

BIOCHEMICAL AND STRUCTURAL INSIGHTS INTO ADAR1 RNA EDITING

by

Beata Maria Kaczmarek

October, 2024

*A thesis submitted to the
Graduate School
of the
Institute of Science and Technology Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy*

Committee in charge:

Carrie Bernecky, Supervisor

Michael Jantsch

Martin Loose

Mario de Bono, Chair



The thesis of Beata Maria Kaczmarek, titled *Biochemical and structural insights into ADAR1 RNA editing*, is approved by:

Supervisor: Carrie Bernecky, ISTA, Klosterneuburg, Austria

Signature: _____

Committee Member: Martin Loose, ISTA, Klosterneuburg, Austria

Signature: _____

Committee Member: Michael Jantsch, Medical University of Vienna, Vienna, Austria

Signature: _____

Defense chair: Mario de Bono, ISTA, Klosterneuburg, Austria

Signature: _____

Signed page is on file

© by Beata Maria Kaczmarek, October 29, 2024

Some Rights Reserved

CC BY 4.0 The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). Under this license, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that you credit the author.

ISTA Thesis, ISSN: 2663-337X

ISBN: 978-3-99078-045-9

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I accept full responsibility for the content and factual accuracy of this work, including the data and their analysis and presentation, and the text and citation of other work.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: _____

Beata Maria Kaczmarek

October, 2024

Signed page is on file

Abstract

ADAR1 is broadly expressed across various tissues and is vital in regulating pathways associated with innate immune responses. ADAR1 marks double-stranded RNA as "self" through its A-to-I editing activity, effectively repressing autoimmunity and maintaining immune tolerance. This editing process has been detected at millions of sites across the human genome. However, the mechanism underlying ADAR1's substrate selectivity properties remains largely unclear, with much of the current knowledge derived from comparisons to its more extensively studied homolog, ADAR2. By studying ADAR1 in complex with its RNA substrates and applying a combination of biochemical techniques and structural studies using CryoEM, we aim to gain a more comprehensive understanding of the substrate selectivity characteristics of ADAR1.

In this thesis, the purification protocol for ADAR1 was successfully optimized, resulting in the first report in the literature to achieve high protein purity and activity. This advancement enabled the investigation of complex formation between ADAR1 and various RNA substrates, leading to the identification of optimal conditions for preparing the cryoEM sample. However, despite comprehensive optimization of the cryo-EM conditions, the resulting data lacked the desired quality, highlighting the need for similar rigorous optimization of the RNA substrates to facilitate structural studies of the ADAR1-RNA complex. The study was complemented by AlphaFold predictions, which provided some insights into this mechanism.

Moreover, during this project I established a collaboration with a research group focused on studying ADAR homologs. Notably ADAR homologs were identified in bivalve species, and it was further demonstrated that ADAR and its A-to-I editing activity are upregulated in Pacific oysters during infections with Ostreid herpesvirus-1—a highly infectious virus that leads to significant losses in oyster populations globally. I successfully purified oyster ADAR and prepared *in vitro* edited RNA for nanopore sequencing—a direct sequencing technology capable of detecting modified nucleotides without the need for reverse transcription. The collaborators initiated optimization of this nanopore-based approach. However, current technological limitations still constrain the reliable detection of modified nucleotides.

The project also examined the impact of RNA editing on RNA binding and filament formation by MDA5, a key cytosolic dsRNA sensor that triggers an interferon response. A primary target of ADAR1's editing activity is RNA derived from repetitive elements present in the genome, particularly Alu elements forming double-stranded RNA. When unedited, these RNA sequences are recognized by MDA5. However, the mechanisms by which MDA5 interacts with Alu RNAs, as well as the role of A-to-I editing in influencing this binding, are still not well understood.

The interaction between MDA5 and Alu elements, was successfully established. This was achieved through the testing of different RNA variants and the evaluation of filament formation using binding techniques and electron microscopy imaging. This groundwork has set the conditions for further evaluation using CryoEM. Furthermore, the effects of A-to-I editing on the binding properties of MDA5 with Alu RNA were investigated. Given the recent research that has provided new insights into MDA5's interaction with dsRNA, it is essential to revise the experimental setup to integrate these findings before moving forward with the CryoEM sample analysis.

Acknowledgments

I would like to express my gratitude to my PhD supervisor, Carrie. Thank you for your guidance, for contributing valuable ideas to the project, and for introducing me to the fields of RNA biology, biochemistry, and structural biology. I greatly value the knowledge and experience you shared with me. I'm also thankful for the supportive work environment you created and your understanding during the challenges I encountered, both within and outside the workplace. Your optimism and encouragement were essential in helping me remain persistent throughout the different stages of the project.

I would also like to thank my PhD committee for their support during the progress review meetings and for contributing new ideas. Martin, I value your expertise in biochemistry and your insightful feedback during these meetings. I am also grateful for the opportunity you gave me as a student intern, which ultimately led to the start of my PhD journey. Michael, I truly appreciate your extensive knowledge of the in vivo aspects of ADAR1 biology and your insights into the MDA5 project, which have been very helpful.

Thank you to all past and present members of the Bernecky group for the valuable feedback you provided during the lab meetings and for creating a positive working atmosphere. I especially want to thank Anita for maintaining a well-organized lab environment and preparing the commonly used reagents for our experiments. Katarina, thank you for your support during the cryoEM sample preparation and data processing phases. David, thank you for the insightful discussions that provided valuable troubleshooting tips. Nora, as an intern, thank you for your help during the optimization of the ADAR1 purification protocol.

I acknowledge the support provided by ISTA infrastructure, specifically the Life Support Facility and the Media Kitchen team. Your support in managing the glassware and preparing various reagents was essential for the experimental phase of this project. Additionally, I would like to thank the Electron Microscope Facility, especially Valentin and Carina, for your assistance during the CryoEM data acquisition process.

I'm genuinely thankful for the support I received from my friends: Anita, Karola, Kristina, Nasim, Victor, Julia, Fleur, and Zane. Each of you has contributed in your unique way, and I cherish the wonderful moments we shared during this time.

Finishing my PhD journey would not have been possible without the invaluable support from my partner, Frederic. You were accompanying me every step of the way providing constant encouragement. Our scientific discussions were invaluable and inspired me with some new ideas.

I'm deeply grateful to my family. Even though you didn't always fully understand the complexities of my work, your support was always there for me.

About the Author

Beata Kaczmarek completed her BSc in Biotechnology at the University of Wrocław, where she also obtained her MSc in Biotechnology specializing in Proteins and Peptides. In 2018, she joined ISTA as a scientific intern and then transitioned to a PhD student. Under the guidance of Carrie Bernecky, she focused her research on exploring different aspects of RNA biology related to innate immunity, using biochemistry and structural biology approaches. During her PhD, she actively presented her findings to the RNA Deco community, participated in conferences, established valuable collaborations, and contributed to various publications.

List of Collaborators and publications

Collaborators

Enrico Bortoletto, Umberto Rosani, Paola Veroni; Department of Biology at the University of Padova

The collaborators identified an ADAR variant that was induced in *Crassostrea gigas* infected with Ostreid herpesvirus-1 (OsHV-1) and subsequently optimized and analyzed nanopore sequencing data (Pacific oyster ADAR chapters: 4.6.1 and 4.6.3)

Publications

Bezeljak, U., Loya, H., **Kaczmarek, B.**, Saunders, T.E., Loose, M., 2020. Stochastic activation and bistability in a Rab GTPase regulatory network. *Proceedings of the National Academy of Sciences*. 117, 6540–6549

Tlučková, K.*, **Kaczmarek, B.***, Salmazo, A., Bernecky, C. Mechanism of mammalian transcriptional repression by noncoding RNA. *Nature Structural & Molecular Biology*. Accepted manuscript NSMB-BC48182

*These authors contributed equally.

Table of Contents

Abstract	<i>I</i>
Acknowledgments.....	<i>II</i>
About the Author	<i>III</i>
List of Collaborators and publications	<i>IV</i>
Table of Contents	<i>V</i>
List of Figures	<i>VIII</i>
List of Tables	<i>X</i>
List of Abbreviations.....	<i>XI</i>
1 Introduction.....	1
1.1 Innate immunity and nucleic acids sensing.....	1
1.2 RNA modifications suppressing the innate immune response.....	2
1.2.1 A-to-I RNA editing	3
1.3 ADAR proteins in humans	6
1.3.1 ADAR1	6
1.3.2 ADAR2	7
1.3.3 Structural characterization of ADAR enzymes.....	8
1.3.4 Characteristics of RNA substrates edited by ADARs.....	13
1.4 dsRNA sensing pathways regulated by ADAR1	15
1.4.1 MDA5 dsRNA recognition and activation	17
1.4.2 Endogenous dsRNAs recognized by MDA5.....	18
1.5 ADAR1 implication in human diseases.....	19
1.5.1 Autoimmune disorders	19
1.5.2 ADAR1 role in cancers.....	21
1.6 ADAR1 evolution	21
1.6.1 ADAR in Pacific oysters	22
2 Aims and objectives	23
2.1 Biochemical and structural insights into ADAR1 substrate selectivity properties....	23
2.2 Investigating the in vitro editing activity of ADAR in Pacific oysters	25
2.3 MDA5 interaction with its endogenous target - irAlu repeats	25
3 Materials and methods	27
3.1 Molecular cloning	27
3.1.1 DNA agarose gel.....	27
3.1.2 InFusion cloning	27
3.1.3 Site-directed mutagenesis	27
3.1.4 Produced plasmids.....	29
3.2 Protein purification methods	30
3.2.1 Protein expression	30
3.2.2 Protein analysis.....	30
3.2.3 Optimization of ADAR1 purification protocol.....	31
3.2.4 ADAR2 purification protocol.....	33
3.2.5 ADAR1 / ADAR2 deaminases purification protocol	33
3.2.6 Pacific oyster ADAR purification protocol	34
3.2.7 MDA5 Δ CARDs purification protocol	34
3.2.8 Testing the activity of various ADAR proteins	34

3.2.9	Checking proteins homogeneity using mass photometry	35
3.3	RNA substrates preparation	35
3.3.1	In vitro transcription, RNA purification and folding	35
3.3.2	Urea-acrylamide denaturing gel	36
3.3.3	RNA fluorescent labeling	36
3.4	ADAR1-RNA complex preparation and characterization	36
3.4.1	Checking complex formation with Native-PAGE	36
3.4.2	Measuring binding affinity using fluorescence anisotropy	36
3.4.3	Complex crosslinking	37
3.5	CryoEM sample preparation	37
3.5.1	Preparing continuously coated carbon grids	37
3.5.2	Preparing graphene-oxide-coated grids	38
3.5.3	Negative staining	38
3.5.4	Sample vitrification	38
3.5.5	CryoEM grid screening and data collection	38
3.6	CryoEM data processing and analysis	39
3.6.1	Data processing	39
3.6.2	Structure prediction using AlphaFold	39
4	Results	40
4.1	ADAR1 purification protocol optimization	40
4.1.1	Choosing the optimal construct for ADAR1 p110 purification	40
4.1.2	Evaluating protein solubility through buffer screening	41
4.1.3	Optimizing protein expression in mammalian cells	42
4.1.4	Designing the optimal ADAR1 p150 construct	42
4.1.5	Issues faced during purification protocol optimization	43
4.1.6	Optimized ADAR1 purification protocol	45
4.2	ADAR1/2 deaminases purification	47
4.3	RNA substrates employed in complex characterization	47
4.4	ADAR - RNA complex characterization	48
4.4.1	Testing ADAR1 enzymatic activity	48
4.4.2	Testing ADAR deaminases enzymatic activity	50
4.4.3	ADAR1-RNA complex formation	51
4.4.4	Measuring binding affinities using fluorescence anisotropy	52
4.5	ADAR1 - RNA structure determination	53
4.5.1	Negative staining of ADAR1 - RNA complex	53
4.5.2	Preparation of cryoEM samples	53
4.5.3	Stabilizing the complex using crosslinkers	57
4.5.4	CryoEM analysis of SPB crosslinked complex	60
4.5.5	Reducing complex aggregation by using dimerization mutant	62
4.5.6	AlphaFold structure prediction	64
4.6	Pacific oyster ADAR	68
4.6.1	C. gigas ADAR protein identification and purification	68
4.6.2	Testing enzymatic activity of CgADAR1v using Sanger sequencing	69
4.6.3	Testing enzymatic activity of CgADAR1v using nanopore sequencing	70
4.7	MDA5 - irAlu complex characterization	72
4.7.1	MDA5 Δ CARDs purification	72
4.7.2	Preparing irAlu RNA substrates	73

