Title

Synthesis of Modified Oligonucleotides for Site-Directed RNA Editing and Structure-Activity Relationships of RNA Binding Enzymes

Author

Jauregui Matos, Victorio

Publication Date

2024

Peer reviewed|Thesis/dissertation

Synthesis of Modified Oligonucleotides for Site-Directed RNA Editing and Structure- Activity Relationships of RNA Binding Enzymes

By

VICTORIO JAUREGUI-MATOS

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Peter A. Beal, Chair

_____

Sheila S. David

_____

Jared T. Shaw

Committee in Charge

2024

ACKNOWLEDGMENTS

I dedicate this dissertation primarily to my Cuban grandfather (Wigberto Jauregui Echemendia) and grandmother (Idolidia Castro Aguilera) who raised me after emigrating from Cuba to the U.S.A. and are looking from above. To my mother, father, and sister for ensuring I have the best education and growth and for their unconditional love. To all my best friends, past and present, wherever they are in the world, for listening to me, grounding me, and keeping me sane. Specifically, to those around me in Davis: Margaret Berrens, Amanda Caceres, Prince Salvador, and Rose Smiley. To all the teachers and professors who have paved the way for me to grow academically and personally.

I want to acknowledge all the support from my supervisor Peter A. Beal and for his guidance in forming all the chapters herein. I want to thank him for his continuous support in my professional career, for believing in me, and for providing the financial support to perform research. I would also like to thank Andrew J. Fisher for his collaborative help in most of these projects and his experience with crystallography. I want to acknowledge my thesis committee Jared Shaw and Sheila David for graduate education and continuous support.

I want to acknowledge all my peers and collaborators for all the synergistic work done on all the chapters and publications. These include Hannah Brinkman, Herra Mendoza, Sukanya Mozumder, Prince Salvador, Aashrita Manjunath, Randall Ouye, Tyson Vu, Natalie Wade, Agya Karki, Erin E. Doherty, Jeff Cheng, Kristen Campbell, and Yanyu Huang. A special thanks to Olivia Jacobs for helping me so much.

TABLE OF CONTENTS

ABSTRACT

Adenosine deaminases acting on RNA (ADAR) are a family of enzymes responsible for the hydrolytic deamination of adenosine (A) to inosine (I), specifically in double-stranded RNA. This editing event can have coding as well as non-coding effects in the transcriptome and proteome, considering that (I) may induce secondary RNA structures and because (I) is translated as guanosine (G) by cellular machinery. The family of ADARs (ADAR1, ADAR2, and ADAR3) are endogenous to humans, but only ADAR1 and ADAR2 are catalytically active. ADARs have multiple crucial roles in RNA processing, thus, dysregulation of ADAR activity has been linked to many diseases, including cancer. While there is abundant work in characterizing the mechanism, structure, and function of these enzymes, there is substantial work in harnessing ADARs to correct mutagenic transcripts; a technique often called Site-Directed RNA Editing (SDRE). In brief, a therapeutic guide oligonucleotide (complementary to the mutated RNA target of interest) is inserted to create a double-stranded RNA substrate and site-specifically recruit ADAR for RNA editing. Knowledge of this enzyme's mechanism, crystal structures, and substrate selectivity will allow us to understand ADAR's role in disease biology and speed up the development of ADAR-based therapeutics.

This dissertation describes the understanding of known high-resolution crystal structures of ADAR2 to rationally design guide oligonucleotides to direct ADAR editing and probe unknown structural features of ADAR1. More specifically, this dissertation details the use of nucleoside analogs and chemical biology to test most of the hypotheses presented herein. Chapter 1 provides a general introduction tailored to the content of the following chapters. It describes the detailed function of ADAR enzymes, their association to diseases, the mechanism, and its use for SDRE. Chapter 2 and Chapter 3 focus on designing guide oligonucleotides to direct therapeutic and highly selective editing in 5'-U<u>A</u> and 5'-A<u>A</u> sequence contexts (<u>A</u> is target adenosine). Chapter 4 provides a structure rationale to design ADAR1 inhibitors and explores methods to covalently cross-link RNA duplexes with modified nucleosides. Chapter 5 details X-ray

crystallography studies to design purine analogs that enhance the ADAR reaction at disfavored sites, like 5'-G<u>A</u>. Lastly, Chapter 6 presents efforts to synthesize a 128 nucleotide (nt) long pegRNA with a guanosine analog for tethering to the reverse transcriptase of a prime editor for complex stabilization and cryo-electron microscopy studies.

SUMMARY OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 4

CHAPTER 5

SUMMARY OF SCHEMES

SUMMARY OF TABLES

CHAPTER 6

SUMMARY OF ABBREVIATIONS

2'-F-U……..2'-fluorouridine

2'-F-ANA-U.....2'-fluoroarabinouridine

2'-*O*-Me……..2'-O-methyl

2'-*O*-4'-*C*-diMeA/U.....2'-O-4'-C-dimethyladenosine/uridine

3-deaza-dA.....3-deaza-2'-deoxyadenosine

4'-*C*-MeA/U.....4'-C-methyladenosine/uridine

5HT$_{2c}$.....5-hydroxytryptamine receptor 2C

7-deaza-dA……7-deaza-2'-deoxyadenosine

8-AN.....8-azanebularine

A.....Adenosine/Adenine

ADAR .....Adenosine Deaminase Acting on RNA

ADAR1/2 RD.....ADAR with one dsRBD

ADAR2d or ADAR1d.....Deaminase domain of ADAR

AGS.....Aicardi-Goutières Syndrome

BDF2.....Bromodomain factor 2

BME……beta-mercaptoetanol

BSA.....Bovine serum albumin

C.....Cytidine/Cytosine

CRISPR.....Clustered regularly interspaced short palindromic repeats

crRNA.....CRISPR RNA

COSY…….Correlation Spectroscopy

dA.....2'-deoxyadenosine

DBU……..……1,8-diazabicyclo(5.4.0)undec-7-ene

dC.....2'-deoxycytidine

DCM……..…Dichloromethane

dG.....2’-deoxyguanosine

dI.....2’-deoxyinosine

DIPEA……..Diisopropylethylamine

DMAP……..N,N-dimethylaminopyridine

DMEM.....Dulbecco’s modified eagle medium

DMSO.....Dimethyl sulfoxide

DMTr/DMT……4’,4’-dimethoxytrityl

DNA.....Deoxyribonucleic acid

DNase.....Deoxyribonuclease

dNTP.....Deoxynucleoside triphosphate

dsRBD.....Double-stranded RNA binding domain

dsRNA.....Double-stranded RNA

dT……...2’-deoxythymine

DTT.....Dithiothreitol

dU…….2’-deoxyuridine

dUTP.....Deoxyuridine triphosphate

dZ.....Benner’s base Z or 6-amino-5-nitro-3-(1’-*β*-D-2’-deoxyribofuranosyl)-2(1H)-pyridone

EDTA.....Ethylenediaminetetraacetic acid

FDA……Federal Drug Administration

FBS.....Fetal bovine serum

G.....Guanosine/Guanine

GLI1.....Glioma-associated oncogene 1

HR.....Homologous recombination

HMBC……...Heteronuclear Multiple Bond Correlation

HSQC……...Heteronuclear Single Quantum Coherence

I.....Inosine/Hypoxanthine

IDUA.....Alpha-L-iduronidase

IFN.....Interferon

LNA……..Locked Nucleic Acid

m1A.....N1-methyladenosine

m6A.....N6-methyladenosine

MALDI-TOF.....Matrix-assisted laser desorption/ionization-time of flight

MECP2.....Methyl CPG binding protein 2

NHEJ.....Non-homologous end joining

NMR……Nuclear Magnetic Resonance

NOE……..Nuclear Overhauser Effect

PAGE.....Polyacrylamide gel electrophoresis

PBS.....Phosphate-buffered saline

PCR….....Polymerase chain reaction

PDB.....Protein Data Bank

PO.....Phosphodiester

PS.....Phosphorothioate

RNA.....Ribonucleic acid

RNase….....Ribonuclease

RNasin.....RNase inhibitor

RT-PCR.....Reverse transcription-Polymerase chain reaction

SD.....Standard deviation

SDS.....Sodium dodecyl sulfate

SEC.....Size exclusion chromatography

SNAr.....Nucleophilic aromatic substitution

SRC.....proto-oncogene protein kinase Src

TBAF……Tetrabutylammonium fluoride

TBS……tert-butyldimethylsilyl

TBDPS……..……tert-butyldiphenylsilyl

TEV.....Tobacco Etch Virus

Tm…….Melting temperature

U.....Uridine/Uracil

UNA……..Unlocked Nucleic Acid

Ura.....Uracil

WT.....Wild-type

CHAPTER 1

Introduction to RNA Editing via Adenosine Deaminases Acting on RNAs (ADARs)

INTRODUCTION

A diverse transcriptome

Advances in sequencing and mass spectrometry have discovered more than 150 chemical modifications of the RNA nucleotides.[1,2] These modifications can be found at the sugar and the nucleobase and have important functions in regulating cellular processes. Common modifications include pseudouridine (ψ) incorporation, and methylation of adenosine and cytidine [N6-methyladenosine (m6A), N1-methyladenosine (m1A), and 5-methylcytosine (5mC)] (Figure 1).[1] RNA editing can be defined as 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. Adenosine to inosine (A-to-I) and cytidine to uracil (C-to-U) deamination are a common form of RNA editing. These post-transcriptional modifications are referred to as the epitranscriptome and allow for the diversification RNAs and their encoded proteins.[3]

Adenosine deamination on dsRNA

Of these post-transcriptional events, the deamination of adenosine to inosine (A-to-I) in double-stranded RNA (dsRNA) is one of the most common RNA editing events in metazoans (Figure 2A, A to I editing).[4] This enzymatic process is catalyzed by a family of deaminases called adenosine deaminases acting on RNA (ADAR). Inosine products are abundant in different types of RNA (noncoding RNAs and mRNA), making this modification crucial for cellular processes.[5] In the context of mRNA editing, ADARs can change the meaning of codons due to inosine's base pairing selectivity to cytidine, causing inosine to be read as guanosine by cellular machinery during translation and other RNA binding enzymes (Figure 2B).[6]

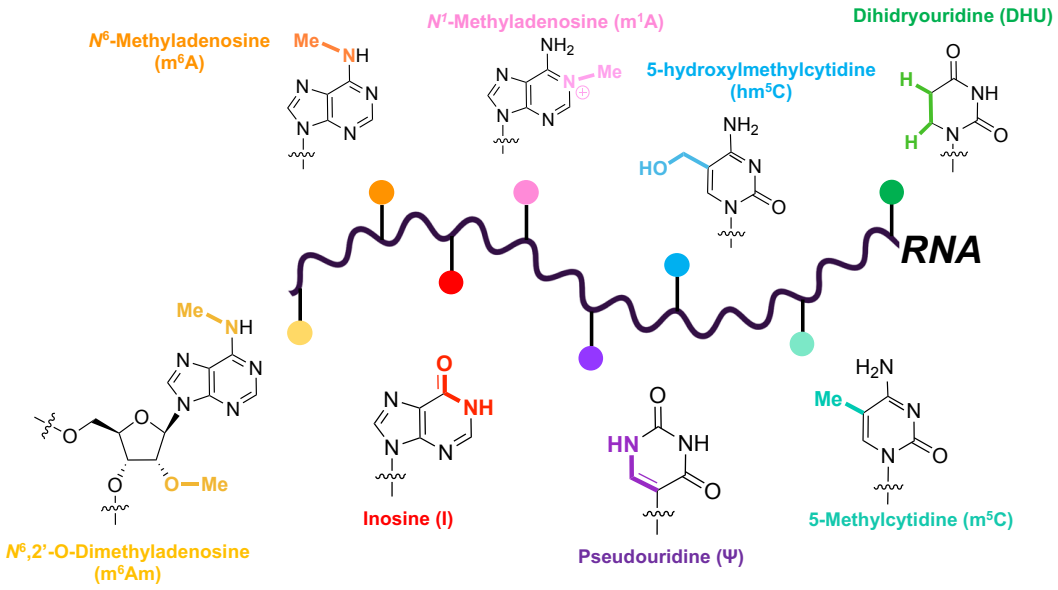


Figure 1. A diverse transcriptome. A subset of chemical structures of different RNA modifications at the sugar and nucleobase level.

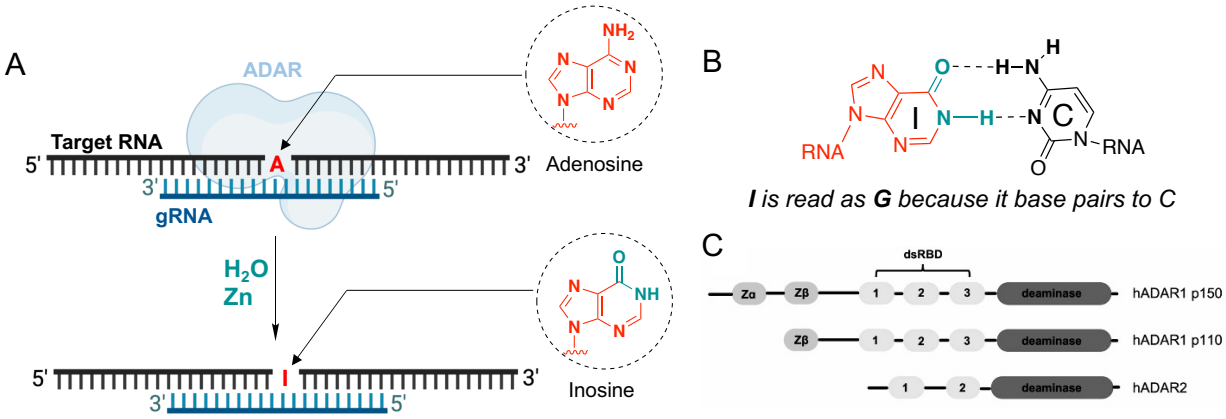


Figure 2. RNA editing by adenosine deaminases acting on RNA. A) Double stranded RNA substrate being edited by ADAR in the presence of water and zinc. Red A: target sit. Structures of adenosine and inosine. B) Base pairing of inosine to cytidine. C) *H. sapien* ADAR isoform's protein domain maps.

The ADAR family of enzymes and their biological importance

The human ADAR family consists of three main members: ADAR1 p110/p150, ADAR2, and ADAR3 (Figure 2C), with only ADAR1 and ADAR2 being catalytically active.[7] Two domains of these enzymes are recurrent within the three isoforms and are the deaminase domain and the double-stranded RNA binding domain (dsRBD), which aid ADARs to recognize their duplex RNA

substrates.[8] In addition, ADAR1 has a Z-DNA binding domain, whereas ADAR3 has an N-terminal arginine-rich domain.[7] The expression of ADAR1 and ADAR2 can be found widespread across many tissues in the human body. However, ADAR3 is highly localized in the brain[9,10]

Understanding the biological importance of ADAR enzymes in humans is an ongoing task in the field. Adenosine to inosine (A-to-I) edits can be categorized into two types of editing events depending on the region of the edited RNA: (1) non-coding and (2) coding. In coding regions of an mRNA, A-to-I editing can produce amino acid substitutions leading to new mRNA sequences and thus modified protein isoforms.[11] Splice donor and acceptor sites can also be created with RNA editing by ADAR, adding complexity to alternative splicing. Non-coding editing events usually occur within the untranslated regions or intergenic regions of a mRNA.–Usually, A-to-I edits in these regions affect RNA secondary structural changes and influence RNA stability and its molecular interactions with other RNA or RNA-binding proteins, such as the RNA interference pathway and proteins involved in RNA degradation and immune response.[12–14]

To date, it is known that ADAR1 and ADAR2 are crucial for the survival of some metazoan organisms.[15] ADAR1 KOs in mice are embryonically lethal; ADAR1 is known to play a major role in innate immune activation in humans and in cancer, by hyper-editing endogenously expressed cytoplasmic dsRNAs to suppress their activation by dsRNA binding proteins.[16–19] To illustrate, in tumor cells, nucleic acid sensors that are activated by specific stimuli turn on the expression of interferon-stimulated genes (ISGs) and type I interferon production, which should lead to tumor cell death due to the activation of the viral sensor melanoma differentiation-associated gene 5 (MDA5).[18]However, expression of ISGs promotes upregulation of ADAR1 expression and hyper-editing of ADAR1. These endogenous dsRNAs then become unrecognizable by MDA5, due to many A-to-I structural changes and effectuate tumor growth. In Chapter 4, we discuss how inhibiting ADAR1 could be used as an immunotherapy tool to promote tumor cell apoptosis with a new class of RNA inhibitors.[20]

ADAR2 is important for brain development. KO of ADAR in mice is lethal due to the absence of editing in the coding region of the glutamate-gated ion channel receptor subunit B (*GluR-B*).[21] This Q/R substitution disrupts GluRB calcium permeability and the presense of the genomically encode glutamine results in neurotoxicity.[21]

Site-directed RNA editing by ADARs

Because ADARs require duplex structure for activity, their binding and subsequent editing can be directed to specific adenosines using complementary guide strands that form duplexes at the target sites.[22] This approach, Site Directed RNA Editing (SDRE), is currently being pursued to develop a new class of therapeutic oligonucleotides that recruit ADARs to correct disease-causing mutations in RNA.[23] Chemical modifications are required for these ADAR-guiding oligonucleotides to facilitate delivery, increase metabolic stability, and increase the efficiency and selectivity of the ADAR reaction.[24–27] Identifying and characterizing ADAR guide strand modification patterns that lead to highly efficient and selective editing at different therapeutically relevant target sites is an ongoing effort by several laboratories in academics and industry.[28–30] Advantages to using SDRE strategy for therapeutics include a transient and tunable effect as RNA is delivered and degraded, no induction of double-stranded DNA breaks (as with DNA editing), no restriction to a particular phase of the cell cycle, and no need for base editor delivery, as ADARs are endogenous to humans and the strategy relies on the guide strand to recruit endogenous ADAR. Yet, the field faces many challenges including the low editing efficiency of endogenous ADAR depending on the cell- and tissue-type, and the continuous administration of the RNA therapeutic. One example of SDRE as a successful therapeutic strategy involves the pathogenesis of Rett syndrome, which can be appreciated in Figure 3. Our lab has pursued reversing the mutagenic stop codon on the W104X site of the human methyl CpG binding protein 2 (*MECP2*) to restore a healthy phenotype in patients suffering from this mutation.[31] W104X has the codon UAG that can be edited into a UGG, restoring protein function.

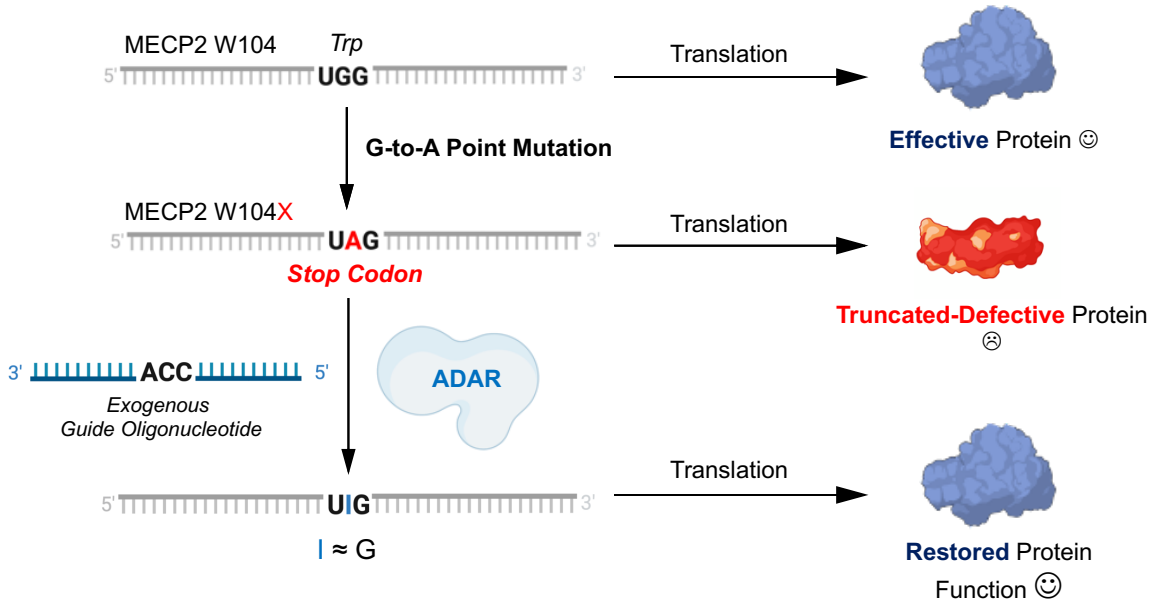


Figure 3. Site-directed editing by ADARs on the methyl CPG binding protein 2 (MeCP2)

Catalytic mechanism and structural characterization of ADAR

In 2000, Stephens and co-workers determined that ADARs induce a base-flipping event on the target adenosine as a key feature of this enzyme's catalytic mechanism.[32] In their study, they used a fluorescent purine analog (2-aminopurine) in an RNA that would fluoresce only when flipped out of the duplex. It was determined that the target A is flipped out as fluorescence signals were recorded when 2-aminopurine was placed at the editing site. Furthermore, this base-flipping mechanism was confirmed in 2016 by the crystal structure of the deaminase of ADAR2 bound to a duplex RNA containing a transition state analog. In this study, Melissa Matthews, Justin Thomas, and co-authors used the modified nucleoside 8-azanebularine (8-AN) to trap the enzyme.[33] The design of 8-AN was inspired by what had been understood of the hydrolytic deamination of adenosine at position C6.[34] It was proposed that the deamination reaction would undergo a nucleophilic aromatic substitution (S_NAr), where water would act as the nucleophile while position C6 in adenosine would act as the electrophile. Water becomes basic by coordinating with zinc while also being deprotonated by residue glutamate 396 in the active site

(Figure 4A). After the nucleophilic attack of water to C6 of adenosine, the tetrahedral-hydrated intermediate forms and briefly interrupts the aromaticity of the purine. Subsequent removal of the exocyclic amine, as ammonia, generates inosine as the reaction product (Figure 4A). Noteworthy, the design of 8-AN arose from this observation. This was achieved by (1) having an additional nitrogen atom at position 8 of the purine and (2) replacing the exocyclic amine to a hydrogen atom (Figure 4B). This in turn (1) increases the electrophilicity of C6 by the additional electronegative effect exerted by the 8-aza position and (2) because the C6-amine was replaced by a hydrogen, the 8-AN hydrated intermediate allows or the mechanistic trapping.[34,35]

Thanks to the crystal structure breakthrough discovery, many exceptional features of the structure of ADAR2 bound to dsRNA have been elucidated over the following 8 years with different chemically modified RNA duplexes.[8,29,33,36,37] In tandem with other enzymes that undergo an analogous base-flipping mechanism, ADAR2 is predicted to stabilize the flipped-out conformation by interacting with its flipping loop via non-covalent interactions with the orphan base (base opposite to the target A upon base-flipping) in the complementary RNA.[38,39] This flipping loop consists of amino acids 486-489. Yet, residue glutamic acid E488 hydrogen bonding to the orphan base [(O4 and N3 for U) or (N4 and N3 for C)] makes up for most of the stabilization needed to withstand the flipped-out conformation (Figure 5A).[29] Interestingly, ADAR's observed preference for editing adenosines across a cytidine (C) or a uridine (U) can be explained by a possible clash between E488 and a purine at the orphan position.[33,40]

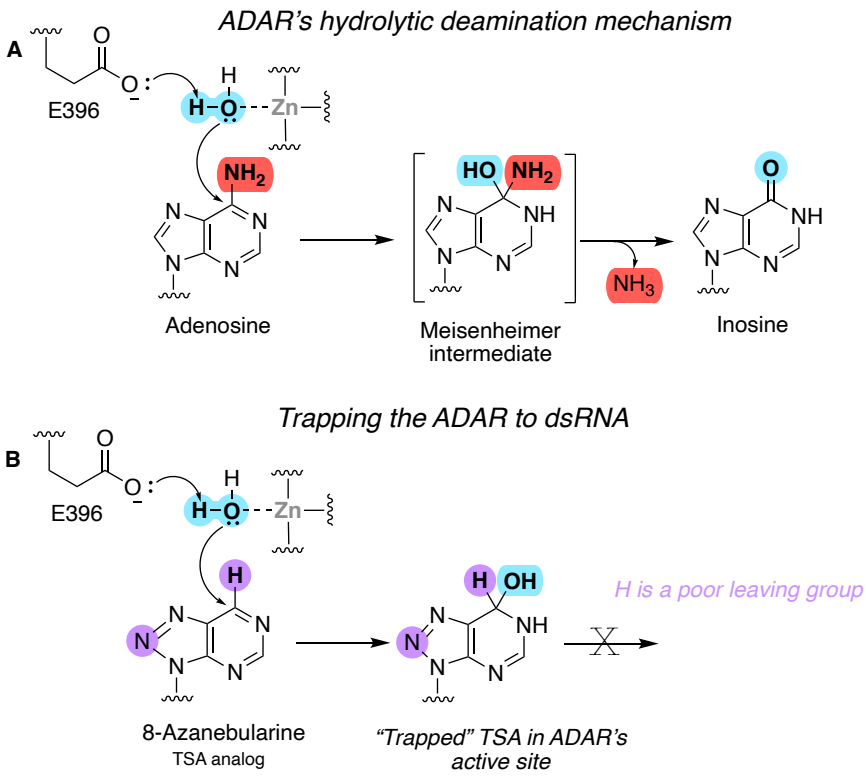


Figure 4. ADAR reaction mechanism and trapping of the enzyme using nucleoside analogs. A) The ADAR reaction mechanism involves hydrolytic deamination of adenosine to inosine in the presence of Zinc and water and is catalyzed by glutamate 396 in the deaminase domain. B) 8-azanebularine serves as a transition state analog by mimicking the transition state of the hydrolytic deamination reaction and traps the dsRNA within the active site considering that the H at C6 is a poor leaving group.

It is noteworthy to highlight several of the conformational distortions that the RNA duplex undergoes after base-flipping. One major change observed in the ADAR bound RNA duplex is its deviation from the typical A-form helix RNA duplex. A widening of the major groove opposite the editing site is observed alongside a kink near the orphan base (Figure 5B). This kink is promoted by the base opposite to the 5' neighbor of the target site, called the -1 nucleotide. Only this nucleotide suffers a change in sugar pucker from C3'-*endo* to C2'-*endo*. In essence, a witty and sleek change from RNA to DNA sugar pucker accommodates the conformational stress induced by the target site being flipped out of the duplex (Figure 5C).[27] Furthermore, Chapters 2 and 3 are dedicated to the -1 position, where ADAR editing is regulated by chemically modifying the -1

nucleotide depending on the substrate sequence context (5'-UA or 5'-AA, with -1 nucleotide as adenosine or uridine, respectively).

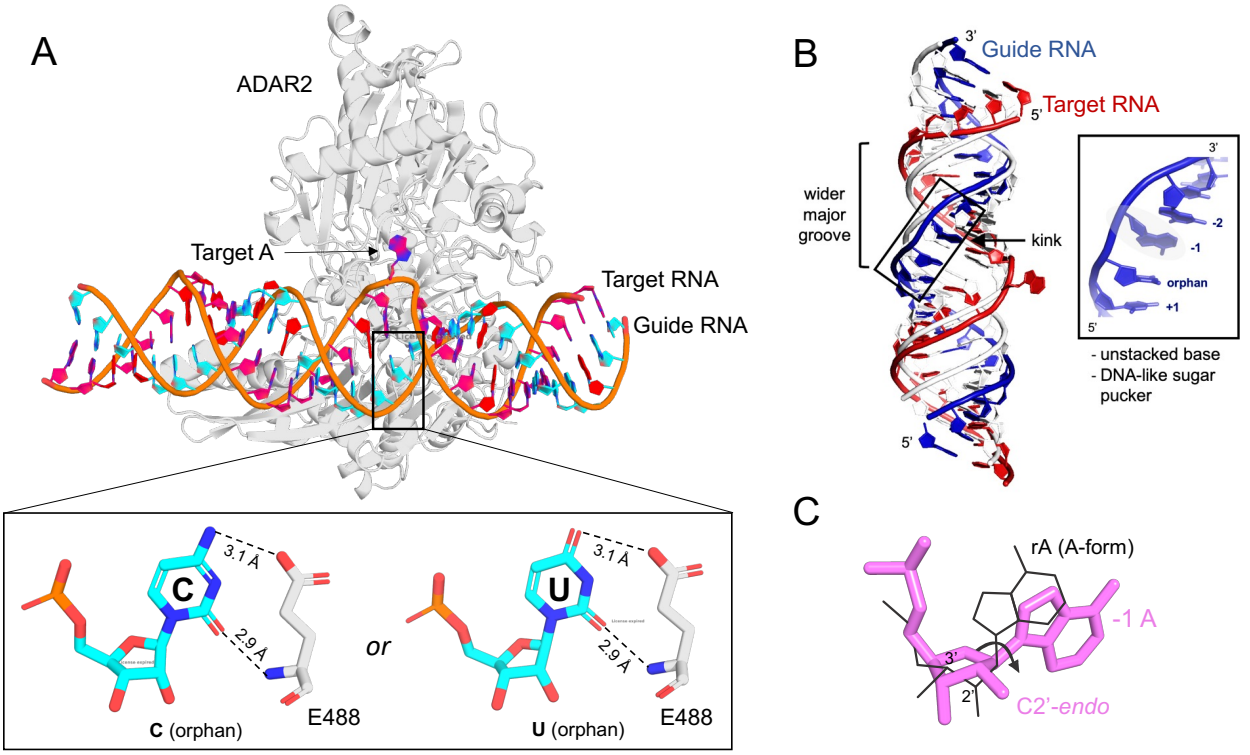


Figure 5.[8] The ADAR2-RNA interface as illuminated by crystal structures of ADAR2-D bound to 8-AN-containing RNA duplexes. (**A**) ADAR reacts with 8-AN in RNA to form a hydrate that mimics the hydrated intermediate of ADAR's adenosine deamination (see Figure 4). Crystal structure of human ADAR2-D (grey) in complex with an 8-AN-containing RNA duplex showing the flipped-out base in the edited strand (red) and intercalating loop bearing E488. E488 hydrogen bonds with the orphan base in the unedited strand with C or U. (**B**) Crystal structure of the RNA duplex from an ADAR2-D-RNA co-crystal structure 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. (**C**) Close up of the -1 sugar pucker and glycosidic bond distortion.

In the same study of Matthews and co-authors, ADAR's nearest-neighbor preferences are detailed. ADARs prefer editing A nucleotides with 5'-nearest neighbor U (or A) and 3' nearest-neighbor G.[33] In addition, the minor-groove edge of this 5'-A::U or 5'-U::A is perfectly juxtaposed with the protein backbone at G489. However, modeling a 5'-G:::C or 5'-C:::G base pair at this neighboring position strongly suggests that the 2-amino group of guanosine in the minor groove

would sterically clash with the protein at residue G489.[37] Hence, this could explain why sequence contexts 5'-GA or 5'-CA are the least preferred by ADAR for efficient editing. In Chapter 5 we will show how purine analogs at the -1 position in the context of 5'-GA targets can restore efficient editing and overcome ADAR's limitation on properly editing adenosines with 5' adjacent guanosines.

Goals and perspectives

Throughout this chapter we have elucidated ADAR and RNA editing, the core function of ADAR enzymes, the structure and mechanism of these enzymes, and their potential as a therapeutic tool for cancer or other diseases with SDRE. The rest of this dissertation will address the use of ADARs as a biotechnological tool employing various chemical modalities to optimize these processes following the dogma of chemical biology. Most of the chapters herein (if not all), have used the high-resolution crystal structures of ADAR2 bound to duplex RNA as muses for the work presented in addition to novel synthetic nucleic acid chemistry. Hence, this dissertation describes the synthesis of modified oligonucleotides for SDRE and structure-activity relationships of ADAR enzymes. Starting with Chapters 2 and 3, we explore how novel nucleoside analogs are used to discover the effects of sugar and nucleobase modifications at the -1 position of ADAR guide strands in the sequence context of 5'-UA and 5'-AA ADAR sites. The discoveries in Chapters 2 and 3 serve as additional tools in the chemical toolbox for companies or academic groups interested in enhancing or site-specifically inhibiting ADAR for therapeutic SDRE. Parts of the findings of Chapter 2 were published in the *RSC Chemical Biology* by Dr. Hannah F. Brinkman and other parts in *Nucleic Acid Research* by the writer of this dissertation. The results of Chapter 3 are being summarized in an ongoing manuscript for future publication. Chapter 4 navigates the design of chemically modified RNA inhibitors for ADAR1 p110 and p150 isoforms to create immunotherapies for cancers where ADAR1 is correlated in its tumorigenesis. The results of this chapter were mainly produced and led by Dr. Herra Mendoza and were published in *Biochemistry*. Lastly, Chapter 5 is a recollection of miscellaneous stories where I had the chance to collaborate

and include (1) the synthesis of 8-azainosine for SDRE in 5'-GA sequence contexts and (2) the synthesis of a guanosine analog containing a C3 methylene thiol linker emanating from the N2 for mechanistic trapping of a pegRNA in the reverse transcriptase domain of a prime editor complex. The findings of 8-azainosine in 5'-GA sites were submitted for publication in *Biomolecules* and were led by current lab member Aashrita Manjunath.

CHAPTER 2

Rational Design and Synthesis of Adenosine Analogs Targeting 5'-UA Sites for RNA Editing

This work originated from a collaborative project led by Dr. Hannah F. Brinkman on the study of multiple chemical modifications at the -1 position of 5'-UA ADAR targets. Observations in that project gave birth to the core project of this chapter which focuses on rationally designing nucleoside analogs to inhibit ADAR site-specifically. This chapter contains excerpts from a manuscript published in Nucleic Acid Research in May 2024, as well as introductory data from the collaboration with Dr. Hannah F. Brinkman[27,41]*.*

INTRODUCTION

Adenosine (A) to inosine (I) is one of the most prevalent RNA modifications in mammals [42,43]. This conversion is catalyzed by a group of enzymes called adenosine deaminases acting on RNA (ADARs) and takes place exclusively in duplex RNA. In humans, ADAR1 and ADAR2 are active adenosine deaminases targeting RNA. Importantly, the A to I reaction can produce codon changes in mRNA because I is read as G during translation[44,45]. In addition, since ADARs require duplex structure for activity, their reaction can be directed to specific adenosines in different transcripts using complementary guide strands that form duplexes at the target sites[22]. This approach is currently being pursued to develop a new class of therapeutic oligonucleotides that recruit ADARs to correct disease-causing mutations in RNA[23]. Chemical modifications are required for these ADAR-guiding oligonucleotides to facilitate delivery, increase metabolic stability, and increase the efficiency and selectivity of the ADAR reaction[24,28,29,46,47]. Identifying and characterizing ADAR guide strand modification patterns that lead to highly efficient and selective editing at different therapeutically relevant target sites is an ongoing effort by several

laboratories in academics and industry[24–29]. In this chapter, we demonstrate the effect of sugar analogs with different conformations and alkylation degrees [2'-O-Methy (2'-OMe), Unlocked nucleic acid (UNA), Locked nucleic acid (LNA) and 2'-deoxy] at a position opposite an editing site's 5' nearest neighbor nucleotide where the neighboring base of the target is a uridine (U), hence 5'-U<u>A</u>[27]. This position in the guide RNA is denominated for the rest of this dissertation as the "-1". Here, we also probe the effect of LNA modification at other positions in ADAR guide strands. Furthermore, we describe efforts to define the mechanism by which ADARs are inhibited by LNA modification at specific positions (-1 and others) in the guide strand. These efforts were facilitated by the synthesis of RNA oligonucleotides bearing 4'-*C*-methylated nucleoside analogs designed to probe the steric effect of a small 4'-alkyl substituent without the conformational restriction of the 4'-*C*, 2'-oxymethylene linkage found in the LNA monomers. Importantly, while LNA modification potently inhibited target site editing at two adjacent sites in the ADAR guide strand (-1 and -2), this effect is seen for the 4'-*C* Me modification only at the -2 position. The ability of a 4'-*C*-methylated nucleoside to retain editing at the -1 site but not at the -2 site provides a novel method to block adjacent bystander editing in 5'-**<u>A</u>**A target sequences. Finally, these results are rationalized considering the different effects of steric and conformational flexibility at different positions in ADAR guide strands.

RESULTS

Crystal structures of ADAR2-RNA duplexes reveal unique RNA duplex conformational features at the sugar, base and backbone level

An abundant amount of work and research in the ADAR therapeutic and basic science field has been enabled by the discoveries of the crystal structures of different ADAR2 constructs trapped and bound to modified RNA duplexes[8,29,33,37]. In all the known crystal structures of ADAR2-RNA complexes, the nucleotide that is to be edited flips out of the RNA duplex into the enzyme's active site where the deamination reaction occurs. Once this target nucleotide is

stabilized and trapped within the active site, many conformational changes are observed in the guide RNA to, presumably, compensate for the highly ordered flipped-nucleotide conformation. One of these remarkable RNA duplex conformational changes is concentrated at the nucleotide that is opposite/base-paired to the 5' neighbor of the target adenosine in the guide RNA. This nucleotide position is denominated as the -1 and it is responsible for a kink in the backbone within the guide RNA (Figure 1A and 1C). This kink is a result of significant deviations from standard A-form conformation for the nucleotide paired with the editing site's 5' nearest neighbor (guide -1 position). This kink is possibly attributed to the drastic change in the sugar pucker of the -1 sugar where it adopts a C2'-endo pucker typical of a DNA nucleotide, and not the C3'-endo sugar pucker found in A-form RNA (Figure 2B)[48]. Hence, all these observations make the -1 position of an ADAR guide strand a good position to be studied. This includes a comprehensive understanding of these enzymes' mechanism and the inhibition or enhancement of ADAR editing utilizing structure rationale.

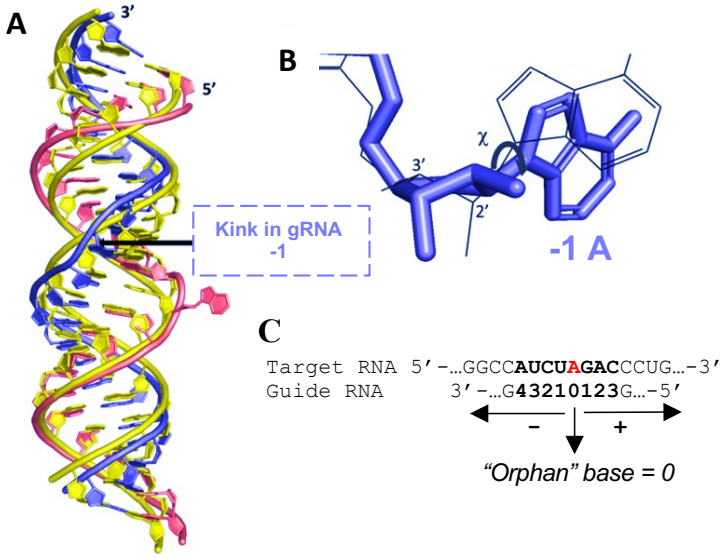


Figure 1. Unusual conformation of guide strand near the -1 nucleotide indicated by X-ray crystallography[33]. (A) Superposition of typical A-form duplex RNA (yellow strands) with RNA duplex conformation present in complex with ADAR2 (blue (guide) and pink (target) strands). (B) The -1 A has a C2'-endo sugar pucker with a high anti glycosidic bond angle (χ = -90.61 degrees). Blue indicates conformation found in ADAR2-RNA structures, superimposed blue wire structure is A-form conformation. (C) Schematic of target:guide

strand duplex formed in site directed RNA editing indicating nomenclature and numbering scheme used in this work. The -1 position was optimized for target A (red) editing.

With the structural information of ADARs bound to dsRNA substrates and the detail in which the -1 position can be observed, Brinkman and co-workers embarked on testing the effect of sugar analogs in this position, in a 5'-UA sequence context, where the -1 harbors adenine as the base[27]. Considering the kink formed upon ADAR's base editing, we chose different sugars considering their variability in sugar conformations and ring strain and contribution to RNA hybridization. To attest this, we subjected ADAR to edit the human methyl CPG binding protein 2 (MECP2) W104X site, a G to A mutation that causes Rett syndrome[31]. Oligonucleotides of 29 nt were synthesized by solid-phase synthesis to incorporate the following sugar modifications in addition to the control – adenosine (rA): 2'-O-Me A, 2'-deoxyA/(dA), LNA A and UNA A. Each of the duplexes bearing different nucleoside analogs at guide position -1 were allowed to react *in vitro* with purified human ADAR2. The effect of each analog was determined by measuring the percent editing of the target site at the reaction end point and by comparing the rate of reaction under single turnover conditions (Table 1). The rate constants (k_{obs}) for editing were then plotted for each analog tested (Figure 2).

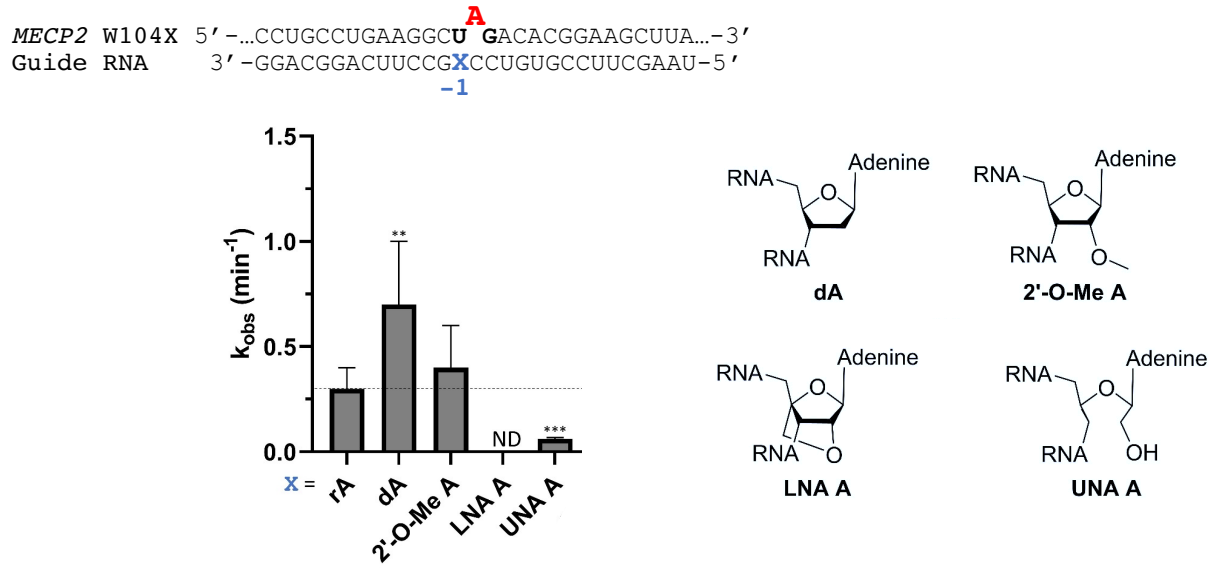


Figure 2. Effect of guide strand -1 nucleotide sugar modification on ADAR2 deamination rate for the MECP2 W104X site.

From the observed plotted rates in Figure 2, -1 dA displayed a significant 2.3 fold increase, while the deamination rate with -1 2'-OMe A remained efficient and not significantly different from rA (Table 1). This 2 fold increase in rate with the use of -1 dA is likely attributed to the fact that the -1 position undergoes an RNA-to-DNA sugar pucker conformation upon deamination. Thus, having already a 2'-deoxy/DNA sugar pucker positioned at the -1 can be interpretated as a facilitation for the enzyme's base flipping for efficient editing. Interestingly, when UNA A was placed -1 to the target site in the MECP2 transcript in the gRNA, there was a significant decrease in rate (five-fold decrease) compared to rA. Furthermore, there was no detected editing at any reaction timepoint when the -1 position was substituted to LNA A. These results prompted us to explore why extreme changes in sugar conformations (acyclic/UNA and hyper-constrained/LNA sugars) at the -1 nucleotide renders wickedly low deamination rates and site-specific ADAR inhibition.

Table 1. The effect of -1 gRNA modifications on fitted end points and observed rate constants for ADAR2 deamination reactions with *MECP2*.[a]

-1 modification [a]	Fitted end point (%)[b]	k_{obs} (min^{-1})[b]	k_{rel}[c]
Adenosine (rA)	96.2 ± 0.6	0.3 ± 0.1	1
dA	95.2 ± 0.1	0.7 ± 0.3	2.3
2'-O-Me A	93.8 ± 1.0	0.4 ± 0.2	1.3
LNA A	no reaction	no reaction	no reaction
UNA A	87.2 ± 0.7	0.061 ± 0.006	0.2

[a] Reactions were carried out with 5 nM of target RNA and 15 nM ADAR2 WT
[b] Reactions were fitted to the equation $[P]_t = \alpha[1 - \exp(-k_{obs} \cdot t)]$
[c] $k_{rel} = k_{obs}$ for modification/k_{obs} for adenosine

Inhibition of the ADAR reaction by LNA substitution is highly position-dependent

Previously, we showed that an LNA monomer located at the -1 position on an ADAR guide strand inhibited target editing and that strategic positioning of LNA could block unwanted bystander editing when the off-target site was at least six nucleotides away from the target[27]. Since LNA modifications in RNA therapeutics could be useful for enhancing target RNA binding affinity and resistance to nuclease degradation, we tested how ADAR responded to LNA

modification at other positions in the guide RNA[49]. We allowed ADAR2 to react with an RNA duplex bearing the sequence present in the c-Src mRNA. c-Src is a protein kinase frequently overexpressed in cancers[50]. Specifically, we targeted for RNA editing the K295 codon (5'-A**A**A-3'). Editing of the central adenosine will convert this to an arginine codon resulting in a K295R mutation that renders the oncogenic protein kinase inactive[51]. The sequence surrounding the target site is 5'-AUCA**A**AA-3', where the target adenosine is directly flanked by 2 potential bystander editing sites (Figure 3A) with two additional bystander sites nearby. We first determined the percent editing by human ADAR2 at 60 min for the desired target and the four bystander sites employing an unmodified guide RNA as a control for comparison with the LNA-containing guide RNAs (Figure 3B). We then modified each sugar independently starting from the -4 position to the +3 position of the guide relative to the target adenosine and measured the percent editing for the target and bystander sites (Figure 3C-I). A schematic of the numbering of positions relative to the target A in the guide RNA is shown in Figure 1A, including the "orphan" position, which is the base on the guide opposite to the target adenosine. Without LNA modifications, we observed significant editing of bystander adenosines four nucleotides and one nucleotide 5' and one nucleotide 3' of the target A. However, certain patterns of inhibition were observed depending on whether the LNA monomer was placed in the guide 5' or 3' of the orphan position. For instance, LNA placement 3' of the orphan position at positions -4 and -3 in the guide RNA reduced by >50% or blocked editing entirely at the bystander sites, while also reducing editing at the target site by 43% and 46%, respectively (Figure 3C and 3F). Moreover, when the LNA monomers were placed closer to the target site in positions -2 and -1, editing at the target site and bystander sites decreased by > 95% for each site (Figure 3D-E). A very different editing profile was observed when LNA monomers were positioned 5' of the orphan nucleotide. When LNA monomers were placed +1, +2, and +3 in the guide RNA, either no inhibition (+2 and +3) or less than two-fold reduction (+1) was observed. Yet, sites opposite the 3' base flanking the LNA monomer or directly across it showed reduced or inhibited editing (Figure 3I).

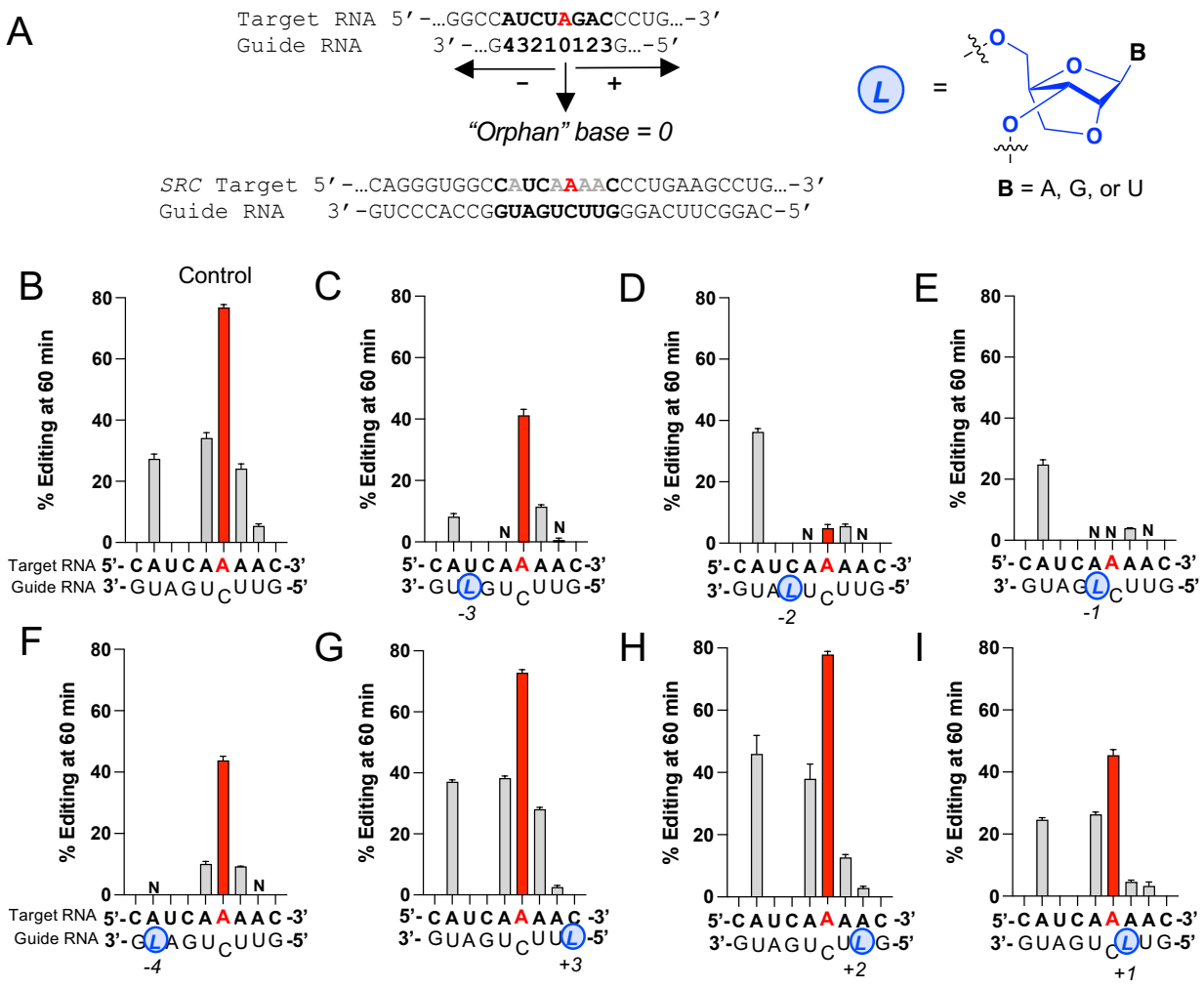


Figure 3. Site-specific LNA modification can reduce or block editing at various off-target sites by ADAR2 on the *SRC* K295R transcript. A) Left Top: Schematic of target:guide strand duplex indicating nomenclature and numbering scheme used in this work; Bottom Left: Target RNA and guide RNA hybrid construct of the *SRC* K295R; red A: adenosine target; grey: bystander sites; bold: fragment studied for the LNA modifications. Right: Structure of LNA in blue. B) Editing yield at 60 min with no LNA modification. Red bars indicate editing for target A; off-target sites are displayed in light grey bars. C-I) Editing yield at 60 min for different guide oligonucleotides bearing LNA (blue and *itallics L*) at different sites on the guide RNA (-3, -2, -1, -4, +3, +2, and +1, respectively). Reactions for each timepoint were carried out with a ratio of 5:50 nM of RNA hybrid to ADAR2 FL WT. Error bars represent the standard deviation of three technical replicates. N: No detected editing.

High-resolution structures reveal possible origin of LNA inhibition of the ADAR reaction

Intrigued by the potent inhibitory effect of LNA monomers at the -1 and -2 positions of a guide RNA, we sought to understand the mechanistic origin of this phenomenon. The available

high-resolution structures of ADAR2 bound to duplex RNA show that the minor groove edge of the

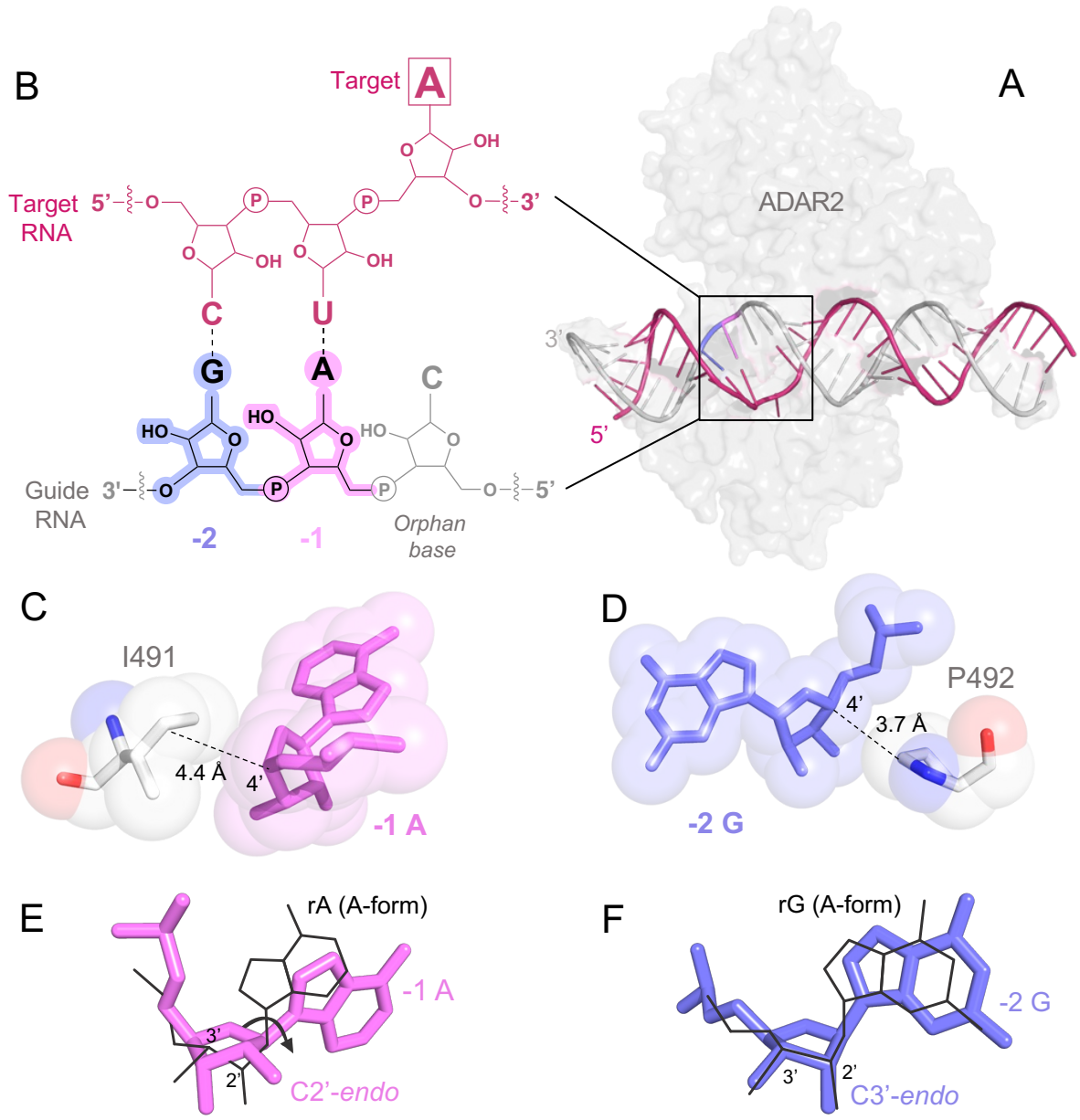


Figure 4. Structure of hADAR2-R2D E488Q bound to the GLI1 32 bp RNA at 2.8 Å resolution.[8] A) View of the structure perpendicular to the dsRNA helical axis, the pink and slate blue region shows the kink in the guide RNA and the widening of the major groove opposite the editing site induced by ADAR2. The warm pink shows the target RNA and the light grey the guide RNA. B) Schematic of target and guide RNA hybrid for ADAR editing in a 5'-UA-3' sequence context indicating nomenclature and numbering scheme used in this work. -1 position highlighted in pink and -2 in blue. C) Space filling model of the -1 nucleotide 4'-*C* closely approaching the γ and δ carbons of I491 in ADAR's flipping loop. D) Space filling model of the -2 nucleotide 4'-*C* closely approaching P492's side chain. E) The -1 A (pink) adopts a C2'-*endo* sugar pucker

with a high anti glycosidic bond angle (χ = -90.6°) compared to an ideal A-form nucleotide (black). F) The -2 G (slate blue) adopts an RNA canonical C3'-*endo* sugar pucker (black).

ribose 4' carbon (4'-C) at the -1 position is in proximity (4.4 Å) to the γ and δ carbons of isoleucine 491 (Figure 4) [8]. This residue lies on a loop involved in stabilizing the flipped-out conformation required for adenosine deamination by ADARs (Figure 4C) [29,33]. Likewise, these structures also reveal that the 4'-C of the -2 nucleotide in the guide RNA approaches even more closely (3.7 Å) the side chain of proline 492, a highly conserved amino acid in the ADAR family of enzymes while retaining an RNA-like C3'-endo sugar pucker (Figure 4D and 4F) [4].

These observations suggested that inhibition by the substitution of the -1 and -2 nucleotides in the guide RNA with an LNA monomer could result from steric or conformational effects (or both). In the case of each of the -1 and -2 LNA nucleotides, one can predict a steric clash between the side chains of I491 and P492 with the 4'-*C*-2'-oxymethylene linkage present in the minor groove at an LNA sugar. Considering that the -1 nucleotide undergoes an RNA-to-DNA sugar pucker conversion, inhibition when LNAs are placed at this position, could arise from the restricted/locked conformation. To distinguish the possible origins of the inhibitory effect of LNA substitution in ADAR guide strands (i.e. restricted conformation or steric clash), we compared LNA substitution to two nucleoside analogs with different sugar methylation sites. We envisioned these analogs arising from cleaving either the bond between 2'-O and 6'-C to afford 4'-*C*-methyladenosine (4'-*C*-MeA) or cleavage of the 4'-C and 6'-C bond to furnish 2'-*O*-methyladenosine (2'-*O*-MeA) (Figure 5). Each of these analogs retain the atoms present in LNA A, but lack the conformational constraint imposed by the 4'-C to 2'-O bridge. Thus, if the steric demand of the additional methylene present in LNA A (compared to rA) were responsible for the inhibitory effect, at least one of these two analogs would be expected to inhibit the ADAR reaction when placed at the corresponding position in the guide strand. However, if each were well tolerated by ADAR, the LNA effect on conformational flexibility would be the more likely origin of inhibition. Since ADAR's preferred 5' nearest neighbor nucleotide is U or A (i.e. 5'-U**A** or 5'-A**A**),

the -1 nucleotide in a guide RNA targeting these sequences is uridine or adenosine. Yet, in this project, we tested sugar analogs of adenosine for studies of mostly 5'-U**A** targets. In addition to the 2'-*O*-MeA, LNA A, and 4'-*C*-MeA, we also tested the unlocked nucleic acid version of adenosine (UNA A). Since UNA A is highly flexible, this analog could tell what extent flexibility is beneficial at the -1 position [52].

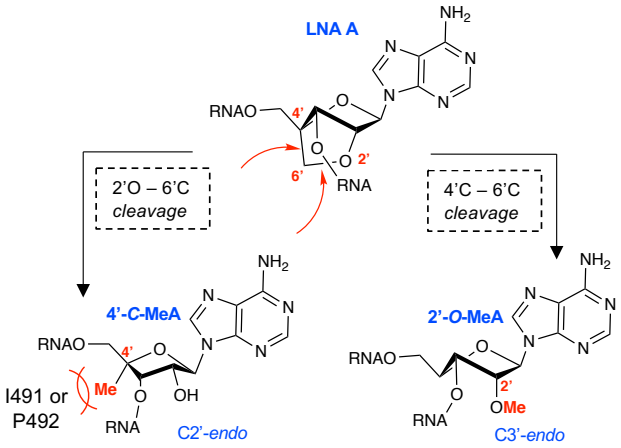


Figure 5. Proposed structures of the possible methyl groups rearranged from the 2'-*C*-4'-*C*-oxymethylene linkage cleavage of LNA at the -1 position. LNA is fixed to an RNA C3'-*endo* sugar pucker conformation. Right: 2'-*O*-Me A; Left: 4'-*C*-Me A and proposed steric clash with I491 or P492. 4'-*C*-Me A adopts a DNA-like C2'-*endo* sugar pucker conformation.

4'-C-methyladenosine phosphoramidite and RNA synthesis

To incorporate each of these modifications into RNA, their corresponding phosphoramidites were needed. However, the 4'-*C*-MeA phosphoramidite had not been previously reported. Waga and co-workers developed a synthetic strategy to incorporate 4'-*C*-alkyl substituents into free ribonucleosides to test for effects on HIV replication [53,54]. Furthermore, Koizumi and Kano synthesized 4'-*C*-aminoalkyl uridine modified RNA and tested for effects on siRNA delivery in RNA interference [55,56]. Lastly, Guo et al. synthesized a 4'-*C*-methyluridine phosphoramidite and modified RNA to understand the effects of 4'-*C*-alkylation on ribose conformation during base-catalyzed RNA cleavage via internal transesterification [57]. Here, we used a 7-step synthesis to furnish the 4'-*C*-MeA phosphoramidite starting from the previously

synthesized 4'-α-*C*-methylribose **1** in reasonable yields (Scheme 1-Methods) [58,59]. Initially, regioselective N^9 ribosylation with N^6-benzoyladenine and tin tetrachloride ($SnCl_4$) afforded the protected adenosine analog **2**. The desired regioisomer was identified by heteronuclear multiple bond correlation (HMBC), heteronuclear single bond correlation (HSQC), and homonuclear correlation (COSY) NMR experiments. In these experiments, the N^7 analog displayed an interaction between the glycosyl anomeric hydrogen and C^5 and C^8 of the purine, while for the N^9 analog, the anomeric hydrogen interacts with carbons C^4 and C^8 [60]. In addition, the selective formation of the β-ribonucleoside is supported by Höfler and Vorbrüggen's mechanistic explanation of glycosylations employing Friedel-Craft catalysts, such as $SnCl_4$ [61]. Subsequent deacylation of the 2'-OH without the removal of the N^6-benzoyl group was accomplished with potassium carbonate (K_2CO_3) and methanol in high yields, followed by the desilylation of the 5'-OH with 1M tetrabutylammonium fluoride (TBAF). Debenzylation of the 3'-OH was achieved by employing 1M boron trichloride (BCl_3) in CH_2Cl_2 to afford **5** in a form ready for advancement to the phosphoramidite. Hence, tritylation of the 5'-OH with *4'4'*-dimethoxytrityl chloride (DMTr-Cl) and pyridine, followed by the regioselective silylation of the 2'-OH in the presence of silver nitrate ($AgNO_3$), pyridine, and *tert*-butyldimethylsilyl chloride (TBS-Cl) provided compound **7** in good yields. The selective silylation of the 2'-OH in ribonucleoside phosphoramidites often produces low yields due to the poor selectivity between the 2' and 3' hydroxyl groups [62]. However, here we propose that the 2'-silylation reaction furnishes higher than usual yields and selectivity due to the inhibitory/steric effect of the 4'-C methyl group for the 3'-*O*-silylation. The desired isomer was identified by NMR homonuclear correlation of the 1'-2' protons and one distinct signal from the ^{29}Si HMBC correlation of the 2'-O**Si**TBDM – 2'-H. Lastly, phosphitylation of the 3'-hydroxyl provided phosphoramidite **8**, suitable for automated RNA synthesis.

Guide RNAs containing 4'-C-MeA at the -1 position do not inhibit ADAR

To test the effect of the sugar modifications at the -1 position, we hybridized chemically modified guide strands to an RNA transcript bearing the sequence from the human methyl CpG binding protein 2 (*MECP2*) containing the W104X site caused by a G-to-A mutation that causes Rett syndrome [31]. The resulting duplexes were then allowed to react with purified human ADAR2. The effect of each analog was measured by determining the percent editing at the target site at 60 min (Figure 6A). In addition, the effect of 4'-*C*-MeA was evaluated by measuring deamination rates under single turnover conditions and compared to duplexes with canonical adenosine (rA) (Figure 6A). Single turnover condition were achieved by reacting duplex RNA with excess enzyme (3 or more fold excess, depending on ADAR isoform). We also determined the effect of these analogs in an ADAR1 reaction using a target RNA bearing the mouse *IDUA* W392X nonsense mutation and purified human ADAR1 p110 (Figure 6B). Mutations within the human *IDUA* gene, which encodes *α-L*-iduronidase, can cause Hurler syndrome [63].

