Title

Probing ADAR-RNA Interactions for the Rational Design of Human ADAR Family-Specific Modulators

Author

Mendoza, Herra Grajo

Publication Date

2023

Peer reviewed|Thesis/dissertation

Probing ADAR-RNA Interactions for the Rational Design
of Human ADAR Family-Specific Modulators

By

HERRA GRAJO MENDOZA
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Peter A. Beal, Chair

_____

Andrew J. Fisher

_____

Kit S. Lam

Committee in Charge

2023

For from Him and through Him and to Him are all things. To Him be glory forever. Amen.

Romans 11:36 (ESV)

Probing ADAR-RNA Interactions for the Rational Design of Human ADAR Family-Specific Modulators

Abstract

An important class of enzymes involved in RNA editing is the ADAR family (adenosine deaminases acting on RNA), which facilitate the deamination of adenosine (A) to inosine (I) in double-stranded RNA (dsRNA). Inosines are decoded as guanosines (G) in most cellular processes; hence, A-to-I editing can be considered an A-to-G substitution. Among the RNA editing enzymes, ADARs are of particular interest because a large portion of RNA editing events are due to A-to-I editing by the two catalytically active human ADARs (ADAR1 and ADAR2). ADARs have diverse roles in RNA processing, gene expression regulation, and innate immunity; and mutations in the ADAR genes and dysregulated ADAR activity have been associated with cancer, autoimmune diseases, and neurological disorders. A-to-I editing is also currently being explored for correcting disease-causing mutations in the RNA, where therapeutic guide oligonucleotides complementary to the target transcript are used to form a dsRNA substrate and site-specifically direct ADAR editing. Knowledge of the mechanism of ADAR-catalyzed reaction and the origin of its substrate selectivity will allow understanding of ADAR's role in disease biology and expedite the process of developing ADAR-targeted therapeutics.

This dissertation describes some biochemical and structural studies that were performed to gain more detailed insights into substrate recognition by both ADAR1 and ADAR2 leading to the informed design of ADAR-selective inhibitors and editing-enabling guide oligonucleotides for directed editing applications with ADARs. Chapter 1 gives a general introduction to human ADAR1 and ADAR2, their association with human diseases, and the search for ADAR inhibitors. It also describes the potential of ADARs for the emerging field of therapeutic site-directed RNA

editing and expounds on what is currently known about ADAR mechanism, substrate specificity and selectivity. Chapter 2 details some biochemical experiments aimed to characterize the ADAR2 dimerization interface and illuminate the function of ADAR dimerization on substrate recognition and editing. The information derived from these studies were then utilized to design and test protein and peptide blockers of ADAR dimerization for editing inhibition in Chapter 3. Given the lag in the molecular understanding of ADAR1 compared to ADAR2, Chapter 4 presents efforts to probe substrate recognition, obtain high resolution structure, and develop selective inhibitors of ADAR1 using 8-azanebularine-modified RNA duplexes. Finally, Chapter 5 details X-ray crystallography studies conducted to characterize a chemical modification and unique sequence motifs in guide oligonucleotides that enabled editing at non-sequence-preferred therapeutic target sites for directed editing with ADARs.

Table of Contents

List of Figures

Figure

List of Tables

Table

List of Abbreviations

5-HT$_{2C}$	5-Hydroxytryptamine receptor 2C
8-azaN	8-Azanebularine
A1dx	Inactive ADAR1 deaminase domain
A2dx	Inactive ADAR2 deaminase domain
ADA	Adenosine deaminase
ADAR	Adenosine deaminase acting on RNA
ADAR1-D	ADAR1 deaminase
ADAR2-D	ADAR2 deaminase
ADAR2-R2D	A fragment of ADAR2 containing deaminase domain + double-stranded RNA binding domain 2
AGS	Aicardi Goutières syndrome
ALRN-6924	Also known as sulanemadlin; a stapled peptide developed by Aileron Therapeutics
ALS	Amyotrophic lateral sclerosis
anti-anti	Antibiotic-antimycotic
APS	Advanced photon source
AR	Antigen receptor
AZIN1	Antizyme inhibitor 1
BLCAP	Bladder cancer-associated protein
BME	ß-Mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
c-Myc	Member of the Myc family of transcription factors
CD	Circular dichroism
COG3	Conserved oligomeric Golgi complex subunit 3
Cryo-EM	Cryogenic electron microscopy
CS	Complement strand
Cy3	Cyanine 3
CYFIP2	Cytoplasmic FMR1-interacting protein 2
DAPI	4′,6-Diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
dN	2'-Deoxynebularine
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's phosphate buffered saline
dsRBD	Double-stranded RNA binding domain
dsRNA	Double-stranded RNA
DTT	Dithiothreitol
dZ	2'-Deoxy Benner's base
E3	Ubiquitin ligase
EDTA	Ethylenediaminetetraacetic acid
EEF2K	Eukaryotic elongation factor-2 kinase
EMERGe	En Masse Evaluation of RNA Guides

FBS	Fetal bovine serum
FDA	Food and drug administration
FLNA	Filamin A
GABRA3	Gamma-aminobutyric acid receptor subunit alpha-3
GFP	Green fluorescent protein
GLI1	Glioma-associated oncogene 1
gRNA	Guide RNA
GSY1	Glycogen synthase 1
HEK293T	Human embryonic kidney 293 cells
HeLa	Human cervical cancer cell line
HER1	Hmg2p ER remodeling
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatography
IC_{50}	Half maximal inhibitory concentration
IPTG	Isopropyl β-*D*-1-thiogalactopyranoside
K_D	Dissociation constant
LNA	Locked nucleic acid
MALDI-TOF	Matrix assisted laser desorption/ionization - time of flight
MBP	Maltose binding protein
MDA5	Melanoma differentiation-associated protein 5
MDM2	Mouse double minute 2 homolog
MDM4	Member of the MDM2 family
MECP2	Methyl CpG binding protein 2
MPD	2-Methyl-2,4-pentanediol
MW	Molecular weight
NCS	Non-crystallographic symmetry
NEIL1	Nei-like DNA glycosylase 1
NES	Nuclear export signal
NGS	Next generation sequencing
NLS	Nuclear localization signal
NMR	Nuclear magnetic resonance
NP-40	Nonidet P-40
nt	Nucleotide
NUP43	Nucleoporin 43
p53	Tumor protein 53
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PPI	Protein-protein interaction
PROTAC	Proteolysis targeting chimera
PS	Phosphorothioate
Rev	Transactivating protein essential to the regulation of HIV-1 and other lentiviral protein expression
RISC	RNA-induced silencing complex
rmsd	Root mean square deviation

RNA	Ribonucleic acid
RT-PCR	Reverse transcription - polymerase chain reaction
SAR	Structure-activity relationship
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
TBE	Tris-borate EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TEG	Triethylene glycol
TEV	Tobacco etch virus
T_M	Thermal melting temperature
TMEM63B	Transmembrane protein 63B
TS	Target strand
Tudor SN	Tudor Staphylococcal nuclease-like
U87	Human glioblastoma cell line
WT	Wild-type
ZNA	Z nucleic acids (*e.g.* Z-DNA or Z-RNA)

Chapter 1

General introduction to adenosine deaminases acting on RNA (ADARs)

This chapter contains excerpts from a recently accepted article in Accounts of Chemical Research[1].

1.1. The human ADAR family

RNA editing refers to any structural change in an RNA molecule (*e.g.* insertion, deletion, or base modification) that changes its coding properties and is not a result of splicing. One of the most prevalent forms of RNA editing is the conversion of adenosine (A) to inosine (I) or A-to-I editing catalyzed by a group of duplex RNA-specific enzymes called ADARs (adenosine deaminases acting on RNA)[2]. ADARs facilitate the hydrolytic deamination at the C6 position of adenosine in double-stranded RNA (dsRNA) substrates, leading to the formation of the non-canonical nucleobase, inosine (Fig. 1.1a)[3]. Because of similar base pairing properties, inosine functions analogously to guanosine (G) in many cellular processes, including splicing, translation, and reverse transcription[4,5].

The two catalytically active human ADARs, ADAR1 and ADAR2, are comprised of a deaminase domain and dsRNA-binding domains (dsRBDs)[3], typical of dsRNA binding or modifying proteins (Fig. 1.1b)[6]. However, existing experimental data support that the deaminase domain of both enzymes also require a duplex RNA structure for substrate binding[7,8]. ADAR1 also contains an additional nucleic acid binding domain for Z-DNA/RNA binding (Zα and Zß) and exists primarily as the two isoforms, p110 and p150 (Fig. 1.1b)[3]. The Zß domain lacks some crucial residues for effective Z nucleic acid (ZNA) binding, but is suggested to play a role in ADAR1 dimerization[9]. On the other hand, the Zα domain can bind both Z-DNA and Z-RNA and is known to regulate gene expression[10] and to facilitate editing[11].

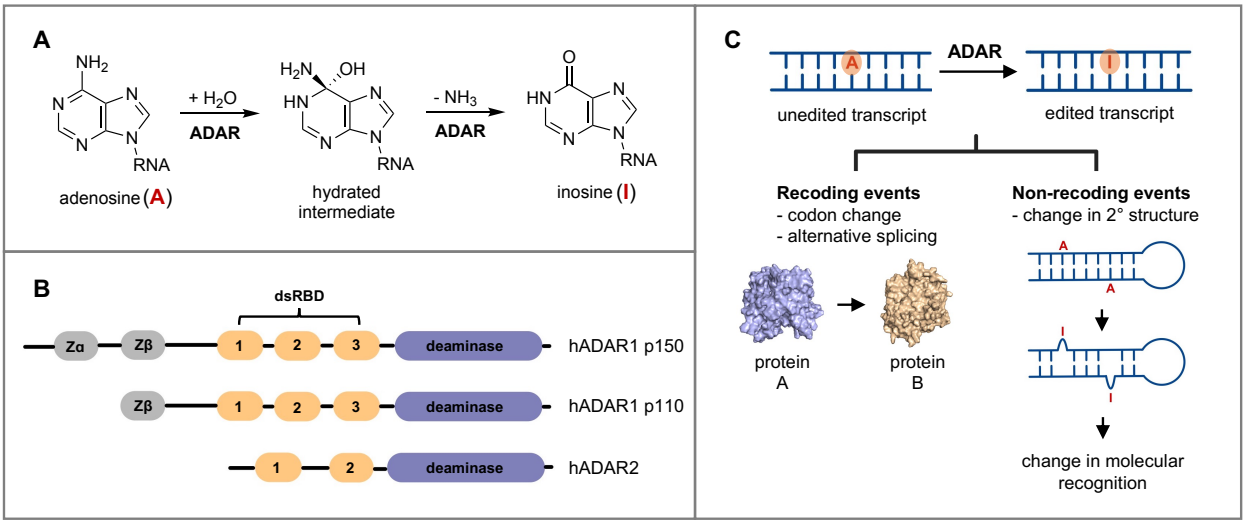


Figure 1.1. (**A**) ADAR catalyzes the hydrolytic deamination of adenosine to inosine in dsRNA via a tetrahedral reaction intermediate. (**B**) Protein domain maps of human ADAR1 and ADAR2. (**C**) General consequences of A-to-I editing by ADARs.

The consequences of A-to-I editing by ADARs are context-dependent, and vary across different RNA molecules, tissues, and developmental stages. In general, A-to-I editing can result in amino acid substitutions as well as disruption or creation of splice sites. These recoding events contribute to the formation of novel protein isoforms (Fig. 1.1c)[2]. A-to-I editing can also lead to non-recoding events, which are collectively due to alterations in the RNA secondary structure. These structural changes can impact RNA stability and its interaction with other molecules, such as the RNA interference machinery and proteins involved in RNA degradation and immune responses (Fig. 1.1c)[2,12,13]. For example, ADAR-mediated A-to-I editing of the miRNA-151 precursor at two specific positions within its dsRNA foldback structure inhibits Dicer cleavage efficiency and RISC (RNA-induced silencing complex) loading leading to a reduction in mature miRNAs[14]. Inosine-containing RNAs, such as those that are hyper-edited by ADAR, are also recognized and degraded by the human Tudor SN (Staphylococcal nuclease-like)[15], a member of the RNase III family. However, the primary and likely original function of ADAR editing,

particularly of ADAR1, appears to be the inhibition of dsRNA-induced activation of the innate immune response mediated by MDA5 (melanoma differentiation-associated protein 5). MDA5 binds long and highly complementary dsRNAs similar to those formed in viruses[16]. A-to-I editing by ADARs on endogenous or "self" dsRNAs causes duplex destabilization, preventing their misrecognition as "non-self" by MDA5 and false activation of the innate immune system[2].

1.2. Role of ADARs in human diseases

Mutations in the ADAR genes as well as aberrant ADAR-mediated A-to-I editing have been associated with various human diseases, including neurological disorders and cancer[17,18], highlighting the functional significance of this tightly-regulated RNA modification process. Mutations in the *ADAR1* gene have been linked to several type I interferenopathies and the autoimmune disorder Aicardi Goutières syndrome (AGS)[19,20] while studies suggest the association of *ADAR2* gene mutations with excitability, psychiatric, and neurodegenerative conditions such as epilepsy, bipolar disorders, and ALS (amyotrophic lateral sclerosis)[18,21]. Increased editing of the AZIN1 transcript by ADAR1 leads to increased levels of two oncoproteins (ornithine decarboxylase and cyclin D1) in human hepatocellular carcinoma[22]. In addition, ADAR1 binding and editing have demonstrated roles in synergizing immune checkpoint blockade and in the survival of cancer cells characterized by increased expression of interferon-stimulated genes[23,24]. A recent study has also shown the involvement of ADAR1 p150's Zα domain in cancer cell survival by sequestering ZNAs that are important ligands of the ZNA binding protein 1 (ZBP1) in order to stimulate a necroptotic response[25].

The growing list of ADAR's diverse roles in various cancer types has led to the increasing interest in the development of ADAR inhibitors as novel anti-cancer drugs[8,20,26]. Small-molecule inhibitors that are structural mimics of adenosine, *e.g.* 8-azaadenosine (8-aza-A) and 8-

chloroadenosine (8-Cl-A), have been identified to suppress ADAR activity in various cancers[27–29]. It was suggested that 8-aza-A targets the catalytic domain of ADAR enzymes, preventing their interaction with RNA substrates and impeding the A-to-I deamination process[30]. 8-Cl-A, on the other hand, does not directly inhibit deaminase activity but reduces ADAR expression[29,30]. However, a recent study has extensively shown the non-specificity of these two small molecules for ADAR inhibition[30]. Another approach that has been utilized for ADAR editing inhibition is the precise suppression of an editing event using antisense oligonucleotides that specifically bind to and change the secondary structures of ADAR transcript targets (*e.g.* AZIN1, NEIL1, 5-HT_{2C})[31,32].

1.3. ADARs for site-directed RNA editing

An emerging therapeutic modality called site-directed RNA editing harnesses the natural process of A-to-I editing by ADARs to correct disease-causing mutations at the transcript level[33,34]. In this approach, an antisense or guide oligonucleotide complementary to the target transcript is introduced to form the required duplex substrate, recruiting ADAR to the desired edit site (Fig. 1.2). ADAR performs the corrective A-to-I edit, resulting in a transcript that would now translate a functional protein (Fig. 1.2). Current directed RNA editing methods either employ endogenous human ADARs or overexpress engineered ADAR proteins with improved editing efficiency[35,36]. The use of endogenous human ADARs offers significant advantages over ectopically expressed engineered proteins, including comparatively lower levels of off-target editing and reduced likelihood of immune stimulation[33,37]. However, directed editing with native ADARs often suffers from inefficient editing of desired target sites due to intrinsic local sequence preferences which will be discussed in detail in the succeeding sections. This has stimulated efforts to optimize the guide oligonucleotides being used in this approach[35,38–40].

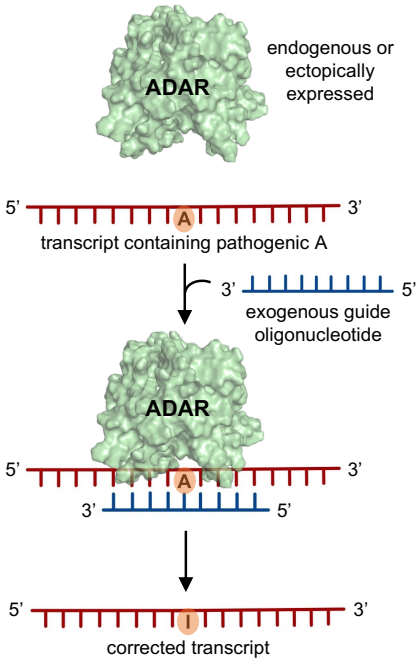


Figure 1.2. ADARs for site-directed RNA editing applications.

1.4. Catalytic mechanism and structural characterization

The crystal structure of ADAR2 deaminase domain (ADAR2-D) reported by the Bass lab revealed the presence of an inositol hexakisphosphate (IP6) molecule buried within the enzyme core and important features of the active site, including a catalytic Zn ion[41] common to other adenosine and cytidine deaminases[42,43]. However, a major breakthrough in the study of ADAR-RNA interactions came with the first set of high resolution crystal structures of ADAR2-D bound to 8-azaN-containing RNA substrates[44]. The nucleoside analog 8-azanebularine (8-azaN) was critical in the Beal lab's earlier efforts to develop ADAR editing substrate mimics which could then be used to probe the ADAR A-to-I editing reaction[45,46]. Upon hydration by ADAR, 8-azaN in RNA is converted to a tetrahedral covalent hydrate that is a structural analog of the predicted ADAR reaction intermediate (Fig. 1.3a, Fig. 1.1a) and the proposed hydrolytic deamination transition state[46]. Aza substitution at C8 of nebularine increases susceptibility of hydration across

N1-C6[45], while the absence of a good leaving group at C6 allows for the mechanistic trapping of the covalent hydrate[46]. Therefore, 8-azaN-modified RNA duplexes facilitate the formation of active site-directed and high affinity ADAR-RNA complexes suitable for biochemical and biophysical investigations with ADARs[7,8,46].

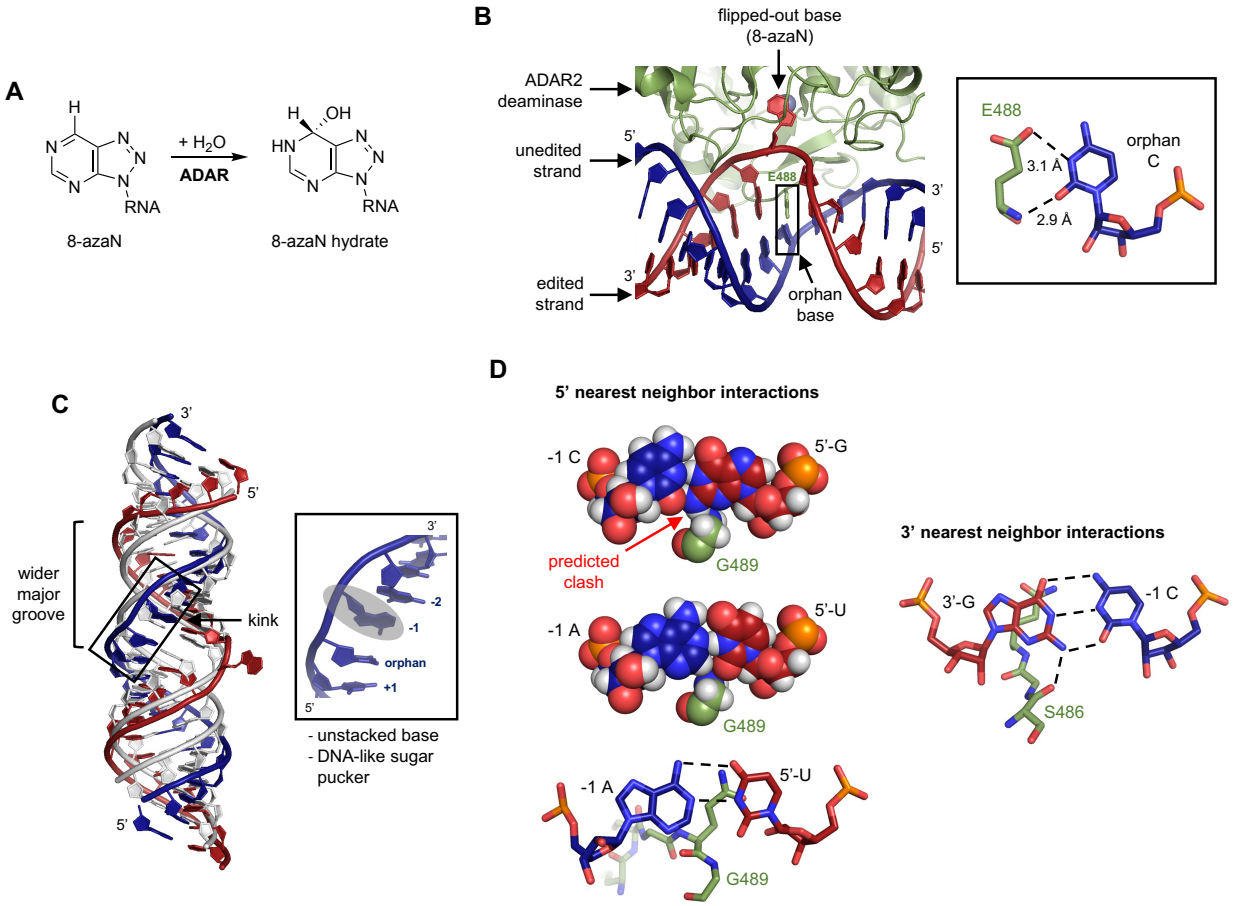


Figure 1.3. The ADAR2-RNA interface as illuminated by crystal structures of ADAR2-D bound to 8-azaN-containing RNA duplexes. (**A**) ADAR reacts with 8-azaN in RNA to form a hydrate that mimics the hydrated intermediate of ADAR's adenosine deamination (see Fig. 1.1a). (**B**) Crystal structure of human ADAR2-D (green) in complex with an 8-azaN-containing RNA duplex showing the flipped-out base in the edited strand (red) and intercalating loop bearing E488 (PDB 5hp3). E488 hydrogen bonds with the orphan base in the unedited strand (blue). (**C**) Crystal structure of the RNA duplex from an ADAR2-D-RNA co-crystal structure (PDB 5hp2) overlayed with an ideal A-form RNA duplex (silver). The unstacked base and DNA-like sugar pucker at the -1 position are highlighted in grey in the inset. (**D**) Predicted steric clash between the G489 backbone and 2-amino group of a 5'-G (top left). A 5'-U avoids this steric clash as observed in crystal structures (middle and bottom left, PDB 5ed2). The backbone carbonyl of residue S486 hydrogen bonds with the 2-amino group of a 3'-G (right, PDB 5ed2).

The mechanism of ADAR deamination clearly requires the edited adenosine to be flipped-out of the RNA duplex into the enzyme active site. Indeed, substituting the edited adenosine by 8-azaN in the duplex substrates used for crystallography studies allowed for the mechanistic trapping of the flipped-out base in the ADAR2 active site (Fig. 1.3b)[44]. Like several other enzymes that invoke this base-flipping mechanism[47,48], ADAR2 stabilizes the flipped-out conformation by utilizing a base-flipping loop. This loop, consisting of amino acid residues 486-489, intercalates into the space vacated by the flipped-out base and forms hydrogen bonding interactions with the orphaned base via residue E488 (Fig. 1.3b). Flanking E488 are two glycine residues, G487 and G489, which allow for the loop's conformational flexibility required for helix penetration. Interestingly, ADAR's observed preference for editing adenosines across a cytidine (C) or a uridine (U)[49] can be explained by a possible clash between E488 and a purine at the orphan position.

From the crystal structures, the intercalating loop is observed to approach from the minor groove side at the edited site. To accommodate this approach, the RNA duplex undergoes other conformational changes resulting in a structure with considerable deviations from the regular A-form helix adopted by dsRNAs (Fig. 1.3c). In particular, the -1 base (position across the nucleotide directly 5' to edited site) appears to be unstacked, pointing towards the helical axis in the orphan base direction. Accompanying this is a change in the -1 ribose pucker from the usual C3'-*endo* in A-form RNA helices to C2'-*endo* typical of B-form DNA helices. Overall, these conformational variations at the -1 position cause the formation of a kink in the unedited strand and a widening of the major groove opposite the editing site (Fig. 1.3c). The observations just described can also rationalize the inability of ADARs to modify deoxyadenosines in duplex DNA. The minor groove in a B-form helix is narrower and would not permit helix penetration and occupancy of the edited

base position by ADAR's base-flipping loop[44]. Indeed, the Beal lab has shown that ADARs can deaminate 2'-deoxyadenosines when present on the DNA strand of DNA/RNA hybrid duplexes consistent with the idea that ADARs require the A-form duplex conformation for reactivity[50].

1.5. Substrate specificity and selectivity

The observed contacts between ADAR2-D and the base pairs adjacent to the editing site provided a rationale for the known nearest neighbor preference of ADARs. In general, ADARs favor editing adenosines flanked by a 5'-U or A and a 3'-G[51,52]. Based on the crystal structures, it was predicted that the 2-amino group of a 5'-G (or -1 G in the case of a 5'-C) in the minor groove would clash with the protein at G489 (Fig. 1.3d). However, a U or A should not experience this clash, explaining why these nucleotides are favored at the 5' position (Fig. 1.3d). Indeed, deamination of an RNA substrate with the 5'-U:A pair replaced with a U:2AP (2-aminopurine) pair resulted in an five-fold reduction in deamination rate[44], validating the negative effect of a 2-amino group in the minor groove at this site.

The preference for a G at the 3' position arises from the hydrogen bonding contact between the 2-amino group of G and the backbone carbonyl group of S486 (Fig. 1.3d). Of the four canonical nucleobases in RNA, only G presents a hydrogen bond donor in the minor groove, hence the predilection for this base at the 3' position. Indeed, deamination of an RNA substrate with the 3'-G replaced with I (no 2-amino group) or N^2-methylguanosine (N^2MeG, blocked access to 2-amino group) resulted in a 50% decrease in rate to no observed editing, respectively[44]. These results clearly indicate the importance of a stabilizing hydrogen bonding interaction at this position.

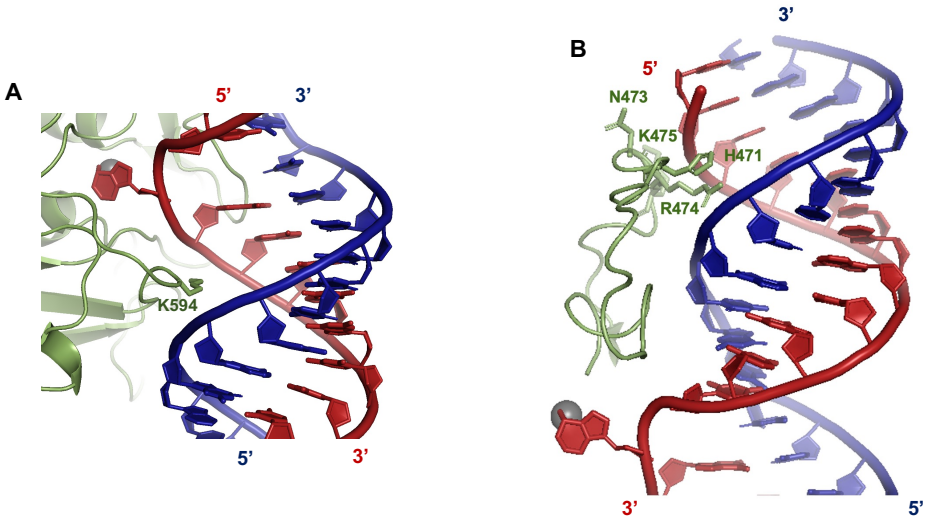


Figure 1.4. (**A**) RNA interaction facilitated by a conserved lysine residue 3' to the edit site and (**B**) the 5' binding loop of ADAR2 as observed in crystal structure (PDB 5ed2).

Aside from the interactions near the edited site, two other regions in the RNA's 5' and 3' direction relative to the flipped-out base show sites for ADAR2-D engagement. At 7-8 bp in the 3' direction, a lysine residue (K594) makes a hydrogen bonding contact with the phosphodiester backbone of the unedited strand (Fig. 1.4a). Importantly, mutating K594 to alanine (K594A) appreciably reduced ADAR2-D's editing activity, suggesting the significance of this contact for substrate recognition and editing[44]. A lysine residue is also conserved at this position in ADAR1 (K1120)[53] and experiments using known ADAR1 and ADAR2 substrates with shortened duplex 3' to the edit site (≤ 6 bp) displayed a substantial loss in binding or editing activity for both enzymes[54]. These observations imply that both ADARs require an RNA duplex with at least 8 bp in the 3' direction relative to the edited base for efficient editing.

Another region of protein-RNA contact is observed ~ 10 bp in the 5' direction from the edited base facilitated by a stretch of amino acid residues (454 to 477) comprising ADAR2's 5' binding loop (Fig. 1.4b). Interestingly, the deaminase domains of ADAR1 and ADAR2 have a relatively high sequence homology except in their 5' binding loops, providing one possible origin of the

difference in substrate selectivity of the two enzymes. Indeed, experiments using a chimeric protein with ADAR2's 5' binding loop grafted onto ADAR1-D exhibited ADAR2-like substrate selectivity[55].

1.6. Goals and perspectives

Paramount to the success of utilizing ADARs for directed editing as well as the development of therapeutic interventions for ADAR-associated diseases is a deeper understanding of ADAR structure, substrate interaction, and selectivity. The deaminase domain of ADAR2 and its interaction with an RNA duplex substrate have been well characterized due availability of high-resolution crystal structures. However, the role of the dsRBDs in dimerization, substrate recognition, and editing activity still needed to be explored. Multiple studies have also shown evidence of ADAR dimerization but the location of the dimer interface was not yet well understood prior to the studies described in Chapter 2 of this dissertation. Given the crucial roles of ADAR1 in disease biology and potential for site-directed RNA editing, interrogation of its structure and substrate recognition will help accelerate the development of targeted inhibitors and the rational design of ADAR1-specific guide oligonucleotides. Finally, guide oligonucleotide optimization through discovery and structural characterization of editing-enabling chemical modifications or secondary structures at non-optimal sites will expand the scope of directed RNA editing using endogenous ADARs.

This dissertation describes efforts to fill in some of the existing knowledge gaps, starting with the biochemical studies probing the ADAR dimerization interface in Chapter 2. The results of these experiments were published in *Nucleic Acids Chemistry* in 2020. Chapter 3 discusses some preliminary studies on the development of inhibitors of the ADAR dimer interaction, while Chapter 4 covers biochemical and biophysical experiments using 8-azaN-containing duplexes to

interrogate ADAR1 structure and substrate recognition. It also describes the use of these modified duplexes for the development of ADAR1-selective inhibitors. Most of the work in Chapter 4 was published in *Biochemistry* in 2023. Finally, Chapter 5 details efforts to structurally characterize a chemical modification in guide RNA and unique sequence motifs that enhance editing at disfavored sites. Part of this chapter was published in *Nucleic Acids Research* in 2022.

1.7. References

(1) Mendoza, H. G.; Beal, P. A. Chemical Modifications in RNA : Elucidating the Chemistry of DsRNA-. *Acc. Chem. Res.* **2023**. https://doi.org/10.1021/acs.accounts.3c00390.

(2) Eisenberg, E.; Levanon, E. Y. A-to-I RNA Editing - Immune Protector and Transcriptome Diversifier. *Nat. Rev. Genet.* **2018**, *19* (8), 473–490. https://doi.org/10.1038/s41576-018-0006-1.

(3) Wang, Y.; Zheng, Y.; Beal, P. A. Adenosine Deaminases That Act on RNA (ADARs). *Enzymes* **2017**, *41*, 215–268. https://doi.org/10.1016/bs.enz.2017.03.006.

(4) 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.

(5) 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.

(6) St Johnston, D.; Brown, N. H.; Gall, J. G.; Jantsch, M. A Conserved Double-Stranded RNA-Binding Domain. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (22), 10979–10983. https://doi.org/10.1073/pnas.89.22.10979.

(7) 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.

(8) Mendoza, H. G.; Matos, V. J.; Park, S. H.; Pham, K. M.; Beal, P. A. Selective Inhibition of ADAR1 Using 8-Azanebularine-Modified RNA Duplexes. *Biochemistry* **2022**, *62* (8), 1376–1387. https://doi.org/10.1021/acs.biochem.2c00686.

(9) Athanasiadis, A.; Placido, D.; Maas, S.; Brown, B. A.; Lowenhaupt, K.; Rich, A. The Crystal Structure of the Zβ Domain of the RNA-Editing Enzyme ADAR1 Reveals Distinct Conserved Surfaces among Z-Domains. *J. Mol. Biol.* **2005**, *351* (3), 496–507. https://doi.org/10.1016/j.jmb.2005.06.028.

(10) Oh, D. B.; Kim, Y. G.; Rich, A. Z-DNA-Binding Proteins Can Act as Potent Effectors of Gene Expression in Vivo. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (26), 16666–16671. https://doi.org/10.1073/pnas.262672699.

(11) Koeris, M.; Funke, L.; Shrestha, J.; Rich, A.; Maas, S. Modulation of ADAR1 Editing Activity by Z-RNA in Vitro. *Nucleic Acids Res.* **2005**, *33* (16), 5362–5370. https://doi.org/10.1093/nar/gki849.

(12) Levanon, E. Y.; Eisenberg, E.; Yelin, R.; Nemzer, S.; Hallegger, M.; Shemesh, R.;

Fligelman, Z. Y.; Shoshan, A.; Pollock, S. R.; Sztybel, D.; Olshansky, M.; Rechavi, G.; Jantsch, M. F. Systematic Identification of Abundant A-to-I Editing Sites in the Human Transcriptome. *Nat. Biotechnol.* **2004**, *22* (8), 1001–1005. https://doi.org/10.1038/nbt996.

(13) Lamers, M. M.; van den Hoogen, B. G.; Haagmans, B. L. ADAR1: "Editor-in-Chief" of Cytoplasmic Innate Immunity. *Front. Immunol.* **2019**, *10* (July), 1763. https://doi.org/10.3389/fimmu.2019.01763.

(14) Kawahara, Y.; Zinshteyn, B.; Chendrimada, T. P.; Shiekhattar, R.; Nishikura, K. RNA Editing of the MicroRNA-151 Precursor Blocks Cleavage by the Dicer - TRBP Complex. *EMBO Rep.* **2007**, *8* (8), 763–769. https://doi.org/10.1038/sj.embor.7401011.

(15) Li, C. L.; Yang, W. Z.; Chen, Y. P.; Yuan, H. S. Structural and Functional Insights into Human Tudor-SN, a Key Component Linking RNA Interference and Editing. *Nucleic Acids Res.* **2008**, *36* (11), 3579–3589. https://doi.org/10.1093/nar/gkn236.

(16) Weber, F.; Wagner, V.; Rasmussen, S. B.; Hartmann, R.; Paludan, S. R. Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses. *J. Virol.* **2006**, *80* (10), 5059–5064. https://doi.org/10.1128/jvi.80.10.5059-5064.2006.

(17) Slotkin, W.; Nishikura, K. Adenosine-to-Inosine RNA Editing and Human Disease. *Genome Med.* **2013**, *5* (11), 1–13. https://doi.org/10.1186/gm508.

(18) Gallo, A.; Vukic, D.; Michalík, D.; O'Connell, M. A.; Keegan, L. P. ADAR RNA Editing in Human Disease; More to It than Meets the I. *Hum. Genet.* **2017**, *136* (9), 1265–1278. https://doi.org/10.1007/s00439-017-1837-0.

(19) Rice, G. I.; Kasher, P. R.; Forte, G. M. A.; Mannion, N. M.; Greenwood, S. M.; Szynkiewicz, 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.; Heiberg, 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.

(20) Baker, A. R.; Slack, F. J. ADAR1 and Its Implications in Cancer Development and Treatment. *Trends Genet.* **2022**, *38* (8), 821–830. https://doi.org/10.1016/j.tig.2022.03.013.

(21) 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.

(22) Chen, L.; Li, Y.; Lin, C. H.; Chan, T. H. M.; Chow, R. K. K.; Song, Y.; Liu, M.; Yuan, Y. F.; Fu, L.; Kong, K. L.; Qi, L.; Li, Y.; Zhang, N.; Tong, A. H. Y.; Kwong, D. L. W.; Man, K.; Lo, C. M.; Lok, S.; Tenen, D. G.; Guan, X. Y. Recoding RNA Editing of AZIN1 Predisposes to Hepatocellular Carcinoma. *Nat. Med.* **2013**, *19* (2), 209–216. https://doi.org/10.1038/nm.3043.

(23) Ishizuka, J. J.; Manguso, R. T.; Cheruiyot, C. K.; Bi, K.; Panda, A.; Iracheta-Vellve, A.;

Miller, B. C.; Du, P. P.; Yates, K. B.; Dubrot, J.; Buchumenski, I.; Comstock, D. E.; Brown, F. D.; Ayer, A.; Kohnle, I. C.; Pope, H. W.; Zimmer, M. D.; Sen, D. R.; Lane-Reticker, S. K.; Robitschek, E. J.; Griffin, G. K.; Collins, N. B.; Long, A. H.; Doench, J. G.; Kozono, D.; Levanon, E. Y.; Haining, W. N. Loss of ADAR1 in Tumours Overcomes Resistance to Immune Checkpoint Blockade. *Nature* **2019**, *565* (7737), 43–48. https://doi.org/10.1038/s41586-018-0768-9.

(24) 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), 1–10. https://doi.org/10.1038/s41467-018-07824-4.

(25) Zhang, T.; Yin, C.; Fedorov, A.; Qiao, L.; Bao, H.; Beknazarov, N.; Wang, S.; Gautam, A.; Williams, R. M.; Crawford, J. C.; Peri, S.; Studitsky, V.; Beg, A. A.; Thomas, P. G.; Walkley, C.; Xu, Y.; Poptsova, M.; Herbert, A.; Balachandran, S. ADAR1 Masks the Cancer Immunotherapeutic Promise of ZBP1-Driven Necroptosis. *Nature* **2022**, *606* (7914), 594–602. https://doi.org/10.1038/s41586-022-04753-7.

(26) Choudhry, H. High-Throughput Screening to Identify Potential Inhibitors of the Zα Domain of the Adenosine Deaminase 1 (ADAR1): High-Throughput Screening to Identify Potential Inhibitors. *Saudi J. Biol. Sci.* **2021**, *28* (11), 6297–6304. https://doi.org/10.1016/j.sjbs.2021.06.080.

