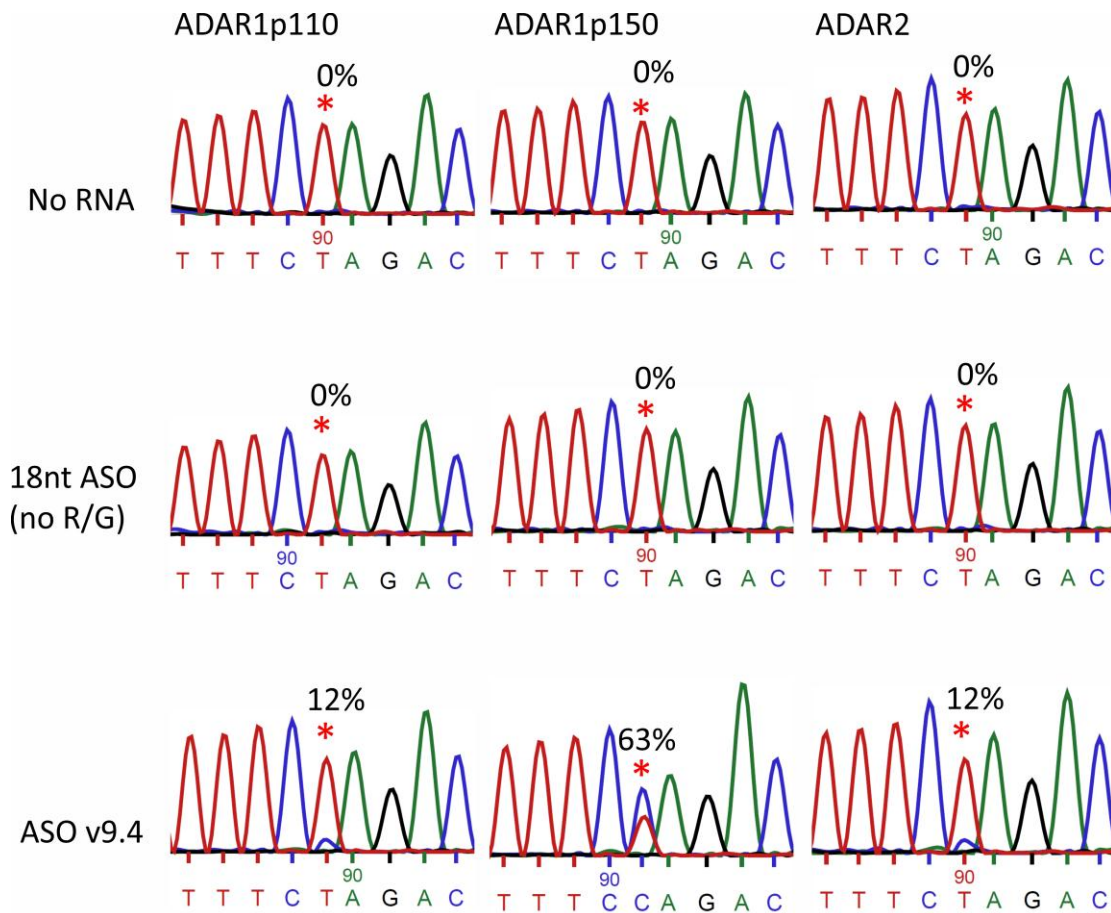
Effect of cotransfection of a nontargeting ASO v9.5 or the chemically stabilized ADAR-recruiting domain v9.5 alone on the *GAPDH* 3'-UTR editing with ASO v9.5, but for the recruitment of endogenous ADAR in HeLa cells without IFN- α .

This control experiment is the exact copy of the experiment shown in the preceding Supplementary Figure but was carried out in HeLa cells, recruiting endogenous ADAR. Exactly the same results have been observed and the same conclusions can be drawn.

A**B****Supplementary Figure 8**

Editing of 5' UAG codons in the ORF of GAPDH versus 3' UTR in ADAR-expressing 293 Flp-In T-REx cells

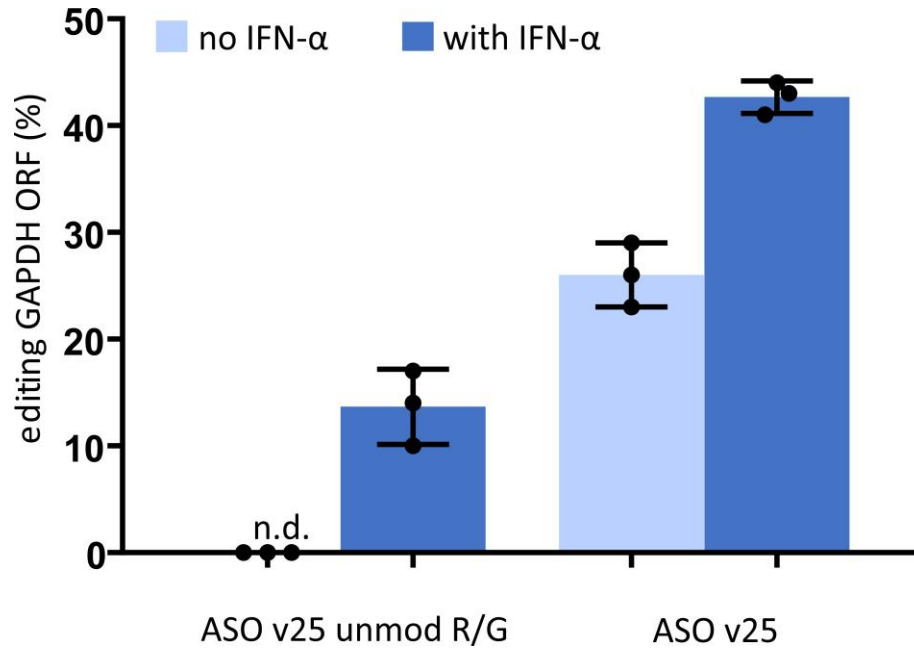
Editing of two different 5'-UAG codons in the ORF of GAPDH in 293 Flp-In T-REx ADAR cells (ORF #1 and #2). A) ORF site #2; here the comparison was made to the editing of the 5'-UAG codon in the 3'-UTR; and all three ADAR-expressing 293 Flp-In T-REx cell lines are included. B) The editing of the ORF site #1 was only tested in ADAR1-expressing Flp-In T-REx cell lines. The results are very similar. Further editing experiments, as shown in Figure 3B, target ORF site #1. Data in A) and B) are shown as the mean \pm SD, N=3 independent experiments. A1p110 = ADAR1p110; A1p150 = ADAR1p150.



Supplementary Figure 9

Sequencing traces for editing of a 5' UAG site (ORF site 2) in the ORF of *GAPDH* in 293 Flp-In T-REx ADAR cells

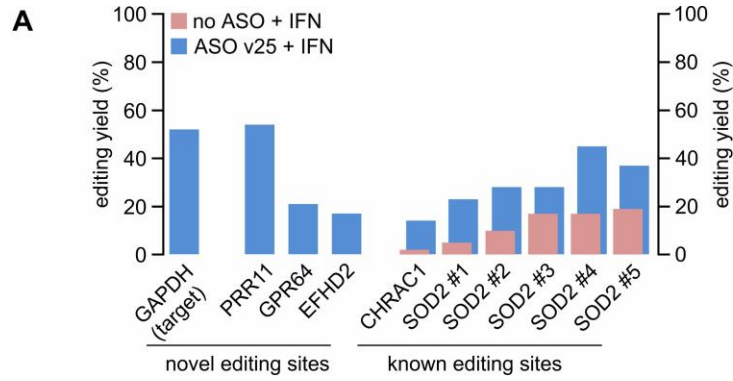
Exemplary editing traces for the editings shown in Supplementary Figure 8A, but including additional controls ("No RNA" = empty transfection; "18nt ASO no R/G" = ASO lacking the ADAR-recruiting domain. Red asterisks indicate the editing sites. A reverse primer was used for sequencing.



Supplementary Figure 10

Editing yields for the editing of a 5' UAG codon in the ORF of *GAPDH* in HeLa cells with ASO v25 containing a chemically unmodified versus modified ADAR-recruiting domain

Here, an ASO v25 with a chemically unmodified ADAR-recruiting domain (unmod R/G), was compared to an ASO of the same sequence with additional chemical modification (all pyrimidine nucleotides in the ADAR-recruiting domain are backbone 2'-O-methylated). ASOs were transfected in HeLa cells. Data are shown as the mean \pm SD, N=3 independent experiments.



B

| Edited site | Editing level | | Sequence |
|-------------|---------------|-------------|--|
| | ASO v25+IFN | no gRNA+IFN | |
| GAPDH | 53% | 0% | CCAACTGCTTAGCAGCCCTGGCCAAGGTCATCCATGACAA |
| PRR11 | 55% | 0% | TGCCAACTCTGCCTCACTCCATGAGATGATCCATAACAAATACAGATAAAAAACCCAGCTGGGTGCAG |
| GPR64 | 21% | 0% | TGCACCAAGATGTTAGCTTTATACCTTGGCCACAGAGAGGGATGAAGTG |
| EFHD2 | 17% | 0% | GAGACACCGCTGCTTAGCACCCCGAGCCAGAACACCCCTGAGGGTCTCGGG |
| CHRC1 | 14% | 2% | ATTTTGTATTTTTAGACGGGGTTCAACATGTTGGCCAGCCTTGCT |
| SOD2#1 | 23% | 5% | ATGCTACCATGCCTGGCTACTTTTTGTAATTTTTAGCAGACAGGGTTTCAACATGTTGGCCAGGGTG |
| SOD2#2 | 28% | 10% | TGGCTCAGCCTCCCAAGTGCTGGGATTAACAGGTGTGAGCCACTGCACCT |
| SOD2#3 | 28% | 17% | GGCCAAGGCAGGCACTTGAAGGAGTTCAAGACCAGGTGGCCAACTGGTGAAACCCCTGCT |
| SOD2#4 | 45% | 17% | GGCTCAGCCTCCCAAGGTGCTGGGATTAACAGGTGTGAGCCACTGCACCT |
| SOD2#5 | 37% | 19% | TGATCCGTCTGGCTCAGCTCCAAAGTGCTGGGATTACAGGTGTGAGCC |

C

GAPDH CCAACTGCTTAGCAGCCCTGGCCAAGGTCATCCATGACAA
 PRR11 TGCCAACCTCTGCCTCACTCCATG---AGATGATCCATACAAATACAGATAAAAAACCCAGCTGGGTGCAG

GAPDH CCAACTGCTTAGCAC----CCCTGGCCAAGGTCATCCATG-ACAA
 GPR64 TGCACCAAGATGTTTAGCTTTATACCTTGGCCACAGAGAGGGATGAAGTG

GAPDH CCAACTGCTTAGCAGCCCTGGCCAAGGTCATCCATGACAA
 EFHD2 GAGACACCGCTGCTTAGCACCCCGAGCC-AGAACACCC-TGAGGGTCTCGGG

GAPDH CCAACTGCTTAG-----CACCC--CTGGCCAAGGTCATCCATGACAA
 CHRC1 ATTTTGTATTTTTAGACGGGGTTCAACATGTTGGCCAGCCTTGCT

GAPDH CCAACTGC--TTAGCACCCCTGGCCAAG-----GTCCATCCATGACAA
 SOD2#1 ATGCTACCATGCCTGGCTACTTTTTGTAATTTTTAGCAGACAGGGTTTCAACATGTTGGCCAGGGTG

GAPDH CCAACTGCTTAGCACCC-----TGGCCAAGGTCATCCATGACAA
 SOD2#2 TGGCTCAGCCTCCCAAGTGCTGGGATTAACAGGTGTGAGCCACTGCACCT

GAPDH CCA-----ACT-----GCTTAGCACC--CTGGCCAAGGTCAT-----CCATGACAA
 SOD2#3 GGCCAAGGCAGGCACTTGAAGGAGTTCAAGACCAGCGTGGCCAACGTGGTGAAACCCCTGCT

GAPDH CCAACTGCTTAGCACCC-----TGGCCAAGGTCATCCATGACAA
 SOD2#4 GGCTCAGCCTCCCAAGGTGCTGGGATTAACAGGTGTGAGCCACTGCACCT

GAPDH CCAACTGCTTAGCAGCCCTGGCCAAGGT-----CATCCATGACAA
 SOD2#5 TGATCCGTCTGGCTCAGCTCCAAAGTGCTGGGATTACAGGTGTGAGCC

D

GAPDH CCAA-----CTGCTTAGCACCC-----
 SOD2 nt2100f . AGGAAACTCAAGATTCTCCTTTATTTCTGTGCTTGTGGGAATCCCTTGGCACACCCAAAGAGGGG

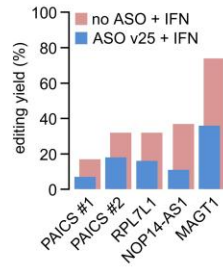
GAPDH -----TGGCCAAG-GTCATCCATGACAA
 SOD2 nt2100f . TCCCTGCTCCGTCTCACAGGGATCTTTTTGTATATTTGGCTTAGCATCATACA-----

Supplementary Figure 11

Analysis of off-target editing sites with increased editing yield upon ASO treatment

A) Besides the targeted site in GAPDH, 9 off-target editing sites were identified in ASO + IFN- α -treated cells compared to the control (no ASO + IFN- α). Six of them (CHARC1, SOD2 #1-#5) were known editing sites and found to be already edited in the control, N=2 independent experiments. **B)** and **C)** The regions around the off-target sites were aligned to the ASO-interacting region (40 nt) of the GAPDH transcript using MUSCLE (ebi.ac.uk/tools/msa/muscle/). The red A indicates the edited site and nucleotides matching to the target sequence of the ASO in GAPDH are highlighted in turquoise. The sequence alignment suggests that the editing at the three novel editing sites (PRR11, GPR64, EFHD2) is clearly induced by misguiding through the ASO. Notably, the strongest off-target (PRR11) might be controllable by further chemical modification of the specificity domain of the ASO. Four of the nine off-target sites (SOD2 #2-5) lack any strong homology to target region, also the edited codon is different from 5'-UAG. This makes it very unlikely that the off-target editing at such sites was induced by the ASO via direct binding to the off-target site, also because those sites all reside in secondary RNA structure (Alu elements). However, we found a potential ASO binding site in the 3'-UTR of SOD2 (panel **D)**) around nt 2100ff. (referring to NM_000636.3) that resides around 300 nt 5' to the first Alu element (nt 2380-2670) and around 1300 nt 5' to the second Alu element (nt 3400-3525). Since all SOD2 off-target sites reside in the two Alu elements one could imagine an ASO-dependent induction of the editing by either increase of the local ADAR concentration or by assisting the formation of an editable RNA secondary structure in the Alu element.

A



B

| Edited site | Editing level | | Sequence |
|-------------|---------------|------------|--|
| | ASO v25+IFN | no ASO+IFN | |
| GAPDH | 53% | 0% | CCAACTGCTTGCACCCCTGGCCAAGGTCATCCATGACAA |
| PAICS#1 | 7% | 17% | TAGCTGGGCACAGTGGCTCACACCTTCGATCACACACTTGGGAGGCTGAAGCAAGCAGATC |
| PAICS#2 | 18% | 32% | TGACAGGCACCTGTAATCCAGCTACTCGGAGGCTGAGGCATGAGAATTGAA |
| RPL7L1 | 16% | 32% | AGAAATGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAAGCCGAGATCACCATGCACTCCCTCTGGGCAACAAAG |
| NOP14-AS1 | 11% | 37% | ACCAAGCTGGGCAACACAGGAAGACCCCTCTCTACCATAAAAAACATAAAAATTAGCCAAGTGTGGTGGCAC |
| MAGT1 | 36% | 74% | GAGGCCAGGAGTTAGAGACCAGCCTGGCCAAACAAGGCATAACCCCGTCTCTACTAAATATGAAAATTAGCTGG |

C

```

GAPDH          CCAACTGCTTGCACCCCTGGCCAAGGTCATCCATGACAA-----CAA
PAICS#1        TAGCTGGGCACAGTGGCTCACACCTTCGATCACACACTTGGGAGGCTGAAGCAAGCAGATC

GAPDH          CCAACTGCTTGCACCCCTGGCCAAGGTCATCCATGACAA
PAICS#2        TGACAGGCACCTGTAATCCAGCTACTCGGAGGCTGAGGCATGAGAATTGAA

GAPDH          CCAA-CTGCTTGCAG-----CCTGGCC-----AAGGTCATCCA-----TGACAA
RPL7L1         AGAAATGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAAGCCGAGATCACCATGCACTCCCTCTGGGCAACAAAG

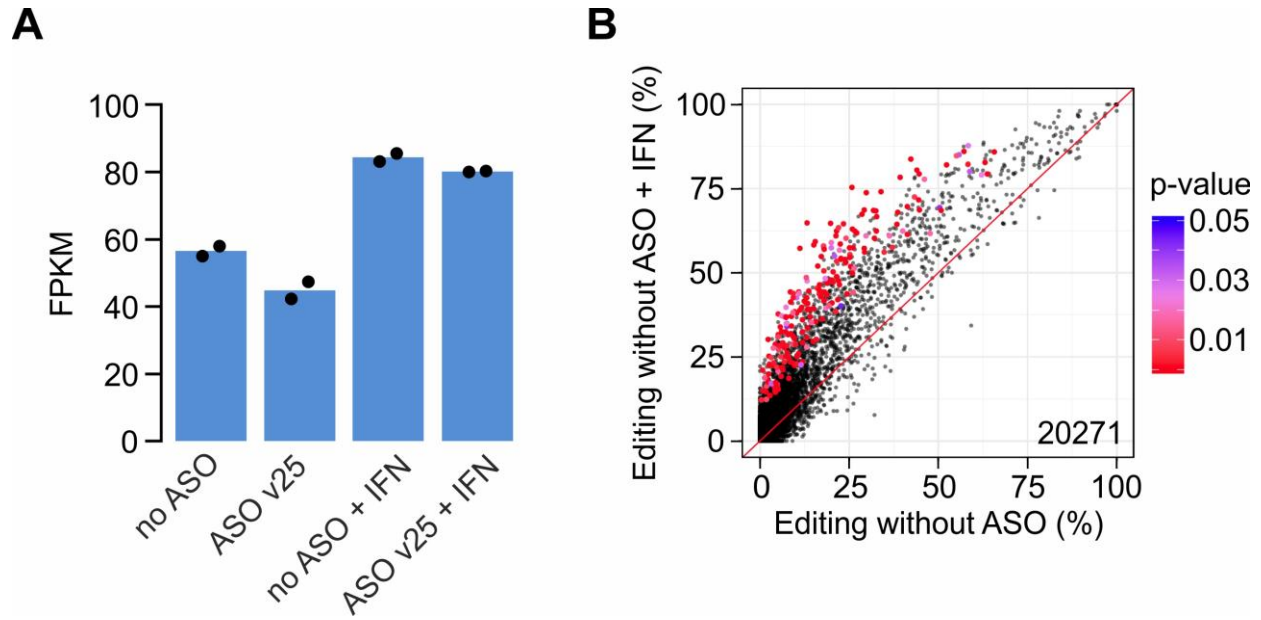
GAPDH          CCAA-CTGCTTGCAG-----CCTGGCC-----AAGGTCATCCA-----TGACAA
NOP14-AS1     ACCAAGCTGGGCAACACAGGAAGACCCCTCTCTACCATAAAAAACATAAAAATTAGCCAAGTGTGGTGGCAC

GAPDH          CCAA-CTGCTTGCAG-----CCTGGCC-----AAGGTCATCCA-----TGACAA
MAGT1         GAGGCCAGGAGTTAGAGACCAGCCTGGCCAAACAAGGCATAACCCCGTCTCTACTAAATATGAAAATTAGCTGG
  
```

Supplementary Figure 12

Analysis of off-target editing sites with attenuated editing upon ASO treatment

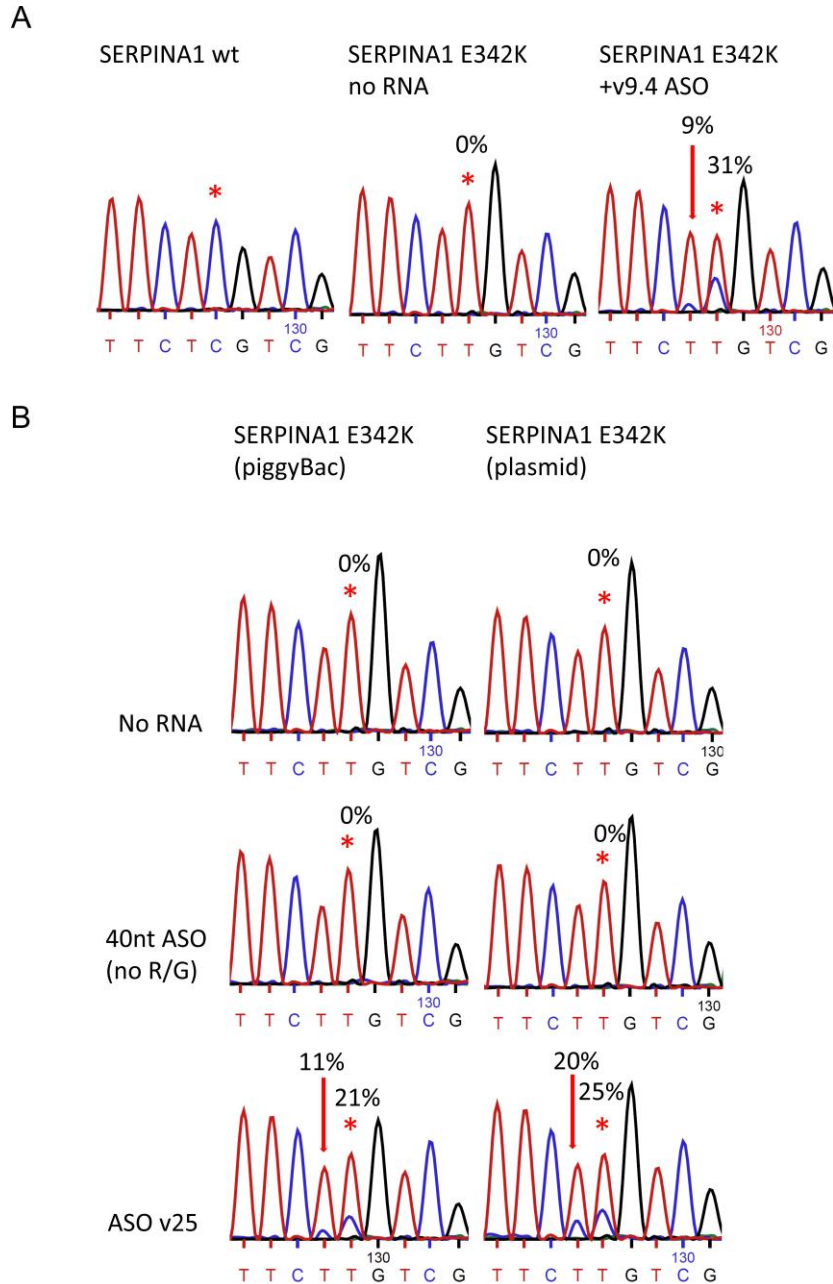
A) Five editing sites, all located in Alu sequences, were found to be significantly less edited in ASO-transfected, IFN- α -treated cells compared to the control lacking ASO transfection (but treated with IFN- α), N=2 independent experiments. **B)** and **C)** The regions around the off-target sites were aligned to the ASO-interacting region (40 nt) of the GAPDH transcript using MUSCLE (ebi.ac.uk/tools/msa/muscle/). The red A indicates the respective edited site and nucleotides matching with the ASO target sequence on GAPDH are highlighted in turquoise. For the most strongly affected site (MAGT1), but also for the other four sites, the ASO seems to be able to bind tightly in proximity to the respective editing sites and therefore interrupt the dsRNA secondary structure of the Alu repeat, which is required for editing. This suggests that the attenuated editing found at those sites is caused by direct interaction of the ASO with the off-target transcript and is not due to a global sequestering of the ADAR enzyme by the ASO.