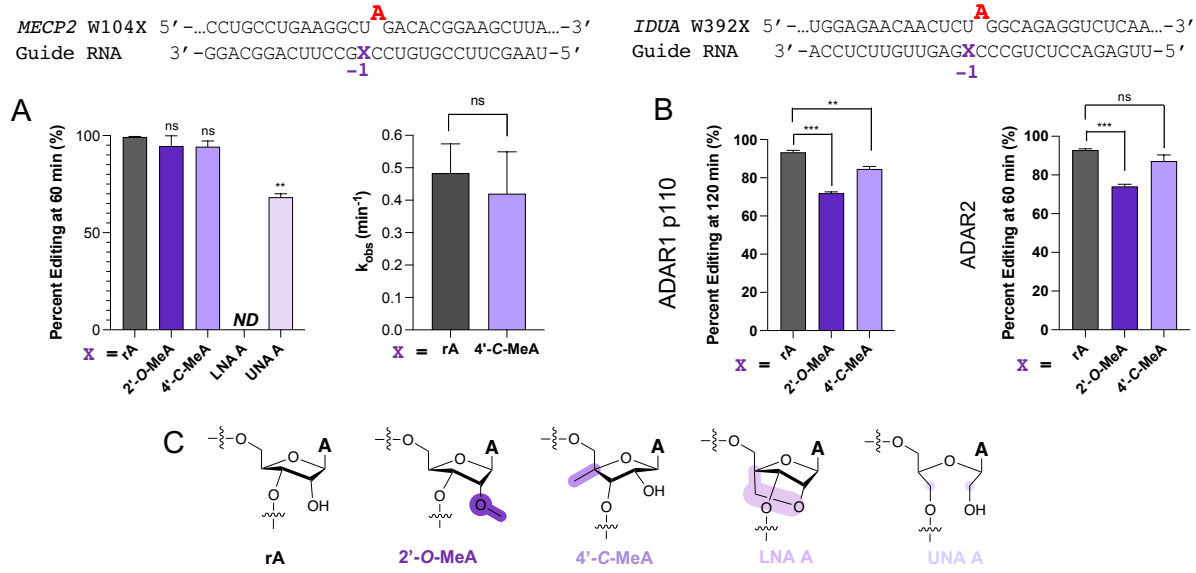


Figure 6. A) Top: Target sequence derived from the mRNA for human *MECP2* proximal to the W104X mutation associated with Rett syndrome and guide RNA used here to direct corrective editing at the premature stop codon. The target adenosine is indicated in red. X indicates the -1 position of the guide strand relative to the target adenosine. Right: Comparison of different ribose modifications at the -1 nucleotide on selective editing at the target site plotted as editing percent at 60 min. Left: Deamination rate comparison of canonical A and 4'-*C*-MeA. Reactions for each timepoint were carried out with a ratio of 5:15

nM of RNA hybrid to ADAR2 WT. B) Top: Target sequence derived from the mRNA for mouse *IDUA* proximal to the W392X truncation mutation. Bottom: Effect of different ribose modifications at the -1 nucleotide on ADAR1p110 and ADAR2 WT editing at the target site plotted as editing percent at 120 and 60 min, respectively. Reactions for each timepoint were carried out with a ratio of 15:150 nM of RNA to ADAR1p110 and 5:15 nM for ADAR2 WT. Error bars represent the standard deviation of three technical replicates. A two-tailed Welch's *t* test was conducted, where *$p < 0.05$, **$p < 0.01$, ***$p< 0.001$ from rA -1 gRNA; ns: no significant difference. C) Structures of rA, 2'-*O*-MeA, 4'-*C*-MeA, LNA A, and UNA A. ND: No detected editing.

The mouse *IDUA* W392X mutation mimics the Hurler-causing human W402X mutation that can be reversed by ADAR-mediated editing and is known to be a good substrate for ADAR1 p110 [29]. The effects of having 2'-*O*-MeA, 4'-*C*-MeA, LNA A, and UNA A at the -1 position for ADAR2 and ADAR1 p110 editing are shown in Figure 6. Interestingly, 2'-*O*-MeA and 4'-*C*-MeA behaved similarly to rA and do not show any significant difference in the level of ADAR2 editing at 60 min. Yet, as previously shown by Brinkman et. al, having LNA A at the -1 position inhibited the ADAR2 at the *MECP2* W104X site [27]. However, UNA A does show a significant difference in the level of editing compared to rA. Importantly, we determined the ADAR2 deamination rate constant for the 4'-*C*-MeA-containing guide and found it to be indistinguishable from the adenosine-containing guide (Figure 6A).

Guide RNAs containing 4'-C-MeA at the -2 position inhibit the ADAR reaction

RNA editing experiments analogous to those described above were carried out to examine the impact of 4'-*C*-MeA on ADAR reactivity when it is placed at the -2 position of the gRNA. First, a modified *MECP2* W104X* transcript was designed and generated where the -2 base 5' of the target was changed from a cytidine to uracil to base pair with 4'-*C*-MeA. Then, guide RNAs containing rA, LNA A, 2'-*O*-MeA and 4'-*C*-MeA at the -2 position were synthesized. *In vitro* deaminations with purified human ADAR2 were carried out and the effect of each analog was measured by determining the percent editing of the target site at the reaction endpoint at 60 min (Figure 7A). Notably, no editing was detected when 4'-*C*-MeA was positioned -2 relative to the target, suggesting that the methyl group of the 4'-*C*-Me in the -2 sugar does indeed clash with

P492 (Figure 7C). Likewise, when LNA A was placed at the -2 position, editing was reduced by 86% when compared to the control (-2 rA), displaying similar editing results to those seen with the c-Src target (Figure 3D). No significant difference was observed when 2'-*O*-MeA was positioned -2 relative to the target A, suggesting that this sugar modification is well tolerated for ADAR editing.

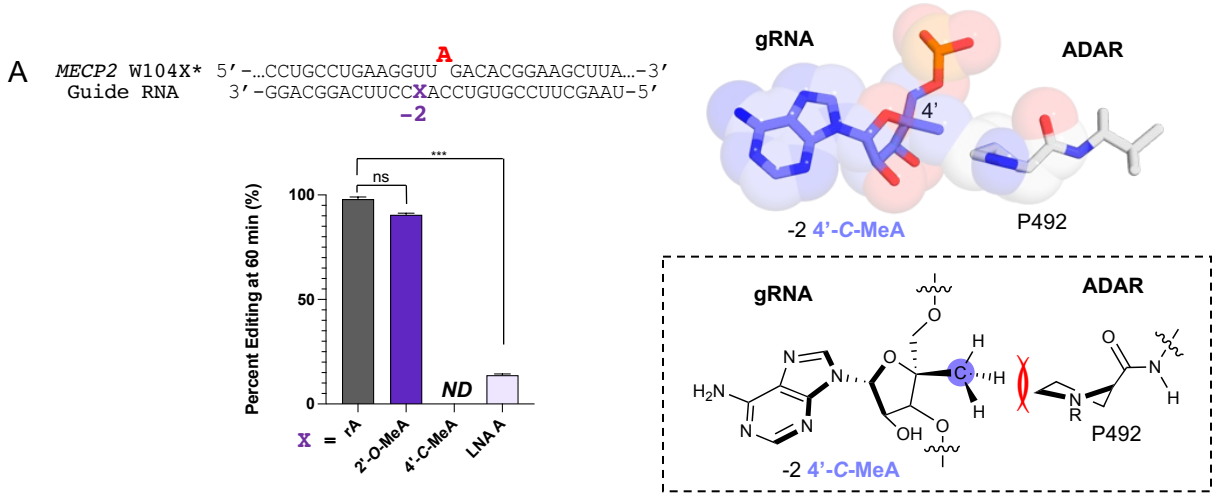


Figure 7. A) Comparison of different ribose modifications at the -2 nucleotide on editing at the human modified MECP2* mRNA target site plotted as editing percent at 60 min. Left: Reactions for each timepoint were carried out with a ratio of 5:15 nM of RNA hybrid to ADAR2 WT. Error bars represent the standard deviation of three technical replicates. A two-tailed Welch's *t* test was conducted, where *$p < 0.05$, **$p < 0.01$, ***$p< 0.001$ from rA -1 gRNA; ns: no significant difference. ND: No detected editing. B) Stick and space filling model overlay of the -2 nucleotide 4'-*C*-methyl closely approaching the methylene group in Pro 492 of ADAR's flipping loop (24). C) Chemical structure of Pro 492 and -2 4'-*C*-Meguanosine in the guide RNA highlighting the location of the predicted steric clash.

Effect on ADAR editing of 4'-C-MeA at various positions of the gRNA

Our previous experiments focused on highlighting the effect of 4'-*C*-MeA when it is placed at the -1 and -2 positions of the gRNA for the MeCP2 and IDUA targets. However, we sought to discover the effect of this analog in another sequence context while comprehensively understanding its effect when it is placed -2, -3, +1 and +4 relative to ADAR editing sites. The A-rich c-Src provides a holistic framework to achieve this. For this, we placed 4'-*C*-MeA in a gRNA at position that would be -2 relative to a 5' adjacent off-target site (blue), -3 relative to the target (red) and -4 and +1 relative to other bystander sites (grey) (Figure 8A). In vitro deaminations show

that 4'-*C*-MeA is still wickedly inhibitory when placed -2 to the 5' bystander site, to the point of no detectable editing when compared to a non-modified gRNA (Figure 8A). Interestingly, 4'-*C*-MeA is significantly inhibitory (3 fold) when placed -3 to the target site. This phenomenon can be attributed to a potential steric clash with ADAR's flipping loop amino acid cluster R481, I456 and F457 and the proximal 4'-*C* facade from the -3 nucleotide in the gRNA (Figure 8B). Yet, no significant difference in editing was observed when this analog was placed +1 and -4 relative to other bystander sites.

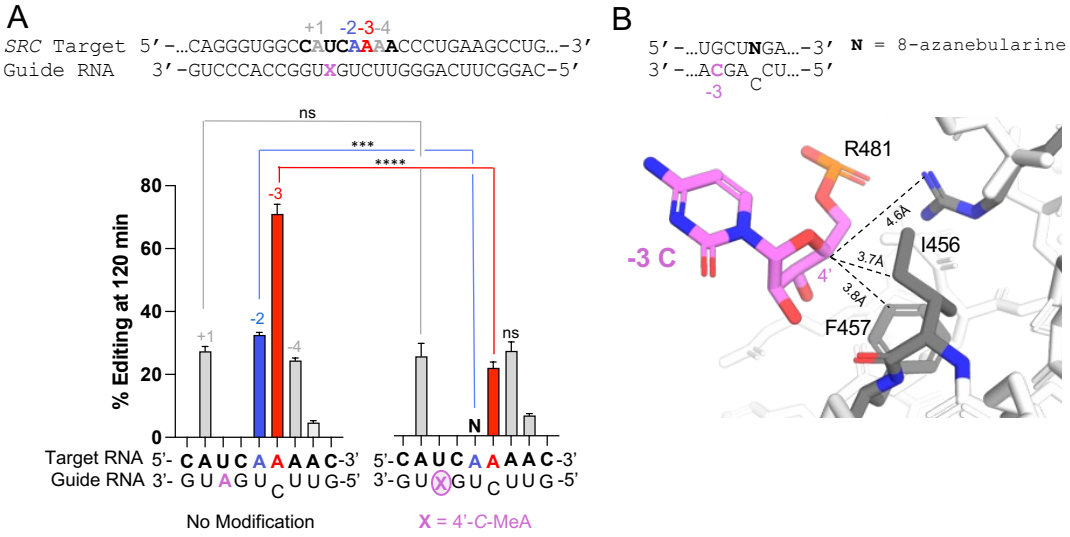


Figure 8. Site-specific 4'-*C*-MeA localization in the guide RNA affects ADAR editing of target and proximal off-target sites on the *SRC* K295R transcript. A) Top: Schematic of target:guide strand duplex indicating target (red), 5' adjacent off-target site (blue) and other bystander sites (grey). Guide RNA with blue X (violet) indicating -2 position relative to the 5' adjacent off-target site (blue), -3 relative to the target (red) and -4 and +1 relative to other bystander sites. Bottom: ADAR editing yield at 120 min when X = no modification (left) and when X = 4'-*C*-methyladenosine. Red bars indicate editing for target A; -2 off-target site editing is displayed in blue bars and remaining bystander sites editing in grey bars. Reactions for each timepoint were carried out with a ratio of 5:50 nM of RNA hybrid to ADAR2 WT. Error bars represent the standard deviation of three technical replicates. N: No detected editing. B) Structure of hADAR2-R2D E488Q bound to the GLI1 32 bp RNA with 8-azanebularine (N), an ADAR transition state adenosine analog at 2.8 Å resolution; Stick model of the -3 nucleotide 4'-*C* closely approaching the side chains of R481, I456 and F457 in ADAR's flipping loop (24).

4'-C-Me-modified guide RNAs can selectively block adjacent bystander sites while retaining on-target editing efficiency

Since it appeared that 4'-*C*-methyl modification was tolerated at the -1 position but not at the -2 position of an ADAR guide strand, we decided to test the impact of the modification on a substrate with two adjacent adenosines. Directing ADARs to edit one of two adjacent adenosines can be challenging, particularly if the desired edit is at an adenosine with a non-optimal nearest neighbor nucleotide [64,65]. Therefore, we generated a substrate with the target sequence 5'-U**A**AG-3', with the desired edit to take place at the first of the two adjacent adenosines and in a non-optimal sequence 5'-U**A**A-3' with the adjacent bystander site in the near-optimal 5'-A**A**G-3' sequence context. ADAR guide strands were synthesized that vary the nucleotide at the -1 position and orphan position relative to the target adenosine. The duplexes formed with these guide strands were then incubated with human ADAR2 E488Q, a hyperactive mutant of the enzyme or wild-type human ADAR1 p110 [5]. Interestingly, with a guide strand of unmodified RNA and uridine at the orphan position (Figure 9AB, X= rA and Y = rU) we find nearly equal editing levels at the two adjacent adenosines with both proteins. With unmodified RNA and cytidine at the orphan position (Figure 9AB, X = rA, Y= rC), the reaction becomes selective for the target A but with significant editing for both ADAR2 E488Q (24 ± 1 %) and ADAR1 p110 (21 ± 3 %) at the bystander A at the 120 min time point, respectively. However, with both proteins, the guide strand with 4'-*C*-methyladenosine at the -1 position relative to the target site (i.e., -2 position relative to the bystander site) reduces bystander editing to an undetectable level while maintaining efficient editing at the target adenosine (Figure 9AB, X = 4'-*C*-MeA, Y = rC). To further attest the site-specific inhibitory effect of 4'-*C*-methylene sugars, we subjected for the first time a native transcript with a potential therapeutic editing site (5'-U***A****AG*) in the human protein phosphatase 2 regulatory subunit B'delta (PPP2R5D). Many point mutations in the *PPP2R5D* gene can lead to Jordan Syndrome, a neurodevelopmental disorder[66]. Mutation of PPP2R5D's E200 codon to

E200K (Glu-to-Lys) has been suspected to be one of the many point mutations responsible for pathogenic phenotypes in Jordan Syndrome patients [66].

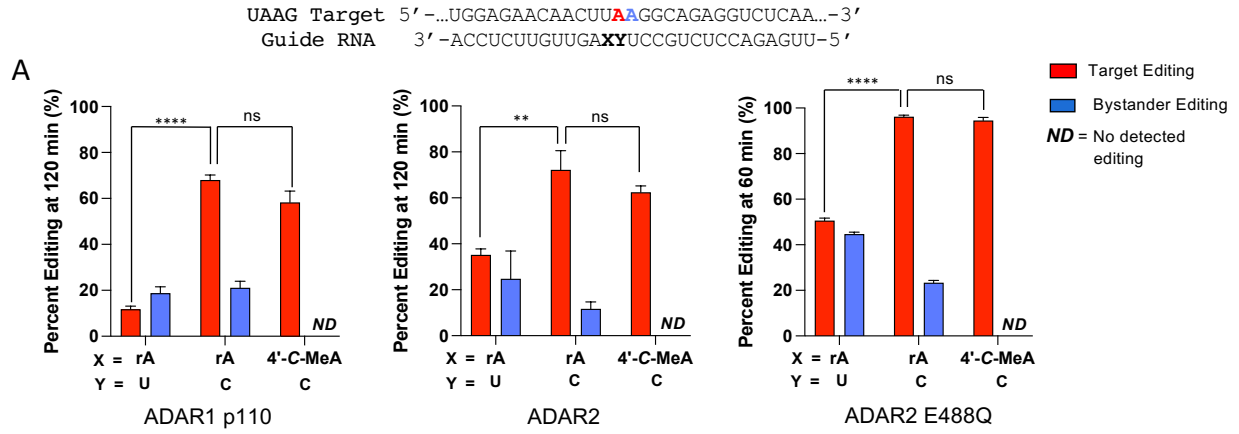


Figure 9. Site-specific inhibition of 3' adjacent bystander editing of the modified mouse IDUA W392X site (5'-UAA-3') employing 4'-*C*-methyladenosine (4'-*C*-MeA) modification at the -1 nucleotide with varying orphan bases and ADAR isoforms. A) Top: 5'-UAA-3' target RNA from modified IDUA with target adenosine (red) and 3' flanking bystander edit site (blue). X: -1 sugar modification (rA or 4'-*C*-MeA), Y: orphan base (rC or rU). Bottom: ADAR1p110, ADAR2 WT and hyperactive mutant ADAR2 E488Q editing at RNA:enzyme ratios of 15:150, 5:15 and 5:15 nM, respectively. Red bars represent % edited of target A and blue bars % editing of bystander site. Error bars represent the standard deviation of three technical replicates. A two-tailed Welch's *t* test was conducted, where $*p < 0.05$, $**p < 0.01$, $***p < 0.001$; ns: no significant difference. ND: No detected editing.

Reversion of this lysine codon (5'-U***AAG***) to glutamic acid (5'-U***GAG)*** might restore proper protein function and revert the pathogenic phenotypes seen with the E200K mutation. Because this site is a 5'-UAAG site, where the target adenosine lies within a 5'-U neighbor and a 3'-adjacent potential bystander, we generated guide RNAs to stimulate efficient editing at the target site (underlined A), while also blocking undesired editing at the proximal bystander site. To achieve this, we synthesized guide oligonucleotides with A-C mismatches at the target site while placing 4'-*C*-methyladenosine or LNA A at the -2 nucleotide position relative to the bystander site and performed in vitro deaminations with three ADAR isoforms (ADAR1p110, ADAR2 and ADAR2 E488Q) for 120 min reactions (Figure 10). Similar to previous results, we observed that placement of -1 4'-*C*-MeA strongly inhibits the 3'-adjacent bystander site exclusively, while retaining sufficient

on-target editing with all ADAR proteins compared to adenosine. Yet, as predicted, LNA A placement -1 to the target inhibits both sites for all ADAR isoforms.

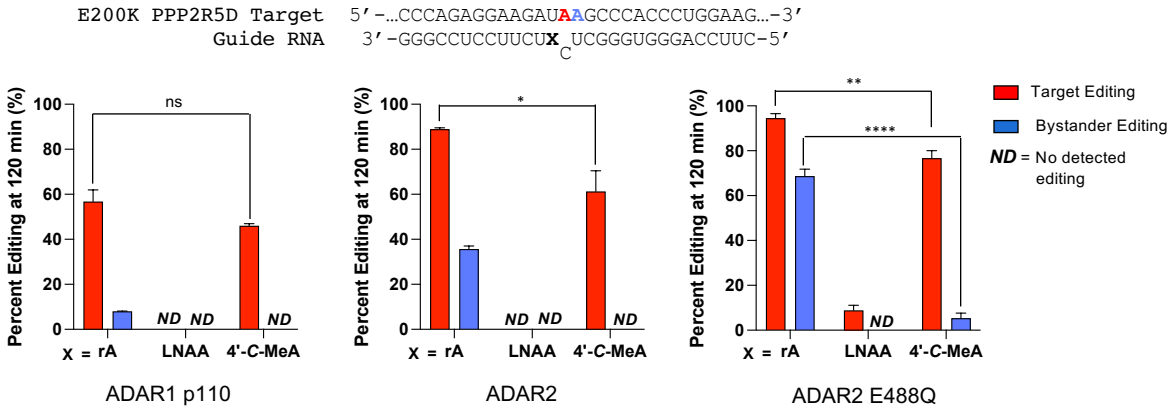


Figure 10. Site-specific inhibition of 3' adjacent bystander editing in a therapeutic transcript employing 4'-*C*-methyladenosine (4'-*C*-MeA) and locked nucleic acid (LNAA). Top: 5'-UAA-3' target RNA from the human protein phosphatase 2 regulatory subunit B'delta (PPP2R5D) E200K point mutation with the target adenosine (red) and 3' flanking bystander edit site (blue). X: -1 sugar modification (rA, 4'-*C*-MeA or LNAA). Bottom: ADAR1p110, ADAR2 WT and hyperactive mutant ADAR2 E488Q editing at RNA:enzyme ratios of 5:50 nM at 120 min. Red bars represent % edited of target A and blue bars % editing of bystander site. Error bars represent the standard deviation of three technical replicates. A two-tailed Welch's *t* test was conducted, where *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns: no significant difference. ND: No detected editing.

4'-C-MeA-modified oligonucleotides selectively direct editing in human cells with ADAR2

So far, in these studies, we have shown how positioning selectively 4'-*C*-MeA and LNA A monomers in guide RNAs can block off-targets in vitro. To determine if the inhibitory effects of these sugars can be translated in mammalian cells, we synthesized metabolically stable guide oligonucleotides to direct editing by full-length ADAR2 to an endogenous target representing a region of the 3'-UTR of β-Actin mRNA [36,67]. The guide oligonucleotide was designed to target a 5'-UUAG-3' sequence where the -1 (X) and -2 (Y) nucleotides in the guide are adenosine analogs. Hence, we tested the effect of having either X or Y = 4'-*C*-MeA or LNA A in these positions, in addition to a positive control (X, Y = rA) (Figure 11A). Additionally, we stabilized the guide oligonucleotides by replacing most riboses with 2'-*O*-methylriboses and phosphorothioate linkages, as described previously for directed editing in cells with chemically synthesized guide strands (Figure 11B) [29,47]. Expression of wild-type human ADAR2 in HEK293T cells transfected

with the positive control guide (X, Y = rA) led to (30 ± 4)% editing, while the negative control (no guide RNA transfection) rendered no editing. In the presence of a -1 4'-*C*-MeA guide, (26 ± 1)% editing was detected, with no statistically significant difference compared to the positive control. Yet, when cells were exposed to the -1 LNA A guide, no editing was detected, suggesting complete inhibition of this site. Inhibition was also achieved when the -2 position of the guide was placed with either a LNA A or 4'-*C*-MeA monomer, supporting our in vitro observations of inhibition patterns. Editing of the editing site was reduced by a 3.0 and 4.4 fold, respectively.

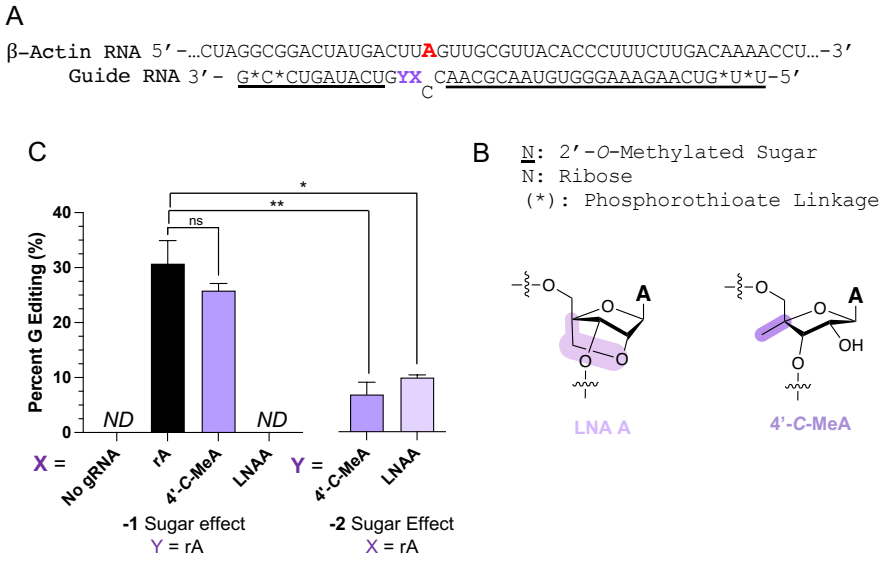


Figure 11. Effect of -1 and -2 adenosine analogs on mammalian cells on an endogenous target as measured by Sanger sequencing from RNA isolation. A) Modification pattern for guide oligonucleotides used in cellular editing and sequence of the endogenous 3'-UTR β-Actin transcript and editing site (red) where X (-1) and Y (-2) are modified. B) Structure and denominations of chemical modifications used. C) Percent editing of endogenous β-Actin in HEK 293T cells with overexpressed hADAR2 with varying sugar modifications at the -1 and -2 position of the guide RNA after 48h. Error bars, SD (n = 3 technical replicates). ND: No detected editing. A two-tailed Welch's *t* test was conducted, where $*p < 0.05$, $**p < 0.01$, $***p < 0.001$ from rA -1 gRNA; ns: no significant difference.

The effects of 4'-*C*-Me A on duplex stability and nuclease sensitivity are similar to adenosine

Lastly, we determined the melting temperature (T_M) and ribonuclease sensitivity of RNAs bearing 4'-*C*-MeA. We conducted thermal denaturation assays to elucidate the impact of 4'-*C*-MeA on duplex stability and base pairing selectivity. For this purpose, we synthesized 12 nt RNAs containing 4'-*C*-MeA or adenosine and hybridized them with complementary 12 nt RNAs for T_M measurements (Figure 12A). The

nuclease sensitivity of 4'-*C*-MeA-containing RNA was assessed using a 15 nt long oligonucleotide with this analog at the 7th position flanked by all 2'-*O*-methyl nucleotides and subjecting it to 80% human serum over a 24 h time course. For comparison, a control oligonucleotide was designed the same, but with rA at the 7th position (Figure 13). The T_M analysis showed that 4'-*C*-MeA has an almost identical profile to that of adenosine, suggesting that 4'-methylation of the sugar does not influence the thermal stability of an RNA duplex or pairing selectivity of the nucleobase. However, compared to rA, a slight decrease in serum stability was observed for the 4'-*C*-MeA-containing oligonucleotide.

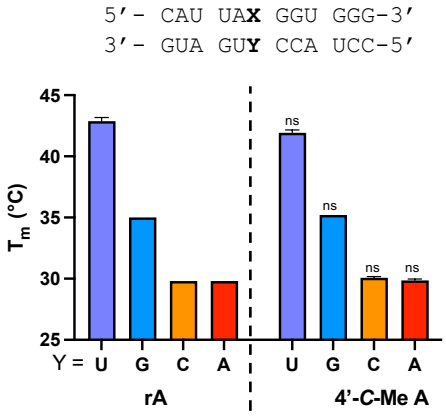


Figure 12. Thermal denaturation (T_m) of 12-mer duplexes containing adenosine and the 4'-*C*-Me Adenosine opposite the four natural bases. Error bars represent the standard deviation of three technical replicates. A two-tailed Welch's *t* test was conducted, where **p* < 0.05, ***p* < 0.01, ****p*< 0.001 from X = rA RNA; ns: no significant difference.

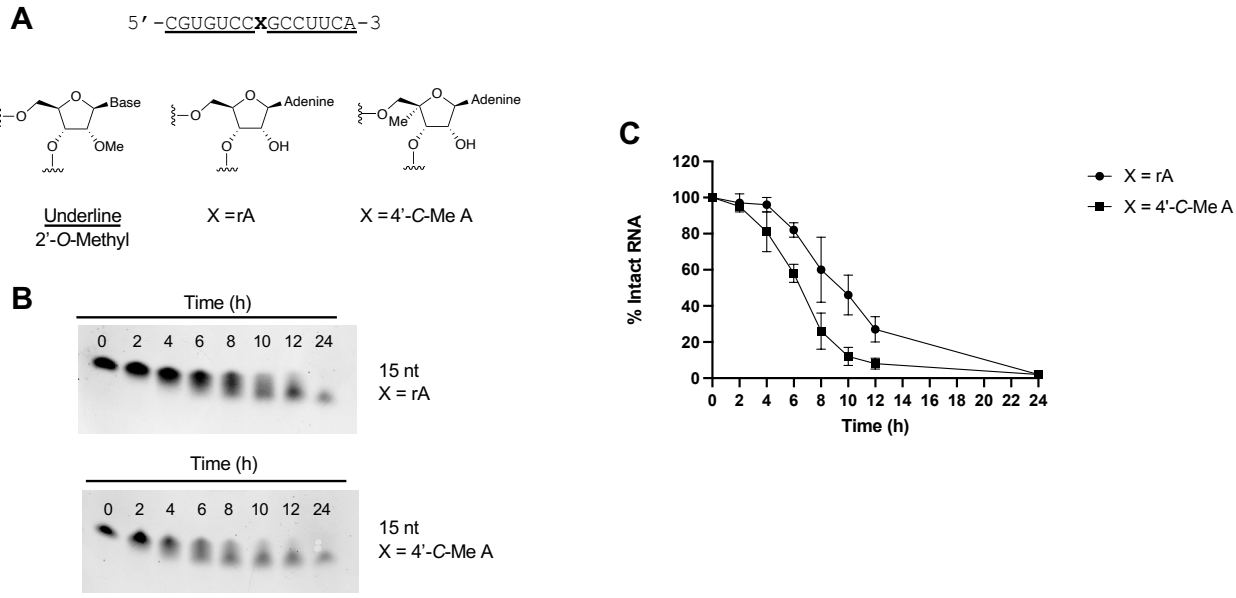


Figure 13. Human serum stability of oligonucleotides containing either adenosine (rA) or 4'-*C*-Methyladenosine (4'-*C*-Me A). A) Sequence and chemical modification structures of the 15 nt oligonucleotide tested in serum stability assay. B) 15% Denaturing PAGE at indicated timepoints in 80% human serum. C) Plot of % intact RNA over time in presence of 80% human serum (one replicate shown).

DISCUSSION

Oligonucleotide therapeutics is a maturing field with over 15 FDA-approved drugs at the time of this writing [68,69]. ADAR-catalyzed therapeutic RNA editing, guided to specific locations in the transcriptome by oligonucleotides, is a novel mechanism of action for oligonucleotide therapeutics [23]. Safe and effective drugs that work by this mechanism are currently under development with clinical trials underway [70,71]. ADAR-guiding oligonucleotides for therapeutic RNA editing require extensive modification of the parent RNA to facilitate metabolic stability, cell uptake, and deamination efficiency and selectivity. Several groups have reported efforts in tailoring guide RNAs with chemical modifications (2'-*O*-methyl, 2'-fluoro, phosphorothioate with stereopure backbones, LNA, etc.) for this purpose [14,24,26,28]. In this study, we demonstrate the effect of specific sugar-modified nucleotides at different positions in ADAR guide strands varying the proximity to the editing site. Importantly, the effects of different sugar modifications were found to be highly dependent on the position modified and rationalized by analysis of existing ADAR2-RNA crystal structures.

Generally, LNA placement in the guide strand 3' of the orphan position inhibits editing at the target A. The degree of ADAR inhibition is directly proportional to the proximity of LNA to the editing site. As shown in Figure 1C-E, when LNA is placed between the -4 and -1 positions, the inhibition significantly increases as it approaches the -1 nucleotide. Furthermore, editing of specific adenosines can be entirely inhibited when LNA is placed at either the -1 or -2 positions of the gRNA. In contrast, placing LNA monomers 5' of the orphan position in the gRNA has substantially less impact on the ADAR reaction. Our previous structural studies showed that the RNA conformation deviates significantly from A-form at the editing site and the adjacent nucleotide on the 5' side (3' direction on the guide strand) whereas the RNA conformation 3' to

the editing site is A-form (5' direction on the guide strand) [33]. Thus, the greater severity of ADAR inhibition by LNA substitution on the 3' side of the guide strand can be explained by the asymmetry of the RNA conformational changes that take place during the ADAR reaction and the innate rigidity of LNA monomers and their influence on neighboring nucleotides [52,72,73]. Lastly, it was observed that LNAs placed at the orphan position render complete inhibition (Figure 1F). This observation is also explained by the inspection of ADAR2-RNA crystal structures. The orphan nucleotide hydrogen bonds with glutamic acid 488 in ADAR2's flipping loop to stabilize the base-flipped conformation while also adopting an unusual sugar pucker that falls in between C2'-*endo* and C3'-*endo* conformations [29]. An LNA monomer is likely unable to adopt this conformation and inhibits the reaction at this position. Also, if an LNA monomer is located at the position immediately on the 5' side of the orphan (Figure 1I, +1 LNA), it is likely that the orphan nucleotide will resist the formation of alternate sugar puckers due to the previously reported influence of the vicinal LNA [74].

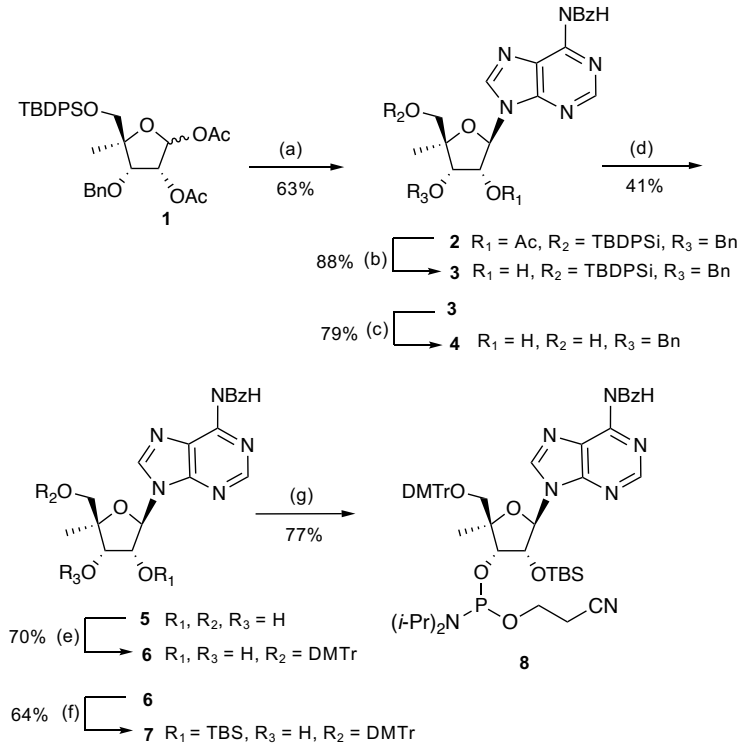
While LNA substitution potently inhibited the ADAR reaction at the 1- and -2 positions of the guide, we found that 4'-*C*-MeA inhibited the ADAR reaction only when placed at position -2. One possible explanation for the lack of inhibition with 4'-*C*-MeA at -1 is that ADAR's I491 side chain may adopt a different conformation to accommodate the 4'-*C*-methyl group emanating from the RNA's minor groove. Interestingly, based on previous work on the characterization of 4'-*C*-methyluridine monomers in RNA, this type of aliphatic modification promotes a DNA-like C2'-*endo* sugar pucker [57]. This structural attribute of the sugar in the 4'-*C*-MeA monomer can be beneficial for ADAR editing when placed at the -1 position based on previous literature that reveals that analogs that favor C2'-*endo* usually promote more efficient deamination than C3'-*endo* nucleotides at this position [27]. Based on the biochemical and structural data presented herein, we conclude that LNA monomers positioned at -1 on a gRNA inhibit ADAR by hindering the sugar re-puckering required during the ADAR reaction. Finally, it can be inferred that the presence of 4'-*C*-MeA or LNA A at the -2 position leads to inhibition of the ADAR enzyme due to steric clashes,

since the sugar pucker of the -2 nucleotide does not change upon deamination (Figure 5C). Unlike isoleucine's flexible side chain, proline lacks the conformational flexibility to accommodate the 4'-*C*-Me or methylene bridge emanating from the -2 nucleotide's minor groove of the 4'-*C*-MeA or LNA, respectively. This effect can be extended with a lesser inhibitor potency when 4'-*C*-MeA is placed -3 relative to an editing site in the gRNA (Figure 6C). This can be rationalized by the crystal structures where this -3 nucleotide 4'-*C* facade resides closely to a cluster of amino acids (4.6Å – Arg481, 3.7Å – Ile456, and 3.8Å – Phe457). Hence, considering the proximity and the cluster, placement of a methyl group in the 4'-C could give rise to another form of steric clash, therefore inhibiting the reaction at this site.

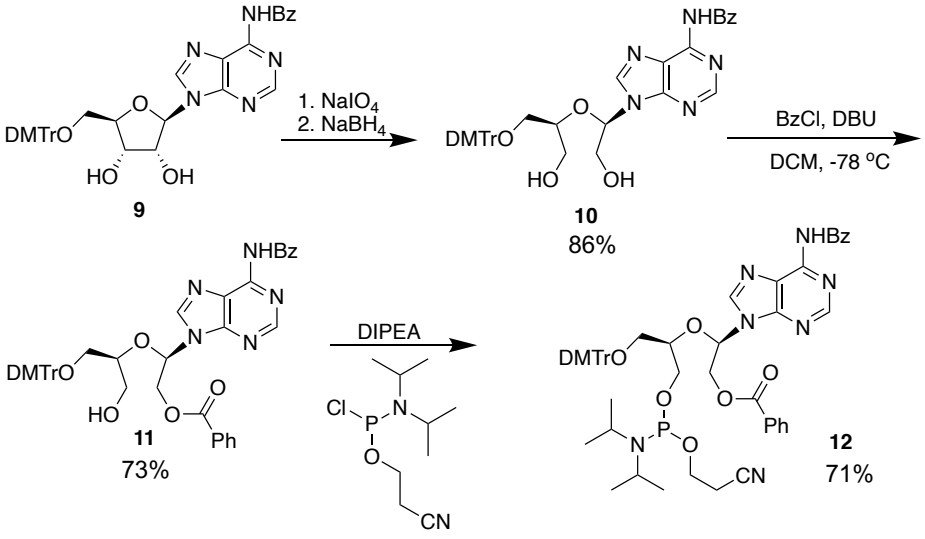
Considering that the field of therapeutic RNA editing is emergent, there is still a need to regulate bystander editing. The strategic positioning of 4'-*C*-Me monomers in ADAR-guiding oligonucleotides, as described here, is a new way to achieve this goal. While LNA monomers can block off-target editing at sites that are distal to the target site (i.e. approximately six nucleotides away surrounding the target), using LNA modification to block adjacent bystander editing in a 5'-AA target sequence will likely cause a substantial loss of efficiency at the target site. Nevertheless, as we have shown in an RNA model as well as in a therapeutic context that that placement of a 4'-*C*-Me monomer at the -1 position relative to the target adenosine in a 5'-**A**A sequence can eliminate editing at the adjacent bystander adenosine while still maintaining high editing efficiency at the target site (Figure, 7-8). Finally, to bridge the inhibitory effects of LNA and 4'-C-Me monomers seen in vitro to cellular systems, we show that the same patterns of inhibition can be translated in mammalian cells with highly metabolically stable guide oligonucleotides. These patterns can be summarized as (1) 4'-*C*-MeA monomers having the selective duality of allowing or inhibiting RNA editing when they are placed -1 or -2, -3 relative to a target, respectively, and (2) LNA as being substantially inhibitory and non-selective when placed -1, -2, and -3 relative to targets.

These structural observations along with biochemical results hint at further chemical manipulations that can be furnished at the minor groove of -1 and -2 nucleotides in the gRNA for the optimization of more selective site-directed editing systems with ADARs. For instance, bulkier 4'-*C* aliphatic nucleotides can be synthesized, as well as engineered ADARs with I491X and P492X mutants to generate new combinations of shape complementarity (i.e. bump-hole method) between the interface of ADAR's flipping loop and the RNA's minor groove. Such combinations could add to the growing list of directed editing systems capable of highly selective corrections of disease-causing mutations at the RNA level [28,29,31,44,75,76].

In summary, we describe a synthesis of 4'-*C*-MeA phosphoramidite and its use to generate modified ADAR guide strands with novel properties. We found that 4'-*C*-MeA behaves similarly to adenosine with respect to its thermal stability in oligoribonucleotides and is slightly more labile to nucleases. 4'-*C*-MeA-modified guides facilitated the identification of the mechanistic origin of inhibition of ADAR using site-specific LNA modification, while also shedding additional light on ADAR's base-flipping mechanism (i.e. illustrating the importance of conformational changes happening at the -1, -2, and orphan nucleotides). ADAR inhibition arises for different reasons depending on the placement of the LNA in the guide oligonucleotides relative to the target. An LNA monomer at the -1 position relative to the target renders no editing due to the inhibition of the re-puckering that the -1 sugar must undertake to stabilize the base-flipped complex. Whereas LNA and 4'-*C*-methyl modification at -2 inhibits ADAR due to steric clash with P492. In conclusion, we have provided a framework that describes guide oligonucleotides with different patterns of chemical modifications containing LNAs and 4'-*C*-Me monomers that can be used to achieve efficient ADAR editing while also blocking editing at bystander sites, including within challenging 5'-AA sequences.



Scheme 1. Synthesis of 4'-*C*-methyladenosine phosphoramidite building block. Scheme 1. Synthesis of the 4'-C-MeA phosphoramidite. Reagents and conditions: (a) $SnCl_4$, N^6-Benzoyladenine, CH_2Cl_2, 0 – 25 °C, 63 %; (b) K_2CO_3, MeOH, 25 °C, 88 %; c) 1M Tetrabutylammonium fluoride/THF, 25 °C, 79 %; (d) 1M BCl_3/CH_2Cl_2, -78 °C, 41 %; (e) DMTr-Cl, pyridine, DMAP, 25 °C, 70 %; (f) *t*-Bu(Me)$_2$Si-Cl, pyridine, $AgNO_3$, THF, 0 – 25 °C, 64 %; (g) 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite, (*i*-Pr)$_2$NEt, 25 °C, 77 %;



Scheme 2. Synthesis of unlocked nucleic acid adenosine phosphoramidite building block [73]

Synthesis General Information

Reagents for the synthesis of 4'-*C*-methyladenosine phosphoramidite were purchased from Ambeed, Sigma-Aldrich, Fisher Scientific, Thomas Scientific, or TCI, and were used without further purification. Solvents were purchased from Sigma-Aldrich and Fisher Scientific. Anhydrous solvents were obtained from in-house solvent purification systems (dried by passage through activated alumina columns under Ar pressure) or drying 24 hr with 200 °C activated 3 Å molecular sieves. Deuterated NMR solvents were purchased from Sigma-Aldrich. All reactions considered air/water sensitive were carried on flame-dried vessels and were sealed under argon atmospheres using Schlenk techniques. Analytical thin-layer chromatography was carried out on glass-backed silica gel plates (Merck/EMD Millipore, 60 Å pore size), and spots were visualized with UV irradiation (254 nm) or stained with potassium permanganate or 4-anisaldehyde. Flash column chromatography was accomplished using silica gel 60 Å (40-60 μm particle size) purchased from Fisher Scientific. NMR spectra were obtained on 400 MHz Bruker Avance III HD Nanobay Spectrometer and 300 MHz Varian Mercury Plus Spectrometer. Chemical shifts (δ) are reported in ppm; coupling constants, J, are reported in Hz. Signals are reported as singlet (s), doublet (d), triplet (t), quartet (q), and multiplet (m). Chemical shifts are reported relative to tetramethylsilane (TMS) and referenced to the appropriate residual solvent peaks for 1H and $^{13}C\{^{1}H\}$ NMR. For the solid-phase oligonucleotide synthesis, an ABI394 Synthesizer was employed from Biolytic. Canonical ribose U, *i*-Pr-Pac-G, A-Pac, Ac-C phosphoramidites and controlled pore glass solid supports were purchased from Glen Research, as well as the activator, cap A mix, cap B mix, oxidizing solution, and deblocking mix ancillary reagents. Modified nucleotide phosphoramidites such as 2'-*O*-methylated A, G, U and C and Locked Nucleic Acid A, G, and T were also purchased from Glen Research. Unlocked Nucleic Acid A phosphoramidite was synthesized by previously reported strategies[73]. Starting material was synthesized by previously reported synthesis [58,59]. 3 Å molecular sieves trap to dry the activator reagent and acetonitrile were purchased from Biolytic. Anhydrous DNA Synthesis

grade acetonitrile and HPLC grade dichloromethane were purchased from Fisher Scientific and Millipore Sigma, respectively.

Mass Spectrometry General Information

For High Resolution MS analysis, samples were analyzed by flow-injection analysis into a Thermo Fisher Scientific QE Exactive HF (Bremen, Germany) operated in the centroided mode. Samples were injected into a mixture of 50 % MeOH and 0.1 % Formic Acid/H_2O at a flow of 200 µL/min. Source parameters were 4 kV spray voltage, capillary temperature of 275 °C, and sheath gas setting of 15. Spectral data were acquired at a resolution setting of 100,000 FWHM with the lockmass feature which typically results in a mass accuracy < 2 ppm.

Oligonucleotides were analyzed by mass spectrometry on a Bruker UltraFlextreme MALDI mass spectrometer (Bruker) in linear positive mode. Samples were first mixed in a 1:1 ratio with a saturated solution of alpha-hydroxycinnamic acid (Milipore Sigma) in nuclease-free water:MeCN (35%:65%) before being spotted on the sample plate and allowed to air dry. The sample plate was then loaded into the high-vacuum region of the MALDI source. Samples were analyzed using the minimum laser fluence to obtain adequate signal (s/n > 20), generally requiring 2000 shots per sample. Data was analyzed in FlexAnalysis. Oligonucleotide masses were determined with Mongo Oligo Calculator v2.08. Oligonucleotides for sequencing and PCR were purchased from Integrated DNA Technologies. Some RNAs containing 2'-*O*-Me A or rA were purchased from Azenta Life Sciences and purified by PAGE. All other oligonucleotides were synthesized via solid-phase synthesis using an ABI 394.

Synthesis of 4'-C-methyladenosine (4'-C-MeA) phosphoramidite

***N*^{6}-Benzoyl-2'-*O*-acetyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-tert-butyldiphenylsilyladenosine (2)**

To a suspension of *N*^{6}-Benzoyladenine (2.07 g, 8.66 mmol, 2 eq.) and an anomeric mixture of the protected and methylated sugar **1** (2.5 g, 4.33 mmol) in anhydrous CH_2Cl_2 (31 mL) at 0 °C was added tin tetrachloride in one portion (0.5 mL, 4.33 mmol, 1 eq.) under an argon atmosphere. After 16 h of stirring at room temperature, the reaction was quenched with 2 mL of saturated $NaHCO_3$ dropwise at 0 °C. The organic layer was diluted with 200 mL of CH_2Cl_2 and extracted twice with 100 mL of a saturated solution of $NaHCO_3$. The combined organics were dried over Na_2SO_4, filtered, and reduced under pressure to afford a white and oily precipitate. The crude mixture was chromatographed with silica gel with 0.3 % MeOH in $CHCl_3$ (R_f = 0.22, 3 % MeOH in CH_2Cl_2) to afford **2** as a light colorless foam in 63 % yield (2.05 g). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.11 (s, 1H), 8.70 (s, 1H), 8.20 (s, 1H), 8.04 (d, *J* = 7.1 Hz, 2H), 7.67 – 7.31 (m, 20H), 6.28 (d, *J* = 5.1 Hz, 1H), 5.98 (t, *J* = 5.5 Hz, 1H), 4.69 (d, *J* = 5.7 Hz, 1H), 4.62 (q, *J* = 38.9 Hz, 2H), 3.88 (d, *J* = 11.0 Hz, 1H), 3.59 (d, *J* = 11.0 Hz, 1H), 2.09 (s, 3H), 1.38 (s, 3H), 1.08 (s, 9H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.95, 169.90, 152.72, 151.55, 149.56, 141.79, 137.55, 135.65, 135.53, 133.73, 132.77, 132.71, 132.60, 129.96, 129.92, 128.86, 128.48, 128.04, 127.92, 127.88, 127.84, 123.37, 87.58, 86.01, 75.01, 74.25, 68.06, 30.94, 26.95, 20.72, 19.24, 18.53. Orbitrap-HRMS calculated for $C_{43}H_{45}N_5O_6Si$ $[M+H]^+$ 756.31 m/z, found 756.3234 m/z.

***N*^{6}-Benzoyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine (3)** To a solution of **2** (0.214 g, 0.283 mmol) in 11 mL of MeOH was added K_2CO_3 (0.473 g, 3.42 mmol, 12 eq.) at 0 °C and stirred at this temperature for 1.5 h. After completion of the reaction, the mixture was neutralized with 10 mL of aqueous 10 % HCl (v/v). The aqueous phase was separated and extracted three times with 10 mL of CH_2Cl_2. The combined organics were dried over Na_2SO_4, filtered, and reduced under pressure to afford a colorless syrup. The crude mixture was purified by silica gel chromatography with a gradient of 1.75 to 2.25 % MeOH in dichloromethane (R_f = 0.43, 4 % MeOH in CH_2Cl_2) to afford **3** as a colorless foam in 88 % yield (178 mg). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.18 (s, 1H), 8.68 (s, 1H), 8.08 (s, 1H), 8.03 (d, *J* = 6.9 Hz, 1H), 7.71 – 7.25 (m,

20H), 6.01 (d, *J* = 5.8 Hz, 1H), 4.92 (t, *J* = 5.8 Hz, 1H), 4.74 (q, 2H), 4.41 (d, *J* = 5.8 Hz, 1H), 3.86 (d, *J* = 10.8 Hz, 1H), 3.88 (s, 1H), 3.58 (d, *J* = 10.8 Hz, 1H), 1.44 (s, 3H), 1.08 (s, 9H). ^{13}C NMR (76 MHz, $CDCl_3$) δ 152.47, 151.36, 149.56, 141.86, 137.13, 135.63, 135.53, 133.70, 132.77, 132.71, 132.60, 130.03, 128.84, 128.73, 128.40, 128.15, 127.94, 127.90, 127.88, 123.27, 89.00, 87.52, 79.34, 77.51, 77.08, 76.66, 74.68, 74.51, 68.46, 26.95, 19.27, 19.01. Orbitrap-HRMS calculated for $C_{41}H_{43}N_5O_5Si$ $[M-H]^-$ 712.29 m/z, found 712.2974 m/z.

***N*^{6}-Benzoyl-3'-*O*-benzyl-4'-*C*-methyladenosine (4) 3** (0.178 g, 0.249 mmol) was co-evaporated twice with anhydrous THF. Subsequently, the silyl-protected nucleoside was dissolved in an anhydrous 1M TBAF solution in THF (0.5 mL) and stirred for 24 h at room temperature under an argon atmosphere. After complete conversion of the starting material, the mixture was reduced to an orange-colored oil and chromatographed with silica gel with a 2-3 % MeOH in CH_2Cl_2 eluent to afford **4** as a white solid in 79 % (93 mg). (R_f = 0.35, 5 % MeOH in CH_2Cl_2). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.43 (s, 1H), 8.67 (s, 1H), 7.96 (d, 2H), 7.75 (s, 1H), 7.58 – 7.25 (m, 8H), 6.14 (s, 1H), 5.73 (d, *J* = 7.6 Hz, 1H), 5.13 (t, *J* = 6.6 Hz, 1H), 4.75 (q, 2H), 4.30 (d, *J* = 5.6 Hz, 2H), 3.89 – 3.69 (m, 2H), 3.58 (d, *J* = 12.5 Hz, 1H), 1.86 (q, 1H), 1.36 (s, 3H). ^{13}C NMR (76 MHz, $CDCl_3$) δ 164.59, 151.94, 150.12, 143.19, 137.27, 133.22, 132.84, 128.75, 128.70, 128.41, 128.25, 127.86, 123.92, 91.14, 89.11, 80.78, 75.18, 74.29, 69.02, 67.99, 25.62, 18.67. Orbitrap-HRMS calculated for $C_{25}H_{25}N_5O_5$ $[M+H]^+$ 476.19 m/z, found 476.1935 m/z.

***N*^{6}-Benzoyl-4'-*C*-methyladenosine (5)** Debenzylation procedure was adapted from Zhou *et al.* [77] **4** (0.230 g, 0.484 mmol) was co-evaporated 3x with anhydrous dichloromethane and dried overnight under high vacuum with a stir bar. The 3' benzylated nucleoside was then dissolved in 20 mL of anhydrous dichloromethane and stirred 5 min at -78 °C under an argon atmosphere, followed by the dropwise addition of 1M BCl_3 in CH_2Cl_2 (3.6 mL, 3.6 mmol, 7.5 eq.) over 10 min. After stirring for 4 h at the same temperature, the reaction was quenched with 16 mL of MeOH followed by 3.2 g of $NaHCO_3$. The mixture was left stirring for 4 h at room temperature. The insoluble solid was then filtered through a silica gel pad (2.8 g) and washed with 49 mL of 2:1

DCM:MeOH. The filtrate was evaporated to dryness and chromatographed with silica gel using a 3-8% gradient MeOH in CH_2Cl_2 to afford **5** as a white solid in 41% yield (75 mg) R_f = 0.15 (10 % MeOH in CH_2Cl_2). ^{1}H NMR (400 MHz, DMSO) δ 11.10 (s, 1H), 8.72 (d, *J* = 14.0 Hz, 2H), 8.09 – 8.03 (m, 2H), 7.65 (t, *J* = 7.3 Hz, 1H), 7.55 (t, *J* = 7.6 Hz, 2H), 6.02 (d, *J* = 7.2 Hz, 1H), 5.46 (d, *J* = 6.3 Hz, 1H), 5.30 (t, *J* = 5.8 Hz, 1H), 5.21 (d, *J* = 4.7 Hz, 1H), 4.91 (d, *J* = 6.5 Hz, 1H), 4.11 (t, *J* = 4.6 Hz, 1H), 3.57 (dd, *J* = 11.6, 4.8 Hz, 1H), 3.44 – 3.35 (m, 3H), 1.20 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 166.39, 152.72, 152.00, 143.75, 134.17, 132.80, 128.98, 128.90, 126.40, 87.94, 87.89, 87.36, 74.44, 72.27, 67.39, 40.60, 40.39, 40.18, 39.97, 39.76, 39.55, 39.34, 19.30. Orbitrap-HRMS calculated for $C_{18}H_{19}N_5O_5$ $[M+H]^+$ 386.14 m/z, found 386.1475 m/z.

5'-*O*-(*4,4'*-Dimethoxytrityl)-N^6-benzoyl-4'-*C*-methyladenosine (6) After co-evaporating **5** (0.118 g, 0.300 mmol) 3x with anhydrous pyridine (3x2 mL) and drying it overnight under high vacuum, the nucleoside was dissolved with distilled and anhydrous pyridine (2.0 mL). DMTr-Cl (0.124 g, 0.367 mmol, 1.2 eq.) was added in one portion to the solution, followed by the catalytic addition of 4-Dimethylaminopyridine (DMAP) (1.7 mg, 0.014 mmol, 0.046 eq.). The solution was stirred for 24 h under argon and at room temperature and quenched with 1 mL of MeOH. All the volatiles were evaporated and the resulting crude was diluted with 30 mL of CH_2Cl_2, washed 1x with 5 % aqueous $NaHCO_3$ (10 mL), 1x with brine (10 mL), dried over Na_2SO_4 and filtered. The crude was purified via silica gel chromatography (base washed with 1% TEA) with a gradient of 0 – 2.75 % MeOH in CH_2Cl_2 to give **6** as a light yellow foam in 70 % yield (144 mg). R_f = 0.18 (4 % MeOH in CH_2Cl_2). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.18 (s, 1H), 8.71 (s, 1H), 8.22 (s, 1H), 8.03 (d, *J* = 7.5 Hz, 2H), 7.58 (dt, *J* = 28.4, 7.4 Hz, 3H), 7.37 – 7.09 (m, 10H), 6.74 (d, *J* = 8.3 Hz, 4H), 6.04 (d, *J* = 6.2 Hz, 1H), 5.93 (s, 1H), 5.01 (t, *J* = 5.8 Hz, 1H), 4.34 (d, *J* = 5.2 Hz, 1H), 3.76 (s, 6H), 3.28 (d, *J* = 10.0 Hz, 1H), 3.14 (d, *J* = 10.0 Hz, 1H), 1.43 (s, 3H), 0.89 (d, *J* = 9.2 Hz, 1H). ^{13}C NMR (76 MHz, $CDCl_3$) δ 158.51, 152.10, 150.93, 149.61, 144.33, 141.55, 135.38, 135.28, 133.50, 132.96, 130.00, 129.16, 128.94, 127.97, 127.90, 127.84, 126.92, 123.11, 113.13, 90.46,

89.13, 86.71, 77.47, 77.05, 76.63, 76.18, 73.71, 68.47, 55.23, 29.73, 18.98. Orbitrap-HRMS calculated for $C_{39}H_{37}N_5O_7$ $[M+H]^+$ 688.27 m/z, found 688.2744 m/z.

5′-*O*-(*4,4′*-Dimethoxytrityl)-2′-*O*-(*tert*-butyldimethylsilyl)-*N*^{6}-benzoyl-4'-*C*-methyladenosine (7) Prior to silylation, **6** was co-evaporated 3x with anhydrous pyridine and dried under high vacuum overnight. To a solution of **6** (0.124 g, 0.180 mmol) in anhydrous pyridine and THF (2.25 mL, 1:1) was added silver nitrate (0.035 g, 0.200 mmol, 1.16 eq.) under an argon atmosphere at 0 °C. After stirring for 5 min, TBS-Cl (0.030 g, 0.200 mmol, 1.16 eq.) was added in one portion at the same temperature. The reaction was allowed to stir overnight at room temperature. After the starting material was consumed, the reaction was diluted with 10 mL of ethyl acetate and filtered. The filtrate was collected and evaporated to a yellow oil, diluted with 10 mL of CH_2C_2, and washed once with 5 % aqueous $NaHCO_3$. After drying and filtering the organic phase with Na_2SO_4, the crude was chromatographed with a base-washed (0.5 % TEA) silica gel column with a gradient of 0-40 % ethyl acetate in *n*-hexanes (R_f = 0.41, 8:3 *n*-hexanes:EtOAc). Purification afforded **7** as a colorless foam in 64 % yield (92 mg). ^{1}H NMR (300 MHz, CD_2Cl_2) δ 9.16 (s, 1H), 8.68 (s, 1H), 8.18 (s, 1H), 8.03 (d, *J* = 7.6 Hz, 2H), 7.73 – 7.15 (m, 12H), 6.98 – 6.79 (m, 4H), 6.09 (d, *J* = 5.8 Hz, 1H), 5.19 (t, *J* = 5.5 Hz, 1H), 4.42 (t, *J* = 3.6 Hz, 1H), 3.82 (s, 6H), 3.45 (d, *J* = 9.9 Hz, 1H), 3.25 (d, *J* = 9.9 Hz, 1H), 2.87 (d, *J* = 3.6 Hz, 1H), 2.04 (s, 1H), 1.44 (s, 3H), 0.88 (s, 9H), 0.01 (s, 3H), -0.17 (s, 3H). ^{29}Si NMR (60 MHz, CD_2Cl_2) δ 24.61. ^{13}C NMR (76 MHz, CD_2Cl_2) δ 164.42, 158.70, 152.43, 151.61, 149.65, 144.82, 142.14, 135.65, 135.60, 134.06, 132.63, 130.14, 130.11, 128.82, 128.12, 127.90, 127.76, 126.89, 123.49, 113.15, 88.16, 86.83, 86.58, 75.80, 72.63, 68.01, 55.21, 29.69, 25.31, 18.86, 17.73, -5.33, -5.44. Orbitrap-HRMS calculated for $C_{45}H_{51}N_5O_7Si$ $[M+H]^+$ 802.36 m/z, found 802.3658 m/z.

5′-*O*-(*4,4′*-Dimethoxytrityl)-3′-*O*-[(2-cyanoethoxy)(*N,N*-diisopropylamino)phosphino]-2′-*O*-(*tert*-butyldimethylsilyl)-*N*^{6}-benzoyl-4'-*C*-methyladenosine (8) **7** was co-evaporated 3x with anhydrous CH_2Cl_2 (8 mL) and dried overnight under P_2O_5 and high vacuum. To a solution of **7** (0.092 g, 0.114 mmol), DMAP (2.7 mg, 0.0228 mmol, 0.2 eq.), distilled over CaH_2 *N,N*-

diisopropylethylamine (60 µL, 0.342 mmol, 3 eq.), and 0.45 mL of anhydrous CH_2Cl_2 was added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (50 µL, 0.230 mmol, 2 eq.) under an argon atmosphere. The reaction was allowed to stir overnight at room temperature. The reaction was quenched with 0.1 mL of MeOH and further diluted with dichloromethane. The organic phase was then washed with 5 % aqueous $NaHCO_3$ once, dried over Na_2SO_4 and filtered to afford a yellow-orange oil. The residue was purified over a 0.5 % TEA-washed silica gel column with a gradient of 30 – 45 % ethyl acetate in *n*-hexanes to furnish the phosphoramidite **8** as a colorless foam in 77 % yield (88 mg). (Diastereomers of **8**, R_f = 0.64, 0.52 in 54% Ethyl acetate in *n*-hexanes). ^{31}P NMR (122 MHz, CD_2Cl_2) δ 150.19, 149.87. Orbitrap-HRMS calculated for $C_{54}H_{68}N_7O_8PSi$ $[M+H]^+$ 1002.47 m/z, found 1002.4713 m/z.

Synthesis of Unlocked Nucleic Acid A (UNA A) phosphoramidite. The synthesis of the UNA A was performed as previously described [73].

Solid-phase Oligonucleotide Synthesis and Purification. All oligonucleotides were synthesized in 0.2 µmol scale based on the phosphoramidite chemistry and standard DNA/RNA reagents and cycles. DMTr-Off and 2'-*O*-TBDMSi protecting strategy was employed for all RNA oligonucleotide synthesis. The coupling time for all phosphoramidites was between 3-10 min. For incorporation of the 4'-*C*-MeA phosphoramidite, an extended coupling time of 45 min was used and was dissolved with a 3:1 mixture of MeCN:DCM to 0.1 M. Before its incorporation, the purified phosphoramidite was co-evaporated 3x with anhydrous DCM, filtered using syringe filter and was dried under high vacuum overnight over P_2O_5. Coupling efficiency of the 4'-*C*-MeA was measured by means of the conductivity released by the trityl cation during detritylation of the growing strand with 3% trichloroacetic acid in CH_2Cl_2 and was comparable with commercially available phosphoramidites detritylations. Post synthesis, columns were dried for 2 h under high vacuum. Cleavage from the solid support and deprotection of the nucleobases were accomplished using

a fresh aliquot of ammonium hydroxide in ethanol at 55 °C for 12 h (1 mL of 28-30% NH_4OH/EtOH in 0.5 mL EtOH). After cleavage, the solid support was allowed to cool to room temperature and centrifuged for 5 min at 13,000 x g. The supernatants were collected and lyophilized. 2'-*O*-silyl ether deprotection was done using 250 μL of anhydrous 1M Tetrabutylammonium fluoride (TBAF) in THF for 24 h at room temperature. Crude oligonucleotides were desalted by precipitation with the addition of 1 mL of 1-butanol and 50 μL of 3 M NaOAc and cooled for 2 h at -70 °C. The solution was centrifuged at 13,000 × g for 30 min, supernatants were removed, and the pellets were washed twice with cold 95% ethanol.

The oligonucleotides were purified by loading a 1:1 mixture of 10-20 nmol of crude RNA to 80% formamide in 1X Tris-Borate EDTA (TBE) and 1mM EDTA into a denaturing polyacrylamide gel. Depending on their size, oligonucleotides were loaded into different acrylamide matrices with different percentages and electrophoresed in 1X TBE. After visualization by UV shadowing on glass-backed silica gel plates, gel bands were excised, crushed, and soaked overnight at 4 °C in 0.5 mL of 0.5 M NH_4OAc and 0.1 mM EDTA. Gel particles were removed from the extracted RNA by filtering the suspension for 5 min, 13,000 x g on 0.2 μm cellulose sterile membrane filter tubes (Corning). The oligonucleotides were precipitated from a solution of 75% EtOH at −70 °C for 2 h. The oligonucleotides were then pelleted from the solution by centrifugation (13,000 × g for 30 min), with the supernatant being discarded. The pellets were lyophilized to dryness, resuspended in nuclease-free water, and quantified by absorbance at 260 nm. Oligonucleotide mass was confirmed by MALDI-TOF (Table 1.).

Target RNA Synthesis*.* The gene fragment of the corresponding human MECP2, mouse IDUA, human SRC, modified IDUA and MECP2*,* and human PPP2R5D were purchased from Integrated DNA Technologies. These were amplified with New England BioLabs (NEB) Q5 Hot Start high-fidelity DNA polymerase and purified by extraction from 1% agarose gels. Target RNA was

synthesized with NEB's HiScribe T7 RNA polymerase according to the protocol. The target RNAs were purified by 8% denaturing polyacrylamide gel electrophoresis.

Hybridization of guide and target RNA. The target and guide RNA were combined in a 1:10 ratio of target:guide to a final concentration of 180 nM of hybrids in 1X TE and 100 mM NaCl. This solution was heated to 95 °C for 5 min then it was allowed to cool to room temperature for 2 h.

