

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Rational Design of Oligonucleotide Guide Strands for Site-Directed RNA Editing

Permalink

<https://escholarship.org/uc/item/1rb49872>

Author

Doherty, Erin Elizabeth

Publication Date

2022

Peer reviewed|Thesis/dissertation

Rational Design of Oligonucleotide Guide Strands for Site-Directed RNA Editing

By

ERIN ELIZABETH DOHERTY
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Peter A. Beal

Sheila S. David

Jared T. Shaw

Committee in Charge

2022

Dedicated to all who have served as my teachers, mentors, supporters, and friends.

TABLE OF CONTENTS

ABSTRACT	vii
SUMMARY OF FIGURES	ix
SUMMARY OF TABLES	xiii
ABBREVIATIONS	xiv
CHAPTERS	
1. Introduction to Therapeutic RNA Editing via Adenosine Deaminases Acting on RNA	1
RNA modification	1
Adenosine Deaminases Acting on RNA (ADARs)	3
Structural biology and mechanism of ADARs	5
Therapeutic RNA editing	8
2. Mechanistic Studies of pH-Responsive Changes in RNA Editing	13
Introduction	13
Results	16
The RNA editing reaction is intrinsically pH sensitive	16
Protonation of a conserved glutamate residue partially accounts for increases in RNA editing at acidic pH	18
Discussion	20
Methods	24
3. Chemically Modified Guide Oligonucleotides for RNA Editing via Endogenous ADARs	29

Introduction	29
Results	30
Orphan base analogs that mimic a hyperactive ADAR2 mutant	30
A fluorescence-based assay shows guide RNA-induced changes in base-flipping	32
Cytidine Analogs increase rate of A-to-I editing in vitro.	34
High-resolution structure reveals ADAR-dZ contacts	38
Directed editing with endogenous ADAR is enhanced by the dZ orphan nucleotide	41
Discussion	44
Methods.....	47
4. Enabling RNA Editing at Disfavored Sites through a Purine-Purine Pairing Interaction	59
Introduction	59
Results	60
ADAR2 editing of difficult target sites is enhanced by a G:G mismatch.....	60
G:G editing efficiency approaches that of ADAR's preferred substrate	61
A G:G pair supports editing in disease-relevant sequence contexts	63
Bystander edits within 5' G targets are eliminated by 2'-O-methyl modification	65
At least one dsRBD is essential for efficient ADAR2 editing of G:G pair-containing substrates	68
Purines are favored by ADAR1 in the guide strand -1 position	70
Guanosine in the -1 position enhances editing of an endogenous RNA in human cells ...	71

Structure-activity experiments support $G_{\text{syn}}:G_{\text{anti}}$ structure hypothesis.....	73
3-Deaza-2'-deoxyadenosine enhances editing of adenosines flanked by a 5' G.....	77
Discussion	78
Methods.....	79
5. Rosetta-based Structural Modeling Predicts Unique Features of the ADAR1 Catalytic Domain.....	87
Introduction	87
Results	89
ADAR1 operates through a base-flipping mechanism	89
Modeling supports discovery of a second metal binding site in the ADAR1 deaminase domain.....	91
Homology modeling of hADAR1 catalytic domain	94
Comparison of ADAR1d homology model and AlphaFold structure	101
Homology model of hADAR3 deaminase domain.....	104
Discussion	106
Methods.....	108
6. Efforts toward the Optimization of Oligonucleotides for Therapeutic RNA Editing	111
Introduction	111
I. RNA Editing as a Potential Chemotherapeutic	111
Editing of the Catalytic Lysine Codon in SRC mRNA is Supported by ADAR2.....	113
Cellular Editing of the SRC mRNA in HEK293T Cells with Overexpressed ADAR2 ..	114

II. A Chimera Allows for High-Throughput Screening of Sequences that Enable Editing in DNA/RNA Hybrids	114
Design of a DNA/RNA Chimera as an ADAR Substrate	116
A High-Throughput Screen identifies RNA Guide Sequences that Support Editing of DNA	117
III. A Phosphorothioate Walk Demonstrates ADARs Tolerance to Backbone Modifications.....	118
A Fully Phosphorothioated Guide Oligonucleotide Hinders Editing	119
ADAR2 is Tolerant of Short Stretches of Phosphorothioate Modifications Throughout the Guide Oligonucleotide	120
Methods.....	121
References	129
Resuse Permissions	163

Rational Design of Oligonucleotide Guide Strands for Site-Directed RNA Editing

ABSTRACT

Diseases of genetic origin can be treated by correcting underlying errors at the nucleic acid level. Genome editing has become a widely used approach, although there are limitations to current methods. The potential for off-target edits, delivery barriers, immune stimulation, and concerns over the feasibility of re-dosing limit the effectiveness of genome editing technologies. More recently, interest has grown in employing endogenous human enzymes for the correction of pathogenic mutations. The Adenosine Deaminase Acting on RNA (ADAR) family of human enzymes offer therapeutic potential due to their ability to convert adenosine to inosine in double stranded RNA. Human ADARs can be directed to predetermined target sites in the transcriptome by complementary guide strands, allowing for the correction of disease-causing mutations at the RNA level. A particular advantage of RNA editing as opposed to DNA editing is that any potential off-target transcript editing does not result in a permanent change to the genome. These studies aim to develop a mechanistic understanding of the protein-nucleic acid contacts that enable ADAR to edit RNA, and to exploit this knowledge in the design of optimized guide strands for site-directed RNA editing. *Chapter 1* gives a broad introduction to ADAR enzymes and their use in directed RNA editing.

In *Chapter 2*, we explore a specific ADAR-RNA contact that has a large influence on the rate of reaction. This understanding helps us to rationalize pathophysiologic conditions that are associated with dysregulated RNA editing. Moreover, understanding the mechanistic basis of this important interaction allows for the design of modifications to the guide strand in this position in *Chapter 3*. The single nucleobase modification we identified leads to over a 3-fold increase in the directed editing yield via endogenous ADARs. This result advances the approach of recruiting endogenous ADARs for site-directed RNA editing.

It is typical for directed editing with ADARs to focus on sites within a 5'-UAG-3' sequence context, as this is the natural substrate preference of ADARs. However, in *Chapter 4* we expand upon previous work

to support ADAR activity within editing-resistant sequence contexts containing a 5' guanosine. We found that pairing this 5' guanosine across a purine in the guide strand led to between an 8- and 60-fold rate enhancement. In addition, through *in vitro* studies with modified purine analogs we identified positions on the purine base that have a strong influence on the ADAR reaction, to allow for the design of chemically modified nucleosides that promote even greater ADAR activity. This expands the scope of disease-causing mutations that can be effectively targeted by ADARs.

The crystal structure of the ADAR2 deaminase domain bound to dsRNA inspired the work done in *Chapters 2 through 5*. However, there is no available high-resolution crystal structure of the other catalytically active enzyme in the family: ADAR1. Therefore, we generated Rosetta-based molecular models with constraints from biochemical data to define structural features unique to ADAR1. These models support the discovery of a novel zinc binding site present on the surface of the ADAR1 deaminase domain but absent in ADAR2. Furthermore, the models explain previously observed properties of the ADAR1 deaminase domain and suggest roles for specific residues present in a binding loop that is partially responsible for substrate selectivity. Given the success of the ADAR1 deaminase domain model in making predictions about the roles of specific residues in the protein, a model of the ADAR3 deaminase domain was also generated.

Lastly, *Chapter 6* details ongoing efforts to improve ADAR editing of non-ideal targets of disease relevance, while expanding our understanding of ADAR's tolerance for chemical modifications within the guide strand that are critical for cellular editing.

SUMMARY OF FIGURES

CHAPTER 1

1.1. RNA modifications	1
1.2. Biological outcomes of RNA editing	2
1.3. Adenosine deamination	3
1.4. Domain map of ADARs	4
1.5. 2-Aminopurine base-flipping assay	5
1.6. ADAR reaction mechanism	6
1.7. ADAR2d crystal structure and RNA contacts	7
1.8. Site-directed RNA editing	10
1.9. Chemical modification of therapeutic RNAs	12

CHAPTER 2

2.1. The effect of varying pH on RNA editing	17
2.2. Residue 488 is implicated in the pH dependence of RNA editing.....	19
2.3. Intracellular acidification leads to enhanced RNA editing	23

CHAPTER 3

3.1. ADAR2 E488 contact with guide RNA	30
3.2. ADAR2 E488 and E488Q interaction with guide RNA	30
3.3. Crystal structures of flipping residue interactions with guide RNA	31
3.4. Base flipping with orphan analogs	34

3.5. <i>In vitro</i> deamination of substrates with orphan analogs	36
3.6. Bystander editing with orphan analogs and mutant ADAR2.....	37
3.7. Crystal structure of ADAR2d bound to RNA containing an orphan analog	38
3.8. Confirming the presence of the dZ base in the crystal structure.....	39
3.9. Density Functional Theory calculations of the orphan analog	40
3.10. Structure of 2'-deoxy orphan base bound to ADAR2d	41
3.11. Cellular editing of the <i>RAB7A</i> target with overexpressed ADAR2	42
3.12. Editing of the <i>IDUA</i> target with endogenous ADAR	43

CHAPTER 4

4.1. Structure of the base 5' to the edit site.....	59
4.2. Deamination of a 5' G substrate with ADAR2	61
4.3. Comparison of a 5' G and 5' U substrate in reaction with ADAR2	62
4.4. Deamination of the <i>MECP2</i> R168X substrate with ADAR2.....	64
4.5. Deamination of the <i>MECP2</i> R255 substrate with ADAR2	65
4.6. Off-target editing of substrates containing a 5' guanosine	66
4.7. 2'-O-methylated guides reduce off-target editing by ADAR2	67
4.8. Deamination of a 5' G substrate with truncated ADAR2 constructs	69
4.9. ADAR1 deamination of a 5' G containing substrate	71
4.10. Cellular editing of a 5' G containing substrate	72
4.11. Proposed conformation of the purine:purine pair	74

4.12. Purine analogs for structure-activity studies	75
4.13. Rate of deamination for ADAR2 with purine analogs	76
4.14. Rate of deamination for ADAR1 with purine analogs	78

CHAPTER 5

5.1. ADAR1 base-flipping with Duplex B and C	90
5.2. ADAR1 base-flipping with Duplex A	91
5.3. Sat-FACS-seq analysis of ADAR1d	92
5.4. Prediction of metal binding residues in ADAR1	94
5.5. Sequence alignment of the ADAR1 and ADAR2 binding loops	95
5.6. Sat-FACS-Seq of the ADAR1 and ADAR2 binding loops	96
5.7. Representation of values used for metal binding constraints	97
5.8. Top ten lowest energy structures of ADAR1d	98
5.9. Comparison of modeling constraints with ADAR2 structure	99
5.10. ADAR1d model overlaid with ADAR2d structure	99
5.11. Comparison of two lysine residues in the ADAR1 binding loop	100
5.12. ADAR1 full length Alpha Fold model	102
5.13. Comparison of contacts in the Alpha Fold and Rosetta models	103
5.14. Ten lowest energy structures of ADAR3d overlaid	105
5.15. Alignment of the binding loops of ADAR1, ADAR2, and ADAR3	106

CHAPTER 6

6.1. ADAR editing of protein kinase mRNA scheme	112
6.2. ADAR editing of the <i>SRC</i> mRNA target	113
6.3. Cellular editing of the <i>SRC</i> mRNA target.....	114
6.4. ADAR2 contacts with the target strand of RNA	115
6.5. Design of a DNA/RNA chimera as an ADAR substrate	117
6.6. Winning sequences from the EMERGE screen of DNA/RNA hybrid editing.....	117
6.7. Phosphorothioate walk diagram.....	119
6.8. Editing with modified vs unmodified guide RNAs	120
6.9. Result of the phosphorothioate walk.....	121

SUMMARY OF TABLES

CHAPTER 3

3.1. Single turnover rate constants of ADARs with orphan analog substrates	35
--	----

CHAPTER 4

4.1. Rate constants for ADAR2 deamination of 5' G <i>IDUA</i> substrate	61
4.2. Rate constants for ADAR2 deamination of 5' G and 5' U <i>IDUA</i> substrates	63
4.3. Rate constants for ADAR2 deamination of R168X <i>MECP2</i> substrate.....	64
4.4. Rate constants for ADAR2 deamination of R255 <i>MECP2</i> substrate	65
4.5. Rate constants for ADAR2d E488Q and ADAR2d deamination of <i>IDUA</i> substrate	70
4.6. Rate constants for ADAR1 p110 deamination of <i>IDUA</i> substrate.....	71
4.7. Rate constants for ADAR2 deaminations with purine analogs	77

CHAPTER 5

5.1. Measurements used as constraints for Rosetta modeling	97
--	----

SUMMARY OF ABBREVIATIONS

2-AP	2-Aminopurine
2'-F-ANA-G	2'-Fluoroarabinoguanosine
3-deaza-dA	3-Deaza-2'-deoxyadenosine
5HT _{2c}	5-Hydroxytryptamine receptor 2C
5mC	5-Methylcytosine
7-deaza-dG	7-Deaza-2'-deoxyguanosine
8-AN	8-Azanebularine
8-Br-dG	8-Bromo-2'-deoxyguanosine
A	Adenosine/Adenine
AAG	Alkyladenine DNA glycosylase
AAV	Adeno-associated virus
ADAR	Adenosine Deaminase Acting on RNA
ADAR1/2 RD	ADAR with one dsRBD
ADAR2d or ADAR1d	Deaminase domain of ADAR
AGS	Aicardi-Goutières Syndrome
ANOVA	Analysis of variance
ARPE-19	Human retinal pigment epithelium
ATCC	American Type Culture Collection
BDF2	Bromodomain factor 2
BSA	Bovine serum albumin
C	Cytidine/Cytosine
CCP4	Collaborative Computational Project Number 4
CM	Complete minimal
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
dA	2'-Deoxyadenosine
dC	2'-Deoxycytidine
ddPCR	Digital droplet PCR
DFT	Density functional theory
dG	2'-Deoxyguanosine

dI	2'-Deoxyinosine
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleoside triphosphate
DSH	Dyschromatosis Symmetrica Hereditaria
dsRBD	Double-stranded RNA binding domain
dsRNA	Double-stranded RNA
DTT	Dithiothreitol
dUTP	Deoxyuridine triphosphate
dZ	Benner's base Z or 6-amino-5-nitro-3-(1'- β -D-2'-deoxyribofuranosyl)-2(1H)-pyridone
EDTA	Ethylenediaminetetraacetic acid
EMERGe	En Masse Examination of RNA Guides
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FI	Fluorescence intensity
G	Guanosine/Guanine
GalNAc	N-Acetylgalactosamine
GLI1	Glioma-associated oncogene homologue 1
GLUA2	GRIA2-encoded AMPA receptor ligand binding core
GluRB or GRIA2	Glutamate-gated ion channel receptor subunit B
GRIA2	Ionotropic AMPA glutamate receptor subunit 2 gene
HEK	Human embryonic kidney
HOGG1	Human 8-oxoguanine DNA N-glycosylase 1
HR	Homologous recombination
I	Inosine/Hypoxanthine
IDUA	Alpha-L-iduronidase
IFN	Interferon
LOWESS	Locally weighted scatterplot smoothing
m1A	N1-methyladenosine

m6A	N6-methyladenosine
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight
MBP	Maltose binding protein
MELAS	Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes
MIB	Metal Ion-Binding site prediction and docking server
miRNA	micro RNA
MTase	Methyltransferase
NES	Nuclear export signal
NGS	Next-Generation Sequencing
NHEJ	Non-homologous end joining
NHP	Non-human primate
NLS	Nuclear localization signal
NMWL	Normal molecular weight limit
NT	No transfection
NTA	Nitroloacetic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PHENIX	Python-based Hierarchical Environment for Integrated Xtallography
piC	Pseudoisocytidine
PO	Phosphodiester
PS	Phosphorothioate
RAB7A	Ras-related protein Rab-7a
RIG-I	Retinoic acid-Inducible Gene I
RMSD	Root-mean square deviation
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	RNase inhibitor
RNP	Ribonucleoprotein
RosettaCM	Rosetta Comparative Modeling

RT-PCR.....Reverse transcription-Polymerase chain reaction
 Sat-FACS-Seq.....Saturation Mutagenesis-FACS-Sequencing
 Scr Scrambled
 SD..... Standard deviation
 SDSSodium dodecyl sulfate
 SEC Size exclusion chromatography
 SINE.....Short interspaced nuclear elements
 SNAr Nucleophilic aromatic substitution
 SRCproto-oncogene protein kinase Src
 TALEN.....Transcription activator-like effector nucleases
 TEVTobacco Etch Virus
 U Uridine/Uracil
 UDGUracil-DNA glycosylase
 Ura.....Uracil
 WT Wild-type
 XDS.....X-ray detector software
 XL-MS Chemical crosslinking-mass spectrometry
 ZFNZinc finger nuclease
 ΨPseudouridine

CHAPTER 1

Introduction to Therapeutic RNA Editing via Adenosine Deaminases Acting on RNA

This chapter contains excerpts from a commentary published in Molecular Therapy in May 2022.¹ © 2022 The American Society of Gene and Cell Therapy.

RNA Modifications

Many cellular processes are regulated by RNA sequence and structure, and dysregulation of these modifications is implicated in pathophysiologic conditions. The extent to which RNA modifications, as opposed to the four canonical ribonucleosides, contribute to cellular regulation is being further realized with recent advances in sequencing and mass spectrometry that allow for more sensitive detection.² There are over 150 RNA modifications including base modifications seen in eukaryotic mRNAs.³ RNA base modifications include adenosine to inosine (A-to-I) and cytidine to uracil (C-to-U) deamination, pseudouridine (ψ), N6-methyladenosine (m6A), N1-methyladenosine (m1A), and 5-methylcytosine (5mC) (Figure 1.1).³

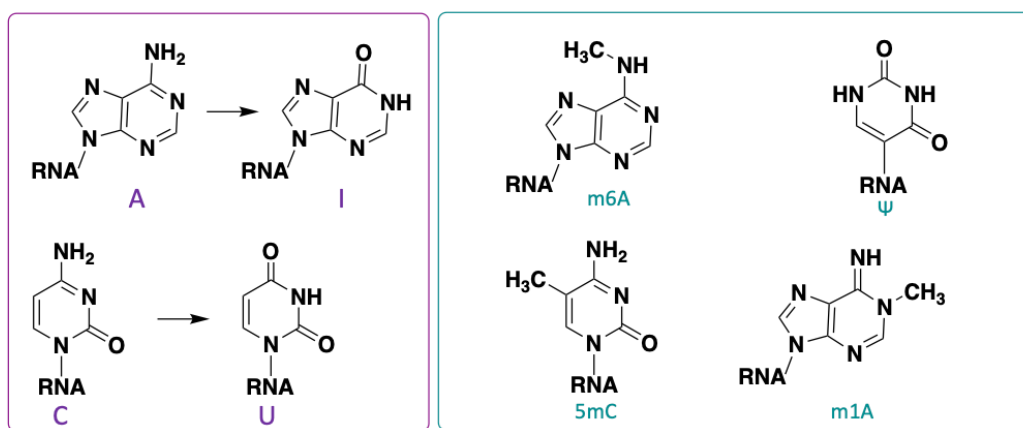


Figure 1.1. RNA Nucleobase Modifications. A) RNA Modifications resulting from deamination. B) Modified bases found in the epitranscriptome. RNA editing is a subset of these modifications comprised of enzymatic processes which insert, delete, or modify ribonucleotides, thus changing the mRNA coding properties.⁴ RNA editing can result in an mRNA that gives rise to an amino acid sequence different than what is encoded by the genomic DNA. Post transcriptional modifications are referred to as the ‘epitranscriptome’ and allow for diversity in RNA and protein sequences beyond the limited number of DNA genes.^{3,5} Modification at the RNA level also allows for spatial and temporal changes based on stages of development or cellular

conditions.⁵ However, the biological implications of RNA editing extend beyond mRNA re-coding. This process serves to regulate alternative splicing, transcription, RNA interference, innate and adaptive immunity, and microRNA (miRNA) recognition (Figure 1.2).⁶ . Specificity of dsRNA binding proteins for their substrates is based largely on the A-form helical structure.⁷ This overall structure can be disrupted by A-to-I edits, which contributes to the observed changes that edits can induce in various pathways. Dysregulated RNA editing has also been associated with autoimmune diseases and cancers.⁸

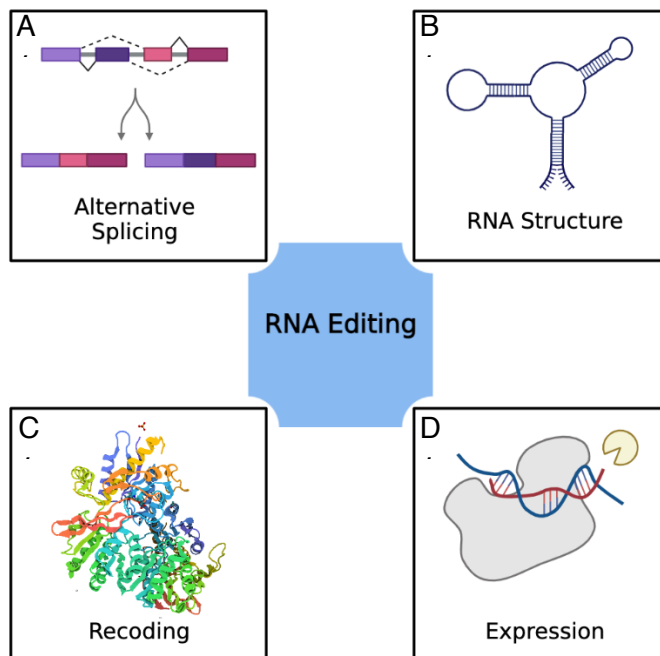


Figure 1.2. Biological outcomes of RNA Editing. A) Editing of mRNA can introduce or delete splice sites leading to differences in alternative splicing. B) RNA editing changes the base pairing properties of RNA which can affect RNA secondary structure. C) mRNA editing can re-code proteins through codon changes. D) Edits in microRNA recognition sites can affect miRNA binding and therefore regulate gene silencing.

Of these enzymatic RNA editing events, one of the most common in metazoans is adenosine to inosine (A-to-I) deamination.⁹ Inosine is abundant in many different types of RNA, including pre-mRNA, mRNA, and noncoding RNA.³ The Watson-Crick-Franklin face of inosine has a pattern of hydrogen bond donors and acceptors that resembles that of guanosine, allowing it to pair selectively with cytidine (Figure 1.3). Therefore, it is recognized by cellular machinery (including ribosomes during translation) as guanosine resulting in functional differences in the RNA.⁵ A-to-I editing is highly conserved and widespread throughout the transcriptome. This A-to-I reaction is catalyzed by the adenosine deaminase acting on RNA (ADAR) family of enzymes.

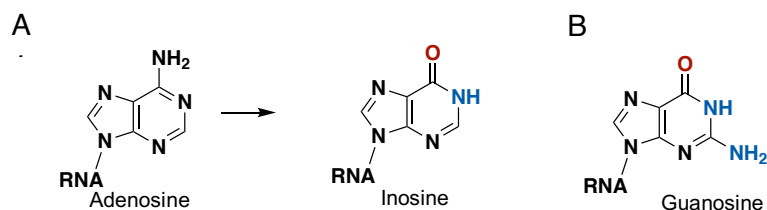


Figure 1.3. A) Adenosine deamination resulting in an inosine product. B) Guanosine has a similar pattern of hydrogen bond donors (blue) and acceptors (red) as inosine.

Adenosine Deaminases Acting on RNA (ADARs)

The human ADAR family contains three members: ADAR1 (p150 and p110 isoforms), ADAR2, and ADAR3.¹⁰ Of this family, only ADAR1 and ADAR2 are catalytically active.¹¹ ADAR enzymes consist of a catalytic deaminase domain, and double stranded RNA binding domains (dsRBDs) which contribute to substrate recognition (Figure 1.4).¹² Additionally, ADAR1 has a Z-DNA binding domain and ADAR3 has an N-terminal arginine rich domain.¹¹ ADAR1 and ADAR2 are ubiquitously expressed across many tissues, while ADAR3 is primarily found in the brain.^{11,13} The differential subcellular localization of ADARs is determined by nuclear localization and nuclear export signals. Both isoforms of ADAR1 contain a nuclear localization signal (NLS), however the p150 isoform also contains a nuclear export signal (NES) which leads to its activity as a nucleocytoplasmic shuttling protein.¹⁰ Despite lacking an NES, ADAR1 p110 also shuttles between the nucleus and cytoplasm due to its interaction with export factor exportin-5.¹⁰ ADAR2 is found to localize to the nucleolus due to a putative NLS at its N-terminus.¹⁴ Subcellular localization of the ADARs may be a mechanism by which their editing activity is regulated.

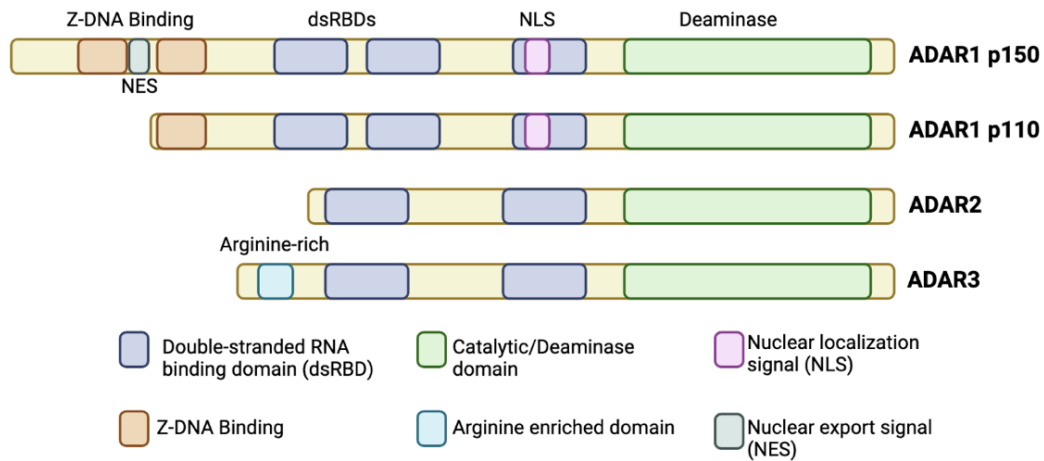


Figure 1.4. Domain maps of ADAR1 (p110 and p150), ADAR2, and ADAR3.

Our understanding of ADAR's biological role is continuously developing. A-to-I edits can be found in both coding and non-coding regions of the RNA.¹⁰ ADAR1 or ADAR2 deficiency in mice leads to abnormal development and is incompatible with life.¹⁵ ADAR2 knockout is lethal in mice due to the lack of an essential edit in the coding region of the glutamate-gated ion channel receptor subunit B (*GluR-B* or *GRIA2*).¹⁶ This editing event decreases the calcium permeability of a class of glutamate receptors responsible for modulating cell excitability, meaning that ADAR2 activity is essential for proper function of neurotransmitter receptors.^{16,17} In contrast, there is no single known ADAR1-dependent essential editing event.¹⁰ However, knockout of ADAR1 is even more detrimental in mice, leading to embryonic lethality due to liver damage and inflammation.¹⁸ ADAR1 plays a major role in innate immunity, by editing RNAs to suppress their activation of RNA binding proteins such as Retinoic acid-Inducible Gene I (RIG-I).¹⁹ Activation of RIG-I by cytoplasmic double-stranded RNA (dsRNA) leads to hyper-activation of dsRNA sensing pathways triggering excessive type I interferon (IFN) production, which accounts for the embryonic lethal phenotype.¹⁹ Editing deficient ADAR1 mutants also result in high levels of interferon production such as seen in Aicardi-Goutières Syndrome (AGS) which is an autoimmune disease associated with mutations in the ADAR1 gene.⁹ More recently, it has been found that ADAR1 activity is critical to the survival of a subset of cancers.²⁰ These cancers have an interferon stimulated gene signature, and are

dependent on ADAR1 editing to avoid detection by the innate immune system.^{21,22} The importance of ADAR1 in disease states and as a potential therapeutic target is discussed further in *Chapter 5*.

Structural biology and mechanism of ADARs

ADAR enzymes operate through a base-flipping mechanism where the edited base is flipped out of the duplex and into the enzyme active site.²³ This was first evidenced in an assay that uses a fluorescent nucleoside analog 2-aminopurine (2-AP) as a probe for base-flipping by ADAR2. 2-AP has been used as a probe for nucleic acid structure because its fluorescence properties respond to changes in its local environment.²⁴ When stacked as part of a duplex, 2-AP fluorescence is quenched, but when unstacked (such as through a base-flipping event) there is a restoration of fluorescence (Figure 1.5). This allows 2-aminopurine to indicate the extent of base-flipping by ADARs when it is included in the edited position of a dsRNA substrate.²³ Initial studies demonstrated a large enhancement in fluorescence when ADAR2 was combined with a substrate containing 2-aminopurine.²³ This suggested that a base-flipping event where the adenosine is removed from the helix is likely a part of the reaction mechanism. Thus far, this fluorescence restoration has only been demonstrated for substrates in the presence of ADAR2. In *Chapter 5* we use this same assay to provide evidence for a base-flipping mechanism by ADAR1.

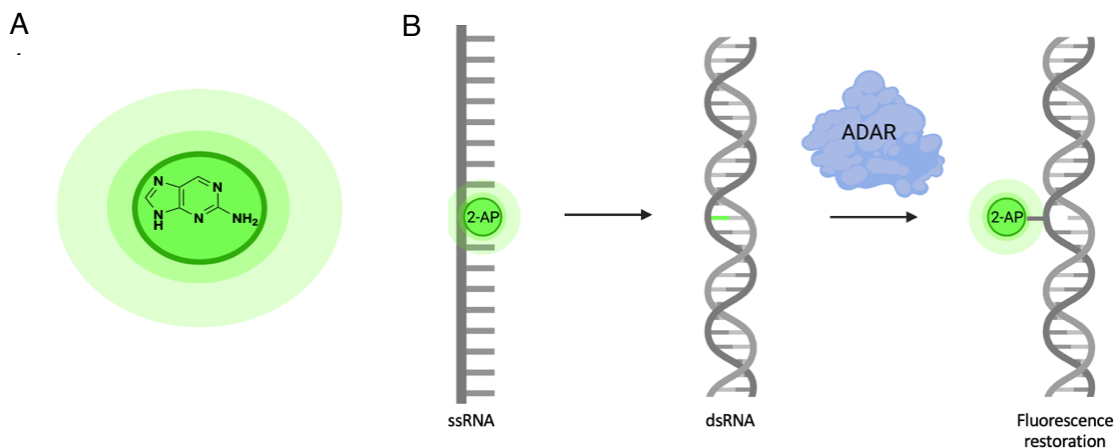


Figure 1.5. 2-aminopurine base-flipping assay. A) The 2-aminopurine (2-AP) nucleotide serves as a probe for base-flipping by the ADAR2 enzyme. B) 2-AP is quenched as part of a duplex, but unstacked by ADAR2 which restores fluorescence.

The crystal structure of the ADAR2 catalytic domain (ADAR2d) bound to dsRNA was solved in 2016, which confirmed that the enzyme operates through a base-flipping mechanism and allowed for the contacts enabling this conformation to be observed.²⁵ Crystallizing the enzyme bound to its substrate was dependent upon the identification of substrates that offered high binding affinity to ADAR2 without initiating turnover and substrate release. This relied on the incorporation of a modified nucleoside that served as a transition state analog.^{25,26} The design of this transition state analog was based on a mechanistic understanding of the hydrolytic deamination of adenosine. It was proposed that ADARs use a nucleophilic aromatic substitution (S_NAr) type mechanism, where a metal-bound hydroxyl attacks the 6-position of adenosine, leading to a high energy Meisenheimer complex in which aromaticity is briefly lost and then inosine is formed through the loss of ammonia (Figure 1.6 A).²⁶ A transition state analog, 8-azanebularine (8-AN) uses two features to trap the enzyme during this mechanism: First, the extra nitrogen in the 8-position is known to increase the electrophilicity of the 6 position, rendering it more susceptible to covalent hydration.²⁷ In addition, the replacement of the exocyclic amine with a hydrogen eliminates the presence of a leaving group at the 6-position. These features allow 8-AN to serve as a high-affinity substrate without productive turnover, thus the enzyme remains bound to its substrate (Figure 1.6 B).

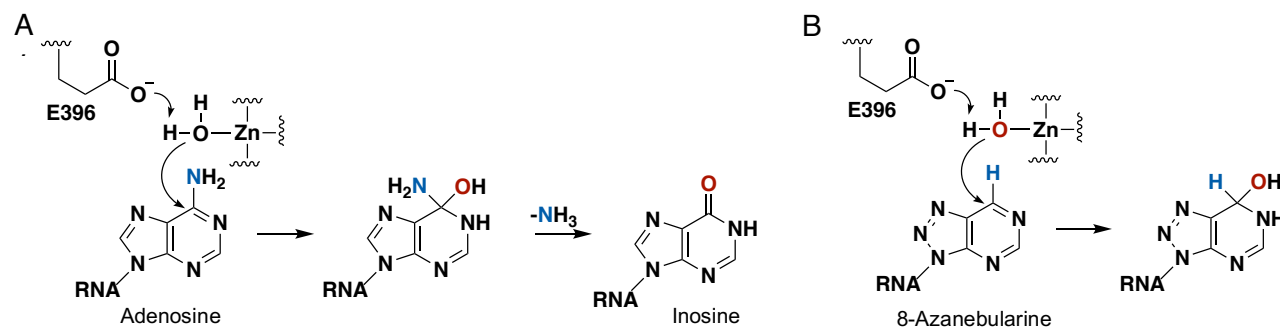


Figure 1.6. ADAR reaction mechanism. A) The ADAR reaction mechanism involves hydrolytic deamination of adenosine to inosine. B) 8-azanebularine serves as a transition state analog by mimicking the transition state of the hydrolytic deamination reaction.

The crystal structure of ADAR2d bound to dsRNA containing 8-AN allowed for observation of the contacts enabling this base-flipped conformation of the enzyme-RNA complex (Figure 1.7 A). When the edited base is flipped into the enzyme active site, residue 488 inserts into the duplex from the minor groove side to occupy the vacated space (Figure 1.7 B).^{12,25} It contacts the base that was previously paired against

the edited base, which is referred to as the orphan base. Residue 488 is part of the base-flipping loop of ADAR2 (aa 487 – 489) which includes the E488 helix-penetrating residue with neighboring glycines.^{23,25} This contact is reminiscent of the mechanism used by other nucleic acid modifying enzymes such as Hha I DNA methyltransferase (MTase) and DNA repair glycosylases (HOGG1, UDG, and AAG).^{28–32} However, as opposed to these B-form DNA-targeting enzymes which accommodate the intercalating residue by bending the duplex at the edited site, ADAR2 induces a conformational change in the RNA duplex across from the edited site (Figure 1.7 C).²⁵ This mechanism, which allows for the intercalating residue (E488) to contact the orphan base, gives rise to its specificity for A-form helices as opposed to B-form DNA. B-form helices have deeper grooves that would not be accessible without bending of the helix.²⁵ The nature of the contact between residue 488 in ADAR2 and the orphan base is discussed further in *Chapters 2 and 3*.

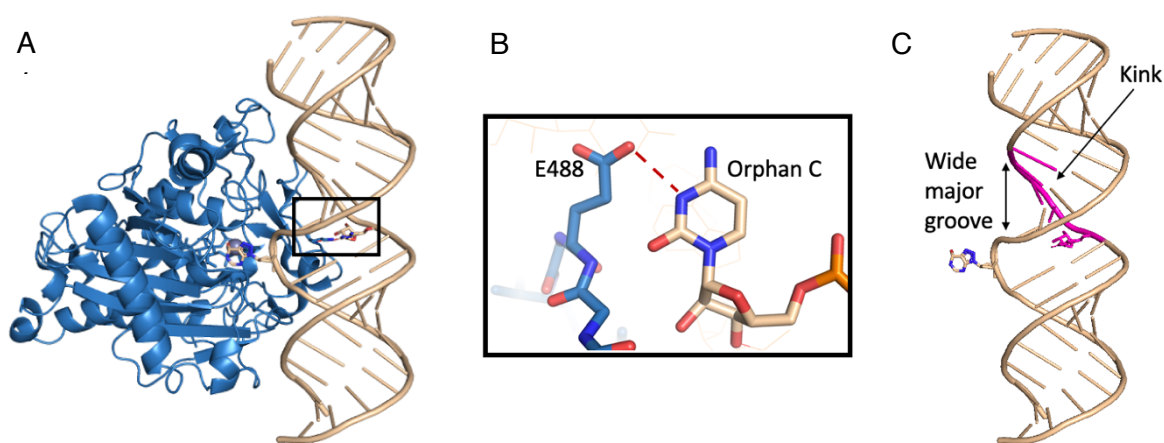


Figure 1.7. Structures of hADAR2d bound to dsRNA. A) The crystal structure of ADAR2d bound to dsRNA²⁵ illustrates the base-flipping mechanism. B) Residue 488 intercalates the duplex to contact the cytosine opposite the edited base. C) Conformational change in the RNA duplex induced by ADAR2d.

Other protein-RNA contacts in the structure of ADAR2d bound to dsRNA helped to rationalize known sequence preferences of ADARs.²⁵ RNA characteristics such as dsRNA length, and secondary structure—including mismatches, bulges, and loops—contribute to editing selectivity by ADARs.³³ ADAR1 and ADAR2 exhibit distinct yet overlapping substrate preferences, which primarily originates from the catalytic domain of either enzyme—in particular their divergent 5' binding loops.^{34,35} While ADARs target dsRNA of many different sequences, large datasets of endogenous editing sites have revealed that editing by ADAR1 and ADAR2 is influenced by the identity of the 3' base, and to a greater extent by the

identity of the 5' neighboring base.^{36,37} Both ADARs prefer to edit adenosines with a 5' neighboring U or A, and a 3' neighboring G (5'-UAG-3').^{25,36,37} ADAR1 and ADAR2 differ only slightly in their nearest neighbor preferences; ADAR1 prefers U=A>C>G in the 5' position versus U≈A>C=G for ADAR2, and in the 3' position ADAR1 shows no distinct preference while ADAR2 prefers U=G>C=A.³³ In ADAR2, the 2-amino group of the 3'-G accepts a hydrogen bond from S486 according to the structure, contributing to the observed selectivity.²⁵ Modelling a 5' G or C into the ADAR2d structure suggests that the 2-amino group of the guanosine in either pair may clash with the G489 in the flipping loop of the protein. In *Chapter 4*, base pair preferences in the 5' position are further explored.

While there is a high-resolution structure of ADAR2d bound to dsRNA, no structure of ADAR1 bound to its substrate has yet been published, leaving remaining questions about the basis for its substrate recognition and preferences. However, using biochemical data to inform the design, we generated a homology model of ADAR1 deaminase domain that helped to make predictions about the roles of specific residues. This work is detailed in *Chapter 5*.

Therapeutic RNA editing

Recent approvals of antisense, exon skipping, and siRNA therapeutics, along with the highly successful mRNA-based COVID-19 vaccines, have invigorated the field of nucleic acid therapeutics and stimulated interest in other mechanisms by which oligonucleotides might elicit therapeutic effects.³⁸ Oligonucleotides can be used to guide ADARs to correct disease-causing mutations, opening up the possibility of therapeutic RNA editing.³⁹⁻⁴²

In a therapeutic context, deamination of adenosine would allow for the recoding of mRNAs. This is significant because most human genetic variants associated with disease are point mutations, and a majority of these could be reversed by A-to-G transitions.^{43,44} While there are over 6,000 known disorders stemming from genetic origins, estimates are that only 10% of these diseases are currently treatable.² Therefore, since the advent of therapeutic gene editing and even more recently therapeutic transcriptome editing, there has been considerable interest in treating the mutations underlying these diseases. Since the

1990s, an array of gene editing tools has been developed to target defective genetic material at the genome level. This includes zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats) nucleases.^{45–48} Apart from the potential to repair mutations, recent reports detailing genomic editing of the pathogenic sickle cell hemoglobin allele to another naturally occurring variant has demonstrated the potential to reverse disease phenotypes without restoration of the original codon.⁴⁹ In fact, single transversion mutations can have implications beyond codon substitutions; to alter translation, skip or induce exons, alter epitranscriptomic modifications, and activate or inactivate enzymes.⁵⁰ This idea broadens the scope of potential targets for therapeutic RNA editing.

ADARs are uniquely suited for directed editing applications, as they only act on double stranded substrates. This provides the opportunity to design oligonucleotides that hybridize to a target sequence and thereby guide enzyme activity to a desired adenosine (Figure 1.8). In addition, the use of endogenous human enzymes circumvents potential issues that occur with nucleic acid editing technologies requiring the use of exogenous or bacterial-derived enzymes, such as delivery barriers or immune stimulation. One potential pitfall of programmable systems derived from prokaryotic origin is that immunogenicity may limit the ability to re-dose a patient. If the editing efficiency is not great enough with an initial dose, treatment options are limited.⁵¹ Directing edits at the transcript level has some advantages for clinical applications compared to genome editing. There is no reliance on the efficiency of homologous recombination (HR) to make directed edits as opposed to non-homologous end joining (NHEJ).⁵¹ However, the advent of DNA base editors—typically a deaminase fused to an inactive Cas protein—has allowed for the correction of genome-level mutations without dependence on HR.^{52–55} A particular advantage of RNA editing as opposed to DNA editing is that any potential off-target transcript editing does not result in a permanent change to the genome.

Aside from their function, delivery is an important consideration in devising therapeutic reagents. Systems may require efficient delivery of multiple components (both a guide oligonucleotide and enzyme) and must be specific to the target tissue.⁵¹ This issue is largely solved for delivery of RNA therapeutics to

the liver. The FDA approved siRNA therapeutic from Alnylam, givosiran, employs the N-acetylgalactosamine (GalNAc) modification which is known to improve oligonucleotide delivery to hepatocytes via uptake by the asialoglycoprotein receptor.^{56,57} However, for protein or ribonucleoprotein complexes (RNPs) delivery remains a prominent challenge. Lipid- or adeno-associated virus- (AAV) mediated delivery are some of the most common strategies in gene therapy.⁵¹ Cationic lipids trigger endocytosis, allowing for the delivery of nucleic acids or negatively charged proteins. Still, lipid-mediated delivery exhibits low transfection efficiency in most tissues, low serum stability, and varying levels of toxicity.⁵¹ Alternatively, the natural ability of viruses to inject genetic material into host cells has been engineered as a delivery vehicle in the case of AAVs. This allows for the delivery of ssDNA with low pathogenicity and immunogenicity.⁵⁸ The challenges are that cargo size is limited (4.7 kb), genome integration is possible, and patients may have pre-existing immunity or immunogenicity against these engineered viruses.^{51,58} Editing systems employing ADAR proteins have the advantage of using endogenous enzymes and a therapeutic RNA, to avoid the issue of enzyme delivery altogether.

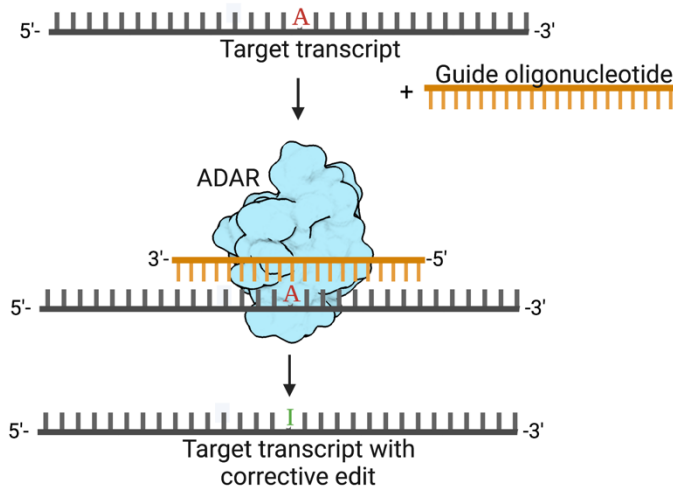


Figure 1.8. Hybridization of a guide oligonucleotide to a target transcript creating the secondary structure necessary for deamination by ADARs, to allow for a disease-corrective edit.

The use of ADARs as a programmable tool for RNA editing was first put forth when Woolf et al. used a 52 mer guide RNA to induce the correction of a premature stop codon in *Xenopus* embryos.^{51,59} Initial strategies for site-directed RNA editing via ADAR enzymes relied on enzyme overexpression or modification to achieve significant editing levels. Enzyme engineering has been employed to boost

efficiency of editing systems. This has primarily involved the generation of a tagged protein and modified RNA to provide high affinity between the enzyme and guide. These strategies have employed the SNAP tag, λ N- boxB, and MS2-MCP tagging systems—Each of which are naturally occurring high affinity interactions in human DNA repair, λ -phages, and MS2 bacteriophages respectively^{60–63} Recruitment of ADARs has also been facilitated via fusion of ADAR1 or ADAR2 deaminase domain to a catalytically inactive Cas13b (dCas13b) and a guide RNA linked to a CRISPR RNA (crRNA).⁶⁴ However, this system led to an abundance of off-target edits and did not appear dependent on the crRNA portion of the guide for its activity.^{61,64} Other systems focused on engineering of the guide RNA alone to recruit full length ADAR2. By fusing part of a natural substrate, the *GRIA2* hairpin, to a guide RNA that is complementary to their target they encouraged recruitment via binding of the dsRBDs.^{65,66} However, these approaches resulted in prominent off-target editing and still required exogenous enzyme delivery.

As an alternative, recent reports have shown success in employing endogenous ADAR enzymes for directed editing. These systems operate primarily by optimizing long guide oligonucleotides (> 100 nt) to induce efficient editing in cells.^{41,42,67} These systems only require use of a guide RNA and reduces the potential for immunogenicity. However, recruitment and editing efficiency of endogenous ADARs is a prominent challenge. The RESTORE (Recruiting Endogenous ADAR to Specific Transcripts for Oligonucleotide-mediated RNA Editing) approach uses chemically modified guide RNAs containing an antisense and a *GRIA2* hairpin motif—quite similar to previous studies.^{41,65,66} This strategy deviates from previous reports via the use of 2'-O-methyl, phosphorothioate, and locked nucleic acid modifications to bolster both stability and editing efficiency via endogenous ADARs (Figure 1.9).⁴¹ An alternative guide oligonucleotide design was seen in the LEAPER (Leveraging Endogenous ADAR for Programmable Editing of RNA) system which uses a long unmodified RNA that is complementary to the target and fully genetically encodable.⁶⁷ Both systems include a cytidine opposite the target site, which is well-known to enhance editing at the target site.³⁴ Complications observed in cellular studies have included limited editing efficacy and specificity, delivery efficiency, and siRNA-like effects due to the requirement for longer RNAs

to achieve desired editing levels.^{41,67} Recently, the recruitment of endogenously expressed ADARs with short, GalNAc-conjugated, chemically modified oligonucleotides in non-human primate (NHP) liver was described for the first time.⁶⁸ Thus, the stage has been set for translation of oligonucleotide-directed RNA editing technologies to the clinic. However, challenges remain for the development of guide oligonucleotides of shorter length that can induce efficient directed editing via endogenous ADARs. These targets must include editing sites within coding sequences, outside the liver, and with non-ideal nearest neighbor contexts for ADARs before the full scope and limitations of this approach can be known.

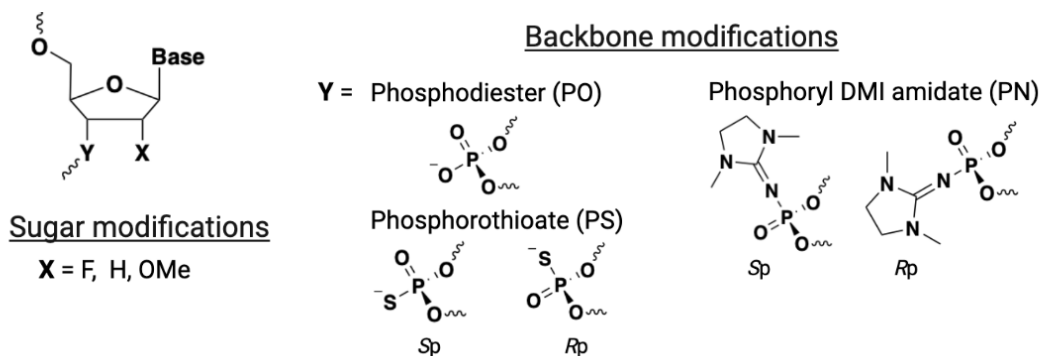


Figure 1.9. Chemical modifications used in guide strands for site-directed RNA editing include sugar and backbone modifications.

It is known that oligonucleotide chemical modifications can alter enzyme activity and provide guide strands with desirable therapeutic properties such as nuclease resistance and cell-specific localization.^{69–71} As described above, modifications have been used in guides to enhance ADAR activity and increase cell stability. This includes sugar modifications such as 2'-fluoro, 2'-deoxy, and 2'-O-methyl, as well as backbone modifications: phosphorothioate and phosphoryl DMI amidate linkages.^{41,68} Other design features can influence activity as well, such as the inclusion of an A-C mismatch at the target site to encourage editing, or A-G mismatches to prevent off-target editing in the region complementary to the guide.^{41,67,72,73} The rational design of guide strands to enhance therapeutic RNA editing via ADARs will be discussed further in *Chapters 3, 4, and 6*.

CHAPTER 2

Mechanistic Studies of pH-responsive Changes in RNA Editing

This work was part of a joint project with Dr. Turnee Malik in the Emeson laboratory at Vanderbilt University. Our observation of the pH dependence of ADAR activity in vitro and their simultaneous observation of changes in cellular editing upon culture medium acidification led to this highly collaborative effort. This chapter contains excerpts from the full manuscript which was published in Nucleic Acids Research in April 2021.⁷⁴

Introduction.

Dynamic and rapidly coordinated gene expression relies upon various post-transcriptional mechanisms that modify RNA sequence, structure, and stability.⁴ Prevalent among more than one hundred RNA processing events that shape the transcriptional landscape is the conversion of adenosine to inosine (A-to-I) by RNA editing.^{75,76} It has been predicted that the human transcriptome contains as many as one-hundred million A-to-I editing sites, comprising selectively edited adenosines in protein-coding regions as well as hyper-edited adenosine clusters in non-coding, repetitive sequences such as short interspersed nuclear elements (SINEs).^{75,77,78} A-to-I editing is generally identified as adenosine to guanosine (A-to-G) discrepancies during comparisons of genomic and cDNA sequences due to the base-pairing of cytidine to inosine (like guanosine) during reverse transcriptase-mediated first-strand cDNA synthesis. Many cellular machines also recognize inosine as guanosine, indicating that A-to-I editing constitutes functional A-to-G substitutions that can modulate diverse pathways involved in innate immunity, RNA splicing, RNA interference, and protein recoding.^{79,80}

The specificity and frequency of A-to-I editing are dictated by both *cis*- and *trans*-acting regulatory elements. *Cis*-acting factors such as RNA sequence context influence the extent of site-specific A-to-I conversion, but the formation of an extended region of dsRNA by intramolecular base-pairing is paramount for editing.^{33,81,82} The major *trans*-regulatory factors are the editing enzymes themselves, referred to as

adenosine deaminases acting on RNA (ADARs), which catalyze the deamination of adenosine residues within dsRNA substrates.⁸⁰ Two active members of the vertebrate ADAR family, ADAR1 and ADAR2, each contain multiple copies of a dsRNA binding domain (dsRBD) and a carboxyl-terminal adenosine deaminase domain.⁸³ The RNA editing reaction involves three main steps: 1) ADAR binding to the dsRNA substrate, 2) flipping the targeted adenosine out of the RNA duplex into the enzyme active site, and 3) hydrolytic deamination at position 6 of the purine ring.⁸⁴

Expression of ADAR1 and ADAR2 is ubiquitous but enriched in the brain along with inosine-containing mRNAs.^{85,86} Interestingly, many editing-dependent recoding events in mRNAs occur within transcripts critical for nervous system function. The extent of editing for these RNAs varies spatiotemporally and carries functional consequences for encoded proteins including changes in calcium permeability through GluA2 subunit-containing AMPA receptors, alterations in inactivation dynamics for the Kv1.1-subtype of voltage-gated potassium channel, and modulation of G-protein coupling efficacy and constitutive activity for the 2C-subtype of serotonin receptor (5HT_{2C}).^{78,87-89} Furthermore, alterations in ADAR1 or ADAR2 expression have been shown to result in neurobehavioral phenotypes, as well as embryonic or early postnatal lethality, in animal models.^{16,90-93} Dysregulation of RNA editing in humans has been implicated in disorders of innate immunity and nervous system function including Aicardi-Goutières syndrome, epilepsy, suicide, amyotrophic lateral sclerosis, and schizophrenia.⁹⁴⁻⁹⁸ Taken together, these data not only demonstrate an important role for A-to-I editing in numerous physiological systems, but also suggest that editing may be regulated to produce transcriptional plasticity that can endow biological systems with adaptive capacities in the face of changing environmental or physiologic conditions.^{99,100}

In mammals, several proteins regulate ADAR stability and subsequent RNA editing. The coordinate action of a positive regulator, Pin1, and a negative regulator, WWP2, modulate ADAR2 expression through post-translational interactions.¹⁰¹ AIMP2 also inhibits editing by decreasing ADAR protein levels.⁷⁸ Although these *trans*-acting regulators provide a mechanism for editing regulation by modulating ADAR

stability, studies spanning the past decade have consistently concluded that changes in ADAR protein expression do not fully account for differences in the extent of A-to-I conversion.^{78,101–106} Accordingly, it is likely that other factors modulate ADAR activity, rather than protein expression, to alter the extent of editing for ADAR substrates.

Consistent with this idea, it was observed that ADAR editing was enhanced in culture when the intracellular proton concentration was increased beyond that normally observed under control conditions.⁷⁴ When pH of the cell culture medium was adjusted to either pH 7.4 (control) or pH 6.7 (acidic) by manipulation of the bicarbonate concentration this acidification significantly increased ADAR1 and ADAR2-mediated editing of numerous sites within the examined transcripts (*5HT_{2C}*, *GluA2*, and *GLII*), although the magnitude of the effect was both site-dependent and dependent upon the specific ADAR acting upon it. Notably, sites preferentially recognized by a specific ADAR underwent robust increases in editing under acidic conditions (up to ~40%) when acted upon by that ADAR. The magnitude of observed changes in editing levels are comparable to the largest changes in editing that have been observed, such as in response to ADAR1 induction by interferon treatment.⁴¹

These observations led to our investigation of the mechanism of such enhancements in editing. Recent studies of the structural basis for ADAR base-flipping have revealed the importance of a highly conserved glutamate (E1008 in ADAR1 and E488 in ADAR2) residing in the deaminase domain of the enzyme.¹⁰⁷ This residue stabilizes the flipped-out conformation of the RNA duplex, presumably by occupying the space vacated by the flipped-out adenosine and hydrogen bonding with the complementary-strand orphaned base. Consistent with the idea that this glutamate requires protonation to stabilize the altered nucleic acid conformation, an enhancement in base-flipping and deamination rate is observed when this glutamate is mutated to a glutamine (E1008Q and E488Q in ADAR1 and ADAR2, respectively), which is fully protonated at physiologically relevant pH.^{84,108} This indicates that protonation of this glutamate may be critical for optimal catalytic activity and may alter the overall rate of ADAR catalysis during pH shifts

in the cell. Thus, we sought to examine the effect of acidification on *in vitro* deamination and base-flipping abilities of both ADAR1 and ADAR2.

Results.