Supplementary Figure 13

Effect of IFN- α and ASO treatment on ADAR1 expression and the natural editing homeostasis

A) FPKM values describing overall ADAR1 (p110+p150) expression following IFN- α treatment and ASO administration. IFN- α treatment induced ADAR1 expression in HeLa cells in a similar manner independent of ASO transfection. N=2 independent experiments. **B**) Analysis of significantly differently edited sites after IFN- α treatment in HeLa cells (no ASO transfection). Editing appears globally increased following IFN- α treatment. Significance of 20271 edited sites was tested using Fisher's exact test (two-sided, $p < 0.01$, $N \geq 50$); 116 sites were detected as significantly differently edited. The NGS experiment was done in independent duplicate.

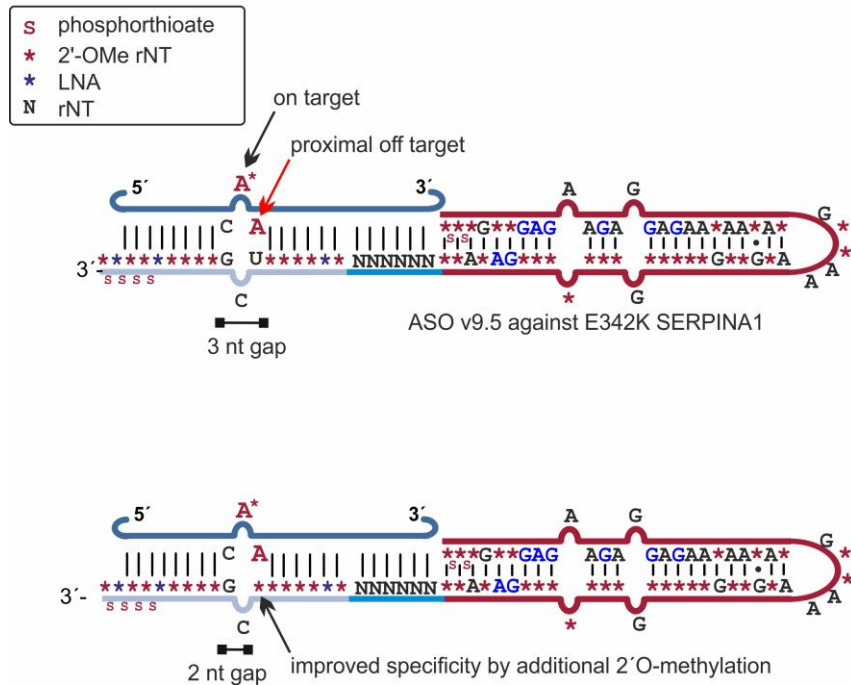


Supplementary Figure 14

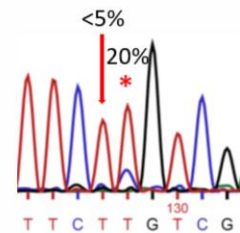
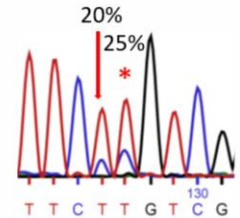
Sequencing traces of editing the PiZZ mutation in *SERPINA1*, showing on- and off-target editing in the A-rich 5' CAA codon

Exemplary sequencing traces of the 3 experimental conditions shown in Figure 3E of the manuscript. Red arrows indicate off-target, red asterisks indicate on-target editing sites, a reverse primer was used for sequencing. Shown are additional controls for empty transfection, and for transfecting an ASO lacking the ADAR-recruiting domain (no R/G). A) Editing in ADAR1p150-expressing 293 Flp-In T-REx cells; B) Editing of *SERPINA1* PiZZ in HeLa cells expressing *SERPINA1* PiZZ either genomically integrated (piggyBac) or transiently overexpressed (plasmid). In particular ASO v25 gave substantial off-target editing with the proximal adenosine in the targeted 5'-CAA codon.

ASO design



editing result



Supplementary Figure 15

Improvement of editing specificity in the ASO:mRNA hybrid

Editing of the 5'-CAA codon to restore the E342K mutation in SERPINA1 (a reverse primer was used here!) comes along with off target editing at the nearest neighboring adenosine (see also preceding Supplementary Figure, panel B). To reduce proximal off-target editing, the 3 nt gap in the modification pattern of the ASO was reduced to a 2 nt gap by putting an additional chemical modification (2'-O-methyl uridine) opposite the off-target nucleotide. Two representative sequencing traces were selected from three very similar replicates that show that the additional chemical modification strongly reduces the proximal off-target edit while only modestly influencing the on target editing. (editing was performed in HeLa cells with SERPINA1 PiZZ cDNA overexpressed from a plasmid, 5 pmol/96 well ASO was transfected)

Supplementary Notes and Tables

Supplementary Table 1: List of guideRNAs, ASOs and further oligonucleotides used in this study.

(N)=RNA base, [N]=2'-OMe RNA base, *=Phosphorothioate linkage, {N}=LNA base.

| R/G guide RNAs expressed from plasmid | 5'-3' sequence | Figure |
|--|--|---|
| Luciferase R/G-v1 | (GUGGAAUAGUAUAACAUAUGCUAAAUGUUGUUAUAGUAUCCACGUGCAGC CAGCCGUCCUCUAGAGGGCCUGAAGAGGGCCC) | SI1 |
| Luciferase R/G-v4 | (GUGGAAGAGGAGACAUAUGCUAAAUGUUGUUCUCGUCUCCACGUGCAGC CAGCCGUCCUCUAGAGGGCCUGAAGAGGGCCC) | SI1 |
| Luciferase R/G-v9.4 | (GUGGUCGAGAAGAGGAGAACAUAUGCUAAAUGUUGUUCUCGUCUCCUCGAC CACGUGCAGCCAGCCGUCUAGAGGGCCUGAAGAGGGCCC) | SI1 |
| Chemically synthesized ASOs | 5'-3' sequence | Figure |
| ACTB 3'UTR 18nt | [GC AAU G] (CCA) [UC AC] [C*][U*][C*][C] Propandiol | 2A |
| ACTB 3'UTR ASO v1 | (GGUGA AUAGUAUAC AAUAGCUAA AUGUUGUUAU AGUAUCCACC) [GC AAU G] (CCA) [UC AC] [C*][U*][C*][C] Propandiol | 1C,2A |
| ACTB 3'UTR ASO v4 | (GGUGAAG AGGAGAACA UAUGCUAAA GUUGUUCUCG UCUCACC)[GC AAU G](CCA) [UC AC][C*][U*][C*][C] Propandiol | 1C,2A |
| ACTB 3'UTR ASO v9.4 | (GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC) [GC AAU G] (CCA) [UC AC] [C*][U*][C*][C] Propandiol | 1C,2A |
| GAPDH 3'UTR 18nt | [AG GGG U] (CCA) [CA UG] [G*][C*][A*][A*][C] Propandiol | 2A, SI2 |
| GAPDH 3'UTR ASO v1 | (GGUGA AUAGUAUAC AAUAGCUAA AUGUUGUUAU AGUAUCCACC) [AG GGG U] (CCA) [CA UG] [G*][C*][A*][A*][C] Propandiol | 1C,2A, SI2 |
| GAPDH 3'UTR ASO v4 | (GGUGAAG AGGAGAACA UAUGCUAAA GUUGUUCUCG UCUCACC)[AG GGG U](CCA)[CA UG][G*][C*][A*][A*][C] Propandiol | 1C,2A, SI2 |
| GAPDH 3'UTR ASO v9.4 | (GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC) [AG GGG U] (CCA) [CA UG] [G*][C*][A*][A*][C] Propandiol | 1C,2A,2B, SI2, SI5 |
| GAPDH 3'UTR ASO v9.5 | [G*][G*][U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][G][UUCUC](G)[UCUCCUC](G A)[C](A) [CCAGGGU] (CCA) [CAUG][G*][C*][A*] [A*][C] | 2B, 2C, 2E, 2F, 2G, 2H, SI5, SI6, SI7, SI8A |
| GAPDH ORF1 ASO 18nt | [GGG GUG] (CCA) [AG CA] [G*][U*][U*][G*][G] Propandiol | SI8B |
| GAPDH ORF1 ASO v9.4 | (GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC)[GGG GUG](CCA)[AG CA] [G*][U*][U*][G*][G] Propandiol | SI8B |
| GAPDH ORF2 ASO 18nt | [GGG GUG](CCA)[AG CA] [G*][U*][U*][G*][G] Propandiol | SI9 |
| GAPDH ORF2 ASO v9.4 | (GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC)[GU UUU U] (CCA) [GA CG] [G*][C*][A*][G*][G] Propandiol | SI8A, SI9 |
| GAPDH ORF1 ASO v25 | [G*][G*][U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C] [U][C][C] [U][C](G A)[C](A) [C][C] (UUGUCAUGGAUGACCUU GGCCA) [G] {G} [GG UG] (CCA) [AGCA] {G*}[U*][U*]{G*}[G] Aminolinker | 3B,3C, SI10-13 |
| GAPDH ORF1 ASO R/G unmod v25 | [G*][G*][U] (GUCGAG AAG AGG AGA ACAUAUGCUA AAUGUUGUUCUCGUCUCCUCG ACACC UUGUCAUGGAUGACCUU GGCCA) [G] {G} [GG UG] (CCA) [AGCA] {G*}[U*][U*]{G*}[G] Aminolinker | SI10 |
| SERPINA ASO v9.4 | (GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CC) [CCU UUC] (UCG) [UCG A] [U*][G*][G*][U*][C] Propandiol | 3E, SI6, SI7, SI14A |
| SERPINA ASO 40nt | (CAUGGCCCCAGCAGCUUCAGUC) [C] {C}[UUUC] (UCG) [UCGA]{T*}[G*] [G*] {T*} [C] Aminolinker | SI14B |
| SERPINA ASO v25 | [G*] [G*] [U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C] [U][C][C] [U][C](G A C A C C CAUGGCCCCAGCAGCUUCAGUC) [C] {C}[UUUC] (UCG) [UCGA]{T*}[G*] [G*] {T*} [C] Aminolinker | 3E, SI14B, SI15 |
| STAT1 ASO v25 | [G*] [G*] [U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C][U][C][C] [U][C](GACACCCA GACACAGAAAUAACUCAGU) [C] {T} [UGAU] (ACA) [UCCA] {G*} [U*] [C*][C] Aminolinker | 3D |
| GAPDH 3'UTR unmod ASO v1 | (GGUGA AUAGUAUAC AAUAGCUAA AUGUUGUUAU AGUAUCCACC AG GGG UCCACA UG GCAAC) | SI2, SI3 |
| GAPDH 3'UTR unmod ASO v4 | (GGUGAAG AGGAGAACA UAUGCUAAA GUUGUUCUCG UCUCACCAG GGG UCCACA UGCAAC) | SI2, SI3 |
| GAPDH 3'UTR unmod ASO v9.4 | (GGU GUC GAG AAG AGG AGA AC AAU AUG CUA AAU GUU GUU CUCGUC UCC UCG ACA CCAG GGG U CCACA UGCAAC) | SI2, SI3 |
| SERPINA ASO v25 2nt gap also called ASO v25.1 in Fig. 3E | [G*] [G*] [U] (G)[U][C] (GAG AAG AGG AGA A)[C] (AA) [U] (A)[U](G) [C][U](A AA)[U] (G)[U][U] (G)[U][U] [C][U][C](G)[U][C] [U][C][C] [U][C](G A C A C C CAUGGCCCCAGCAGCUUCAGUC) [C] {C}[UUUCU] (CG) [UCGA]{T*}[G*] [G*] {T*} [C] Aminolinker | 3E, SI15 |
| Sense guideRNAs for RT PCR | 5'-3' sequence | Figure |
| GAPDH sense | (GGACCAACUGCUUGGCACCCUGGCCAAGGUCAUCCAUGACAACUUGGUAUCUGGGAAGGACC) | 3B, 3C |

| | | |
|----------------|--|----------------|
| STAT1 sense | (GGGAACUGGAUCUAUCAAGACUGAGUUGAUUUUCUGUGUCUGAAGUGUAAGUGAACACAGAA) | 3D |
| SERPINA1 sense | (GGACCATCGACGAGAAAGGGACTGAAGCTGCTGGGGCCATGTTTTAGAGGCCATACCCAT) | 3E,SI14B, SI15 |

Supplementary Table 2: List of the antibodies used to generate the western blot illustrated in figure 2D.

| Antibody | Target Protein | Produced in | Immunoglobulin Class | Dilution used | Supplier | Order # | Against | Validation |
|----------------|----------------------|-------------|----------------------|---------------|---------------|----------|--|--|
| ADAR1 antibody | α -ADAR1 | Mouse | monoclonal IgG | 1:1000 | Santa Cruz | sc-73408 | amino acids 440-826 corresponding to the middle region of ADAR1 of human origin | Validated in our lab via siRNA KO and Western Blot PMID: # 28669490 PMID: # 28278381 PMID: # 27573237 |
| ADAR2 antibody | α -ADAR2 | Mouse | monoclonal IgG | 1:1000 | Santa Cruz | sc-73409 | N-terminal region corresponding to amino acids 2-179 of ADAR2 of human origin | Validated in our lab via overexpression and Western Blot PMID: # 26601943 PMID: # 24345557 PMID: # 27907896 |
| Clone AC-15 | α -Beta-Actin | Mouse | monoclonal IgG | 1:40.000 | Sigma Aldrich | A5441 | Actin N-terminal peptide, Ac-Asp-Asp-Asp-Ile-Ala-Ala-Leu-Val-Ile-Asp-Asn-Gly-Ser-Gly-Lys | PMID: # 15809369 PMID: # 15048076 PMID: # 21217779 |

Supplementary Note 1. List of gene sequences and target sequences

Sequence of dual Luciferase Renilla 2A Firefly W417X reporter cDNA with chosen editing site (Firefly-Luciferase W417X, yellow).

```

1         10         20         30         40         50         60
1   ATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGG
1   M T S K V Y D P E Q R K R M I T G P Q W

        70         80         90        100        110        120
61  TGGGCCAGATGTAACAAATGAATGTTCTTGATTCAATTATTAATTATTATGATTCAGAA
21  W A R C K Q M N V L D S F I N Y Y D S E

        130        140        150        160        170        180
121 AAACATGCAGAAAATGCTGTTATTTTTTTACATGGTAACGCGGCCTCTTCTTATTTATGG
41  K H A E N A V I F L H G N A A S S Y L W

        190        200        210        220        230        240
181 CGACATGTTGTGCCACATATTGAGCCAGTAGCGCGGTGTATTATACCAGACCTTATGGT
61  R H V V P H I E P V A R C I I P D L I G

        250        260        270        280        290        300
241 ATGGGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATAT
81  M G K S G K S G N G S Y R L L D H Y K Y

        310        320        330        340        350        360
301 CTTACTGCATGGTTTGAACCTTCTTAATTTACCAAAGAAGATCATTTTGTGCGCCATGAT
101 L T A W F E L L N L P K K I I F V G H D

        370        380        390        400        410        420
361 TGGGGTCTTGTGGTGGCATTTCATTATAGCTATGAGCATCAAGATAAGATCAAAGCAATA
121 W G A C L A F H Y S Y E H Q D K I K A I

        430        440        450        460        470        480
421 GTTCACGCTGAAAGTGTAGTAGATGTGATTGAATCATGGGATGAATGGCCTGATATGAA
141 V H A E S V V D V I E S W D E W P D I E

        490        500        510        520        530        540
481 GAAGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTC
161 E D I A L I K S E E G E K M V L E N N F

        550        560        570        580        590        600
541 TTCGTGAAACCATGTTGCCATCAAAAATCATGAGAAAGTTAGAACCAGAAGAATTGCA
181 F V E T M L P S K I M R K L E P E E F A

        610        620        630        640        650        660
601 GCATATCTTGAACCATCAAAGAGAAAAGGTGAAGTTCGTCCGTCACATTATCATGGCCT
201 A Y L E P F K E K G E V R R P T L S W P

        670        680        690        700        710        720
661 CGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAAATGTTAGGAATTAT
221 R E I P L V K G G K P D V V Q I V R N Y

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

        790        800        810        820        830        840
781 TTCTTTTCCAATGCTATGTTGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAA
261 F F S N A I V E G A K K F P N T E F V K

        850        860        870        880        890        900
841 GTAAAAGTCTTCATTTTTTCGCAAGAAGATGCACCTGATGAAATGGGAAAATATATCAA
281 V K G L H F S Q E D A P D E M G K Y I K

        910        920        930        940        950        960
901 TCGTTCGTTGAGCGAGTTCTCAAAAATGAACAAGGAAGCGGAGCTACTAAGTTCAGCCTG
301 S F V E R V L K N E Q G S G A T N F S L

        970        980        990        1000       1010       1020
961 CTGAAGCAGGCTGGAGACGTGGAGGAGAACCCTGGACCTATGGAAGATGCCAAAACATT
```


321 L K Q A G D V E E N P G P M E D A K N I
 1030 1040 1050 1060 1070 1080
 1021 AAGAAGGGCCAGCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCAC
 341 K K G P A P F Y P L E D G T A G E Q L H
 1090 1100 1110 1120 1130 1140
 1081 AAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATC
 361 K A M K R Y A L V P G T I A F T D A H I
 1150 1160 1170 1180 1190 1200
 1141 GAGGTGGACATTACCTACGCCGAGTACTTCGAGATGAGCGTTTCGGCTGGCAGAAGCTATG
 381 E V D I T Y A E Y F E M S V R L A E A M
 1210 1220 1230 1240 1250 1260
 1201 AAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCAGCGAGAATAGCTTGCAG
 401 K R Y G L N T N H R I V V C S E N S L Q
 1270 1280 1290 1300 1310 1320
 1261 TTCTTCATGCCCGTGTGGGTGCCCTGTTTCATCGGTGTGGCTGTGGCCCCAGCTAACGAC
 421 F F M P V L G A L F I G V A V A P A N D
 1330 1340 1350 1360 1370 1380
 1321 ATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCACCCTCGTATTC
 441 I Y N E R E L L N S M G I S Q P T V V F
 1390 1400 1410 1420 1430 1440
 1381 GTGAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAA
 461 V S K K G L Q K I L N V Q K K L P I I Q
 1450 1460 1470 1480 1490 1500
 1441 AAGATCATCATATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTC
 481 K I I I M D S K T D Y Q G F Q S M Y T F
 1510 1520 1530 1540 1550 1560
 1501 GTGACTTCCCATTGGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAGCTTCGAC
 501 V T S H L P P G F N E Y D F V P E S F D
 1570 1580 1590 1600 1610 1620
 1561 CGGGACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCCAAGGGC
 521 R D K T I A L I M N S S G S T G L P K G
 1630 1640 1650 1660 1670 1680
 1621 GTAGCCCTACCGCACCGCACCCTTGTGTCCGATTCAAGTATGCCCCGACCCCATCTTC
 541 V A L P H R T A C V R F S H A R D P I F
 1690 1700 1710 1720 1730 1740
 1681 GGCAACAGATCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTACCACGGCTTC
 561 G N Q I I P D T A I L S V V P F H H G F
 1750 1760 1770 1780 1790 1800
 1741 GGCATGTTACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTACCGC
 581 G M F T T L G Y L I C G F R V V L M Y R
 1810 1820 1830 1840 1850 1860
 1801 TTCGAGGAGGACTATCTTGCGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTGCTG
 601 F E E E L F L R S L Q D Y K I Q S A L L
 1870 1880 1890 1900 1910 1920
 1861 GTGCCCACACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTACGACCTAAGC
 621 V P T L F S F F A K S T L I D K Y D L S
 1930 1940 1950 1960 1970 1980
 1921 AACTTGCACGAGATCGCCAGCGGGCGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCCGTG
 641 N L H E I A S G G A P L S K E V G E A V
 1990 2000 2010 2020 2030 2040
 1981 GCCAAACGCTTCCACCTACCAGGCATCCGCCAGGGCTACGGCTGACAGAAACAACAGC
 661 A K R F H L P G I R Q G Y G L T E T T S
 2050 2060 2070 2080 2090 2100
 2041 GCCATTCTGATCACCCCGAAGGGGACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGCCC
 681 A I L I T P E G D D K P G A V G K V V P

```

2110      2120      2130      2140      2150      2160
2101    TTCTTCGAGGCTAAGGTGGTGGACTTGGACACCGGTAAGACACTGGGTGTGAACCAGCGC
701      F F E A K V V D L D T G K T L G V N Q R

2170      2180      2190      2200      2210      2220
2161    GGCGAGGTGTGCGTCCGTGGCCCATGATCATGAGCGGCTACGTTAACAACCCCGAGGCT
721      G E L C V R G P M I M S G Y V N N P E A

2230      2240      2250      2260      2270      2280
2221    ACAAACGCTCTCATCGACAAGGACGGCTAGCTGCACAGCGCGACATCGCCTACTGGGAC
741      T N A L I D K D G * L H S G D I A Y W D

2290      2300      2310      2320      2330      2340
2281    GAGGACGAGCACTTCTTCATCGTGGACCGGCTGAAGAGCCTGATCAAATACAAGGGCTAC
761      E D E H F F I V D R L K S L I K Y K G Y

2350      2360      2370      2380      2390      2400
2341    CAGGTAGCCCCAGCCGAAGTGGAGAGCATCCTGCTGCAACACCCCAACATCTTCGAGGCC
781      Q V A P A E L E S I L L Q H P N I F D A

2410      2420      2430      2440      2450      2460
2401    GGGGTGCGCCGGCCTGCCCGACGACGATGCCGGCGAGCTGCCCGCCGAGTCGTCGTGCTG
801      G V A G L P D D D A G E L P A A V V V L

2470      2480      2490      2500      2510      2520
2461    GAACACGGTAAAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTCAACA
821      E H G K T M T E K E I V D Y V A S Q V T

2530      2540      2550      2560      2570      2580
2521    ACCGCCAAGAAGTGC CGCGGTGGTGTGTTGCTGGACGAGGTGCCTAAAGGACTGACC
841      T A K K L R G G V V F V D E V P K G L T

2590      2600      2610      2620      2630      2640
2581    GGCAAGTTGGACGCCCGCAAGATCCGCGAGATTCTCATTAAAGCCAAGAAGGGCGGCAAG
861      G K L D A R K I R E I L I K A K K G G K

2650
2641    ATCGCCGTGTAA
881      I A V *