ADAR1 p110 Overexpression and Purification*.* Human ADAR1 p110 (UniProtKB P55265-5) with a C-terminal intein fusion protein with chitin binding domain (CBD) was overexpressed in *S. cerevisiae* BCY123 as previously described [78]. Human ADAR1 p110 was purified by lysing cells in 20 mM Tris−HCl pH 8.0, 5% (v/v) glycerol, 750 mM NaCl, 50 mM imidazole, 1 mM tris(2-carboxyethyl)phosphine-HCl (TCEP-HCl), 50 μM $ZnCl_2$ and EDTA-free Protease Inhibitor using a microfluidizer. The lysate was centrifuged at 39,000 x g for 45 min at 4 ˚C. The clarified lysate was passed over an equilibrated chitin binding column using gravity flow and was incubated for 2 h at 4 ˚C for binding. The chitin column was then washed with 10 column volumes (CV) of lysis buffer and subsequently, the column was quickly washed with 5 bed volumes of the cleavage buffer (20 mM Tris−HCl pH 8.0, 5% (v/v) glycerol, 350 mM NaCl, 50 mM imidazole, 50 mM DTT and 50 μM $ZnCl_2$. To ensure cleavage of the CBD domain, the resin was incubated in 2 CVs of cleavage buffer overnight at room temperature for 18 h. The protein was eluted by washing columns with an additional 5 CVs of cleavage buffer. Finally, the eluted target protein was pooled, concentrated to 0.3 mg/ml, and then dialyzed against a storage buffer containing 20 mM Tris−HCl pH 8.0, 20% (v/v) glycerol, 350 mM NaCl, 50 mM imidazole, 1 mM TCEP-HCl. Protein concentration was determined by running the sample alongside bovine serum albumin (BSA) standards in a sodium dodecyl sulfate-PAGE (SDS-PAGE) gel followed by SYPRO Orange (Invitrogen) staining. The protein was stored at -80˚C in aliquots for further use.

ADAR2 Wild-type (WT) and ADAR2 E488Q Overexpression and Purification. ADAR2 wild-type and E488Q mutant overexpression and purification were performed as previously described [79].

***In vitro Deamination Reaction and Sanger Sequencing*.** For ADAR2 WT and ADAR2 E488Q deaminations, hybridized target and guide RNA were diluted to 5 nM in 1X ADAR2 reaction buffer (15 mM Tris-HCl pH 7.5, 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, and 3 mM $MgCl_2$), 0.5 mM dithiothreitol (DTT), 160 units/mL RNAse inhibitor, 1 μg/mL yeast tRNA, and 15 nM ADAR2 protein. The reactions were conducted at 30 °C. For two time points reactions analysis, as well as for kinetic analysis (1, 3, 5, 10, 15, 30, 60 min), aliquots of 8 μL were quenched with 190 μL of 95 °C heated nuclease-free water for 5 min. For ADAR1 p110 deaminations, hybridized target and guide RNA were diluted to 15 nM in 1x ADAR1p110 reaction buffer (15 mM Tris-HCl pH 7.5, 3% glycerol, 26 mM KCl, 40 mM potassium glutamate, 1.5 mM EDTA, 0.003% Nonidet P-40, and 4% glycerol), 0.5 mM DTT, 160 units/mL RNAse inhibitor, 1 μg/mL yeast tRNA, and 150 nM ADAR1 p110. Due to the lower enzymatic activity of overexpressed and purified ADAR1 p110, a higher concentration of molar ratio of ADAR1 than ADAR2 was used to optimize deamination reactions. The reactions were conducted the same way as ADAR2. A 5 μL aliquot of each time point quench solution was reverse transcribed with Access Reverse Transcription (RT) Polymerase Chain Reaction (RT-PCR) system purchased from Promega. DNA was cleaned and concentrated using Zymo's purification kit and protocol. For Sanger sequencing with Azenta Life Sciences, purified DNA was diluted to 0.66 ng/μL in 1.6 μM of forward or reverse sequencing primer. Sanger sequencing traces were observed with SnapGene Viewer and statistical analysis along with nonlinear fits were conducted with Microsoft Excel and GraphPad Prism. The editing level for the corresponding zero time point was subtracted from each data point as background subtraction.

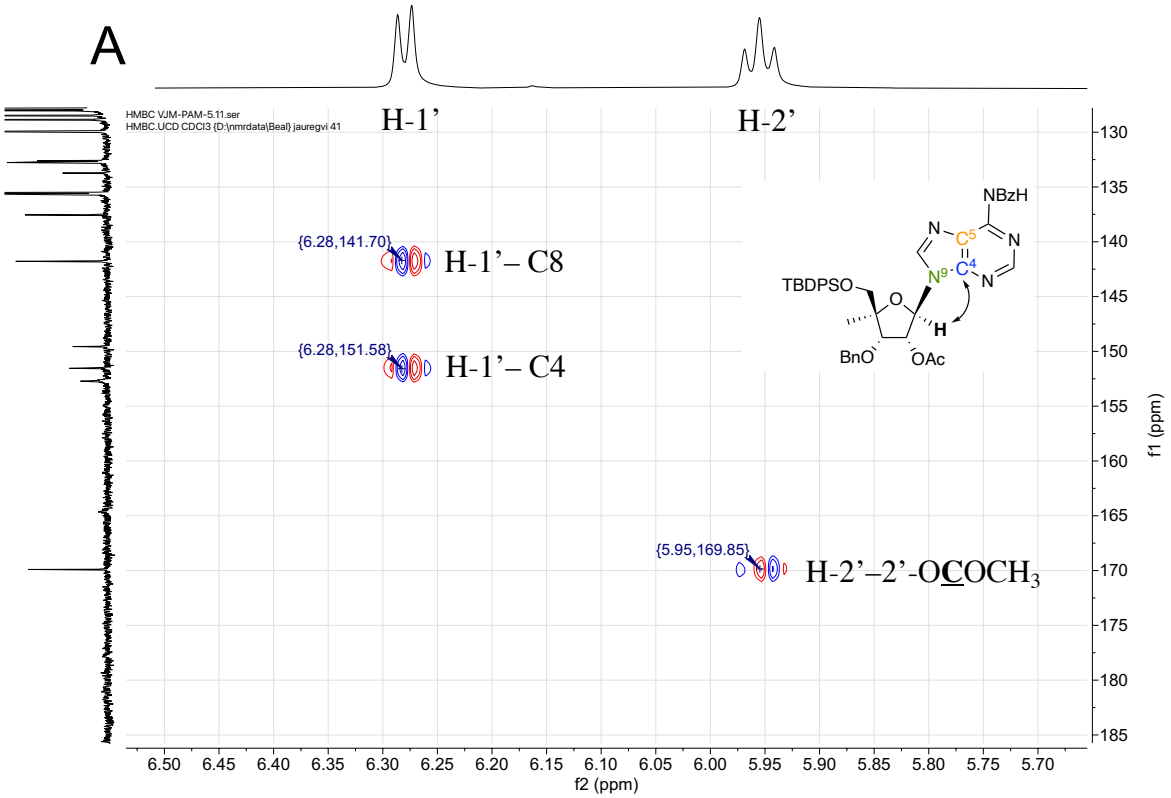
Oligonucleotide Metabolic Stability Assay. A 60 µM stock solution of oligonucleotide was made in IDT Nuclease-Free Buffer (30 mM HEPES, pH 7.5; 100 mM potassium acetate). In a 50 µL reaction volume, the oligonucleotide sample was diluted to a final concentration of 8 µM with Dulbecco's Phosphate Buffer Saline (DPBS, Gibco 14190144) in 80% human serum (Millipore Sigma H4522). Each reaction was incubated at 37 °C and 5 uL aliquots were taken out at 0, 2, 4, 6, 8, 10, 12, and 16 hours and combined with 5 uL of 80% formamide in 1X TBE and 1 mM EDTA, vortexed and immediately stored at -80 °C. Each aliquoted time point containing 40 pmol of oligonucleotide was loaded into a 15% denaturing Novex TBE Gels (Invitrogen, EC62752BOX or EC63152BOX) and electrophoresed at 200 V for 60 min in 1X TBE. The gel was then stained in SYBR Gold Nucleic Acid Gel Stain (Invitrogen S11494) for 30 min and imaged on the GelDoc Go Imaging System (Bio-Rad) under SYBR Gold settings on a UV/Stain-free tray. The densitometry of RNA bands was determined using Image Lab and plotted by a function of time using GraphPad Prism.

***Oligonucleotide Melting Temperature Assay*.** Melting temperature analysis of oligonucleotides was performed as previously described [20].

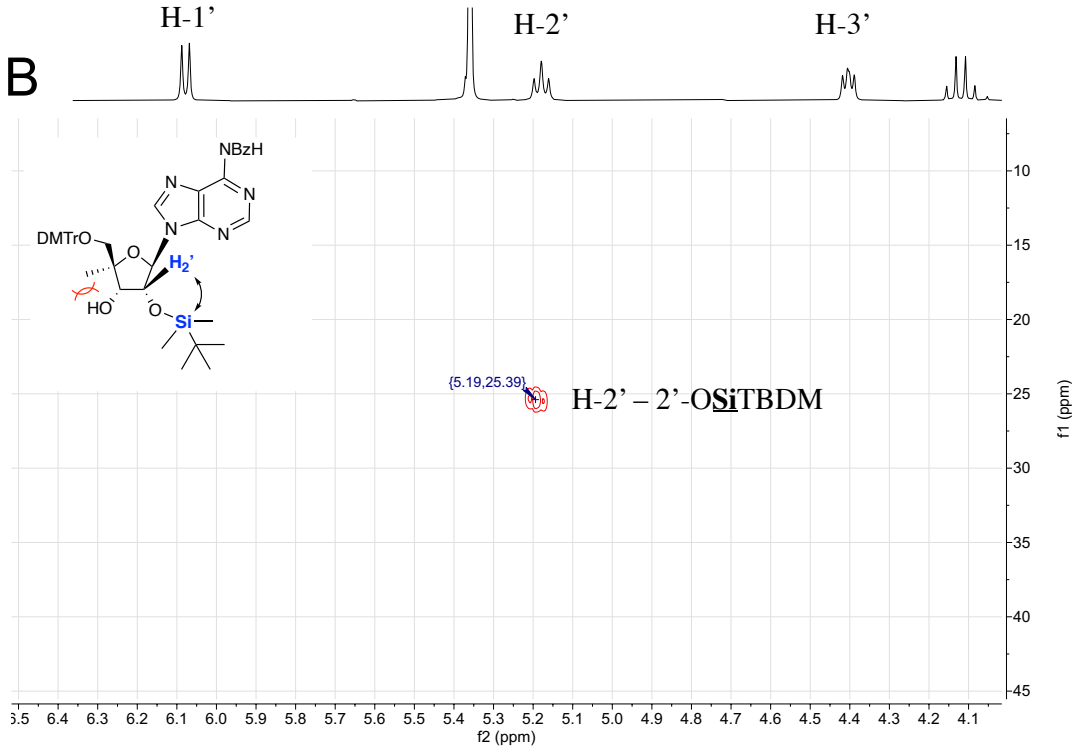
Directed Editing on Endogenous β-actin in HEK293T Cells and Analysis. The procedure was performed as previously described [36].

2-D NMR Correlations

(2) N^6-Benzoyl-2'-*O*-acetyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine ^{1}H-^{13}C HMBC – (400 MHz, $CDCl_3$)



(7) 5'-*O*-(4,4'-Dimethoxytrityl)-2'-*O*-(*tert*-butyldimethylsilyl)-N^6-benzoyl-4'-*C*-methyladenosine ^{1}H-^{21}Si HMBC– (300-76 MHz, CD_2Cl_2)



Oligonucleotide Sequences and Masses

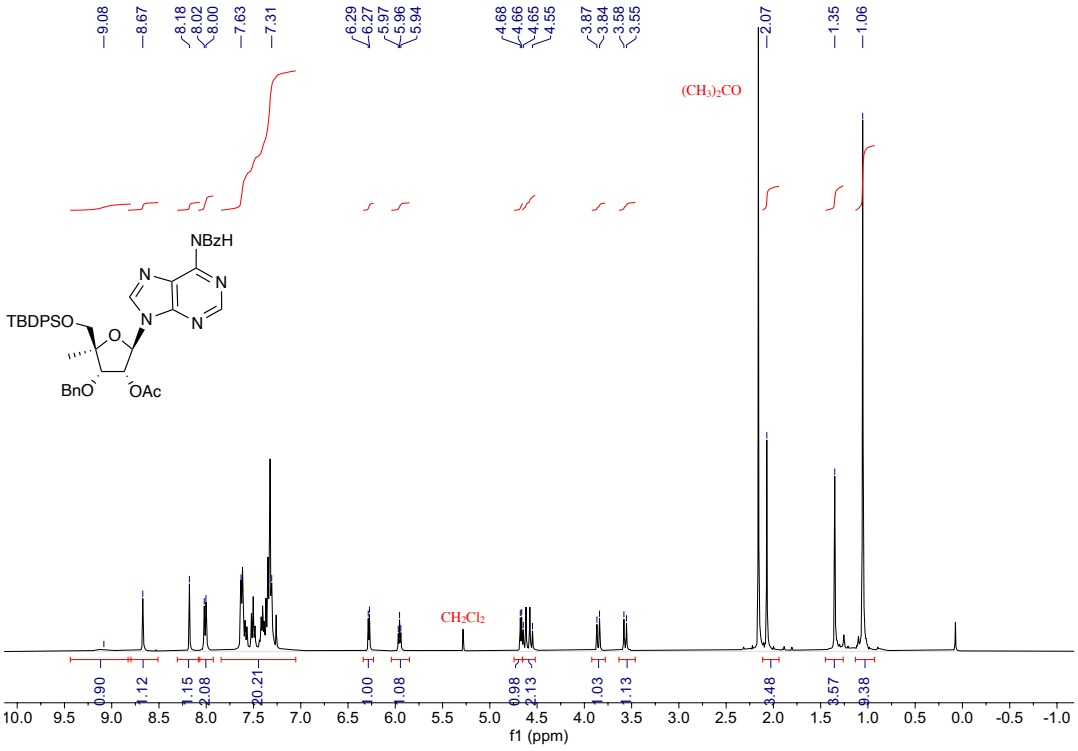
Table 1. Sequences of oligonucleotides. Primers are 2'-deoxynucleotides, all others are ribonucleotides unless otherwise stated. Underlined nucleotides indicate 2'-*O*-Me, target adenosines are indicated in red, nucleotides in (parenthesis) are modified and represent the -1 or -2 nucleotide. Subscript $_L$ indicates LNA sugar. Oligonucleotides were measured in linear positive mode analysis which typically results in mass accuracies of 0.1% or better. Standard oligonucleotides (commercially bought from Azenta Life Sciences and PAGE purified) were measured and showed mass accuracies better than 0.1% error.

Oligonucleotide	Sequence
Human *MECP2* W104X target sequence	5'GGGAGCCCGCAGAGGCAGGCAAAGCAGAGACAUCAGAAGGGUCA GGCUCCGCCCCGGCUGUGCCGGAAGCUUCUGCCUCCCCCAAACAG CGGCGCUCCAUCAUCCGUGACCGGGGACCCAUGUAUGAUGACCCC ACCCUGCCUGAAGGCUAGACACGGAAGCUUAAGCAAAGGAAAUCUG GCCGCUCUGCUGGGAAGUAUGAUGUGUAUUUGAUCAAUCCCCAGG GAAAAGCCUUUCGCUCUAAAGUGGAGUUGAUUGCGUACUUCGAAAA GGUAGGCGACACAUCCCUGGACCCUAAUGA-3'
Modified Human *MECP2* W104X target sequence (modification in blue)	5'GGGAGCCCGCAGAGGCAGGCAAAGCAGAGACAUCAGAAGGGUCA GGCUCCGCCCCGGCUGUGCCGGAAGCUUCUGCCUCCCCCAAACAG CGGCGCUCCAUCAUCCGUGACCGGGGACCCAUGUAUGAUGACCCC ACCCUGCCUGAAGGUUAGACACGGAAGCUUAAGCAAAGGAAAUCUG GCCGCUCUGCUGGGAAGUAUGAUGUGUAUUUGAUCAAUCCCCAGG GAAAAGCCUUUCGCUCUAAAGUGGAGUUGAUUGCGUACUUCGAAAA GGUAGGCGACACAUCCCUGGACCCUAAUGA-3'
Mouse *IDUA* W392X target sequence	5'GGGCUCCUCCCAUCCUGUGGGCUGAACAGUAUAACAGACUCCCA GUAUACAAAUGGUGGGAGCUAGAUAUUAGGGUAGGAAGCCAGAUGC UAGGUAUGAGAGAGCCAACAGCCUCAGCCCUCUGCUUGGCUUAUAG AUGGAGAACAACUCUAGGCAGAGGUCUCAAAGGCUGGGGCUGUGU UGGACAGCAAUCAUACAGUGGGUGUCCUGGCCAGCACCCAUCACCC UGAAGGCUCCGCAGCGGCCUGGAGUACCACAGUCCUCAUCUACACU AGUGAUGACACCCACGCACACCCCGGAUCC-3'
Modified Mouse *IDUA* UAAG target sequence (modification in blue)	5'GGGCUCCUCCCAUCCUGUGGGCUGAACAGUAUAACAGACUCCCA GUAUACAAAUGGUGGGAGCUAGAUAUUAGGGUAGGAAGCCAGAUGC UAGGUAUGAGAGAGCCAACAGCCUCAGCCCUCUGCUUGGCUUAUAG AUGGAGAACAACUUAAGGCAGAGGUCUCAAAGGCUGGGGCUGUGU UGGACAGCAAUCAUACAGUGGGUGUCCUGGCCAGCACCCAUCACCC UGAAGGCUCCGCAGCGGCCUGGAGUACCACAGUCCUCAUCUACACU AGUGAUGACACCCACGCACACCCCGGAUCC-3'
Human *SRC* K295R target sequence	5'GGGCCACGUCCAAGCCGCAGACUCAGGGCCUGGCCAAGGAUGCC UGGGAGAUCCCUCGGGAGUCGCUGCGGCUGGAGGUCAAGCUGGGC CAGGGCUGCUUUGGCGAGGUGUGGAUGGGGACCUGGAACGGUACC ACCAGGGUGGCCAUCAAAACCCUGAAGCCUGGCACGAUGUCUCCAG AGGCCUUCCUGCAGGAGGCCCAGGUCAUGAAGAAGCUGAGGCAUG AGAAGCUGGUGCAGUUGUAUGCUGUGGUUUCAGAGGAGCCCAUUU ACAUCGUCACGGAGUACAUGAGCAAGGGGAG3'
Human *MECP2* W104X gRNAs (-1 varied)	5'-UAAGCUUCCGUGUCC(A)GCCUUCAGGCAGG-3'
Modified Human *MECP2* W104X gRNAs	5'-UAAGCUUCCGUGUCCA(A)CCUUCAGGCAGG-3'
Mouse *IDUA* W392X gRNAs (-1 varied)	5'-UUGAGACCUCUGCCC(A)GAGUUGUUCUCCA-3'
Modified Mouse *IDUA* UAAG gRNA	5'-UUGAGACCUCUGCCU(U)(A)AGUUGUUCUCCA-3
SRC gRNAs	5'-CAGGCUUCAGGGUUC(U)GAUGGCCACCCUG-3'
Standard DNA oligo sequence for MALDI	5'-TTACGCCAGAATGCGTTCGCACAGCCGCCA-3'
MECP2 RT-PCR and sequencing forward primer	5'-AGAGGCAGGCAAAGCAGA-3'
MECP2 amplification and RT-PCR reverse primer	5'-TCATTAGGGTCCAGGGAT-3'

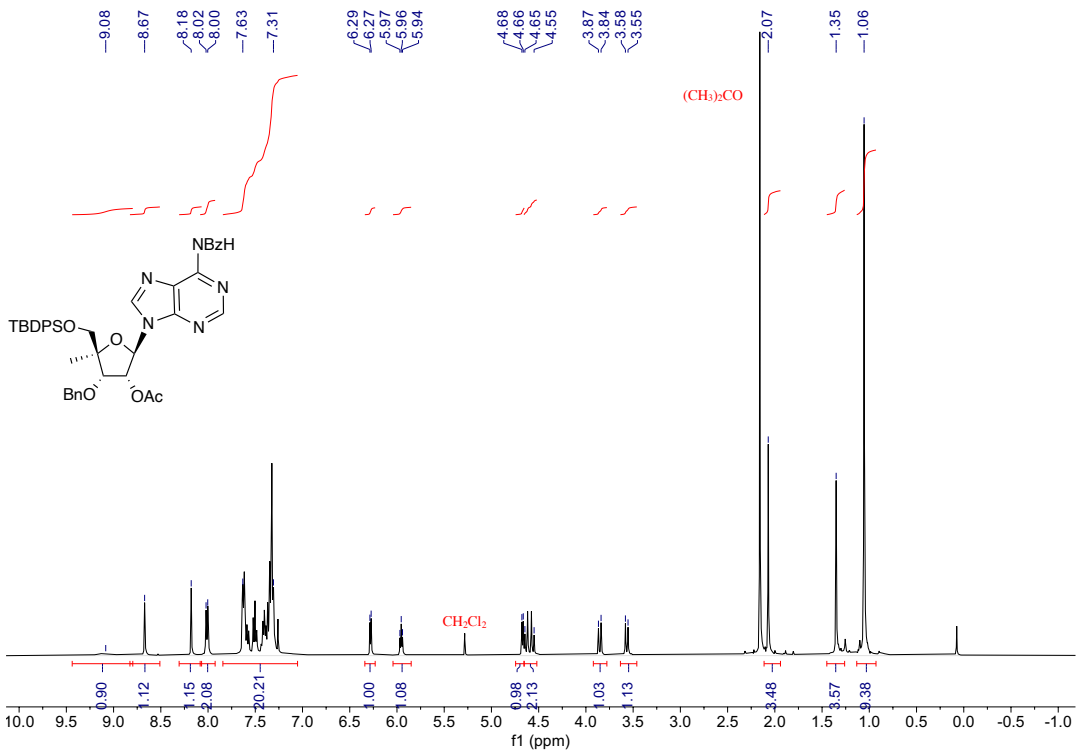
MECP2 T7 DNA amplification forward primer	5’-TAATACGACTCACTATAGGGAGC-3’
IDUA RT-PCR primer forward	5’-CTCCCATCCTGTGGGCT-3’
IDUA amplification, RT-PCR and sequencing primer reverse	5’-GGATCCGGGGTGTGC-3’
IDUA T7 DNA amplification forward primer	5’-TAATACGACTCACTATAGGGCTC-3’
SRC RT-PCR primer forward	5’-TCCAAGCCGCAGACTCAG-3’
SRC amplification, RT-PCR and sequencing primer reverse	5’-CTCCCCTTGCTCATGTACTCC-3’
SRC T7 DNA amplification forward primer	5’-TAATACGACTCACTATAGGGCCAC-3’
+1 LNA T *SRC* gRNA	5’-CAGGCUUCAGGGU**T**$_L$CUGAUGGCCACCCUG-3’
+2 LNA T *SRC* gRNA	5’-CAGGCUUCAGGG**T**$_L$UCUGAUGGCCACCCUG-3’
+3 LNA G *SRC* gRNA	5’-CAGGCUUCAGG**G**$_L$UUCUGAUGGCCACCCUG-3
-2 LNA G *SRC* gRNA	5’-CAGGCUUCAGGGUUCU**G**$_L$AUGGCCACCCUG-3’
-3 LNA A *SRC* gRNA	5’-CAGGCUUCAGGGUUCUG**A**$_L$UGGCCACCCUG-3
-4 LNA T *SRC* gRNA	5’-CAGGCUUCAGGGUUCUGA**T**$_L$GGCCACCCUG-3
Top Control rA Tm	5’-CAUUAAGGUGGG-3’
Top -1 4’-*C*-MeA Tm	5’-CAUUA**A**$_{4'\text{-Me}}$GGUGGG-3’
Bottom rU Tm	5’-CCUACCUUGAUG-3’
Bottom rC Tm	5’-CCUACCCUGAUG-3’
Bottom rG Tm	5’-CCUACCGUGAUG-3’
Bottom rA Tm	5’-CCUACCAUGAUG-3’
RNAse Control rA	5’-CGUGUCC**A**$_{4'\text{-Me}}$GCCUUCA-3
RNAse 4’-*C*-MeA	5’-CGUGUCCAGCCUUCA-3
PPP2R5D RT-PCR and sequencing forward primer	5’-GGGGGAAAGGACACAGGG-3’
PPP2R5D amplification and RT-PCR reverse primer	5’-CACGAGCTATAGGGGAAAAGGCAC-3’
PPP2R5D T7 DNA amplification forward primer	5’-TAATACGACTCACTATAGGGGGAAAGGA-3’
PPP2R5D E200K Target sequence	5’GGGGGAAAGGACACAGGGGGGACUGGUGAGGGGCUCUG GAGAAGCCCAGGUGGAGCUCUAACUGGCCCUACCCCUCAG UUUUCAGUGAACCUCUUCCGGACGCUGCCACCUUCAUCGA AUCCCACAGGGGCUGAGUUUGACCCAGAGGAAGAU**A**AGCC CACCCUGGAAGCUGCUUGGCCACAUCUCCAGGUACCAGGG CAAGGGGGCAGAUUGGCCGUGGCUGCAGGGAGUGGGGCA CUUGGAGGCCUGCAAGUCCUUGGGAACAUCCCCUCAGUGG CGUGCCUUUUCCCCUAUAGCUCGUG3’
β-Actin 3’-UTR Target sequence	5’GGGUCCUCCUGAGCGCAAGUACUCCGUGUGGAUCGGCG GCUCCAUCCUGGCCUCGCUGUCCACCUUCCAGCAGAUGUG GAUCAGCAAGCAGGAGUAUGACGAGUCCGGCCCCUCCAUC GUCCACCGCAAAUGCUUCUAGGCGGACUAUGACUU**A**GUUG CGUUACACCCUUUCUUGACAAAACCUAACUUGCGCAGAAAA CAAGAUGAGAUUGGCAUGGCUUUAUUUGUUUUUUUUGUUU UGUUUUGGUUUUUUUUUUUUUUUUGGCUUGACUCAGGAUU UAAAAACUGGAACGGUGAAGGUGAC-3’
β-Actin Modified gRNA for cells (X and Y sugar modification sites), * P-S Linkage, N 2’-OMe, N Ribose	3’-G*C*CUGAUACUG**YX**CCAACGCAAUGUGGGAAAGAACUG*U*U-5’
Endogenous β-Actin RT Forward	5’-CAGCAGATGTGGATCAGCAAGCAGGAG-3’
Endogenous β-Actin RT Reverse	5’-GGAAGGGGGGGCACGAAGGCTCATC-3’
Endogenous β-Actin Nested Forward	5'-TATGACGAGTCCGGCCCCTCCATCGT-3'
Endogenous β-Actin Nested Reverse	5'-GCAATGCTATCACCTCCCCTGTGTGGACT-3'
Endogenous β-Actin Sequencing	5’-AACAAATAAAGCCATGCCAATCTCATCTTGTT-3

NMR Spectra

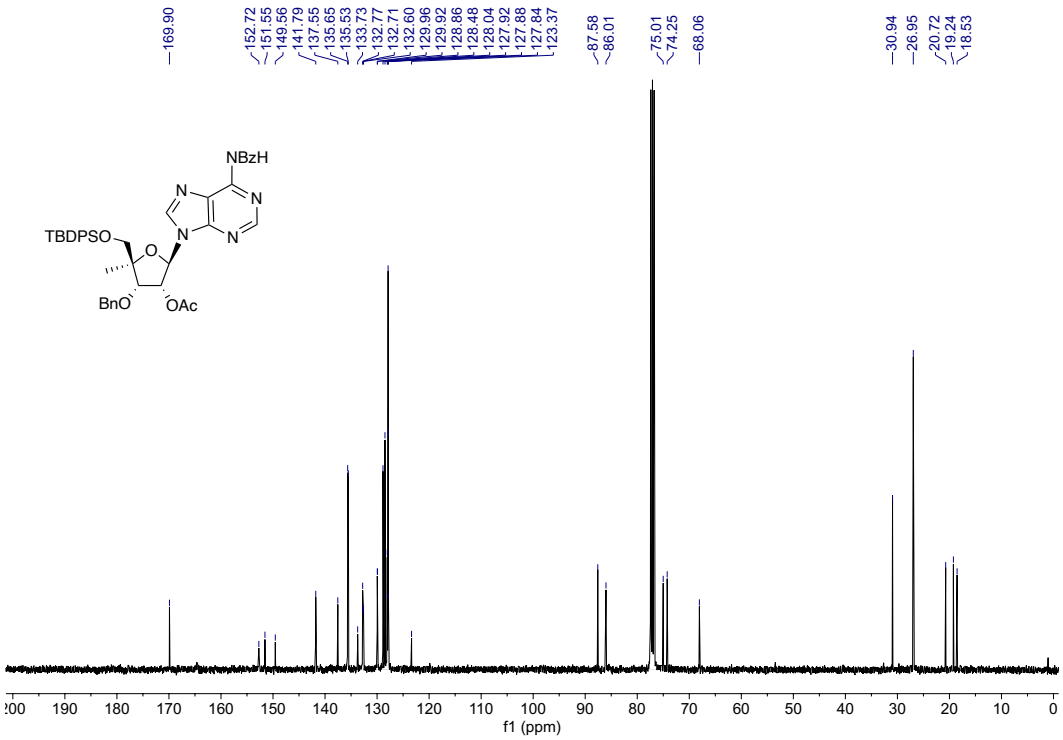
(2) N^6-Benzoyl-2'-*O*-acetyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine ^{1}H – (400 MHz, $CDCl_3$)



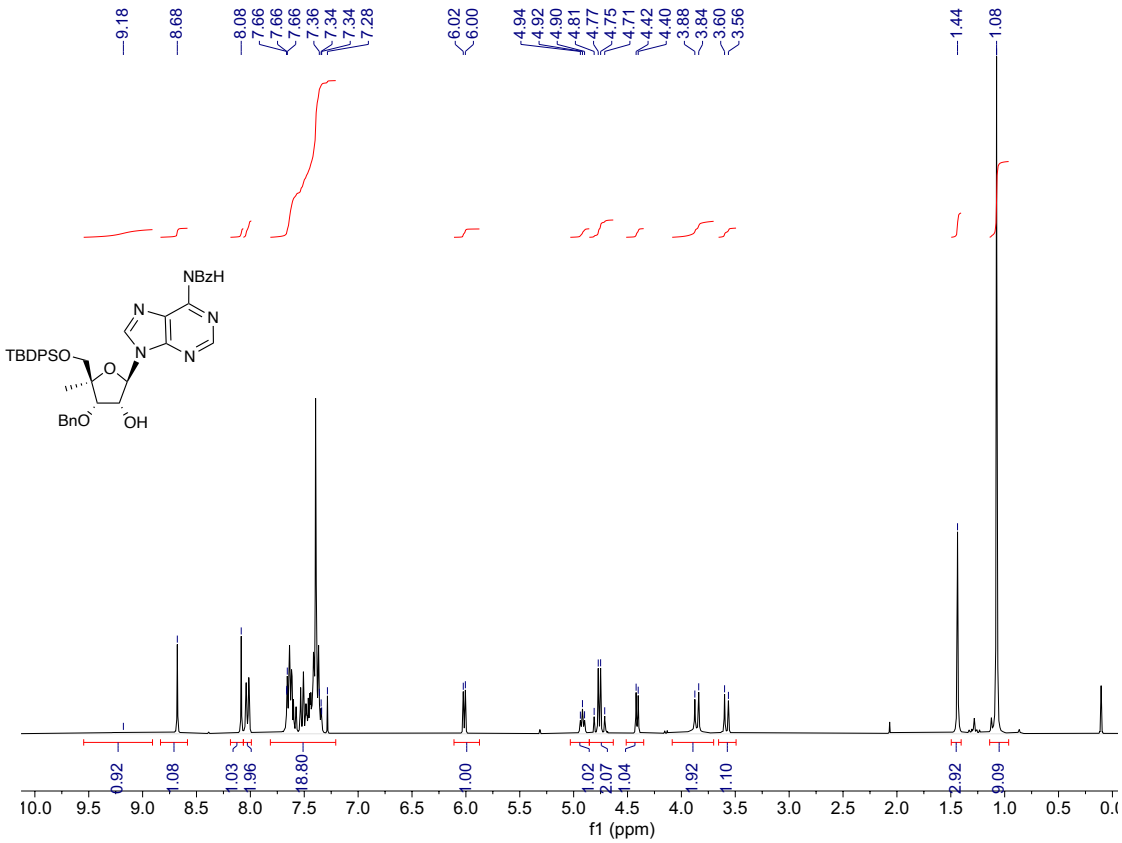
(2) N^6-Benzoyl-2'-*O*-acetyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine ^{1}H – (400 MHz, $CDCl_3$)



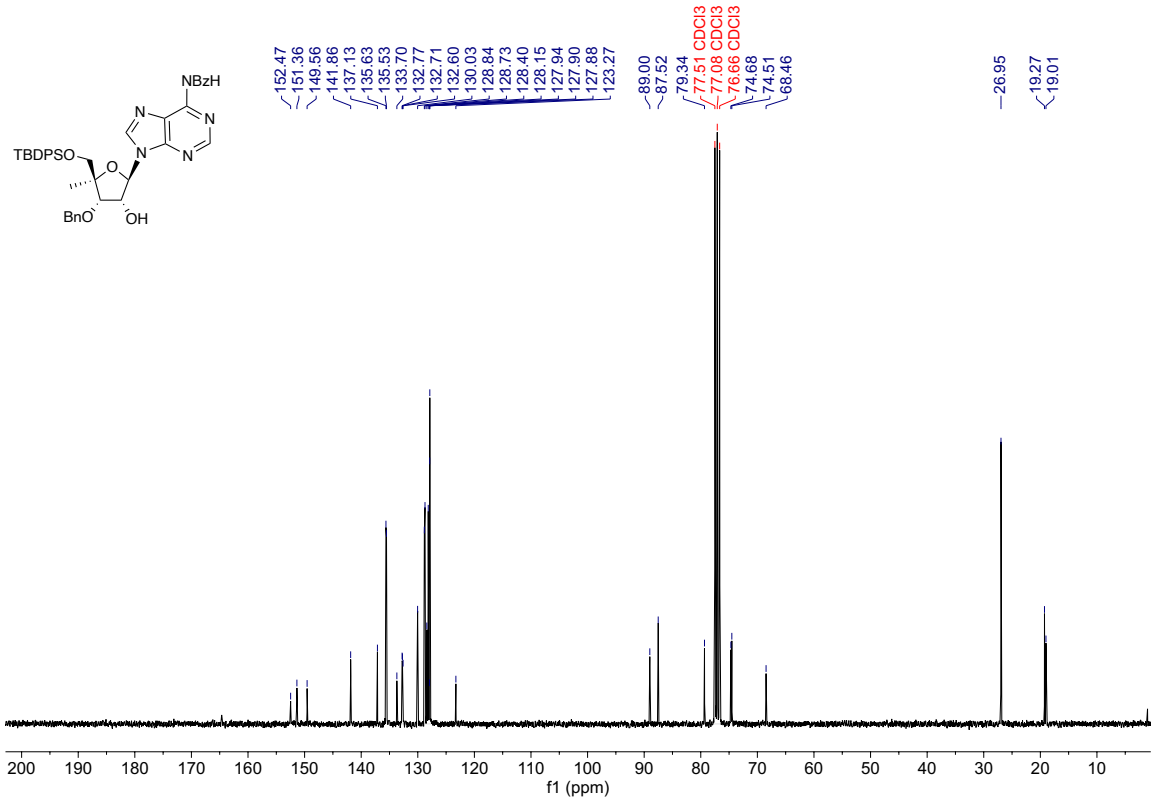
(2) N^6-Benzoyl-2'-*O*-acetyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine $^{13}C\{^1H\}$ – (101 MHz, $CDCl_3$)



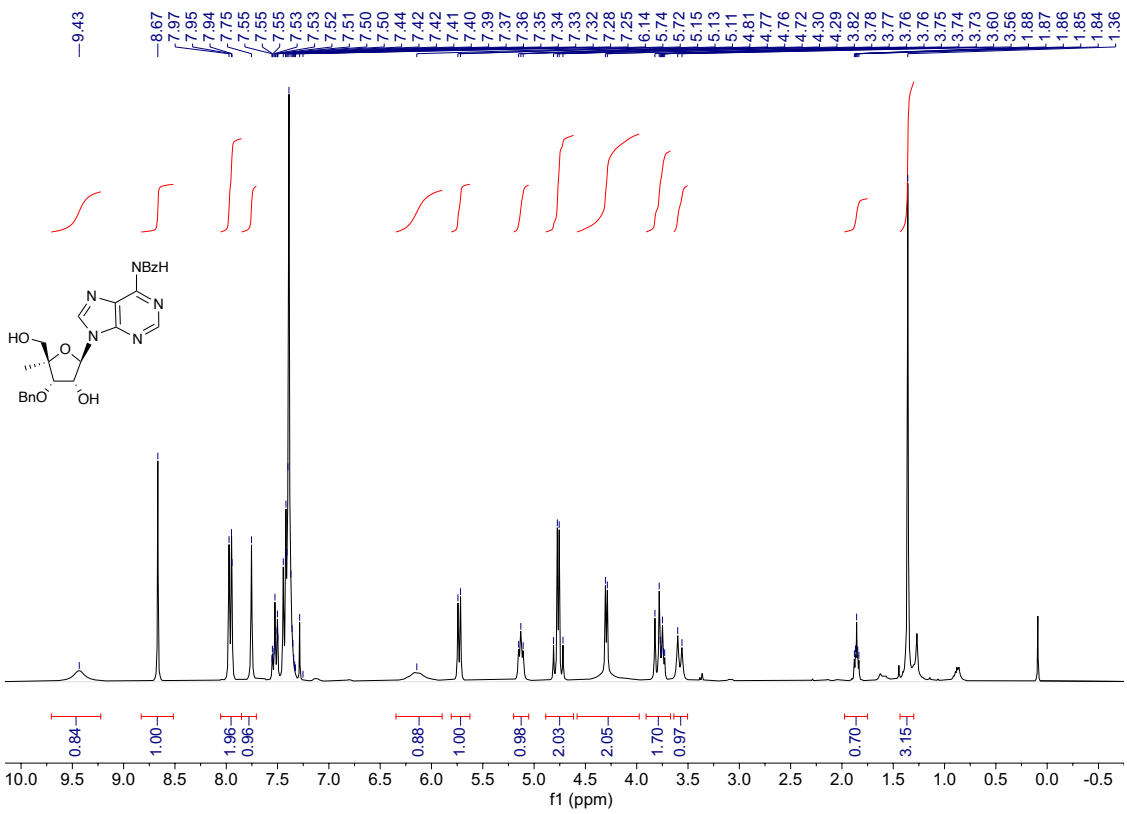
(3) N^6-Benzoyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine ^1H – (300 MHz, $CDCl_3$)



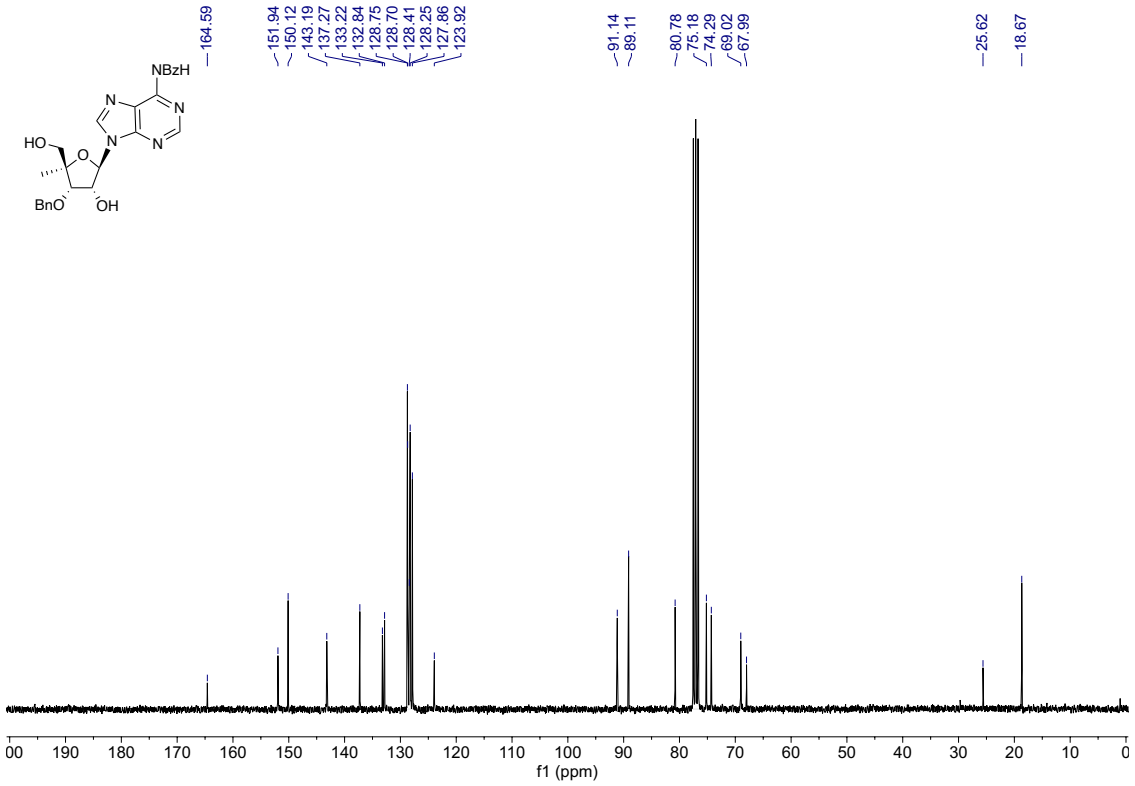
(3) N^6-Benzoyl-3'-*O*-benzyl-4'-*C*-methyl-5'-*O*-(*tert*-butyldiphenylsilyl)adenosine $^{13}C\{^1H\}$ – (76 MHz, $CDCl_3$)



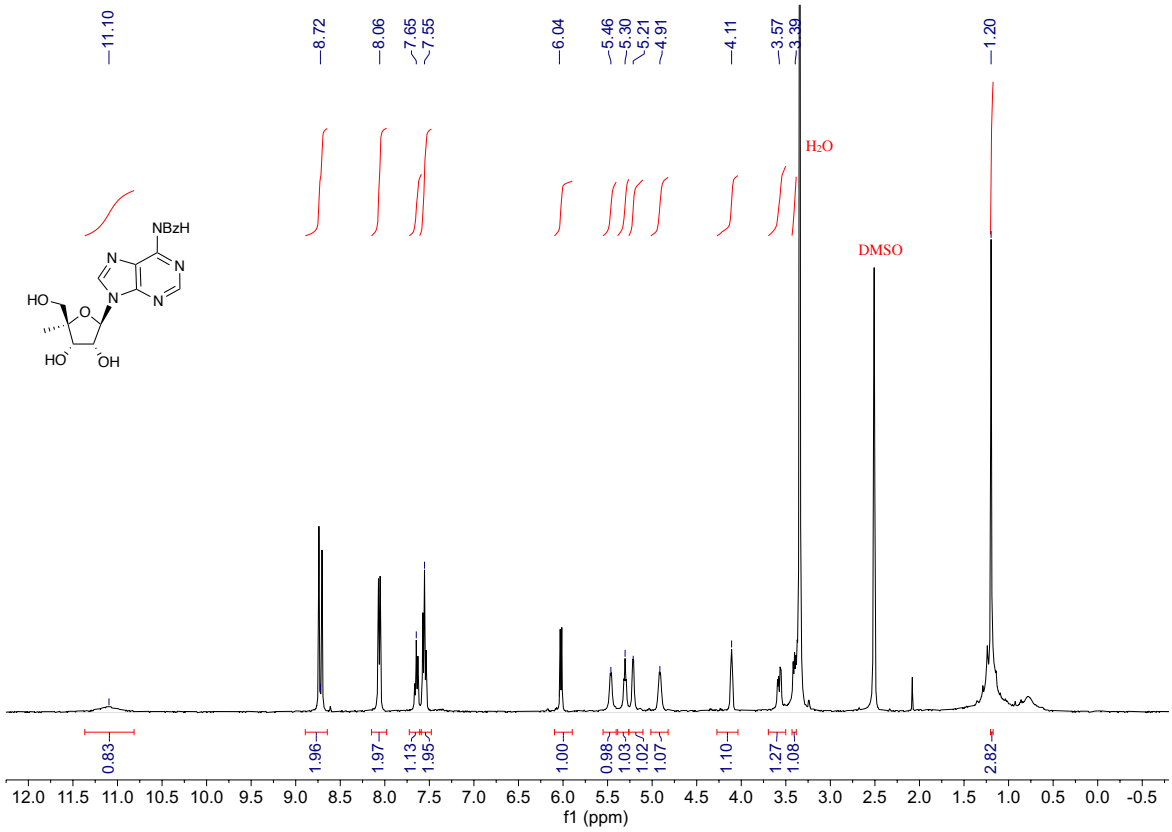
(4) N^6-Benzoyl-3'-*O*-benzyl-4'-*C*-methyladenosine ^1H – (300 MHz, $CDCl_3$)



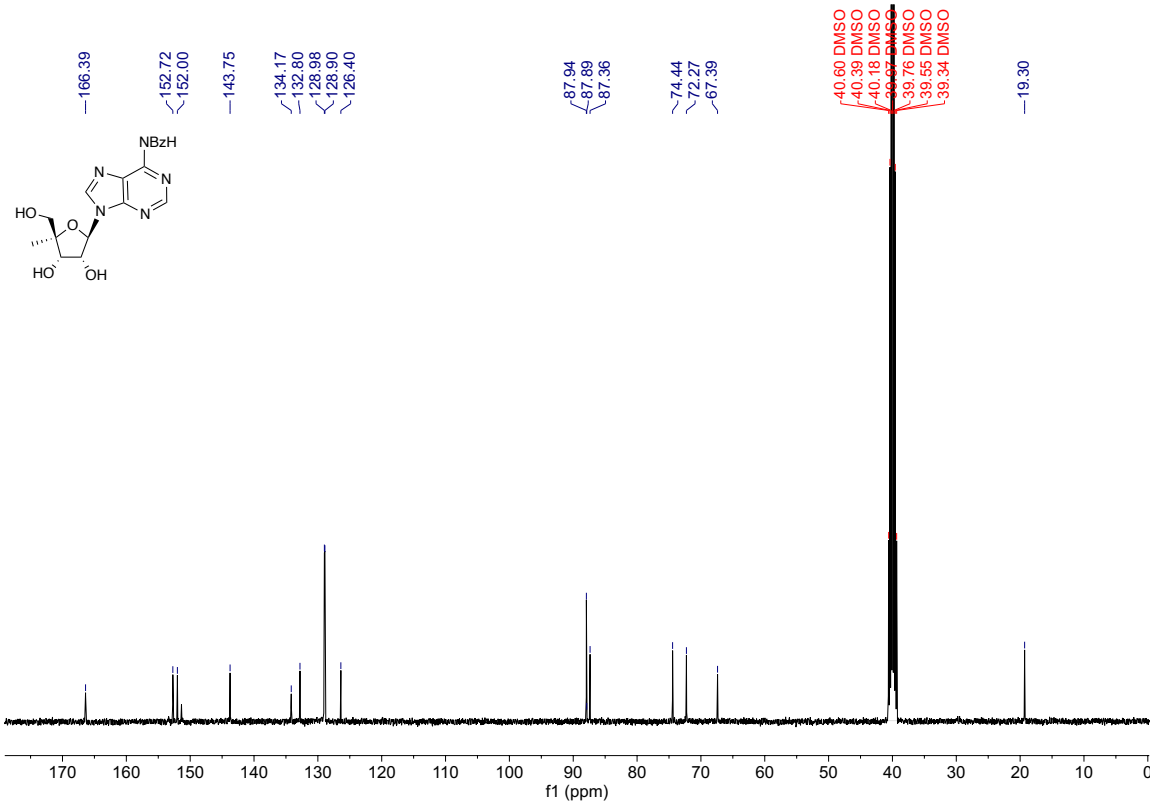
(4) N^6-Benzoyl-3'-*O*-benzyl-4'-*C*-methyladenosine $^{13}C\{^1H\}$ – (76 MHz, $CDCl_3$)



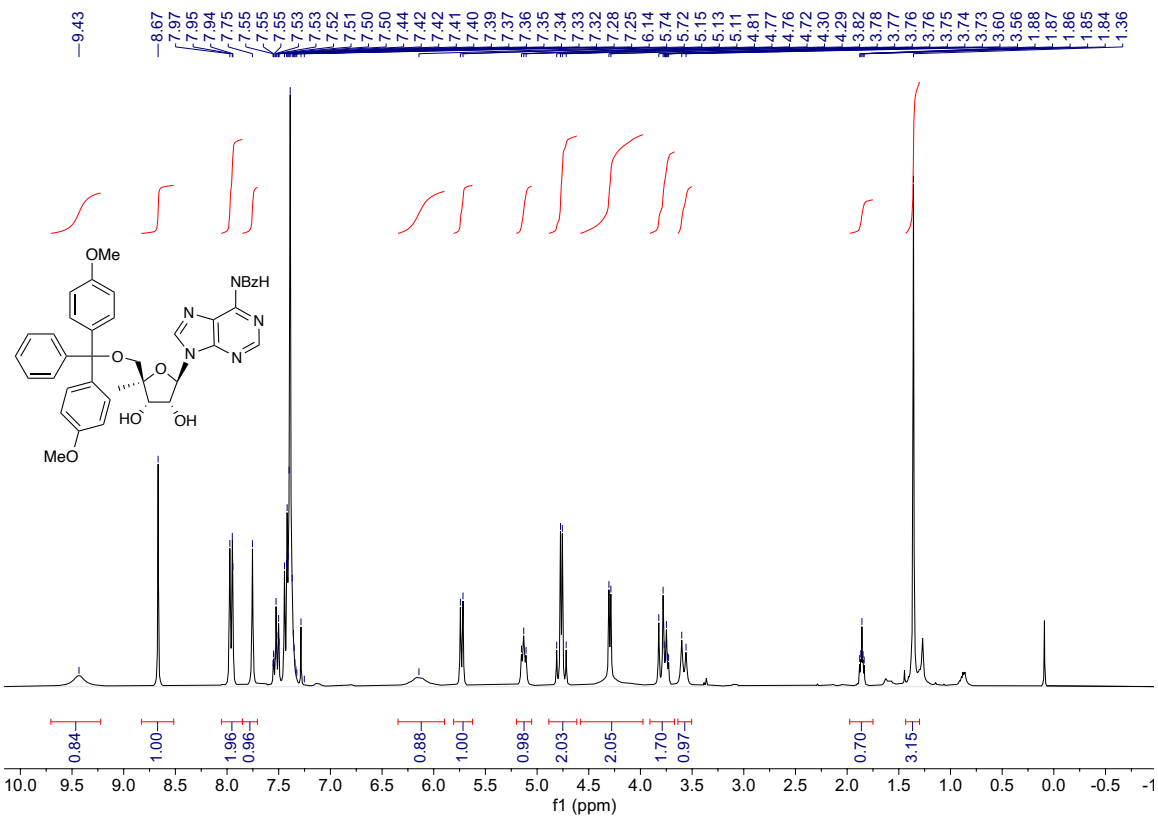
(5) N^6-Benzoyl-4'-*C*-methyladenosine ^1H – (400 MHz, DMSO-d_6)



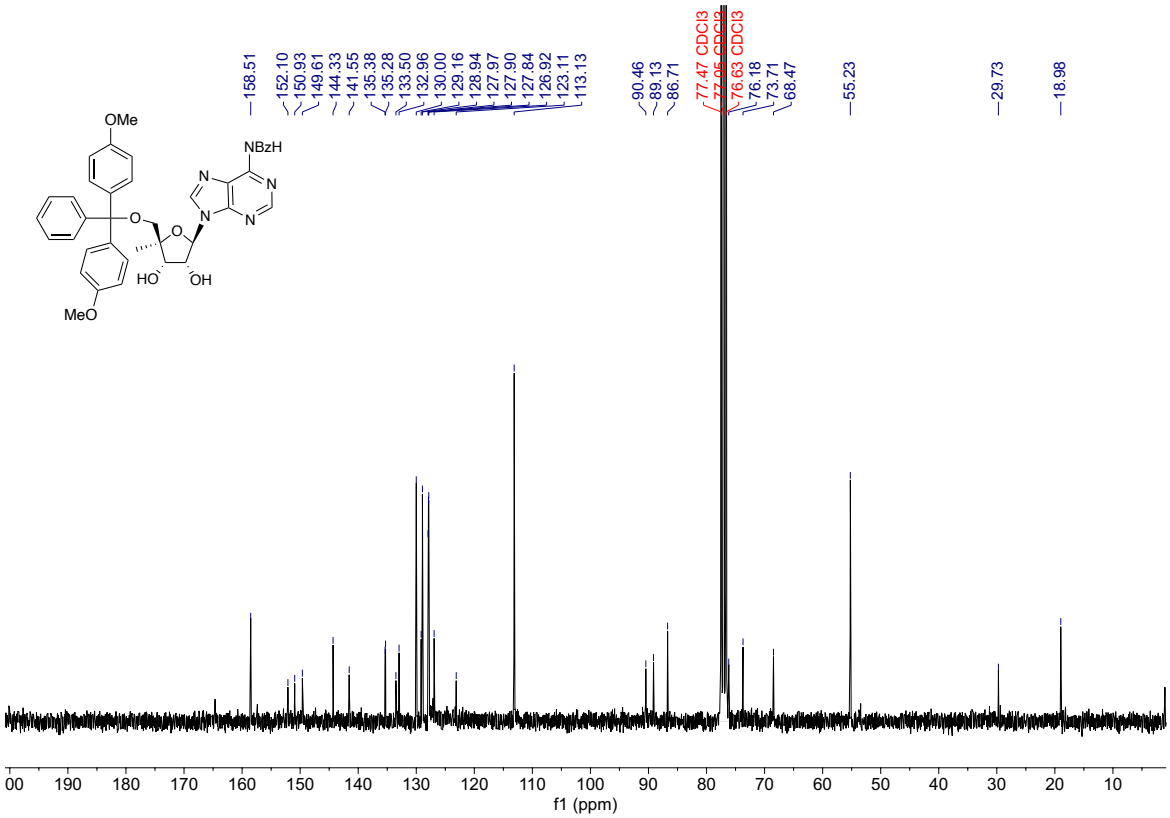
(5) N^6-Benzoyl-4'-*C*-methyladenosine $^{13}C\{^1H\}$ – (76 MHz, DMSO-d_6)



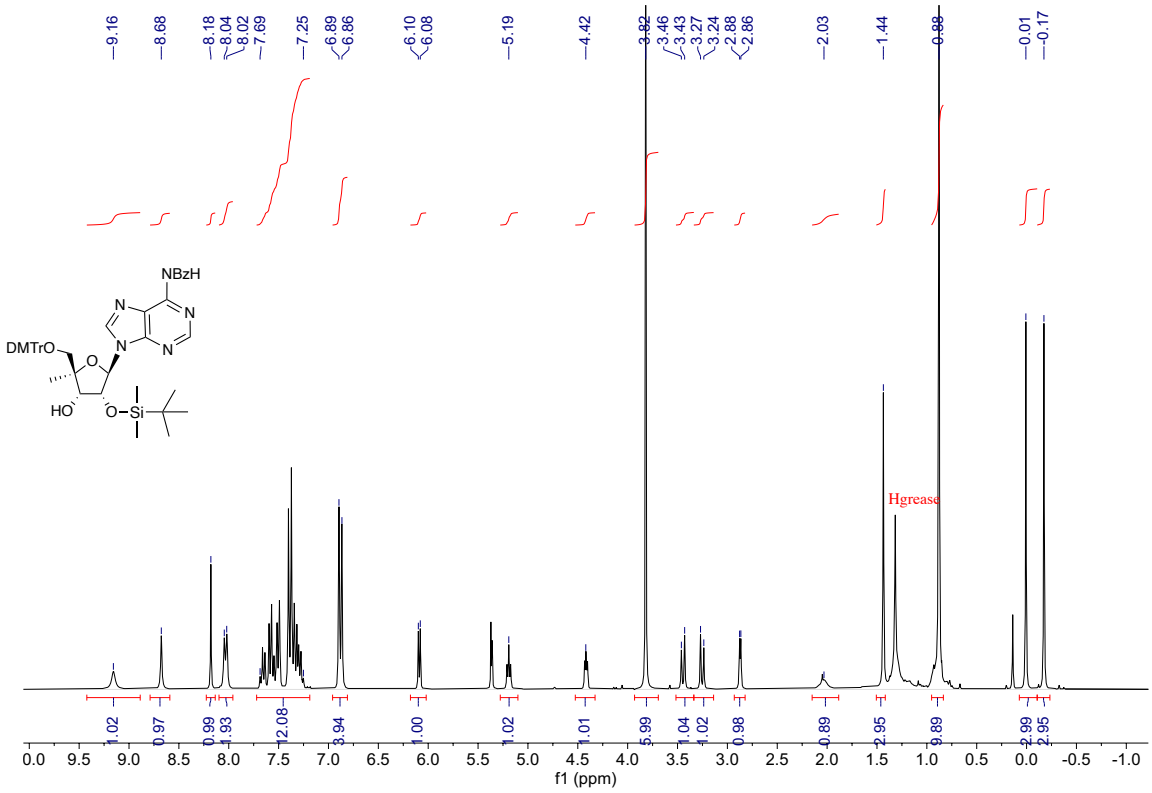
(6) 5'-*O*-(4,4'-Dimethoxytrityl)-N^6-benzoyl-4'-*C*-methyladenosine ^1H – (400 MHz, $CDCl_3$)



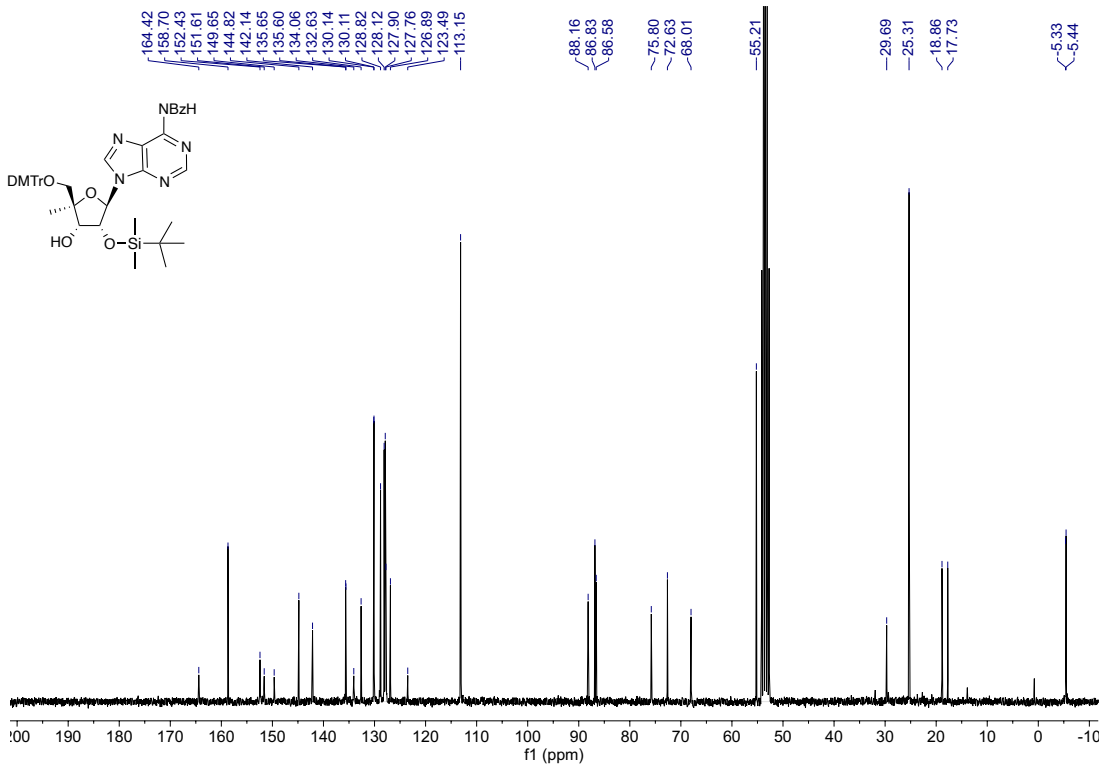
(6) 5'-*O*-(4,4'-Dimethoxytrityl)-N^6-benzoyl-4'-*C*-methyladenosine $^{13}C\{^1H\}$ – (76 MHz, $CDCl_3$)



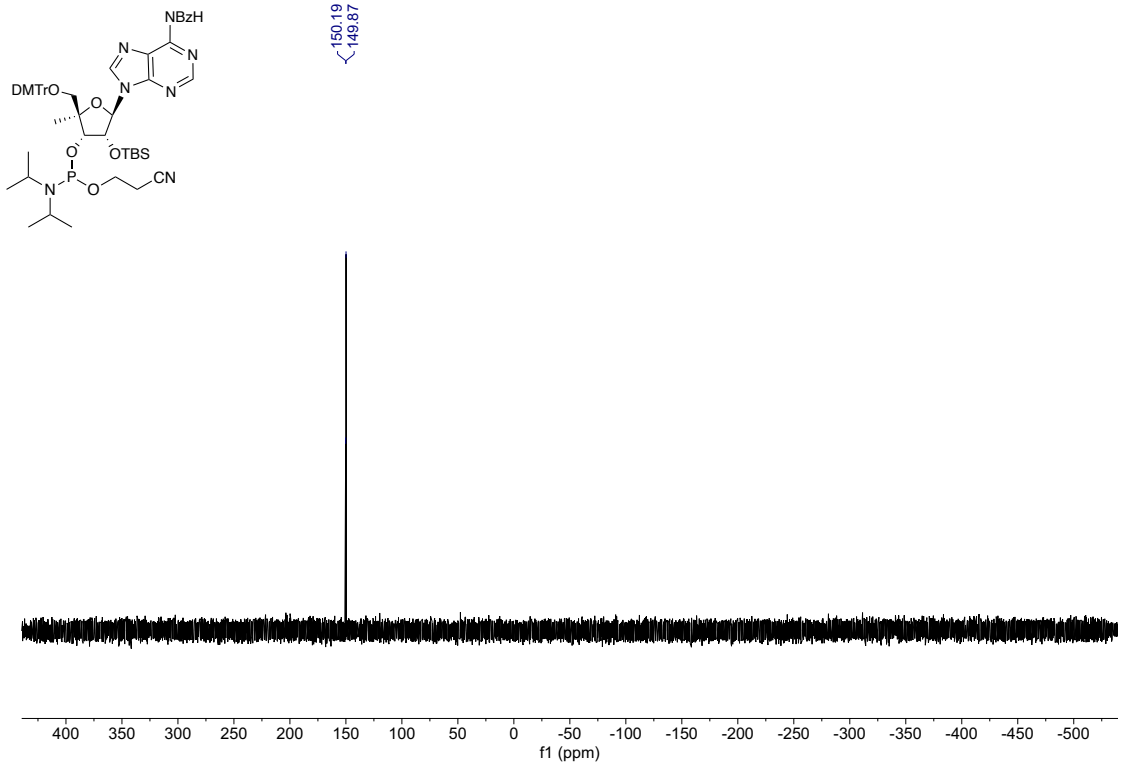
(7) 5'-*O*-(4,4'-Dimethoxytrityl)-2'-*O*-(*tert*-butyldimethylsilyl)-N^6-benzoyl-4'-*C*-methyladenosine ^1H – (300 MHz, CD_2Cl_2)



(7) 5′-*O*-(4,4′-Dimethoxytrityl)-2′-*O*-(*tert*-butyldimethylsilyl)-*N*^{6}-benzoyl-4’-*C*-methyladenosine $^{13}C\{^1H\}$ – (76 MHz, CD_2Cl_2)



(8) 5′-*O*-(4,4′-Dimethoxytrityl)-3′-*O*-[(2-cyanoethoxy)(*N*,*N*-diisopropylamino)phosphino]-2′-*O*-(*tert*-butyldimethylsilyl)-*N*^{6}-benzoyl-4’-*C*-methyladenosine



CHAPTER 3

Synthesis of 2′-*O*-4′-*C*-dimethyl ribonucleoside analogs and their properties in RNA and ADAR editing

Some of the work presented herein on 5'-AA targets was originated by Dr. Hannah F. Brinkman and was replicated to complete the story. The findings of this chapter will be submitted to ACS Chemical Biology for article publication in December 2024.

INTRODUCTION

In Chapter 2 we have shown the synthesis of 4'-*C*-methyladenosine (4'-*C*-MeA) (Figure 1) and its effect on ADAR editing when inserted into a guide oligonucleotide targeting 5'-UA sites (underline indicates target A). We show how different sugar conformations influenced by methyl or methylene groups regulate the editing of ADAR sites depending on the position of the sugar analogs. Additionally, even though the impact of 4'-*C*-methyladenosine rendered efficient site-specific regulation of ADAR, its nuclease resistance performed poorly. Not much work has been done on the effect of ribose analogs in 5'-AA sites, where uridine analogs would modulate ADAR editing considering the base pair selectivity of the 5' neighbor. In this chapter, we sought to synthesize a new adenosine analog phosphoramidite for the nuclease protection of the 2'-OH of 4'-*C*-methyladenosine, namely 2'-*O*-4'-*C*-methyladenosine (24dMeA) (Figure 1). We also subjected ADAR for editing a 5'-AA with various sugar analogs to evaluate the impact of sugar pucker in such unexplored sites. This included the synthesis of 4'-*C*-methyluridine (4MeU) and its dimethylated analog 2'-*O*-4'-*C*-dimethyluridine (24dMeU) phosphoramidites and their incorporation into ADAR guide oligonucleotides for editing a 5'-AA target. The model substrate we chose in this chapter is the c-Src messenger RNA which encodes for a protein kinase frequently overexpressed in cancers and is therapeutically relevant.[1] Lastly, in addition to these

in vitro deamination experiments, we determined how 4' or 2' methylation impact the melting temperature of modified RNA.