The RNA editing reaction is intrinsically pH-sensitive

While increases in RNA editing produced by acidification in HEK293T cells could result from increased ADAR protein expression, such a mechanism cannot account for the acidification-induced increases in editing in HeLa cells. As such, the observed increases in RNA editing at reduced pH also could be explained by other molecular mechanisms including activation of various pH-regulated signaling pathways or the effects of pH on RNA structure and the intrinsic pH sensitivity of ADAR activity. To directly examine this last possibility, *in vitro* editing assays were performed to quantify deamination rate constants (k_{obs}) at varying pH (from pH 6.0 to pH 8.5) using purified, recombinant ADAR protein and an *in vitro* transcribed *5HT*_{2C} substrate. For most sites examined, an inverse correlation between deamination rate and pH between pH 6.5 and 8.5 was observed. For example, ADAR1 deamination of the A-site was most efficient at pH 6.5 (Figure 2.1 A). Similarly, ADAR2 deamination of the A-, B-, and C-sites also was most efficient at pH 6.5 and pH 7.0 (Figure 2.1 B). The rate of editing for these sites is about 1.5-fold less efficient at pH 7.5 than at either pH 7.0 or pH 6.5. To assess whether the observed differences in catalytic rate resulted from changes in pH-dependent stability of ADAR proteins, the melting temperature of recombinant ADAR2 protein was quantified using SYPRO Orange, a dye that increases in fluorescence intensity upon thermal denaturation of the protein.¹⁰⁹ Results from this ThermoFluor analysis revealed that ADAR2 was relatively stable from pH 6.5 to pH 8.5, with a melting temperature of ~53°C across this range (Figure 2.1 C). At pH 6.0 however, the melting temperature of the ADAR2 protein was significantly decreased, an instability that paralleled the observed reduction in catalytic rate for the *5HT*_{2C} editing sites (Figure 2.1 A-C).

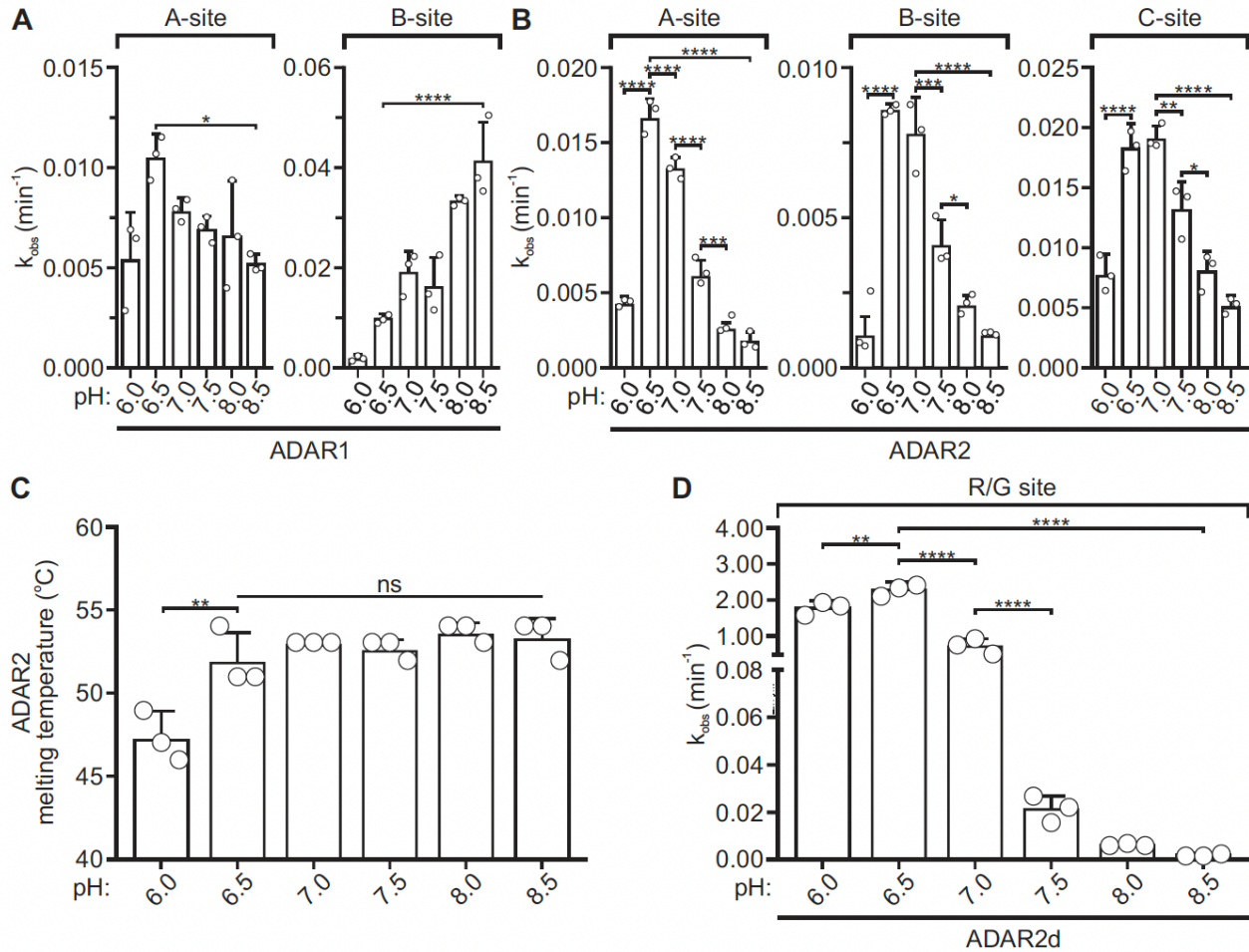


Figure 2.1. Effect of varying pH on in vitro A-to-I editing. (a) Rate of in vitro ADAR1- or (b) ADAR2-mediated *5HT2C* editing at half-pH intervals from 6.0 to 8.5 at 30 °C. Plotted values represent the means of three technical replicates (○) ± SD. Statistical significance between groups was determined using one-way ANOVA with Tukey's multiple comparisons test; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$. (c) Quantification of ADAR2 melting temperature across the pH range used for in vitro editing experiments. Plotted values represent the means of three technical replicates (○) ± SD. Statistical significance between groups was determined using one-way ANOVA with Tukey's multiple comparisons test; ** $p \leq 0.01$; ns, not significant. (d) Rate of in vitro ADAR2d-mediated *GLI1* R/G site editing. Plotted values represent the means of three technical replicates (○) ± SD. Statistical significance between groups was determined using one-way ANOVA with Tukey's multiple comparisons test; ** $p \leq 0.01$; **** $p \leq 0.0001$.

Taken together, these *in vitro* analyses demonstrate that the RNA editing reaction is intrinsically pH-sensitive, a property that may arise through changes in ADAR-substrate binding or catalysis at reduced pH, but not by changes in ADAR protein expression. To exclude the possibility that the observed increase in editing rates resulted from enhanced ADAR2-substrate binding via the double-stranded RNA binding domains (dsRBDs), the pH-sensitivity of the ADAR catalytic domain alone was assessed by taking advantage of the efficient editing of *GLI1* by the ADAR2 deaminase domain (ADAR2d) lacking dsRBDs.¹¹⁰

ADAR2d deamination of the *GLII* R/G site was 40- to 130-fold more efficient under acidic conditions than at pH 7.5, with $k_{\text{obs}} = 3.1 \text{ min}^{-1} \pm 0.2$ at pH 6.5 and $k_{\text{obs}} = 0.023 \text{ min}^{-1} \pm 0.006$ at pH 7.5 (Figure 2.1 D). These data indicate that RNA editing can be facilitated under acidic conditions independently of substrate interactions with the ADAR dsRBDs.

Protonation of a conserved glutamate residue in the ADAR base-flipping loop partially accounts for increases in RNA editing at acidic pH

A recent investigation of the structural basis for base-flipping by ADAR2 revealed the importance of a highly conserved glutamate, E488 (corresponding to E1008 in ADAR1), residing in the deaminase domain of the enzyme.¹⁰⁷ This residue stabilizes the flipped-out conformation of the RNA duplex, presumably by occupying the space vacated by the flipped-out adenosine and hydrogen bonding with the complementary-strand orphaned base (Figure 2.2 A,B). Mutant ADAR proteins bearing a glutamate-to-glutamine substitution at this residue (ADAR1 E1008Q and ADAR2 E488Q) exhibit increased catalytic activity via enhanced base-flipping^{84,101,102} As this glutamine is fully protonated under normal physiologic conditions at pH 7.4, these observations are consistent with the idea that the corresponding glutamate residue in wild-type ADARs requires protonation to support RNA stabilization during the base-flipping step in catalysis (Figure 2.2 A,B). To further examine how base-flipping is modulated by ADAR protonation, we compared base-flipping for ADAR2 and ADAR2 E488Q proteins as a function of pH using a 2-aminopurine (2-AP)-modified GluA2 transcript to measure 2-AP fluorescence intensity, which has been shown previously to correlate with base-flipping.^{84,111} The fluorescence intensity observed with the ADAR2 E488Q mutant was greater than that exhibited with the wild-type ADAR2 protein at each pH, confirming that ADAR2 E488Q has enhanced base-flipping abilities (Figure 2.2 C). However, a differential pH dependence between wild-type ADAR2 and the ADAR2 E488Q mutant enzyme was observed where the fluorescence intensity with the wild-type enzyme increased with decreasing pH, whereas fluorescence intensity with the ADAR2 E488Q mutant was maximal at pH 7.0, but dropped off significantly with increasing or decreasing pH (Figure 2.2 C). These data indicate that ADAR base-flipping is intrinsically

pH-dependent and that RNA editing by the ADAR E488Q mutant enzyme is less affected by acidification than the wild-type ADAR2 protein.

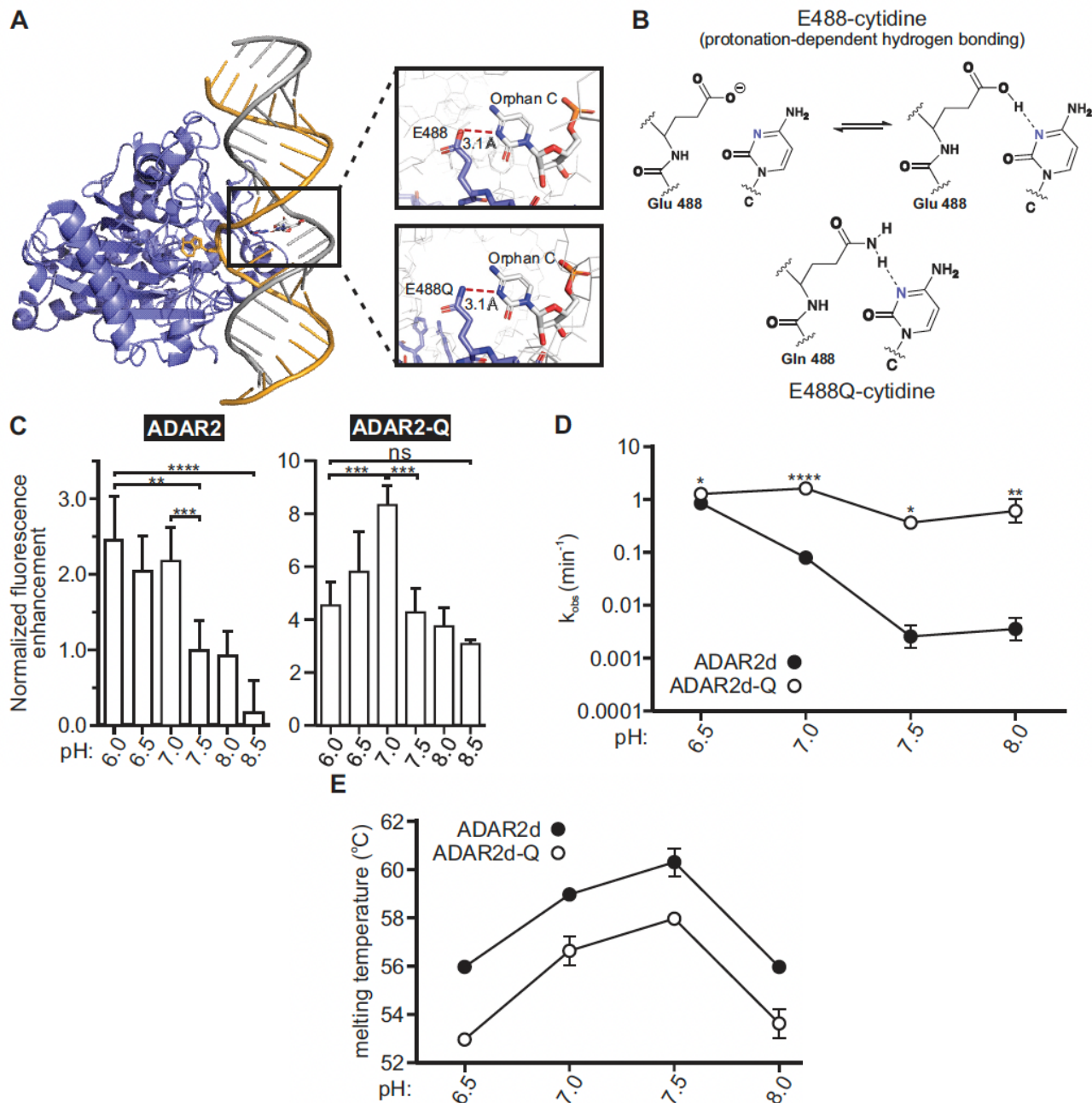


Figure 2.2. Effects of a glutamate-to-glutamine substitution on the pH-dependency of ADAR base-flipping and editing. (a) The crystal structure of ADAR2d bound to dsRNA (PDB ID: 5HP3 & 5ED1) shows the base-flipped conformation stabilized by contacts between residue 488 and the orphan base. (b) An illustration of the hydrogen bonding contact between ADAR2 and the orphan base showing protonation-dependent hydrogen bonding for wild-type ADAR2. (c) Normalized fluorescence enhancement from a dsRNA substrate containing 2-aminopurine in the edited position, corresponding to base-flipping by the ADAR2 enzyme. Plotted values represent the means of three

technical replicates \pm SD. Statistical significance between groups was determined using one-way ANOVA with Tukey's multiple comparisons test; ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$; ns, not significant. (d) Rates of *in vitro* ADAR2d- and ADAR2d E488Q mediated *GLII* (+23 site) editing. Plotted values represent the means of three technical replicates (\circ) \pm SD. Statistical significance between groups was determined using the Holm-Sidak t-test for multiple comparisons; * $p \leq 0.05$; ** $p \leq 0.01$; **** $p \leq 0.0001$. (e) Quantification of ADAR2d and ADAR2d E488Q melting temperatures across the pH range used for *in vitro* editing experiments. Plotted values represent the means of three technical replicates (\circ) \pm SD.

To further compare the relative pH sensitivity of wild-type and mutant (ADAR E488Q) proteins, we measured the rate of *GLII* deamination (+23 site) by ADAR2d and ADAR2d E488Q enzymes from pH 6.5-8.0. Interestingly, the deamination rates for the wild-type and mutant deaminase domains are similar at pH 6.5, with $k_{\text{obs}} = 0.91 \text{ min}^{-1} \pm 0.2$ and $k_{\text{obs}} = 1.3 \text{ min}^{-1} \pm 0.09$, respectively. However, the efficiency of wild-type ADAR2d deamination decreased over 450-fold with increasing pH, while the efficiency of ADAR2d E488Q deamination only decreased between 2.5- and 5-fold with increasing pH (Figure 2.2 D). Moreover, ADAR2d E488Q deaminated *GLII* more efficiently than ADAR2d at each pH despite decreased stability of the mutant protein across the entire pH range (Figure 2.2 E). These results show that the pH-dependence of deamination is much greater for wild-type ADAR2d than for ADAR2d E488Q. Taken together, these data suggest that protonation of ADAR1 at E1008 or ADAR2 at E488 partially accounts for the increases in RNA editing observed at acidic pH.

Discussion.

The mechanisms underlying changes in RNA editing profiles in response to physiologic signals are not well-defined. Most of the known *trans*-acting regulators of editing modulate ADAR protein levels, yet steady-state ADAR protein expression fails to fully account for the observed spatiotemporal variations in A-to-I conversion.^{78,101–106} Though dynamic regulation of ADAR activity—rather than ADAR expression—could explain significant changes in RNA editing in a cellular context, the mechanisms regulating such activity remain elusive. Previous structural and biochemical characterization of ADARs and ADAR mutants has suggested pH-sensitive deamination of these enzymes^{84,107,108} Therefore, our studies focused on how changes in pH regulate RNA editing. Analysis of *in vitro* biochemical systems revealed significant increases in ADAR1 and ADAR2-mediated editing under acidic conditions relative to editing at a

physiologic pH of ~ 7.4 (Figure 2.1). While acidification could alter the rate of editing by affecting ADAR binding and/or catalytic efficiency, pH-sensitive editing of *GLII* by ADAR2d (lacking the dsRBDs) showed that increased ADAR activity did not result solely from increased dsRBD-mediated binding under acidic conditions (Figure 2.1 D and Figure 2.2 D). Rather, our data show that the ADAR deaminase domain—and specifically the base-flipping loop—is fundamental to the intrinsic pH-sensitivity of ADAR catalysis (Figure 2.1 D and Figure 2.2).

During base-flipping, a highly conserved glutamate residue within the ADAR base-flipping loop invades the vacated space and hydrogen bonds with the base opposite the flipped-out adenosine to stabilize the strained nucleic acid conformation (Figure 2.2 A,B).¹⁰⁷ Consistent with the hypothesis that this interaction may be sensitive to changes in proton concentration, mutation of this critical glutamate to a glutamine, which is fully protonated at pH 7.4, results in increased catalytic efficiency via enhanced base-flipping of the ADAR glutamine mutants.^{84,108} Our analysis suggests that the ADAR2 reaction is accelerated by low pH regardless of the identity of the orphan base. Protonation of E488 enables this residue to donate a hydrogen-bond to contact an orphan cytidine (Figure 2.2 A,B). In addition, protonation of this residue neutralizes the negative charge on the side chain, decreasing charge repulsion with RNA during the flipping step for either adenosine-uridine base pairs or adenosine-cytidine mismatches (Figure 2.2 B). Our studies further show that base-flipping and editing activity for wild-type ADARs increase with decreasing pH, yet the pH-dependency of these properties is diminished for ADAR glutamine mutants (Figure 2.2 C-E). While the pK_a of ADAR1 E1008 and ADAR2 E488 have not been determined, our data suggest that protonation of these single amino acid residues within the ADARs enhances base-flipping to increase RNA editing at acidic pH and is the primary determinant that accounts for the pH-dependence of deamination (Figure 2.2 C-E). Though our studies using ADAR glutamine mutants also suggest that these enzymes are resistant but not completely insensitive to reductions in pH, the residual pH-sensitivity exhibited by ADAR glutamine mutants *in vitro* may be attributed to protonation of other amino acid residues involved in other steps of the

editing reaction (e.g. RNA binding, deprotonation of the reactive water molecule, etc.), as well as pH-dependent increases in the thermodynamic stability of targeted RNA duplexes.

The pH optima for most enzymes coincide with the pH of the subcellular compartments in which they reside.^{112,113} For example, the pH optimum of many proteases involved in prohormone processing coincides with an intragranular pH of ~5.5, whereas the pH optimum of many cytoplasmic enzymes is ~7.4.^{114,115} Surprisingly, ADARs function optimally between pH 6.5-7.0, well below the normal pH of the nucleus or cytoplasm.¹¹⁶ For many *5HT*_{2C} sites, large stepwise changes in editing were observed between pH 7.5 and 7.0 *in vitro* (Figure 2.1), suggesting that ADARs could serve as pH sensors to modulate RNA editing patterns in response to physiologically-driven pH shifts.

Cytoplasmic acidification often results from the accumulation of acid equivalents produced as metabolic byproducts.¹¹⁶ Although cells engage various regulatory mechanisms to maintain pH homeostasis, these systems can be overwhelmed during periods of unusually high metabolic activity, leading to intracellular acidification. Previous studies have reported increases in editing under hypoxic conditions but lacked evidence to suggest that the observed increases in RNA editing are driven by parallel increases in ADAR protein levels or through modulation of other hypoxia-sensitive cellular pathways.¹¹⁷⁻¹¹⁹ The present studies support the hypothesis that metabolic stress during hypoxia triggers intracellular acidification, which in turn enhances base-flipping activity to increase the overall ADAR catalytic rate and editing (Figure 2.3). Similar pH-dependent regulatory mechanisms may exist in conditions such as inflammation and epilepsy, which also are associated with both intracellular acidification as well as increased RNA editing^{18,94,120-125}

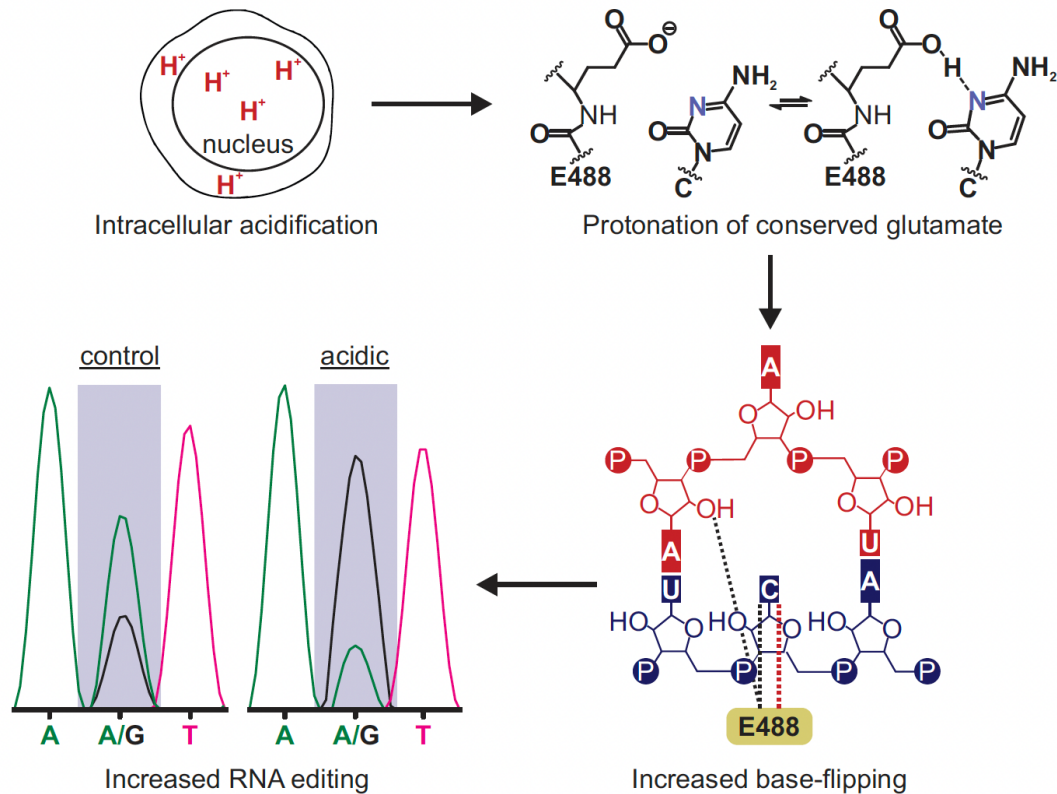


Figure 2.3 Intracellular acidification leads to protonation of the conserved glutamate (E488 in ADAR2), which increases base-flipping and leads to enhancements in RNA editing.

During hypoxia/ischemia, epilepsy, MELAS (Mitochondrial Encephalopathy, Lactic Acidosis, and Stroke-like episodes), and other pathophysiologic conditions in which pH homeostasis is disrupted, electrically-active cells such as neurons and cardiomyocytes experience broad disturbances in ion dynamics, often resulting in increased intracellular Ca^{2+} concentration to induce cytotoxicity and membrane hyperexcitability.¹¹⁶ Though the present studies of pH-dependent increases in RNA editing were limited to several model transcripts, it is likely that editing of many sites in the transcriptome increases upon intracellular acidification since editing is favored under such conditions. While it is unknown how such global increases in editing might influence overall physiology, studies using *Drosophila* model systems have shown that ADAR overexpression or knockdown results in decreased or increased neuronal excitability, respectively.¹²⁶ Since many editing-dependent recoding events affect the function of proteins involved in membrane excitability, it is intriguing to speculate that increased editing of various RNA targets

may serve as a pH-dependent homeostatic mechanism to limit membrane hyperexcitability and protect against excitotoxic damage.^{126,127}

The observation that acidic pH enhances base-flipping and increases the rate of deamination by ADARs has implications beyond our understanding of mechanisms of natural regulation of A-to-I editing. Several recent reports have described efforts to direct ADAR reactions for therapeutic benefit.^{64,128} One approach uses a guide RNA to form a duplex at a target site that can recruit endogenous ADAR enzymes for deamination of a specific adenosine.^{41,67} Thus, this directed RNA editing approach can “repair” G-to-A mutations associated with genetic disease. Optimization of ADAR guide RNAs requires a comprehensive understanding of factors that control RNA editing efficiency. The results described here stimulated our efforts to develop modifications to guide RNAs that mimic the effects of low pH. This is discussed in detail in Chapter 3.

Methods.

General biochemical procedures

Unless otherwise stated, all reagents were purchased from commercial sources (Fisher Scientific or Sigma Aldrich) and were used without further purification. Reagents for in vitro transcription, in vitro editing, and PCR amplification were purchased from: Promega: Access RT-PCR kit, RQ1 DNase free RNase; Qiagen: Gel Extraction kit; Zymo Research: DNA Clean & Concentrator kit; Syd Labs: Spin columns for PCR product clean up; New England BioLabs: Molecular-biology-grade bovine serum albumin (BSA), and RNase inhibitor. SDS-polyacrylamide gels were visualized with a Molecular Dynamics 9400 Typhoon phosphorimager. Data were analyzed with Molecular Dynamics ImageQuant 5.2 software. All MALDI analyses were performed at the University of California, Davis Mass Spectrometry Facilities using a Bruker ultrafleXtreme MALDI TOF/TOF mass spectrometer. Oligonucleotide masses were determined with Mongo Oligo Calculator v2.08. Unless otherwise noted, oligonucleotides were purchased from either Dharmacon or Integrated DNA Technologies.

Purification of oligonucleotides

All oligonucleotides for biochemical experiments were purified by denaturing polyacrylamide gel electrophoresis (PAGE) and visualized by UV shadowing. Oligonucleotides extracted from the gel using the crush and soak method for 16 h at 4 °C into 0.5 M NH₄OAc and 1 mM EDTA. Polyacrylamide fragments were removed using a 0.2 μm pore size cellulose acetate membrane filter (Corning). Oligonucleotides were precipitated from 75% ethanol containing 75 mM NaOAc at -70°C for 2 hours. The resulting pellet was dried under vacuum and resuspended in nuclease free water.

In vitro transcription and preparation of editing target RNA

Target RNA was transcribed from a DNA template with the MEGAScript T7 Kit (ThermoFisher). DNA digestion was performed using RQ1 RNase-free DNase (Promega). The DNase-treated RNA product was purified by 4% PAGE as described above. Purified *5HT2c* target RNA (180nM) was added to 1X TE buffer and 100 mM NaCl, heated to 95 °C for 5 min, and slowly cooled to room temperature.

Protein overexpression and purification of ADAR2 constructs

Human ADAR2 deaminase domain (ADAR2d), human ADAR2d-E488Q, wild-type human ADAR2, and human ADAR2-E488Q were expressed and purified as previously described.¹²⁹ Purification was carried out by lysing cells in buffer containing 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM 2-mercaptoethanol, 750 mM NaCl, 35 mM imidazole, and 0.01% Nonidet P-40 using a French press. Cell lysate was clarified by centrifugation (39,000 x g for 1 h). Lysate was passed over a 3 mL Ni-NTA column, which was then washed in 3 steps with 20 mL lysis buffer, wash I buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM 2-mercaptoethanol, 750 mM NaCl, 35 mM imidazole, 0.01% Nonidet P-40), wash II buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1mM 2-mercaptoethanol, 35 mM imidazole, 500 mM NaCl), and eluted with 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM 2-mercaptoethanol, 400 mM imidazole, 100 mM NaCl. Fractions containing the target protein were pooled and concentrated to 30-80 μM for use in biochemical assays. Purified wild-type ADAR2 was stored in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 20%

glycerol and 1 mM 2-mercaptoethanol at -70°C. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels

Protein overexpression and purification of ADAR1 p110.

MBP-tagged human ADAR1 p110 construct was cloned into a pSc vector using standard PCR techniques. The generated construct (yeast codon optimized) consisted of an N-terminal MBP-tag, a tobacco etch virus (TEV) protease cleavage site followed by the human ADAR1 p110 gene. The construct was transformed in *S. cerevisiae* BCY123 cells and overexpressed as described previously.¹²⁹ Purification was carried out by lysing cells in lysis/binding buffer containing 50 mM Tris-HCl, pH 8.0, 5% glycerol, 5 mM 2-mercaptoethanol, 1000 mM KCl, 0.05% NP-40 and 50 μ M ZnCl₂ using a microfluidizer. Cell lysate was clarified by centrifugation (39,000 x g for 50 min). Lysate was passed over a 2 mL NEB amylose column (pre-equilibrated with binding buffer), which was then washed in 2 steps with 50 mL binding buffer followed by 100 mL wash buffer (50 mM Tris-HCl, pH 8.0, 5% glycerol, 5 mM 2-mercaptoethanol, 500 mM KCl, 0.01% NP-40 and 50 μ M ZnCl₂) and eluted with buffer containing 50 mM Tris-HCl, pH 8.0, 10% glycerol, 5 mM 2-mercaptoethanol, 500 mM KCl, 0.01% NP-40, 50 μ M ZnCl₂, and 20 mM maltose. Fractions containing the target protein were pooled and dialyzed against a storage buffer containing 50 mM Tris-HCl, pH 8.0, 400 mM KCl, 0.5 mM EDTA, 0.01% NP-40, 10% glycerol and 1 mM tris(2-carboxyethyl)phosphine. Dialyzed protein was concentrated to 2-50 μ M and stored as aliquots at -70 °C until further use in biochemical assays. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels.

Deamination assay with ADAR2d, ADAR2d-E488Q, ADAR2, and ADAR1 p110

Deamination assays were performed under single-turnover conditions in 15 mM Tris-HCl (pH 7.0 to 8.5) or 15 mM Bis-Tris-HCl (pH 6.0 to 7.0), 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, 3 mM MgCl₂, 160 U/mL RNasin, 1.0 μ g/mL yeast tRNA, 10 nM RNA, and nM ADAR2d, ADAR2d-E488Q, or wild-type ADAR2. Each reaction solution was incubated at 30 °C for 30 min before

the addition of enzyme. Reactions were then incubated at 30 °C for varying times prior to quenching with 190 µL 95°C water and heating at 95 °C for 5 min. Reaction products were used to generate cDNA using RT-PCR (Promega Access RT-PCR System). DNA was purified using a DNA Clean & Concentrator kit (Zymo) and subjected to Sanger Sequencing using an ABI Prism 3730 Genetic Analyzer at the UC Davis DNA Sequencing Facility. The sequencing peak heights were quantified in 4Peaks v1.8. Data were fit to the equation $[P]_t = 0.9[1 - e^{-(k_{obs} * t)}]$ for ADAR2 and $[P]_t = 0.6[1 - e^{-(k_{obs} * t)}]$ for ADAR1 p110 where $[P]_t$ is percent edited at time t, and k_{obs} is the observed rate constant. Each experiment was carried out in triplicate where the k_{obs} reported is the average of each replicate \pm standard deviation (SD). Statistical significance between groups was determined by one-way ANOVA using Prism software (GraphPad). For the ADAR1 p110 enzyme, deamination reactions were performed as above with the following modifications: The final reaction solution for ADAR1 p110 contained 15 mM Tris-HCl (pH 7.0 to 8.5) or 15 mM Bis-Tris-HCl (pH 6.0 to 7.0), 4% glycerol, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% Nonidet P-40, 160 U/mL RNasin, 1.0 µg/mL yeast tRNA, and 10 nM RNA, and 150 nM ADAR1 p110.

ThermoFluor melting temperature analysis of recombinant ADAR2, ADAR2d, or ADAR2d-E488Q protein

Spectra were obtained using a Bio-Rad CFX Connect Real-Time PCR Detection System. Solutions contained 15 mM Tris-HCl (pH 7.0 to 8.5) or 15 mM Bis-Tris-HCl (pH 6.0 to 7.0), 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, 3 mM MgCl₂, and 2X SYPRO orange dye, with or without protein. Samples containing protein included 3 µM wild-type ADAR2, or 4 µM of either deaminase domain protein (ADAR2d or ADAR2d E488Q). To a 96-well clear bottom PCR plate was added 20 µL of each solution. Wells were sealed with PCR plate sealing film. Fluorescence was measured as the solutions were heated from 5 °C to 90 °C at a rate of 2 °C per minute. The derivative of fluorescence signal as a function of temperature ($-dF/dT$) was exported, and the background values of the buffered solution without protein was subtracted from each sample. Melting temperature was determined as the temperature where the derivative of fluorescence signal was at a minimum. Measurements were performed in triplicate. Melting

temperature values reported are the average of each replicate \pm standard deviation (SD). Statistical significance between groups was determined by one-way ANOVA using Prism software (Graph Pad).

Preparation of duplex substrates for base-flipping analysis

Oligonucleotides previously described for use in ADAR2 base-flipping analyses were purchased from Dharmacon.⁸⁴ RNAs were purified by 18% PAGE as previously described. PAGE purified top and bottom strands were annealed for a final concentration of 30 μ M edited strand, 45 μ M guide strand, 30 mM Tris-HCl, 6% glycerol, 120 mM KCl, 3 mM EDTA, 0.006% NP-40, and 0.6mM DTT. The mixture was heated to 95°C for 5 min, and slowly cooled to room temperature.

Base-flipping assay using a fluorescent RNA substrate.

Fluorescence measurements were performed using a CLARIOstar microplate reader and a Nunc MaxiSorp 384-well black bottom plate. Excitation was at 320 nm and fluorescence emission was scanned from 340 to 430 nm with 0.2 nm resolution. Spectra were obtained for solutions containing 2.5 μ M RNA, with or without 10 μ M ADAR2, in 36 mM Tris-HCl (pH 7.0 to 8.5) or Bis-Tris-HCl (pH 6.0 to 7.0), 7% glycerol, 142 mM KCl, 3.6 mM EDTA, 0.007% NP-40, and 0.7 mM DTT at room temperature.¹³⁰ The background fluorescence of the enzyme buffered at each pH was subtracted from the spectrum of the complex, and the background fluorescence of the buffer alone at each pH was subtracted from the RNA. Each spectrum is an average of three independent measurements that were LOWESS fit using Graphpad Prism software. The fluorescence intensity values at λ_{max} were used to determine the fluorescence enhancement by ADAR in the formula $FE = (FI_{ADAR-RNA} - FI_{RNA})/FI_{RNA}$ where FE is fluorescence enhancement, and FI values are the fluorescence intensity of samples containing either RNA or RNA in the presence of ADAR2. Fluorescence enhancement values were normalized so that the value of $FI_{ADAR-RNA}$ at pH 7.5 corresponds to 1.

CHAPTER 3

Chemically Modified Guide Oligonucleotides for RNA Editing via Endogenous ADARs

Reprinted with permission from the J. Am. Chem. Soc. 2021, 143 (18), 6865–6876. Copyright 2021 American Chemical Society.¹³¹

Introduction.

Enzymes that can be programmed to site-specifically alter nucleic acids offer a promising platform for the correction of disease-causing mutations.^{132–134} Adenosine Deaminases Acting on RNA (ADARs) have been used to deaminate target adenosines to inosine in double stranded RNAs.¹³⁵ Inosine is recognized as guanosine during translation, meaning that the ADAR reaction has therapeutic potential for the correction of pathogenic G-to-A point mutations. Commonly, RNA guided enzymes are localized to a target sequence by nucleic acid hybridization.^{45,136–138} Since ADARs only act on double stranded RNA, they can be directed to target adenosines using a guide oligonucleotide. This technology boasts the same programmable nature as Cas-mediated DNA editing, but holds some possible advantages for clinical applications. The effects of RNA editing are transient, meaning that any off-target edits do not result in a permanent change to the genome. On the other hand, unlike DNA editing, RNA editing therapeutics would require continuous administration.⁶⁷ However, RNA editing provides a reversible and tunable solution. In addition, base-editing by ADARs does not induce double strand breaks, and is not restricted to a particular phase of the cell cycle.¹³⁴ Also, immunogenicity of bacterial-derived Cas enzymes and delivery vectors has led to concerns over the safety and efficacy of their use in gene therapy.^{139–141}

These delivery barriers and immune stimulation issues have sparked interest in systems that employ endogenous human enzymes, including ADARs, to edit disease-relevant nucleic acids.^{41,42,67} Until recently, the primary approach to site-directed RNA editing by ADARs was reliant on the overexpression of engineered enzymes and co-delivery of a guide RNA.^{62,64,142,143} However, ectopic expression of ADARs leads to increased levels of off-target editing,¹³ and has been shown to induce toxicity in mice.¹⁴ Additional strategies have been reported that establish the use of endogenous ADARs for RNA editing.^{41,42,67} However, such strategies have required long guide RNAs, and still offer room for improvement in efficiency and

specificity. Here, we describe efforts to design nucleobase modifications for short guide oligonucleotides to increase A-to-I editing efficiency at target sites using endogenous ADARs.

Results.

Orphan base analogs that mimic a hyperactive ADAR2 mutant.

When ADARs edit target adenosines, the edited base is flipped into the enzyme active site leaving behind an “orphan base.”¹⁰⁷ In the preferred pairing context for ADARs, an A-C mismatch, this creates an unpaired cytidine.^{72,144} Structures of ADAR2 deaminase domain bound to RNA revealed that glutamic acid residue 488, in the highly conserved flipping loop of the enzyme, occupies the space vacated by the reacting adenosine.¹⁰⁷ The enzyme stabilizes this base-flipped conformation through a suspected hydrogen bond between the protonated side chain of E488 and N3 of the orphan cytidine (Figure 3.1 A; Figure 3.2 B).¹⁰⁷

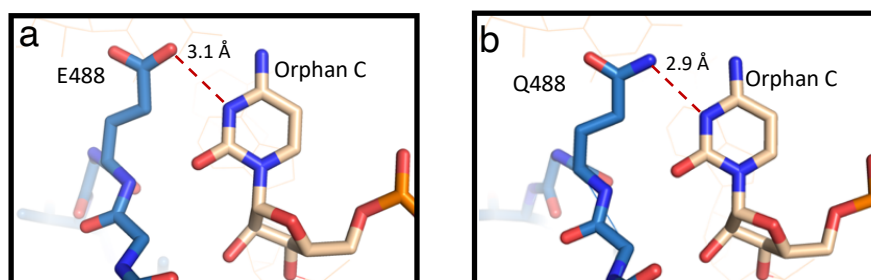


Figure 3.1. (a) The intercalating E488 residue of the ADAR2 flipping loop is engaged in a suspected hydrogen-bonding interaction with N3 of the orphan cytidine (PDB ID: 5HP3). (b) The structure of the mutant enzyme ADAR2 Q488 (PDB ID: 5ED1) also shows a hydrogen bond with N3 of the orphan cytidine.

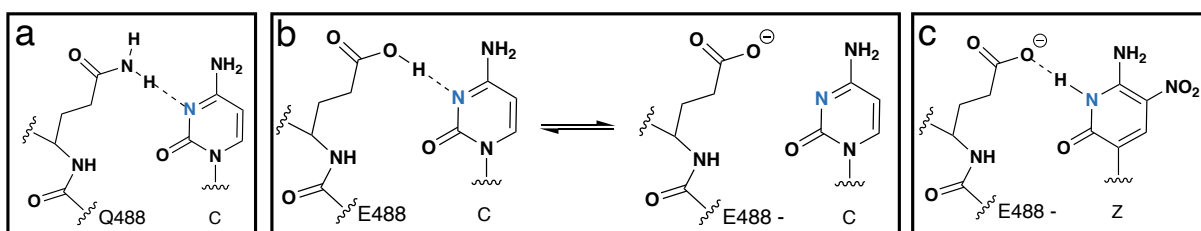


Figure 3.2. (a) The protonation-independent hydrogen bond between ADAR2 Q488 and the orphan cytidine. (b) The hydrogen bond between E488 and the orphan cytidine requires protonation to occur. (c) Benner's base Z in the orphan position creates a protonation-independent hydrogen bond with wild-type ADAR2.

Previously, it had been shown that a single mutation at residue 488 from glutamic acid to glutamine leads to enzyme hyperactivity, with up to a 60-fold increase in deamination rate constant for some substrates.⁸⁴ The structures of the deaminase domain of both wild-type and hyperactive ADAR2 bound to dsRNA have been solved, and reveal that the positioning of residue 488 is nearly superimposable between

the two enzymes (Figure 3.1 A,B).¹⁰⁷ It appears that ADAR2 E488Q makes the same hydrogen bonding contact between the side chain and N3 of the orphan cytidine. Importantly, the carboxamide of the glutamine side chain can function as a hydrogen bond donor independent of pH in the physiologically relevant range. This is thought to be the reason for the hyperactivity of ADAR2 E488Q (Figure 3.2 A).^{84,107} Indeed, our studies on the pH effects on the deamination rates for wild-type and E488Q mutant ADAR2 support this hypothesis, as discussed in Chapter 2.⁷⁴

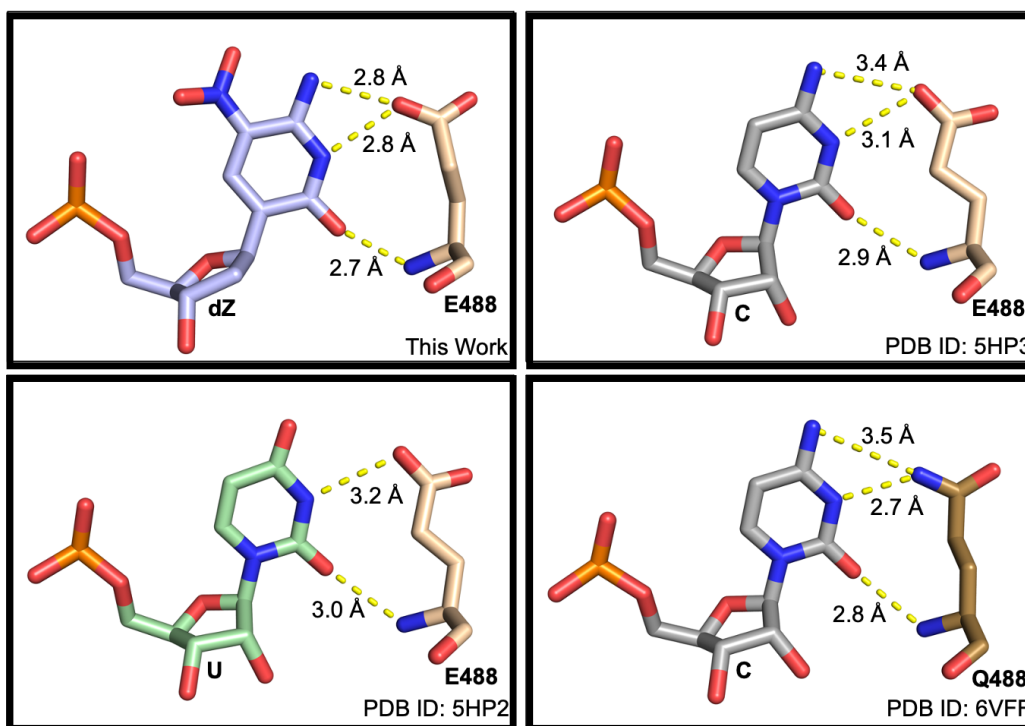


Figure 3.3. Comparison of interactions between the hADAR2 flipping residue 488 and the orphan base. (a) residue E488 and orphan base 2'-deoxy Benner's base Z (dZ) (PDB ID: 7KFN), (b) residue E488 and orphan cytidine (C) (PDB ID: 5HP3), (c) residue E488 and orphan base uridine (U) (PDB ID: 5HP2), and (d) residue Q488 and orphan base cytidine (C) (PDB ID: 6VFF).

These observations inspired us to test orphan base analogs capable of hydrogen bonding with the wild-type E488 residue in a protonation-independent manner. By stabilizing the same contact made by the mutant enzyme, this may allow wild-type ADAR2 to elicit a similar hyperactive effect. Chemically modifying the guide RNA to enhance this base-flipping contact, instead of using an enzyme mutation, could enhance substrate reactivity toward endogenous ADAR enzymes. To test this idea, we incorporated cytidine analogs with hydrogen bond donors at N3 into the orphan position of guide RNAs. The orphan base analogs

8-oxo-2'-deoxyadenosine (8-oxodA), 2'-deoxypseudoisocytidine (dpiC),¹⁴⁵⁻¹⁴⁷ and 6-amino-5-nitro-3-(1'- β -D-2'-deoxyribofuranosyl)-2(1H)-pyridone (referred to as 2'-deoxy Benner's base Z or dZ)¹⁴⁸ were chosen due to their hydrogen bond *donor-donor-acceptor* patterns (Table 3.1); retaining structural features involved in ADAR2-RNA contacts, while altering the hydrogen bonding capability at N3 (Figure 3.2 C, 3.3).¹⁰⁷ The use of 2'-deoxynucleotides in these studies was due to their ready availability and metabolic stability in cell-based directed editing assays compared to their ribonucleotide counterparts. The use of 2'-deoxy, 2'-O-methyl and phosphorothioate modifications in therapeutic oligonucleotides is a common strategy to increase cellular stability.^{69,70}

A fluorescence-based assay shows guide RNA-induced changes in base-flipping.

To study the effect of these orphan base analogs on base-flipping by ADAR2, we used a previously described assay that monitors changes in the fluorescence of a 2-aminopurine (2-AP)-labelled RNA.^{84,111} 2-AP is highly fluorescent as part of a single-stranded oligonucleotide, but its fluorescence is quenched when it is part of a duplex structure. This has made 2-AP useful as a probe for base-flipping by nucleic acid modifying enzymes, including ADAR2.^{111,149,150} By incorporating 2-AP at the edited position of a target RNA, fluorescence intensity (FI) can be monitored allowing for quantification of base-flipping by ADAR2.¹¹¹ Previously, this assay has been used to examine the effect of ADAR2 mutations on base-flipping, where it was shown that ADAR2 E488Q has a 2.1-fold increase in FI compared to ADAR2 WT.⁸⁴ In binding assays, the affinity of ADAR2 WT and E488Q for target RNA was shown to be similar, suggesting that the increase in deamination rate was attributed to the enhanced base-flipping ability of the enzyme.⁸⁴ In this study, instead of comparing the effect of ADAR2 mutations, we examined the effect of guide strand modifications on base-flipping.

We incorporated 2-AP into the R/G editing site of a 28-nt RNA derived from the *GRIA2* pre-mRNA, a well-studied ADAR2 substrate.^{84,111} The 2-AP containing target RNA was paired with 28-nt orphan base-modified guide RNAs, where the editing site nucleotide was paired across either dC, dZ, or dpiC (Figure 3.4 A). FI measurements were taken of the RNA substrate in the absence of protein, and the

change in fluorescence upon the addition of ADAR2 was examined. The enzyme-dependent change in fluorescence observed with different guide RNAs indicated that base-flipping was preferentially enhanced when ADAR2 wild-type (WT) was paired with an RNA containing dZ in the orphan position (Figure 3.4). The normalized fluorescence enhancement increased 3.9-fold when ADAR2 WT was paired with dZ containing RNA as opposed to the unmodified orphan base dC (Figure 3.4 B). However, 2'-deoxypseudoisocytidine (dpiC) failed to show an enhancement in base-flipping compared to dC (Figure 3.4 C,D)

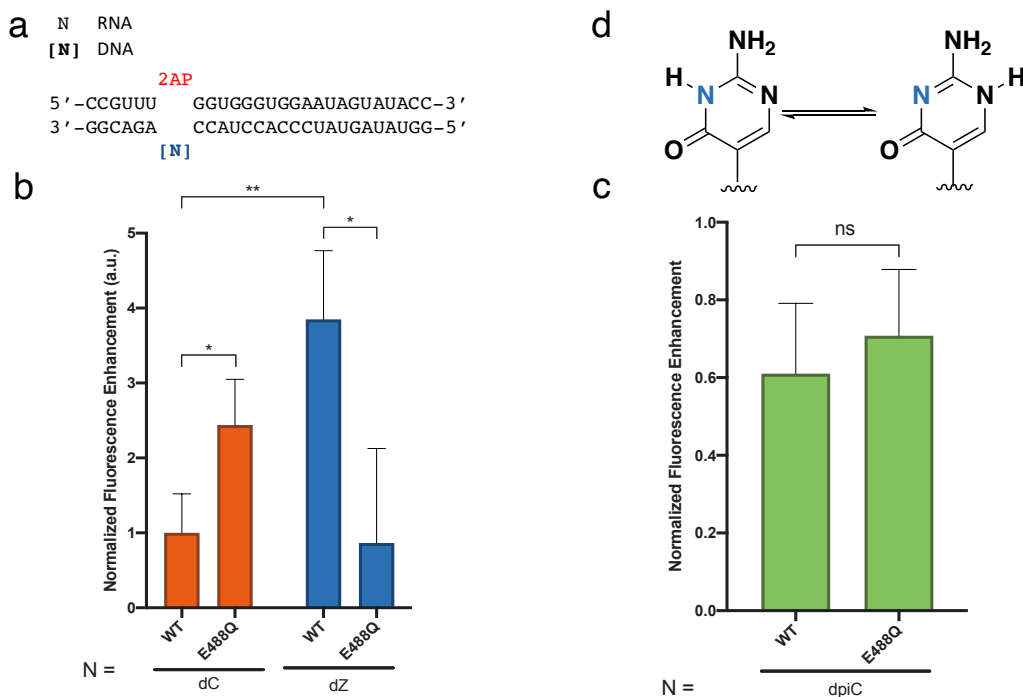


Figure 3.4. (a) The substrate used in the base-flipping assay, containing 2-aminopurine (2AP) in the edited position. 2'-Deoxynucleotides are shown in brackets, and all others are ribonucleotides. The orphan base is shown as the blue N, and is comprised of either 2'-deoxycytidine (dC) or 2'-deoxy Benner's base Z (dZ). (b) Plot showing normalized fluorescence enhancement for combinations of 10 μ M ADAR2 wild-type or E488Q paired with 2.5 μ M dsRNA containing either orphan base. Error bars represent the standard deviation of three technical replicates. Statistical significance between groups was determined using an unpaired t-test with Welch's correction; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$. (c) Plot showing normalized fluorescence enhancement for combinations of 10 μ M ADAR2 wild-type or E488Q paired with 2.5 μ M dsRNA containing either orphan base. Error bars represent the standard deviation of three technical replicates (* $p \leq 0.05$). (d) Pseudoisocytidine undergoes tautomerization which allows for a hydrogen bond donor or acceptor at N3. All nucleotides are RNA unless in brackets ([N]) to indicate DNA.

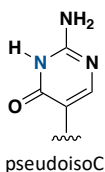
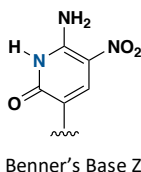
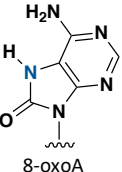
We also examined the change in 2-AP fluorescence for substrates in the presence of ADAR2 E488Q. It was previously shown that ADAR2 E488Q exhibits a greater extent of base-flipping than the wild-type enzyme.⁸⁴ We found that this held true for guide RNAs containing 2'-deoxycytidine in the orphan

position. For guide RNAs containing dC, a 2.4-fold greater increase in fluorescence was observed when ADAR2 E488Q was added to the fluorescent substrate compared to wild-type (Figure 3.4). However, we found that the enzyme preference was reversed for RNAs containing dZ. For guide RNAs containing dZ, base-flipping was reduced by over 4-fold in the presence of ADAR2 E488Q (Figure 3.4). Interestingly, guide RNAs containing dpiC did not show a preference for either enzyme (Figure 3.4 C). This could be due to the ability of the pseudoisocytidine base to tautomerize such that N3 would be either a hydrogen bond donor or acceptor (Figure 3.4 D).¹⁵¹ These results suggest that base-flipping by ADAR2 is most efficient when a complementary hydrogen bond donor-acceptor pair exists between residue 488 and N3 of the orphan base.

Cytidine Analogs increase rate of A-to-I editing in vitro.

While examining how amino acids within the ADAR2 enzyme mediate editing preferences, Kuttan and Bass observed a positive correlation between catalytic rate and enhanced base-flipping ability.⁸⁴ Therefore, orphan analogs that stabilize the base-flipped conformation of the protein-RNA complex should also stimulate catalysis. To test this idea, cytidine analogs examined in the base-flipping assay were incorporated into 29-nt guide RNAs and hybridized to a 320-nt RNA sequence derived from a mouse model of Hurler syndrome (Figure 3.5a).^{152,153} Hurler syndrome is the most severe type of α -l-iduronidase enzyme deficiency, caused by mutations in the *IDUA* gene. The most common of these is the nonsense mutation W402X, which is modelled by the W392X mutation in the mouse *Idua* gene.¹⁵² This TGG-to-TAG mutation can be reversed by site-directed RNA editing with ADAR enzymes. Qu et. al have shown that RNA guides can recruit endogenous ADARs to edit the *IDUA* pre-mRNA and mRNA, restoring α -l-iduronidase catalytic activity in patient-derived W402X fibroblasts.⁶⁷ The base-flipping study described above suggests that guide RNAs containing dZ may enhance ADAR2 activity. Therefore, we compared duplexes where the orphan base was either dC, dZ, dpiC, or 8-oxodA in an *in vitro* deamination reaction with ADAR2 wild-type under single-turnover conditions (Figure 3.5). The combination of ADAR2 wild-type and an RNA substrate containing 2'-deoxycytidine (dC), led to a deamination reaction where $k_{\text{obs}} = 1.2 \pm 0.1 \text{ min}^{-1}$ ($k_{\text{rel}} =$

1), while the same protein deaminated the substrate containing an orphan dZ 3-fold faster ($k_{\text{obs}} = 3.6 \pm 0.3 \text{ min}^{-1}$, $k_{\text{rel}} = 3.0$) (Figure 3.5 B; Table 3.1). This result demonstrates the important role that the E488-orphan base contact performs in enabling efficient deamination by the wild-type enzyme. The A-dpiC substrate also showed a significant increase in reaction rate ($k_{\text{obs}} = 1.6 \pm 0.1 \text{ min}^{-1}$, $k_{\text{rel}} = 1.3$). The rate enhancement was not as large as with the A-dZ substrate, which could be due to its ability to present either a hydrogen bond donor or acceptor at N3 through tautomerization (Figure 3.4 D; Table 3.1). However, when the RNA substrate contained 8-oxodA in the orphan position, the rate of reaction significantly decreased to $k_{\text{obs}} = 0.4 \pm 0.2 \text{ min}^{-1}$ ($k_{\text{rel}} = 0.3$) (Table 3.1). This may be a result of 8-oxodA adopting the *anti* conformation instead of the *syn* conformation that is required to present the hydrogen bond *donor-donor-acceptor* face to contact residue 488.¹⁵⁴

N	RNA				
[N]	DNA				
		A			
	5' - ...GGAGAACAACUCU	GGCAGAGGUCUCAAA...-3'			
	3' - CCUCUUGUUGAGA	CUGUCUCCAGAGUUU -5'			
		[N]			
N =		OR		OR	
	pseudoisoC		Benner's Base Z		8-oxoA

enzyme	substrate [N]	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b
ADAR2	C	1.2 ± 0.1	1
	piC	1.6 ± 0.1	1.3
	Z	3.6 ± 0.3	3.0
ADAR1 p110	8-oxodA	0.4 ± 0.2	0.3
	C	0.060 ± 0.008	1
	Z	0.19 ± 0.02	3.2

Table 3.1. Single turnover rate constants for reaction of ADAR1/2 with guide RNAs containing cytosine analogs at the orphan position. Reactions with ADAR2 were carried out with 0.8 nM RNA and 2 nM enzyme and for ADAR1 p110 were carried out with 5 nM RNA and 50 nM enzyme.^a Data were fitted to the equation $[P]_t = \alpha[1 - \exp(-k_{\text{obs}}t)]$.^b $k_{\text{rel}} = k_{\text{obs}}$ for analog / k_{obs} for cytosine. All nucleotides are RNA unless in brackets ([N]) to indicate DNA.