(27) Ramírez-Moya, J.; Baker, A. R.; Slack, F. J.; Santisteban, P. ADAR1-Mediated RNA Editing Is a Novel Oncogenic Process in Thyroid Cancer and Regulates MiR-200 Activity. *Oncogene* **2020**, *39* (18), 3738–3753. https://doi.org/10.1038/s41388-020-1248-x.

(28) Zipeto, M. A.; Court, A. C.; Sadarangani, A.; Delos Santos, N. P.; Balaian, L.; Chun, H. J.; Pineda, G.; Morris, S. R.; Mason, C. N.; Geron, I.; Barrett, C.; Goff, D. J.; Wall, R.; Pellecchia, M.; Minden, M.; Frazer, K. A.; Marra, M. A.; Crews, L. A.; Jiang, Q.; Jamieson, C. H. M. ADAR1 Activation Drives Leukemia Stem Cell Self-Renewal by Impairing Let-7 Biogenesis. *Cell Stem Cell* **2016**, *19* (2), 177–191. https://doi.org/10.1016/j.stem.2016.05.004.

(29) Ding, H. Y.; Yang, W. Y.; Zhang, L. H.; Li, L.; Xie, F.; Li, H. Y.; Chen, X. Y.; Tu, Z.; Li, Y.; Chen, Y.; Yang, S. Y. 8-Chloro-Adenosine Inhibits Proliferation of MDA-MB-231 and SK-BR-3 Breast Cancer Cells by Regulating ADAR1/P53 Signaling Pathway. *Cell Transplant.* **2020**, *29*, 1–9. https://doi.org/10.1177/0963689720958656.

(30) Cottrell, K. A.; Soto-Torres, L.; Dizon, M. G.; Weber, J. D. 8-Azaadenosine and 8-Chloroadenosine Are Not Selective Inhibitors of ADAR. *Cancer Res. Commun.* **2021**, *1* (2), 56–64. https://doi.org/10.1158/2767-9764.crc-21-0027.

(31) Mizrahi, R. A.; Schirle, N. T.; Beal, P. A. Potent and Selective Inhibition of A-to-I RNA Editing with 2′-O-Methyl/Locked Nucleic Acid-Containing Antisense Oligoribonucleotides. *ACS Chem. Biol.* **2013**, *8* (4), 832–839. https://doi.org/10.1021/cb300692k.

(32) Tay, D. J. T.; Song, Y.; Peng, B.; Toh, T. B.; Hooi, L.; Toh, D. F. K.; Hong, H. Q.; Tang, S. J.; Han, J.; Gan, W. L.; Chan, T. H. M.; Krishna, M. S.; Patil, K. M.; Maraswami, M.; Loh, T. P.; Dan, Y. Y.; Zhou, L.; Bonney, G. K.; Chow, P. K. H.; Chen, G.; Chow, E. K. H.; Le, M. T. N.; Chen, L. Targeting RNA Editing of Antizyme Inhibitor 1: A Potential Oligonucleotide-Based Antisense Therapy for Cancer. *Mol. Ther.* **2021**, *29* (11), 3258–

3273. https://doi.org/10.1016/j.ymthe.2021.05.008.
(33) Khosravi, H. M.; Jantsch, M. F. Site-Directed RNA Editing: Recent Advances and Open Challenges. *RNA Biol.* **2021**, *18* (S1), 41–50. https://doi.org/10.1080/15476286.2021.1983288.
(34) Diaz Quiroz, J. F.; Siskel, L. D.; Rosenthal, J. J. C. Site-Directed A → I RNA Editing as a Therapeutic Tool: Moving beyond Genetic Mutations. *Rna* **2023**, *29* (4), 498–505. https://doi.org/10.1261/rna.079518.122.
(35) Monian, P.; Shivalila, C.; Lu, G.; Shimizu, M.; Boulay, D.; Bussow, K.; Byrne, M.; Bezigian, A.; Chatterjee, A.; Chew, D.; Desai, J.; Favaloro, 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**, *40* (7), 1093–1102. https://doi.org/10.1038/s41587-022-01225-1.
(36) Katrekar, D.; Xiang, Y.; Palmer, N.; Saha, A.; Meluzzi, D.; Mali, P. Comprehensive Interrogation of the ADAR2 Deaminase Domain for Engineering Enhanced RNA Editing Activity and Specificity. *Elife* **2022**, *11*, 1–19. https://doi.org/10.7554/eLife.75555.
(37) 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.
(38) 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.
(39) Doherty, E. E.; Karki, A.; Wilcox, X. E.; Mendoza, H. G.; Manjunath, A.; Matos, V. J.; Fisher, A. J.; Beal, P. A. ADAR Activation by Inducing a Syn Conformation at Guanosine Adjacent to an Editing Site. *Nucleic Acids Res.* **2022**, 1–12. https://doi.org/10.1093/nar/gkac897.
(40) Brinkman, H. F.; Jauregui Matos, V.; Mendoza, H. G.; Doherty, E. E.; Beal, P. A. Nucleoside Analogs in ADAR Guide Strands Targeting 5′-U<u>A</u> Sites. *RSC Chem. Biol.* **2022**, *4* (1), 74–83. https://doi.org/10.1039/d2cb00165a.
(41) Macbeth, M. R.; Schubert, H. L.; VanDemark, A. F.; Lingam, A. T.; Hill, C. P.; Bass, B. L. Structural Biology: Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science (80-.).* **2005**, *309* (5740), 1534–1539. https://doi.org/10.1126/science.1113150.
(42) Betts, L.; Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W. Cytidine Deaminase. the 2.3 Å Crystal Structure of an Enzyme: Transition-State Analog Complex. *Journal of Molecular Biology*. 1994, pp 635–656. https://doi.org/10.1006/jmbi.1994.1018.
(43) Kuratani, M.; Ishii, R.; Bessho, Y.; Fukunaga, R.; Sengoku, T.; Shirouzu, M.; Sekine, S. I.; Yokoyama, S. Crystal Structure of TRNA Adenosine Deaminase (TadA) from Aquifex Aeolicus. *J. Biol. Chem.* **2005**, *280* (16), 16002–16008. https://doi.org/10.1074/jbc.M414541200.
(44) 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.
(45) Véliz, E. A.; Easterwood, L. H. M.; Beal, P. A. Substrate Analogues for an RNA-Editing Adenosine Deaminase: Mechanistic Investigation and Inhibitor Design. *J. Am. Chem. Soc.* **2003**, *125* (36), 10867–10876. https://doi.org/10.1021/ja029742d.
(46) 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.
(47) 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.
(48) Hoang, C.; Ferré-D'Amaré, A. R. Cocrystal Structure of a TRNA Ψ55 Pseudouridine Synthase: Nucleotide Flipping by an RNA-Modifying Enzyme. *Cell* **2001**, *107* (7), 929–939. https://doi.org/10.1016/S0092-8674(01)00618-3.
(49) 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.
(50) 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.
(51) 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.
(52) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 1–9. https://doi.org/10.1038/ncomms1324.
(53) 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), 1–13. https://doi.org/10.1038/s41467-020-18862-2.
(54) Fukuda, M.; Oyama, Y.; Nishitarumizu, A.; Omura, M.; Nose, K.; Deshimaru, M. Identification of an RNA Element for Specific Coordination of A-to-I RNA Editing on HTR2C Pre-MRNA. *Genes to Cells* **2015**, *20* (10), 834–846. https://doi.org/10.1111/gtc.12272.
(55) Wang, Y.; Park, S. H.; Beal, P. A. Selective Recognition of RNA Substrates by ADAR Deaminase Domains. *Biochemistry* **2018**, *57* (10), 1640–1651. https://doi.org/10.1021/acs.biochem.7b01100.

Chapter 2

Investigating the role of ADAR dimerization on substrate recognition

The biochemical studies described in this chapter were part of a joint project with Beal lab alumnus Dr. Alexander Thuy-Boun. This chapter contains excerpts from the full manuscript which was published in Nucleic Acids Research in June 2020[1]*. It also contains excerpts from a recently accepted article in Accounts of Chemical Research*[2]*.*

2.1. Introduction

The crystal structures of ADAR2-D in complex with 8-azaN-bearing substrates provided highly valuable insights into the deaminase domain's function on RNA binding and editing. However, questions on the role of the dsRBDs on ADAR activity still remained. For instance, the deaminase domain of ADAR2 alone has been shown to be sufficient in recognizing and deaminating adenosines in certain substrates (*e.g.* GLI1 mRNA)[3–5], but requires the presence of dsRBDs in editing others (*e.g.* 5-HT_{2C} pre-mRNA, D-site)[6]. Furthermore, multiple studies have provided evidence of ADAR dimerization, its possible importance in ADAR activity, and the likely role of dsRBDs in dimer formation[7–9]. All these accounts further emphasized the need for uncovering the detailed function of dsRBDs on RNA recognition and editing.

Dr. Alexander Thuy-Boun then sought to crystallize a fragment of ADAR2 bearing the deaminase domain and dsRBD2 (ADAR2-R2D) bound to an 8-azaN-modified RNA duplex (Fig. 2.1a)[1]. From the resulting X-ray crystal structure, we found that the RNA-bound protein forms an asymmetric homodimer (Fig. 2.1a-b), with the deaminase domain of one monomer (monomer A) interacting with the RNA in the same fashion observed in ADAR2-D-RNA crystal structures. On the other hand, the deaminase domain of the other monomer (monomer B) is not directly involved in RNA binding. Instead, its dsRBD2 contacts the RNA at a site found in the 3' direction relative to the flipped-out base (Fig. 2.1a), consistent with results from protein-RNA footprinting

experiment with ADAR2-R2D[1]. While dsRBD2 from monomer B is shown to interact with the RNA, there is no clear electron density to model the location of dsRBD2 from monomer A.

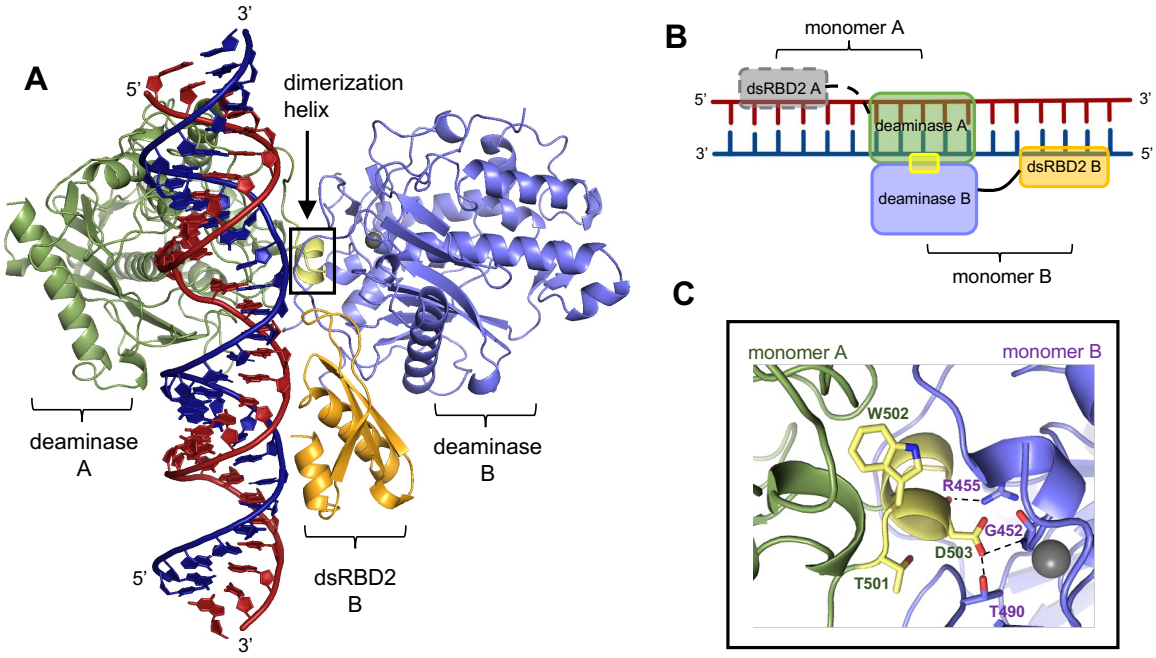


Figure 2.1. (**A**) Crystal structure of ADAR2-R2D bound to an 8-azaN-containing RNA duplex (PDB 6vff) and (**B**) cartoon representation of the asymmetric homodimer (RNA binding contacts not drawn to scale). Broken lines represent regions that were not resolved in the crystal structure. (**C**) Inset shows the α-helical dimerization interface (yellow) in the deaminase domain of monomer A and relevant hydrogen bonding contacts between the deaminase domains of monomers A and B facilitated by D503 in the dimerization helix.

As previously specified, the deaminase domain of monomer B does not directly contact the RNA. However, it interacts with the RNA-bound deaminase domain A via an extensive network of protein-protein interactions centered around an α-helical dimerization interface (dimerization helix; Fig. 2.1a and c). The dimerization helix (residues 501-509) is presented by deaminase A to a binding site in deaminase B, which consists of similar residues involved in RNA duplex binding around the catalytic pocket of deaminase A. Interestingly, three amino acid residues in the dimerization helix (T501, W502, and D503) are highly conserved in ADARs across different

species[1]. Mutating these residues to alanine led to disruption of ADAR2-R2D dimer formation, with the severity of effect following the trend D503A > W502A > T501A as shown through gel shift experiments[1]. In this chapter, the effect of these mutations in the dimer interface on ADAR editing activity *in vitro* and in cells are presented. The protein expression and purification steps as well as the *in vitro* deamination assays were performed in collaboration with Dr. Alexander Thuy-Boun.

2.2. Results

2.2.1. Mutation of the dimer interface inhibits RNA editing *in vitro*. Human ADAR2-R2D E488Q proteins containing T501A, W502A, or D503A mutations were first overexpressed and purified as described in Section 2.4. ADAR2-R2D E488Q was also similarly expressed and purified as a background control. *In vitro* deamination assays were carried out using the 5-HT_{2C} pre-mRNA, a well-studied substrate of ADAR1 and ADAR2[10–16]. In normal human cells, the D-site of this RNA transcript is edited nearly to completion by ADAR2[11]. Furthermore, Eggington and Bass reported that D-site editing is inefficient *in vitro* with the ADAR2 deaminase domain alone compared to full length ADAR2, indicating a requirement for dsRBD binding[6]. To assess the effect of dimerization helix mutation on RNA editing, *in vitro* editing assays were carried out using a transcribed 332 nt RNA fragment of the 5-HT_{2C} pre-mRNA and the ADAR2-R2D E488Q protein along with the corresponding T501A, W502A and D503A mutants (Fig. 2.2a). Interestingly, we found that mutation of the dimerization helix affects 5-HT_{2C} D-site editing to a varying degree dependent on the mutant tested. For instance, the D503A mutation reduced the observed rate constant for D-site editing by approximately 1000-fold compared to the protein bearing the wild type dimerization helix. Complete inhibition for all other known ADAR2 editing sites in 5-HT_{2C} was also observed for this mutant. The W502A mutation reduced the rate of D-site

editing by 17-fold, whereas the T501A mutation was found to be slightly stimulatory (Table 2.1). Indeed, the trend observed for changes in rate for *in vitro* editing of the 5-HT_{2C} D-site follows the same trend seen for disruption of dimer formation described in Section 2.1 (D503A>W502A>T501A).

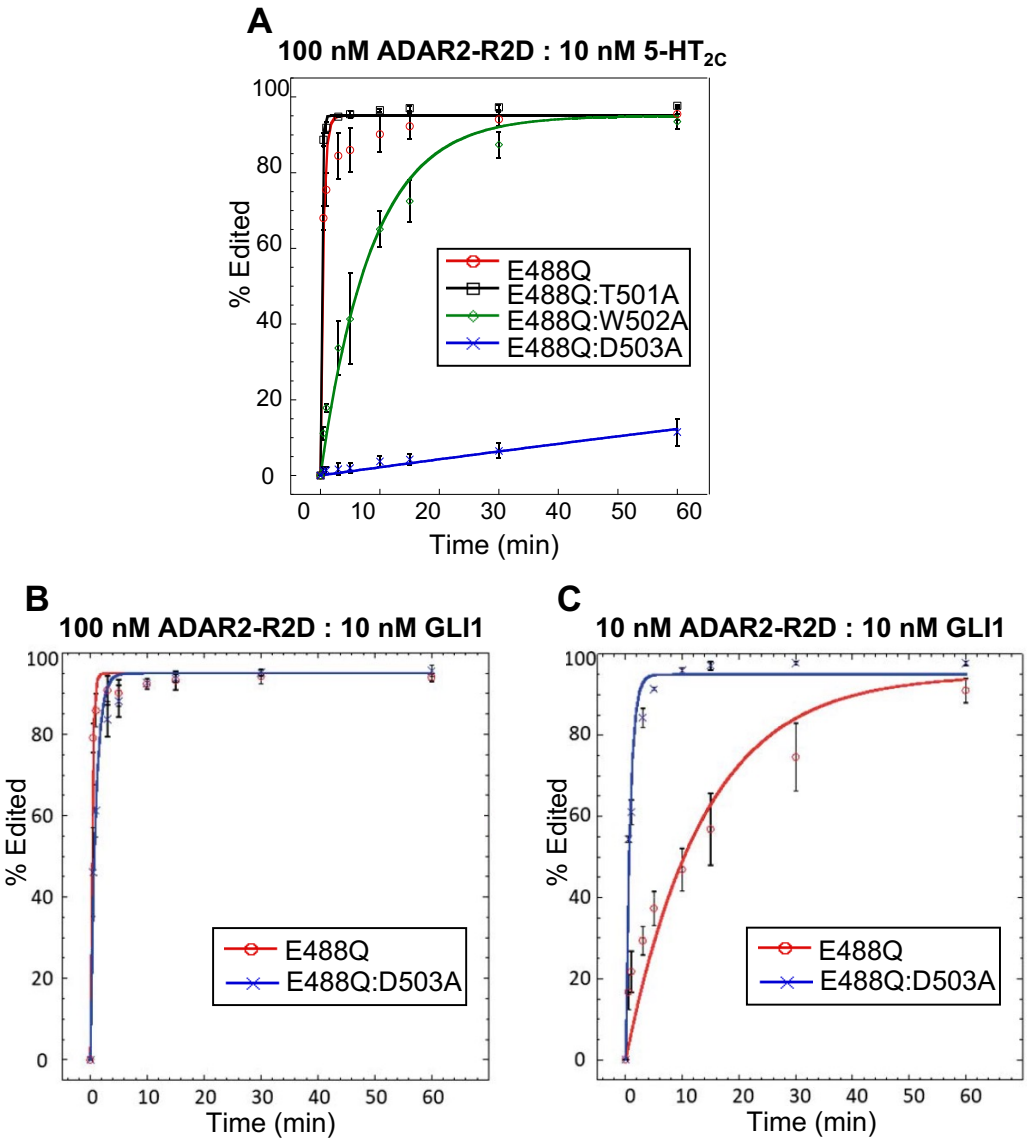


Figure 2.2. (**A**) Deamination kinetics of ADAR2-R2D E488Q and dimerization mutants (100 nM) on 5-HT_{2C} D-site (10 nM) and (**B**-**C**) GLI1 substrate (10 nM) at (**B**) 100 nM or (**C**) 10 nM enzyme. Conditions for *in vitro* deaminations are stated in Section 2.4.4. Error bars represent standard deviation (n = 3 technical replicates).

Table 2.1. Single-turnover deamination kinetics of ADAR2-R2D E488Q and dimerization mutants (100 nM) on D-site of the 5-HT_{2C} substrate (10 nM). Conditions for *in vitro* deaminations are stated in Section 2.4.4.

Enzyme	k_{obs} (min^{-1})[a]	k_{rel}[b]
ADAR2-R2D E488Q	2.13 ± 0.30	1
ADAR2-R2D E488Q:T501A	> 3	-
ADAR2-R2D E488Q:W502A	0.12 ± 0.03	0.056
ADAR2-R2D E488Q:D503A	$2.32 \times 10^{-3} \pm 0.76 \times 10^{-3}$	0.001

[a] k_{obs} was calculated by fitting the time course to the equation: $[P]_t = \alpha[1 - e^{-k_{obs}t}]$ where $[P]_t$ is percent edited at time t, α is the end point fitted to 95%, and k_{obs} is the observed rate constant.
[b] $k_{rel} = k_{obs}$ for mutant/k_{obs} for ADAR2-R2D E488Q.

For comparison, the rate constants for editing of a 147 nt RNA duplex resembling the GLI1 site was also measured (Fig. 2.2b-c). dsRBD binding has been shown to be less important for editing efficiency of this substrate[5]. Indeed, the D503A mutation is not inhibitory and instead has a slight stimulatory effect under the conditions of the *in vitro* assay used (Fig. 2.2c). At higher enzyme concentration (100 nM, Fig. 2.2b), the WT protein showed ~ 3-fold higher editing rate than the D503A mutant while at lower enzyme concentration (10 nM, Fig. 2.2c), the D503A mutant exhibited ~ 18-fold higher deamination activity than the WT protein.

2.2.2. Mutation of the dimer interface inhibits RNA editing in cells. To evaluate the effect of the dimerization mutants on the activity of full length ADAR2 in the context of human cells, these proteins were overexpressed in HEK293T cells and editing was quantified at known ADAR2 sites (Fig. 2.3a)[17]. The effect of these mutations in this experimental context showed a similar trend to that observed in the *in vitro* editing experiments using the ADAR2-R2D mutants. Once again, D503A exhibited the greatest perturbation on editing followed by W502A and T501A (Figure 2.3a). We also observed that the magnitude of the effect varied depending on the substrate analyzed, with the GLI1 and COG3 (site 2) affected much less by mutation in the dimerization helix than other targets, such as the CYFIP2 and TMEM63B transcripts. Western blot analysis of

proteins in whole cell lysate was performed for each of the ADAR transfections and confirmed similar expression levels for these ADAR mutants (Fig. 2.3b).

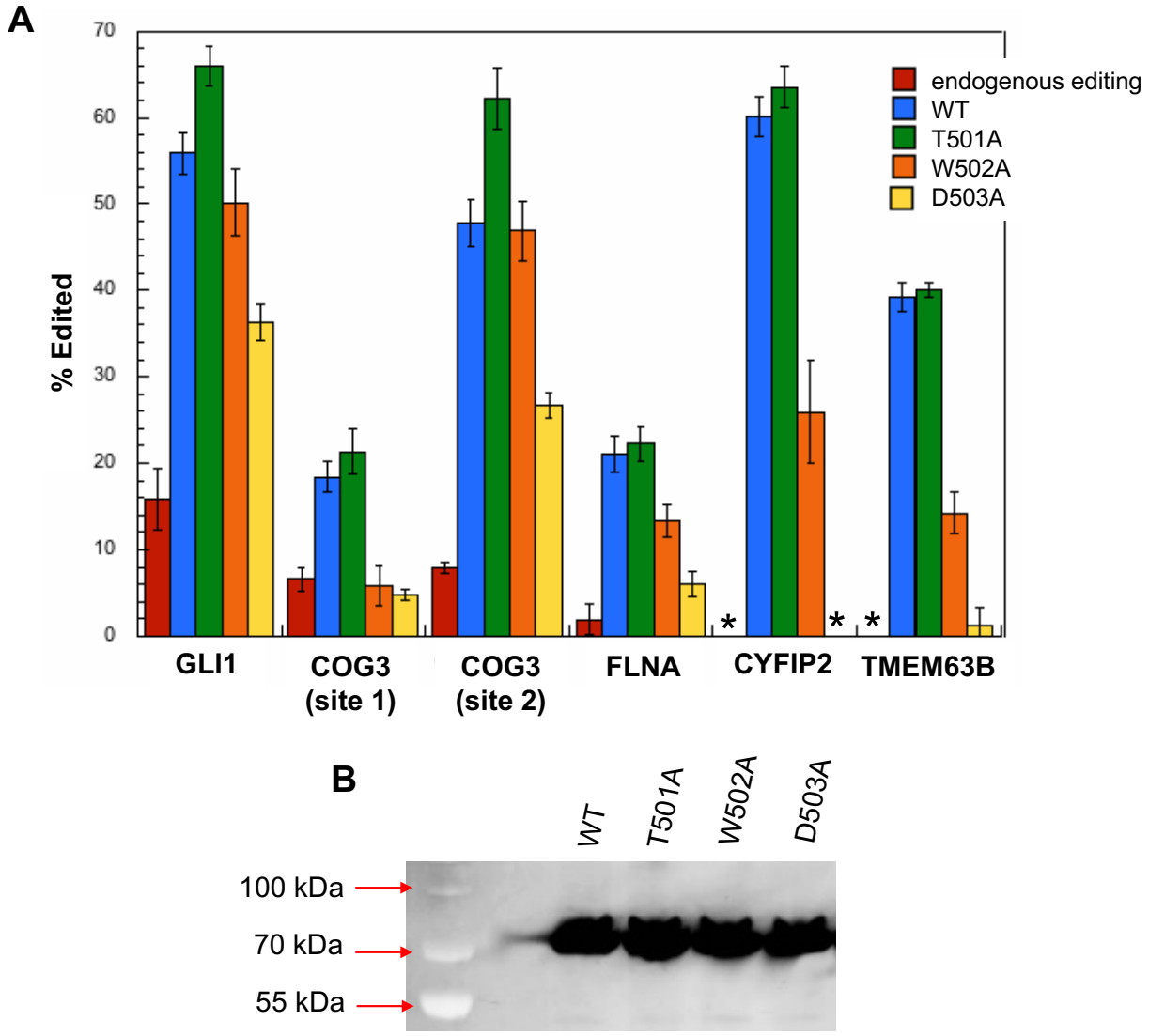


Figure 2.3. (**A**) Editing levels of endogenous substrates in HEK293T cells with or without overexpression of full length human ADAR2 wild-type (WT) and dimerization mutants. Endogenous editing was measured from cells transfected with empty vector. Asterisks indicate no detected editing. Error bars represent standard deviation (n = 3 biological replicates). (**B**) Western blot analysis of whole cell lysate from HEK293T cells expressing full length human ADAR2 WT and dimerization mutants. Expected molecular weight (MW) of full length human ADAR2 is approximately 75 kDa.

2.3. Discussion

The results of the *in vitro* and cellular deamination experiments with the dimerization mutants clearly indicate the importance of dimerization on editing of most substrates tested. The D503A mutant, in particular, exhibited the most detrimental effect on dimerization and editing, consistent with observed key protein-protein interactions facilitated by D503 between monomers A and B (Fig. 2.1c). Notably, overexpression of ADAR1 p110 with the corresponding aspartic acid residue mutation (D1023A) also displayed decreased editing of most endogenous substrates studied in HEK293T cells[1] as demonstrated by previous lab member, Dr. SeHee Park.

Altogether, the reported crystal structure of the ADAR2-R2D homodimer revealed both protein-RNA and protein-protein interactions but did not exactly disclose the involvement of dsRBD2 binding in dimer formation. However, gel shift experiments using duplex substrates with or lacking sufficient base pairs 3' to edit site for dsRBD2 contact clearly showed dependence of ADAR2-R2D dimer formation on dsRBD2-RNA interaction in this region[1]. These results, coupled with the structural and biochemical data described in this chapter, validate the essential role played by dsRBDs in dimerization, substrate recognition, and editing. Specifically, dimerization enables the simultaneous binding of a deaminase domain and a dsRBD on the same RNA molecule. For substrates with target adenosines situated in an ideal position in the duplex (mostly base-paired) and in the preferred orphan base and nearest neighbor context (*e.g.* GLI1, Fig. 2.4a), dimerization or additional binding of dsRBDs might not be necessary for efficient editing. However, the opposite is true for substrates characterized by mismatches, bulges, and loops, such as the 5-HT_{2C} RNA (Fig. 2.4b). For these substrates, the deaminase domain is inadequate for effective binding and deamination, hence dimerization is essential for supplementary RNA contact via the dsRBDs.

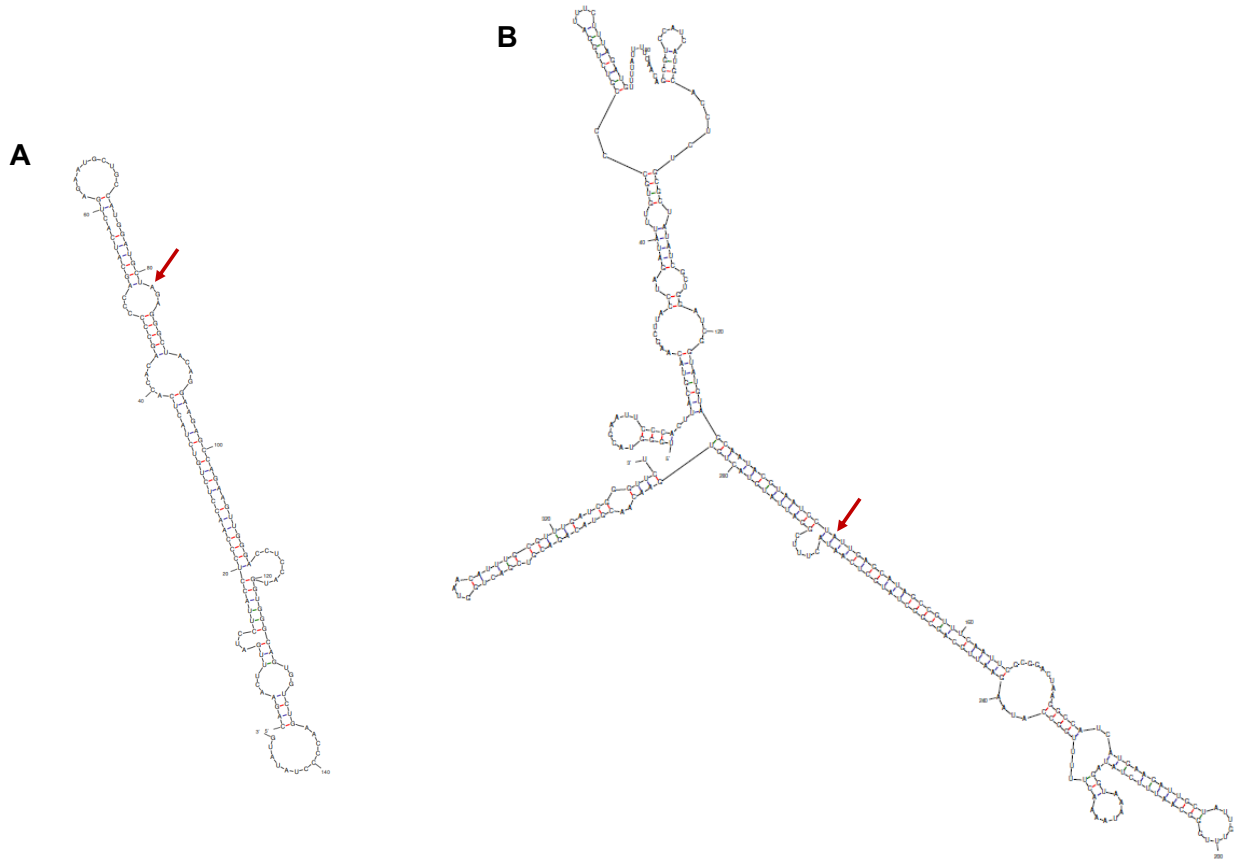


Figure 2.4. Secondary structure predicted by mFold[18] of the *in vitro* transcribed (**A**) GLI1 147mer and (**B**) 5-HT_{2C} 332mer RNA substrates used for *in vitro* deamination assays. Deaminated adenosines (D-site for 5-HT_{2C}) are indicated by red arrows.

It remains to be seen what fraction of endogenous ADAR substrates are sensitive to inhibition of dimerization. For the six ADAR2 sites analyzed here from HEK293T cells, four (CYFIP2, TMEM63B, FLNA and COG3 (site 1)) showed a substantial decrease in editing with the ADAR2 dimer interface mutants whereas two sites (GLI1 and COG3 (site 2)) were less dependent. Given the apparent significance of ADAR dimerization in editing, inhibition of dimerization presents an alternative mechanism for ADAR editing inhibition. Hence, identification of molecules that can block ADAR dimerization could lead to a new class of potential therapeutics for cancer and other ADAR-associated disorders. Efforts towards the development of ADAR dimerization inhibitors are presented in the following chapter.

2.4. Methods

2.4.1. Site-directed mutagenesis. Mutagenesis of the human ADAR2-R2D and full-length constructs was carried out via PCR site-directed mutagenesis using QuikChange XL Site-Directed Mutagenesis kit (Agilent) with the primers listed in Section 2.5. All primers were purchased from IDT. Sequences for mutant plasmids were confirmed by Sanger sequencing.

2.4.2. Expression and purification of human ADAR2-R2D. N-terminal His_{10}-tagged human ADAR2-R2D mutant proteins were overexpressed in *Saccharomyces cerevisiae* strain BCY123 as previously described[19]. Cells were lysed in lysis buffer (20 mM Tris-HCl pH 8.0, 750 mM NaCl, 35 mM imidazole, 5% (v/v) glycerol, 0.01% (v/v) Triton X-100, 1 mM ß-mercaptoethanol (BME)) using a microfluidizer and the cell lysate was purified by a 0.45 μm filter after centrifugation at 39,000 × g for 25 min. The clarified lysate was loaded through a pre-equilibrated 5 mL Ni-NTA Superflow column (Qiagen) at 2.5 mL/min. The column was washed with 100 mL of lysis buffer, followed by 50 mL washes of buffer A (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 35 mM imidazole, 5% (v/v) glycerol, 0.01% (v/v) Triton X-100, 1 mM BME) and buffer B (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 35 mM imidazole, 5% (v/v) glycerol, 0.01% (v/v) Triton X-100, 1 mM BME) at 2.5 mL/min. Elution was performed by a 45 mL gradient from buffer B to buffer C (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 300 mM imidazole, 5% (v/v) glycerol, 0.01% (v/v) Triton X-100, 1 mM BME) at 1 mL/min. Fractions were analyzed by sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) and selected fractions were pooled and loaded at 1.5 mL/min on a pre-equilibrated 5 mL HiTrap Heparin HP column (GE). The column was washed with 30 mL of buffer D (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 5% (v/v) glycerol, 1 mM BME) at 1.5 mL/min and eluted with a 45 mL gradient from buffer D to buffer E (20 mM Tris-HCl pH 8.0, 1000 mM NaCl, 5% (v/v) glycerol, 1 mM BME). Fractions were analyzed by SDS-PAGE and selected

fractions were pooled and concentrated to < 300 µL in a 30,000 MW cutoff Amicon Ultra 4 centrifugal filter at 5000 × g and 4 ºC. Buffer exchange was accomplished via three rounds of concentration to < 300 µL and subsequent addition of 3 mL of storage buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 20% (v/v) glycerol, 1 mM BME). After final centrifugation, protein concentrations were determined with bovine serum albumin (BSA) standards, as visualized by SYPRO Orange (Invitrogen) staining on SDS-PAGE gels. The purified proteins were stored at –70 ºC.

2.4.3. Preparation of substrates for *in vitro* deamination kinetics. ADAR editing substrates (5-HT_{2C} and GLI1) were *in vitro* transcribed from a DNA template using HiScribe T7 RNA Synthesis kit (NEB). The RNA products were purified using 8% urea PAGE gels and visualized by UV shadowing. Bands were excised, crushed, and soaked overnight at 4 °C in a buffer containing 0.5 M NH_4OAc and 1 mM ethylenediaminetetraacetic acid (EDTA). Gel fragments were removed using a 0.2 µm cellulose acetate filter, and the oligonucleotides were precipitated from the supernatant by ethanol precipitation at −70 °C. The oligonucleotides were then dried under vacuum, resuspended in nuclease-free water, and quantified by measuring the absorbance at 260 nm. ADAR editing substrates were allowed to self-anneal in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 100 mM NaCl by heating at 95 °C for 5 min and slow cooling to ≤ 30 °C.

2.4.4. *In vitro* deamination assay. Deamination reactions had a final volume of 10 µL with concentrations of 10 nM RNA and 10 or 100 nM ADAR2-R2D. The final reaction solution contained 17 mM Tris pH 7.4, 60 mM KCl, 16 mM NaCl, 2 mM EDTA, 0.003% Nonidet P-40 (NP-40), 0.5 mM dithiothreitol (DTT), 1.0 µg/mL yeast tRNA, and 160 units/mL RNase inhibitor (NEB). The reaction was incubated at 30 ºC. Time points include 0, 0.5, 1, 3, 5, 10, 15, 30, and 60

min. The zero time point consisted of the same deamination reaction components without the protein. Reactions were quenched by addition of 190 µL of 95 ºC nuclease free H_2O and incubation for 5 min at 95 ºC. cDNA was generated from RNA via reverse transcription - polymerase chain reaction (RT-PCR) using Access RT-PCR kit (Promega) for 24 cycles. PCR product was purified using DNA Clean and Concentrator kit (Zymo). Purified samples were subjected to Sanger sequencing and sequence traces were analyzed by 4Peaks (Nucleobytes) to quantify percent editing. Rate constants were calculated by fitting the time course to the equation: $[P]_t = \alpha[1 - e^{-k_{obs}t}]$ where $[P]_t$ is percent edited at time t, α is the end point fitted to 95%, and k_{obs} is the observed rate constant. See Section 2.5 for cDNA sequencing primers.

2.4.5. Cellular editing of endogenous substrates in HEK293T cells. HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco), 10% fetal bovine serum (FBS, Gibco) and 1% antibiotic-antimycotic (anti-anti, Gibco) at 37 ºC and 5% CO_2. At 70–90% confluency, 8×10^4 cells were seeded into a 96-well plate. After 24 h, cells were transfected with 750 ng pcDNA3.1 plasmid containing full length human ADAR2 WT or mutant with HA tag. Transfection was carried out using Lipofectamine 2000 according to manufacturer protocol (Invitrogen). After incubation of transfection reagent and plasmid in Opti-MEM (Gibco), the solution was added to designated well and incubated at 37 ºC and 5% CO_2 for 48 h. Total RNA was isolated using RNAqueous Total RNA Isolation kit (Invitrogen) and DNase treated with RQ1 RNase-free DNase (Promega). Nested PCR was performed by first using Access RT-PCR kit (Promega) for 24 cycles followed by a second PCR using Phusion Hot Start DNA Polymerase (Thermo) for 19 cycles. The PCR products were purified by agarose gel and Qiagen Gel Extraction kit. The purified product was subjected to Sanger sequencing and sequence traces were analyzed

by 4Peaks (Nucleobytes) to measure percent editing. See Section 2.5 for RT-PCR and nested PCR/sequencing primers.