4.7.3 Complex binding	74
4.8 Filament formation.....	75
4.8.1 CryoEM analysis	76
4.8.2 irAlu editing by ADAR enzymes.....	77
4.8.3 MDA5 ΔCARDs interaction with edited irAlu.....	78
5 Discussion.....	79
5.1 Biochemical and structural insights into ADAR1 substrate selectivity properties....	79
5.2 Investigating the in vitro editing activity of ADAR in Pacific oysters	83
5.3 MDA5 interaction with its endogenous target - irAlu repeats	84
References	88
Appendix data.....	102
List of generated RNAs.....	102
AlphaFold3 models	106

List of Figures

Figure 1 Type I interferon system.	2
Figure 2 RNA modification present on mRNA.	2
Figure 3 The conversion of adenosine to inosine catalyzed by ADAR.	4
Figure 4 A-to-I editing marks cellular dsRNA as “self” preventing activation of innate immune responses.	5
Figure 5 ADAR1 and ADAR2 domains organization.	6
Figure 6 Organization of ADAR1 protein and genomic domains.	6
Figure 7 Structure of ADAR2 E488Q bound to an RNA substrate, illustrating the deamination catalytic mechanism.	10
Figure 8 Structural data showing ADAR dimers.	12
Figure 9 Structures of ADAR1 domains.	13
Figure 10 RNA structural features recognized by ADARs.	15
Figure 11 dsRNA sensing pathways regulated by ADAR1.	17
Figure 12 RLRs receptors and their domain organization.	17
Figure 13 Model of MDA5 filament formation and signal activation.	18
Figure 14 AGS mutations mapped in ADAR1 deaminase model bound to dsRNA.	20
Figure 15 Buffers composition used in the protein solubility by thermal unfolding.	32
Figure 16 SDS-PAGE analysis following various affinity purification strategies.	40
Figure 17 Buffer solubility test.	41
Figure 18 Optimization of protein expression in mammalian cells.	42
Figure 19 Designing the optimal construct for ADAR1 p150 expression and purification.	43
Figure 20 Assessment of various tubes for tag removal using 3C protease.	44
Figure 21 Purification of ADAR1 p110 or 150 from Expi293F mammalian cells.	46
Figure 22 Purification of the ADAR1 or ADAR2 deaminase domains.	47
Figure 23 RNA substrates generated for CryoEM studies.	48
Figure 24 Editing patterns for ADAR1 and ADAR2.	50
Figure 25 . Comparison of editing patterns for full-length enzymes and isolated deaminase domains.	51
Figure 26 ADAR1 complex formation with different RNA substrates.	52
Figure 27 Fluorescence anisotropy data for the ADAR1 p110 - NEIL1 complex.	53
Figure 28 . Negative staining of the Δ 133 ADAR1 p150, both in isolation and in complex with NEIL1 RNA.	53
Figure 29 Scheme of the CryoEM grid.	54
Figure 30 CryoEM analysis of the ADAR1 p110 - RNA complex.	55
Figure 31 CryoEM analysis of the Δ 133 ADAR1 p150 - RNA complex prepared on ice.	56
Figure 32 CryoEM analysis of the Δ 133 ADAR1 p150-Gria2 R/G complex.	57
Figure 33 Crosslinking of Δ 133 ADAR1 p150 with RNA substrates using BS3.	58
Figure 34 Crosslinking of Δ 133 ADAR1 p150 with RNA substrates using glutaraldehyde.	59
Figure 35 Crosslinking of Δ 133 ADAR1 p150 - TTYH2 RNA complex using SPB.	60
Figure 36 CryoEM analysis of the Δ 133 ADAR1 p150 - TTYH2 RNA complex crosslinked with SPB.	62
Figure 37 Analysis of the ADAR1 dsRBD3 dimer mutant.	63
Figure 38 AlphaFold2 ADAR1 structure prediction.	64
Figure 39 AlphaFold3 Predictions of the ADAR - TTYH2 RNA Complex.	66
Figure 40 . Identification of ADAR variant in <i>C. gigas</i> induced in OsHV-1 antiviral response.	68

Figure 41 Purification of the CgADAR1v.	69
Figure 42 Comparison of editing patterns for cgADAR1v and human ADARs.....	70
Figure 43 Nanopore sequencing results.	71
Figure 44 MDA5 Δ CARDs purification.	73
Figure 45 Generating irAlu RNA Substrates.	74
Figure 46 Binding of MDA5 Δ CARDs to various irAlu RNAs analyzed by EMSA.....	75
Figure 47 Negative staining analysis of the MDA5 Δ CARDs - irAlu RNAs filaments.	76
Figure 48 CryoEM analysis of the MDA5 Δ CARDs - irAlu RNAs filaments.	77
Figure 49 Editing of irAlu NICN1 FL by ADAR enzymes and isolated deaminase domains.....	78
Figure 50 Binding of MDA5 Δ CARDs with irAlu NICN1 A, non-modified or edited with ADAR2.	78
Figure 51 Models of ADAR - TTYH2 RNA complex including IDRs.	106

List of Tables

Table 1 List of used bacteria strains.....	28
Table 2 List of used vectors.....	29
Table 3 List of the accession number for protein-coding sequences.....	29
Table 4 List of produced plasmids.	30
Table 5 Buffer list used for the initial protein solubility screen.	32
Table 6 List of generated RNAs.....	104

List of Abbreviations

8-azaN	8-azanebularine
A-to-I	Adenosine to Inosine
ADAR	Adenosine Deaminase acting on double-stranded RNA
AGS	Aicardi-Goutières syndrome
AJUBA	RNA derived from Ajuba LIM Protein gene
BS3	bis(sulfosuccinimidyl)suberate
CARDs	caspase activation and recruitment domains
CDS	Coding sequence
Cg	Crassostrea gigas
cis-NATs	cis-natural antisense transcripts
CryoEM	Cryogenic electron microscopy
CTD	carboxy-terminal domains
dsRBD	Double-stranded RNA binding domain
dsRNA	Double-stranded RNA
DTT	Dithiothreitol
E. Coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
GO	Graphene oxide
Gria2 R/G	RNA derived from mRNA of Glutamate ionotropic receptor AMPA type subunit 2 gene
HDV	hepatitis delta virus
IDRs	Intrinsically disordered regions
IFNs	Interferons
IP6	inositol-hexakisphosphate
irAlu	Inverted repeat Alu
ISGs	Interferon stimulated genes
LGP2	Laboratory of Genetics and Physiology 2
m6A	N6-methyladenosine
MAVS	Mitochondrial antiviral signaling protein
MBP	Maltose binding protein
MDA5	Melanoma differentiation-associated protein 5
mtRNA	Mitochondrial RNA
nanoDSF	Nanodifferential scanning fluorimetry
NEIL1	RNA derived from DNA repair glycosylase NEIL1 gene
ON	Overnight
OsHV-1	Ostreid herpesvirus-1
PAMPs	pathogen-associated molecular patterns
PKR	protein kinase R
PRRs	pattern recognition receptors
RESCUE	RNA Editing for Specific C to U Exchange
RIG-I	Retinoic acid-inducible gene I
RLRs	RIG-I-like receptors

RPM	Round per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SINEs	Short interspersed nuclear elements
SPB	succinimidyl-[4-(psoralen-8-yloxy)]-butyrate
SPG	succinic acid, sodium dihydrogen phosphate, and glycine buffer
ssRNA	Single-stranded RNA
TLRs	Toll-like receptors
TTYH2	RNA derived from Tweety family member 2 gene
UTRs	untranslated regions
ZBP1	Z-DNA-binding protein 1
ψ	Pseudouridine

1 Introduction

1.1 Innate immunity and nucleic acids sensing

Innate immunity serves as the body's initial line of defense against pathogens, providing a rapid yet non-specific response. This system is capable of distinguishing between harmful pathogens and beneficial commensal organisms. It detects foreign molecules primarily by recognizing pathogen-associated molecular patterns (PAMPs), which are common molecular signatures found on pathogens. This recognition process is mediated by pattern recognition receptors (PRRs), specialized receptors that specifically bind to these PAMPs. Through this targeted interaction, the innate immune system can effectively identify and mount a response against potential threats, playing a vital role in the body's defense mechanisms (Janeway, 1989; Medzhitov, 2009).

Different PRR families vary in ligand recognition capabilities, signal transduction mechanisms, and subcellular localization. They are expressed in diverse cellular locations to effectively respond to a wide range of stimuli, being present on the cell surface as well as in intracellular compartments such as the cytosol or nucleus. PAMPs are characteristic molecules that are typically absent in the host organism. These can include distinctive lipids and carbohydrates, such as lipopolysaccharides and peptidoglycans, as well as unique protein features specific to bacteria, viruses, and other microbes. PAMPs also include certain nucleic acid structures, like double-stranded RNA (dsRNA) or single-stranded RNA (ssRNA). (Carpenter and O'Neill, 2024; Janeway and Medzhitov, 2002).

Nucleic acid sensors can be categorized based on their expression patterns and subcellular localization. Toll-like receptors (TLRs) can recognize different forms of nucleic acids derived from bacteria or viruses and they are located in endosomes of specialized immune cells, like dendritic cells, B-cells, or macrophages (Lind et al., 2022). Alternatively in the cytoplasm of almost all cell types DNA can be recognized through cyclic GMP-AMP synthase (cGAS) (Sun et al., 2013) and interferon-gamma-inducible protein 16 (Unterholzner et al., 2010) while RNA can be detected by RIG-I like receptors (RLRs).

RLRs like retinoic acid-inducible gene I (RIG-I) (Yoneyama et al., 2004) and its homolog melanoma differentiation-associated gene 5 (MDA5) (Kang et al., 2004) are part of a class of cytoplasmic RNA helicases. These receptors particularly sense ssRNA or dsRNA from viral genomes or RNA accumulated during viral replication, and they do so without requiring sequence specificity. Their activation triggers a downstream signaling cascade, leading to the production of pro-inflammatory cytokines or type I interferons (type I IFNs). The secreted type I IFNs then induce the transcription of interferon-stimulated genes (ISGs), which have a direct effect on fighting against the invading pathogens. This versatile and rapid response mechanism is a vital component of the innate immune system's defense against viral infections (Fig. 1) (Wu and Chen, 2014).