Notably, the ADAR2 E488Q enzyme shows an opposite substrate preference from ADAR2 WT. The wild-type enzyme more efficiently deaminates adenosines paired across dZ, while ADAR2 E488Q prefers substrates containing the orphan base dC (Figure 3.5). This illustrates that k_{obs} is greater when a hydrogen-bond *donor-acceptor* pair is present between N3 of the orphan base and residue 488. This also mirrors the result of the base-flipping assay, showing that orphan base modifications that enhance base-flipping also provide an increase in editing rate.

Although there are currently no published high-resolution structures of ADAR1 to inform guide RNA design, it is known that ADAR1 E1008 corresponds to E488 in ADAR2.¹⁵⁵ Similarly, a glutamate to glutamine mutation in this residue of ADAR1 (E1008Q) shows higher activity than the wild-type enzyme.¹⁰⁸ Therefore, replacing the orphan base with an N3 hydrogen bond donor may increase reaction rate for ADAR1 as well as ADAR2. To test this idea, we compared duplexes where the orphan base was either dC or dZ in an *in vitro* deamination reaction with ADAR1 p110 wild-type under single-turnover conditions (Figure 3.5 C,D). The combination of ADAR1 p110 and the A-dC substrate led to a reaction where $k_{\text{obs}} = 0.060 \pm 0.008 \text{ min}^{-1}$ ($k_{\text{rel}} = 1$) (Table 3.1). When the RNA substrate instead contained dZ in the orphan position, the rate of reaction increased to $k_{\text{obs}} = 0.19 \pm 0.02 \text{ min}^{-1}$ ($k_{\text{rel}} = 3.2$). This result indicates that the rate of reaction for the ADAR1 enzyme can also be enhanced by including an N3 hydrogen bond donor analog in the orphan position of the guide RNA.

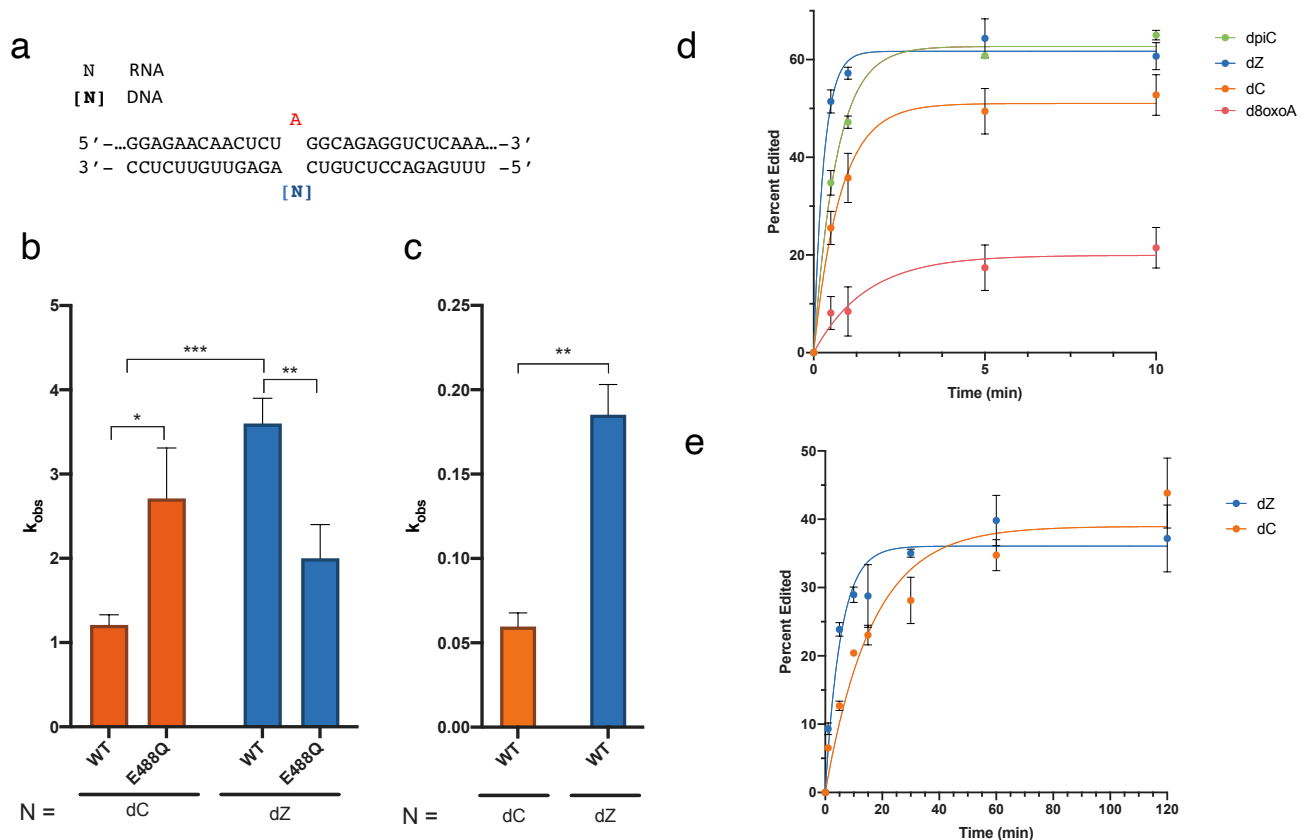


Figure 3.5. (a) Partial sequence of the *IDUA* target substrate used for *in vitro* deaminations, shown opposite the guide sequence. The target adenosine is shown in red, 2'-deoxynucleotides are shown in brackets, and all others are ribonucleotides. The orphan base is shown as the blue N, and is either 2'-deoxycytidine (dC) or 2'-deoxy Benner's base Z (dZ). (b) Observed rate constants for the reaction of 2 nM ADAR2 wild-type (WT) or E488Q paired with 0.8

nM dsRNA substrate containing either dC or dZ in the orphan position. (c) Observed rate constants for the reaction of 50 nM ADAR1 p110 wild-type paired with 5 nM RNA substrate containing either dC or dZ in the orphan position. Error bars represent the standard deviation of three technical replicates. Statistical significance between groups was determined using an unpaired t-test with Welch's correction; * $p \leq 0.05$; ** $p \leq 0.01$. (d) Deamination product versus time for 2 nM wild-type ADAR2 with RNA guides containing varying orphan bases (2'-deoxypseudoisocytidine, dpiC; 2'-deoxy Benner's base Z, dZ; 2'-deoxycytidine, dC; 8-oxo-2'-deoxyadenosine, 8-oxodA). (e) Deamination product versus time for 50 nM wild-type ADAR1 with RNA guides containing either 2'-deoxy Benner's base Z or 2'-deoxycytidine.

As seen by others, the use of ADAR2 E488Q *in vitro* resulted in significantly greater levels of off-target editing (Figure 3.6).^{42,43,64,73,128} The presence of glutamate at the 488 position in wild-type ADAR2 is likely important in editing regulation, providing a compromise between editing efficiency and specificity.⁸⁴ In contrast with ADAR2 E488Q, which enhances the rate of reaction but also results in editing promiscuity,^{42,64,73,128} the dZ modification did not lead to higher levels of bystander editing (Figure 3.6).

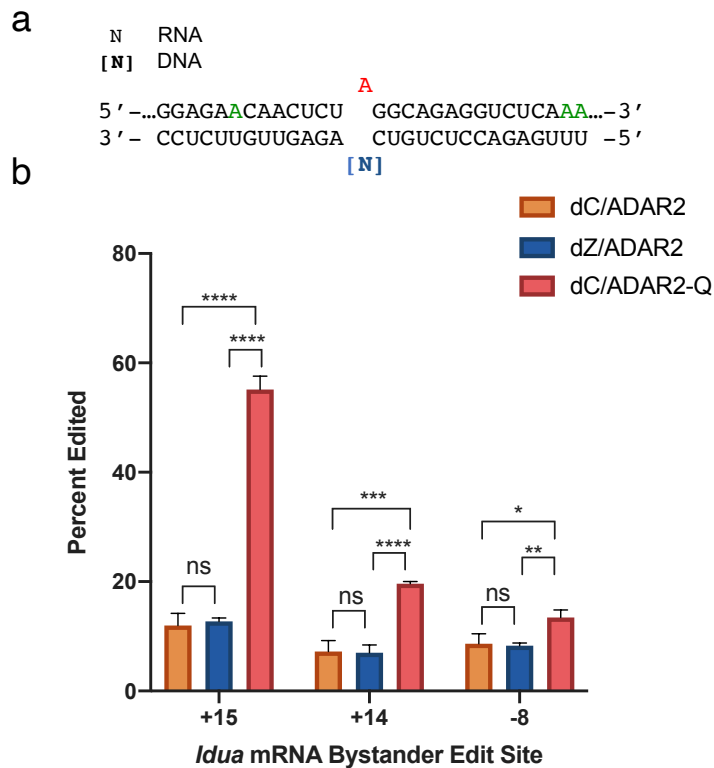


Figure 3.6. Bystander editing observed within the *IDUA* mRNA sequence. (a) Partial sequence of the *IDUA* target substrate used for *in vitro* deaminations, shown opposite the guide sequence. The target adenosines are shown in red, bystander edit sites are shown in green, 2'-deoxynucleotides are shown in brackets, and all others are ribonucleotides. The orphan base is shown as the blue N, and is comprised of either 2'-deoxycytidine (dC) or 2'-deoxy Benner's base Z (dZ). (b) Percent editing of *IDUA* *in vitro* substrate with combinations of RNA guide and ADAR2 WT or ADAR2 E488Q (ADAR2 E488Q) after 60 minutes.

High-resolution structure reveals ADAR-dZ contacts.

The enhanced base-flipping and *in vitro* reaction rate stimulated by dZ containing substrates inspired the use of X-ray crystallography to examine the specific contacts made by the dZ base. The nucleoside analog 8-azanebularine has been shown to enable these structural studies, by trapping the ADAR2-RNA complex in the base-flipped conformation.^{26,107,156} We synthesized a target RNA containing 8-azanebularine in place of the reactive adenosine, paired across a guide RNA containing dZ at the orphan position (Figure 3.7 A). The crystal structure of human ADAR2 deaminase domain WT (hADAR2d WT)/dZ was determined at 2.5Å resolution in collaboration with Xander Wilcox from Professor Andy Fisher's lab.

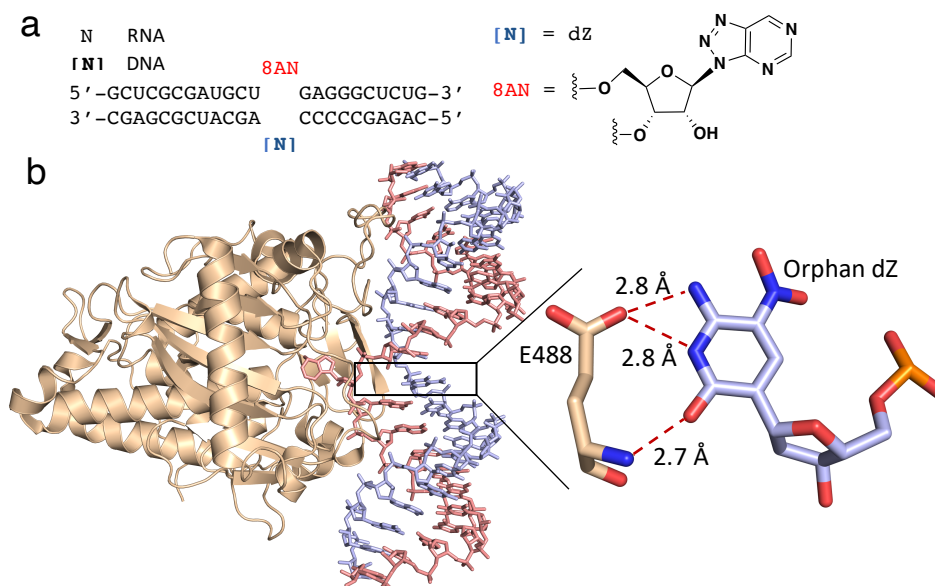


Figure 3.7. (a) The *GLII* sequence duplex RNA used for crystallization. The Z base ([N]) is paired across from 8-azanebularine (8AN). (b) The crystal structure of ADAR2 deaminase domain bound to an RNA containing Benner's base dZ as the orphan base (PDB ID: 7KFN) shows enhanced hydrogen bonding contacts between the orphan base and E488.

As observed in previous structures, the asymmetric unit contained one deaminase domain bound to dsRNA (Chain A) dimerized with a dsRNA-free deaminase domain (Chain D).¹² Expectedly, the 8-azanebularine is flipped out of the duplex into the protein active site and the base-flipping residue E488 interacts with the orphan base dZ through hydrogen bonds from both sidechain and main-chain atoms. The presence of the dZ base was confirmed by deleting the nitro group at position 5 from the structural model, running a round of refinement, and analyzing the 2Fo-Fc and Fo-Fc maps. Deleting the nitro group resulted in a strong peak of positive Fo-Fc difference electron density where the nitro group would lie (Figure 3.8 A). Reintroduction of the nitro group results in a loss of the positive Fo-Fc difference density (Figure 3.8 B).

Interestingly, the E488-dZ pair makes an additional interaction where the carboxylate of E488 forms a 2.80 Å bifurcated hydrogen bond with the exocyclic amine at the 4-position and the N3 of the orphan dZ (Figure 3.7 B). Given that dZ provides an additional hydrogen bond donor to the E488 acceptor that cytidine lacks, the E488-dZ pair likely stabilizes the base-flipped conformation. In previous structures this stabilization required either protonation at N3 (Figure 3.2 B) or a mutant Q488 flipping residue (Figure 3.2 A) to provide a hydrogen bond donor.

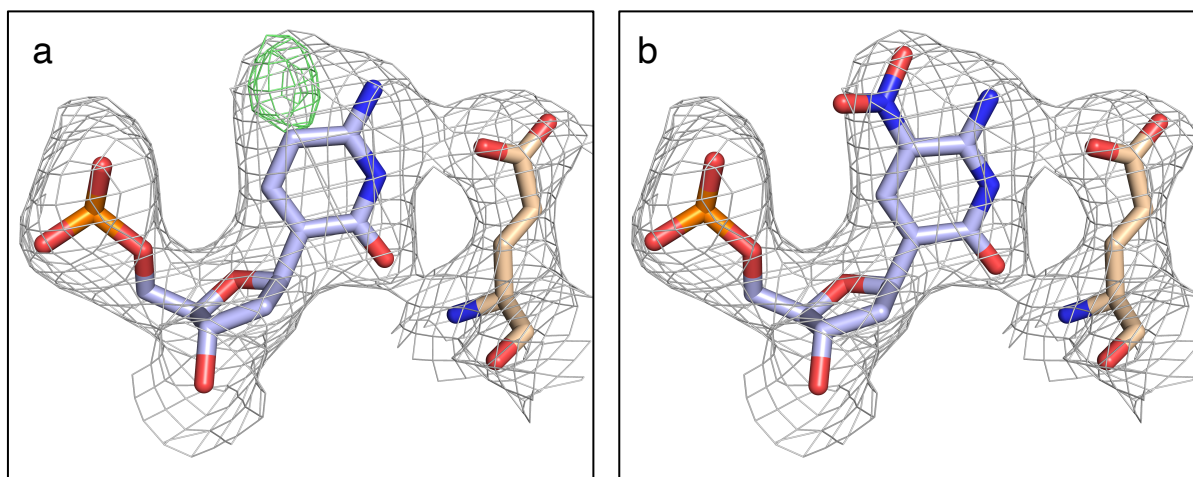


Figure 3.8. Related to Figure 2.7. Confirmation of the presence of the dZ base. (a) Deleting the nitro group of the dZ base resulted in a strong peak of positive Fo-Fc difference density shown in green (contoured at 3σ) together with 2Fo-Fc electron density (contoured at 1σ). (b) Reintroduction of the nitro group eliminates the positive difference density. Electron densities are colored and contoured as in (a).

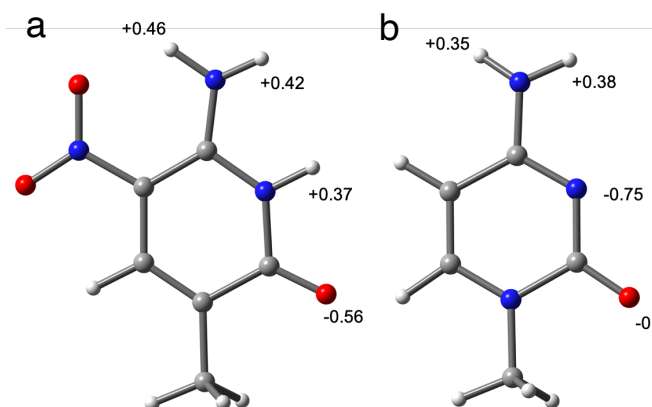


Figure 3.9. Density Functional Theory (DFT) charge density calculations of the Watson-Crick-Franklin faces of (a) Benner's base Z and (b) cytosine. The M06-2X/6-311+G(2d,p) level of theory with CHelpG charges was used.

The presence of the nitro group on the dZ base prompted us to question if the charge distribution of the Watson-Crick-Franklin face was contributing to enhanced stability of the flipped conformation. To investigate this, density functional theory (DFT) calculations of Benner's base Z and cytosine were carried out by Professor Dean Tantillo using the M06-2X/6-311+g(2d,p) method. These calculations suggest that the partial charge of the *donor-donor-acceptor* face of the dZ base adopts a more positive charge (Figure 3.9 A) than the *donor-acceptor-acceptor* face of cytosine (Figure 3.9 B). This suggests that the electrostatic interaction between the partial positive face of dZ and the carboxylate of E488 may provide additional favorable interaction energy to the base-flipped conformation.

The deaminase domain of ADAR2 is known to make several contacts with the 2'-hydroxyl of the ribose sugars. Given that the orphan dZ nucleotide lacks a 2'-hydroxyl, we sought to observe the consequence of losing this interaction. Typically, R510 forms a hydrogen bond with the 2'-hydroxyl of the orphan nucleotide and a previous 2.80 Å resolution structure (PDB ID: 6VFF) shows an additional hydrogen bond to a nearby water molecule (Figure 3.10 A).¹² While direct contact to the orphan dZ was not observed in this structure, R510 interacts with the neighboring cytosine through a water-mediated hydrogen bonding network. The hydrogen bonding network links R510 to two successive water molecules (at 2.6 and

3.3 Å, respectively) and ultimately to the carbonyl oxygen at the 2-position of the neighboring cytosine at 3.0 Å (Figure 3.10 B)

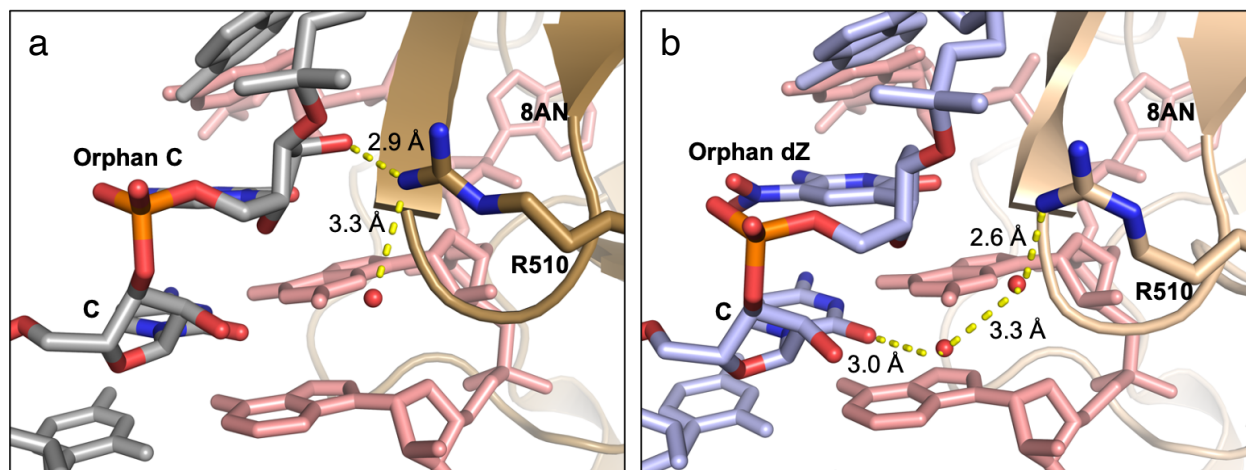


Figure 3.10. Related to Figure 3.7. (a) Previous structures (PDB ID: 6VFF) observe R510 making a hydrogen bond to a water molecule and the 2'-hydroxyl of the orphan cytidine. (b) The use of dZ at the orphan base removes the hydrogen bonding capability between R510 and the orphan base sugar but a water mediated hydrogen bonding network links R510 to the exocyclic carbonyl oxygen at the 2-position of the cytidine of the neighboring base. Site-directed RNA editing in human cells with overexpressed ADAR2 is increased by the dZ modification.

Directed editing with endogenous ADAR is enhanced by the dZ orphan nucleotide.

Our *in vitro* findings suggested that directed A-to-I editing by wild-type human ADARs could be enhanced with the use of dZ in guide oligonucleotides. To determine if this modification could provide editing enhancement in human cells, we synthesized a 39-nt guide oligonucleotide to direct editing by ADAR2 wild-type to the 3'-UTR of the endogenous *RAB7A* RNA present in HEK293T cells. This target was used previously in directed-editing experiments, and was chosen due to the presence of an adenosine within the optimal 5'-UAG-3' flanking sequence for ADAR2 (Figure 3.11 A).^{37,65,156} The guide strand contained 2'-O-methyl and phosphorothioate modifications to increase cellular stability and prevent bystander editing,^{61,157} as well as either dC or dZ opposite the target adenosine.^{59,156-158} HEK293T cells were transfected with either dC- or dZ- containing guide oligonucleotide, and a plasmid for ADAR2 overexpression.

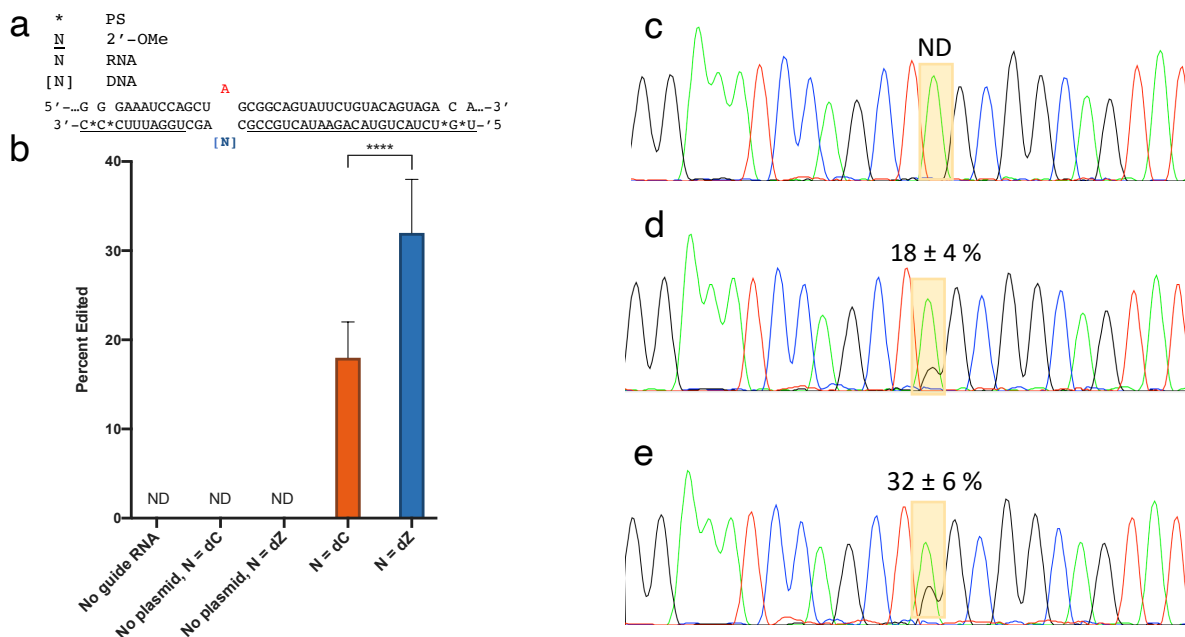


Figure 3.11. (a) Partial sequence of the endogenous *RAB7A* transcript paired with the guide RNA. The target adenosine is shown in red. Underlining indicates 2'-O-methyl nucleotides, asterisks indicate phosphorothioate linkages, and nucleotides in brackets are 2'-deoxynucleotides, all others are ribonucleotides. N = cytosine or Benner's base Z. (b) Percent editing of the 3'-UTR of *RAB7A* with guide RNA and overexpression of ADAR2, overexpression of ADAR2 with no guide RNA, or guide RNA without ADAR2 overexpression. ND indicates no detected editing. Error bars represent standard deviation (n = 12; 4 biological and 3 technical replicates), statistical analysis was performed using the unpaired t-test with Welch's correction; ****p < 0.0001. (c-e) Sequencing traces from directed editing of endogenous *RAB7A* target. Color coding is as follows: Green, A; Blue, C; Red, T; Black, G. (c) Control (no guide RNA) condition with overexpression of hADAR2 WT. (d) Directed editing with overexpression of hADAR2 WT using a guide RNA bearing the orphan base 2'-deoxycytidine. (e) Directed editing with overexpression of hADAR2 WT using a guide RNA bearing the orphan base 2'-deoxy Benner's base Z.

Guide oligonucleotide was required to observe editing at the *RAB7A* target, and no editing was observed without ADAR2 overexpression (Figure 3.11 B). The lack of editing without ADAR2 overexpression is likely due to the relatively low levels of ADAR expression in HEK293T cells.¹⁰⁵ Efficient directed editing in these cells has previously required either plasmid transfection or stable overexpression (e.g. Flip-In-T-Rex).^{41,156,159} Transfection with the orphan dC guide oligonucleotide resulted in editing levels of 18 % ± 4 %, while transfection of the dZ guide oligonucleotide increased target editing levels to 32 % ± 6 % (Figure 3.11). This single nucleotide change was able to induce an approximately 1.8-fold increase in editing of an endogenous target.

The enhancement of cellular editing using overexpressed ADAR2 suggested that use of the dZ orphan nucleotide in engineered guide strands might also amplify editing by endogenous ADARs (i.e. without ADAR overexpression). To test this idea, we synthesized guide oligonucleotides containing either dC or dZ, in addition to 2'-O-methyl and phosphorothioate modifications, to be tested in different sequence contexts with endogenous ADARs (Figure 3.12).

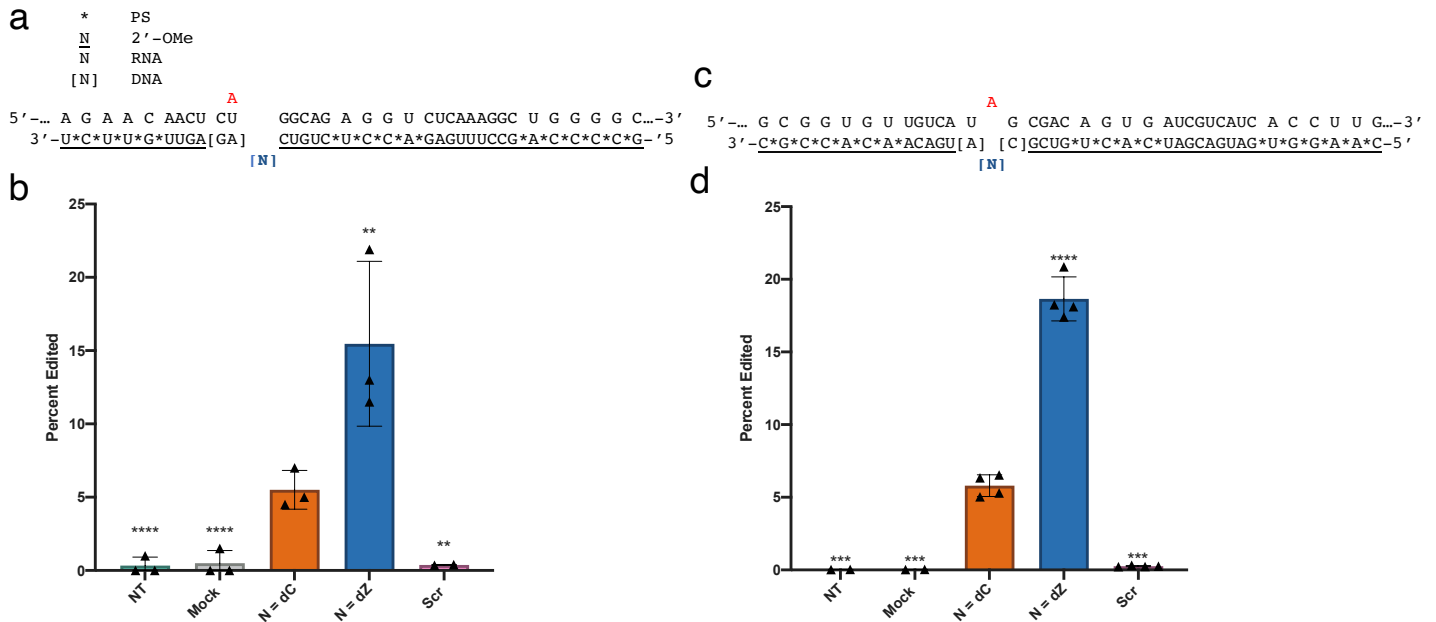


Figure 3.12. (a) Partial sequence of the endogenous *IDUA* mRNA target sequence paired with the guide oligonucleotide. The guide modifications are annotated as described in Figure 3.11 (b) Directed editing on the *IDUA* RNA transcript in mouse primary liver fibroblasts. NT is no transfection, and the scrambled sequence (Scr). Plotted values are the means of three biological replicates (\blacktriangle) and two technical replicates \pm standard deviation. The scrambled sequence (Scr) consisted of $n = 4$ measurements (2 biological and 2 technical replicates). (c) Partial sequence of the endogenous *APP* mRNA target sequence paired with the guide oligonucleotide. (d) Directed editing on the *APP* RNA transcript in ARPE-19 cells. Plotted values are the means of two sets of two identical biological replicates (\blacktriangle) and two technical replicates \pm standard deviation. The NT and Mock conditions consisted of $n = 4$ measurements (2 biological and 2 technical replicates). Statistical significance was determined by comparing each value to the N = dC condition using the unpaired t-test with Welch's correction; $n = 6$ or 8 (3 biological and 2 technical replicates; 2 sets of 2 biological and 2 technical replicates); ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p < 0.0001$.

These studies were carried out in collaboration with scientists at ProQR Therapeutics. Editing was performed in mouse primary liver fibroblast cells isolated from homozygous *Idua-W392X* mice on a C57BL/6 background. These cells served as a mouse model for Hurler Syndrome, carrying the disease-associated G-to-A mutation in the *Idua* gene.¹⁵² Editing was quantified by digital droplet PCR (ddPCR) and no editing was seen in conditions without transfection, with a transfection lacking guide oligonucleotide,

or using a scrambled sequence oligonucleotide. Transfection with 100 nM control guide strand bearing dC at the orphan position induced $6 \pm 2\%$ editing at the *IDUA* target site whereas a similar guide with dZ at the orphan position induced $15 \pm 5\%$ editing (Figure 3.12 B) Thus, compared to a guide oligonucleotide that was otherwise identical, the dZ containing guide produced 2.5-fold higher editing levels with endogenous ADAR.

Finally, directed editing was carried out on the *APP* transcript in human retinal pigment epithelium (ARPE-19) cells. The *APP* transcript was chosen as a model due to its ubiquitous expression in wild-type cells and the presence of an adenosine within a preferred sequence context for ADAR editing. We synthesized guide oligonucleotides containing either dC or dZ to target an adenosine in an AUA codon of Ile 712 thus creating the Ile to Met amino acid change within the γ - secretase cleavage site of *APP*. We again used ddPCR to quantify the editing of *APP* mRNA. As shown in Figure 2.12 D, the $19 \pm 2\%$ *APP* RNA editing was achieved in ARPE-19 cells by transfection of 100 nM guide strand containing dZ orphan nucleotide, while only $5.8 \pm 0.8\%$ was measured for the guide strand containing dC with otherwise identical sequence and chemical modifications (Figure 3.12 C). Thus, again 3.3-fold increase in editing was achieved by dZ containing guide strand when compared to an identical but dC containing guide strand. No editing of *APP* was observed after transfection of the scrambled guide strand, no guide strand, or without transfection. Together, these experiments demonstrate that the enhancement of directed editing via endogenous ADARs by the dZ orphan nucleotide is not target RNA, sequence, or cell specific.

Discussion.

The use of endogenous ADARs represents a potentially safer strategy for targeting RNA and has been shown to substantially reduce levels of off-target edits, when compared to enzyme overexpression.⁴² However, native ADARs typically exhibit lower levels of editing, inspiring strategies to engineer guide RNAs that overcome this barrier. Previously, Merkle et. al appended a hairpin derived from a native ADAR2 substrate to create ~60-95 nt chemically modified guides that recruit endogenous ADARs.⁴¹ Katrekar and co-workers also utilized a recruitment domain strategy, which they linked to long antisense

domains.⁴² In the LEAPER system, Qu et. al used 70-151 nt guide RNAs with A-C mismatches at the target site to direct editing by endogenous ADARs to a specific position within the resulting dsRNA.⁶⁷ Both Katrekar and Qu opted to genetically encode their guide RNAs as opposed to transfecting chemically modified guide oligonucleotides.^{42,67} These studies have established the feasibility of an endogenous ADAR editing platform that only requires delivery of a guide RNA. However, limitations to each of these approaches still exist. The length of these guide oligonucleotides can lead to off-targets within the long double stranded region formed on the target transcript, or through partial base pairing of the guide to off-target transcripts, although the guide oligonucleotides were restricted to these lengths as to overcome insufficient editing levels.⁶⁷ Long antisense domains have also been associated with a decrease in target mRNA levels through an RNAi-like effect.^{42,43} In addition to optimizing the length and sequence of guide oligonucleotides, chemical modification of ADAR guide RNAs (2'-O-methylations, phosphorothioates, locked nucleic acid) has been employed to improve cellular stability, binding kinetics, and to reduce off-target editing.⁴¹ Previous efforts have focused on recruitment and binding of endogenous ADARs, but the need remains for guide RNA modifications that specifically enhance catalysis. Here, we described nucleotide modifications that can enhance *in vitro* deamination kinetics when paired across a target adenosine (dZ and dpiC), and a single nucleotide modification (dZ) that ultimately more than doubled editing yield via endogenous ADARs.

The orphan position of guide oligonucleotides is known to play an important role in ADAR editing. Positioning a cytidine opposite the target A to generate an A-C mismatch is a widely used technique, and is shown to increase ADAR reaction rates *in vitro*.^{72,111} This rate enhancement has been attributed to the stability of the resultant base pair (I-C > I-U) post-deamination.⁷² However, we sought to examine how orphan base modifications can influence base-flipping to alter the reaction rate. The ADAR2 E488Q mutant is able to dramatically increase the deamination rate for an abundance of substrates without a significant change in binding affinity, owing to an enhancement in base-flipping.⁸⁴ Our *in vitro* studies confirm that this enhancement is due to an important hydrogen bonding contact between N3 of the orphan base and

residue 488 (Figure 2.2 A,B). Therefore, we sought to increase editing efficiency of wild-type ADAR2 by engineering the guide RNA to fulfill this same contact (Figure 2.2 C). We identified cytidine analogs that retained other ADAR2-RNA contacts, while offering hydrogen bond donation at N3. Yang et al. had previously reported an artificially expanded genetic information system (AEGIS) in which they designed purine-pyrimidine nucleobase pairs that retain standard geometry, while rearranging hydrogen bond donor-acceptor groups.¹⁴⁸ Our interest was in the “Z” nucleobase reported by the Benner group: a pyrimidine *donor-donor-acceptor*. In contrast with cytosine—a pyrimidine *donor-acceptor-acceptor*—this base offers hydrogen bond donation at N3.

Our analysis of base-flipping with different combinations of orphan base and either ADAR2 or ADAR2 E488Q, furthered the understanding of protein-RNA contacts required for base-flipping. Base-flipping efficiency was greatest when there was a hydrogen bond *donor-acceptor* pair between N3 of the orphan base, and the side chain of residue 488. Consequently, base-flipping for wild-type ADAR2 was greatest when paired with a substrate containing dZ (Figure 3.4). Kuttan and Bass had previously demonstrated a positive correlation between catalytic rate and base-flipping ability of ADAR2 mutants.⁸⁴ Correspondingly, we observed that the dZ modification which enhanced base-flipping, also resulted in greater *in vitro* deamination rates (Figure 3.5). The high-resolution crystal structure of ADAR2 deaminase domain bound to a dZ-containing RNA allowed us to explore the structural basis of this improvement. As intended, the E488 residue formed a close hydrogen bonding interaction with the dZ base (Figure 3.7).

We have shown that the dZ modification can be used to increase editing efficiency in mouse and human cells. A potential advantage of using orphan C analogs in guide oligonucleotides designed to promote directed RNA editing resides in the avoidance of off-target editing associated with the use of hyper-editing ADAR2 mutants. Ultimately, we showed that guide oligonucleotides containing dZ enhance editing by endogenous ADARs for two targets (up to 3.3-fold) when compared to an otherwise identical guide containing dC at the orphan position (Figure 3.12). The use of dZ in the orphan position of guide RNAs represents a general approach to increase ADAR editing efficiency, having demonstrated success in

multiple sequence contexts. In addition, it is compatible with a range of other chemical modifications (2'-O-methyl, phosphorothioate). Therefore, this approach could be combined with other methods used for endogenous ADAR recruitment and editing efficiency (chemical modifications for cellular stability, recruitment domains, lengthening of the guide oligonucleotide, etc.) to provide additional increases in editing efficiency.^{41,42,67} In addition, positioning of the chemical modifications and the relative location of the orphan base within the guide RNA can be further optimized. Our method faces similar limitations as other chemically modified ADAR guides, in that they are not genetically encodable and must be delivered to target tissues. However, targeted delivery of oligonucleotides is becoming more feasible, as demonstrated by recent successes in the delivery of chemically modified oligonucleotides to the liver.¹⁶⁰ In addition, different RNA substrates demonstrate varying levels of rate enhancement with the use of ADAR2 E488Q.⁸⁴ Since the mechanism of rate enhancement by the dZ base likely mimics that of the mutant enzyme, it remains to be seen whether this strategy will be useful for substrates and sequence contexts that are not affected by mutation of the base-flipping residue to glutamine. Furthermore, we have yet to explore fully the utility of this orphan modification when the target adenosine is not within the context of the ideal nearest neighbors for ADARs.³⁷ This work is ongoing in our laboratory.

Methods.

Sequences of Oligonucleotides.

a) RNA sequences for AD2d/dZ crystallography.

<i>GLII</i> top strand containing 8-azaN (N)	5'-GCUCGCGAUGCUNGAGGGCUCUG-3'
<i>GLII</i> bottom strand containing 2'-deoxy Benner's base Z (Z)	5'-CAGAGCCCCC[Z]AGCAUCGCGAGC-3'

b) Sequences for base-flipping assay (nucleotides in brackets are 2'-deoxy. All others are ribonucleotides).

<i>GRIA2</i> top strand containing 2-aminopurine (2AP)	5'-CCGUUU[2AP]GGUGGGUGGAAUAGUAUACC-3'
--	---------------------------------------

<i>GRIA2</i> bottom strand containing 2'-deoxycytidine	5'-GGUAUAGUAUCCCACCUACC[C]AGACGG-3'
<i>GRIA2</i> bottom strand containing 2'-deoxypseudoisocytidine (piC)	5'-GGUAUAGUAUCCCACCUACC[piC]AGACGG-3'
<i>GRIA2</i> bottom strand containing 2'-deoxy benner's base Z (Z)	5'-GGUAUAGUAUCCCACCUACC[Z]AGACGG-3'

c) Sequences for *in vitro* kinetics (nucleotides in brackets are 2'-deoxy. All others are ribonucleotides). All PCR primers are 2'-deoxynucleotides.

<i>IDUA</i> bottom strand containing 2'-deoxy cytidine	5'-UUUGAGACCUCUGUC[C]AGAGUUGUUCUCC-3'
<i>IDUA</i> bottom strand containing 2'-deoxy pseudoisocytidine (piC)	5'-UUUGAGACCUCUGUC[piC]AGAGUUGUUCUCC-3'
<i>IDUA</i> bottom strand containing 2'-deoxy benner's base Z (Z)	5'-UUUGAGACCUCUGUC[Z]AGAGUUGUUCUCC-3'
<i>IDUA</i> bottom strand containing 2'-deoxy-8-oxoadenosine (8oA)	5'-UUUGAGACCUCUGUC[8oA]AGAGUUGUUCUCC-3'
<i>IDUA</i> RT-PCR forward and sequencing primer	5'-GCTCCTCCCATCCTGTGGGCTGAACAGT-3'
<i>IDUA</i> RT-PCR reverse primer	5'-CGGGGTGTGCGTGGGTGTCATCACT-3'

d) Sequences for directed editing in HEK293T cells (phosphorothioate modification is marked with an asterisk, 2'-O-methylated nucleotides are underlined, 2'-deoxynucleotides are in brackets; all others are ribonucleotides). All PCR primers are 2'-deoxynucleotides.

<i>RAB7A</i> endogenous target guide RNA	5'- <u>U*G*UCUACUGUACAGAAUACUGCCGC</u> [C]AGCUGGAUUU <u>C*C*C-</u> '3
<i>RAB7A</i> endogenous target guide RNA containing 2'-deoxy Benner's base Z (Z)	5'- <u>U*G*UCUACUGUACAGAAUACUGCCGC</u> [Z]AGCUGGAUUU <u>C*C*C-</u> '3
Endogenous <i>RAB7A</i> RT-PCR forward primer	5'-GCAACCAATTTAAATGTATAAATTAGTGTAAGAAATT-3'
Endogenous <i>RAB7A</i> RT-PCR reverse primer	5'-GCTACAATGCAGGGGCAGATCCTAGGAAG-3'

Endogenous <i>RAB7A</i> nested PCR forward and sequencing primer	5'-CTTGGATTATGTGTTTAAGTCCTGTAATGCAGGCC-3'
Endogenous <i>RAB7A</i> nested PCR reverse primer	5'-GGAGCAGAAGCTGCCAGGGTTCCAACC-3'

d) Sequences for directed editing in mouse primary liver fibroblasts and an oligonucleotide with a scrambled sequence (phosphorothioate modification is marked with an asterisk, 2'-O-methylated nucleotides are underlined, 2'-deoxynucleotides are in brackets; all others are ribonucleotides). All PCR primers and probe sequences are 2'-deoxynucleotides.

Endogenous <i>IDUA</i> guide RNA	5'- <u>G*C*C*C*C*A</u> *GCCUUUGAG*A*C*C*U*CUGUC[<u>CAG</u>]AGU U*G*U*U*C*U-'3
Endogenous <i>IDUA</i> guide RNA containing 2'-deoxy Benner's base Z (Z)	5'- <u>G*C*C*C*C*A</u> *GCCUUUGAG*A*C*C*U*CUGUC[<u>ZAG</u>]AGU U*G*U*U*C*U-'3
Scrambled Sequence (Scr)	5'- <u>G*G*C*U*C*U</u> *CCCAAUCUC*C*C*G*C*AGUGC[<u>TGC</u>]CUG A*U*A*U*G*U-'3
Endogenous <i>IDUA</i> sequence Forward primer	5'- CTCACAGTCATGGGGCTC -3'
Endogenous <i>IDUA</i> sequence Reverse primer	5'- CACTGTATGATTGCTGTCCAAC -3'
Fluorescin, non- fluorescent quencher labeled wild-type probe	5'-AGAACAACCTCTGGGCAGAGGTCTCA -3'
Hexachlorofluorescin, non-fluorescent quencher labeled mutant probe	5'- AGAACAACCTCTAGGCAGAGGTCTCA-3'

e) Sequences for directed editing in human retinal pigment epithelium cells (ARPE-19) and an oligonucleotide with a scrambled sequence (phosphorothioate modification is marked with an asterisk, 2'-O-methylated nucleotides are underlined, 2'-deoxynucleotides are in brackets; all others are ribonucleotides). All PCR primers and probe sequences are 2'-deoxynucleotides.

Endogenous <i>APP</i> guide RNA	5'- <u>C*A*A*G*G*U</u> *GAUGACGAU*C*A*C*U*GUCG[<u>CCA</u>]UGACA*A*C*A* C*C*G*C -'3
Endogenous <i>APP</i> guide RNA containing 2'- deoxy Benner's base Z (Z)	5'- <u>C*A*A*G*G*U</u> *GAUGACGAU*C*A*C*U*GUCG[<u>CZA</u>]UGACA*A*C*A* C*C*G*C -'3
Scrambled Sequence (Scr) for <i>APP</i> target	5'- <u>G*C*A*U*U*G</u> *AAAGCCACU*C*C*U*C*CAGG[<u>TCA</u>]CAGCU*G*G*C* A*G*A*A -'3

Endogenous <i>APP</i> sequence Forward primer	5'- CATTGGACTCATGGTGG -3'
Endogenous <i>APP</i> sequence Reverse primer	5'- CAGCATCACCAAGGTG -3'
Fluorescein, non-fluorescent quencher labeled wild-type probe	5'- /56-FAM/TGTTGTCAT+G+GCGACAGT/3IABkFQ/ -3'
Hexachlorofluorescein, non-fluorescent quencher labeled mutant probe	5'- /5HEX/TGTT+GTCAT+A+G+CGACAGT/3IABkFQ/ -3'

General Biochemical Procedures.

Molecular-biology-grade bovine serum albumin (BSA), and RNase inhibitor were purchased from New England BioLabs. SDS-polyacrylamide gels were visualized with a Molecular Dynamics 9400 Typhon phosphorimager. Data were analyzed with Molecular Dynamics ImageQuant 5.2 software. All MALDI analyses were performed at the University of California, Davis Mass Spectrometry Facilities using a Bruker UltraFlex extreme MALDI TOF/TOF mass spectrometer. Oligonucleotide masses were determined with Mongo Oligo Calculator v2.08. Oligonucleotides for sequencing and PCR were purchased from Integrated DNA Technologies. RNA containing 2-aminopurine was purchased from Dharmacon and purified as described below. All other oligonucleotides were synthesized as described below.

Synthesis of oligonucleotides.

Chemical synthesis for all oligonucleotides was performed using an ABI 394 synthesizer. Pseudoisocytidine (piC) deoxyribonucleoside phosphoramidite and 8-azanebularine ribonucleoside phosphoramidite were purchased from Berry & Associates, Inc. Benner's base Z was purchased from FireBird Biomolecular Sciences, LLC as a deoxyribonucleoside phosphoramidite. All other bases were purchased from Glen Research. Nucleosides were incorporated during the appropriate cycle on a 0.2 μ mol scale; See Methods for sequences. Upon completion of the synthesis, columns were evaporated under

reduced pressure for 4 h. Deprotection of dZ containing oligos was initiated by treating the solid support with 1M 1,8-diazabicyclo[5.4.0]undec-7-ene in acetonitrile at room temperature for 10 h. After 10 h, the mixture was centrifuged at 13,000 rpm for 5 min. The supernatant was discarded, and the support was dried under reduced pressure before proceeding with cleavage. All oligonucleotides were cleaved from the solid support by treatment with 1:3 ethanol/ 30% NH₄OH at 55 °C for 12 h. The supernatant was transferred to a new screw-cap tube and evaporated under reduced pressure. Desilylation was performed by resuspending the pellets in anhydrous DMSO, and treating with 55% (v/v) Et₃N-3HF at room temperature overnight. To each reaction was added 75 mM sodium acetate in butanol. The oligonucleotides were then precipitated from a solution of 65% butanol at -70 °C for 2 h. The solution was centrifuged at 13,000 rpm for 20 min, supernatant was removed, and the pellet was washed twice with cold 95% ethanol. The RNA pellets were then desalted using a Sephadex G-25 column and purified as described below.

Purification of oligonucleotides.

Single-stranded RNA oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis and visualized by UV shadowing. Bands were excised from the gel, crushed and soaked overnight at 4 °C in 0.5 M NaOAc, 0.1% sodium dodecyl sulfate (SDS), and 0.1 mM EDTA. Polyacrylamide fragments were removed with a 0.2 µm filter, and the RNAs were precipitated from a solution of 75% EtOH at -70 °C for 4 h. The solution was centrifuged 13,000 rpm for 20 min and supernatant was removed. The RNA solutions were lyophilized to dryness, resuspended in nuclease-free water, and quantified by absorbance at 260 nm. dZ and 8-azanebularine containing oligonucleotides used for crystallography were additionally desalted by four rounds of concentration and subsequent addition of nuclease-free water in a 3000 NMWL Amicon Ultra 0.5 mL centrifugal filter. Oligonucleotide mass was confirmed by MALDI-TOF.

In vitro transcription of editing target RNA.

Target RNA was transcribed from a DNA template with the MEGAScript T7 Kit (ThermoFisher). DNA Digestion was performed using RQ1 RNase-free DNase (Promega). DNase treated RNA product was purified as described above.

Preparation of Duplex Substrates for Analysis of ADAR Deamination Kinetics.

Purified guide and transcribed RNA were added in a 10:1 ratio to hybridization buffer (180 nM transcribed RNA target, 1.8 μ M guide, 1X TE Buffer, 100 mM NaCl), heated to 95 °C for 5 min, and slowly cooled to room temperature.

Preparation of Duplex Substrates for Crystallography.

dZ and 8-azanebularine containing oligonucleotides were hybridized in a 1:1 ratio in nuclease-free water by heating the mixture to 95°C for 5 min and slow cooling to room temperature.

Protein Overexpression and Purification of ADAR2 full length constructs.

hADAR2 wild-type (hADAR2 WT) was expressed and as previously described.¹²⁹ Purification of hADAR2 was carried out by lysing cells in buffer containing 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 750 mM NaCl, 35 mM imidazole, and 0.01% Nonidet P-40 using a French press. Cell lysate was clarified by centrifugation (19,000 rpm for 1 hour). Lysate was passed over a 3 mL Ni-NTA column, which was then washed in three steps with 20 mL lysis buffer, wash I buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 750 mM NaCl, 35 mM imidazole, 0.01% Nonidet P-40), wash II buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1mM BME, 35 mM imidazole, 500 mM NaCl), and eluted with 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 400 mM imidazole, 100 mM NaCl. Fractions containing the target protein were pooled and concentrated to 30-80 μ M for use in biochemical assays. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels. Purified hADAR2 WT was stored in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 20% glycerol and 1 mM BME at -70 °C.

Protein Overexpression and Purification of ADAR1 p110.

MBP-tagged human ADAR1 p110 wild-type (hADAR1 p110 WT) was expressed as previously described.⁷⁴ Purification of hADAR1 p110 was carried out by lysing cells in buffer containing 50 mM Tris-HCl, pH 8.0, 5% glycerol, 5 mM BME, 1000 mM KCl, 50 μ M ZnCl₂, and 0.05% NP-40 using a microfluidizer. Cell lysate was clarified by centrifugation (18,000 rpm for 50 min). Lysate was passed over a 2 mL NEB amylose column, which was then washed in two steps with 50 mL lysis buffer, 100 mL wash I buffer (50 mM Tris-HCl, pH 8.0, 5% glycerol, 5 mM BME, 500 mM KCl, 50 μ M ZnCl₂, 0.01% Nonidet P-40), and eluted with 50 mM Tris-HCl, pH 8.0, 10% glycerol, 5 mM BME, 500 mM KCl, 50 μ M ZnCl₂, 0.01% Nonidet P-40, and 20 mM maltose. Fractions containing the target protein were pooled and concentrated to 5-50 μ M for use in biochemical assays. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels. Purified hADAR1 p110 WT was stored in 50 mM Tris-HCl pH 8.0, 400 mM KCl, 0.5 mM EDTA, 0.01% NP-40, 10% glycerol, and 1 mM tris(2-carboxyethyl)phosphine at -70 °C.

In Vitro Deamination Kinetics.

Deamination assays were performed under single-turnover conditions. Reactions with ADAR2 wild-type (WT) or E488Q contained 15 mM Tris-HCl pH 7.5, 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, 3 mM MgCl₂, 160 U/mL RNAsin, 1.0 μ g/mL, 0.8 nM RNA, and 2 nM enzyme. ADAR1 reactions contained 15 mM Tris-HCl pH 7.5, 4% glycerol, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% Nonidet P-40, 160 U/mL RNAsin, 1.0 μ g/mL yeast tRNA, 5 nM RNA and 50 nM ADAR1 p110 wild-type. Each reaction solution was incubated at 30 °C for 30 min before adding enzyme, and allowed to incubate at 30 °C for varying times prior to stopping with 190 μ L 95°C water and heating at 95 °C for 5 min. RT-PCR (Promega Access RT-PCR System) was used to generate cDNA from deaminated RNA. The resulting cDNA was purified using the DNA Clean & Concentrator kit from Zymo, and subjected to Sanger Sequencing using an ABI Prism 3730 Genetic Analyzer at the UC Davis DNA Sequencing Facility with the forward PCR primers. The sequencing peak heights were quantified in 4Peaks

v1.8. Each experiment was carried out in triplicate. The editing level for the corresponding zero time point was subtracted from each data point as a background subtraction.

Preparation of Duplex Substrates for Base-flipping Analysis.

PAGE purified top and bottom strands were annealed for a final concentration of 30 μ M edited strand, 45 μ M guide strand, 30 mM Tris-HCl, 6% glycerol, 120 mM KCl, 3 mM EDTA, 0.006% NP-40, and 0.6mM DTT. The mixture was heated to 95 °C for 5 min, and slowly cooled to room temperature.

Plate-based assay using a fluorescent RNA substrate to monitor base-flipping by ADAR2.

Fluorescence measurements were performed using a CLARIOstar microplate reader and a 384-well black bottom plate. Excitation was at 320 nm and fluorescence emission was scanned from 340 nm to 430 nm. Spectra were obtained for solutions of a total volume of 20 μ L containing 2.5 μ M RNA, with or without 10 μ M ADAR2, in 36 mM Tris-HCl at pH 7.5, 7% glycerol, 142 mM KCl, 3.6 mM EDTA, 0.007% NP-40, and 0.7 mM DTT at room temperature. The background fluorescence of the enzyme solution was subtracted from the spectrum of the complex, and the background fluorescence of the buffer alone was subtracted from the RNA.

Expression and Purification of hADAR2 deaminase domain (hADAR2d) for Crystallography.