2.4.6. Detection of full length human ADAR protein in transfected HEK293T cells. HEK293T cells were cultured in DMEM (Gibco), 10% FBS (Gibco), and 1% anti-anti (Gibco) at 37 °C and 5% CO_2. At 70-90% confluency, 8×10^4 cells were seeded into a 96 well plate. After 24 h, cells were transfected with 750 ng pcDNA3.1 plasmid containing full length human ADAR2 WT or mutant with HA tag. Transfection was carried out using Lipofectamine 2000. After incubation of transfection reagent and plasmid in Opti-MEM, the solution was added to designated well and incubated at 37 °C and 5% CO_2 for 48 h. Cells were lysed with 50 µL of lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% (v/v) NP-40) supplemented with Halt protease inhibitor cocktail (Thermo) with shaking at 4 °C for 30 min. Samples were resolved on an SDS-PAGE gel alongside Page Ruler Prestained Plus Protein Ladder (ThermoFisher). Western blotting was carried out using an HA tag monoclonal antibody (ThermoFisher) as primary antibody at 1:10,000 dilution and anti-mouse IgG with alkaline phosphatase-conjugated secondary antibody (Santa Cruz Biotechnology) at 1:2,000 dilution. ADAR proteins were detected using an ECF substrate (GE Healthcare) on a Typhoon Trio Variable Mode Imager (GE Healthcare).

2.5. Table of oligonucleotides

Table 2.2. Sequences of oligonucleotides used for work presented in Chapter 2.

Oligonucleotide	Sequence (5' to 3')
Mutagenesis primers	
E488Q FWD	GACCAAAATAGAGTCTGGTCAGGGGACGATTCCAGTGCGCTC
E488Q RVS	GAGCGCACTGGAATCGTCCCCTGACCAGACTCTATTTTGGTC
T501A FWD	AATGCGAGCATCCAAGCTTGGGACGGGGTGCTG
T501A RVS	CAGCACCCCGTCCCAAGCTTGGATGCTCGCATT
W502A FWD	GCGAGCATCCAAACGGCTGACGGGGTGCTGCAA
W502A RVS	TTGCAGCACCCCGTCAGCCGTTTGGATGCTCGC
D503A FWD	GAGCATCCAAACGTGGGCTGGGGTGCTGCAAGGGG
D503A RVS	CCCCTTGCAGCACCCCAGCCCACGTTTGGATGCTC
RT-PCR/sequencing primers for in vitro deamination substrates	
GLI1 FWD	CAGAACTTTGATCCTTACCTC
GLI1 RVS	GCAGTGGTCTGAACCCCTATATG
5-HT_{2C} FWD	TGGGTACGAATTCCCACTTACGTACAAGCTT
5-HT_{2C} RVS	GTAACATTTGCGTTTGATCGGGTTCT
RT-PCR and nested PCR/sequencing primers for cellular editing studies	
TMEM63B RT FWD	CCGCTGGCTCTTTGATAAGAAATTCTTGGCTGAGG
TMEM63B RT RVS	AGCCAGAAGAGGCAGAGGATGGGCG
TMEM63B Nest FWD	CAGCTATTCGGTTTGAGTGTGTGTTCC
TMEM63B Nest RVS	CGGCCACCACCTGGTTCACAGCCC
CYFIP2 RT FWD	TCCTGGCCAACCACAACAGGATCACCCAGTGTC
CYFIP2 RT RVS	TAGGTCGAAGAGCTCGCGATACTCCTCGTCTG
CYFIP2 Nest FWD	TCCACCAGCAACTTGAAGTGATCCCAGGCTATGA
CYFIP2 Nest RVS	ACTTCTGGCTGTCCAGCCCTGAGCCCG
FLNA RT FWD	TCAGTATCTGGACCCGGGAAGCTGGTGC
FLNA RT RVS	TGCCGTTGAACTTGACGTCAATCAGGTAAACGCC
FLNA Nest FWD	TGGAGGCCTGGCCATTGCTGTCGAGGG
FLNA Nest RVS	ATTCTCCCGAGGGATGAAGCGCACAGC
COG3 RT FWD	CAGATGCATAGATAGGGCAGTGTTCCAAGGA
COG3 RT RVS	ACCTTTGTCATGAACTCCTCCAGCTGTTC
COG3 Nest FWD	TTATCACAGGAAGCATTGTCTGCCTGCATTCAGTC
COG3 Nest RVS	TACAAACAGCTTGGTCTGCTGCTGAAT
GLI1 RT FWD	CGAGCCGAGTATCCAGGATACAAC
GLI1 RT RVS	CCCATATCCCAGAGTATCAGTAGGTGG
GLI1 Nest FWD	CCCAATGCAGGGGTCACCCGGAGGG
GLI1 Nest RVS	GAAGTCCATATAGGGGTTCAGACCACTGCCCAC

2.6. References

(1) Thuy-Boun, A. S.; Thomas, J. M.; Grajo, H. L.; Palumbo, C. M.; Park, S.; 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), 7958–7952. https://doi.org/10.1093/nar/gkaa532.

(2) Mendoza, H. G.; Beal, P. A. Chemical Modifications in RNA : Elucidating the Chemistry of DsRNA-. *Acc. Chem. Res.* **2023**. https://doi.org/10.1021/acs.accounts.3c00390.

(3) Macbeth, M. R.; Schubert, H. L.; VanDemark, A. F.; Lingam, A. T.; Hill, C. P.; Bass, B. L. Structural Biology: Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science (80-.).* **2005**, *309* (5740), 1534–1539. https://doi.org/10.1126/science.1113150.

(4) 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.

(5) 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.

(6) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 1–9. https://doi.org/10.1038/ncomms1324.

(7) Gallo, A.; Keegan, L. P.; Ring, G. M.; O'Connell, M. A. An ADAR That Edits Transcripts Encoding Ion Channel Subunits Functions as a Dimer. *EMBO J.* **2003**, *22* (13), 3421–3430. https://doi.org/10.1093/emboj/cdg327.

(8) Cho, D. S. C.; Yang, W.; Lee, J. T.; Shiekhattar, R.; Murray, J. M.; Nishikura, K. Requirement of Dimerization for RNA Editing Activity of Adenosine Deaminases Acting on RNA. *J. Biol. Chem.* **2003**, *278* (19), 17093–17102. https://doi.org/10.1074/jbc.M213127200.

(9) Poulsen, H.; Jorgensen, R.; Heding, A.; Nielsen, F. C.; Bonven, B.; Egebjerg, J. Dimerization of ADAR2 Is Mediated by the Double-Stranded RNA Binding Domain. *Rna* **2006**, *12* (7), 1350–1360. https://doi.org/10.1261/rna.2314406.

(10) 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.

(11) Burns, C. M.; Chu, H.; Rueter, S. M.; Hutchinson, 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. *Nature* **1997**, *387115* (May), 303–308.

(12) Fitzgerald, L. W.; Iyer, G.; Conklin, D. S.; Krause, C. M.; Marshall, A.; Patterson, J. P.; Tran, D. P.; Jonak, G. J.; Hartig, P. R. Messenger RNA Editing of the Human Serotonin 5-HT2creceptor. *Neuropsychopharmacology* **1999**, *21* (99), 82S-90S. https://doi.org/10.1016/S0893-133X(99)00004-4.

(13) Flomen, R.; Knight, J.; Sham, P.; Kerwin, R.; Makoff, A. Evidence That RNA Editing Modulates Splice Site Selection in the 5-HT2C Receptor Gene. *Nucleic Acids Res.* **2004**, *32* (7), 2113–2122. https://doi.org/10.1093/nar/gkh536.

(14) Rula, E. Y.; Emeson, R. B. *Mouse Models to Elucidate the Functional Roles of Adenosine-to-Inosine Editing*; Elsevier Masson SAS, 2007; Vol. 424. https://doi.org/10.1016/S0076-6879(07)24016-9.
(15) Wang, Q.; Miyakoda, M.; Yang, W.; Khillan, J.; Stachura, D. L.; Weiss, M. J.; Nishikura, K. Stress-Induced Apoptosis Associated with Null Mutation of ADAR1 RNA Editing Deaminase Gene. *J. Biol. Chem.* **2004**, *279* (6), 4952–4961. https://doi.org/10.1074/jbc.M310162200.
(16) Yang, W.; Wang, Q.; Kanes, S. J.; Murray, J. M.; Nishikura, K. Altered RNA Editing of Serotonin 5-HT2C Receptor Induced by Interferon: Implications for Depression Associated with Cytokine Therapy. *Mol. Brain Res.* **2004**, *124* (1), 70–78. https://doi.org/10.1016/j.molbrainres.2004.02.010.
(17) Vallecillo-Viejo, I. C.; Liscovitch-Brauer, N.; Montiel-Gonzalez, M. F.; Eisenberg, E.; Rosenthal, J. J. C. Abundant Off-Target Edits from Site-Directed RNA Editing Can Be Reduced by Nuclear Localization of the Editing Enzyme. *RNA Biol.* **2018**, *15* (1), 104–114. https://doi.org/10.1080/15476286.2017.1387711.
(18) Zuker, M. Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res.* **2003**, *31* (13), 3406–3415. https://doi.org/10.1093/nar/gkg595.
(19) 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* (07), 319–331. https://doi.org/10.1016/S0076-6879(07)24015-7.

Chapter 3

Inhibition of ADAR editing by blocking dimerization

3.1. Introduction

Targeting protein-protein interactions (PPIs) has emerged as a promising strategy for studying protein functions and for the development of novel therapeutic interventions for a wide range of medical conditions[1–3]. PPIs play crucial roles in various cellular pathways, and their perturbation can have significant impacts on cellular signaling and processes[4]. Small molecules, peptides, antibodies, and other therapeutic agents can be designed to specifically interfere with or stabilize PPIs[1,2]. For example, the stapled peptide ALRN-6924 has been shown to block the interaction of tumor suppressor p53 and its inhibitor MDM2 or MDM4, thereby exhibiting antileukemic effects[2,5]. A small molecule inhibitor that targets the androgen receptor (AR) homodimerization interface has also displayed growth suppression of AR-dependent prostate cancer cell lines[6,7]. In addition, molecular glues that enhance the interaction of binding mutant ß-catenin with its cognate E3 ligase represent a unique approach to target tumors with elevated levels of this cancer-associated transcription factor[8,9].

The crystal structure of ADAR2-R2D - RNA complex described in the last chapter revealed that ADAR2 homodimerizes via an α-helical interface and mutations in this dimerization surface, particularly in the D503 residue, disrupt dimer formation and RNA editing. The observed dependence on dimerization for recognition and editing of some substrates therefore presents a plausible mechanism for ADAR editing inhibition, which could lead to the development of targeted therapeutics against cancer and other ADAR-related metabolic and neurological disorders. In this chapter, preliminary studies on the design and testing of protein- and peptide-based blockers of ADAR dimerization for editing inhibition are presented and discussed.

3.2. Results

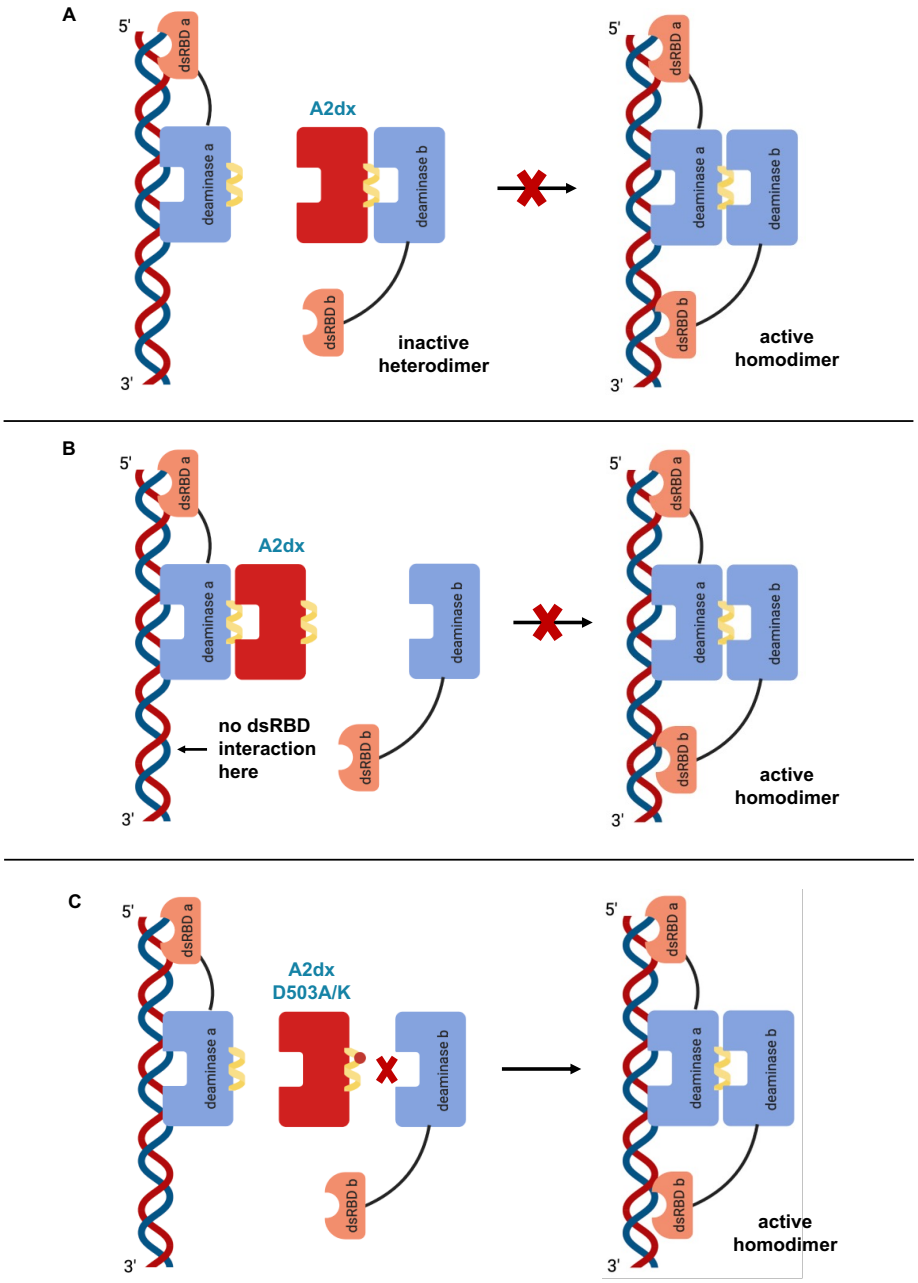


Figure 3.1. Proposed modes of inhibition by the A2dx protein inhibitor (red). Only the deaminase domain (blue) and dsRBD1 (orange) of the full-length ADAR2 protein are shown. The α-helical ADAR2 dimerization interface is shown in yellow and the edited and unedited strands of the RNA substrate are shown in dark red and dark blue, respectively. (**A**) A2dx forms an inactive heterodimer by presenting its dimer helix with ADAR2, blocking the formation of an active ADAR2 homodimer. (**B**) A2dx forms a heterodimer with ADAR2 that lacks the crucial dsRBD-RNA interaction 3' to the edited site. (**C**) A D503A/K mutation in A2dx's dimer helix (red circle) should prevent the formation of the inactive heterodimer described in A.

3.2.1. Protein inhibitor design and *in vitro* inhibition studies

To begin our investigations, we imagined that the catalytically dead deaminase domain of ADAR2 (A2dx) can act as an ADAR2 homodimerization inhibitor by forming an inactive heterodimer with the full length ADAR2 protein (Fig. 3.1a). Heterodimer formation is possible because A2dx bears the α-helical dimer surface found in ADAR2 and hence can be a protein binding partner to the full length protein. Moreover, the heterodimer is inactive because A2dx contains an inactivating mutation (E396A) to the glutamate residue that is critical for the deamination reaction[10–13]. However, since A2dx also has the residues comprising the dimerization helix receptor site, it could also engage with the full length protein to form a heterodimer that lacks the stabilizing dsRBD-RNA interaction 3' to the edited site[14] (Fig. 3.1b). To check the importance of A2dx's dimer helix in the described mode of inhibition in Fig. 3.1a, mutations in the key dimerization residue (A2dx D503A or D503K) should lead to decrease in A2dx's potency of inhibition (Fig. 3.1c).

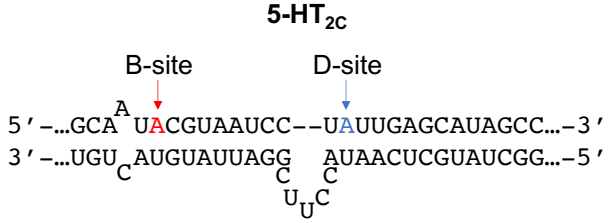


Figure 3.2. Predicted secondary structures around the 5-HT$_{2C}$ B and D editing sites. Structure predicted by mFold[15].

To test all these hypotheses, *in vitro* deamination reactions with ADAR2 and a 332 nt fragment of the human 5-HT$_{2C}$ pre-mRNA (Fig. 3.2) were set up, to which 0 to 6 μM of A2dx and A2dx D503A/K controls were added. Indeed, we observed that titration of A2dx in a full length ADAR2 reaction resulted in reduced editing of 5-HT$_{2C}$ at the dimerization dependent D-site[14] with an estimated IC$_{50}$ of 1.3 ± 0.5 μM (Fig. 3.3a). This inhibition appears to rely on A2dx-ADAR2

interaction via A2dx's dimer helix, as titration of the two A2dx control proteins showed no inhibition of editing. Interestingly, titrating A2dx in a full length ADAR1 p110 reaction did not show any reduction in 5-HT_{2C} editing at the B-site (Fig. 3.3b), an ADAR1 edit site that was also shown to require dsRBD (and hence dimerization) for recognition[16]. The ADAR2-selective inhibition by A2dx suggests that ADAR1 and ADAR2 may have unique dimerization interfaces that do not allow for heterodimerization of the two proteins.

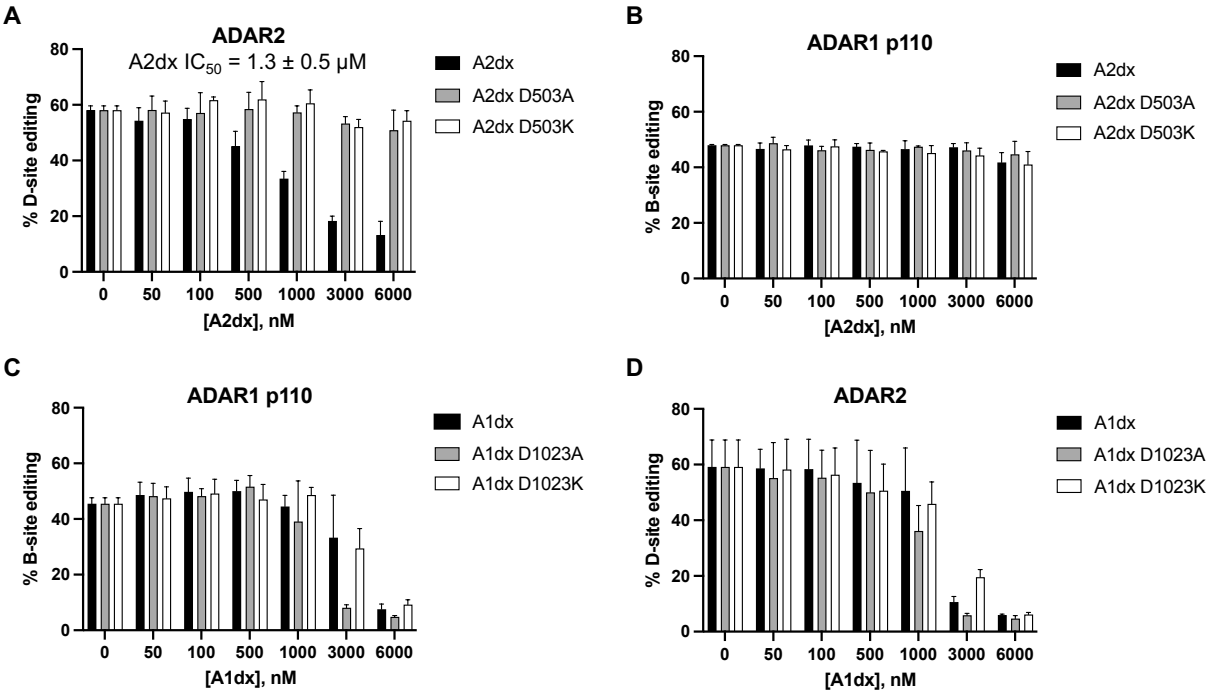


Figure 3.3. *In vitro* inhibition studies with the protein inhibitor. Titration of A2dx and dimer helix mutant controls to (**A**) ADAR2/5-HT_{2C} or (**B**) ADAR1 p110/5-HT_{2C} reaction. The IC_{50} value reported in A is the average of three independent measurements ± standard deviation. Titration of A1dx and dimer helix mutant controls to (**C**) ADAR1 p110/5-HT_{2C} or (**D**) ADAR2/5-HT_{2C} reaction. Reactions were performed at the following conditions: 0 to 6 μM protein inhibitor, 100 nM ADAR2 or ADAR1 p110, 5 nM substrate, 10 min (for ADAR2) or 15 min (for ADAR1 p110), at 30 °C. Error bars represent standard deviation from n = 3 technical replicates.

As discussed in the previous chapter, ADAR1 also contains the dimer surface residues found in ADAR2 and a charge reversal mutation in the corresponding key dimerization residue in

ADAR1 (D1023A) resulted in inhibition of editing of several endogenous substrates in HEK293T[14]. With this observation and other existing evidence that ADAR1 can also homodimerize[10,12], the catalytically dead deaminase domain of ADAR1 (ADAR1-D E912A or A1dx) was also tested for full length ADAR1 p110 inhibition along with the corresponding dimer helix mutant controls (A1dx D1023A or D1023K). However, inhibition of 5-HT_{2C} B-site editing was observed for all A1dx proteins, including the dimer helix mutant controls (Fig. 3.3c). Therefore, the editing reduction seen is not dependent on the A1dx dimer helix interface and could instead be due to competitive substrate binding. The inhibition experiment with the A1dx proteins was also performed with ADAR2 and the same inhibition pattern was seen (Fig. 3.3d), further supporting the idea that the A1dx proteins may be competitively binding to the RNA substrate.

3.2.2. Peptide inhibitor design and *in vitro* inhibition studies

The promising results from the *in vitro* inhibition studies with the A2dx inhibitor inspired us to try lower MW mimics of the ADAR2's dimerization interface such as peptides. Like the protein inhibitor, the idea is for the peptide to bind to the dimerization receptor site, hence blocking formation of an active ADAR2 homodimer (Fig. 3.4a). Four linear peptides (8, 10, 12, and 14mer) were designed from the sequence of amino acid residues making up ADAR2's dimerization helix as observed in the ADAR2-R2D - RNA crystal structure[14] (Fig. 3.4b). A 12mer cyclic peptide was also designed from the crystal structure, where the side chains of Q500 and L511 are shown to be at an ideal distance for covalent linkage. To facilitate cyclization, a 12mer peptide with Q500E and L511K mutations was used to allow for side chain to side chain amidation (Fig. 3.4b). The linear and cyclic peptides were then synthesized by and purchased from GenScript.

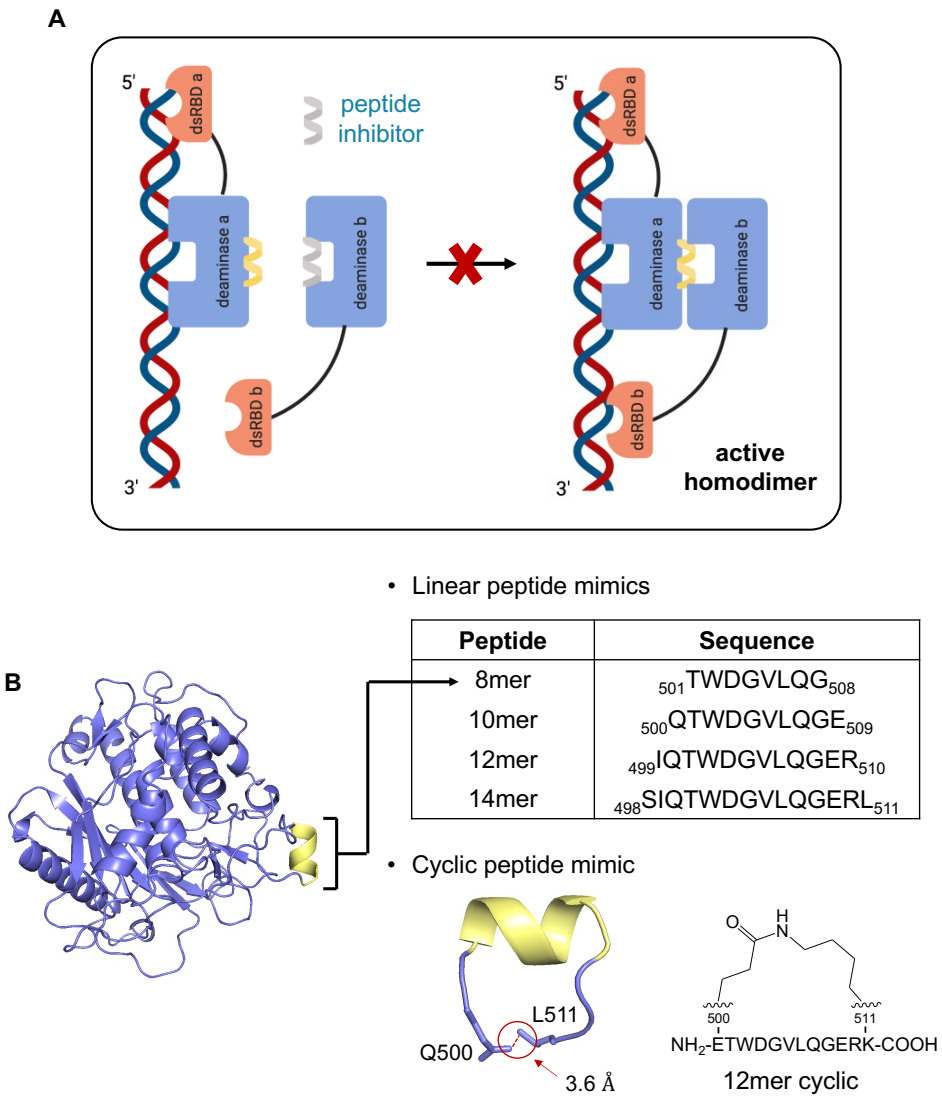


Figure 3.4. Proposed mode of inhibition of the peptide inhibitor and structure-based design. (**A**) The peptide inhibitor (gray) would putatively block ADAR2 homodimerization by competitively binding to the dimer helix receptor site. (**B**) The linear peptide sequences (written from N- to C-terminal) are derived from the sequence of amino acid residues making up the dimerization helix (yellow) in ADAR2 deaminase domain (blue). A 12mer cyclic peptide was designed based on the observed distance between Q500 and L511 side chains in the crystal structure (PDB 6vff)[14], which is ideal for side chain to chain amidation upon Q500E and L511K conversions.

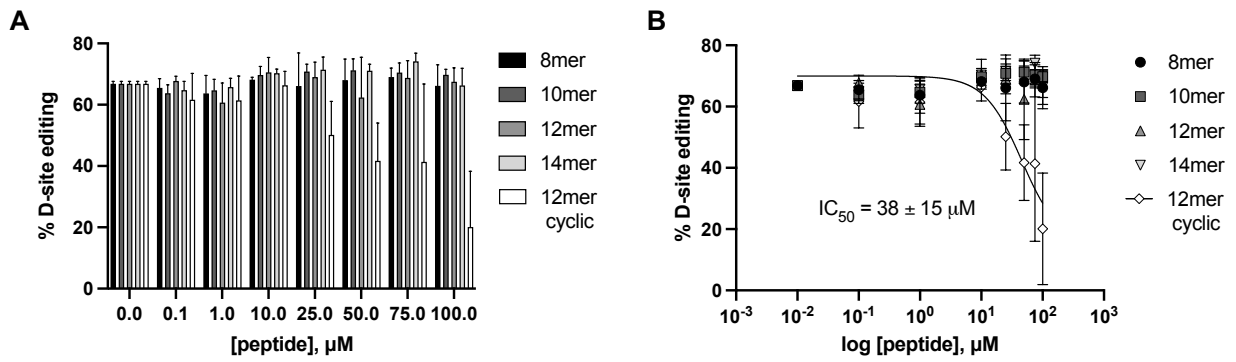


Figure 3.5. *In vitro* inhibition studies with the peptide inhibitor. (**A**) Editing of 5-HT_{2C} by ADAR2 in the presence of 0 to 100 μM peptide. (**B**) IC_{50} plot for the inhibition experiment in A. The reported IC_{50} value is the average of three independent measurements ± standard deviation. Reactions were performed at the following conditions: 100 nM ADAR2, 5 nM substrate, 10 min, at 30 °C. Error bars represent standard deviation from $n = 3$ technical replicates.

To test the peptides for ADAR2 inhibition, 0 to 100 μM of the linear or cyclic peptides were titrated into an *in vitro* deamination reaction with full length ADAR2 and 5-HT_{2C}. Figure 3.5a shows that the cyclized peptide mimic of the ADAR2 dimer helix reduced 5-HT_{2C} editing by ADAR2 with an estimated IC_{50} of 38 ± 15 μM (Fig. 3.5b). On the other hand, no inhibition was observed for the linear peptides. Peptide cyclization has been utilized as a strategy to create biostable and conformationally constrained protein domain mimics[17,18]. Therefore, it could be inferred from the inhibition experiments with the linear and cyclic peptides that rigidifying the conformation of the dimerization helix mimics improves binding and potency for blocking dimerization.

3.2.3. Cellular inhibition studies with the A2dx protein inhibitor

Efforts to test the efficacy of the A2dx protein inhibitor in inhibiting ADAR2 in mammalian cells were first led by Beal lab alumnus Dr. Kevin Pham. Using HEK293T, a cell line that has relatively low levels of endogenous ADAR2, Dr. Pham overexpressed the full length ADAR2 protein by transfecting 500 ng of plasmid containing the *ADAR2* gene. He also co-

expressed the A2dx protein inhibitor and dimer mutant controls by transfecting varying quantities of expression plasmid (0 to 750 ng; empty vector was used to keep the amount of plasmid transfected constant per treatment). Dr. Pham then evaluated the editing profile of the endogenous, dimerization-dependent substrate TMEM63B[14] 48 h post-transfection; however, no editing reduction was observed under the conditions used for his cellular inhibition studies.

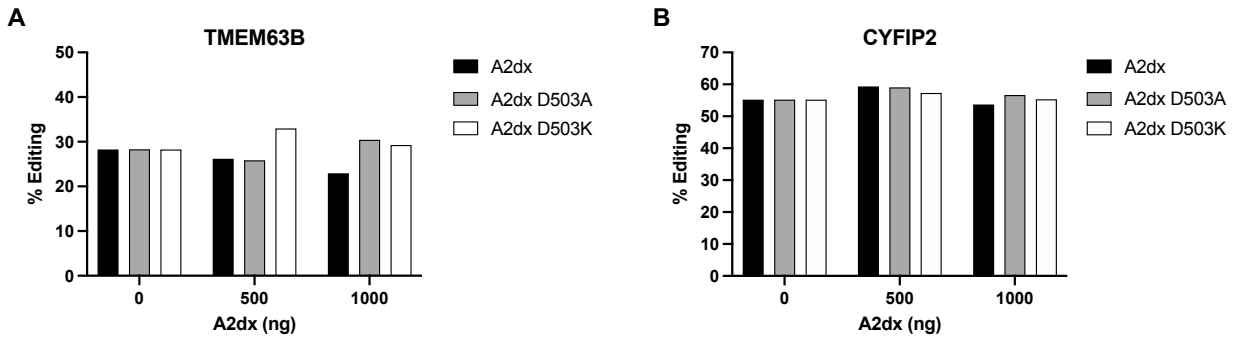


Figure 3.6. Cellular inhibition studies with the A2dx proteins. Editing levels of endogenous substrates, (**A**) TMEM63B and (**B**) CYFIP2, in HEK293T cells 48 h after transfection with 100 ng ADAR2 and 0 to 1000 ng A2dx expression plasmids.

It is important to note that for the *in vitro* experiments, editing inhibition was only observed at greater than 5-fold excess of the A2dx protein to ADAR2 (Fig. 3.3a). Hence, to continue Dr. Pham's work, the amount of expression plasmid was changed to 100 ng for the full length ADAR2 protein and up to 1000 ng for the A2dx proteins. The editing profiles for TMEM63B (Fig. 3.6a) and another endogenous, dimerization-dependent substrate CYFIP2[14] (Fig. 3.6b) were then studied after transfection. Unfortunately, no editing inhibition was observed for both transcripts even after using 10-fold excess of the A2dx protein to ADAR2 in this cellular inhibition study.

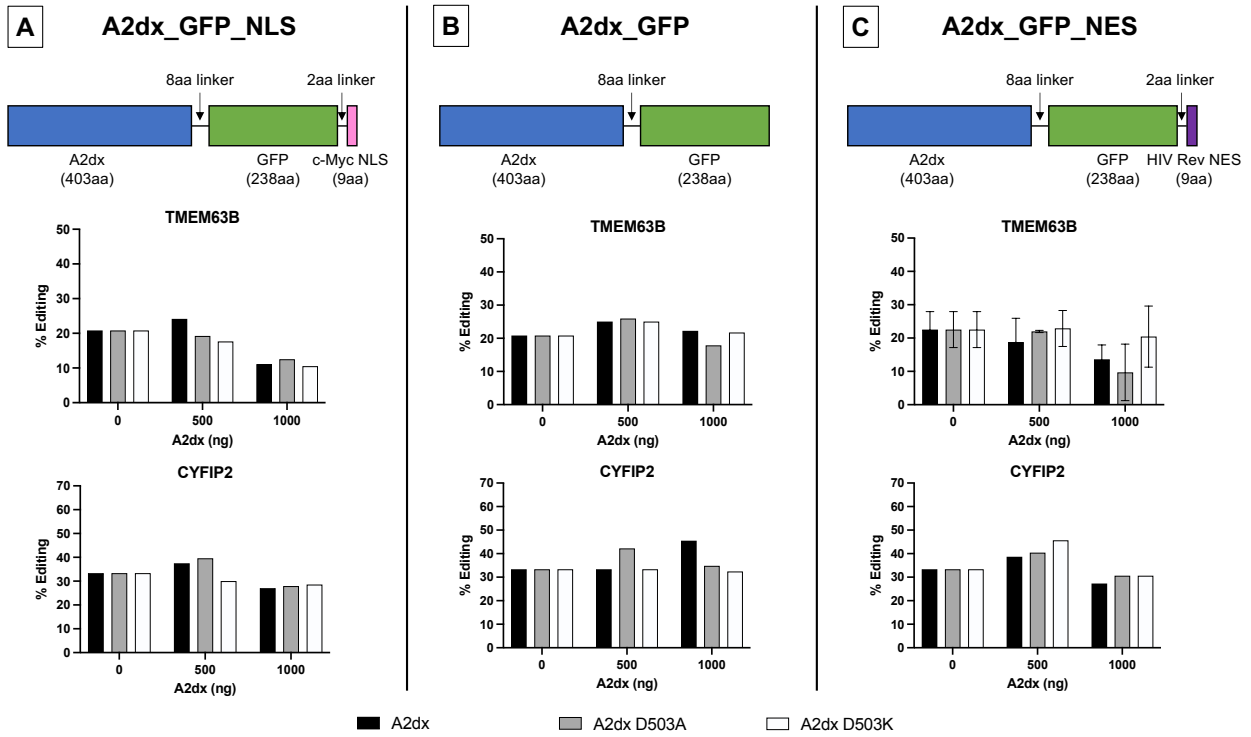


Figure 3.7. Cellular inhibition studies with the A2dx fusion proteins. Editing levels of endogenous substrates, TMEM63B and CYFIP2, in HEK293T cells 48 h after transfection with 100 ng ADAR2 and 0 to 1000 ng A2dx fusion protein expression plasmids: (**A**) A2dx_GFP_NLS, (**B**) A2dx_GFP, (**C**) A2dx_GFP_NES. Error bars represent standard deviation from n = 3 technical replicates. Protein construct maps are drawn from N- to C-terminal.

Careful consideration of where the overexpressed full length ADAR2 and A2dx proteins might localize in the cell made us consider appending cellular localization signals to the A2dx protein. A nuclear localization signal (NLS) at the N-terminus of ADAR2, particularly near its dsRBD2, allows the protein to localize in the nucleus[19,20]. This region of the protein is clearly not present in A2dx; hence, a new A2dx fusion protein construct was designed to include a C-terminal NLS derived from human c-Myc protein[21] and a green fluorescent protein (GFP) from *Aequorea victoria* to visualize and confirm cellular localization (A2dx_GFP_NLS; Fig. 3.7a). Two control A2dx fusion protein constructs, one that lacks the NLS signal (A2dx_GFP; Fig. 3.7b) and another that contains a nuclear export signal (NES) from HIV Rev protein[22,23] (A2dx_GFP_NES; Fig.

3.7c), were also prepared. The cellular inhibition experiment described above was then carried out using these new protein constructs (Fig. 3.7). A slight reduction in editing of the TMEM63B substrate was observed for the A2dx_GFP_NLS and A2dx_GFP_NES fusion proteins (Fig. 3.7a,c) but this inhibition does not appear to be dependent on the A2dx-ADAR2 interaction via A2dx's dimer helix surface, as the dimer mutant control setup also showed inhibition of substrate editing.

3.3. Discussion

The *in vitro* experiments with the A2dx protein inhibitor displayed an encouraging, dimerization interface-specific inhibition of ADAR2 editing with an IC_{50} value of 1.3 ± 0.5 μM under the conditions tested (Fig. 3.3a). Indeed, the Nishikura group has shown that a heterodimer complex made between a wild-type and a non-functional ADAR mutant, or between a wild-type and a dsRNA binding ADAR mutant, has a lower enzymatic activity compared to a wild-type homodimer[10,12]. However, the *in vitro* studies presented in this chapter indicate that this editing reduction is due to protein-protein interactions via the α-helical dimerization interface. The ADAR2-selective inhibition by A2dx also offers a clear evidence of unique dimerization interfaces which do not allow for heterodimerization between ADAR1 and ADAR2 despite their similarity in amino acid sequence in this region.