Nucleic acid sensors play a crucial role in detecting invading pathogens, yet they are also capable of detecting endogenous nucleic acids. When there are defects in the sensing mechanism, it can lead to dysregulation, resulting in inappropriate immune responses and the onset of inflammatory and autoimmune diseases (Khan et al., 2019). To effectively distinguish self-nucleic acids from foreign ones, these sensors rely on specific modifications

that mark self-nucleic acids. These modifications enable the sensors to recognize and distinguish endogenous nucleic acids from those that are foreign, ensuring a balanced immune response (Karikó et al., 2005; Mannion et al., 2014; Yang et al., 2014).

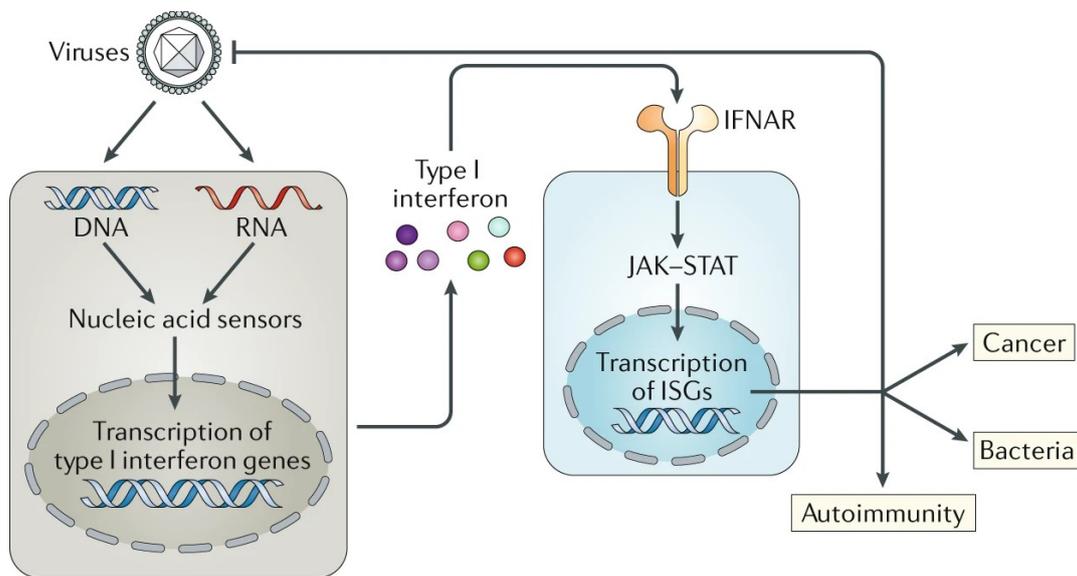


Figure 1 Type I interferon system.
Figure was downloaded from (Rehwinkel and Gack, 2020)

1.2 RNA modifications suppressing the innate immune response

RNA undergoes numerous modifications, with approximately 150 different types identified to date, including several examples found within RNA coding sequences (Fig. 2a) (Boccaletto et al., 2022). They are involved in influencing RNA metabolism and structural properties, which impact various aspects such as stability, splicing, translation, localization or nuclear export. The modifications can occur in all four nucleotides and the most common being methylation, pseudouridylation, and adenosine-to-inosine (A-to-I) editing (Delaunay et al., 2024).

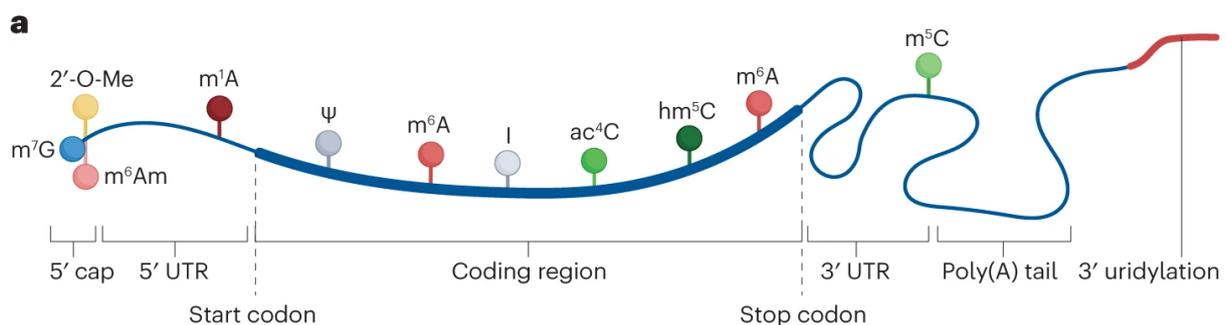


Figure 2 RNA modification present on mRNA.
Figure was downloaded from (Delaunay et al., 2024)

Numerous RNA modifications have been identified that can modulate the function of PRRs in various ways. Some modifications hinder PRRs recognition by allowing RNA to evade detection, while others positively influence the immune response by preventing unwanted activation.

One notable example of how mRNA modifications can shield against recognition by RLRs is the modification of the 5' cap structure. Specifically, the 7-methylguanosine cap (m7G, also known as cap0) and 2'-O-methylation (2'-O-Me) of the first and second nucleotides of the cap (cap1 or cap2) have been shown to be protective. Research indicates that these cap modifications help protect self-RNAs from detection by RIG-I, which recognizes uncapped RNAs with 5'-triphosphate or 5'-diphosphate groups that are typically found on viral RNAs (Ren et al., 2019; Schuberth-Wagner et al., 2015). Additionally, mRNAs lacking the 2'-O-Me modification have been observed to trigger the production of type I interferons through the Mda5 pathway (Züst et al., 2011).

m6A methylations are one of the most prevalent RNA modifications and can modulate innate immune sensing. This modification impacts the sensing by TLRs; specifically, *in vitro*, transcribed RNA containing m6A is a less potent activator than its unmodified counterparts (Karikó et al., 2005). The loss of the methyltransferase METTL3, the catalytic component of the m6A writer complex, in hematopoietic stem cells results in the abnormal formation of long dsRNA m6A-modified transcripts. Consequently, this leads to the induction of dsRNA-mediated innate immune responses, including the activation of the RIG-I and MDA5 pathways. The mechanisms through which METTL3 and the m6A modification suppress endogenous dsRNA levels remain unclear (Gao et al., 2020). Additionally, m6A modification in viral RNA helps it evade recognition by RIG-I. In particular, a study on human metapneumovirus (HMPV) demonstrated that m6A-deficient HMPV RNA triggers significantly higher type I interferon responses (Lu et al., 2020).

Pseudouridine (Ψ), a C5-glycoside isomer of uridine, significantly impacts PRRs by reducing their activation. Its most notable function is to maintain the structure and stability of various RNAs (Boo and Kim, 2020). This modification plays a significant role in host-virus interactions, as it has been demonstrated to help viral RNA in evading detection by host RLRs (Shen and Zhang, 2023). The pseudouridine modification can impair RIG-I's ability to form filaments and bind effectively to RNA ([Peisley et al., 2013](#)). While RIG-I is capable of recognizing the modification, it's unable to initiate the conformational changes required for activation (Durbin et al., 2016). Additionally, pseudouridine has been effectively utilized in the development of mRNA vaccines, such as the SARS-CoV-2 vaccine, to improve RNA stability and reduce undesired immune responses (Liu and Wang, 2022).

1.2.1 A-to-I RNA editing

Adenosine-to-inosine (A-to-I) RNA editing is one of the most prominent and irreversible modifications that can occur in both coding and non-coding regions of RNA primarily taking place in dsRNA and is mediated by a group of enzymes known as ADARs (adenosine deaminases acting on dsRNA). ADARs catalyze the hydrolytic deamination reaction at the C6 position of adenine, converting it into inosine (Fig. 3). The chemical structure of inosine closely resembles that of guanosine, differing only by the absence of the amino group attached to the C2 position. This structural similarity allows inosine to be recognized and interpreted as guanosine by the translational machinery (Bass, 2002).

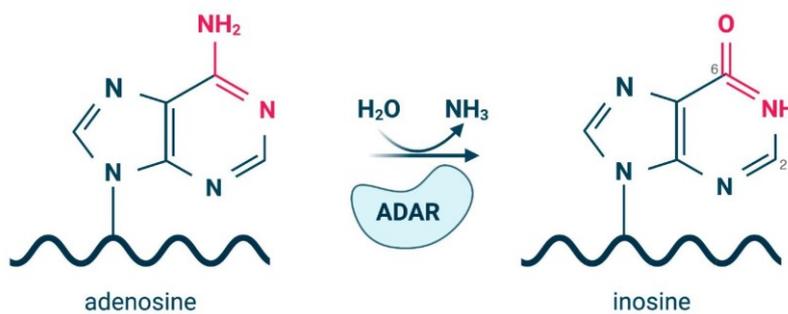


Figure 3 The conversion of adenosine to inosine catalyzed by ADAR. Chemical changes resulting from hydrolytic deamination are highlighted in red. Figure was prepared with BioRender.

A-to-I RNA editing was first identified as a puzzling enzymatic activity that caused the unwinding of injected dsRNA in the nucleus of *Xenopus laevis* oocytes (Bass, 1987; Rebagliati and Melton, 1987). Subsequent studies involving molecular cloning and biochemical purifications identified ADARs as enzymes catalyzing the editing reaction. The protein was initially called Double-stranded RNA-specific adenosine deaminase (DRADA or dsRAD) (Bass et al., 1997; Bass and Weintraub, 1988; Wagner et al., 1989)

Next-generation sequencing techniques have been utilized to globally identify inosines by detecting A-to-G variations through comparisons with the genomic sequence, leading to the discovery of millions of sites (Bazak et al., 2014; Ramaswami and Li, 2014). To minimize false-positive results, several bioinformatics pipelines have been developed (Bortoletto and Rosani, 2024). Despite the identification of numerous sites and the continuous discovery of new ones, the regulatory mechanisms underlying the editing process remain poorly understood. Editing levels exhibit variability across different cell types and change both developmentally and temporally, suggesting that regulation occurs at multiple levels. Small levels of these variations can be attributed to differences in ADAR expression levels or the abundance of dsRNA (GTEx Consortium et al., 2017; Lv et al., 2023; Porath et al., 2017).

The majority of identified RNA editing events are found in transcripts derived from Alu elements (Bazak et al., 2014; Ramaswami and Li, 2014), which are short interspersed nuclear elements (SINEs) abundant in the human genome, with over 1 million copies present, each approximately 300 nucleotides in length (Deininger, 2011). Alu elements are typically located within introns and 3' untranslated regions (UTRs) and can be arranged in an inverted orientation within the same transcript, forming inverted-repeat Alu (IR-Alu) structures. These structures often exhibit hairpin configurations with extended regions of dsRNA, making them prime targets for ADAR-mediated editing (Athanasiadis et al., 2004; Grover et al., 2004; Kim et al., 2004). Consequently, there are over 100 million potential ADAR-edited sites within the human transcriptome. However, the frequency of RNA editing at any individual Alu element is relatively low, typically less than 1% (Bazak et al., 2014).