```

Sequence of GAPDH mRNA isoform 1 (NM_002046.5) with chosen editing sites (yellow).

```

10      20      30      40      50      60
1      GCCTCAAGACCTTGGGCTGGGACTGGCTGAGCCTGGCGGGAGGCGGGTCCGAGTCACCG
1

70      80      90      100     110     120
61     CCTGCCGCCGCGCCCCGGTTTCTATAAAATTGAGCCCGCAGCCTCCCGCTTCGCTCTCTG
20

130     140     150     160     170     180
121    CTCTCTGTTCGACAGTCAGCCGCATCTTCTTTTGGCTGCCAGCCGAGCCACATCGCT
40

190     200     210     220     230     240
181    CAGACACCATGGGGAAGGTGAAGGTCGGAGTCAACGGATTTGGTCGTATTGGGCGCCTGG
60      M G K V K V G V N G F G R I G R L

250     260     270     280     290     300
241    TCACCAGGCTGCTTTTAACCTCTGGTAAAGTGGATATGTTGCCATCAATGACCCCTTCA
80      V T R A A F N S G K V D I V A I N D P F

310     320     330     340     350     360
301    TTGACCTCAACTACATGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATCCATG
100     I D L N Y M V Y M F Q Y D S T H G K F H

370     380     390     400     410     420
361    GCACCGTCAAGGCTGAGAACGGGAAGCTTGTCATCAATGGAATCCCATCACCATCTTCC
120     G T V K A E N G K L V I N G N P I T I F

430     440     450     460     470     480
421    AGGAGCGAGATCCCTCCTCAAAAATCAAGTGGGGCGATGCTGGCGCTGAGTACGTCGTGGAGT
140     Q E R D P S K I K W G D A G A E Y V V E

490     500     510     520     530     540
481    CCACTGGCGTCTTACCACCATGGAGAAGGCTGGGGCTCATTGTCAGGGGGGAGCCAAAA
160     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  GGGTCATCATCTCTGCCCCCTCTGCTGATGCCCCCATGTTTCGTCATGGGTGTGAACCATG
601          610          620          630          640          650          ORF site #1
200  R V I I S A P S A D A P M F V M G V N H
    AGAAGTATGACAACAGCCTCAAGATCATCAGCAATGCCTCCTGCACCACCAACTGCTTAG
661          670          680          690          700          710          720
220  E K Y D N S L K I I S N A S C T T N C L
    CACCCCTGGCCAAGGTATCCATGACAACCTTTGGTATCGTGAAGGACTCATGACCACAG
721          730          740          750          760          770          780
240  A P L A K V I H D N F G I V E G L M T T
    TCCATGCCATCACTGCCACCCAGAAGACTGTGGATGGCCCTCCGGGAAACTGTGGCGTG
781          790          800          810          820          830          840
260  V H A I T A T Q K T V D G P S G K L W R
    ATGGCCGCGGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAGGCTGTGG
841          850          860          870          880          890          900
280  D G R G A L Q N I I P A S T G A A K A V
    GCAAGTTCATCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCGCTGTCCCCACTG
901          910          920          930 ORF site #2          960
300  G K V I P E L N G K L T G M A F R V P T
    CCAACGTGTCAAGTGGTGGACCTGACCTGCCGCTCTAGAAAAACCTGCCAAATATGATGACA
961          970          980          990          1000          1010          1020
320  A N V S V V D L T C R L E K P A K Y D D
    TCAAGAAGGTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGCATCCTGGGCTACACTG
1021          1030          1040          1050          1060          1070          1080
340  I K K V V K Q A S E G P L K G I L G Y T
    AGCACCAGGTGGTCTCCTCTGACTTCAACAGCGACACCCACTCCTCCACCTTTGACGCTG
1081          1090          1100          1110          1120          1130          1140
360  E H Q V V S S D F N S D T H S S T F D A
    GGGCTGGCATTGCCCTCAACGACCACCTTTGTCAAGCTCATTTCTGGTATGACAACGAAT
1141          1150          1160          1170          1180          1190          1200
380  G A G I A L N D H F V K L I S W Y D N E
    TTGGTACAGCAACAGGGTGGTGGACCTCATGGCCACATGGCCTCCAAGGAGTAAGACC
1201          1210          1220          1230          1240          1250          1260
400  F G Y S N R V V D L M A H M A S K E *
    CCTGGACCACCAGCCCCAGCAAGAGCACAAGAGGAAGAGAGACCCCTCACTGCTGGGGA

          1270          1280          1290          1300          1310          3'-UTR site
1261  GTCCCTGCCACACTCAGTCCCCACCCACACTGAATCTCCCCTCCTCACAGTTGCCATGTA
420
          1330          1340          1350          1360          1370          1380
1321  GACCCCTTGAAGAGGGGAGGGGCTAGGGAGCCGACCTTGTTCATGTACCATCAATAAAG
440
          1390          1400          1410          1420
1381  TACCCTGTGCTCAACCAGTTAAAAAAAAAAAAAAAAAAAAAAAAA

```

Sequence of β -actin mRNA (NM_001101.3) with chosen editing site (yellow).

```

          10          20          30          40          50          60
1  ACCGCCGAGACCGCGTCCGCCCCGCGAGCACAGAGCCTCGCCTTTGCCGATCCGCCGCC
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  GTCCACACCCGCCGCCAGCTCACCATGGATGATGATATCGCCGCGCTCGTTCGACAAC
21  V H T R R Q L T M D D D I A A L V V D N
          130          140          150          160          170          180
121  GGCTCCGGCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGGCCGTCTTCCCC
41  G S G M C K A G F A G D D A P R A V F P
          190          200          210          220          230          240
181  TCCATCGTGGGCGCCCCAGGCACCAGGGCGTGATGGTGGGCATGGGTGAGAAGGATTCC
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  TATGTGGGCGACGAGGCCAGAGCAAGAGGCATCCTCACCTGAAGTACCCCATCGAG
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 AAGCCAAACCGCGAGAAGATGACCCAGATCATGTTTGTAGACCTTCAACACCCAGCCATG
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 TACGTTGCTATCCAGGCTGTGTATCCCTGTACGCCTCTGGCCGTACCACTGGCATCGTG
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 ATGGACTCCGGTGACGGGGTCAACCACTGTGCCATCTACGAGGGGTATGCCCTCCC
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 CATGCCATCCTGCGTCTGGACCTGGCTGGCCGGACCTGACTGACTACCTCATGAAGATC
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 CTCACCGAGCGCGCTACAGCTTACCACCAGCGCCGAGCGGAAATCGTGCCTGACATT
221 L T E R G Y S F T T A E R E I V R D I
730 740 750 760 770 780
721 AAGGAGAAGCTGTGCTACGTGCGCTGGACTTCGAGCAAGAGATGGCCACGGCTGCTTCC
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 GAGCGGTTCCGCTGCCCTGAGGCACTTCCAGCCTTCCCTTCTGGGCATGGAGTCCCTGT
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 CTGTACGCCAACACAGTGTCTGTGGGGCCACCACCATGTACCCTGGCATTGCCGACAGG
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 ATGCAGAAGGAGATCACTGCCCTGGCACCCAGCACAAATGAAGATCAAGATCATGTGCTCCT
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 CAGCAGATGTGGATCAGCAAGCAGGAGTATGACGAGTCCGGCCCCCTCCATCGTCCACCGC
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 AAATGCTTCTAGGCGGACTATGACTTAGTTGCGTTACACCCTTCTTGTACAAAACCTAAC
401 K C F *
1270 1280 1290 1300 1310 1320
1261 TTGCGCAGAAAACAAGATGAGATTGGCATGGCTTTATTGTTTTTTTTGTGTTTTGG
421
1330 1340 1350 1360 1370 1380
1321 TTTTTTTTTTTTTTTGGCTTACTCAGGATTTAAAACTGGAACGGTGAAGGTGACAGC
441
1390 1400 1410 1420 1430 1440
1381 AGTCGGTTGGAGCGAGCATCCCCAAAGTTCACAATGTGGCCGAGGACTTTGATTGCACA
461
1450 1460 1470 1480 1490 1500
1441 TTGTTGTTTTTTTAAATAGTCATTCCAAATATGAGATGCGTTGTACAGGAAGTCCCTTGC
481 C
1510 1520 1530 1540 1550 1560
1501 CATCCTAAAAGCCACCCACTTCTCTCTAAGGAGAATGGCCCAGTCTCTCCAAGTCCA
501
1570 3'-UTR site 1600 1610 1620
1561 CACAGGGGAGGTGATAGCATTGCTTTTCGTGTAATAATGTAATGCAAAATTTTTTAATC
521
1630 1640 1650 1660 1670 1680
1621 TTCGCCTTAATACTTTTTTATTGTTTTATTTTGAATGATGAGCCTTCGTGCCCCCCT
541 P
1690 1700 1710 1720 1730 1740
1681 TCCCCCTTTTTTGTCCCCCAACTTGAGATGTATGAAGGCTTTTGGTCTCCCTGGGAGTGG
561
1750 1760 1770 1780 1790 1800
1741 GTGGAGGCAGCCAGGGCTTACCTGTACACTGACTTGAGACCAGTTGAATAAAAGTGACA
581
1810 1820 1830 1840 1850
1801 CCTTAAAATGAAA
601

Sequence of STAT1 mRNA (NM_007315.3) with chosen editing site Y701 (yellow).

```

10      20      30      40      50      60
1      GCTGAGCGCGGAGCCGCCCGGTGATTGGTGGGGCGGAAGGGGGCCGGGCGCCAGCGCTG
1
70      80      90      100     110     120
61     CCTTTTCTCCTGCCGGGTAGTTTCGCTTTCCTGCGCAGAGTCTGCGGAGGGGCTCGGCTG
21
130     140     150     160     170     180
121    CACCGGGGGGATCGCGCTGGCAGACCCAGACCCAGAGCGAGGCGACCCAGCGCGCTCGG
41
190     200     210     220     230     240
181    GAGAGGCTGCACCGCCCGCCCGCCTAGCCCTTCCGGATCCTGCGCGCAGAAAAGTTT
61
250     260     270     280     290     300
241    CATTGTGTATGCCATCCTCGAGAGCTGTCTAGGTTAACGTTGCGACTCTGTGTATATA
81
310     320     330     340     350     360
301    ACCTCGACAGTCTTGGCACCTAACGTGCTGTGCGTAGCTGCTCCTTTGGTTGAATCCCCA
101
370     380     390     400     410     420
361    GGCCCTTGTGGGGCACAAAGGTGGCAGGATGTCTCAGTGGTACGAACTTCAGCAGCTTGA
121
430     440     450     460     470     480
421    CTCAAAATTCCTGGAGCAGGTTCCACCAGCTTTATGATGACAGTTTTCCCATGGAAATCAG
141    S K F L E Q V H Q L Y D D S F P M E I R
490     500     510     520     530     540
481    ACAGTACTGGCAGTGGTTAGAAAAGCAAGACTGGGAGCACGCTGCCAATGATGTTTC
161    Q Y L A Q W L E K Q D W E H A A N D V S
550     560     570     580     590     600
541    ATTTGCCACCATCCGTTTTCATGACCTCCTGTCACAGCTGGATGATCAATATAGTCGCTT
181    F A T I R F H D L L S Q L D D Q Y S R F
610     620     630     640     650     660
601    TTCTTTGGAGAATAACTTCTTGCTACAGCATAACATAAGGAAAAGCAAGCGTAATCTTCA
201    S L E N N F L L Q H N I R K S K R N L Q
670     680     690     700     710     720
661    GGATAATTTTCAGGAAGACCCCAATCCAGATGTCTATGATCATTACAGCTGTCTGAAGGA
221    D N F Q E D P I Q M S M I I Y S C L K E
730     740     750     760     770     780
721    AGAAAGAAAATTCTGGAAAACGCCCAGAGATTTAATCAGGCTCAGTCGGGGAATATTCA
241    E R K I L E N A Q R F N Q A Q S G N I Q
790     800     810     820     830     840
781    GAGCACAGTGTGTTAGACAAACAGAAAGAGCTTGACAGTAAAGTCAGAAATGTGAAGGA
261    S T V M L D K Q K E L D S K V R N V K D
850     860     870     880     890     900
841    CAAGGTTATGTGTATAGAGCATGAAATCAAGAGCCTGGAAGATTTACAAGATGAATATGA
281    K V M C I E H E I K S L E D L Q D E Y D
910     920     930     940     950     960
901    CTTCAAATGCAAAACCTTGCAGAACAGAGAACACGAGACCAATGGTGTGGCAAAGAGTGA
301    F K C K T L Q N R E H E T N G V A K S D
970     980     990     1000    1010    1020
961    TCAGAAACAAGAACAGCTGTTACTCAAGAAGATGTATTTAATGCTTGACAATAAGAGAAA
321    Q K Q E Q L L L K K M Y L M L D N K R K
1030    1040    1050    1060    1070    1080
1021   GGAAGTAGTTCACAAAATAAGAGTTGCTGAATGTCACTGAACTTACCCAGAATGCCCT
341    E V V H K I I E L L N V T E L T Q N A L
1090    1100    1110    1120    1130    1140
1081   GATTAATGATGAACTAGTGGAGTGAAGCGGAGACAGCAGAGCGCCTGTATTGGGGGGCC
361    I N D E L V E W K R R Q Q S A C I G G P
1150    1160    1170    1180    1190    1200
1141   GCCCAATGCTTGTGGATCAGCTGCAGAACTGGTTCACTATAGTTGCGGAGAGTCTGCA
381    P N A C L D Q L Q N W F T I V A E S L Q
1210    1220    1230    1240    1250    1260
1201   GCAAGTTCGGCAGCAGCTTAAAAAGTTGGAGGAATTGGAACAGAAATACACCTACGAACA

```

401 Q V R Q Q L K K L E E L E Q K Y T Y E H
 1270 1280 1290 1300 1310 1320
 1261 TGACCCTATCACAACAAAACAAGTGTATGGGACCGCACCTTCAGTCTTTCCAGCA
 421 D P I T K N K Q V L W D R T F S L F Q Q
 1330 1340 1350 1360 1370 1380
 1321 GCTCATTGAGCTCGTTTGTGGTGGAAAGACAGCCCTGCATGCCAACGCACCCTCAGAG
 441 L I Q S S F V V E R Q P C M P T H P Q R
 1390 1400 1410 1420 1430 1440
 1381 GCCGCTGGTCTGAAGACAGGGTCCAGTTCCTGTGAAGTTGAGACTGTTGGTGAATTT
 461 P L V L K T G V Q F T V K L R L L V K L
 1450 1460 1470 1480 1490 1500
 1441 GCAAGAGCTGAATTATAATTTGAAAGTCAAAGTCTTATTTGATAAAGATGTGAATGAGAG
 481 Q E L N Y N L K V K V L F D K D V N E R
 1510 1520 1530 1540 1550 1560
 1501 AAATACAGTAAAAGGATTTAGGAAGTTCAACATTTTGGGCACGCACACAAAAGTGATGAA
 501 N T V K G F R K F N I L G T H T K V M N
 1570 1580 1590 1600 1610 1620
 1561 CATGGAGGAGTCCACCAATGGCAGTCTGGCGGCTGAATTTCCGCACCTGCAATTGAAAGA
 521 M 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 ACAGAAAATGCTGGCACCAGAACGAATGAGGGTCCCTCATCGTTACTGAAGAGCTTCA
 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 CTCCCTTAGTTTTGAAACCAATTTGTCCAGCCTGGTTGGTAATTGACCTCGAGACGAC
 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 CTCTCTGCCCGTTGTGGTGTATCTCAACGTCAGCCAGCTCCCGAGCGGTTGGCCCTCCAT
 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 CCTTTGGTACAACATGCTGGTGGCGGAACCCAGGAATCTGTCTTCTTCTGACTCCACC
 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 ATGTGCACGATGGGCTCAGCTTTCAGAAGTGTGAGTTGGCAGTTTTCTTCTGTACCAA
 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 AAGAGGTCTCAATGTGGACAGCTGAACATGTTGGGAGAGAAGCTTCTGGTCTTAACGC
 641 R G L N V D Q L N M L G E K L L G P N A
 1990 2000 2010 2020 2030 2040
 1981 CAGCCCCGATGGTCTCATTCCGTGGACGAGGTTTTGTAAGGAAAATATAAATGATAAAAA
 661 S P D G L I P W T R F C K E N I N D K N
 2050 2060 2070 2080 2090 2100
 2041 TTTTCCCTTCTGGCTTTGGATTGAAAGCATCCTAGAACTCATTAAAAAACCTGCTCCC
 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 TCTCTGGAATGATGGGTGCATCATGGGCTTCATCAGCAAGGAGCGAGAGCGTCCCTGTT
 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 GAAGGACCAGCAGCCGGGACCTTCCCTGCTGCGGTTGAGTGAAGCCTCCCGGGAAGGGGC
 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 CATCACATTACATGGGTGGAGCGGTCCAGAACGGAGGCGAACCTGACTTCCATGCGGT
 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 TGAACCTACAGAAAGAAAGAACTTTCTGCTGTACTTCCCTGACATCATTCGCAATTA
 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 GGAACCTGATGGCCCTAAAGGAACTGGATATCAAGACTGAGTTGATTTCTGTGTCTGA
 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 AGTTACACCTTCTAGACTTCAGACCACAGACAACCTGCTCCCATGTCTCCTGAGGAGTT
 841 V H P S R L Q T T D N L L P M S P E E F
 2590 2600 2610 2620 2630 2640
 2581 TGACGAGGTGTCTCGGATAGTGGGCTCTGTAGAATTCGACAGTATGATGAACACAGTATA
 861 D E V S R I V G S V E F D S M M N T V *
 2650 2660 2670 2680 2690 2700
 2641 GAGCATGAATTTTTTTCATCTTCTCTGGCAGAGTTTTCTTCTCATCTGTGATCCCTC
 881

Sequence SERPINA1 mature peptide cDNA (NM_001127707.1) with chosen editing site (PiZZ E342K, yellow).

```

      10      20      30      40      50      60
1  GAGGATCCCCAGGGAGATGCTGCCAGAAAGACAGATACATCCCACCATGATCAGGATCAC
1  E D P Q G D A A Q K T D T S H H D Q D H

      70      80      90     100     110     120
61  CCAACCTTCAACAAGATCACCCCAACCTGGCTGAGTTCGCCTTCAGCCTATACCGCCAG
21  P T F N K I T P N L A E F A F S L Y R Q

      130     140     150     160     170     180
121 CTGGCACACCAGTCCAACAGCACCAATATCTTCTTCCCCAGTGAGCATCGCTACAGCC
41  L A H Q S N S T N I F F S P V S I A T A

      190     200     210     220     230     240
181 TTTGCAATGCTCTCCCTGGGGACCAAGGCTGACACTCACGATGAAATCCTGGAGGGCCTG
61  F A M L S L G T K A D T H D E I L E G L

      250     260     270     280     290     300
241 AATTTCAACCTCACGGAGATTCCGGAGGCTCAGATCCATGAAGGCTTCCAGGAACCTCTC
81  N F N L T E I P E A Q I H E G F Q E L L

      310     320     330     340     350     360
301 CGTACCCTCAACCAGCCAGACAGCCAGCTCCAGCTGACCACCGCAATGGCCTGTTCTCTC
101 R T L N Q P D S Q L Q L T T G N G L F L

      370     380     390     400     410     420
361 AGCGAGGGCCTGAAGCTAGTGGATAAGTTTTTGGAGGATGTTAAAAAGTTGTACCCTCA
121 S E G L K L V D K F L E D V K K L Y H S

      430     440     450     460     470     480
421 GAAGCCTTCACTGTCAACTTCGGGGACCCGAAGAGCCAAAGAAACAGATCAACGATTAC
141 E A F T V N F G D T E E A K K Q I N D Y

      490     500     510     520     530     540
481 GTGGAGAAGGGTACTCAAGGGAAAATGTGGATTTGGTCAAGGAGCTTGACAGAGACACA
161 V E K G T Q G K I V D L V K E L D R D T

      550     560     570     580     590     600
541 GTTTTTGCTCTGGTGAATTACATCTTCTTTAAAGGCAAATGGGAGAGACCCTTTGAAGTC
181 V F A L V N Y I F F K G K W E R P F E V

      610     620     630     640     650     660
601 AAGGACACCGAGGAAGAGGACTTCCACGTGGACCAGGTGACCACCGTGAAGGTGCCTATG
201 K D T E E E D F H V D Q V T T V K V P M

      670     680     690     700     710     720
661 ATGAAGCCTTTAGGCATGTTTAAACATCCAGCACTGTAAGAAGCTGTCCAGCTGGGTGCTG
221 M K R L G M F N I Q H C K K L S S W V L

      730     740     750     760     770     780
721 CTGATGAAATACCTGGGCAATGCCACCGCCATCTTCTTCTCCTGCCTGATGAGGGGAAACTA
241 L M K Y L G N A T A I F F L P D E G K L

      790     800     810     820     830     840
781 CAGCACCTGGAATAACTCAACCCACGATATCATACCAAGTTCCCTGGAAAATGAAGAC
261 Q H L E N E L T H D I I T K F L E N E D

      850     860     870     880     890     900
841 AGAAGGCTGCGCAGCTTACATTTACCCAAACTGTCCATTACTGGAACCTATGATCTGAAG
281 R R S A S L H L P K L S I T G T Y D L K

      910     920     930     940     950     960
901 AGCGTCTGGGTCAACTGGGCATCACTAAGGTCTTCAGCAATGGGGCTGACCTCTCCGGG
301 S V L G Q L G I T K V F S N G A D L S G
```

```
          970          980          990          1000          1010          1020
961      GTCACAGAGGAGGCACCCCTGAAGCTCTCCAAGGCCGTGCATAAGGCTGTGCTGACCATC
321      V T E E A P L K L S K A V H K A V L T I

          1030          1040          1050          1060          1070          1080
1021     GACAAAGAAAGGACTGAAGCTGCTGGGGCCATGTTTTTAGAGGCCATACCCATGCTATC
341      D K K G T E A A G A M F L E A I P M S I

          1090          1100          1110          1120          1130          1140
1081     CCCCCGAGGTCAAGTTCAACAAACCCCTTGTCTTCTTAATGATTGAACAAAATACCAAG
361      P P E V K F N K P F V F L M I E Q N T K