Protein expression and purification was carried out according to previously reported protocols.¹⁰⁷ In short, BCY123 yeast cells were transformed with a pSc-ADAR plasmid encoding hADAR2d-WT using a high lithium transformation. Cells were plated on yeast minimal media minus uracil plates (CM-Ura+Glucose). An isolated colony was used to inoculate 15 mL of CM-Ura+Glucose media and incubated at 30°C with shaking at 300 rpm overnight. One liter of CM-Ura+Glycerol/Lactate yeast growth media was inoculated with 12 mL of starter culture. After 24 h, cells were induced by adding 110 mL sterile 30% galactose solution per liter of culture and protein was expressed for 6 h. Cells were collected by centrifugation at 5,000 \times g for 15 minutes and cell pellets were stored at -80 °C until purification. A cell pellet corresponding to a 2 liter growth was resuspended in cold Buffer A (20 mM Tris-HCl pH 8.0, 5% glycerol, 35 mM

imidazole, 0.01% Triton x 100, 1 mM β -mercaptoethanol) with 750 mM NaCl. Cells were lysed on a microfluidizer and lysate clarified by centrifugation at 39,000 $\times g$ for 45 min. Lysate was filtered through 0.45 μm filter and passed over a 1 mL Ni-NTA column and washed with 20 mL Buffer A + 750 mM NaCl. Loaded Ni-NTA resin was then washed with 60 mL Wash Buffer I (Buffer A + 400 mM NaCl) followed by 30 mL Wash Buffer II (Buffer A + 100 mM NaCl). Target protein was eluted with a 35-300 mM imidazole elution gradient in Wash Buffer II over 60 min at a flow rate of 1 mL/min. Fractions were analyzed on SDS-PAGE and fractions containing the target protein were pooled and loaded onto a 1 mL GE Healthcare Lifesciences Hi-Trap Heparin HP column at a flow rate of 0.5 mL/min. The loaded heparin resin was washed with 30 mL Wash Buffer II (Buffer A + 100 mM NaCl) and target protein was eluted with a gradient of 100-1000 mM NaCl in Wash Buffer II (Buffer A + 100 mM NaCl) over 30 min at a flow rate of 1 mL/min. Fractions containing the target protein were pooled and the 10xHis tag was cleaved with a 1:1 mass ratio of TEV to target protein for 2 h at room temperature. The cleaved protein was collected by passing the cleavage reaction over another Ni-NTA column at a flow rate of 0.5 mL/min. The flowthrough and wash were collected and dialyzed against SEC Buffer (20 mM Tris pH 8.0, 200 mM NaCl, 5% glycerol, and 1 mM β -mercaptoethanol). Dialyzed protein was then concentrated to 500 μL using an Amicon Ultra centrifugation filter with a 30,000 NMWL. Concentrated protein was loaded onto a GE Healthcare HiLoad 16/600 Superdex PG column equilibrated with SEC Buffer. Fractions containing purified protein were pooled and concentrated for crystallography trials using centrifugal filtration.

hADAR2d-GLII-dZ Complex Crystallization.

Crystals of the human ADAR2 deaminase domain WT (hADAR2d WT)/dZ complex were grown at room temperature using the sitting-drop vapor diffusion method. A 1 μL solution of 4.5 mg/mL hADAR2d WT and 100 μM of *GLII*-dZ 23mer RNA (1:1 hADAR2d WT: RNA molar ratio) were mixed with 1 μL of 0.1 M MES:NaOH pH 6.5 and 10% PEG 20,000. Crystals appeared within two weeks as a cluster of plates. A single crystal was broken off the cluster, soaked in a solution of mother liquor and 30% glycerol, and flash-cooled in liquid nitrogen. Data were collected *via* fine-phi slicing using 0.2° oscillations at beamline 24-

ID-E at the Advanced Photon Source at Argonne National Laboratory. X-ray diffraction data were collected to 2.50 Å resolution.

Processing and Refinement of Crystallographic Data.

Crystallographic data were processed using XDS¹⁶¹ and scaled using Aimless (CCP4 1994).¹⁶² The structure of a hADAR2d mutant E488Y bound to dsRNA (PDB ID: 6D06)¹⁵⁶ was used as a phasing model for molecular replacement using PHENIX.¹⁶³ The structure was refined using PHENIX¹⁶⁴ including TLS parameters, simulated annealing, and Zn coordinate restraints. The model was built and adjusted using COOT.¹⁶⁵ Ideal Zn-ligand distances were determined using average distances found in previously deposited ADAR structures and in the MetalPDB database.¹⁶⁶ As indicated by the Matthews coefficient (CCP4)¹⁶⁷, molecular replacement successfully placed two protein monomers one of which was bound to dsRNA as observed in previous structures.¹² In each protein monomer (chains A and D), the C-terminal proline (Pro701) was disordered and thus not included in the model. Geometry restraints from previous hADAR2d-dsRNA structures were used to model in the hydrated 8-azanebularine.¹⁰⁷ Geometry restraints for the dZ base were calculated by Monomer Library Sketcher (CCP4).¹⁶² The structure refined to a final R_{factor} and R_{free} of 17.30% and 22.60%, respectively. Atomic coordinates and structure factors have been deposited to the Protein Data Bank (PDB ID: 7KFN)

Density Functional Theory Calculations of Benner's Base Z and Cytosine.

Structures of Benner's Base Z and cytosine were fully optimized using the M06-2X/6-311+G(2d,p) method¹⁶⁸ as implemented in GAUSSIAN16.¹⁶⁹ Frequency analysis confirmed the structures to be minima. Partial charges were computed using the CHelpG model.¹⁷⁰

Directed Editing on the endogenous RAB7A Target in HEK293T Cells.

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11995-065) with 10% fetal bovine serum (FBS) (Thermo Fisher, 26140-087) and additionally supplemented with 1X antibiotic-antimycotic (Thermo Fisher, 15240-062) at 37 °C, 5 % CO₂. Once cells reached 70-90%

confluency, cells were seeded into 96 well plates (6.4×10^3 cells per well). After 24 h, cells were co-transfected with 500 ng ADAR plasmid and 50 nM guide oligonucleotide using Lipofectamine 2000 (Thermo Fisher, 11668-019). After incubation of transfection reagent, plasmid, and guide oligonucleotide in Opti-MEM Reduced Serum Media (Thermo Fisher, 31985-062), the solution was added to designated wells and incubated at 37 °C, 5 % CO₂. After 48 h, total RNA was isolated using RNAqueous Total RNA Isolation Kit (Thermo Fisher, AM1912) and DNase treated with RQ1 RNase-free DNase (Promega, M6101). Nested RT-PCR was performed in triplicate using Access RT-PCR kit (Promega, A1280) for 20 cycles and then followed by Phusion Hot Start II DNA Polymerase (Thermo Fisher, F-549L) for the second PCR of 30 cycles with target specific primers. PCR product was purified by agarose gel and QIAquick Gel Extraction kit (Qiagen, 28706). Product was submitted for Sanger Sequencing and sequence traces were analyzed by 4Peaks (Nucleobytes) to quantify percent editing.

Directed Editing on the endogenous WT APP target in human retinal pigment epithelium cells (ARPE-19).

Spontaneously arising human retinal pigment epithelium cell line (ARPE-19) carrying the wild-type *APP* gene was obtained from ATCC (ATCC® CRL-2302™, Lot. 70013110). Briefly, 1.5×10^5 ARPE-19 cells per well were seeded in a 12 well plate 24 h before transfection. Transfection was performed with 100 nM guide strand and Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions (at a ratio of 1:2, 1 μ g guide strand to 2 μ l Lipofectamine 2000). A non-transfection (NT), a mock transfection and a scrambled guide strand were used as negative controls. Total RNA was extracted from cells 48 h after transfection using the Direct-zol RNA MicroPrep (Zymo Research) kit according to the manufacturer's instructions, and ~500 ng of total RNA was used to prepare cDNA using the Maxima reverse transcriptase kit (Thermo Fisher) according to the manufacturer's instructions, with a combination of random hexamer and oligo-dT primers. Digital droplet PCR (ddPCR) was performed using 1 μ l of a 10X dilution of cDNA. The ddPCR assay for absolute quantification of nucleic acid target sequences was performed using BioRad's QX-200 Droplet Digital PCR system. Diluted cDNA obtained from RT-PCR was used in a total mixture of 21 μ l of reaction mix, including the ddPCR Supermix for Probes no dUTP (Bio Rad), a Taqman

SNP genotype assay with forward and reverse primers combined with the gene-specific probes. The PCR mix was filled in the middle row of a ddPCR cartridge (BioRad) using a multichannel pipette. The replicates were divided by two cartridges. The bottom rows were filled with 70 μ l of droplet generation oil for probes (BioRad). After the rubber gasket replacement, droplets were generated in the QX200 droplet generator. An oil emulsion volume of 42 μ l from the top row of the cartridge was transferred to a 96-wells PCR plate. The PCR plate was sealed with tin foil for 4 s at 170 °C using the PX1 plate sealer, followed by the following PCR program: 1 cycle of enzyme activation for 10 min at 95 °C, 40 cycles denaturation for 30 sec at 95 °C and annealing/extension for 1 min at 53.8 °C, 1 cycle of enzyme deactivation for 10 min at 98 °C, followed by storage at 8 °C. After PCR, the plate was read and analyzed with the QX200 droplet reader.

Directed Editing on the endogenous IDUA target in mouse primary liver fibroblasts.

Homozygous *Idua-W392X* mice on a C57BL/6 background were obtained under a license agreement from the UAB Research Foundation (UABRF). Mouse primary liver fibroblasts were isolated from homozygous *Idua-W392X* mice using a Liver Dissociation kit, mouse (Miltenyi Biotech, 130-105-807). Cells were seeded with 3.0×10^5 cells per well in a collagen coated 6 well plate 24 h prior to transfection. Transfection, RNA extraction, and analysis were carried out using the protocol described above, with the following changes: Direct-zol RNA MiniPrep Kit (Zymo Research) was used for RNA extraction, 1 μ l of a 3X dilution of cDNA was used as a template for ddPCR, and the annealing/extension step of ddPCR was carried out at 63.8 °C.

CHAPTER 4

Enabling RNA Editing at Disfavored Sites through a Purine:Purine Pairing Interaction

Introduction.

Endogenous human enzymes can be programmed to site-specifically alter disease-causing sequences in nucleic acids through the use of complementary guide oligonucleotides. Editing via the Adenosine Deaminase Acting on RNA (ADAR) family of enzymes deaminates adenosine (A) to inosine (I) in double-stranded RNA. Inosine is recognized by cellular machinery as guanosine, effectively replacing adenosine with guanosine in RNA. This strategy has been shown to restore functional proteins in cells, and to edit with up to 50% efficiency in non-human primates.^{68,133,171}

Thus far, the field has worked within the natural substrate preferences of ADARs to inform what sequences to target. ADARs edit most effectively within a 5'-UAG-3' sequence context, with the 5' position having a more prominent influence on editing efficiency than the identity of the 3' base.³⁷ The structural basis for this preference was rationalized via the crystal structure of the catalytic domain of ADAR2 (ADAR2d) bound to its double-stranded RNA (dsRNA) target.²⁵ While a uracil in the 5' position relative to the edit site is accommodated by the enzyme, modeling a guanosine into this position indicated that this could induce a steric clash between the exocyclic amine of the guanosine and a close-approaching glycine residue (G489) (Figure 4.1).

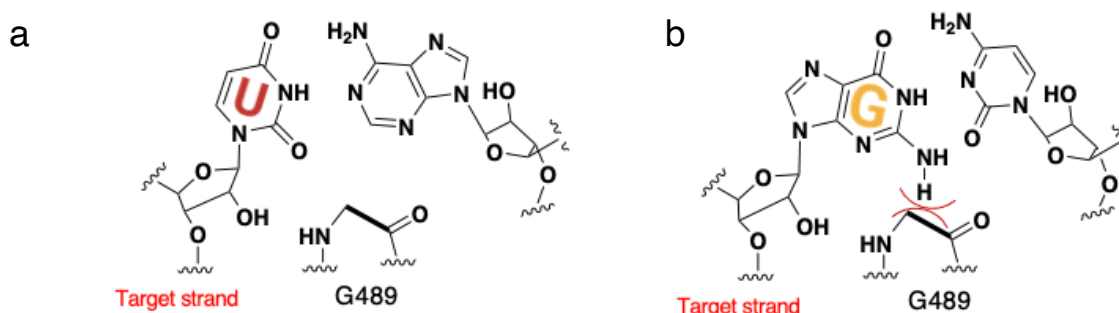


Figure 4.1. A line-angle representation of the conformation surrounding the 5' base as observed in the crystal structure of ADAR2d bound to dsRNA.¹⁰⁷ A) Uracil in the position 5' to the edited adenosine. B) Guanosine modeled into the position 5' to the edited base shows a potential steric clash with G489.

Previous *in vitro* studies with fusion proteins of the ADAR deaminase domain had shown that altering the guide strand to include a guanosine across from the 5' G (what we have defined as the -1 position of the guide oligonucleotide) helped to increase editing yields within this context.¹⁵⁷ Because this strategy had not yet been demonstrated with a full length or wild-type protein, it was unclear if this observation was an inherent property of ADARs or an artifact of the fusion protein construct. Furthermore, the mechanistic basis for this improvement was not explained. We sought to define how the identity of the -1 base might influence the rate of deamination by ADAR. Moreover, we wanted to understand the structural and mechanistic basis of any observed differences in activity to aid in our design of more effective guide strands for sequences that are resistant to ADAR editing.

Results.

ADAR2 Editing of difficult target sites is enhanced by a G:G pair

To determine if the base paired across from the 5' guanosine influences editing by full-length ADAR2, we designed guide strands containing each canonical base (A, C, G, U) in the -1 position of the guide (Figure 4.2 A). These oligonucleotides were designed to target a mutant of the *IDUA* (α -1-iduronidase) sequence with a guanosine 5' to the edit site as opposed to the wild-type 5' uracil. This sequence was used as a model substrate due to its relatively rapid rate of deamination by ADAR2 in our previous work.¹³¹ These results showed that the identity of the -1 base had a prominent effect on the observed rate of reaction. Interestingly, when a purine:purine pair (A or G across from the 5' G) was present, a higher rate of editing was supported as opposed to a purine:pyrimidine pairing (Figure 4.2 B-C). In fact, the slowest observed rate was found for the canonical G:C pair, supporting a rate of $k_{\text{obs}} = 0.010 \pm 0.003$. In comparison, the most efficiently edited substrate, the G:G pair, was 30 times faster at $k_{\text{obs}} = 0.3 \pm 0.1$ (Table 4.1).

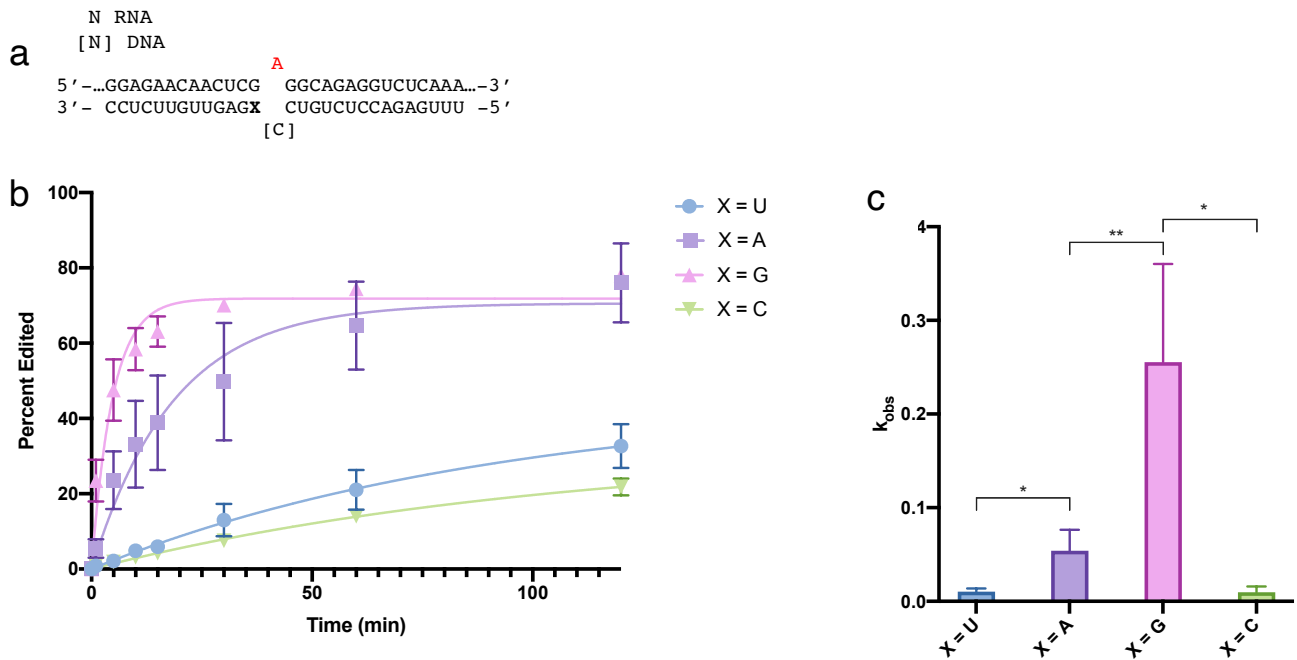


Figure 4.2. ADAR2 deamination with varying -1 base. A) The mutant *IDUA* substrate used for *in vitro* deaminations under single turnover conditions. Reactions with ADAR2 were carried out with 10 nM RNA and 100 nM enzyme. B) Reaction progress curve for duplexes containing each canonical base (X) in the -1 position. C) A plot of the observed rate of reaction (k_{obs}) for each substrate.

N RNA
[N] DNA

5' -...GGAGAACAACUCY ^AGGCAGAGGUCUCAA...-3'
3' - CCUCUUGUUGAGX CUGUCUCCAGAGUUU -5'
[C]

enzyme	Y:X	k_{obs} min ⁻¹ ^a	k_{rel} ^b
ADAR2 ^a	G:G	0.3 ± 0.1	33
	G:C	0.009 ± 0.006	1
	G:A	0.05 ± 0.02	5.5
	G:U	0.010 ± 0.003	1.1

Table 4.1. Rates of *in vitro* deamination for 100 nM ADAR2 acting on 10 nM mutant *IDUA* substrate under single turnover conditions. Y:X indicates the base pairing in the -1 position.^a Data were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{obs} \cdot t)]$. ^b $k_{rel} = k_{obs}$ for analog / k_{obs} for cytosine.

G:G editing efficiency approaches that of ADAR's preferred substrate

Having observed a 30-fold rate enhancement when changing from the canonical G:C pair to the G:G pair, we wanted to understand how the rate of editing for a G:G pair compared to ADAR's preferred sequence context: 5'-UA-3'. Using a mutant (5'-GA-3') and wild-type (5'-UA-3') *IDUA* sequence, we

compared the *in vitro* rate of deamination for either substrate. These deaminations were performed at 1 nM duplex RNA and 10 nM ADAR2, a lower concentration than used previously due to the rapid rate of deamination of the wild-type *IDUA* sequence. The G:G substrate again demonstrated a clear improvement over the G:C pair, and approached the rate of the preferred 5' U substrate. While the G:C substrate was deaminated at $k_{obs} = 0.03 \pm 0.04$; $k_{rel} = 0.06$, the G:G pair gave rise to a rate of 0.29 ± 0.02 ($k_{rel} = 0.63$) (Figure 4.3, Table 4.2). Each of these rates was contrasted to that of the preferred substrate under the same conditions ($k_{obs} = 0.47 \pm 0.03$) at a relative rate (k_{rel}) of 1.0 (Table 4.2). These results showed the impact of the G:G substitution. This single pair enhanced the rate from 2% to 63% that of the preferred 5'-UA-3' sequence. This puts into context the benefit of this pair, that we may now achieve levels of editing in the most challenging sequence context (5'-GA) that rival that of the preferred sequence.

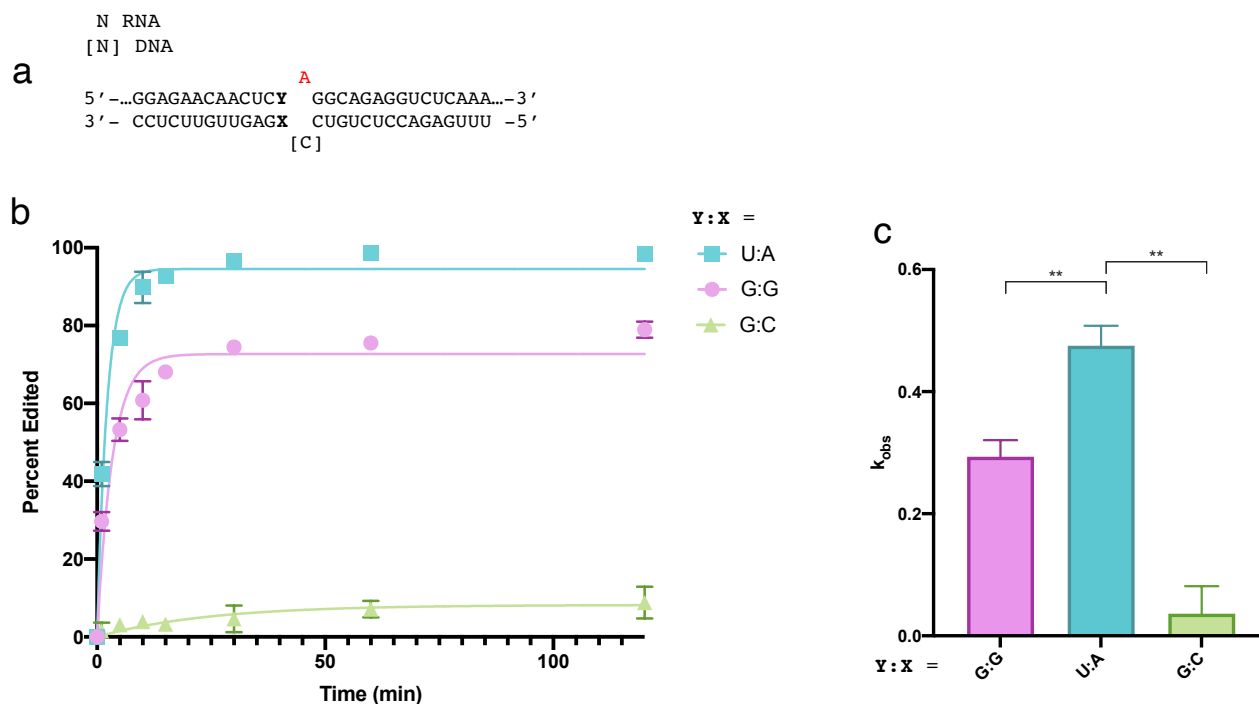


Figure 4.3. Comparison of ADAR2 preferred and non-preferred substrates. A) The mutant *IDUA* substrate used for *in vitro* deaminations under single turnover conditions. Y:X is defined in the subsequent figure legends. Reactions with ADAR2 were carried out with 1 nM RNA and 10 nM enzyme. B) Reaction progress curve for duplexes containing each canonical base in the -1 position. C) A plot of the observed rate of reaction (k_{obs}) for each substrate.

N	RNA		
[N]	DNA		
		A	
	5' -...GGAGAACAACUCY	GGCAGAGGUCUCAAA...-3'	
	3' - CCUCUUGUUGAGX	CUGUCUCCAGAGUUU -5'	
		[C]	

enzyme	Y:X	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b
ADAR2	G:G	0.29 ± 0.02	10
	G:C	0.03 ± 0.04	1
	U:A	0.47 ± 0.03	16

Table 4.2. Rates of *in vitro* deamination for 10 nM ADAR2 acting on 1 nM mutant *IDUA* substrate under single turnover conditions. Y:X indicates the base pairing in the -1 position. ^a Data were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{\text{obs}}t)]$. ^b $k_{\text{rel}} = k_{\text{obs}}$ for analog / k_{obs} for cytosine.

A G:G pair supports editing in disease-relevant sequence contexts

Due to the sizeable increase in editing afforded to our model substrate when a G:C pair was replaced with a G:G pair, we sought to determine if this observation could be used to enhance editing for disease-relevant substrates containing a 5' G. We chose to target mutations in the Methyl-CpG binding Protein 2 (*MECP2*) sequence. Mutations to *MECP2* cause the neurological disease Rett Syndrome. While there are many mutations to *MECP2* seen in human populations, the most common adenosine point mutations are within a 5' G sequence context.¹⁷² Additionally, recent *in vivo* studies have shown that RNA editing can result in up to 50% correction of a nonsense mutation in the *MECP2* sequence, equating to 50% restoration in protein levels.¹³³ For this reason, the *MECP2* sequence is a promising target that would benefit from enhanced editing at adenosines flanked by a 5' G. We chose two T-to-C mutations, R168X and R255X, to test the effect of the G:G pair (Figure 4.4 A, 4.5 A). The disease-associated C to T mutations leads to UGA termination codons in the *MECP2* transcript. While ADAR editing is not capable of restoring the wild-type sequence, it can convert each termination codon to one for tryptophan leading to expression of full length R168W or R255W mutant MeCP2 protein. These substrates were hybridized to a complementary 29 mer guide RNA containing either a cytidine or guanosine at the -1 position and deaminated at a concentration of 10 nM, with 100 nM ADAR2 enzyme. For the R168X mutation, the observed rate of reaction was

enhanced >50 times when replacing the canonical G:C pair with a G:G pair (Figure 4.4 B-C). This enhancement in rate is even greater than the 30-fold increase seen for the model substrate (Table 4.3).

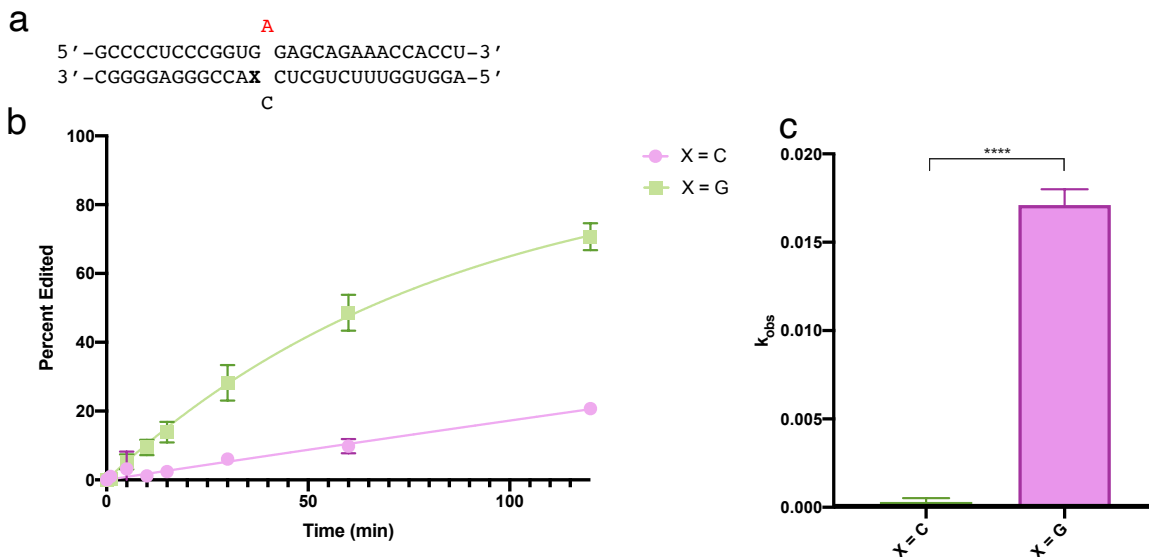
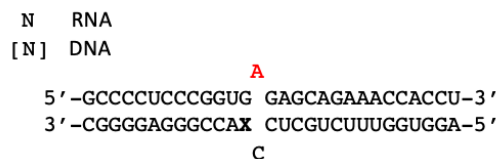


Figure 4.4. *In vitro* deamination of the R168X mutation in *MECP2* under single turnover conditions. A) Partial sequence of the target, and sequence of the guide RNA with varying -1 position (X). B) Progress curve of the deamination of substrates with either base in the -1 position of the guide. C) A comparison of the observed rate of reaction (k_{obs}) for either guide RNA showing a 56-fold enhancement in the rate of reaction for the X = G substrate.



enzyme	X	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b
	G	0.0171 ± 0.0008	57
	C	0.0003 ± 0.0002	1.0

Table 4.3. Rates of *in vitro* deamination for 100 nM ADAR2 acting on 10 nM R168X *MECP2* substrate under single turnover conditions. X indicates the base pairing in the -1 position. ^aData were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{\text{obs}} \cdot t)]$. ^b $k_{\text{rel}} = k_{\text{obs}}$ for analog / k_{obs} for cytosine.

The R255 wild-type substrate was also deaminated *in vitro* at a concentration of 10 nM with 100 nM ADAR2. The guide strand again contained either the canonical pair or the G:G pair in the -1 position. The complementary -1 guide containing cytidine was deaminated at a rate of $k_{\text{obs}} = 0.015 \pm 0.009$, while the guide containing guanosine had an observed rate of $k_{\text{obs}} = 0.117 \pm 0.002$ (Figure 4.5, Table 4.4). Therefore, the G:G pair provided a 7.8-fold increase in the rate of reaction for this substrate. This demonstrated that

the enhancement in rate was not sequence-specific and that making a single nucleobase change across from the 5' guanosine could induce between a 7.8- and 57-fold increase in the rate of deamination.

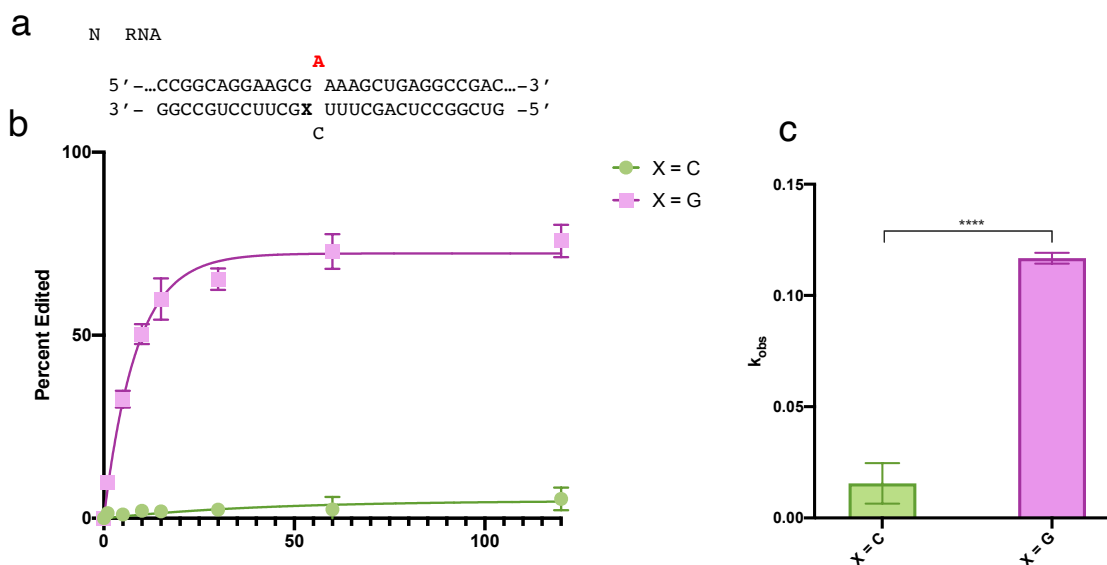


Figure 4.5. *In vitro* deamination of the R255 site in *MECP2* under single turnover conditions. A) Partial sequence of the target, and sequence of the guide RNA with varying -1 position (X). B) Progress curve of the deamination of substrates with either base in the -1 position of the guide. C) A comparison of the observed rate of reaction (k_{obs}) for either guide RNA showing an 8-fold enhancement in the rate of reaction for the X = G substrate.

N RNA

[N] DNA

5' -...CCGGCAGGAAGCG **A** AAAGCUGAGGCCGAC...-3'

3' - GGCCGUCCUUCG**X** UUUCGACUCCGGCUG -5'

C

enzyme	X	$k_{obs} \text{ min}^{-1a}$	k_{rel}^b
	G	0.117 ± 0.002	7.8
	C	0.015 ± 0.009	1.0

Table 4.4. Rates of *in vitro* deamination for 100 nM ADAR2 acting on 10 nM R255 *MECP2* substrate under single turnover conditions. X indicates the base pairing in the -1 position. ^aData were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{obs} \cdot t)]$. ^b $k_{rel} = k_{obs}$ for analog / k_{obs} for cytosine.

Bystander edits within 5' G target sequences are eliminated by 2'-O-Methyl modification

Although we observed an encouraging enhancement in editing of both *MECP2* substrates, we also observed off-target edits within the bystander region complementary to the guide strand in both the *IDUA* and *MECP2* substrates (Figure 4.6). Previous studies had showed that chemical modification of the guide strand by 2'-O-methylation could mitigate bystander editing.¹⁵⁸ Therefore, we examined how bystander editing of the *MECP2* R255 substrate was affected by chemical modification. The guide strand for R255

(Figure 4.5 B) was modified to include 2'-O-methyl modifications in every position except for the three bases across from the edit site (Figure 4.7 A). Compared to the unmodified guide, the 2'-O-methylated guide showed less resolution between the G:C and G:G pair. However, the 2'-O-methyl G:G guide still offered a 3-fold enhancement in the rate of editing, while nearly abolishing off-target edits (Figure 4. 7).

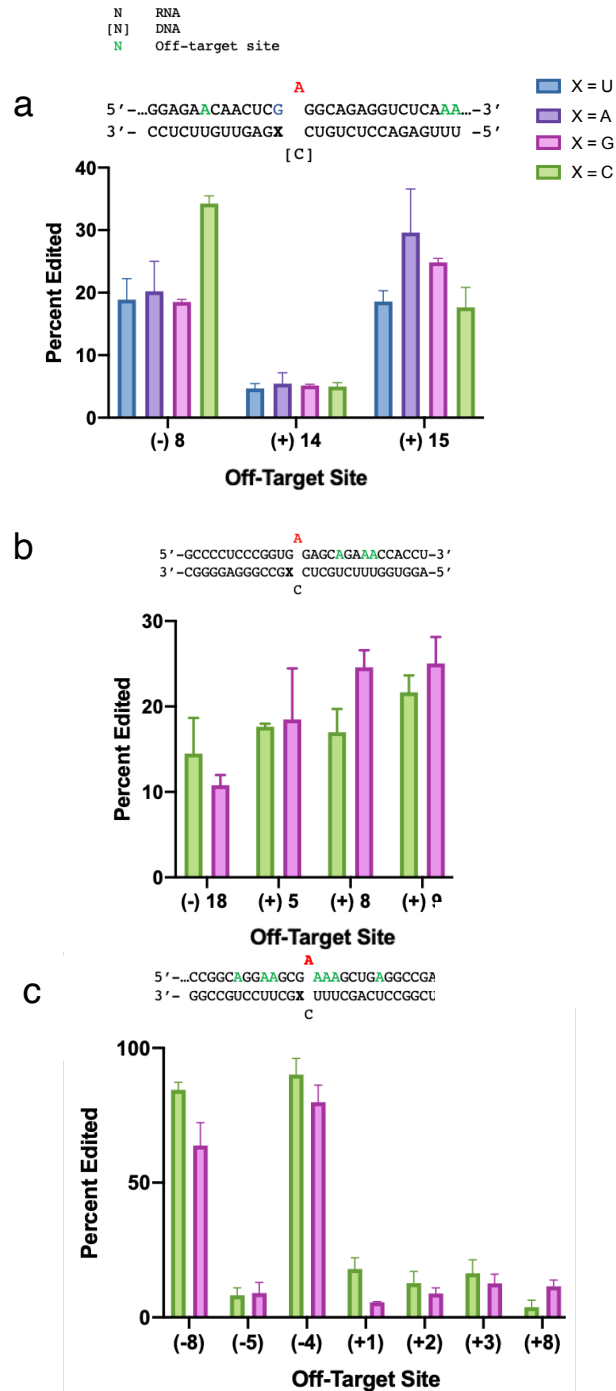


Figure 4.6. Off-target editing from the *in vitro* deamination of 10 nM *IDUA* and *MECP2* substrates with 100 nM ADAR2. Substrates in each plot: A) *IDUA* B) *MECP2* R168X, C) *MECP2* R255.

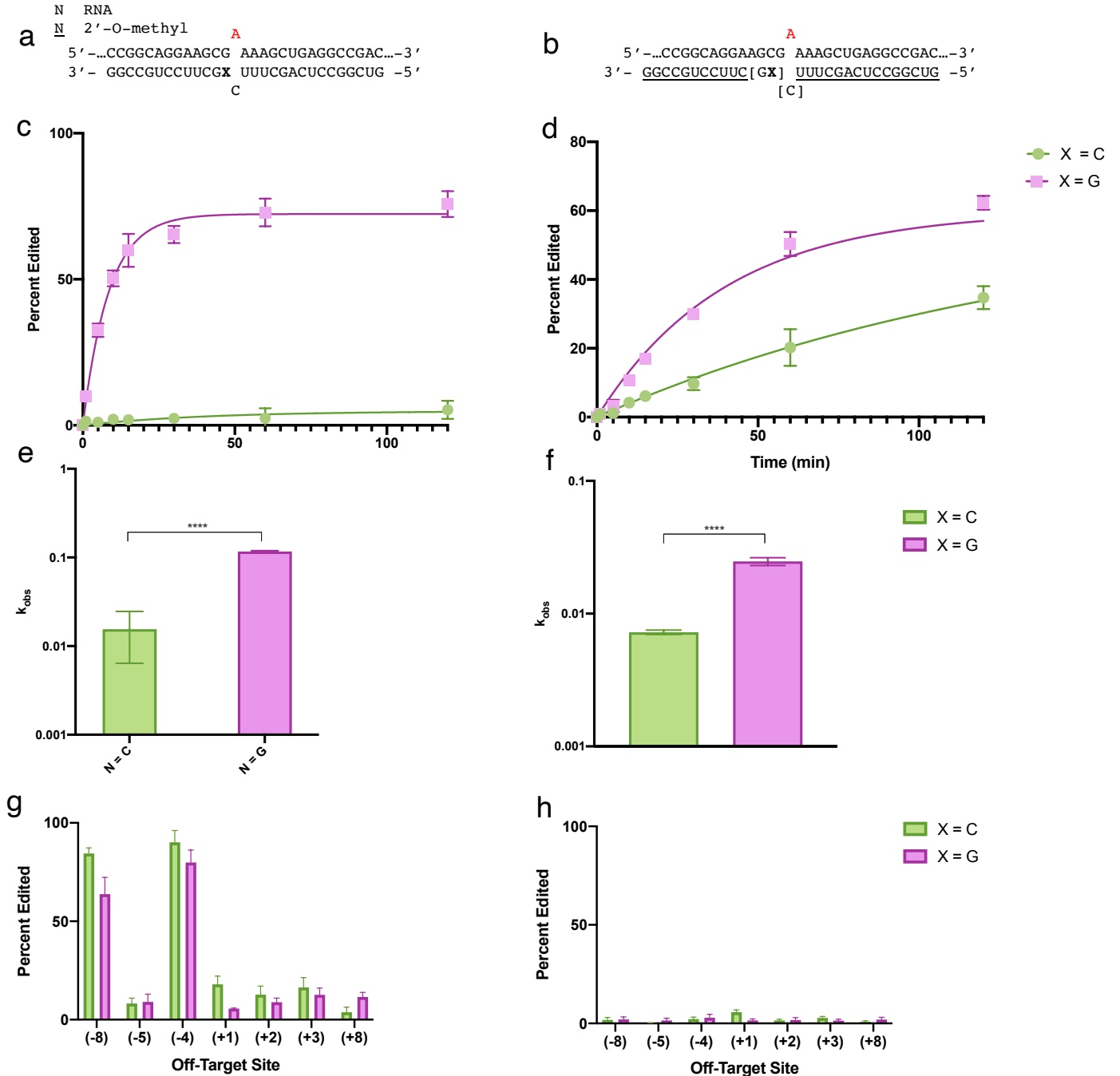


Figure 4.7. *In vitro* deamination of 10 nM of the R255 site in *MECP2* with 100 nM ADAR2 under single turnover conditions. A) Partial sequence of the target RNA, and sequence of the guide RNA with varying -1 position (X). B) Partial sequence of the target RNA, and sequence of the 2'-O-methyl guide oligonucleotide with varying -1 position (X). C) A comparison of the observed rate of reaction (k_{obs}) for either guide RNA showing an 8-fold enhancement in the rate of reaction for the N = G substrate. D) A comparison of the observed rate of reaction (k_{obs}) for either guide RNA showing an 8-fold enhancement in the rate of reaction for the N = G substrate with 2'-O-methyl modification. G) Off-target editing within the bystander region of the R255 *MECP2* target RNA. G) Off-target editing within the bystander region of the R255 *MECP2* target RNA across from the 2'-O-methyl modified guide oligonucleotide.

At least one dsRBD is essential for efficient ADAR2 editing of G:G-containing substrates

The enhanced *in vitro* reaction rate stimulated by G:G containing substrates inspired the use of X-ray crystallography to examine the specific contacts made by the G:G pair. Previously, X ray crystallography of ADAR2 deaminase domain (ADAR2d) and ADAR2 with a single double-stranded RNA binding domain (ADAR2 RD) has been enabled by high-affinity binding substrates.^{12,25,131,156} Typically, substrates that have been identified as high-affinity binders also have a relatively rapid rate of deamination *in vitro*. Therefore, to see if either ADAR2d E488Q or ADAR2 RD E488Q would be good candidates for crystallography, we tested the rate of *in vitro* reaction for either enzyme with a G:G pair-containing RNA. The E488Q mutation was used in both constructs due to its ability to promote efficient reaction, especially for substrates in non-ideal sequence contexts.⁸⁴ The rate of deamination for each enzyme was compared against the preferred sequence context (5'-UA-3'), which is known to crystallize with both constructs. In addition, each was assessed against a 5' guanosine-containing substrate matched against a canonical cytidine, to illustrate the rate enhancement from the G:G pair in these conditions. ADAR2d E488Q showed a 13-fold rate enhancement for the G:G substrate as opposed to G:C (Figure 4.8, Table 4.5). However, the rate of the G:G substrate was substantially slower than the preferred U:A substrate (G:G $k_{\text{obs}} = 0.004 \pm 0.002$, U:A $k_{\text{obs}} = 0.3 \pm 0.1$) (Figure 4.8 B-C, Table 4.5). Therefore, it was determined that ADAR2d E488Q likely would not be a promising candidate for crystallography due to its poor interaction with the G:G-containing substrate due to its lack of dsRBDs. ADAR2 RD E488Q was next used to deaminate the three substrates: G:G, U:A, and G:C. With the ADAR2 RD E488Q enzyme, the rate of deamination for the G:G substrate ($k_{\text{obs}} = 1.83$ $k_{\text{rel}} = 7$) again demonstrated a significant enhancement from the G:C pair ($k_{\text{obs}} = 0.248 \pm 0.002$, $k_{\text{rel}} = 1$) (Figure 4.8 D-E). However, the deamination rate of G:G now approached that of the preferred substrate ($k_{\text{obs}} = 3 \pm 1$, $k_{\text{rel}} = 12$) (Figure 4.8 D-E, Table 4.5). ADAR2 RD E488Q has been crystallized bound to dsRNA in the U:A sequence context. Thus, due to the similarity in activity for ADAR2 E488Q with G:G versus U:A it was determined that this construct was a promising candidate for X-ray crystallography with a G:G-containing substrate. In fact, collaborators Agya Karki, Xander Wilcox, and

Dr. Andrew Fisher were able to obtain a crystal structure of ADAR2d E488Q bound to a G:G containing RNA at high enough resolution to view the specific conformation of the G:G pair. This structure is in the final stages of refinement.

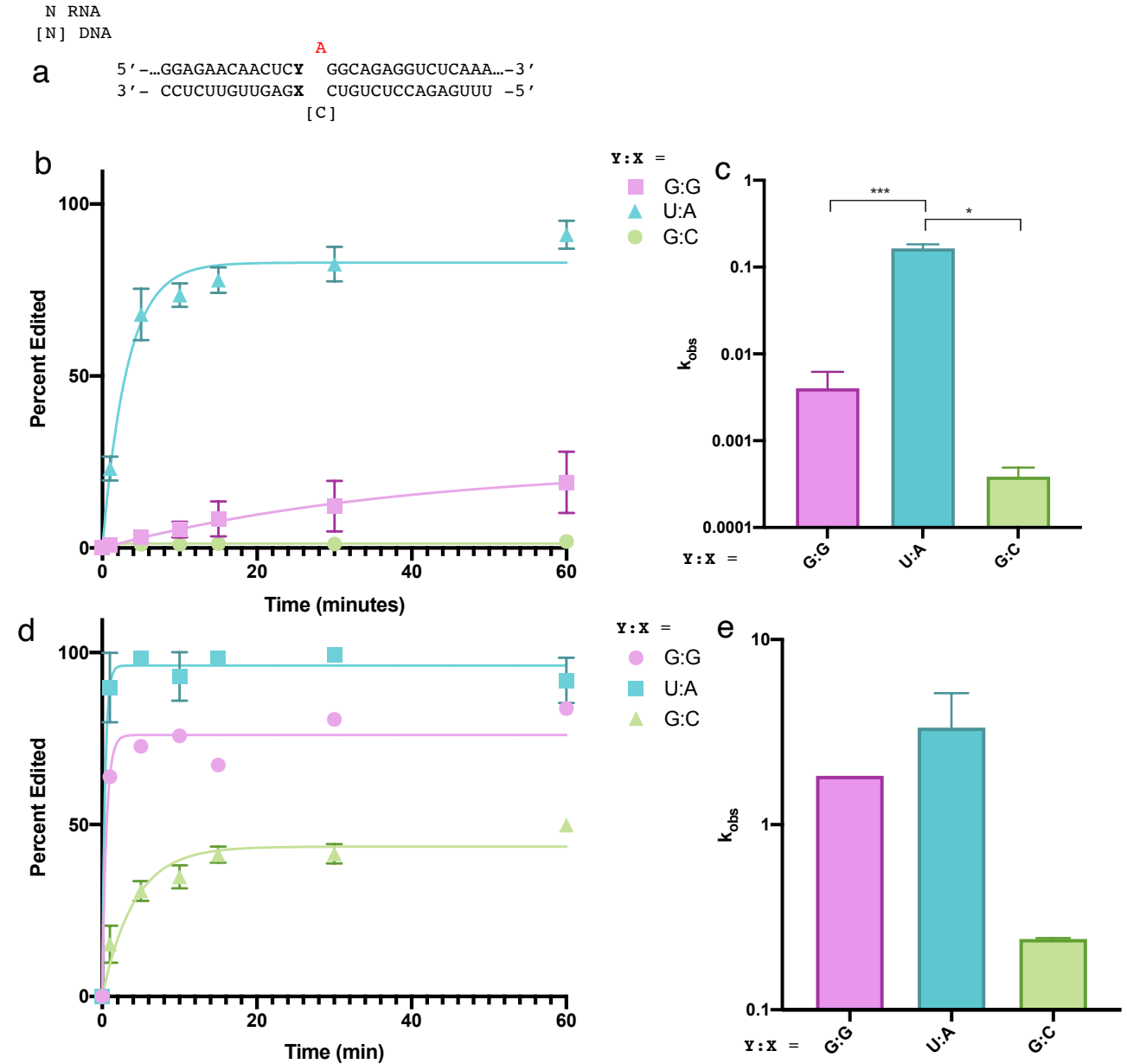


Figure 4.8. *In vitro* deamination of 10 nM of the mutant *IDUA* substrate with 100 nM ADAR2d E488Q or ADAR2 RD E488Q under single turnover conditions. A) Partial sequence of the target RNA, and sequence of the guide RNA with varying bases in the -1 position (Y:X). B) Reaction progress curves for each sequence context in reaction with ADAR2d E488Q. C) A comparison of the observed rate of reaction (k_{obs}) for each guide RNA with ADAR2 RD E488Q. D) Reaction progress curves for each sequence context in reaction with ADAR2d E488Q. E) A comparison of the observed rate of reaction (k_{obs}) for each guide RNA with ADAR2 RD E488Q.

enzyme	Y:X	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b	enzyme	Y:X	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b
ADAR2d E488Q	G:G	0.004 ± 0.002	13	ADAR2 RD E488Q	G:G	1.83	7
	G:C	0.0003 ± 0.001	1		G:C	0.248 ± 0.002	1
	U:A	0.3 ± 0.1	1000		U:A	3 ± 1	12

Table 4.5. Rates of *in vitro* deamination for 100 nM ADAR2d E488Q or ADAR2 RD E488Q acting on 10 nM mutant *IDUA* substrate under single turnover conditions. Y:X indicates the base pairing in the -1 position. ^aData were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{\text{obs}} \cdot t)]$. ^b $k_{\text{rel}} = k_{\text{obs}}$ for analog / k_{obs} for cytosine.

Purines are favored by ADAR1 in the guide strand -1 Position

While engineering guide oligonucleotides we typically test initial designs with ADAR2, due to the ease of purification and *in vitro* assays as compared to ADAR1. However, ADAR1 is more ubiquitously expressed than ADAR2, and depending on the target cell type it would be important for guide strands to promote editing with either catalytically active ADAR.^{13,86} Due to the similarity in the sequence and proposed structure of ADAR2d and ADAR1d in the region that would be in close proximity to the residue likely involved in determining sequence preference at the 5' position (G489 in ADAR2 and G1009 in ADAR1) we hypothesized that ADAR1 would also exhibit an enhancement in rate for 5' guanosine substrates in a G:G pair sequence context.¹⁵⁵ Therefore, we tested ADAR1 against the mutant 5'-GA-3' *IDUA* substrate with each canonical base paired across the -1 guanosine (Figure 4.9 A). These results demonstrated that the identity of the -1 base had a prominent effect on the observed rate of reaction of ADAR1, similarly to ADAR2. Interestingly, ADAR1 demonstrated a similar rate for either purine:purine pair (A or G across from the 5' G) as opposed to favoring a the G:G pair as had been observed with ADAR2 (Figure 4.9 B-C). The rate of deamination for A:G was $k_{\text{obs}} = 0.027 \pm 0.003$ ($k_{\text{rel}} = 13$) and for G:G was $k_{\text{obs}} = 0.020 \pm 0.009$ ($k_{\text{rel}} = 10$) (Table 4.6). However, ADAR1 also exhibited a strong preference for either purine:purine pair over both purine:pyrimidine pairs (Figure 4.9 B-C, Table 4.6). Thus, the purine:purine pair strategy represents a promising approach for either catalytically active ADAR enzyme.

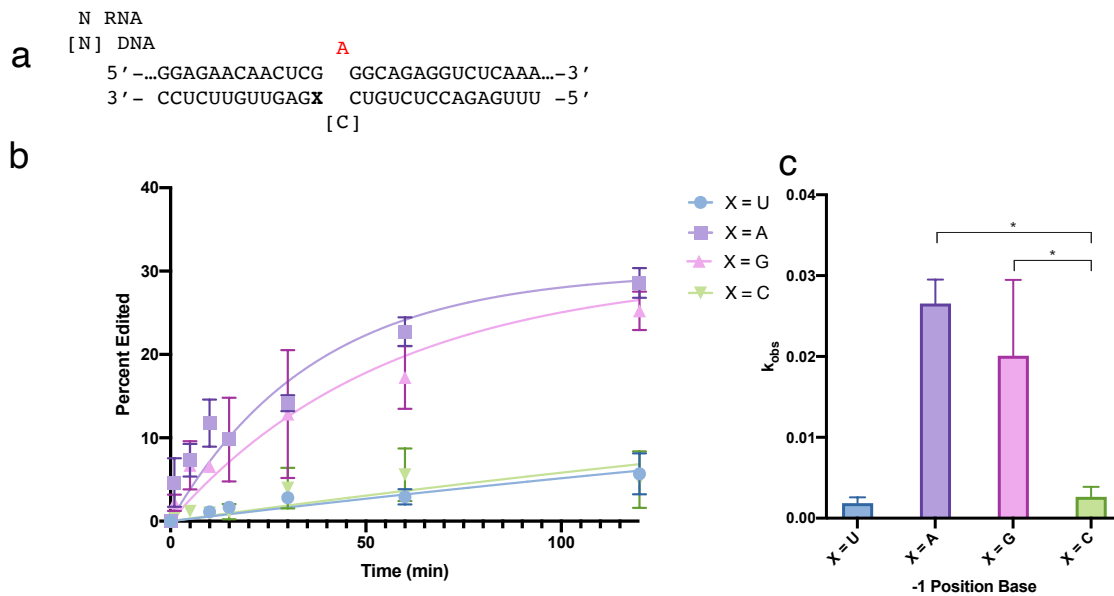
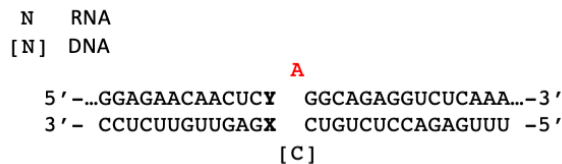


Figure 4.9. A) The mutant *IDUA* substrate used for *in vitro* deaminations under single turnover conditions. Reactions with ADAR1 were carried out with 10 nM RNA and 250 nM enzyme. B) Reaction progress curve for duplexes containing each canonical base in the -1 position. C) A plot of the observed rate of reaction (k_{obs}) for each substrate.



enzyme	Y:X	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b
ADAR1 p110	G:U	0.0018 ± 0.0007	2
	G:A	0.0076 ± 0.0007	8.4
	G:G	0.008 ± 0.003	8.8
	G:C	0.0009 ± 0.0004	1

Table 4.6. Rates of *in vitro* deamination for 250 nM ADAR1 p110 acting on 10 nM mutant *IDUA* substrate under single turnover conditions. Y:X indicates the base pairing in the -1 position. ^a Data were fitted to the equation $[P]_t = 0.3*[1-\exp(-k_{\text{obs}} \cdot t)]$. ^b $k_{\text{rel}} = k_{\text{obs}}$ for analog / k_{obs} for cytosine.

Guanosine in the -1 position enhances editing of an endogenous RNA in human cells.

Our *in vitro* findings suggested that directed A-to-I editing of 5' guanosine by ADARs could be enhanced with the use of a guide oligonucleotide containing guanosine in the -1 position. Therefore, to determine if the G:G pair could provide editing enhancement in human cells, we used a 39-nt guide oligonucleotide to direct editing by ADAR2 wild-type to endogenous β -actin mRNA present in HEK293T cells. This target mRNA was used previously in directed-editing experiments, and we chose an alternate editing site containing a 5'-GA-3' flanking sequence.^{37,65,156} The guide strand contained 2'-O-methyl and

phosphorothioate modifications to increase cellular stability and prevent bystander editing,^{61,157} as well as either cytidine or guanosine in the -1 position opposite the 5' guanosine (Figure 4.10 A).^{59,156–158} HEK293T cells were transfected with either C- or G- containing guide oligonucleotide, and a plasmid for ADAR2 overexpression. Transfection with the -1 C guide oligonucleotide resulted in editing levels of 8.6 % ± 0.9 %, while transfection of the -1 G guide oligonucleotide increased target editing levels to 13 % ± 2 % (Figure 4.10 B). This single nucleotide change induced an approximately 1.5-fold increase in editing of an endogenous target. However, transfection of a scrambled sequence oligonucleotide also led to levels of editing comparable to the -1 C guide oligonucleotide (6 % ± 0.9 %), suggesting that the editing seen under this condition is not guide-dependent and minor amounts of editing already occur at this position. Therefore, there may be secondary structure already existing in the mRNA at this position which impedes the guide from forming a duplex and could explain relatively low levels of editing at this site. To determine the extent of improvement that the G:G pair can offer in cells, experiments should be done for 5' G containing endogenous targets that do not have editing in the presence of a scrambled sequence guide or transfection of only ADAR2 overexpression plasmid.