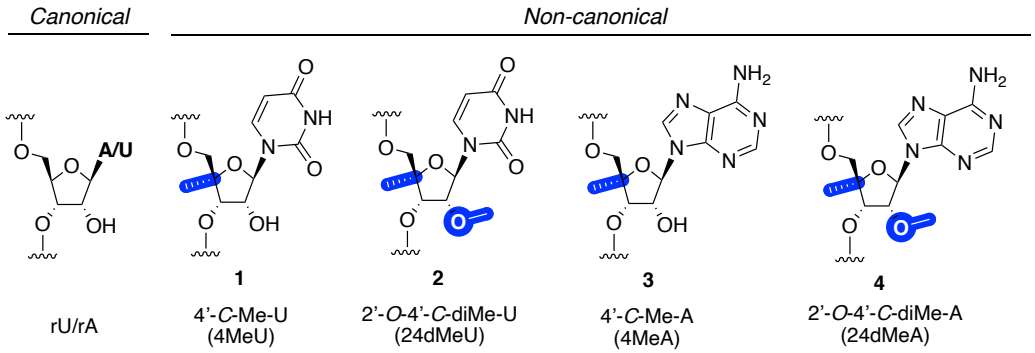


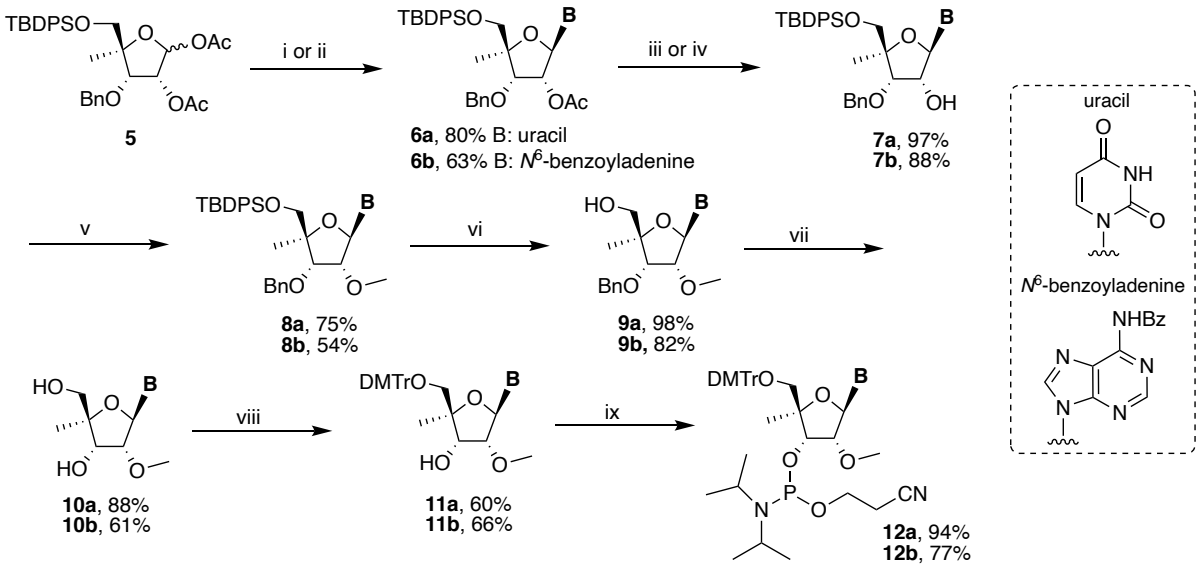
Figure 1. Chemical structures of the methylated nucleic acid analogs used in this work and their abbreviations. rU: uridine, rA : adenosine, 1: 4'-*C*-methyluridine, 2: 2'-*O*-4'-*C*-dimethyluridine, 3: 4'-*C*-methyladenosine, 4: 2'-*O*-4'-*C*-dimethyladenosine.

RESULTS AND DISCUSSION

Synthesis of 2'-O-4'-C-dimethyladenosine and uridine phosphoramidites

The synthesis of the 4'-C-MeU and 4'-C-MeA phoshoramidites has been reported before.[2,3] Yet, there is no synthetic data on the synthesis of both dimethylated U and A phosphoramidites. Herein, we present the synthesis of 2′-*O*-4′-*C*-dimethyladenosine and uridine phosphoramidites starting from the previously synthesized intermediates **6a**[4] and **7b**[3] (Scheme 1). Initially, after glycosylation of the nucleobases, the 2'-hydroxy for both analogs was selectively deprotected for further regioselective 2'-*O*-methylation. For uridine, 1M NaOH in MeOH was employed to remove the acetyl group in 97% yield, whereas for adenosine, a milder base (K_2CO_3 in MeOH) was used to remove the acetyl group in **7b** in 88% yield, while preventing the hydrolysis of the exocyclic amine N^6-benzoyl group. For the 2'-*O*-methylation, we initially tried utilizing MeI and NaH in DMF. Yet, for both analogs methylation was primarily favored at the ring nitrogens of the nucleobases. To circumvent the regioselectivity of methylation, procedures from Kajino et al.[5] and Koizumi et al.[6] were employed, where the same equivalents of MeI and NaH were used, but the solvent differed. In both uridine (**8a**) and adenosine (**8b**) reactions, the substitution of DMF for

THF rendered regioselective 2'-*O*-methylation with reasonable yields (75% for **8a** and 54% for **8b**). Subsequent deprotection of the 5'-*O*-silyl ether group afforded **9a** and **9b** in high yields with *n*-tetrabutylammonium fluoride. Debenzylation of both **9a** and **9b** were carried out using BCl_3 in DCM at – 78 °C to produce **10a** at 88% and **10b** at 61%. Dimethoxytritylation of the 5'-OH was achieved with *4,4'*-dimethoxytrityl chloride (DMTrCl) and *N,N*-dimethylaminopyrine (DMAP) in pyridine to furnish **11a** and **11b** in 66% and 60% yields, respectively. Lastly, to produce the phosphoramidite building blocks, the 3'-OH group of **11a** and **11b** were phosphitylated with 2-(cyanoethyl)phosphoramidite chloride to afford **12a** and **12b** in good yields.



Scheme 1. Synthesis of 24dMeA and 24dMeU phosphoramidites and solid support. *Synthesis and reagents.* (i) BSA, uracil, TMSOTf, MeCN, 60 °C; (ii) N^6-benzoyladenine, $SnCl_4$, 0-25 °C, MeCN; (iii) 1M NaOH, MeOH, 25 °C; (iv) K_2CO_3, MeOH, 0 °C; (v) Iodomethane, THF, NaH, 0 °C; (vi) 1M tetrabutylammonium fluoride, THF, 25 °C; (vii) 1M boron trichloride, DCM, -78 °C; (viii) DMTr-Cl, pyridine, DMAP, 25 °C; (ix) chloro(diisopropylamino)-β-cyanoethoxyphosphite, DIPEA, 1-methylimidazole, 25 °C.

Effect of 2'-O-4'-C-dimethylation in RNA: Duplex stability.

To characterize the thermodynamic and metabolic properties of 2'-*O*-4'-*C*-dimethyluridine and adenosine, we determined their melting temperature (Tm) and ribonuclease sensitivity in human serum. We wanted to explore to what extent dual methylation affects the hybridization in RNA duplexes. In addition, we sought to examine if methylating the 2'-OH of these 4'-C-methyl

analogs can increase their metabolic stability, considering that previously we showed that 4'-*C*-methylation in adenosine with a free 2'-OH is as equally susceptible to canonical adenosine in human serum.[3] Additionally, these novel nucleic acid analogs might interest the RNA therapeutic field, where nuclease resistance is pivotal for efficient RNA drugs.[7] We conducted thermal denaturation assays to elucidate the impact of 2'-*O*-4'-*C*-dimethyluridine and adenosine on duplex stability and base pairing selectivity. For this purpose, we synthesized 12 nt RNAs containing 2'-*O*-4'-*C*-diMeA or U and hybridized them with complementary 12 nt RNAs for Tm measurements (Table 1). The nuclease sensitivity of RNAs with 2'-*O*-4'-*C*-dimethyluridine or adenosine was assessed using a 15 nt long oligonucleotide with 2'-*O*-4'-*C*-diMeA or U being at the 7th position flanked by all 2'-*O*-methyl nucleotides and subjected to 80% human serum over a time course of 24 h. For both experiments, two controls were employed to compare the effect of 4'-C-methyl and 2'-*O*-4'-*C*-dimethyl moieties: 2'-O-methylA or U and canonical rA and rU. The melting curves for duplexes containing methylated adenosine analogs paired with uridine showed that 4'-*C*-MeA, as well as 2'-O-MeA, decreased the Tm of the parent RNA (rA) by 1.4 °C and 1.9 °C, respectively. Interestingly, it seems that both 2' and 4' effects combined to decrease the Tm of 2'-*O*-4'-*C*-diMeA, considering that it decreased by 3.8 °C (Table 1). A different effect was observed when methylated uridine analogs were paired to rA, where 4'-*C*-MeU decreased the Tm by -1.1 °C and 2'-*O*-4'-*C*-diMeU by only -0.5 °C. These Tm values suggest that 2'-methylation in uridine increases the Tm (+0.6 °C) whereas 4'-methylation decreases it (-1.1 °C). Lastly, when these adenosine and uridine analogs were paired to the rest of the canonical RNA bases individually, the pairing selectivity of the nucleobases did not seem to be influenced (Figure 2).

Entry	Sequence[a]	*Tm*[b] (°C)	Δ*Tm*[c](°C)
	Adenosine Analogs		
1	5′-r(CAUUAAGGUGGG)-3′ 3′-r(GUAGUUCCAUCC)-5′	46.1 ± 0.1	NA
2	5′-r(CAUUA<u>A</u>GGUGGG)-3′ 3′-r(GUAGUUCCAUCC)-5′	45.0 ± 0.2	-1.9
3	5′-r(CAUUA**A**GGUGGG)-3′ 3′-r(GUAGUUCCAUCC)-5′	42.2 ± 0.0	-1.4

4	5′-r(CAUUA**A**GGUGGG)-3′ 3′-r(GUAGUUCCAUCC)-5′	43.1 ± 1.0	-3.8
	Uridine Analogs		
1b	5′-r(CAUUAUGGUGGG)-3′ 3′-r(GUAGUACCAUCC)-5′	45.4 ± 0.1	NA
2b	5′-r(CAUUA**U**GGUGGG)-3′ 3′-r(GUAGUACCAUCC)-5′	46.1 ± 0.1	+0.6
3b	5′-r(CAUUA**U**GGUGGG)-3′ 3′-r(GUAGUACCAUCC)-5′	44.3 ± 0.1	-1.1
4b	5′-r(CAUUA**U**GGUGGG)-3′ 3′-r(GUAGUACCAUCC)-5′	45.0 ± 0.0	-0.5

Table 1. Melting Temperatures of RNA Duplexes Modified with Uridine and Adenosine Analogs. [a]Uppercase letters indicate unmodified RNA. Underline is 2'-methoxy RNA. **U** and **A** are 4'-*C*-methyluridine and adenosine, respectively. **U** and **A** are 2'-*O*-4'-*C*-dimethyluridine and adenosine, respectively. [b]*T*m values were calculated in triplicates. The RNA was brought to 95 °C and held at 95 °C for 3 min, 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). *Tm* was determined as the temperature where -dF/dT is the largest. Buffer conditions: 1 mM EDTA, 100 mM NaCl and 10 mM Tris HCl buffer (pH 7.5) using 1 μM concentrations of each strand in 1x Eva Green fluorescent dye. [c]Δ*Tm* is the difference in melting temperature between the modified duplex and the reference RNA duplex (A or U).

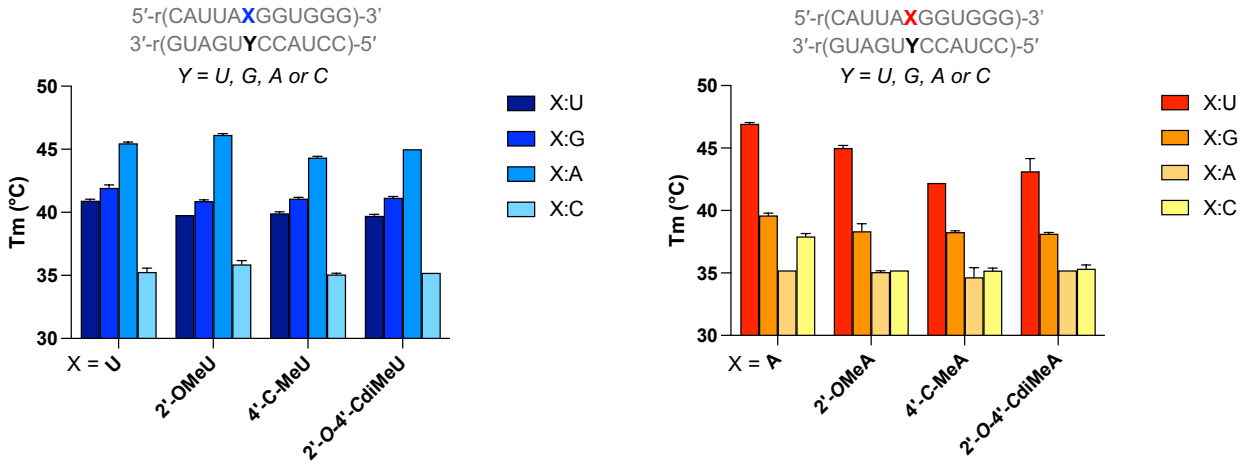


Figure 2. Melting Temperatures of RNA Duplexes Modified with Uridine and Adenosine Analogs Paired To RNA Bases. Uppercase letters indicate unmodified RNA. *T*m values were calculated in triplicates.

Effect of different uridine analogs in ADAR editing of a 5'-AAA site

Previously, Brinkman et al. reported a holistic understanding of the effect that different sugar analogs (UNA, LNA, 2'-F, 2'-OMe, 2'-deoxy, etc.) on the gRNA have on ADAR editing of 5'-UA targets where the -1 nucleobase is adenosine.[8] On another note, we reported the effect of 4'-C-methylated sugars to site-specifically block bystander editing sites in 5'-UA targets.[3] However, no data has been reported for the effect of different sugars on RNA editing and off-target mitigation

where the 5' nearest neighbor of the target adenosine is another adenosine (5'-AA). To do this, we subjected an RNA duplex bearing the sequence present in the c-Src mRNA to an ADAR2 reaction *in vitro*. c-Src is a protein kinase frequently overexpressed in cancers.[1] Specifically, we targeted for RNA editing the K295 codon (5'-A**A**A-3'). Editing of the central adenosine will convert this to an arginine codon resulting in a K295R mutation that renders the oncogenic protein kinase inactive, rendering this coding event as a therapeutic approach.[9] The sequence surrounding the target site is 5'-A**A**A-3', where the target adenosine is directly flanked on the 5' and 3' end by 1 potential bystander editing site (Figure 3B). Initially, we determined the percent editing by ADAR2 at 60 min at the target and off-target sites with an unmodified gRNA as a control. We then modified the -1 position of the gRNA with various uridine and thymine analogs, considering that the 5' neighbor nucleobase is adenosine, and it is paired to uridine in the gRNA design (Figure 3B). Considering the noteworthy impact of 2'-deoxy-2'-fluoro, 2'-*O*-methoxy, 4'-*C*-methyl, and 2'-deoxy chemistries on the ribose in ADAR guide strands for RNA editing in 5'-UA sites, we synthesized various gRNAs targeting the SRC mRNA with uridine analogs with varying beforementioned modified sugar moieties. The structure of these modifications can be found in Figure 3A and includes methylated sugars (2'-*O*-MeU, 4'-*C*-MeU, and 2'-*O*-4'-*C*-diMeU), fluorinated sugars (2'-F-U and 2'-F-Arabino-U) and deoxy sugars (2'-deoxyU and 2'-deoxyT).

2'-*O*-Methylation of uridine at the -1 position of the gRNA did not affect the editing of the target site in a statistically significant manner when compared to the no-modification gRNA. However, a complete loss of editing (26 to 0%) was observed for the 5'-adjacent bystander adenosine, which is the base opposite to the -1 position where 2'-*O*-MeU lies (i.e. its "orphan" base). In contrast, the 3'-adjacent bystander site remained edited (14%). Conversely, when 4'-*C*-MeU was positioned -1 to the target site, a 26.8% decrease in editing was observed at the target site. When comparing the effect of 2'-*O*-MeU, the 3-adjacent bystander site was reduced to 0%, while the 5'-adjacent bystander site remained edited (18%) (Figure 3C). Finally, when 2'-*O*-4'-*C*-diMeU was placed at the -1, both 5' and 3' adjacent off-target sites were reduced to

no detectable % editing while having little effect on the target site editing. To further attest to the impact of these methylated uridine analogs, we performed deamination kinetics *in vitro* on the SRC mRNA K295R with the methylated gRNAs within 0 to 120 min time course. The results indicate that the rate of reaction for all three methylated analogs do not have a significant difference (Figure 4, Table 2). The pattern of site-specific inhibition of bystander sites relative to the site-specific methylation of the riboses can be explained using the high-resolution crystal structures of ADAR2 bound to dsRNA.

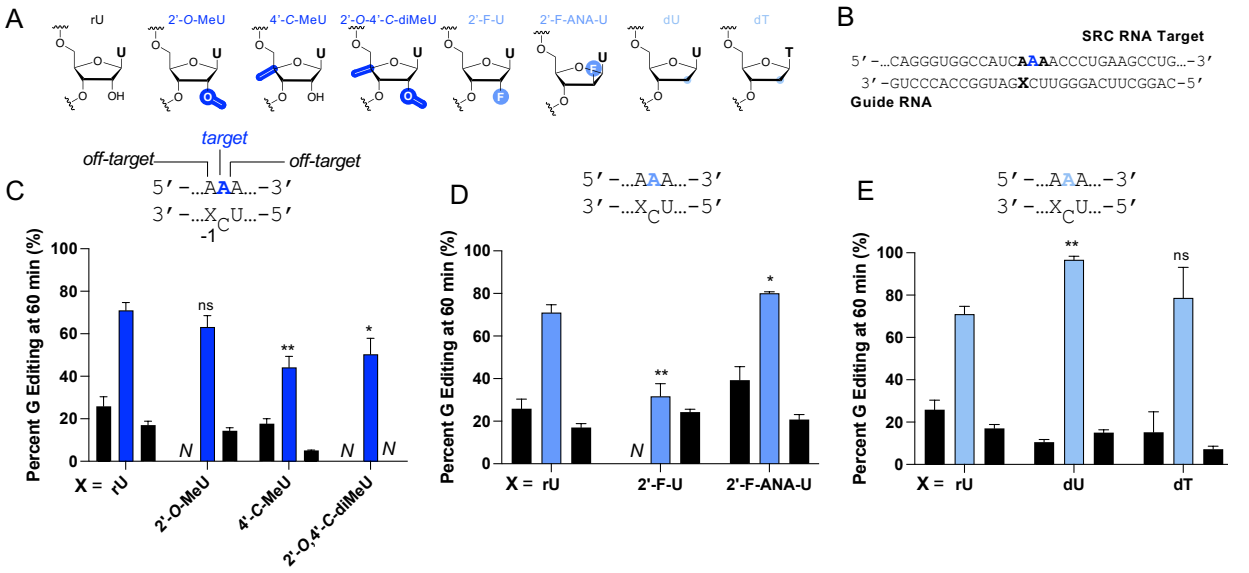


Figure 3. In vitro adjacent off-target RNA editing regulation with various ribose and 2'-deoxyribose analogs. A) Structure and abbreviation of uridine and thymine analogs used in this study. B) Target sequence derived from the mRNA of the proto-oncogene tyrosine-kinase Src (SRC) proximal to the K295R mutation associated with cancer and guide RNA used here to direct corrective editing at the blue adenosine. The target adenosine is indicated in blue. X in bold indicates the -1 position of the guide strand relative to the target adenosine. C-E) 60 min timepoint comparison of % G editing of the 5' and 3' adjacent off-target sites in the SRC mRNA K295R mutation region. Colored bars represent the % G editing of the target A whereas black bars represent the % G editing of the adjacent off-target adenosines. C) Dark blue-Effect on editing using 2' and/or 4' methylated riboses. D) Light blue-Effect on editing using 2'-fluorinated riboses. E) Sky blue-Effect on ADAR editing using 2'-deoxyribose analogs. A two-tailed Welch's *t* test was conducted, where **p* < 0.05, ***p* < 0.01, ****p*< 0.001 from X= rU guide RNA; ns: no significant difference.

Considering that the kink seen in the RNA duplex is partly stabilized by the orphan base's 2'-OH group contact to the side chain of R510, we hypothesize that when 2'-OMeU is placed orphan to an editing site, editing is drastically inhibited because methylation of the 2'-OH abolishes any potentially stabilizing hydrogen bond to R510 (Figure 5A). Interestingly, the

impact of 4'-C-MeU in a 5'-AAA-3' target can be correlated to the inhibitory effect seen with 4'-C-MeA in 5'-UAA-3' targets, where positioning this ribose analog -2 relative to the editing site blocks ADAR editing.[3] These results suggest that this form of bystander site inhibition is independent of the nucleobase, considering that the origin of this form of inhibition lies in a steric clash between the minor groove of the 4'-C-Me group at the sugar level and the side chain of P492 (Figure 5B). Remarkably, when both methylating effects are synergized in one sugar (e.g. 2'-*O*-4'-*C*-dimethyluridine), efficient mitigation of off-target editing can be achieved selectively at both the 5' and 3' adjacent sites flanking the target adenosine (Figure 3C).

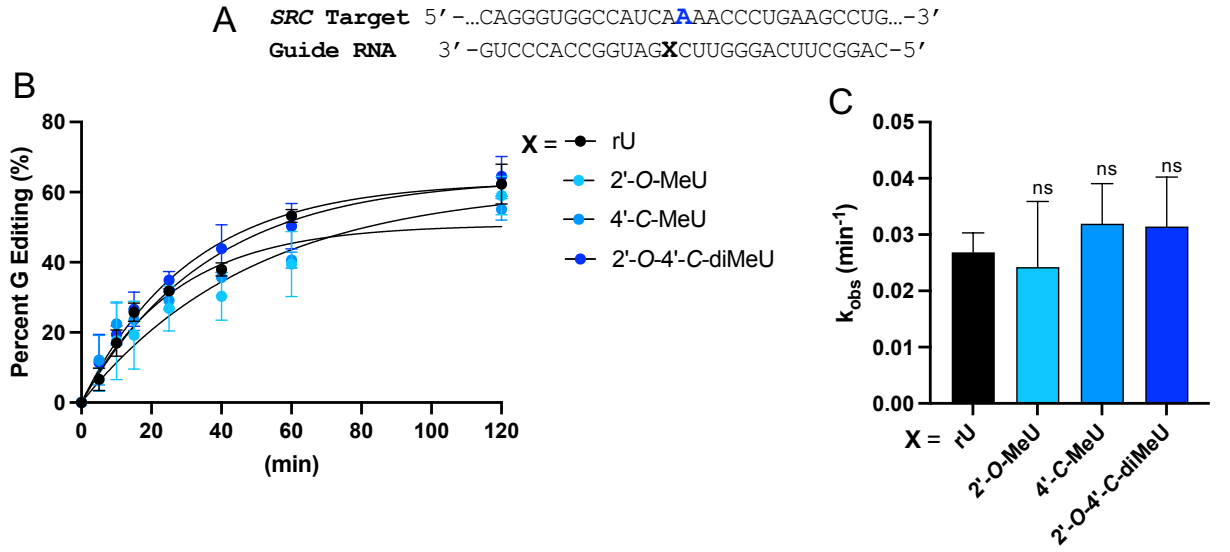


Figure 4. In vitro deamination kinetics for ADAR2 varying the 5' nearest neighbor ribose. A) Top: Target sequence derived from the mRNA of the proto-oncogene tyrosine-kinase Src (SRC) proximal to the K295R mutation associated with cancer and guide RNA used here to direct corrective editing at the blue adenosine. The target adenosine is indicated in blue. X in bold indicates the -1 position of the guide strand relative to the target adenosine. B) % G editing at different timepoints when X is varied. B) Comparison of rate constants for reaction with 50 nM ADAR2 and 5 nM target RNA. Error bars represent the standard deviation of three technical replicates. A two-tailed Welch's *t* test was conducted, where *$p < 0.05$, **$p < 0.01$, ***$p< 0.001$ from X= rU guide RNA; ns: no significant difference.

2'-Fluorinated and 2'-deoxy sugars in the -1 position of the gRNA rendered interesting yet expected results depending on the configuration at the fluorination site of the ring. Previously, our lab reported that 2'-deoxy-2'-F-A at the -1 position of 5'-UA sites decreased the rate of deamination in ADAR2 and ADAR1 p110, whereas substitution to 2'-deoxy-2'-F-ANA -A (arabino = ANA) increased the rate of reaction significantly, compared to rA.[8]

Table 2. Rate constants for ADAR2 in vitro deamination of the SRC K295R mRNA with varying the -1 position in the guide RNA.

-1 modification[a]	k_{obs} (min^{-1})[b]	k_{rel}[c]
rU	0.029 ± 0.004	1
2'-*O*-MeU	0.02 ± 0.01	0.7
4'-*C*-MeU	0.024 ± 0.005	0.8
2'-*O*-4'-*C*-diMeU	0.034 ± 0.009	1.2

[a] Reactions were carried out with 5 nM target RNA and 50 nM ADAR2
[b] Reactions were fitted to the equation $[P]t = \alpha[1 - \exp(-kobs \cdot t)]$
[c] $k_{rel} = k_{obs}$ for modification/k_{obs} for uridine (rU)

Analogously, when dA was placed at the -1, the rate of deamination increased significantly compared to rA. These results can be correlated to those observed herein with the SRC mRNA (5'-AA site) and their uridine analogs. Thus, 2'-deoxy-2'-F-U performed poorly compared to rU (41% decrease), while dU, dT and 2'-deoxy-2'-F-ANA-U displayed higher and faster editing percentages at 60 min (25%, 7%, and 9% higher than rU, respectively) (Figure 3D-E). Additionally, only 2'-deoxy-2'-F-U displayed off-target mitigation of the 5'-adjacent site (orphan to 2'-deoxy-2'-F-U) with no detectable editing (Figure 3D-E). Yet, target editing was also compromised with a loss of 41% editing. A common ribose denominator in uridine analogs that favor ADAR editing (dT, dU, 2'-deoxy-2'-F-ANA-U) is the ring conformation that these analogs adopt. dT, dU, and 2'-deoxy-2'-F-ANA-U adopt a C2'-*endo* sugar pucker, while 2'-deoxy-2'-F-U adopts a C3'-*endo* RNA sugar pucker.[10–12] These observations concur with previous analyses where C2'-*endo* sugars at the -1 position of the gRNA highly favor the ADAR reaction considering that the conformation of this ribose in the crystal structures is the only ribose in the guide strand that adopts an unusual C2'-*endo* conformation (Figure 5C).[13–17]

In conclusion, the *in vitro* data presented here suggests that the before seen effects of these ribose analogs in 5'-UA sites can be translated to 5'-AA sites. This adds to the chemistry toolbox for enhancing the selectivity and efficacy of potential guide oligonucleotides for therapeutic ADAR editing.

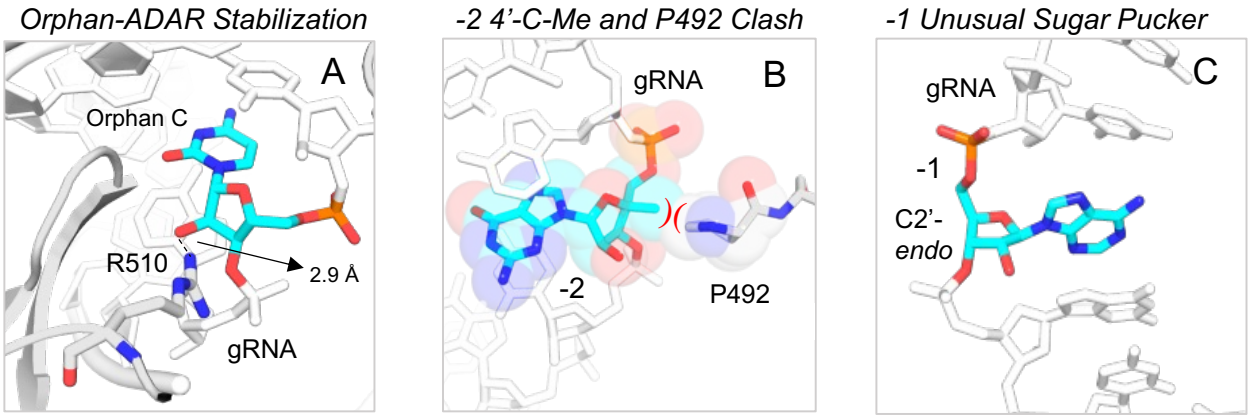


Figure 5. Structure of hADAR2-R2D E488Q bound to the GLI1 32 bp RNA at 2.8 Å resolution.[8] A) Zoom in view of the orphan base C 2'-OH group contacting the side chain of R510 in ADAR2 via hydrogen-bonding. B) Stick and space filling model overlay of the -2 nucleotide 4'-*C*-methyl closely approaching the methylene group in Pro 492 of ADAR's flipping loop. C) The -1 nucleotide adopts a C2'-*endo* sugar pucker with a high anti glycosidic bond angle compared to an ideal A-form nucleotide.

On-going experiments that are not included here include solving the crystal structure of RNA duplexes bearing 4'-C-MeA and 4'-C-MeU, in vitro editing data of the effect of 2'-O-4'-C-diMeA in a 5'-UAA site, synthesis of the solid supports of the dimethylated A and U analogs and the human serum stability/nuclease resistance of the newly synthesized analogs in RNA.

MATERIALS AND METHODS

Synthesis of phosphoramidites 12a and 12b. **1-((2*R*,3*R*,4*S*,5*R*)-4-(benzyloxy)-5-(((*tert*-butyldiphenylsilyl)oxy)methyl)-3-hydroxy-5-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1*H*,3*H*)-dione (7a).** To a solution of **6a** (0.962 g, 1.5 mmol) in 34 mL of MeOH was added 3.86 mL of 1M aqueous NaOH, sealed and stirred at room temperature for 1.25 h. After completion of the reaction, 2 mL of 50% (v/v) AcOH was added, and the volatiles were evaporated. The crude mixture was washed with water and extracted three times with 50 mL of CH_2Cl_2. The combined organics were dried over Na_2SO_4, filtered, and reduced under pressure to afford a colorless syrup. The crude mixture was purified by silica gel chromatography with a 50% ethyl acetate in *n*-hexanes to afford **3** as a colorless syrup in 97% yield (853 mg, 1.45 mmol). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.49 (s, 1H), 7.73 (d, *J* = 8.2 Hz, 1H), 7.69 – 7.56 (m, 4H), 7.54 – 7.31 (m, 10H), 6.02

(d, *J* = 4.9 Hz, 1H), 5.38 (dd, *J* = 8.1, 2.0 Hz, 1H), 4.73 (d, *J* = 11.5 Hz, 1H), 4.62 (d, *J* = 11.5 Hz, 1H), 4.33 (dd, *J* = 6.0, 4.9 Hz, 1H), 4.19 (d, *J* = 6.0 Hz, 1H), 3.79 (d, *J* = 11.2 Hz, 1H), 3.60 (d, *J* = 11.2 Hz, 1H), 1.26 (s, 3H), 1.11 (s, 9H). ^{13}C NMR (76 MHz, $CDCl_3$) δ 163.25, 150.87, 139.95, 137.12, 135.64, 135.38, 132.72, 132.16, 130.23, 130.16, 128.65, 128.28, 128.11, 128.07, 128.03, 127.87, 102.54, 88.90, 86.77, 77.85, 75.26, 74.06, 68.45, 27.08, 19.36, 18.87. Orbitrap-FTMS calculated for $C_{33}H_{38}N_2O_6Si$ $[M+H]^+$ 587.25 m/z, found 587.2533 m/z.

1-((2*R*,3*R*,4*S*,5*R*)-4-(benzyloxy)-5-(((*tert*-butyldiphenylsilyl)oxy)methyl)-3-methoxy-5-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1*H*,3*H*)-dione (8a). The experimental procedure was adapted from Koizumi et al.[6] NaH (60% in mineral oil) (0.211 g, 5.28 mmol, 3 eq.) was added to a solution of **7a** (1.00 g, 1.76 mmol) in THF (11.2 mL) at 0 °C under argon atmosphere and the mixture was stirred 10 min at 0 °C. Iodomethane (MeI, 0.657 mL, 10.56 mmol, 6 eq.) was added dropwise to the mixture at 0 °C and stirred 16 h at 25 °C. Saturated $NaHCO_3$ solution was added to the mixture at 0 °C. The mixture was extracted with saturated $NaHCO_3$ solution and ethyl acetate. The organic layer was washed with brine, dried over Na_2SO_4 and concentrated in vacuo. The crude yellow oil residue was purified by chromatography on silica gel 30-45% ethyl acetate in *n*-hexanes to give **8a** as a white foam in 75% (0.762 g, 1.27 mmol). ^{1}H NMR (400 MHz, CD_2Cl_2) δ 9.32 (s, 1H), 7.89 (d, *J* = 8.1 Hz, 1H), 7.83 – 7.30 (m, 15H), 6.11 (d, *J* = 2.6 Hz, 1H), 5.17 (d, *J* = 2.0 Hz, 1H), 4.73 (d, *J* = 11.6 Hz, 1H), 4.57 (d, *J* = 11.6 Hz, 1H), 4.32 (d, *J* = 5.9 Hz, 1H), 3.92 – 3.79 (m, 2H), 3.74 (d, *J* = 11.4 Hz, 1H), 3.41 (s, 3H), 1.27 (s, 3H), 1.14 (s, 9H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 163.21, 150.12, 140.03, 137.79, 135.67, 135.58, 135.45, 135.38, 133.07, 132.28, 130.11, 130.09, 130.01, 129.98, 128.60, 128.44, 127.99, 127.96, 127.91, 127.76, 102.05, 87.67, 86.69, 84.02, 75.61, 72.84, 67.22, 58.97, 26.88, 19.24, 18.50. Orbitrap-FTMS calculated for $C_{34}H_{40}N_2O_6Si$ $[M+H]^+$ 601.27 m/z, found 601.2733 m/z.

1-((2*R*,3*R*,4*S*,5*R*)-4-(benzyloxy)-5-(hydroxymethyl)-3-methoxy-5-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1*H*,3*H*)-dione (9a). 8a (0.762 g, 1.27 mmol) was co-evaporated twice with

anhydrous THF. Subsequently, the silyl-protected nucleoside was dissolved in an anhydrous 1M tetrabutylammonium fluoride solution in THF (2 mL, 2 mmol) and stirred for 24 h at room temperature under an argon atmosphere. After complete conversion of the starting material, the mixture was reduced to a light yellow oil and chromatographed with silica gel with a 2-3 % MeOH in CH_2Cl_2 eluent to afford **9a** as a white solid in 98 % (0.450 g, 1.23 mmol). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.43 (s, 1H), 8.67 (s, 1H), 7.96 (d, 2H), 7.75 (s, 1H), ^{1}H NMR (300 MHz, CD_2Cl_2) δ 9.80 (s, 1H), 7.79 (d, *J* = 8.1 Hz, 1H), 7.54 – 7.23 (m, 5H), 5.86 (d, *J* = 3.5 Hz, 1H), 5.73 (d, *J* = 8.1 Hz, 1H), 4.74 (d, *J* = 11.5 Hz, 1H), 4.62 (d, *J* = 11.5 Hz, 1H), 4.25 (d, *J* = 5.8 Hz, 1H), 4.07 (dd, *J* = 5.9, 3.6 Hz, 1H), 3.71 (d, *J* = 11.9 Hz, 1H), 3.59 (d, *J* = 11.9 Hz, 1H), 3.51 (s, 3H), 3.02 (s, 1H), 1.30 (s, 3H). ^{13}C NMR (76 MHz, CD_2Cl_2) δ 163.82, 150.45, 141.87, 138.09, 128.37, 127.83, 127.73, 102.22, 90.07, 87.04, 83.15, 76.55, 73.25, 66.43, 58.93, 18.38. Orbitrap-FTMS calculated for $C_{18}H_{22}N_2O_6$ $[M+H]^+$ 363.15 m/z, found 363.1557 m/z.

1-((2*R*,3*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)-3-methoxy-5-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1*H*,3*H*)-dione (10a). Debenzylation procedure was adapted and modified from Zhou *et al*.[18] **9a** (0.200 g, 0.522 mmol) was co-evaporated 3x with anhydrous dichloromethane and dried overnight under high vacuum with a stir bar. The 3' benzylated nucleoside was then dissolved in 22 mL of anhydrous dichloromethane and stirred 5 min at -78 °C under an argon atmosphere, followed by the dropwise addition of 1M BCl_3 in CH_2Cl_2 (3.46 mL, 3.46 mmol, 6.8 eq.). After stirring for 8 h at the same temperature, the reaction was quenched with 6 mL of MeOH:DCM (1:1) at -78 °C followed by 3 g of $NaHCO_3$. The mixture was left stirring for 30 min at room temperature. The insoluble solid was then filtered through and washed with 20 mL of 1:2 DCM:MeOH. The filtrate was evaporated to dryness and chromatographed with silica gel using a 5-8% gradient of MeOH in CH_2Cl_2 to afford **10a** as a white solid in 88% yield (136 mg, 0.5 mmol). ^{1}H NMR (400 MHz, MeOD) δ 8.13 (d, *J* = 8.1 Hz, 1H), 6.01 (d, *J* = 4.8 Hz, 1H), 5.72 (d, *J* = 8.1 Hz, 1H), 4.32 (d, *J* = 5.7 Hz, 1H), 4.01 (t, *J* = 5.3 Hz, 1H), 3.60 (q, 2H), 3.50 (s, 3H), 1.22 (s, 3H).

^{13}C NMR (101 MHz, MeOD) δ 164.76, 150.89, 141.42, 101.37, 87.23, 86.88, 84.20, 69.74, 66.13, 57.61, 17.23. Orbitrap-FTMS calculated for $C_{11}H_{16}N_2O_6$ $[M+H]^+$ 273.10 m/z, found 273.1081 m/z.

1-((2*R*,3*R*,4*S*,5*R*)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxy-3-methoxy-5-methyltetrahydrofuran-2-yl)pyrimidine-2,4(1*H*,3*H*)-dione (11a). After co-evaporating **10a** (0.106 g, 0.392 mmol) 3x with anhydrous pyridine (3x2 mL) and drying it overnight under high vacuum, the nucleoside was dissolved with distilled and anhydrous pyridine (1.4 mL). *4*',*4*'-Dimethoxytrityl chloride (0.198 g, 0.588 mmol, 1.5 eq.) was added in one portion to the solution and stirred for 24 h under argon and at room temperature and quenched with 1 mL of MeOH. All the volatiles were evaporated, and the resulting crude was diluted with 30 mL of CH_2Cl_2, washed 1x with 5 % aqueous $NaHCO_3$ (10 mL), 1x with brine (10 mL), dried over Na_2SO_4 and filtered. The crude was purified via silica gel chromatography (base washed with 1% TEA) with a gradient of 2 % MeOH in CH_2Cl_2 to give **11a** as a yellow foam in 60 % yield (202 mg, 0.35 mmol). ^{1}H NMR (300 MHz, CD_2Cl_2) δ 9.24 (s, 1H), 7.92 (d, *J* = 8.2 Hz, 1H), 7.49 – 7.24 (m, 9H), 6.96 – 6.84 (m, 4H), 6.04 (d, *J* = 2.8 Hz, 1H), 5.27 (dd, *J* = 8.2, 1.9 Hz, 1H), 4.59 (s, 1H), 3.91 (dd, *J* = 6.2, 2.9 Hz, 1H), 3.83 (s, 6H), 3.62 (s, 3H), 3.30 (q, 2H), 2.89 (s, 1H), 1.21 (s, 3H). ^{13}C NMR (76 MHz, CD_2Cl_2) δ 163.23, 158.82, 158.80, 150.25, 144.53, 140.26, 135.36, 135.07, 130.19, 130.14, 128.14, 127.99, 127.06, 113.24, 113.22, 102.10, 87.03, 86.85, 86.67, 84.98, 69.86, 66.96, 59.21, 55.25, 18.33. Orbitrap-FTMS calculated for $C_{32}H_{34}N_2O_8$ $[M+H]^+$ 575.23 m/z, found 575.2394 m/z.

(2*R*,3*S*,4*R*,5*R*)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)-4-methoxy-2-methyltetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (12a). 6 was co-evaporated 3x with anhydrous MeCN. To a solution of **11a** (0.082 g, 0.142 mmol), distilled over CaH_2 *N,N*-diisopropylethylamine (116 µL, 0.852 mmol, 6 eq.), and 1 mL of anhydrous CH_2Cl_2 was added 2-cyanoethyl *N,N*-

diisopropylchlorophosphoramidite (80 µL, 0.365 mmol, 2.6 eq.) under an argon atmosphere. The reaction was allowed to stir at room temperature for 2 h. The reaction was quenched with 0.05 mL of MeOH and further diluted with ethyl acetate. The organic phase was then washed with saturated aqueous $NaHCO_3$ once, dried over Na_2SO_4 and filtered to afford a colorless oil. The residue was purified over a 0.5 % TEA-washed silica gel column with a gradient of 50 – 70 % ethyl acetate in *n*-hexanes to furnish the phosphoramidite **12a** as a colorless foam in 94 % yield (103 mg, 0.132 mmol). ^{31}P NMR (122 MHz, CD_2Cl_2) δ 151.16, 150.31. Orbitrap-FTMS calculated for $C_{41}H_{51}N_4O_9P$ $[M+H]^+$ 775.35 m/z, found 775.3502 m/z.

***N*-(9-((2*R*,3*R*,4*S*,5*R*)-4-(benzyloxy)-5-(((*tert*-butyldiphenylsilyl)oxy)methyl)-3-methoxy-5-methyltetrahydrofuran-2-yl)-9*H*-purin-6-yl)benzamide (8b).** Synthetic procedures were adapted from Kajino et al.[5] NaH (60% suspension in oil) (84 mg, 2.1 mmol) was added to a solution of compound **7b** (0.5 g, 0.70 mmol) in dry THF (6 mL) at 0 °C. Then, CH_3I (0.238 mL, 3.85 mmol) was added dropwise and stirred 8 hours at 0 °C. A saturated $NaHCO_3$ aqueous solution was added, then the mixture was extracted with ethyl acetate and saturated $NaHCO_3$ aqueous solution. The organic layer was washed with brine, dried over Na_2SO_4, filtered, and concentrated. The crude material was purified by column chromatography (50-60% ethyl acetate in *n*-hexanes) to afford **8b** as a light yellow foam (0.42 g, 0.58 mmol, 82%).^1H NMR (400 MHz, DMSO-d_6) δ 11.20 (s, 1H), 8.60 (d, *J* = 10.5 Hz, 2H), 8.04 (dt, *J* = 7.1, 1.3 Hz, 2H), 7.67 – 7.52 (m, 7H), 7.48 – 7.27 (m, 12H), 6.22 (d, *J* = 4.7 Hz, 1H), 4.79 (dd, *J* = 10.8, 5.8 Hz, 1H), 4.74 (s, 1H), 4.65 (d, *J* = 11.8 Hz, 1H), 4.52 (d, *J* = 5.4 Hz, 1H), 3.83 (d, *J* = 10.8 Hz, 1H), 3.61 (d, *J* = 10.8 Hz, 1H), 3.34 (s, 3H), 1.34 (s, 3H), 0.96 (s, 9H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.66, 151.78, 151.55, 150.52, 143.49, 138.15, 135.11, 135.08, 133.35, 132.68, 132.49, 132.46, 129.93, 128.52, 128.49, 128.31, 128.05, 127.91, 127.83, 127.74, 127.69, 126.02, 86.23, 86.11, 81.43, 76.88, 72.50, 67.88, 58.34, 26.64, 18.81, 18.55. Orbitrap-FTMS calculated for $C_{42}H_{45}N_5O_5Si$ $[M+H]^+$ 728.33 m/z, found 728.3272 m/z.

***N*-(9-((2*R*,3*R*,4*S*,5*R*)-4-(benzyloxy)-5-(hydroxymethyl)-3-methoxy-5-methyltetrahydrofuran-2-yl)-9*H*-purin-6-yl)benzamide (9b). 8b** (0.90 g, 1.25 mmol) was co-evaporated twice with anhydrous THF. Subsequently, the silyl-protected nucleoside was dissolved in 8.6 mL of THF. 1M tetrabutylammonium fluoride solution in THF (1.9 mL, 1.9 mmol) was added to the solution and stirred for 5 h at room temperature under an argon atmosphere. After complete conversion of the starting material, the mixture was reduced to a light yellow oil and chromatographed with silica gel with a 2-3 % MeOH in CH_2Cl_2 eluent to afford **9b** as a white solid in 82% (0.5 g, 1.02 mmol). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.80 (s, 1H), 8.10 (s, 1H), 8.05 – 7.97 (m, 2H), 7.68 – 7.57 (m, 1H), 7.53 (dd, *J* = 8.3, 6.8 Hz, 2H), 7.46 – 7.28 (m, 5H), 5.96 (d, *J* = 7.7 Hz, 1H), 4.85 – 4.77 (m, 2H), 4.70 (d, *J* = 11.6 Hz, 1H), 4.31 (d, *J* = 5.0 Hz, 1H), 3.80 (d, *J* = 12.6 Hz, 1H), 3.56 (d, *J* = 12.6 Hz, 1H), 3.28 (s, 3H), 1.32 (s, 3H).^{13}C NMR (101 MHz, $CDCl_3$) δ 164.89, 152.56, 150.87, 150.69, 143.76, 138.25, 133.87, 133.45, 129.42, 128.97, 128.79, 128.48, 128.42, 128.34, 124.86, 90.11, 89.61, 83.81, 74.97, 69.44, 59.57, 19.32.Orbitrap-FTMS calculated for $C_{26}H_{27}N_5O_5$ $[M+H]^+$ 490.20 m/z, found 490.2094 m/z.

***N*-(9-((2*R*,3*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)-3-methoxy-5-methyltetrahydrofuran-2-yl)-9*H*-purin-6-yl)benzamide (10b).** The synthetic procedure was adapted from Zhou et al.[18] and Jauregui-Matos et al.[3] **9b** (0.5 g, 1 mmol) was dissolved in 42 mL of anhydrous DCM and cooled to -78 °C. To this solution was added 1M of boron trichloride solution in DCM (6 mL, 6 mmol). The reaction was left stirring for 6 h under the same temperature. The reaction was quenched by adding 26.6 mL of MeOH and 5.3 g of $NaHCO_3$. The suspension was left stirring for 4 h at 25 °C. Afterwards, the volatiles were removed under reduced pressure and the crude mixture was purified by silica gel chromatography (2-6% MeOH in DCM) to afford **10b** in 61% yield (0.247 g, 0.617 mmol). ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.21 (s, 1H), 8.76 (d, *J* = 2.0 Hz, 2H), 8.10 – 7.99 (m, 2H), 7.69 – 7.59 (m, 1H), 7.56 (dd, *J* = 8.2, 6.9 Hz, 2H), 6.15 (d, *J* = 6.9 Hz, 1H), 5.29 (t, *J* =

5.7 Hz, 1H), 5.23 (d, *J* = 5.9 Hz, 1H), 4.65 (dd, *J* = 7.0, 5.0 Hz, 1H), 4.32 (t, *J* = 5.4 Hz, 1H), 3.58 (d, *J* = 5.3 Hz, 1H), 3.41 (dd, *J* = 11.6, 6.2 Hz, 1H), 3.32 (s, 3H), 1.19 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.63, 152.20, 151.69, 150.47, 143.16, 133.31, 132.47, 128.49, 128.47, 125.81, 88.02, 84.95, 83.00, 69.62, 66.60, 57.41, 54.91, 18.74.Orbitrap-FTMS calculated for $C_{19}H_{21}N_5O_5$ $[M+H]^+$ 400.15 m/z, found 400.1621 m/z.

***N*-(9-((2*R*,3*R*,4*S*,5*R*)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxy-3-methoxy-5-methyltetrahydrofuran-2-yl)-9*H*-purin-6-yl)benzamide (11b).** To a solution of compound **10b** (0.245 g, 0.61 mmol) in pyridine (4 mL) were added 4,4-dimethoxytrityl chloride (0.271 g, 0.80 mmol) and DMAP (62.16 mg, 508.81 µmol) and the mixture was stirred at 25°C for 16h. The reaction mixture was diluted with EtOAc, washed with saturated $NaHCO_3$ (aq.) and brine, and dried over Na_2SO_4. The organic phase was evaporated to dryness and purified by silica gel column chromatography using 1-2% MeOH in a mixture of 1:1 EtOAc and hexanes. Purification afforded **11b** as a light yellow foam in 66% yield (0.285 g, 0.405 mmol). ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.21 (s, 1H), 8.63 (s, 1H), 8.54 (s, 1H), 8.08 – 8.00 (m, 2H), 7.65 (d, *J* = 7.5 Hz, 1H), 7.56 (dd, *J* = 8.3, 6.9 Hz, 2H), 7.39 (dt, *J* = 8.3, 1.2 Hz, 2H), 7.33 – 7.19 (m, 7H), 6.92 – 6.81 (m, 4H), 6.12 (d, *J* = 6.0 Hz, 1H), 5.32 (d, *J* = 6.1 Hz, 1H), 4.63 (t, *J* = 5.6 Hz, 1H), 4.38 (t, *J* = 5.6 Hz, 1H), 3.73 (d, *J* = 1.1 Hz, 6H), 3.31 (s, 3H), 3.24 (d, *J* = 9.7 Hz, 1H), 3.12 (d, *J* = 9.8 Hz, 1H), 1.30 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 158.05, 151.52, 150.50, 149.61, 144.86, 143.65, 135.60, 135.38, 132.46, 129.74, 128.48, 128.46, 127.79, 127.69, 126.67, 123.90, 113.15, 86.32, 85.48, 81.70, 69.90, 67.65, 57.73, 55.00, 40.15, 39.94, 39.73, 39.52, 39.31, 39.10, 38.89, 19.13. Orbitrap-FTMS calculated for $C_{40}H_{39}N_5O_7$ $[M+H]^+$ 702.28 m/z, found 702.2956 m/z.

(2*R*,3*S*,4*R*,5*R*)-5-(6-benzamido-9*H*-purin-9-yl)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-methoxy-2-methyltetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (12b). To a stirred solution of **11b** (0.280 g, 0.40

mmol) in DCM (4 mL) was added DIPEA (0.38 mL, 2.2 mmol) , N-methylimidazole (32.69 µL, 0.4 mmol) and 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.175 mL, 0.78 mmol). The reaction was left to stir for 45 min at 25˚C and then quenched with a saturated aqueous solution of sodium bicarbonate. The organic layer was extracted with excess DCM and dried over sodium sulfate, filtered, and then evaporated under reduced pressure. The crude mixture was then purified by silica gel chromatography using 70% EtOAc in hexanes to furnish **12b** in 77% (.277 g, 0.30 mmol). ^{31}P NMR (122 MHz, CD_3CN) δ 149.98, 149.90. Orbitrap-FTMS calculated for $C_{49}H_{56}N_7O_8P$ $[M+H]^+$ 902.39 m/z, found 902.4015 m/z.

Oligonucleotide synthesis and purification. All oligonucleotides were synthesized in 0.2 µmol scale based on the phosphoramidite chemistry and standard DNA/RNA reagents and cycles. DMTr-Off and 2'-*O*-TBDMSi protecting strategy was employed for all RNA oligonucleotide synthesis. The coupling time for all phosphoramidites was between 3-10 min. For incorporation of the 4'-*C*-MeA, 4'-C-MeU, **2′-*O*-4′-*C*-diMeA, and 2′-*O*-4′-*C*-diMeU** phosphoramidite, an extended coupling time of 45 min was used and was dissolved with a 4:1 mixture of MeCN:DCM to 0.1 M.[2] Coupling efficiency of all novel amidites was measured using the conductivity released by the trityl cation during detritylation of the growing strand with 3% trichloroacetic acid in CH_2Cl_2 and was comparable with commercially available phosphoramidites detritylations. Post synthesis, columns were dried for 0.5 h under high vacuum. Cleavage from the solid support and deprotection of the nucleobases were accomplished using a fresh aliquot of ammonium hydroxide in ethanol at 55 °C for 10 h (1 mL of 28-30% NH_4OH/EtOH in 0.5 mL EtOH). After cleavage, the solid support was allowed to cool to room temperature and centrifuged for 5 min at 13,000 x g. The supernatants were collected and lyophilized. 2'-*O*-silyl ether deprotection was done using 125 µL of hydrogen fluoride-triethylamine cocktail in 100 µL of DMSO for 24 h at room temperature. Crude oligonucleotides were desalted by precipitation with the addition of 1 mL of

1-butanol and 25 μL of 3 M NaOAc and cooled for 2 h at -70 °C. The solution was centrifuged at 13,000 × g for 30 min, supernatants were removed, and the pellets were washed twice with cold ethanol.

The oligonucleotides were purified by loading a 1:1 mixture of 10-20 nmol of crude RNA to 80% formamide in 1X Tris-Borate EDTA (TBE) and 1mM EDTA into a denaturing polyacrylamide gel. Depending on their size, oligonucleotides were loaded into different acrylamide matrices with different percentages and electrophoresed in 1X TBE. After visualization by UV shadowing on glass-backed silica gel plates, gel bands were excised, crushed, and soaked overnight at 4 °C in 0.5 mL of 0.5 M NH_4OAc and 0.1 mM EDTA. Gel particles were removed from the extracted RNA by filtering the suspension for 5 min, 13,000 x g on 0.2 μm cellulose sterile membrane filter tubes (Corning). The oligonucleotides were precipitated from a solution of 75% EtOH at −70 °C for 2 h. The oligonucleotides were then pelleted from the solution by centrifugation (13,000 × g for 30 min), with the supernatant being discarded. The pellets were lyophilized to dryness, resuspended in nuclease-free water, and quantified by absorbance at 260 nm. Oligonucleotide mass was confirmed by MALDI-TOF (Table S2).

ADAR2 overexpression and purification and melting temperature of oligonucleotides were performed as previously described. [3,19]

Target RNA Synthesis*.* The gene fragment of the corresponding human SRC were purchased from Integrated DNA Technologies. These were amplified with New England BioLabs (NEB) Q5 Hot Start high-fidelity DNA polymerase and purified by extraction from 1% agarose gels. Target RNA was synthesized with NEB's HiScribe T7 RNA polymerase according to the protocol. The target RNAs were purified by 8% denaturing polyacrylamide gel electrophoresis.

Hybridization of guide and target RNA. The target and guide RNA were combined in a 1:10 ratio of target:guide to a final concentration of 180 nM of hybrids in 1X TE and 100 mM NaCl. This solution was heated to 95 °C for 5 min then it was allowed to cool to room temperature for 2 h.

In vitro Deamination Reaction and Sanger Sequencing. For ADAR2 WT deaminations, hybridized target and guide RNA were diluted to 5 nM in 1X ADAR2 reaction buffer (15 mM Tris-HCl pH 7.5, 3% glycerol, 60 mM KCl, 1.5 mM EDTA, 0.003% Nonidet P-40, and 3 mM $MgCl_2$), 0.5 mM dithiothreitol (DTT), 160 units/mL RNAse inhibitor, 1 μg/mL yeast tRNA, and 15 nM ADAR2 protein. The reactions were conducted at 30 °C. For two time points reactions analysis, as well as for kinetic analysis (1, 3, 5, 10, 15, 30, 60 min), aliquots of 8 μL were quenched with 190 μL of 95 °C heated nuclease-free water for 5 min. A 5 μL aliquot of each time point quench solution was reverse transcribed with Access Reverse Transcription (RT) Polymerase Chain Reaction (RT-PCR) system purchased from Promega. DNA was cleaned and concentrated using Zymo's purification kit and protocol. For Sanger sequencing with Azenta Life Sciences, purified DNA was diluted to 0.66 ng/μL in 1.6 μM of forward or reverse sequencing primer. Sanger sequencing traces were observed with SnapGene Viewer and statistical analysis along with nonlinear fits were conducted with Microsoft Excel and GraphPad Prism. The editing level for the corresponding zero time point was subtracted from each data point as background subtraction.

Oligonucleotide Sequences and Masses

Table S1. Sequences of oligonucleotides. Primers are 2'-deoxynucleotides, all others are ribonucleotides unless otherwise stated. Underlined nucleotides indicate 2'-*O*-Me, target adenosines are indicated in red, nucleotides in (parenthesis) are modified and represent the -1 or -2 nucleotide. X = 4'-C-MeA; Y = 4'-C-MeU; W = 2',4'-diMeA; Q = 2',4'-diMeU.

Oligonucleotide	Sequence
SRC K295R target RNA sequence	5'GGGCCACGUCCAAGCCGCAGACUCAGGGCCUGGCCAAGGAUGC CUGGGAGAUCCCUCGGGAGUCGCUGCGGCUGGAGGUCAAGCUG GGCCAGGGCUGCUUUGGCGAGGUGUGGAUGGGGACCUGGAACG GUACCACCAGGGUGGCCAUCA**A**AACCCUGAAGCCUGGCACGAUG UCUCCAGAGGCCUUCCUGCAGGAGGCCCAGGUCAUGAAGAAGCU

	GAGGCAUGAGAAGCUGGUGCAGUUGUAUGCUGUGGUUUCAGAG GAGCCCAUUUACAUCGUCACGGAGUACAUGAGCAAGGGGAG3’
SRC gRNAs	5’-CAGGCUUCAGGGUUC(U)GAUGGCCACCCUG-3’
Standard DNA oligo sequence for MALDI	5’-TTACGCCAGAATGCGTTCGCACAGCCGCCA-3’
SRC RT-PCR primer forward	5’-TCCAAGCCGCAGACTCAG-3’
SRC amplification, RT-PCR and sequencing primer reverse	5’-CTCCCCTTGCTCATGTACTCC-3’
SRC T7 DNA amplification forward primer	5’-TAATACGACTCACTATAGGGCCAC-3’
Top Control rA Tm	5’-CAUUAAGGUGGG-3’
Top Control rU Tm	5’-CAUUAUGGUGGG-3’
Top Control 2’OMeA Tm	5’-CAUUAAGGUGGG-3’
Top Control 2’OMeU Tm	5’-CAUUAUGGUGGG-3’
Top 4’-*C*-MeA Tm	5’-CAUUA**X**GGUGGG-3’
Top 4’-*C*-MeU Tm	5’-CAUUA**Y**GGUGGG-3’
Top 2’,4’-diMeA Tm	5’-CAUUA**W**GGUGGG-3’
Top 2’,4’-diMeU Tm	5’-CAUUA**Q**GGUGGG-3’
Bottom rU Tm	5’-CCUACCUUGAUG-3’
Bottom rC Tm	5’-CCUACCCUGAUG-3’
Bottom rG Tm	5’-CCUACCGUGAUG-3’
Bottom rA Tm	5’-CCUACCAUGAUG-3’

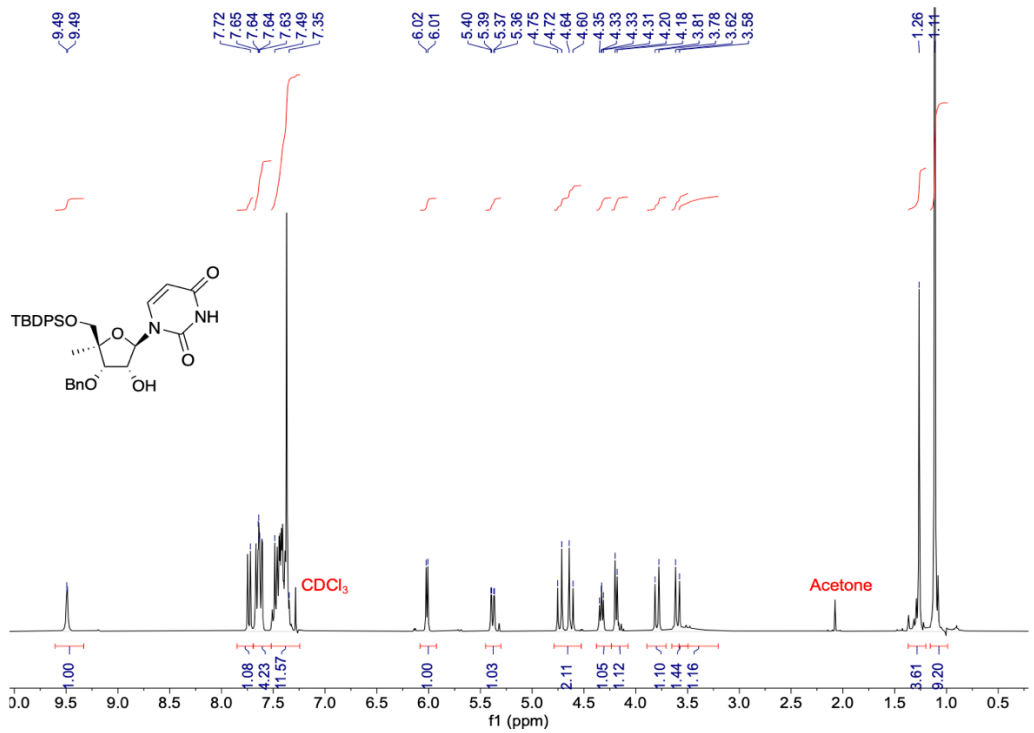
Table S2. Oligonucleotides were measured in linear positive mode analysis which typically results in mass accuracies of 0.1% or better. Standard oligonucleotides (commercially bought from Azenta Life Sciences or IDT and PAGE purified) were measured and showed mass accuracies better than 0.1% error.