The presence of endogenous dsRNAs can activate the cytoplasmic sensor MDA5, initiating an innate immune response. This activation can lead to the aberrant production of type I IFNs, resulting in potentially pathological consequences. To mitigate this unwanted immune response, A-to-I RNA editing serves as a vital modification that effectively marks cellular

dsRNA as "self" thereby repressing autoimmunity and preserving immune tolerance (Fig. 4) (Ahmad et al., 2018).

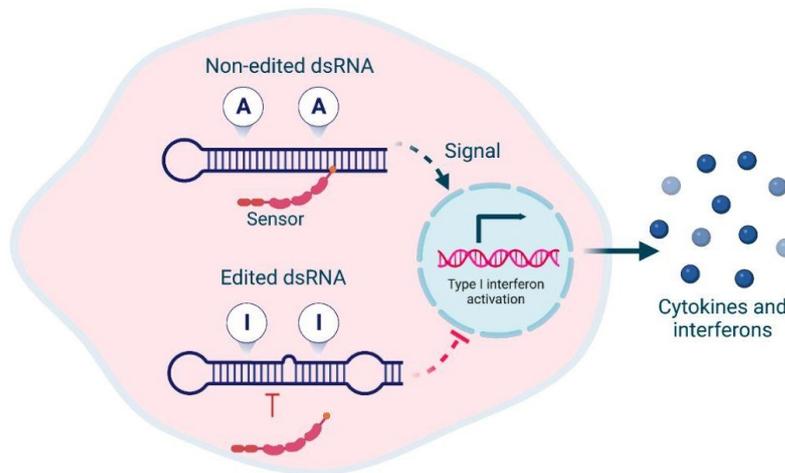


Figure 4 A-to-I editing marks cellular dsRNA as "self" preventing activation of innate immune responses. Figure was prepared with BioRender.

A-to-I editing can also have various consequences on RNA metabolism, including alterations in RNA stability, changes in splicing patterns or effects on gene recoding.

Introducing inosine in mRNA can have consequences in sequence recoding, leading to changes in the protein's amino acid sequence and function. One of the first reported examples was the substitution of glutamine for arginine (Q/R site) in the glutamate receptor GRIA2 (also known as GluR-B or GluR2). This change results in reduced permeability of the Ca^{2+} channel, which has been associated with early-onset epilepsy (Brusa et al., 1995; Higuchi et al., 1993; Sommer et al., 1991). However, it's important to note that the number of genes edited within their coding regions is relatively low (Gabay et al., 2022).

Inosine incorporation into RNA can alter its stability by affecting base pairing. Specifically, when inosine substitutes for adenine (A), it can form more stable pairs with cytosine (C) due to the presence of two hydrogen bonds in the I-C pair, compared to the single hydrogen bond in the A-C mismatch. Conversely, the inosine-uracil (I-U) mismatch is less stable than the adenine-uracil (A-U) pair. Overall, introducing inosine can enhance RNA stability in certain configurations while making others less stable (Wong et al., 2001; Wright et al., 2007, 2018).

1.3 ADAR proteins in humans

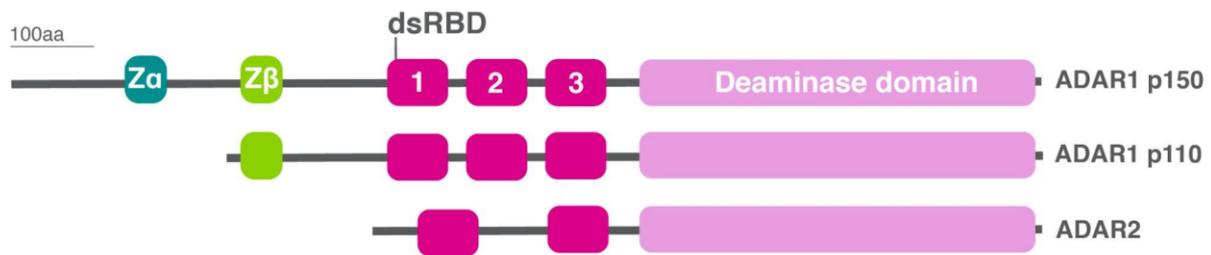


Figure 5 ADAR1 and ADAR2 domains organization.

A-to-I RNA editing in humans is facilitated by two catalytically active enzymes: ADAR1 and ADAR2. Both enzymes possess a modular domain organization, which includes a C-terminal deaminase domain that serves as the catalytic site responsible for the editing process. They also contain two to three double-stranded RNA binding domains (dsRBDs) that promote interaction with dsRNA. Additionally, ADAR1 features Z-DNA/RNA binding domains and one of these domains is capable of recognizing nucleic acids in an unusual left-handed conformation (Fig. 5) (Nishikura, 2016).

Humans also express a third ADAR enzyme, ADAR3, which is catalytically inactive (Chen et al., 2000). While ADAR3 does not possess RNA editing activity, it can bind to dsRNA and negatively regulate the RNA editing activities of ADAR1 and ADAR2 (Raghava Kurup et al., 2022).

1.3.1 ADAR1

ADAR1 is widely expressed in various tissues (GTEx Consortium et al., 2017) and has two major isoforms, ADAR1 p110 and ADAR1 p150, which are transcribed from distinct promoter regions. The ADAR1 p110 isoform is produced from constitutive promoters upstream of exon 1B and exon 1C, resulting in a protein of approximately 110 kDa. In contrast, the larger ADAR1 p150 isoform is generated from an interferon-inducible promoter upstream of exon 1A, leading to a protein of around 150 kDa (Fig. 6) (George et al., 2005; George and Samuel, 1999). Interestingly, the ADAR1 p110 isoform can also be translated from the mRNA that encodes the p150 variant. This occurs through a mechanism known as ribosome skipping, which bypasses the p150 start codon located in exon 1A (Sun et al., 2021).

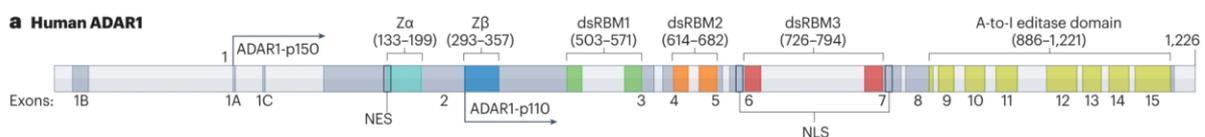


Figure 6 Organization of ADAR1 protein and genomic domains.

Exons 1B and 1C lack a methionine codon in the reading frame, which leads to mRNA variants with an extended 5' UTR that intrudes into exon 2. This results in the formation of the p110 variant, beginning at methionine 296. Exon 1A is interferon-inducible and gives rise to the p150 variant. Figure adapted from (De Reuver and Maelfait, 2024).

The ADAR1 isoforms can be localized to the nucleus due to the presence of a nuclear localization signal (NLS) in their dsRBD3 domain (Strehblow et al., 2002). ADAR1 p150 has an

additional Z-alpha (Z α) domain that features a nuclear export signal (NES), allowing it to shuttle between the cytoplasm and the nucleus, with its primary localization in the cytoplasm (Poulsen et al., 2001). In contrast, the ADAR1 p110 isoform is predominantly found in the nucleus. However, under stressful conditions such as UV radiation and heat shock, ADAR1 p110 becomes phosphorylated and transported to the cytoplasm where it binds to irAlus, which serve as binding sites for the Staufen1 protein, a key player in mRNA degradation (Gong and Maquat, 2011). By binding to irAlus, ADAR1 p110 competitively inhibits Staufen1, thereby preventing the degradation of RAD51, a vital DNA repair protein. This mechanism provides an anti-apoptotic effect, promoting cell survival during periods of cellular stress (Sakurai et al., 2017).

ADAR1 p150 contains a Z α domain, which is known to bind to left-handed double-stranded helical Z-DNA and Z-RNA (Herbert et al., 1997; Placido et al., 2007; Schwartz et al., 1999). Both isoforms of ADAR1 also possess a Z β domain; however, this domain is unable to bind to Z-DNA (Athanasiadis et al., 2005; Kim et al., 2003). The occurrence of proteins with a Z α domain is relatively rare, with only the Z-DNA-binding protein 1 (ZBP1) identified in mammals, which also plays a significant role in innate immunity. Additionally, this domain has been found in the fish PKR-protein like kinase (PKZ) and in certain viruses (Karki and Kanneganti, 2023). It is possible that additional proteins may possess Z α domain, as suggested by structural predictions, although they have not been yet experimentally validated (Bartas et al., 2022).

The enzymatic activity of ADAR1 is predominantly associated with the editing of noncoding RNA sequences located in the 5' and 3' UTRs as well as introns. Notably, ADAR1 primarily targets mobile elements, particularly those formed by irAlu repeats (Bazak et al., 2014; Ramaswami and Li, 2014). Research focusing on isoform-specific contributions has shown that the ADAR1 p150 is preferentially associated with the 3' UTRs, whereas the ADAR1 p110 isoform tends to target intronic sites (Kleinova et al., 2023).

ADAR1 knockout mouse models result in embryonic lethality, primarily due to significant liver damage, defects in organ development, and increased expression of ISGs. Importantly, the longer isoform of ADAR1, p150, is essential for its crucial functions. (Hartner et al., 2004; Pestal et al., 2015; Wang et al., 2004; Ward et al., 2011). ADAR1-mediated functions also have significant implications for autoinflammatory diseases and cancer (Song et al., 2022). A detailed explanation of the mechanisms underlying embryonic lethality and associated diseases will be presented later in the introduction.

1.3.2 ADAR2

ADAR2, also known as ADAR1b, exhibits the highest expression levels in the brain and central nervous system in both humans and rodents. Additionally, in humans, ADAR2 expression is significantly elevated in arteries, lungs, and the bladder (GTEx Consortium et al., 2017). This expression is regulated by multiple constitutive promoters and is influenced by alternative splicing of transcripts (Gerber et al., 1997; Slavov and Gardiner, 2002). Notably, ADAR2 can autoregulate its expression by editing its pre-mRNA, leading to the production of splicing variants with reduced levels of functional ADAR2 (Dawson et al., 2004). Furthermore, ADAR2 can shuttle between the nucleolus and nucleoplasm, a function which is influenced by its expression and substrate binding (Sansam et al., 2003).

ADAR2 activity is primarily linked to editing within the coding sequence, particularly in the nervous system, where it targets the majority of neural recoding sites (Gabay et al., 2022; GTEx Consortium et al., 2017). Gene knockout studies in mice have demonstrated that the only essential target of ADAR2 is the Q/R recoding site within the GRIA2 gene. Mice lacking the *Adar2* gene experience seizures, show significantly reduced survival and generally succumb within three weeks. Interestingly, this lethal phenotype can be rescued by the knock-in of the already edited *Gria2* Q/R site (Higuchi et al., 2000).

In patients with bi-allelic variants in the ADAR2 gene, mutations across various domains are associated with a spectrum of neurodevelopmental disorders, including microcephaly, intellectual disability, and seizures. These identified variants lead to reduced ADAR2 editing activity, which may contribute to the under-editing of the GRIA2 gene, thereby exacerbating the neurological symptoms observed in affected individuals (Tan et al., 2020).

1.3.3 Structural characterization of ADAR enzymes

Most of the experimental research characterizing the structural aspects of ADAR enzymes has primarily focused on ADAR2. This is largely due to the relative ease of working with ADAR2 in laboratory settings. ADAR2 can be readily overexpressed in yeast or insect expression systems, allowing for the purification of sufficient quantities of the protein, which also exhibits high stability (Cho et al., 2003; Keegan et al., 2007; Macbeth and Bass, 2007). Additionally, ADAR2 features a simpler domain organization compared to ADAR1, further facilitating its structural characterization. Currently, only the isolated domains of ADAR1 have been successfully purified, and a limited number of these domains have been structurally characterized.