The deaminase domain of ADAR1 is characterized by more surface cysteine residues than that of ADAR2, making this protein much more prone to aggregation and more challenging to purify[24]. The inability of A1dx to inhibit ADAR1 in a dimer helix dependent manner in the *in vitro* inhibition experiment could be due to more instability issues brought about by the catalytic site and/or dimer surface mutations to the A1dx proteins. Indeed, the A1dx proteins were observed to have a higher tendency to aggregate compared to the wild-type ADAR1-D protein when purified under the same conditions. Since protein stability issues may be circumvented when we need not

have to isolate or purify the protein out of its native environment, it may be more interesting to test the A1dx proteins for ADAR1 inhibition in human cells. However, the results of the cellular inhibition experiments with the A2dx protein inhibitor are still preliminary and quite inconclusive. Immunofluorescent imaging experiments must be conducted to determine the subcellular localization of the overexpressed A2dx fusion proteins and whether there is a necessity to optimize the cellular localization signals used and fusion protein constructs designed.

In vitro inhibition experiments with the 12mer cyclic peptide showed a promising IC_{50} value of 38 ± 15 μM (Fig. 3.5b). Although this value is about 30-fold higher than that of the protein inhibitor (Fig. 3.3a), the 12mer cyclic peptide offers a good starting point for potency optimization. Additional helix stabilization methods[25] or alternative cyclization strategies[26] may be used in future optimization studies for this inhibitor. Methods such as circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy should confirm the bioactive structure of the cyclic peptide inhibitor. It is also important to validate its mode of inhibition by performing a control *in vitro* inhibition experiment with a peptide having the same modifications and sequence except containing a D503A or D503K mutation. This mutation should prevent binding of the control peptide to the dimer helix receptor site in ADAR2 and confirm that the inhibition observed with the 12mer cyclic peptide is not due to non-specific peptide-ADAR2 interactions.

3.4. Methods

3.4.1. General molecular biology procedures. Mutagenesis of the human ADAR constructs was carried out via PCR site-directed mutagenesis using QuikChange XL Site-Directed Mutagenesis kit (Agilent) with the primers listed in Table 3.1. Molecular cloning of mammalian expression plasmids for ADAR fusion proteins were performed using the Gibson Assembly Cloning kit (NEB). Plasmid sequences were confirmed by Sanger sequencing. All primers were purchased

from IDT. The linear and cyclic peptides were purchased from GenScript (see Table 3.2 for sequences, purity, and masses).

3.4.2. Protein overexpression and purification. The full length wild-type protein or deaminase domain mutants of human ADAR2 and ADAR1 p110 with an N- (all proteins except ADAR1 p110) or C-terminal (ADAR1 p110 only) His_{10}-tag were overexpressed in *S. cerevisiae* BCY123 as previously described[27]. Full length ADAR2 WT and A2dx proteins were purified by lysing cells in lysis buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 750 mM NaCl, 35 mM imidazole, 0.01% (v/v) Triton X-100, 1 mM BME) using a microfluidizer. The clarified lysate was then passed over a Ni-NTA column using an ÄKTA pure 25 FPLC system. The column was washed with the lysis buffer and then with wash buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 300 mM NaCl, 35 mM imidazole, 0.01% (v/v) Triton X-100, 1 mM BME). Bound proteins were eluted by gradient elution with imidazole (30 to 400 mM). Fractions containing the target protein were pooled, concentrated, and then dialyzed against a storage buffer containing 20 mM Tris-HCl pH 8.0, 20% (v/v) glycerol, 100 mM NaCl, 1 mM BME. Protein concentration was determined by running the sample alongside BSA standards in an SDS-PAGE gel followed by SYPRO Orange (Invitrogen) staining.

Full length ADAR1 p110 WT was purified as above except using the following buffers: (1) lysis buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 1000 mM KCl, 30 mM imidazole, 1 mM tris(2-carboxyethyl)phosphine-HCl (TCEP-HCl), 0.05% (v/v) Triton X-100, and 50 μM $ZnCl_2$); (2) wash buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 500 mM KCl, 30 mM imidazole, 1 mM TCEP-HCl, and 50 μM $ZnCl_2$); and (3) storage buffer (50 mM Tris-HCl pH 8.0, 10% (v/v) glycerol, 400 mM KCl, 50 mM imidazole, 1 mM TCEP-HCl, and 0.01% (v/v) NP-40).

The A1dx proteins were purified as above except using the following buffers: (1) lysis buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 750 mM NaCl, 30 mM imidazole, 1 mM TCEP-HCl, 0.05% (v/v) Triton X-100, and 100 μM $ZnCl_2$); (2) wash buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 350 mM NaCl, 30 mM imidazole, 1 mM TCEP-HCl, and 100 μM $ZnCl_2$); and (3) storage buffer (50 mM Tris-HCl pH 8.0, 10% (v/v) glycerol, 200 mM KCl, 50 mM imidazole, 1 mM TCEP-HCl, 0.01% (v/v) NP-40).

3.4.3. Preparation of 5-HT_{2C} substrate for *in vitro* deamination. A 332 nt fragment of the 5-HT_{2C} transcript was *in vitro* transcribed from a DNA template using HiScribe T7 RNA Synthesis kit (NEB). The RNA product was purified using 8% urea PAGE gels and visualized by UV shadowing. Bands were excised, crushed, and soaked overnight at 4 °C in a buffer containing 0.5 M NH_4OAc and 1 mM ethylenediaminetetraacetic acid (EDTA). Gel fragments were removed using a 0.2 μm cellulose acetate filter, and the oligonucleotides were precipitated from the supernatant by ethanol precipitation at −70 °C. The oligonucleotides were then dried under vacuum, resuspended in nuclease-free water, and quantified by measuring the absorbance at 260 nm. ADAR editing substrates were allowed to self-anneal in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 100 mM NaCl by heating at 95 °C for 5 min and slow cooling to ≤ 30 °C.

3.4.4. *In vitro* deamination assay. Deamination assays were performed under single-turnover conditions. Samples containing 100 nM ADAR2 and 0 to 6 μM protein inhibitor (or 0 to 100 μM peptide inhibitor) were incubated in 15 mM Tris-HCl pH 7.5, 60 mM KCl, 3 mM $MgCl_2$, 1.5 mM EDTA, 0.003% (v/v) NP-40, 3% glycerol, 0.5 mM DTT, 1 μg/mL yeast tRNA, and 0.16 U/μL RNAse inhibitor at 30 °C for 10 min. The reaction was commenced by addition of 5 nM 5-HT_{2C} substrate and allowed to incubate at 30 °C for 10 min before quenching with 95 °C water,

vortexing, and heating at 95 °C for 5 min. Oligonucleotides were then purified using DNA Clean & Concentrator kit (Zymo) before conversion to cDNA using Access RT-PCR System (Promega). PCR amplicons were purified using 1% agarose gel and submitted for Sanger sequencing (Azenta). Sequencing peak heights at the edit site were quantified using 4Peaks (Nucleobytes). The sequencing traces also served to validate the sequences of the editing substrate. Data were plotted (% editing = [peak height G/(peak height A + peak height G)] × 100%) and analyzed using Microsoft Excel and GraphPad Prism. For IC_{50} calculations, data were plotted to the equation: $y = m1+(m2-m1)/[1+(x/m4)^{m3}]$, where y = % editing; x = log of RNA duplex concentration; m1 = minimum % editing; m2 = maximum % editing; m3 = slope factor or Hill slope; and m4 = log of IC_{50} value.

Deamination assays with ADAR1 p110 were performed as above except for the 15 min reaction time in the following buffer conditions: 15 mM Tris-HCl pH 7.5, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% (v/v) NP-40, 4% glycerol, 0.5 mM DTT, 1 μg/mL yeast tRNA, and 0.16 U/μL RNAse inhibitor.

3.4.5. Cellular editing of endogenous substrates in HEK293T cells. HEK293T cells were cultured in DMEM (Gibco), 10% FBS (Gibco) and 1% anti–anti (Gibco) at 37 ºC and 5% CO_2. At 70–90% confluency, 8×10^4 cells were seeded into a 96-well plate. After 24 h, cells were transfected with 100 ng pcDNA3.1 plasmid containing HA-tagged full length human ADAR2 WT and 0 to 1 μg pcDNA3.1 plasmid containing HA-tagged A2dx proteins. pcDNA3.1 empty vector was also transfected to keep the amount of plasmid transfected equal per sample. Transfection was carried out using Lipofectamine 2000. After incubation of transfection reagent and plasmid in Opti-MEM, the solution was added to designated well and incubated at 37 ºC and 5% CO_2 for 48 h. Total RNA was isolated using RNAqueous Total RNA Isolation kit (Invitrogen) and DNase

treated with RQ1 RNase-free DNase (Promega). Nested PCR was performed by first using Access RT-PCR kit (Promega) for 24 cycles followed by a second PCR using Phusion Hot Start DNA Polymerase (Thermo) for 19 cycles. The PCR products were purified by agarose gel and Qiagen Gel Extraction kit. The purified product was subjected to Sanger sequencing and sequence traces were analyzed by 4Peaks (Nucleobytes) to measure percent editing. See Table 3.1 for RT-PCR and nested PCR/sequencing primers.

3.5. Table of oligonucleotides and peptides

Table 3.1. Sequences of oligonucleotides used for work presented in Chapter 3.

Oligonucleotide	Sequence (5' to 3')
Mutagenesis primers	
E396A FWD	GCATTAAATGACTGCCATGCAGCTATAATATCTCGGAGATCC
E396A RVS	GGATCTCCGAGATATTATAGCTGCATGGCAGTCATTTAATGC
D503A FWD	GAGCATCCAAACGTGGGCTGGGGTGCTGCAAGGGG
D503A RVS	CCCCTTGCAGCACCCCAGCCCACGTTTGGATGCTC
D503K FWD	GAGCATCCAAACGTGGAAAGGGGTGCTGCAAGGGG
D503K RVS	CCCCTTGCAGCACCCCTTTCCACGTTTGGATGCTC
E912A FWD	CTGTTAACGACTGTCACGCTGCTATCATCTCTAGAAGAGGTTTC
E912A RVS	GAAACCTCTTCTAGAGATGATAGCAGCGTGACAGTCGTTAACAG
D1023A FWD	CTGACATCGTTCCAACTTGGGCTGGTATCAGATTGGGTGAAAG
D1023A RVS	CTTTCACCCAATCTGATACCAGCCCAAGTTGGAACGATGTCAG
D1023K FWD	CTGACATCGTTCCAACTTGGAAAGGTATCAGATTGGGTGAAAG
D1023K RVS	CTTTCACCCAATCTGATACCTTTCCAAGTTGGAACGATGTCAG
RT-PCR/sequencing primers for in vitro deamination substrates	
5-HT_{2C} FWD	TGGGTACGAATTCCCACTTACGTACAAGCTT
5-HT_{2C} RVS	GTAACATTTGCGTTTGATCGGGTTCT
RT-PCR and nested PCR/sequencing primers for cellular editing studies	
TMEM63B RT FWD	CCGCTGGCTCTTTGATAAGAAATTCTTGGCTGAGG
TMEM63B RT RVS	AGCCAGAAGAGGCAGAGGATGGGCG
TMEM63B Nest FWD	CAGCTATTCGGTTTGAGTGTGTGTTCC
TMEM63B Nest RVS	CGGCCACCACCTGGTTCACAGCCC
CYFIP2 RT FWD	TCCTGGCCAACCACAACAGGATCACCCAGTGTC
CYFIP2 RT RVS	TAGGTCGAAGAGCTCGCGATACTCCTCGTCTG
CYFIP2 Nest FWD	TCCACCAGCAACTTGAAGTGATCCCAGGCTATGA
CYFIP2 Nest RVS	ACTTCTGGCTGTCCAGCCCTGAGCCCG

Table 3.2. Sequences, purity (from HPLC purification), and masses (calculated and observed from mass spectroscopy) of peptides used for work presented in Chapter 3.

Peptide	Sequence (N- to C-terminal)	Purity (%)	Calculated Mass (g/mol)	Observed Mass (m/z)
8mer	TWDGVLQG	97.193	874.94	874.2
10mer	QTWDGVLQGE	95.296	1132.19	1131.8
12mer	IQTWDGVLQGER	97.552	1401.53	1401.2
14mer	SIQTWDGVLQGERL	95.390	1601.77	1601.4
12mer cyclic	ETWDGVLQGERK[a]	95.287	1399.52	1399.2

[a] cyclized at the 1 (E) and 12 (K) locations via side chain to side chain lactam bridge formation

3.6. References

(1) Bakail, M.; Ochsenbein, F. Targeting Protein-Protein Interactions, a Wide Open Field for Drug Design. *Comptes Rendus Chim.* **2016**, *19* (1–2), 19–27. https://doi.org/10.1016/j.crci.2015.12.004.

(2) Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J. Recent Advances in the Development of Protein–Protein Interactions Modulators: Mechanisms and Clinical Trials. *Signal Transduct. Target. Ther.* **2020**, *5* (1). https://doi.org/10.1038/s41392-020-00315-3.

(3) Wang, Z. Z.; Shi, X. X.; Huang, G. Y.; Hao, G. F.; Yang, G. F. Fragment-Based Drug Discovery Supports Drugging 'Undruggable' Protein–Protein Interactions. *Trends Biochem. Sci.* **2023**, *48* (6), 539–552. https://doi.org/10.1016/j.tibs.2023.01.008.

(4) Cheng, F.; Zhao, J.; Wang, Y.; Lu, W.; Liu, Z.; Zhou, Y.; Martin, W. R.; Wang, R.; Huang, J.; Hao, T.; Yue, H.; Ma, J.; Hou, Y.; Castrillon, J. A.; Fang, J.; Lathia, J. D.; Keri, R. A.; Lightstone, F. C.; Antman, E. M.; Rabadan, R.; Hill, D. E.; Eng, C.; Vidal, M.; Loscalzo, J. Comprehensive Characterization of Protein–Protein Interactions Perturbed by Disease Mutations. *Nat. Genet.* **2021**, *53* (3), 342–353. https://doi.org/10.1038/s41588-020-00774-y.

(5) Carvajal, L. A.; Ben Neriah, D.; Senecal, A.; Benard, L.; Thiruthuvanathan, V.; Yatsenko, T.; Narayanagari, S. R.; Wheat, J. C.; Todorova, T. I.; Mitchell, K.; Kenworthy, C.; Guerlavais, V.; Annis, D. A.; Bartholdy, B.; Will, B.; Anampa, J. D.; Mantzaris, I.; Aivado, M.; Singer, R. H.; Coleman, R. A.; Verma, A.; Steidl, U. Dual Inhibition of MDMX and MDM2 as a Therapeutic Strategy in Leukemia. *Sci. Transl. Med.* **2018**, *10* (436). https://doi.org/10.1126/scitranslmed.aao3003.

(6) Dalal, K.; Ban, F.; Li, H.; Morin, H.; Roshan-Moniri, M.; Tam, K. J.; Shepherd, A.; Sharma, A.; Peacock, J.; Carlson, M. L.; LeBlanc, E.; Perez, C.; Duong, F.; Ong, C. J.; Rennie, P. S.; Cherkasov, A. Selectively Targeting the Dimerization Interface of Human Androgen Receptor with Small-Molecules to Treat Castration-Resistant Prostate Cancer. *Cancer Lett.* **2018**, *437* (August), 35–43. https://doi.org/10.1016/j.canlet.2018.08.016.

(7) Saraç, H.; Cherkasov, A.; Lack, N. A. Development of Novel Androgen Receptor Inhibitors to Overcome Castrate-Resistant Prostate Cancer. *Biol. Mech. Adv. Approaches to Overcoming Cancer Drug Resist.* **2020**, 23–46. https://doi.org/10.1016/B978-0-12-821310-0.00007-3.

(8) Simonetta, K. R.; Taygerly, J.; Boyle, K.; Basham, S. E.; Padovani, C.; Lou, Y.; Cummins, T. J.; Yung, S. L.; von Soly, S. K.; Kayser, F.; Kuriyan, J.; Rape, M.; Cardozo,

M.; Gallop, M. A.; Bence, N. F.; Barsanti, P. A.; Saha, A. Prospective Discovery of Small Molecule Enhancers of an E3 Ligase-Substrate Interaction. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-09358-9.

(9) Weagel, E. G.; Foulks, J. M.; Siddiqui, A.; Warner, S. L. Molecular Glues: Enhanced Protein-Protein Interactions and Cell Proteome Editing. *Med. Chem. Res.* **2022**, *31* (7), 1068–1087. https://doi.org/10.1007/s00044-022-02882-2.

(10) Cho, D. S. C.; Yang, W.; Lee, J. T.; Shiekhattar, R.; Murray, J. M.; Nishikura, K. Requirement of Dimerization for RNA Editing Activity of Adenosine Deaminases Acting on RNA. *J. Biol. Chem.* **2003**, *278* (19), 17093–17102. https://doi.org/10.1074/jbc.M213127200.

(11) Macbeth, M. R.; Schubert, H. L.; VanDemark, A. F.; Lingam, A. T.; Hill, C. P.; Bass, B. L. Structural Biology: Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science (80-.).* **2005**, *309* (5740), 1534–1539. https://doi.org/10.1126/science.1113150.

(12) Valente, L.; Nishikura, K. RNA Binding-Independent Dimerization of Adenosine Deaminases Acting on RNA and Dominant Negative Effects of Nonfunctional Subunits on Dimer Functions. *J. Biol. Chem.* **2007**, *282* (22), 16054–16061. https://doi.org/10.1074/jbc.M611392200.

(13) Singh, M.; Kesterson, R. A.; Jacobs, M. M.; Joers, J. M.; Gore, J. C.; Emeson, R. B. Hyperphagia-Mediated Obesity in Transgenic Mice Misexpressing the RNA-Editing Enzyme ADAR2. *J. Biol. Chem.* **2007**, *282* (31), 22448–22459. https://doi.org/10.1074/jbc.M700265200.

(14) Thuy-Boun, A. S.; Thomas, J. M.; Grajo, H. L.; Palumbo, C. M.; Park, S.; 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), 7958–7952. https://doi.org/10.1093/nar/gkaa532.

(15) Zuker, M. Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res.* **2003**, *31* (13), 3406–3415. https://doi.org/10.1093/nar/gkg595.

(16) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 1–9. https://doi.org/10.1038/ncomms1324.

(17) Hill, T. A.; Shepherd, N. E.; Diness, F.; Fairlie, D. P. Constraining Cyclic Peptides to Mimic Protein Structure Motifs. *Angew. Chemie - Int. Ed.* **2014**, *53* (48), 13020–13041. https://doi.org/10.1002/anie.201401058.

(18) Huang, H.; Damjanovic, J.; Miao, J.; Lin, Y. S. Cyclic Peptides: Backbone Rigidification and Capability of Mimicking Motifs at Protein-Protein Interfaces. *Phys. Chem. Chem. Phys.* **2021**, *23* (1), 607–616. https://doi.org/10.1039/d0cp04633g.

(19) Desterro, J. M. P.; Keegan, L. P.; Lafarga, M.; Berciano, M. T.; O'Connell, M.; Carmo-Fonseca, M. Dynamic Association of RNA-Editing Enzymes with the Nucleolus. *J. Cell Sci.* **2003**, *116* (9), 1805–1818. https://doi.org/10.1242/jcs.00371.

(20) Sansam, C. L.; Wells, K. S.; Emeson, R. B. Modulation of RNA Editing by Functional Nucleolar Sequestration of ADAR2. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (SUPPL. 2), 14018–14023. https://doi.org/10.1073/pnas.2336131100.

(21) Dang, C. V.; Lee, W. M. F. Identification of the Human C- Myc Protein Nuclear Translocation Signal . *Mol. Cell. Biol.* **1988**, *8* (10), 4048–4054. https://doi.org/10.1128/mcb.8.10.4048-4054.1988.

(22) Malim, M. H.; McCarn, D. F.; Tiley, L. S.; Cullen, B. R. Mutational Definition of the

Human Immunodeficiency Virus Type 1 Rev Activation Domain. *J. Virol.* **1991**, *65* (8), 4248–4254. https://doi.org/10.1128/jvi.65.8.4248-4254.1991.

(23) Meyer, B. E.; Meinkoth, J. L.; Malim, M. H. Nuclear Transport of Human Immunodeficiency Virus Type 1, Visna Virus, and Equine Infectious Anemia Virus Rev Proteins: Identification of a Family of Transferable Nuclear Export Signals. *J. Virol.* **1996**, *70* (4), 2350–2359. https://doi.org/10.1128/jvi.70.4.2350-2359.1996.

(24) 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), 1–13. https://doi.org/10.1038/s41467-020-18862-2.

(25) Sawyer, N.; Watkins, A. M.; Arora, P. S. Protein Domain Mimics as Modulators of Protein-Protein Interactions. *Acc. Chem. Res.* **2017**, *50* (6), 1313–1322. https://doi.org/10.1021/acs.accounts.7b00130.

(26) Vu, Q. N.; Young, R.; Sudhakar, H. K.; Gao, T.; Huang, T.; Tan, Y. S.; Lau, Y. H. Cyclisation Strategies for Stabilising Peptides with Irregular Conformations. *RSC Med. Chem.* **2021**, *12* (6), 887–901. https://doi.org/10.1039/d1md00098e.

(27) 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* (07), 319–331. https://doi.org/10.1016/S0076-6879(07)24015-7.

Chapter 4

Probing ADAR1 substrate recognition and selective inhibition of ADAR1 using 8-azanebularine-modified RNA duplexes

Most of the work presented in this chapter was published in Biochemistry in March 2023[1]*. The ADAR1 crystallography studies described at the end of the chapter was a collaborative effort with Dr. Alexander Thuy-Boun.*

4.1. Introduction

In Chapter 1, the role of ADARs in human diseases and how these enzymes are currently being utilized for directed RNA editing applications were discussed. In particular, studies have shown that ADAR1 deletion results in increased tumor sensitivity to immunotherapy and in eventual cell death for some cancers that are characterized by elevated levels of interferon-stimulated genes[2,3]. These findings strongly suggest ADAR1 inhibitors as promising candidates for anticancer drugs. However, there are still no FDA-approved inhibitors of ADARs to date. Multiple reports have also indicated that ADAR1 is more ubiquitously expressed than the other members of the human ADAR family[4–7], making it a highly valuable tool for site-directed RNA editing using endogenous enzymes.

Whether as a therapeutic or as a therapeutic target, understanding ADAR's structure, reaction chemistry, substrate recognition, and selectivity is important to control its activity. Several crystal structures of ADAR2 alone or bound to duplex RNA have been made available[8–13]; however, our current molecular knowledge of ADAR1 and its interaction with an RNA substrate is lagging. This clearly limits the structure-guided design of ADAR1 inhibitors and ADAR1-specific guide oligonucleotides.

Much of the Beal lab's success in characterizing the ADAR2 structure, reaction mechanism, and substrate selectivity can be attributed to the use of 8-azaN[9,10,14–17] as described in

Chapter 1. The use of dsRNAs containing 8-azaN at the reactive site allowed for the mechanistic trapping of ADAR2-RNA complexes for binding and structural studies[9–14]. However, whether this nucleoside analog alone or integrated into an RNA structure can likewise be used for ADAR1 investigations is still unexplored. In this chapter, we sought to employ a panel of short RNA duplexes bearing 8-azaN to probe the ADAR1 catalytic domain substrate recognition via *in vitro* binding and structure-activity relationship (SAR) studies and X-ray crystallography. In an effort to develop ADAR1-specific inhibitors, we also attempted to utilize these 8-azaN-modified duplexes as substrate decoys for competitive inhibition of ADAR1 *in vitro* and in cells.

4.2. Results

4.2.1. A short RNA duplex bearing 8-azanebularine binds tightly to ADAR1. From a previous screen for human ADAR substrates in yeast, former lab member Dr. Yuru Wang identified a number of RNA substrates that are preferentially edited by the human ADAR1 deaminase (ADAR1-D) over the human ADAR2-D[18]. One of these substrates, *S. cerevisiae* HER1 mRNA, was relatively small in size and efficiently edited by ADAR1-D, deeming it practical and suitable for further investigations of RNA binding by the ADAR1 catalytic domain using chemically modified duplexes. Dr. SeHee Park then designed a 16 bp intermolecular RNA duplex (H_{16}) from a short hairpin stem containing the edit site in the yeast HER1 RNA (Fig. 4.1) for ADAR1-D binding studies (Fig. 4.2). The Beal lab has previously shown that the E488Q mutation, which increases ADAR2's catalytic rate[19], also enhances binding of ADAR2-D to 8-azaN-containing RNA duplexes and leads to substantially less smearing in gel shift experiments[14]. This could be due to slower dissociation rates of the complexes formed between the E488Q protein and the 8-azaN duplexes compared to the WT protein. Nevertheless, a similar preference for the 8-azaN-modified duplexes over the unmodified duplexes was observed for both ADAR2-D WT and

E488Q proteins[14]. In collaboration with Dr. Park, we then proceeded to use the corresponding hyperactive ADAR1 deaminase protein (ADAR1-D E1008Q) for gel shift assays with H_{16} 8-azaN or A (Fig. 4.2). Indeed, we observed a significantly tighter binding with the 8-azaN containing duplex compared to the duplex with adenosine at the editing site with $K_D = 21 \pm 11$ nM for H_{16} 8-azaN and $K_D > 300$ nM for H_{16} A at the conditions tested. These results clearly demonstrate that 8-azaN can be used to generate high-affinity RNA ligands for the ADAR1 deaminase domain and enable studies of ADAR1-RNA interactions.

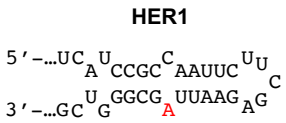


Figure 4.1. Predicted secondary structure around the yeast HER1 editing site. Structure predicted by mFold[20]. Edited A is shown in red.

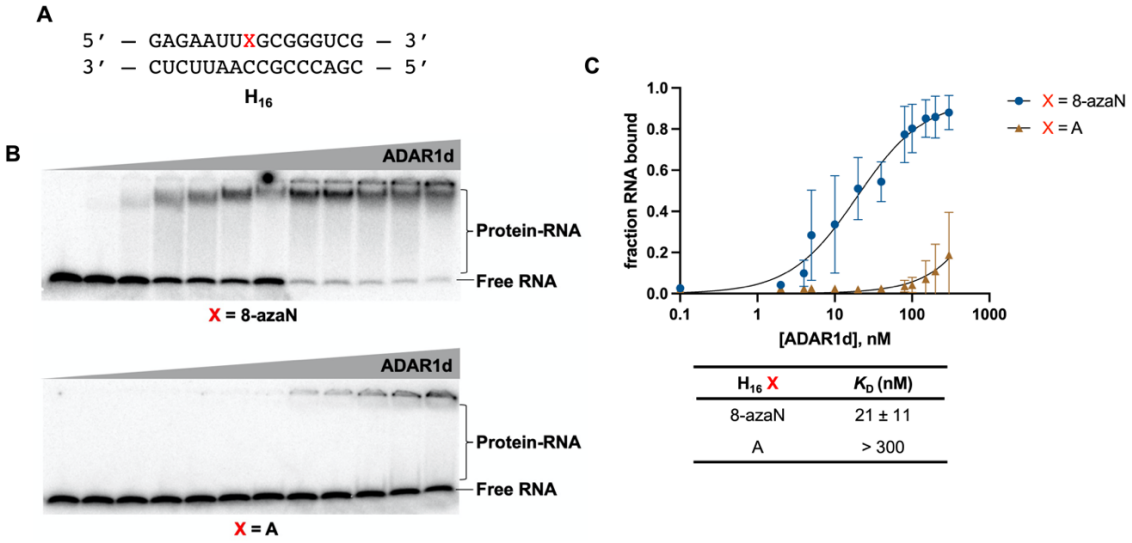


Figure 4.2. A short RNA duplex bearing 8-azaN binds tightly to ADAR1. (**A**) Sequence of the 16 bp duplex derived from yeast HER1 RNA. X is A in native target sequence but is varied (8-azaN, A, or G) in the succeeding experiments. (**B**) Gel shifts of ADAR1-D with H_{16} 8-azaN or A at 0, 2, 4, 8, 10, 20, 40, 80, 100, 150, 200, and 300 nM ADAR1-D E1008Q and 5 nM H_{16} RNA. (**C**) Fitted plots of fraction H_{16} RNA bound *vs* protein concentration. Data were plotted to the equation: $y = A \times [x/(K_D + x)]$ where y is fraction H_{16} RNA bound, x is [ADAR1-D]; A is binding endpoint; and K_D is dissociation constant. Error bars represent standard deviation from $n = 3$ technical replicates.

Our findings from the gel shift experiments also led us to consider the use of 8-azaN-modified duplexes as substrate decoys to potentially suppress ADAR1 editing of an RNA target. For this experiment, the full-length human ADAR1 p110 was used to gain more biologically relevant insights into ADAR1 inhibition by these modified duplexes. *In vitro* deamination reactions with ADAR1 p110 and a fragment of the human 5-HT_{2C} pre-mRNA were first set up, to which 0 to 3 μM of H_{16} 8-azaN or A/G control duplexes were then added (Fig. 4.2). Indeed, the results show that H_{16} 8-azaN inhibited 5-HT_{2C} editing by ADAR1 p110 at the B-site (preferred edit site of ADAR1 in 5-HT_{2C}[21], Fig. 4.3) in a concentration- and 8-azaN-dependent manner with an estimated IC_{50} of 13 ± 2 nM (Fig. 4.4a, Table 4.1). Titration of the H_{16} 8-azaN duplex also inhibited editing by ADAR1 p110 at a preferred edit site[22] (Site 1, Fig. 4.3) in the human NEIL1 pre-mRNA substrate with an estimated IC_{50} of 8.9 ± 0.8 nM (Fig. 4.4a, Table 4.1). This assay was also tested on another full-length human ADAR1 isoform, ADAR1 p150, and a similar concentration and 8-azaN-dependent inhibition of 5-HT_{2C} B-site editing was observed (IC_{50} = 28 ± 3 nM) (Fig. 4.4b, Table 1). It is important to note that the A- or G-containing H_{16} controls also brought about a small concentration-dependent decrease in 5-HT_{2C} and NEIL1 editing by ADAR1 p110 at greater than or equal to 200-fold excess to editing substrate ([H_{16} A/G] ≥ 1 μM). However, this slight inhibition by the unmodified duplex is not seen in ADAR1 p150 for H_{16} A. We hypothesize that the additional Z-DNA/RNA binding domain (Zα) in ADAR1 p150 contributes to selective binding to editing substrate RNA over the H_{16} duplex. This can also explain the two-fold higher IC_{50} measured for H_{16} 8-azaN in the ADAR1 p150/5-HT_{2C} *vs* the ADAR1 p110/5-HT_{2C} reaction.

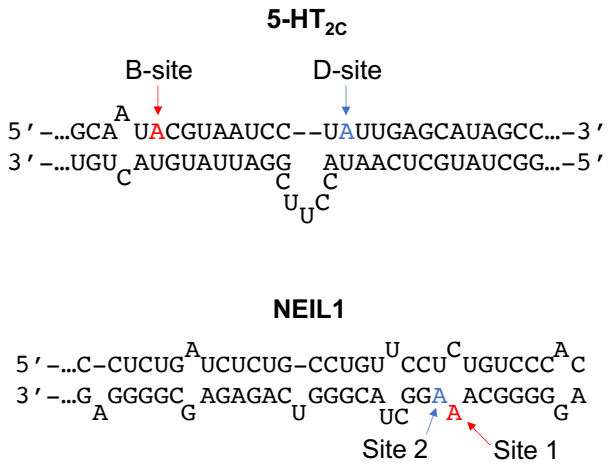


Figure 4.3. Predicted secondary structures around (top) 5-HT$_{2C}$ B and D and (bottom) NEIL1 Site 1 and 2 editing sites. Structure predicted by mFold[20].

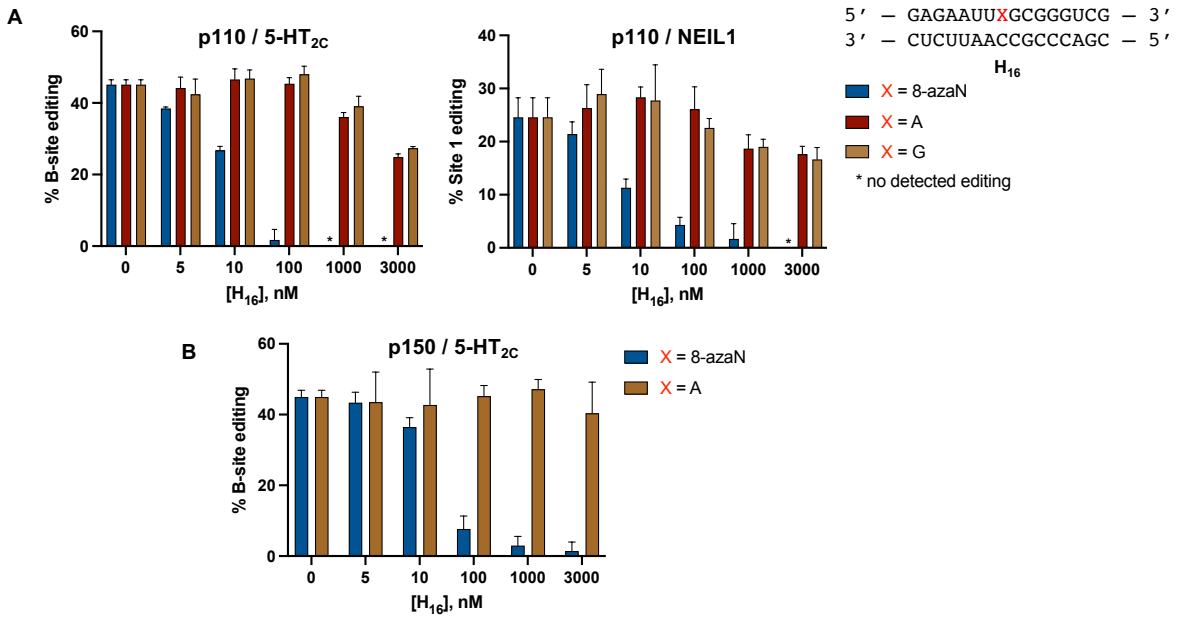


Figure 4.4. 8-azaN-containing duplex inhibits 5-HT$_{2C}$ and NEIL1 editing by ADAR1. (**A**) Inhibition of ADAR1 p110 by H$_{16}$ 8-azaN tested on two different editing substrates, 5-HT$_{2C}$ (left) and NEIL1 (right). (**B**) Inhibition of ADAR1 p150 by H$_{16}$ 8-azaN tested on 5-HT$_{2C}$. *In vitro* deaminations were performed at the following conditions: 0–3 μM H$_{16}$ 8-azaN or A/G control, 100 nM ADAR1 p110 or p150, 5 nM substrate, 15 min (for 5-HT$_{2C}$) or 30 min (for NEIL1), at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

Table 4.1. Estimated IC_{50} values from inhibition experiments with 8-azaN duplexes.

8-azaN Duplex	Enzyme	Editing Substrate	IC_{50} (nM)[a]
H_{16}	ADAR1 p110	5-HT_{2C}	13 ± 2
H_{16}	ADAR1 p110	NEIL1	8.9 ± 0.8
H_{16}	ADAR1 p150	5-HT_{2C}	28 ± 3
5'-3' swap	ADAR1 p110	5-HT_{2C}	15 ± 1
GC-rich	ADAR1 p110	5-HT_{2C}	18 ± 8
AU-rich	ADAR1 p110	5-HT_{2C}	74 ± 3
2'-OMe	ADAR1 p110	5-HT_{2C}	11.4 ± 0.7
H_{14} A[b]	ADAR1 p110	5-HT_{2C}	18 ± 4
H_{14} B[b]	ADAR1 p110	5-HT_{2C}	> 3000
H_{12} A[b]	ADAR1 p110	5-HT_{2C}	> 3000
H_{12} B[b]	ADAR1 p110	5-HT_{2C}	> 3000

[a] Data from inhibition experiments were plotted to the equation: $y = m1+(m2-m1)/[1+(x/m4)^{m3}]$, where y = % editing; x = log of RNA duplex concentration; m1 = basal response; m2 = maximal response; m3 = slope factor or Hill slope; and m4 = log of IC_{50} value. Values reported are the average of three independent measurements ± standard deviation. [b] Duplex sequences defined in Figure 4.7

4.2.2. Secondary structure requirement for ADAR1 engagement. To test the importance of the duplex RNA structure for ADAR1 binding, the inhibition experiment described above was repeated with 8-azaN as a free nucleoside and with the 8-azaN-containing single strand component of the H_{16} 8-azaN duplex (Fig. 4.5). We observed no reduction in 5-HT_{2C} B-site editing by ADAR1 p110 even at the highest concentration tested for 8-azaN nucleoside (1 mM) and 8-azaN ssRNA (3 μM). These results clearly demonstrate the need for 8-azaN to be in a duplex context to effect ADAR1 binding and inhibition.

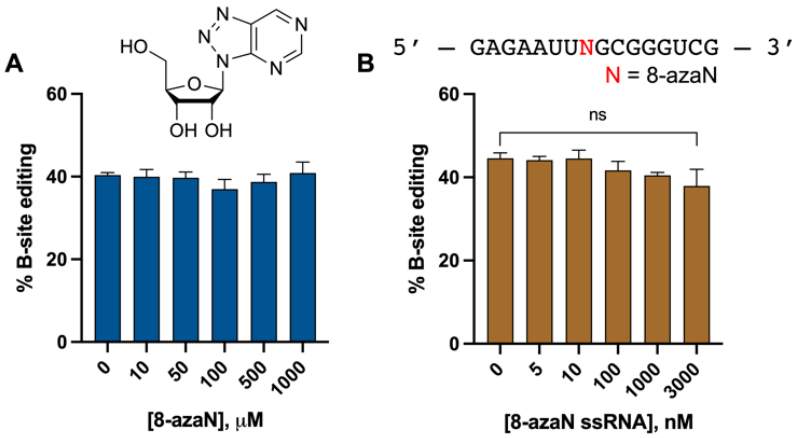


Figure 4.5. 8-azaN as a free nucleoside (**A**) and an 8-azaN-modified ssRNA (**B**) do not inhibit ADAR1. *In vitro* deaminations were performed at the following conditions: 0−1 mM 8-azaN or 0−3 μM 8-azaN ssRNA, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates. A two-tailed Welch's t-test was conducted between indicated groups, ns, not significant.