Oligonucleotide Sample	Expected Mass (m/z)	Observed Mass (m/z)
SRC -1 rU	9251	9252
SRC -1 2’-O-MeU	9265	9266
SRC -1 4’-C-MeU	9265	9265
SRC -1 2’-O-4’-C-diMeU	9279	9279
SRC -1 dU	9240	9241
SRC -1 dT	9254	9254
SRC -1 2’-F-U	9253	9254
SRC -1 2’-F-ANA-U	9253	9254
Top Control rA Tm	3875	3875
Top Control rU Tm	3852	3852
Top Control 2’OMeA Tm	3889	3889
Top Control 2’OMeU Tm	3866	3866
Bottom rU Tm	3732	3732
Bottom rC Tm	3731	3731
Bottom rG Tm	3771	3771
Bottom rA Tm	3755	3755
Top 4’-*C*-MeA Tm	3889	3889
Top 4’-*C*-MeU Tm	3866	3866

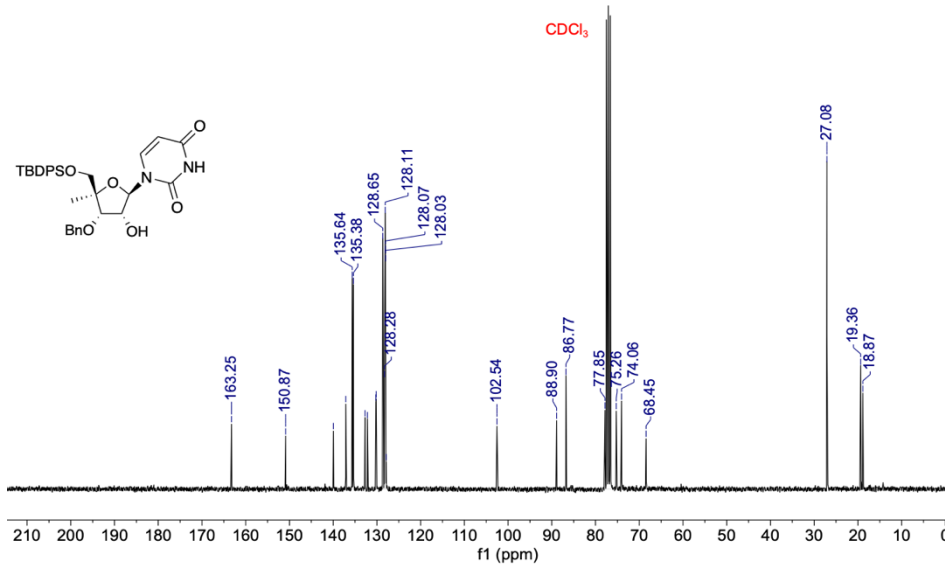
Top 2',4'-diMeA Tm	3903	3903
Top 2',4'-diMeU Tm	3880	3880

NMR Spectra

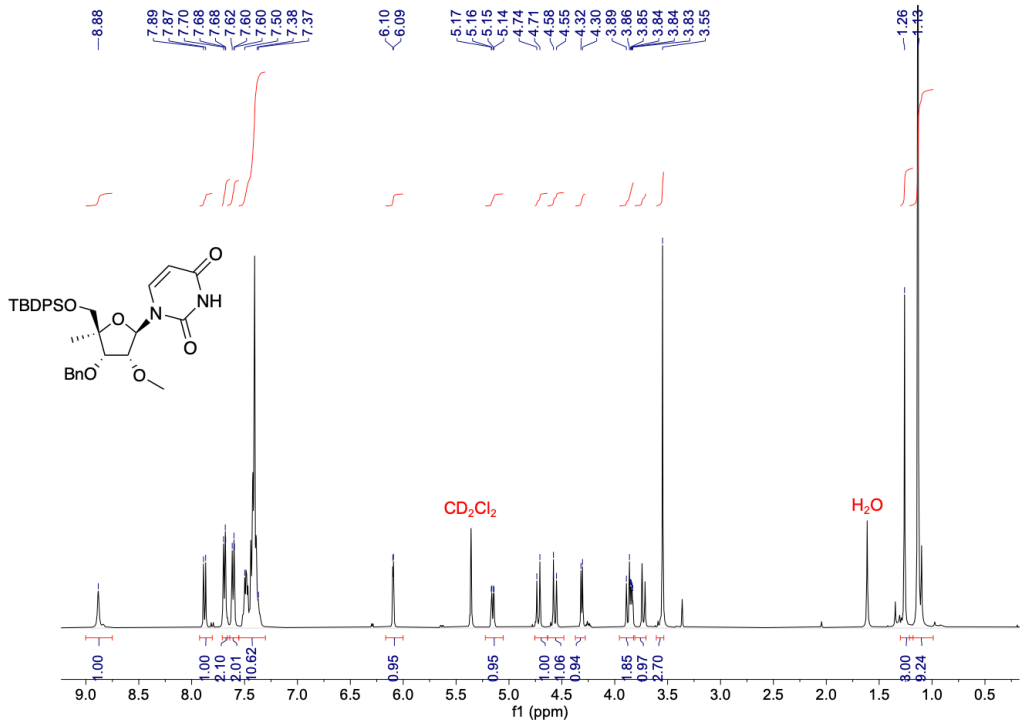
^{1}H NMR (300 MHz, $CDCl_3$) of **7a**



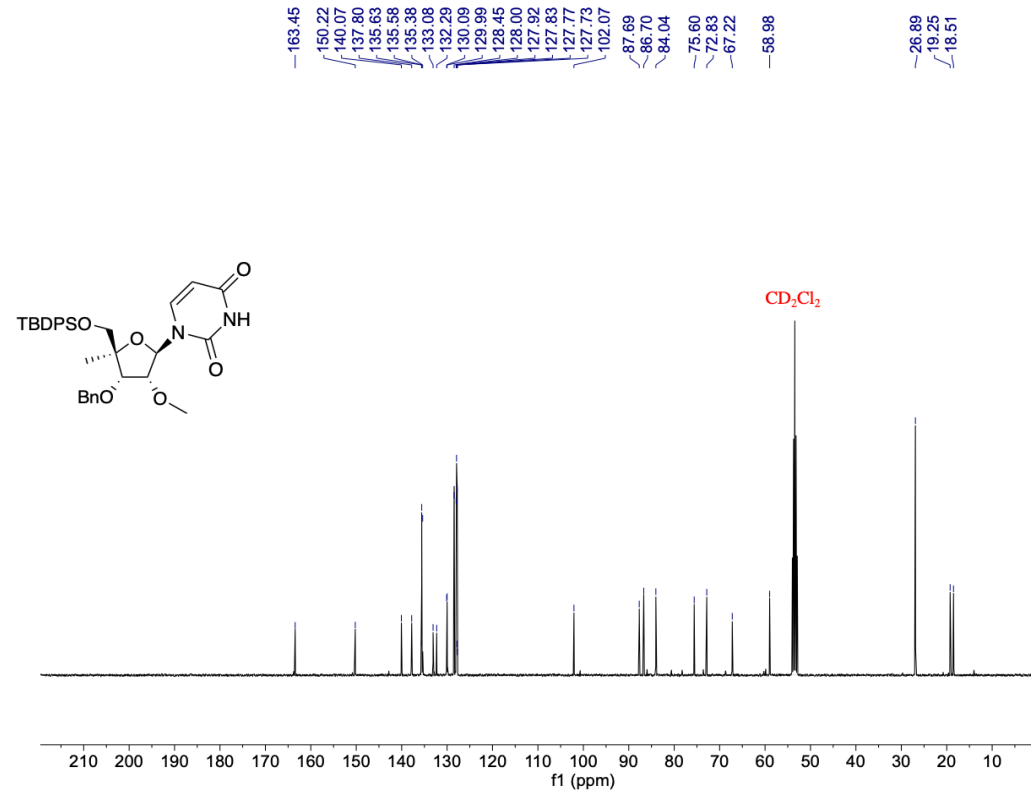
^{13} C NMR (76 MHz, $CDCl_3$) of **7a**



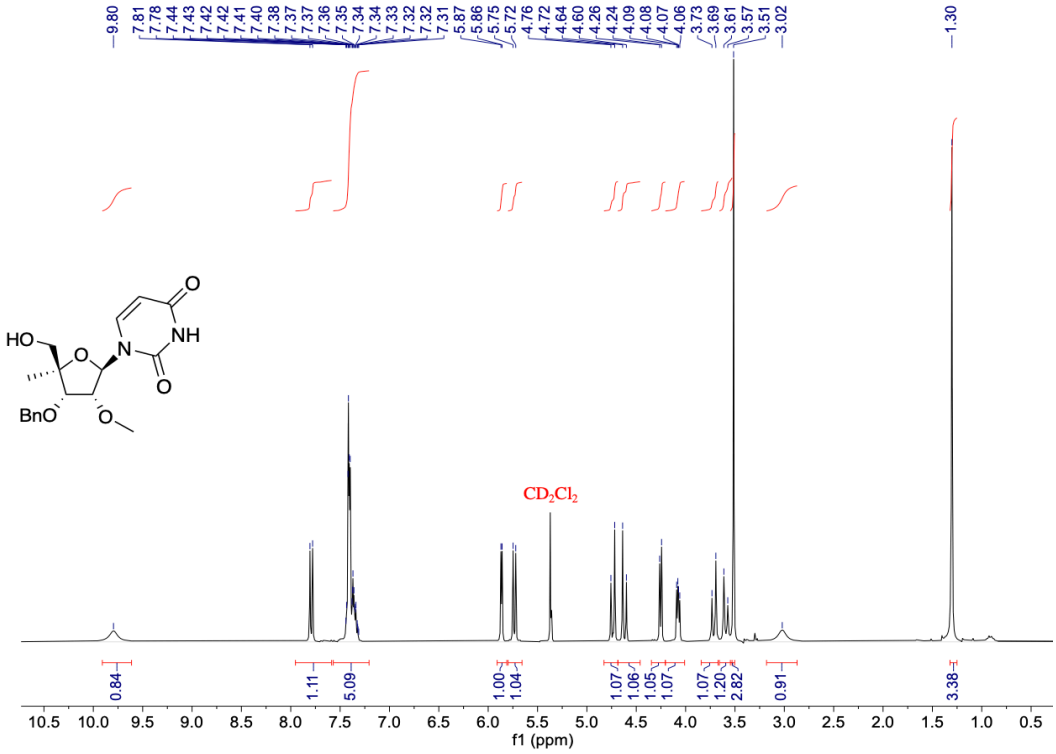
^{1}H NMR (400 MHz, CD_2Cl_2) of **8a**



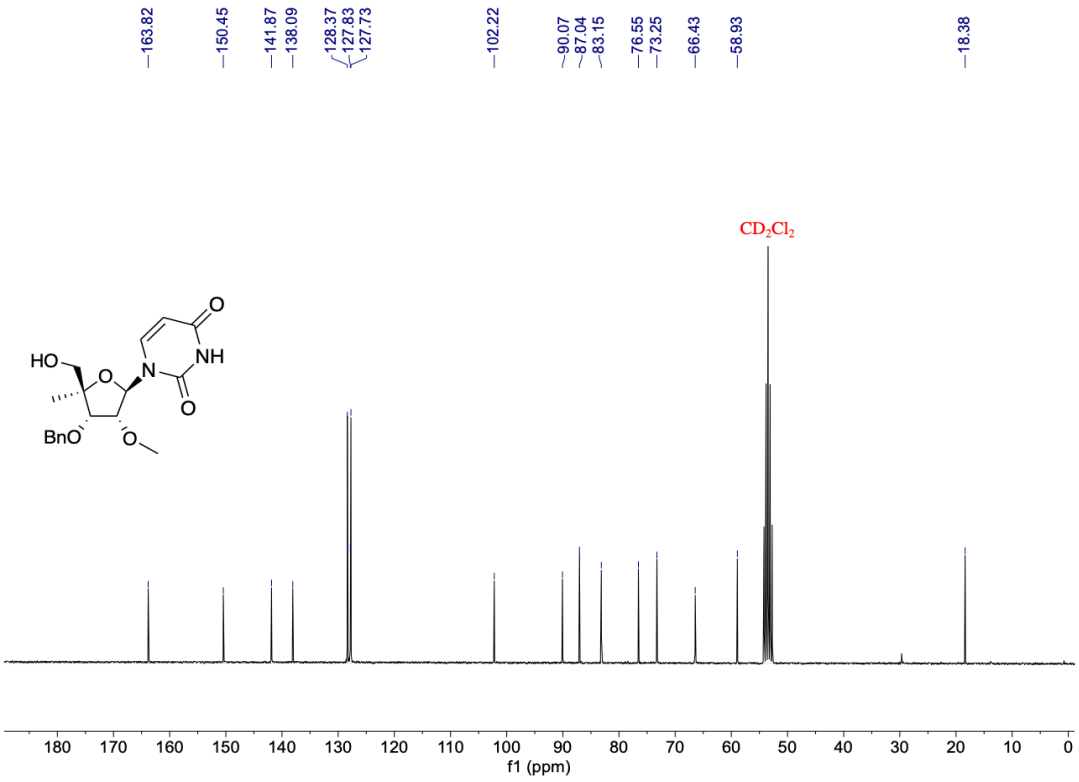
^{13}C NMR (76 MHz, CD_2Cl_2) of **8a**



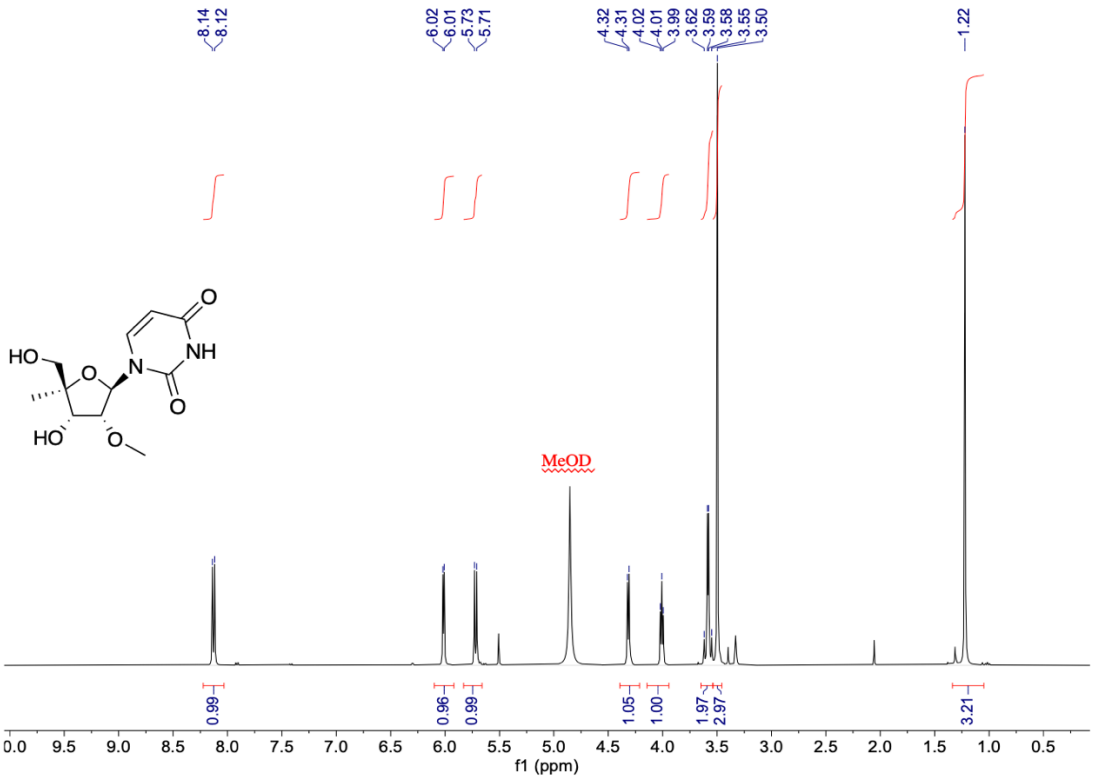
^{1}H NMR (300 MHz, CD_2Cl_2) of **9a**



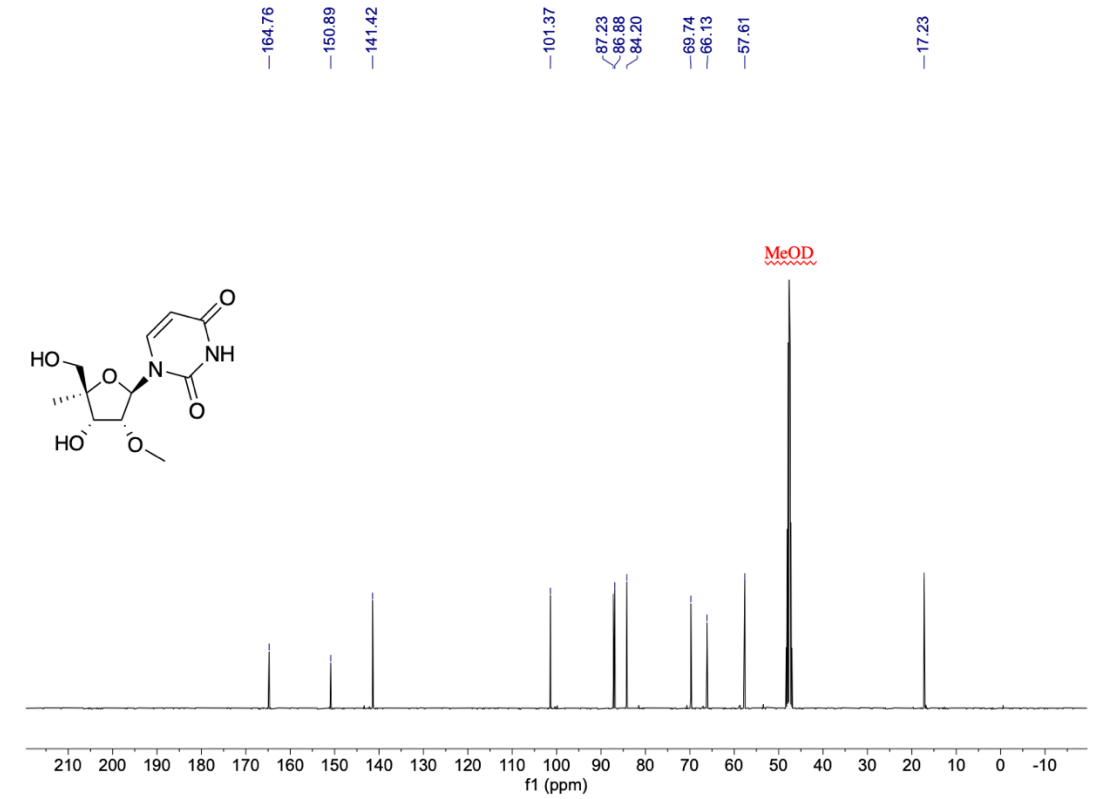
^{13}C NMR (76 MHz, CD_2Cl_2) of **9a**



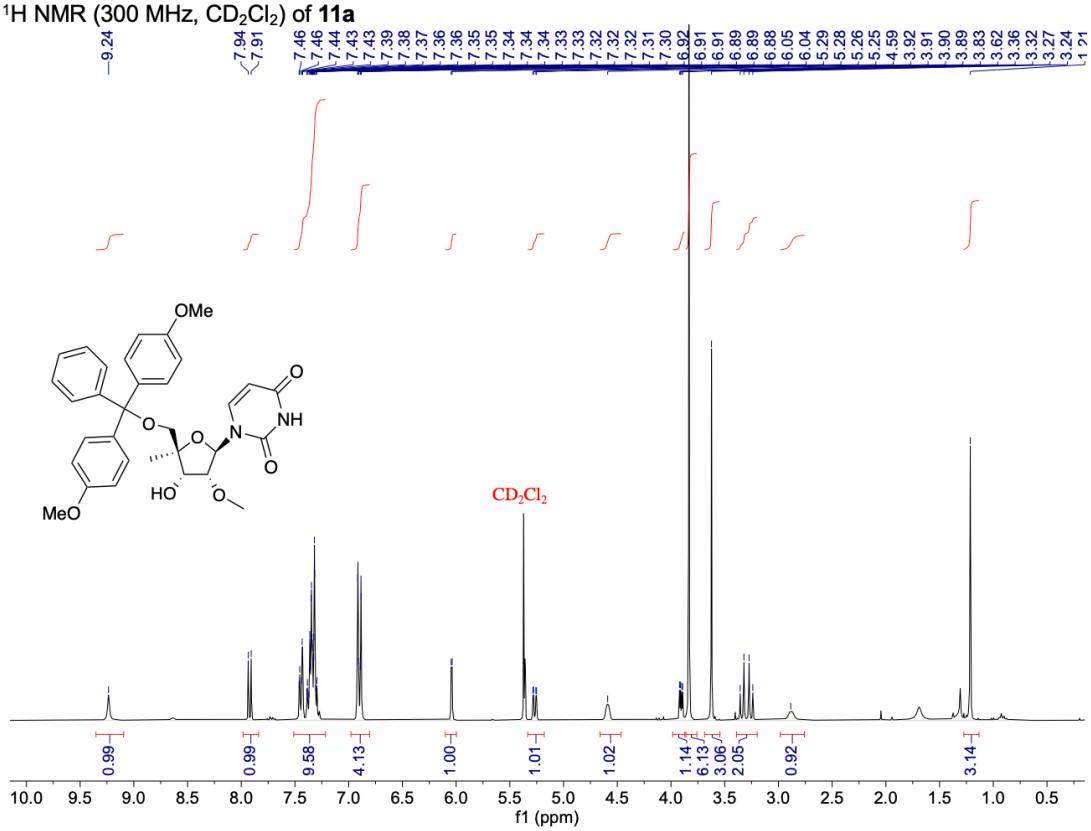
^{1}H NMR (400 MHz, MeOD) of **10a**



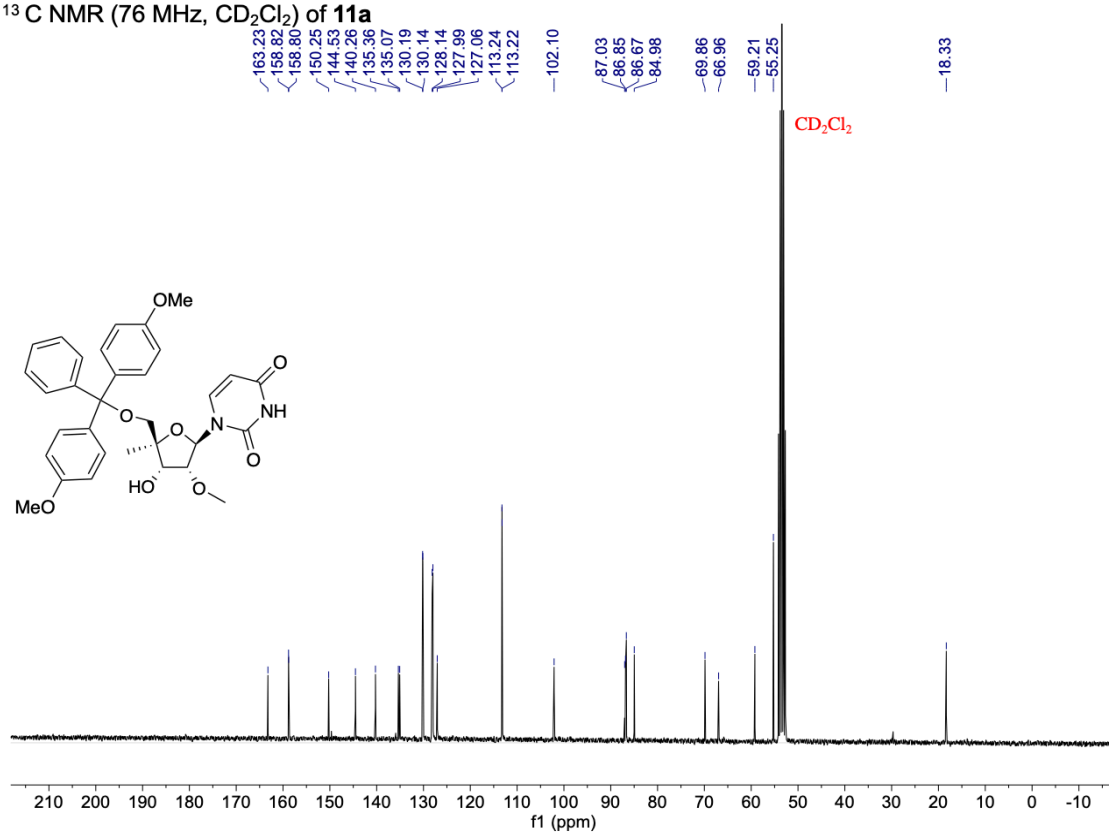
^{13}C NMR (101 MHz, MeOD) of **10a**



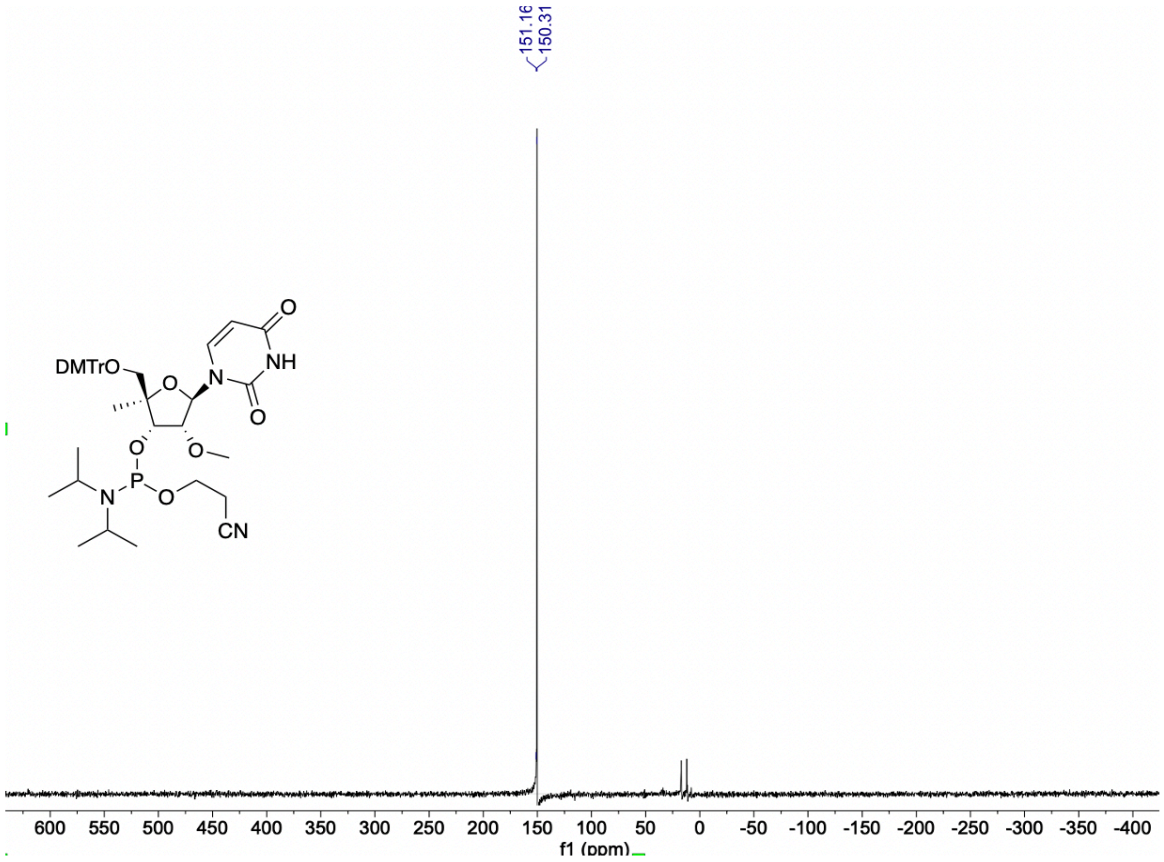
^{1}H NMR (300 MHz, CD_2Cl_2) of **11a**

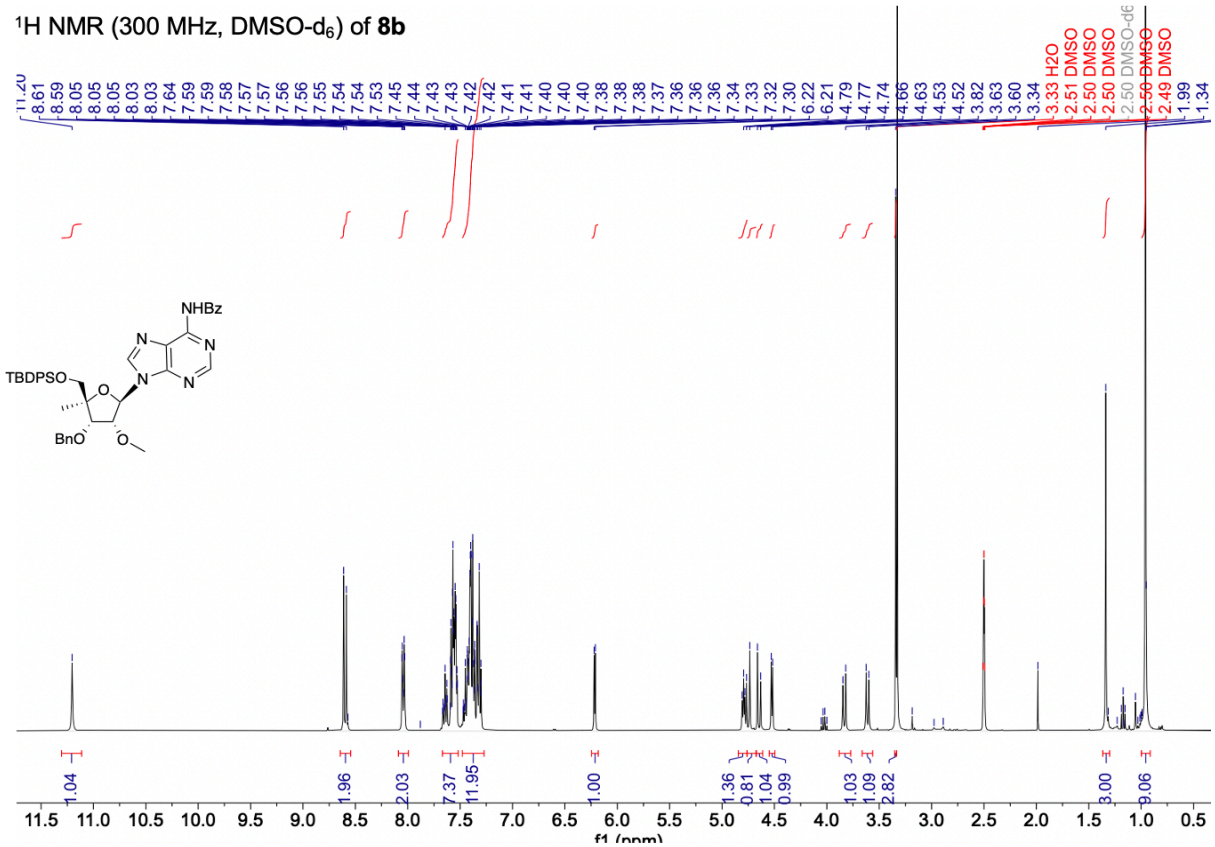


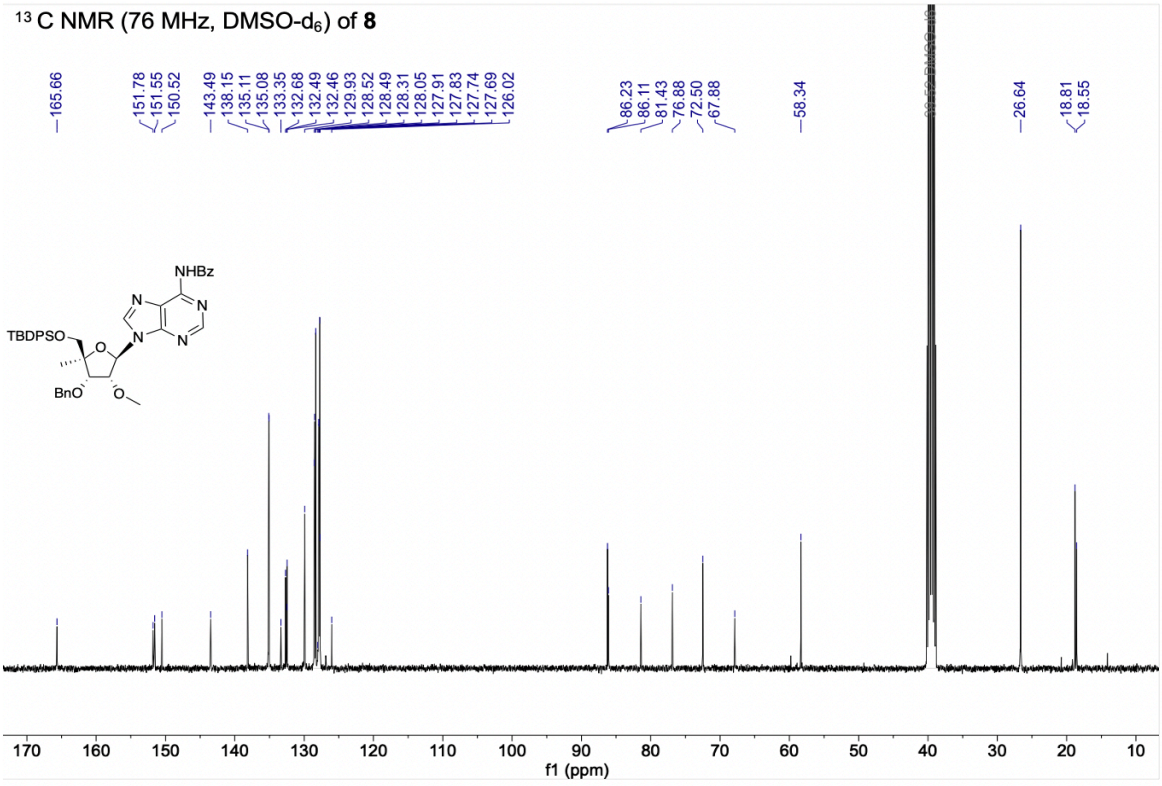
^{13}C NMR (76 MHz, CD_2Cl_2) of **11a**



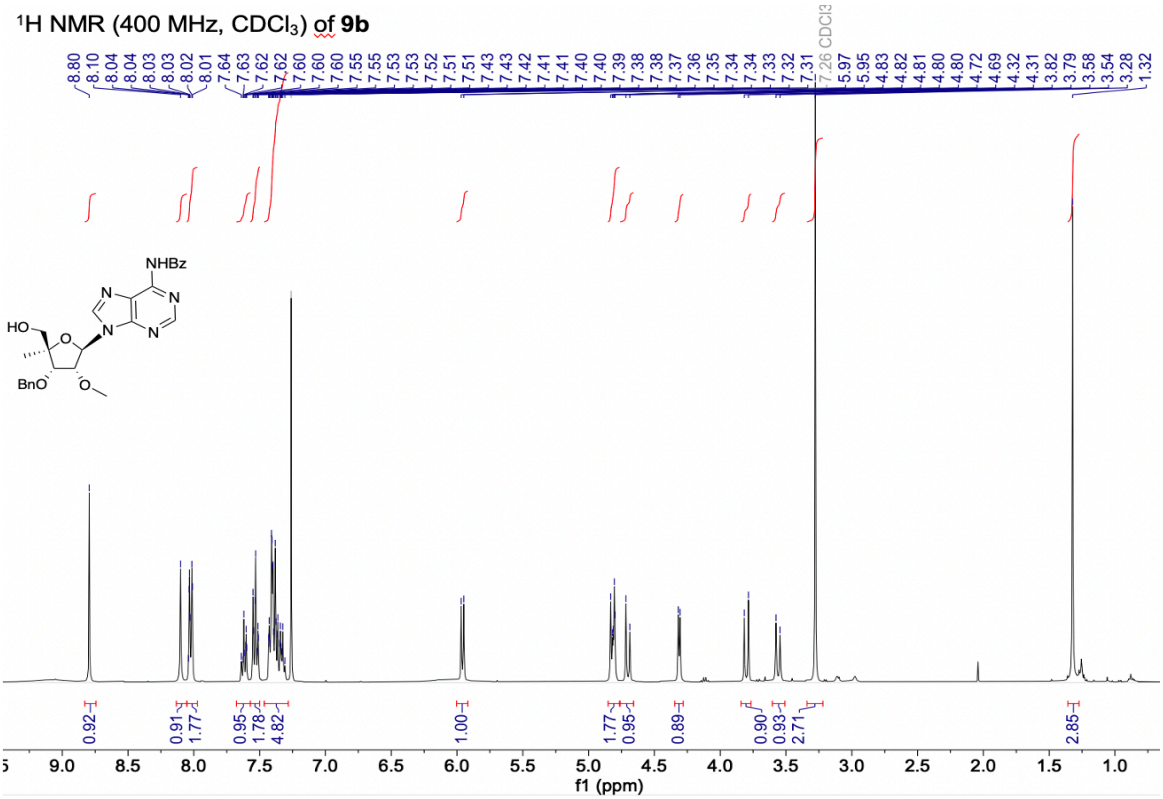
^{31}P NMR (76 MHz, CD_2Cl_2) of **12a**



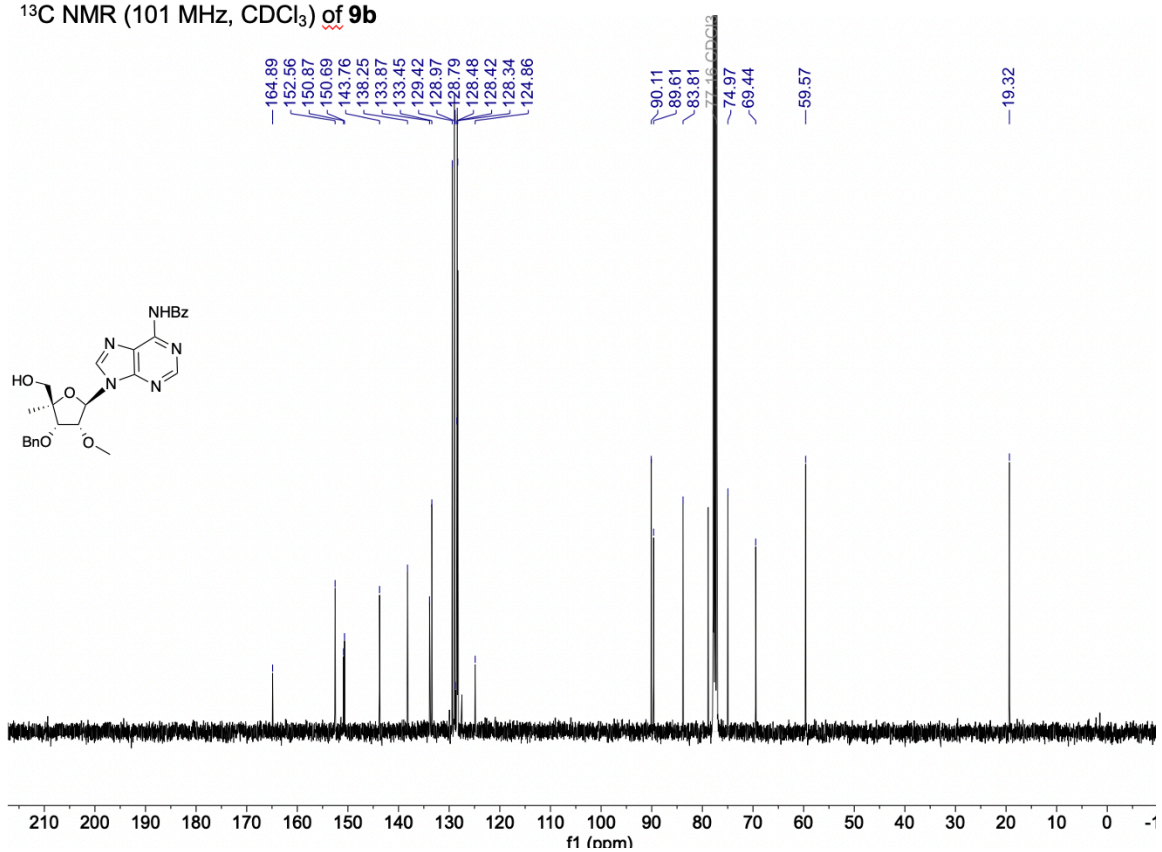

1H NMR (300 MHz, DMSO-d6) of 8b
NHBz
TBDPSO
BnO
O
3.33 H2O
2.51 DMSO
2.50 DMSO
2.50 DMSO
2.50 DMSO
2.50 DMSO
2.49 DMSO
f1 (ppm)


13 C NMR (76 MHz, DMSO-d6) of 8
NHBz
TBDPSO
BnO
O
f1 (ppm)

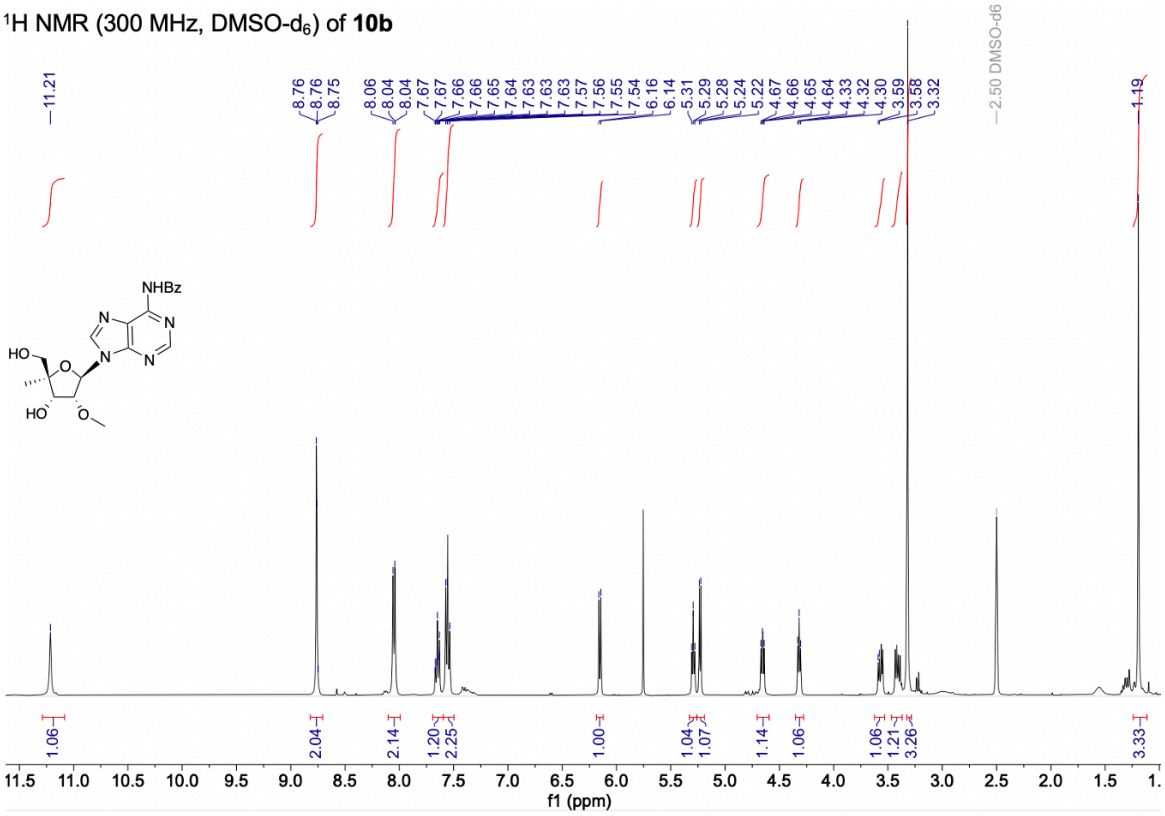
^{1}H NMR (400 MHz, $CDCl_3$) of **9b**



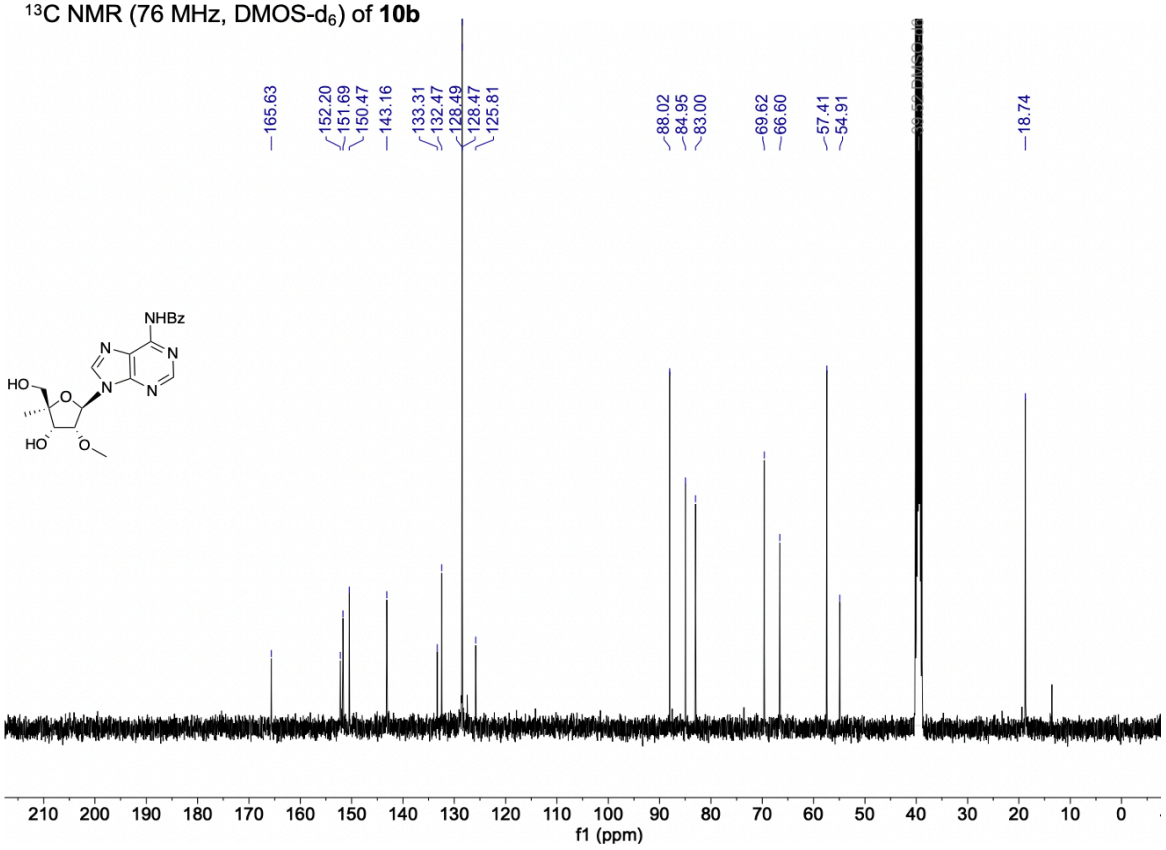
^{13}C NMR (101 MHz, $CDCl_3$) of **9b**



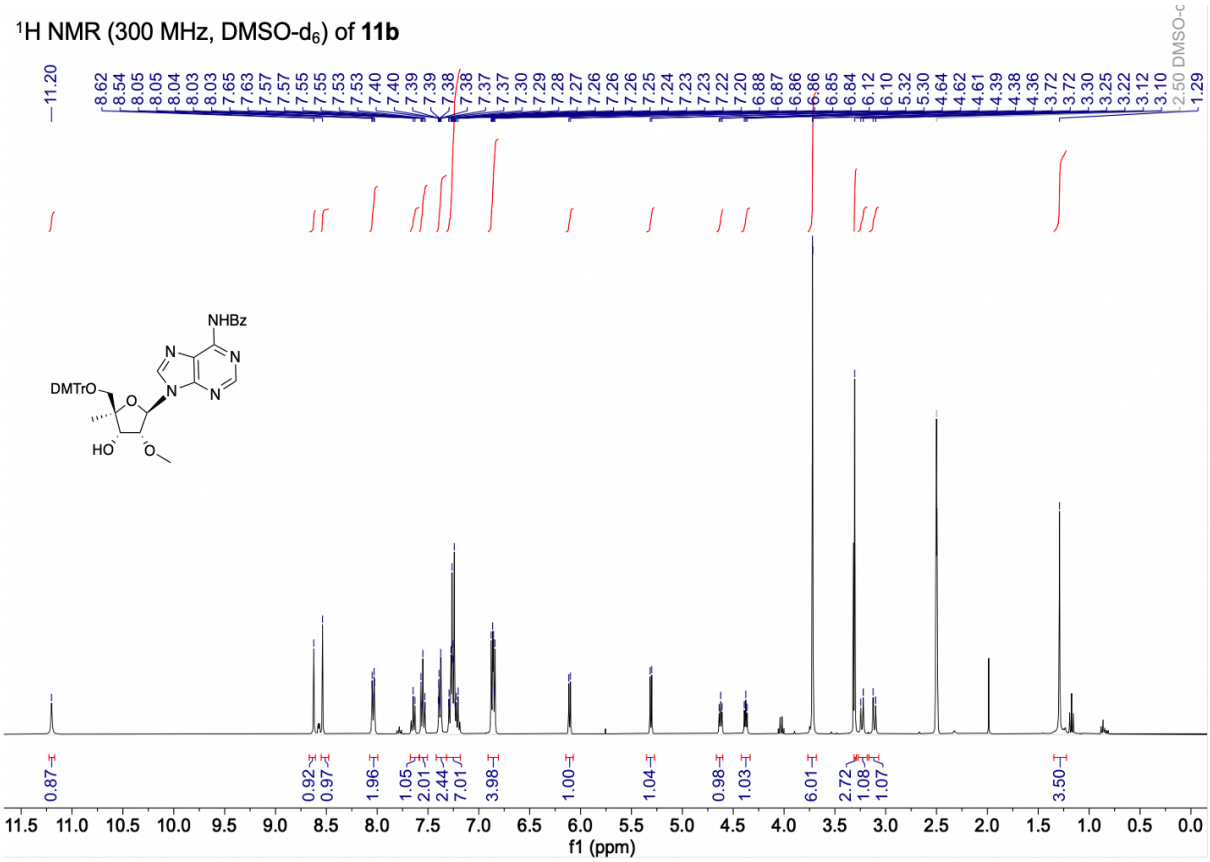
^{1}H NMR (300 MHz, DMSO-d_6) of **10b**



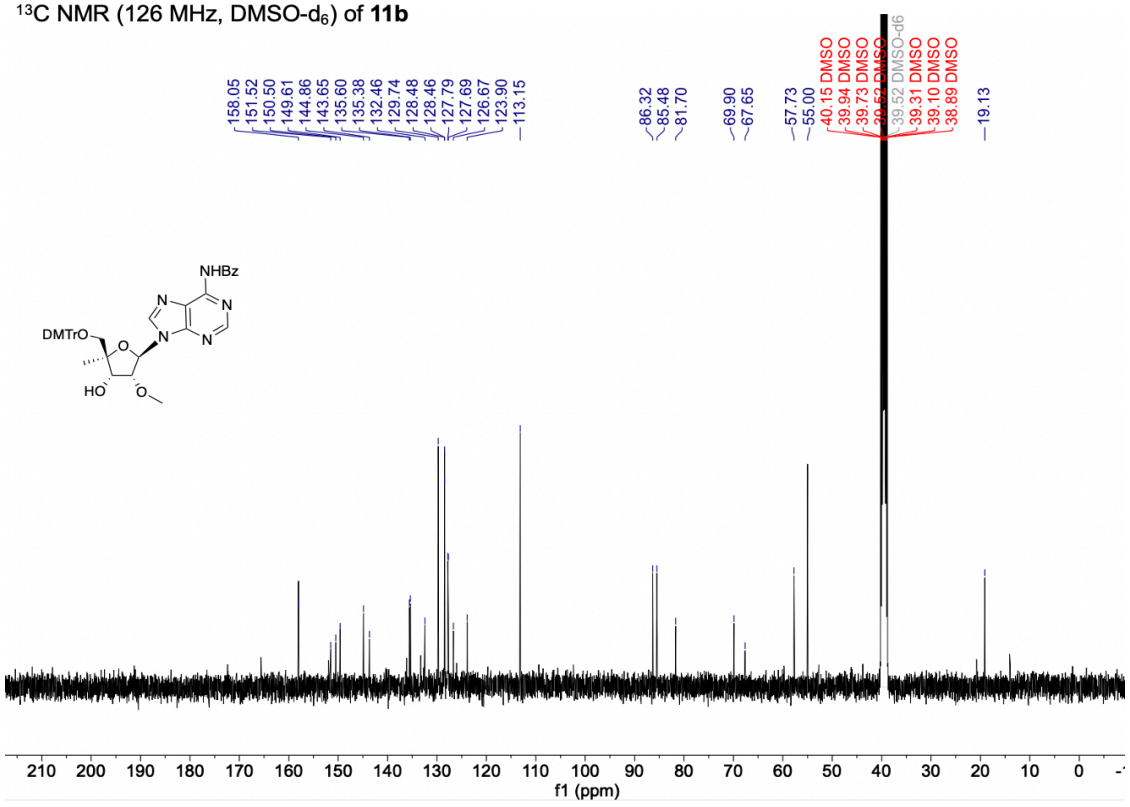
^{13}C NMR (76 MHz, DMOS-d_6) of **10b**



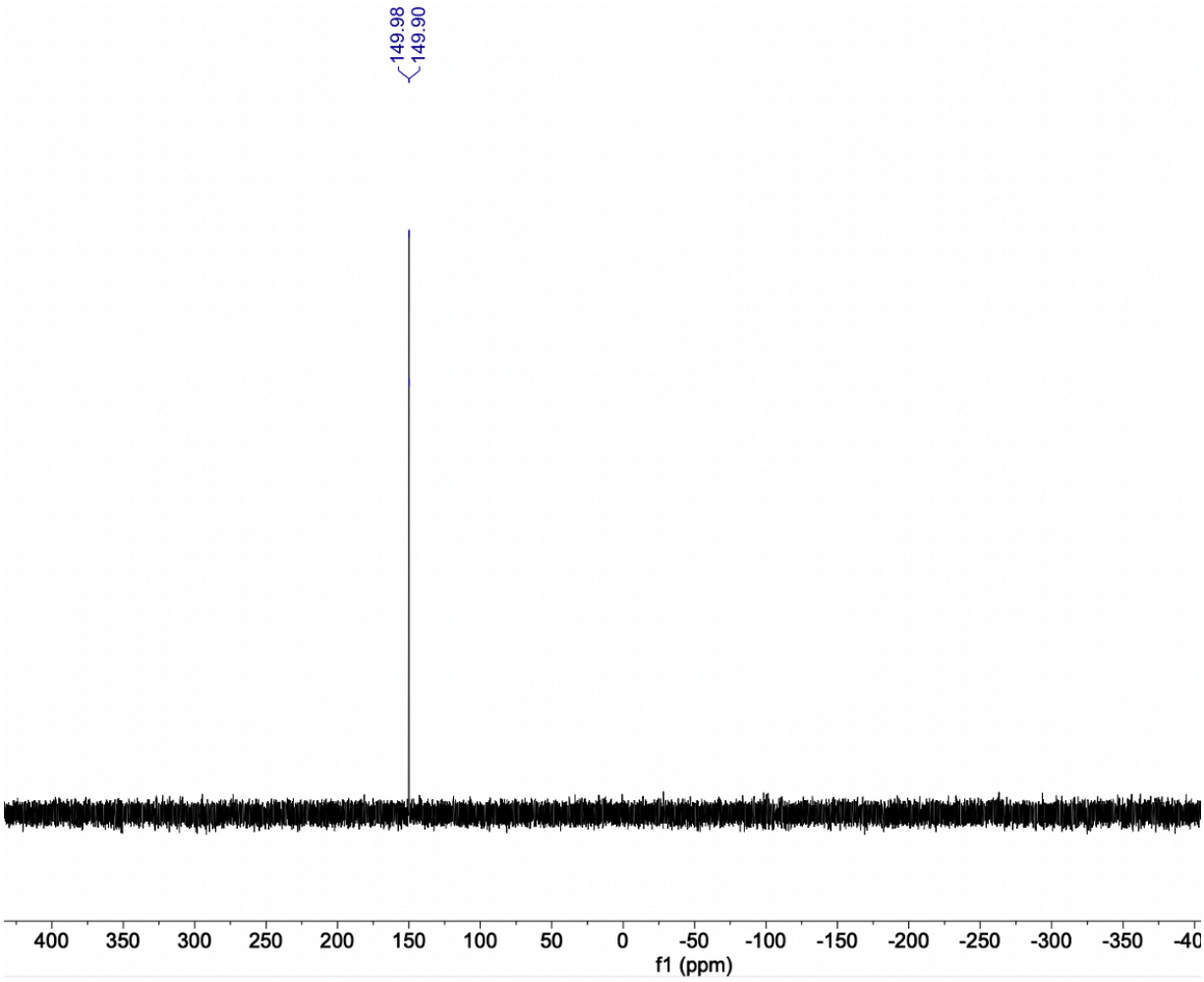
^{1}H NMR (300 MHz, DMSO-d_6) of **11b**



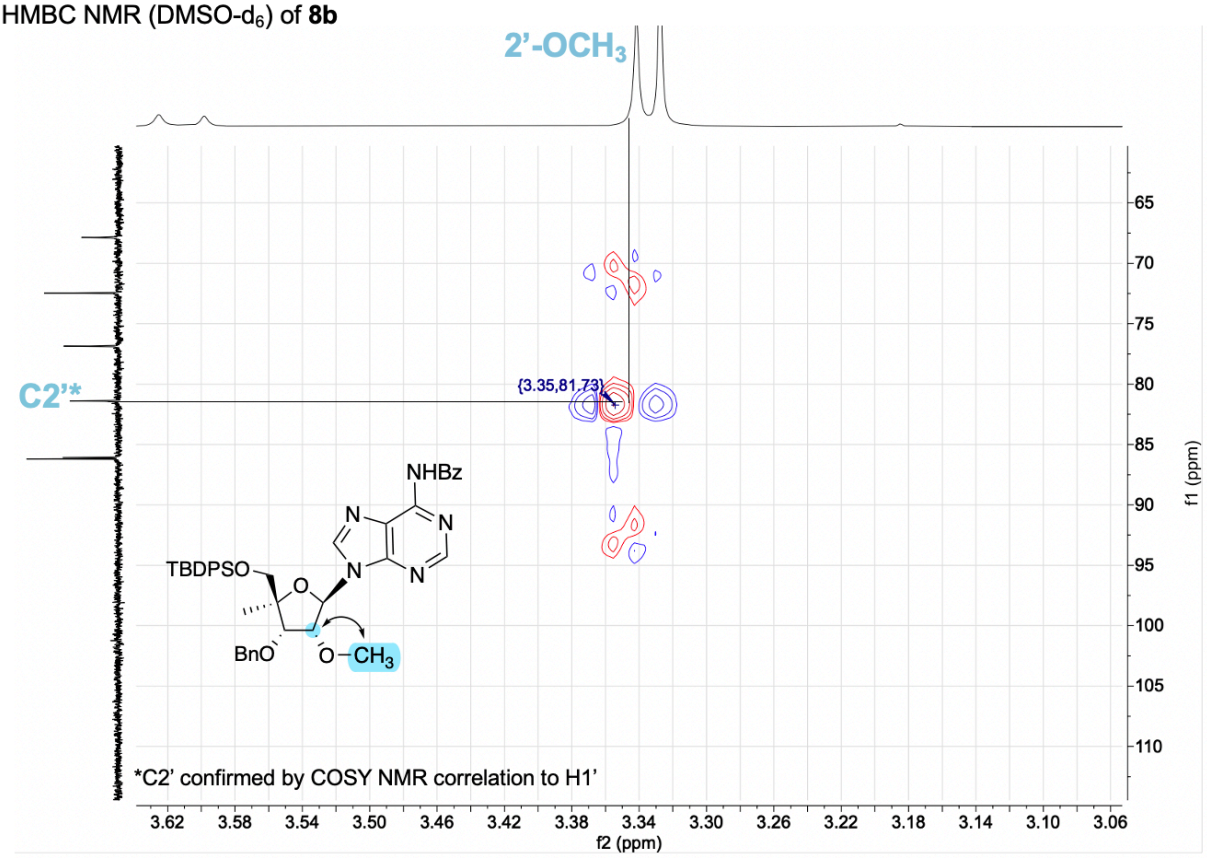
^{13}C NMR (126 MHz, DMSO-d_6) of **11b**



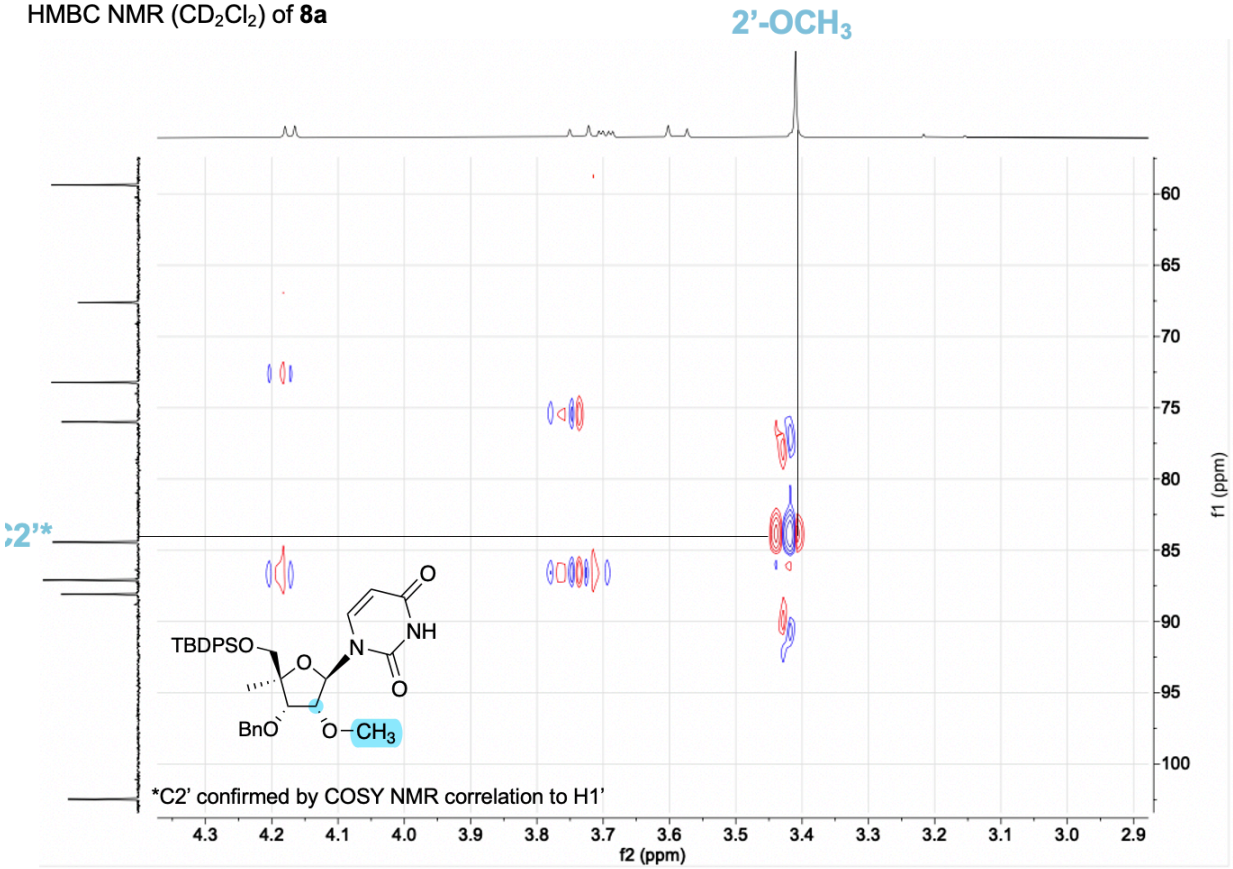
^{31}P NMR (76 MHz, CD_2Cl_2) of **12b**



HMBC NMR (DMSO-d_6) of **8b**



HMBC NMR (CD_2Cl_2) of **8a**



CHAPTER 4

Synthesis and thermodynamic properties of cross-linked RNAs containing 8-azanebularine for ADAR1 Structure Probing

This chapter originated as a collaboration for an ADAR1 project led by Dr. Herra G Mendoza. The scientific findings of this chapter are published in an article in Biochemistry[20] and contain excerpts from the publication. However, it will focus on the synthesis, characterization, and in vitro data of 8-azanebularine containing cross-linked oligonucleotides. All the in vitro data and processing was performed by Dr. Herra G. Mendoza.

INTRODUCTION

In Chapters 2 and 3, the usage of nucleoside analogs to optimize or inhibit the ADAR reaction site-specifically for therapeutic and mechanistic purposes was discussed. Most of the work conducted so far in this dissertation has focused on ADAR2 enzymes, with little work on ADAR1. Previous work in the tumorigenesis field has uncovered that for some cancers, ADAR1 is known to play a role in the development of cancer cells.[17] In particular, loss of function of ADAR1 results in an increase in tumor sensitivity to immunotherapy and in eventual apoptosis for some cancers that are known to have high levels of interferon (IFN) stimulated genes.[19] Understanding the role of ADAR1 in cancer progression can present an advantage to the field of immunotherapy.[18] These findings strongly suggest that inhibiting ADAR1 cell-specifically during tumorigenesis can become an anticancer therapy. Yet, there are still no FDA-approved ADAR1 inhibitors up to date.

One way to regulate ADAR1 editing is by understanding how this enzyme binds to its substrate and catalyzes the conversion of A-to-I, in addition to understanding its reaction

chemistry, and target selectivity. Several crystal structure of ADAR2 alone or bound to dsRNA have been elucidated.[8,29,33,37] Nevertheless, a high-resolution crystal structure of ADAR1 is still lacking, and hence are specific and detailed understanding of how atom per atom this enzyme binds and acts upon duplex RNA. Most of the work done in the Beal and Fisher groups to elucidate the structure of ADAR2 has been attributed mainly to the use of 8-azanebularine (8-AN) and its analogs within RNA molecules as described in Chapter 1.[34,35,80,81] Strategically, RNA molecules with this chemical modification at the active site allows trapping of the ADAR2-RNA complex. Analogous studies with ADAR1 remain unexplored. In this chapter, we designed a set of experiments that made use of ADAR1 tight-binding inhibitors containing 8-AN in the following form: within one strand of RNA in hybridized duplexes, and within a duplex RNA that is covalently cross-linked. In addition, in efforts to understand the binding of ADAR1's RNA inhibitors, we sought to explore novel biophysical properties of RNA duplexes by incorporating internal cross-links via copper click chemistry. More specifically, this chapter will focus on (1) the synthesis, methodology and properties of novel covalently cross-linked RNAs containing 8-AN and (2) the effect of internal cross-links in ADAR1 binding efficiency. All these 8-AN-containing molecules were designed to shed light into ADAR1's catalytic domain substrate recognition, binding, and structure-activity relationship studies (SAR) while also trying to develop strong RNA inhibitors through unique RNA chemistries.

RESULTS

Determination of an RNA inhibitor standard and substrate model for ADAR1

Previous work in the Beal lab has led to the core purpose of this chapter: finding a tight binding RNA inhibitor for ADAR1. For instance, in 2018 Dr. Yuru Wang identified a *Saccharomyces cerevisiae* short RNA (stemming from the HER1 mRNA) hairpin that binds strongly only to the deaminase domain of ADAR1 (hADAR1d) (Figure 1A).[82] This short RNA hairpin was found via the screening of the transcriptome of *S. cerevisiae* that was subjected to ADAR. With the use of RNA-seq, they found common secondary structures of duplexes that favor

hADAR1d binding and editing. Later, this short hairpin RNA was transformed into a two-stranded duplex version (HER1 16 bp) that was developed by Dr. Sehee Park by chemically modifying it with 8-AN (H_{16}) (Figure 1B). Furthermore, from gel shift experiments conducted by Dr. Park and Dr. Wang, a significantly tighter binding was observed with the 8-AN containing duplex compared to the duplex with adenosine at the editing site with a K_D = 21 ± 11 nM for H_{16} 8- AN and K_D > 300 nM for H_{16} A at the conditions tested (Figure 1C-D). These results clearly demonstrated that 8-AN can be used to generate high-affinity RNA ligands for the ADAR1 deaminase domain and enable studies of ADAR1-RNA interactions.