          1150          1160          1170          1180
1141     TCTCCCTCTTCATGGGAAAAGTGGTGAATCCACCCAAAAATAA
381      S P L F M G K V V N P T Q K *
```


New frontiers for site-directed RNA editing - Harnessing endogenous ADARs

Tobias Merkle and Thorsten Stafforst

Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany. *Correspondence should be addressed to T.S. (thorsten.stafforst@uni-tuebingen.de).

Abstract

RNA editing activity can be exploited for the restoration of disease-causing nonsense and missense mutations and as a tool to manipulate the transcriptome in a simple and programmable way. The general concept is called site-directed RNA editing and has high potential for translation into the clinics. Due to its different mode of action RNA editing may well complement gene editing and other gene therapy options. In this method paper, we particularly highlight RNA editing strategies that harness endogenous ADARs. Such strategies circumvent the delivery and expression of engineered editases and are notably precise and simple. This is particularly true if endogenous ADARs are recruited with chemically modified antisense oligonucleotides, an approach we call RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing). To foster the research and development of RESTORE we now report a detailed protocol for the procedure of editing reactions, and a protocol for the generation of partly chemically modified RESTORE ASOs with a combination of in-vitro transcription and ligation.

Keywords

RNA editing – antisense oligonucleotide – RNA ligation – in-vitro transcription – therapeutic RNA editing – site-directed RNA editing – ADAR – RESTORE

Introduction

Genome editing has advanced into an indispensable tool for the generation of transgenic cell lines and animals. Furthermore, it has potential to be developed into novel gene therapies. However, the manipulation of genetic information at the RNA level is an attractive alternative to genome editing and may overcome some of the major limitations of gene editing [1]. These include permanent off-target edits, a lack of efficiency (in particular in postmitotic tissue), and the technically demanding delivery of several components, which are of large size, bacterial origin and of different chemical entity (protein and nucleic acids). However, an RNA editing approach differs principally from genome editing in the fact that the induced change is reversible. For an enduring repair of an inherited, disease-causing loss-of-function mutation this might be a disadvantage as it may require repeated dosing. However, the drawbacks might be compensated by a better safety profile and a better control over adverse effects by the dosing scheme. With respect to tool development, the reversibility of RNA editing may offer an additional advantage. It could be beneficial for the manipulation of essential signaling cues where a permanent change would be lethal or quickly compensated and thus inaccessible at the genome level [2].

Consequently, various programmable editases have been developed [2,3]. Most of them depend on the targeting of ADAR's deaminase domain towards arbitrary RNAs and allow for the site-specific deamination of specific adenosines yielding inosine (A-to-I editing). As inosine is biochemically read as guanosine this leads to a formal A-to-G substitution at the target site. The field was pioneered in 2012

by the Rosenthal lab [4] and our group [5]. In the meanwhile numerous variations of the theme have been presented [1-3]. Such engineered editing systems achieve notable editing yields, a broad codon scope, and have been proven to change cellular phenotypes. However, they have two main drawbacks. First, they technically require the ectopic expression of an editase. Second, ectopic expression of the engineered editase typically comes along with substantial off-target editing all over the transcriptome. This is particularly severe if the expression level of the editase is high and if a hyperactive ADAR deaminase mutant with extended codon scope is applied [2].

In consequence of the latter, an exciting new frontier for site-directed RNA editing is the harnessing of endogenous ADAR [6]. ADAR editing activity seems to play an important role in the dampening of the innate immune response against the double-stranded RNA species in the human transcriptome [7]. Thus ADAR is expressed and editing-active in most human tissues. In 2016, we presented a guide RNA that enables the recruitment of full length human ADAR2 for site-directed RNA editing, and demonstrated the editing of several endogenous transcript and the restoration of a mitophagy signaling pathway by repair of a PINK1 mutation [8]. The approach was based on a modular design combining a rationally programmable specificity domain antisense to the target with an invariant hairpin structure which recruits ADAR. The ADAR recruiting domain was derived from a natural, cis-acting intronic motif which naturally recruits ADAR2 to the R/G site of the GluR2 transcript for editing. This concept was very soon confirmed by a similar design developed by the Fukuda lab [9]. Interestingly, notable editing yields always required the co-expression of natural ADARs. Even though we could detect some restoration of mitophagy in absence of overexpressed ADAR, the amount of endogenous ADAR was typically insufficient to obtain editing yields clearly detectable by Sanger sequencing [8]. This was recently confirmed by the Mali lab, who applied our guide RNA design in an AAV-driven format for the repair of missense and nonsense point mutations in vivo in murine disease models [10]. Again, co-expression of natural ADARs was required to obtain notable editing yields. However, co-expression of ADAR activity is not free of risk. It will induce off-target editing, and indeed, the Mali lab reported severe toxicity under certain circumstances [10].

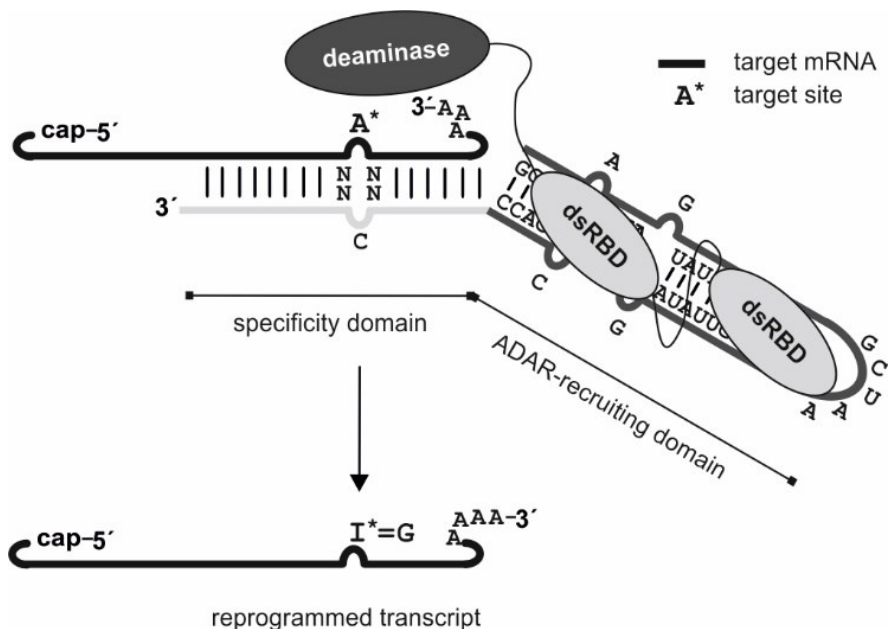


Figure 1: General principle of RESTORE: ASOs comprise an invariant ADAR-recruiting domain to attract endogenous ADAR via its double-stranded RNA-binding domains (dsRBD) and a programmable specificity domain that mediates mRNA binding and editing by the deaminase domain of ADAR. As a result, site-specific deamination of the target adenosine to inosine occurs. During translation inosine is functionally equivalent to guanosine, thus a formal adenosine-to-guanosine mutation is inserted.

A better way to recruit endogenous ADARs might be the administration of chemically stabilized antisense oligonucleotides [2,6]. The field of therapeutic oligonucleotides made significant progress during the last few years and led to the recent market approval of new oligo drugs [11,12]. We had already shown in 2014 that the human ADAR deaminase accepts highly chemically modified oligonucleotides in substrate complexes and that the chemical modification was not only accepted but even improved editing efficiency [13]. With this in mind, we developed antisense oligonucleotides (ASO) that enable the harnessing of endogenous ADARs, a strategy we called RESTORE [6]. We were able to obtain decent editing yields when targeting endogenous transcripts in various cell lines and primary human cells. Most significant, the RESTORE approach was markedly precise. On one hand, bystander editing was successfully suppressed by chemical modification of the ASO, even in challenging adenosine-rich sequence context. On the other hand, there was no sign of global off-target editing nor was there evidence that the natural editing homeostasis was perturbed. In the meanwhile, further genetically encoded guide RNAs have been presented that enable the harnessing of endogenous ADARs [14]. These largely depend on the expression of unstructured guide RNAs that basepair to the target RNAs on long stretches, e.g. 70-150 bp, and omit the usage of specific ADAR recruiting domains.

Interestingly, we found that ASOs are typically more efficacious than genetically encoded or transfected guide RNAs of the exact same sequence lacking chemical modification [6]. Thus, chemical modification can largely improve the pharmacological properties [15]. Furthermore, the ASO approach does not require any transgene expression and thus is not a gene therapy. All this makes the approach particularly promising for translation into medicine. However, the screening of large pools of long and highly chemically modified ASOs is cumbersome and expensive. In our initial RESTORE approach we have been using a ligation strategy to enable the attachment of various ADAR recruiting domains to the same specificity domain – and vice versa – the attachment of various specificity domains to the same ADAR recruiting domain [6]. We provide here a detailed protocol for this ligation strategy to foster research and development in the field.

Materials

Reagents and Buffers

Protocol I

- HeLa cells (ATCC CCL-2)
- Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher)
- fetal bovine serum (FBS, Thermo Fisher)
- 1% penicillin/streptomycin (P/S, Thermo Fisher)
- trypsin/EDTA (Sigma Aldrich)
- Phosphate-buffered saline (PBS)
- Trypan blue (Sigma Aldrich)
- RESTORE ASOs, HPLC or PAGE purified (self-made/Biospring/Eurogentec)
- Lipofectamine 2000 (Thermo Fisher)
- Lipofectamine RNAiMAX (Thermo Fisher)
- FuGENE6 (Promega)
- Opti-MEM (Thermo Fisher)
- RNeasy minElute (Qiagen)
- DNase I (NEB)
- M-MuLV reverse transcriptase (NEB)

- RNase inhibitor, murine (NEB)
- Taq DNA polymerase (NEB)
- NucleoSpin gel and PCR clean-up kit (Machery-Nagel)
- Agarose NEEO ultra-quality (Carl Roth GmbH)
- Roti-GelStain (Carl Roth GmbH)

Protocol II

- T4 Polynucleotide Kinase (T4 PNK, NEB)
- T4 RNA Ligase 1 (NEB)
- nuclease free water (Sigma Aldrich)
- 50% PEG 8000 (NEB)
- materials for 20%Urea-PAGE (all obtained from Roth, Germany)
- sodium acetate (Sigma Aldrich)
- 100% EtOH (HPLC grade, Sigma Aldrich)
- RNA loading dye (1xTBE, 7M Urea, bromophenol blue, xylene cyanol)

Protocol III

- T7 RNA Polymerase, HC (200 U/ μ L) (Thermo Fisher)
- Nuclease-free water (Sigma Aldrich)
- DMSO (Merck)
- NTP bundle, 100 mM single solutions (Jena Bioscience)
- materials for 20% Urea-PAGE (all obtained from Roth, Germany)
- sodium acetate (Sigma Aldrich)
- 100% EtOH (HPLC grade, Sigma Aldrich)
- DNA templates (Sigma Aldrich)
- RNA loading dye (1xTBE, 7M Urea, bromophenol blue, xylene cyanol)
- Magnesium chloride (Sigma Aldrich)

Equipment and Consumables

Protocol I

- T75 cell culture flasks (Sarstedt)
- 24-/96-well plates (Greiner Bio-One)
- hemocytometer (Neubauer improved, Precicolor HBG)
- incubator suitable for human cell culture (5% CO₂)
- sterile working bench
- bench-top centrifuge
- PCR Thermocycler
- Agarose gel chamber Mini-Sub[®] Cell GT Cell (Bio-Rad)
- UV transilluminator (365 nm wavelength)

Protocol II

- heating block (e.g. Eppendorf ThermoMixer C)
- sequencing gel chamber (e.g. Analytik Jena Biometra Model S2)
- TLC Plates (Merck TLC Silica Gel 60 F₂₅₄)
- UV hand lamp (e.g. UVP UVGL-58)

- cooled bench-top centrifuge (e.g. Hettich Mikro 220R)

Protocol III

- heating block (e.g. Eppendorf ThermoMixer C)
- sequencing gel chamber (e.g. Analytik Jena Biometra Model S2)
- TLC Plates (Merck TLC Silica Gel 60 F₂₅₄)
- UV hand lamp (e.g. UVP UVGL-58)
- cooled bench-top centrifuge (e.g. Hettich Mikro 220R)

Protocols (with Notes)

Protocol I. Editing procedure with RESTORE ASOs

For the transfection of the ASOs forward and reverse transfection is possible. While forward transfection is generally better tolerated by the cells and thus more suitable for sensitive cells, one can save time and reagents using reverse transfection which is especially suitable for ASO screening.

Reverse Transfection

1. Grow and subculture HeLa cells (works as well for U2OS, SH-SY5Y, SK-N-BE(2), U87MG, Huh7, HepG2, AKN-1, A549, HEK293T) in DMEM supplemented with 10% FBS and 1% P/S at 37°C and 5% CO₂ in a 75cm² cell culture flask to 70-90% confluency.
2. Remove culture medium. Wash cells with 10 mL PBS and trypsinize with 1 mL trypsin/EDTA for 3-5 min at 37°C until cells detach.
3. Stop trypsin digest by adding 9 mL DMEM+10% FBS and resuspend cells. Transfer cell suspension in a 15 mL centrifugation tube and spin for at 200xg for 5 min.
4. Remove supernatant and carefully resuspend cells in 5 mL DMEM+10% FBS. Determine the cell number of sample of the cell suspension with a hemocytometer (dilute the sample in trypan blue to exclude dead cells)
5. Dilute cell suspension with DMEM+10% FBS to 500,000 cells/mL (and add Interferon- α to 6000 units/mL if applicable)
6. Prepare a dilution of 5 pmol of each ASO with Opti-MEM to a final volume of 10 μ L/well of a 96-well plate (it is recommended to use at least 2 wells per ASO and prepare 10% excess) in a reaction tube. In a separate reaction tube, prepare a master mix of 0.5 μ L Lipofectamine 2000 in 9.5 μ L Opti-MEM per 96-well. After 5 min incubation, mix 10 μ L Lipofectamine 2000 master mix with 10 μ L of each ASO dilution and incubate the transfection mix for another 20 min. Then pipette 20 μ L transfection mix per well in a 96-well plate.

Note: ASOs should be HPLC- or PAGE purified.

7. Add 100 μ L/well of the previously diluted cell suspension.
8. After 24h, remove the medium, wash with 100 μ L PBS, trypsinize the cells (20 μ L trypsin/EDTA), resuspend them in 100 μ L DMEM+10%FBS and centrifuge at 200xg for 5 min. Remove the supernatant, wash with 200 μ L PBS, centrifuge again and add 350 μ L RLT lysis buffer (alternatively cell pellets can be frozen in liquid nitrogen and stored at -80°C)

Note: Alternatively, wash well-adherent cells with PBS in the well and add the RLT lysis buffer directly on top of the cells.

9. After lysis, RNA is purified with the RNeasy minElute kit according to the manufacturer's protocol.
10. Measure RNA concentration and use not more than 2 μ g RNA for DNase I digest
11. Dilute 2 μ g of RNA in nuclease-free water to a final volume of 26 μ L, add 3 μ L of 10x DNase I buffer and 1 μ L of DNase I (2U/ μ L). Incubate for 30min at 37°C.
12. Inactivate DNase by adding 3 μ L of a 25 mM EDTA solution and incubate at 75°C for 10 min.

Note: If you edit an endogenous mRNA and choose primers binding to the cDNA in different exons then DNase digestion is not mandatory in your PCR.

13. Add 1 μL of a 10 μM reverse primer to 15 μL of DNase digested RNA and incubate for 3 min at 70°C.

Note: For ASOs that bind with high affinity to the mRNA (e.g. locked nucleic acid-containing oligonucleotides) it is recommended to add 1 μL of a 5 μM *sense oligo* and incubate for 3 min at 95°C to capture the ASO which otherwise can block reverse transcription. *Sense oligos* are reverse complement to the ASOs and either consist of 2'-OMe RNA nucleotides or DNA. However, if it is a DNA oligo, additional 3 non-binding nucleotides must be added to the 3' end so that the DNA oligo is unable to serve as a primer in the following PCR.

14. Cool down the RNA-primer mix on ice and add 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). Incubate for 42°C for 2 h. Afterwards, inactivate M-MuLV reverse transcriptase heating to 90°C for 10 min.
15. Set up a Taq PCR for cDNA amplification as follows: Mix 5 μL of the reverse transcriptase reaction mix with 5 μL ThermoPol buffer (10 x), 2.5 μL forward primer (10 μM), 2.5 μL reverse primer (10 μM), 1.25 μL dNTPs (10 mM each) and 0.5 μL Taq DNA polymerase (5 U/ μL) and add nuclease free water to a final volume of 50 μL .

Note: Control PCRs are highly recommended to avoid misinterpretation due to DNA contamination in the reagents (use nuclease-free water instead of cDNA template) or incomplete DNase I digestion (use DNase I-digested mRNA instead of cDNA template).

Note: Alternatively, there are commercially available kits which combine reverse transcription and PCR in one reaction, e.g. the NEB OneTaq® One-Step RT-PCR Kit, which simplify the procedure.