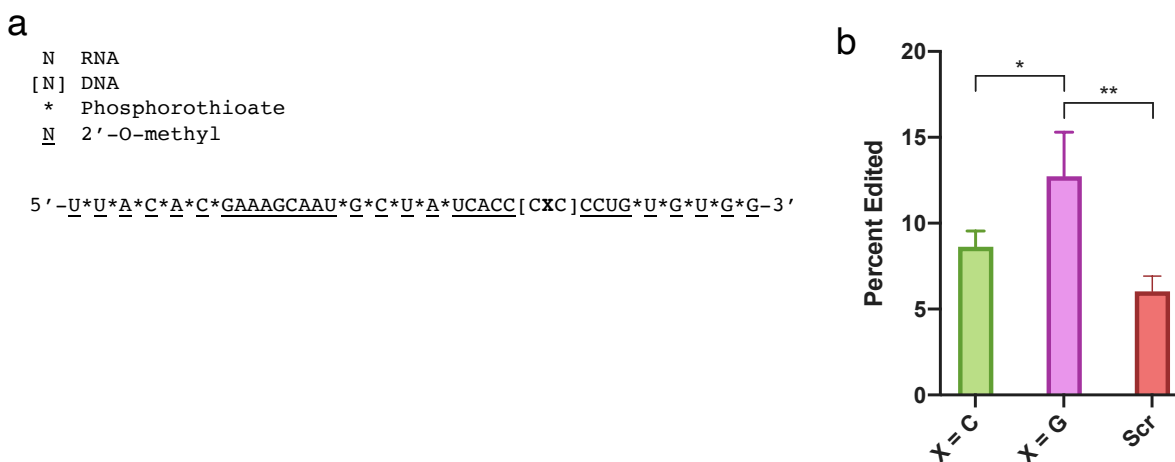


Figure 4.10. A) Partial sequence of the endogenous β -actin transcript paired with the guide RNA. The target adenosine is shown in red. Underlining indicates 2'-O-methyl nucleotides, asterisks indicate phosphorothioate linkages, and nucleotides in brackets are 2'-deoxynucleotides, all others are ribonucleotides. X = C or G, Scr indicates the scrambled sequence. B) Percent editing of the β -actin target with guide RNA and overexpression of ADAR2.

Structure-activity experiments support $G_{\text{syn}}:G_{\text{anti}}$ structure hypothesis

It is apparent from the *in vitro* experiments that editing at an adenosine flanked by a 5' guanosine is supported by including a purine:purine pair between the guide and target RNA at this position. For editing by ADAR2, including a 5' G:G pair yields the greatest rate of editing (Figure 4.2). The crystal structure of ADAR2d bound to dsRNA helps to explain why editing within a 5' guanosine sequence context is challenging, based on a clash with the G489 residue.²⁵ However, we sought to understand more about the conformation that this G:G pair is adopting, and the specific contacts enabling this rate enhancement. It seems that the G:G pair may be adopting a conformation that moves the 5' guanosine away from the conflicting G489 residue. Our hypothesis about the structure of this interaction could be either a purine:purine pair creating a larger diameter of the helix in this position to allow more space from the G489 residue, or the existence of a specific $G_{\text{syn}}:G_{\text{anti}}$ base pairing interaction (Figure 4.11 A). This proposed hydrogen bonding interaction between the two guanosines would relieve the proposed steric clash of the exocyclic amine, by flipping it into the syn conformation which points the amine away from the G489 residue. This syn-anti conformation would also explain the rate enhancement seen when including an adenosine in the 5' position (a G:A pair) due to adenosines ability to adopt a $G_{\text{syn}}:AH^+_{\text{anti}}$ conformation which would similarly allow for editing (Figure 4.11 B). Therefore to determine if the editing enhancement is due to this $\text{purine}_{\text{syn}}:\text{purine}_{\text{anti}}$ conformation—specifically with the guanosine in the edited strand occupying the $\text{purine}_{\text{syn}}$ conformation—we used guanosine analogs to probe the importance of specific contacts between the 5' guanosine and its partner. If specific contacts involved in a $\text{purine}_{\text{syn}}:\text{purine}_{\text{anti}}$ conformation were shown to influence to the rate of editing, then we could suggest that this conformation exists.

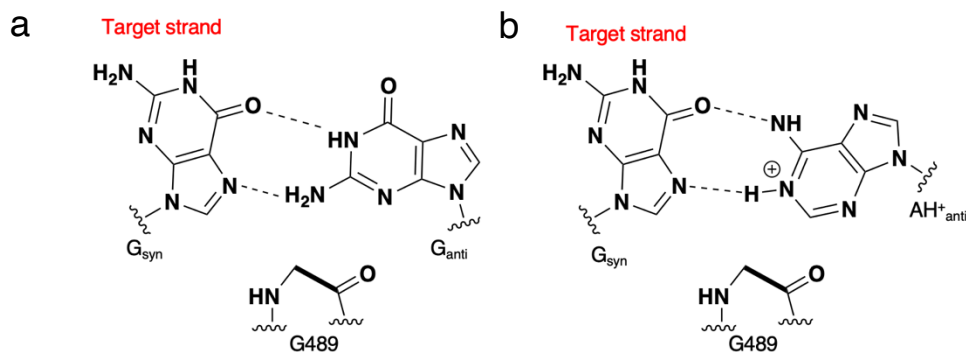


Figure 4.11. Proposed conformations of the purine_{syn}:purine_{anti} conformation 5' to the edited base. A) The G:G pair in the -1 position occupying a G_{syn}:G_{anti} conformation which draws the exocyclic amine on the target strand away from the conflicting G489 residue. B) The G:A pair in the -1 position occupying a G_{syn}:AH⁺_{anti} conformation which draws the exocyclic amine on the target strand away from the conflicting G489 residue.

We identified analogs that would manipulate either the hydrogen bonding face or sugar pucker of the purine in the guide strand. Changing either of these features would affect its ability to adopt a purine_{syn}:purine_{anti} conformation. The non-canonical bases 7-deaza-2'-deoxyguanosine (7-deaza-dG), 3-deaza-2'-deoxyadenosine (3-deaza-dA), 8-bromo-2'-deoxyguanosine (8-Br-dG), 2'-fluoroarabino-guanosine (2'-F-ANA-G), and 2'-deoxy-inosine (dI) were incorporated into the -1 position of guide oligonucleotides complementary to the *MECP2* R255X substrate (Figure 4.12 A). Guides were purchased from Dharmacon if commercially available, and otherwise synthesized and purified by Beal lab member Aashrita Manjunath. These modified bases were also compared with 2'-deoxyguanosine (dG), adenosine (A), 2'-deoxyadenosine (dA), cytidine (C), and guanosine (G). Comparing each 2'-deoxynucleoside to the corresponding ribonucleoside would offer insight into the sugar pucker adopted by the G-G pair since the 2'-deoxynucleosides prefer the C2'-*endo* conformation as opposed to C3'-*endo* of the ribose. Structural information of G:G pairs has shown that their sugar puckers remain similar to that of the native duplex.¹⁷³ The inclusion of 2'-F-ANA-G in this screen also perturbs the conformation of the sugar, as it prefers to adopt an O4'-*endo* sugar pucker.¹⁷⁴ Analogs affecting the hydrogen bonding face of the purine in the -1 position are detailed in Figure 4.12. Impacting the pK_a of the N1 nitrogen in the case of 7-deaza-dG and 3-deaza-dA, removing a hydrogen bond donor (dI), or including sterically bulky groups that force the -1 base into a syn conformation should each affect the rate of editing if the G:G pair exists as a G_{syn}:G_{anti} structure (Figure 4.12 B).

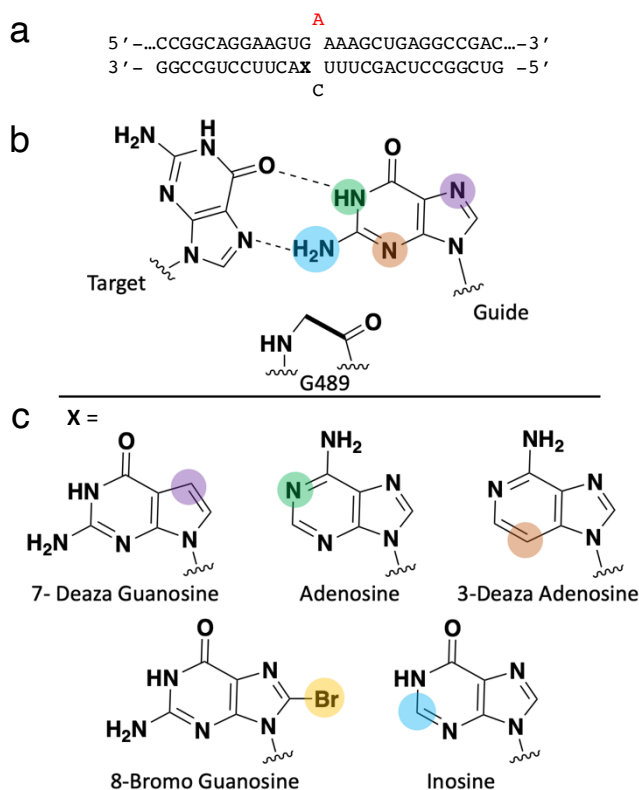


Figure 4.12. Substrates for the *in vitro* deamination of the 5' guanosine containing *MECP2* R255X target. A) Partial sequence of the target RNA, and sequence of the guide RNA with varying -1 position (X). B) The G:G pair in the -1 position occupying a $G_{syn}:G_{anti}$ conformation with highlighting on each position that is probed via modifications. C) Guanosine analogs containing modifications (highlighted) to probe each position's affect on the rate of reaction.

Therefore, we compared duplexes where the -1 base (X) was either C, G, A, dA, dG, 7-deaza-dG, 8-Br-dG, or 3-deaza-dA in an *in vitro* deamination reaction with ADAR2 under single-turnover conditions. Guanosine in the -1 position was slightly favored over -1 dG ($k_{obs} = 0.4 \pm 0.1$, $k_{rel} = 0.5$) indicating that the sugar pucker may stay in the C-3'-*endo* conformation adopted by RNA A-form helices (Figure 4.13, Table 4.7). However, in contrast there was no significant difference between -1 A and -1 dA. In addition, there was a significant decrease in rate for -1 7-deaza-dG and 8-Br-dG. However, there was a significant decrease for duplexes containing 2'-F-ANA-G ($k_{obs} = 0.18$, $k_{rel} = 0.2$), indicating that the conformation of the sugar is important to this structure. The pronounced decrease in rate when the duplex contains 7-deaza-dG ($k_{obs} = 0.028 \pm 0.006$, $k_{rel} = 0.04$) could be attributed to the elevated pKa of the N1 nitrogen (Figure 4.13, Table 4.7). This would form a weaker hydrogen bond at the N1 position which would affect the strength of a $G_{syn}:G_{anti}$ pairing interaction (Figure 4.12 B). The rate of deamination for 8-Br-dG was reduced to $k_{obs} = 0.05$

± 0.02 ($k_{rel} = 0.07$), which supports our structural hypothesis because the bromo group in the 8-position precludes the base from adopting a syn conformation during the ADAR reaction. The most significant result was the enhancement in rate when 3-deaza-dA or dI is included in the -1 position of the guide. This led to a 1.6-fold rate increase for 3-deaza-dA as compared to guanosine, and an 8-fold increase in rate as compared to dA (Figure 4.13, Table 4.7). In order to adopt the $\text{purine}_{syn}:\text{purine}_{anti}$ conformation, adenosine would have to become protonated (Figure 4.11). 3-Deaza-dA offers a higher pK_a at the N1 position, encouraging this protonation event. However, the protonated adenosine would have a slightly different conformation than the guanosine, which may be favorable (Figure 4.11). Guide strands containing dI at the -1 position are deaminated $k_{obs} = 1.3 \pm 0.2$ ($k_{rel} = 1.8$), a rate similar to that of 3-deaza-dA (Figure 4.13, Table 4.7). dI loses one of the hydrogen bonds present in the $\text{purine}_{syn}:\text{purine}_{anti}$ interaction, but it removes the 2-amino from the minor groove. Collectively, the results here demonstrate that specific hydrogen bonding contacts are important between the 5'adenosine and the -1 position base, as opposed to just the presence of a purine widening the helix diameter. This supports our hypothesis of the $\text{purine}_{syn}:\text{purine}_{anti}$ as the structural basis for improvement of editing. In addition, the discovery of a base (3-deaza-dA) that promotes editing better than guanosine or adenosine could prove useful in directed editing applications.

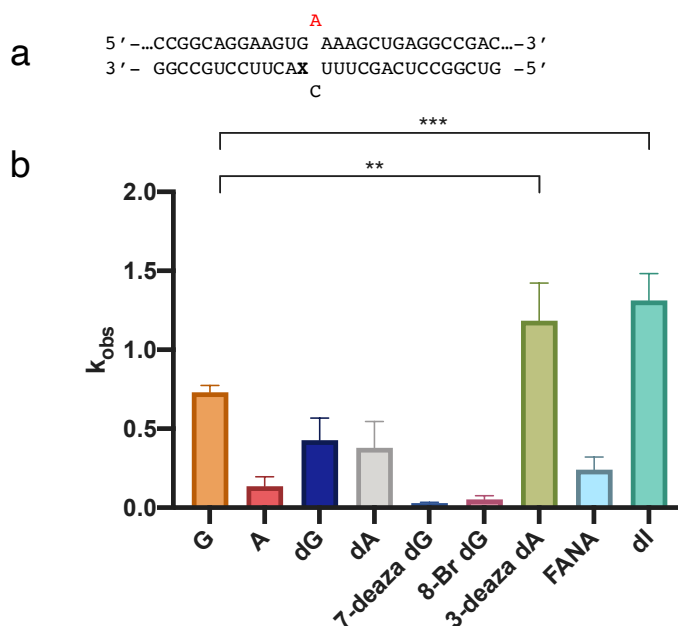
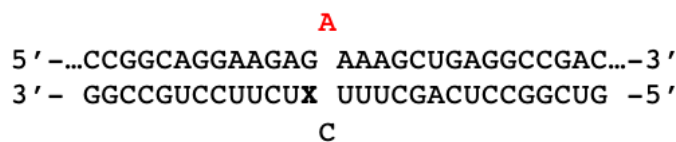


Figure 4.13. A) The R255X *MECP2* substrate used for *in vitro* deaminations under single turnover conditions. Reactions with ADAR2 were carried out with 10 nM RNA and 100 nM enzyme. B) A plot of the observed rate of reaction (k_{obs}) for each substrate.



enzyme	X	$k_{\text{obs}} \text{ min}^{-1a}$	k_{rel}^b
ADAR2	G	0.73 ± 0.04	1
	A	0.14 ± 0.06	0.2
	dG	0.4 ± 0.1	0.5
	dA	0.4 ± 0.2	0.5
	7-deaza dG	0.028 ± 0.006	0.04
	8-Br dG	0.05 ± 0.02	0.07
	3-deaza dA	1.2 ± 0.2	1.6
	FANA	0.24 ± 0.08	0.3
	dI	1.3 ± 0.2	1.8

Table 4.7. Rates of *in vitro* deamination for 100 nM ADAR2 acting on 10 nM substrate under single turnover conditions. X indicates the base pairing in the -1 position. ^a Data were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{\text{obs}}t)]$. ^b $k_{\text{rel}} = k_{\text{obs}}$ for analog / k_{obs} for cytosine.

3-Deaza-2'-deoxyadenosine enhances editing by ADAR1 p110 of adenosines flanked by a 5' G

The improvement in editing seen by ADAR2 with guide strands containing 3-deaza-dA compared to guanosine prompted us to test if editing by ADAR1 would also be improved by this analog. Therefore, we compared duplexes where the -1 base (X) was either dA, dG, or 3-deaza-dA in an *in vitro* deamination reaction with ADAR2 under single-turnover conditions (Figure 4.14). Since the rate of deamination was similar between ribonucleoside and 2'-deoxyribonucleosides, and 2'-deoxyribonucleosides are favored at the -1 position for cellular oligos, ribonucleosides were not tested with ADAR1. We observed that the rate of deamination for either canonical purine was similar, where -1 dG was $k_{\text{obs}} = 0.011 \pm 0.0002$ ($k_{\text{rel}} = 1$), -1 dA was 0.007 ± 0.002 ($k_{\text{rel}} = 0.6$). However, the 3-deaza-dA base exhibited an increased rate of reaction at $k_{\text{obs}} = 0.019 \pm 0.002$ ($k_{\text{rel}} = 1.7$). This result suggests that the same $\text{purine}_{\text{syn}}:\text{purine}_{\text{anti}}$ conformation is enabling editing of 5' guanosine targets via ADAR1, and that the 3-deaza-dA base can enhance therapeutic RNA editing.

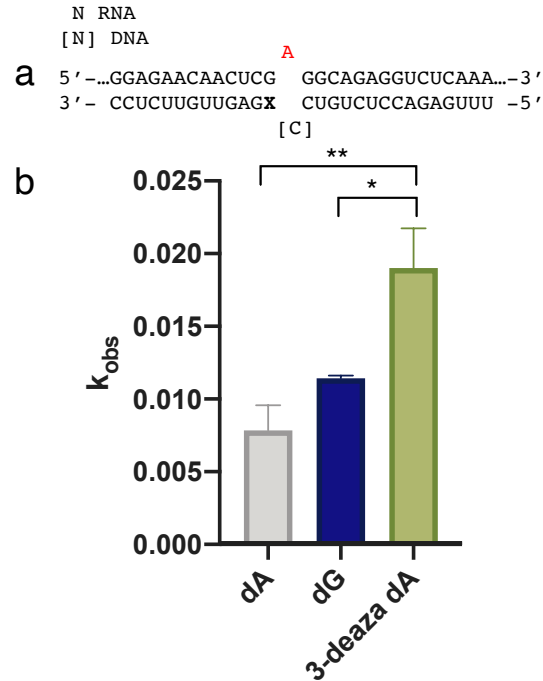


Figure 4.14. A) The *IDUA* substrate used for in vitro deaminations under single turnover conditions. Reactions with ADAR1 p110 were carried out with 10 nM RNA and 250 nM enzyme. B) A plot of the observed rate of reaction (k_{obs}) for each substrate.

Discussion.

This work has demonstrated how a very simple guide strand alteration can significantly improve ADAR editing within the most challenging sequence context (5'-GA). Expanding on the work done by the Stafforst group with SNAP-ADARs, we confirmed that for full-length ADARs incorporating a purine across from the 5' guanosine (in the -1 position of the guide) demonstrated between a 7.8- and 57-fold increase in the rate of reaction *in vitro* when compared to a guide oligonucleotide containing the canonical G:C pair in this position. In addition, this strategy was applied to an endogenous substrate in human cells and achieved a 1.5-fold enhancement in the editing yield. These *in vitro* and cellular experiments showed that this approach is not sequence or RNA specific and is compatible with chemical modifications commonly used for directed editing *in vivo* (2'-O-methyl, phosphorothioate, 2'-deoxy). This expands the scope of disease-causing mutations that can be effectively targeted by ADARs.

In addition, through the use of *in vitro* studies with modified purine analogs we have identified positions on the purine base that have a strong influence on the ADAR reaction. This serves to provide support for our structural hypothesis, that it is unlikely to be a non-specific contact between the -1 base and the 5' guanosine. Rather than the purine:purine pair widening the diameter of the duplex at this position to accommodate the conflicting glycine residue, we believe that there is a purine_{syn}:purine_{anti} conformation at this position. Flipping the 5' guanosine into a syn conformation would relieve the steric clash between its exocyclic amine and the glycine. This was determined through *in vitro* assays which showed that when the -1 purine was not able to adopt an anti conformation, or when specific hydrogen bond donors or acceptors were removed, the rate of reaction slowed. Knowledge of the structure at this position aids in the design of chemically modified nucleosides that could promote reaction better than guanosine in the -1 position.

Adenosine has to become protonated at the N1 position to adopt a purine_{syn}:purine_{anti} conformation. Therefore, we incorporated an adenosine analog, 3-deaza 2'-deoxyadenosine, with an elevated N1 pK_a to the -1 position of guide RNAs. *In vitro*, this base performed better in the guide strand than cytosine, guanine, or adenine for both ADAR1 and ADAR2. Removal of the 2-amino group from the minor groove of the -1 base could also potentially alleviate a steric clash with the G489 residue. Therefore, 2'-deoxyinosine was incorporated into the -1 position of guide RNAs, and indeed showed an enhanced rate of deamination *in vitro* with ADAR2. These chemical modifications hold the potential to further bolster editing yields in cells. Cellular studies of guide strands featuring -1 purines and purine analogs are ongoing in our lab.

Methods.

General Biochemical Procedures.

Molecular-biology-grade bovine serum albumin (BSA), and RNase inhibitor were purchased from New England BioLabs. SDS-polyacrylamide gels were visualized with a Molecular Dynamics 9400 Typhon phosphorimager. Data were analyzed with Molecular Dynamics ImageQuant 5.2 software. All MALDI analyses were performed at the University of California, Davis Mass Spectrometry Facilities using

a Bruker UltraFlex extreme MALDI TOF/TOF mass spectrometer. Oligonucleotide masses were determined with Mongo Oligo Calculator v2.08. Oligonucleotides for sequencing and PCR were purchased from Integrated DNA Technologies or Dharmacon. All other oligonucleotides were synthesized as described below.

Synthesis of oligonucleotides.

Chemical synthesis for all oligonucleotides was performed using an ABI 394 synthesizer. All bases were purchased from Glen Research. Nucleosides were incorporated during the appropriate cycle on a 0.2 μ mol scale; See Methods for sequences. Upon completion of the synthesis, columns were evaporated under reduced pressure for 4 h. All oligonucleotides were cleaved from the solid support by treatment with 1:3 ethanol/ 30% NH_4OH at 55 °C for 12 h. The supernatant was transferred to a new screw-cap tube and evaporated under reduced pressure. Desilylation was performed by resuspending the pellets in anhydrous DMSO and treating TBAF-THF at room temperature overnight. To each reaction was added 75 mM sodium acetate in butanol. The oligonucleotides were then precipitated from a solution of 65% butanol at -70 °C for 2 h. The solution was centrifuged at 13,000 rpm for 20 min, supernatant was removed, and the pellet was washed twice with cold 95% ethanol. The RNA pellets were then desalted using a Sephadex G-25 column and purified as described below.

Purification of oligonucleotides.

Single-stranded RNA oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis and visualized by UV shadowing. Bands were excised from the gel, crushed and soaked overnight at 4 °C in 0.5 M NaOAc, 0.1% sodium dodecyl sulfate (SDS), and 0.1 mM EDTA. Polyacrylamide fragments were removed with a 0.2 μ m filter, and the RNAs were precipitated from a solution of 75% EtOH at -70 °C for 4 h. The solution was centrifuged 13,000 rpm for 20 min and supernatant was removed. The RNA solutions were lyophilized to dryness, resuspended in nuclease-free water, and quantified by absorbance at 260 nm. Oligonucleotide mass was confirmed by MALDI-TOF.

In vitro transcription of editing target RNA.

Target RNA was transcribed from a DNA template with the MEGAScript T7 Kit (ThermoFisher). DNA digestion was performed using RQ1 RNase-free DNase (Promega). DNase treated RNA product was purified as described above.

Preparation of Duplex Substrates for Analysis of ADAR Deamination Kinetics.

Purified guide and transcribed RNA were added in a 10:1 ratio to hybridization buffer (180 nM transcribed RNA target, 1.8 μ M guide, 1X TE Buffer, 100 mM NaCl), heated to 95 °C for 5 min, and slowly cooled to room temperature.

Protein Overexpression and Purification of ADAR2 constructs

hADAR2 (hADAR2) was expressed and as previously described.¹²⁹ Purification of hADAR2 was carried out by lysing cells in buffer containing 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 750 mM NaCl, 35 mM imidazole, and 0.01% Nonidet P-40 using a French press. Cell lysate was clarified by centrifugation (19,000 rpm for 1 hour). Lysate was passed over a 3 mL Ni-NTA column, which was then washed in three steps with 20 mL lysis buffer, wash I buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 750 mM NaCl, 35 mM imidazole, 0.01% Nonidet P-40), wash II buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 35 mM imidazole, 500 mM NaCl), and eluted with 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 400 mM imidazole, 100 mM NaCl. Fractions containing the target protein were pooled and concentrated to 30-80 μ M for use in biochemical assays. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels. Purified hADAR2 WT was stored in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 20% glycerol and 1 mM BME at -70 °C.

Protein overexpression and purification of ADAR1 p110

MBP-tagged human ADAR1 p110 construct was cloned into a pSc vector using standard PCR techniques. The generated construct (yeast codon optimized) consisted of an N-terminal MBP-tag, a tobacco etch virus (TEV) protease cleavage site followed by the human ADAR1 p110 gene. The construct was transformed in *S. cerevisiae* BCY123 cells and overexpressed as described previously.¹²⁹ Purification was carried out by lysing cells in lysis/binding buffer containing 50 mM Tris-HCl, pH 8.0, 5% glycerol, 5 mM 2-mercaptoethanol, 1000 mM KCl, 0.05% NP-40 and 50 μ M ZnCl₂ using a microfluidizer. Cell lysate was clarified by centrifugation (39,000 x g for 50 min). Lysate was passed over a 2 mL NEB amylose column (pre-equilibrated with binding buffer), which was then washed in 2 steps with 50 mL binding buffer followed by 100 mL wash buffer (50 mM Tris-HCl, pH 8.0, 5% glycerol, 5 mM 2-mercaptoethanol, 500 mM KCl, 0.01% NP-40 and 50 μ M ZnCl₂) and eluted with buffer containing 50 mM Tris-HCl, pH 8.0, 10% glycerol, 5 mM 2-mercaptoethanol, 500 mM KCl, 0.01% NP-40, 50 μ M ZnCl₂, and 20 mM maltose. Fractions containing the target protein were pooled and dialyzed against a storage buffer containing 50 mM Tris-HCl, pH 8.0, 400 mM KCl, 0.5 mM EDTA, 0.01% NP-40, 10% glycerol and 1 mM tris(2-carboxyethyl)phosphine. Dialyzed protein was concentrated to 2-50 μ M and stored as aliquots at -70 °C until further use in biochemical assays. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels.

Deamination assay with ADAR2d, ADAR2d-E488Q, hADAR2d RD, ADAR2, and ADAR1 p110

Deamination assays were performed under single-turnover conditions in 15 mM Tris-HCl pH 7.5 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, 3 mM MgCl₂, 160 U/mL RNAsin, 1.0 μ g/mL yeast tRNA, 10 nM RNA, and 75 nM ADAR2d, ADAR2d-E488Q, or wild-type ADAR2. Each reaction solution was incubated at 30 °C for 30 min before the addition of enzyme. Reactions were then incubated at 30 °C for varying times prior to quenching with 190 μ L 95°C water and heating at 95 °C for 5 min. Reaction products were used to generate cDNA using RT-PCR (Promega Access RT-PCR System). DNA was purified using a DNA Clean & Concentrator kit (Zymo) and subjected to Sanger Sequencing through GeneWiz (Azenta). The sequencing peak heights were quantified in SnapGene (Domotics). Data

were fit to the equation $[P]_t = P_f * [1 - e^{-(k_{obs} * t)}]$ for ADAR2 and $[P]_t = 0.4 * [1 - e^{-(k_{obs} * t)}]$ for ADAR1 p110 where $[P]_t$ is percent edited at time t, $[P]_f$ is the final endpoint of editing, and k_{obs} is the observed rate constant. Each experiment was carried out in triplicate where the k_{obs} reported is the average of each replicate \pm standard deviation (SD). Statistical significance between groups was determined by one-way ANOVA using Prism software (GraphPad). For the ADAR1 p110 enzyme, deamination reactions were performed as above with the following modifications: The final reaction solution for ADAR1 p110 contained 15 mM Tris-HCl, pH 7.0 4% glycerol, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% Nonidet P-40, 160 U/mL RNasin, 1.0 μ g/mL yeast tRNA, and 10 nM RNA, and 250 nM ADAR1 p110.

Directed Editing on the endogenous β -actin Target in HEK293T Cells.

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11995-065) with 10% fetal bovine serum (FBS) (Thermo Fisher, 26140-087) and additionally supplemented with 1X antibiotic-antimycotic (Thermo Fisher, 15240-062) at 37 °C, 5 % CO₂. Once cells reached 70-90% confluency, cells were seeded into 96 well plates (6.4 x 10³ cells per well). After 24 h, cells were co-transfected with 500 ng ADAR plasmid and 50 nM guide oligonucleotide using Lipofectamine 2000 (Thermo Fisher, 11668-019). After incubation of transfection reagent, plasmid, and guide oligonucleotide in Opti-MEM Reduced Serum Media (Thermo Fisher, 31985-062), the solution was added to designated wells and incubated at 37 °C, 5 % CO₂. After 48 h, total RNA was isolated using RNAqueous Total RNA Isolation Kit (Thermo Fisher, AM1912) and DNase treated with RQ1 RNase-free DNase (Promega, M6101). Nested RT-PCR was performed in triplicate using Access RT-PCR kit (Promega, A1280) for 20 cycles and then followed by Phusion Hot Start II DNA Polymerase (Thermo Fisher, F-549L) for the second PCR of 30 cycles with target specific primers. PCR product was purified by agarose gel and QIAquick Gel Extraction kit (Qiagen, 28706). Product was submitted for Sanger Sequencing and sequence traces were analyzed by SnapGene (Domatics) to quantify percent editing.

Sequences of Oligonucleotides.

a) Sequences for *in vitro* kinetics of the *IDUA* target (nucleotides in brackets are 2'-deoxy. All others are ribonucleotides). All PCR primers are 2'-deoxynucleotides. (3dA) indicates 3-deaza adenosine.

<i>IDUA</i> guide strand -1 A	5'- UUUGAGACCUCUGUC[C]AGAGUUGUUCUCC -3'
<i>IDUA</i> guide strand -1 G	5'- UUUGAGACCUCUGUC[C]GGAGUUGUUCUCC -3'
<i>IDUA</i> guide strand -1 C	5'- UUUGAGACCUCUGUC[C]CGAGUUGUUCUCC -3'
<i>IDUA</i> guide strand -1 C	5'- UUUGAGACCUCUGUC[C]UGAGUUGUUCUCC -3'
<i>IDUA</i> guide strand -1 2'-deoxy G	5'- UUUGAGACCUCUGUC[CG]GAGUUGUUCUCC -3'
<i>IDUA</i> guide strand -1 2'-deoxy A	5'- UUUGAGACCUCUGUC[CA]GAGUUGUUCUCC -3'
<i>IDUA</i> guide strand -1 2'-3-deaza adenosine	5'- UUUGAGACCUCUGUC[C(3dA)]GAGUUGUUCUCC -3'
<i>IDUA</i> RT-PCR forward and sequencing primer	5'-GCTCCTCCCATCCTGTGGGCTGAACAGT-3'
<i>IDUA</i> RT-PCR reverse primer	5'-CGGGGTGTGCGTGGGTGTCATCACT-3'

b) DNA template sequence for *in vitro* kinetics of the *IDUA* 5'-UA target. Grey indicates the T7 promoter, underline is the region complementary to the guide strands, and the red A is the target adenosine.

TAATACGACTCACTATAGGGctcctcccatcctgtgggctgaacagtataacagactcccagtatacaaatggtgggagctagatattaggtaggaagccagatgctaggtatgagagagccaacagcctcagccctcgttgcttagATGGAGAACAACTCTAGGCAGAGGTCTCAAAGGCTGGGGCTGTGTTGGACAGCAATCATACAGTGGGTGTCCTGGCCAGCACCATCACCCCTGAAGGCTCCGCAGCGGCCTGGAGTACCACAGTCCTCATCTACTAGTGATGACACCCACGCACACCCCGGATCC

c) DNA template sequence for *in vitro* kinetics of the *IDUA* 5'-GA target. Grey indicates the T7 promoter, underline is the region complementary to the guide strands, and the red A is the target adenosine.

TAATACGACTCACTATAGGGctcctcccatcctgtgggctgaacagtataacagactcccagtatacaaatggtgggagctagatattaggtaggaagccagatgctaggtatgagagagccaacagcctcagccctcgttgcttagATGGAGAACAACTCGAGGCAGAGGTCTCAAAGGCTGGGGCTGTGTTGGACAGCAATCATACAGTGGGTGTCCTGGCCAGCACCATCACCCCTGAAGGCTCCGCAGCGGCCTGGAGTACCACAGTCCTCATCTACTAGTGATGACACCCACGCACACCCCGGATCC

d) Sequences for *in vitro* kinetics of the *MECP2* R168X target. All guides are ribonucleotides. All PCR primers are 2'-deoxynucleotides.

<i>MECP2</i> R168X guide strand -1 C	5'- UUUGAGACCUCUGUC[C]AGAGUUGUUCUCC -3'
<i>MECP2</i> R168X guide strand -1 G	5'- UUUGAGACCUCUGUC[C]GGAGUUGUUCUCC -3'
<i>MECP2</i> R168X RT-PCR forward and sequencing primer	5'- GGGATCAATCCCCAG-3'

<i>MECP2</i> R168X RT-PCR reverse primer	5'- CTTTTCACCTGCACAC-3'
--	-------------------------

e) DNA template sequence for *in vitro* kinetics of the *MECP2* R168X target. Grey indicates the T7 promoter, underline is the region complementary to the guide strands, and the red A is the target adenosine.

TAATACGACTCACTATAGGGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTG
ATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGT
AACTGGGAGAGGGAGCCCCTCCCGGTGAGAGCAGAAAACCACCTAAGAAGCCCAAATCTCCC
AAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCAAAGGGAGCGGCACCACGAGACCC
AAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAG

f) Sequences for *in vitro* kinetics of the *MECP2* R255 target. All guides are ribonucleotides. All PCR primers are 2'-deoxynucleotides.

<i>MECP2</i> R255 guide strand -1 C	5'- GUCGGCCUCAGCUUUCGCUUCCUGCCGG -3'
<i>MECP2</i> R255 guide strand -1 G	5'- GUCGGCCUCAGCUUUCGGCUUCCUGCCGG-3'
<i>MECP2</i> R255 RT-PCR forward and sequencing primer	5'-GTGCAGGTGAAAAGGGTC-3'
<i>MECP2</i> R255 RT-PCR reverse primer	5'-TACGGTCTCCTGCACAGATCG-3'

g) DNA template sequence for *in vitro* kinetics of the *MECP2* R255 target. Grey indicates the T7 promoter, underline is the region complementary to the guide strands, and the red A is the target adenosine.

TAATACGACTCACTATAGGGGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCT
CCTTGTCAAGATGCCTTTTCAAACCTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACC
ACATCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACC
CTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGC
CGAGGCCAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTA

h) Sequences for *in vitro* kinetics of the *MECP2* R255X target. All guides are ribonucleotides. All PCR primers are 2'-deoxynucleotides. (8BrG) is 8-bromo-2'-deoxyguanosine, (7dA) is 7-deaza-2'-deoxyadenosine, (3dA) is 3-deaza-2'-deoxyadenosine

<i>MECP2</i> R255X guide strand -1 C	5'- GUCGGCCUCAGCUUUCACUUCCUGCCGG -3'
<i>MECP2</i> R255X guide strand -1 G	5'- GUCGGCCUCAGCUUUCGACUUCCUGCCGG -3'
<i>MECP2</i> R255X guide strand -1 A	5'- GUCGGCCUCAGCUUUCAACUUCCUGCCGG -3'
<i>MECP2</i> R255X guide strand -1 2'-deoxy A	5'- GUCGGCCUCAGCUUUC[A]ACUUCCUGCCGG -3'

<i>MECP2</i> R255X guide strand -1 2'-deoxy G	5'- GUCGGCCUCAGCUUUC[G]ACUUCCUGCCGG -3'
<i>MECP2</i> R255X guide strand -1 2'-deoxy 8-bromo guanosine	5'- GUCGGCCUCAGCUUUC[(8BrG)]ACUUCCUGCCGG -3'
<i>MECP2</i> R255X guide strand -1 2'-deoxy 7-deaza guanosine	5'- GUCGGCCUCAGCUUUC[(7dG)]ACUUCCUGCCGG -3'
<i>MECP2</i> R255X guide strand -1 2'-deoxy 3-deaza adenosine	5'- GUCGGCCUCAGCUUUC[(3dA)]ACUUCCUGCCGG -3'
<i>MECP2</i> R255X RT-PCR forward and sequencing primer	5'-GGGTGTGCAGGTGAAAAGG-3'
<i>MECP2</i> R255X RT-PCR reverse primer	5'-TCTTGATGGGGAGTACGGTC-3'

i) DNA template sequence for *in vitro* kinetics of the *MECP2* R255X target. Grey indicates the T7 promoter, underline is the region complementary to the guide strands, and the red A is the target adenosine.

CACGATTAATACGACTCACTATAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTG
GGAAGCTCCTTGTCAAGATGCCTTTTCAAACCTCGCCAGGGGGCAAGGCTGAGGGGGGTGG
GGCCACCACATCCACCCAGGTCATGGTGATCAAACGCCCGGCAGGAAGTGAAAAGCTGAG
GCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCG
CTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGAC
CGTACTCCCCATCAAGAA

j) Sequences for directed editing of β -Actin in HEK293T cells (phosphorothioate modification is marked with an asterisk, 2'-O-methylated nucleotides are underlined, 2'-deoxynucleotides are in brackets; all others are ribonucleotides). All PCR primers are 2'-deoxynucleotides.

β -actin endogenous target guide RNA -1 G	5'-U*U*A* C*A*C* GAAAGCAAU* G*C*U *A*UCACC[CCC]C CUG*U*U* U*G*G-3'
β -actin endogenous target guide RNA -1 C	5'-U*U*A* C*A*C* GAAAGCAAU* G*C*U *A*UCACC[CGC]C CUG*U*U* U*G*G-3'
Scrambled sequence	5'- A*U*G*U*C*U*AAGGCGCGA*C*A*C*C*GUCUU[ACG]UGA C*A*C*C*A*U-3'
Endogenous β -actin RT-PCR forward primer	5'-CAGCAGATGTGGATCAGCAAGCAGGAG-3'
Endogenous β -actin RT-PCR reverse primer	5'-GGAAGGGGGGGCACGAAGGCTCATC-3'
Endogenous β -actin nested PCR forward and sequencing primer	5'-TATGACGAGTCCGGCCCCCTCCATCGT-3'

Endogenous β -actin
nested PCR reverse
primer

5'-GCAATGCTATCACCTCCCCTGTGTGGACT-3'

CHAPTER 5

Rosetta-based Structural Modeling Predicts Unique Features of the ADAR1 Catalytic Domain

*The work presented in this chapter was enabled by biochemical studies done by Dr. SeHee Park (Beal lab alumna). This chapter focuses on the structural modeling contributions to this collaborative effort and contains excerpts from the full manuscript which was published in Nature Communications in October 2020.*¹⁵⁵

Introduction.

The different enzymes known to carry out A-to-I conversion in humans are ADAR1 (p150 and p110) and ADAR2. ADAR activity is required for nervous system function and altered editing has been linked to neurological disorders.¹⁷⁵⁻¹⁷⁷ Furthermore, ADAR1 plays an important role in innate immunity.¹⁷⁸⁻¹⁸⁰ Mutations in the *ADAR1* gene are known to cause the autoimmune disease Aicardi-Goutières Syndrome (AGS) and the skin disorder Dyschromatosis Symmetrica Hereditaria (DSH).^{98,181,182} In addition, ADAR1 mediated A-to-I editing of dsRNAs prevents overactivation of dsRNA sensing pathways, thus avoiding an auto inflammatory response.^{178,179} In fact, a recent study has shown that ADAR1 knockout sensitizes certain types of tumors to immunotherapy by activating dsRNA sensing pathways in tumor cells, which in turn leads to an innate immune response.²⁰ Thus, investigating how ADAR1 activity is regulated and how ADAR1 differentiates self RNAs from non-self RNAs is crucial to understanding the immune response pathways regulated by ADAR1. Additionally, recent studies show that the loss of ADAR1 function is lethal in a specific subset of cancers that display an interferon-stimulated gene signature, identifying ADAR1 as a potential chemotherapeutic target.^{21,22,180,183}

The ADAR proteins have a modular structure with double stranded RNA binding domains (dsRBDs) and a C-terminal deaminase domain.⁸³ ADARs require duplex secondary structure in their substrate RNAs and use a base-flipping mechanism to place the reactive adenosine into a zinc-containing active site.^{25,83} Indeed, surface loops present on the ADAR deaminase domain have been identified that bind

RNA on the 5' side of an editing site (i.e. 5' binding loop), bind RNA on the 3' side of the editing site (i.e. 3' binding loop) and are directly involved in base-flipping (i.e. flipping loop).²⁵ ADARs selectively edit certain adenosines over others in an RNA molecule and ADAR1 and ADAR2 have overlapping yet distinct selectivity.^{36,144,184} However, our understanding of the basis for ADAR-specific selectivity is limited. Domain swapping experiments demonstrated that selectivity differences between the ADARs primarily originate from differences in their deaminase domains and earlier work indicated that differences in sequence of the 5' binding loops of the deaminase domains were important determinants for ADAR-specific selectivity.⁷²³⁵ The structure and RNA recognition properties of the deaminase domain of human ADAR2 (hADAR2d) have been extensively studied.^{25,35,185} Less is known about the ADAR1 deaminase domain (hADAR1d).^{35,36,157} This is, in part, due to the lack of structural information for the ADAR1 deaminase domain alone or in complex with RNA. Although the deaminase domains of human ADAR1 and ADAR2 (hADAR1d and hADAR2d, respectively) have 59% sequence similarity, these two proteins show different A-to-I editing efficiencies on different substrate RNAs.^{7,36} There have been various studies directed at understanding the differences between these two proteins.^{35,36,144} Thus, studies that advance our understanding of structural features of the ADAR1 deaminase domain, particularly those that are unique to ADAR1, are important. Here we describe the generation of Rosetta-based molecular models based on high throughput mutagenesis/functional screening data to define structural features unique to ADAR1.

Model generation is supported by constraints from biochemical information obtained via high throughput mutagenesis and functional screening carried out by Dr. SeHee Park. These models support the discovery of a novel zinc binding site present on the surface of the ADAR1 deaminase domain but absent in ADAR2. We identify amino acids that make up the ligand environment for this second zinc site which supports the biochemical results showing these residues are important for efficient deamination by the ADAR1 deaminase domain *in vitro* and by full length ADAR1 p110 in cultured human cells. Furthermore, the models explain previously observed properties of the ADAR1 deaminase domain and suggest roles for specific residues present in the ADAR1 5' binding loop, including one mutated in Aicardi-Goutières

Syndrome. Given the success of the ADAR1d model in making predictions about the roles of specific residues in the protein, a model of ADAR3d was also generated.

Results.

ADAR1 operates through a base-flipping mechanism

Although it is generally accepted that ADAR1 operates through a base-flipping mechanism like ADAR2, there is not yet experimental evidence to directly support this. In the case of ADAR2, base-flipping assays have been performed that indicate unstacking of the target base, and eventually the crystal structure was solved allowing the base-flipped conformation to be visualized.^{23,107} In order to use the base-flipped ADAR2d structure as a template for modeling ADAR1d, we sought to obtain evidence of ADAR1's base-flipping mechanism.

An assay that uses the fluorescent nucleoside 2-aminopurine (2-AP) to detect flipping of the edited base was used.^{23,84} When incorporated into a duplex, 2-AP is quenched. However, ADARs are able to restore fluorescence of the base by flipping 2-AP out of the duplex into the enzyme active site. We identified positions within the *HER1* RNA, a known substrate of ADAR1d, that would become unstacked during the base-flipping mechanism based on our knowledge of the ADAR2d structure. In the target strand this included the edited position, 5' and 3' adjacent to the edited position, and on the complementary strand 3' to the base paired across the edited base (Figure 5.1-2). Each of these positions experience ADAR2-induced conformational change that lead to at least partial unstacking.¹⁰⁷ All duplexes contained 8-azanebularine (8-AN) in the edited position of the duplex, with the exception of the duplex with 2-AP in the edited position. 8-Azanebularine is a transition state analog that is known to trap ADAR2 in the base flipped conformation.^{26,107}

We chose to use the deaminase domain of ADAR1 with the E1008Q mutation, which renders the enzyme hyperactive.¹⁸⁶ The analogous mutation in ADAR2 (E488Q) enhances flipping of the edited base, which makes it useful to amplify signal in this assay.^{84,186} When ADAR1d E1008Q was added to a buffered solution of Duplex C, containing 2-aminopurine in the active site, there was no significant change in fluorescence maximum observed (Figure 5.1 B,D). Unlike ADAR2, the local environment of the ADAR1 active site may quench 2-AP, making it difficult to observe the fluorescence enhancement due to unstacking. In the presence of Duplex B, where the 2-AP was placed in the complementary strand adjacent to the base opposite the edited base, there was fluorescence quenching observed (Figure 5.1 A,C). The quantum yield of 2-AP is highly sensitive to its electronic environment, so this decrease could indicate conformational change due to ADAR1d E1008Q binding.

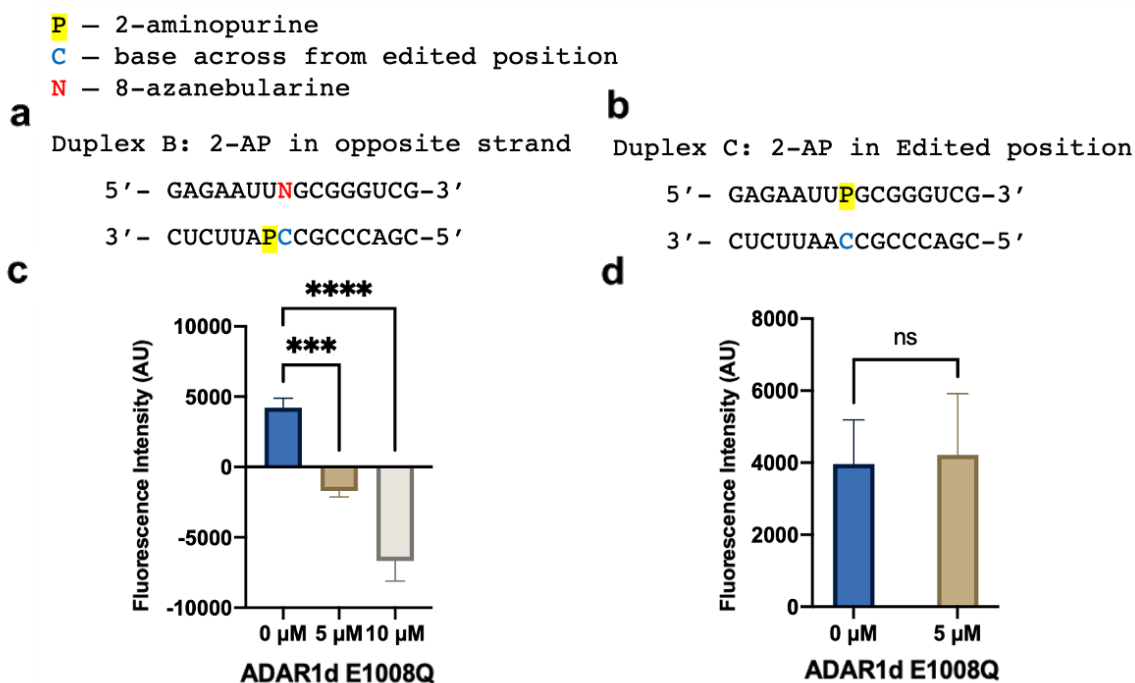


Figure 5.1. Base-flipping assay with ADAR1d E1008Q in the presence of a 16mer HER1 duplex containing 2-aminopurine in different positions. A) Duplex B containing 2-AP in the strand opposite the target strand. B) Fluorescence change observed for Duplex B with varying concentrations of ADAR1d E1008Q. C) Duplex C containing 2-AP in the edited position of the target strand. D) Fluorescence change observed for Duplex C upon the addition of ADAR1d E1008Q. Plotted values represent the means of three technical replicates \pm standard deviation. Statistical significance between groups was determined using one-way ANOVA with Tukey's multiple comparisons test; *** $p \leq 0.001$; **** $p \leq 0.0001$.

In the presence of Duplex A, containing 2-AP directly 5' to the edited position, we observed a fluorescence enhancement that is typical for base-flipping enzymes.^{23,149,150} The observed increase in

fluorescence was time-dependent, stabilizing after approximately 30 minutes, and was dependent upon the presence of active enzyme (Figure 5.2 A,C). When the time dependence of complex formation was fitted to a pseudo-first order rate equation, the observed rate of formation was consistent with the rate of deamination for *HER1* in previous reports (Figure 5.2 C).³⁵ This result offers support for the mechanism of ADAR1 also containing a base-flipping event similar to ADAR2, making the ADAR2d crystal structure a promising model for ADAR1d.

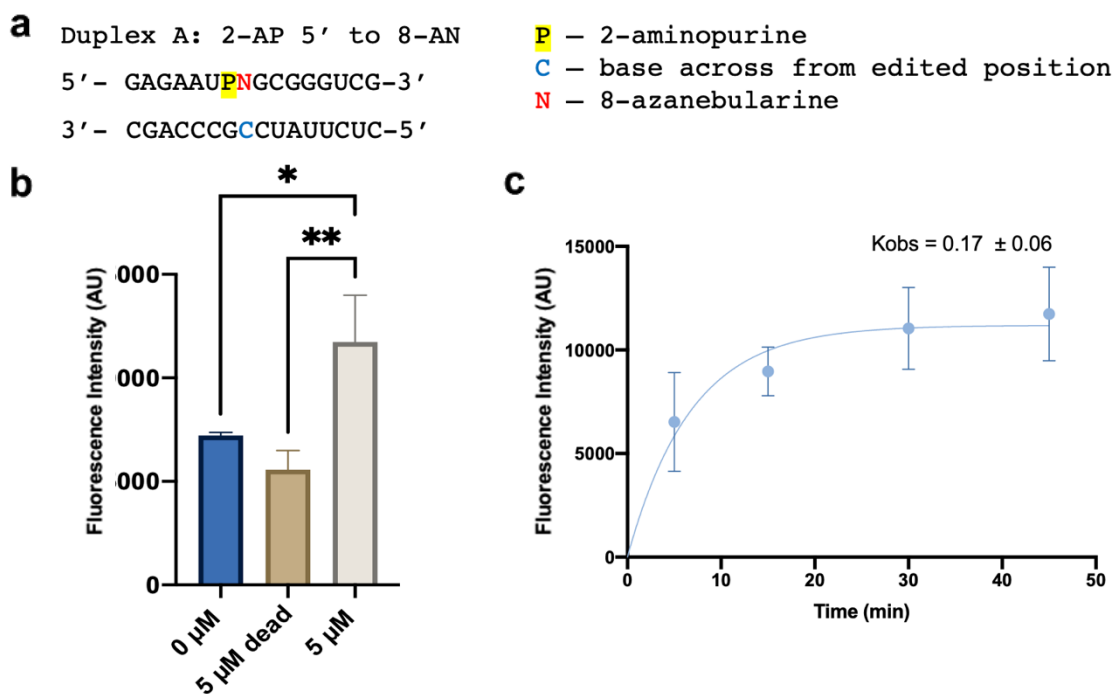


Figure 5.2. Base-flipping assay with ADAR1d E1008Q in the presence of a 16mer *HER1* duplex containing 2-aminopurine in different positions. A) Duplex A containing 2-AP directly 5' to the transition state analog 8-AN in the edited position. B) Fluorescence enhancement of the duplex with denatured (dead) ADAR1d E1008Q versus active ADAR1d E1008Q. C) The change in fluorescence intensity over time for Duplex A in the presence of ADAR1d E1008Q.

Modeling supports discovery of a second metal binding site in the ADAR1 deaminase domain.

Previous work done in the Beal lab by Dr. SeHee Park employed the Sat-FACS-Seq screening method to investigate the importance of cysteine residues present in hADAR1d.¹⁵⁵ This method utilizes Saturation mutagenesis (Sat) to incorporate codons for the 20 different amino acids at specific positions of interest, Fluorescence Activated Cell Sorting (FACS) to sort yeast cells based on various levels of fluorescence from a fluorescent activity reporter, and next-generation Sequencing (Seq) to identify ADAR library mutants with different levels of editing activity.¹⁸⁶ Most of the cysteine positions were found to

tolerate mutations to many amino acids other than the wild-type residue, except for two positions: C1081 and C1082. The C1081 position showed a strong preference for the wild-type cysteine residue. The C1082 position also has a high preference for cysteine, although it can be replaced with other amino acids and retain moderate activity. Thus, this Sat-FACS-Seq data indicated that C1081 and C1082 play an important role in the function of hADAR1d. Interestingly, we noted that the amino acids that are well tolerated at C1082X are histidine, glutamic acid, and aspartic acid (Figure 5.3). These amino acids are commonly found in metal binding sites in proteins along with cysteine.¹⁸⁷ Because there are two adjacent cysteine residues at positions 1081 and 1082 and both positions showed a high preference for amino acids that are often associated with a metal binding, we speculated that an additional metal binding site other than the catalytic zinc binding site could exist in hADAR1d.

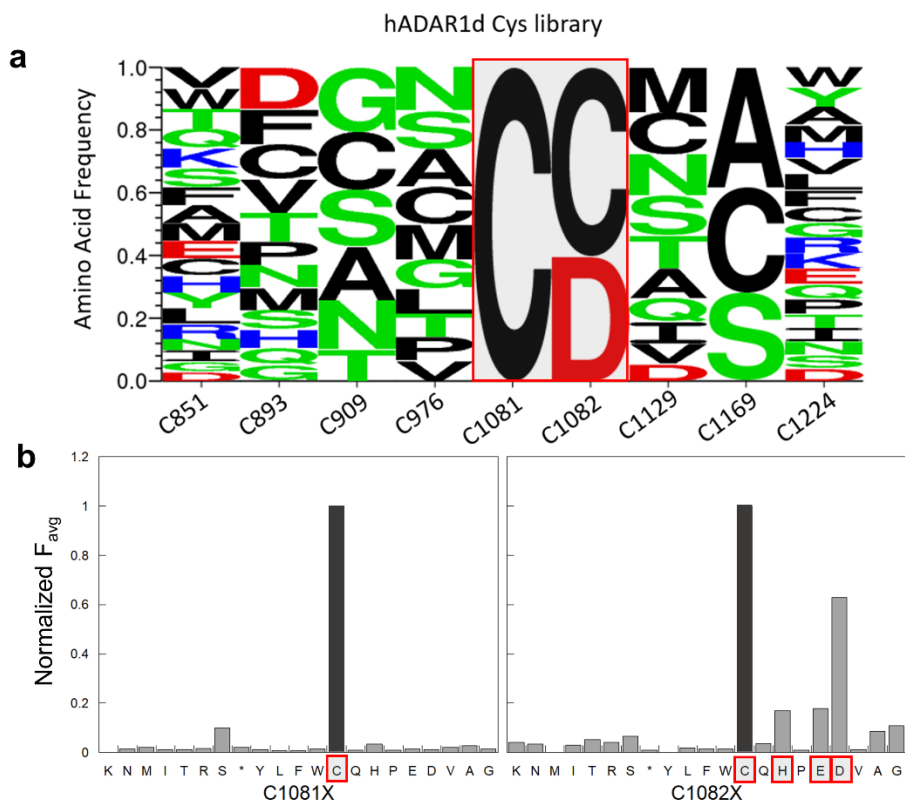


Figure 5.3. Sat-FACS-seq analysis of hADAR1d cysteine libraries. (a) Logo plot result of Sat-FACS-Seq analysis of native Cys residues in hADAR1 catalytic domain. Positions where native Cys residues is highly preferred are boxed in red. (b) Normalized averaged fluorescence of position C1081 and C1082. Amino acids that are commonly involved in metal binding are boxed in red.