ADARs are known to have a preference for editing adenosines in an A:C mismatch and within a nearest neighbor 5′-U and 3′-G sequence context[21,23,24]. However, how the flanking base sequences beyond this 5′-UAG-3′ triplet play a role in ADAR selectivity and recognition is still unclear. Interestingly, the *S. cerevisiae* GSY1 mRNA, another human ADAR1-D preferred substrate identified from the yeast screen, shares a common secondary structure surrounding the edit site with HER1 and are both comprised of a short A-U rich (5 bp) hairpin stem on the 5′ side of the edit site[18]. To test whether this A-U rich region is crucial for binding, the sequences present on the 5′ and 3′ sides of the H_{16} 8-azaN duplex (5′-3′ swap, Fig. 4.6a) were swapped and the resulting duplex was titrated into an ADAR1 p110/5-HT_{2C} reaction (Fig. 4.6b). However, the sequence swapping did not significantly affect binding of the 8-azaN duplex (IC_{50} = 15 ± 1 nM), suggesting that the short stretch of A-U base pairs 5′ to the edit site is not necessarily a requirement for ADAR1 recognition.

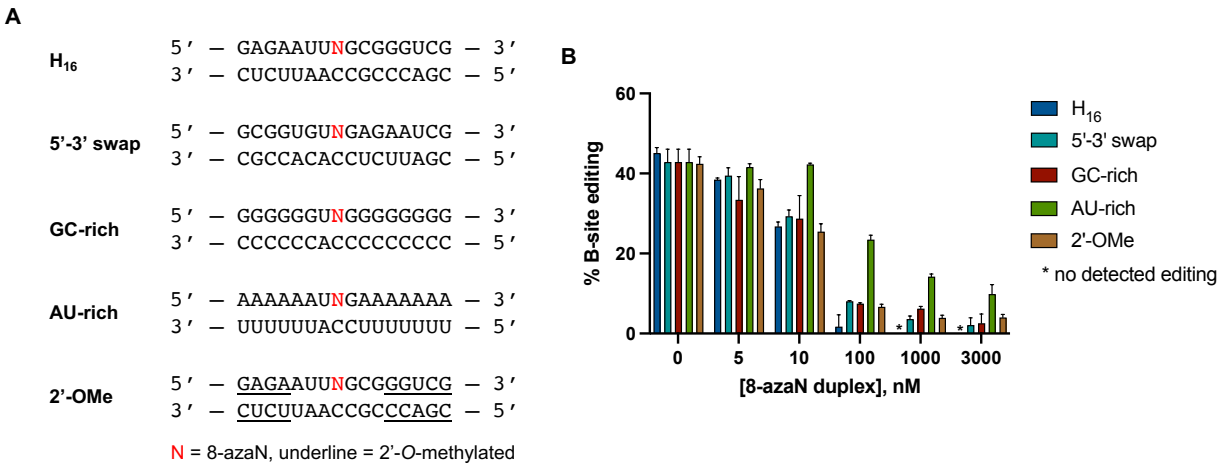


Figure 4.6. Effect of flanking sequence, G-C content, and 2′-*O*-methylation on ADAR1 binding. (**A**) Sequences of 8-azaN duplexes tested. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0−3 μM 8-azaN duplex, 100 nM ADAR1 p110, 5 nM 5-HT$_{2C}$, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

The effect of the 8-azaN duplex's G-C content on ADAR1 binding and subsequent inhibition of target editing was also determined. We imagined that increasing duplex stability by increasing G-C bp composition may increase 8-azaN duplex binding and potency. Having an 8-azaN duplex with low G-C content might then have the reverse effect. Therefore, two 8-azaN duplexes with either a higher G-C content (GC-rich) or a higher A-U content (AU-rich) were designed and tested for inhibition of ADAR1 p110/5-HT$_{2C}$ reaction (Fig. 4.6). Interestingly, the GC-rich duplex inhibited the 5-HT$_{2C}$ substrate deamination with similar potency as the original H$_{16}$ 8-azaN duplex (IC$_{50}$ = 18 ± 8 nM, Table 4.1). The AU-rich duplex, on the other hand, suffered a six-fold decrease in binding affinity (IC$_{50}$ = 74 ± 3 nM, Table 1). We rationalize that the additional duplex stability brought about by higher G-C content might not exhibit a substantial effect on ADAR1 binding under the conditions used for the deamination assay. The experimental T_M values for H$_{16}$, GC-rich, and AU-rich duplexes (62.5 ± 0.1, 88 ± 1, and 25.5 ± 0.1 °C, respectively; Table 4.2) suggest that the H$_{16}$ and GC-rich duplexes should predominantly exist in

duplex form at 30 °C, while some of the AU-rich duplex species would be single-stranded. It is also possible that the expected positive effect of the GC-rich duplex is masked by hindrance of necessary local melting in the duplex for ADAR recognition and reaction (*e.g.*, to facilitate base-flipping around the editing site).

Table 4.2. Experimental thermal melting temperatures (T_M) for 8-azaN duplexes.

8-azaN Duplex	T_M (°C)[a]
H_{16}	62.5 ± 0.1
GC-rich	88 ± 1
AU-rich	25.5 ± 0.1
H_{14} A	54.6 ± 0.9
H_{14} B	55.1 ± 0.1
X_A	58.0 ± 0.7
X_{A+}	81.3 ± 0.1
X_B	20.9 ± 0.5
X_{B+}	58.4 ± 0.0

[a] T_M values were measured at the following conditions: 1 μM duplex, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl, and 1.25 μM EvaGreen. Values reported are the average of three independent measurements ± standard deviation.

Finally, the effect of 2′-*O*-methyl modification at nucleotides distal to the editing site was also studied (Figure 4.6). This modification at the ribose backbone has been shown to increase resistance of oligonucleotides to nucleases[25–27]. The results clearly indicate that 2′-*O*-methylation at the specified positions does not affect ADAR1 binding (Fig. 4.6b). This is further validated by the comparable IC_{50} values between the parent H_{16} duplex and 2′-OMe duplex (IC_{50} = 11.4 ± 0.7 nM, Table 4.1).

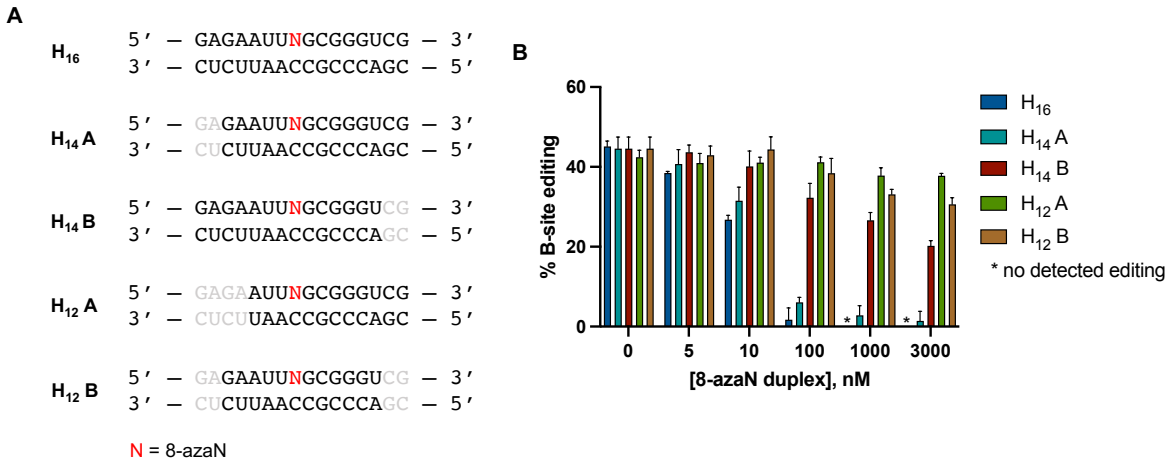


Figure 4.7. Minimum duplex length for ADAR1 binding. (**A**) 8-azaN-modified duplexes of different lengths tested for inhibition. Bases omitted from the original H_{16} 8-azaN duplex sequence are in gray. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0−3 μM 8-azaN duplex, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

4.2.3. Minimum duplex length for ADAR1 binding. The use of short oligonucleotides for biochemical and biophysical studies of nucleic acid binding or modifying enzymes such as ADARs is advantageous for ease of synthesis and incorporation of nucleoside analogs at specific positions in the strand[14,28]. Therefore, the minimal 8-azaN duplex length to engage ADAR1 in order to effect editing inhibition was determined. A panel of shortened 8-azaN duplexes derived from the original H_{16} duplex was designed and tested for inhibition of the ADAR1 p110/5-HT_{2C} reaction (Fig. 4.7). Interestingly, we observed that removing two base pairs from the 5′ end of the H_{16} duplex (5′ and 3′ defined by the 8-azaN-containing strand) did not significantly affect potency of inhibition (IC_{50}, H_{14} A = 18 ± 4 nM, Table 4.1). However, removing two base pairs from the 3′ end of H_{16} resulted in a remarkably weaker inhibitor (IC_{50}, H_{14} B > 3000 nM, Table 4.1). Two base pairs from the 5′ end of H_{14} A were further removed to generate H_{12} A, which resulted in a substantial loss of potency (IC_{50} > 3000 nM, Table 4.1). These findings suggest a need for at least 5 bp 5′ and 8 bp 3′ to editing site for an efficient ADAR1-RNA interaction. Finally, removing two base pairs from the

3′ end of H_{14} A to generate H_{12} B also led to a notable loss in binding and potency (IC_{50} > 3000 nM, Table 4.1). This suggests that the 2 bp at the 3′ end of the parent H_{16} duplex harbors an important RNA contact site for the protein, consistent with our observation with H_{14} B. It is also important to note that the difference in potency of the two H_{14} duplexes is not due to dissimilarity in duplex stability as shown by the comparable experimental T_M values for H_{14} A and H_{14} B (54.6 ± 0.9 and 55.1 ± 0.1 °C, respectively; Table 4.2).

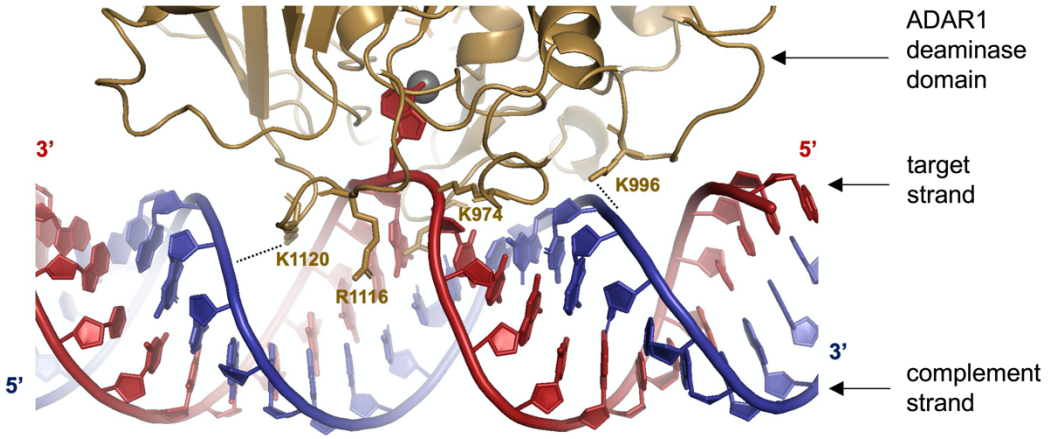


Figure 4.8. Predicted contacts between ADAR1 deaminase (gold) and an RNA duplex (red = target strand; blue = complement strand) based on a previously reported Rosetta homology model of ADAR1 catalytic domain[29]. Active site Zn metal (gray sphere).

Previously, Beal lab alumna Dr. Erin Doherty generated a Rosetta homology model of the ADAR1 catalytic domain bound to a dsRNA substrate[29]. This model suggested that the ADAR1 deaminase domain contacts the RNA duplex mostly at positions proximal to edit site through both target and complement strands (Fig. 4.8). However, at distal positions, the ADAR1 catalytic domain mostly interacts with the duplex through the complement strand. The 8-azaN-containing strand components of H_{12} A and H_{12} B were then hybridized to the H_{14} A complement strand to generate two 12 bp duplexes with 2 nt overhang (H_{12} A overhang and H_{12} B overhang) (Fig. 4.9a). We imagined that the 2 nt overhangs in the complement strand will provide the necessary contact

for better ADAR1 engagement than the corresponding blunt-ended 12 bp duplexes. Indeed, titration of H_{12} A overhang or H_{12} B overhang resulted in better inhibition of ADAR1 p110/5-HT_{2C} reaction compared to H_{12} A or H_{12} B (Fig. 4.9b). However, both duplexes with overhangs are not as potent as H_{14} A, suggesting that a duplex structure is still needed for more effective contact, especially on the 5′ end of the duplex.

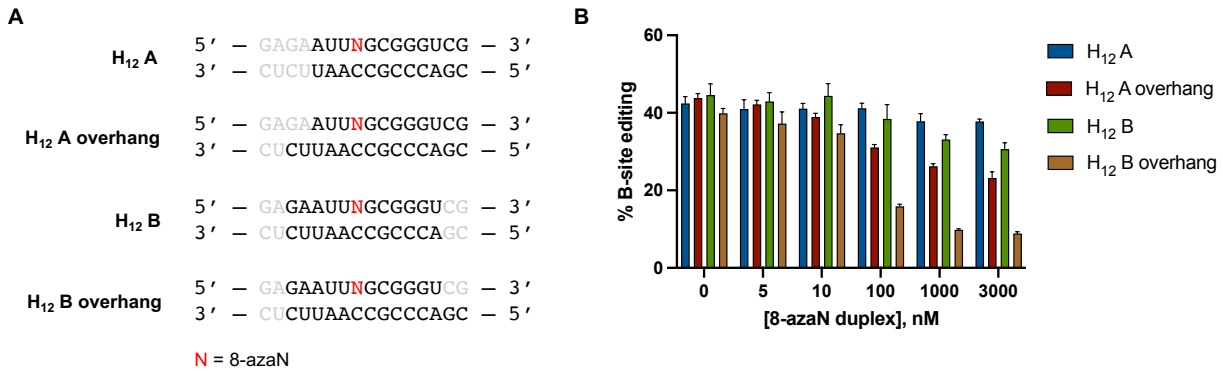


Figure 4.9. (**A**) 8-azaN-modified duplexes with overhangs tested for ADAR1 inhibition. Bases omitted from the original H_{16} 8-azaN duplex sequence are in gray. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0−3 μM 8-azaN duplex, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

The effect of stabilization of these duplexes was also tested by generating two covalently cross-linked forms with overhangs (X_{A+} and X_{B+}) and testing them for inhibition of ADAR1 p110/5-HT_{2C} reaction (Fig. 4.10). These cross-linked duplexes were synthesized by current lab member Victorio Jauregui Matos. We hypothesized that creating an intramolecular duplex via a covalent cross-link would increase duplex stability and subsequent ADAR1 binding, especially for the duplex with shorter 8-azaN-bearing strand (X_{B+}). We expect that the covalent cross-link may not meaningfully cause irregularities in the duplex structure and affect ADAR1 binding as its estimated length from C1′ to C1′ (~13.5 Å, Fig. 4.11) is very similar to that of a canonical base pair in duplex RNA (~11 Å). While cross-linking indeed drastically increased the experimental

T_M values for both duplexes (X_A to X_{A+} = 58.0 ± 0.7 to 81.3 ± 0.1 °C; X_B to X_{B+} = 20.9 ± 0.5 to 58.4 ± 0.0 °C; Table 4.2), we only found a slight difference in potency of inhibition between X_A and X_{A+} at duplex concentrations ≥ 100 nM and no significant difference in potency of inhibition between X_B and X_{B+} at the conditions tested (Fig. 4.10b). The minimal increase (X_A to X_{A+}) to no increase (X_B to X_{B+}) in potency that we observed upon crosslinking could also be explained by a decrease in ADAR1 binding due to the reduction in the duplexes' conformational flexibility around the site of chemical cross-linking. Nevertheless, these observations agree with the results from the intermolecular duplexes with overhangs in Fig. 4.9; that while modeling suggests that the ADAR1 deaminase domain contacts the RNA duplex at distal positions to the edit site via the complement strand's phosphodiester backbone, a duplex structure at the terminal ends of the duplex is still required for efficient interaction.

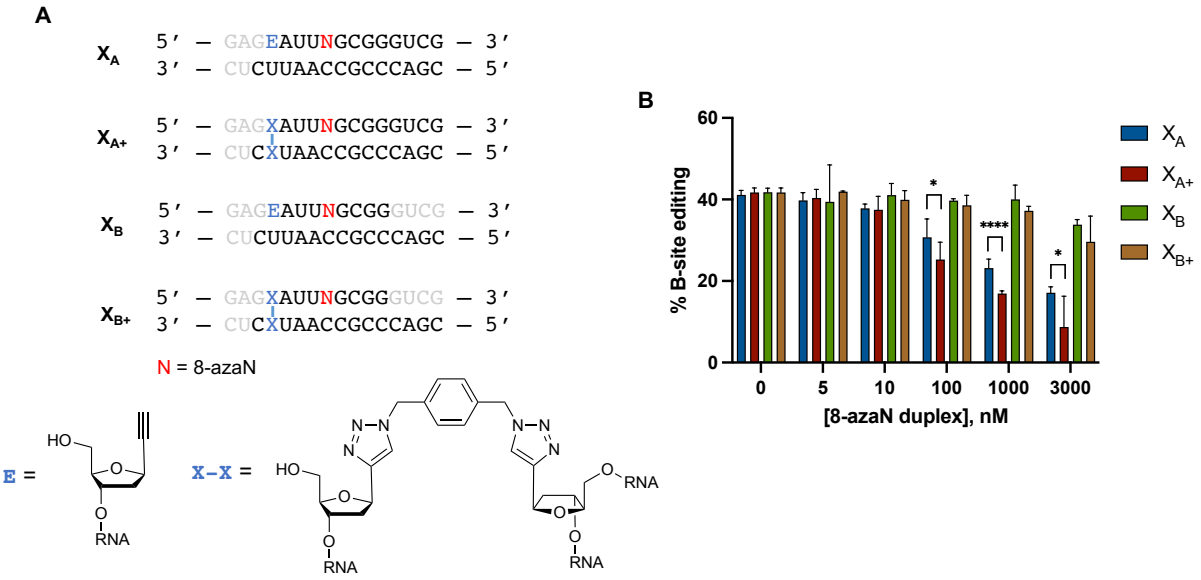


Figure 4.10. (**A**) Cross-linked RNA duplexes tested for inhibition. Bases omitted from the original H16 8-azaN duplex sequence are in gray. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0−3 µM 8-azaN duplex, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates. A two-tailed Welch's t-test was conducted between indicated groups, * $p < 0.05$, **** $p < 0.0001$.

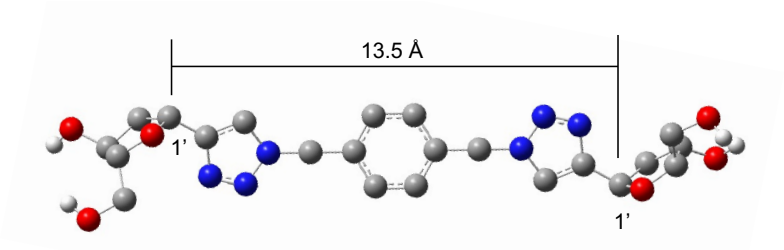


Figure 4.11. Estimated length of the covalent cross-link from C1' to C1'. Molecular model generated using Gaussian 16. C (grey sphere, N (blue sphere), O (red sphere), H (white sphere). Analysis courtesy of current lab member Randall Ouye.

4.2.4. Selective inhibition of ADAR1. Earlier, two human ADAR1-D preferred substrates identified from a screen of human ADAR1-D substrates in yeast were described (HER1 and GSY1)[18]. These two mRNAs have relatively short (5 bp) stem 5′ to the editing site, which from previous observations is not long enough for efficient binding to the deaminase domain of human ADAR2[18]. This suggested that the H_{16} 8-azaN duplex may be an ADAR1-specific inhibitor and not inhibit the reaction of ADAR2. This is indeed the case as shown in Fig. 4.12. We found that concentrations of the H_{16} 8-azaN that clearly inhibit ADAR1 do not affect editing by human ADAR2 of the 5-HT_{2C} D-site (preferred edit site of ADAR2 in 5-HT_{2C}[21], Fig. 3.3) or NEIL1 Site 2 (preferred edit site of ADAR2 in NEIL1[22], Fig. 3.3). However, we did observe a reduction of 5-HT_{2C} D-site editing at $[H_{16}] = 3$ μM and of NEIL1 Site 2 editing at $[H_{16}] \geq 0.1$ μM. This decrease in editing, however, does not appear to be 8-azaN-dependent and may be attributed to nonspecific binding of the enzyme to the short RNA duplex, hindering access to the editing substrate. Overall, these findings reveal a new strategy for designing effective human ADAR family-specific modulators of ADAR1.

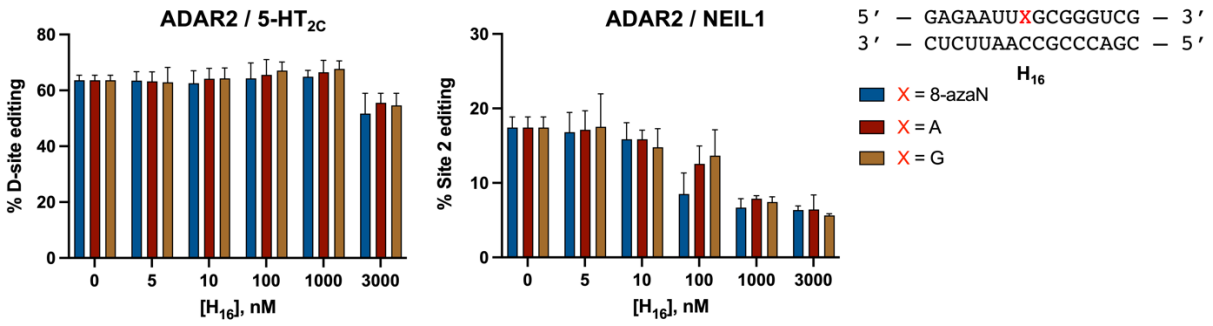


Figure 4.12. ADAR1-selective inhibition. H_{16} 8-azaN duplex does not inhibit 5-HT_{2C} (left) nor NEIL1 (right) editing by ADAR2. *In vitro* deaminations were performed at the following conditions: 0–3 μM H_{16} 8-azaN or A/G control, 100 nM ADAR2, 5 nM substrate, 10 min (for 5-HT_{2C}) or 30 min (for NEIL1), at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

4.2.5. Crystallography studies with ADAR1-D-RNA complexes. Earlier attempts to obtain high resolution X-ray crystal structures of ADAR1-D were conducted by Dr. SeHee Park. Previous work by Dr. Park led to the determination of optimal purification conditions to obtain sufficient amounts of pure ADAR1-D protein of high concentration suitable for crystallization trials. Dr. Park was able to set up multiple crystallization screens using different buffer and RNA duplex conditions. A few of the conditions yielded crystals that were suitable for mounting and X-ray diffraction; unfortunately, most crystals were of dsRNA alone and the highest resolution data obtained for an RNA-bound ADAR1-D crystal was only up to 5 Å. Nonetheless, these results motivated us to continue Dr. Park's efforts in solving high resolution X-ray crystal structures of ADAR1-D in complex with dsRNA. In collaboration with Dr. Alexander Thuy-Boun, we conducted screens using a larger number of buffer conditions, different RNA duplex sequences and lengths, and a broader range of protein to RNA ratios. Fig. 4.13 lists the 8-azaN-modified duplexes that were used for the crystallography screens under the conditions specified in Table 4.3. The RNA duplex sequences were derived from the yeast HER1 (H_{17}, H_{17} B, H_{17} C, H_{18}) and human GABRA3 (G_{18} A, G_{18} B) RNA substrates. As in the yeast HER1, the human GABRA3

transcript has a similar hairpin stem structure on the 5' end of the editing site and is also preferentially edited by human ADAR1-D over ADAR2-D[18].

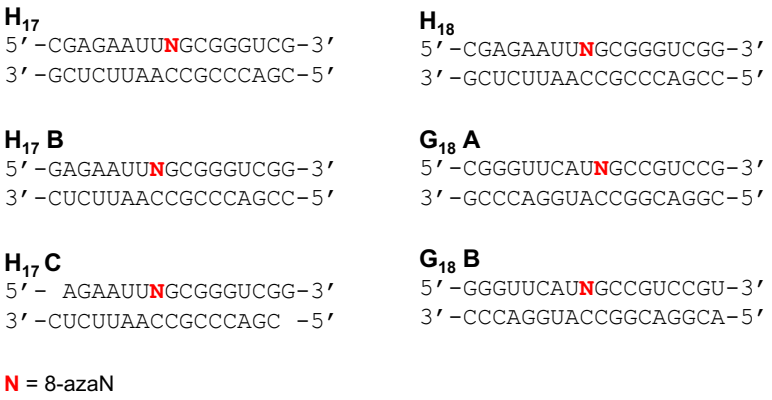


Figure 4.13. 8-azaN-modified RNA duplexes used for ADAR1-D-RNA crystallization trials.

Table 4.3. Screening conditions used for ADAR1-D-RNA crystallization trials.

dsRNA	H_{17}	H_{17} B	H_{17} C	H_{18}	G_{18} A	G_{18} B
ADAR1-D E1008Q to dsRNA ratios screened for crystallization	0.7:1, 1:1, 1.5:1	0.7:1, 1:1	0.7:1, 1:1, 1.5:1	0.7:1, 1:1, 1.5:1	0.7:1, 1:1, 1.5:1	0.7:1, 1:1, 1.5:1
96-well screen (25 ºC)	MCSG2, Nuc-Pro HTS, Natrix HT	MCSG2, Nuc-Pro HTS, Natrix HT	MCSG2, Nuc-Pro HTS, Natrix HT	MCSG1, Nuc-Pro HTS, Wizard Cryo 1 & 2	MCSG1, Nuc-Pro HTS, Wizard Cryo 1 & 2	MCSG1, Nuc-Pro HTS, Wizard Cryo 1 & 2
Optimization screen (25 ºC)	Nuc-Pro HTS F5, G4			Nuc-Pro HTS F5, G4		

A number of the conditions listed in Table 4.3 encouraged crystal formation; and even though most crystals are of the ADAR1-D-RNA complex, the best data that we have collected only diffracted to approximately 4.2 Å (Table 4.4). This came from a crystal derived from the following conditions: 150 µM ADAR1-D E1008Q, 100 µM H_{17} 8-azaN, 200 mM NaCl, and 50 mM Bis-Tris propane pH 7.0 (Fig. 4.14). It is important to note that several other crystals with

similar morphology were produced under the same conditions above at a range of pH (6.3-7.0) and salt concentrations (50-200 mM NaCl).

Table 4.4. Data processing and refinement statistics for ADAR1-D E1008Q bound to H_{17}.

Synchrotron (beamline)	APS (24-ID-C)
Wavelength (Å)	0.9792
Space group	*C*2
Unit cell parameters	a = 192.2 Å; b = 86.3 Å; c = 88.9 Å α = 90º; ß = 95.44º; γ = 90º
Resolution range (Å)	95.69-4.2 (4.62-4.2)
No. of observed reflections	73,506 (7364)
No. of unique reflections	10,705 (1060)
Completeness (%)	97.91 (92.25)
I/σ (I)	6.46 (0.34)
R_{merge}[a] (%)	10.7 (476.6)
$CC_{1/2}$	99.7 (25.6)
Refinement statistics	
R_{factor}[b] (%)	0.2896 (0.6011)
R_{free}[b] (%)	0.3545 (0.6527)
RMS bond length (Å)	0.005
RMS bond angle (º)	1.27
Ramachandran plot statistics[c]	
Favored (%)	94.43
Allowed (%)	2.72
Outliers (%)	2.85
No. of atoms	
Macromolecules	6860
Inositol hexakisphosphate	72
B factors	
Macromolecules	330.89
Ligands	315.60

[a] $R_{merge} = [\Sigma_h\Sigma_i|I_h - I_{hi}|/\Sigma_h\Sigma_i I_{hi}]$ where I_h is the mean of I_{hi} observations of reflection h. Numbers in parenthesis represent highest resolution shell.
[b] R_{factor} and $R_{free} = \Sigma||F_{obs}| - |F_{calc}|| / \Sigma|F_{obs}|$ x 100 for 95% of recorded data (R_{factor}) or 5% data (R_{free}).
[c] Ramachandran plot statistics from Molprobity[30].

Figure 4.14. Image of the crystals that grew from the conditions that gave the low resolution H_{17}-bound ADAR1-D E1008Q crystal structure. Growth conditions: 150 μM ADAR1-D E1008Q, 100 μM H_{17}, 200 mM NaCl, and 50 mM Bis-Tris propane, pH 7.0.

Despite being low resolution, a few key features can be inferred from the H_{17}-bound ADAR1-D crystal structure (Fig. 4.15). The electron density around the duplex RNA is well defined and 8-azaN is clearly shown to flip out of the duplex and into the enzyme's active site (Fig. 4.15a). As previously hypothesized, ADAR1's 5' binding loop appears to associate with fewer RNA base pairs 5' to the flipped-out base[18] (~4-5 bp, Fig. 4.15b) and a protein-RNA contact seems to be present ~6-7 bp 3' to 8-azaN (Fig. 4.15c). The electron density map also suggests the presence of a dimerization interface, consisting of amino acids analogous to those in ADAR2[10] (Fig. 4.15d). Finally, the data demonstrate that the four amino acids forming the second zinc binding site, as identified by Dr. Park[29], are within suitable distance to allow zinc chelation (Fig. 4.15e-f). However, the lack of strong characteristic zinc signal in this dataset suggests that the site may not be fully occupied.

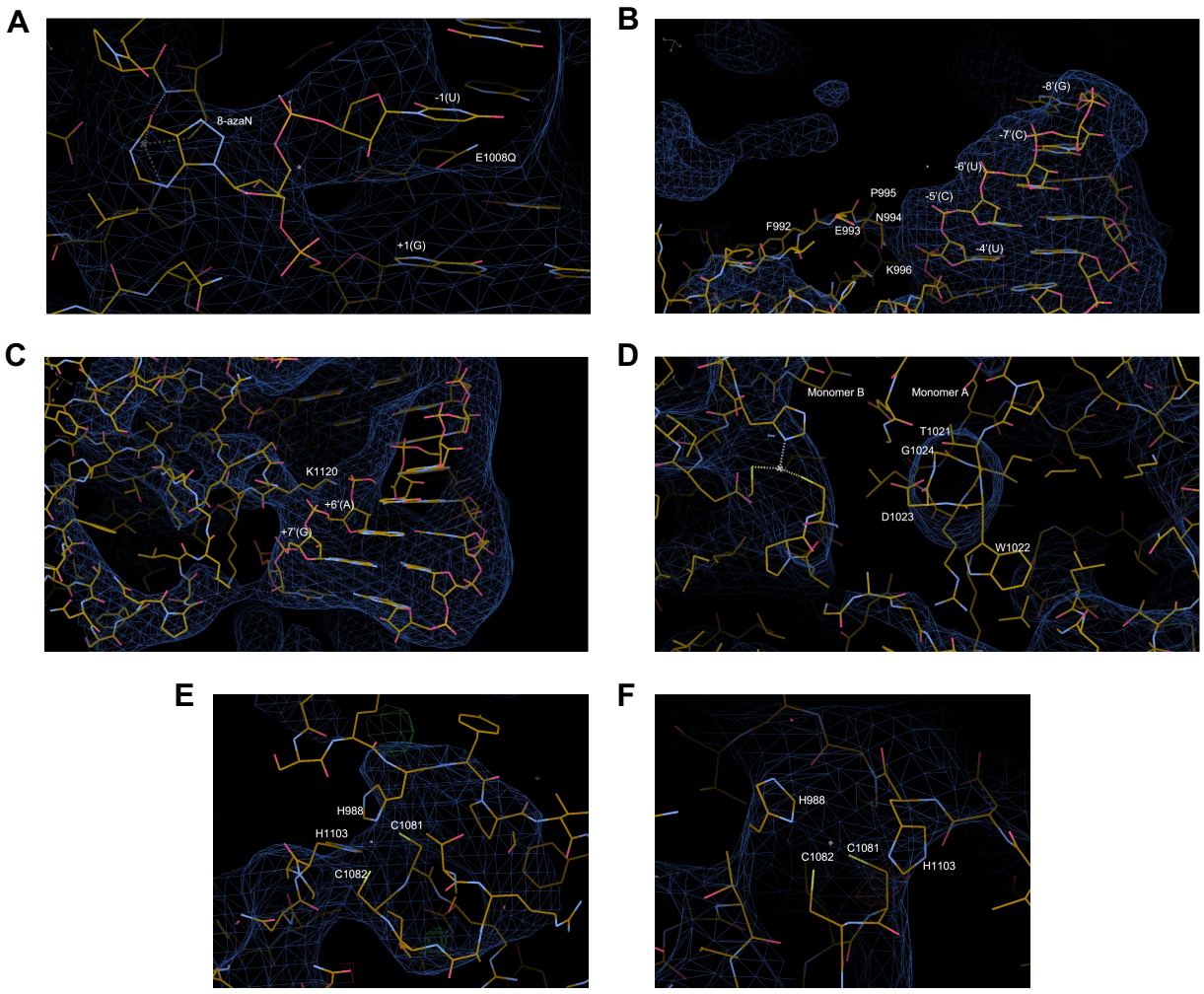


Figure 4.15. Electron density maps showing various regions in the low resolution H_{17}-bound ADAR1-D E1008Q crystal structure: (**A**) flipped-out 8-azaN; RNA binding contacts (**B**) 5' and (**C**) 3' to 8-azaN; (**D**) dimerization interface; and the second zinc binding site in (**E**) monomer A and (**F**) B.

4.2.6. Inhibition of ADAR1 in mammalian cells. The *in vitro* inhibition data discussed earlier in the chapter have encouraged us to test the efficacy of the 8-azaN-modified RNA duplexes in suppressing ADAR1 editing in mammalian cells. Dr. Kevin Pham had first spearheaded these efforts starting with the original H_{16} 8-azaN, 2'-OMe (Fig. 4.6), and gen2 duplexes, with the latter being a highly modified H_{16} 8-azaN duplex containing 2'-*O*-methyl, locked nucleic acid (LNA), phosphorothioate (PS), 2'-deoxynebularine (dN), and 2'-deoxy Benner's base (dZ) modifications.

LNA modification to the ribose sugar can help improve the duplex's thermodynamic stability[31,32], while PS modification to the sugar-phosphate backbone can aid in nuclease resistance[33–35]. When paired across a 5'-U nearest neighbor, dN has been shown to increase ADAR1 editing[36], while dZ at the orphan position has previously elicited the same favorable effect on ADAR1 activity[12]. The dN and dZ modifications were then utilized to potentially enhance ADAR1 recognition and selectivity towards the duplex inhibitor, thereby increasing its potency. Indeed, *in vitro* inhibition of the ADAR1 p110/5-HT_{2C} reaction with gen2 duplex revealed that all the modifications described did not substantially affect the IC_{50} (IC_{50} = 35 ± 5 nM; data not shown). Dr. Pham then proceeded to assess the inhibitory activity of H_{16} 8-azaN, 2'-OMe, and gen2 duplexes against endogenous and ectopically expressed ADAR1 by looking at specific editing sites on various substrates (COG3, EEF2K, BLCAP, GLI1, AZIN1, NUP43, NEIL1, 5-HT_{2C}) across different cell lines (HeLa, HEK293T, U87) and using either lipid nanoparticle or electroporation for oligonucleotide transfection. However, under all the conditions tested, Dr. Pham was not able to see any apparent reduction on substrate editing due to ADAR1 inhibition.

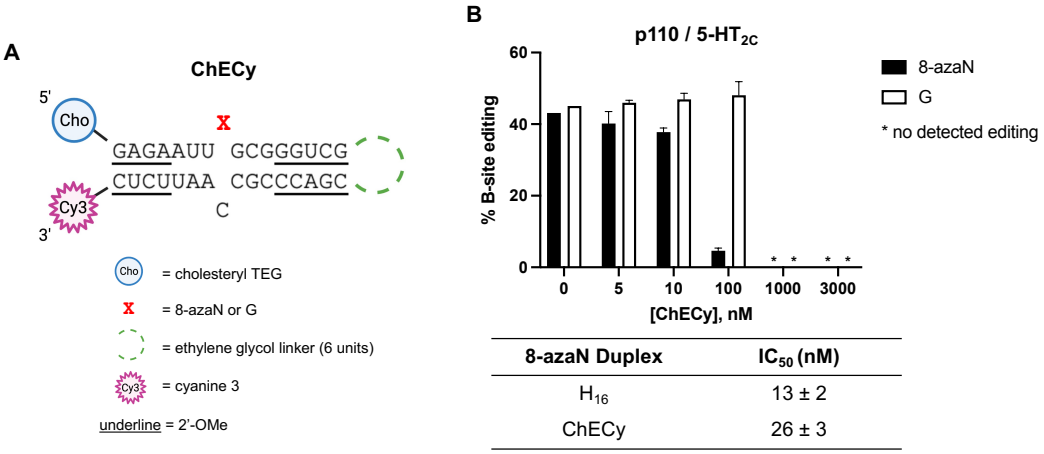


Figure 4.16. (**A**) Sequence of ChECy 8-azaN or G control hairpin duplex. TEG, triethylene glycol. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0–3 μM ChECy 8-azaN or G control duplex, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from n = 3 technical replicates.

As a continuation of Dr. Pham's work, a new heavily modified H_{16} 8-azaN duplex (ChECy 8-azaN) was designed and synthesized to study ADAR1 inhibition in cells. Apart from 8-azaN and 2'-*O*-methyl groups, the new oligo contains a cholesterol moiety at the 5' end to aid in delivery[37–39] and a Cy3 fluorophore at the 3' end to visualize transfection and subcellular localization (Fig. 4.16a). In addition, ChECy forms a hairpin or intramolecular duplex structure via a nuclease-resistant polyethylene glycol loop linking the two complementary strands of the parent H_{16} duplex (Fig. 4.16a). These modifications were aimed to address the potential delivery, compartmentalization, and metabolic stability issues encountered by Dr. Pham.