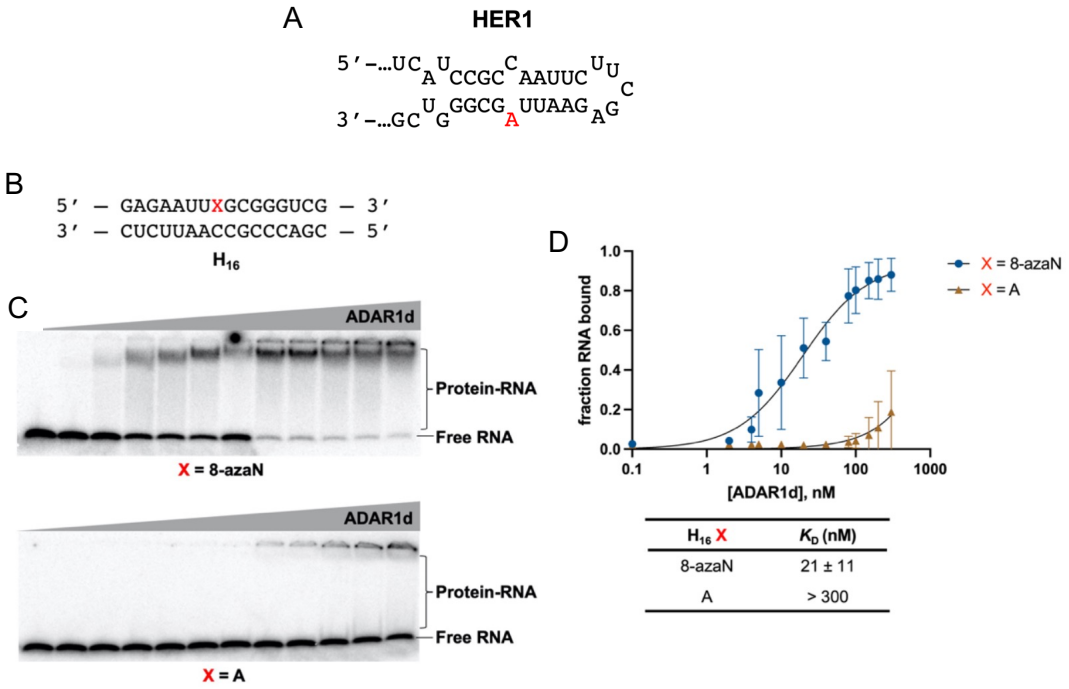


Figure 1. (**A**) Predicted secondary structure around the yeast HER1 editing site. Structure predicted by mFold[83]. Edited A is shown in red. **(B**) Sequence of the 16 bp duplex derived from yeast HER1 RNA. X is A in native target sequence but is varied (8-azaN, A, or G) in the succeeding experiments. (**C**) Gel shifts of ADAR1-D with H16 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 × [x/(K_D + x)] where y is fraction H_{16} RNA bound, x is [hADAR1d]; A is binding endpoint; and K_D is dissociation constant. Error bars represent standard deviation from *n* = 3 technical replicates. Data reference from Wang, et al. Selective recognition of RNA substrates by ADAR deaminase domains **2018**, *57*, https://doi.org/10.1021/acs.biochem.7b01100.

To evaluate the strength of the inhibitors developed in this chapter *in vitro*, we sought to establish transcript models suitable for ADAR1 experiments. So far, most studies with 8-AN

duplexes have been done with the deaminase domain of ADAR1.[82] In this chapter, the ADAR1 p110 isoform was used in the following experiments, considering its biological relevance in tumorigenesis. To do this, Dr. Herra G. Mendoza conducted a set of basal experiments to determine the ideal concentration range for the titration of the RNA inhibitor (H_{16} w/ 8-AN, which is the positive control, while H_{16} w/ G the negative control). The two editing targets chosen for these experiments were the B-site of the pre-mRNA of the 5-hydroxytryptamine receptor 2C (5-HT_{2C}) and site 1 of the Nei-like glycolsylase I (NEIL1) mRNA, both to be known to be edited efficiently by ADAR1 p110 (Figure 2).[84–86] Initially, a concentration range from 0 – 3 μM was tested to determine the half-maximal inhibitory concentration (IC_{50}) for both targets titrating H_{16}. The results show that H_{16} 8-AN inhibited 5-HT_{2C} editing by ADAR1 p110 at the B-site in a concentration- and 8-AN-dependent manner with an estimated IC_{50} of 13 ± 2 nM.[20] Titration of the H_{16} 8-AN duplex also inhibited editing by ADAR1 p110 at the preferred edit site in the human NEIL1 pre-mRNA substrate with an estimated IC_{50} of 8.9 ± 0.8 nM.[20]

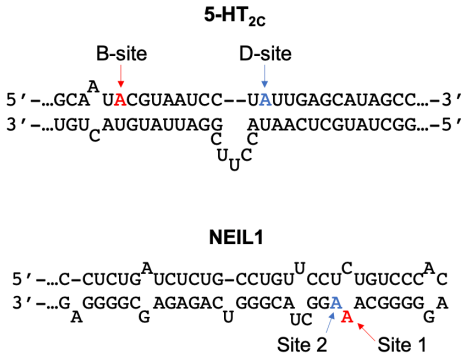
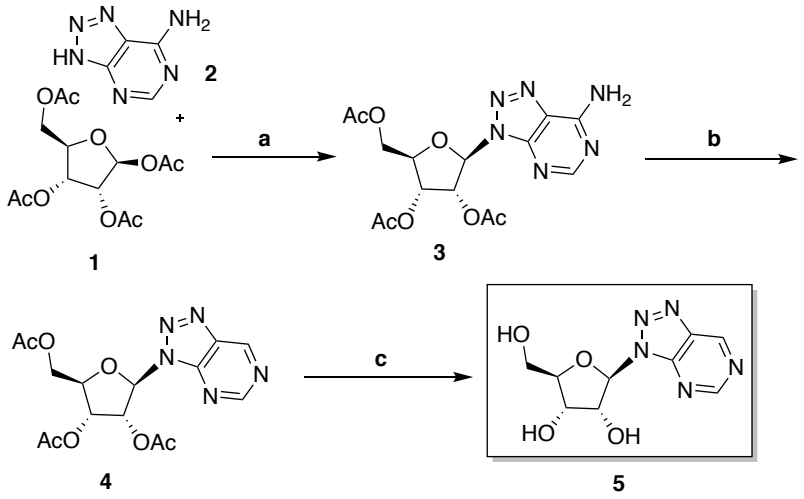


Figure 2. Predicted secondary structures around (top) 5-HT_{2C} B and D and (bottom) NEIL1 Site 1 and 2 editing sites. Structure predicted by mFold.[83]

H_{16} 8-AN inhibition requires secondary structures for ADAR1 p110 engagement

To confirm that inhibition of ADAR1 p110 is dependent on the secondary structure of the H_{16} 8-AN duplex, titration (0 to 1000 μM) of the 8-azanebularine nucleoside was performed in the presence of ADAR1 p110 and the 5-HT_{2C} transcript. To perform this experiment, the modified

nucleoside had to be synthesized as described in Scheme 1, following the strategies of Haudenschild et al.[34] Noteworthy, no inhibition was observed at all at the highest concentration of 8-azanebularine (Figure 3A). Additionally, in a different experiment, the single top RNA strand of the H_{16} was titrated in a similar manner with a concentration range of 0 to 1000 nM to determine if duplex structure dictates the strength of the H_{16} 8-AN duplex inhibitor. Similarly, by Sanger sequencing data, no reduction of the B-site of 5-HT_{2C} was observed, indicating that no inhibition took place (Figure 3B). Both results suggest that 8-AN has to be within a RNA duplex for efficient ADAR1 p110 inhibition.



Scheme 1. Synthesis of 8-azanebularine nucleoside.[34] a) $SnCl_4$, MeCN, 25 ºC; b) *t*-BuONO, TMSBr, MeCN, 0 °C; c) NH_3, MeOH, 0 °C.

In the following section, we will discuss what is the required minimum length of the H_{16} 8-AN inhibitor for efficient inhibition as well as how the thermodynamic stability of the duplex exerts an effect on inhibition. In addition, we will provide structure-activity relationship analyses based on the homology model of the structure of ADAR1 published in 2020.[87] These following experiments were enabled by the synthesis of several modified H_{16} duplexes containing 8-AN and different lengths constructs as well as novel chemistries that internally cross-link two RNA strands for improved hybridization properties.

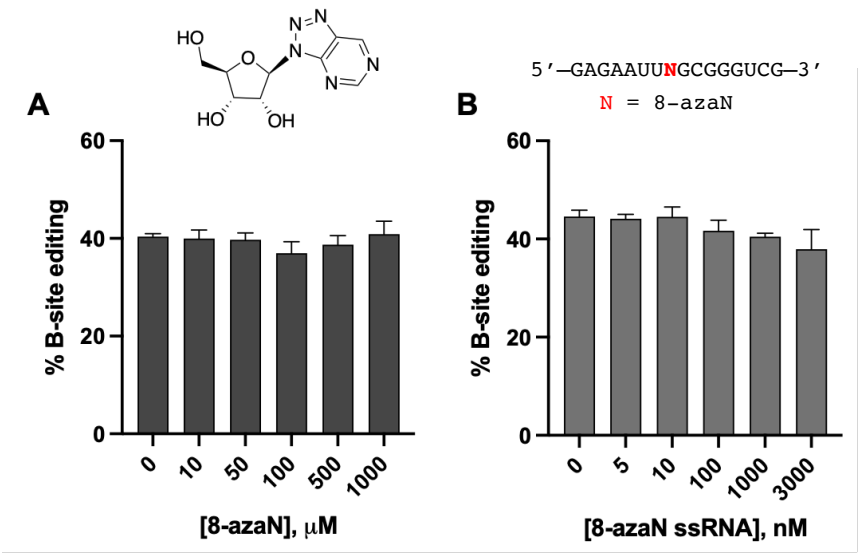


Figure 3. 8-azanebularine free nucleoside and single strand do not inhibit ADAR1 p110. 8-AN as a free nucleoside (**A**) and an 8-AN-modified ssRNA (**B)** do not inhibit ADAR1. *In vitro* deaminations were performed at the following conditions: 0 – 1 mM 8-AN or 0 – 3 μM 8-AN ssRNA, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

The minimum duplex length for ADAR1 binding is 14 bp, with 5 bp 5' and 8 bp 3' to editing site

In our 2023 publication in *Biochemistry*, alumna Dr. Herra G. Mendoza conducted a series of experiments that would elucidate what is the minimum length and form of the H_{16} 8-AN duplex. This was done in order to reduce the amount of nucleotides for the ease of synthesis and if ever this inhibitor was used as a therapeutic, considering that lengthy exogenous RNAs present a hurdle during delivery due to their high anionic charge.[88] From the homology model of ADAR1 generated by Dr. Doherty in collaboration with the Siegel lab, we proposed that the ADAR1 deaminase domain contacts the RNA duplex mostly at positions proximal to the editing site, through both target and complement strands (Figure 4A). However, at distal positions, the ADAR1 catalytic domain mostly interacts with the duplex through the target strand's complement (K1120 and K996 with the phosphodiester linkages of the complementary strand). To validate these contacts, we generated four 8-AN-containing duplexes with different 2 nt overhangs at different 5' and 3' terminals (namely: H_{12} A, H_{12} A overhang, H_{12} B, and H_{12} B overhang) (Figure 5A). 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 (Figure 5B).

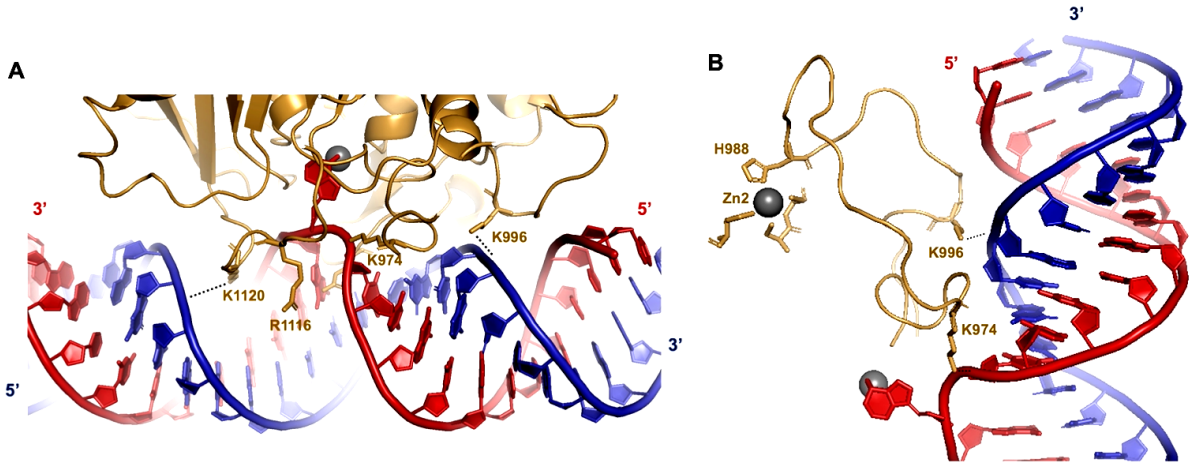


Figure 4. Modeled and predicted contacts between ADAR1 deaminase and RNA duplex. (**A**) Predicted contacts between ADAR1 deaminase (gold) and RNA duplex (red = target strand; blue = complement strand) based on a previously reported Rosetta homology model of ADAR1 catalytic domain.[87] Active site Zn metal (grey sphere). (**B**) Predicted contacts between the 5' binding loop of ADAR1 with RNA duplex. ADAR1's second Zn metal (Zn2, grey sphere) binding site is also shown.

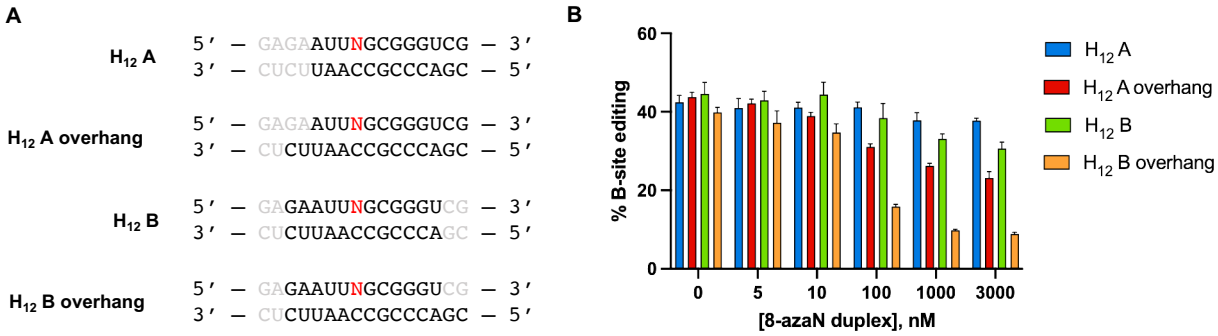


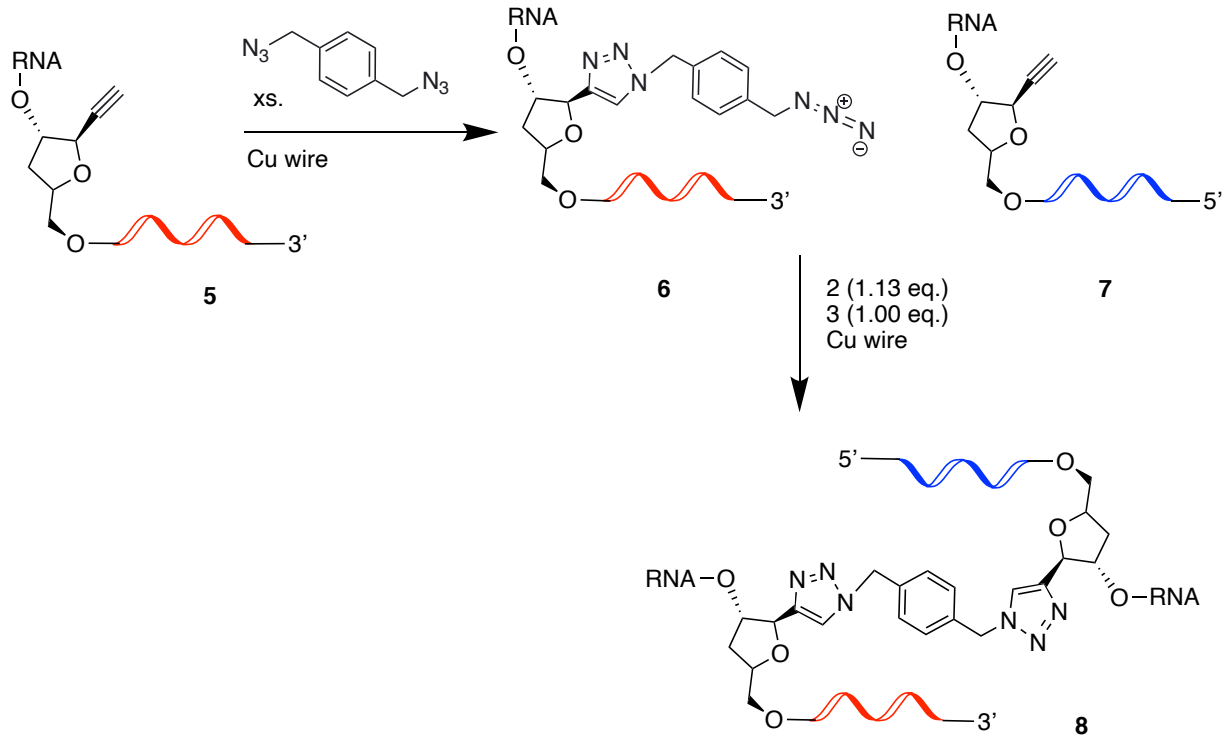
Figure 5. (**A**) 8-AN-modified duplexes with overhangs tested for ADAR1 inhibition. N = 8-AN. Bases omitted from the original H_{16} 8-AN duplex sequence are in grey. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0 – 3 μM 8-AN duplex, 100 nM ADAR1 p110, 5 nM 5-HT_{2C},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

Synthesis of RNA duplexes containing 8-AN with intermolecular cross-links

To understand the effect of the 5' and 3' overhangs seen in Figure 5, we decided to make the duplex extremely stable by inserting internal covalent cross-links. However, some pilot

experiments had to be done to determine which positions can be cross-linked without compromising the inhibitory effect of the HER1 duplex. For the synthetic methodology of internally cross-linked RNAs, we were inspired by other groups that have cross-linked DNA and RNA molecules through different parts of the sugar moiety in both nucleic acids. For instance, Pujari, et al. and Seela's group have developed several methods to cross-link oligonucleotides via 2'-propargylated riboses or alkynylated nucleobases, where a bi-functional azide linker is used to join both strands via CuAAC (copper(I)-catalyzed azide-alkyne cycloaddition).[89,90] On another hand, Gubu et al. developed a method to cross-link one RNA strand via an intramolecular covalent bond (circular RNA) for specific RNase-H-Mediated microRNA inhibition and off-target mitigation.[91] For instance, the authors in the beforementioned publication furnished an intramolecular cross-link utilizing a C8-alkyne-dT analog and a bisazide linker no bigger than 6 Å, again employing CuAAC chemistry.[91,92] However, in this dissertation we followed a different and novel approach of cross-linking the RNAs, in respect to the placement of the internal adduct. To illustrate, we envisioned cross-linking the two strands to mimic a canonical base pair. This was considered due to the fact that for ADAR1 inhibition, a natural RNA duplex structure is required, and we wanted to avoid this by connecting the RNAs via the sugar level, considering that this might change the canonical form of the duplex.[90] In general, we followed a stepwise series of click reactions to avoid covalent homodimerization of RNAs. To yield a mimic of a base pair cross-link, we utilized RNA containing 1'-ethynyl-2'-deoxyribose phosphoramidites, which are commercially available and do not require monomer synthesis efforts. The linker we chose was the 1,4-*bis*(azidomethyl)benzene (bisazide) that was prepared by Gubu and co-workers in one step from 1,4-*bis*(bromomethyl)benzene and sodium azide. Additionally, Randall Ouye, a current lab Beal member, kindly generated a model of the predicted structure of the internal cross-link and its estimated length from each C1' to C1' to be ~13.5 Å. The model shows that the adduct is very similar to that of a canonical base pair in duplex RNA (~11 Å) (Figure 6). The general synthesis of the cross-linked 8-AN-containing RNAs synthesized here is depicted in Scheme 2. Generally,

the first stepwise reaction involved functionalizing one strand of RNA with excess of the bisazide linker (Scheme 2, **6**). After PAGE purification, the mono-azide containing RNA was cross-linked to the complementary strand in a 1 to 1.13 where the 1'-ethynyl-containing RNA **7** would be in slight excess to ensure full product formation **8**.



Scheme 2. General synthesis of the stepwise intermolecular cross-link of two complementary oligonucleotides utilizing 1,4-*bis*(azidomethyl)benzene as a linker and CuAAc chemistry. Blue and red strand are complementary.

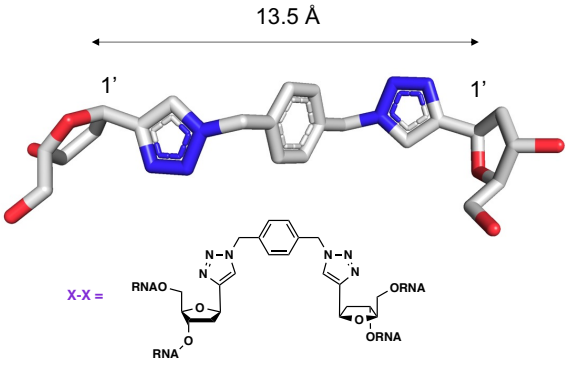


Figure 6. The predicted length of the covalent cross-link from C1' to C1'. Molecular model generated using Gaussian 16. C (grey), N (blue), and O (red). Analysis courtesy of current lab member Randall Ouye.

Cross-link of RNAs bearing 8-AN placed 2 bases 5' of the target site do not inhibit ADAR1

Chapter 2 of this dissertation was dedicated to the synthesis of nucleoside analogs to rationalize ADAR's mechanism in a structure-activity relationship method (in addition to inhibit site-specifically these enzymes). In this section we will discuss the results of the effects of cross-linking HER1 16 bp at different sites in the duplex relative to the target site and their effect inhibiting ADAR1 p110 editing. We performed these control experiments as a proof-of-concept to determine what region within the HER1 duplex is tolerable for covalently cross-linking without impeding ADAR1 p110 inhibition, considering that they will be needed in the following sections. In the following schemes of this section, we will denominate the duplexes based on the placement of the cross-link relative to the target site (Figure 7). The first stepwise synthesis of these oligonucleotides (functionalization of only one strand) can be detailed in Scheme 3, where different bottom strands of the HER1 duplex were synthesized with one 1'-ethynyl-2'-deoxyribose monomer according to the positions where the cross-links would be present. With an excess of the bisazide linker in THF:H_2O, a flat copper wire and the alkynylated bottom strands (VO10A-C), the mono-azide strands were furnished in quantitative yields (Figure 8D). Subsequently, the purified mono-azide strands (VO10AN-CN) were subjected to the complementary RNA strand containing 8-AN in the presence of another copper wire (Figure 9C). Interestingly, each of the resulting cross-linked RNA duplexes displayed a unique migrating pattern on 15% denaturing PAGE (Figure 9D). However, the masses of the oligonucleotides were confirmed by MALDI-TOF (masses in Table 4). Once the cross-linked duplexes were synthesized and fully characterized, we subjected them to the regular *in vitro* ADAR1 p110 inhibition experiments against 5-HT_{2C} pre-mRNA (Figure 10). From these, results, we can strongly infer that placement of the cross-link two bases 5' of the target site is detrimental for efficient ADAR1 p110 binding (Figure 10A, VO11A). Hence, no trapping is taking place, and this can be attributed by the high levels of the B-site editing. Interestingly, if the cross-link is placed four bases upstream of the target site (VO11B), inhibition is restored, shown by the plummeting levels of B-site editing.

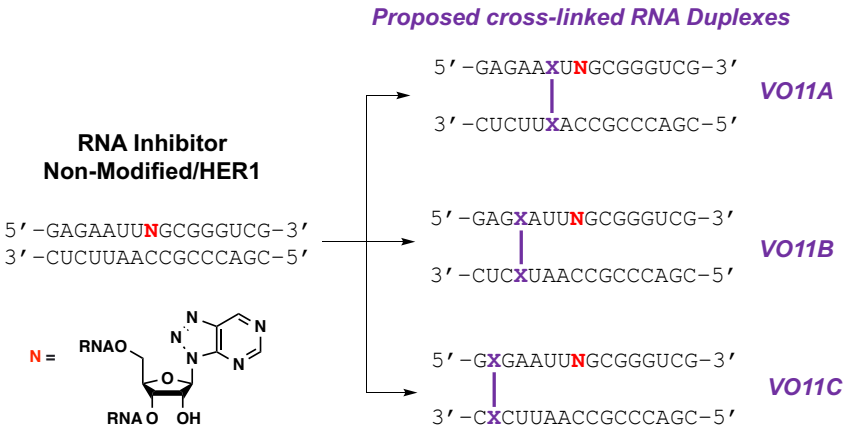


Figure 7. Proposed cross-linked RNA duplexes based on the sequence of HER1.

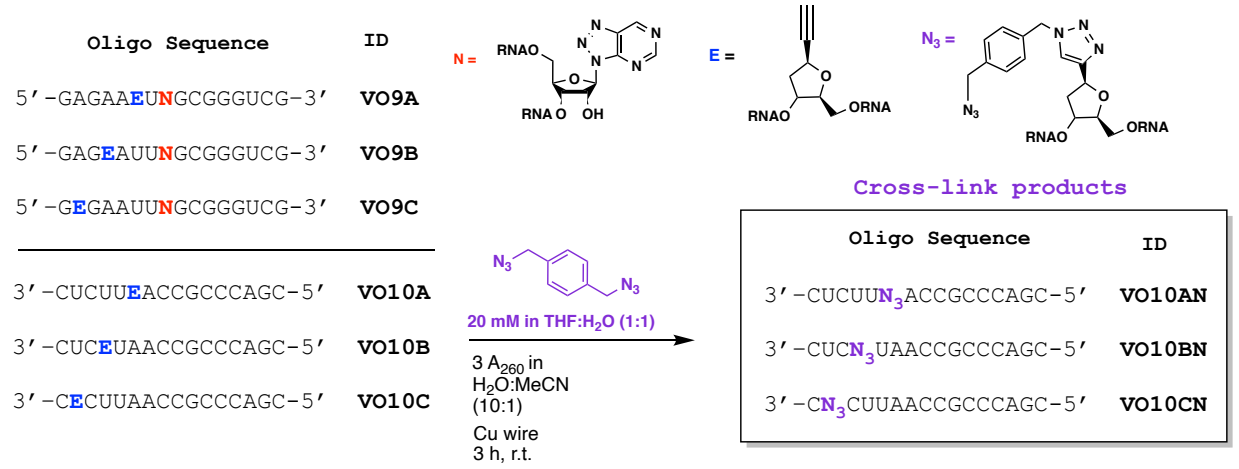


Figure 8. Stepwise synthesis 1 of the bottom strands of HER1 – mono-functionalization. (A) Sequences of the top and bottom strands of HER1 with chemical modifications. (B) Chemical structure of the chemical modifications used for cross-linking (E in blue: 1'-ethynyl-2'-deoxyribose; N in red: 8-azanebularine, N_3 in violet: monofunctionalized azide linker). (C) Reaction conditions for the CuAAC reaction including the structure of the 1,4-*bis*(azidomethyl)benzene in violet. (D) 15% denaturing PAGE gel detailing the slight change in band mobility of reacted with non-reacted RNA strands.

Lastly, when the cross-link is placed 6 bases 5' of the target, the inhibition profile behaves similarily to that of the control HER1 (no cross-link) (Figure 10C). From these editing results, an inhibition-response curve was generated to retrieve the IC_{50} values for the molecules that did inhibit ADAR1 p110, which where VO11B-C.

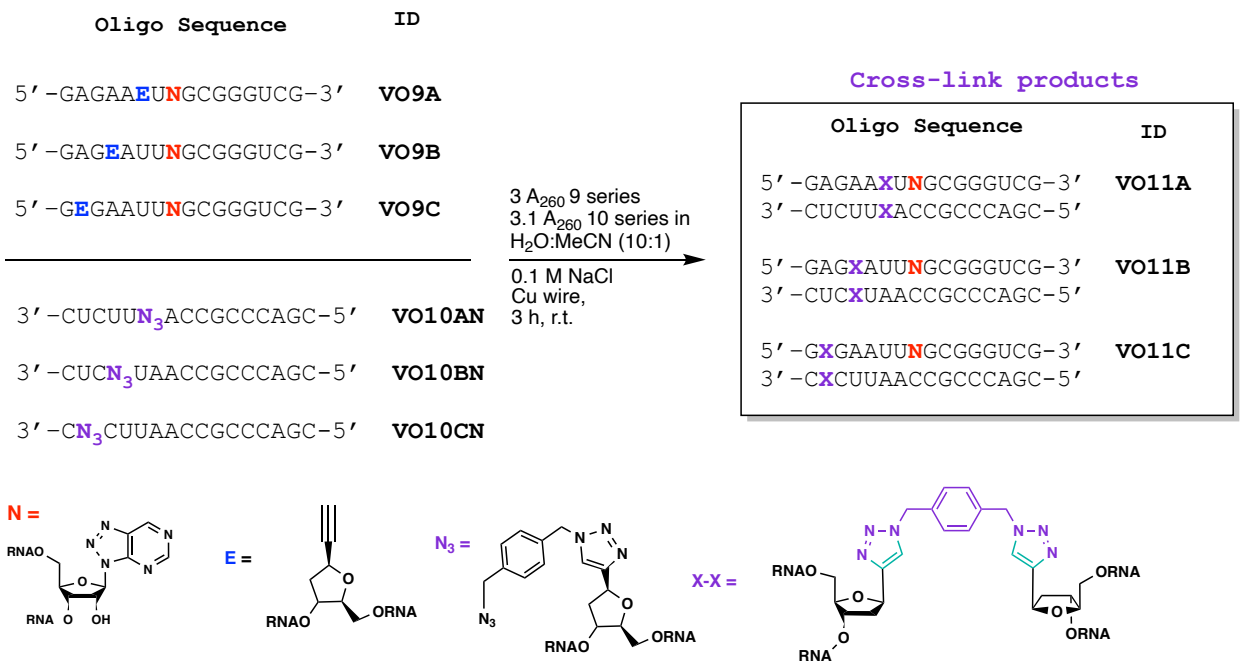
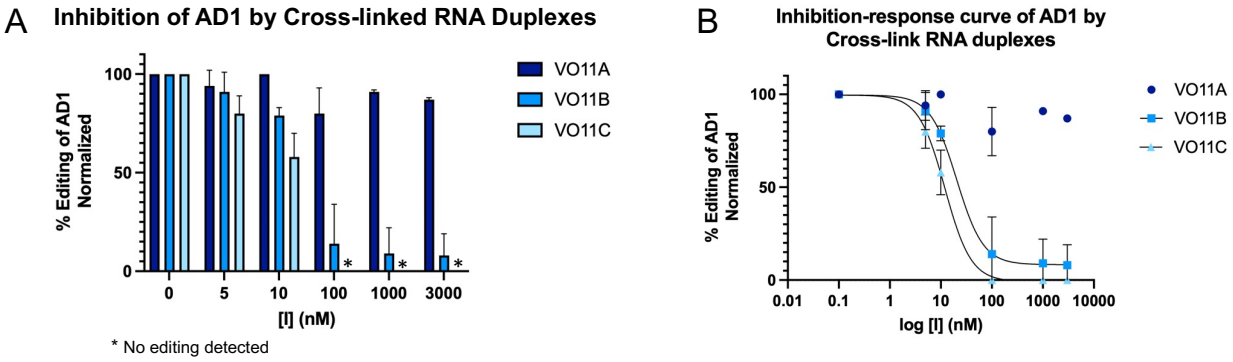


Figure 9. Stepwise synthesis 2 of the top and bottom strands or HER1-cross-linkining the two RNA strands. (A) Sequences of the top and bottom strands of HER1 with chemical modifications. (B) Chemical structure of the chemical modifications used for cross-linking (E in blue: 1'-ethynyl-2'-deoxyribose, N in red: 8-azanebularine, N_3 in violet: monofunctionalized azide linker, X-X in violet: structure of the dual cross-link product from each RNA (C) Reaction conditions for the CuAAC reaction and prodcuts. (D) 15% denaturing PAGE gel detailing the noticeable changes in band mobility of reacted with non-reacted RNA strands.



C

Inhibitor	IC_{50} (nM)
RNA (Non-Modified)	13.2 ± 2.1
VO11A	No Inhibition
VO11B	22.2 ± 4.6
VO11C	12.1 ± 3.4

Figure 10. *In vitro* inhibition results of ADAR1 p110 with different HER1 cross-linked duplexes at different positions against 5-HT_{2C}. (A) % Editing of ADAR1 p110/AD1 normalized, where * is no detected editing. (B) Inhibition-response curve of ADAR1 p110/AD1 by the different cross-linked duplexes. (C) IC_{50} (nM) values from the inhibition-response curves for the modified and non-modified RNA inhibitors. n = 3 replicates for all *in vitro* experiments.

RNA duplex secondary structures are crucial for ADAR1 p110 binding regardless of hybridization strength

Considering our previous results indicating that 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 in Figure 5, we sought to test the effect of stabilization of these duplexes 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. 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-AN-bearing strand (X_{B+}). Specifically, the synthesis scheme of the two covalently cross-linked duplexes with overhangs (X_{A+} and X_{B+}) can be found in Scheme 3. In addition to determining the effect of this cross-link for ADAR1 binding and inhibition, we also sought to explore how the duplex stability/thermodynamic properties were affected due to this internal and covalent cross-link. In effect, cross-linking indeed drastically increased the experimental Tm 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 1).

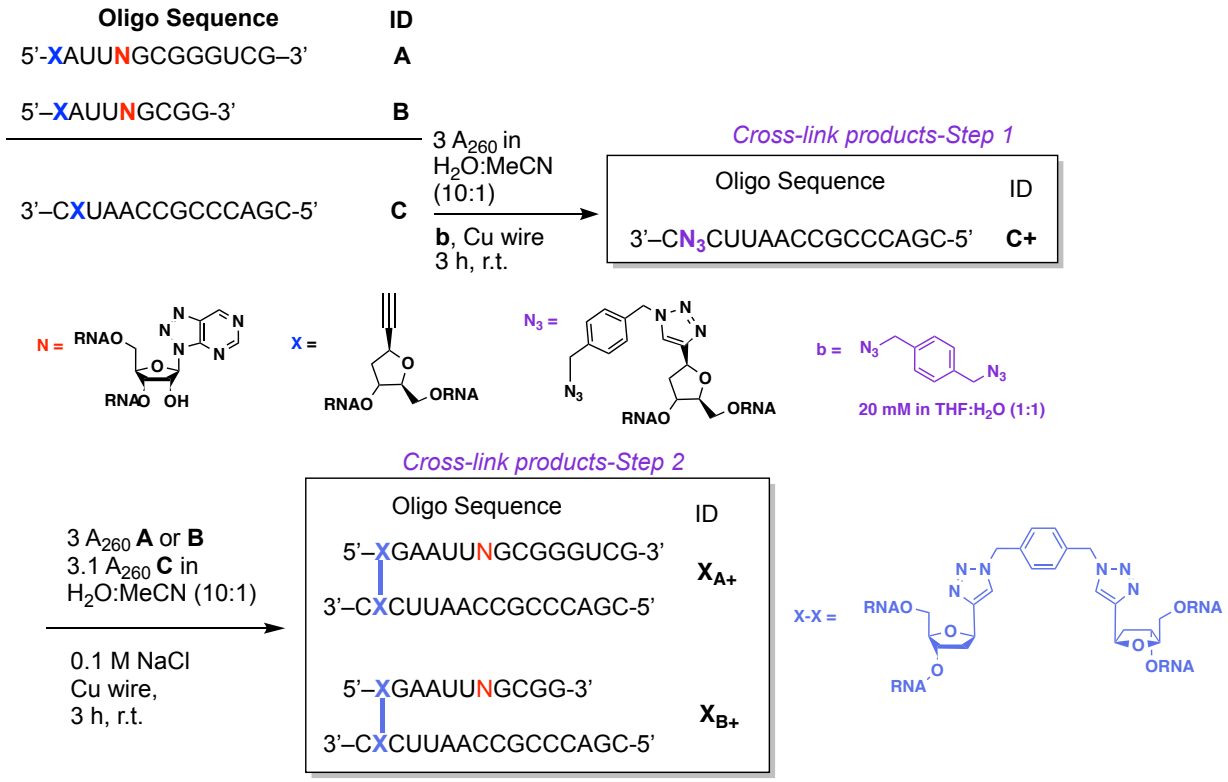
Table 1. Experimental thermal melting temperatures (Tm) for 8-AN duplexes

8-AN Duplex	Tm (°C)[a]
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]Tm 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.

Regarding the inhibition effect in the presence of the cross-links, 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 (Figure 11B)



Scheme 3. Stepwise synthesis and sequences of the two covalently cross-linked forms with overhangs (X_{A+} and X_{B+}).

The minimal increase (X_A to X_{A+}) to no increase (X_B to X_{B+}) in potency that we observed upon cross-linking 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 Figure 5; 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.

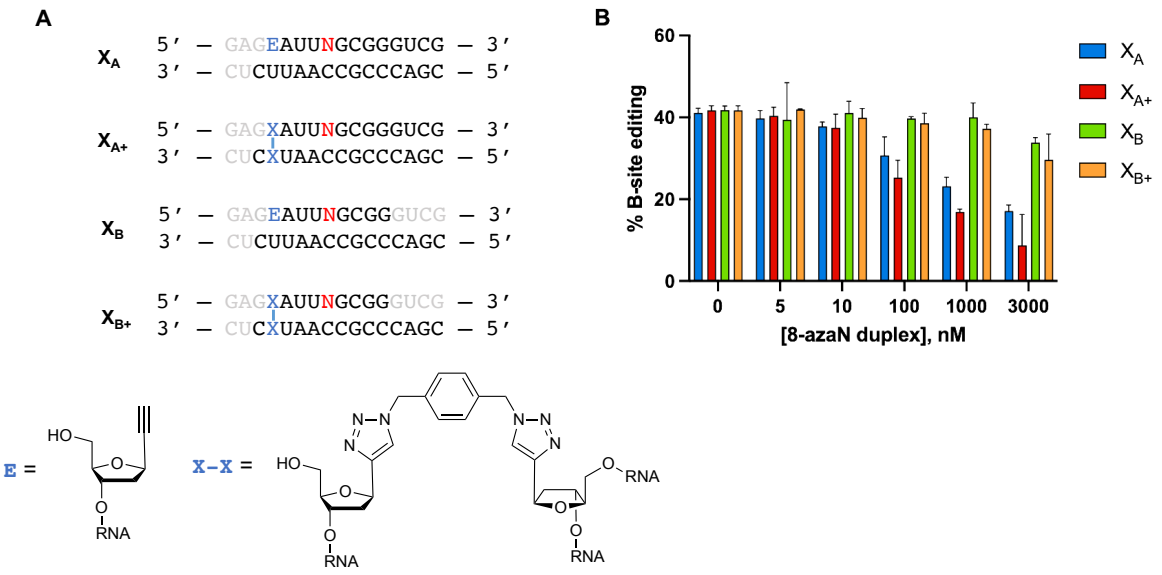


Figure 11. (**A**) Crosslinked RNA duplexes tested for inhibition. N = 8-AN. Bases omitted from the original H_{16} 8-AN duplex sequence are in grey. (**B**) Data from *in vitro* deaminations performed at the following conditions: 0 – 3 μM 8-AN duplex, 100 nM ADAR1 p110, 5 nM 5-HT_{2C}, 15 min, at 30 °C. Error bars represent standard deviation from $n \geq 3$ technical replicates.

DISCUSSION

Designing and synthesizing transition state analogs to trap enzymes bound to their substrates in the active site is an efficient way to elucidate the structure and function of these systems. In the case for adenosine deaminases, the design of adenosine analogs started with adenosine deaminases (ADA).[93] These enzymes catalyze the deamination of the C6 of adenosine nucleosides, whereas ADARs edit adenosine nucleotides strictly in double-stranded RNA.[93] This inspired the Beal lab to synthesize adenosine analogs in the context of an RNA duplex, considering that both enzymes' reaction rates are increased with the use of 8-aza substitution at adenosine. To this end, the Beal lab tried to design inhibitors for ADAR2 with 8-azanebularine as a free ribonucleoside. However, the effort was unsuccessful, considering the weak binding to the deaminase domain and high IC_{50} value of 15 mM.[35] Thus, considering ADAR2s strict substrate preference for RNA duplexes, they tried incorporating 8-aza analogs to a fragment of the mRNA of the glutamate-gated ion channel B-subunit (GluR-B), a substrate known to be edited very

efficiently by ADAR2.[34] Positively, incorporation of 8-AN into the RNA duplex via the phosphoramidite method resulted in a dramatically increased binding affinity of K_D = 2 nM.[34] Yet, most published, and successful studies were focused utilizing the ADAR2 isoform. Consequently, in this chapter, we show that we can perform analogous experiments with 8-AN-modified RNA duplexes with the ADAR1 isoform.

Not surprisingly, it was reported that for ADAR1, the free 8-azaA nucleoside and 8-chloroA nucleoside do not inhibit the enzyme.[94] Here, we sought to incorporate the 8-AN analog into HER1 RNA, a substrate that ADAR1 is known to edit very well. With further synthetic optimization of this 8-AN-modified HER1 duplex, we discovered that for proper and efficient ADAR1 inhibition, the duplex must have a minimum length of 14 bp with 5 bp 5' and 8 bp 3' relative to the editing site. Herein too, we validated with the failed inhibition of a ssRNA containing 8-AN and 8-AN ribonucleoside that an ADAR1 inhibitor must follow a duplex structure for proper binding. Furthermore, we demonstrate that the IC_{50} for ADAR1 inhibition is not highly affected by the melting temperature of HER1 duplexes, whereas by duplex length it is highly influenced. In this chapter we also detail a novel way to covalently cross-link two RNA duplexes via copper-click chemistry. We show a facile way to cross-link RNA inhibitors and demonstrate that cross-links are tolerable for HER1 inhibition starting 4 bp 5' of the editing site. In conclusion, in this chapter we have demonstrated with 8-aza and copper click chemistry that ADAR1 p110 can indeed be inhibited with a new class of oligonucleotide inhibitors.

METHODS

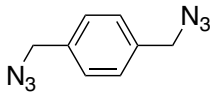
RNA synthesis and purification. The synthesis of transcripts and HER1 related duplexes were done in collaboration with Dr. Herra G. Mendoza, Dr. Park, and Dr. Pham. ADAR editing substrates (5-HT_{2C} and NEIL1) were transcribed from a DNA template using HiScribe T7 RNA synthesis kit (NEB). Unmodified and 2'-O-methylated RNA oligonucleotides for generation of short duplexes ($\leq$ 16 bp) were purchased from IDT. RNA 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.[34] 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, then the pellet was resuspended in anhydrous 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 follow (see Table 2 and 3 for sequences).

Oligonucleotides were purified using denaturing polyacrylamide gel electrophoresis (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 0.1 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 MALDI-TOF mass spectrometry using a Bruker UltraFlextreme MALDI TOF/TOF mass spectrometer at the UC Davis Mass Spectrometry Facility (see Table 4).

Synthesis of cross-linked RNA duplexes. Two ssRNA oligonucleotides each containing 1'-ethynyl-2'-deoxyribose were crosslinked via a bisazide linker, 1,4-bis(azidomethyl)benzene (**1**), by copper(I)-catalyzed azide-alkyne cycloaddition or click chemistry. Bisazide linker **1** was synthesized according to literature.[91] The ssRNA containing just the 1'-ethynyl-2'-deoxyribose modification was first monofunctionalized with an excess of the bisazide linker **1** to afford the first

click product. For this, the ssRNA in a 10:1 nuclease-free H_2O:MeCN mixture (30 μM) was added **1** in a 1:1 nuclease-free H_2O:THF mixture (5.5 mM). After adding a flat copper bar to this mixture, the reaction was stirred in the dark at room temperature for 1 h. Then, the copper bar was removed, and the reaction was allowed to continue at room temperature for 2 h. The oligonucleotides were filtered through a 3K MW cutoff filter, washed twice with nuclease-free water, then dried under vacuum. The monofunctionalized azide-containing ssRNA was then purified from the starting material by denaturing PAGE.

1 

To furnish the final cross-link RNA duplex, the unmodified ssRNA containing 8-azaN in a 10:1 nuclease-free H_2O/MeCN mixture (20 μM) were added NaCl (33 mM) and the complementary azide-monofunctionalized ssRNA (21 μM) . After adding a flat copper bar to this mixture, the reaction was stirred in the dark at room temperature for 1 h. Then, the copper bar was removed, and the reaction was allowed to continue at room temperature for 2 h. The oligonucleotides were passed through a 3K MW cutoff filter, washed twice with nuclease-free water, then dried under vacuum. The cross-linked RNA duplex was then purified from the starting material by denaturing PAGE.

Protein overexpression and purification. The synthesis of proteins was done by Dr. Herra G. Mendoza and Dr. Park. Human ADAR1 p110, ADAR2, and ADAR1d E1008Q with a C- (ADAR1 p110) or N-terminal (ADAR2 and ADAR1d E1008Q) His_{10}-tag were overexpressed in S. cerevisiae BCY123 as previously described.[95] Human ADAR1 p150 with an N-terminal FLAG-tag was purchased from BPS Bioscience. Human ADAR1 p110 was purified by lysing cells in 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 $ZnCl_2$ 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, 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, 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) Nonidet P-40 (NP-40). Protein concentration was determined by running the sample alongside bovine serum albumin (BSA) standards in a sodium dodecyl sulfate-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 β-mercaptoethanol (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 ADAR1d 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).

Preparation of RNA duplex substrates for gel shift assay. The 16mer 8-azaN or A containing RNA strand was end-labeled with ^{32}P using T4 polynucleotide kinase (NEB). Excess [□-^{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 1X TE buffer pH 7.5 and 100 mM NaCl by heating at 95 °C for 5 min and slow cooling to ≤ 30 °C.

Gel shift assay. 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 dithiothreitol (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/bis-acrylamide gel and electrophoresed under non-denaturing conditions in 1X TBE at 4 °C for 1.5 h. The gels were dried under reduced pressure at 80 °C for 1.5 h, 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 by ImageQuant (GE Healthcare). Data were plotted and analyzed using Microsoft Excel and GraphPad Prism.

Hybridization of duplexes. ADAR editing substrates and cross-linked RNA duplexes were allowed to self-anneal in 1X TE buffer pH 7.5 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-AN-containing or A/G-containing strand) to complement strand in the same buffer and conditions as above.

In vitro deamination assay. 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 is 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). Data were plotted 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.

Thermal denaturation studies. Equimolar amounts (1 μM) of intermolecular duplex strand components or 1 μM solution of crosslinked duplex were prepared in 1X TE buffer pH 7.5 and 100 mM NaCl. Samples (19 μL) were placed in a 96-well plate and mixed with 1 μL 20X EvaGreen (Biotium). The plate was then sealed 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, 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). *Tm* was determined as the temperature where -dF/dT is the largest. Data were plotted and analyzed using Microsoft Excel and GraphPad Prism.

Table 2. Sequences of oligonucleotides used in the study. All sequences are written from 5' to 3'. All oligos are made up of ribonucleotides except the primers (which are made up of deoxyribonucleotides) or when specified. TS = target strand, CS = complement strand, N = 8-azanebularine, underline = 2'-*O*-methylated, E = 1'-ethynyl-2'-deoxyribose.

Oligo Name	Sequence
For generating short duplexes	

H_{16} 8-azaN TS	GAGAAUUNGCGGGUCG
H_{16} A TS	GAGAAUUAGCGGGUCG
H_{16} G TS	GAGAAUUGGCGGGUCG
H_{16} 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

Table 2. Sequences of oligonucleotides used in the study. *Cont'd…*

Editing substrates	
5-HT_{2C}[a]	UGGGUACGAAUUCCCACUUACGUACAAGCUUACCUAGAUAUUU GUGCCCCGUCUGGAUUUCUUUAGAUGUUUUAUUUUCAACAGCG UCCAUCAUGCACCUCUGCGCUAUAUCGCUGGAUCGGUAUGUAG CAAUACGUAAUCCUAUUGAGCAUAGCCGUUUCAAUUCGCGGAC UAAGGCCAUCAUGAAGAUUGCUAUUGUUUGGGCAAUUUCUAUA GGUAAAUAAAACUUUUUGGCCAUAAGAAUUGCAGCGGCUAUGC UCAAUACUUUCGGAUUAUGUACUGUGAACAACGUACAGACGUC GACUGGUAACAUUUGCGUUUGAUCGGGUUCU
NEIL1[b]	CCUGAGCCUGCCCUCUGAUCUCUGCCUGUUCCUCUGUCCCACA GGGGGCAAAGGCUACGGGUCAGAGAGCGGGGAGGAGGACU
RT-PCR and sequencing primers	
5-HT_{2C} Fwd	TGGGTACGAATTCCCACTTACGTACAAGCTT
5-HT_{2C} Rvs	AGAACCCGATCAAACGCAAATGTTAC
NEIL1 Fwd	TAATACGACTCACTATAGGGAACTGACTAAACTAGGTGCCACGTC GTGAAAGTCTGACAACCTGAGCCTGCCCTCTGA
NEIL1 Rvs	AGTCCTCCTCCCCGCTCTCTGAC

[a] red A = B-site, blue A = D-site
[b] red A = Site 1, blue A = Site

Table 3. Calculated and observed masses of oligonucleotides. TS = target strand, CS = complement strand.

Oligonucleotide Name	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
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

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-d and a G:G pair were primarily conducted with Beal lab members Dr. Agya Karki and Dr. Herra Mendoza and were part of a highly collaborative project published in Nucleic Acids Research in October 2022.[37] This chapter contains excerpts from the published manuscript as well as from a manuscript published in Biomolecules in 2024 led by Aashrita Manjunath, which includes my work with 8-azainosine at 5'-GA ADAR sites and the G:dI structure solved by Jeff Cheng and Kristen Campbell. My contribution to these projects were mainly on the synthesis of oligonucleotides or phosphoramidites related to 8-azaN, 3-deazadA or 8-azaI for crystallization or in vitro deaminations.

INTRODUCTION

In Chapters 2 and 3, we witnessed the effects of nucleoside analogs in ADAR sites where the 5'-neighboring nucleotide is U or A (*i.e.* 5'-UA or 5'-AA). Much of this work has been facilitated by the known high-resolution crystal structures of ADAR2-RNA complexes, where structure-activity relationships can be established to predict editing enhancement, optimization or site-specific inhibition. In this chapter, we explore chemical modifications at the -1 position in 5'-GA ADAR editing sites, where, unlike canonical base pairings (e.g. 5'-U:A or 5'-A:U seen in Chapter 2 and 3), we explore how G:purine "mismatches" actually enhance editing at these sites. Previously, 5'-GA sites were predicted to be disfavored by ADAR because the crystal structures of ADAR2 bound to dsRNA 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 proper adenosine deamination (Figure 1A-B).[33] In 2014, Schneider M. F., and co-

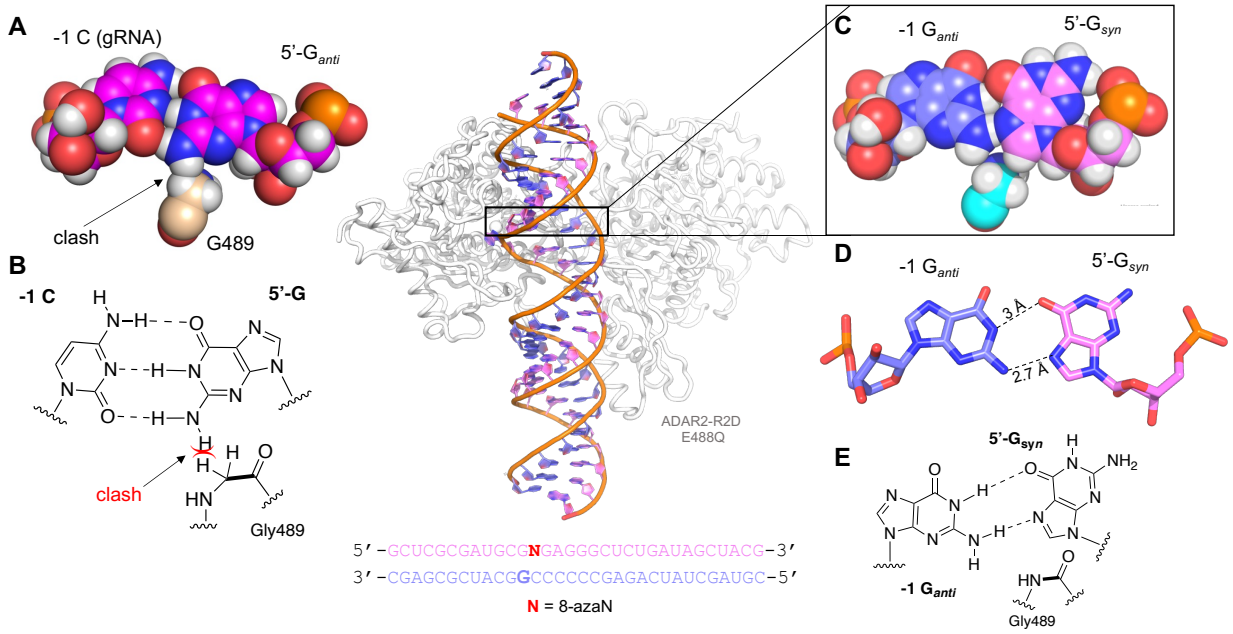


Figure 1. The G*syn*:G*anti* pair accommodates G489 in the minor groove. (**A**) Predicted steric clash space filling model from the amide proton of Gly 489 and the 2-amino group in the 5'-G adjacent to the target A. (**B**) Chemical structure of the predicted clash in A. (**C**) Space filling representation of the G*syn*:G*anti* pair and its location relative to G489 in the ADAR complex. (**D**) Stick model of the G*syn*:G*anti* pair displaying intermolecular H-bond distances in Angstroms. **E**) Chemical structure of G*syn*:G*anti* pair highlighting the location and accommodation of the 2-amino group of the 5'-G relative to Gly 489.

workers reported that editing at 5'-GA sites by fusion proteins bearing ADAR deaminase domains can be improved by pairing the 5'-G with a G or A.[96] These observations led our group to discover the nature of these mismatches. In 2022, Dr. Erin E. Doherty led a simple study, where the -1 position of a 5'-GA therapeutic site was base-paired with all the four canonical bases of RNA (U, A, G and C).[37] In this study, it was 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 pairs. We originally envisioned that a purine:purine mismatch near the editing site could destabilize the duplex and facilitate the base-flipping required for the editing reaction. However, since purine:purine base pairs have also been shown to exist as stable H-bonded pairs in RNA through different combinations of Hoogsteen and Watson-Crick base pairings, 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.[97]

In our 2022 *Nucleic Acids Research* publication and in collaboration with the Fisher lab, we reported the first crystal structure of ADAR2 bound to a dsRNA containing a 5'-G:G mismatch near the editing site to understand the enhancement of G:G pairs at the -1 position of 5'-GA sites. Effectively, in the reported complex structure (ADAR2-R2D E488Q – GLI1 G:G), we saw that the 5'-G:G mismatch adopts a stable G_{syn}:G_{anti} conformation where the 5'-G of the target A displays its Hoogsteen face (G_{syn}) while base pairing to the Watson Crick face of the -1 G in the gRNA in its *anti*-conformation (Figure 1C). This G_{syn}:G_{anti} conformation is induced to presumably have the 5'-G 2-amino group at the target strand avoid the minor groove clash with G489 and enable an efficient deamination reaction (Figure 1E). Considering the success seen with a G:G base pair in ADAR editing of 5'-GA sites, chemically modifying the -1 position with other purine analogs could further enhance and modulate the catalytic efficiency of ADAR editing at these sites. These hypotheses led to the results of this chapter where we inserted several purine analogs at the -1 nucleotide of a 5'-GA site related to Rett Syndrome.[31] These include 3-deaza-2'-deoxyadenosine (3-deazadA), 2'-deoxyinosine (dI) and 8-azainosine (8-azaI) (Figure 2).

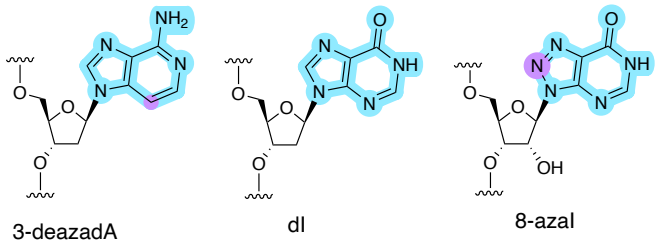


Figure 2. Structure various purine analogs. Structure of 3-deaza-2'-deoxyadenosine (3-deazadA), 2'-deoxyinosine (dI) and 8-azainosine (8-azaI). Violet: highlights chemical modification.

RESULTS and DISCUSSION

Previously, we showed that a G:G base pair drastically enhances the deamination rate of the ADAR reaction in 5'-GA sites. This prompted us to discover what chemical modifications can be made in purines to further enhance RNA editing at 5'-GA sites, considering that several therapeutic RNA sites are in this sequence context.[79] This led to the discovery that adenosine analogs (3-deazadA, inosine) at this position also enhance drastically the ADAR reactions.

Deaza-adenosine analogs enhance the ADAR reaction at 5'-GA sites

3-Deaza-2'-deoxyadenosine (3-deazadA) had faster deamination rates compared to G:G base pairs for both ADAR2 and ADAR1 p110.[37] To characterize the nature of this effect, a 32 bp duplex derived from GLI1 RNA was designed bearing 8-azaN adjacent to a 5'-G:3-deazadA pair for X-ray crystallography with ADAR2-R2D E488Q (Figure 3A). Crystals of this complex formed readily and diffracted X-rays beyond 2.8 Å. The overall structure of the asymmetric homodimeric protein bound to RNA is very similar to the ADAR2-R2D E488Q - GLI G:G structure with well-resolved electron density for the purine:purine pair adjacent to the 8-azaN (Figure 3B). 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-deazadA 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-deazadA N1 to G N7 (2.8 Å) and 3-deazadA 6-NH_2 to G O6 (3.0 Å) (Figure 3C). This interaction causes the base of 3-deazadA to shift slightly towards the minor groove, while also pushing the 5'-G*syn* out towards the major groove. This orientation strongly suggests the 3-deazadA N1 is protonated to donate a hydrogen bond to N7 forming a G*syn*:AH^+ anti pair (Figure 3C). In addition, it is likely that this N1 is protonated at ADAR's reaction conditions considering that the N1 pKa of this nitrogen is 6.8, while for canonical 2'-deoxyadenosine it is 3.62.[98–100] In addition, 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).

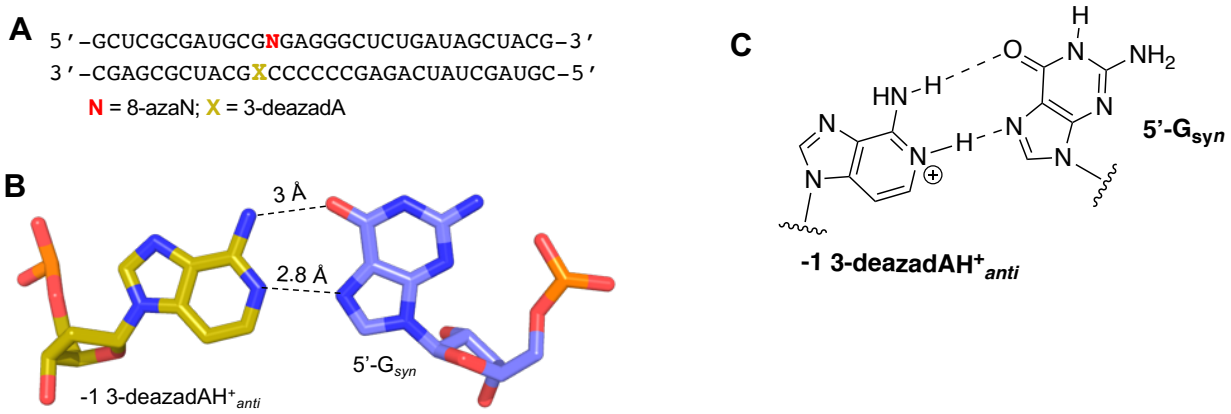


Figure 3. X-ray crystal structure of a complex formed between ADAR2-R2D E488Q and (**A**) a GLI1 32 bp 8-AN containing RNA duplex with G:3-deaza-dA pair adjacent to 8-AN. (**B-C**) Stick model and chemical structure of the and the G*syn*:3-deaza-dAH$^+$*anti* pair.

Inosine analogs enhance the ADAR reaction at 5'-GA sites

Another analog that stood out in further studies led by current lab member Aashrita Manjunath was inosine. To illustrate, we subjected ADAR to edit the mutated transcript of the methyl-CpG-binding protein 2 (MeCP2) R255X. This arginine to stop nonsense mutation results in a truncated MECP2 protein linked to Rett Syndrome.[31] Upon ADAR editing at this site, a new sense codon is created (R255W) that allows for translation of a full-length protein. We subjected this target to ADAR editing by introducing ADAR2 to a duplex substrate containing the R255X sequence (Figure 4A) hybridized to our guide oligonucleotide bearing different analogs at the -1 position. As for ADAR1 p110 experiments, we made hybrids bearing a sequence derived from the mouse IDUA gene modified to contain a 5'- G<u>A</u> target (Figure 4C). To holistically compare the effects of inosine and previous analogs, we tried the following analogs: 3-deazadA, 7-deaza-2'-deoxyadenosine (7-deazadA), G, inosine (rI), 2'-deoxyinosine (dI), and 8-azainosine (8-azaI).

While 3-deazadA, 7-deazadA, G, rI, and dI phosphoramidites were commercially available, the 8-azaI phosphoramidite had to be synthesized.

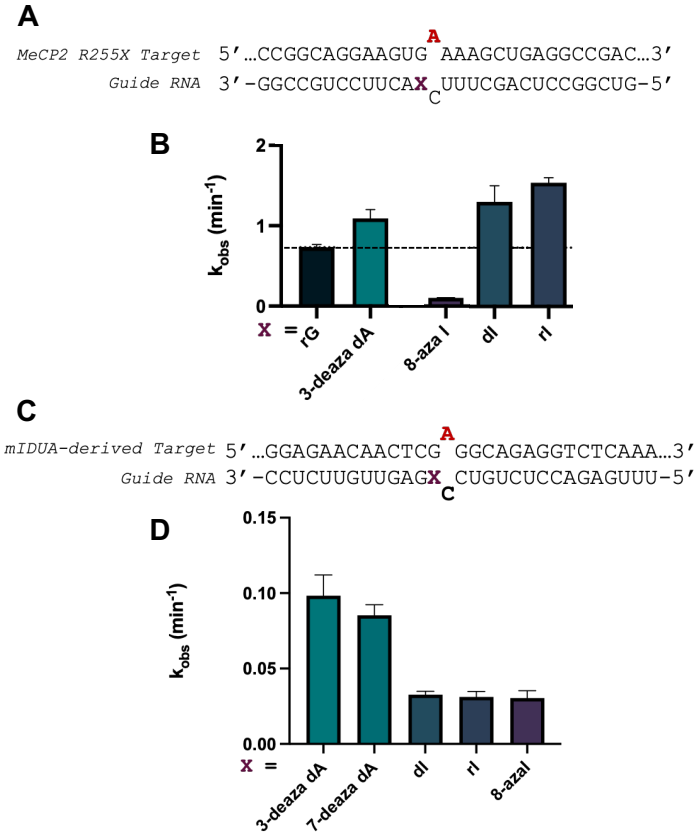


Figure 4. Effect of purine analogs at -1 position on deamination rate at two 5'-GA target sites. (A) Target RNA strand encoding MECP2 R255X mutation and guide oligonucleotide bearing various tested nucleoside analogs at -1 position denoted as X. Target adenosine is in red. (B) Observed rate constants for ADAR2 deamination of targets hybridized to guides containing each respective nucleoside analog. Reactions were carried out with 10 nM hybrid and 100 nM ADAR2 (C) Target RNA strand encoding a modified m*IDUA* mutation containing a 5'-GA site and guide oligonucleotide bearing various tested analogs at -1 position denoted as X. Bold indicates a 2'-deoxynucleotide. Target adenosine is in red. Orphan cytidine is a 2'-deoxyribonucleotide (D) Observed rate constants for ADAR1 p110 deamination of targets hybridized to guides containing each respective analog. Reactions were carried out with 10 nM hybrid and 100 nM ADAR1p110 and in triplicates for both ADAR isoforms.