To test this hypothesis further, we generated a homology model of hADAR1d using the SWISS-MODEL web server¹⁸⁸ based on the crystal structure of hADAR2d bound to dsRNA.²⁵ The resulting homology model suggested that H1103 is in close proximity to C1081 and C1082, consistent with a metal binding site involving ligation by the side chains of these three amino acids (Figure 5.3 A). Indeed, C1081, C1082 and H1103 are conserved among ADAR1 proteins from different organisms suggesting functional significance for all three (Figure 5.3 C). In addition, we used the Metal Ion-Binding Site Prediction and Docking Server (MIB) to further assess the possibility of a second metal binding site in hADAR1d.¹⁸⁹ The homology model for hADAR1d was used as a template for the metal binding residue prediction and docking for various metal ions with MIB. The C1081, C1082, and H1103 site showed a high binding score for zinc, suggesting that these residues are involved in zinc binding. While the Sat-FACS-seq data were consistent with C1081 and C1082 being involved in metal binding, H1103 was predicted to be a part of the metal binding site based on the homology model and the MIB prediction. To further test the role of H1103, Dr. Park mutated this residue to various amino acids (H1103A, H1103F, H1103Q, H1103S, and H1103C) and determined the activity level for each mutant by monitoring fluorescence intensities for the fluorescent activity reporter that was used for Sat-FACS-Seq and normalizing to the fluorescence generated by WT hADAR1d. Interestingly, only H1103C, a mutant that could still bind metal ions, showed activity (Figure 5.3). Taken together these results provide strong support for a second zinc metal binding site within the deaminase domain of hADAR1 involving C1081, C1082 and H1103.

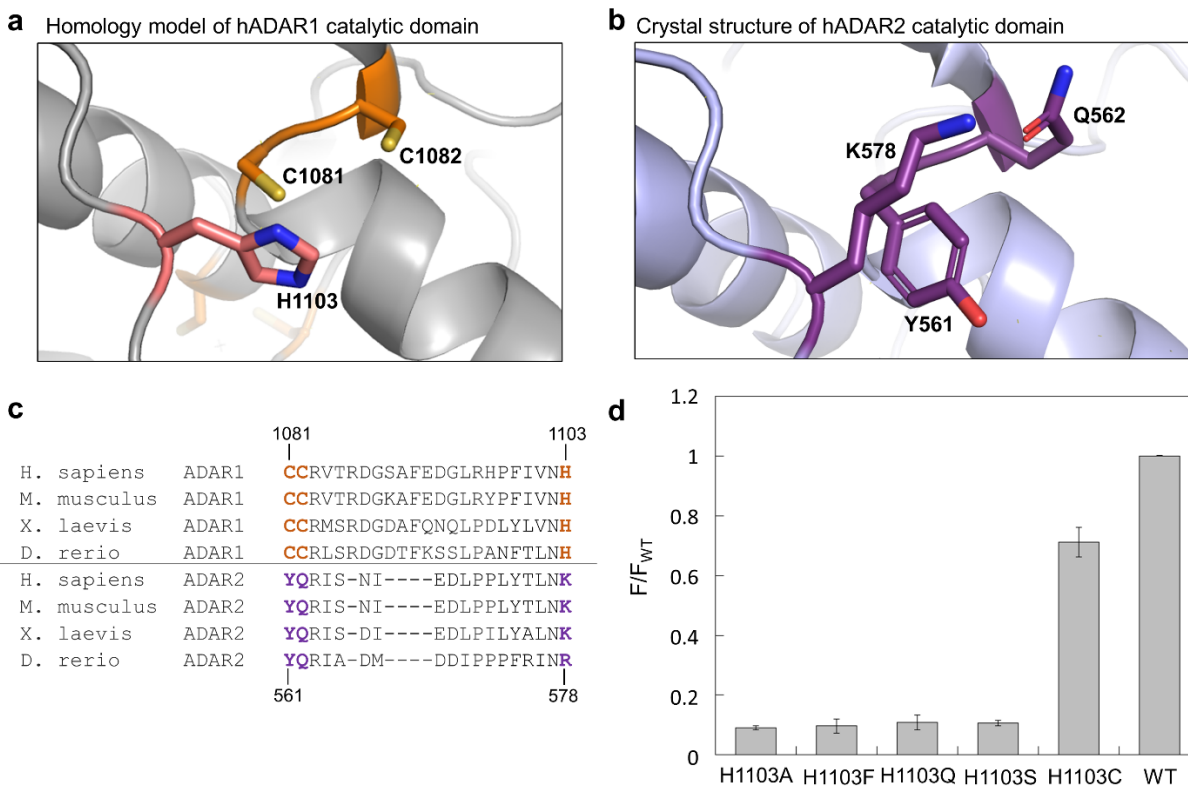


Figure 5.4. Prediction of metal binding residues through a homology model and mutagenesis study. (a) Close-up view of possible second metal binding site from a homology model of hADAR1d generated by SWISS-MODEL web server where C1081, C1082 and H1103 are likely to be involved in a metal binding. (b) Crystal structure of hADAR2d (PDB: 5HP3) with residues corresponding to possible metal binding residues in hADAR1d.^{107,188} (c) Sequence alignment of ADAR1 and ADAR2 from different organisms showing C1081, C1082 and H1103 residues (highlighted in orange) and K561, Q562, and K578 (highlighted in purple) that are corresponding to possible metal binding residues of hADAR1d WT. (d) Fluorescence activated activity assay of hADAR1d H1103 mutants. BDF2-derived fluorescence reporter was used. F/F_{WT} is the ratio of hADAR1d H1103 mutant fluorescence over hADAR1d WT fluorescence.¹⁵⁵

In hADAR2d, Y561, Q562, and K578 residues correspond to hADAR1d C1081, C1082, and H1103 (Figure 5.4 B). In ADAR2, these residues stabilize the fold of the protein by a combination of H-bonding of the K578 ammonium group and the Q562 carboxamide along with hydrophobic interactions between the K578 methylenes and the Y561 phenyl ring (Figure 5.4 B). Therefore, it appears the second metal site in hADAR1d binding is responsible for stabilization of the protein fold using metal binding instead of H-bonding and hydrophobic interactions as seen in the hADAR2d structure.

Homology modeling of hADAR1 catalytic domain.

A high-resolution structure of ADAR1 is not yet available. Nevertheless, Dr. SeHee Park's discovery of a second zinc site and Dr. Yuru Wang's high throughput mutagenesis study of the ADAR1 5'

binding loop have provided useful constraints for modeling the ADAR1 deaminase domain structure.³⁵ Therefore, in collaboration with Tiffany Yue Zhang and Dr. Justin Siegel, we generated a structural model of hADAR1d, using RosettaCM and the experimentally derived constraints from these biochemical experiments.

		969		999
H. sapiens	ADAR1	CGD	GALFDKSCSDRAMESTESRHYPVFENPKQ	GKLRT
M. musculus	ADAR1	CGD	GALFDKSCSDRAVESTESRHYPVFENPKQ	GKLRT
X. laevis	ADAR1	CGD	GALFDKSCSDQPSAEGDNKHCP	FENVKQ
D. rerio	ADAR1	CGD	GALFDKSCSEAAE-LNGSGHMPLFEN	IKQ
<hr/>				
H. sapiens	ADAR2	CGD	ARIFSPHEPILE-----EPADRHPNRKARG	QLRT
M. musculus	ADAR2	CGD	ARIFSPHEPVLE-----EPADRHPNRKARG	QLRT
X. laevis	ADAR2	CGD	ARIFSPHEVGQE-----DQGDRHPNRKARG	QLRT
D. rerio	ADAR2	CGD	ARIFSPHEAGAE-----DQGDRHPNRKARG	QLRT
		454		479

Figure 5.5. Sequence alignment of 5' binding loop of ADAR1 and ADAR2 from different organisms. Green: conserved in both ADAR1 and ADAR2, Red: conserved in ADAR1, Blue: conserved in ADAR2, Black: not conserved.

The deaminase domains of hADAR1 and hADAR2 have high sequence similarity so structures of hADAR2d provide a good starting point for modeling hADAR1d. However, the sequence of the protein loop that interacts with the 5' side of an RNA substrate (i.e. the 5' binding loop) are substantially different in ADAR1 and ADAR2 (Figure 5.5). Due to the size of the loop, roughly 30 amino acids, lack of structural homologs, and no significant predicted secondary structural elements additional data was desirable to reduce the structural search space and direct protein modeling efforts. In Dr. Yuru Wang's previously published Sat-FACS-Seq studies of the ADAR1 and ADAR2 5' binding loops, we noted a striking similarity in selectivity for specific amino acids at three common positions in the two ADARs (i.e. F457, D469 and R477 in ADAR2 and F972, D973 and K996 in ADAR1) suggesting these residues play similar roles in the two deaminase domains (Figure 5.6 A)^{35,190} In the ADAR2 5' binding loop, D469 contacts R477 via an ion pair between the side chains and the phenyl ring of F457 provides a platform onto which the guanidinium group of R477 stacks in an apparent cation- π interaction. Thus, the side chain positions of F972, D973 and K996 in ADAR1 were constrained in our Rosetta modeling to match the distances found between F457, D469 and R477 in structures of ADAR2 (Figure 5.6 B).

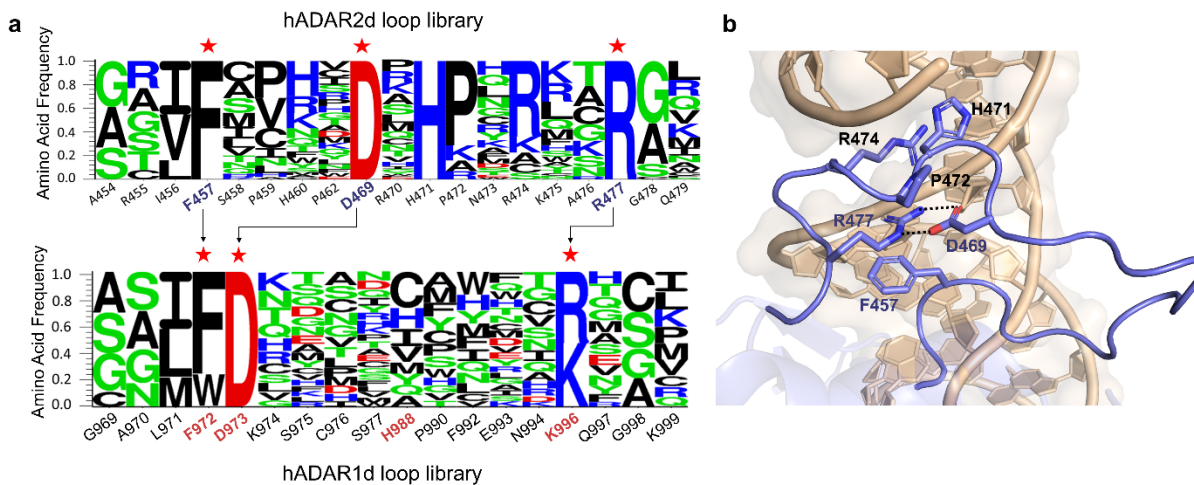


Figure 5.6. Comparison of Sat-FACS-Seq results for the 5' binding loops of hADAR1 and hADAR2. (a) Arrows indicate functionally similar residues in the different loops.^{35,190} (b) Close-up view of interactions within the 5' binding loop observed in the crystal structure of hADAR2d with dsRNA.¹⁰⁷

In addition to structurally conserved residues between hADAR1 and hADAR2, additional constraints were used in modeling to enforce canonical Zinc coordination at the predicted binding site. Zinc ions are typically coordinated by four ligands forming tetrahedral structures.^{191,192} However, the analysis described above only identified three hADAR1 residues involved in the second zinc binding site (C1081, C1082 and H1103), leading to the question of the identity of a potential fourth metal binding residue. Interestingly, Dr. Yuru Wang's previous study of the 5' binding loop of hADAR1 also provided evidence that a residue within that loop could serve this role.³⁵ We observed that mutation of H988 to cysteine within the 5' binding loop of hADAR1d increased deaminase activity in our fluorescent reporter assay in yeast whereas mutation of this residue to each of the other common amino acids reduced activity (Figure 5.6 A).³⁵ Thus, H988 is a promising candidate for the fourth metal ligand given the fact that histidine and cysteine are common zinc-binding amino acids and the 5' binding loop is near residues already implicated in the second zinc site.¹⁹³ Therefore, Rosetta modeling was performed with the four hADAR1 residues C1081, C1082, H1103, H998 constraints in addition to the constraints derived from hypothesized structural homology to hADAR2 between residues F972, D973 and K996. For the residues constrained to the zinc

ion, average distances from known structures containing a zinc ion bound by two cysteine and two histidine residues were used (Table 5.1, Figure 5.7).

Atom 1	Atom 2	Average Distance (Å)	Standard Deviation (Å)
Cys SG 1	Cys SG 2	3.81	0.24
Cys SG 2	His ND 2	3.61	0.33
His ND 1	His ND 2	3.47	0.41
His ND 2	Cys SG 1	3.66	0.23

Table 5.1. Measurements used as constraints for Rosetta modeling. The constraints were defined using average measurements from 20 different PDBs (PDB: 1A1H, 1BBO, 1FRE, 1GUP, 1IA6, 1K6Y, 1LLM, 1ODH, 1RMD, 1SVM, 1WIR, 1WJP, 1X3C, 1X6F, 1X3I, 1UN6, 1V5N, 1YUI, 1ZW8, 2A25) each containing zinc bound by two Cys and two His residues.

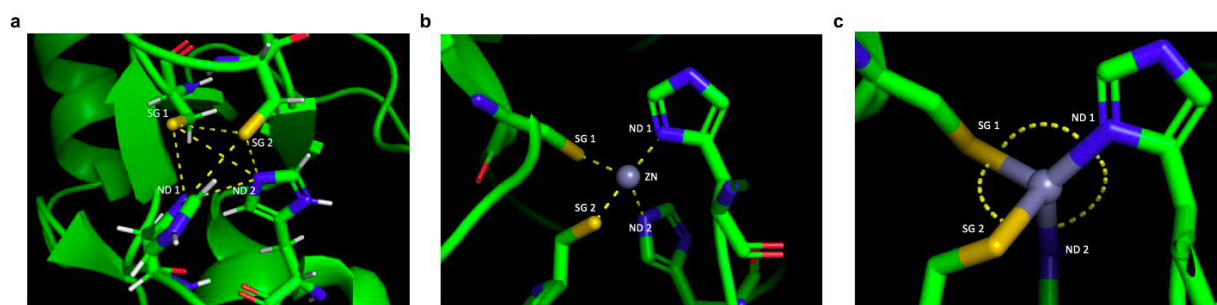


Figure 5.7. Representation of measurement values used as constraints for Rosetta modeling. (a) Distances measured between metal binding residues. (b) Zinc binding distances of metal binding residues. (c) Zinc binding angles of metal binding residues.

A three-dimensional model of the human ADAR1 deaminase domain was then generated using the RosettaCM protocol including the constraints described above.^{191,194–197} The crystal structure of ADAR2 deaminase domain bound to double stranded RNA (PDB ID: 5HP3) was used as the template for generating the model.²⁵ During the modeling, zinc was treated as a ligand and incorporated as the homology models were constructed. A total of five-thousand models were constructed from which the ten lowest energy structures were evaluated (Figure 5.8 showing all 10 overlaid).

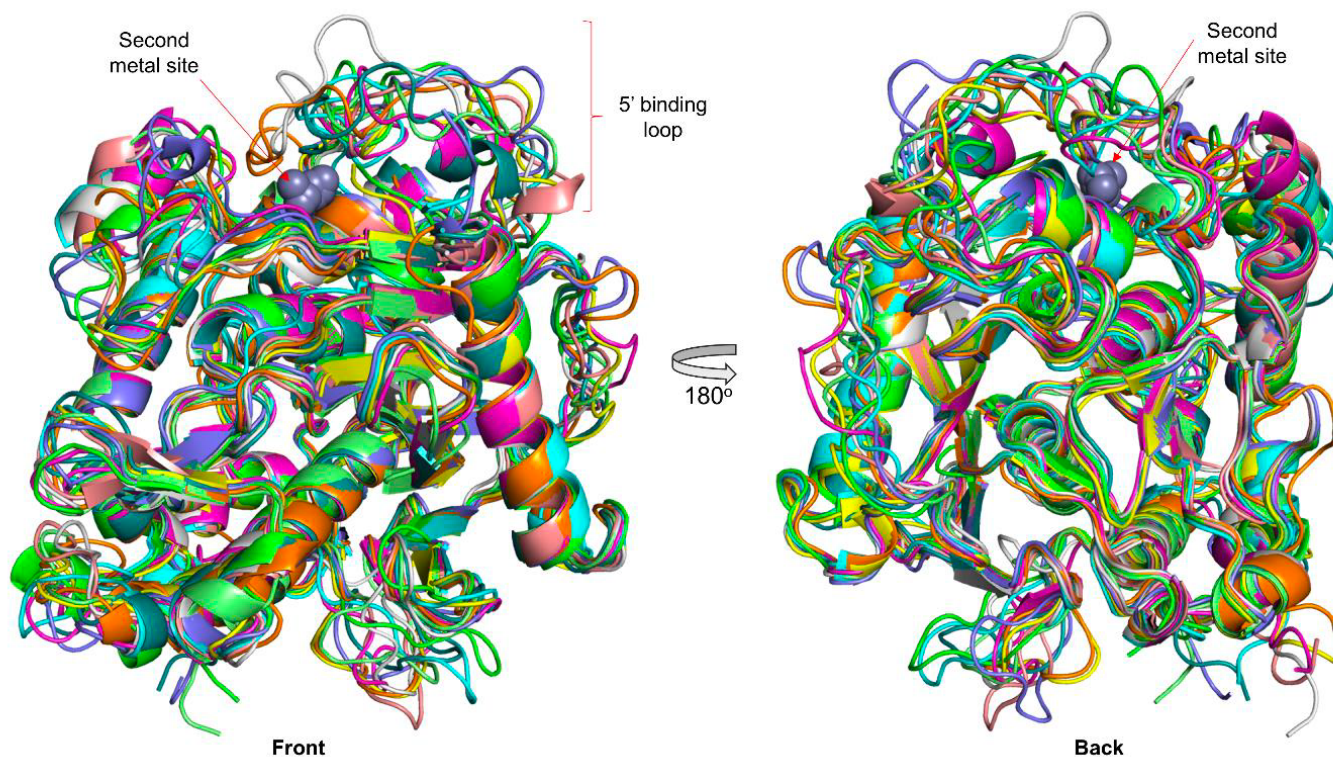


Figure 5.8. An overlay of the top ten lowest energy structures generated by structural modeling. The 5' binding loop region shows divergent conformations, whereas most of core structures show a high convergence. All input files and the lowest 10 energy structures are available in GitHub.

Figure 5.9 shows a representative hADAR1d model superimposed on the previously reported crystal structure of an hADAR2d-RNA complex.²⁵ Among the top 10 lowest energy models, this model shows the most similar conformation of the 5' binding loop interactions observed in hADAR2d crystal structure that were used as modeling constraints (Figure 5.6 B).

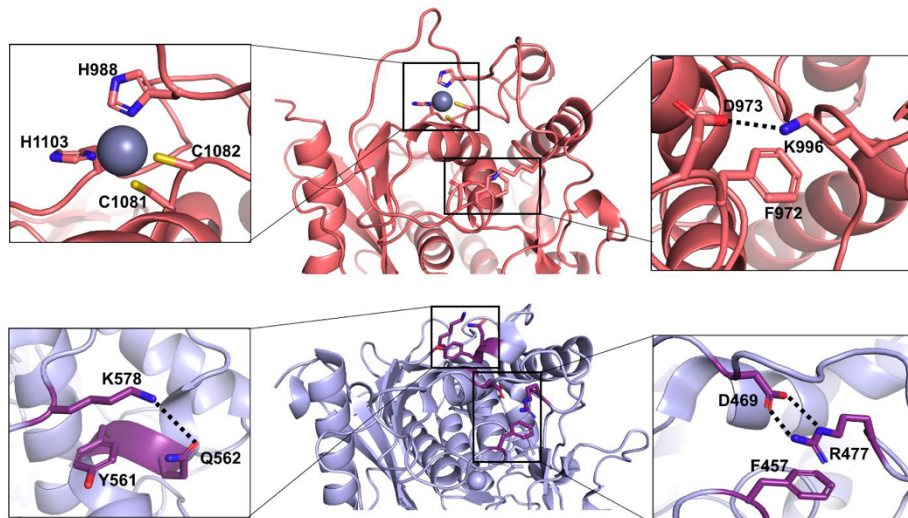


Figure 5.9. Comparison of constraints used to generate structural models of ADAR1d (Top structure in Salmon) to high-resolution structure of ADAR2d 1 (Bottom structure in Purple).¹⁰⁷ (Right) Comparison of interactions of three residues that stabilize the 5' binding loop conformation. (Left) Comparison of the second metal binding site and corresponding residues in hADAR2d crystal structure.

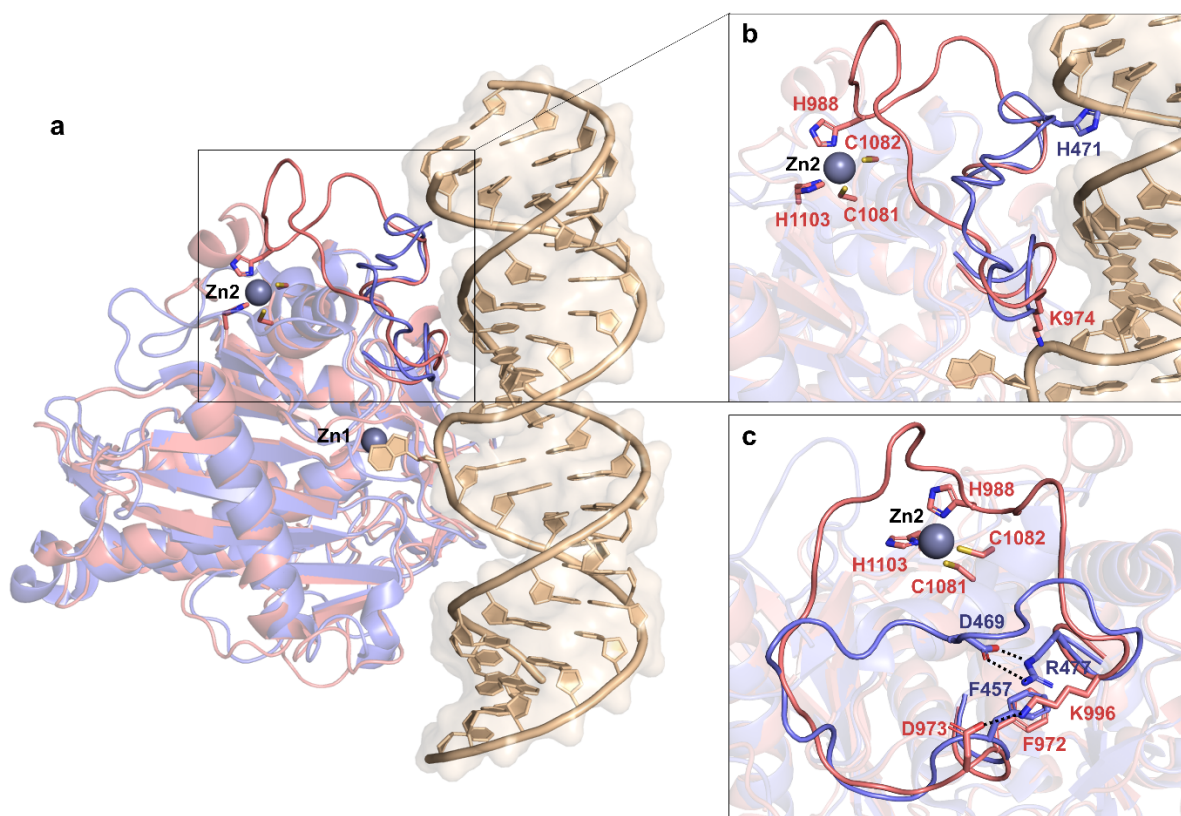


Figure 5.10. Comparison of a hADAR1d model with the crystal structure of an hADAR2d-RNA complex.¹⁰⁷ (a) A representative hADAR1d model (in salmon) is superimposed on crystal structure of a hADAR2d-RNA complex (in blue and RNA in wheat), showing 5' binding loop of each protein and the second metal binding site of hADAR1d. Zn1 is a catalytic zinc at the active site and Zn2 is bound in the second metal site reported here. (b) Close-up of 5'

binding loop of each protein and the second metal binding site of hADAR1d. H471 residue in hADAR2d is an RNA contacting residue seen in the crystal structure of hADAR2d. K974 is a possible RNA contacting residue predicted from the hADAR1d model structure. (c) Close-up comparison of 5' binding loop conformations of hADAR1 and hADAR2 and ionic and cation- π interactions of residues within each of 5' binding loop.

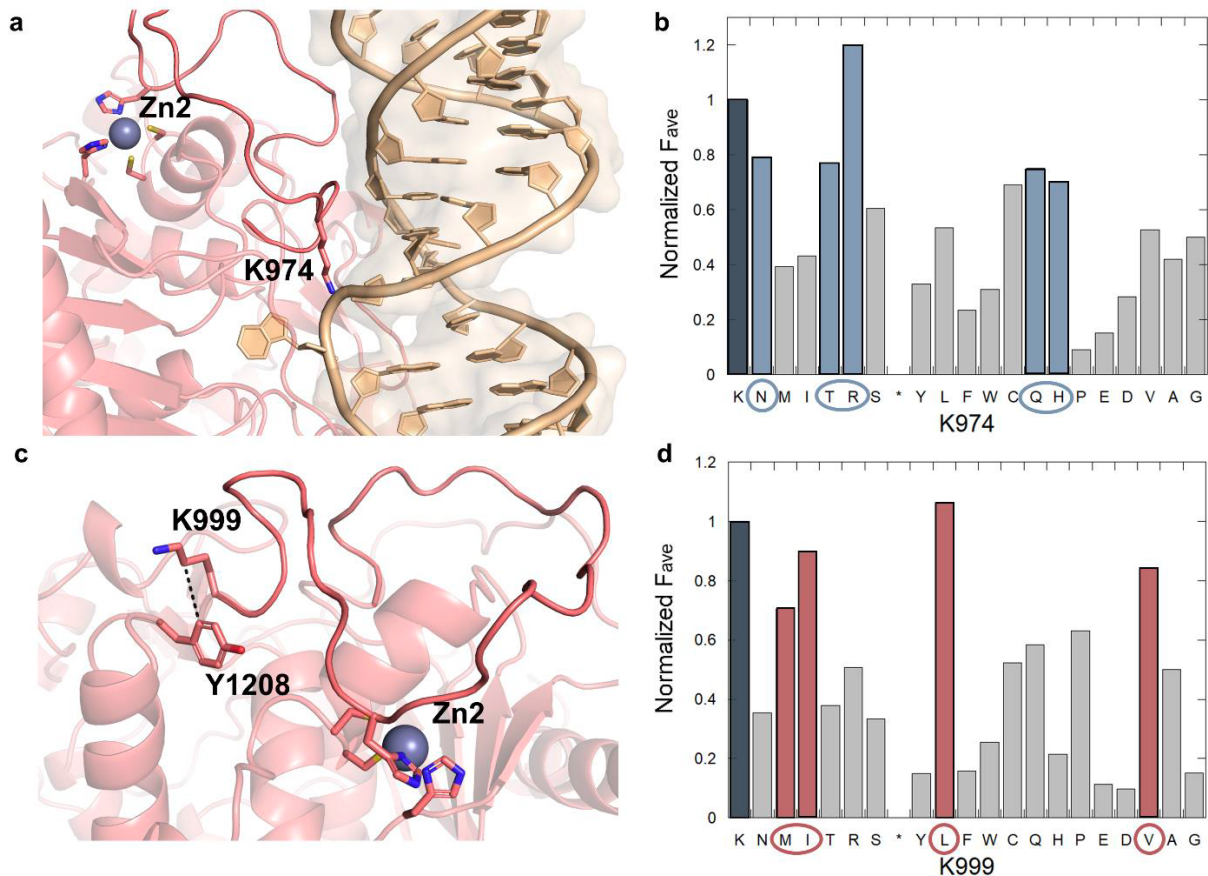


Figure 5.11. Comparison of two lysine residues within the 5' binding loop of hADAR1d. (a) hADAR1d model structure suggests that K974 is a potential RNA contact residue. (b) Previously reported Sat-FACS-Seq data also supports that K974 is an RNA contacting residue because the mutation with polar residues (highlighted in blue) shows a comparable activity to WT.³⁵ (c) hADAR1d model structure suggesting a hydrophobic interaction between K999 and Y1208, stabilizing the 5' binding loop fold. (d) Sat-FACS-Seq data also supports this by showing a preference of hydrophobic amino acids highlighted in salmon.³⁵ Thus, one of Aicardi-Goutières Syndrome (AGS)-associated mutations (K999N) could disrupt the protein structure by introducing a short polar side chain into this hydrophobic site.⁹

This model is consistent with known properties of hADAR1d. For instance, given the positioning of the hADAR1 5' binding loop in the model, most of the amino acids present on this loop are directed away from the likely RNA binding site, unlike the hADAR2 5' binding loop where the side chains of H471 and R474 contact substrate RNA on the edited strand approximately 10 nt 5' to the editing site (Figure 5.10). The hADAR1 deaminase domain readily deaminates RNA substrates with short 5' duplexes (< 10 bp) whereas a longer 5' duplex (> 10 bp) is required for the efficient editing of the hADAR2 deaminase

domain.³⁵ In addition, the 5' binding residues that are not evolutionally conserved among ADAR1 proteins from different species (corresponding to aa978-aa987 in human, Figure 5.6) are located away from the region predicted to be the RNA binding surface in our model. This hADAR1 deaminase domain model also suggests likely roles for specific residues in the 5' binding loop. For instance, K974 within the 5' binding loop is a good candidate for an RNA-binding residue given its proximity to the predicted RNA binding surface (Figure 5.11 B). Also, our model predicts methylenes present on the side chain of K999, a residue mutated in Aicardi-Goutières Syndrome (K999N), is involved in hydrophobic contacts to Y1208, resulting in stabilization of the 5' binding loop fold (Figure 5.11).

Although the model described here provides valuable information on the general course of the ADAR1 5' binding loop and features of the second metal binding site, it should be noted that the ten lowest energy models do not converge to a single conformation throughout the loop (Figure 5.8). The uncertainty for the exact positions for certain residues in this loop highlights the need for additional studies to define high-resolution structures for ADAR1 bound to RNA.

Comparison of ADAR1d homology model and AlphaFold¹⁹⁸ structure.

The ability to predict the 3-D conformation that a protein will adopt solely based on its amino acid sequence has been a prominent unresolved challenge for over 50 years.¹⁹⁹ While we have generated a structural model of ADAR1 based on a homologous protein (ADAR2), many programs attempt to computationally solve structures without fitting to any particular homologous structure. Of note is a recently published program, AlphaFold, which has received considerable attention due to the accuracy of its structural predictions. AlphaFold greatly outperformed existing methods and its accuracy was comparable to experimental structures in many cases.^{198,200} This is a machine learning approach that uses evolutionary, physical, and geometric data about proteins to design its deep learning algorithm.¹⁹⁸ AlphaFold offers an online structure database (<https://alphafold.ebi.ac.uk>) to view nearly one million proteins of interest, and includes a calculated structure of ADAR1 (Figure 5.12). The structure of ADAR1 has obvious weaknesses; the linker regions between domains are represented as large, disordered loops and are indicated by the

software to be parts of the model with low confidence. However, the domains of the protein—the deaminase domain in particular—appear reasonably structured and are ranked at a higher confidence interval by the software.

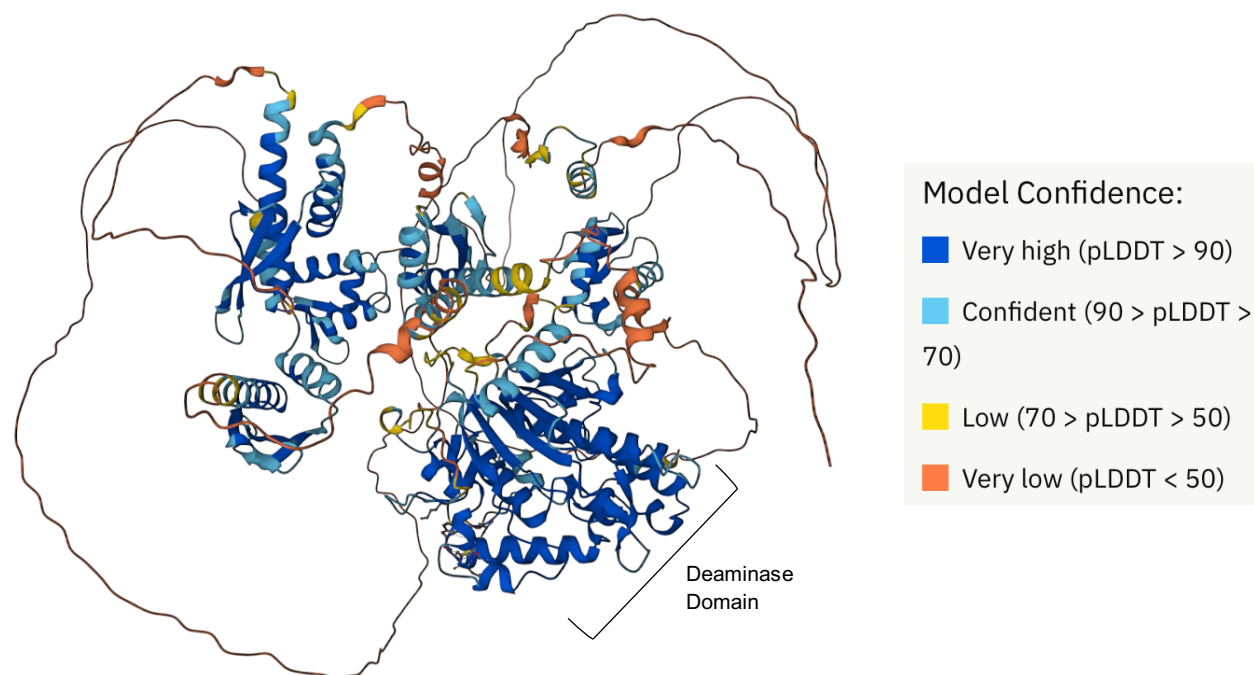


Figure 5.12. The AlphaFold predicted structure for full length ADAR1 p150. Colors indicate the confidence score of the local structure as noted in the legend.¹⁹⁸ Figure obtained from <https://alphafold.ebi.ac.uk> Entry # P55265

With this recent advancement in structural prediction, we sought to compare the outcomes of our homology modeling approach with this sequence-based machine learning system. The full structure of ADAR1 (Figure 5.12) was truncated to include just the deaminase domain residues present in the homology model, and the two structures were overlaid in PyMol (Figure 5.13). The global overlay of both structures is very similar, so we focused on analyzing any notable differences in the regions of the structure that are thought to be important for activity and specificity. During the homology modeling process the low energy models did not converge to a single configuration in the 5' binding loop, so it is unsurprising that there is variation in the 5' binding loop of the homology model and the AlphaFold structure. Despite the variation, there is similarity in the placement of the constrained residues from the homology model: H988, H1103, C1081 and C1082 (Figure 5.13 D). The similarity in the placement of H988 is particularly notable because it is part of the 5' binding loop and is the feature of the loop that orients it back away from the RNA in the

5' direction in our homology model. However, the overall orientation of the loop in the AlphaFold structure is different than that of the representative model (Figure 5.13 B). In our model, the second zinc site is a key feature providing structural constraints to orient the 5' binding loop. AlphaFold does not predict structures for protein-ligand complexes, and therefore the orientation of the 5' binding loop in our homology model of ADAR1d may be better informed.^{198,200}

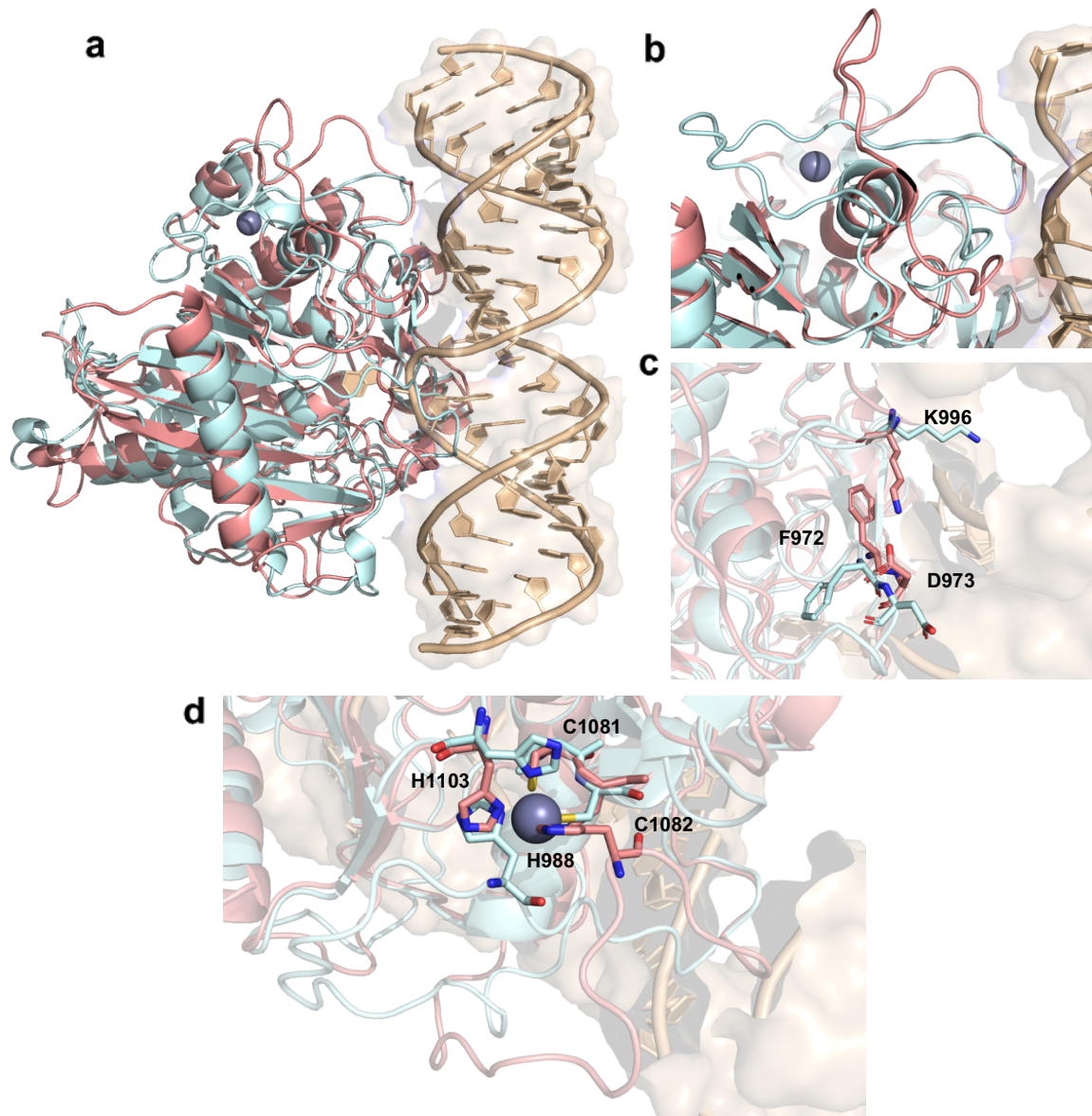


Figure 5.13 A) The AlphaFold prediction of the ADAR1d structure (light blue) overlaid with the homology model of ADAR1d (salmon).^{155,198} B) The 5' binding loop of both models. C) Residues F972, D973, and K996 in either model overlaid. D) Residues H988, C1801, and C1082 in either model overlaid.

There is also deviation between the structures in the other area of the homology model that was informed by biochemical results. The conserved residues in the 5' binding loop identified through Sat-FACS-Seq were constrained to mimic an ionic interaction (D973 and K996) and cation-pi stacking interaction (K996 and F972) observed in the ADAR2 catalytic domain structure.^{107,186} Each of these residues have significantly different conformations in the AlphaFold model versus the homology model (Figure 5.13 C). The AlphaFold model has D973 pointing into the solvent, which is contradictory to Dr. Yuru Wang's activity screening results which indicated that there is an absolute requirement for aspartic acid in this position (Figure 5.6, 5.13 C). Our understanding of the role of these specific residues based on previous biochemical screens likely accounts for the deviation in the structures at this position. Due to the requirement for aspartic acid, the ionic interaction of D973 seen in the homology model seems more likely than a solvent exposed residue as seen in the Alpha Fold model. These results indicate that while machine learning approaches such as AlphaFold have seen dramatic improvements in accuracy in the past few years, biochemical data can still be used to improve predictions as computational constraints.

Homology model of hADAR3 deaminase domain

The homology model of ADAR1d proved useful in making predictions about the roles of specific residues, which led to our interest in modeling the other member of the ADAR family that does not currently have high-resolution structural data available: ADAR3. ADAR3 has a catalytically inactive deaminase domain but may serve a role in regulation of ADAR1 and ADAR2 activity by competitively binding their targets^{78,201,202} Structural information for ADAR3 may help us better predict its substrate preferences. In addition, the ADAR3 deaminase domain shares 50% identity and 70% similarity in amino acid sequence with ADAR2d, making ADAR2d a good model for ADAR3d.¹¹ The Beal and Toney labs previously published a homology model of ADAR3 using the SWISS-MODEL fully automated comparative protein modeling server.^{11,188} However, in direct comparisons Rosetta has been shown to produce more accurate models than SWISS-MODEL based on the root-mean-square deviation (RMSD) values, where lower backbone RMSD values indicate more precise models.²⁰³ Also, the Rosetta method yields the option to add

biochemical constraints to improve model accuracy, as opposed to a fully automated system such as SWISS-MODEL.^{155,188} Therefore, a three-dimensional model of the human ADAR3 deaminase domain was generated using the RosettaCM protocol without additional constraints.^{191,194–197} The crystal structure of ADAR2 deaminase domain bound to double stranded RNA (PDB ID: 5HP3) was used as the template for generating the model.²⁵ A total of five-thousand models were constructed from which the ten lowest energy structures were evaluated (Figure 5.14, showing all 10 overlaid)

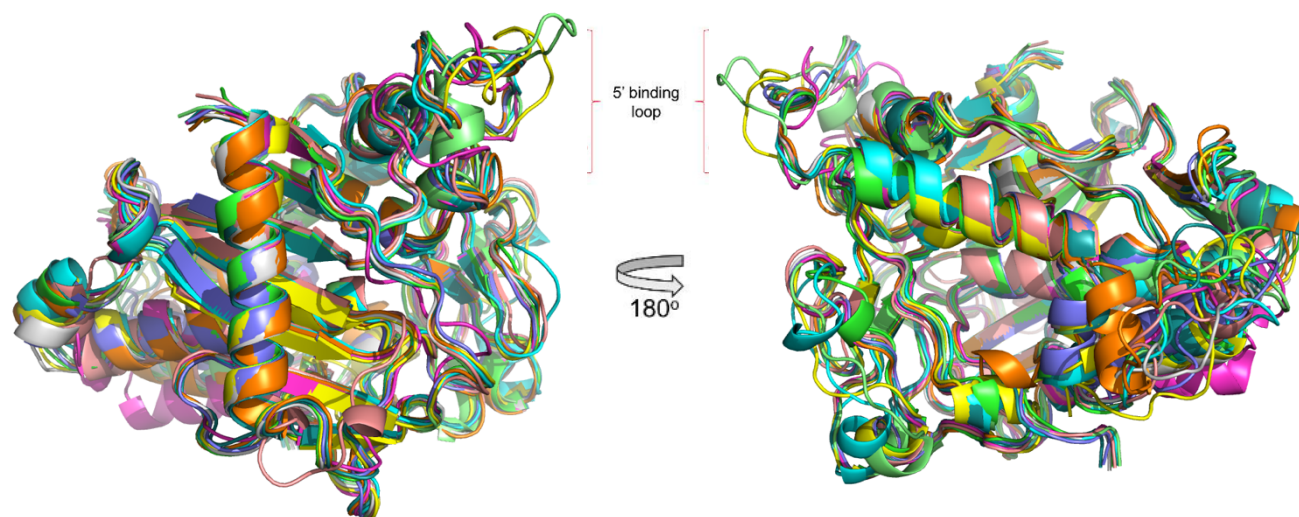


Figure 5.14. An overlay of the top ten lowest energy structures generated by structural modeling of ADAR3d. The overall structure including the 5' binding loop region shows conformations that are less divergent than the ADAR1d lowest energy structural models (Figure 5.7).

The ten lowest energy models show convergence throughout most of the structure (Figure 5.14). In contrast to the model generated of ADAR1d, the 5' binding loop of ADAR3d is not as divergent through the ten low energy models (Figure 5.14). This observation is consistent with the greater similarity of the ADAR3d 5' binding loop with the template structure, ADAR2d (Figure 5.15). The 5' binding loops of ADAR2d and ADAR3d are each comprised of 26 amino acids, while the loop of ADAR1d contains an

extra 5 amino acids. This model may be used as a tool to generate new insight into the structure of ADAR3, especially through further biochemical studies which may provide constraints for more precise modeling.

```

ADAR1      GALFDKSCSDRAMESTESRHYPVFENPKQGK
ADAR2      ARI FSPHEPILE--EPADRH---PNRKARGQ
ADAR3      ARLHSPYEITTD--LHSSKH---LVRKFRGH
           . : . . . . . : * . : * :

```

Figure 5.15. Clustal Omega²⁰⁴ alignment of the 5' binding loop regions of ADARs 1,2, and 3. “*” Indicates complete conservation between all sequences, “:” indicates conservation, “.” is for semi-conservative positions, and unmarked spaces indicate no conservation.

Discussion.

Metal ions are important cofactors that can serve essential roles in protein structure and function.^{205,206} In this work, we carried out molecular modeling using Rosetta and structures of the hADAR2 deaminase domain bound to RNA along with constraints defined from high throughput mutagenesis/functional screening by Dr. SeHee Park (i.e. Sat-FACS-Seq).¹⁵⁵ This supported the discovery of a previously undisclosed zinc binding site within the catalytic domain of human ADAR1 that, along with the zinc ion present in the active site, makes for a total of two known zinc sites in this domain. Given the positioning of the second zinc site, we believe it most likely that this site plays a structural role like other zinc sites found on the surface of proteins.^{207,208} Indeed, our model suggests the second zinc site is important for maintaining the conformation of the 5' binding loop suitable for RNA substrate recognition (Figure 5.9).

Beyond the apparent structural role, the second zinc site in hADAR1 could be involved in additional functions. For instance, it is possible that when each ligating amino acid (H988, C1081, C1082 and H1103) is bound to the second zinc site, the fully active conformation of the 5' binding loop is populated. However, if H988 binding is transient and it releases from the metal site, the 5' binding loop could adopt lower activity conformations. In addition, it is well known that cysteines in proteins are susceptible to oxidation under oxidative stress conditions.²⁰⁹ Cysteine oxidation can lead to protein conformational changes, unfolding, and degradation.²⁰⁹⁻²¹³ Finally, the hADAR1 second zinc site could serve as an interface for protein-protein interactions. Given that zinc ions are often found in metal-mediated protein complexes and the second zinc site in hADAR1 is located on the surface of the protein, this metal

site could stabilize or mediate protein-protein interactions.²¹⁴ Additional studies are required to determine if the zinc site identified here has other regulatory roles or is involved in the formation of protein complexes.

It is important to note the assumptions we made in defining our modeling constraints. First, based on the remarkable similarity in amino acid preference at three positions in the two different ADAR 5' binding loops (e.g F457, D469 and R477 in ADAR2; F972, D973 and K996 in ADAR1) (Figure 5.6), we assumed these amino acids play similar roles in stabilizing the 5' binding loop conformation in each protein.³⁵ Second, we assumed that the side chain of H988 is directly involved in binding to the second zinc in hADAR1. This is based on the Sat-FACS-Seq result at this position (Figure 5.6) previously reported by Dr. Yuru Wang indicating that mutation to cysteine enhanced hADAR1d activity whereas mutation to other residues reduced activity.³⁵ Rosetta modeling is often supplemented with experimentally derived constraints.²¹⁵ For example, chemical crosslinking followed by mass spectrometric analysis (XL-MS) has been used to generate constraints for modeling as well as other structural related experimental data such as NOE measurements from NMR and EPR/DEER measurements.²¹⁶⁻²¹⁹ Our modeling approach is unique in that our constraints were obtained from comparative high-throughput mutagenesis/functional screening instead of structural data alone. The model we arrived at not only supports the biochemical data reported previously, but also allows for analysis of possible roles for different 5' binding loop residues. Understanding interactions of the 5' binding loops present in the ADAR proteins is crucial to understanding ADAR selectivity. However, it has been challenging to predict the RNA binding residues present in the hADAR1 5' binding loop given only hADAR2-RNA structural data. Indeed, our model suggests the 5' binding loop structures diverge from each other substantially. Part of the 5' binding loop of hADAR2 is in position for direct contact to the RNA substrate ~ 10 nt from the editing site, whereas a similar contact with hADAR1 seems unlikely given that a large portion of its 5' binding loop is directed away from the RNA (Figure 5.9). However, the model does predict K974 is a likely RNA contact residue given its proximity to the putative RNA binding site (Figure 5.10 A).³⁵ While additional studies will be necessary to confirm this prediction, the previous observation that hADAR1 editing activity is maintained when this site is mutated

to polar residues (e.g. R, N, Q, T and H) is consistent with K974 functioning in direct RNA binding (Figure 5.10 B). On the other hand, our model predicts the side chain methylenes of K999 contact Y1208, stabilizing the 5' binding loop fold (Figure 5.10 C). This is consistent with Dr. Yuru Wang's previously published mutagenesis data indicating that K999 can be mutated to several different large hydrophobic residues (e.g. I, L, M and V) and editing activity is maintained (Figure 5.10 D). This also suggests that the AGS-associated mutation (K999N) disrupts the protein structure by introducing a short polar side chain into this hydrophobic site.

Methods.

Homology modeling of hADAR1d WT.

A three-dimensional model of hADAR1 catalytic domain was generated using the RosettaCM protocol.¹⁹⁴ The crystal structure of hADAR2 deaminase domain bound to double stranded RNA (PDB: 5HP3) was used as the template for generating the model.²⁵ Promals3D was used to generate a sequence alignment of the template and query sequence to correlate sequence position to structure.¹⁹⁶ First, the query sequence was threaded onto the template sequence, resulting in a threaded partial model. To fill in unaligned regions, Monte Carlo sampling was used to generate Rosetta de novo fragments.¹⁹⁵ The scoring function of this sampling was a combination of the Rosetta low-resolution energy function and distance constraints from the template structure. Second, Monte Carlo sampling was again used for full backbone minimization. During the third stage, the structure underwent full-atom refinement for side-chain optimization and refinement of the side-chain and backbone conformations. During the third stage, side-chain constraints were used to enhance sampling.¹⁹⁷ The conserved residues in the 5' binding loop identified through Sat-FACS-Seq were constrained to mimic an ionic interaction (D973 and K996) and cation-pi stacking interaction (K996 and F972) observed in the ADAR2 catalytic domain structure^{25,35} Furthermore, zinc was treated as a ligand and H988 was defined as the fourth metal binding residue when the structural models were constructed. The metal binding constraints were defined using average distances and angles from 20 different PDBs found from MetalPDB each containing zinc bound by two Cys and two His residues (Table

5.1, Figure 5.7).^{166,220} Using each protocol, 5000 decoys of hADAR1d structure were generated and the 10 lowest energy structures were evaluated. All input files and the lowest 10 energy structures are available in GitHub at https://github.com/siegel-lab-ucd/Publication_Tiffany/tree/master/High-throughput%20Mutagenesis%20Reveals%20Unique%20Structural%20Features%20of%20Human%20ADAR1

Base-Flipping Assay with ADAR1d E1008Q.

Oligonucleotides of comprised of canonical bases were purchased from Dharmacon. Oligonucleotides containing the 8-azanebularine modified base were synthesized on an ABI 394 oligonucleotide synthesizer, and deprotected and purified by 18% PAGE as previously described.¹³¹ PAGE purified top and bottom strands were annealed for a final concentration of 15 μ M edited strand, 30 μ M guide strand, 30 mM Tris-HCl, 6% glycerol, 120 mM KCl, 3 mM EDTA, 0.006% NP-40, and 0.6mM DTT. The mixture was heated to 95°C for 5 min, and slowly cooled to room temperature.

Fluorescence measurements were performed using a CLARIOstar microplate reader and a Nunc MaxiSorp 384-well black bottom plate. Excitation was at 320 nm and fluorescence emission was scanned from 340 to 430 nm with 0.2 nm resolution. Spectra were obtained for solutions containing 1 μ M RNA, with or without 5 μ M ADAR2, in 15 mM Tris-HCl pH 7.5, 4% glycerol, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% Nonidet P-40, 160 U/mL RNAsin, 1.0 μ g/mL yeast tRNA, and 0.7 mM DTT at room temperature.¹³⁰ The background fluorescence of the buffered enzyme solution was subtracted from the spectrum of the complex, and the background fluorescence of the buffer alone at each pH was subtracted from the RNA. Each spectrum is an average of three independent measurements that were LOWESS fit using Graphpad Prism software. The fluorescence intensity values at λ_{max} were used to determine the fluorescence enhancement by ADAR in the formula $FE = (FI_{ADAR-RNA} - FI_{RNA})/FI_{RNA}$ where FE is fluorescence enhancement, and FI values are the fluorescence intensity of samples containing either RNA or RNA in the presence of ADAR2.

Homology modeling of hADAR3d WT.

A three-dimensional model of hADAR3 catalytic domain was generated using the RosettaCM protocol.¹⁹⁴ The crystal structure of hADAR2 deaminase domain bound to double stranded RNA (PDB: 5HP3) was used as the template for generating the model.²⁵ Promals3D was used to generate a sequence alignment of the template and query sequence to correlate sequence position to structure.¹⁹⁶ First, the query sequence was threaded onto the template sequence, resulting in a threaded partial model. To fill in unaligned regions, Monte Carlo sampling was used to generate Rosetta de novo fragments.¹⁹⁵ The scoring function of this sampling was a combination of the Rosetta low-resolution energy function and distance constraints from the template structure. Second, Monte Carlo sampling was again used for full backbone minimization. During the third stage, the structure underwent full-atom refinement for side-chain optimization and refinement of the side-chain and backbone conformations. This allowed for the generation of 5000 decoys of the hADAR3d structure and the 10 lowest energy structures were evaluated.

CHAPTER 6

Efforts toward the Optimization of Oligonucleotides for Therapeutic RNA Editing

The following chapter details ongoing efforts in the laboratory. The work on phosphorothioate positioning is in collaboration with Synthego Corporation who synthesized the guide oligonucleotides.

Introduction.

RNA editing enzymes such as Adenosine Deaminases Acting on RNA (ADARs) can be directed to edit specific sequences via hybridization of a guide oligonucleotide to a target sequence.^{45,136-138} In double-stranded RNA (dsRNA) ADARs deaminate adenosine to inosine: a base with hydrogen bonding properties similar to guanosine.⁷⁹ Therefore, ADAR has the ability to affect a functional A-to-G change in RNA. This activity can be directed to a target sequence with the use of a guide RNA that is complementary to the target sequence. The effects of RNA editing by ADAR could be used to repress or enhance translation, skip or induce an exon, alter epitranscriptomic modifications, and activate or inactivate enzymes.⁶ Expanding the domain of sequences that can be efficiently targeted by RNA editing will aid in achieving these numerous possibilities. Our preliminary work described in *Chapters 2 and 3* probed a specific contact with the guide RNA that was important for editing efficiency of the ADARs. *Chapter 4* applies the use of guide RNA chemical modifications for enabling editing within the most challenging sequence contexts: Those that contain a guanosine 5' to the editing site. This chapter will focus on other non-ideal targets that are disease relevant, while expanding our understanding of ADAR's tolerance for chemical modifications within the guide strand that are critical for cellular editing.