1.3.3.1 Deamination catalytic mechanism

In early studies, the crystal structure of the deaminase domain of ADAR2 was elucidated, revealing the presence of an inositol-hexakisphosphate (IP6) molecule situated within a protein cavity primarily formed by conserved lysine and arginine residues. Binding of IP6 cofactor is crucial for the enzyme's activity and it is incorporated during protein expression. Furthermore, the structural analysis unveiled a zinc ion located at the active site of ADAR2. This zinc ion is tetrahedrally coordinated by two cysteine residues (Cys451 and Cys516), one histidine residue (His394), and a water molecule serving as the fourth ligand (Fig. 7B) (Macbeth et al., 2005).

Subsequent studies characterized the crystal structure of ADAR2 bound to an RNA substrate, providing deeper insights into the molecular mechanisms underlying the deamination reaction (Fig. 7A) (Matthews et al., 2016). The structure was determined using a 23-mer RNA duplex containing the nucleotide analog 8-azanebularine (8-azaN) at the targeted adenosine with opposing cytosine. This analog is unable to complete the deamination reaction because it lacks the 6-amino group. Consequently, the enzyme-substrate complex becomes trapped in the transition state, enabling the formation of a stable and tight complex between ADAR2 and the RNA interacting with nanomolar affinity (Haudenschild et al., 2004).

The structure reveals that the targeted adenosine analog, 8-azaN, is flipped out of the duplex RNA into the zinc-containing active site of ADAR2. The base-flipping process is facilitated by the conserved base-flipping loop, which approaches the minor groove of the RNA. Within this loop, the conserved sequence Gly-Glu-Gly allows the glutamate residue (Glu488 in ADAR2,

Glu1008 in ADAR1 p150) to penetrate the duplex RNA, effectively flipping the adenosine analog into the active site while forming two hydrogen bonds with the orphaned base opposite it. The flanking glycines (Gly487 and Gly489) provide the necessary flexibility, enabling Glu488 to stabilize the flipped-base conformation.

When the adenosine is positioned in the active site, the catalytic glutamate (Glu396 in ADAR2, Glu912 in ADAR1) deprotonates water molecule that is bound to the zinc ion. This deprotonated water molecule then acts as a nucleophile, attacking carbon-6 of the adenine ring and displacing the 6-amino group, resulting in the conversion of adenine to inosine. The flipped-out base is further stabilized by interactions with other active site residues, including Val35, Glu396 and Cys451, which make direct contact with the RNA ribose and phosphodiester backbone (Fig. 7C).

Additionally, the RNA interacts with ADAR2 through binding loops at both the 3' and 5' ends, which make contacts that are distal from the editing sites. Notably, the 5' loop (residues 454-477) was found to be unstructured in the absence of RNA but becomes ordered upon binding to the RNA substrate. This loop primarily interacts with the RNA through the phosphodiester backbone via positively charged residues. Mutagenesis studies identified that 6 out of 18 conserved residues within this loop are critical for the editing activity of ADAR2 (Wang and Beal, 2016). Interestingly, the 5' loop of ADAR1 differs significantly from that of ADAR2 in both amino acid composition and loop size. To investigate the functional implications of these differences, loop swapping experiments were conducted where the 5' loop of ADAR1 was exchanged with that of ADAR2. The results demonstrated that the editing activity of ADAR1 was affected, by changing the substrate selectivity (Wang et al., 2018).

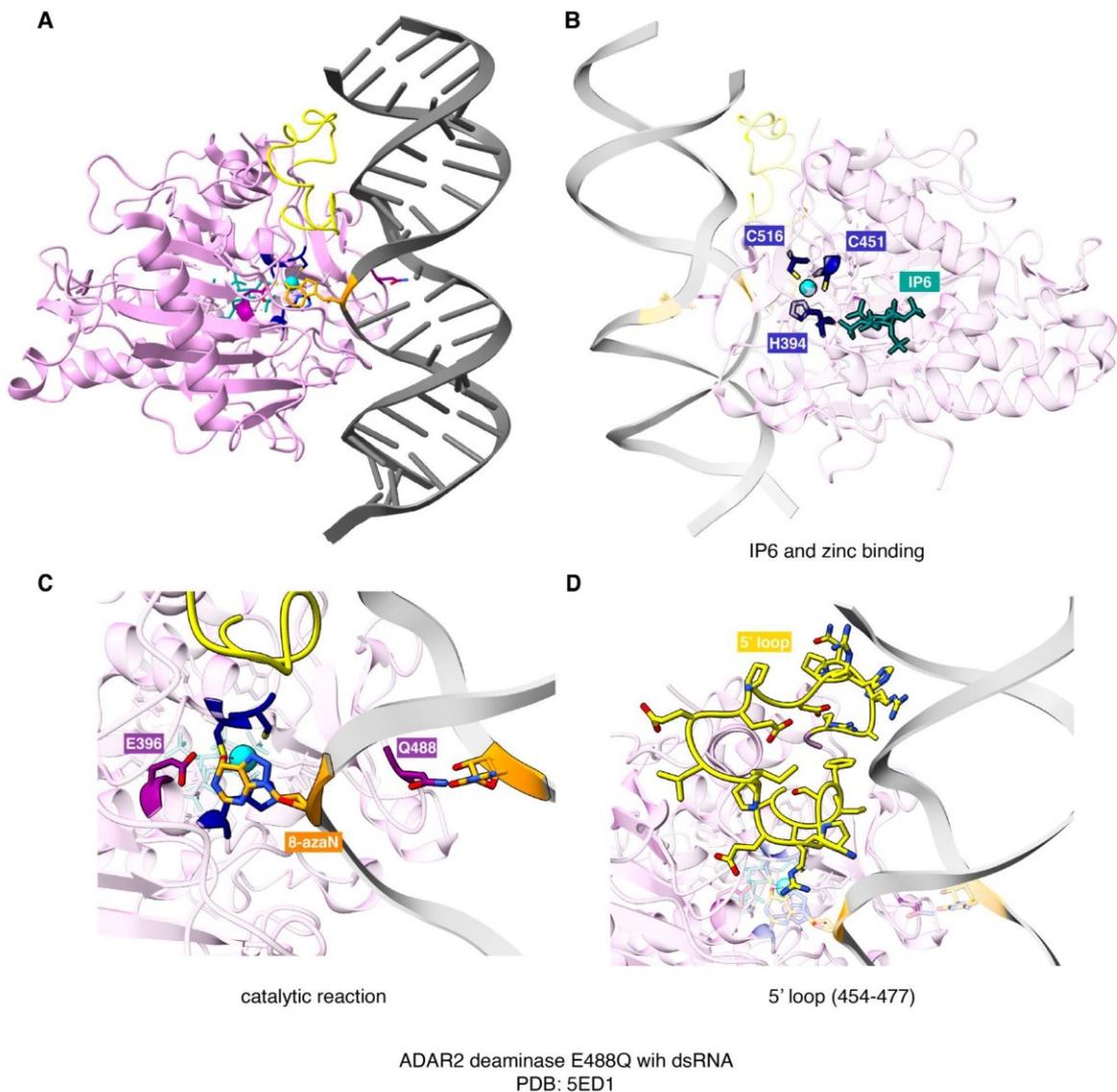


Figure 7 Structure of ADAR2 E488Q bound to an RNA substrate, illustrating the deamination catalytic mechanism.

- The overall structure of the deaminase domain (indicated in plum) alongside the RNA helix (shown in dim gray). PDB: 5ED1.
 - Highlighted IP6 and zinc binding with the zinc coordinating residues (indicated in dark blue)
 - Highlighted catalytic reaction with catalytic residues (indicated in purple) and nucleotides: 8-azaN and orphaned cytosine (indicated in orange)
 - Highlighted 5' loop (454-477) (indicated in yellow) displaying the charges of residues
- Color naming is based on the ChimeraX color palette.

The enzymatic activity of ADAR2 can be significantly enhanced by mutating Glu488 to glutamine (E488Q) (Kuttan and Bass, 2012). This mutation results in the enzyme exhibiting a higher affinity for RNA duplexes containing 8-azaN (Phelps et al., 2015). Under physiological conditions, glutamine is fully protonated, in contrast to glutamic acid, which can be deprotonated. This protonation state can improve hydrogen bonding interactions with the orphaned nucleotide. These observations are further supported by experiments investigating the effects of varying pH on reaction efficiency. When the pH is lowered, Glu488 becomes more protonated, which may facilitate the base-flipping mechanism (Malik et al., 2021). The

enhanced activity of ADAR2 can also be modulated by substituting the orphan cytosine with modified analogs that are capable of forming hydrogen bonds with the wild-type Glu488 residue in a manner that is independent of protonation state. This observation was confirmed by structural studies accompanied by biochemical assays (Doherty et al., 2021).

1.3.3.2 ADAR dimers

Various studies have reported that ADARs can form dimers in an RNA-dependent or -independent manner (Gallo, 2003; Poulsen et al., 2006; Valente and Nishikura, 2007). The formation of ADAR1-ADAR2 heterodimers has been observed in human cells (Cenci et al., 2008) and demonstrated using FRET analysis (Chilibeck et al., 2006). The requirement of dimerization for the editing activity of ADARs has been a subject of debate. However, the publication of a crystal structure of the ADAR2 construct comprising the deaminase domain and dsRBD2 bound to an RNA substrate has provided more insights into the role of dimerization (Thuy-Boun et al., 2020).

The structure revealed the formation of asymmetric dimers, in which the deaminase domain of one monomer directly engages with RNA and performs the catalytic reaction as observed previously. In contrast, the other monomer interacts with RNA through its dsRBD2. Importantly, the second deaminase domain does not interact with RNA; rather, it establishes a dimerization interface with the catalytic deaminase. The dimerization interface is established by a conserved TWDG motif, which is also found in ADAR1 (Fig. 8A). Disruption of this dimerization interface through targeted mutations led to a reduction in editing activity for both ADARs. This was assessed through in vitro editing experiments on selected substrates and further evaluated in HEK293T cells (Thuy-Boun et al., 2020).

Structural data has revealed an additional dimerization interface for ADAR1. The structural analyses have shown that the dsRBD3 domains are capable of dimerizing in an RNA-independent manner. Importantly, this dimerization interface does not preclude the individual dsRBD3 domains from binding to dsRNA (Fig. 8B). Further investigation has demonstrated that by mutating the dimerization interface, the RNA editing activity was reduced at selected target sites, while having no effect on some other sites, as tested in an in vivo cellular system. This suggests that the dimerization of the dsRBD3 domains plays a role in modulating the RNA editing activity of ADAR1, but the specific effects are dependent on the target site (Mboukou et al., 2023).