Originally, Véliz and co-workers developed a procedure to synthesize 8-azaI phosphoramidite by deaminating 8-azaadenosine with adenosine deaminase (ADA) at the nucleoside level (Scheme 1).[35] However, this procedure had to be optimized considering that ADA is no longer sold as a solid, but as a suspension in ammonium sulfate, which complicated the

purification of 8-azaI (see Methods). The effects of the different analogs at the -1 position for ADAR2 and ADAR1 p110 can be found in Figure 4B and 4D, respectively. We found that each of the purine analogs tested increased ADAR2 editing efficiency compared to a canonical G:C pair at this position (Table 1), which renders a substantially slow rate of reaction.[37]

Table 1. Rate constants for ADAR2 *in vitro* deamination of MECP2 R255X substrate with varying -1 gRNA modifications at the -1 position.[a]

-1 modification[a]	k_{obs} (min^{-1})[b]	k_{rel}[c]
dC	0.028 ± 0.003	1
rG	0.73 ± 0.03	26
3-deaza dA	1.1 ± 0.1	39
7-deaza dA	2.4 ± 0.3	85
dI	1.3 ± 0.2	46
rI	1.54 ± 0.06	54
8-aza I	0.103 ± 0.003	3.6

[a]Reactions were carried out with 10 nM of target RNA and 100 nM ADAR2
[b]Reactions were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{obs} \cdot t)]$
[c]k_{rel}= k_{obs} for modification/k_{obs} for deoxycytidine

Table 2. Rate constants for ADAR1 *in vitro* deamination of the sequence modified IDUA substrate with varying -1 gRNA modifications at the -1 position.[a]

-1 modification[a]	k_{obs} (min^{-1})[b]	k_{rel}[c]
dC	0.0072 ± 0.0009	1
3-deaza dA	0.10 ± 0.01	14
7-deaza dA	0.085 ± 0.007	12
dI	0.033 ± 0.002	4.6
rI	0.031 ± 0.004	4.3
8-aza I	0.030 ± 0.005	4.2

[a]Reactions were carried out with 10 nM of target RNA and 100 nM ADAR1 p110
[b]Reactions were fitted to the equation $[P]t = \alpha[1 - \exp(-k_{obs} \cdot t)]$
[c]k_{rel}= k_{obs} for modification/k_{obs} for deoxycytidine

7-deaza-2'-deoxyadenosine (7-deazadA), 2'-deoxyinosine (dI), inosine (rI), enhanced editing beyond a G:G, G:A, and G:3-deazadA pairing as previously reported by Doherty et al.[37] For ADAR1 p110, we observed that the general rate-enhancing patterns seen in ADAR2 were also seen with ADAR1 p110, with the deazaadenosine analogs accelerating editing rate more than the inosine analogs (Table 2). Interestingly, all the inosine analogs exhibited the same deamination rate in ADAR1 p110, suggesting that sugar differences have little effect on the overall rate-enhancing properties of inosine. However, it was noteworthy to see that 8-azaI significantly decreased the rate of reaction for ADAR1. This observation will be rationalized further with the use of a crystal structure detailing ADAR2 bound to a 32 bp duplex with a G:dI pair.

Inosine's ability to enhance the ADAR reaction more than all the analogs discussed until this point for 5'-GA sites prompted us to crystallize an ADAR complex with a G:dI pair to understand the nature of this interaction. Based on the previous structures of G:G and G:3-deazadA, we hypothesized that (1) the 5'-G adjacent to the target site would adopt a *syn* conformation and (2) that inosine would remain in *anti*-conformation providing a bifurcated H-bond between the N1-H of inosine and the N7and the O6 of the 5'-G_{syn}. Considering its likelihood for the formation of well-ordered crystals, we used the same enzyme-RNA construct used for the crystallization trials used in the G:3-deazadA structure ADAR2-R2D E488Q with the GLI1 derived 32 bp duplex with 8-azaN and dI at the -1 position (Figure 5A). The crystals that grew diffracted X-rays to 2.8 Å (Figure 5B). Indeed, we saw that the 5'-G of the target A adopts the *syn* conformation and portrays its Hoogsteen face towards the Watson-Crick face of dI on the opposite side. Evidently, the H-bond network seen between the G:dI pair can be classified as a bifurcated

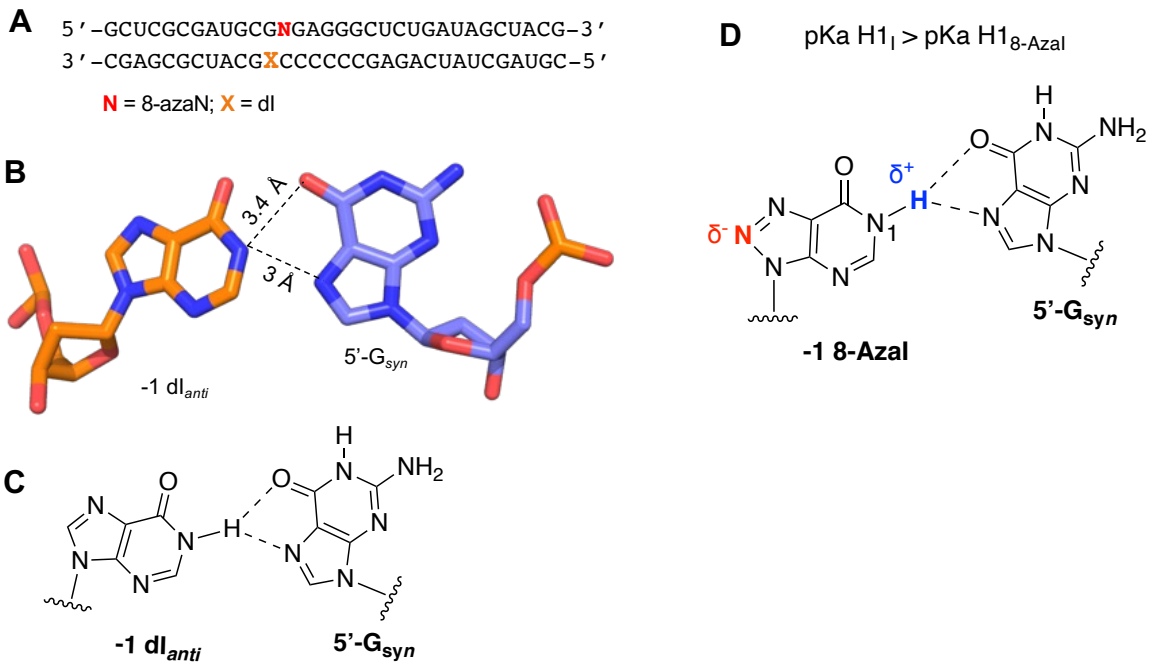


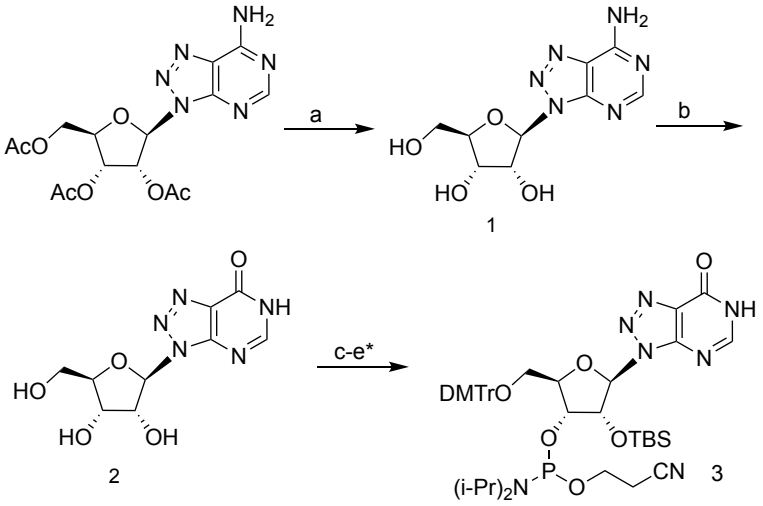
Figure 5. (**A**) A GLI1 32 bp 8-AN containing RNA duplex with G:dI pair adjacent to 8-AN and the chemical structure of the G*syn*:dI*anti* pair. (**B-C**) Stick model and chemical structure of the and the G*syn*:dI*anti* pair. (**D**) Predicted H-bond interaction with 5'-G*syn* and 8-azaI.

H-bond where the N1 proton of inosine is interacting with the electrons of N7 (3.4 Å) and O6 (3.0 Å) of the 5'-G*syn* (Figure 5C). In addition, we hypothesize this too because we also observe that due to the bifurcated hydrogen bond, dI shifts in the direction of the major groove and away from the Gly489 position relative to the position of 3-deazadA. Finally, considering that inosine is very similar to G but has no 2-amino group, the absence of this exocyclic amine clearly enables space between the minor groove of this nucleotide and ADAR's flipping loop region. Previously, we suspected that 8-azaI would act the same or better relative to inosine, considering that the 8-aza electron withdrawing effect might make the N1-H more disposed to make a bifurcated H-bond (Figure 5D). However, this was not the case for ADAR2, where positioning 8-azaI at the -1 position decreased the reaction rate significantly (Figure 4B). This could be rationalized as the 8-azaI H-bond bifurcation being more sensitive to the reaction conditions of ADAR2, where having a lower pKa and a stronger N1 acid might prevent higher populations of the protonation state and hence the H-bond availability.

In this chapter we have outlined how several purine analogs (3-deazadA, dI, rI, 8-azaI, and 7-deazadA) greatly enhance the ADAR reaction at disfavored sites. We present an optimized route for the synthesis of 8-azainosine phosphoramidite and show how purine analogs act significantly better than canonical 5'-GC base pairs adjacent to the editing site, rationalized by steric clashes with ADAR's flipping loop. Three crystals structures of ADAR2-R2D bound to 32 bp 8-azaN modified duplexes with various chemical modifications were shown here detailing how noncanonical purines at the -1 position consistently induce a *syn* confirmation of the 5'-G neighboring base in 5'-G<u>A</u> ADAR sites.[37] This allows atypical hydrogen bond interactions where the 5'-G_{syn} base consistently interacts with its Hoogsteen face, whereas the -1 $base_{anti}$ faces this H-bond network through its Watson-Crick-Franklin façade. In conclusion, the mechanistic elucidations of ADAR editing at disfavored sites alongside the use of chemically modified oligonucleotides add to the toolbox of SDRE techniques in the growing field of ADAR therapeutics, specifically for 5'-G<u>A</u> sites.

METHODS

Synthesis of 8-Azainosine phosphoramidite optimization



Scheme 1. Synthesis of 8-Azainosine from 8-Azaadenosine. (a) 7N NH_3/MeOH, 25 ^oC, 95%; (b) Adenosine deaminase, 200 U/mg, 0.1 M Na_3PO_4, 85%; (c-e)* as previously published: Véliz, E. A., et al, *J. Am. Chem. Soc*. 2003, 125, 10867-10876.[35]

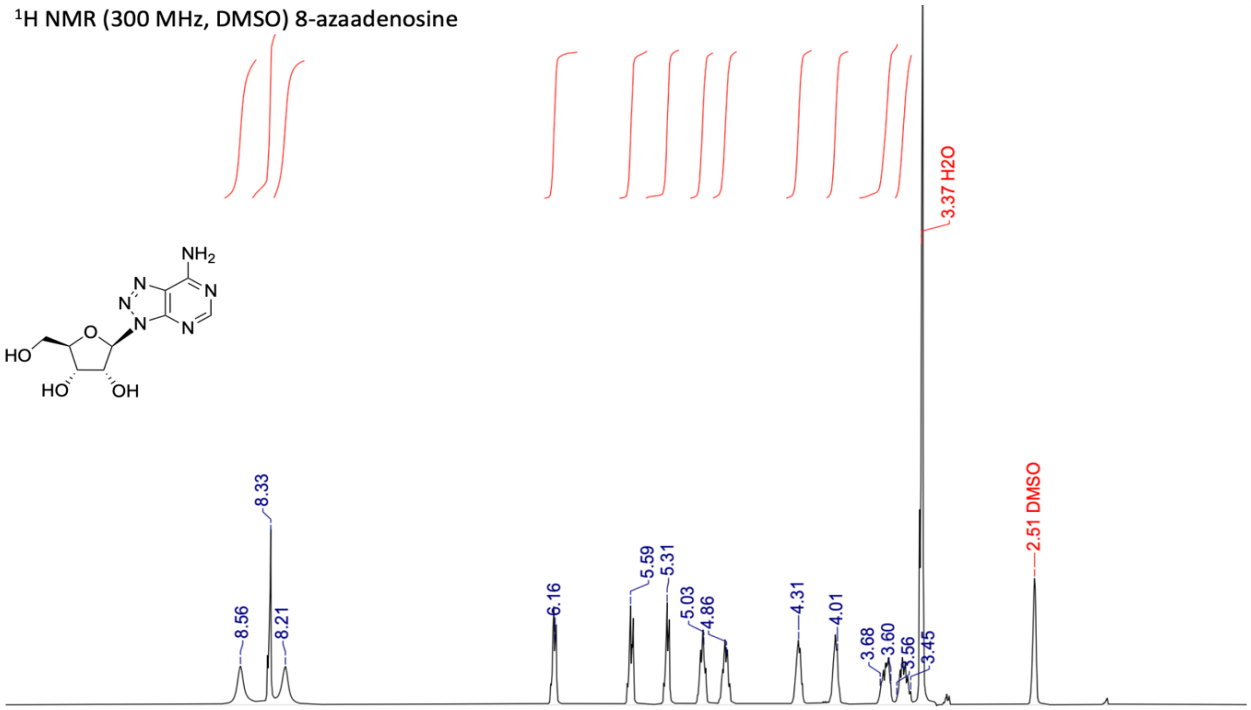
Synthetic procedures. **8-Azaadenosine (2) .** 2',3',5'-Tri-*O*-acetyl-8-azaadenosine (2.3 g, 5.7 mmol) was dissolved in 27 mL of methanolic ammonia 7N. The mixture was sealed and stirred for 16 h at room temperature. The resulting white solid was resuspended in dichloromethane and concentrated under reduced pressure. The resulting solid was then resuspended in *n*-hexanes, filtered, and washed with 30 mL of a mixture of 3:1 Ethyl acetate:*n*-hexanes and 30 mL of 1% MeOH in dichloromethane. The resulting filter cake was dried down under reduced pressure to afford 8-azaadenosine (1.44 g, 95%) as a white solid and used without further purification. ^1H NMR (300 MHz, DMSO) δ 8.55 (s, 1H), 8.32 (s, 1H), 8.21 (s, 1H), 6.16 (d, *J* = 5.2 Hz, 1H), 5.58 (d, *J* = 6.0 Hz, 1H), 5.30 (d, *J* = 5.3 Hz, 1H), 5.03 (dd, *J* = 6.7, 5.2 Hz, 1H), 4.86 (q, *J* = 5.3 Hz, 1H), 4.30 (q, *J* = 4.8 Hz, 1H), 4.01 (p, *J* = 5.0 Hz, 1H), 3.68 – 3.60 (m, 1H), 3.56 – 3.45 (m, 1H).^{13}C NMR (76 MHz, DMSO) δ 157.41, 156.74, 149.33, 124.65, 90.23, 86.75, 73.42, 71.26, 62.42. HRMS calc. [M-H]: 267.08, obsd. [M-H]: 267.0840

8-Azainosine (1). ADA (368 µL ; Sigma, 200 units/mg protein, calf intestinal mucosa from a 2 mL stock) was added to a suspension of 8-azaadenosine (369 mg, 1.38 mmol) in aqueous sodium phosphate buffer (0.1 M, 33 mL; pH 7.4) and stirring was continued at room temperature overnight. The reaction mixture was concentrated under reduced pressure and the residue was purified by flash column chromatography (10-20% MeOH/$CHCl_3$). The chromatographed solid was then washed and filtered with 10 mL of a mixture of H_2O:MeOH:EtOH (0.5:5.5:4) at 0 °C to afford 8-azainsoine (295 mg, 80%) as a white solid. ^1H NMR (400 MHz, DMSO) δ 12.83 (s, 1H), 8.31 (s, 1H), 6.11 (d, *J* = 4.7 Hz, 1H), 5.63 (d, *J* = 5.9 Hz, 1H), 5.31 (d, *J* = 5.5 Hz, 1H), 4.79 (dt, *J* = 17.6, 5.3 Hz, 2H), 4.30 (q, *J* = 4.7 Hz, 1H), 4.03 – 3.99 (m, 1H), 3.61 (dt, *J* = 11.8, 4.8 Hz, 1H), 3.48 (dt, *J* = 11.7, 5.7 Hz, 1H).^{13}C NMR (76 MHz, MeOD:H_2O 3:1) δ 156.59, 149.97, 149.09,

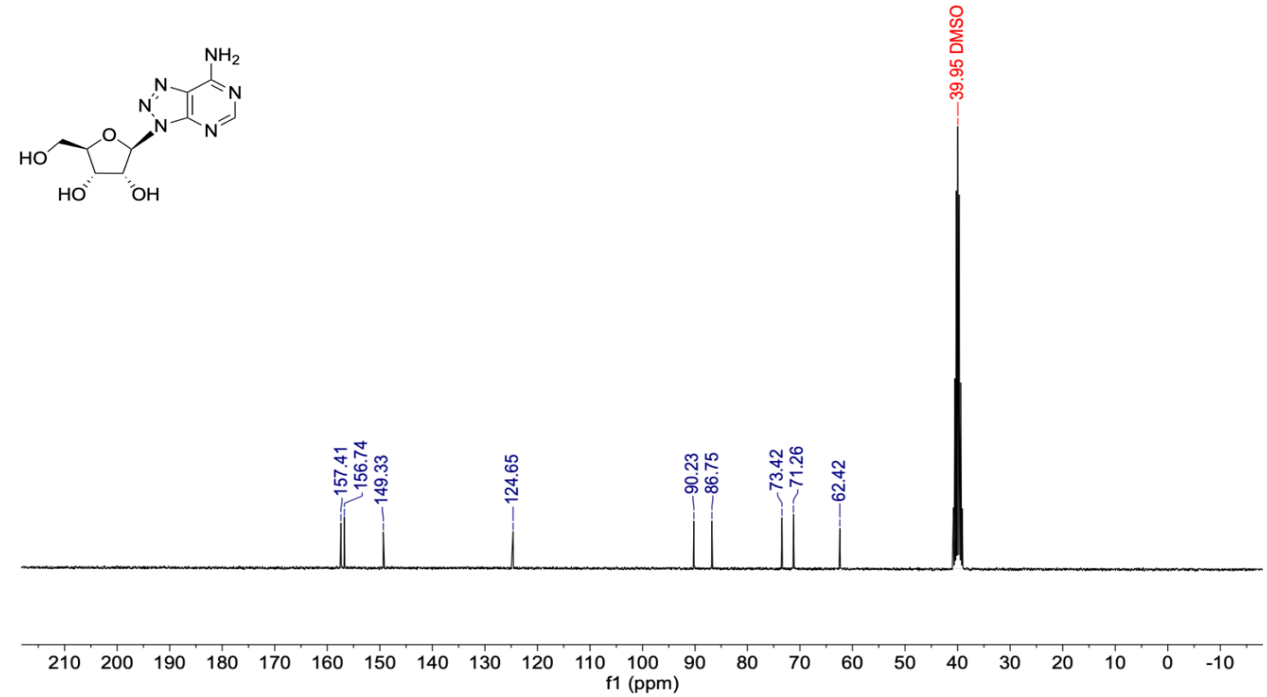
129.94, 90.21, 86.05, 74.13, 70.96, 61.90, 47.93. HRMS calc, [M-H]: 268.07, obsd. [M-H]: 268.0680

NMR Spectra

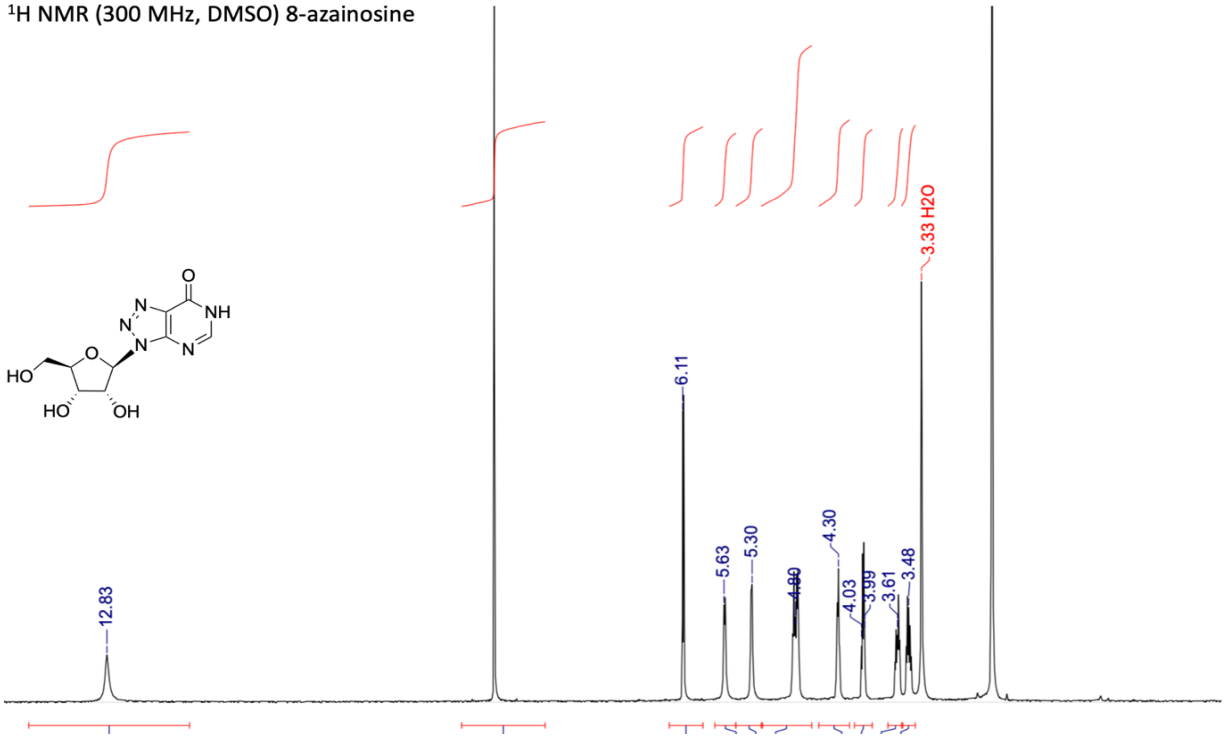
^{1}H NMR (300 MHz, DMSO) 8-azaadenosine



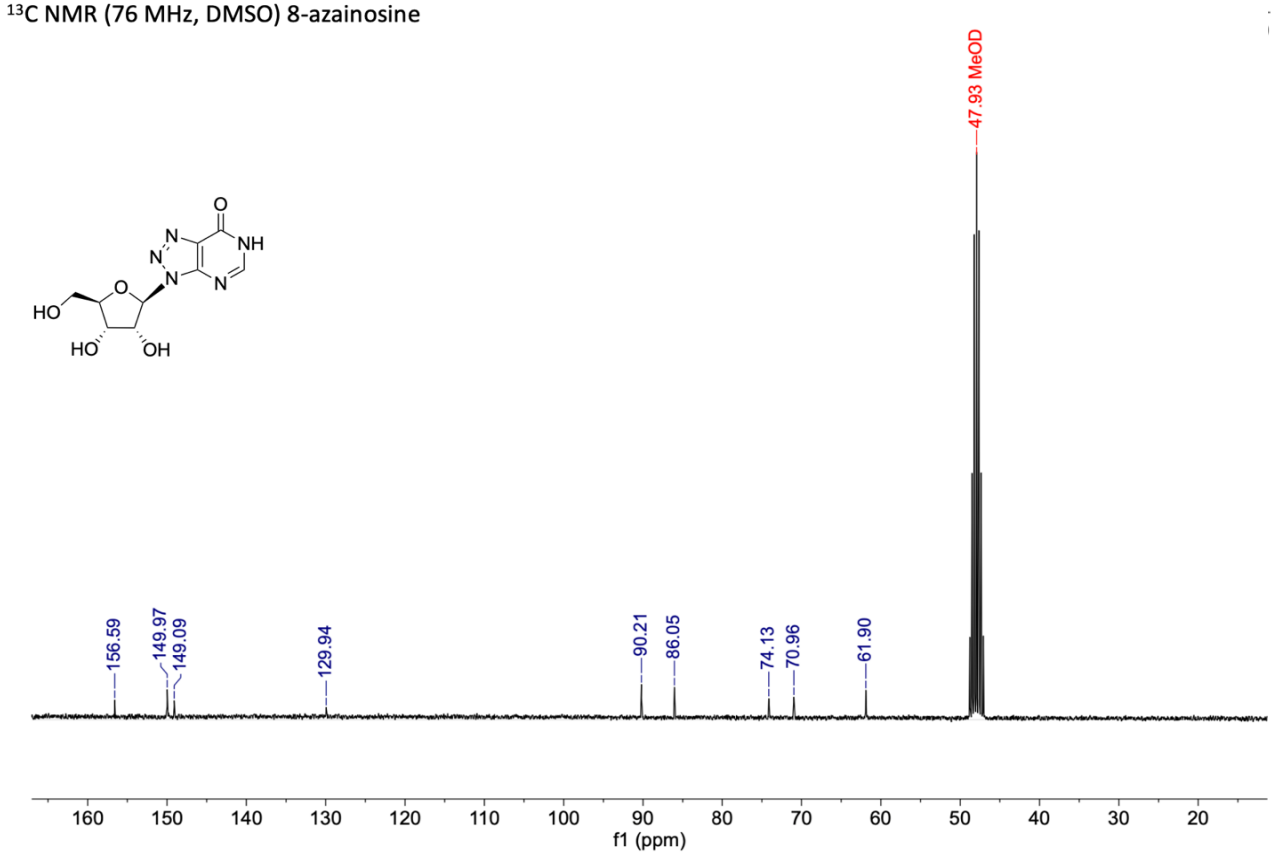
^{13}C NMR (76 MHz, DMSO) 8-azaadenosine



^{1}H NMR (300 MHz, DMSO) 8-azainosine



^{13}C NMR (76 MHz, DMSO) 8-azainosine



Synthesis and purification of oligonucleotides. Chemical synthesis for all oligonucleotides was performed using an ABI 394 synthesizer in collaboration with Agya Karki, 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 as previously described.[34]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 °C 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.[34] To each reaction was added 75 mM sodium acetate in butanol. The oligonucleotides were then precipitated from a solution of 65% butanol at -70 °C for 2 h. The solution was centrifuged at 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. Single-stranded RNA oligonucleotides were purified by denaturing PAGE and visualized by UV shadowing. Bands were excised from the gel, crushed, and soaked overnight at 4 °C in 0.5 M NaOAc, 0.1% (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 °C 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.

Expression, purification, and crystallization of enzymes were carried in collaboration with Dr. Herra Mendoza, Dr. Agya Karki, Dr. Xander Wilcox, Jeff Cheng, Kristen Campbell, Dr. Erin Doherty, and Prof. Andrew Fisher.

Expression and purification of TEV protease. 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. 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 °C 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.

Expression and purification of human ADAR2-R2D for crystallography. Protein expression and purification were carried as previously reported protocol.[78] *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 °C 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 °C. 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 °C 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.

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 °C.

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.

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 and scaled with AIMLESS.[101,102] The structures were determined by molecular replacement using PHENIX. The previous ADAR2-R2D E488Q GLI1-bound crystal structure (PDB 6vff) was used as the phasing model.[8] 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.[103] 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.

Table 3. Sequences for *in vitro* kinetics of the *MECP2* R255X target. All PCR primers are 2'-deoxynucleotides. All guides are ribonucleotides. [X] denotes varied nucleobases including 3-deaza-2'-deoxyadenosine, 7-deaza-2'-depxyadenosine, I, 2'-deoxyinosine, and 8-azainosine.

Human *MECP2* R255X target DNA template sequence (target A indicated in red, T7 promoter highlighted in grey)	CACGATTAATACGACTCACTATAGGGTGTGCAGGTGAAAAGGGTCCTGGA GAAAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAG GGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCATGG TGATCAAACGCCCCGGCAGGAAGTGAAAAGCTGAGGCCGACCCTCAGGC CATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGC TGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCT GTGCAGGAGACCGTACTCCCCATCAAGAA
MECP2 R255X guide strands	5'-GUCGGCCUCAGCUUUC[X]ACUUCCUGCCGG -3'
MECP2 R255X RT-PCR forward and sequencing primer	5'-GGGTGTGCAGGTGAAAAGG-3'
MECP2 R255X RT-PCR reverse primer	5'-TCTTGATGGGGAGTACGGTC-3'

Table 4. Sequences for *in vitro* kinetics of the *IDUA* target. All PCR primers are 2'-deoxynucleotides. All guides are ribonucleotides except for the following: bold indicates 2'-deoxynucleotide. [X] denotes varied nucleobases including 3-deaza-2'-deoxyadenosine, 7-deaza-2'-depxyadenosine, I, 2'-deoxyinosine, and 8-azainosine.

Mouse IDUA target DNA template sequence (target A indicated in red, T7 promoter highlighted in grey)	TAATACGACTCACTATAGGGCTCCTCCCATCCTGTGGGCTGAACAGT ATAACAGACTCCCAGTATACAAATGGTGGGAGCTAGATATTAGGGTA GGAAGCCAGATGCTAGGTATGAGAGAGCCAACAGCCTCAGCCCTCT GCTTGGCTTATAGATGGAGAACAACTCGAGGCAGAGGTCTCAAAGG CTGGGGCTGTGTTGGACAGCAATCATACAGTGGGTGTCCTGGCCAG

	CACCCATCACCCTGAAGGCTCCGCAGCGGCCTGGAGTACCACAGTC CTCATCTACACTAGTGATGACACCCACGCACACCCCGGATCC
IDUA guide strands	5'-UUUGAGACCUCUGUC**C**[X]GAGUUGUUCUCC-3'
Idua RT-PCR forward and sequencing primer	5'-GCTCCTCCCATCCTGTGGGCTGAACAGT-3'
Idua RT-PCR reverse primer	5'-CGGGGTGTGCGTGGGTGTCATCACT-3'

Table 5. Sequences for crystallography. All bases are ribonucleotides. (N) is 8-azanebularine.[37]

GLI1 32mer top with 8-azanebularine	5'- GCUCGCGAUGCG(N)GAGGGCUCUGAUAGCUACG -3'
GLI1 (X) 32mer bottom; X = dI, G or 3-deazadA	5'- CGUAGCUAUCAGAGCCCCCC(X)GCAUCGCGAGC -3'

Table 6. Oligonucleotide mass spectrometry data. Oligonucleotides were measured in linear positive mode analysis which typically results in mass accuracies of 0.1% or better. Standard oligonucleotides (commercially bought from Azenta Life Sciences and PAGE purified) were measured and showed mass accuracies better than 0.1% error.

Oligo	Calculated Mass (Da)	Observed Mass (m/z)
MECP2 R255X -1 3-deaza dA	9134.47	9138.8
MECP2 R255X -1 7-deaza dA	9127.192	9124.2
MECP2 R255X -1 dI	9129.177	9129.9
MECP2 R255X -1 rI	9145.177	9152.3
MECP2 R255X -1 rG	9164.5	9169.3
MECP2 R255X -1 dC	9108.5	9108.4
MECP2 R255X -1 8-azaI	9178.0	9179.2
IDUA -1 3-deaza dA	9086.1	9089.7
IDUA -1 8-azaI	9121.4	9122.6
IDUA -1 7-deaza dA	9086.1	9081.1
IDUA -1 dI	9091.1	9085.0
IDUA -1 rI	9107.1	9113.6
GLI1 8-aza top strand	10328.3	10320.6
GLI1 dI bottom strand	10176.4	10176.8
GLI1 3-deazadA bottom strand	10174.4	101773.2
GLI rG bottom strand	10204.2	10206.3

CHAPTER 6

Synthesis of an alkanethiol-containing guanosine analog for prime editor complex stabilization and structure elucidation

This chapter will briefly focus on the synthetic efforts of (1) synthesizing a novel alkanethiol tether—containing guanosine analog phosphoramidite and (2) it's incorporation and deprotection in RNA strands.

INTRODUCTION

More than 22 drugs have been approved for gene therapy utilizing CRISPR/Cas9 (**c**lustered **r**egularly **i**nterspaced **s**hort **p**alindromic **r**epeats/CRISPR-associated protein 9).[104,105] Thus, gene editing has become a very prominent tool in the growing field of oligonucleotide therapeutics and has sparked many academics and pharmaceuticals to further develop efficient systems for DNA editing, whether for off-target mitigation, editing efficiency, or delivery. Since its discovery, success, and Nobel recognition of Dr. Doudna and Dr. Charpentier[106], many other scientists in the field have perfected and modulated the original DNA editing platform with various systematic modifications (*e.g.* Cas9 mutants, length of sgRNA, fused proteins, etc.).[107–112] Many of these gene editing variants led to a novel way of gene editing: prime editing. In comparison to CRISPR/Cas9, prime editing mediates targeted insertions, deletions, and base editing without the need of double-stranded breaks (DSBs) or donor DNA templates.[112] To understand the basic science behind these editing systems, many scientists have devoted efforts in solving the structures of these DNA/RNA/enzyme complexes. In 2015, Neruki's lab solved the crystal structure of full-length *S. pyogenes* Cas9 (residues 1–1368; D10A/C80L/C574E/H840A), in complex with a 98-nt sgRNA and a 23-nt target DNA, at 2.5 Å resolution to understand the

molecular and atomic mechanism of CRISPR/Cas9 systems.[113] However, considering that prime editing is rather a newer technology, mechanistic studies of the prime editor complex need to be further analyzed. This final chapter focuses on the efforts made to try to stabilize a prime editor complex for structure elucidations utilizing novel phosphoramidites and disulfide cross-linking. This work originated as a collaboration between Dr. Lapinaite, Dr. Liu and graduate student Collin Hemez from the Broad Institute.

RESULTS and DISCUSSION

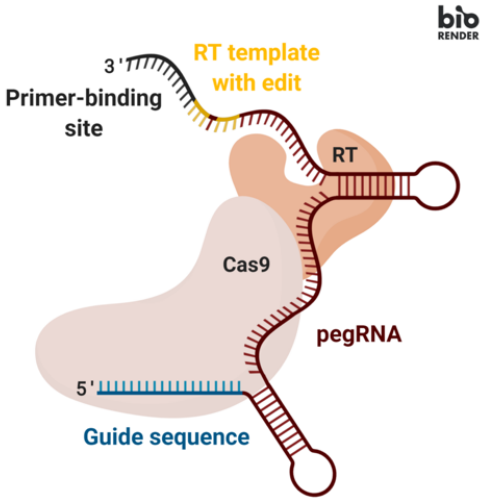


Figure 1. Components of a prime editor. PegRNA in multiple colors: black region for the primer binding site (PBS), yellow for the RT template with desired sequence, blue: guide sequence; M-MLV reverse transcriptase in orange fused to Cas9 H840A in beige. Images generated using Biorender.

Basics of prime editing

In general, the prime editor consists of four major components for proper DNA editing. These are: (1) a prime editing guide RNA (pegRNA), a (2) *Streptococcus pyogenes*-endogenous Cas9 H840A nickase fused to (3) a Moloney Murine Leukemia Virus reverse transcriptase (M-MLV) and (4) the DNA region to be edited containing a protospacer adjacent motif (PAM), which is a short specific sequence of 2-6 nt long downstream of the DNA sequence targeted by the guide RNA where the Cas9 cleaves 3-4 nt upstream of it (Figure 1).[112] In essence, the pegRNA

is responsible in identifying the target DNA site to be edited while containing new or "corrected" genetic information to be inserted. This 128-nt RNA contains a primer binding site (PBS) and a reverse transcriptase (RT) template sequence region. From the fused proteins, the mutant Cas9 enzyme contains two nuclease domains meant to cleave DNA sequences. The RuvC domain enables cleavage of the non-target strand while HNH cleaves the target DNA strand. However, because prime editing only requires cleaving at the non-target strand, the mutation of this enzyme at amino acid 840 (histidine to alanine/H840A) renders catalytic impairment of the HNH domain, inducing a single strand nick--hence the denomination "nickase".[114] Lastly, the M-MLV RT synthesizes new DNA from the single stranded RNA template (RT region).

The mechanism by which the prime editor edits DNA can be categorized in 5 main steps: (1) nicking of the PAM strand, (2) hybridization of the primer binding site to the PAM strand, (3) reverse transcription, (4) hybridization of DNA strands and flap cleavage, and finally (5) ligation and mismatch repairing (Figure 2). Briefly, in step 1, the fusion Cas9 protein nicks the target DNA sequence, exposing the 3'-OH group that can be used to initiate/prime the reverse transcription of the RT template region of the pegRNA. This results in a branched DNA structure that contains two DNA flaps: a 3'-flap that contains the newly synthesized/edited sequence, and a 5'-flap that contains a dispensable and unedited DNA sequence. The 5'-flap is cleaved by a structure-specific endonuclease or 5' exonuclease. Cleaving of this 5' flap permits the DNA ligation of the 3' flap into the native DNA and creates a heteroduplex DNA composed of one edited strand and one unedited. The re-annealed DNA usually contains nucleotide mismatches at the region where the RT synthesized new DNA. To correct these post-editing mismatches, the cells use native mismatch repair mechanism with two possible outcomes: (1) the new/corrected DNA sequence is perfectly copied and hybridized to the complementary DNA strand, resulting in efficient editing or (2) the original sequence is re-incorporated again into the edited strand, hence no editing.[112]

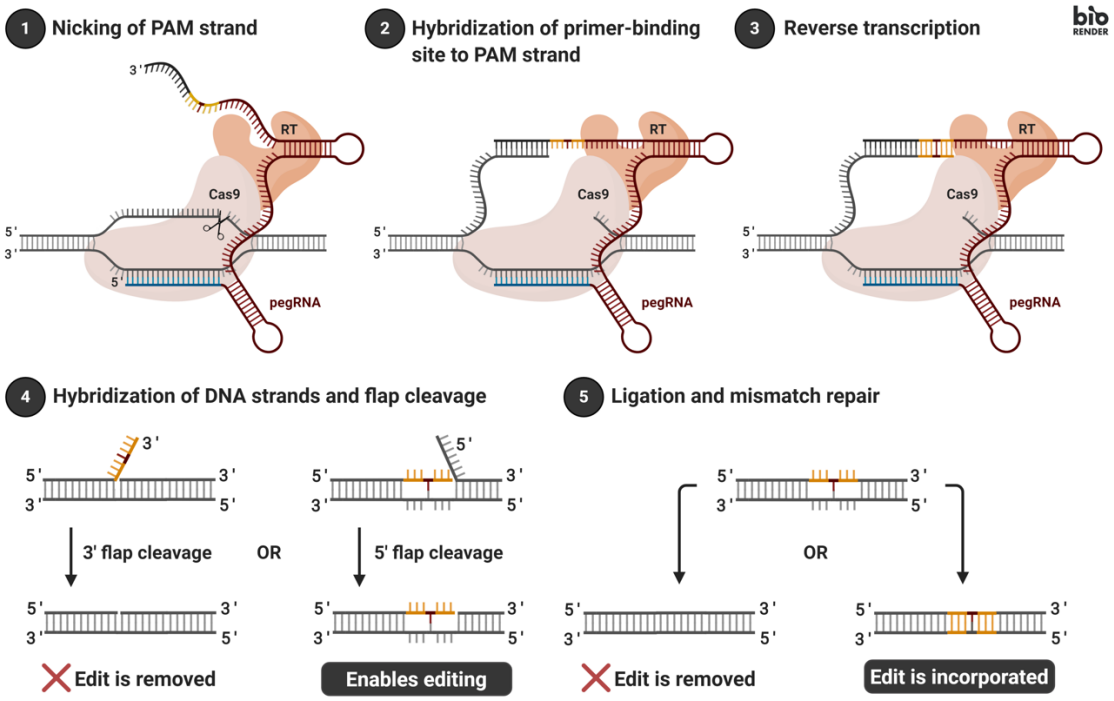


Figure 2. General mechanism of prime editing. (1) nicking of the PAM strand, (2) hybridization of the primer binding site to the PAM strand, (3) reverse transcription, (4) hybridization of DNA strands and flap cleavage and finally (5) ligation and mismatch repairing. Images generated using Biorender.

Strategies for determining the prime editor cryo-electron microscopy structure

Recently, in 2024 Shuto et al. published the cryo-electron microscopy (Cryo-EM) structure of a prime editor complex composed of S. *pyogenes* Cas9–M-MLV RTΔRNaseH–pegRNA–target DNA complex in multiple states.[115] This discovery was achieved due to the complex formation of rationally engineered pegRNA variants and prime-editor variants in which M-MLV RT is fused within SpCas9 (Figure 3). However, in this chapter we focus in solving the cryo-electron microscopy structure of the prime editor complex utilizing rather a different modality to stabilize prime editor complexes via covalent cross-linking strategies. The chemical strategies utilized in this chapter were inspired by the work of Verdine, Allerson, and co-workers on their efforts to crystallize the HIV reverse transcriptase via structural stabilization using disulfide cross-linkings.[116–118] One way to achieve structural stabilization is using modified oligonucleotides that contain thiol tethers known to react with sulfur in cysteine residues due to proximity-dependent

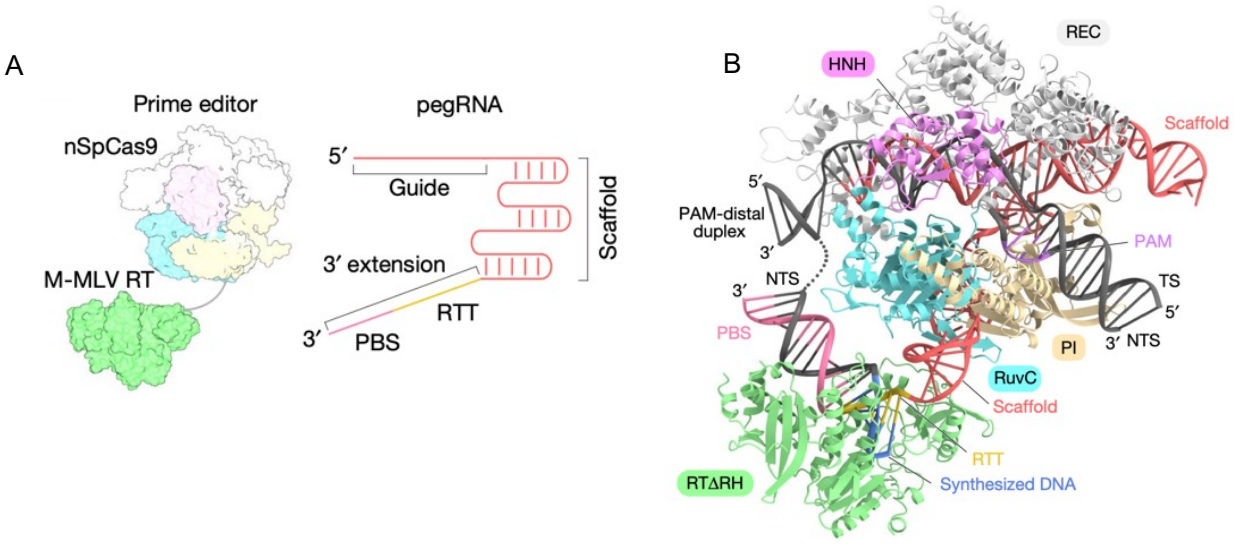


Figure 3. Cryo-EM structure of a prime editor.[115] A) Two components of the prime editing system. The prime editor is composed of nSpCas9 and M-MLV RT, and the pegRNA comprises the sgRNA region and the 3′ extension region. The 3′ extension region consists of the PBS and the RTT. The sgRNA region, PBS and RTT are colored red, pink, and yellow respectively. B) Overall structure of the SpCas9–RTΔRH–pegRNA–target DNA complex in the termination state. The disordered regions are indicated as dotted lines.

reactivity between both sites (Figure 4). This in hand requires the synthesis of thiol tether containing building blocks for the synthesis of thiol-modified oligonucleotides. Several labs have invested time in developing nucleoside analogs of this nature for the sake of covalently trapping DNA/RNA binding proteins to RNA/DNA duplexes. To illustrate, Tran et al. and Peacock et al. in the Beal lab have developed phosphoramidites for tethering in RNA with different thiol-containing purine analogs: (a) 2-*N*-(2-tritylthioethylamino)purine[119], (b) 6-*N*-(2-tritylthioethylamino)adenosine[120], (c) 2-*N*-(2-tritylthioethylamino)guanosine (Dr. Kiet Tran dissertation). The structure of these monomers can be found in Figure 5a-c. For instance, other groups (Xu et al. in Swann lab) have incorporated thiol tethers in DNA post solid phase synthesis using a reactive 4-*N*-triazoly-thymidine phosphoramidite that selectively reacts to nucleophilic sulfur or nitrogen ligands (*e.g.* cysteamine, methylamine, etc.) to afford thymine/cytidine analogs modified at position 4 of the pyrimidine (Figure 5e-f). Even though compounds **5a-c**, and **5e** have been published or synthesized before, compound **5d**, has not.

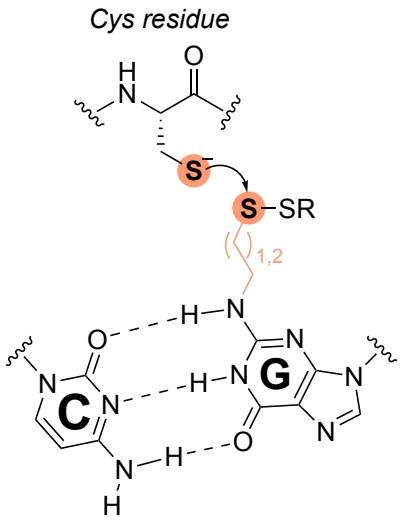


Figure 4. Disulfide cross-linking model and reaction structure used here. A modified guanosine analog with an alkanethiol tether in the major groove with various lengths (C3 or C2) within a G:C base pair proximal to a nucleophilic cysteine residue of an HIV RT for disulfide cross-linking.[117,118]

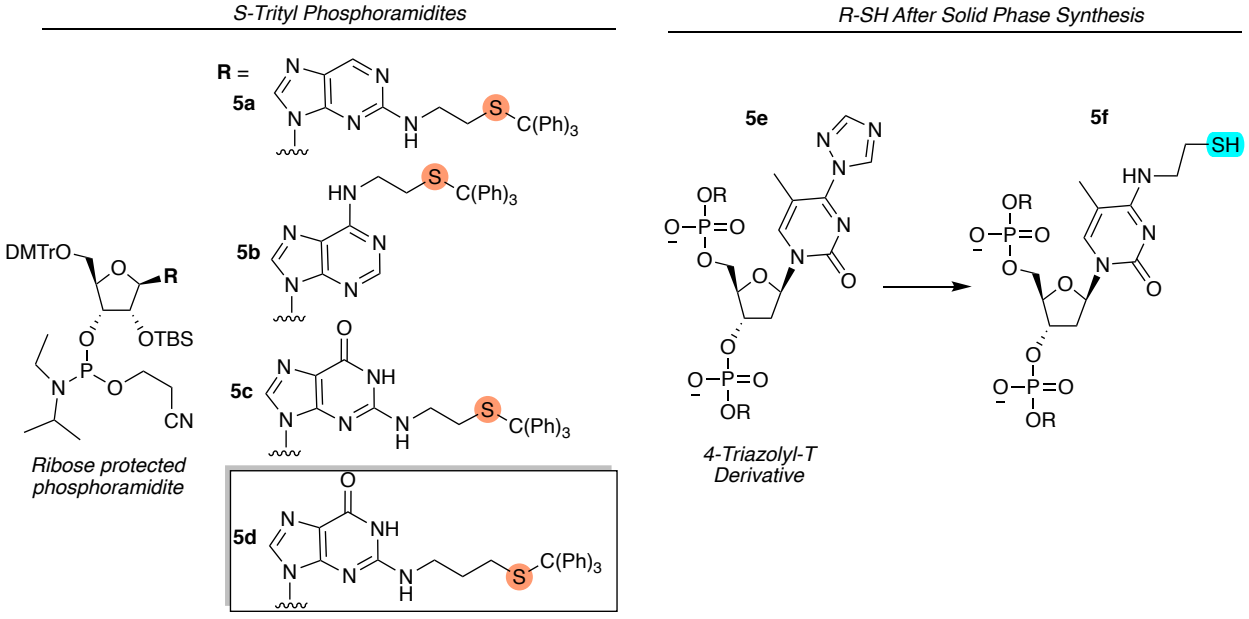


Figure 5. Sulfur containing monomers for disulfide cross-linkings. (a-d) RNA phosphoramidites with trityl-protected thiols suitable for oligonucleotide solid phase synthesis. (a) 2-*N*-(2-tritylthioethylamino)purine[119], (b) 6-*N*-(2-tritylthioethylamino)adenosine[120], (c) 2-*N*-(2-tritylthioethylamino)guanosine, (d) 2-*N*-(2-tritylthiopropylamino)guanosine. 4-*N*-triazoly-thymidine monomer[121] for post-oligonucleotide solid phase synthesis derivatization. (e) Structure 4-*N*-triazolyl-thymidine monomer in DNA reactive to nucleophiles for nucleophilic aromatic substitutions (NAS); (f) Structure of DNA product post NAS to afford 4-*N*-(2-tritylthioethylamino)-5-methylcytidine monomer.

Utilizing disulfide cross-linking with a novel nucleoside analog to stabilize the pegRNA-RT complex

Because this collaboration originated to stabilize the prime editor complex for structure determination, we rationally designed a nucleoside analog that would covalently cross-link the reverse transcriptase with the pegRNA in the primer binding site of the RT. We rationalized this by using an analogous model that Verdine and co-workers published where they covalently trap the RT of the human immunodeficiency virus (HIV) to a primer:template oligonucleotide with a modified G analog, all in efforts to crystallize the HIV RT to its substrate.[117,122] This G analog has an alkanethiol tethered to it that depending in its proximity to cysteine residues, it can furnish disulfide adducts. Based on the structural analogy of amino acid sequences proximal to a guanosine in the primer:template of the HIV RT, the amino acid residue R298 of M-MLV was mutated to cysteine (R298C), which is near a minor groove guanosine in a position distal to the dNTP-adding end of the protein. Thus, we reasoned that a modified guanosine three bases into the primer binding site of the pegRNA would have the highest chance of cross-linking, as it would allow the RT to add 2-3 bases before getting trapped. To make this possible, a 128-nt pegRNA was designed with 2-*N*-(2-tritylthiopropylamino)guanosine (Figure 5d) positioned at the 7th base from the 3' end of the pegRNA in the PBS. We chose this analog considering that a three carbon thiol linker would be ideal considering its proposed proximity to cysteine residue R298C in the M-MLV RT.

Strategy to synthesize a 128-nt modified pegRNA

Synthesizing longer than 40-nt RNAs via solid phase synthesis becomes a challenge considering the low coupling efficiencies and yields as the RNA strand grows longer due to the steric demand between each silyl-protected internucleotide linkage. Hence, to make a 128-nt RNA, other strategies had to be employed. We envisioned synthesizing two RNA fragments, one with the modified G analog and the other one with canonical RNA. After synthesis and purification,

both would be fused using a T4 DNA ligase (Figure 6). Therefore, we divided the pegRNA into fragment 1 (96-nt RNA) and fragment 2 (32-nt with modified G-tether) (see Table 1 for sequences). Fragment 1 was commercially available by Integrated DNA Technologies (IDT), whereas fragment 2 was synthesized in house.

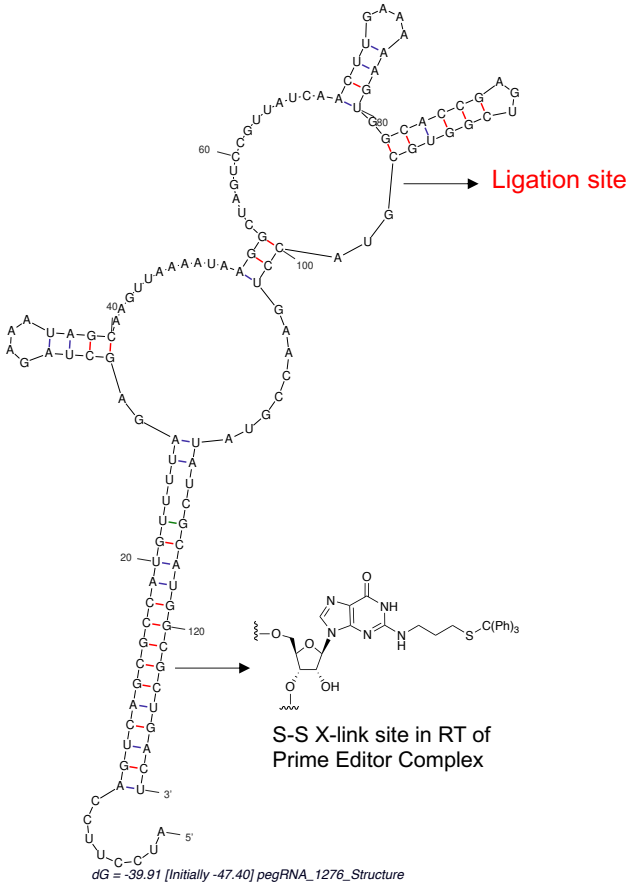
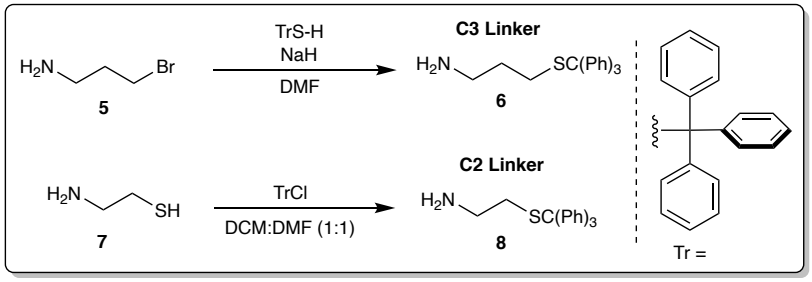


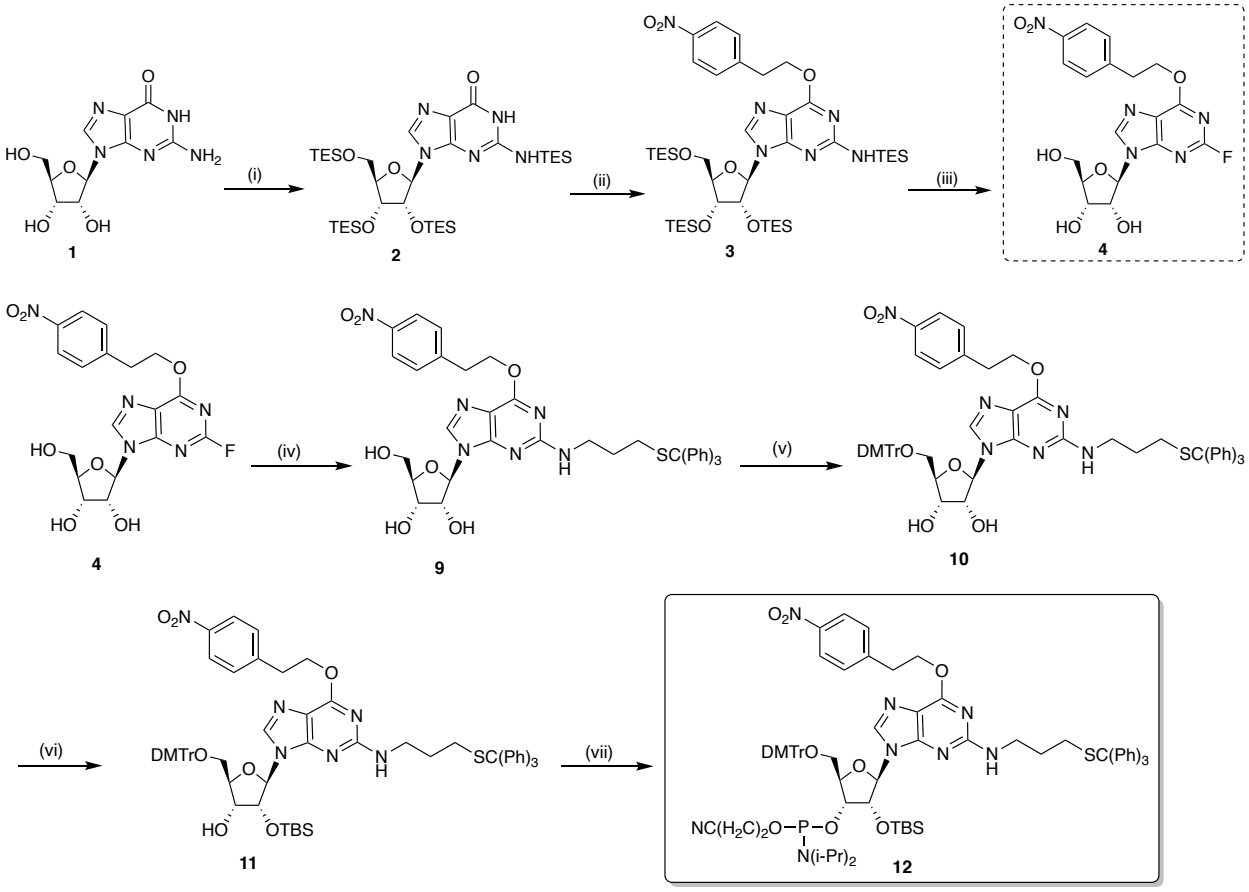
Figure 6. mfold model of the 128-nt pegRNA.[83]

In order to synthesize the C3-linker G analog phosphoramidite (Figure 5d), a novel synthetic route had to be employed. Previously, Dr. Kiet Tran synthesized the C2-linker G analog phosphoramidite (Figure 5c) utilizing the previously published linker tritylthiolethylamino **8** (Scheme 1).[120] Here, we synthesize the C3-linker G analog phosphoramidite utilizing the propyl version of this linker **6** (Scheme 1).[116] We synthesized phosphoramidite **12** in 7-steps starting with commercially available and inexpensive guanosine (**1**) (Scheme 2). In 1997, Allerson et al.

synthesized the 2-F inosine intermediate (**4**) for post solid phase synthesis 2-position derivatizations.[118]



Scheme 1. Synthesis of ethyl and propyl aminothiol linkers.[116,120] Reagents and conditions: (i): 46%; (ii): 41%



Scheme 2. Synthesis of 2-*N*-(3-tritylthiopropylamino)guanosine phosphoramidite.[118] *Reagents and conditions:* (i): TESCl, imidazole, DMF, 85%; (ii): DIAD, $P(OPh)_3$, 1,4-dioxane, 4-nitrophenethyl ethanol 80%; (iii): HF/pyridine, *t*-BuONO, 75%; (iv) $H_2N(CH_2)_3STr$, DMF:DCM (1:1), 89%; (v): DMTrCl, pyridine, DMAP, 90%; (vii): TBDMS-OTf, pyridine, 37%; (viii): DIPEA, DCM, 2-cyanoethyl-(N,N-diisopropylamino)chlorophosphite, 79%.

The synthesis consisted in initially protecting all the free amino and hydroxy groups in guanosine with triethylsilyl chloride (TESCl) and imidazole. Post silyl-protection, the O-6 in guanosine **5** was protected with 4-nitrophenethyl ethanol (NPE) using diisopropyl azodicarboxylate and triphenyl phosphite to furnish fully protected G **3** in reasonable yields. Subsequent diazotization with *tert*-butyl nitrite in the presence of 60% HF/pyridine resulted in fluorination at the 2-position and parallel removal of the triethylsilyl ethers, thus affording the nucleoside **4**. To incorporate the thiol-amino tether, we subjected **4** to react with linker **6** via a nucleophilic aromatic substitution in DMF:DCM (1:1) at room temperature for 8 h to afford the alkanethiol-tethered G **9** in good yields. Further 5' dimethoxytritylation in pyridine afforded **10** in 90% yield. Silylation of the 2'-OH was originally attempted with standards conditions employing tert-butyldimethylsilyl chloride (TBSCl) and triethylamine. However, silver nitrate was avoided considering that removal of the trityl group may occur due to silver-sulfur strong-binding complexes. However, without silver nitrate, barely any product was formed. To overcome this, we employed pyridine and *tert*-butyldimethylsilyl triflate (TBSOTf), which has a better leaving group than TBSCl, and a stronger electrophilic silyl ether. Hence, we successfully and selectively isolated the 2'-O silylated intermediate **11** with 37%, which is usual for competitive 2',3'-diol protections. Lastly, phosphitylation of the 3'-OH afforded phosphoramidite **12** in 79% yield. Incorporation of amidite **12** into fragment 2 of the pegRNA followed standard coupling RNA procedures, except for prolonged coupling time (30 min). Furthermore, incorporation deemed fruitful and fragment 2 was successfully synthesized via solid phase synthesis (Figure 7A). After cleavage, deprotection, and purification of the tritylated thiol-containing RNA (fragment 2) utilizing standard conditions, subjection of this RNA to silver nitrate and dithiothreitol (DTT) in triethylammonium acetate afforded the free thiol in reasonable yields (62%) (Figure 7B). After buffer exchange and precipitation, the 32-nt RNA fragment 2 containing the free thiol tether was confirmed by mass spectrometry (Figure 7C).

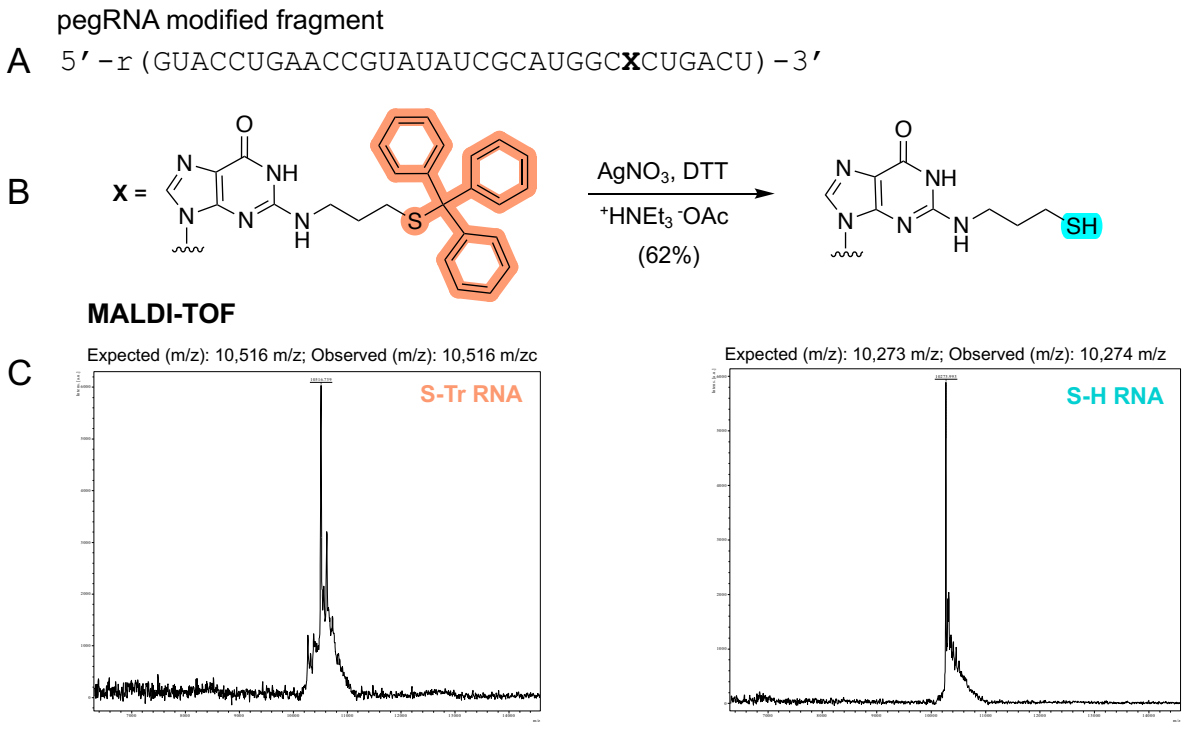


Figure 7. Deprotection of trityl-protected thiols in RNA strands. (A) Sequence of the 32-nt fragment of the pegRNA model that corrects a 4 bp insertion in the *HEXA* gene that causes Tay-Sachs disease used for complex stabilization complex with prime editors. X indicates position of modification for disulfide cross-linking in the prime binding site (PBS) of the reverse transcriptase (RT). (B) Detritylation reaction of 2-*N*-(3-thiotritylpropyl)guanine in an RNA strand employing silver nitrate and DTT in triethylammoniun acetate buffer. (C) MS (MALDI-TOF) characterization of the detritylated 32-nt RNA product post purification and buffer exchange.

The work done in this project has been halted due to the recent discovery of the prime editor structure solved by the Neruki lab in Japan. However, I hope this information is helpful for future scientists utilizing nucleoside analogs for tethering RNA/DNA to proteins.