16. Purify the PCR product with a TAE agarose gel with Roti-GelStain for visualization (1.4% is suitable for fragments between 300 and 3000 bp). Add 10 μL 6xPurple loading dye to the PCR reaction and load it on the gel. And 8 μL 2-log DNA ladder (NEB) containing loading dye on one lane.
17. Excise the bands from the Agarose gel under an UV transilluminator (365 nm wavelength) and extract the DNA from the gel pieces with the NucleoSpin gel and PCR clean-up kit according to the manufacturer's protocol.
18. Send purified PCR products for commercial Sanger sequencing. To determine editing rates measure the peak heights at the target site in the sequencing trace and calculate the ratio of the product to the substrate value.

Note: If possible, use a reverse primer for sequencing since the ratio of C to T peak heights is typically more precise than the respective G to A ratio.

Forward Transfection

Step 1-4 are identical to reverse transfection (see above)

5. Dilute cell suspension with DMEM+10% FBS to 200,000 cells/mL and seed 100,000 cells/well in a 24-well plate by distributing 0.5 mL cell suspension in each well.
6. After 24h, prepare a dilution of 25 pmol of each ASO with Opti-MEM to a final volume of 50 μL /well of a 24-well plate in a reaction tube. In a separate reaction tube, prepare a master mix of 1.5 μL Lipofectamine RNAiMAX in 48.5 μL Opti-MEM per 24-well. After 5 min incubation, mix 50 μL Lipofectamine RNAiMAX master mix with 50 μL of each ASO dilution and incubate the transfection mix for another 20 min.

7. Replace the culture medium with 500 μ L fresh medium (optional: with 6000 units/mL Interferon- α) and add 100 μ L/well of the transfection mix.
8. After 24h, remove the medium, wash with 500 μ L PBS, trypsinize the cells (100 μ L trypsin/EDTA), resuspend them in 500 μ L DMEM+10%FBS and centrifuge at 200xg for 5 min. Remove the supernatant, wash with 500 μ L PBS, centrifuge again and add 350 μ L RLT lysis buffer (alternatively cell pellets can be frozen in liquid nitrogen and stored at -80°C)

The following steps are the same as in the reverse transfection protocol.

Note: If the editing target is not an endogenous transcript but is transiently expressed from a plasmid then adjust the protocol as follows:

Step 1-4 are identical to reverse transfection (see above)

5. Dilute cell suspension with DMEM+10% FBS to 100,000 cells/mL and seed 50,000 cells/well in a 24-well plate by distributing 0.5 mL cell suspension in each well.
6. After 24h, prepare a dilution of 300 ng of each plasmid with Opti-MEM to a final volume of 50 μ L/well of a 24-well plate in a reaction tube. In a separate reaction tube, prepare a master mix of 0.9 μ L FuGENE6 in 49.1 μ L Opti-MEM per 24-well. After 5 min incubation, mix 50 μ L FuGENE6 master mix with 50 μ L of each plasmid dilution and incubate the transfection mix for another 20 min.

Note: Lipofectamine 2000 or Lipofectamine 3000 can also be used for forward transfection of plasmids. However, they show slightly higher toxicity than FuGENE6.

7. Replace the culture medium with 500 μ L fresh medium and add 100 μ L/well of the transfection mix.
8. 24h later, remove the medium, wash with 500 μ L PBS, trypsinize the cells (100 μ L trypsin/EDTA), resuspend them in 500 μ L DMEM+10%FBS and centrifuge at 200xg for 5 min. Remove supernatant and carefully resuspend cells in 300 μ L DMEM+10% FBS. Determine the cell number as described above.

Continue with the reverse transfection protocol at step 5

Note: If forward transfection is desired after plasmid transfection, seed only 50,000 cells/well, skip step 8 and continue at step 6 in the forward transfection protocol.

Note: RNA editing of endogenous transcripts is always preferred over editing of transiently overexpressed transcripts. Plasmid overexpression leads to cell toxicity, very high expression levels and uneven expression patterns, resulting in artefacts [2]. If editing of an exogenous transcript is desired, stable overexpression e.g. using the Flp-In-T-REx or piggyBac system is highly recommended [6,16].

Protocoll II. Assembly of RESTORE ASOs via Ligation

For longer RESTORE ASOs it is recommended to produce two shorter RNA pieces and to ligate them together. Especially, if one part of the ASO can be made by in-vitro-transcription (e.g. the ADAR-recruiting domain) and the other one has to be synthesized due to dense chemical modification. Generally, there are two common ways to ligate two RNA pieces enzymatically, using either T4 RNA ligase 1 or T4 DNA ligase 1. Since T4 DNA ligase 1 requires a double-stranded template, an additional DNA oligo that serves as a splint to connect both RNA strands is needed [17]. However, in our hands T4 RNA ligase was the enzyme of choice due to higher yields and easier handling when generating RESTORE ASOs. In general, both enzymes require a 5' phosphorylated *donor* RNA strand and an *acceptor* RNA strand with a free 3' hydroxyl group. 5' Phosphorylation can either be achieved during chemical synthesis or enzymatically after synthesis with T4 polynucleotide kinase.

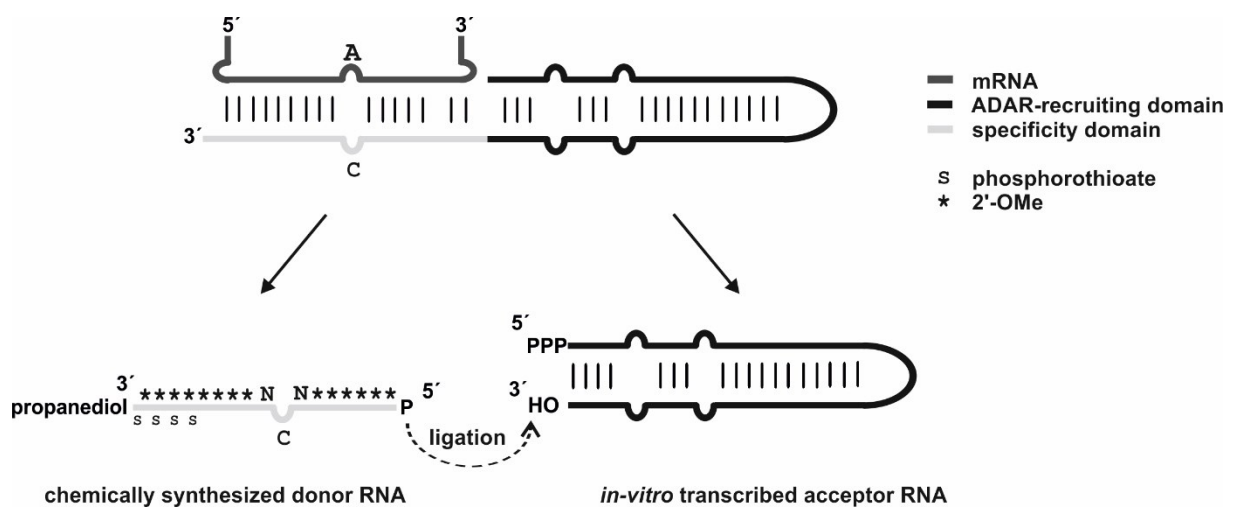


Figure 2: Ligation scheme of a RESTORE ASO. The chemically synthesized specificity domain is ligated by T4 RNA ligase 1 to an in-vitro-transcribed acceptor RNA to form a RESTORE ASO for RNA editing. For ligation the 5' end of the donor RNA must be phosphorylated and the 3' end of the acceptor RNA must have a free hydroxyl group.

Phosphorylation with T4 polynucleotide kinase

1. To 6 nmol of a donor RNA strand, add 50 μL of 10xPNK buffer, 20 μL of T4 PNK (10U/ μL) and 50 μL of 10 mM ATP and adjust the reaction volume to 500 μL with nuclease-free water.
2. Incubate at 37°C for 2 hours.
3. Optional heat inactivation: incubate for 20 min at 65°C.

Note: The 5' end of the *donor* RNA must have a free hydroxyl group. This is usually the case for chemically synthesized oligonucleotides. When using in-vitro-transcribed RNA oligonucleotides that do not utilize a 5' ribozyme to generate uniform 5' ends, then the 5' end carries a triphosphate that needs to be removed prior to the T4 PNK protocol by treatment with a phosphatase (e.g. calf intestine alkaline phosphatase or antarctic phosphatase).

Note: Heat inactivation of T4 PNK (step 3) is only recommended if the *acceptor* RNA strand has a free 5' hydroxyl group that could be phosphorylated subsequently and cause byproduct formation.

Addition of EDTA to chelate divalent cations prior to heat inactivation is strongly recommended. To remove EDTA before ligation, the phosphorylated donor RNA can be purified by ethanol precipitation.

Note: T4 PNK accepts also DNA [18] and 2' OMe RNA nucleotides as substrates.

Note: To avoid byproducts in the ligation it is recommended to block the 3' end of the *donor* RNA strand. If a chemically synthesized oligo is used, the incorporation of 3' linkers (e.g. propanediol or amino-C6) can fully prevent such byproducts.

Ligation with T4 RNA ligase 1

4. Add 6 nmol of acceptor RNA strand, 50 μ L 10x ligation buffer, 50 μ L 10 mM ATP and 20 μ L T4 RNA ligase 1 (NEB), 200 μ L 50% PEG 8000 and nuclease-free water directly to the phosphorylation mix to a final volume of 1 mL.
5. Incubate overnight at 25°C.
6. Precipitate the oligonucleotides from the ligation mix by adding 100 μ L (0.1 volumes) 3 M sodium acetate solution and 3 mL (3 volumes) 100% ethanol to the reaction mix (distribute equally on 3x1.5 mL reaction tubes)
7. Incubate at -20°C for at least 1 hour
8. Centrifuge at 12,000xg at 4°C for 1 hour
9. Remove supernatant and wash with 500 μ L ice cold 70% ethanol
10. Centrifuge again at 12,000g at 4°C for 20 min
11. Discard supernatant and dissolve the pellets in a total volume of 60 μ L of RNA loading dye (1xTBE, 7M Urea, bromophenol blue, xylene cyanol)
12. Load the sample in 3 pockets on a 15% Urea-PAGE sequencing gel (0.8mm thickness)
13. Run PAGE at 1200 V and 65 W for 4-6 h.
14. To visualize the nucleic acid bands, place the gel on a TLC plate wrapped with several layers of saran wrap and illuminate the gel with a UV hand lamp with low intensity 254 nm UV light. The ligated product migrates slowest on the gel.
15. Excise the ligated product bands with a scalpel and transfer the gel slices in a 2.0 mL reaction tube.
16. Add 700 μ L of nuclease-free water to the tube and shake at 1100 rpm overnight at 4°C.
17. Transfer the solution to a new reaction tube and remove remaining gel pieces by short centrifugation (12,000xg, 2 min).
18. Precipitate the ligated oligo by adding 3 volumes 100% ethanol and 0.1 volume 3 M sodium acetate.
19. Incubate at -20°C for at least 1 hour
20. Centrifuge at 12,000xg at 4°C for 1 hour
21. Remove supernatant and wash with 500 μ L ice cold 70% ethanol
22. Centrifuge again at 12,000g at 4°C for 20 min
23. Discard supernatant and dissolve the pellet in a total volume of 20-50 μ L in nuclease-free water.
24. Measure the absorbance at 260 nm in a NanoDrop spectrophotometer and determine concentration with the predicted molar extinction coefficient of the ligated ASO.
25. Store the ASO until further use at -20°C.

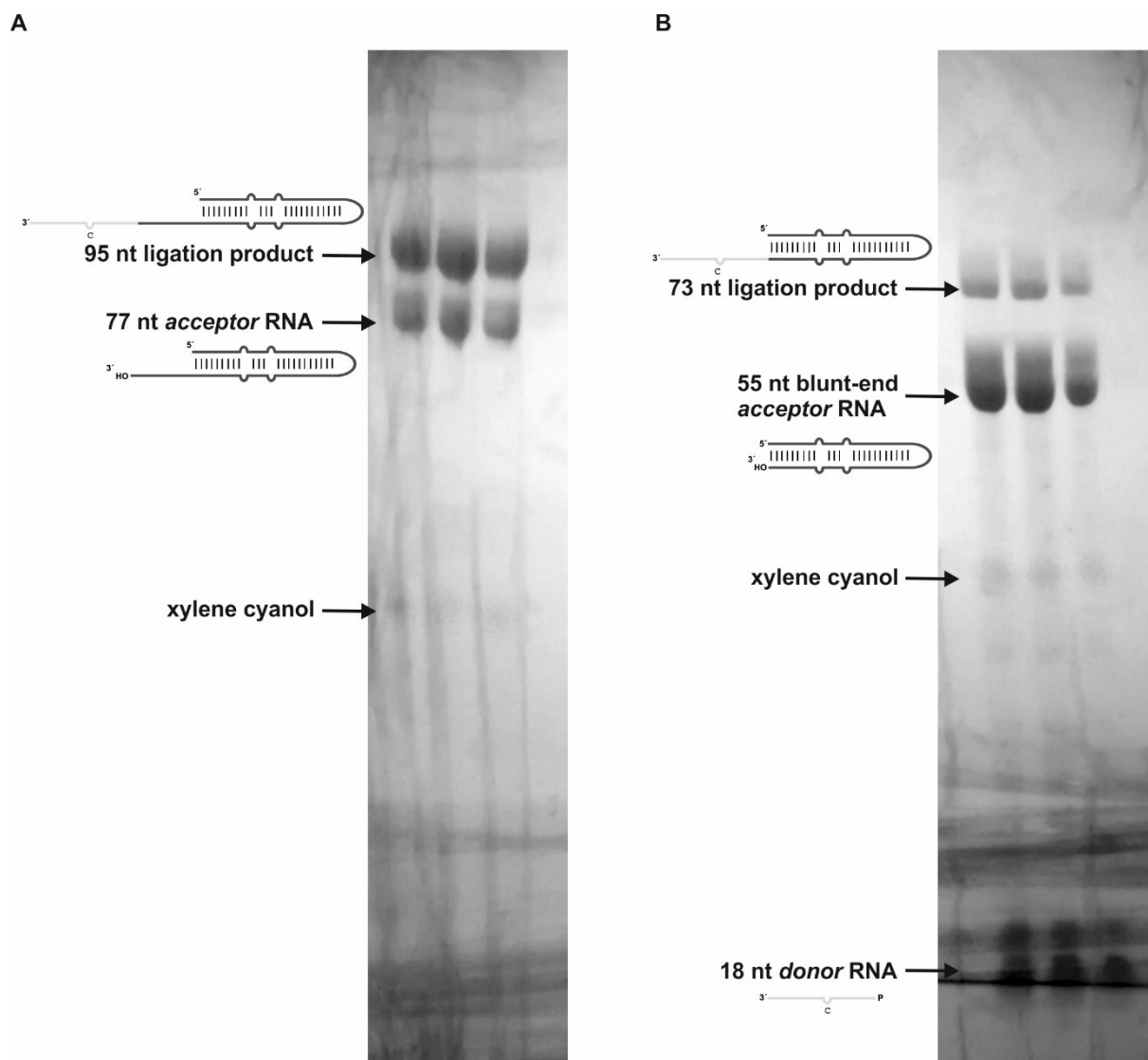


Figure 3: Exemplary PAGEs. In **A**, a 18 nt chemically modified donor RNA was ligated to a 77 nt acceptor RNA. While the 18 nt donor RNA is fully converted (thus not visible on the gel), an excess of the 77 nt acceptor RNA is still visible under the more intense 95 nt band of the ligation product. In **B**, however, the 18 nt donor RNA and the 55 nt blunt-end acceptor RNA are still clearly visible, and the band of the 73 nt ligation product appears more faint. This is likely due to the less efficient blunt-end ligation.

Note: The *acceptor* RNA strand must have a free 3' hydroxyl group. However, a 5' phosphate can produce byproducts by unwanted ligation of two acceptor strands. If required a 5'-phosphorylation can be removed by a prior phosphatase treatment.

Note: If the *acceptor* RNA strand is not limiting a 1.5-fold excess of the acceptor over the donor strand can be used to improve yields.

Note: To obtain optimal yields it is recommended that the ligation site is not at the blunt end of a double-stranded RNA but either some nucleotides before or after that (see Figure 3).

Note: Ligation yields depend largely on the sequences used and vary between 3% and 45%.

Table 1: Representative examples for RESTORE ligations from our lab. Yields are final yields after PAGE purification and recovery. All the non-modified RNA acceptor oligos were in-vitro-transcribed. All donor oligos were phosphorylated using the T4 PNK protocol above. (N) = RNA base, [N] = 2'-OMe RNA base, * = phosphorothioate linkage, {N} = LNA base.

| Acceptor oligo | Donor oligo | Blunt end | Yield [%] |
|--|--|-----------|-----------|
| (GGUGAAGAGGAGAACAAUAUGC UAAAUGUU GUUCUCGUCUC) | (CACC)[CACUGC](CCA)[GGCA U*C*G*C] | no | 24.3 |
| (GGUGAAGAGGAGAACAAUAUGC UAAAUGUU GUUCUCGUCUC) | (CACC)[CACUGC](CCA)[GGCA UCAGCCU*G*C*U*G] | no | 45.3 |
| (GGUGAAGAGGAGAACAAUAUGC UAAAUGUU GUUCUCGUCUC) | (CACC)[CCACUG](CCG)[AGGC A*U*C*A*G] | no | 35.5 |
| (GGUGAAGAGGAGAACAAUAUGC UAAAUGUU GUUCUCGUCUC) | (CACC)[CCACUG](CCG)[AGGC AUCAGCCU*U*G*C*U] | no | 38.7 |
| (GGUGAAUAGUAUAACAAUAUGC UAAAUGUU GUUAUAGUAUCCACC) | [AGGGGU](CCA)[CAUGG*C* A*A*C] propanediol | yes | 18.3 |
| (GGUGAAGAGGAGAACAAUAUGC UAAAUGUU GUUCUCGUCUCCACC) | [AGGGGU](CCA)[CAUGG*C* A*A*C] propanediol | yes | 25.5 |
| (GGUGUCGAGAAGAGGAGAACAUAUGC UAAAUGUU GUUUGUUCUCGUCUCCUCGACACC) | [AGGGGU](CCA)[CAUGG*C* A*A*C] propanediol | yes | 16.8 |
| (GGUGAAUAGUAUAACAAUAUGC UAAAUGUU GUUAUAGUAUCCACC) | [GCAAUG](CCA)[UCACC*U*C *C*C] propanediol | yes | 22.6 |
| (GGUGAAGAGGAGAACAAUAUGC UAAAUGUU GUUCUCGUCUCCACC) | [GCAAUG](CCA)[UCACC*U*C *C*C] propanediol | yes | 42.5 |
| (GGUGUCGAGAAGAGGAGAACAUAUGC UAAAUGUU GUUUGUUCUCGUCUCCUCGACACC) | [GCAAUG](CCA)[UCACC*U*C *C*C] propanediol | yes | 16.1 |
| (GGUGUCGAGAAGAGGAGAACAUAUGC UAAAUGUU GUUUGUUCUCGUCUCCUCGACACC) | [CCUUUC](UCG)[UCGAU*G* G*U*C] propanediol | yes | 7.7 |
| [G*G*U](G)[UC](GAGAAGAGGAGAA)[C](AA)[U](A)[U](G)[CU](AAA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GACACC) | [CCUUUC](UCG)[UCGAU*G* G*U*C] propanediol | yes | 3.3 |
| [G*G*U](G)[UC](GAG AAG AGG AGA A)[C](AA)[U](A)[U](G)[CU](AAA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GACACC) | (CAUGGCCCCAGCAGCUUCA GUC)[C]{C}[UUUC](UCG)[UC GA]{T*}[G*][G*]{T*}[C]amino C6 | yes | 3.1 |
| [G*G*U](GUCGAGAAGAGGAGAACAUAUGC UAAAUGUU GUUUGUUCUCGUCUCCUCGACACCUUGU CAUGGAUGACCUUGGCCA) | [G]{G}[GGUG](CCA)[AGCA]{G*}[U*U*]{G*}[G] amino C6 | no | 8.6 |
| [G*G*U](G)[UC](GAGAAGAGGAGAA)[C](AA)[U](A)[U](G)[CU](AAA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GA)[C](A)[CC](UUGUCAUGGAUGA CCUUGGCCA) | [G]{G}[GGUG](CCA)[AGCA]{G*}[U*U*]{G*}[G] amino C6 | no | 6.3 |

Protocol III. In-vitro-transcription of RESTORE ASOs and PAGE purification

RESTORE ASOs or parts of it, e.g. the ADAR recruitment domain and the specificity domain, can be obtained without chemical modification by in vitro transcription using T7 RNA polymerase. Especially for screening length, sequence, and placement of a RESTORE ASOs, in-vitro-transcription represents a fast and cheap alternative to chemically synthesized ASOs. Similarly, this also applies to the ADAR-recruiting moiety that can be in-vitro-transcribed and later be ligated to chemically modified specificity domains. In vitro transcription is DNA template-dependent. The template can be linearized plasmids, PCR products or synthetic DNA oligonucleotides. The minimal requirement for a DNA template is a double-stranded T7 promoter on the 5' end (5'-dTAAATACGACTCACTATAGGGAGA-3'), where the transcription starts with 5'-rGGGAGA-3'. However, the rest of the transcribed DNA template can be single-stranded. For the synthesis of short RNAs the transcription initiation is limiting [19]. Thus, we prefer to anneal two synthetic DNA strands as templates to generate short RNAs in high yields, e.g. to screen ASOs. These comprise of a constant T7 promoter DNA-oligo and a DNA-oligo with a reverse complement T7 promoter and the desired RNA sequence.

Note: Only the first guanosine of the transcription start site is mandatory. The second base can be chosen freely, however resulting in lower yields. For high yields a guanosine in position 2 is recommended.

1. Mix 3 μL of 100 μM T7 promoter DNA strand with 3 μL of 100 μM DNA template strand, 4.8 μL 1 M magnesium chloride, 94.2 μL nuclease-free water and 30 μL DMSO.

Note: Magnesium chloride and DMSO increase the RNA yield for short strands (<100bp) dramatically [20].

2. Incubate at 70°C for 5 min and let cool down slowly.
3. Add 50 μL of 5x transcription buffer (Thermo Fisher), a total of 20 μL NTPs (ATP, CTP, GTP, UTP, each 100 mM) and 5 μL T7 RNA polymerase (200 U/ μL).

Note: For optimal yields, it is recommended to add single nucleotides in the ratio they occur in the desired RNA product rather than in an equimolar mix.

4. Incubate overnight at 37°C.
5. Remove white pyrophosphate precipitate by short centrifugation (12,000xg, 2min)

Note: To avoid pyrophosphate precipitation inorganic pyrophosphatase (e.g. NEB) can be added.

6. Transfer supernatant in a new 1.5 mL reaction tube and add 3 volumes 100% ethanol

Note: No additional sodium acetate is needed for precipitation due to the high salt concentration in the reaction mix.

7. Incubate at -20°C for at least 1 hour.
8. Centrifuge at 12,000xg at 4°C for 1 hour
9. Remove supernatant and wash with 500 μL ice cold 70% ethanol
10. Centrifuge again at 12,000g at 4°C for 20 min
11. Discard supernatant and dissolve the pellet in a total volume of 20 μL of RNA loading dye (1xTBE, 7 M Urea, bromophenol blue, xylene cyanol)

12. Load the sample in a pocket on a 15% Urea-PAGE sequencing gel (0.8 mm thickness)