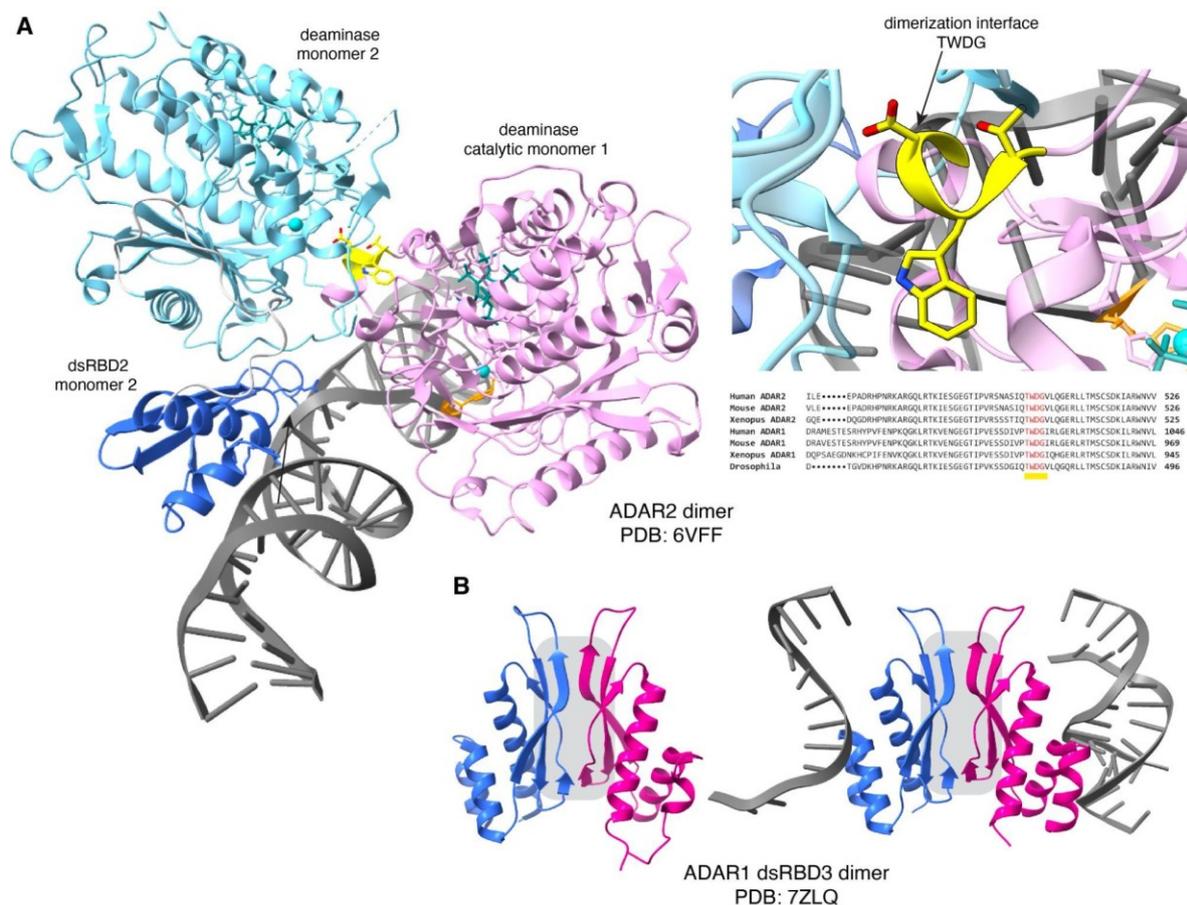


Figure 8 Structural data showing ADAR dimers.

- ADAR2 asymmetric dimer bound to dsRNA (PDB: 6VFF). Color labeling: the catalytic deaminase of monomer 1 is shown in plum, deaminase of monomer 2 in sky blue, and dsRBD2 from monomer 2 in royal blue. The dimerization interface TWDG is indicated in yellow.
- ADAR1 dsRBD3 dimer, either alone or bound to dsRNA. The dsRBD3 domains are indicated in medium violet red and royal blue. The dimerization interface between the β -sheets is highlighted with a gray square.

1.3.3.3 ADAR1 structural characteristics

The ADAR1 deaminase domain harbors an additional zinc-binding motif that is absent in ADAR2. This was identified in a study that utilized high-throughput mutagenesis to investigate the importance of cysteine residues in ADAR1, combined with functional screening and biochemical assays. Mass spectrometry metal analysis confirmed the presence of an additional zinc ion, which is coordinated by the His988, Cys1081, Cys1082, and His1103 residues. When the Cys1082 residue was mutated to aspartic (C1082D) or glutamic (C1082E) acid, the resulting protein showed significantly reduced editing activity in *in vitro* assays, and no activity was detected when tested in cellular systems (Park et al., 2020) (Fig. 9A)

As previously mentioned, ADAR1 possesses additional Z-domains - Z α and Z β . Several crystal structures have been resolved for the Z α domain in complex with Z-DNA (Ha et al., 2009, 2005; Schwartz et al., 1999) or Z-RNA (Placido et al., 2007), and the crystal structure of the free Z-beta domain has also been determined (Athanasiadis et al., 2005). The Z α domain is capable of recognizing the zig-zag conformation of phosphodiester backbones, which allows it to bind to both DNA and RNA. Among the various conserved residues in the Z α domain, Asn173,

Tyr177, and Trp195 are completely conserved in other reported Z-domains and play critical roles in binding (Fig. 9B)

Structurally, the Z β domain is closely related to the Z α domain. However, it lacks the Tyr177 residue and instead has an Ile335 residue, which prevents it from binding to Z-nucleic acids. Interestingly, when Ile335 residue is mutated to a tyrosine, the domain exhibits the ability to bind to Z-DNA (Kim et al., 2003) (Fig. 9B).

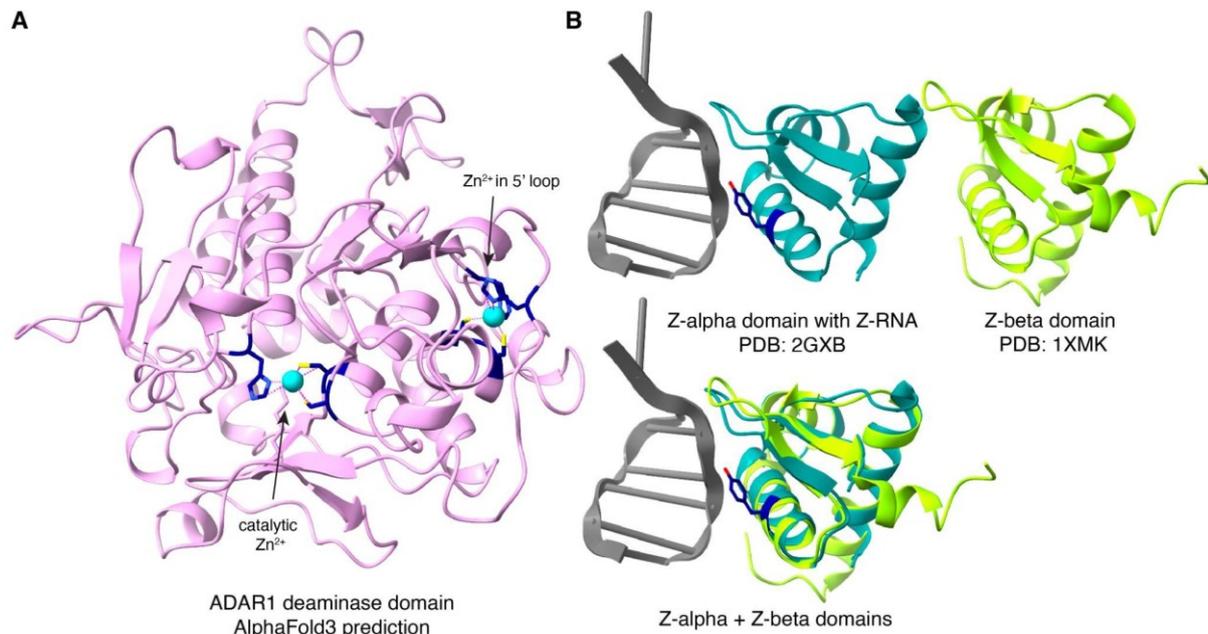


Figure 9 Structures of ADAR1 domains.

- Predicted structure of the ADAR1 deaminase domain, displaying two zinc ions (indicated in dark turquoise) and their coordinating residues (labeled in dark blue).
- Structure of the Z-alpha domain bound to Z-RNA (dark cyan, PDB: 2GXB) and the Z-beta domain (yellow-green, PDB: 1XMK). The two structures are aligned with Tyr177, highlighted in dark blue, which is present only in the Z-alpha domain.

1.3.4 Characteristics of RNA substrates edited by ADARs

ADARs have some defined preferences for modifying adenosines at specific positions within RNA substrates. They can target dsRNA duplexes longer than 15 base pairs and their specificity is influenced by factors like the identity of opposing and neighboring bases and the overall secondary structural features of the RNA molecule.

Previous studies have systematically examined the preferences of ADARs for opposing bases within different RNA configurations (Kallman, 2003; Wong et al., 2001). The findings indicate that ADARs preferentially edit A-C mismatches and A-U pairs, while showing little to no activity on A-A and A-G mismatches. This observed substrate specificity can be explained based on structural insights. The Glu488 residue which occupies the space vacated by the flipped-out target adenosine could clash with a purine base which is bigger than preferred pyrimidine (Matthews et al., 2016).

Different studies have revealed the distinct preferences of ADARs regarding the neighboring nucleotides of their target adenosines. For the 5' nearest nucleotide, both ADAR1 and ADAR2 exhibit a preference for U and A, while C and G are less tolerated. Furthermore, ADAR2 has been found to favor the 3' nearest neighbors in the order: G = U > C = A. In contrast, ADAR1 shows a preference for 3' neighbors as follows: G > C ≈ A > U. Notably, the identity of the 5' nucleotide exerts a more significant influence on the editing capabilities of ADARs, while the deaminase domain plays a crucial role in selecting adenosines within various sequence contexts (Eggington et al., 2011; Eifler et al., 2013; Lehmann and Bass, 2000).

The preference for a 5' nucleotide can also be elucidated through structural studies. Experimental findings indicate that substituting 5'U with C or G significantly reduces enzymatic activity by approximately 80%. This reduction can be further attributed to a minor clash involving the Gly489 residue in the penetrating loop of the enzyme. While this clash is not severe, it suggests that the enzyme may still have some capacity to accommodate 5' C or G (Matthews et al., 2016). Interestingly, the accommodation of 5'G and the enhancement of editing rates can be further modulated by introducing a G-purine (G or A) mismatch at the -1 position within the guide strand (Doherty et al., 2022).

The preference for 3' G can be attributed to the interaction between the 2-amino group of guanine and the Ser486 residue in the base-flipping loop of the enzyme. When the 2-amino group is eliminated by replacing G with I - which can still form a base pair with C - the enzymatic activity is notably reduced by approximately 50% (Matthews et al., 2016).

Editing specificity of ADARs can also be influenced by deviations from perfect base pairing within the RNA structure. Such structural deviations, like mismatches, bulges, and loops, can affect the overall secondary structure of the RNA molecule and influence the ADAR's targeting preferences (Fig. 10) (Eggington et al., 2011; Lehmann and Bass, 1999, 2000).