The ChECy hairpin duplex was first tested for inhibition of ADAR1 p110/5-HT_{2C} reaction *in vitro*. Despite all the modifications described, ChECy 8-azaN reduced 5-HT_{2C} editing by ADAR1 p110 at a remarkably similar potency as H_{16} 8-azaN (IC_{50} = 26 ± 3 nM; Fig. 4.16b). Notably, the ChECy G control also inhibited the reaction at concentrations higher than or equal to 1 μM, possibly due to non-specific binding brought about by the cholesterol, Cy3, and/or ethylene glycol modifications. Nevertheless, 8-azaN-dependent inhibition was clearly observed at [ChECy] = 5 to 100 nM. These results stimulated us to progress into cellular studies, beginning with immunofluorescence imaging experiments of U87 cells transfected with the ChECy hairpin duplexes (8-azaN and G control). The data in Fig 4.17 suggest that the ChECy duplexes were getting into the cells even at the lowest concentration tested (50 nM), albeit not conclusively in a concentration-dependent manner particularly for the ChECy 8-azaN duplex. Like the Cy3-labeled anti-miR control, the transfected duplexes appear as pink puncta against the green fluorescent dye-labeled cytoplasm, and hence are mainly cytoplasmic. While some of the duplexes are dispersed throughout the cytoplasm, some also seem to cluster around the DAPI-stained nucleus (shown in

blue; Fig. 4.17). The appearance of distinct puncta implies that the oligonucleotides were likely internalized by endocytosis[40].

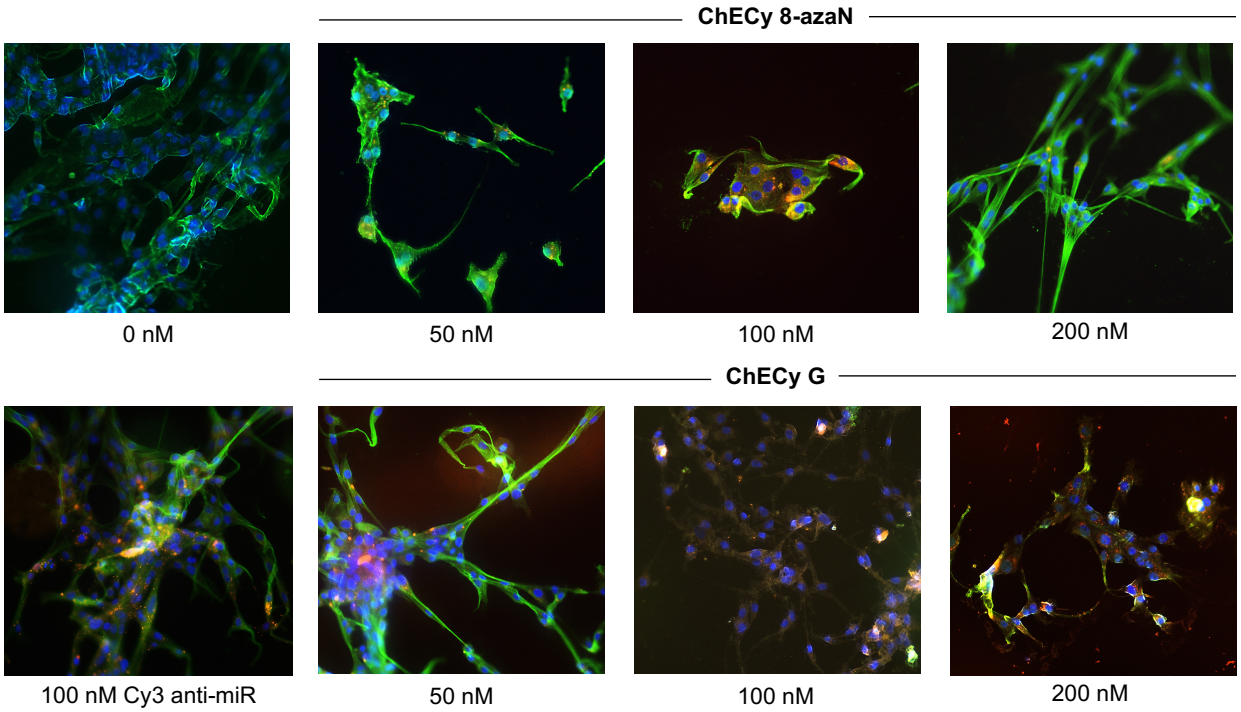


Figure 4.17. Immunofluorescence imaging of U87 cells transfected with ChECy 8-azaN or G hairpin duplexes (0 to 200 nM). Cy3-labeled anti-miR (100 nM) was used as a transfection control. Transfected oligonucleotides appear as pink in green (cytoplasm) or blue (nucleus) background.

Next, the effect of duplex inhibitor transfection on NUP43 and NEIL1 transcript editing by endogenous ADAR1 was explored. Dr. Kevin Pham had previously established that NUP43 is a preferred editing transcript by endogenous ADAR1 in HEK293T cells, whereas NEIL1 is also a preferred substrate by ADAR1[22] as discussed earlier in the chapter. HEK293T cells were then transfected with the H_{16} 8-azaN or G and ChECy 8-azaN or G duplexes and editing of NUP43 and NEIL1 substrates were quantified by Sanger sequencing post cDNA amplification from total RNA. Unfortunately, we did not observe any decrease in editing of NUP43 at Sites 1 to 4 (Fig. 4.18) and NEIL1 at Sites 1 and 2 (Fig. 4.19) for any or the inhibitor duplexes tested. However, the siRNA

targeting ADAR1 (siRNA) showed reduction of NUP43 and NEIL1 editing at all sites (Fig. 4.18-4.19), indicating successful introduction into cells and knockdown of endogenous ADAR1 transcripts. It is important to note that these experiments have also been done in U87 cells and like in HEK293T cells, we did not observe any apparent inhibition of NEIL1 editing (data not shown).

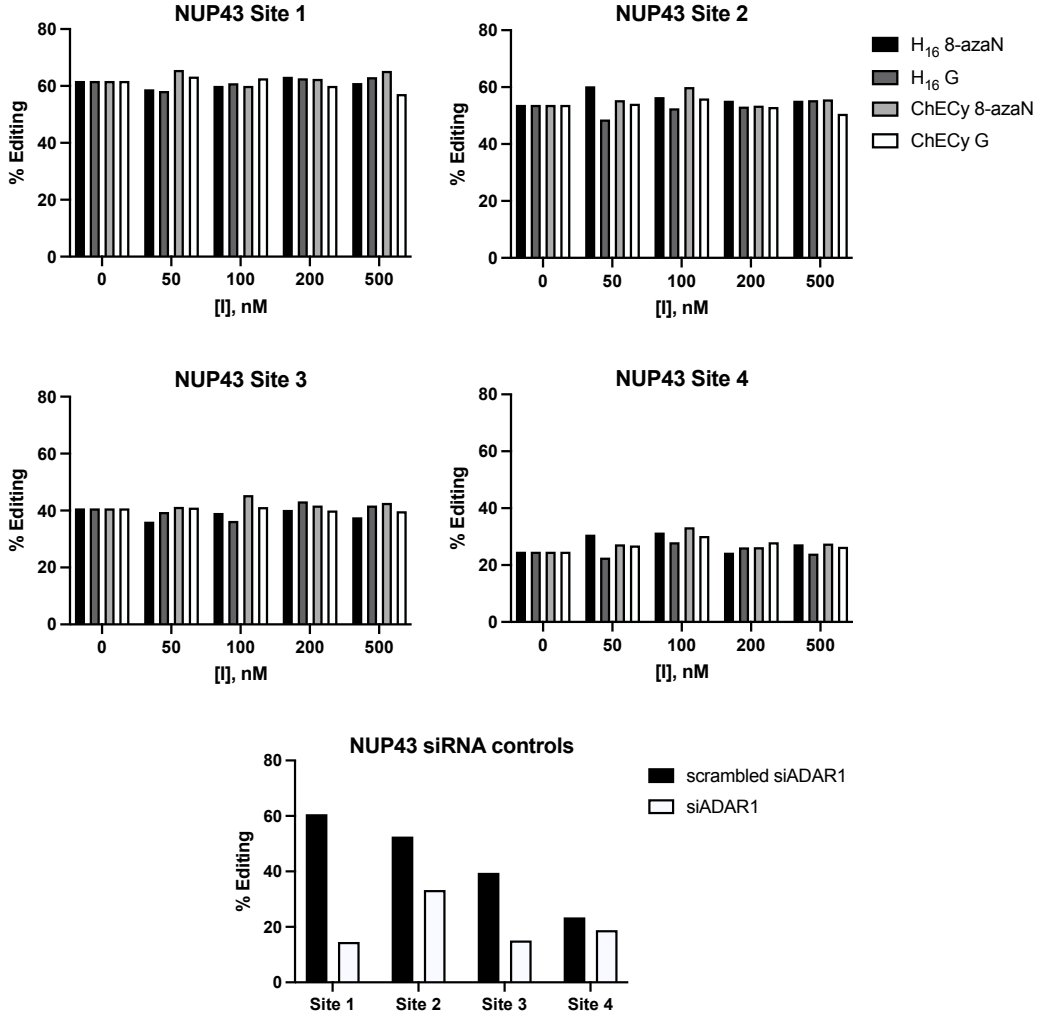


Figure 4.18. Editing of NUP43 transcript (Sites 1 to 4) in HEK293T cells upon transfection with 0-500 nM H_{16} 8-azaN or G and ChECy 8-azaN or G. siADAR1 and scrambled siADAR1 (100 nM) were used as positive and negative controls for ADAR1 editing inhibition.

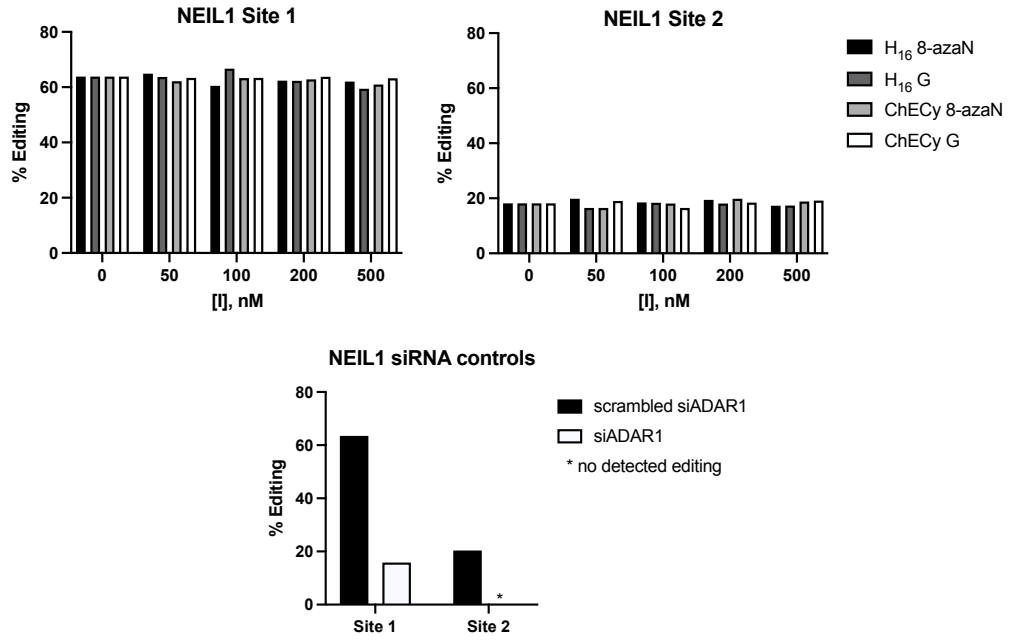


Figure 4.19. Editing of NEIL1 transcript (Sites 1 and 2) in HEK293T cells upon transfection with 0-500 nM H_{16} 8-azaN or G and ChECy 8-azaN or G. siADAR1 and scrambled siADAR1 (100 nM) were used as positive and negative controls for ADAR1 editing inhibition.

4.3. Discussion

The nucleoside analog, 8-azaN, is a known inhibitor (IC_{50} = 1.5 μM) of the nucleoside modifying enzyme, adenosine deaminase (ADA)[41]. Although structurally unrelated, ADAs and ADARs can both displace different leaving groups from C6 and the deamination rates of both enzymes are enhanced by 8-aza substitution[16]. This had driven the Beal group to test 8-azaN for ADAR2 inhibition in an earlier study where 8-azaN as a free ribonucleoside was found to inhibit the ADAR2 reaction poorly with a disappointingly high IC_{50} value of 15 mM[16]. Incorporating the nucleoside in a short RNA duplex, on the other hand, dramatically increased binding affinity to a K_D = 2 nM[17]. Since then, this nucleoside analog has been routinely used by our lab for acquiring a deeper understanding of ADAR2-RNA recognition[9,10,14,15]. However, prior to this work, whether 8-azaN could similarly enable studies of the related ADAR1 enzyme was unknown. Here, we show that like ADAR2, ADAR1 forms high-affinity complexes with 8-azaN-modified duplexes (Figs.

4.2 and 4.4). It has also been established that the 8-azaN as a free ribonucleoside does not inhibit ADAR1 and requires a duplex structure for effective binding (Fig. 4.5).

Previously, the Beal group has used gel-based assays such as gel shift and endonuclease footprinting experiments to probe ADAR2-RNA interactions and give insight into potential substrates for X-ray crystallography screens[9,10,13,14]. However, the use of these techniques for ADAR1 studies has been complicated by nonspecific binding from protein aggregates, especially for longer ADAR1 constructs (containing one or more dsRNA or Z-DNA/RNA binding domains). Estimation of relative binding affinities by means of IC_{50} values in the *in vitro* deamination assay described here offers an alternative method for biochemical and biophysical investigations of ADAR1-RNA interactions. It is also noteworthy to mention that this assay is compatible with both full-length ADAR1 isoforms p110 and p150 (Fig. 4.4) and does not necessitate the use of the hyperactive mutant to give quantifiable binding measurements with the 8-azaN-modified duplexes. The ADAR1 inhibition results, especially of the ADAR1 p150 isoform, are also highly relevant given recent findings suggesting p150 as the ADAR1 isoform mainly responsible for cancer progression[42–44].

The results from SAR studies with the panel of 8-azaN duplexes of different lengths suggest important ADAR1-RNA interactions ∼4−5 bp 5′ and ∼7−8 bp 3′ to the editing site (Fig. 4.7). Importantly, these observations agree with predicted RNA contacts between the ADAR1 deaminase and an RNA duplex from our previously published model of the ADAR1 catalytic domain-RNA complex[29] (Fig. 4.8). The model shows two lysine residues (K996 and K1120) that are in position to contact the phosphodiester backbone of the target strand's complement. The K996 side chain approaches the duplex between the third and fourth base pair 5′ from the editing site while the side chain of K1120 appears to approach between the seventh and eighth base pair

3′ from the edit site. Additionally, Fig. 4.20a illustrates the predicted 5′ binding loop contacts of ADAR1 with the RNA duplex 5′ to the edit site. Here, the model suggests that the ADAR1's 5′ binding loop only makes notable contacts (via K974 and K996) proximal to the editing site (up to ~4 bp 5′ from the reactive nucleotide) as the rest of the loop folds away from the RNA to stabilize the second Zn binding site (via H988)[29]. A longer duplex in the 3' direction may also allow for ADAR1's dsRBD3-RNA contacts if we infer from the structural information available for ADAR2's dsRBD2-RNA interactions described in Chapter 2. This could have also contributed to the higher potency for ADAR1 inhibition observed for $H_{14}A$ compared to $H_{14}B$.

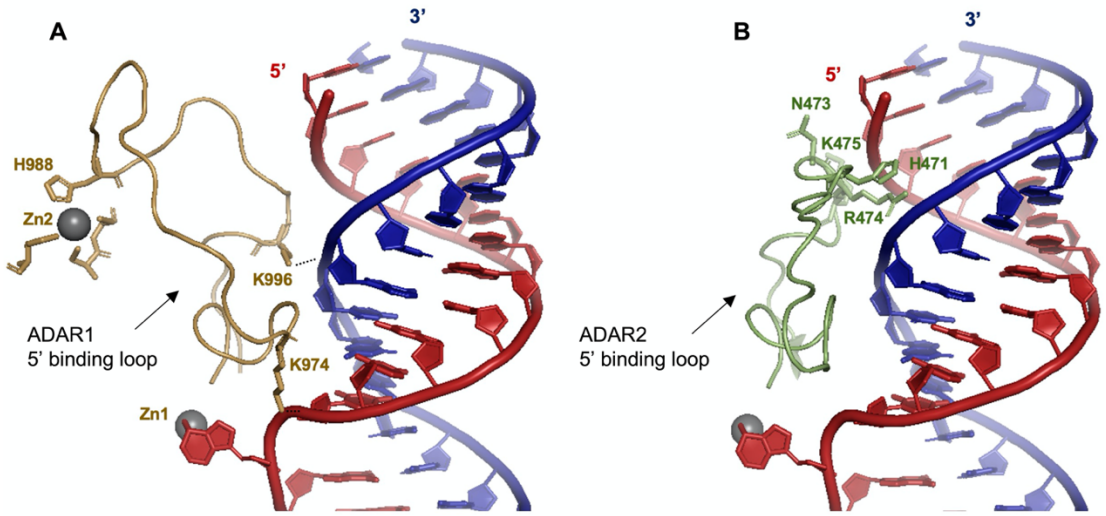


Figure 4.20. (**A**) Predicted contacts between the 5′ binding loop of ADAR1 (gold) and an RNA duplex (red = target strand; blue = complement strand) based on a previously reported Rosetta homology model of ADAR1 catalytic domain[29]. Active site Zn metal (Zn1, gray sphere). ADAR1's second Zn metal (Zn2, gray sphere) binding site is also shown. (**B**) Contacts between the 5′ binding loop of ADAR2 (green) and an RNA duplex as shown by X-ray crystallography (PDB 6vff[10]). Active site Zn metal (gray sphere).

It is also important to note that the minimum duplex length that was presented in this chapter is in agreement with two previously described short ADAR1 substrates by Dr. Yuru Wang[18] and Herbert and Rich[45]. Dr. Wang was able to show that a 33 nt hairpin RNA duplex with

5 bp 5′ and 8 bp 3′ to edit site is efficiently edited by human ADAR1-D[18]. Herbert and Rich also reported that a 15 bp RNA duplex with the edit site at the center and with a single base mismatch at the edit site is sufficient for ADAR1 editing[45]. These findings are consistent with ADAR1 requiring only a short 5′ duplex to facilitate substrate recognition and reaction. The inability of the H_{16} 8-azaN duplex to effectively bind and inhibit ADAR2 (Fig. 4.12) is explained by ADAR2 deaminase domain's requirement for a longer duplex on the 5′ side of the editing site for full contact. This is clearly seen in crystal structures of ADAR2 in complex with an RNA duplex where a cluster of residues (*e.g.*, H471, N473, R474, and K475) in the 5′ binding loop of ADAR2 contacts the edited strand of the duplex ~10 bp from the edit site (Fig. 4.20b)[9,10,46].

Although the low resolution ADAR1-D-RNA crystal structure we obtained cannot give us atomic details of protein-RNA interactions, it is remarkable how the apparent RNA binding contacts at the 5' and 3' ends of the duplex coincide with our data from SAR studies and the Rosetta homology model. Albeit longer than ADAR2's, ADAR1's 5' binding loop only associates with fewer base pairs 5' to the edited base as it folds away from the RNA to form part of the second Zn binding site that is unique to ADAR1[29]. It is possible that full occupancy of this Zn binding site is necessary to restrict the 5' binding loop's conformational flexibility, which could favor ordered crystallization. Indeed, the 5' binding loop of ADAR2 was only well resolved in co-crystal structures of the protein with RNA and was disordered in the RNA-free crystal structure[8], demonstrating the dynamic nature of these binding loops. However, we found that supplementing $ZnCl_2$ (100 μM) in the optimization screens discouraged crystal growth and adding it in the purification buffers encouraged concentration-dependent protein aggregation, making it difficult to get to a concentration high enough for crystallography. Although these experiments warrant testing the effect of supplementing lower $ZnCl_2$ concentrations, cryogenic electron microscopy

(cryo-EM) studies are now being pursued by our current postdoctoral researcher Dr. Sukanya Mozumder to obtain high resolution structures of ADAR1. Aside from this approach not requiring high concentrations of protein, Dr. Mozumder's use of a purification tag different from His-tag could also be advantageous as it is possible that the high amounts of imidazole used during the Ni-NTA elution step could be stripping off some of the Zn metal ions from the protein.

The use of shorter RNA substrates allows for routine incorporation of nucleoside analogs such as 2-aminopurine[12,19,47–49] and 8-azanebularine[14,15,17,28] for ADAR mechanistic and biophysical studies. It also offers an easily modifiable framework for other chemical modifications to enhance targeted delivery, metabolic stability, and binding affinity for cellular studies involving ADARs[4,12,50–52]. The results from *in vitro* inhibition experiments with the 2′-OMe (Fig. 4.6), gen2 (data not shown), cross-linked (Fig. 4.10), and CheCy duplexes (Fig. 4.16) demonstrate the flexibility for further modification of the 8-azaN duplexes. Although ADAR1 inhibition by these duplexes has yet to be established in cells, the modifications used appear to serve their purpose as shown by the increased nuclease stability of the 2'-OMe and gen2 duplexes (Dr. Kevin Pham's unpublished work) and improved thermodynamic stability of the 2'-OMe and cross-linked duplexes (Table 4.2).

A few points have to be considered for further studies on cellular inhibition of ADAR1 by the 8-azaN-modified duplexes. The effective concentration of ADAR1, editing transcripts, and transfected inhibitor duplexes at different cellular compartments should greatly influence editing outcomes and whether any editing inhibition can be detected or observed. It is likely that the internalized oligonucleotides are being sequestered by intracellular vesicles like endosomes and are undergoing gradual endosomal escape[53,54]. This issue can be potentially addressed by increasing the amount of duplex inhibitor transfected; however, using higher concentrations could

result in immune stimulation and cellular toxicity[55,56]. Chemical modifications that could improve ADAR1 binding and inhibitor potency could also be employed and should be pursued. In addition, the use of these 8-azaN-containing duplexes as PROTACs (proteolysis targeting chimeras) can enhance potency through an alternative inhibition mechanism[57]. In this approach, the ADAR1-binding oligonucleotide can be tethered to a ubiquitin ligase to facilitate proximity-induced ADAR1 degradation via the ubiquitin-proteasome pathway.

It is also important to note that the cells used in the cellular inhibition studies were harvested for total RNA extraction 48 h post duplex inhibitor transfection. Evaluation of ADAR1 editing inhibition for this length of time could be challenging, especially for sites that are well edited by ADAR1 and hence would not be as sensitive to ADAR1 inhibition. Future experiments should then consider optimizing the treatment time and the use of transcriptome-wide analyses (*e.g.* via RNA-seq) as a more sensitive and robust method for evaluating editing of multiple transcripts upon inhibitor treatment.

4.4. Methods

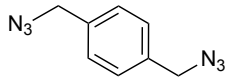
4.4.1. Oligonucleotide synthesis and purification. Oligonucleotides were synthesized and purified in collaboration with Victorio Jauregui Matos, Dr. Kevin Pham, and Dr. SeHee Park. ADAR editing substrates (5-HT_{2C} and NEIL1) were transcribed from a DNA template using HiScribe T7 RNA synthesis kit (NEB). Unmodified and 2′-*O*-methylated oligonucleotides for generation of short duplexes were purchased from IDT. Oligonucleotides containing 8-azaN were chemically synthesized on a 0.2 μmol scale using an ABI 394 synthesizer. The 8-azaN ribonucleoside phosphoramidite was either purchased from Berry & Associates, Inc. or synthesized in-house[17] by Victorio Jauregui Matos. All other phosphoramidites were purchased from Glen Research. Upon synthesis completion, columns were dried overnight under vacuum.

The oligonucleotides were then cleaved from the solid support by treatment with 1:3 ethanol/30% NH_4OH at 55 °C for 12 h. The supernatant was evaporated to dryness under vacuum, and then the pellet was resuspended in anhydrous dimethyl sulfoxide (DMSO). Deprotection was performed by treatment with 55% (v/v) Et_3N-3HF at room temperature overnight. The oligonucleotides were then precipitated from the solution by butanol precipitation at −70 °C, desalted using a Sephadex G-25 column, and purified as follows (see Table 4.5 for sequences).

All oligonucleotides (purchased or synthesized in-house) were purified using denaturing PAGE and visualized by UV shadowing. Bands were excised, crushed, and soaked overnight at 4 °C in a buffer containing 0.5 M NH_4OAc and 1 mM EDTA. Gel fragments were removed using a 0.2 μm cellulose acetate filter, and the oligonucleotides were precipitated from the supernatant by ethanol precipitation at −70 °C. The oligonucleotides were then dried under vacuum, resuspended in nuclease-free water, and quantified by measuring the absorbance at 260 nm. Oligonucleotide masses were confirmed by matrix assisted laser desorption/ionization - time of flight (MALDI-TOF) mass spectrometry using a Bruker Ultra-Flextreme MALDI TOF/TOF mass spectrometer at the UC Davis Mass Spectrometry Facility (see Table 4.6 for masses).

4.4.2. Synthesis of cross-linked RNA duplexes. The following syntheses were performed by Victorio Jauregui Matos. Two ssRNA oligonucleotides each containing 1′-ethynyl-2′-deoxyribose were cross-linked via a bis-azide linker, 1,4-bis(azidomethyl)-benzene (**1**), by copper(I)-catalyzed azide-alkyne cycloaddition or click chemistry[58]. Bis-azide linker 1 was synthesized according to the literature[59]. The ssRNA containing just the 1′-ethynyl-2′-deoxyribose modification (30 μM) was first clicked with an excess of 1 (5.5 mM) in 10:1 H_2O/MeCN to afford an azide-functionalized product. A flat copper bar was added to the mixture, and then the reaction was stirred in the dark at room temperature for 1 h. The copper bar was then removed, and the reaction was allowed to

continue at room temperature for 2 h. The oligonucleotides were passed through a 3,000 MW cutoff filter, washed twice with water, and then dried under vacuum. The azide-functionalized ssRNA was then purified from the starting material by denaturing PAGE.

1 

The ssRNA containing both 1′-ethynyl-2′-deoxyribose and 8-azaN ribose modifications (20 μM) was then clicked with the complementary azide-functionalized ssRNA (21 μM) in 10:1 H_2O/MeCN and 33 mM NaCl to afford the final crosslinked product. A flat copper bar was added to the mixture, and then the reaction was stirred in the dark at room temperature for 1 h. The copper bar was then removed, and the reaction was allowed to continue at room temperature for 2 h. The oligonucleotides were passed through a 3,000 MW cutoff filter, washed twice with water, and then dried under vacuum. The cross-linked RNA duplex was then purified from the starting material by denaturing PAGE.

4.4.3. Protein overexpression and purification. Human ADAR1 p110, ADAR2, and ADAR1-D E1008Q with a C- (ADAR1 p110) or N-terminal (ADAR2 and ADAR1-D E1008Q) His_{10}-tag were overexpressed in *S. cerevisiae* BCY123 as previously described[60]. Human ADAR1 p150 with an N-terminal FLAG-tag was purchased from BPS Bioscience. Human ADAR1 p110 was purified by lysing cells in lysis buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 1000 mM KCl, 30 mM imidazole, 1 mM TCEP-HCl, 0.05% (v/v) Triton X-100, and 50 μM $ZnCl_2$) using a microfluidizer. The clarified lysate was then passed over a Ni-NTA column using an ÄKTA pure 25 FPLC system. The column was washed with the lysis buffer and then with wash buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 500 mM KCl, 30 mM imidazole, 1 mM TCEP-HCl, and 50 μM $ZnCl_2$). Bound proteins were eluted by gradient elution with imidazole (30 to 400 mM). Fractions containing the

target protein were pooled, concentrated, and then dialyzed against a storage buffer containing 50 mM Tris-HCl pH 8.0, 10% (v/v) glycerol, 400 mM KCl, 50 mM imidazole, 1 mM TCEP-HCl, and 0.01% (v/v) NP-40. Protein concentration was determined by running the sample alongside BSA standards in an SDS-PAGE gel followed by SYPRO Orange (Invitrogen) staining.

Human ADAR2 was purified as above except using the following buffers: (1) lysis buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 750 mM NaCl, 35 mM imidazole, 0.01% (v/v) Triton X-100, 1 mM BME); (2) wash buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 300 mM NaCl, 35 mM imidazole, 0.01% (v/v) Triton X-100, 1 mM BME); and (3) storage buffer (20 mM Tris−HCl pH 8.0, 20% (v/v) glycerol, 100 mM NaCl, 1 mM BME).

Human ADAR1-D E1008Q was purified as above except using the following buffers: (1) lysis buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 750 mM NaCl, 30 mM imidazole, 1 mM TCEP-HCl, 0.05% (v/v) Triton X-100); (2) wash buffer (20 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 350 mM NaCl, 30 mM imidazole, 1 mM TCEP-HCl); and (3) storage buffer (50 mM Tris-HCl pH 8.0, 10% (v/v) glycerol, 200 mM KCl, 50 mM imidazole, 1 mM TCEP-HCl, 0.01% (v/v) NP-40).

4.4.4. Gel shift assay. Gel shift assays were performed in collaboration with Dr. SeHee Park. The 16mer 8-azaN or A containing RNA strand was end-labeled with ^{32}P using T4 polynucleotide kinase (NEB). Excess [γ-^{32}P]-ATP was removed by passing the reaction mixture through a Sephadex G-25 column. The labeled oligonucleotides were purified by denaturing PAGE gel as described above, except visualized by storage phosphor autoradiography. The labeled oligonucleotides were then hybridized at a 1:3 ratio to their complement in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 100 mM NaCl by heating at 95 °C for 5 min and slow cooling to ≤ 30 °C.

Samples containing 5 nM labeled duplex RNA and 0 to 300 nM enzyme were incubated in 15 mM Tris-HCl pH 7.5, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% (v/v) NP-40, 4% glycerol, 0.5 mM DTT, 1 μg/mL yeast tRNA, 0.16 U/μL RNAse inhibitor, and 0.2 mg/mL BSA at 30 °C for 30 min. Samples were loaded onto a 6% 80:1 acrylamide/bisacrylamide gel and electrophoresed under nondenaturing conditions in 1× TBE at 4 °C for 1.5 h. The gels were dried under reduced pressure at 80 °C for 1.5 h and then exposed to storage phosphor imaging plates (Kodak) in the dark. Plates were scanned using a Typhoon imaging system (GE Healthcare), and band intensities were quantified using ImageQuant (GE Healthcare). Data were plotted (fraction RNA bound = protein-RNA/(protein-RNA + free RNA)) and analyzed using Microsoft Excel and GraphPad Prism. The band intensities arising from protein-RNA aggregation in the wells were excluded in final data processing as it does not appreciably change the K_D values derived from the binding curves, whether included or excluded from the calculation for fraction of RNA bound to the protein.

4.4.5. *In vitro* deamination assay. ADAR editing substrates and cross-linked RNA duplexes were allowed to self-anneal in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 100 mM NaCl by heating at 95 °C for 5 min and slow cooling to ≤ 30 °C. The short intermolecular RNA duplexes were hybridized at a 1:3 ratio of target strand (8-azaN-containing or A/G-containing strand) to complement strand in the same buffer and conditions as above.

Deamination assays were performed under single-turnover conditions. Samples containing 100 nM ADAR1 and 0 to 3000 nM short RNA duplex were incubated in 15 mM Tris-HCl pH 7.5, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% (v/v) NP-40, 4% glycerol, 0.5 mM DTT, 1 μg/mL yeast tRNA, and 0.16 U/μL RNAse inhibitor at 30 °C for 30 min. The reaction was commenced by addition of 5 nM editing substrate and allowed to incubate at 30 °C for the

following times: 15 min for 5-HT_{2C} and 30 min for NEIL1 before quenching with 95 °C water, vortexing, and heating at 95 °C for 5 min. Oligonucleotides were then purified using DNA Clean & Concentrator kit (Zymo) before conversion to cDNA using Access RT-PCR System (Promega). PCR amplicons were purified using 1% agarose gel and submitted for Sanger sequencing (Azenta). Sequencing peak heights at the edit site were quantified using 4Peaks (Nucleobytes). The sequencing traces also served to validate the sequences of the editing substrates. Data were plotted (% editing = [peak height G/(peak height A + peak height G)] × 100%) and analyzed using Microsoft Excel and GraphPad Prism.

Deamination assays with ADAR2 were performed as above except in the following buffer conditions: 15 mM Tris-HCl pH 7.5, 60 mM KCl, 3 mM $MgCl_2$, 1.5 mM EDTA, 0.003% (v/v) NP-40, 3% glycerol, 0.5 mM DTT, 1 μg/mL yeast tRNA, and 0.16 U/μL RNAse inhibitor and for the following reaction times: 10 min for 5-HT_{2C} and 30 min for NEIL1.

4.4.6. Thermal denaturation studies. Equimolar amounts (1 μM) of intermolecular duplex strand components or 1 μM solution of cross-linked duplex were prepared in 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl, and 1.25 μM EvaGreen (Biotium). Samples were placed in a sealed 96-well plate and loaded onto CFX Connect Real-Time PCR Detection System (Bio-Rad). The plate was brought to 95 °C and held at 95 °C for 3 min, and then was cooled down to 15 °C at 0.1 °C/s and held at 15 °C for 5 min. The temperature was then slowly brought to 100 °C with fluorescence (F) measured every 0.2 °C (T). T_M was determined as the temperature where -dF/dT is the largest. Data were plotted and analyzed using Microsoft Excel and GraphPad Prism.

4.4.7. X-ray crystallography studies. The following were conducted in collaboration with Dr. Alexander Thuy-Boun. ADAR1-D E1008Q was overexpressed and purified as described above

except without $ZnCl_2$ in all purification buffers and including the following additional steps after the initial Ni-NTA purification. Fractions containing ADAR1-D E1008Q protein were pooled and buffer exchanged to TEV buffer (20mM Tris-HCl pH 8.0, 350 mM NaCl, 50 mM imidazole, 5% (v/v) glycerol, 1 mM TCEP). His_{10}-tag cleavage was performed using a 1:2 mass ratio of ADAR1-D to TEV protease for 3 h at room temperature. The protein mixture was passed through a second Ni-NTA column and the column was washed with wash buffer. The flowthrough and wash were pooled, concentrated, and buffer exchanged to heparin wash buffer (20mM Tris-HCl pH 8.0, 100 mM NaCl, 50 mM imidazole, 5% (v/v) glycerol, 1 mM TCEP) using a 30,000 MW cutoff filter (Amicon) at 5,000 × g to a volume of 10 mL. The sample was then loaded onto a 5 mL HiTrap Heparin HP column (GE Healthcare), washed with heparin wash buffer, and eluted with a linear NaCl gradient (0.1 to 1 M). Fractions containing the target protein were pooled, concentrated to < 1.5 mL, and ran through a HiLoad 16/600 Superdex 200pg column (GE Healthcare) using size exclusion buffer (same as TEV buffer). Fractions containing the target protein were then pooled and concentrated to 8-11 mg/mL for crystallization trials.

8-azaN-containing oligonucleotides were synthesized and purified as detailed above. The unmodified complement strands were purchased from IDT and also purified as described. Duplex RNA was hybridized in a 1:1 ratio in water by heating at 95 °C for 5 min and slow cooling to ≤ 30 °C. Crystals of ADAR1-D E1008Q bound to H_{17} were grown at room temperature by the hanging-drop vapor-diffusion method. A 200 nL solution containing 6.8 mg/mL ADAR1-D protein (150 μM) and 100 μM H_{17} RNA (1.5:1 molar ratio) was mixed with 200 nL of 200 mM NaCl in 50 mM Bis-Tris propane, pH 7.0. Crystals took approximately two weeks to grow. A single thick triangular crystal approximately 150 μm in size was soaked in a solution of mother liquor plus 30% dextrose before flash cooling in liquid N_2. Data were collected with 0.2º oscillations at

Argonne National Laboratory's Advanced Photon Source (APS) beam line 24-ID-C. Crystals diffracted to approximately 4.2 Å.

Crystallographic data processing and refinement were performed by Dr. Alexander Thuy-Boun. X-ray diffraction data were processed with XDS and scaled with XSCALE[61]. The structure was determined by molecular replacement using PHENIX[62]. The GLI1-bound ADAR2-D E488Q crystal structure (PDB 5ed2)[9] was used as a model for molecular replacement. The structure was refined using PHENIX including Zn coordination restraints[63]. Ideal Zn-ligand distances were determined with average distances found for similar coordination models in the PDB database.

4.4.8. Immunofluorescence imaging of transfected 8-azaN-containing duplex in U87 cells. U87 cells were cultured in DMEM, 10% FBS and 1% anti–anti at 37 ºC and 5% CO_2. At 70–90% confluency, 5×10^4 cells were seeded into a 24-well plate with coverslips. After 24 h, cells were transfected with 0 to 200 nM ChECy 8-azaN or G hairpin duplexes. Transfection was carried out using Lipofectamine RNAiMax. After incubation of transfection reagent and duplexes in Opti-MEM, the solution was added to designated well and incubated at 37 ºC and 5% CO_2 for 48 h. The growth media were aspirated from each well and the cells were fixed with 0.5 mL Image-iT Fixative Solution (Invitrogen) for 15 min. The cells were washed three times for 5 min each with Dulbecco's phosphate buffered saline (DPBS) and were treated with 0.5 mL permeabilization solution (0.5% Triton X-100 in phosphate buffered saline (PBS)) for 15 min. The cells were washed three times for 5 min each with DPBS, then 0.5 mL blocking buffer (3% BSA in PBS) was added and incubated for 1 h. After another three washes for 5 min each with PBS, cells were stained overnight at 4 °C with 0.5 mL mouse IgG alpha tubulin monoclonal primary antibody (Invitrogen) diluted to 1 μg/mL in 1% BSA in PBS. After 24 h, the primary antibody staining solution was removed and the cells were washed three times for 5 min each using DPBS. The cells

were stained for 1 h at room temperature in the dark with 0.5 mL secondary antibody solution containing rabbit anti-mouse IgG (H+L) cross-adsorbed Alexa-Fluor 488 conjugate (Invitrogen) diluted to 1 μg/mL in 1% BSA in PBS. The secondary antibody staining solution was removed and cells were washed three times for 5 min each with DPBS. One to two drops of NucBlue™ Live ReadyProbes™ Reagent (Hoechst 33342, Invitrogen) was added to each well of cells in 0.5 mL DPBS and incubated for 20. Finally, the cover slip containing fixed cells was mounted onto a microscope slide and imaged under DAPI, GFP, and/or RFP fluorescent cube settings at 20X magnification using an EVOS M7000 Imaging System (Invitrogen).