13. Run PAGE at 1200 V and 65 W for 4-6 h.
14. To visualize the nucleic acid band, place the gel on a TLC plate wrapped with several layers of saran wrap and illuminate the gel with a UV hand lamp with low intensity 254 nm UV light. One major band should be visible at

Note: In a PAGE sequencing gel the resolution is usually high enough to see the minor bands of typical byproducts, e.g. with 1-2 additional terminal nucleotides. The formation of those byproducts should already be strongly reduced by addition of DMSO (in step 1) [20]. It can be further reduced, if necessary, by using a DNA template strand with two consecutive 2'-OMe nucleotides at its 5' end [21].

Note: At the running front of the gel, short break-off products from failed transcription initiation and free nucleotides can usually be seen.

Note: The DNA template is separated from the RNA strand during PAGE purification and is sometimes visible as a faint band.

15. Excise the RNA band with a scalpel and transfer the gel slice in a 2.0 mL reaction tube.
16. Add 350 μ L of nuclease-free water to the tube and shake at 1100 rpm overnight at 4°C.
17. Transfer the solution to a new reaction tube and remove remaining gel pieces by short centrifugation (12,000xg, 2min).
18. Precipitate the RNA by adding 3 volumes 100% ethanol and 0.1 volume 3 M sodium acetate.
19. Incubate at -20°C for at least 1 hour
20. Centrifuge at 12,000xg at 4°C for 1 hour
21. Remove supernatant and wash with 500 μ L ice cold 70% ethanol
22. Centrifuge again at 12,000g at 4°C for 20min
23. Discard supernatant and dissolve the pellet in a total volume of 20-50 μ L in nuclease-free water.
24. Measure the absorbance at 260 nm in a NanoDrop spectrophotometer and determine concentration with the predicted molar extinction coefficient of the RNA.
25. Store the ASO until further use at -20°C.

Note: Yields are very sequence specific, but typical yields are between 0.6 and 1.2 nmol per 200 μ L reaction mix.

Note: It may also be possible to statistically incorporate nucleotides with modifications during in-vitro-transcription, e.g. 2'-Fluoro or 2'-OMe. However, except for phosphorothioates we were unable to generate chemically modified ASO parts in yields sufficient for our ASO screens, even when using described T7 RNA polymerase mutants (e.g. Y639F [22], RGVG-M5 [23]). Furthermore, the modified nucleotide triphosphates are significantly more expensive.

Note: To obtain uniform 5' and 3' ends, a hammerhead ribozyme can be added 5' to the RNA sequence and a Hepatitis delta virus ribozyme 3' to the RNA sequence [24]. However, quantitative cleavage of the ribozymes is not always given, and the yields after cleavage and PAGE purification drop drastically. Furthermore, the nature of the resulting 5' and 3' termini must be considered for the subsequent ligation. The 5' end is a free hydroxyl (instead of a triphosphate) and the 3' end consist of a 2'-3'-cyclic phosphate (instead of a hydroxyl group).

Acknowledgement

We gratefully acknowledge support from the University of Tübingen. This work was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) – STA 1053/3-2, STA 1053/7-1, STA 1053/11-1 (430214260).

References

1. H.A. Rees, D.R. Liu. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Gen.* 19, 770-788 (2018).
2. P. Vogel, T. Stafforst. Critical review on engineering deaminases for site-directed RNA editing. *Curr. Opin. Biotech.* 55, 74-80 (2019).
3. M.F. Montiel-Gonzalez, J.F. Diaz Quiroz, J.J.C. Rosenthal. Current strategies for Site-Directed RNA Editing using ADARs. *Methods* (2018); doi: 10.1016/j.ymeth.2018.11.016.
4. Montiel-Gonzalez, M. F., Guillermo, I., Yudowski, A., Rosenthal, J. J. C. *Proc. Natl. Acad. Sci. USA* 110, 18285-290 (2013).
5. Stafforst, T., Schneider, M.F. *Angew. Chem. Int. Ed.* 51, 11166-9 (2012).
6. T. Merkle, S. Merz, P. Reautschnig, A. Blaha, Q. Li, P. Vogel, J. Wettengel, J. B. Li, T. Stafforst. Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nature Biotech.* 37, 133-138 (2019).
7. Meng, H.T., et al. Dynamic landscape and regulation of RNA editing in mammals. *Nature* 550, 249–254 (2017).
8. Wettengel, J., Reautschnig, J., Geisler, S., Kahle, P. J., Stafforst, T. Harnessing human ADAR2 for RNA repair – Recoding a PINK1 mutation rescues mitophagy. *Nucl. Acids Res.* 45, 2797-2808 (2017).
9. Fukuda M. et al., *Sci. Rep.* 7, 41478 (2017).
10. D. Katrekar et al. In vivo RNA targeting of point mutations via RNA-guided adenosine deaminases. *Nature Methods*, in press (2019).
11. Adams, D., et al. Patisiran, an RNAi Therapeutic, for Hereditary Transthyretin Amyloidosis. *New Engl. J. Med.* 379, 11-21 (2018).
12. Benson, M.D., et al. Inotersen Treatment for Patients with Hereditary Transthyretin Amyloidosis. *New Engl. J. Med.* 379, 22-31 (2018).
13. Vogel, P., Schneider, M.F., Wettengel, J., Stafforst, T. *Angew. Chem. Int. Ed.* 53, 6267-6271 (2014).
14. Qu, L. et al. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nat. Biotech.* (2019). doi: 10.1038/s41587-019-0178-z
15. 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).
16. P. Vogel, M. Moschref, Q. Li, T. Merkle, K. D. Selvasarayanan, J. B. Li, T. Stafforst. Efficient and Precise Editing of Endogenous Transcripts with SNAP-tagged ADARs. *Nature Methods* 15, 535-8 (2018).
17. Paredes, E., Evans, M. & Das, S. R. RNA labeling, conjugation and ligation. *Methods* 54, 251-259 (2011).
18. Richardson, C. C. Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage-infected *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 54, 158-165 (1965)
19. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* 15, 8783-8798, (1987)
20. Helmling, C. et al. Rapid NMR screening of RNA secondary structure and binding. *Journal of Biomolecular NMR* 63, 67-76 (2015).

-
21. Kao, C., Zheng, M. & Rüdiger, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *Rna* 5, 1268-1272 (1999)
 22. Padilla, R. & Sousa, R. Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2'-groups using a mutant T7 RNA polymerase (RNAP). *Nucl. Acids Res.* 27, 1561-1563 (1999)
 23. Meyer, A. J. et al. Transcription yield of fully 2'-modified RNA can be increased by the addition of thermostabilizing mutations to T7 RNA polymerase mutants. *Nucl. Acids Res.* (2015)
 24. Walker, S. C., Avis, J. M. & Conn, G. L. General plasmids for producing RNA in vitro transcripts with homogeneous ends. *Nucl. Acids Res.* 31, e82 (2003)

Improved antisense oligonucleotides for efficient and precise RNA editing with endogenous ADARs

Tobias Merkle, Carolin Schlitz, Laura Pfeiffer, Clemens Lochmann and Thorsten Stafforst

Affiliations

Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany

Abstract

Recruiting endogenous ADARs for site-directed RNA editing opens an innovative and potentially safer option for the treatment of some genetic disorders beyond CRISPR. With our previously published RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing) approach we could demonstrate the feasibility of precise RNA editing with endogenous ADAR using special antisense oligonucleotides (ASOs). Here we report a novel improved design of our RESTORE ASOs with significantly reduced size and optimized chemical modification pattern with enhanced editing efficacy up to 80% without interferon treatment and improved stability in plasma and cerebrospinal fluid of several days. Not only could we observe high editing yields in a panel of human cell lines and primary cells, but the improved design allowed gymnotic uptake. We could also demonstrate the application of these new ASOs for disease relevant targets like MECP2, SERPINA1 and IDUA. In fibroblasts from patients with Hurler syndrome a restoration of more than 6-fold of enzyme activity observed in the much milder Scheie syndrome was possible, emphasizing the therapeutic potential of these improved RESTORE ASOs.

Introduction

Site-directed RNA editing has been proposed as an alternative to CRISPR-mediated DNA editing^{1, 2}. A major advantage of RNA over DNA editing is the dose-dependency and reversibility of the treatment, which may allow for fine-tuning the therapeutic outcome. As off-target edits are reversible, the danger of devastating side-effects may be less likely, and a therapy could be stopped and reverted if necessary. Due to the potentially better safety profile, the temporary and limited manipulation of human genetic information at the RNA might become more broadly applicable and expanded to less severe medical indications as compared to genome editing. Various RNA editing approaches have been reported by our lab and others, recently³⁻⁷. Most of these approaches rely on the overexpression of artificial, engineered editing enzymes, e.g. SNAP-ADAR, λ N-ADAR, and Cas-ADAR, which makes the approaches technically difficult with respect to multi-component delivery and tight control of global off-target editing. With regard to therapy, the harnessing of the endogenous human protein ADAR (adenosine deaminase acting on RNA) is most promising. ADAR enzymes are widely expressed across human tissues and enable the conversion of adenosine to inosine (A-to-I RNA editing). As inosine is biochemically read as guanosine, A-to-G substitutions are formally introduced when redirecting ADAR activity. Recently, we discovered the recruitment of human endogenous ADAR by simple administration of moderately chemically modified antisense oligonucleotides into various human cells, including primary cells from several tissues. We called the strategy RESTORE (recruiting endogenous ADAR to specific transcripts for oligonucleotide-mediated RNA editing).⁸ Qu et al. could independently demonstrate the recruitment of endogenous ADARs with plasmid- or virus-borne, genetically encoded guideRNAs, a closely related approach called LEAPER (leveraging endogenous ADAR for programmable editing of RNA)⁹. However, both strategies offer plenty of space for improvement, as the applied ASOs and guideRNA designs are still very long (95 nt for RESTORE and 111 nt-151 nt for LEAPER), not fully stabilized, and give comparably low editing yields, in particular in absence of ADAR1 p150 induction

with interferon. Here, we report on a new design of RESTORE ASOs, which are a) significantly shorter, b) have improved efficacy, c) are independent of IFN- α treatment, and d) are stabilized against nuclease digestion. We demonstrate RESTORE v2 ASOs to induce correction of pathogenic point mutations and efficacy after gymnotic uptake into human primary cells.

Results and Discussion

Identification of two new lead designs

For practical application it is desirable to shorten RESTORE ASOs. The old RESTORE ASOs comprised of a 40 nt long specificity domain (mediating programmable binding to the target mRNA) plus a 55 nt highly structured ADAR-recruiting domain (Fig. 1A). Both parts could be optimized by means of sequence and chemical modification. The ADAR recruitment domain, which consists of a 25 bp RNA helix with bulges and wobble base pairs, apparently offered room for shortening. The design idea was simply to form the structured ADAR-recruiting domain not inside the ASO itself, but rather upon binding of the ASO with the target RNA (Fig. 1B). By this it should be possible to remove 30 nt of the ASO. We initially started with a 59 nt unstructured ASO which was end-blocked (3x 2'-OMe at each terminus) and was strongly modified with phosphorothioates, based on a sequence targeting a 5'-UAG codon in the ORF of human GAPDH. We optimized the phosphorothioate content and pattern and kept a symmetric gap of eight unmodified nucleotide linkages around the cytosine that mismatches with the target adenosine (SI Fig. 1). A symmetric 59 nt ASO (v117.19) gave already remarkably good editing yields on the endogenous GAPDH transcript in HEK-293 cells stably overexpressing ADAR isoforms when compared to our original RESTORE v1 ASO (v25), Fig. 1C. However, shortening of the symmetric ASO from 59 nt down to 47 nt resulted in a strong decline of editing. Alternatively, we started from an asymmetric 59 nt ASO (5'-47-3-9, v119.4), putting the edited adenosine more towards the 5'-end of the ASO/target hybrid. Surprisingly, this design could be shortened down to 40 nt (5'-28-3-9, v121.1) without major loss in efficacy, but further shortening (down to 35 nt, 5'-23-3-9, v122.1) reduced the editing yield notably.

Next, we challenged the new designs by testing them for the recruitment of endogenous ADAR in HeLa cells. Under these conditions three out of eight designs performed clearly better than the original RESTORE v1 ASO (v25), Fig. 1D. Importantly, editing yields up to 70% have been achieved in the absence of IFN- α . Before, editing yields above 25% were unfeasible. Notably, the presence of dense phosphorothioate modification was essential. The symmetric 59 nt design v117.19 and the asymmetric 45 nt design v120.2 raised our interest in particular. Both were notably shorter (36 and 50 nt) than the old RESTORE v1 ASO (v25). Both ASOs were independent of IFN- α treatment. Both ASOs did not contain LNA (locked nucleic acid) building blocks, which have been indispensable in the old ASO design (v25). Based on this, we defined v117.19 and v120.2 as our new lead designs and further characterized their properties. We tested them side-by-side in a panel of eight immortalized cell lines (HeLa, U2OS, SH-SY5Y, SK-N-BE, Huh7, HepG2, A549, THP-1, Fig. 1E) in presence and absence of IFN- α . We obtained editing yields ranging between 74% and 15%. A dependency on IFN- α was virtually absent. In some cell lines, the short design (v120.2) matched the editing levels of the long design (v117.19). Sometimes, when the editing yields were low, the long design was superior to the short one. We further characterized ASO activity in three different primary human cells (NHA, NHBE, RPE) in the absence of IFN- α (Fig. 1F). Editing levels between 61% and 88% were obtained. The longer design was typically superior to the shorter. Compared to the old RESTORE v1 ASO, the editing yields were substantially better. In RPE cells, the old v25 ASO gave editing yields below 10% in absence of IFN- α . The new designs gave yields >70%, a seminal improvement.

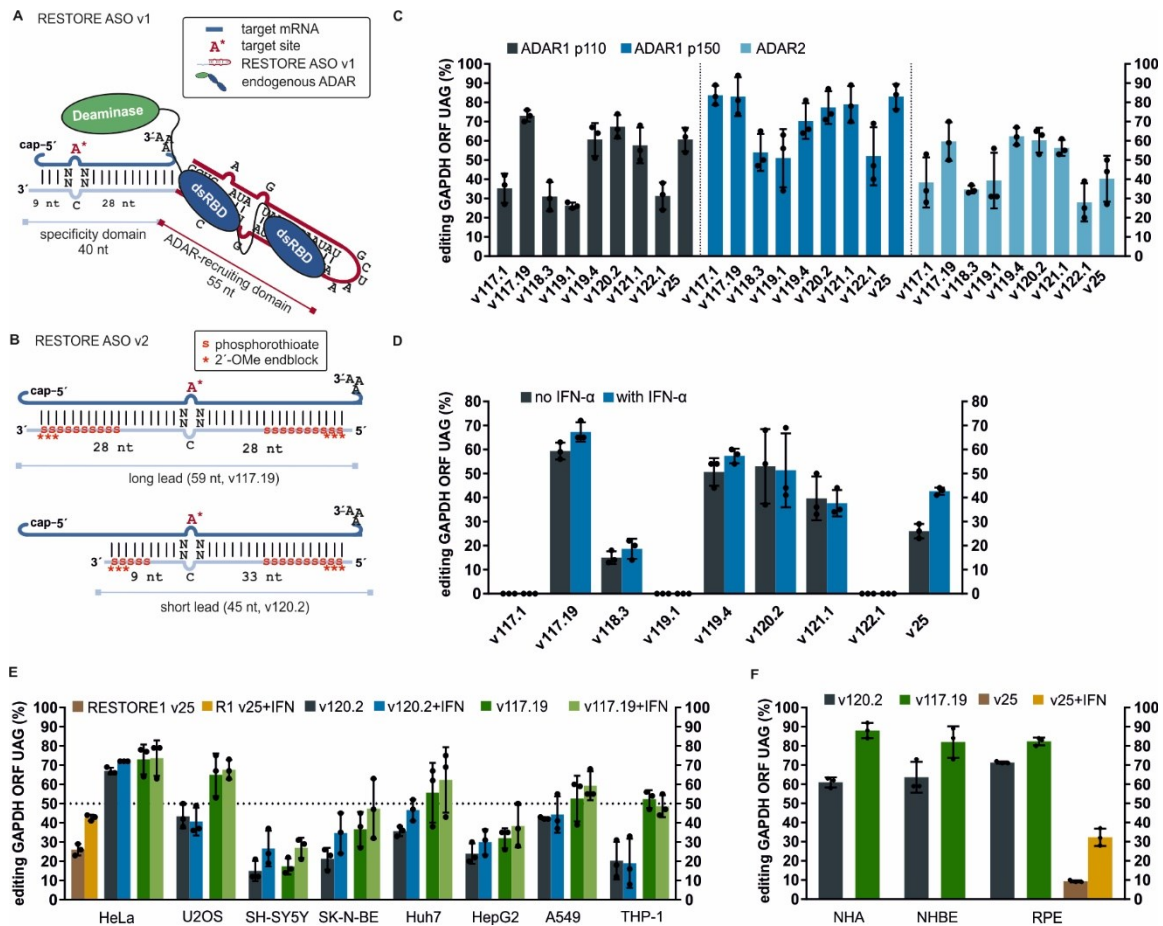


Figure 1. ASO screening. **A)** Scheme of the old RESTORE ASO v1, comprising of specificity and ADAR-recruiting domain. **B)** RESTORE v2: Schematic view of the two new lead designs, symmetric and long (v117), and asymmetric and short (v120). **C)** Effect of shorting symmetric and asymmetric ASOs for the recruitment of stably overexpressed ADARs (ADAR1p110, ADAR1p150 or ADAR2 have been overexpressed from transgenic 293 Flp-IN-T-Rex cell lines, as described before). **D)** Activity of ASOs to recruit endogenous ADAR in HeLa cells, with vs. without IFN- α treatment. **E)** Cell line screen of the two lead designs (long and short). **F)** Activity of ASOs in primary human cells. The complete sequence and modification pattern can be found in SI Table 2. NHA = normal human astrocytes, NHBE = normal human bronchial epithelium, RPE = retinal pigment epithelium; Data are shown as the mean \pm s.d., N = 3 independent experiments.

Optimization of both lead designs by further chemical modification