The initial studies focused on a limited selection of known RNA substrates and lacked a systematic approach. However, high-throughput screening combined with mutagenesis of selected natural substrates provided a deeper understanding of the influence of RNA secondary structure. These investigations confirmed previous findings regarding the roles of opposing and neighboring bases and demonstrated that both general and substrate-specific characteristics collaboratively affect editing levels. The degree to which each feature influences editing varies across different RNAs (Liu et al., 2021). In another study, addressing high-throughput screening of secondary structures in long dsRNA substrates, ADAR1 editing activity was mapped 35 nucleotides upstream of disruptions in secondary structure, demonstrating a periodic pattern in the editing process (Uzonyi et al., 2021). Furthermore, the irCLASH approach (ADAR RNA co-immunoprecipitation with RNA sequencing) revealed that ADAR proteins bind dsRNA substrates in tandem with a 50-base pair footprint (Song et al., 2020).

Collectively, these studies suggest a complex and context-sensitive regulation of the RNA editing landscape that cannot be defined by simple parameters alone. The various factors, including secondary structure, neighboring bases, and substrate-specific characteristics, work

synergistically to influence editing levels in a manner that varies across different RNAs.

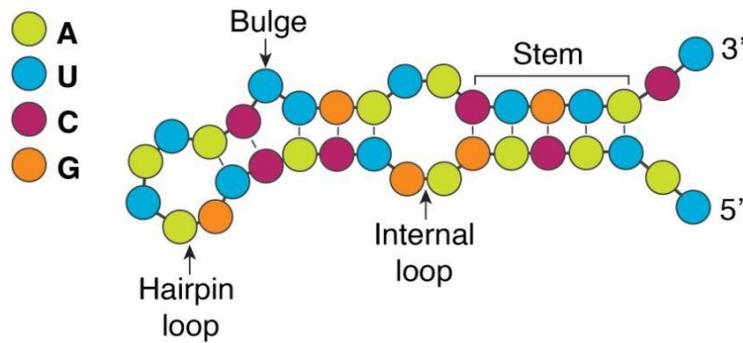


Figure 10 RNA structural features recognized by ADARs.

1.4 dsRNA sensing pathways regulated by ADAR1

The crucial role of ADAR1 in regulating dsRNA sensing pathways has been well-established through experimental evidence from various mouse models. These studies have also provided valuable insights into the lethality associated with ADAR1 knockout.

The loss of function of both isoforms of ADAR1, p110 and p150 (*Adar*^{-/-}), in mice leads to embryonic lethality at day 12.5 due to extensive liver damage, defects in organs development and elevated ISGs expression (Hartner et al., 2004; Pestal et al., 2015; Wang et al., 2004). Mice with a specific deletion of the longer isoform ADAR1 p150 phenocopy the full deletion of ADAR1, demonstrating that the essential functions of ADAR1 rely on the longer isoform, either due to its cytosolic RNA editing or RNA binding activities (Ward et al., 2011).

Further research has shown that the embryonic lethality associated with ADAR1 deletion is directly connected to its RNA editing activity. Mice with catalytically deficient ADAR1 mutant (*Adar*^{E861A/E861A}), where both isoforms can bind but not edit RNA, survive to a comparable lifespan of 13.5 days, similar to those with a full deletion of ADAR1 (Liddicoat et al., 2015).

The lethality caused by ADAR1 deletion and elevated ISGs expression could be rescued when animals were crossed with mice lacking the dsRNA sensor MDA5 (*Ifih1*^{-/-}) or its adaptor protein MAVS (*Mavs*^{-/-}), but not when crossed with mice lacking protein kinase R (PKR, *Eif2ak2*^{-/-}) or RIG-I (*Ddx58*^{-/-}) (Bajad et al., 2020; Liddicoat et al., 2015; Mannion et al., 2014; Pestal et al., 2015; Wang et al., 2004). These findings highlight the critical role of ADAR1's editing activity in preventing the spontaneous activation of the MDA5 signaling pathway. Notably, the degree of rescue varied among different ADAR1 variants. While ADAR1 null mutants could only be rescued to the point of birth, the editing-deficient mutants were able to survive until adulthood (Heraud-Farlow et al., 2017; Liddicoat et al., 2015).

The observation that the rescued editing-deficient ADAR1 mutant exhibits a longer lifespan compared to the complete ADAR1 knockout highlights the significance of other RNA editing independent functions of ADAR1. Despite lacking the capacity for A-to-I editing, the editing-deficient ADAR1 mutant retains the ability to bind to various nucleic acids. This RNA-binding activity may lead to the sequestration of dsRNA, thereby preventing the downstream activation of other interferon-inducible nucleic acid sensors, such as ZBP1 and PKR.

ADAR1 mediated suppression of PKR was reported in virus infected cells as well as in human and mice cell models (Chung et al., 2018; Nie et al., 2007; Wang et al., 2004). PKR is a protein kinase that becomes activated upon binding to dsRNA through dimerization and autophosphorylation. Subsequently, it phosphorylates the protein synthesis initiation factor eIF2 α , leading to translational cell shutdown and the integrated stress response (Hur, 2019).

Earlier reports of individual knockouts of either PKR or MDA5/MAVS alone were unable to effectively compensate for the loss of ADAR1 and rescue the mice to adulthood. Recent studies have shown that ADAR1 or ADAR1 p150 mutant mice can reach 40% to adulthood or full rescue respectively, when both PKR and either MDA5 or MAVS are removed (Hu et al., 2023; Sinigaglia et al., 2024). Utilizing a human cell model, it was shown that ADAR1 p150 competitively binds to dsRNA substrates alongside PKR, thereby inhibiting PKR from inducing translational arrest. Notably, this competitive interaction is associated with the cytoplasmic localization of ADAR1 p150 (Hu et al., 2023). Consequently, to achieve a complete rescue of the ADAR1 p150 knockout phenotype, it is crucial to suppress both the PKR and the MDA5/MAVS signaling pathways (Hu et al., 2023; Sinigaglia et al., 2024).

ZBP1 is known to activate a form of inflammatory cell death called PANoptosis, which involves a combination of pyroptosis, apoptosis, and necroptosis (Wang and Kanneganti, 2021). Given that ADAR1 and ZBP1 are the only known proteins in mammals harboring the Za domain, it was speculated that ZBP1 might have a role in the immunopathology associated with ADAR1 knockouts. Deletion of ZBP1 (*Zbp1*^{-/-}) alone wasn't able to rescue ADAR1 mutants, by combining it with subsequent deletion of MAVS (*Zbp1*^{-/-}/*Mavs*^{-/-}) the life span of mice was prolonged, but it did not achieve the same level of rescue as observed when both the MDA5 and PKR pathways were inactivated. Further investigation by independent research groups has revealed that the postnatal lethality and immunopathology seen in mouse models with homozygous mutations in the ADAR1 p150 Za domain can be prevented by the subsequent deletion of ZBP1 or the ZBP1 Za domain mutant (De Reuver et al., 2022; Hubbard et al., 2022; Jiao et al., 2022).

Overall the immunosuppressive function of ADAR1 p150 might be attributed to its ability to bind and potentially destabilize immunostimulatory dsRNA in the cytoplasm. This hypothesis is supported by experiments using overexpression systems, in which ADAR1 p110 was redirected to the cytoplasm, enabling it to alter its specificity to mimic that of ADAR1 p150 (Kleinova et al., 2023).

In summary, ADAR1 plays a critical role in safeguarding dsRNA sensing pathways and mitigating inflammation through multiple mechanisms. Its A-to-I editing activity modulates MDA5/MAVS signaling pathways, thereby inhibiting the activation of type I IFNs. Additionally, ADAR1 prevents translational shutdown by competing with PKR for binding to dsRNA. Furthermore, the distinctive Za domain of the ADAR1 p150 isoform is important in preventing cell death through the process of PANoptosis (Fig. 11).

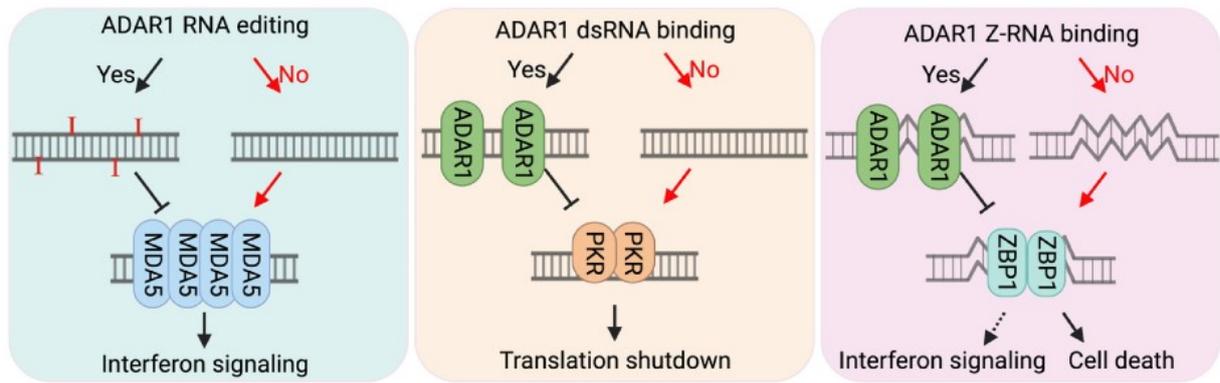


Figure 11 dsRNA sensing pathways regulated by ADAR1.

Figure was downloaded from (Hu and Li, 2024)

1.4.1 MDA5 dsRNA recognition and activation

MDA5 is part of the RLR family (Fig. 12), which also includes RIG-I and Laboratory of Genetics and Physiology 2 (LGP2). These receptors are ATP-dependent helicases that bind to immunostimulatory dsRNA in the cytosol through their helicase and carboxy-terminal domains (CTD) (Rehwinkel and Gack, 2020). RIG-I and MDA5 both possess two caspase activation and recruitment domains (CARDs) that oligomerize upon binding to dsRNA and subsequently interact with MAVS (Mitochondrial Antiviral Signaling Protein) on the mitochondria. This interaction initiates downstream signaling pathways that lead to the production of type I IFNs and pro-inflammatory cytokines (Hou et al., 2011; Kawai et al., 2005; Wu et al., 2013a)

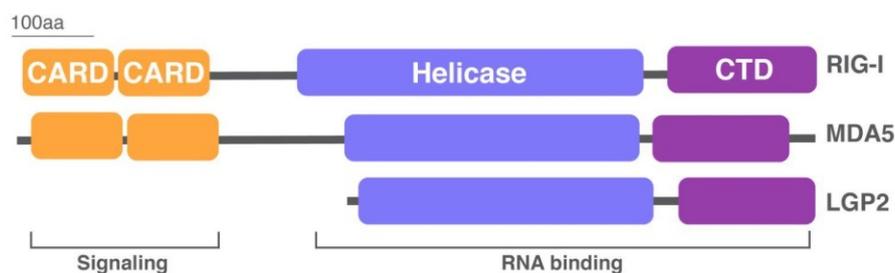


Figure 12 RLRs receptors and their domain organization.

RIG-I and MDA5 can recognize distinct types of RNA. RIG-I specifically binds to blunt-ended dsRNA with di- or triphosphate groups, which are characteristic of viral RNA. In contrast, MDA5 interacts with long strands of dsRNA without any sequence specificity (Kato et al., 2008, 2006; Rehwinkel and Gack, 2020). MDA5 employs its helicase and CTD domains to assemble into a ring-like structure that can bind to the phosphodiester backbone of dsRNA. This structural arrangement allows MDA5 to interact with dsRNA in a sequence-independent manner, as it recognizes the overall dsRNA backbone rather than specific nucleotide sequences (Wu et al., 2013a; Yu et al., 2018a).