4.4.9. Cellular editing of endogenous substrates in HEK293T and U87 cells. HEK293T and U87 cells were cultured in DMEM (Gibco), 10% FBS (Gibco) and 1% anti–anti (Gibco) at 37 ºC and 5% CO_2. At 70–90% confluency, 6.4×10^3 HEK293T cells or 2.13×10^4 cells were seeded into a 96-well plate. After 24 h, cells were transfected with 0 to 500 nM H_{16} 8-azaN or G and ChECy 8-azaN or G duplexes. Transfection was carried out using Lipofectamine RNAiMax. After incubation of transfection reagent and plasmid in Opti-MEM, the solution was added to designated well and incubated at 37 ºC and 5% CO_2 for 48 h. Total RNA was isolated using RNAqueous Total RNA Isolation kit (Invitrogen) and DNase treated with RQ1 RNase-free DNase (Promega). Nested PCR was performed by first using Access RT-PCR kit (Promega) for 24 cycles followed by a second PCR using Phusion Hot Start DNA Polymerase (Thermo) for 19 cycles. The PCR products were purified by agarose gel and Qiagen Gel Extraction kit. The purified product was subjected to Sanger sequencing and sequence traces were analyzed by 4Peaks (Nucleobytes) to measure percent editing. See Section 4.3 for RT-PCR and nested PCR/sequencing primers.

4.5. Table of oligonucleotides

Table 4.5. Sequences of oligonucleotides used for work presented in Chapter 4. All oligonucleotides are made up of ribonucleotides except the primers (which are made up of 2'-deoxyribonucleotides) or when specified. TS = target strand, CS = complement strand, N = 8-azanebularine, underline = 2'-*O*-methylated, E = 1'-ethynyl-2'-deoxyribose, X = cholesteryl TEG, Y = ethylene glycol linker (6 units), Z = cyanine 3.

Oligonucleotide	Sequence (5' to 3')
For generating short duplexes	
H_{16} 8-azaN TS	GAGAAUUNGCGGGUCG
H_{16} A TS	GAGAAUUAGCGGGUCG
H_{16} G TS	GAGAAUUGGCGGGUCG
H_{16} CS	CGACCCGCCAAUUCUC
5'-3' swap TS	GCGGUGUNGAGAAUCG
5'-3' swap CS	CGAUUCUCCACACCGC
GC-rich TS	GGGGGGUNGGGGGGGG
GC-rich CS	CCCCCCCCCACCCCCC
AU-rich TS	AAAAAAUNGAAAAAAA
AU-rich CS	UUUUUUUCCAUUUUUU
2'-OMe TS	GAGAAUUNGCGGGUCG
2'-OMe CS	CGACCCGCCAAUUCUC
H_{14} A TS	GAAUUNGCGGGUCG
H_{14} A CS	CGACCCGCCAAUUC
H_{14} B TS	GAGAAUUNGCGGGU
H_{14} B CS	ACCCGCCAAUUCUC
H_{12} A TS	AUUNGCGGGUCG
H_{12} A CS	CGACCCGCCAAU
H_{12} B TS	GAAUUNGCGGGU
H_{12} B CS	ACCCGCCAAUUC
X_A TS	EAUUNGCGGGUCG
X_B TS	EAUUNGCGG
X CS	CGACCCGCCAAUEC
ChECy 8-azaN	XGAGAAUUNGCGGGUCGYCGACCCGCCAAUUCUCZ
ChECy G	XGAGAAUUGGCGGGUCGYCGACCCGCCAAUUCUCZ
RT-PCR and sequencing primers	
5-HT_{2C} FWD	TGGGTACGAATTCCCACTTACGTACAAGCTT
5-HT_{2C} RVS	AGAACCCGATCAAACGCAAATGTTAC
NEIL1 FWD	TAATACGACTCACTATAGGGAACTGACTAAACTAGGTGCCACGTCG TGAAAGTCTGACAACCTGAGCCTGCCCTCTGA
NEIL1 RVS	AGTCCTCCTCCCCGCTCTCTGAC
For X-ray crystallographic studies	
H_{17} TS	CGAGAAUUNGCGGGUCG
H_{17} CS	CGACCCGCCAAUUCUCG
H_{17} B TS	GAGAAUUNGCGGGUCGG
H_{17} B CS	CCGACCCGCCAAUUCUC
H_{17} C TS	AGAAUUNGCGGGUCGG
H_{17} C CS	CGACCCGCCAAUUCUC
H_{18} TS	CGAGAAUUNGCGGGUCGG

H_{18} CS	CCGACCCGCCAAUUCUCG
G_{18} A TS	CGGGUUCAUNGCCGUCCG
G_{18} A CS	CGGACGGCCAUGGACCCG
G_{18} B TS	GGGUUCAUNGCCGUCCGU
G_{18} B CS	ACGGACGGCCAUGGACCC
RT-PCR and nested PCR/sequencing primers for cellular editing studies	
NUP43 RT FWD	TCTTTCCAAACTTTGTTAGATTTTAAATGTATTATTGACCTGAGA
NUP43 RT RVS	CGGTGTTACATTACATAAAGCTTAGTTTCTTATTAAAATTGGCAA
NUP43 Nest FWD	CCTTAATGACAAATCACTGCTATTAGACAATTG
NUP43 Nest RVS	AACTTGTCAATTGGCATGAAATGTAATGCT
NEIL1 RT FWD	TCCAGACCTGCTGGAGCTAT
NEIL1 RT RVS	TGGCCTTGGATTTCTTTTTG
NEIL1 Nest FWD	CCCAAGGAAGTGGTCCAGTTGG
NEIL1 Nest RVS	CTGGAACCAGATGGTACGGCC

Table 4.6. Calculated and observed masses of oligonucleotides used for work presented in Chapter 4. TS = target strand, CS = complement strand.

Oligonucleotide	Calculated Mass (g/mol)	Observed Mass (m/z)
H_{16} 8-azaN TS	5183.7	5182.4
H_{16} A TS	5197.7	5196.4
H_{16} G TS	5213.7	5214.3
H_{16} CS	4973.7	4972.7
5’-3’ swap TS	5183.7	5186.8
5’-3’ swap CS	4973.7	4973.4
GC-rich TS	5389.8	5385.9
GC-rich CS	4842.7	4833.2
AU-rich TS	5181.8	5175.2
AU-rich CS	4855.5	4848.0
2’-OMe TS	5309.9	5312.2
2’-OMe CS	5099.9	5097.6
H_{14} A TS	4509.6	4506.7
H_{14} A CS	4362.6	4360.8
H_{14} B TS	4533.6	4530.7
H_{14} B CS	4323.6	4319.6
H_{12} A TS	3835.5	3834.0
H_{12} A CS	3751.6	3751.8
H_{12} B TS	3859.5	3857.1
H_{12} B CS	3712.5	3712.7
X_A TS	4038.5	4040.2
X_B TS	2737.4	2736.4
X CS	4245.6	4249.4
X-azide CS	4433.7	4435.6
X_{A+}	8472.2	8472.2
X_{B+}	7171.0	7171.2
H_{17} TS	5488.8	5486.4

H_{17} CS	5318.7	5319.7
H_{17} B TS	5528.7	5531.3
H_{17} B CS	5278.7	5279.3
H_{17} C TS	5183.7	5184.0
H_{17} C CS	4973.7	4975.1
H_{18} TS	5833.8	5836.7
H_{18} CS	5623.8	5625.9
G_{18} A TS	5706.7	5707.4
G_{18} A CS	5781.8	5782.9
G_{18} B TS	5707.7	5710.1
G_{18} B CS	5765.8	5766.7
ChECy 8-azaN	12010.5	12010.3
ChECy G	12040.6	12040.2

4.6. References

(1) Mendoza, H. G.; Matos, V. J.; Park, S. H.; Pham, K. M.; Beal, P. A. Selective Inhibition of ADAR1 Using 8-Azanebularine-Modified RNA Duplexes. *Biochemistry* **2022**, *62* (8), 1376–1387. https://doi.org/10.1021/acs.biochem.2c00686.

(2) Ishizuka, J. J.; Manguso, R. T.; Cheruiyot, C. K.; Bi, K.; Panda, A.; Iracheta-Vellve, A.; Miller, B. C.; Du, P. P.; Yates, K. B.; Dubrot, J.; Buchumenski, I.; Comstock, D. E.; Brown, F. D.; Ayer, A.; Kohnle, I. C.; Pope, H. W.; Zimmer, M. D.; Sen, D. R.; Lane-Reticker, S. K.; Robitschek, E. J.; Griffin, G. K.; Collins, N. B.; Long, A. H.; Doench, J. G.; Kozono, D.; Levanon, E. Y.; Haining, W. N. Loss of ADAR1 in Tumours Overcomes Resistance to Immune Checkpoint Blockade. *Nature* **2019**, *565* (7737), 43–48. https://doi.org/10.1038/s41586-018-0768-9.

(3) 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), 1–10. https://doi.org/10.1038/s41467-018-07824-4.

(4) Monian, P.; Shivalila, C.; Lu, G.; Shimizu, M.; Boulay, D.; Bussow, K.; Byrne, M.; Bezigian, A.; Chatterjee, A.; Chew, D.; Desai, J.; Favaloro, 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**, *40* (7), 1093–1102. https://doi.org/10.1038/s41587-022-01225-1.

(5) Nishikura, K. A-to-I Editing of Coding and Non-Coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **2016**, *17* (2), 83–96. https://doi.org/10.1038/nrm.2015.4.

(6) Kim, U.; Wang, Y.; Sanford, T.; Zeng, Y.; Nishikura, K. Molecular Cloning of CDNA for Double-Stranded RNA Adenosine Deaminase, a Candidate Enzyme for Nuclear RNA Editing. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (24), 11457–11461. https://doi.org/10.1073/pnas.91.24.11457.

(7) Doherty, E. E.; Beal, P. A. Oligonucleotide-Directed RNA Editing in Primates. *Mol. Ther.* **2022**, *30* (6), 2117–2119. https://doi.org/10.1016/j.ymthe.2022.04.005.
(8) Macbeth, M. R.; Schubert, H. L.; VanDemark, A. F.; Lingam, A. T.; Hill, C. P.; Bass, B. L. Structural Biology: Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science (80-.).* **2005**, *309* (5740), 1534–1539. https://doi.org/10.1126/science.1113150.
(9) 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.
(10) Thuy-Boun, A. S.; Thomas, J. M.; Grajo, H. L.; Palumbo, C. M.; Park, S.; 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), 7958–7952. https://doi.org/10.1093/nar/gkaa532.
(11) 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.
(12) 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.
(13) Doherty, E. E.; Karki, A.; Wilcox, X. E.; Mendoza, H. G.; Manjunath, A.; Matos, V. J.; Fisher, A. J.; Beal, P. A. ADAR Activation by Inducing a Syn Conformation at Guanosine Adjacent to an Editing Site. *Nucleic Acids Res.* **2022**, 1–12. https://doi.org/10.1093/nar/gkac897.
(14) 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.
(15) Maydanovych, O.; Beal, P. A. C6-Substituted Analogues of 8-Azanebularine: Probes of an RNA-Editing Enzyme Active Site. *Org. Lett.* **2006**, *8* (17), 3753–3756. https://doi.org/10.1021/ol061354j.
(16) Véliz, E. A.; Easterwood, L. H. M.; Beal, P. A. Substrate Analogues for an RNA-Editing Adenosine Deaminase: Mechanistic Investigation and Inhibitor Design. *J. Am. Chem. Soc.* **2003**, *125* (36), 10867–10876. https://doi.org/10.1021/ja029742d.
(17) 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.
(18) Wang, Y.; Park, S. H.; Beal, P. A. Selective Recognition of RNA Substrates by ADAR Deaminase Domains. *Biochemistry* **2018**, *57* (10), 1640–1651. https://doi.org/10.1021/acs.biochem.7b01100.
(19) Kuttan, A.; Bass, B. L. Mechanistic Insights into Editing-Site Specificity of ADARs. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (48), 3295–3304. https://doi.org/10.1073/pnas.1212548109.
(20) Zuker, M. Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Res.* **2003**, *31* (13), 3406–3415. https://doi.org/10.1093/nar/gkg595.
(21) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-

Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 1–9. https://doi.org/10.1038/ncomms1324.
(22) Yeo, J.; Goodman, R. A.; Schirle, N. T.; David, S. S.; Beal, P. A. RNA Editing Changes the Lesion Specificity for the DNA Repair Enzyme NEIL1. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (48), 20715–20719. https://doi.org/10.1073/pnas.1009231107.
(23) 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.
(24) Li, J. B.; Levanon, E. Y.; K., Y. J.; Aach, J.; Xie, B.; LeProust, E.; Zhang, K.; Gao, Y.; Church, G. M.; 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* (May), 1210–1213.
(25) Choung, S.; Kim, Y. J.; Kim, S.; Park, H. O.; Choi, Y. C. Chemical Modification of SiRNAs to Improve Serum Stability without Loss of Efficacy. *Biochem. Biophys. Res. Commun.* **2006**, *342* (3), 919–927. https://doi.org/10.1016/j.bbrc.2006.02.049.
(26) Cummins, L. L.; Owens, S. R.; Risen, L. M.; Lesnik, E. A.; Freier, S. M.; Mc Gee, D.; Cook, C. J.; Cook, P. D. Characterization of Fully 2'-Modified Oligoribonucleotide Hetero-and Homoduplex Hybridization Andnuclease Sensitivity. *Nucleic Acids Res.* **1995**, *23* (11), 2019–2024. https://doi.org/10.1093/nar/23.11.2019.
(27) Monia, B. P.; Johnston, J. F.; Sasmor, H.; Cummins, L. L. Nuclease Resistance and Antisense Activity of Modified Oligonucleotides Targeted to Ha-Ras. *J. Biol. Chem.* **1996**, *271* (24), 14533–14540. https://doi.org/10.1074/jbc.271.24.14533.
(28) Maydanovych, O.; Easterwood, L. H. M.; Cui, T.; Véliz, E. A.; Pokharel, S.; Beal, P. A. Probing Adenosine-to-Inosine Editing Reactions Using RNA-Containing Nucleoside Analogs. *Methods Enzymol.* **2007**, *424* (07), 369–386. https://doi.org/10.1016/S0076-6879(07)24017-0.
(29) 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), 1–13. https://doi.org/10.1038/s41467-020-18862-2.
(30) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci.* **2018**, *27* (1), 293–315. https://doi.org/10.1002/pro.3330.
(31) Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J. I.; Morio, K. I.; Doi, T.; Imanishi, T. Stability and Structural Features of the Duplexes Containing Nucleoside Analogues with a Fixed N-Type Conformation, 2'-O,4'-C- Methyleneribonucleosides. *Tetrahedron Lett.* **1998**, *39* (30), 5401–5404. https://doi.org/10.1016/S0040-4039(98)01084-3.
(32) Koshkin, A.; Singh, S. K.; Nielsen, P.; Rajwanshi, V. K.; Kumar, R.; Meldgaard, M.; Olsen, C. E.; Wengel, J. LNA (Locked Nucleic Acids): Synthesis of the Adenine, Cytosine, Guanine, 5-Methylcytosine, Thymine and Uracil Bicyclonucleoside Monomers, Oligomerisation, and Unprecedented Nucleic Acid Recognition. *Tetraheron* **1998**, *54*, 3607–3630. https://doi.org/10.1080/07328319908044718.
(33) Hoke, G. D.; Draper, K.; Freier, S. M.; Gonzalez, C.; Driver, V. B.; Zounes, M. C.; Ecker, D. J. Effects of Phosphorothioate Capping on Antivense Oligonucleotide Stability, Hybridization and Antiviral Efficacy versus Herpes Simplex Virus Infection. *Nucleic Acids Res.* **1991**, *19* (20), 5743–5748. https://doi.org/10.1093/nar/19.20.5743.

(34) Shaw, J. pyng; Kent, K.; Bird, J.; Fishback, J.; Froehler, B. Modified Deoxyoligonucleotides Stable to Exonuclease Degradation in Serum. *Nucleic Acids Res.* **1991**, *19* (4), 747–750. https://doi.org/10.1093/nar/19.4.747.
(35) Woolf, T. M.; Jennings, C. G. B.; Rebagliati, M.; Melton, D. A. The Stability, Toxicity and Effectiveness of Unmodified and Phosphorothioate Antisense Oligodeoxynucleotides in Xenopus Oocytes and Embryos. *Nucleic Acids Res.* **1990**, *18* (7), 1763–1769.
(36) Brinkman, H. F.; Jauregui Matos, V.; Mendoza, H. G.; Doherty, E. E.; Beal, P. A. Nucleoside Analogs in ADAR Guide Strands Targeting 5′-UA Sites. *RSC Chem. Biol.* **2022**, *4* (1), 74–83. https://doi.org/10.1039/d2cb00165a.
(37) Bijsterbosch, M. K.; Rump, E. T.; De Vrueh, R. L. A.; Dorland, R.; Van Veghel, R.; Tivel, K. L.; Biessen, E. A. L.; Van Berkel, T. J. C.; Manoharan, M. Modulation of Plasma Protein Binding and in Vivo Liver Cell Uptake of Phosphorothioate Oligodeoxynucleotides by Cholesterol Conjugation. *Nucleic Acids Res.* **2000**, *28* (14), 2717–2725. https://doi.org/10.1093/nar/28.14.2717.
(38) Manoharan, M.; Tivel, K. L.; Andrade, L. K.; Mohan, V.; Condon, T. P.; Bennett, C. F.; Dan Cook, P. Oligonucleotide Conjugates: Alteration of the Pharmacokinetic Properties of Antisense Agents. *Nucleosides and Nucleotides* **1995**, *14* (3–5), 969–973. https://doi.org/10.1080/15257779508012513.
(39) Lorenz, C.; Hadwiger, P.; John, M.; Vornlocher, H. P.; Unverzagt, C. Steroid and Lipid Conjugates of SiRNAs to Enhance Cellular Uptake and Gene Silencing in Liver Cells. *Bioorganic Med. Chem. Lett.* **2004**, *14* (19), 4975–4977. https://doi.org/10.1016/j.bmcl.2004.07.018.
(40) Bhingardeve, P.; Madhanagopal, B. R.; Naick, H.; Jain, P.; Manoharan, M.; Ganesh, K. Receptor-Specific Delivery of Peptide Nucleic Acids Conjugated to Three Sequentially Linked N-Acetyl Galactosamine Moieties into Hepatocytes. *J. Org. Chem.* **2020**, *85* (14), 8812–8824. https://doi.org/10.1021/acs.joc.0c00601.
(41) Shewach, D. S.; Krawczyk, S. H.; Acevedo, O. L.; Townsend, L. B. Inhibition of Adenosine Deaminase by Azapurine Ribonucleosides. *Biochem. Pharmacol.* **1992**, *44* (9), 1697–1700.
(42) Zhang, T.; Yin, C.; Fedorov, A.; Qiao, L.; Bao, H.; Beknazarov, N.; Wang, S.; Gautam, A.; Williams, R. M.; Crawford, J. C.; Peri, S.; Studitsky, V.; Beg, A. A.; Thomas, P. G.; Walkley, C.; Xu, Y.; Poptsova, M.; Herbert, A.; Balachandran, S. ADAR1 Masks the Cancer Immunotherapeutic Promise of ZBP1-Driven Necroptosis. *Nature* **2022**, *606* (7914), 594–602. https://doi.org/10.1038/s41586-022-04753-7.
(43) Kung, C. P.; Cottrell, K. A.; Ryu, S.; Bramel, E. R.; Kladney, R. D.; Bao, E. A.; Freeman, E. C.; Sabloak, T.; Maggi, L.; Weber, J. D. Evaluating the Therapeutic Potential of ADAR1 Inhibition for Triple-Negative Breast Cancer. *Oncogene* **2021**, *40* (1), 189–202. https://doi.org/10.1038/s41388-020-01515-5.
(44) Teoh, P. J.; Chung, T. H.; Chng, P. Y. Z.; Toh, S. H. M.; Chng, W. J. IL6R-STAT3-ADAR1 (P150) Interplay Promotes Oncogenicity in Multiple Myeloma with 1q21 Amplification. *Haematologica* **2020**, *105* (5), 1391–1404. https://doi.org/10.3324/HAEMATOL.2019.221176.
(45) Herbert, A.; Rich, A. The Role of Binding Domains for DsRNA and Z-DNA in the in Vivo Editing of Minimal Substrates by ADAR1. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (21), 12132–12137. https://doi.org/10.1073/pnas.211419898.
(46) Thomas, J. M.; Beal, P. A. How Do ADARs Bind RNA? New Protein-RNA Structures

Illuminate Substrate Recognition by the RNA Editing ADARs. *BioEssays* **2017**, *39* (4), 1–8. https://doi.org/10.1002/bies.201600187.
(47) 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.
(48) 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.
(49) Yi-Brunozzi, H. Y.; Stephens, O. M.; Beal, P. A. Conformational Changes That Occur during an RNA-Editing Adenosine Deamination Reaction. *J. Biol. Chem.* **2001**, *276* (41), 37827–37833. https://doi.org/10.1074/jbc.m106299200.
(50) 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.
(51) 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.
(52) Nose, K.; Hidaka, K.; Yano, T.; Tomita, Y.; Fukuda, M. Short-Chain Guide RNA for Site-Directed A-to-I RNA Editing. *Nucleic Acid Ther.* **2021**, *31* (1), 58–67. https://doi.org/10.1089/nat.2020.0866.
(53) Juliano, R. L. Intracellular Trafficking and Endosomal Release of Oligonucleotides: What We Know and What We Don't. *Nucleic Acid Ther.* **2018**, *28* (3), 166–177. https://doi.org/10.1089/nat.2018.0727.
(54) Shete, H. K.; Prabhu, R. H.; Patravale, V. B. Endosomal Escape: A Bottleneck in Intracellular Delivery. *J. Nanosci. Nanotechnol.* **2014**, *14* (1), 460–474. https://doi.org/10.1166/jnn.2014.9082.
(55) Juliano, R. L.; Wang, L.; Tavares, F.; Brown, E. G.; James, L.; Ariyarathna, Y.; Ming, X.; Mao, C.; Suto, M. Structure-Activity Relationships and Cellular Mechanism of Action of Small Molecules That Enhance the Delivery of Oligonucleotides. *Nucleic Acids Res.* **2018**, *46* (4), 1601–1613. https://doi.org/10.1093/nar/gkx1320.
(56) Janas, M. M.; Jiang, Y.; Schlegel, M. K.; Waldron, S.; Kuchimanchi, S.; Barros, S. A. Impact of Oligonucleotide Structure, Chemistry, and Delivery Method on in Vitro Cytotoxicity. *Nucleic Acid Ther.* **2017**, *27* (1), 11–22. https://doi.org/10.1089/nat.2016.0639.
(57) Samarasinghe, K. T. G.; An, E.; Genuth, M. A.; Chu, L.; Holley, S. A.; Crews, C. M. OligoTRAFTACs: A Generalizable Method for Transcription Factor Degradation. *RSC Chem. Biol.* **2022**, *3* (9), 1144–1153. https://doi.org/10.1039/d2cb00138a.
(58) Pujari, S. S.; Leonard, P.; Seela, F. Oligonucleotides with "Clickable" Sugar Residues: Synthesis, Duplex Stability, and Terminal versus Central Interstrand Cross-Linking of 2′-O-Propargylated 2-Aminoadenosine with a Bifunctional Azide. *J. Org. Chem.* **2014**, *79* (10), 4423–4437. https://doi.org/10.1021/jo500392j.
(59) Gubu, A.; Su, W.; Zhao, X.; Zhang, X.; Fan, X.; Wang, J.; Wang, Q.; Tang, X. Circular Antisense Oligonucleotides for Specific RNase-H-Mediated MicroRNA Inhibition with Reduced Off-Target Effects and Nonspecific Immunostimulation. *J. Med. Chem.* **2021**, *64*

(21), 16046–16055. https://doi.org/10.1021/acs.jmedchem.1c01421.
(60) 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* (07), 319–331. https://doi.org/10.1016/S0076-6879(07)24015-7.
(61) Kabsch, W. Biological Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *D66*, 125–132. https://doi.org/10.1107/S0907444909047337.
(62) 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.
(63) 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 Phenix.Refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68* (4), 352–367. https://doi.org/10.1107/S0907444912001308.

Chapter 5

Structural basis for ADAR editing enhancement at disfavored sites

The crystallography studies with ADAR2 bound to an RNA containing a G:3-deaza-dA pair were conducted with Beal lab member Agya Karki and were part of a highly collaborative project published in Nucleic Acids Research in October 2022[1]*. This chapter contains excerpts from the published manuscript and from a recently accepted article in Accounts of Chemical Research*[2]*. Lastly, the crystallography studies with editing-enabling gRNA sequence motifs discovered through the EMERGe screen*[3] *were also performed with Agya Karki.*

5.1. Introduction

As discussed in Chapter 1, ADARs have sequence preferences that make certain adenosines disfavored for editing, which clearly limits the scope of utilizing these enzymes for therapeutic directed RNA editing applications. For instance, the nearest neighbor nucleotide preferences for ADARs show a strong bias against reaction at adenosines with a guanosine 5' to the target adenosine (5'-G<u>A</u> sites)[4]. This preference is explained by the crystal structures of ADAR2 bound to 8-azaN-containing RNAs that suggest a clash between the 2-amino group of the 5'-G in the minor groove and G489 of the ADAR2 loop involved in stabilizing the flipped-out conformation required for the adenosine deamination reaction[5].

Previously, the Stafforst lab reported that editing at 5'-G<u>A</u> sites by fusion proteins bearing ADAR deaminase domains can be improved by pairing the 5'-G with a G or A[6]. Indeed, our group observed that both 5'-G:G and G:A pairs led to faster *in vitro* deamination rates of both full-length ADAR2 and ADAR1 p110 in comparison to G:U and G:C[1]. We imagined that a purine:purine mismatch near the edit site could destabilize the duplex and facilitate the base-flipping required for the editing reaction. However, since purine:purine combinations have also been shown to exist as stable H-bonded pairs in RNA[7], we also considered that the 5'-G:G and G:A pairs could activate

ADAR by assuming a hydrogen bonded structure that avoids the detrimental steric clash in the minor groove.

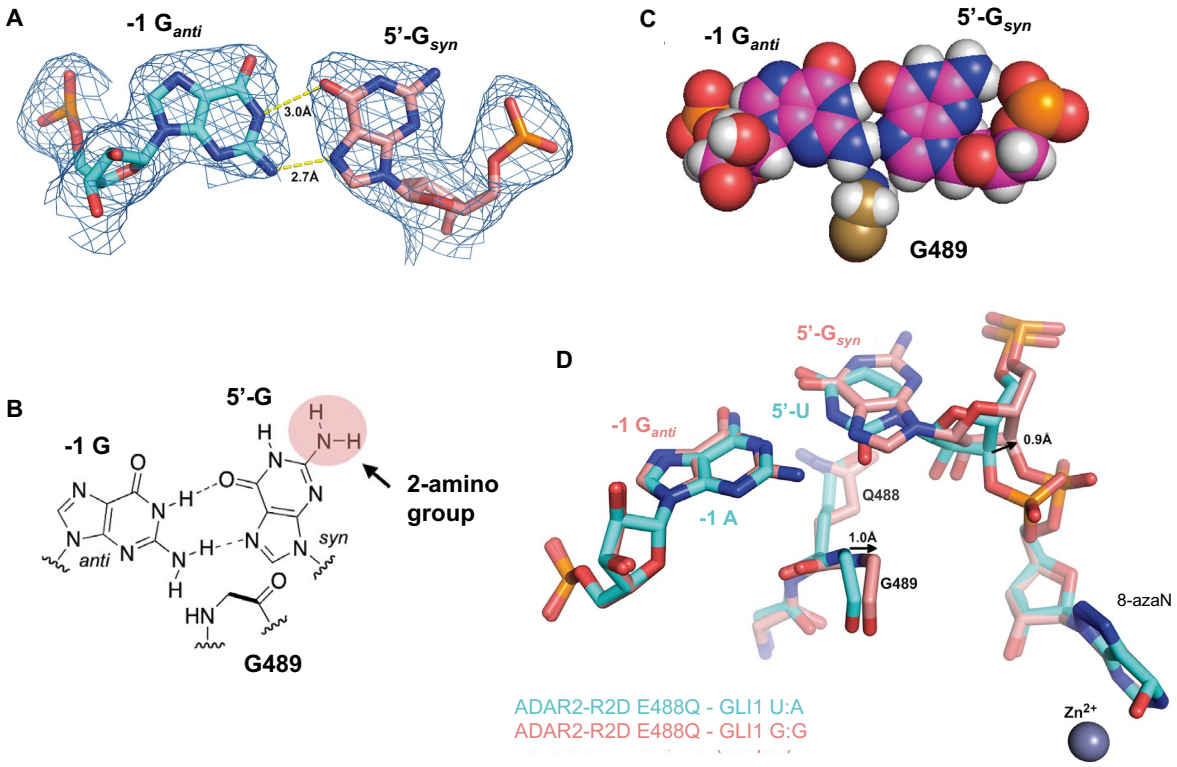


Figure 5.1. The G_{syn}:G_{anti} pair accommodates G489 in the minor groove. (**A**) Fit of a G_{syn}:G_{anti} base pair in the $2Fo - Fc$ electron density map contoured at 1σ. (**B**) Chemical structure of G_{syn}:G_{anti} pair highlighting the location of the 2-amino group of the 5'-G relative to G489. (**C**) Space filling representation of the G_{syn}:G_{anti} pair and its location relative to G489 in the complex. (**D**) Overlay of ADAR2-R2D E488Q structures with RNA bearing either 5'-U paired with A (cyan; PDB 6vff)[8] or 5'-G paired with G (salmon; PDB 8e0f)[1].

In collaboration with the Fisher lab, we turned to X-ray crystallography to test the above hypotheses by introducing a G:G pair adjacent to an editing site in an ADAR2-RNA complex that previously crystallized (ADAR2-R2D E488Q - GLI1 U:A)[8]. Interestingly, in the new structure (ADAR2-R2D E488Q - GLI1 G:G), we found that the 5'-G:G pair adopts a G_{syn}:G_{anti} conformation that is H-bonded between the Hoogsteen face of 5'-G_{syn} and the Watson-Crick face of -1 G_{anti} (Fig. 5.1a)[1]. For simplicity, we call the position across the 5' nucleotide as the -1 position. The *syn*

conformation assumed by 5'-G places its 2-amino group in the major groove, hence eliminating the minor groove clash and enabling the deamination reaction (Fig. 5.1b-c). In addition, a slight shift of the base-flipping loop accommodates the 2-amino group of -1 G_{anti} in the minor groove, compared to previous ADAR2-RNA crystal structures with a 5'-U:A nearest neighbor pair[5,8,9] (Fig. 5.1d).

The observation of a well-defined G_{syn}:G_{anti} pair in the ADAR-RNA complex suggested chemical modifications to the nucleoside paired with the 5'-G might further modulate deamination efficiency. To test this idea, we generated a series of ADAR guide RNAs (gRNAs) that varied at the -1 position to include several different nucleoside analogs. Interestingly, 2'-deoxy-3-deazaadenosine (3-deaza-dA) gave the greatest enhancement in both ADAR2 and ADAR1 p110 deamination rates among the other analogs tested[1]. This inspired us to solve the crystal structure of ADAR2-R2D bound to an RNA duplex containing a 5'-G:3-deaza-dA pair, which is presented in this chapter.

This chapter also contains our group's effort to structurally characterize some unique gRNA sequence motifs that enable editing at 5'-G<u>A</u> sites. These sequences were discovered through En Masse Evaluation of RNA Guides (EMERGe) – a high throughput, next generation sequencing (NGS)-based, and cell-free screening method for ADAR gRNAs for site-directed RNA editing[3]. Beal lab member Casey Jacobsen developed the EMERGe screen, and using this method, he and fellow lab member Prince Salvador identified novel gRNA sequence motifs that induce a potentially corrective edit at the premature termination codon generated by the R255X and R270X mutations in the *MECP2* gene associated with Rett Syndrome[10].

5.2. Results

5.2.1. Crystallography studies with ADAR2 bound to RNA bearing a G:3-deaza-dA pair

As described earlier, the efficient reactions with RNA substrates bearing 3-deaza-dA paired with a 5'-G for both ADAR2 and ADAR1 p110[1] stimulated us to characterize this ADAR-RNA interaction further. Therefore, we prepared a 32 bp duplex derived from GLI1 RNA bearing 8-azaN adjacent to a 5'-G:3-deaza-dA pair for X-ray crystallography with ADAR2-R2D E488Q (Fig. 5.2a). Crystals of this complex formed readily and diffracted X-rays beyond 2.8 Å (Table 5.1). The overall structure of the asymmetric homodimeric protein bound to RNA is very similar to the ADAR2-R2D E488Q - GLI G:G structure (rmsd of 0.201 Å for 698 equivalent α-carbons) with well-resolved electron density for the purine:purine pair adjacent to the 8-azaN (Fig. 5.2b). The G on the 5' side of the 8-azaN is in a *syn* conformation with its Hoogsteen face directed toward the Watson-Crick face of the 3-deaza-dA on the opposing strand. The position of the substrate strand is shifted slightly from that of the structure described above such that, in this structure, the base pair hydrogen bonding involves 3-deaza-dA N1 to G N7 (2.8 Å) and 3-deaza-dA 6-amino to G O6 (3.0 Å) (Fig. 5.2c). This interaction causes the base of 3-deaza-dA to shift slightly towards the minor groove, while also pushing the 5'-G_{syn} out towards the major groove (Fig. 5.2c). This orientation suggests the 3-deaza-dA N1 is protonated to donate a hydrogen bond to N7 forming a G_{syn}:AH^+ anti pair[11] (Fig. 5.2d). This base pair conformation may be further stabilized by the hydrogen bond between the 2-amino group of the 5'-G_{syn} and its 5' phosphodiester oxygen with a 2.6 Å distance compared to 3.4 Å in the G_{syn}:G_{anti} structure (Fig. 5.2b-c).

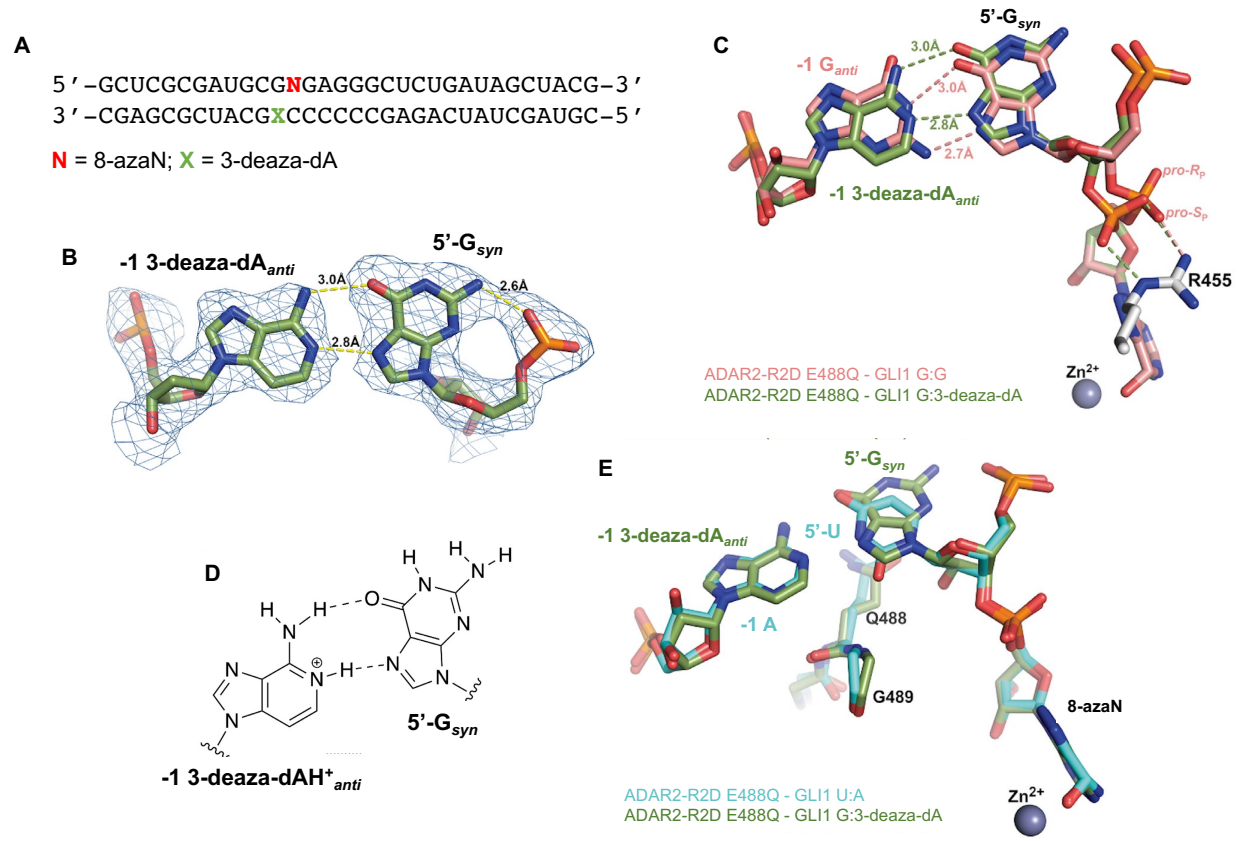


Figure 5.2. X-ray crystal structure of a complex formed between ADAR2-R2D E488Q and (**A**) a 32 bp 8-azaN containing RNA duplex with G:3-deaza-dA pair adjacent to 8-azaN. (**B**) Fit of a G_{syn}:3-deaza-dA_{anti} base pair in the $2Fo - Fc$ electron density map contoured at 1σ. (**C**) Overlay of ADAR2-R2D E488Q structures with RNA bearing either 5'-G paired with G (salmon colored carbons) or 5'-G paired with 3-deaza-dA (green colored carbons). R455 in both structures is identical and shown with white colored carbons. (**D**) The G_{syn}:3-deaza-dAH^{+}_{anti} pair[11]. (D) Overlay of ADAR2-R2D E488Q structures with RNA bearing either 5'-U paired with A (cyan colored carbons) or 5'-G paired with 3-deaza-dA (green colored carbons).