METHODS

Synthesis of 2-N-(3-Tritylthiolpropylamino)guanosine phosphoramidite

(2*R*,3*S*,4*R*,5*R*)-2-(hydroxymethyl)-5-(6-(4-nitrophenethoxy)-2-((3-(tritylthio)propyl)amino)-9*H*-purin-9-yl)tetrahydrofuran-3,4-diol (9). The 2-F-inosine intermediate (**4**) (0.266 mmol, 116 mg) was dissolved in 6 mL of DCM:DMF (1:1). The solution was then treated with the S-Tr C3 linker **6** (0798 mmol, 266 mg) and the reaction was left stirring at room temperature for 16 h. The

reaction mixture was quenched by the addition of a saturated aqueous solution of sodium bicarbonate and the organic phase was extracted using excess DCM. The organic phase was washed 3x with brine to remove DMF. The organic phase was then dried over sodium sulfate and filtered to afford a syrup. The crude was purified by column chromatography using a 4-7% gradient of MeOH in DCM to afford **9** as a yellow pale solid in 89% yield. ^{1}H NMR (300 MHz, MeOD) δ 8.19 – 8.06 (m, 2H), 8.03 (s, 1H), 7.51 (d, *J* = 8.4 Hz, 2H), 7.42 – 7.01 (m, 17H), 5.84 (d, *J* = 6.0 Hz, 1H), 4.68 (dt, *J* = 10.4, 6.0 Hz, 3H), 4.30 (dd, *J* = 5.2, 3.1 Hz, 1H), 4.10 (t, *J* = 3.0 Hz, 1H), 3.82 (dd, *J* = 12.3, 2.9 Hz, 1H), 3.70 (dd, *J* = 12.3, 3.3 Hz, 1H), 3.36 (t, *J* = 6.0 Hz, 2H), 3.21 (t, *J* = 6.6 Hz, 2H), 2.19 (t, *J* = 7.4 Hz, 2H), 1.66 (h, *J* = 7.1 Hz, 2H). ^{13}C NMR (76 MHz, MeOD) δ 161.65, 160.51, 154.65, 148.12, 147.80, 146.30, 140.30, 131.32, 131.24, 130.70, 128.76, 127.62, 124.50, 115.61, 90.68, 87.49, 74.96, 72.51, 67.76, 67.24, 63.43, 41.73, 36.06, 30.61, 30.08.

(2*R*,3*S*,4*R*,5*R*)-2-((Bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-5-(6-(4-nitrophenethoxy)-2-((3-(tritylthio)propyl)amino)-9*H*-purin-9-yl)tetrahydrofuran-3,4-diol (10). 4,4-Dimethoxytrityl chloride (DMTrCl) (217 mg, 0.640 mmol, 1.2 eq.) and 4-Dimethylaminopyridine (DMAP) (32 mg, 0.26, 0.5 eq) were added to a solution of **2** (400 mg, 1.21 mmol) in 4.4 mL of anhydrous and distilled pyridine. The reaction mixture was stirred at room temperature for 24 h. Co-evaporation with 20 mL of toluene 2x was performed to remove pyridine. Afterwards, the mixture was diluted with dichloromethane (15 mL) and washed with a saturated aqueous solution of $NaHCO_3$ (2 x 10 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to afford an orange-to-amber oil. The crude was then purified by column chromatography (0-3% methanol in dichloromethane with 1% triethylamine). Excess triethylamine in column fractions were removed via azeotrope formation with acetonitrile, yielding compounding **3** as a pale yellow foam (2.02 g, 90%).^{1}H NMR (300 MHz, $CDCl_3$) δ 8.04 (d, 2H), 7.67 (s, 1H), 7.21 (s, 26H), 6.72 – 6.60 (m, 4H), 5.70 (d, *J* = 5.9 Hz, 1H), 4.79 (s, 1H), 4.60 (dt, *J* = 7.4, 6.0 Hz, 3H), 4.26 (t, *J* = 4.1 Hz, 2H), 3.65 (s, 6H), 3.34 – 3.04 (m, 4H), 2.15 (t, *J* = 7.1 Hz, 2H), 1.88 (s, 1H), 1.56 (s, 3H). ^{13}C NMR (76 MHz, CD_2Cl_2) δ 160.53, 158.65, 158.56,

153.32, 146.84, 146.22, 144.88, 144.71, 137.22, 135.66, 135.57, 130.02, 129.96, 129.55, 128.03, 127.86, 126.82, 126.65, 123.60, 114.94, 113.10, 89.76, 86.43, 85.42, 75.27, 72.26, 66.69, 66.07, 63.85, 55.18, 54.19, 53.83, 53.47, 53.11, 52.75, 40.90, 35.12, 29.30, 28.55. ESI-HRMS: calcd for $C_{61}H_{58}N_6O_9S$ $(M + H)^+$: 1051.40, obsd: 1051.4007.

(2*R*,3*R*,4*R*,5*R*)-2-((Bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-((*tert*-butyldimethylsilyl)oxy)-5-(6-(4-nitrophenethoxy)-2-((3-(tritylthio)propyl)amino)-9*H*-purin-9-yl)tetrahydrofuran-3-ol (11). After co-evaporating **10** (320 mg, 0.300 mmol) 3x in anhydrous pyridine, the intermediate was dissolved in 3.9 mL of anhydrous and distilled pyridine and cooled to -42 °C. *tert*-butyldimethylsilyl trifluoromethanesulfonate (TBDMS-OTf) (31 µL, 0.135 mmol, 0.45 eq.) was added dropwise and stirred for 5 min at -42 °C followed by 10 min at room temperature. This cycle of TBDMS-OTf addition was repeated 3 more times (1.8 eq. total). Upon completion of the reaction, the mixture was diluted with EtOAc (20 mL) and washed with saturated $NaHCO_3$ (2 x 20 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to afford an orange-to-amber oil. The crude reaction mixture was purified by column chromatography (20 → 30% EtOAc in *n*-hexanes) to afford compound **4** as a white foam (130 mg, 37%). ^{1}H NMR (400 MHz, CD_2Cl_2) δ 8.21 – 8.14 (m, 2H), 7.72 (s, 1H), 7.59 – 7.11 (m, 26H), 6.87 – 6.79 (m, 4H), 5.89 – 5.84 (m, 1H), 4.97 (s, 1H), 4.74 (t, *J* = 6.8 Hz, 2H), 4.64 (s, 1H), 4.38 (q, *J* = 4.6 Hz, 1H), 4.19 (q, *J* = 4.0 Hz, 1H), 3.79 (s, 6H), 3.53 – 3.11 (m, 5H), 2.69 (d, *J* = 4.6 Hz, 1H), 2.20 (t, *J* = 7.2 Hz, 2H), 1.57 (d, *J* = 7.2 Hz, 2H), 0.88 (s, 9H), 0.03 (s, 3H), -0.10 (s, 3H).^{13}C NMR (101 MHz, CD_2Cl_2) δ 160.48, 158.81, 158.68, 154.10, 146.83, 146.33, 144.93, 144.86, 135.66, 130.10, 130.06, 129.93, 129.54, 129.03, 128.12, 127.85, 127.82, 126.85, 126.60, 123.58, 115.23, 113.10, 88.06, 86.41, 83.66, 71.43, 66.59, 65.82, 63.65, 55.19, 40.81, 35.17, 29.29, 28.64, 25.36, 25.30, 17.80, 0.76, -5.11, -5.39. ESI-HRMS: calcd for $C_{67}H_{72}N_6O_9SSi$ $(M + H)^+$: 1165.49, obsd: 1165.4885.

(2*R*,3*R*,4*R*,5*R*)-2-((Bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-((*tert*-butyldimethylsilyl)oxy)-5-(6-(4-nitrophenethoxy)-2-((3-(tritylthio)propyl)amino)-9*H*-purin-9-

yl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (12). To a flame-dried 25 mL round bottom flask, distilled and anhydrous diisopropylethylamine (117 uL, 0.670 mmol, 6 eq.) and 2-cyanoethyl-(N,N-diisopropylamino)chlorophosphite (60 uL, 0.222 mmol, 2 eq.) were added to a solution of compound **4** (130 mg, 0.111 mmol) in 1.8 mL of anhydrous dichloromethane. The reaction mixture was stirred for 1.5 h at room temperature. Upon completion of the reaction, the mixture was diluted with EtOAc (10 mL) and washed with 5% (w/v) aqueous $NaHCO_3$ (2 x 10 mL). The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to afford a pale yellow syrup. Purification by column chromatography (20-40% EtOAc in *n*-hexanes) yielded compound **5** as a white foam (119 mg, 79%). ^{31}P NMR (122 MHz, CD_2Cl_2) δ 150.73, 149.01. ESI-HRMS: calcd for $C_{76}H_{89}N_8O_{10}PSSi$ $(M + H)^+$: 1365.59, obsd: 1365.6007.

Synthesis and purification of oligonucleotides. Chemical synthesis for all oligonucleotides was performed using an ABI 394 synthesizer. All protected phosphoramidites were purchased from Glen Research. 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 °C for 12 h. The supernatant was transferred to a new screw-cap tube and evaporated under reduced pressure. Desilylation was performed by treating the pellets with 250 μL of 1M TBAF–THF at room temperature overnight. To each reaction was added 75 mM sodium acetate in butanol. The oligonucleotides were then precipitated from a solution of 65% butanol at -70 °C for 2 h. The solution was centrifuged at 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. Single-stranded RNA oligonucleotides were purified by denaturing PAGE and visualized by UV shadowing. Bands were excised from the gel, crushed, and soaked overnight at 4 °C in 0.5 M NaOAc, 0.1% (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 °C 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.

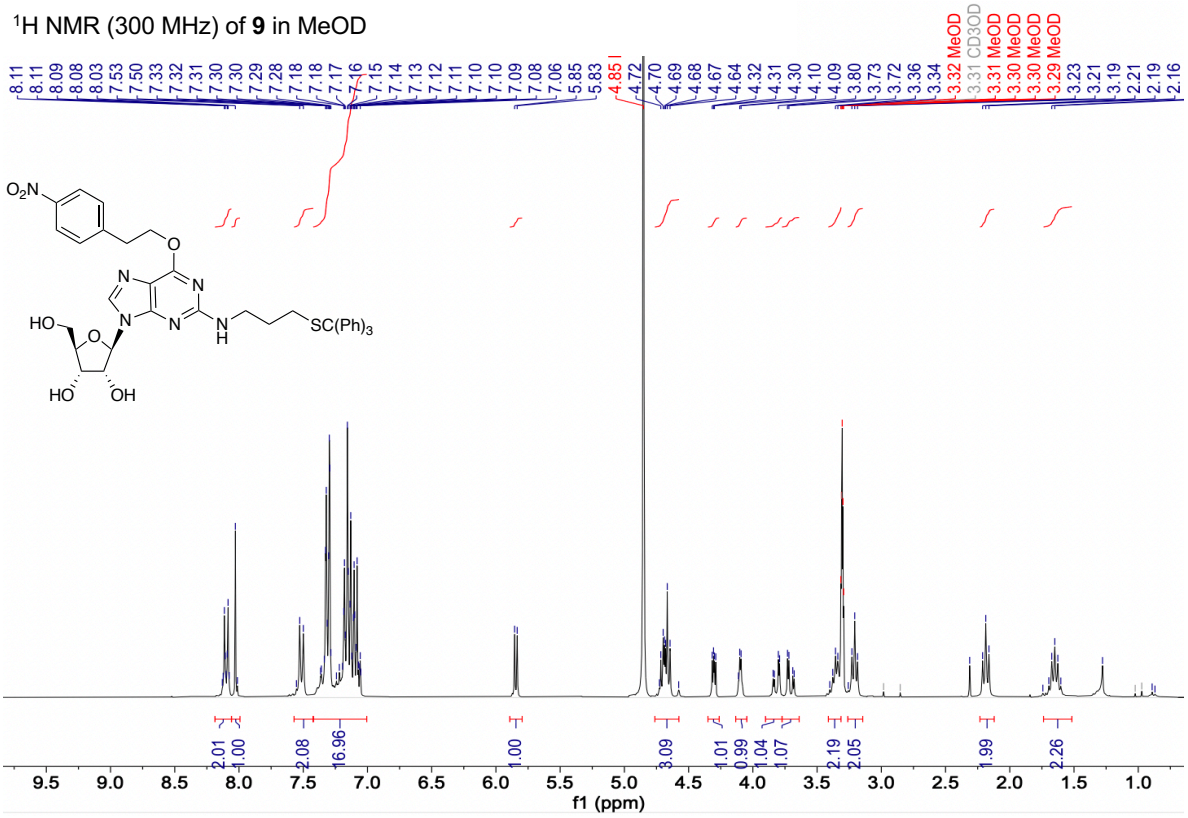
S-Trityl RNA Deprotection. To remove the S-Trityl protecting group, purified RNA (940 pmol) was dissolved in aqueous triethylammonium acetate (6 uL, 0.1 M, pH 6.5) and treated with $AgNO_3$ (1 uL, 1 M in H_2O). After 30 min of gentle agitation at room temperature, DTT was added (2 uL, 1 M H_2O), and the reaction was allowed to proceed for a further 5 min. The mixture was diluted with aqueous triethylammonium acetate (50 uL, 0.1 M, pH 6.5) and centrifuged (16 000 x g, 4 °C, 5 min) to separate the precipitate. The supernatant was collected and reserved. The pellet was washed twice with 50 uL aqueous triethylammonium acetate, to minimize the loss of the RNA. During each wash, the pellet was agitated vigorously by vortexing for several minutes, followed by centrifuging. The combined supernatants (150 uL) were loaded in to a 3000 MWCO centrifugal concentrator (Microcon-3, Millipore) and centrifuged (12 000 x g, 4 °C, approx. 40 min). The concentrated solution was diluted with storage/reaction buffer (50 mM Bis-Tris•HCl, 100 mM NaCl, 10 mM $MgCl_2$, pH 7.0 containing 1 mM tris(2-carboxyethyl)phosphine•HCl (TCEP)) and re-concentrated. Dilution and re-concentration was performed two further times to achieve complete buffer exchange. Deprotected RNA was quantified by the absorbance of the solution at 260 nm. Yields ranged from 330 pmol (35%) to 583 pmol (62%). MALDI-MS confirmed quantitative conversion to thiol-RNA, and the purity of the product. The deprotected RNA was stored in buffer at -80 °C and was stable for several months under these conditions.

Table 1. Sequences of oligonucleotides used for the proposal of DNA T4 ligation for pegRNA synthesis. All sequences displayed are 5' to 3'. All A, G, C and U are ribose except DNA splint.

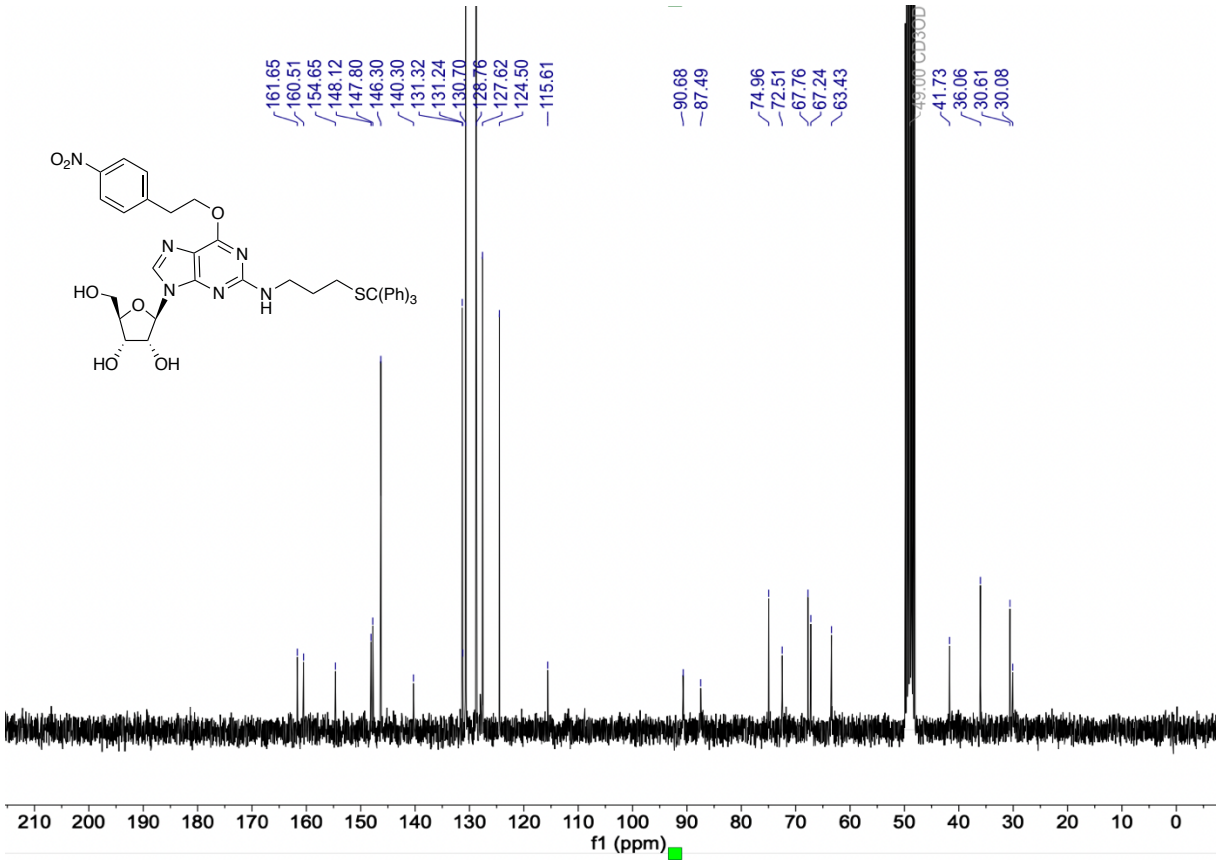
pegRNA model that corrects a 4 bp insertion in the *HEXA* gene that causes Tay-Sachs disease used here	AUCCUUCCAGUCAGCGCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGUACCUGAACCGUAUAUCGCAUGGC**X**CUGACU
Fragment 1 of pegRNA (96-nt) bought from IDT	AUCCUUCCAGUCAGCGCCAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
Fragment 2 of pegRNA (32-nt) with 5'-PO_4^- (control)	**Phos/**GUACCUGAACCGUAUAUCGCAUGGC**G**CUGACU
Fragment 2 of pegRNA (32-nt) with 5'-PO_4^- (S = N^2-(3-thiopropyl)G)	**Phos/**GUACCUGAACCGUAUAUCGCAUGGC**S**CUGACU
DNA splint for T4 DNA Ligase (all 2'-Deoxy)	AGTCAGCGCCATGCGATATACGGTTCAGGTACGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACATGGCGCTGACTGGAAGGAT

NMR Spectra

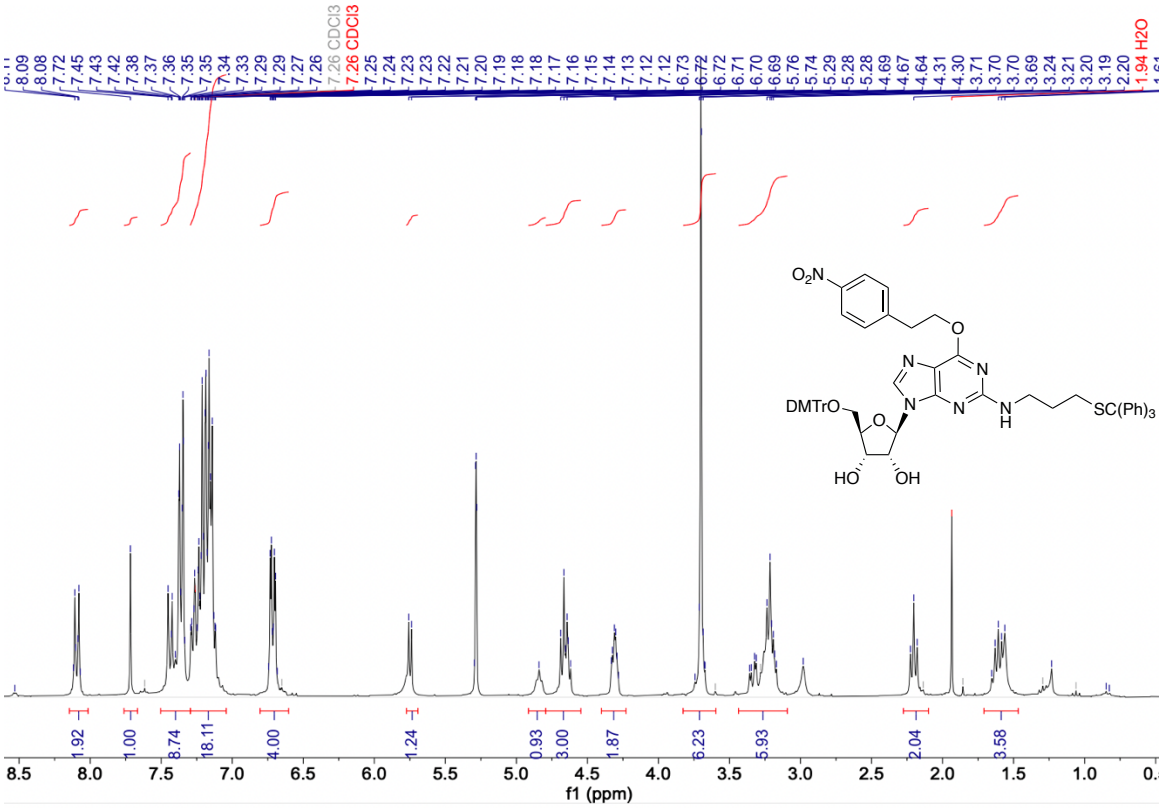
^{1}H NMR (300 MHz) of **9** in MeOD



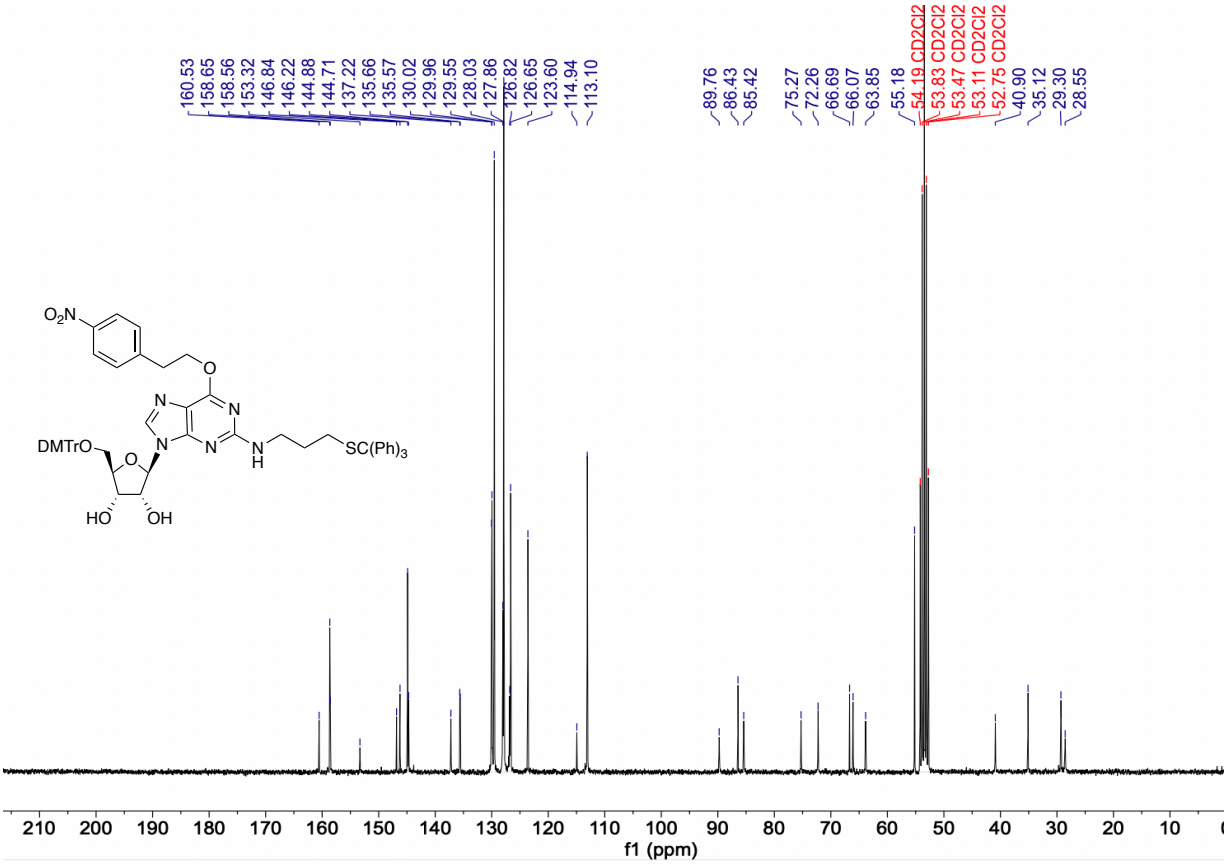
^{13}C NMR (76 MHz) of **9** in MeOD



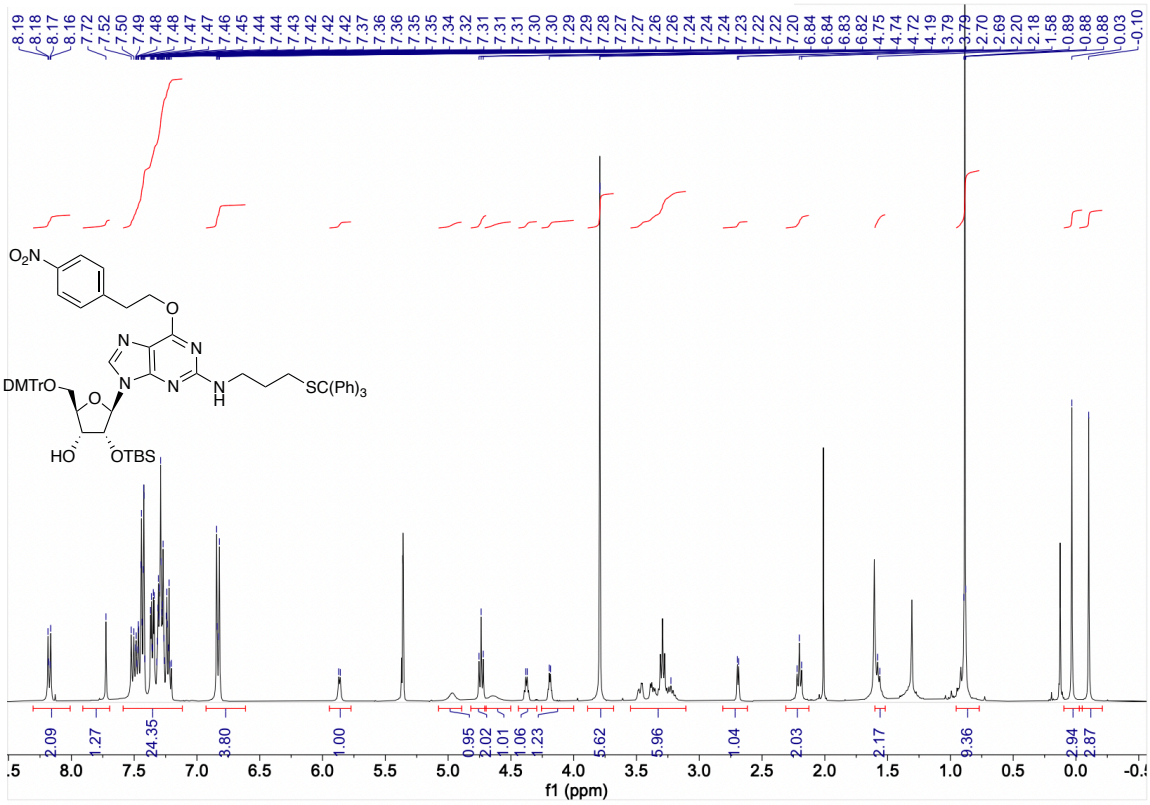
^{1}H NMR (300 MHz) of **10** in $CDCl_3$



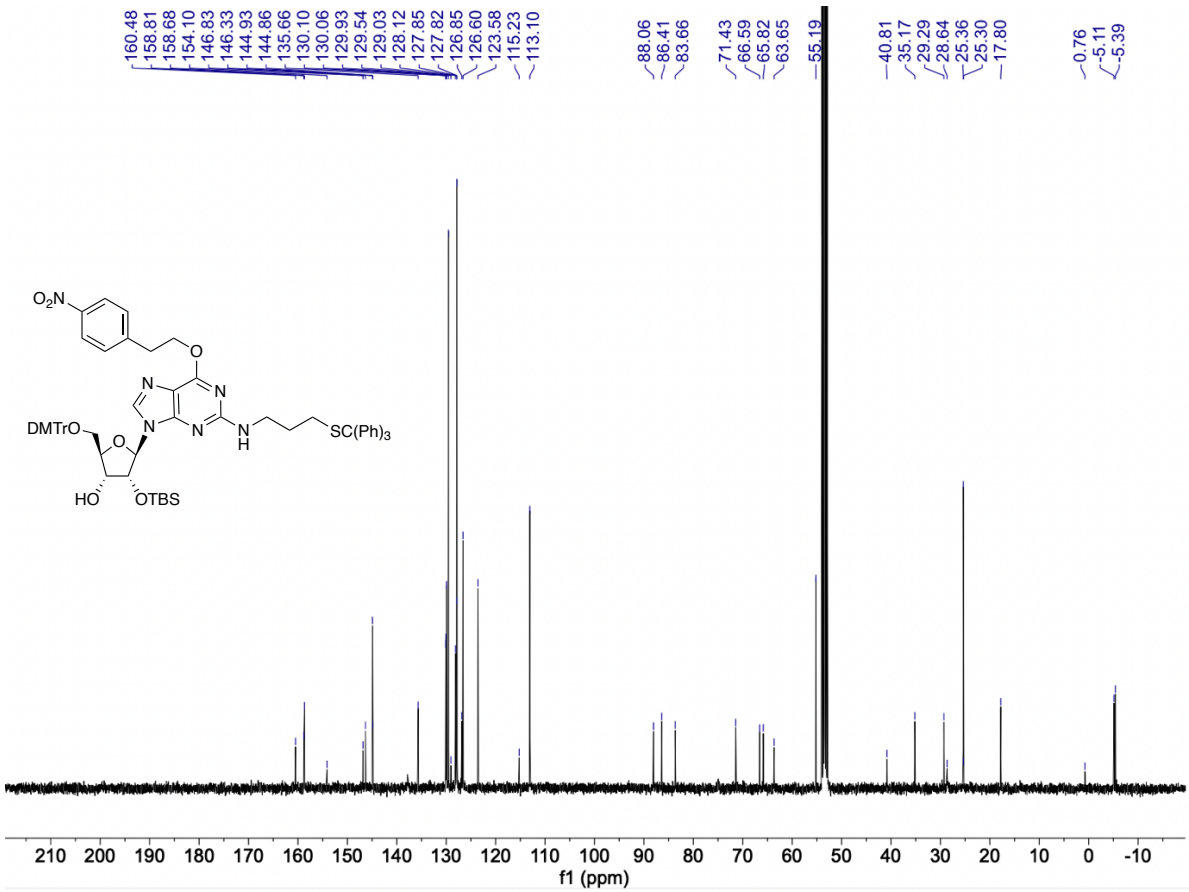
^{13}C NMR (76 MHz) of **10** in $CDCl_3$



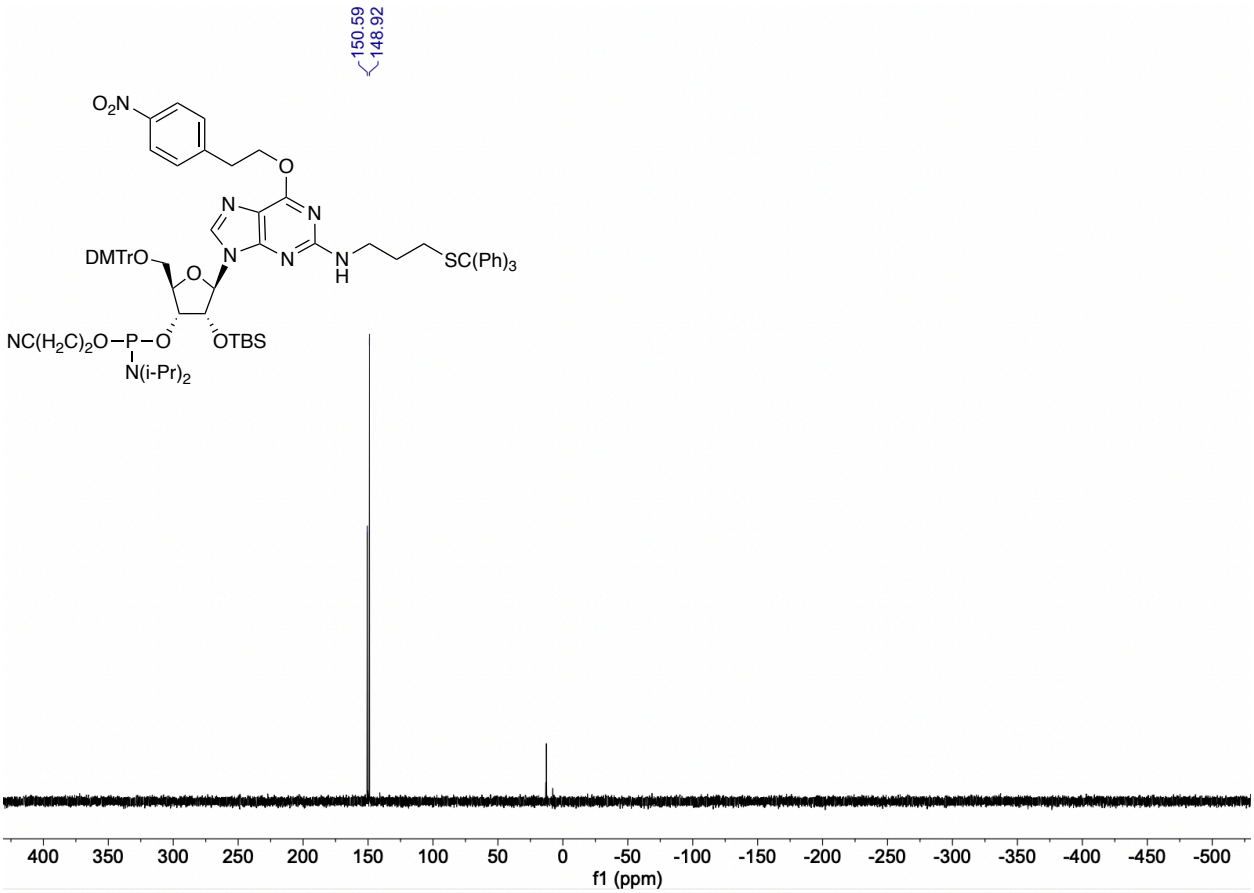
^{1}H NMR (300 MHz) of **11** in CD_2Cl_2



^{13}C NMR (76 MHz) of **11** in CD_2Cl_2



^{31}P NMR (122 MHz) of **12** in CD_2Cl_2



REFERENCES

(1) Quin, J.; Sedmík, J.; Vukić, D.; Khan, A.; Keegan, L. P.; O'Connell, M. A. ADAR RNA Modifications, the Epitranscriptome and Innate Immunity. *Trends Biochem. Sci.* **2021**, *46* (9), 758–771. https://doi.org/10.1016/j.tibs.2021.02.002.

(2) Khosravi, H. M.; Jantsch, M. F. Site-Directed RNA Editing: Recent Advances and Open Challenges. *RNA Biol.* **2021**, *18* (sup1), 41–50. https://doi.org/10.1080/15476286.2021.1983288.

(3) Zinshteyn, B.; Nishikura, K. Adenosine-to-Inosine RNA Editing. *WIREs Systems Biology and Medicine* **2009**, *1* (2), 202–209. https://doi.org/https://doi.org/10.1002/wsbm.10.

(4) Fisher, A. J.; Beal, P. A. Effects of Aicardi-Goutières Syndrome Mutations Predicted from ADAR-RNA Structures. *RNA Biol.* **2017**, *14* (2), 164–170. https://doi.org/10.1080/15476286.2016.1267097.

(5) Kuttan, A.; Bass, B. L. Mechanistic Insights into Editing-Site Specificity of ADARs. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (48). https://doi.org/10.1073/pnas.1212548109.

(6) Bass, B. L. RNA Editing by Adenosine Deaminases That Act on RNA. *Annu. Rev. Biochem.* **2002**, *71* (Volume 71, 2002), 817–846. https://doi.org/https://doi.org/10.1146/annurev.biochem.71.110601.135501.

(7) Hogg, M.; Paro, S.; Keegan, L. P.; O'Connell, M. A. 3 - RNA Editing by Mammalian ADARs. In *Adv. Genet.*; Friedmann, T., Dunlap, J. C., Goodwin, S. F., Eds.; Academic Press, 2011; Vol. 73, pp 87–120. https://doi.org/https://doi.org/10.1016/B978-0-12-380860-8.00003-3.

(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 Acid Res.* **2020**, *48* (14), 7958–7972. https://doi.org/10.1093/nar/gkaa532.

(9) Wang, Y.; Chung, D. hee; Monteleone, L. R.; Li, J.; Chiang, Y.; Toney, M. D.; Beal, P. A. RNA Binding Candidates for Human ADAR3 from Substrates of a Gain of Function Mutant Expressed in Neuronal Cells. *Nucleic Acids Res.* **2019**, *47* (20), 10801–10814. https://doi.org/10.1093/nar/gkz815.

(10) Savva, Y. A.; Rieder, L. E.; Reenan, R. A. The ADAR Protein Family. *Genome Biol.* **2012**, *13* (12), 252. https://doi.org/10.1186/gb-2012-13-12-252.

(11) Wang, Y.; Zheng, Y.; Beal, P. A. Chapter Seven - Adenosine Deaminases That Act on RNA (ADARs). In *The Enzymes*; Chanfreau, G. F., Ed.; Academic Press, 2017; Vol. 41, pp 215–268. https://doi.org/https://doi.org/10.1016/bs.enz.2017.03.006.

(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*. https://doi.org/10.3389/fimmu.2019.01763.

(14) Mendoza, H. G.; Beal, P. A. Chemical Modifications in RNA: Elucidating the Chemistry of DsRNA-Specific Adenosine Deaminases (ADARs). *Acc. Chem. Res.* **2023**, *56* (18), 2489–2499. https://doi.org/10.1021/acs.accounts.3c00390.

(15) Riedmann, E. M.; Schopoff, S.; Hartner, J. C.; Jantsch, M. F. Specificity of ADAR-Mediated RNA Editing in Newly Identified Targets. *RNA* **2008**, *14* (6), 1110–1118. https://doi.org/10.1261/rna.923308.

(16) Yang, S.; Deng, P.; Zhu, Z.; Zhu, J.; Wang, G.; Zhang, L.; Chen, A. F.; Wang, T.; Sarkar, S. N.; Billiar, T. R.; Wang, Q. Adenosine Deaminase Acting on RNA 1 Limits RIG-I RNA Detection and Suppresses IFN Production Responding to Viral and Endogenous RNAs. *J. Immunol.* **2014**, *193* (7), 3436–3445. https://doi.org/10.4049/jimmunol.1401136.

(17) Gannon, H. S.; Zou, T.; Kiessling, M. K.; Gao, G. F.; Cai, D.; Choi, P. S.; Ivan, A. P.; Buchumenski, I.; Berger, A. C.; Goldstein, J. T.; Cherniack, A. D.; Vazquez, F.; Tsherniak, A.; Levanon, E. Y.; Hahn, W. C.; Meyerson, M. Identification of ADAR1 Adenosine Deaminase Dependency in a Subset of Cancer Cells. *Nat. Commun.* **2018**, *9* (1), 5450. https://doi.org/10.1038/s41467-018-07824-4.

(18) Bhate, A.; Sun, T.; Li, J. B. ADAR1: A New Target for Immuno-Oncology Therapy. *Mol.* dsRNA *Cell*. Cell Press March 7, 2019, pp 866–868. https://doi.org/10.1016/j.molcel.2019.02.021.

(19) 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.

(20) Mendoza, H. G.; Jauregui-Matos, V.; Park, S.; Pham, K. M.; Beal, P. A. Selective Inhibition of ADAR1 Using 8-Azanebularine-Modified RNA Duplexes. *Biochemistry* **2023**, *62* (8), 1376–1387. https://doi.org/10.1021/acs.biochem.2c00686.

(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) Khosravi, H. M.; Jantsch, M. F. Site-Directed RNA Editing: Recent Advances and Open Challenges. *RNA Biol.* **2021**, *18* (sup1), 41–50. https://doi.org/10.1080/15476286.2021.1983288.

(23) 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.

(24) 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.

(25) Booth, B. J.; Nourreddine, S.; Katrekar, D.; Savva, Y.; Bose, D.; Long, T. J.; Huss, D. J.; Mali, P. RNA Editing: Expanding the Potential of RNA Therapeutics. *Mol. Ther.* **2023**, *31* (6), 1533–1549. https://doi.org/10.1016/j.ymthe.2023.01.005.

(26) Diaz Quiroz, J. F.; Ojha, N.; Shayhidin, E. E.; De Silva, D.; Dabney, J.; Lancaster, A.; Coull, J.; Milstein, S.; Fraley, A. W.; Brown, C. R.; Rosenthal, J. J. C. Development of a Selection Assay for Small Guide RNAs That Drive Efficient Site-Directed RNA Editing. *Nucleic Acid Res.* **2023**, *51* (7), e41–e41. https://doi.org/10.1093/nar/gkad098.

(27) 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.* **2023**, *4* (1), 74–83. https://doi.org/10.1039/D2CB00165A.

(28) 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.

(29) 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.

(30) Ojha, N.; Diaz Quiroz, J. F.; Rosenthal, J. J. C. In Vitro and in Cellula Site-Directed RNA Editing Using the ΛNDD-BoxB System. In *Method. Enzymol.*; 2021; Vol. 658, pp 335–358. https://doi.org/10.1016/bs.mie.2021.06.009.
(31) Sinnamon, J. R.; Kim, S. Y.; Corson, G. M.; Song, Z.; Nakai, H.; Adelman, J. P.; Mandel, G. Site-Directed RNA Repair of Endogenous Mecp2 RNA in Neurons. *Proc. Natl. Acad. Sci. U.S.A* **2017**, *114* (44). https://doi.org/10.1073/pnas.1715320114.
(32) 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.
(33) 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.
(34) 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.
(35) Véliz, E. A.; Easterwood, L. 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.
(36) 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.
(37) 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.
(38) 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.
(39) Hoang, C.; Ferré-D'Amaré, A. R. Cocrystal Structure of a TRNA 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.
(40) Kee Wong, S.; Sato, S.; Lazinski, D. W. *Substrate Recognition by ADAR1 and ADAR2*; 2001.
(41) Jauregui-Matos, V.; Jacobs, O.; Ouye, R.; Mozumder, S.; Salvador, P. J.; Fink, K. D.; Beal, P. A. Site-Specific Regulation of RNA Editing with Ribose-Modified Nucleoside Analogs in ADAR Guide Strands. *Nucleic Acids Res.* **2024**, *52* (12), 6733–6747. https://doi.org/10.1093/nar/gkae461.
(42) Pullirsch, D.; Jantsch, M. F. Proteome Diversification by Adenosine to Inosine RNA-Editing. *RNA Biol.* **2010**, *7* (2), 205–212. https://doi.org/10.4161/rna.7.2.11286.
(43) Yang, Y.; Okada, S.; Sakurai, M. Adenosine-to-Inosine RNA Editing in Neurological Development and Disease. *RNA Biol.* **2021**, *18* (7), 999–1013. https://doi.org/10.1080/15476286.2020.1867797.
(44) Montiel-Gonzalez, M. F.; Diaz Quiroz, J. F.; Rosenthal, J. J. C. Current Strategies for Site-Directed RNA Editing Using ADARs. *Methods* **2019**, *156*, 16–24. https://doi.org/10.1016/j.ymeth.2018.11.016.
(45) 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.
(46) 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.

(47) Vogel, P.; Schneider, M. F.; Wettengel, J.; Stafforst, T. Improving Site-Directed RNA Editing In Vitro and in Cell Culture by Chemical Modification of the GuideRNA. *Angew. Chem., Int. Ed.* **2014**, *53* (24), 6267–6271. https://doi.org/10.1002/anie.201402634.

(48) Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. New Description Using the Concept of Pseudorotation. *J. Am. Chem. Soc.***1972**, *94* (23), 8205–8212. https://doi.org/10.1021/ja00778a043.

(49) Hagedorn, P. H.; Persson, R.; Funder, E. D.; Albæk, N.; Diemer, S. L.; Hansen, D. J.; Møller, M. R.; Papargyri, N.; Christiansen, H.; Hansen, B. R.; Hansen, H. F.; Jensen, M. A.; Koch, T. Locked Nucleic Acid: Modality, Diversity, and Drug Discovery. *Drug Discov.Today* **2018**, *23* (1), 101–114. https://doi.org/10.1016/j.drudis.2017.09.018.

(50) Ferrando, I. M.; Chaerkady, R.; Zhong, J.; Molina, H.; Jacob, H. K. C.; Herbst-Robinson, K.; Dancy, B. M.; Katju, V.; Bose, R.; Zhang, J.; Pandey, A.; Cole, P. A. Identification of Targets of C-Src Tyrosine Kinase by Chemical Complementation and Phosphoproteomics. *Mol. Cell. Proteom.* **2012**, *11* (8), 355–369. https://doi.org/10.1074/mcp.M111.015750.

(51) Xu, W.; Doshi, A.; Lei, M.; Eck, M. J.; Harrison, S. C. Crystal Structures of C-Src Reveal Features of Its Autoinhibitory Mechanism. *Mol. Cell* **1999**, *3* (5), 629–638. https://doi.org/10.1016/S1097-2765(00)80356-1.

(52) Campbell, M. A.; Wengel, J. Locked vs. Unlocked Nucleic Acids (LNA vs. UNA): Contrasting Structures Work towards Common Therapeutic Goals. *Chem. Soc. Rev.* **2011**, *40* (12), 5680. https://doi.org/10.1039/c1cs15048k.

(53) Waga, T.; Nishizaki, T.; Miyakawa, I.; Ohrui, H.; Meguro, H. Synthesis of 4′- *C* -Methylnucleosides. *Biosci. Biotechnol. Biochem.* **1993**, *57* (9), 1433–1438. https://doi.org/10.1271/bbb.57.1433.

(54) Waga, T.; Ohrui, H.; Meguro, H. Synthesis and Biological Evaluation of 4′- *C* -Methyl Nucleosides. *Nucleosides Nucleotides* **1996**, *15* (1–3), 287–304. https://doi.org/10.1080/07328319608002385.

(55) Koizumi, K.; Maeda, Y.; Kano, T.; Yoshida, H.; Sakamoto, T.; Yamagishi, K.; Ueno, Y. Synthesis of 4′-C-Aminoalkyl-2′-O-Methyl Modified RNA and Their Biological Properties. *Bioorg. Med. Chem.* **2018**, *26* (12), 3521–3534. https://doi.org/10.1016/j.bmc.2018.05.025.

(56) Kano, T.; Katsuragi, Y.; Maeda, Y.; Ueno, Y. Synthesis and Properties of 4′-C-Aminoalkyl-2′-Fluoro-Modified RNA Oligomers. *Bioorg. Med. Chem.* **2018**, *26* (15), 4574–4582. https://doi.org/10.1016/j.bmc.2018.08.001.

(57) Guo, F.; Yue, Z.; Trajkovski, M.; Zhou, X.; Cao, D.; Li, Q.; Wang, B.; Wen, X.; Plavec, J.; Peng, Q.; Xi, Z.; Zhou, C. Effect of Ribose Conformation on RNA Cleavage via Internal Transesterification. *J. Am. Chem. Soc.* **2018**, *140* (38), 11893–11897. https://doi.org/10.1021/jacs.8b06313.

(58) Rangam, G.; Rudinger, N. Z.; Müller, H. M.; Marx, A. Synthesis and Application of 4′-C-Alkylated Uridines as Probes for Uracil-DNA Glycosylase. *Synthesis (Stuttg)* **2005**, No. 9, 1467–1472. https://doi.org/10.1055/s-2005-865328.

(59) Nacro, K.; Lee, J.; Barchi, J. J.; Lewin, N. E.; Blumberg, P. M.; Marquez, V. E. Conformationally Constrained Analogues of Diacylglycerol (DAG). Part 19: Asymmetric Syntheses of (3R)- and (3S)-3-Hydroxy-4,4-Disubstituted Heptono-1,4-Lactones as Protein Kinase C (PK-C) Ligands with Increased Hydrophilicity. *Tetrahedron* **2002**, *58* (26), 5335–5345. https://doi.org/10.1016/S0040-4020(02)00477-5.

(60) Tranová, L.; Stýskala, J. Study of the N^7 Regioselective Glycosylation of 6-Chloropurine and 2,6-Dichloropurine with Tin and Titanium Tetrachloride. *J. Org. Chem.* **2021**, *86* (19), 13265–13275. https://doi.org/10.1021/acs.joc.1c01186.

(61) Vorbrüggen, H.; Höfle, G. Nucleoside Syntheses, XXIII [1] On the Mechanism of Nucleoside Synthesis. *Chem. Ber.* **1981**, *114* (4), 1256–1268. https://doi.org/10.1002/cber.19811140405.

(62) Somoza, Á. Protecting Groups for RNA Synthesis: An Increasing Need for Selective Preparative Methods. *Chem. Soc. Rev.* **2008**, *37* (12), 2668. https://doi.org/10.1039/b809851d.
(63) Wang, D.; Shukla, C.; Liu, X.; Schoeb, T. R.; Clarke, L. A.; Bedwell, D. M.; Keeling, K. M. Characterization of an MPS I-H Knock-in Mouse That Carries a Nonsense Mutation Analogous to the Human IDUA-W402X Mutation. *Mol. Genet. Metab.* **2010**, *99* (1), 62–71. https://doi.org/10.1016/j.ymgme.2009.08.002.
(64) Vogel, P.; Moschref, M.; Li, Q.; Merkle, T.; Selvasaravanan, K. D.; Li, J. B.; Stafforst, T. Efficient and Precise Editing of Endogenous Transcripts with SNAP-Tagged ADARs. *Nat. Methods* **2018**, *15* (7), 535–538. https://doi.org/10.1038/s41592-018-0017-z.
(65) 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.
(66) Ka, S.; Cm, P.; Swingle, M. R.; A, M.; C, L.; Ad, C.; Re, H.; An, K. Quantitative Proteomics and Phosphoproteomics of PPP2R5D Variants Reveal Deregulation of RPS6 Phosphorylation through Converging Signaling Cascades. *bioRxiv* **2023**. https://doi.org/10.1101/2023.03.27.534397.
(67) Wettengel, J.; Reautschnig, P.; Geisler, S.; Kahle, P. J.; Stafforst, T. Harnessing Human ADAR2 for RNA Repair - Recoding a PINK1 Mutation Rescues Mitophagy. *Nucleic Acids Res.* **2017**, *45* (5), 2797–2808. https://doi.org/10.1093/nar/gkw911.
(68) Al Musaimi, O.; Al Shaer, D.; Albericio, F.; de la Torre, B. G. 2022 FDA TIDES (Peptides and Oligonucleotides) Harvest. *Pharmaceuticals* **2023**, *16* (3), 336. https://doi.org/10.3390/ph16030336.
(69) Egli, M.; Manoharan, M. Chemistry, Structure and Function of Approved Oligonucleotide Therapeutics. *Nucleic Acid Res.* **2023**, *51* (6), 2529–2573. https://doi.org/10.1093/nar/gkad067.
(70) Wave Life Sciences USA, Inc. *Wave Life Sciences Announces Submission of First Clinical Trial Application for WVE-006, the First-Ever RNA Editing Clinical Candidate, and Plans for Upcoming Virtual "R&D Day.";* Cambridge, MA, 2023.
(71) Sheridan, C. Shoot the Messenger: RNA Editing Is Here. *Nat. Biotechnol.* **2023**, *41* (3), 306–308. https://doi.org/10.1038/s41587-023-01709-8.
(72) Wengel, J.; Koshkin, A.; Singh, S. K.; Nielsen, P.; Meldgaard, M.; Rajwanshi, V. K.; Kumar, R.; Skouv, J.; Nielsen, C. B.; Jacobsen, J. P.; Jacobsen, N.; Olsen, C. E. Lna (Locked Nucleic Acid). *Nucleosides Nucleotides* **1999**, *18* (6–7), 1365–1370. https://doi.org/10.1080/07328319908044718.
(73) Langkjær, N.; Pasternak, A.; Wengel, J. UNA (Unlocked Nucleic Acid): A Flexible RNA Mimic That Allows Engineering of Nucleic Acid Duplex Stability. *Bioorg. Med. Chem.* **2009**, *17* (15), 5420–5425. https://doi.org/10.1016/j.bmc.2009.06.045.
(74) Petersen, M.; Bondensgaard, K.; Wengel, J.; Jacobsen, J. P. Locked Nucleic Acid (LNA) Recognition of RNA: NMR Solution Structures of LNA:RNA Hybrids. *J. Am. Med. Chem.* **2002**, *124* (21), 5974–5982. https://doi.org/10.1021/ja012288d.
(75) Reautschnig, P.; Wahn, N.; Wettengel, J.; Schulz, A. E.; Latifi, N.; Vogel, P.; Kang, T.-W.; Pfeiffer, L. S.; Zarges, C.; Naumann, U.; Zender, L.; Li, J. B.; Stafforst, T. CLUSTER Guide RNAs Enable Precise and Efficient RNA Editing with Endogenous ADAR Enzymes in Vivo. *Nat. Biotechnol.* **2022**, *40* (5), 759–768. https://doi.org/10.1038/s41587-021-01105-0.
(76) Abudayyeh, O. O.; Gootenberg, J. S.; Essletzbichler, P.; Han, S.; Joung, J.; Belanto, J. J.; Verdine, V.; Cox, D. B. T.; Kellner, M. J.; Regev, A.; Lander, E. S.; Voytas, D. F.; Ting, A. Y.; Zhang, F. RNA Targeting with CRISPR–Cas13. *Nature* **2017**, *550* (7675), 280–284. https://doi.org/10.1038/nature24049.
(77) Zhou, S.; Kern, E. R.; Gullen, E.; Cheng, Y.-C.; Drach, J. C.; Tamiya, S.; Mitsuya, H.; Zemlicka, J. 9-{[3-Fluoro-2-(Hydroxymethyl)Cyclopropylidene]Methyl}adenines and -Guanines. Synthesis and

Antiviral Activity of All Stereoisomers. *J. Med. Chem.* **2006**, *49* (20), 6120–6128. https://doi.org/10.1021/jm0607404.

(78) Macbeth, M. R.; Bass, B. L. Large-Scale Overexpression and Purification of ADARs from Saccharomyces Cerevisiae for Biophysical and Biochemical Studies. In *Method Enzymol.;* Gott. M. Jonatha, Ed.; Elsevier, 2007; Vol. 424, pp 319–331. https://doi.org/10.1016/S0076-6879(07)24015-7.

(79) 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**, *18* (10), 2188–2199. https://doi.org/10.1021/acschembio.3c00107.

(80) 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.

(81) 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.

(82) Wang, Y.; Park, S.; 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.

(83) 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.

(84) 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.

(85) 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. *Brain Res. Mol.* **2004**, *124* (1), 70–78. https://doi.org/https://doi.org/10.1016/j.molbrainres.2004.02.010.

(86) Eggington, J. M.; Greene, T.; Bass, B. L. Predicting Sites of ADAR Editing in Double-Stranded RNA. *Nat. Commun.* **2011**, *2* (1), 319. https://doi.org/10.1038/ncomms1324.

(87) Park, S.; Doherty, E. E.; Xie, Y.; Padyana, A. K.; Fang, F.; Zhang, Y.; Karki, A.; Lebrilla, C. B.; Siegel, J. B.; Beal, P. A. High-Throughput Mutagenesis Reveals Unique Structural Features of Human ADAR1. *Nat. Commun.* **2020**, *11* (1), 5130. https://doi.org/10.1038/s41467-020-18862-2.

(88) Whitehead, K. A.; Langer, R.; Anderson, D. G. Knocking down Barriers: Advances in SiRNA Delivery. *Nat. Rev. Drug. Discov.* **2009**, *8* (2), 129–138. https://doi.org/10.1038/nrd2742.

(89) Pujari, S. S.; Seela, F. Cross-Linked DNA: Propargylated Ribonucleosides as “Click” Ligation Sites for Bifunctional Azides. *J. Org. Chem.* **2012**, *77* (9), 4460–4465. https://doi.org/10.1021/jo300421p.

(90) 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.

(91) 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.

(92) Meldal, M.; Tornøe, C. W. Cu-Catalyzed Azide−Alkyne Cycloaddition. *Chem. Rev.* **2008**, *108* (8), 2952–3015. https://doi.org/10.1021/cr0783479.

(93) Shewach, D. S.; Krawczyk, S. H.; Acevedo, O. L.; Townsend, L. B. Rapid Communication: Inhibition of Adenosine Deaminase by Azapurine Ribonucleosides. *Biochem. Pharmacol.* **1992**, *44* (9), 1697–1700. https://doi.org/https://doi.org/10.1016/0006-2952(92)90061-M.

(94) 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.

(95) Macbeth, M. R.; Bass, B. L. Chapter 15 - Large-Scale Overexpression and Purification of ADARs from Saccharomyces Cerevisiae for Biophysical and Biochemical Studies. In *Methods Enzymol.*; Gott, J. M., Ed.; Academic Press, **2007**; Vol. 424, pp 319–331. https://doi.org/https://doi.org/10.1016/S0076-6879(07)24015-7.

(96) 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), e87–e87. https://doi.org/10.1093/nar/gku272.

(97) 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.

(98) Bereiter, R.; Himmelstoß, M.; Renard, E.; Mairhofer, E.; Egger, M.; Breuker, K.; Kreutz, C.; Ennifar, E.; Micura, R. Impact of 3-Deazapurine Nucleobases on RNA Properties. *Nucleic Acids Res.* **2021**, *49* (8), 4281–4293. https://doi.org/10.1093/nar/gkab256.

(99) Harkins, T. R.; Freiser, H. Adenine-Metal Complexes1,2. *J. Am. Chem. Soc.* **1958**, *80* (5), 1132–1135. https://doi.org/10.1021/ja01538a030.

(100) Mlotkowski, A. J.; Schlegel, H. B.; Chow, C. S. Calculated PKa Values for a Series of Aza- and Deaza-Modified Nucleobases. *J. Phys. Chem.* **2023**, *127* (15), 3526–3534. https://doi.org/10.1021/acs.jpca.3c01358.

(101) Kabsch, W. Biological Crystallography. *research papers Acta Cryst.* **2010**, *66*, 125–132. https://doi.org/10.1107/S0907444909047337.

(102) Evans, P. R.; Murshudov, G. N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr. D. Biol. Crystallogr.* **2013**, *69* (7), 1204–1214. https://doi.org/10.1107/S0907444913000061.

(103) 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.

(104) Ma, C.-C.; Wang, Z.-L.; Xu, T.; He, Z.-Y.; Wei, Y.-Q. The Approved Gene Therapy Drugs Worldwide: From 1998 to 2019. *Biotechnol. Adv.* **2020**, *40*, 107502. https://doi.org/https://doi.org/10.1016/j.biotechadv.2019.107502.

(105) Porto, E. M.; Komor, A. C.; Slaymaker, I. M.; Yeo, G. W. Base Editing: Advances and Therapeutic Opportunities. *Nat. Rev. Drug. Discov.* **2020**, *19* (12), 839–859. https://doi.org/10.1038/s41573-020-0084-6.

(106) Doudna, J. A.; Charpentier, E. The New Frontier of Genome Engineering with CRISPR-Cas9. *Science (1979)* **2014**, *346* (6213), 1258096. https://doi.org/10.1126/science.1258096.

(107) Anzalone, A. V; Gao, X. D.; Podracky, C. J.; Nelson, A. T.; Koblan, L. W.; Raguram, A.; Levy, J. M.; Mercer, J. A. M.; Liu, D. R. Programmable Deletion, Replacement, Integration and Inversion of Large DNA Sequences with Twin Prime Editing. *Nat. Biotechnol.* **2022**, *40* (5), 731–740. https://doi.org/10.1038/s41587-021-01133-w.

(108) Adikusuma, F.; Lushington, C.; Arudkumar, J.; Godahewa, G. I.; Chey, Y. C. J.; Gierus, L.; Piltz, S.; Geiger, A.; Jain, Y.; Reti, D.; Wilson, L. O. W.; Bauer, D. C.; Thomas, P. Q. Optimized Nickase- and Nuclease-Based Prime Editing in Human and Mouse Cells. *Nucleic Acids Res.* **2021**, *49* (18), 10785–10795. https://doi.org/10.1093/nar/gkab792.

(109) Chen, P. J.; Hussmann, J. A.; Yan, J.; Knipping, F.; Ravisankar, P.; Chen, P.-F.; Chen, C.; Nelson, J. W.; Newby, G. A.; Sahin, M.; Osborn, M. J.; Weissman, J. S.; Adamson, B.; Liu, D. R. Enhanced Prime Editing Systems by Manipulating Cellular Determinants of Editing Outcomes. *Cell* **2021**, *184* (22), 5635-5652.e29. https://doi.org/10.1016/j.cell.2021.09.018.

(110) Jang, H.; Jo, D. H.; Cho, C. S.; Shin, J. H.; Seo, J. H.; Yu, G.; Gopalappa, R.; Kim, D.; Cho, S.-R.; Kim, J. H.; Kim, H. H. Application of Prime Editing to the Correction of Mutations and Phenotypes in Adult Mice with Liver and Eye Diseases. *Nat. Biomed. Eng.* **2022**, *6* (2), 181–194. https://doi.org/10.1038/s41551-021-00788-9.

(111) Liu, P.; Liang, S.-Q.; Zheng, C.; Mintzer, E.; Zhao, Y. G.; Ponnienselvan, K.; Mir, A.; Sontheimer, E. J.; Gao, G.; Flotte, T. R.; Wolfe, S. A.; Xue, W. Improved Prime Editors Enable Pathogenic Allele Correction and Cancer Modelling in Adult Mice. *Nat. Commun.* **2021**, *12* (1), 2121. https://doi.org/10.1038/s41467-021-22295-w.

(112) Anzalone, A. V; Randolph, P. B.; Davis, J. R.; Sousa, A. A.; Koblan, L. W.; Levy, J. M.; Chen, P. J.; Wilson, C.; Newby, G. A.; Raguram, A.; Liu, D. R. Search-and-Replace Genome Editing without Double-Strand Breaks or Donor DNA. *Nature* **2019**, *576* (7785), 149–157. https://doi.org/10.1038/s41586-019-1711-4.

(113) Nishimasu, H.; Ran, F. A.; Hsu, P. D.; Konermann, S.; Shehata, S. I.; Dohmae, N.; Ishitani, R.; Zhang, F.; Nureki, O. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **2014**, *156* (5), 935–949. https://doi.org/10.1016/j.cell.2014.02.001.

(114) Ran, F. A.; Hsu, P. D.; Lin, C.-Y.; Gootenberg, J. S.; Konermann, S.; Trevino, A. E.; Scott, D. A.; Inoue, A.; Matoba, S.; Zhang, Y.; Zhang, F. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **2013**, *154* (6), 1380–1389. https://doi.org/10.1016/j.cell.2013.08.021.

(115) Shuto, Y.; Nakagawa, R.; Zhu, S.; Hoki, M.; Omura, S. N.; Hirano, H.; Itoh, Y.; Zhang, F.; Nureki, O. Structural Basis for PegRNA-Guided Reverse Transcription by a Prime Editor. *Nature* **2024**, *631* (8019), 224–231. https://doi.org/10.1038/s41586-024-07497-8.

(116) Erlanson, D. A.; Wolfe, S. A.; Chen, L.; Verdine, G. L. Selective Base-Pair Destabilization Enhances Binding of a DNA Methyltransferase. *Tetrahedron* **1997**, *53* (35), 12041–12056. https://doi.org/https://doi.org/10.1016/S0040-4020(97)00767-9.

(117) Verdine, G. L.; Norman, D. P. G. Covalent Trapping of Protein-DNA Complexes. *Annu. Rev. Biochem.* **2003**, *72* (Volume 72, 2003), 337–366. https://doi.org/https://doi.org/10.1146/annurev.biochem.72.121801.161447.

(118) Allerson, C. R.; Chen, S. L.; Verdine, G. L. A Chemical Method for Site-Specific Modification of RNA: The Convertible Nucleoside Approach. *J. Am. Chem. Soc.* **1997**, *119* (32), 7423–7433. https://doi.org/10.1021/ja962858n.

(119) Peacock, H.; Bachu, R.; Beal, P. A. Covalent Stabilization of a Small Molecule–RNA Complex. *Bioorg. Med. Chem. Lett.* **2011**, *21* (17), 5002–5005. https://doi.org/https://doi.org/10.1016/j.bmcl.2011.04.136.

(120) Tran, K.; Arkin, M. R.; Beal, P. A. Tethering in RNA: An RNA-Binding Fragment Discovery Tool. *Molecules* **2015**, *20* (3), 4148–4161. https://doi.org/10.3390/molecules20034148.

(121) Xu, Y. Z.; Zheng, Q.; Swann, P. F. Synthesis of DNA Containing Modified Bases by Post-Synthetic Substitution. Synthesis of Oligomers Containing 4-Substituted Thymine: O4-Alkylthymine, 5-Methylcytosine, N4-Dimethylamino-5-Methylcytosine, and 4-Thiothymine. *J. Org. Chem.* **1992**, *57* (14), 3839–3845. https://doi.org/10.1021/jo00040a024.

(122) Jacobo-Molina, A.; Ding, J.; Nanni, R. G.; Clark, A. D.; Lu, X.; Tantillo, C.; Williams, R. L.; Kamer, G.; Ferris, A. L.; Clark, P.; Hizi, A.; Hughes, S. H.; Arnold, E. Crystal Structure of Human Immunodeficiency Virus Type 1 Reverse Transcriptase Complexed with Double-Stranded DNA at 3.0 A Resolution Shows Bent DNA. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90* (13), 6320–6324. https://doi.org/10.1073/pnas.90.13.6320.