The sparse content of chemical modifications in lead designs v117.19 and v120.2 was insufficient to prevent their fast degradation in fetal bovine serum (FBS) and human cerebrospinal fluid (CSF). Antisense oligonucleotide drugs that are currently used in the clinics, e.g. splice switching ASOs, RNaseH-recruiting ASOs and therapeutic siRNAs, are densely chemically modified and achieve stability in vivo for weeks to months¹⁰. Thus, we included additional chemical modifications in our new RESTORE ASO designs, all based on a sequence targeting a 5'-UAG codon in the ORF of human GAPDH. We initially focused on the modification pattern at the 5'-half of the ASO, which is distal to the editing site and which was newly introduced into the ASO design to replace the former ADAR recruiting domain. We tested various patterns, including mixtures of 2'-Fluoro (2'-F), 2'-OMe, 2'-deoxy, and unmodified 2'-ribo in the context of the short lead design. We found strong interference with editing activity for any chemical moiety when it was introduced at all nucleotides in the 5'-half. However, putting 2'-Fluoro only at the pyrimidine nucleotides was comparably well accepted. We extended this concept over the entire ASO sequence with exception of the pyrimidine nucleotides around the editing site.

Here we chose 2'-desoxy, in accordance with data from our SNAP-ADAR system and recent observations by others¹¹. This modification pattern was similarly well working for the long lead design. The 2'-Fluoro modification at the pyrimidines could also be substituted by 2'-OMe with only slight loss in editing yield. Notably, the analog 2'-MOE (methoxyethyl) modification at pyrimidine bases fully blocked editing (SI Fig. 2).

While the weakly modified lead designs (end-blocked, phosphorothioate) were degraded in 10% FBS or 100% human CSF within seconds, the 2'-F/DNA pyrimidine-modified ASOs were stable for up to several days (Fig. 2A). We characterized the new modification pattern on our leads targeting endogenous GAPDH in HeLa cells (Fig. 2B) and in three human primary cells (RPE, NHA, NHBE, Fig. 2C). While there was only a minor effect in HeLa cells, the editing yields in the primary cells suffered more from the additional modifications, however, still obtaining editing levels in the range of 30% to 68% (without IFN- α). The chemically stabilized long ASO design (v117.28) was capable of inducing editing in human primary cells upon gymnotic uptake (Fig. 2D).

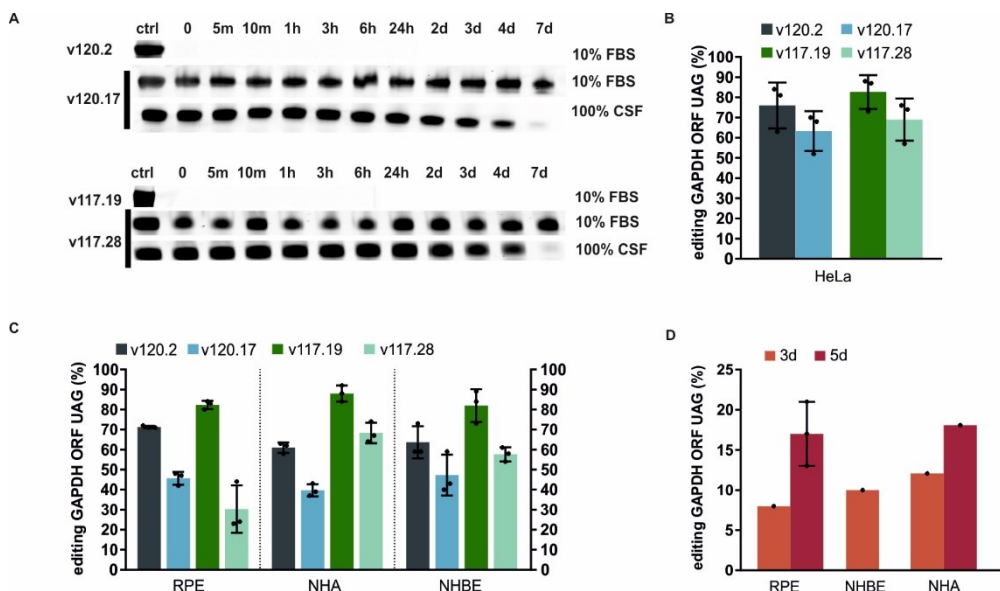


Figure 2. Further optimization of the lead ASOs. **A)** The inclusion of additional backbone modifications at all pyrimidine bases (2'-F/DNA) achieved effective stabilization of both lead ASOs in FBS and CSF. ASOs targeting the ORF of GAPDH. **B)** The stabilized ASOs are highly active in HeLa cells, and **C)** in primary cells. **D)** Chemical stabilization further enabled gymnotic uptake in primary cells. The sequences and modification patterns of all ASO are given in SI Table 3. Data are shown as the mean \pm s.d, where applicable. N=1-3 independent experiments as indicated by data points.

Editing of disease-relevant targets

We tested both our leads on several disease-relevant targets. First we tested the editing of the essential phosphorylation site tyrosine 701 in endogenous STAT1. In immortalized cell lines (HeLa, Huh7), we obtained editing levels ranging from 19% to 66% (Fig. 3A). The long design was superior to the short one. In context of the long design, stabilization of the ASO with 2'-F/DNA was possible with only minor loss in editing activity. Beside nuclease stabilization (SI Fig.3), the additional chemical modifications suppressed bystander editing at five sites (SI Fig.4). We did not find dependency on IFN- α treatment at all. We also tested the long, weakly modified design in two primary human cells (NHA, RPE, Fig. 3A) and obtained editing levels between 65% and 78%. This was remarkable, as we did not

obtain editing yields above 2% in RPE before (in absence of IFN- α) highlighting the power of the new lead designs.

We then tested the editing of several known, disease-relevant inherited mutations, e.g. IDUA W392X (murine context), MeCP2 W104X (murine context), and Serpina1 E342K (human context, PiZ allele). To gain cell models, the respective cDNAs were either overexpressed from plasmids or were inserted into the genome of the HeLa cell by the PiggyBac transposon system¹², as indicated. For the MeCP2 mutation, we tested the short lead with and without further chemical modification. Editing yields ranged between 25% and 59 \pm % (Fig. 3B). No IFN- α dependency was detected. Additional chemical modifications (2'-F/DNA) hardly affected editing levels. Editing could be run free from bystander editing (SI Fig. 5), even though this reported to be problematic before¹³. We confirmed the repair and nuclear localization of MeCP2-GFP fusion protein by fluorescence microscopy in the respective transgenic HeLa cell line (SI Fig.6 and SI Fig.7).

We tested the repair of murine IDUA W392X on basis of the short lead (v120). We found editing levels between 55% (weakly modified) and 32% (strongly modified, e.g. 2'-F/DNA), Fig. 3C. In this case, additional chemical modifications (2'-F/DNA) reduced the editing yield and affected the enzyme activity negatively as measured by a fluorogenic cleavage assay (Fig. 3D). We found that the modified (2'-F/DNA) but not the lead itself (v120.2) did negatively affect the wildtype control in the enzyme assay, which suggests that the ASO might interfere with translation (SI Fig.8). However, we could solve the problem by alternating the 2'-modification on the pyrimidine bases between 2'-Fluoro and 2'-OMe, which had little effect on the editing yield, but significantly improved restoration of enzyme activity (Fig. 3C,D).

We tested the repair of the PiZ allele in Serpina1 with the long lead design. The weakly modified ASO (v117.19) gave a comparably low editing yield (8-16%) and notable bystander editing at several sites including the direct 5'-neighboring adenosine base in this adenosine-rich codon (SI Fig. 9), a problem that was repeatedly seen before. By inclusion of five 2'-OMe groups, this bystander editing could be reduced but not fully abrogated. However, the further modification of the pyrimidines (2'-F/DNA) abolished bystander editing. Importantly, on-target editing could be enhanced to 47% by incorporation of an desoxyinosine opposite the cytosine of the 5'-CAA target codon (Fig. 3E). The rationale behind this base modification was to reduce steric hindrance in the active site of the enzyme. It was shown before that there is a clash between the minor groove face of a G:C base pair 5'-proximal to the edited adenosine and the backbone of ADAR (Gly489 in ADAR2)¹⁴. We speculate that removal of the amino group at C-2 of the purine (inosine) might help to relax that clash. An α 1-antritysin (AAT) ELISA revealed more AAT secretion of HeLa cells overexpressing the SERPINA PiZ mutation upon ASO v117.25 treatment (Fig. 3F).

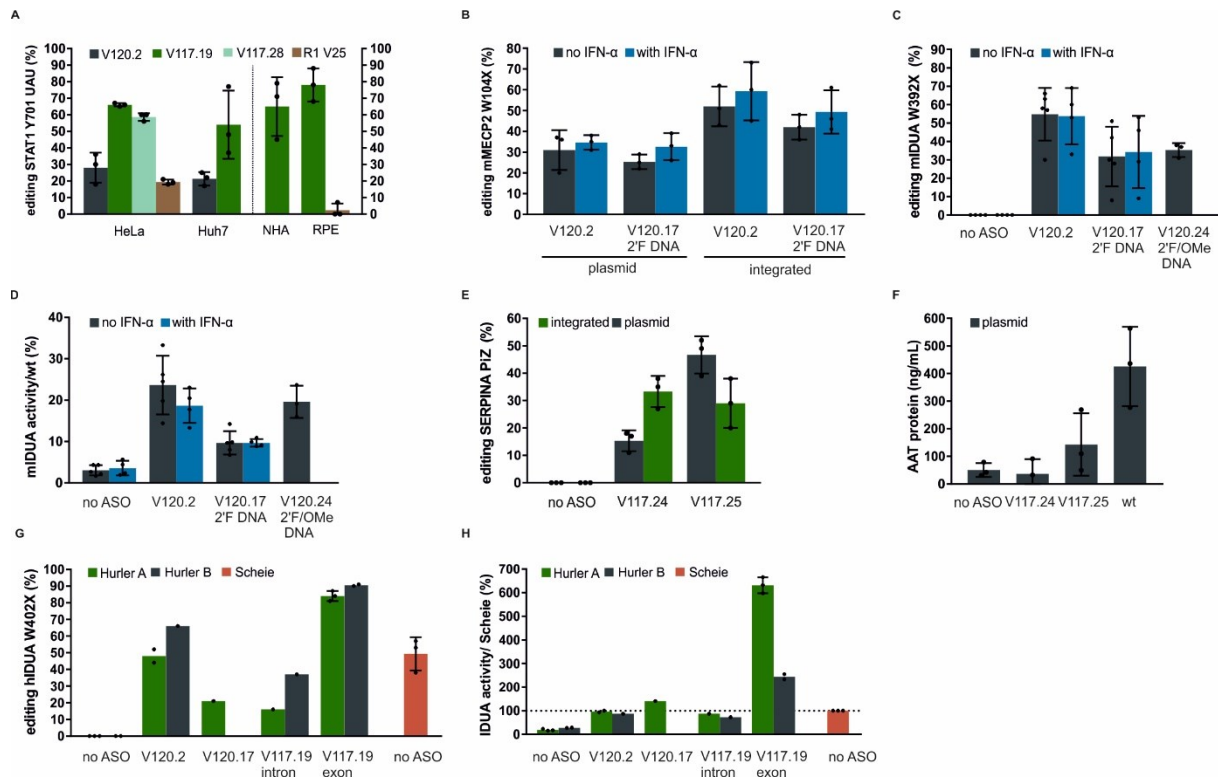


Figure 3. Application of RESTORE v2 ASO. A) Application of RESTORE v2 ASO. A) Editing of Y701 in endogenous STAT1. B) Editing of murine MeCP2 W104X in transgenic HeLa cells (PiggyBac) or under cDNA transfection (plasmid). C) Editing of murine IDUA W392X in HeLa cells or under cDNA transfection. D) Restoration of IDUA enzyme activity after editing. E) Editing of human Serpina1 E342K in transgenic HeLa cells or under cDNA transfection. F) Restoration of α -1-antitrypsin secretion after editing. G) Editing of endogenous human IDUA W402X in two different patient fibroblasts (A, B). Long ASOs are either targeting the pre-mRNA (intron) or the mature mRNA (exon). H) Restoration of IDUA enzyme activity after editing, normalized to IDUA activity of the residual activity from a patient suffering from the more benign Scheie phenotype. The exact sequences and modifications pattern of all ASOs are given in SI Table 4. Data are shown as the mean \pm s.d, where applicable. N =1-5 independent experiments as indicated by data points.

Editing in patient fibroblasts (Hurler syndrome)

To get more relevant data, we tested ASOs for the repair of an IDUA mutation directly in primary patient fibroblasts. The chosen mutation is highly relevant for human pathophysiology. Patients that carry the W402X mutation on both alleles suffer from Hurler syndrome, a very severe type I mucopolysaccharidosis that causes damage to various organs, including the nervous system, the eyes, the skeletal system, and the brain and results in premature death during infancy¹⁵. Patients with compound heterozygous mutation, e.g. which carry the W402X mutation in combination with a less severe mutation that still has residual IDUA activity, develop a more benign phenotype with a normal lifespan (Scheie syndrome)¹⁵. We treated two different Hurler fibroblasts with various ASO designs, including the short and the long lead, with and without additional modifications. The long lead (v117.19) was either designed to base-pair into a full-length duplex with the pre-mRNA only (covering the intron), or with the mature mRNA only (spanning two exons). We analyzed editing yield by Sanger sequencing and restoration of enzyme activity by a fluorogenic cleavage assay. The latter was normalized to activity in Scheie (Fig. 3H) or wildtype fibroblasts (SI Fig.10). Editing was only detectable in presence of an ASO. Both lead designs gave considerable editing, with the long design (v117.19) being superior reaching up to 90% editing (Fig. 3G). The exon-spanning variant was clearly better than

the intron/exon variant. The very high editing levels observed here, are likely overestimated and result from the stabilization of the edited transcript which can escape from nonsense-mediated decay. Accordingly, the observed restoration of enzyme activity lacked behind that of the RNA editing trace. However, the long lead design enabled a restoration of more than 6-fold of activity in the Scheie fibroblast (Fig. 3H), which was obtained from a patient with the milder disease phenotype, indicating that the obtained editing yield could be therapeutically relevant once obtained in a patient.

Conclusion

In summary, our data demonstrate the high potential to optimize ADAR-recruiting antisense oligonucleotides for therapy. With the new design rules, we achieved a strong reduction in ASO size, a strong increase of their stabilities in body fluids like FBS and CSF, and a substantial improvement of their editing yields in various cell lines including primary cells. These new properties enabled RESTORE v2 ASOs to fully overcome the former requirement for ADAR1 p150 induction, and to harness endogenous ADARs under gymnotic uptake of the ASO. Furthermore, we repeatedly demonstrate improved activity and specificity in the human disease-related context. Notably, we demonstrate improvement of bystander editing by backbone modification and a clear improvement of editing yield by nucleobase modification in the difficult *Serpina1* context. All these findings are highly instructive to pave the way for ADAR-recruiting ASOs in clinical settings. However, we screened only a manageable space of base and backbone modifications, and additional modifications, including cET, LNA, or 2'-MOE are likely to enhance their performance even further

Acknowledgements

We gratefully acknowledge the donation of the U2OS Flp-In cell line from E. Schiebel (Universität Heidelberg, Germany). We gratefully acknowledge support from the Deutsche Forschungsgemeinschaft to T.S. (STA 1053/3-2; STA 1053/7-1). This work is supported by the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63) with an intramural innovation grant for Jacqueline Wettengel.

Methods

Antisense oligonucleotides. All ASOs used in this study were purchased HPLC purified from Eurogentec (Belgium) or Biospring (Germany) and were directly used. The sequences and chemical modifications can be found in SI Tables 1-4.

Analysis of RNA editing. RNA editing was analyzed as reported earlier⁸: Briefly, total RNA was extracted, treated with DNase I, reverse transcribed and cDNA was amplified by Taq PCR. The PCR product was purified on an agarose gel and sent for Sanger sequencing. Editing yields were calculated by dividing peak height of the guanosine by the sum of the guanosine and adenosine peak heights at the target site (cytidine and thymidine peak heights when reverse primer were used for sequencing).

Editing procedure with ASOs in ADAR-expressing 293 cells. Procedure was performed as reported earlier⁸: Results are reported in Fig. 1C and SI Fig.1A,B.

Editing procedure for ASO screen in HeLa cells. Procedure was performed as reported earlier⁸: Results are reported in Fig. 1D and SI Fig.1C.

Editing procedure in immortalized cell lines. All cells were cultured in DMEM plus 10% FBS plus P/S, except for THP-1 which were cultured in RPMI plus 10% FBS. 1×10^5 cells/well (HeLa cells (cat. no. ATCC CCL-2), U2OS-Flp-In T-REx32 (kind donation from Elmar Schiebel), SK-NBE(2) (cat. no. ATCC CRL-2271), U87MG (cat. no. ATCC HTB-14), Huh7 (CLS GmbH, Heidelberg, cat. no. 300156), HepG2 (DSMZ, Braunschweig, Germany, cat. no. ACC180), SH-SY5Y (cat. no. ATCC CRL-2266), and A549 (European Collection of Authenticated Cell Cultures ECACC 86012804)) were seeded in a 24-well plate. After 24 h medium was changed (plus 3,000 U IFN- α) and cells were forward transfected with a transfection mix of 25 pmol ASO/well in 50 μ L OptiMEM and 1.5 μ L/well Lipofectamine RNAiMAX Reagent (ThermoFisher Scientific) in 50 μ L OptiMEM. Both solutions were combined after 5 min incubation and incubated for an additional 20 min before the transfection mix was distributed evenly into one well. After 24 h cells were harvested for RNA isolation and sequencing. THP-1 were transfected the same way after 3×10^5 cells/well of a 24-well plate were differentiated for 3 days in RPMI plus 10%FBS plus PMA (200nM) and cultured for 5 days in RPMI+10%FBS afterwards. Results are reported in Fig.1E,2B,3A and SI Fig.2,4.

Editing procedure in human primary cells. All primary cells were purchased from Lonza. Normal human astrocytes (NHA, Lonza cat. no. CC-2565) were cultured in ABM Basal Medium (Lonza cat. no. CC-3187) with AGM SingleQuot Kit Supplementary & Growth Factors (Lonza cat. no. CC-4123), human retinal pigment epithelial cells (H-RPE, Lonza cat. no. 00194987) were cultured in RtEBM Basal Medium (Lonza cat. no. 00195406) supplemented with RtEGM Retinal Epithelial Cell Growth Medium SingleQuots Supplements and Growth Factors (Lonza ca. no. 00195407) without FBS (for seeding FBS was added and after 24 h medium was changed to FBS-free medium), and normal human bronchial epithelial cells (NHBE, Lonza cat. no. CC-2540) were cultured in BEGM Bronchial Epithelial Cell Growth Basal Medium (Lonza cat. no. CC-3171) supplemented with BEGM Bronchial Epithelial Cell Growth Medium SingleQuots Supplements and Growth Factors (CC-4175). The transfection procedure was performed the same way as for immortalized cell lines with 1×10^5 cells/well seeded and 25 pmol ASO transfected. Results are reported in Fig.1F,2C,3A.

Gymnotic uptake. 10^4 cells/well RPE, NBE and NHA were seeded as were seeded into 24-well plates in the respective medium and after 24 h the medium was replaced with 250 μ L medium and 50 μ L ASO in OptiMEM was added with a final ASO concentration of 5 μ M. Cells were harvested for RNA isolation after three or five days. Results are reported in Fig.2D.

MECP2 editing and fluorescence microscopy: 3.5×10^3 HeLa with integrated mMECP2 (W104X or wild-type) were seeded in 24-well plates on glass coverslips. After 24 h medium was changed and cells were transfected with 10 pmol ASO as described in the editing procedure for immortalized cell lines. After 24 h cells were either harvested and RNA editing was analyzed or cells were fixated with 3.7% formaldehyde in PBS, washed, incubated with NucBlue Live ReadyProbes Reagent (ThermoFisher Scientific), washed and mounted with fluorescent mounting Medium (Dako). Microscopy was performed with a Zeiss CellObserverZ1 microscope. Results are reported in Fig.3B and SI Fig.5,6,7.