I. RNA Editing as a Potential Chemotherapeutic

ADARs prefer to edit within A-form helices at an adenosine within a 5'-UAG-3' sequence.²²¹ For this reason, the field of directed editing via ADARs has chosen to focus on disease-relevant targets within this context. However, editing of RNA can be observed in other non-preferred sequences with reduced efficiency. There are numerous factors that govern ADAR editing. In the position directly 5' to the editing

site ADARs prefer U=A>C>G.³³ While it is particularly challenging to edit when a guanosine is present at the 5' position (discussed further in *Chapter 4*), editing is commonly observed when the codon contains an adenosine 5' to the editing site. Therefore, we chose to identify novel therapeutically relevant targets containing an adenosine in the 5' position.

Of relevance to oncology, RNA editing of lysine codons (AAA and AAG) can be used to induce an inactivating mutation in protein kinases (Figure 6.1). The development of new strategies to inhibit protein kinase activity for the treatment of cancers is critical, since the efficacy of small molecule inhibitors is commonly compromised by drug resistance mechanisms.²²² Conventional small molecule therapeutics also suffer from selectivity challenges and usually require that the target kinase has deep grooves for binding.⁸ RNA editing may expand the scope of druggable protein kinase targets, since the mRNA sequence is being targeted, rather a unique binding site in a given protein. Therefore, RNA editing offers a unique platform for targeting protein kinases that can easily be adapted to different targets. Rational design of small molecule inhibitors requires extensive structural data of the protein kinase target, and high-throughput screening to identify inhibitors is costly and not guaranteed to yield a promising result. However, oligonucleotide-based therapeutics can be reprogrammed to different targets by simply changing the guide sequence. Therefore, we chose the mRNA of the proto-oncogene tyrosine protein kinase Src (*SRC*) as a model protein kinase target within an AAA (lysine codon) sequence context.²²³

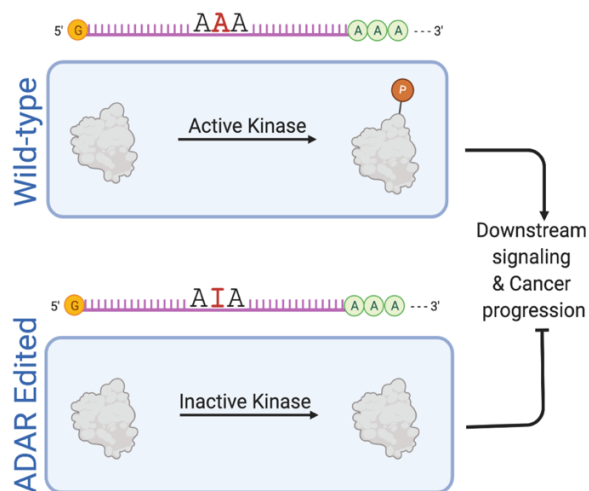


Figure 6.1. ADAR editing of protein kinase mRNA results in inactive kinases and therefore inhibits protein phosphorylation and cell signaling.

Results.

Editing of the Catalytic Lysine Codon in SRC mRNA is Supported by ADAR2

We had previously observed off-target editing within an AAA sequence context at high concentration of ADAR2, which suggested that editing within an AAA (lysine) sequence context would be facilitated. We synthesized a chemically modified 37-mer guide oligonucleotide for *in vitro* studies to target the *SRC* mRNA that contained the necessary chemical modifications for cellular studies (Figure 6.2 A). When 10 nM *SRC* target was hybridized with the chemically modified guide oligonucleotide and reacted with 100 nM wild type ADAR2, deamination was seen to an endpoint of > 40% at $k_{\text{obs}} = 0.13 \pm 0.06$ (Figure 6.2 B). In addition, no bystander editing of other sites within the region complementary to the guide was observed, demonstrating that editing is very specific to the target site (Figure 6.2 C). This rate is comparable or greater than other substrates *in vitro* so we decided to next determine if this guide oligonucleotide can support editing of the *SRC* mRNA in human cells.

a dN DNA
N - RNA
[N] - 2' O-Me
* - Phosphorothioate backbone
A - Edited Adenosine

```
5' - [ G * A * C * A * T * C * G T G C C A G G C * T * T * C * A * G G G T T ] d C d T d G [ A T G G * C * C * A * C * C ] - 3'
3' - C T G T A G C A C G G T C C G A A G T C C C A A C T A C C G G T G G - 5'
```

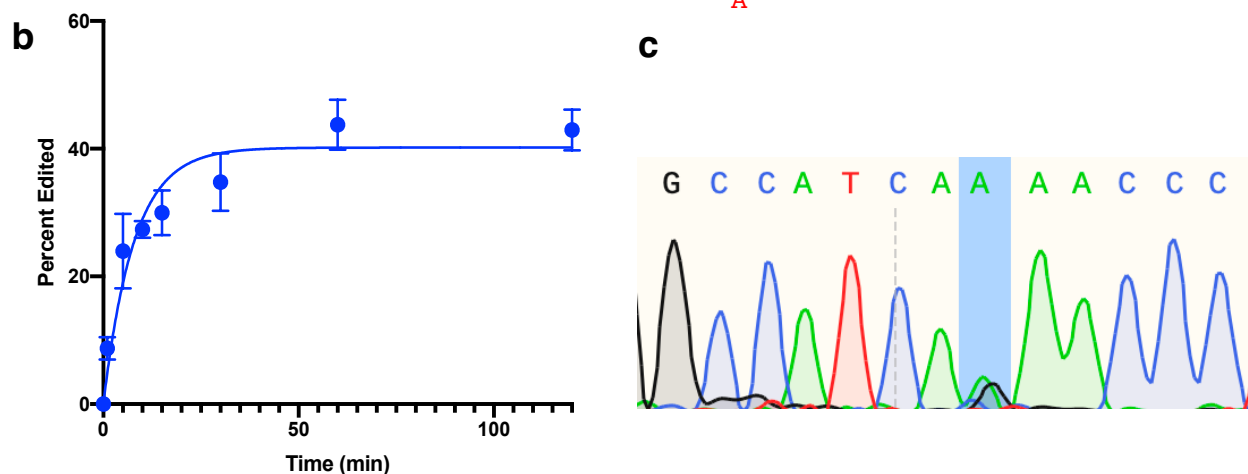


Figure 6.2. Editing of the *SRC* mRNA is supported by ADAR2. A) Guide oligonucleotide design used to hybridize to the *in vitro* target. B) Reaction progress curve showing deamination of 10 nM of the target duplex by 100 nM

ADAR2. C) Electropherogram from Sanger sequencing of cDNA from the deaminated target showing editing only at the desired site.

Cellular Editing of the SRC mRNA in HEK293T Cells with Overexpressed ADAR2

To determine how our chemically modified guide oligonucleotide would support editing via ADARs in a cellular context, we transfected HEK293T cells with *SRC* guide oligonucleotide and a plasmid for ADAR2 overexpression. Although the goal of these studies is to support editing via endogenous ADARs, overexpression allows for amplified editing signal as we optimize our guide oligonucleotide design. Sequencing of the *SRC* cDNA revealed that the target adenosine was edited up to $23\% \pm 15\%$ (Figure 6.3). Although this level of cellular editing is relatively low, it demonstrates that this target can be selectively edited in human cells, and future studies will aim to optimize the guide oligonucleotide design to allow for more efficient editing.

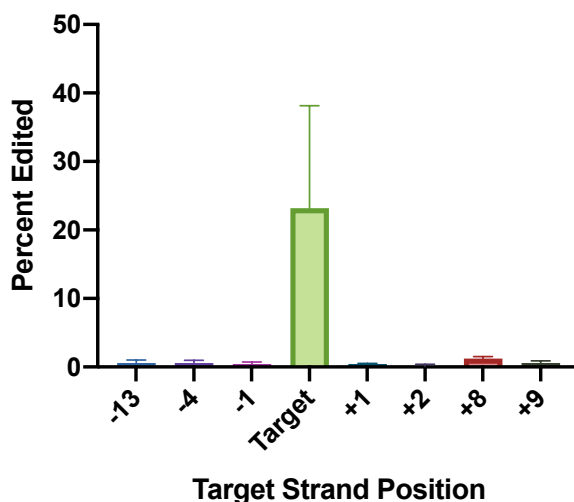


Figure 6.3. Cellular editing of positions in the *SRC* mRNA complementary the guide strand shows selective editing at the target position.

II. A Chimera Allows for High-Throughput Screening of Sequences that Enable Editing in DNA/RNA Hybrids

In addition to the neighboring sequence preference of ADARs, their specificity for dsRNA comes from recognition of an overall A-form helical structure.⁷⁹ Helices comprised of complementary DNA and RNA strands also exhibit an A-form helix structure.¹³⁰ Dr. Eric Zheng, a previous Beal lab member, showed that editing of adenosines within the DNA strand of these DNA/RNA hybrids is possible.¹³⁰ This discovery

opened the door to our developing understanding of the biological relevance of editing within DNA/RNA hybrids. DNA/RNA hybrids are found endogenously as part of R-loops, which result from an invading RNA pairing with one strand of a DNA helix, and ADAR is seen to play a role in R-loop regulation.^{224,225} This understanding also gave rise to the idea of using ADARs to produce therapeutic DNA edits in addition to RNA.

Current approaches to treating disease with genetic origins aim to correct disease-causing mutations in nucleic acid sequences. CRISPR-Cas-mediated editing of DNA (genome editing) is a popular approach used to induce sequence specific changes.^{156,226} However, the effort to move CRISPR from the laboratory into the clinic has been hampered by unforeseen challenges. Two prominent hurdles have been delivery of the editing system and its tendency to elicit an immune response.^{227–229} Although CRISPR research continues unabated, some interest has turned to systems that avoid these issues by utilizing human enzymes to edit disease-relevant nucleic acids.⁴⁰ Therefore, the use of ADARs to produce site-specific single nucleotide changes in DNA is an attractive alternative.^{60,130} However, ADAR editing of DNA is hindered by the removal of specific contacts made between the enzyme and 2'-hydroxyls in the target RNA as seen in the ADAR2d crystal structure, and directed editing yields *in vitro* are low.¹³⁰ Therefore, to produce therapeutically relevant levels of editing within DNA targets, it is necessary to optimize the guide strand of the RNA. Here we employ a new method that uses Next Generation Sequencing (NGS) to identify guide strand sequences that enhance editing in the DNA strand of a DNA/RNA hybrid.

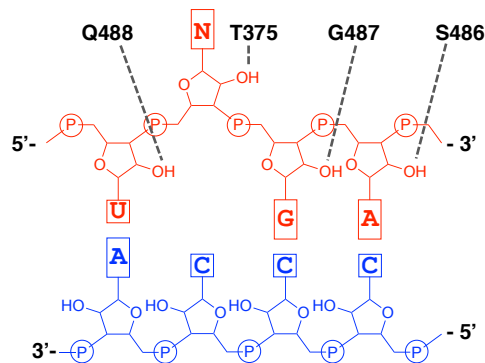


Figure 6.4. ADAR2 contacts with 2'-hydroxyls on the target strand of RNA as seen in the crystal structure of ADAR2d bound to dsRNA.^{107,130}

Results.

Design of a DNA/RNA Chimera as an ADAR Substrate

To screen a large number of guide RNA sequences to determine those that best promote editing within a DNA strand, we used a new method called EMERGE (En Masse Examination of RNA Guides) pioneered by Beal lab member Casey Jacobsen. The EMERGE method uses a chemically synthesized oligonucleotide target containing a randomized region across from the edit site. The region across from the edit site was chosen because it will likely have the most influence on editing. The other important feature of the substrate is that it is a hairpin structure: Folding back on itself so that the randomized region and the target adenosine are within the same strand. This allows editing of the target adenosine to be directly linked to the guide sequence enabling the editing when sequenced via NGS. This screen is being used by Casey Jacobsen to find sequences that enable editing at difficult targets such as those containing a guanosine 5' to the edit site. He typically obtains 6 million NGS reads encompassing over 600,000 possible guide sequences.

We sought to use this method to screen for guide RNA sequences that allow for editing within a DNA strand. Because a central aspect of this method is that the target and guide sequences are linked (allowing for editing to be linked to a particular guide) this would necessitate the use of a DNA/RNA chimeric oligonucleotide. We designed a 90mer DNA/RNA chimera based on the *GLII* sequence which was used by Dr. Eric Zheng (Figure 6.5). This was chosen as a substrate for the initial screening although it is not disease relevant, because patterns of enabling guide sequences could be identified and applied to DNA containing pathogenic mutations. This substrate contained primer binding sites on the 5' and 3' ends, a fixed *GLII* DNA sequence containing the target adenosine, loop residues to allow for the hairpin turn and the complementary *GLII* sequence containing a randomized region of 10 nucleotides (Figure 6.6).

Although the self-complementary (*cis*) substrate enables this EMERGE screen, therapeutic DNA editing would require a *trans* system where a guide RNA is hybridized to the target DNA. Therefore, we designed a *trans* system comprised of a 90 mer DNA and 30 mer guide RNA to test the winning sequences in a more therapeutically relevant context. Efforts to evaluate the selected sequences in this format are ongoing in the lab.

III. A Phosphorothioate Walk Demonstrates ADARs Tolerance to Backbone Modifications

To produce edits for any target sequence in a cellular or *in vivo* context, the guide oligonucleotide must be stable in these conditions. The 2'-hydroxyl of ribose-containing oligonucleotides presents a metabolic liability. Therefore, RNA molecules typically do not exhibit good serum stability, which leads to the use of chemical modifications in many therapeutic RNAs to decelerate their degradation.^{69,70,230} Commonly, this includes 2'-deoxy, and 2'-O-methyl sugar modifications in addition to phosphorothioate backbone modifications. Recent studies harnessing ADAR editing activity in human cells and in non-human primates have focused on using short (~30-38 mer), chemically modified oligonucleotides to achieve editing.^{68,131} Each of these guide designs employed backbone chemical modifications for cellular stability. However, while these modifications contribute to longevity of these oligonucleotides in cells and *in vivo*, replacing phosphodiester groups with backbone modifications can hinder enzyme activity if they are used in positions where the enzyme contacts the backbone of the RNA. Therefore, we sought to systematically examine the effect of phosphorothioate modifications on different positions of guide oligonucleotides.

Oligonucleotides for directed editing of the *APP* target were designed with three phosphorothioate linkages in successive positions down the guide (Figure 6.7). By doing a “walk” down the length of the guide we would be able to determine if specific positions are less tolerant to the chemical modification and could design a mixed-backbone guide that has phosphorothioate only in optimized positions. We were able to make predictions on which positions may be less tolerant to backbone modifications based on the crystal

structure of ADAR2 deaminase domain and one its double stranded RNA binding domains (ADAR2 RD) bound to dsRNA.¹² This structure allowed for the visualization of protein-backbone contacts (Figure 6.7).

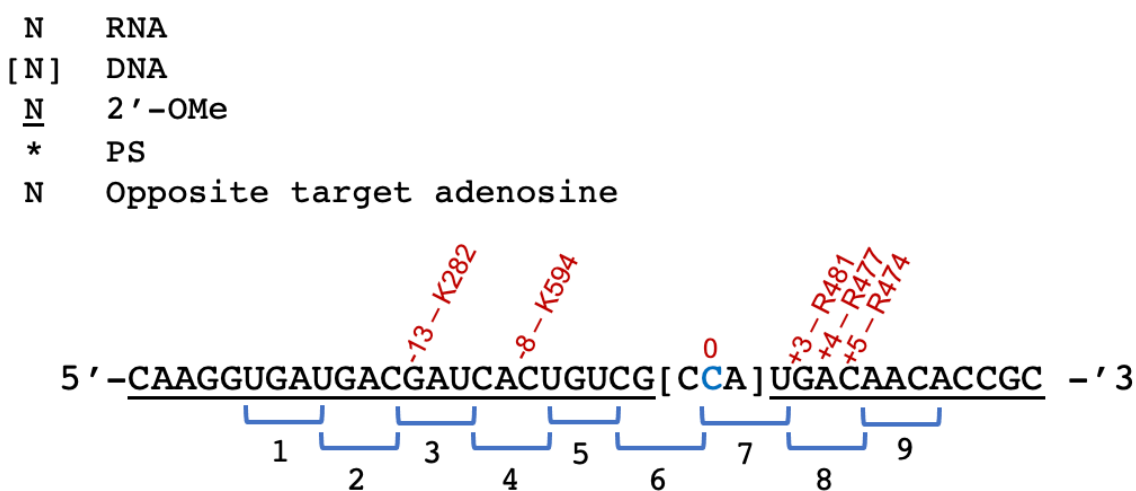


Figure 6.7. The 38mer *APP* guide sequence with 3 nucleotide phosphorothioate positionings annotated. Guides will be synthesized containing a phosphodiester backbone, with phosphorothioate modifications within one of these blue bracketed positions at a time. Red text indicates possible protein-backbone contacts based on the ADAR2 RD crystal structure.

Results.

A Fully Phosphorothioated Guide Oligonucleotide Hinders Editing

This work was a collaboration with Synthego Corporation to optimize guide oligonucleotide chemical modifications that allow for efficient ADAR editing. Synthego made 11 guides to test the effect of phosphorothioates within each of the annotated positions in Figure 6.7. This included a successive block of phosphorothioates within each annotated position (9 guides total), as well as control oligonucleotides that are fully unmodified (containing only phosphodiesters) or completely phosphorothioate modified. First, the two control guide oligonucleotides were tested to determine the magnitude of difference between their reaction with ADAR2 *in vitro*. There was a 1.7-fold difference in the rate of reaction for the more quickly deaminated phosphodiester backbone guide ($k_{\text{obs}} = 1.2$) versus the phosphorothioate guide ($k_{\text{obs}} = 0.70$) (Figure 6.8 A). In addition, the endpoint of deamination for the phosphorothioated oligonucleotide was suppressed to only 14%, compared with 78% for the phosphodiester guide (Figure 6.8 B). We suspect

that the reduced endpoint is due to poor engagement of ADAR2 with the phosphorothioated oligo. This showed that indeed the ADAR reaction is negatively affected by backbone modifications, and there exists a resolution between deamination of the two guides that should allow for the identification of smaller differences caused by phosphorothioates in specific positions.

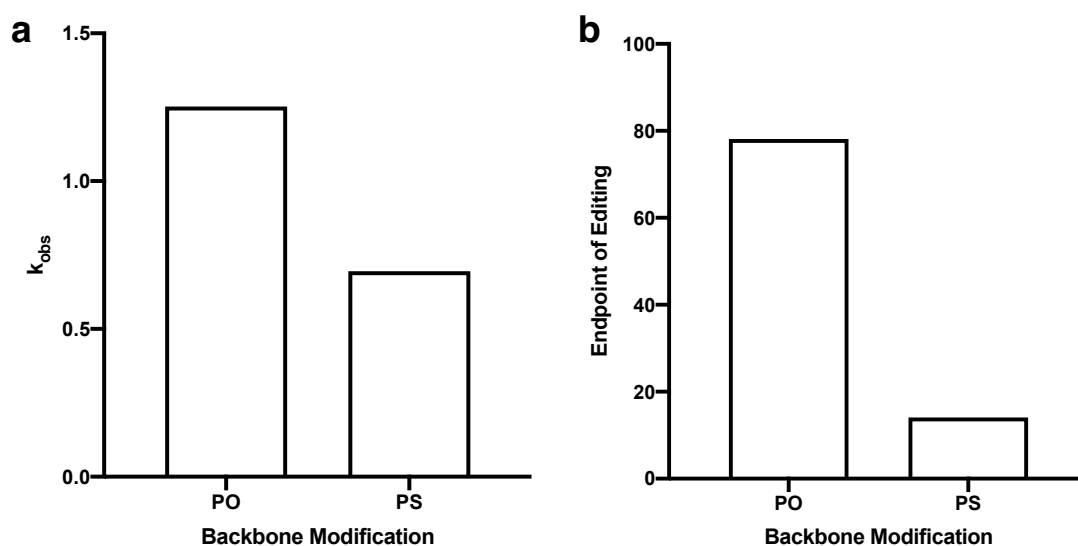


Figure 6.8. Results of the *in vitro* deamination of 10 nM of either control guides either having a fully phosphodiester (PO) backbone or a fully phosphorothioate (PS) backbone with 100 nM ADAR2. A) The observed rate of reaction for either guide. B) The fitted endpoint of editing for either guide.

ADAR2 is Tolerant of Short Stretches of Phosphorothioate Modifications Throughout the Guide Oligonucleotide

To determine which backbone positions are responsible for decreases in editing efficiency when modified to phosphorothioates, we used guides labelled PS 1-9, with the number corresponding to the location of the three phosphorothioates as annotated in Figure 6.7. After subjecting 10 nM of a duplex containing each guide to 100 nM ADAR2 *in vitro* the observed rate of editing and endpoint of editing were plotted (Figure 6.9). It was expected that guide PS 8 would have the largest detrimental effect on editing activity, due to the presence of 3 backbone contacts according to the crystal structure of ADAR2 RD.¹² In addition, both guides PS 3 and PS 4 were anticipated to disrupt protein-backbone contacts. However, our results show that although the fully phosphorothioated guide had a significant effect on editing, small groupings of three phosphorothioates in a row did not have a pronounced effect on the endpoint of editing

in any particular position (Figure 6.9 A). The most significant result is the depression in rate of reaction for PS 1, 8, and 9 (Figure 6.9 B).

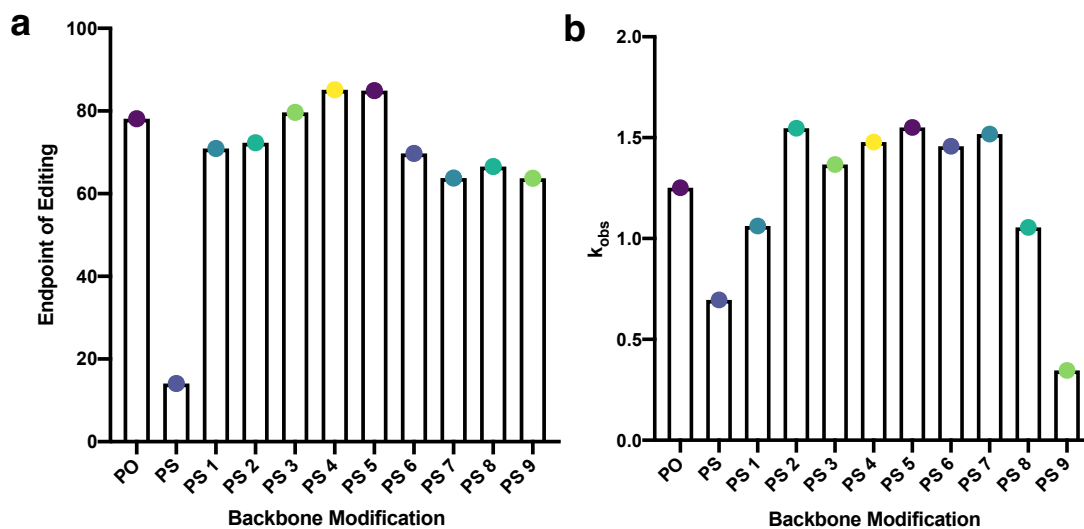


Figure 6.9. Results from the *in vitro* deamination of 10 nM of duplex containing the APP target and guide RNAs with differing backbone modification patterns where PO is phosphodiester, PS is phosphorothioate, and the numbering indicates the position as annotated in Figure 6.7.

This result could indicate that while completely phosphorothioated guide oligonucleotides inhibit ADAR2 editing, there is flexibility in the exact binding register, allowing for ADAR2 to slightly adjust the positioning of the double stranded RNA binding domains to accommodate small regions of phosphorothioates on the guide. It remains to be seen how these modifications affect editing by ADAR1, which is primarily targeted for *in vivo* guide oligonucleotide design due to its more ubiquitous expression in humans.

Methods.

General Biochemical Procedures.

Molecular-biology-grade bovine serum albumin (BSA), and RNase inhibitor were purchased from New England BioLabs. SDS-polyacrylamide gels were visualized with a Molecular Dynamics 9400 Typhon phosphorimager. Data were analyzed with Molecular Dynamics ImageQuant 5.2 software. All MALDI analyses were performed at the University of California, Davis Mass Spectrometry Facilities using

a Bruker UltraFlex extreme MALDI TOF/TOF mass spectrometer. Oligonucleotide masses were determined with Mongo Oligo Calculator v2.08. Oligonucleotides for sequencing and PCR were purchased from Integrated DNA Technologies. Phosphorothioated oligonucleotides were synthesized by Synthego Corporation. The DNA/RNA chimera was purchased from GeneWiz (Azenta). All other oligonucleotides were synthesized as described below.

Synthesis of oligonucleotides.

Chemical synthesis for all other oligonucleotides was performed using an ABI 394 synthesizer. All bases were purchased from Glen Research. Nucleosides were incorporated during the appropriate cycle on a 0.2 μmol scale; See Methods for sequences. Upon completion of the synthesis, columns were evaporated under reduced pressure for 4 h. All oligonucleotides were cleaved from the solid support by treatment with 1:3 ethanol/ 30% NH_4OH at 55 °C for 12 h. The supernatant was transferred to a new screw-cap tube and evaporated under reduced pressure. Desilylation was performed by resuspending the pellets in anhydrous DMSO, and treating with 55% (v/v) Et_3N -3HF at room temperature overnight. To each reaction was added 75 mM sodium acetate in butanol. The oligonucleotides were then precipitated from a solution of 65% butanol at -70 °C for 2 h. The solution was centrifuged at 13,000 rpm for 20 min, supernatant was removed, and the pellet was washed twice with cold 95% ethanol. The RNA pellets were then desalted using a Sephadex G-25 column and purified as described below.

Purification of oligonucleotides.

Single-stranded RNA oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis and visualized by UV shadowing. Bands were excised from the gel, crushed and soaked overnight at 4 °C in 0.5 M NaOAc, 0.1% sodium dodecyl sulfate (SDS), and 0.1 mM EDTA. Polyacrylamide fragments were removed with a 0.2 μm filter, and the RNAs were precipitated from a solution of 75% EtOH at -70 °C for 4 h. The solution was centrifuged 13,000 rpm for 20 min and supernatant

was removed. The RNA solutions were lyophilized to dryness, resuspended in nuclease-free water, and quantified by absorbance at 260 nm. Oligonucleotide mass was confirmed by MALDI-TOF.

In vitro transcription of editing target RNA.

Target RNA was transcribed from a DNA template with the MEGAScript T7 Kit (ThermoFisher). DNA Digestion was performed using RQ1 RNase-free DNase (Promega). DNase treated RNA product was purified as described above.

Preparation of Duplex Substrates for Analysis of ADAR Deamination Kinetics.

Purified guide and transcribed RNA were added in a 10:1 ratio to hybridization buffer (180 nM transcribed RNA target, 1.8 μ M guide, 1X TE Buffer, 100 mM NaCl), heated to 95 °C for 5 min, and slowly cooled to room temperature.

Protein Overexpression and Purification of ADAR2 full length constructs.

hADAR2 wild-type (hADAR2 WT) was expressed and as previously described.¹²⁹ Purification of hADAR2 was carried out by lysing cells in buffer containing 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 750 mM NaCl, 35 mM imidazole, and 0.01% Nonidet P-40 using a French press. Cell lysate was clarified by centrifugation (19,000 rpm for 1 hour). Lysate was passed over a 3 mL Ni-NTA column, which was then washed in three steps with 20 mL lysis buffer, wash I buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 750 mM NaCl, 35 mM imidazole, 0.01% Nonidet P-40), wash II buffer (20 mM Tris-HCl, pH 8.0, 5% glycerol, 1mM BME, 35 mM imidazole, 500 mM NaCl), and eluted with 20 mM Tris-HCl, pH 8.0, 5% glycerol, 1 mM BME, 400 mM imidazole, 100 mM NaCl. Fractions containing the target protein were pooled and concentrated to 30-80 μ M for use in biochemical assays. Protein concentrations were determined using BSA standards visualized by SYPRO orange staining of SDS-polyacrylamide gels. Purified hADAR2 WT was stored in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 20% glycerol and 1 mM BME at -70 °C.

In Vitro Deamination Kinetics.

Deamination assays were performed under single-turnover conditions. Reactions with ADAR2 wild-type (WT) contained 15 mM Tris-HCl pH 7.5, 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, 3 mM MgCl₂, 160 U/mL RNAsin, 1.0 µg/mL, 0.8 nM RNA, and 2 nM enzyme. ADAR1 reactions contained 15 mM Tris-HCl pH 7.5, 4% glycerol, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% Nonidet P-40, 160 U/mL RNAsin, 1.0 µg/mL yeast tRNA, 5 nM RNA and 50 nM ADAR1 p110 wild-type. Each reaction solution was incubated at 30 °C for 30 min before adding enzyme, and allowed to incubate at 30 °C for varying times prior to stopping with 190 µL 95°C water and heating at 95 °C for 5 min. RT-PCR (Promega Access RT-PCR System) was used to generate cDNA from deaminated RNA. The resulting cDNA was purified using the DNA Clean & Concentrator kit from Zymo, and subjected to Sanger Sequencing via GeneWiz (Azenta) with the forward PCR primers. The sequencing peak heights were quantified in SnapGene (Dotmatics). Each experiment was carried out in triplicate. The editing level for the corresponding zero time point was subtracted from each data point as a background subtraction. Deaminations of the chimera substrate deviate from the above protocol through an alternative RT-PCR protocol and sequencing as indicated below.

RT-PCR of DNA/RNA Chimeric Substrate

Reverse transcription of the DNA/RNA chimera was carried out using the M-MLV Reverse Transcriptase from Invitrogen using the forward RT primer according to manufacturer instructions. PCR was then performed using the Phusion Hot Start II DNA Polymerase (Thermo Fisher, F-549L) for 30 cycles. The resulting cDNA was purified using the DNA Clean & Concentrator kit from Zymo. Prior to submitting for NGS, it was subjected to Sanger Sequencing via GeneWiz (Azenta) with the forward and reverse PCR primers.

Processing of NGS data

Next Generation Sequencing was carried out with a 30% chi spike through GeneWiz (Azenta).

Data was obtained in a .fastq format for further processing. Preprocessing was carried out using standard methods. Using R Studio, the output file was selected for sequences of the correct length with sequences matching the *GLII* fixed sequences from the substrate design (Figure 6.5) via the script below.

R Script for Sorting winning Chimeric Sequences

```
template <-
DNAStrng("GGGAATGGCGCTTTGGAGAACAACCTCTAGGCAGAGGTCTCAAATTTCTTTCTTTCTTTGAGAC
CTNNNNNNNNNTTGTTCCTCCATTGTACCCTC")
#Replace sequence with DNA sequence post T7 promoter but with GGG
cjr <- readFastq("CJRE488F_preprocesed_SE.fastq")
#Replace file name
cjr <- cjr[width(cjr) %in% c(93:96)]
#Replace 93:96 with desired sequence length -3:+1

perfect <- xscat(subseq(sread(cjr),21,26), subseq(sread(cjr),30,35),
subseq(sread(cjr),60,65), subseq(sread(cjr),76,81)) ==
DNAStrng("CAACTCGCAGAGAGACCTTTGTTC")
#replace location indicators and string. Location indicators, 21,26 and so
on are the nucleotides at those positions on the initial sequence.
#4 sets of 6 nucleotides should be adjusted and the sequence taken and
replaced to create the new DNAStrng above.
#if analyzed the DNA string is 24 nucleotides coresponding to the 4 sets
of 6 nucleotides
table(perfect)

#perfect
#FALSE TRUE
#5578150 5701494

key_seq1 <- subseq(sread(cjr)[perfect],27,29)
#Replace location indicators for desired edited codon
#key_seq1 <- subseq(sread(cjr),27,29)
table(key_seq1)
ks1 <- as.data.frame(table(key_seq1))
ks1$sample="cjr"
colnames(ks1) <- c("codon", "count", "sample")

mer_seq1 <- subseq(sread(cjr)[perfect],66,75)
#replace location indicators for the N10 region
mk_table <- table((as.character(mer_seq1)),(as.character(key_seq1)))
write.csv(mk_table,"CJRE488F_codon_10mer.csv")
#replace what is in quotes for what your project name is. Keep the .csv at
the end.
```

Directed Editing on the endogenous Src Target in HEK293T Cells.

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11995-065) with 10% fetal bovine serum (FBS) (Thermo Fisher, 26140-087) and additionally supplemented with 1X antibiotic-antimycotic (Thermo Fisher, 15240-062) at 37 °C, 5 % CO₂. Once cells reached 70-90% confluency, cells were seeded into 96 well plates (6.4 x 10³ cells per well). After 24 h, cells were co-transfected with 500 ng ADAR plasmid and 50 nM guide oligonucleotide using Lipofectamine 2000 (Thermo Fisher, 11668-019). After incubation of transfection reagent, plasmid, and guide oligonucleotide in Opti-MEM Reduced Serum Media (Thermo Fisher, 31985-062), the solution was added to designated wells and incubated at 37 °C, 5 % CO₂. After 48 h, total RNA was isolated using RNAqueous Total RNA Isolation Kit (Thermo Fisher, AM1912) and DNase treated with RQ1 RNase-free DNase (Promega, M6101). Nested RT-PCR was performed in triplicate using Access RT-PCR kit (Promega, A1280) for 20 cycles and then followed by Phusion Hot Start II DNA Polymerase (Thermo Fisher, F-549L) for the second PCR of 30 cycles with target specific primers. PCR product was purified by agarose gel and QIAquick Gel Extraction kit (Qiagen, 28706). Product was submitted for Sanger Sequencing and sequence traces were analyzed by SnapGene (Dotmatics) to quantify percent editing.

Oligonucleotide sequences

a) Sequences for *in vitro* kinetics with the *SRC* substrate. Phosphorothioate modification is marked with an asterisk, 2'-O-methylated nucleotides are underlined, 2'-deoxynucleotides are in brackets; all others are ribonucleotides. All PCR primers are 2'-deoxynucleotides.

<i>SRC</i> bottom strand containing 2'-deoxy cytidine	5'-[G*C*C*C*C*A*GCCUUUGAG*A*C*C*U*CUGUC]dCdAdG[AGUU*G*U*U*C*U]-3'
<i>SRC</i> RT-PCR forward and sequencing primer	5'-CCAAGGATGCCTGGGAGA-3'
<i>SRC</i> RT-PCR reverse primer	5'-CTCATGTACTCCGTGACGATG-3'
<i>SRC</i> DNA template PCR forward primer	5'-TAATACGACTCACTATACGTCC-3'
<i>SRC</i> DNA template PCR reverse primer	5'-CAAAC TCCCCTTGCTCATG-3'

b) Target DNA of the *SRC* substrate for reverse transcription via T7 to form the *in vitro* substrate. Red is the target adenosine, grey is the T7 promoter, underline is the region complementary to the guide oligonucleotide.

TAATACGACTCACTATAACGTCCAAGCCGACACTCAGGGCCTGGCCAAGGATGCCTGGGAGATCCCTCGGGAGTCGC
 TCGCGCTGGAGGTCAAGCTGGGCCAGGGCTGCTTTGGCGAGGTGTGGATGGGGACCTGGAACGGTACCACCAGGGTG
 GCCATCAAAACCCTGAAGCCTGGCCACGATGTCTCCAGAGGCCTTCCTGCAGGAGGCCAGGTCATGAAGAAGCTGAG
 GCATGAGAAGCTGGTGCAGTTGTATGCTGTGGTTTCAGAGGAGCCCATTTACATCGTCACGGAGTACATGAGCAAGG
 GGAGTTTG

c) Sequences for *in vitro* deamination of the *GLII* chimeric substrate. 2'-deoxynucleotides are in brackets; all others are ribonucleotides. All PCR primers are 2'-deoxynucleotides.

<i>GLII</i> chimera	5'-[ATCACTGAGAATGCTGCCATGGATGCTAGAGGGCTACAG GAAGGTAA]CUUCCUGUAGCNNNNNNNNNCCAUGGCAGU AGGAUUCGCGUC-3'
<i>SRC</i> RT-PCR forward and sequencing primer	5'-AUCACUGAGAAUGCUGCCA-3'
<i>SRC</i> RT-PCR reverse primer	5'-GACGCGAAUCCUACUGCC-3'

d) Sequences for *in vitro* kinetics with the *APP* substrate. Phosphorothioate modification is marked with an asterisk, 2'-O-methylated nucleotides are underlined, 2'-deoxynucleotides are in brackets; all others are ribonucleotides. All PCR primers are 2'-deoxynucleotides.

<i>APP</i> RT-PCR forward and sequencing primer	5'- CATTGGACTCATGGTGG-3'
<i>APP</i> RT-PCR reverse primer	5'- CAGCATCACCAAGGTG-3'
<i>APP</i> guide PO	5'-CAAGGUGAUGACGAUCACUGUCG[CCA]UGACAAC ACCGC -'3
<i>APP</i> guide PS	5'-C*A*A*G*G*U*G*A*U*G*A*C*G*A*U*C*A*C* U*G*U*C*G*[C*C*A]*U*G*A*C*A*A*C*A*C*G*C -'3
<i>APP</i> guide PS 1	5'-C*A*A*G*G*U*G*A*UGACGAUCACUGUCG[CCA] UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 2	5'-C*A*A*G*G*UGAU*G*A*CGAUCACUGUCG[CCA] UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 3	5'-C*A*A*G*G*UGAUGAC*G*A*UCACUGUCG[CCA] UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 4	5'-C*A*A*G*G*UGAUGACGAU*C*A*CUGUCG[CCA] UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 5	5'-C*A*A*G*G*UGAUGACGAUCAC*U*G*UCG[CCA] UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 6	5'-C*A*A*G*G*UGAUGACGAUCACUGU*C*G[*CCA] UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 7	5'-C*A*A*G*G*UGAUGACGAUCACUGUCG[C*C*A] *UGACAAC*A*C*C*G*C -'3
<i>APP</i> guide PS 8	5'-C*A*A*G*G*UGAUGACGAUCACUGUCG[CCA] U*G*A*CAAC*A*C*C*G*C -'3

APP guide PS 9

5'-C*A*A*G*G*UGAUGACGAUCACUGUCG[CCA]
UGAC*A*A*C*A*C*C*G*C -'3

e) Target DNA of the *APP* substrate for reverse transcription via T7 to form the in vitro substrate. Red is the target adenosine, grey is the T7 promoter, underline is the region complementary to the guide oligonucleotide.

TAATACGACTCACTATAGGGGACGGAGGAGATCTCTGAAGTGAAGATGGATGCAGAATTCCGACATGACTCAGGATA
TGAAGTTCATCATCAAAAATTGGTGTCTTTGCAGAAGATGTGGGTTCAAACAAAGGTGCAATCATTGGACTCATGG
TGGGCGGTGTTGTCATAGCGACAGTGATCGTCATCACCTTGGTGATGCTGAAGAAGAAACAGTACACATCCATTCAT
CATGGTGTGGTGGAGGTTGACGCCGCTGTCACCCAGAGGAGCGCCACCTGTCCAAGATGCAGCAGAACGGCTACGA
AAATCCAACCTAC

References.

- (1) Doherty, E. E.; Beal, P. A. Oligonucleotide-Directed RNA Editing in Primates. *Mol. Ther.* **2022**. <https://doi.org/https://doi.org/10.1016/j.yymthe.2022.04.005>.
- (2) Khosravi, H. M.; Jantsch, M. F. Site-Directed RNA Editing: Recent Advances and Open Challenges. *RNA Biol.* **2021**, *18* (sup1), 41–50. <https://doi.org/10.1080/15476286.2021.1983288>.
- (3) Quin, J.; Sedmík, J.; Vukić, D.; Khan, A.; Keegan, L. P.; O’Connell, M. A. ADAR RNA Modifications, the Epitranscriptome and Innate Immunity. *Trends Biochem. Sci.* **2021**, *46* (9), 758–771. <https://doi.org/https://doi.org/10.1016/j.tibs.2021.02.002>.
- (4) Pan, D. *Fine-Tuning of RNA Functions by Modification and Editing*; Springer: Berlin ; New York, 2004; Vol. 7.
- (5) Zinshteyn, B.; Nishikura, K. Adenosine-to-Inosine RNA Editing. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2009**, *1* (2), 202–209. <https://doi.org/10.1002/wsbm.10>.
- (6) Christofi, T.; Zaravinos, A. RNA Editing in the Forefront of Epitranscriptomics and Human Health. *J. Transl. Med.* **2019**, *17* (1), 1–15. <https://doi.org/10.1186/s12967-019-2071-4>.
- (7) Deffit, S. N.; Hundley, H. A. To Edit or Not to Edit: Regulation of ADAR Editing Specificity and Efficiency. *Wiley Interdiscip. Rev. RNA* **7**, 113–127.
- (8) Roskoski, R. Properties of FDA-Approved Small Molecule Protein Kinase Inhibitors: A 2020 Update. *Pharmacol. Res.* **2020**, *152*, 104609. <https://doi.org/https://doi.org/10.1016/j.phrs.2019.104609>.
- (9) Fisher, A. J.; Beal, P. A. Effects of Aicardi-Goutières Syndrome Mutations Predicted from ADAR-RNA Structures. *RNA Biol.* **2017**, *14* (2), 164–170. <https://doi.org/10.1080/15476286.2016.1267097>.
- (10) Hogg, M.; Paro, S.; Keegan, L. P.; O’Connell, M. A. 3 - RNA Editing by Mammalian ADARs; Friedmann, T., Dunlap, J. C., Goodwin, S. F. B. T.-A. in G., Eds.; Academic Press, 2011; Vol. 73, pp 87–120. <https://doi.org/https://doi.org/10.1016/B978-0-12-380860-8.00003-3>.
- (11) Wang, Y.; Chung, D. hee; Monteleone, L. R.; Li, J.; Chiang, Y.; Toney, M. D.; Beal, P. A. RNA Binding Candidates for Human ADAR3 from Substrates of a Gain of Function Mutant Expressed in Neuronal Cells. *Nucleic Acids Res.* **2019**, *47* (20), 10801–10814. <https://doi.org/10.1093/nar/gkz815>.
- (12) Thuy-Boun, A. S.; Thomas, J. M.; Grajo, H. L.; Palumbo, C. M.; Park, S. H.; Nguyen, L. T.; Fisher, A. J.; Beal, P. A. Asymmetric Dimerization of Adenosine Deaminase Acting on RNA Facilitates Substrate Recognition. *Nucleic Acids Res.* **2020**, *48* (14), 7958–7972. <https://doi.org/10.1093/nar/gkaa532>.
- (13) Savva, Y. A.; Rieder, L. E.; Reenan, R. A. The ADAR Protein Family. *Genome Biol.* **2012**, *13* (12), 252. <https://doi.org/10.1186/gb-2012-13-12-252>.
- (14) Maas, S.; Gommans, W. M. Identification of a Selective Nuclear Import Signal in Adenosine Deaminases Acting on RNA. *Nucleic Acids Res.* **2009**, *37* (17), 5822–5829. <https://doi.org/10.1093/nar/gkp599>.
- (15) Riedmann, E. M.; Schopoff, S.; Hartner, J. C.; Jantsch, M. F. Specificity of ADAR-Mediated RNA Editing in Newly Identified Targets. *RNA* **2008**, *14* (6), 1110–1118. <https://doi.org/10.1261/rna.923308>.

- (16) Higuchi, M.; Maas, S.; Single, F. N.; Hartner, J.; Rozov, A.; Burnashev, N.; Feldmeyer, D.; Sprengel, R.; Seeburg, P. H. Point Mutation in an AMPA Receptor Gene Rescues Lethality in Mice Deficient in the RNA-Editing Enzyme ADAR2. *Nature* **2000**, *406* (6791), 78–81. <https://doi.org/10.1038/35017558>.
- (17) Doble, A. Excitatory Amino Acid Receptors and Neurodegeneration. *Therapie* **1995**, *50* (4), 319–337.
- (18) Wang, G.; Wang, H.; Singh, S.; Zhou, P.; Yang, S.; Wang, Y.; Zhu, Z.; Zhang, J.; Chen, A.; Billiar, T.; Monga, S. P.; Wang, Q. ADAR1 Prevents Liver Injury from Inflammation and Suppresses Interferon Production in Hepatocytes. *Am. J. Pathol.* **2015**, *185* (12), 3224–3237. <https://doi.org/10.1016/j.ajpath.2015.08.002>.
- (19) Yang, S.; Deng, P.; Zhu, Z.; Zhu, J.; Wang, G.; Zhang, L.; Chen, A. F.; Wang, T.; Sarkar, S. N.; Billiar, T. R.; Wang, Q. Adenosine Deaminase Acting on RNA 1 Limits RIG-I RNA Detection and Suppresses IFN Production Responding to Viral and Endogenous RNAs. *J. Immunol.* **2014**, *193* (7), 3436–3445. <https://doi.org/10.4049/jimmunol.1401136>.
- (20) Ishizuka, J. J. Loss of ADAR1 in Tumours Overcomes Resistance to Immune Checkpoint Blockade. *Nature* **565**, 43–48.
- (21) Gannon, H. S.; Zou, T.; Kiessling, M. K.; Gao, G. F.; Cai, D.; Choi, P. S.; Ivan, A. P.; Buchumenski, I.; Berger, A. C.; Goldstein, J. T.; Cherniack, A. D.; Vazquez, F.; Tsherniak, A.; Levanon, E. Y.; Hahn, W. C.; Meyerson, M. Identification of ADAR1 Adenosine Deaminase Dependency in a Subset of Cancer Cells. *Nat. Commun.* **2018**, *9* (1), 5450. <https://doi.org/10.1038/s41467-018-07824-4>.
- (22) Liu, H. Tumor-Derived IFN Triggers Chronic Pathway Agonism and Sensitivity to ADAR Loss. *Nat. Med* **25**, 95–102.
- (23) Stephens, O. M.; Yi-Brunozzi, H. Y.; Beal, P. A. Analysis of the RNA-Editing Reaction of ADAR2 with Structural and Fluorescent Analogues of the GluR-B R/G Editing Site. *Biochemistry* **2000**, *39* (40), 12243–12251. <https://doi.org/10.1021/bi0011577>.
- (24) Vrma, S.; Vaish, N. k.; Eckstein, F. 15 - Applications of Ribonucleotide Analogues in RNA Biochemistry; Söll, D., Nishimura, S., Moore, P. B. T.-R. N. A., Eds.; Pergamon: Oxford, 2001; pp 259–275. <https://doi.org/https://doi.org/10.1016/B978-008043408-7/50036-8>.
- (25) Matthews, M. M.; Thomas, J. M.; Zheng, Y.; Tran, K.; Phelps, K. J.; Scott, A. I.; Havel, J.; Fisher, A. J.; Beal, P. A. Structures of Human ADAR2 Bound to DsRNA Reveal Base-Flipping Mechanism and Basis for Site Selectivity. *Nat. Struct. Mol. Biol.* **2016**, *23* (5), 426–433. <https://doi.org/10.1038/nsmb.3203>.
- (26) Haudenschild, B. L.; Maydanovych, O.; Véliz, E. A.; Macbeth, M. R.; Bass, B. L.; Beal, P. A. A Transition State Analogue for an RNA-Editing Reaction. *J. Am. Chem. Soc.* **2004**, *126* (36), 11213–11219. <https://doi.org/10.1021/ja0472073>.
- (27) Erion, M. D.; Reddy, M. R. Calculation of Relative Hydration Free Energy Differences for Heteroaromatic Compounds: Use in the Design of Adenosine Deaminase and Cytidine Deaminase Inhibitors. *J. Am. Chem. Soc.* **1998**, *120* (14), 3295–3304. <https://doi.org/10.1021/ja972906j>.
- (28) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. HhaI Methyltransferase Flips Its Target Base out of the DNA Helix. *Cell* **1994**, *76* (2), 357–369. [https://doi.org/https://doi.org/10.1016/0092-8674\(94\)90342-5](https://doi.org/https://doi.org/10.1016/0092-8674(94)90342-5).

- (29) Daujotytė, D.; Serva, S.; Vilkaitis, G.; Merkienė, E.; Venclovas, Č.; Klimašauskas, S. HhaI DNA Methyltransferase Uses the Protruding Gln237 for Active Flipping of Its Target Cytosine. *Structure* **2004**, *12* (6), 1047–1055. <https://doi.org/https://doi.org/10.1016/j.str.2004.04.007>.
- (30) Slupphaug, G.; Mol, C. D.; Kavli, B.; Arvai, A. S.; Krokan, H. E.; Tainer, J. A. A Nucleotide-Flipping Mechanism from the Structure of Human Uracil–DNA Glycosylase Bound to DNA. *Nature* **1996**, *384* (6604), 87–92. <https://doi.org/10.1038/384087a0>.
- (31) Bruner, S. D.; Norman, D. P. G.; Verdine, G. L. Structural Basis for Recognition and Repair of the Endogenous Mutagen 8-Oxoguanine in DNA. *Nature* **2000**, *403* (6772), 859–866. <https://doi.org/10.1038/35002510>.
- (32) Lau, A. Y.; Schärer, O. D.; Samson, L.; Verdine, G. L.; Ellenberger, T. Crystal Structure of a Human Alkylbase–DNA Repair Enzyme Complexed to DNA: Mechanisms for Nucleotide Flipping and Base Excision. *Cell* **1998**, *95* (2), 249–258. [https://doi.org/https://doi.org/10.1016/S0092-8674\(00\)81755-9](https://doi.org/https://doi.org/10.1016/S0092-8674(00)81755-9).
- (33) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 319. <https://doi.org/10.1038/ncomms1324>.
- (34) Wong, S. K.; Sato, S.; Lazinski, D. W. Substrate Recognition by ADAR1 and ADAR2. *RNA* **7**, 846–858.
- (35) Wang, Y.; Park, S.; Beal, P. A. Selective Recognition of RNA Substrates by ADAR Deaminase Domains. *Biochemistry* **57**, 1640–1651.
- (36) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 319. <https://doi.org/10.1038/ncomms1324>.
- (37) Li, J. B.; Levanon, E. Y.; Yoon, J.-K.; Aach, J.; Xie, B.; LeProust, E.; Zhang, K.; Gao, Y.; Church, G. M. Genome-Wide Identification of Human RNA Editing Sites by Parallel DNA Capturing and Sequencing. *Science* (80-.). **2009**, *324* (5931), 1210 LP – 1213. <https://doi.org/10.1126/science.1170995>.
- (38) Damase, T. R.; Sukhovshin, R.; Boada, C.; Taraballi, F.; Pettigrew, R. I.; Cooke, J. P. The Limitless Future of RNA Therapeutics. *Front. Bioeng. Biotechnol.* **2021**, *9*, 628137.
- (39) Woolf, T. M.; Chase, J. M.; Stinchcomb, D. T. Toward the Therapeutic Editing of Mutated RNA Sequences. *Proc. Natl. Acad. Sci.* **1995**, *92* (18), 8298 LP – 8302. <https://doi.org/10.1073/pnas.92.18.8298>.
- (40) Qu, L.; Yi, Z.; Zhu, S.; Wang, C.; Cao, Z.; Zhou, Z.; Yuan, P.; Yu, Y.; Tian, F.; Liu, Z.; Bao, Y.; Zhao, Y.; Wei, W. Programmable RNA Editing by Recruiting Endogenous ADAR Using Engineered RNAs. *Nat. Biotechnol.* **2019**, *37* (9), 1059–1069. <https://doi.org/10.1038/s41587-019-0178-z>.
- (41) Merkle, T.; Merz, S.; Reautschnig, P.; Blaha, A.; Li, Q.; Vogel, P.; Wettengel, J.; Li, J. B.; Stafforst, T. Precise RNA Editing by Recruiting Endogenous ADARs with Antisense Oligonucleotides. *Nat. Biotechnol.* **2019**, *37* (2), 133–138. <https://doi.org/10.1038/s41587-019-0013-6>.
- (42) Katrekar, D.; Chen, G.; Meluzzi, D.; Ganesh, A.; Worlikar, A.; Shih, Y.-R.; Varghese, S.; Mali, P. In Vivo RNA Editing of Point Mutations via RNA-Guided Adenosine Deaminases. *Nat. Methods* **2019**, *16* (3), 239–242. <https://doi.org/10.1038/s41592-019-0323-0>.
- (43) Chen, G.; Katrekar, D.; Mali, P. RNA-Guided Adenosine Deaminases: Advances and Challenges

- for Therapeutic RNA Editing. *Biochemistry* **2019**, 58 (15), 1947–1957. <https://doi.org/10.1021/acs.biochem.9b00046>.
- (44) Rees, H. A.; Liu, D. R. Base Editing: Precision Chemistry on the Genome and Transcriptome of Living Cells. *Nat. Rev. Genet.* **2018**, 19 (12), 770–788. <https://doi.org/10.1038/s41576-018-0059-1>.
- (45) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.)*. **2012**, 337 (6096), 816 LP – 821. <https://doi.org/10.1126/science.1225829>.
- (46) Christian, M.; Cermak, T.; Doyle, E. L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A. J.; Voytas, D. F. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **2010**, 186 (2), 757–761. <https://doi.org/10.1534/genetics.110.120717>.
- (47) Urnov, F. D.; Miller, J. C.; Lee, Y.-L.; Beausejour, C. M.; Rock, J. M.; Augustus, S.; Jamieson, A. C.; Porteus, M. H.; Gregory, P. D.; Holmes, M. C. Highly Efficient Endogenous Human Gene Correction Using Designed Zinc-Finger Nucleases. *Nature* **2005**, 435 (7042), 646–651. <https://doi.org/10.1038/nature03556>.
- (48) Prashant, M.; Luhan, Y.; M., E. K.; John, A.; Marc, G.; E., D. J.; E., N. J.; M., C. G. RNA-Guided Human Genome Engineering via Cas9. *Science (80-.)*. **2013**, 339 (6121), 823–826. <https://doi.org/10.1126/science.1232033>.
- (49) Newby, G. A.; Yen, J. S.; Woodard, K. J.; Mayuranathan, T.; Lazzarotto, C. R.; Li, Y.; Sheppard-Tillman, H.; Porter, S. N.; Yao, Y.; Mayberry, K.; Everette, K. A.; Jang, Y.; Podracky, C. J.; Thaman, E.; Lechauve, C.; Sharma, A.; Henderson, J. M.; Richter, M. F.; Zhao, K. T.; Miller, S. M.; Wang, T.; Koblan, L. W.; McCaffrey, A. P.; Tisdale, J. F.; Kalfa, T. A.; Pruett-Miller, S. M.; Tsai, S. Q.; Weiss, M. J.; Liu, D. R. Base Editing of Haematopoietic Stem Cells Rescues Sickle Cell Disease in Mice. *Nature* **2021**, 595 (7866), 295–302. <https://doi.org/10.1038/s41586-021-03609-w>.
- (50) Christofi, T.; Zaravinos, A. RNA Editing in the Forefront of Epitranscriptomics and Human Health. *J. Transl. Med.* **2019**, 17 (1), 319. <https://doi.org/10.1186/s12967-019-2071-4>.
- (51) Chen, G.; Katrekar, D.; Mali, P. RNA-Guided Adenosine Deaminases: Advances and Challenges for Therapeutic RNA Editing. *Biochemistry* **2019**, 58 (15), 1947–1957. <https://doi.org/10.1021/acs.biochem.9b00046>.
- (52) Gaudelli, N. M.; Komor, A. C.; Rees, H. A.; Packer, M. S.; Badran, A. H.; Bryson, D. I.; Liu, D. R. Programmable Base Editing of T to G C in Genomic DNA without DNA Cleavage. *Nature* **2017**, 551 (7681), 464–471. <https://doi.org/10.1038/nature24644>.
- (53) 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* **2016**, 533 (7603), 420–424. <https://doi.org/10.1038/nature17946>.
- (54) Gaudelli, N. M.; Komor, A. C.; Rees, H. A.; Packer, M. S.; Badran, A. H.; Bryson, D. I.; Liu, D. R. Programmable Base Editing of A•T to G•C in Genomic DNA without DNA Cleavage. *Nature* **2017**, 551 (7681), 464–471. <https://doi.org/10.1038/nature24644>.
- (55) Keiji, N.; Takayuki, A.; Nozomu, Y.; Satomi, B.; Mika, K.; Mayura, T.; Masao, M.; Aya, M.; Michihiro, A.; Y., H. K.; Zenpei, S.; Akihiko, K. Targeted Nucleotide Editing Using Hybrid Prokaryotic and Vertebrate Adaptive Immune Systems. *Science (80-.)*. **2016**, 353 (6305), aaf8729. <https://doi.org/10.1126/science.aaf8729>.