Interestingly, the shift in substrate RNA strand position observed in this G_{syn}:3-deaza-dAH^{+}_{anti} structure allows it to adopt a phosphodiester backbone conformation that is nearly identical to that seen in complexes with RNA bearing ADARs' preferred 5' nearest neighbor (5' U paired with A) (Fig. 5.2e). Notably, the phosphate group between the 5'-G and 8-azaN swings back towards the guanidino group of R455 as seen in all other ADAR2-RNA structures. In the G_{syn}:G_{anti} pair structure, only the *pro*-S_P non-bridging oxygen interacts with R455, while in previous ADAR2-RNA complex structures, both non-bridging oxygens interact with the

guanidino group of R455 (Fig. 5.2c and 5.2e). The standard conformation of this phosphate in the G:3-deaza dA pair might contribute to the observed higher deamination rate relative to the G:G pair[1]. The lack of 2-amino group in the minor groove for the G_{syn}:3-deaza dA_{anti} pair may also help improve activity as it allows the base-flipping loop to adopt a structure similar to what is observed for ADAR bound to RNA with the preferred 5' nearest neighbor pair (5'-U:A pair).

Table 5.1. Data processing and refinement statistics for ADAR2-R2D E488Q bound to GLI G:3-deaza-dA RNA (PDB 8e4x).

Synchrotron (beamline)	APS (24-ID-C)
Wavelength (Å)	0.97918
Space group	*C*2
Unit cell parameters	a = 169.91 Å; b = 63.24 Å; c = 142.65 Å ß = 118.07º
Resolution range (Å)	125-2.80 (2.95-2.80)
No. of observed reflections	99,409 (15,009)
No. of unique reflections	31,644 (4,741)
Completeness (%)	95.3 (98.1)
I/σ (I)	12.0 (1.2)
R_{merge}[a] (%)	6.4 (117.5)
$CC_{1/2}$	0.998 (0.524)
Refinement statistics	
R_{factor}[b] (%)	19.50
R_{free}[b] (%)	23.94
RMS bond length (Å)	0.011
RMS bond angle (º)	1.422
Ramachandran plot statistics[c]	
Favored (%)	94.5
Allowed (%)	5.4
Outliers (%)	0.1
No. of atoms	
Protein	6,558
RNA	1,359
Inositol hexakisphosphate	72
Zn	2
Waters	42

[a] $R_{merge} = [\Sigma_h\Sigma_i|I_h - I_{hi}|/\Sigma_h\Sigma_i I_{hi}]$ where I_h is the mean of I_{hi} observations of reflection h. Numbers in parenthesis represent highest resolution shell.
[b] R_{factor} and $R_{free} = \Sigma||F_{obs}| - |F_{calc}|| / \Sigma|F_{obs}| \times 100$ for 95% of recorded data (R_{factor}) or 5% data (R_{free}).
[c] Ramachandran plot statistics from Molprobity[12].

5.2.2. Crystallography studies with editing-enabling gRNA sequence motifs from EMERGe screen

Rett syndrome is a rare neurological disorder caused by mutations in the *MECP2* gene. Two common non-sense mutations in the *MECP2* gene (*i.e.* R255X, R270X) result in the premature termination of translation of the MECP2 protein, leading to disease manifestation[10]. The resulting stop codon in the R255X and R270X mutations is a potential site for a corrective edit using ADARs (5'-UGA). However, the edit site is in a non-ideal 5' nearest neighbor context (5'-G)[5]. To gain insight into improving editing in this poorly edited sequence, Casey Jacobsen and Prince Salvador used the EMERGe screening approach, where they identified unique gRNA sequence motifs (S5 / 3'-GUG, 3'-CGG; Fig. 5.3) that enable ADAR2 editing at the 5'-GA sites in the MECP2 R255X and R270X transcripts[3].

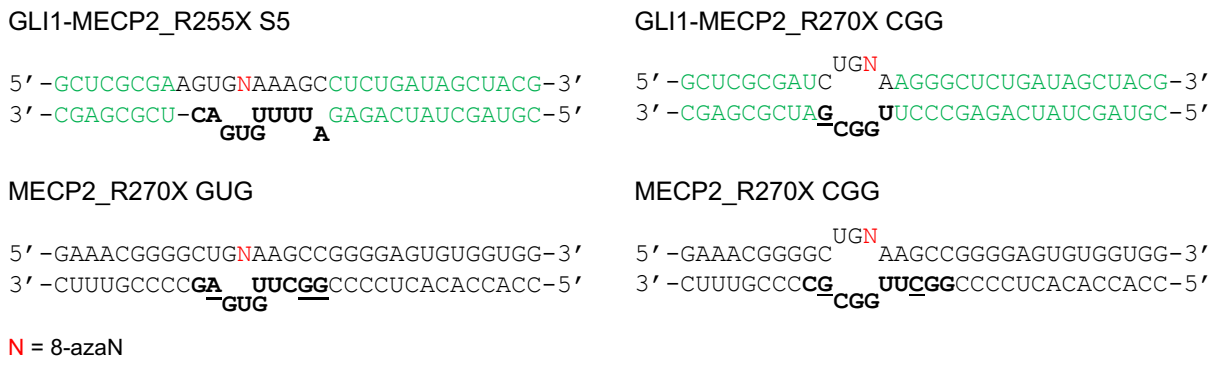


Figure 5.3. Candidate 8-azaN-modified RNA duplexes for EMERGe gRNA-bound ADAR2-R2D crystallization trials. Sequences derived from GLI1 and MECP2 are in green and black, respectively. Sequences derived from the EMERGe screen are bolded. Base changes to achieve higher complementarity are underlined.

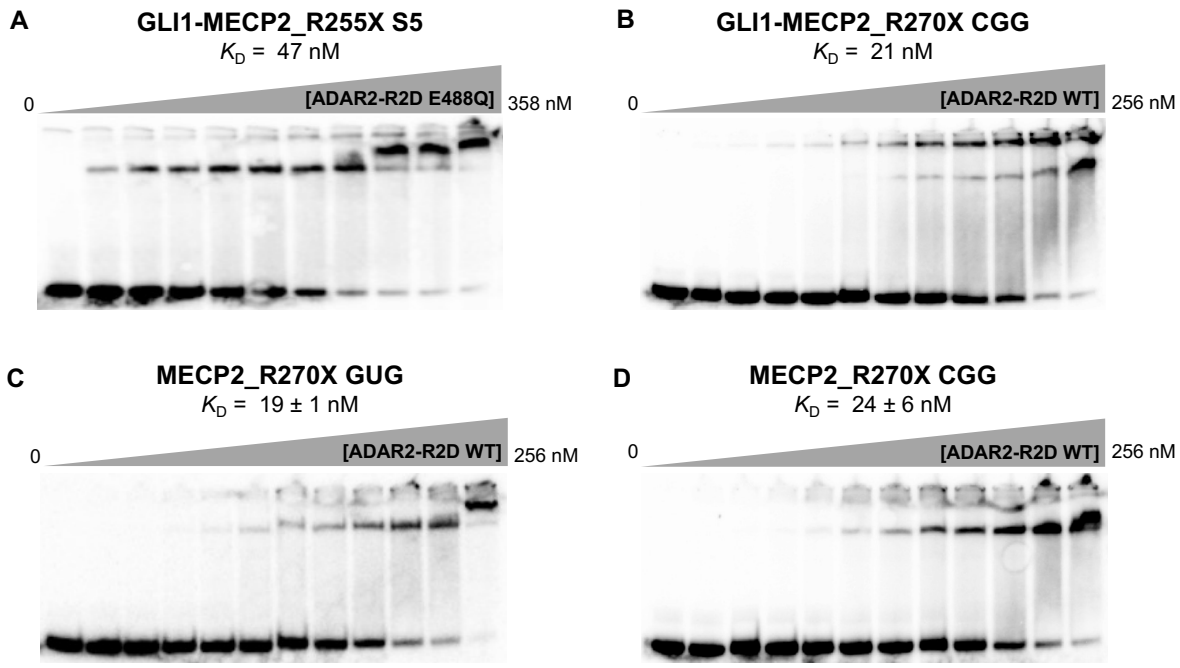


Figure 5.4. Representative gel shifts of ADAR2-R2D with the candidate 8-azaN-modified RNA duplexes for crystallization trials. (**A**) GLI1-MECP2_R255X S5; (**B**) GLI1_MECP2_R270X CGG; (**C**) MECP2_R270X GUG; (**D**) MECP2_R270X GUG. Gel shift assays were performed at 0 to 358 nM ADAR2-R2D E4888Q for A and 0 to 256 nM ADAR2-R2D WT for B to D at 5 nM RNA duplex. Data were plotted to the equation: $y = A \times [x/(K_D + x)]$ where y is fraction RNA bound, x is [ADAR2-R2D]; A is binding endpoint; and K_D is dissociation constant. Values reported are the average of three independent measurements ± standard deviation.

We then attempted to use X-ray crystallography to probe the structural basis for how these sequence motifs enable editing at 5'-G<u>A</u> sites. We first started with setting up crystallization screens with a chimeric sequence (GLI-MECP2_R255X S5) derived from GLI1 RNA, which was previously shown to crystallize[5,8,9,13], and MECP2 R255X RNA. The duplex also contains the 10 nt editing-enabling gRNA sequence (S5) discovered through the EMERGe screen for the MECP2 R255X transcript conducted by Prince Salvador (Fig. 5.3). To gain some insight into how well this duplex will bind to the protein for crystallization, Agya Karki first performed gel shift assays with this duplex and the ADAR2-R2D E488Q protein, which was previously shown to crystallize[1,8] (Fig. 5.4a). The GLI-MECP2_R255X S5 duplex gave a K_D of 47 nM with ADAR2-R2D E488Q,

which is ~40-fold higher than that of the original GLI1 sequence present in our high-resolution ADAR2-RNA crystal structures[1,8]. However, given that the conditions in which the gel shift assays were carried out may not be exactly the same as those used for crystal growth, we still set up crystallization screens for the ADAR2-R2D E488Q - GLI-MECP2_R255X S5 complex (Table 5.2). Since a significant amount of protein was lost during the His-tag cleavage step, only a few crystallization screens were set up for the GLI-MECP2_R255X S5 duplex. Unfortunately, none of the conditions yielded any crystal hits.

Table 5.2. Screening conditions used for EMERGe gRNA-bound ADAR2-R2D crystallization trials.

dsRNA (50 µM)	GLI1-MECP2_R255X S5	MECP2_R270X GUG	MECP2_R270X CGG
ADAR2-R2D (100 µM)	WT	E488Q	E488Q
96-well screen (25 ºC)	Nuc-Pro HTS	Nuc-Pro HTS, Natrix HT, MCSG2, PACT premier HT-96	Nuc-Pro HTS, Natrix HT, MCSG2, PACT premier HT-96
Optimization screen (25 ºC)	Nuc-Pro HTS F5	Nuc-Pro HTS F5	Nuc-Pro HTS F5

Next, we tried three other candidate duplexes for crystallography studies (GLI-MECP2_R270X CGG; MECP2_R270X GUG; MECP2_R270X CGG; Fig. 5.3). Apart from the 3'-CGG motif, an EMERGe screen for the MECP2 R270X transcript conducted by Casey Jacobsen also selected a similar sequence motif as S5 (3'-GUG). Therefore, the new candidate duplexes for crystallography are derived fully or partially from the MECP2 R270X RNA. A few bases in the 10 nt gRNA sequence from the screen were also modified to increase complementarity which could aid in duplex stability and crystallization (Fig. 5.3). In collaboration with Agya Karki, these new duplexes were first examined through gel shift assays with ADAR2-R2D WT (Fig. 5.4b-c). Although this protein has not yet been shown to crystallize, *in vitro* deamination studies

performed by Prince Salvador and Casey Jacobsen displayed significantly better editing rates with the WT protein compared to the E488Q protein on substrates containing the 3'-GUG and 3'-CGG motifs[3]. We found that the K_D values for the GLI-MECP2_R255X CGG, MECP2_R270X GUG, and MECP2_R270X CGG duplexes with ADAR2-R2D WT are comparable (~20 nM) and more promising than that of the ADAR2-R2D E488Q - GLI-MECP2_R255X S5. Moreover, the duplexes whose sequences were derived solely from MECP2_R270X showed less aggregation of the ADAR2-RNA complex in the PAGE wells during the gel shift assay, motivating us to set up crystallization screens with the ADAR2-R2D - MECP2_R270X GUG and - MECP2_R270X CGG combinations.

Table 5.3. Growth conditions for crystal hits and cryo-protectants used for crystal mounting.

dsRNA	MECP2_R270X GUG	MECP2_R270X CGG
Growth conditions + cryo-protectant	(1) Nuc-Pro HTS F4 + 30% ethylene glycol (2) PACT premier HT-96 B2 + 30% ethylene glycol	(1) Nuc-Pro HTS G8 + 30% MPD[a] (2) Nuc-Pro HTS H1 (already contains 50% MPD) (3) Nuc-Pro HTS H4 + 30% ethylene glycol (4) Natrix HT D9 + 30% ethylene glycol

[a] 2-methyl-2,4-pentanediol

For this batch of protein purification, we were able to recover a significantly higher amount of protein using freshly purified TEV protease during the His-tag cleavage step. This allowed us to set up multiple screening conditions for crystallization (Table 5.2) and a few conditions led to crystal growth for both duplexes (Table 5.3). The crystals were harvested and are currently stored in liquid N_2 for data collection scheduled later this year.

5.3. Discussion

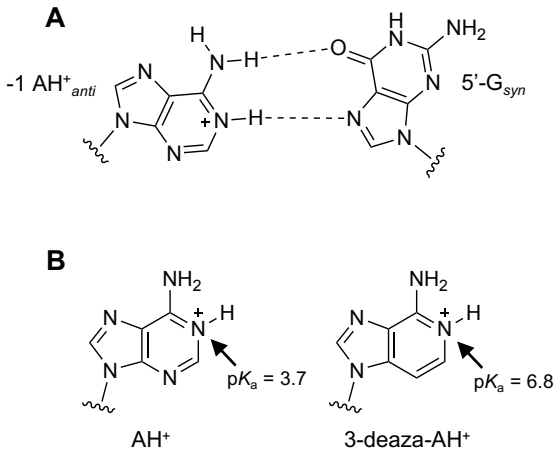


Figure 5.5. (**A**) Predicted base-pairing interaction in a 5'-G_{syn}:AH^+_{anti} pair. (**B**) The pK_a of 3-deaza-AH^+ N1H is higher than that of AH^+. Estimated pK_a values derived from literature[14].

As described earlier, ADAR editing at 5'-G<u>A</u> sites can be activated by pairing the 5'-G with a -1 G or A. The G_{syn}:G_{anti} structure assumed by the 5'-G:G pair in the ADAR2-R2D E488Q - GLI1 G:G crystal structure offered a convincing explanation for the editing enhancement at the disfavored 5'-G<u>A</u> site. The increase in editing rates observed by both the Stafforst group and our group for 5'-G:A pair compared to G:U and G:C pairs can also be rationalized by the formation of a G_{syn}:AH^+_{anti} interaction (Fig. 5.5a). To facilitate this conformation, N1 of -1 A_{anti} should be protonated to effectively donate a hydrogen bond to N7 of 5'-G_{syn}. The importance of this N1 protonation is demonstrated by the significantly improved editing rates observed with 3-deaza-2'-deoxyadenosine (3-deaza-dA) when placed at the -1 position of a gRNA targeting a 5'-G<u>A</u> site[1]. The substantially higher pK_a for N1 protonation in 3-deazaadenosine suggests that this site is more likely protonated at physiologically relevant pH compared to A (Fig. 5.5b). It is important to note that a -1 dA gRNA also led to ~2.5-fold faster ADAR2 editing rates compared to a -1 A gRNA as

expected based on the preferred sugar pucker conformation at this site[5,15]. However, a -1 3-deaza-dA remarkably showed ~8.5-fold higher rate of ADAR2 deamination compared to a -1 A gRNA, highlighting the beneficial effect of the increased N1H pK_a in the base modification.

The crystallography studies with ADAR2-R2D E488Q - GLI1 G:G or G:3-deaza-dA strongly supported the observed deamination kinetics at 5'-G<u>A</u> sites using gRNAs with variable canonical or chemically modified nucleosides at the -1 position[1]. Together, these structural and kinetic studies identified an approach to facilitate ADAR editing at challenging 5'-G<u>A</u> sites. In general, we found that the use of nucleosides capable of hydrogen bonding to the Hoogsteen face of the 5' G and inducing a *syn* conformation at this location in the RNA without also introducing additional sterically demanding groups into the minor groove enables efficient editing at these sites.

Currently, we have not yet solved any high-resolution crystal structures from our crystallography studies with ADAR2-R2D in complex with the EMERGe screen-derived editing-enabling gRNA motifs. However, the presence of a 5'-G:G pair in both 3'-GUG and 3'-CGG sequences suggests that a favorable G_{syn}:G_{anti} interaction 5' to the edit site could be contributing to the editing enhancements observed with these gRNAs. Crystal structures could help identify the orphan base interactions in both motifs, which may be unique given the unexpected trend in deamination efficiency observed between the WT protein and the hyperactive E488Q mutant (better editing rates with WT). Nevertheless, these crystallization attempts showed us that we can obtain reproducible ADAR2-R2D - GLI1 RNA crystals even with the WT protein at 50 mM MOPS pH 7.0, 100 to 200 mM NaCl, and 13 to 20% PEG 4000, as ADAR2-R2D WT - GLI1 G:3-deaza-dA crystals of similar morphology as ADAR2-R2D E488Q - GLI1 G:3-deaza-dA also grew under these specified conditions (control set up). Given that reproducible protein-RNA crystals can be

obtained with the GLI1 RNA, it could be helpful to try setting up crystallization screens with this RNA containing minimal modifications to accommodate the 3'-GUG or 3'-CGG motifs.

In addition, we have also established that using freshly purified TEV protease (stored in -70 °C for less than a week) during the His-tag cleavage step results in higher yields of the cleaved protein. It is also important to note that the cleavage reaction must be performed at NaCl concentrations not exceeding 200 mM as high salt concentrations inhibit TEV protease[16]. Buffer exchanging or diluting the cleavage reaction to bring down glycerol concentration to ~5% (v/v) post-cleavage could also improve separation of cleaved and uncleaved protein, eliminating the need for another round of Ni-NTA purification (see Section 5.4.4). Future crystallization attempts may also consider removing the His-tag cleavage step from the ADAR purification scheme and setting up crystallization screens with the tagged protein. These could substantially shorten the required time for ADAR purification and TEV protease would also no longer have to be expressed and purified. Together, these modifications to the existing procedure could make the ADAR purification process for crystallography studies more higher yielding and more efficient, allowing for faster set up of crystallization screens for multiple ADAR-RNA complexes under multiple screening conditions in one protein purification batch.

5.4. Methods

5.4.1. Synthesis and purification of oligonucleotides. Chemical synthesis for all oligonucleotides was performed using an ABI 394 synthesizer in collaboration with Agya Karki, Victorio Jauregui Matos, Casey Jacobsen, and Prince Salvador. All protected phosphoramidites were purchased from Glen Research except the 8-azaN phosphoramidite which was purchased from Berry & Associates or synthesized by Victorio Jauregui Matos as previously described[17]. Nucleosides were incorporated during the appropriate cycle on a 0.2 or 1.0 μmol scale. Table 5.4 shows the sequences

of all oligonucleotides used in this chapter. Upon completion of the synthesis, columns were evaporated under reduced pressure for 12 h. All oligonucleotides were cleaved from the solid support by treatment with 1.5 mL 1:3 ethanol/30% NH_4OH at 55 ºC for 12 h. The supernatant was transferred to a new screw-cap tube and evaporated under reduced pressure. For all oligonucleotides except the 8-azaN-modified strand, desilylation was performed by treating the pellets with 250 μL of 1M TBAF–THF at room temperature overnight. For the 8-azaN strand, desilylation was carried out in TEA•3HF as previously described[17]. To each reaction was added 75 mM sodium acetate in butanol. The oligonucleotides were then precipitated from a solution of 65% butanol at –70 ºC for 2 h. The solution was centrifuged at 17,000 × g 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.

Oligonucleotide purification was performed in collaboration with Agya Karki, Victorio Jauregui Matos, Casey Jacobsen, and Prince Salvador. Single-stranded RNA oligonucleotides were purified by denaturing PAGE and visualized by UV shadowing. Bands were excised from the gel, crushed and and soaked overnight at 4 ºC in 0.5 M NaOAc, 0.1% (w/v) 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 ºC for 12 h. The solution was centrifuged at 17,000 × g 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. Mass spectrometry data can be found in Table 5.5.

5.4.2. Gel shift assay. The unmodified RNA strands were purchased from IDT or Horizon Dharmacon and purified as described above. The 5'-end of top strands containing 8-azaN were

labeled using [γ-^{32}P]-ATP (6000 Ci/mmol) and T4 polynucleotide kinase (NEB). The labeled reaction was passed through a G-25 column to remove excess ATP and further purified with a 19% denaturing PAGE gel at 10 W for 8 h. The labeled oligonucleotide was visualized by storage phosphor autoradiography and the gel band containing it was excised, crushed, soaked and ethanol precipitated as described for other gel purified oligonucleotides above. The labeled top strand was diluted to a stock solution of approximately 300 nM and hybridized to its compliment at 1:3 ratio in 1X TE buffer, pH 7.5 and 200 mM NaCl by heating at 95 °C for 5 min and slowly cooling to 30 °C to a final concentration of approximately 50 nM. Samples containing ≤ 1 nM RNA and varying concentrations of the protein (260, 130, 65, 32.5, 16.25, 8.1, 4.1, 2.0, 1.0, 0.5, 0.25 and 0 nM) were incubated together in 20 mM Tris-HCl pH 7.0, 3.5% (v/v) glycerol, 0.5 mM DTT, 60 mM KCl, 20 mM NaCl, 0.1 mM BME, 1.5 mM EDTA, 0.003% NP-40, 160 U/μL RNase inhibitor, 100 μg/mL BSA, and 1 μg/mL yeast tRNA for 30 min at 30 °C. Samples were loaded onto a 6% native PAGE gel (79:1 acrylamide:*bis*-acrylamide) and electrophoresed under nondenaturing conditions in 1× TBE buffer at 4 °C for 90 min. The gels were dried on a BioRad gel drier for 90 min at 80 °C under vacuum followed by exposure to storage phosphor imaging plates (Kodak) for 24 h at room temperature in the dark. After exposure, the gels were removed, and the phosphor imaging plates were scanned by Typhoon Trio Variable Mode Imager (GE Healthcare). Dissociation constants were measured by calculating the fraction of RNA bound by the protein and using the equation: $y = A \times [x/(K_D + x)]$ where y is fraction RNA bound, x is [protein]; A is binding endpoint; and K_D is dissociation constant. The human ADAR2-R2D WT or E488Q proteins used for this assay were expressed and purified as described in Section 5.4.4 except without undergoing TEV protease treatment, post Ni-NTA purification, and gel filtration

chromatography. The ^{32}P labeling reaction and gel shift assay with the GLI1-MECP2_R255X S5 and MECP2_R270X GUG RNA duplexes were performed by Agya Karki.

5.4.3. Expression and purification of TEV protease. TEV protease expression and purification were carried out based on a previously reported protocol[18] and were performed in collaboration with Agya Karki. A pLysS vector containing TEV protease catalytic domain S219V with an N-terminal maltose binding protein (MBP) tag, TEV cleavage site, and His_7-tag, was transformed into *E. coli* BL21 (DE3) cells according to published protocol[18]. A single colony was used to inoculate 5 mL LB + ampicillin (100 μg/mL) and shaken at 250 rpm at 37 °C overnight. The 5 mL culture was then used to inoculate 1 L LB + ampicillin at 37 °C until OD_{600} was approximately 0.6 (3-4 h). The culture was cooled to 30 ºC and incubated at this temperature for 30 min. Cells were induced by addition of 1 mL of 1 M isopropyl β-*D*-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and expressed for 3 h at 30 °C while shaking at 250 rpm. Cells were then harvested by centrifugation at 5,000 x g for 30 min and stored at -70 °C prior to purification. Cells were lysed in TEV lysis buffer (20 mM Tris-HCl pH 8.0, 5 mM imidazole, 1000 mM NaCl, 5% (v/v) glycerol) using a microfluidizer. Cell lysate was clarified by centrifugation (39,000 x g for 30 min at 4 °C) and the supernatant was passed through a 0.45 μm filter before loading onto a 5 mL Ni-NTA agarose column (Qiagen) pre-equilibrated in TEV lysis buffer. The column was washed in three steps with 50 mL of TEV wash I buffer (20 mM Tris-HCl pH 8.0, 35 mM imidazole, 1000 mM NaCl, 5% (v/v) glycerol), TEV wash II buffer (20 mM Tris-HCl pH 8.0, 35 mM imidazole, 500 mM NaCl, 5% (v/v) glycerol), and TEV wash III buffer (TEV wash II + 1mM BME). Protein was eluted with a 35-300 mM imidazole gradient in TEV wash III buffer over 80 min at a flow rate of 1 mL/min. Fractions containing TEV protease were pooled and buffer exchanged into 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% (v/v) glycerol, and 1 mM βME using

a 10,000 MW cutoff centrifugal concentrator (Amicon). The sample was concentrated to ~2 mg/mL and diluted to 1 mg/mL with 20 mM Tris-HCl pH 8.0, 350 mM NaCl, and 70% (v/v) glycerol so that the final glycerol concentration is at ~50% (v/v). The sample was aliquoted into 5 mL tubes and stored at -70 °C.

5.4.4. Expression and purification of human ADAR2-R2D for crystallography. Protein expression and purification were carried out in collaboration with Agya Karki by modifying a previously reported protocol[19]. *S. cerevisiae* BCY123 cells were transformed with a pSc-ADAR construct encoding human ADAR2-R2D WT or E488Q (corresponding to residues 214–701) with an N-terminal His_{10}-tag. Cells were streaked on yeast minimal media minus uracil (CM-ura) plates. A single colony was used to inoculate a 20 mL CM-ura starter culture. After cultures were shaken at 300 rpm and 30 ºC overnight, 10 mL of starter culture was used to inoculate each liter of yeast growth medium. After cells reached an OD_{600} between 1-2 (~20-24 h) cells were induced with 110 mL of sterile 30% (w/v) galactose per liter and protein was expressed for 6 h. Cells were collected by centrifugation at 5,000 × g for 10 min and stored at –70 ºC. Cells were lysed in 750 mM NaCl in buffer A (20 mM Tris-HCl pH 8.0, 35 mM imidazole, 5% (v/v) glycerol, 1 mM BME, and 0.01% Triton X-100) with 750 mM NaCl using a microfluidizer. Cell lysate was clarified by centrifugation (39,000 × g for 25 min). Lysate was passed over a 5 mL Ni-NTA column equilibrated with lysis buffer, which was then washed in three steps with 50 mL of lysis buffer, wash I buffer (buffer A + 300 mM NaCl), and wash II buffer (buffer A + 100 mM NaCl). Protein was eluted with a 35-300 mM imidazole gradient in wash II buffer over 80 min at a flow rate of 1 mL/min. Fractions containing target protein were pooled and further purified on a 5 mL GE Healthcare Lifesciences Hi-Trap Heparin HP column in wash II buffer. The protein was washed with 50 mL of wash II buffer and eluted with a 100-1000 mM NaCl gradient over 60 min at a flow

rate of 0.8 mL/min. Fractions containing target protein were pooled and NaCl concentration was brought down to ~200 mM by dilution with wash II buffer without salt. The His_{10}-tag was cleaved from ADAR2-R2D with an optimized ratio of 1 mg of TEV protease per 1 mg of protein. Cleavage was carried out for 4 h at room temperature without agitation, followed by overnight cleavage at 4 ºC before the product was passed over another Ni-NTA column at a flow rate of 0.5 mL/min. The flow through and wash were collected and passed through another Ni-NTA column to remove remaining uncleaved protein. The flow through and wash were collected, dialyzed against 20 mM Tris pH 8.0, 200 mM NaCl, 5% (v/v) glycerol, and 1 mM BME, followed by concentration to just under 1 mL for gel filtration on a GE Healthcare HiLoad 16/600 Superdex 200pg column. Fractions containing purified protein were pooled and concentrated to 7-9 mg/mL for crystallization trials.

5.4.5. Preparation of duplex RNA substrates for crystallography. For crystallography, the unmodified RNA strands were purchased from IDT or Horizon Dharmacon and gel-purified as described above. As in previous structures, the edited strands contained the adenosine analog 8-azaN at the editing site. Duplex RNA was hybridized in water in a 1:1 ratio by heating to 95 ºC for 5 min and slow cooling to 30 ºC. This step was performed in collaboration with Agya Karki.

5.4.6. Crystallization of the human ADAR2-R2D-RNA complex. Crystals of the human ADAR2-R2D E488Q - GLI1 G:3-deaza-dA RNA complex were grown at room temperature by the sitting-drop vapor-diffusion method. A solution of 1.0 µL volume containing 100 µM protein and 50 µM GLI1 G:3-deaza-dA RNA duplex was mixed with 1.0 µL of 50 mM MOPS pH 7.0, 100 mM NaCl, and 13% PEG 4000. Flat rhomboid-shaped crystals took about 8 days to grow to

100 μm. A single crystal was soaked briefly in a solution of mother liquor plus 30% glycerol before flash cooling in liquid nitrogen.

Crystals of EMERGe gRNA-bound ADAR2-R2D were grown in either hanging- or sitting-drop vapor diffusion method under the screening conditions listed in Table 5.2. Crystals took about a month to grow to about 100 μm and were flash cooled in liquid nitrogen under the freezing conditions listed in Table 5.3. Crystal growth trays for ADAR2-R2D E488Q - GLI1 G:3-deaza-dA and - GLI1-MECP2_R255X S5 RNA complexes were set up in collaboration with Agya Karki, with the assistance of Fisher lab member Jeff Cheng. Preparation of freezing buffers were performed with the assistance of Jeff Cheng. Crystal growth trays for ADAR2-R2D WT - MECP2_R270X GUG and - MECP2_R270X CGG RNA complexes were set up in collaboration with Agya Karki, with the assistance of Beal lab members Kristen Campbell and Prince Salvador. Preparation of freezing buffers were performed with Prince Salvador.

5.4.7. Processing and refinement of crystallographic data. X-ray diffraction data for the human ADAR2-R2D E488Q - GLI1 G:3-deaza-dA RNA complex were collected at the Advanced Photon Source (APS). Diffraction data were collected on the 24-ID-C beamline to 2.8 Å resolution. The dataset was processed with XDS[20] and scaled with AIMLESS[21]. The structures were determined by molecular replacement using PHENIX[22]. The previous ADAR2-R2D E488Q GLI1-bound crystal structure[8] (PDB 6vff) was used as the phasing model. The structures were refined with PHENIX including non-crystallographic symmetry (NCS) restraints at the start, but relaxed at the final stages resulting in a lower R_{free}. Additionally, RNA base-stacking and base-pair restraints, when appropriate, were also imposed in refinement. Table 5.1 shows statistics in data processing and model refinement. As observed in the previous ADAR2-R2D E488Q GLI1-bound crystal structure, the asymmetric unit includes two protein monomers assembled as an asymmetric

homodimer complexed with 32 bp RNA duplex. Both complexes displayed similar overall structures. The dsRBD (residues 215–315) of monomer A was disordered and therefore not included in the model. The first 20 residues (215–234) and part of the 5' RNA binding loop, residues 464–475, of monomer B were disordered and not included in the model. The dsRBD of monomer B interacts with the 3' end of the dsRNA relative to the 8-azaN.

5.5. Table of oligonucleotides

Table 5.4. Sequences of oligonucleotides used for work presented in Chapter 5. All oligonucleotides are made up of ribonucleotides except when specified. TS = target strand, CS = complement strand, N = 8-azanebularine, 3dA = 2'-deoxy-3-deazaadenosine.

Oligonucleotide	Sequence (5' to 3')
GLI1 G:3-deaza-dA TS	GCUCGCGAUGCG(N)GAGGGCUCUGAUAGCUACG
GLI1 G:3-deaza-dA CS	CGUAGCUAUCAGAGCCCCCC(3dA)GCAUCGCGAGC
GLI1-MECP2_R255X S5 TS	GCUCGCGAAGUG(N)AAAGCCUCUGAUAGCUACG
GLI1-MECP2_R255X S5 CS	CGUAGCUAUCAGAGAUUUUGUGACUCGCGAGC
MECP2_R270X GUG TS	GAAACGGGGCUG(N)AAGCCGGGGAGUGUGGUGG
MECP2_R270X GUG CS	CCACCACACUCCCCGGCUUGUGAGCCCCGUUUC
GLI1-MECP2_R270X CGG TS	GCUCGCGAUCUG(N)AAGGGCUCUGAUAGCUACG
GLI1-MECP2_R270X CGG CS	CGUAGCUAUCAGAGCCCUUGGCGAUCGCGAGC
MECP2_R270X CGG TS	GAAACGGGGCUG(N)AAGCCGGGGAGUGUGGUGG
MECP2_R270X CGG CS	CCACCACACUCCCCGGCUUGGCGCCCCGUUUC

Table 5.5. Calculated and observed masses of oligonucleotides used for work presented in Chapter 5. TS = target strand, CS = complement strand.

Oligonucleotide	Calculated Mass (g/mol)	Observed Mass (m/z)
GLI1 G:3-deaza-dA TS	10315.4	10320.2
GLI1 G:3-deaza-dA CS	10174.4	10177.9
GLI1-MECP2_R255X S5 TS	10267.4	10272.7
GLI1-MECP2_R255X S5 CS	10236.3	10234.6
MECP2_R270X GUG TS	10537.4	10541.8
MECP2_R270X GUG CS	10347.4	10352.5
GLI1-MECP2_R270X CGG TS	10260.3	10259.8
GLI1-MECP2_R270X CGG CS	10249.4	10249.4
MECP2_R270X CGG TS	10537.4	10541.8
MECP2_R270X CGG CS	10017.3	10012.3

5.6. References

(1) Doherty, E. E.; Karki, A.; Wilcox, X. E.; Mendoza, H. G.; Manjunath, A.; Matos, V. J.; Fisher, A. J.; Beal, P. A. ADAR Activation by Inducing a Syn Conformation at Guanosine Adjacent to an Editing Site . *Nucleic Acids Res.* **2022**, *50* (19), 10857–10868. https://doi.org/10.1093/nar/gkac897.

(2) Mendoza, H. G.; Beal, P. A. Chemical Modifications in RNA : Elucidating the Chemistry of DsRNA-. *Acc. Chem. Res.* **2023**. https://doi.org/10.1021/acs.accounts.3c00390.

(3) Jacobsen, C. S.; Salvador, P.; Yung, J. F.; Kragness, S.; Mendoza, H. G.; Mandel, G.; Beal, P. A. Library Screening Reveals Sequence Motifs That Enable ADAR2 Editing at Recalcitrant Sites. *ACS Chem. Biol.* **2023**. https://doi.org/10.1021/acschembio.3c00107.

(4) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 1–9. https://doi.org/10.1038/ncomms1324.

(5) 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.

(6) Schneider, M. F.; Wettengel, J.; Hoffmann, P. C.; Stafforst, T. Optimal GuideRNAs for Re-Directing Deaminase Activity of HADAR1 and HADAR2 in Trans. *Nucleic Acids Res.* **2014**, *42* (10), 1–9. https://doi.org/10.1093/nar/gku272.

(7) Leontis, N. B.; Westhof, E. Geometric Nomenclature and Classification of RNA Base Pairs. *Rna* **2001**, *7* (4), 499–512. https://doi.org/10.1017/S1355838201002515.

(8) Thuy-Boun, A. S.; Thomas, J. M.; Grajo, H. L.; Palumbo, C. M.; Park, S.; 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), 7958–7952. https://doi.org/10.1093/nar/gkaa532.

(9) 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.

(10) Krishnaraj, R.; Ho, G.; Christodoulou, J. RettBASE: Rett Syndrome Database Update. *Hum. Mutat.* **2017**, *38* (8), 922–931. https://doi.org/10.1002/humu.23263.

(11) Pan, B.; Mitra, S. N.; Sundaralingam, M. Crystal Structure of an RNA 16-Mer Duplex R (GCAGAGUUAAAUCUGC) 2 with Nonadjacent G(Syn). *Biochemistry* **1999**, *38* (9), 2826–2831.

(12) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci.* **2018**, *27* (1), 293–315. https://doi.org/10.1002/pro.3330.

(13) 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.

(14) Krishnamurthy, R. Role of PKa of Nucleobases in the Origins of Chemical Evolution. *Acc. Chem. Res.* **2012**, *45* (12), 2035–2044. https://doi.org/10.1021/ar200262x.

(15) Brinkman, H. F.; Jauregui Matos, V.; Mendoza, H. G.; Doherty, E. E.; Beal, P. A.

Nucleoside Analogs in ADAR Guide Strands Targeting 5′-UA Sites. *RSC Chem. Biol.* **2022**, *4* (1), 74–83. https://doi.org/10.1039/d2cb00165a.
(16) Nallamsetty, S.; Kapust, R. B.; Tözsér, J.; Cherry, S.; Tropea, J. E.; Copeland, T. D.; Waugh, D. S. Efficient Site-Specific Processing of Fusion Proteins by Tobacco Vein Mottling Virus Protease in Vivo and in Vitro. *Protein Expr. Purif.* **2004**, *38* (1), 108–115. https://doi.org/10.1016/j.pep.2004.08.016.
(17) 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.
(18) Tropea, J. E.; Cherry, S.; Waugh, D. S. Expression and Purification of Soluble His6-Tagged TEV Protease. *Methods Mol. Biol.* **2009**, *498* (1), 297–307. https://doi.org/10.1007/978-1-59745-196-3_19.
(19) 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* (07), 319–331. https://doi.org/10.1016/S0076-6879(07)24015-7.
(20) Kabsch, W. Biological Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *D66*, 125–132. https://doi.org/10.1107/S0907444909047337.
(21) Evans, P. R.; Murshudov, G. N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69* (7), 1204–1214. https://doi.org/10.1107/S0907444913000061.
(22) 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.