Editing procedure and protein extraction from mIDUA expressing HeLa: 5×10^4 HeLa cells were seeded in a 24-well plate. After 24 h cells were forward transfected with a plasmid containing either the hSERPINA1 wildtype cDNA or hSERPINA1 E342K cDNA on a plasmid. 300 ng plasmid and 0.9 μ L FuGENE[®] 6 (Promega) were each diluted in 50 μ L Opti-MEM and incubated for 5 min, then combined and incubated for an additional 20 min. The medium was changed, and the transfection mix evenly distributed into one well. 24 h after plasmid transfection, cells were forward transfected with 25 pmol ASO/well and 1.5 μ L/well Lipofectamine RNAiMAX Reagent (ThermoFisher Scientific). After 24 h, cells were harvested for RNA isolation and sequencing or cells were lysed in 100 μ L M-PER buffer (Thermo Scientific) for α -L-iduronidase enzyme assay. Results are reported in Fig.3C,D and SI Fig.8.

Editing procedure and protein extraction from fibroblasts. Fibroblasts from patients with Scheie syndrome (GM01323), Hurler syndrome (GM06214, “Hurler A” and GM00798, “Hurler B”), and from a healthy donor (GM05659) were purchased from the Coriell Institute for Medical Research (USA). Fibroblasts were cultivated in DMEM containing 15 % FBS. 2.5×10^5 cells/well in 2.5 ml DMEM plus 15 % FBS were seeded into 6-well plates, and for each tested condition, two 6-wells were used. Transfection was performed 24 h after seeding with 125 pmol ASO and 7.5 μ l RNAiMAX, each diluted in 250 μ l Opti-MEM. Both solutions were combined after 5 min incubation and incubated for an additional 20 min before the transfection mix was distributed evenly into one well. The medium was changed 24 h after transfection. 48 h after transfection, fibroblasts were detached and washed once with PBS. 40 μ l 0.5 % Triton X-100 in PBS were added to the cell pellet and incubated on ice for 30 min and α -L-iduronidase enzyme assay was performed. Results are reported in Fig.3G,H and SI Fig.10.

A-L-iduronidase enzyme assay. A standard dilution series of 4-methylumbelliferone was prepared. For each concentration, 25 μ l of the standard solution were added to 25 μ l 0.4 M sodium formate buffer (pH 3.5) in a 96-well plate. For the protein samples, 25 μ l of each solution were added to a 96-well and mixed with 25 μ l substrate solution (180 μ M 4-methylumbelliferyl α -L-iduronide in 0.4M sodium formate buffer, pH 3.5). The reaction was incubated at 37°C for 90 min and the enzyme activity stopped by adding 200 μ l glycine carbonate buffer (pH 10.4) to the well. The fluorescence of 4-methylumbelliferone was measured with an excitation wavelength of $\lambda_{ex} = 355$ nm at an emission wavelength of $\lambda_{em} = 460$ nm with a Tecan Spark 10M plate reader. Calculated enzyme activities were referenced to the protein amount (determined by Bradford Assay for HeLa and BCA assay for fibroblasts) and standardized to the enzyme activity of Scheie lysate. Results are reported in Fig. 3G and SI Fig.8,10.

SERPINA1 editing and AAT assay. 2.5×10^4 HeLa cells/well in 500 μ l DMEM plus 10 % FBS were seeded into 24-well plates. After 24 h cells were forward transfected with a plasmid containing either the hSERPINA1 wildtype cDNA or hSERPINA1 E342K cDNA on a plasmid. 300 ng plasmid and 0.9 μ l FuGENE[®] 6 (Promega) were each diluted in 50 μ l Opti-MEM and incubated for 5 min, then combined and incubated for an additional 20 min. The medium was changed and the transfection mix evenly distributed into one well. 24 h after plasmid transfection, cells were forward transfected with 5 pmol ASO/well and 1.5 μ l/well Lipofectamine RNAiMAX Reagent (Thermo Fisher Scientific). Cells were harvested for RNA isolation and sequencing after 48 h. After transfection, cell supernatants were collected, centrifuged and supernatant was frozen every 24 h and tested for their alpha 1 antitrypsin content with the alpha 1 antitrypsin human ELISA Kit (abcam, ab108799) according to the manufacturer’s protocol. Results are reported in Fig.3E,F and SI Fig.9.

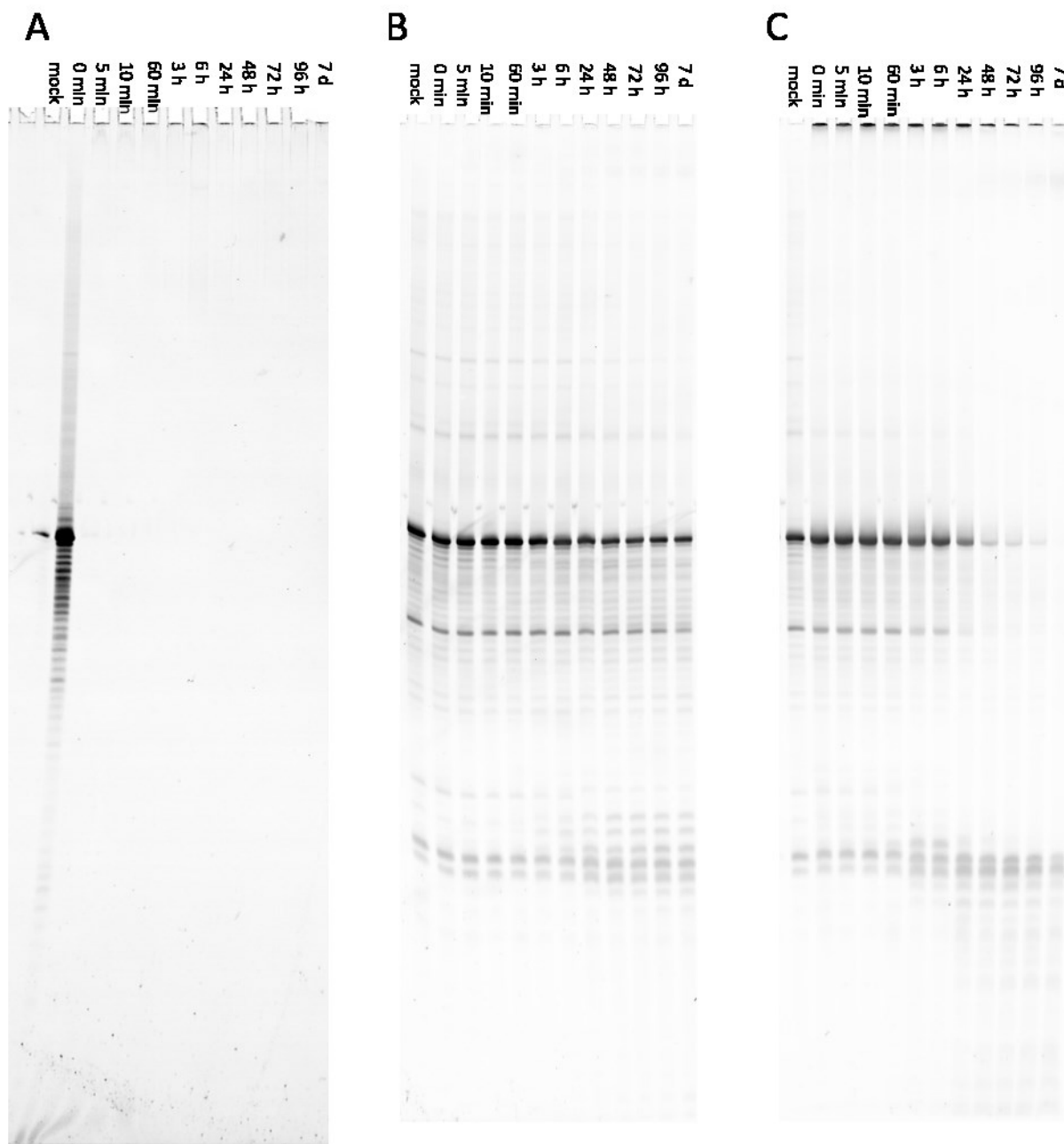
Stability assay of ASOs. 15 pmol of the respective ASO were diluted in 10 μ l PBS plus 10% FBS, 100 % FBS or 100% CSF, as noted. Mock samples contained 15 pmol ASO diluted in PBS only. All samples were incubated at 37°C for the given time points, then frozen and stored immediately at - 80°C. Denaturation of the samples was achieved by adding 7 μ l RNA loading dye (1:10 dilution of Rotiphorese[®] Sequencing gel buffer concentrate in Rotiphorese[®] Sequencing gel diluent, Carl Roth) each and incubation at 70°C for 2 min. For samples containing 100 % FBS, a proteinase K digestion was performed prior to the addition of RNA loading dye by adding 30 mM Tris-HCl (pH 7.5) and 60 μ g proteinase K (20 U/mg, Analytik Jena) to a final sample volume of 15 μ l. The digestion mix was incubated for 5 min at 50°C. Afterwards, 5 μ l RNA loading dye was added and the mix incubated for 2 min at 70°C. Denatured samples were then loaded on a urea (7 M) polyacrylamide (15 %) electrophoresis (PAGE) gel and run for 4 – 6 h at 1200 V in TBE buffer. Bands were visualized through a SYBR[™] Gold Nucleic Acid Gel Stain (Thermo Fisher Scientific) according to manufacturer’s instructions and scanned at the excitation wavelength $\lambda_{ex} = 473$ nm with a Fujifilm FLA-5100 Fluorescent Image Analyzer. Results are reported in Fig. 2A and SI Fig.3.

References

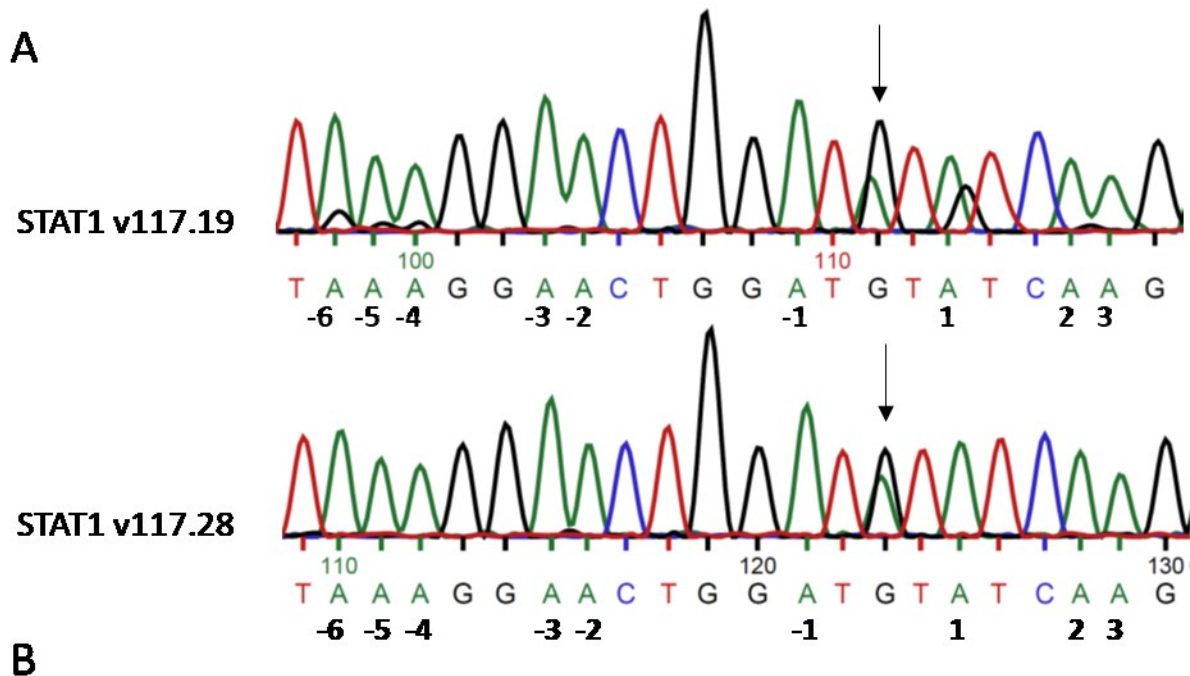
1. Vogel, P. & Stafforst, T. Critical review on engineering deaminases for site-directed RNA editing. *Current Opinion in Biotechnology* **55**, 74-80 (2019).
2. Rees, H.A. & Liu, D.R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature Reviews Genetics* (2018).
3. Stafforst, T. & Schneider, M.F. An RNA-deaminase conjugate selectively repairs point mutations. *Angew Chem Int Ed Engl* **51**, 11166-11169 (2012).
4. 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 Research* (2016).
5. 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. *Proceedings of the National Academy of Sciences* **110**, 18285-18290 (2013).
6. Cox, D.B.T. et al. RNA editing with CRISPR-Cas13. *Science* (2017).
7. Fukuda, M. et al. Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing. *Scientific Reports* **7**, 41478 (2017).
8. Merkle, T. et al. Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nature Biotechnology* **37**, 133-138 (2019).
9. Qu, L. et al. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nature Biotechnology* (2019).
10. Khvorova, A. & Watts, J.K. The chemical evolution of oligonucleotide therapies of clinical utility. *Nature Biotechnology* **35**, 238 (2017).
11. Zheng, Y., Lorenzo, C. & Beal, P.A. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res* (2017).
12. Woodard, L.E. & Wilson, M.H. piggyBac-ing models and new therapeutic strategies. *Trends Biotechnol* **33**, 525-533 (2015).
13. Sinnamon, J.R. et al. Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proceedings of the National Academy of Sciences* (2017).
14. 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).
15. Giugliani, R. et al. Mucopolysaccharidosis I, II, and VI: brief review and guidelines for treatment. *Genetics and Molecular Biology* **33**, 589-604 (2010).

SI Table 2: Sequences of single-stranded ASOs with different lengths and modifications patterns used in Figure 1. The C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, {N}=LNA base * = phosphorothioate linkage.

| GAPDH ORF1 UAG ASO sequences (5' to 3') : | |
|---|--|
| v117. 1 | [UUG](UCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGUUGGUGGUGCAGGAGCAUU)[GCU] |
| v117. 19 | [U*U*G*](U*C*A* U*G*G*A*U*G*A *C*C*U*U*G *G*C*C* A*G*G*G*GUGC C AAGC*A*G*U*U*G*G*U*G*G*U*G*G*C*A*G*G*A*G*G*C*A*U*U*)[G*C*U] |
| v118. 3 | [U*G*G*](A*U*G*A*C*C*U*U*G*G*C*C*A* G*G*G*GUGC C AAGC*A*G*U*U*G*G*U*G*G*U*G*C*A*G*G*A*)[G*G*C] |
| v119. 1 | [UCC](UUCACGAUACCAAGUUGUCAUGGAUGACCUUGGCCAGGGGUGC C AAGCAGU)[UGG] |
| v119. 4 | [U*C*C*](U*U*C*C*A*C*G*A*U*A*C*C*A*A*A*G*U*U*G*U*C*A*U*G*G*A*U*G*A*C*U*U*G*G*C*A*G*G*G*GUGC C AAGC*A*G*U*)[U*G*G] |
| v120. 2 | [C*A*A*](A*G*U*U*G*U*C*A*U*G*G*A*U*G*A*C*U*U*G*G*C*C*A* G*G*G*GUGC C AAGC*A*G*U*)[U*G*G] |
| v121. 1 | [U*U*G*](U*C*A*U*G*G*A*U*G*A*C*C*U*U*G*G*C*C*A* G*G*G*GUG CCA AGC*A*G*U*)[U*G*G] |
| v122. 1 | [A*U*G*](G*A*U*G*A*C*C*U*U*G*G*C*C*A* G*G*G*GUG CCA AGC*A*G*U*)[U*G*G] |
| v25 | [G*G*U](G)[UC](GAGAAGAGGAGAA)[C](AA)[U](A)[U](G)[CU](AAA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GA)[C](A)[C C](UUGUCAUGGAUGACCUUGGCCA)[G]{G}{GGUG}(C C A)[AGCA]{G*}[U*U*]{G*}[G] AminoC6 |



SI Figure 3: Stability assays of STAT1 v117.19. The influence of 2'-chemical modifications on the stability of the ASO was tested by tracing nuclease resistance over the course of one week in FBS at 37°C. Corresponding ASO sequences are depicted in **SI Table 4**. **A)** Stability assay of STAT1 v117.19 in PBS plus 10 % FBS. The ASO is degraded in a few seconds. **B)** The stability assay of STAT1 v117.28 shows nuclease resistance for seven days in PBS plus 10 % FBS. **C)** STAT1 v117.28 resists nuclease degradation in 100 % FBS for 24 h.

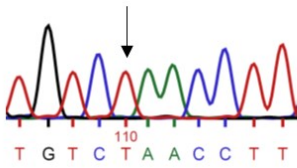


B

| A position | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | target | 1 | 2 | 3 | 4 | 5 | 6 |
|----------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---------------|----------|----------|----------|----------|----------|----------|
| STAT1 v117.19 | 0 | 0 | 0 | 15 | 10 | 14 | 0 | 0 | 0 | 66 | 36 | 9 | 0 | 0 | 0 | 0 |
| STAT1 v117.28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 59 | 0 | 0 | 0 | 0 | 0 | 0 |

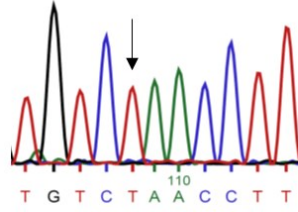
SI Figure 4: Bystander off-target editing with in the STAT1 transcript. HeLa cells were transfected with either v117.19 or its further stabilized counterpart v117.28 to edit endogenous STAT1. **A)** Exemplary sequence traces. Arrows indicate the target site and numbers indicate the potential off-target sites. **B)** The data in the table show the editing at all As that are spanned by the ASO.as the mean of N =3 independent experiments. The exact sequences and modifications pattern of all ASOs are given in **SI Table 4.**

no ASO:



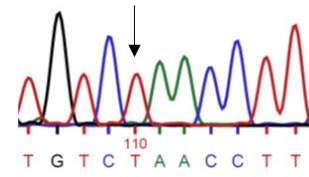
no editing

GAPDH V120.2:



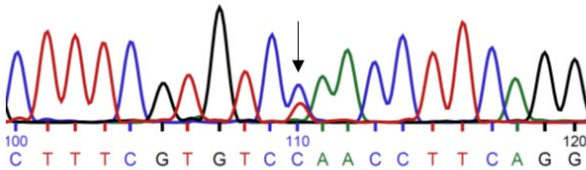
no editing

GAPDH V120.17:



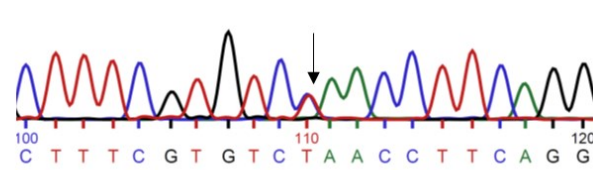
no editing

mMECP2 v120.2



66 % editing

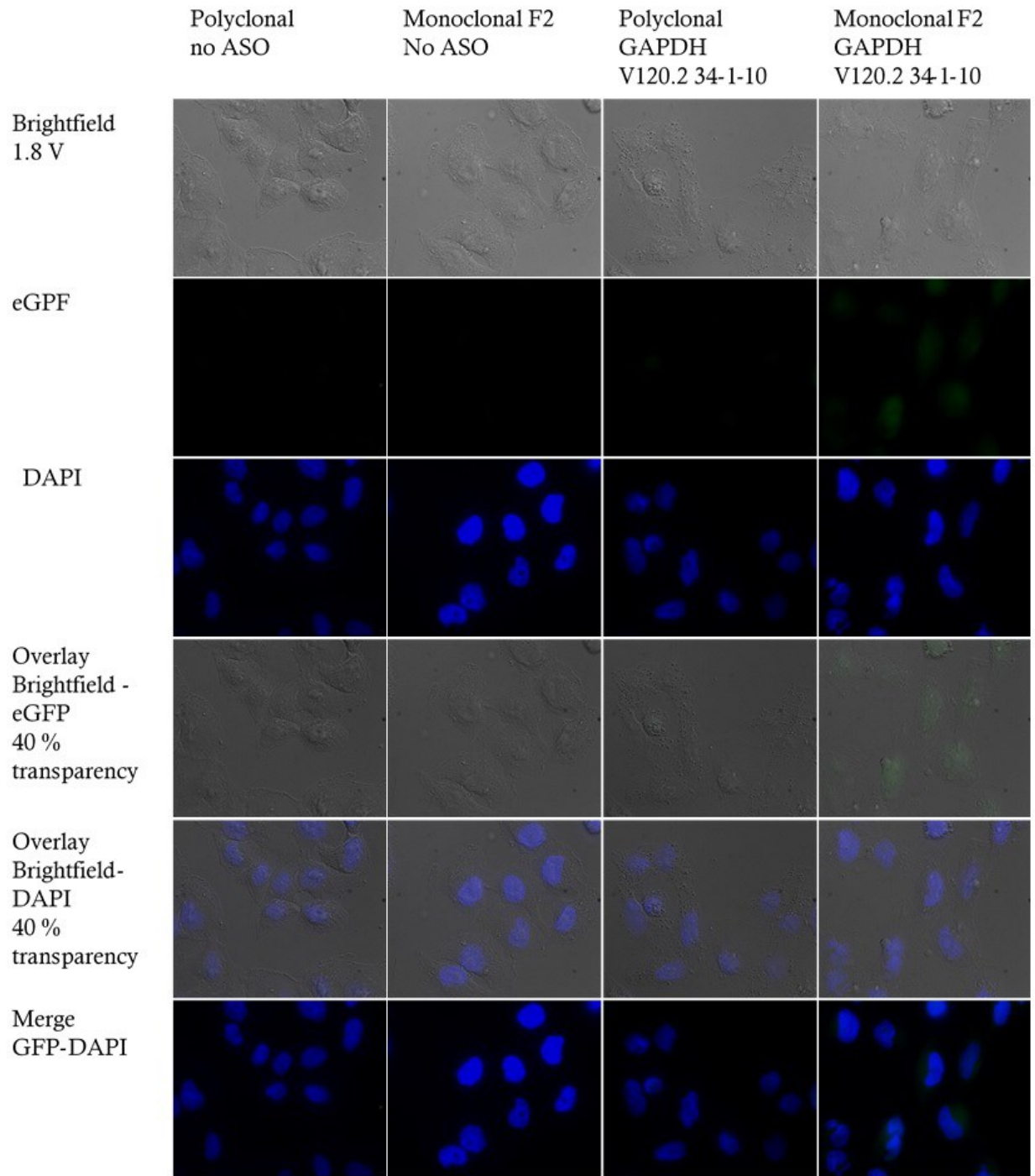
mMECP2 v120.17:



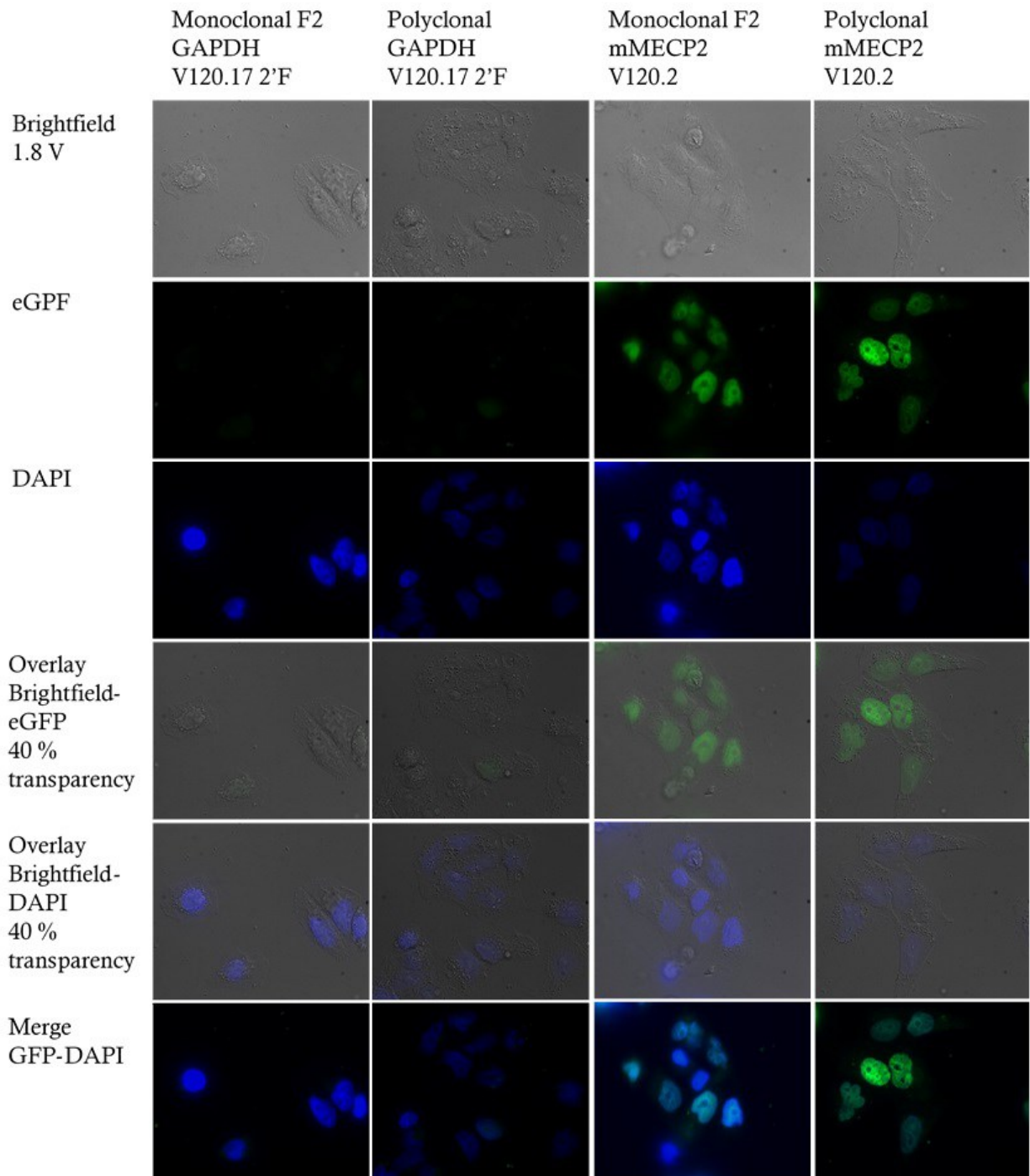
51 % editing

SI Figure 5: mMECP2 W104X editing. Exemplary traces of editing in HeLa cells with polyclonal integrated mMECP2 cDNA. Sequencing was performed with a reverse primer and target site is indicated with an arrow. Transfection of no ASO or ASOs against GAPDH show no editing while ASOs against mMECP2 show good editing yields without any bystander off-target editing.

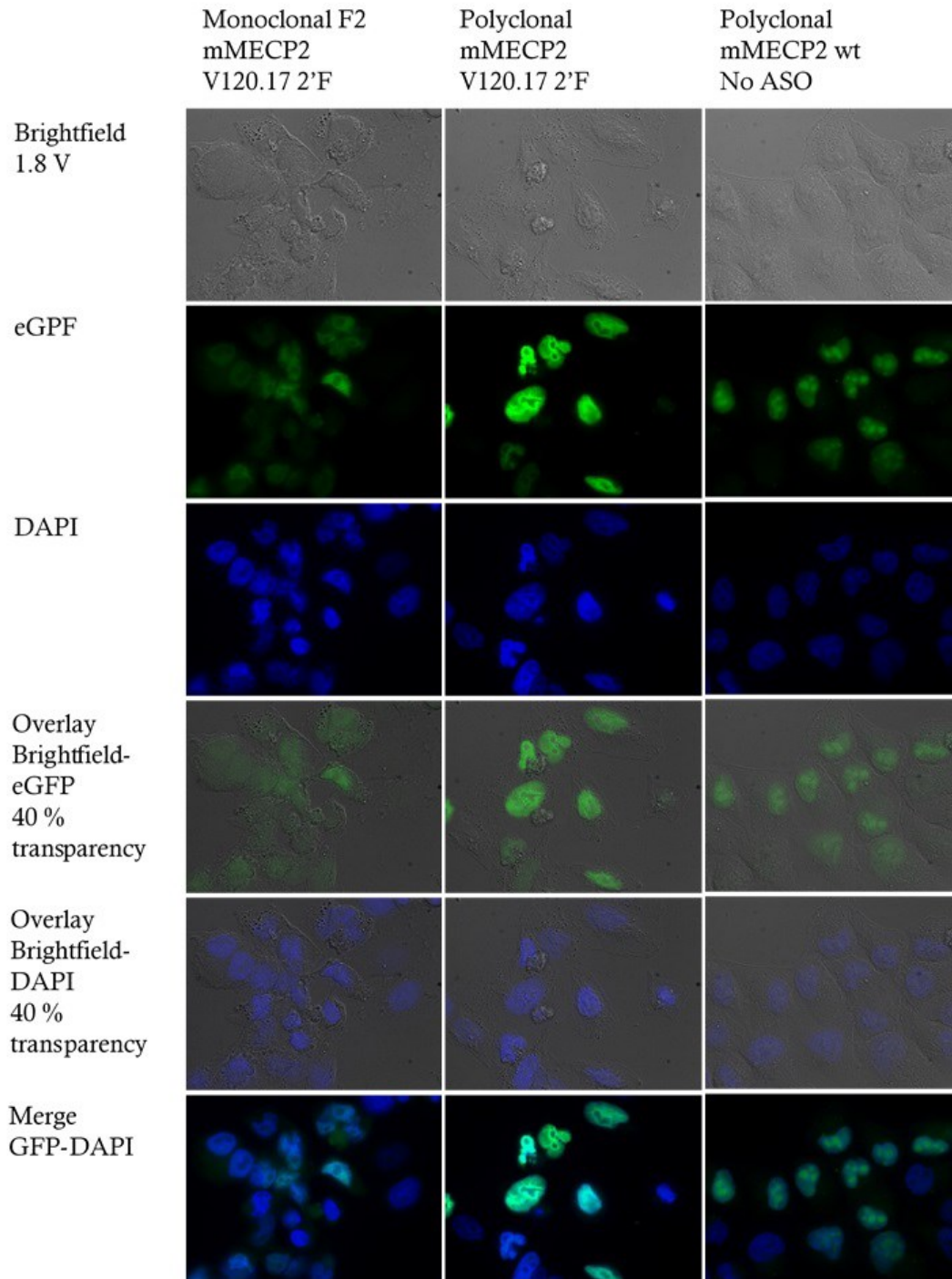
A



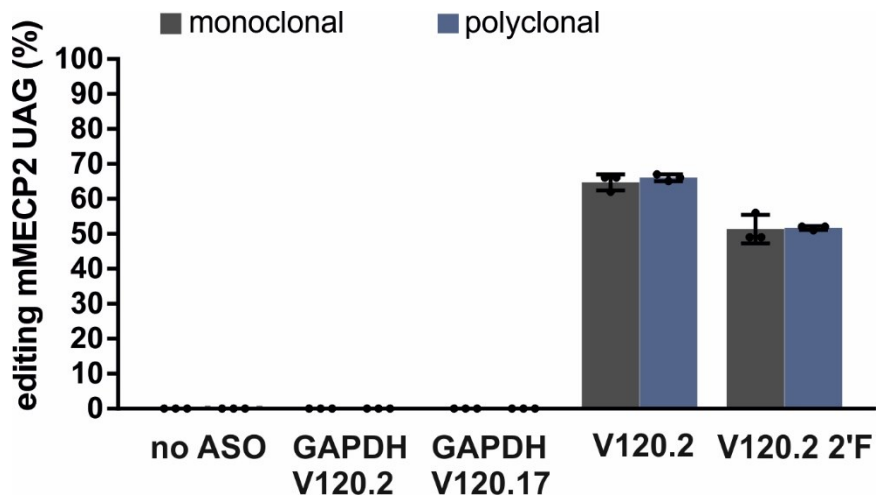
B



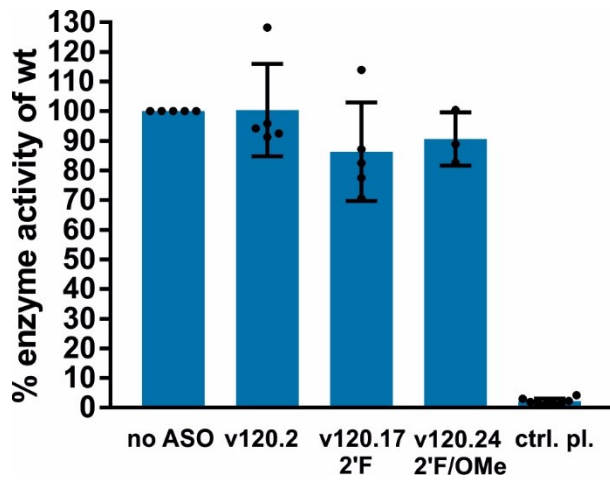
C



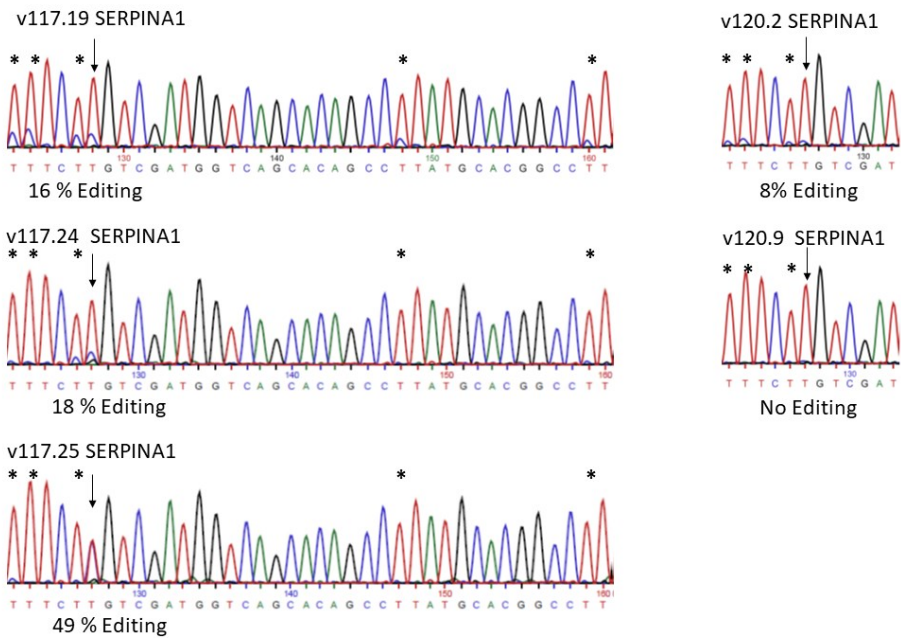
SI Figure 6: Microscopic analysis of HeLa with integrated mMECP2 W104X-EGFP cDNA after ASO transfection. The cDNA of mMECP2-EGFP or mMECP2 W104X-EGFP was stably integrated in HeLa and single clones were selected. In **A-C**) these HeLa were transfected with different ASOs, stained with Hoechst and analyzed under a fluorescence microscope. Cells transfected with ASOs against mMECP2 W104X show clear green fluorescence similar to mMECP2-EGFP cells. Pictures were taken at 630x magnification.



SI Figure 7: Editing analysis of HeLa with stably integrated mMECP2 W104X-EGFP. Editing results correspond to the microscopic pictures in **SI Figure 6**. The exact sequences and modifications pattern of all ASOs are given in **SI Table 4**. Data are shown as the mean \pm s.d of N=3 independent experiments.

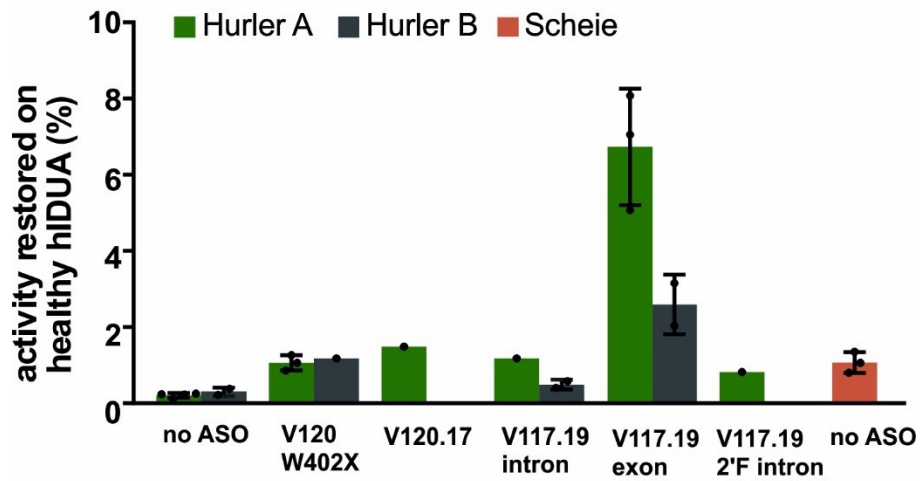


SI Figure 8: Effect of ASOs on translation. In HeLa cells transiently expressing mIDUA wt cDNA different ASOs were transfected and the enzymatic activity of the IDUA protein was analyzed. The exact sequences and modifications pattern of all ASOs are given in **SI Table 4**. Data are shown as the mean \pm s.d, N =3-5 independent experiments as indicated by data points.

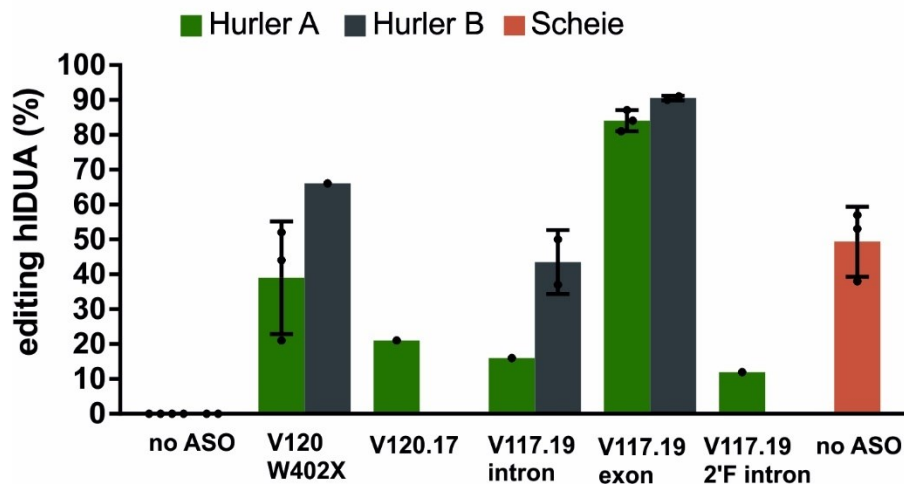


SI Figure 9: SERPINA editing and bystander off-target analysis. Editing experiments were performed in Hela transiently transfected with SERPINA cDNA expressing plasmid. Different ASOs were analyzed to suppress bystander off-target editing with high on-target editing. Arrows indicate target site and asterisks indicate bystander off-target sites. The exact sequences and modifications pattern of all ASOs are given in **SI Table 4**.

Fibroblasts hIDUA assay



Fibroblasts hIDUA editing



SI Figure 10: Restoration of IDUA enzyme activity and corresponding RNA editing. Normalized to IDUA activity of fibroblasts from a healthy donor. The exact sequences and modifications pattern of all ASOs are given in SI Table 4. Data are shown as the mean \pm s.d, where applicable. N =1-3 independent experiments as indicated by data points.

SI Table 4: Sequences of ASOs for disease relevant targets. The C opposite of the target A is highlighted in bold. (N)=RNA base, [N]=2'-OMe RNA base, fN=2'-F RNA base, N=DNA base, {N}=LNA base * = phosphorothioate linkage.

| STAT1 Y701C UAU ASO sequences (5' to 3'): | |
|---|--|
| v120.2 | [A*A*C*](U*U*C*A*G*A*C*A*C*A*G*A*A*A*U*C*A*A*C*U*C*A*G*U*C*U*U*GAUA C AUCC*A*G*U*)[U*C*C] |
| v117.19 | [C*A*G*](A*C*A*C*A*G*A*A*U*C*A*A*C*U*C*A*G*U*C*U*U*GAUA C AUCC*A*G*U*U*C*U*U*A*G*G*G*C*A*U*C*A*G*U*)[U*C*C] |
| v117.28 | [C*A*G*](A*)fC*(A*)fC*(A*G*A*A*A*)fU*fC*(A*A*)fC*fU*fC*(A*G*)fU*fC*fU*fU*(GA)fUA C AfUfCfC*(A*G*)fU*fU*fC*fU*fU*fU*(A*G*G*G*)fC*fC*(A*)fU*fC*(A*A*G*)fU*[U*C*C] |
| v25 | [G*G*U](G)[UC](GAGAAGAGGAGAA)[C](AA)[U](A)[U](G)[CU](A AA)[U](G)[UU](G)[UUCUC](G)[UCUCCUC](GACACCCA GACACAGAAAUCAACUCAGU)[C][T][UGAU](A C A) [UCCA]{G*}[U*U*]{C*}[C] Aminolinker |
| mMECP2 W104X UAG ASO sequences (5' to 3'): | |
| v120.2 | [U*C*G*](G*C*C*A*G*A*C*U*U*C*U*U*U*U*U*A*G*C*U*U*U*C*G*UGUC C AACCU*U*U*)[A*G*G] |
| v120.17 | [U*C*G*](G*)fC*fC*(A*G*A*)fC*fU*fU*fC*fC*fU*fU*fU*(G*)fU*fU*fU*(A*A*G*)fC*fU*fU*fU*fC*(G*)fU(G)fUC C (AA)fC fC*fU*fU*fC*[A*G*G] |
| miDUA W392X UAG ASO sequences (5' to 3'): | |
| v120.2 | [G*U*C*](C*A*A*C*A*C*A*G*C*C*C*A*G*C*U*U*U*G*A*G*A*C*U*U*U*G*CC C AGAG*U*U*G*)[U*U*C] |
| v120.17 | [G*U*C*]fC*(A*A*)fC*(A*)fC*(A*G*)fC*fC*fC*fC*(A*G*)fC*fC*fU*fU*fU*(G*A*G*A*)fC*fC*fU*fC*fU(G)fCC C (AGAG*)fU*fU*(G*)[U*U*C] |
| v120.24 | [G*U*C*C*](A*A*)fC*(A*)[C*](A*G*)fC*[C*]fC*[C*](A*G*)fC*[C*]fU*[U*]fU*(G*A*G*A*)[C*]fC*[U*]fC*[U](G)fCC C (AGAG*)[U*]fU*(G*)[U*U*C] |
| hSERPINA1 E342K CAA ASO sequences (5' to 3'): | |
| v120.2 | [A*A*A*](A*A*C*A*U*G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*U*UUCU C GUCG*A*U*G*)[G*U*C] |
| v120.9 | [A*A*A*](A*A*C*A*U*G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*C*)[U*U](UC)[U] (C GUCG*A*U*G*)[G*U*C] |
| v117.19 | [C*A*U*](G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*U*UUCU C GUCG*A*U*G*G*U*C*A*G*C*A*G*C*U*U*A*U*G*C*A*)[C*G*G] |
| v117.24 | [C*A*U*](G*G*C*C*C*A*G*C*A*G*C*U*U*C*A*G*U*C*C*)[U*UUCU] (C GUCG*A*U*G*G*U*C*A*G*C*A*G*C*)[U*U*](A*U*G*C*A*)[C*G*G] |
| v117.25 | [C*A*U*](G*G*)fC*fC*fC*fC*(A*G*)fC*(A*G*)fC*fU*fU*fC*(A*G*)fU*fC*fC*fC*fU*fUfUfC C IfUfC(G*A*)fU*(G*G*)fU*fC*(A*G*)fC*(A*)fC*(A*G*)fC*fC*fU*fU*(A*)fU*(G*)fC*A*)[C*G*G] |
| hiDUA W402X UAG ASO sequences (5' to 3'): | |
| v120.2 | [G*U*C*](C*A*G*G*A*C*G*U*C*C*G*G*C*U*G*C*G*A*C*A*U*U*C*GGCC C AGAG*C*U*G*)[C*U*C] |
| v120.17 | [G*U*C*]fC*(A*G*G*A*)fC*(G*G*)fU*fC*fC*fC*(G*G*)fC*fC*fU*(G*)fC*(G*A*)fC*(A*)fC*fU*fU*fC*(GG)fCC C (AGAG*)fC*fU*(G*)[C*U*C] |
| v117.19 intron | [G*G*A*](C*G*G*U*C*C*G*G*C*U*G*C*G*A*C*A*U*U*C*GGCC C AGAG*C*U*G*C*U*C*U*C*A*U*U*G*C*G*G*G*C*G*G*) [G*G*G] |
| v117.19 exon | [G*G*A*](C*G*G*U*C*C*G*G*C*U*G*C*G*A*C*A*U*U*C*GGCC C AGAG*C*U*G*C*U*C*U*C*A*U*U*G*C*G*G*C*G*G*)[A*G*C] |