- (56) Nair, J. K.; Willoughby, J. L. S.; Chan, A.; Charisse, K.; Alam, M. R.; Wang, Q.; Hoekstra, M.; Kandasamy, P.; Kelin, A. V.; Milstein, S.; Taneja, N.; O Shea, J.; Shaikh, S.; Zhang, L.; Van Der Sluis, R. J.; Jung, M. E.; Akinc, A.; Hutabarat, R.; Kuchimanchi, S.; Fitzgerald, K.; Zimmermann, T.; Van Berkel, T. J. C.; Maier, M. A.; Rajeev, K. G.; Manoharan, M. Multivalent N - Acetylgalactosamine-Conjugated SiRNA Localizes in Hepatocytes and Elicits Robust RNAi-Mediated Gene Silencing. *J. Am. Chem. Soc.* **2014**, *136* (49), 16958–16961. <https://doi.org/10.1021/ja505986a>.
- (57) Debacker, A. J.; Voutila, J.; Catley, M.; Blakey, D.; Habib, N. Delivery of Oligonucleotides to the Liver with GalNAc: From Research to Registered Therapeutic Drug. *Mol. Ther.* **2020**, *28* (8), 1759–1771. <https://doi.org/https://doi.org/10.1016/j.ymthe.2020.06.015>.
- (58) Wilson, R. C.; Gilbert, L. A. The Promise and Challenge of In Vivo Delivery for Genome Therapeutics. *ACS Chem. Biol.* **2018**, *13* (2), 376–382. <https://doi.org/10.1021/acscchembio.7b00680>.
- (59) Woolf, T. M.; Chase, J. M.; Stinchcomb, D. T. Toward the Therapeutic Editing of Mutated RNA Sequences. *Proc. Natl. Acad. Sci.* **1995**, *92* (18), 8298 LP – 8302. <https://doi.org/10.1073/pnas.92.18.8298>.
- (60) Stafforst, T.; Schneider, M. F. An RNA-Deaminase Conjugate Selectively Repairs Point Mutations. *Angew. Chemie - Int. Ed.* **2012**, *51* (44), 11166–11169. <https://doi.org/10.1002/anie.201206489>.
- (61) Vogel, P.; Moschref, M.; Li, Q.; Merkle, T.; Selvasaravanan, K. D.; Li, J. B.; Stafforst, T. Efficient and Precise Editing of Endogenous Transcripts with SNAP-Tagged ADARs. *Nat. Methods* **2018**, *15* (7), 535–538. <https://doi.org/10.1038/s41592-018-0017-z>.
- (62) Montiel-Gonzalez, M. F.; Vallecillo-Viejo, I.; Yudowski, G. A.; Rosenthal, J. J. C. C. Correction of Mutations within the Cystic Fibrosis Transmembrane Conductance Regulator by Site-Directed RNA Editing. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (45), 18285–18290. <https://doi.org/10.1073/pnas.1306243110>.
- (63) Azad, M. T. A.; Bhakta, S.; Tsukahara, T. Site-Directed RNA Editing by Adenosine Deaminase Acting on RNA for Correction of the Genetic Code in Gene Therapy. *Gene Ther.* **2017**, *24* (12), 779–786. <https://doi.org/10.1038/gt.2017.90>.
- (64) Cox, D. B. T. T.; Gootenberg, J. S.; Abudayyeh, O. O.; Franklin, B.; Kellner, M. J.; Joung, J.; Zhang, F. RNA Editing with CRISPR-Cas13. *Science (80-.)*. **2017**, *358* (6366), 1019 LP – 1027. <https://doi.org/10.1126/science.aaq0180>.
- (65) Wettengel, J.; Reautschnig, P.; Geisler, S.; Kahle, P. J.; Stafforst, T. Harnessing Human ADAR2 for RNA Repair – Recoding a PINK1 Mutation Rescues Mitophagy. *Nucleic Acids Res.* **2017**, *45* (5), 2797–2808. <https://doi.org/10.1093/nar/gkw911>.
- (66) Fukuda, M.; Umeno, H.; Nose, K.; Nishitarumizu, A.; Noguchi, R.; Nakagawa, H. Construction of a Guide-RNA for Site-Directed RNA Mutagenesis Utilising Intracellular A-to-I RNA Editing. *Sci. Rep.* **2017**, *7*, 41478. <https://doi.org/10.1038/srep41478>.
- (67) Qu, L.; Yi, Z.; Zhu, S.; Wang, C.; Cao, Z.; Zhou, Z.; Yuan, P.; Yu, Y.; Tian, F.; Liu, Z.; Bao, Y.; Zhao, Y.; Wei, W. Programmable RNA Editing by Recruiting Endogenous ADAR Using Engineered RNAs. *Nat. Biotechnol.* **2019**, *37* (9), 1059–1069. <https://doi.org/10.1038/s41587-019-0178-z>.
- (68) Monian, P.; Shivalila, C.; Lu, G.; Shimizu, M.; Boulay, D.; Bussow, K.; Byrne, M.; Bezigan, A.;

- Chatterjee, A.; Chew, D.; Desai, J.; Favalaro, F.; Godfrey, J.; Hoss, A.; Iwamoto, N.; Kawamoto, T.; Kumarasamy, J.; Lamattina, A.; Lindsey, A.; Liu, F.; Looby, R.; Marappan, S.; Metterville, J.; Murphy, R.; Rossi, J.; Pu, T.; Bhattarai, B.; Standley, S.; Tripathi, S.; Yang, H.; Yin, Y.; Yu, H.; Zhou, C.; Apponi, L. H.; Kandasamy, P.; Vargeese, C. Endogenous ADAR-Mediated RNA Editing in Non-Human Primates Using Stereopure Chemically Modified Oligonucleotides. *Nat. Biotechnol.* **2022**. <https://doi.org/10.1038/s41587-022-01225-1>.
- (69) Cho, I. S.; Kim, J.; Lim, D. H.; Ahn, H.-C.; Kim, H.; Lee, K.-B.; Lee, Y. S. Improved Serum Stability and Biophysical Properties of SiRNAs Following Chemical Modifications. *Biotechnol. Lett.* **2008**, *30* (11), 1901. <https://doi.org/10.1007/s10529-008-9776-4>.
- (70) Kratschmer, C.; Levy, M. Effect of Chemical Modifications on Aptamer Stability in Serum. *Nucleic Acid Ther.* **2017**, *27* (6), 335–344. <https://doi.org/10.1089/nat.2017.0680>.
- (71) Iwamoto, N.; Butler, D. C. D.; Svrzikapa, N.; Mohapatra, S.; Zlatev, I.; Sah, D. W. Y.; Meena; Standley, S. M.; Lu, G.; Apponi, L. H.; Frank-Kamenetsky, M.; Zhang, J. J.; Vargeese, C.; Verdine, G. L. Control of Phosphorothioate Stereochemistry Substantially Increases the Efficacy of Antisense Oligonucleotides. *Nat. Biotechnol.* **2017**, *35* (9), 845–851. <https://doi.org/10.1038/nbt.3948>.
- (72) Wong, S. K.; Sato, S.; Lazinski, D. W. Substrate Recognition by ADAR1 and ADAR2. *Rna* **2001**, *7* (6), 846–858. <https://doi.org/10.1017/S135583820101007X>.
- (73) Vallecillo-Viejo, I. C.; Liscovitch-Brauer, N.; Montiel-Gonzalez, M. F.; Eisenberg, E.; Rosenthal, J. J. C. C. Abundant Off-Target Edits from Site-Directed RNA Editing Can Be Reduced by Nuclear Localization of the Editing Enzyme. *RNA Biol.* **2018**, *15* (1), 104–114. <https://doi.org/10.1080/15476286.2017.1387711>.
- (74) Malik, T. N.; Doherty, E. E.; Gaded, V. M.; Hill, T. M.; Beal, P. A.; Emeson, R. B. Regulation of RNA Editing by Intracellular Acidification. *Nucleic Acids Res.* **2021**, *49* (7), 4020–4036. <https://doi.org/10.1093/nar/gkab157>.
- (75) Walkley, C. R.; Li, J. B. Rewriting the Transcriptome: Adenosine-to-Inosine RNA Editing by ADARs. *Genome Biol.* **2017**, *18* (1), 205. <https://doi.org/10.1186/s13059-017-1347-3>.
- (76) Roundtree, I. A.; Evans, M. E.; Pan, T.; He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **2017**, *169* (7), 1187–1200. <https://doi.org/10.1016/j.cell.2017.05.045>.
- (77) Bazak, L.; Haviv, A.; Barak, M.; Jacob-Hirsch, J.; Deng, P.; Zhang, R.; Isaacs, F. J.; Rechavi, G.; Li, J. B.; Eisenberg, E.; Levanon, E. Y. A-to-I RNA Editing Occurs at over a Hundred Million Genomic Sites, Located in a Majority of Human Genes. *Genome Res.* **2014**, *24* (3), 365–376. <https://doi.org/10.1101/gr.164749.113>.
- (78) Tan, M. H.; Li, Q.; Shanmugam, R.; Piskol, R.; Kohler, J.; Young, A. N.; Liu, K. I.; Zhang, R.; Ramaswami, G.; Ariyoshi, K. Dynamic Landscape and Regulation of RNA Editing in Mammals. *Nature* **550**, 249–254.
- (79) Bass, B. L. RNA Editing by Adenosine Deaminases That Act on RNA. *Annu. Rev. Biochem.* **2002**, *71*, 817–846. <https://doi.org/10.1146/annurev.biochem.71.110601.135501>.
- (80) Nishikura, K. Functions and Regulation of RNA Editing by ADAR Deaminases. *Annu. Rev. Biochem.* **2010**, *79*, 321–349. <https://doi.org/10.1146/annurev-biochem-060208-105251>.
- (81) Nishikura, K.; Yoo, C.; Kim, U.; Murray, J. M.; Estes, P. A.; Cash, F. E.; Liehaber, S. A. Substrate Specificity of the DsRNA Unwinding/Modifying Activity. *EMBO J.* **1991**, *10* (11),

- 3523–3532. <https://doi.org/10.1002/j.1460-2075.1991.tb04916.x>.
- (82) Phelps, K. J.; Tran, K.; Eifler, T.; Erickson, A. I.; Fisher, A. J.; Beal, P. A. Recognition of Duplex RNA by the Deaminase Domain of the RNA Editing Enzyme ADAR2. *Nucleic Acids Res.* **2015**, *43* (2), 1123–1132. <https://doi.org/10.1093/nar/gku1345>.
- (83) Goodman, R. A.; MacBeth, M. R.; Beal, P. A. ADAR Proteins: Structure and Catalytic Mechanism. *Curr. Top. Microbiol. Immunol.* **2012**, *353* (1), 1–33. https://doi.org/10.1007/82_2011_144.
- (84) Kuttan, A.; Bass, B. L. Mechanistic Insights into Editing-Site Specificity of ADARs. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (48), E3295 LP-E3304. <https://doi.org/10.1073/pnas.1212548109>.
- (85) Paul, M. S.; Bass, B. L. Inosine Exists in mRNA at Tissue-Specific Levels and Is Most Abundant in Brain mRNA. *EMBO J.* **1998**, *17* (4), 1120–1127. <https://doi.org/10.1093/emboj/17.4.1120>.
- (86) Jacobs, M. M.; Fogg, R. L.; Emeson, R. B.; Stanwood, G. D. ADAR1 and ADAR2 Expression and Editing Activity during Forebrain Development. *Dev. Neurosci.* **2009**, *31* (3), 223–237. <https://doi.org/10.1159/000210185>.
- (87) Brusa, R.; Zimmermann, F.; Koh, D. S.; Feldmeyer, D.; Gass, P.; Seeburg, P. H.; Sprengel, R.; Sakmann, B. Early-Onset Epilepsy and Postnatal Lethality Associated with an Editing-Deficient GluR-B Allele in Mice. *Science (80-.)*. **1995**, *270* (5242), 1677–1680. <https://doi.org/10.1126/science.270.5242.1677>.
- (88) Bhalla, T.; Rosenthal, J. J. C.; Holmgren, M.; Reenan, R. Control of Human Potassium Channel Inactivation by Editing of a Small mRNA Hairpin. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 950–956. <https://doi.org/10.1038/nsmb825>.
- (89) Burns, C. M.; Chu, H.; Rueter, S. M.; Hutehinson, L. K.; Canton, H.; Sanders-bush, E.; Emeson, R. B. Regulation of Serotonin-2C Receptor G-Protein Coupling by RNA Editing Sequence Analysis of Clones Isolated from a Rat Striatum cDNA. **1997**, *387115* (May), 303–308.
- (90) Palladino, M. J.; Keegan, L. P.; O’Connell, M. A.; Reenan, R. A. A-to-I Pre-mRNA Editing in Drosophila Is Primarily Involved in Adult Nervous System Function and Integrity. *Cell* **2000**, *102* (4), 437–449. [https://doi.org/10.1016/S0092-8674\(00\)00049-0](https://doi.org/10.1016/S0092-8674(00)00049-0).
- (91) Tonkin, L. A.; Saccomanno, L.; Morse, D. P.; Brodigan, T.; Krause, M.; Bass, B. L. RNA Editing by ADARs Is Important for Normal Behavior in Caenorhabditis Elegans. *EMBO J.* **2002**, *21* (22), 6025–6035. <https://doi.org/10.1093/emboj/cdf607>.
- (92) Hartner, J. C.; Schmittwolf, C.; Kispert, A.; Müller, A. M.; Higuchi, M.; Seeburg, P. H. Liver Disintegration in the Mouse Embryo Caused by Deficiency in the RNA-Editing Enzyme ADAR1. *J. Biol. Chem.* **2004**, *279* (6), 4894–4902. <https://doi.org/10.1074/jbc.M311347200>.
- (93) Wang, Q.; Miyakoda, M.; Yang, W.; Khillan, J.; Stachura, D. L.; Weiss, M. J.; Nishikura, K. Stress-Induced Apoptosis Associated with Null Mutation of ADAR1 RNA Editing Deaminase Gene. *J. Biol. Chem.* **2004**, *279* (6), 4952–4961. <https://doi.org/10.1074/jbc.M310162200>.
- (94) Vollmar, W.; Gloger, J.; Berger, E.; Kortenbruck, G.; Köhling, R.; Speckmann, E. J.; Musshoff, U. RNA Editing (R/G Site) and Flip-Flop Splicing of the AMPA Receptor Subunit GluR2 in Nervous Tissue of Epilepsy Patients. *Neurobiol. Dis.* **2004**, *15* (2), 371–379. <https://doi.org/10.1016/j.nbd.2003.11.006>.
- (95) Niswender, C. M.; Herrick-Davis, K.; Dilley, G. E.; Meltzer, H. Y.; Overholser, J. C.; Stockmeier, C. A.; Emeson, R. B.; Sanders-Bush, E. RNA Editing of the Human Serotonin 5-HT_{2C} Receptor:

- Alterations in Suicide and Implications for Serotonergic Pharmacotherapy. *Neuropsychopharmacology* **2001**, *24* (5), 478–491. [https://doi.org/10.1016/S0893-133X\(00\)00223-2](https://doi.org/10.1016/S0893-133X(00)00223-2).
- (96) Hideyama, T.; Yamashita, T.; Aizawa, H.; Tsuji, S.; Kakita, A.; Takahashi, H.; Kwak, S. Profound Downregulation of the RNA Editing Enzyme ADAR2 in ALS Spinal Motor Neurons. *Neurobiol. Dis.* **2012**, *45* (3), 1121–1128. <https://doi.org/10.1016/j.nbd.2011.12.033>.
- (97) Breen, M.; Dobbyn, A.; Li, Q.; Roussos, P.; Hoffman, G.; Stahl, E.; Chess, A.; Li, J. B.; Devlin, B.; Buxbaum, J. 95 Global Landscape and Genetic Regulation of Rna Editing in Cortical Samples From Individuals With Schizophrenia. *Eur. Neuropsychopharmacol.* **2019**, *29*, S112–S113. <https://doi.org/10.1016/j.euroneuro.2019.07.236>.
- (98) Rice, G. I.; Kasher, P. R.; Forte, G. M. A.; Mannion, N. M.; Greenwood, S. M.; Szykiewicz, M.; Dickerson, J. E.; Bhaskar, S. S.; Zampini, M.; Briggs, T. A.; Jenkinson, E. M.; Bacino, C. A.; Battini, R.; Bertini, E.; Brogan, P. A.; Brueton, L. A.; Carpanelli, M.; De Laet, C.; De Lonlay, P.; Del Toro, M.; Desguerre, I.; Fazzi, E.; Garcia-Cazorla, À.; Heiberger, A.; Kawaguchi, M.; Kumar, R.; Lin, J. P. S. M.; Lourenco, C. M.; Male, A. M.; Marques, W.; Mignot, C.; Olivieri, I.; Orcesi, S.; Prabhakar, P.; Rasmussen, M.; Robinson, R. A.; Rozenberg, F.; Schmidt, J. L.; Steindl, K.; Tan, T. Y.; Van Der Merwe, W. G.; Vanderver, A.; Vassallo, G.; Wakeling, E. L.; Wassmer, E.; Whittaker, E.; Livingston, J. H.; Lebon, P.; Suzuki, T.; McLaughlin, P. J.; Keegan, L. P.; O'Connell, M. A.; Lovell, S. C.; Crow, Y. J. Mutations in ADAR1 Cause Aicardi-Goutières Syndrome Associated with a Type I Interferon Signature. *Nat. Genet.* **2012**, *44* (11), 1243–1248. <https://doi.org/10.1038/ng.2414>.
- (99) Rosenthal, J. J. C. The Emerging Role of RNA Editing in Plasticity. *J. Exp. Biol.* **2015**, *218* (12), 1812–1821. <https://doi.org/10.1242/jeb.119065>.
- (100) Garrett, S.; Rosenthal, J. J. C. RNA Editing Underlies Temperature Adaptation in K⁺ Channels from Polar Octopuses. *Science* (80-.). **2012**, *335* (6070), 848–851. <https://doi.org/10.1126/science.1212795>.
- (101) Wahlstedt, H.; Daniel, C.; Ensterö, M.; Öhman, M. Large-Scale mRNA Sequencing Determines Global Regulation of RNA Editing during Brain Development. *Genome Res.* **2009**, *19* (6), 978–986. <https://doi.org/10.1101/gr.089409.108>.
- (102) Hood, J. L.; Morabito, M. V.; Martinez, C. R.; Gilbert, J. A.; Ferrick, E. A.; Ayers, G. D.; Chappell, J. D.; Dermody, T. S.; Emeson, R. B. Reovirus-Mediated Induction of ADAR1 (P150) Minimally Alters RNA Editing Patterns in Discrete Brain Regions. *Mol. Cell. Neurosci.* **2014**, *61*, 97–109. <https://doi.org/10.1016/j.mcn.2014.06.001>.
- (103) Porath, H. T.; Hazan, E.; Shpigler, H.; Cohen, M.; Band, M.; Ben-Shahar, Y.; Levanon, E. Y.; Eisenberg, E.; Bloch, G. RNA Editing Is Abundant and Correlates with Task Performance in a Social Bumblebee. *Nat. Commun.* **2019**, *10* (1), 1605. <https://doi.org/10.1038/s41467-019-09543-w>.
- (104) Li, J. B.; Church, G. M. Deciphering the Functions and Regulation of Brain-Enriched A-to-I RNA Editing. *Nat. Neurosci.* **2013**, *16* (11), 1518–1522. <https://doi.org/10.1038/nn.3539>.
- (105) Schaffer, A. A.; Kopel, E.; Hendel, A.; Picardi, E.; Levanon, E. Y.; Eisenberg, E. The Cell Line A-to-I RNA Editing Catalogue. *Nucleic Acids Res.* **2020**, *48* (11), 5849–5858. <https://doi.org/10.1093/nar/gkaa305>.
- (106) Sapiro, A. L.; Freund, E. C.; Restrepo, L.; Qiao, H. H.; Bhate, A.; Li, Q.; Ni, J. Q.; Mosca, T. J.; Li, J. B. Zinc Finger RNA-Binding Protein Zn72D Regulates ADAR-Mediated RNA Editing in

- Neurons. *Cell Rep.* **2020**, *31* (7), 107654. <https://doi.org/10.1016/j.celrep.2020.107654>.
- (107) Matthews, M. M.; Thomas, J. M.; Zheng, Y.; Tran, K.; Phelps, K. J.; Scott, A. I.; Havel, J.; Fisher, A. J.; Beal, P. A. Structures of Human ADAR2 Bound to DsRNA Reveal Base-Flipping Mechanism and Basis for Site Selectivity. *Nat. Struct. Mol. Biol.* **2016**, *23* (5), 426–433. <https://doi.org/10.1038/nsmb.3203>.
- (108) Wang, Y.; Havel, J.; Beal, P. A. A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1. *ACS Chem. Biol.* **2015**, *10* (11), 2512–2519. <https://doi.org/10.1021/acscchembio.5b00711>.
- (109) Huynh, K.; Partch, C. L. Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay. *Curr. Protoc. protein Sci.* **2015**, *79*, 28.9.1-28.9.14. <https://doi.org/10.1002/0471140864.ps2809s79>.
- (110) Eifler, T.; Pokharel, S.; Beal, P. A. RNA-Seq Analysis Identifies a Novel Set of Editing Substrates for Human ADAR2 Present in *Saccharomyces Cerevisiae*. *Biochemistry* **2013**, *52* (45), 7857–7869. <https://doi.org/10.1021/bi4006539>.
- (111) Stephens, O. M.; Yi-Brunozzi, H. Y.; Beal, P. A. Analysis of the RNA-Editing Reaction of ADAR2 with Structural and Fluorescent Analogues of the GluR-B R/G Editing Site. *Biochemistry* **2000**, *39* (40), 12243–12251. <https://doi.org/10.1021/bi0011577>.
- (112) Garcia-Moreno, B. Adaptations of Proteins to Cellular and Subcellular PH. *J. Biol.* **2009**, *8* (11), 98. <https://doi.org/10.1186/jbiol199>.
- (113) Talley, K.; Alexov, E. On the PH-Optimum of Activity and Stability of Proteins. *Proteins Struct. Funct. Bioinforma.* **2010**, *78* (12), 2699–2706. <https://doi.org/10.1002/prot.22786>.
- (114) Brett, C. L.; Donowitz, M.; Rao, R. Does the Proteome Encode Organellar PH? *FEBS Lett.* **2006**, *580* (3), 717–719. <https://doi.org/10.1016/j.febslet.2005.12.103>.
- (115) Davidson, H. W.; Peshavaria, M.; Hutton, J. C. Proteolytic Conversion of Proinsulin into Insulin. Identification of a Ca²⁺-Dependent Acidic Endopeptidase in Isolated Insulin-Secretory Granules. *Biochem. J.* **1987**, *246* (2), 279–286. <https://doi.org/10.1042/bj2460279>.
- (116) Casey, J. R.; Grinstein, S.; Orłowski, J. Sensors and Regulators of Intracellular PH. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (1), 50–61. <https://doi.org/10.1038/nrm2820>.
- (117) Nevo-Caspi, Y.; Amariglio, N.; Rechavi, G.; Paret, G. A-to-I RNA Editing Is Induced upon Hypoxia. *Shock* **2011**, *35* (6), 585–589. <https://doi.org/10.1097/SHK.0b013e31820fe4b7>.
- (118) Ben-Zvi, M.; Amariglio, N.; Paret, G.; Nevo-Caspi, Y. F11R Expression upon Hypoxia Is Regulated by RNA Editing. *PLoS One* **2013**, *8* (10), 77702. <https://doi.org/10.1371/journal.pone.0077702>.
- (119) Nigita, G.; Acunzo, M.; Romano, G.; Veneziano, D.; Laganà, A.; Vitiello, M.; Wernicke, D.; Ferro, A.; Croce, C. M. MicroRNA Editing in Seed Region Aligns with Cellular Changes in Hypoxic Conditions. *Nucleic Acids Res.* **2016**, *44* (13), 6298–6308. <https://doi.org/10.1093/nar/gkw532>.
- (120) Tyrtysnaia, A. A.; Lysenko, L. V.; Madamba, F.; Manzhulo, I. V.; Khotimchenko, M. Y.; Kleschevnikov, A. M. Acute Neuroinflammation Provokes Intracellular Acidification in Mouse Hippocampus. *J. Neuroinflammation* **2016**, *13* (1), 283. <https://doi.org/10.1186/s12974-016-0747-8>.

- (121) Yang, J. H.; Luo, X.; Nie, Y.; Su, Y.; Zhao, Q.; Kabir, K.; Zhang, D.; Rabinovici, R. Widespread Inosine-Containing mRNA in Lymphocytes Regulated by ADAR1 in Response to Inflammation. *Immunology* **2003**, *109* (1), 15–23. <https://doi.org/10.1046/j.1365-2567.2003.01598.x>.
- (122) Raimondo, J. V.; Burman, R. J.; Katz, A. A.; Akerman, C. J. Ion Dynamics during Seizures. *Front. Cell. Neurosci.* **2015**, *9* (OCTOBER), 419. <https://doi.org/10.3389/fncel.2015.00419>.
- (123) Raimondo, J. V.; Irkle, A.; Wefelmeyer, W.; Newey, S. E.; Akerman, C. J. Genetically Encoded Proton Sensors Reveal Activity-Dependent PH Changes in Neurons. *Front. Mol. Neurosci.* **2012**, *5* (MAY 2012), 68. <https://doi.org/10.3389/fnmol.2012.00068>.
- (124) Streit, A. K.; Derst, C.; Wegner, S.; Heinemann, U.; Zahn, R. K.; Decher, N. RNA Editing of Kv1.1 Channels May Account for Reduced Ictogenic Potential of 4-Aminopyridine in Chronic Epileptic Rats. *Epilepsia* **2011**, *52* (3), 645–648. <https://doi.org/10.1111/j.1528-1167.2011.02986.x>.
- (125) Srivastava, P. K.; Bagnati, M.; Delahaye-Duriez, A.; Ko, J. H.; Rotival, M.; Langley, S. R.; Shkura, K.; Mazzuferi, M.; Danis, B.; Van Eyll, J.; Foerch, P.; Behmoaras, J.; Kaminski, R. M.; Petretto, E.; Johnson, M. R. Genome-Wide Analysis of Differential RNA Editing in Epilepsy. *Genome Res.* **2017**, *27* (3), 440–450. <https://doi.org/10.1101/gr.210740.116>.
- (126) Li, X.; Overton, I. M.; Baines, R. A.; Keegan, L. P.; O’Connell, M. A. The ADAR RNA Editing Enzyme Controls Neuronal Excitability in *Drosophila Melanogaster*. *Nucleic Acids Res.* **2014**, *42* (2), 1139–1151. <https://doi.org/10.1093/nar/gkt909>.
- (127) Rosenthal, J. J. C.; Seeburg, P. H. A-to-I RNA Editing: Effects on Proteins Key to Neural Excitability. *Neuron* **2012**, *74* (3), 432–439. <https://doi.org/10.1016/j.neuron.2012.04.010>.
- (128) Montiel-González, M. F.; Vallecillo-Viejo, I. C.; Rosenthal, J. J. C. An Efficient System for Selectively Altering Genetic Information within MRNAs. *Nucleic Acids Res.* **2016**, *44* (21), e157–e157. <https://doi.org/10.1093/nar/gkw738>.
- (129) Macbeth, M. R.; Bass, B. L. Large-Scale Overexpression and Purification of ADARs from *Saccharomyces Cerevisiae* for Biophysical and Biochemical Studies. *Methods Enzymol.* **2007**, *424*, 319–331. [https://doi.org/10.1016/S0076-6879\(07\)24015-7](https://doi.org/10.1016/S0076-6879(07)24015-7).
- (130) Zheng, Y.; Lorenzo, C.; Beal, P. A. DNA Editing in DNA/RNA Hybrids by Adenosine Deaminases That Act on RNA. *Nucleic Acids Res.* **2017**, *45* (6), 3369–3377. <https://doi.org/10.1093/nar/gkx050>.
- (131) Doherty, E. E.; Wilcox, X. E.; van Sint Fiet, L.; Kemmel, C.; Turunen, J. J.; Klein, B.; Tantillo, D. J.; Fisher, A. J.; Beal, P. A. Rational Design of RNA Editing Guide Strands: Cytidine Analogs at the Orphan Position. *J. Am. Chem. Soc.* **2021**, *143* (18), 6865–6876. <https://doi.org/10.1021/jacs.0c13319>.
- (132) Villiger, L.; Grisch-Chan, H. M.; Lindsay, H.; Ringnalda, F.; Pogliano, C. B.; Allegri, G.; Fingerhut, R.; Häberle, J.; Matos, J.; Robinson, M. D.; Thöny, B.; Schwank, G. Treatment of a Metabolic Liver Disease by in Vivo Genome Base Editing in Adult Mice. *Nat. Med.* **2018**, *24* (10), 1519–1525. <https://doi.org/10.1038/s41591-018-0209-1>.
- (133) Sinnamon, J. R.; Kim, S. Y.; Corson, G. M.; Song, Z.; Nakai, H.; Adelman, J. P.; Mandel, G. Site-Directed RNA Repair of Endogenous Mecp2 RNA in Neurons. *Proc. Natl. Acad. Sci.* **2017**, *114* (44), E9395 LP-E9402. <https://doi.org/10.1073/pnas.1715320114>.
- (134) Yeh, W.-H.; Chiang, H.; Rees, H. A.; Edge, A. S. B.; Liu, D. R. In Vivo Base Editing of Post-

- Mitotic Sensory Cells. *Nat. Commun.* **2018**, 9 (1), 2184. <https://doi.org/10.1038/s41467-018-04580-3>.
- (135) Bass, B. L. RNA Editing by Adenosine Deaminases That Act on RNA. *Annu. Rev. Biochem.* **2002**, 71 (1), 817–846. <https://doi.org/10.1146/annurev.biochem.71.110601.135501>.
- (136) Zhao, X.; Yu, Y.-T. Targeted Pre-mRNA Modification for Gene Silencing and Regulation. *Nat. Methods* **2008**, 5 (1), 95–100. <https://doi.org/10.1038/nmeth1142>.
- (137) Karijolich, J.; Yu, Y.-T. Converting Nonsense Codons into Sense Codons by Targeted Pseudouridylation. *Nature* **2011**, 474 (7351), 395–398. <https://doi.org/10.1038/nature10165>.
- (138) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and Specific Genetic Interference by Double-Stranded RNA in *Caenorhabditis Elegans*. *Nature* **1998**, 391 (6669), 806–811. <https://doi.org/10.1038/35888>.
- (139) Wilbie, D.; Walther, J.; Mastrobattista, E. Delivery Aspects of CRISPR/Cas for in Vivo Genome Editing. *Acc. Chem. Res.* **2019**, 52 (6), 1555–1564. <https://doi.org/10.1021/acs.accounts.9b00106>.
- (140) Colella, P.; Ronzitti, G.; Mingozzi, F. Emerging Issues in AAV-Mediated *In Vivo* Gene Therapy. *Mol. Ther. - Methods Clin. Dev.* **2018**, 8, 87–104. <https://doi.org/10.1016/j.omtm.2017.11.007>.
- (141) Crudele, J. M.; Chamberlain, J. S. Cas9 Immunity Creates Challenges for CRISPR Gene Editing Therapies. *Nat. Commun.* **2018**, 9 (1), 3497. <https://doi.org/10.1038/s41467-018-05843-9>.
- (142) Stafforst, T.; Schneider, M. F. An RNA–Deaminase Conjugate Selectively Repairs Point Mutations. *Angew. Chemie Int. Ed.* **2012**, 51 (44), 11166–11169. <https://doi.org/10.1002/anie.201206489>.
- (143) Azad, M. T. A. A.; Bhakta, S.; Tsukahara, T. Site-Directed RNA Editing by Adenosine Deaminase Acting on RNA for Correction of the Genetic Code in Gene Therapy. *Gene Ther.* **2017**, 24 (12), 779–786. <https://doi.org/10.1038/gt.2017.90>.
- (144) Lehmann, K. A.; Bass, B. L. Double-Stranded RNA Adenosine Deaminases ADAR1 and ADAR2 Have Overlapping Specificities. *Biochemistry* **2000**, 39 (42), 12875–12884. <https://doi.org/10.1021/bi001383g>.
- (145) Chin, T.-M.; Lin, S.-B.; Lee, S.-Y.; Chang, M.-L.; Cheng, A. Y.-Y.; Chang, F.-C.; Pasternack, L.; Huang, D.-H.; Kan, L.-S. “Paper-Clip” Type Triple Helix Formation by 5′-d-(TC)₃Ta(CT)₃Cb(AG)₃ (a and b = 0–4) as a Function of Loop Size with and without the Pseudoisocytosine Base in the Hoogsteen Strand. *Biochemistry* **2000**, 39 (40), 12457–12464. <https://doi.org/10.1021/bi0004201>.
- (146) Burchenal, J. H.; Ciovacco, K.; Kalaher, K.; O’Toole, T.; Kiefner, R.; Dowling, M. D.; Chu, C. K.; Watanabe, K. A.; Wempfen, I.; Fox, J. J. Antileukemic Effects of Pseudoisocytidine, a New Synthetic Pyrimidine C-Nucleoside. *Cancer Res.* **1976**, 36 (4), 1520 LP – 1523.
- (147) Lu, J.; Li, N.-S.; Koo, S. C.; Piccirilli, J. A. Synthesis of Pyridine, Pyrimidine and Pyridinone C-Nucleoside Phosphoramidites for Probing Cytosine Function in RNA. *J. Org. Chem.* **2009**, 74 (21), 8021–8030. <https://doi.org/10.1021/jo9016919>.
- (148) Yang, Z.; Hutter, D.; Sheng, P.; Sismour, A. M.; Benner, S. A. Artificially Expanded Genetic Information System: A New Base Pair with an Alternative Hydrogen Bonding Pattern. *Nucleic Acids Res.* **2006**, 34 (21), 6095–6101. <https://doi.org/10.1093/nar/gkl633>.

- (149) Holz, B.; Weinhold, E.; Klimasauskas, S.; Serva, S. 2-Aminopurine as a Fluorescent Probe for DNA Base Flipping by Methyltransferases. *Nucleic Acids Res.* **1998**, *26* (4), 1076–1083. <https://doi.org/10.1093/nar/26.4.1076>.
- (150) Jeltsch, A.; Roth, M.; Friedrich, T. Mutational Analysis of Target Base Flipping by the EcoRV Adenine-N6 DNA Methyltransferase. *J. Mol. Biol.* **1999**, *285* (3), 1121–1130. <https://doi.org/https://doi.org/10.1006/jmbi.1998.2389>.
- (151) Hartono, Y. D.; Pabon-Martinez, Y. V.; Uyar, A.; Wengel, J.; Lundin, K. E.; Zain, R.; Smith, C. I. E.; Nilsson, L.; Villa, A. Role of Pseudoisocytidine Tautomerization in Triplex-Forming Oligonucleotides: In Silico and in Vitro Studies. *ACS Omega* **2017**, *2* (5), 2165–2177. <https://doi.org/10.1021/acsomega.7b00347>.
- (152) Wang, D.; Shukla, C.; Liu, X.; Schoeb, T. R.; Clarke, L. A.; Bedwell, D. M.; Keeling, K. M. Characterization of an MPS I-H Knock-in Mouse That Carries a Nonsense Mutation Analogous to the Human IDUA-W402X Mutation. *Mol. Genet. Metab.* **2010**, *99* (1), 62–71. <https://doi.org/10.1016/j.ymgme.2009.08.002>.
- (153) Scott, H. S.; Litjens, T.; Nelson, P. V.; Thompson, P. R.; Brooks, D. A.; Hopwood, J. J.; Morris, C. P. Identification of Mutations in the Alpha-L-Iduronidase Gene (IDUA) That Cause Hurler and Scheie Syndromes. *Am. J. Hum. Genet.* **1993**, *53* (5), 973–986.
- (154) Guschlbauer, W.; Duplaa, A. M.; Guy, A.; Téoule, R.; Fazakerley, G. V. Structure and in Vitro Replication of DNA Templates Containing 7,8-Dihydro-8-Oxoadenine. *Nucleic Acids Res.* **1991**, *19* (8), 1753–1758. <https://doi.org/10.1093/nar/19.8.1753>.
- (155) Park, S. H.; Doherty, E. E.; Xie, Y.; Padyana, A. K.; Fang, F.; Zhang, Y.; Karki, A.; Lebrilla, C. B.; Siegel, J. B.; Beal, P. A. High-Throughput Mutagenesis Reveals Unique Structural Features of Human ADAR1. *Nat. Commun.* **2020**, *11* (1), 5130. <https://doi.org/10.1038/s41467-020-18862-2>.
- (156) Monteleone, L. R.; Matthews, M. M.; Palumbo, C. M.; Thomas, J. M.; Zheng, Y.; Chiang, Y.; Fisher, A. J.; Beal, P. A. A Bump-Hole Approach for Directed RNA Editing. *Cell Chem. Biol.* **2019**, *26* (2), 269–277.e5. <https://doi.org/10.1016/j.chembiol.2018.10.025>.
- (157) Schneider, M. F.; Wettengel, J.; Hoffmann, P. C.; Stafforst, T. Optimal GuideRNAs for Redirecting Deaminase Activity of HADAR1 and HADAR2 in Trans. *Nucleic Acids Res.* **2014**, *42* (10), e87–e87. <https://doi.org/10.1093/nar/gku272>.
- (158) 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. Chemie Int. Ed.* **2014**, *53* (24), 6267–6271. <https://doi.org/10.1002/anie.201402634>.
- (159) Heep, M.; Mach, P.; Reautschnig, P.; Wettengel, J.; Stafforst, T. Applying Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing-G/C Substitution Stabilizes GuideRNAs against Editing. *Genes (Basel)*. **2017**, *8* (1), 34. <https://doi.org/10.3390/genes8010034>.
- (160) Prakash, T. P.; Graham, M. J.; Yu, J.; Carty, R.; Low, A.; Chappell, A.; Schmidt, K.; Zhao, C.; Aghajan, M.; Murray, H. F.; Riney, S.; Booten, S. L.; Murray, S. F.; Gaus, H.; Crosby, J.; Lima, W. F.; Guo, S.; Monia, B. P.; Swayze, E. E.; Seth, P. P. Targeted Delivery of Antisense Oligonucleotides to Hepatocytes Using Triantennary N-Acetyl Galactosamine Improves Potency 10-Fold in Mice. *Nucleic Acids Res.* **2014**, *42* (13), 8796–8807. <https://doi.org/10.1093/nar/gku531>.
- (161) Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66* (Pt 2), 125–132. <https://doi.org/10.1107/S09074444909047337>.

- (162) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67* (Pt 4), 235–242. <https://doi.org/10.1107/S0907444910045749>.
- (163) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674. <https://doi.org/10.1107/S0021889807021206>.
- (164) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with `{\it Phenix.Refine}`. *Acta Crystallogr. Sect. D* **2012**, *68* (4), 352–367. <https://doi.org/10.1107/S0907444912001308>.
- (165) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D* **2004**, *60* (12 Part 1), 2126–2132. <https://doi.org/10.1107/S0907444904019158>.
- (166) Putignano, V.; Rosato, A.; Banci, L.; Andreini, C. MetalPDB in 2018: A Database of Metal Sites in Biological Macromolecular Structures. *Nucleic Acids Res.* **2018**, *46* (D1), D459–D464. <https://doi.org/10.1093/nar/gkx989>.
- (167) Matthews, B. W. Solvent Content of Protein Crystals. *J. Mol. Biol.* **1968**, *33* (2), 491–497. [https://doi.org/https://doi.org/10.1016/0022-2836\(68\)90205-2](https://doi.org/https://doi.org/10.1016/0022-2836(68)90205-2).
- (168) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1), 215–241. <https://doi.org/10.1007/s00214-007-0310-x>.
- (169) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A. Gaussian 16, Rev. C.01. *Gaussian, Inc.* 2016.
- (170) Breneman, C. M.; Wiberg, K. B. Determining Atom-centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *J. Comput. Chem.* **1990**. <https://doi.org/10.1002/jcc.540110311>.
- (171) Katrekar, D.; Yen, J.; Xiang, Y.; Saha, A.; Meluzzi, D.; Savva, Y.; Mali, P. Efficient in Vitro and in Vivo RNA Editing via Recruitment of Endogenous ADARs Using Circular Guide RNAs. *Nat. Biotechnol.* **2022**. <https://doi.org/10.1038/s41587-021-01171-4>.
- (172) Christodoulou, J.; Grimm, A.; Maher, T.; Bennetts, B. RettBASE: The IRSA MECP2 Variation Database—a New Mutation Database in Evolution. *Hum. Mutat.* **2003**, *21* (5), 466–472. <https://doi.org/10.1002/humu.10194>.
- (173) Skelly, J. V.; Edwards, K. J.; Jenkins, T. C.; Neidle, S. Crystal Structure of an Oligonucleotide Duplex Containing G.G Base Pairs: Influence of Mispairing on DNA Backbone Conformation. *Proc. Natl. Acad. Sci.* **1993**, *90* (3), 804–808. <https://doi.org/10.1073/pnas.90.3.804>.
- (174) Li, F.; Sarkhel, S.; Wilds, C. J.; Wawrzak, Z.; Prakash, T. P.; Manoharan, M.; Egli, M. 2'-Fluoroarabino- and Arabinonucleic Acid Show Different Conformations, Resulting in Deviating RNA Affinities and Processing of Their Heteroduplexes with RNA by RNase H. *Biochemistry* **2006**, *45* (13), 4141–4152. <https://doi.org/10.1021/bi052322r>.
- (175) Maas, S.; Kawahara, Y.; Tamburro, K. M.; Nishikura, K. A-to-I RNA Editing and Human

- Disease. *RNA Biol* 3, 1–9.
- (176) Slotkin, W.; Nishikura, K. Adenosine-to-Inosine RNA Editing and Human Disease. *Genome Med* 5, 105.
- (177) Morabito, M. V. Mice with Altered Serotonin 2C Receptor RNA Editing Display Characteristics of Prader-Willi Syndrome. *Neurobiol. Dis* 39, 169–180.
- (178) Lamers, M. M.; Hoogen, B. G.; Haagmans, B. L. A. Editor-in-Chief” of Cytoplasmic Innate Immunity. *Front. Immunol* 10, 1763.
- (179) Chung, H. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. *Cell* 172, 811–824 14.
- (180) Xu, L.-D.; Öhman, M. ADAR1 Editing and Its Role in Cancer. *Genes (Basel)*. 10, 12.
- (181) Miyamura, Y. Mutations of the RNA-Specific Adenosine Deaminase Gene (DSRAD) Are Involved in Dyschromatosis Symmetrica Hereditaria. *Am. J. Hum. Genet* 73, 693–699.
- (182) Zhang, X.-J. Seven Novel Mutations of the ADAR Gene in Chinese Families and Sporadic Patients with Dyschromatosis Symmetrica Hereditaria (DSH). *Hum. Mutat* 23, 629–630.
- (183) Bhate, A.; Sun, T.; Li, J. B. ADAR1: A New Target for Immuno-Oncology Therapy. *Mol. Cell* 73, 866–868.
- (184) Källman, A. M.; Sahlin, M.; Ohman, M. ADAR2 A→I Editing: Site Selectivity and Editing Efficiency Are Separate Events. *Nucleic Acids Res* 31, 4874–4881.
- (185) Macbeth, M. R. Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science (80-.)*. 309, 1534–1539.
- (186) Wang, Y.; Havel, J.; Beal, P. A. A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1. *ACS Chem. Biol.* **2015**, 10 (11), 2512–2519. <https://doi.org/10.1021/acscchembio.5b00711>.
- (187) Barber-Zucker, S.; Shaanan, B.; Zarivach, R. Transition Metal Binding Selectivity in Proteins and Its Correlation with the Phylogenomic Classification of the Cation Diffusion Facilitator Protein Family. *Sci. Rep* 7, 16381.
- (188) Waterhouse, A. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res* 46, 296– 303.
- (189) Lin, Y.-F. MIB: Metal Ion-Binding Site Prediction and Docking Server. *J. Chem. Inf. Model* 56, 2287–2291.
- (190) Wang, Y.; Beal, P. A. Probing RNA Recognition by Human ADAR2 Using a High-Throughput Mutagenesis Method. *Nucleic Acids Res* 44, 9872–9880.
- (191) Kochańczyk, T.; Drozd, A.; Krężel, A. Relationship between the Architecture of Zinc Coordination and Zinc Binding Affinity in Proteins—Insights into Zinc Regulation. *Met. Integr. Biometal Sci* 7, 244–257.
- (192) McCall, K. A.; Huang, C.; Fierke, C. A. Function and Mechanism of Zinc Metalloenzymes. *J. Nutr* 130, 1437–1446.
- (193) Dokmanić, I.; Sikić, M.; Tomić, S. Metals in Proteins: Correlation between the Metal-Ion Type, Coordination Number and the Amino-Acid Residues Involved in the Coordination. *Acta*

Crystallogr. D. Biol. Crystallogr 64, 257–263.

- (194) Song, Y. High-Resolution Comparative Modeling with RosettaCM. *Structure* 21, 1735–1742.
- (195) Gront, D.; Kulp, D. W.; Vernon, R. M.; Strauss, C. E. M.; Baker, D. Generalized Fragment Picking in Rosetta: Design, Protocols and Applications. *PLoS One* 6, 23294.
- (196) Pei, J.; Grishin, N. V. PROMALS3D: Multiple Protein Sequence Alignment Enhanced with Evolutionary and Three-Dimensional Structural Information. *Methods Mol. Biol* 1079, 263–271.
- (197) Thompson, J.; Baker, D. Incorporation of Evolutionary Information into Rosetta Comparative Modeling. *Proteins* 79, 2380–2388.
- (198) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, 596 (7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- (199) B., A. C. Principles That Govern the Folding of Protein Chains. *Science* (80-). **1973**, 181 (4096), 223–230. <https://doi.org/10.1126/science.181.4096.223>.
- (200) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Res.* **2022**, 50 (D1), D439–D444. <https://doi.org/10.1093/nar/gkab1061>.
- (201) Chen, C. X.; Cho, D. S.; Wang, Q.; Lai, F.; Carter, K. C.; Nishikura, K. A Third Member of the RNA-Specific Adenosine Deaminase Gene Family, ADAR3, Contains Both Single- and Double-Stranded RNA Binding Domains. *RNA* **2000**, 6 (5), 755–767.
- (202) Oakes, E.; Anderson, A.; Cohen-Gadol, A.; Hundley, H. A. Adenosine Deaminase That Acts on RNA 3 (ADAR3) Binding to Glutamate Receptor Subunit B Pre-mRNA Inhibits RNA Editing in Glioblastoma. *J. Biol. Chem.* **2017**, 292 (10), 4326–4335. <https://doi.org/10.1074/jbc.M117.779868>.
- (203) Dolan, M. A.; Noah, J. W.; Hurt, D. Comparison of Common Homology Modeling Algorithms: Application of User-Defined Alignments BT - Homology Modeling: Methods and Protocols; Orry, A. J. W., Abagyan, R., Eds.; Humana Press: Totowa, NJ, 2012; pp 399–414. https://doi.org/10.1007/978-1-61779-588-6_18.
- (204) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, 7, 539. <https://doi.org/10.1038/msb.2011.75>.
- (205) Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. Metal Ions in Biological Catalysis: From Enzyme Databases to General Principles. *J. Biol. Inorg. Chem* 13, 1205–1218.
- (206) Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J. Metalloproteins and Metal Sensing. *Nature* 460, 823–830.

- (207) Baraldi, E. Structure of the PH Domain from Bruton's Tyrosine Kinase in Complex with Inositol 1,3,4,5-Tetrakisphosphate. *Structure* 7, 449–460.
- (208) Mackereth, C. D.; Arrowsmith, C. H.; Edwards, A. M.; McIntosh, L. P. Zincbundle Structure of the Essential RNA Polymerase Subunit RPB10 from Methanobacterium Thermoautotrophicum. *Proc. Natl Acad. Sci. USA* 97, 6316.
- (209) Birben, E.; Sahiner, U. M.; Sackesen, C.; Erzurum, S.; Kalayci, O. Oxidative Stress and Antioxidant Defense. *World Allergy Organ. J* 5, 9–19.
- (210) Wolff, S. P.; Dean, R. T. Fragmentation of Proteins by Free Radicals and Its Effect on Their Susceptibility to Enzymic Hydrolysis. *Biochem. J* 234, 399–403.
- (211) Davies, K. J. Protein Damage and Degradation by Oxygen Radicals. I. General Aspects. *J. Biol. Chem* 262, 9895–9901.
- (212) Lyras, L.; Cairns, N. J.; Jenner, A.; Jenner, P.; Halliwell, B. An Assessment of Oxidative Damage to Proteins, Lipids, and DNA in Brain from Patients with Alzheimer's Disease. *J. Neurochem* 68, 2061–2069.
- (213) Briner, W. The Role of Metal Regulatory Proteins in Brain Oxidative Stress: A Tutorial. *Oxid. Med. Cell. Longev* 981561.
- (214) Song, W. J.; Sontz, P. A.; Ambroggio, X. I.; Tezcan, F. A. Metals in Protein–Protein Interfaces. *Annu. Rev. Biophys* 43, 409–431.
- (215) Bender, B. J. Protocols for Molecular Modeling with Rosetta3 and RosettaScripts. *Biochemistry* 55, 4748–4763.
- (216) Kahraman, A. Cross-Link Guided Molecular Modeling with ROSETTA. *PLoS One* 8, e73411.
- (217) Zhang, Z.; Porter, J.; Tripsianes, K.; Lange, O. F. Robust and Highly Accurate Automatic NOESY Assignment and Structure Determination with Rosetta. *J. Biomol. NMR* 59, 135–145.
- (218) Alexander, N. S. RosettaEPR: Rotamer Library for Spin Label Structure and Dynamics. *PLoS One* 8, e72851.
- (219) Hirst, S. J.; Alexander, N.; McHaourab, H. S.; Meiler, J. RosettaEPR: An Integrated Tool for Protein Structure Determination from Sparse EPR Data. *J. Struct. Biol* 173, 506–514.
- (220) Andreini, C.; Cavallaro, G.; Lorenzini, S.; Rosato, A. MetalPDB: A Database of Metal Sites in Biological Macromolecular Structures. *Nucleic Acids Res* 41, 312–319.
- (221) Billy, L. J.; Y., L. E.; Jung-Ki, Y.; John, A.; Bin, X.; Emily, L.; Kun, Z.; Yuan, G.; M., C. G. Genome-Wide Identification of Human RNA Editing Sites by Parallel DNA Capturing and Sequencing. *Science* (80-.). **2009**, 324 (5931), 1210–1213.
<https://doi.org/10.1126/science.1170995>.
- (222) Tang, H.; Shrager, J. B. CRISPR/Cas-Mediated Genome Editing to Treat EGFR-Mutant Lung Cancer: A Personalized Molecular Surgical Therapy. *EMBO Mol. Med.* **2016**, 8 (2), 83–85.
<https://doi.org/10.15252/emmm.201506006>.
- (223) Kamps, M. P.; Sefton, B. M. Neither Arginine nor Histidine Can Carry out the Function of Lysine-295 in the ATP-Binding Site of P60src. *Mol. Cell. Biol.* **1986**, 6 (3), 751–757.
<https://doi.org/10.1128/mcb.6.3.751>.
- (224) Shiromoto, Y.; Sakurai, M.; Minakuchi, M.; Ariyoshi, K.; Nishikura, K. ADAR1 RNA Editing

- Enzyme Regulates R-Loop Formation and Genome Stability at Telomeres in Cancer Cells. *Nat. Commun.* **2021**, *12* (1), 1654. <https://doi.org/10.1038/s41467-021-21921-x>.
- (225) Jimeno, S.; Prados-Carvajal, R.; Fernández-Ávila, M. J.; Silva, S.; Silvestris, D. A.; Endara-Coll, M.; Rodríguez-Real, G.; Domingo-Prim, J.; Mejías-Navarro, F.; Romero-Franco, A.; Jimeno-González, S.; Barroso, S.; Cesarini, V.; Aguilera, A.; Gallo, A.; Visa, N.; Huertas, P. ADAR-Mediated RNA Editing of DNA:RNA Hybrids Is Required for DNA Double Strand Break Repair. *Nat. Commun.* **2021**, *12* (1), 5512. <https://doi.org/10.1038/s41467-021-25790-2>.
- (226) Doudna, J. A.; Charpentier, E. The New Frontier of Genome Engineering with CRISPR-Cas9. *Science* (80-.). **2014**, *346* (6213). <https://doi.org/10.1126/science.1258096>.
- (227) Wang, H. X.; Li, M.; Lee, C. M.; Chakraborty, S.; Kim, H. W.; Bao, G.; Leong, K. W. CRISPR/Cas9-Based Genome Editing for Disease Modeling and Therapy: Challenges and Opportunities for Nonviral Delivery. *Chem. Rev.* **2017**, *117* (15), 9874–9906. <https://doi.org/10.1021/acs.chemrev.6b00799>.
- (228) Wang, D.; Mou, H.; Li, S.; Li, Y.; Hough, S.; Tran, K.; Li, J.; Yin, H.; Anderson, D. G.; Sontheimer, E. J.; Weng, Z.; Gao, G.; Xue, W. Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR/Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. *Hum. Gene Ther.* **2015**, *26* (7), 432–442. <https://doi.org/10.1089/hum.2015.087>.
- (229) Chew, W. L.; Tabebordbar, M.; Cheng, J. K. W.; Mali, P.; Wu, E. Y.; Ng, A. H. M.; Zhu, K.; Wagers, A. J.; Church, G. M. A Multifunctional AAV–CRISPR–Cas9 and Its Host Response. *Nat. Methods* **2016**, *13* (10), 868–874. <https://doi.org/10.1038/nmeth.3993>.
- (230) Kim, Y.-K. RNA Therapy: Current Status and Future Potential. *Chonnam Med. J.* **2020**, *56* (2), 87–93. <https://doi.org/10.4068/cmj.2020.56.2.87>.

REUSE PERMISSIONS

The published work contained within this document is permitted to be reused within an academic dissertation. The rights and permissions pertaining to each journal are listed below:

Molecular Therapy – Cell Press – “As an author, you (or your employer or institution) may do the following: Include the article in full or in part in a thesis or dissertation (provided that this is not to be published commercially).”

Nucleic Acids Research – Oxford University Press – “OUP is pleased to grant this permission for the following uses: inclusion within your thesis or dissertation.”

Journal of the American Chemical Society – ACS – “I am a student writing my thesis. May I use papers I have authored in ACS journals, or material from them, in my thesis without obtaining explicit permission? Yes. You may reuse all or part of the Submitted, Accepted, or Published versions of your ACS papers in your thesis or dissertation.”