MDA5 cooperatively assembles along dsRNA to form filamentous structures, enabling robust binding to long dsRNA molecules (Berke et al., 2012; Peisley et al., 2011; Wu et al., 2013a). The disassembly of MDA5 from dsRNA is regulated in a length-dependent manner through

ATP hydrolysis (Berke and Modis, 2012; Peisley et al., 2012). Structural analyses have revealed that the binding state of ADP/ATP within MDA5 induces conformational changes that control its association and dissociation from dsRNA (Yu et al., 2018a). It has been proposed that ATP hydrolysis facilitates the dissociation of MDA5 from shorter dsRNA molecules, thereby ensuring selective binding to longer dsRNA sequences. This selective binding process functions as a proofreading mechanism, enhancing the specificity and accuracy of MDA5's interactions with its target RNA ligands (Peisley et al., 2012; Yu et al., 2018a). Recent findings revealed that ATP hydrolysis drives MDA5's translocation along dsRNA, potentially contributing to subsequent CARD-CARD oligomerization (Han et al., 2024) which in turn facilitates the nucleation of microfibrils made up of MAVS CARDS. This process leads to the activation of type I IFNs downstream signaling pathway (Hou et al., 2011) (Fig.13).

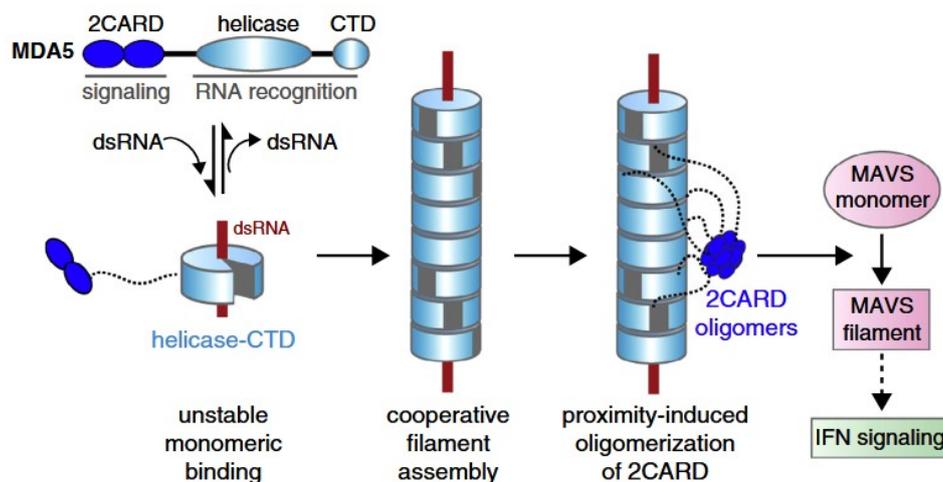


Figure 13 Model of MDA5 filament formation and signal activation.

When MDA5 binds to dsRNA, it forms a ring-like structure through its helicase and CTD domains. The initial monomeric binding is unstable, but MDA5 then cooperatively assembles along the dsRNA, forming filaments. This filament formation brings the neighboring 2CARD domains into close proximity, inducing their oligomerization. The oligomerized 2CARD domains then interact with the MAVS, which initiates the downstream interferon signaling pathways. The figure was downloaded from (Del Toro Duany et al., 2015)

LGP2, which lacks the CARD domains, is unable to initiate the MAVS signaling. However, it functions as a positive regulator of MDA5 activation by facilitating the assembly of MDA5 onto dsRNA (Sato et al., 2010). LGP2 has been shown to accelerate the initial interactions between MDA5 and dsRNA, resulting in the formation of numerous shorter filaments. This process enhances the cooperative assembly and activation of MDA5 on these shorter stretches of dsRNA (Bruns et al., 2014).

1.4.2 Endogenous dsRNAs recognized by MDA5

The connection between ADAR1 deficiency and MDA5 activation has been established through mouse genetic models, as previously described. Biochemical evidence supporting this relationship is provided by RNase protection assays, which demonstrate that MDA5 targets irAlu as its primary endogenous ligand. Normally, MDA5 does not recognize irAlu due to its sensitivity to structural irregularities, such as mismatches, loops, or bulges present in these RNA sequences. Moreover, the ADAR1 mediated A-to-I editing of irAlu RNA can introduce more structural irregularities by weakening Watson-Crick base pairing. This serves as a

regulatory mechanism to prevent MDA5 from inappropriately activating in response to cellular self-RNA. However, in the absence of A-to-I editing by ADAR1, or when MDA5 harbors gain-of-function mutations, it can bypass this regulatory mechanism and recognize irAlu (Ahmad et al., 2018). This aberrant recognition of endogenous RNA by MDA5 can trigger excessive and inappropriate activation of interferon signaling, resulting in undesirable immune responses. Such dysregulation may contribute to the pathogenesis of a wide range of inflammatory disorders (Rice et al., 2014).

In addition to irAlu RNA, MDA5 signaling can also be activated by other endogenous RNA species, including mislocalized mitochondrial dsRNA (mtRNA). The loss of mitochondrial enzymes SUV3 or PNPase, which play a crucial role in the degradation of mtRNA, can result in the abnormal accumulation of these mitochondrial RNA species in the cytoplasm. This accumulation ultimately leads to the aberrant activation of type I interferon signaling via MDA5 activation. Currently, it is not known how mtRNA gains access to the cytosol (Dhir et al., 2018).

Another class of immunogenic dsRNA that activates MDA5 originates from cis-natural antisense transcripts (cis-NATs). These complementary RNAs are transcribed from partially overlapping sequences on opposing DNA strands within the same genomic locus (Rehwinkel and Mehdipour, 2024). Unlike irAlus, cis-NATs form perfectly dsRNA with an average length of 600 base pairs, making them more suitable ligands for MDA5 activation (Li et al., 2022).

Genome-wide association studies have identified cis-NAT editing as a potential mechanism underlying genetic variants linked to common inflammatory conditions. In vitro experiments demonstrated that MDA5 could form filaments on cis-NATs, with filament lengths significantly reduced when the RNA was pre-edited with ADAR1. Transfecting cis-NATs into cells deficient in functional ADAR1 editing led to enhanced expression of ISGs, driven by the increased expression of MDA5. In contrast, transfection of the same cis-NATs into wild-type cells resulted in a reduced immune response. These experiments highlight the potent ability of cis-NATs to induce MDA5-dependent immune activation (Li et al., 2022).

It is important to note that mtRNA and cis-NATs are relatively rare within the genome, typically expressing under specific conditions associated with various diseases. In contrast, irAlus represents the most abundant endogenous dsRNA targeted by MDA5.

1.5 ADAR1 implication in human diseases

1.5.1 Autoimmune disorders

Mutations in the ADAR1 gene are associated with impaired protein function, which can lead to the development of autoimmune disorders. Notable examples include Aicardi-Goutières syndrome, dyschromatosis symmetrica hereditaria, and systemic lupus erythematosus.

Aicardi-Goutières syndrome (AGS) is a rare genetic autoimmune disorder that can present with a diverse range of symptoms, including severe neurological impairments, skin lesions, and systemic autoimmune features. AGS is caused by mutations in various genes associated with type I IFNs signaling pathways, leading to interferonopathies (Liu and Ying, 2023).

Several mutations in the ADAR1 gene have been identified in patients with AGS. ADAR1-mediated A-to-I RNA editing serves as a suppressor of type I IFNs signaling, and it has been

proposed that impaired ADAR1 function results in increased immunostimulatory dsRNA, which activates innate immune response pathways (Rice et al., 2012). Most ADAR1 mutations are found in the deaminase domain, among with the most commonly reported mutation P193A, located in the Z α domain in a heterozygous manner. The P193A mutation is associated with other variants in the deaminase domain. In knock-in mouse models harboring this point mutation (in mouse Adar P195A), the mutation is well tolerated, and the animals are viable. In contrast, when the P195A mutation is combined with a complete loss of ADAR1 gene, the mice exhibit a reduced lifespan and abnormalities in the spleen, kidney, as well as increased expression of ISGs. Notably, these defects can be rescued by deleting either the MDA5 or PKR genes. This observation may help explain the absence of homozygous mutations in the ADAR1 Z α domain found in AGS patients to date (Guo et al., 2023; Liang et al., 2023; Maurano et al., 2021).

The effects of mutations located in the deaminase domain of the ADAR1 enzyme were assessed through biochemical experiments, providing deeper mechanistic insights into the enzyme's properties. These mutations can be classified into different categories based on their impact on enzyme function. The first group comprises residues at the protein-RNA binding interface, including R892H, G1007R, and K999N, the latter being located near the 5' loop. The G1007R mutation, situated in the Gly-Glu-Gly base-flipping loop, has the most detrimental impact on enzyme catalytic activity. Additionally, its inhibitory effect may arise from interference with the dimerization interface, as the presence of arginine may clash with neighboring residues. R892H affects the deamination reaction by impairing the stabilization of the flipped base, as this stabilization may be disrupted by less optimal binding resulting from the insertion of histidine. The second category includes the mutations Y1112F and K999N, also located near the 5' binding loop. Mutations in this category have been shown to have a mild effect on enzyme activity, with observations being substrate- and context-dependent (Karki et al., 2024). The final category encompasses residues A870T and I872T, found in the internal region of the protein. They can shift positioning of the catalytic glutamate residue (Fisher and Beal, 2017) (Fig.14).

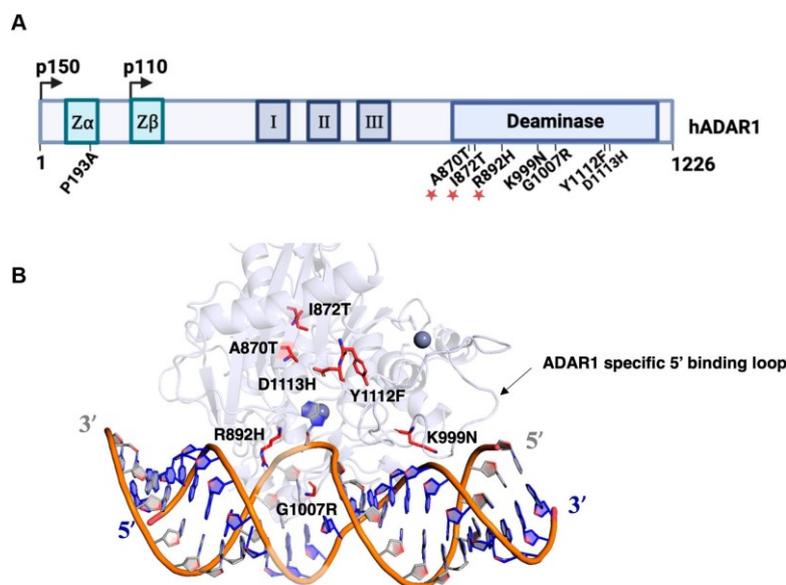


Figure 14 AGS mutations mapped in ADAR1 deaminase model bound to dsRNA. Figure was downloaded from (Karki et al., 2024)

- A. The positions of mutations in the ADAR1 domains found in AGS patients with asterisks denoting mutations found together with P193A
- B. ADAR1 deaminase Rosetta model with dsRNA, showcasing highlighted mutations. This model was generated in a recent study from (Park et al., 2020)

Dyschromatosis symmetrica hereditaria (DSH) is a skin pigment disorder characterized by the presence of both hypopigmented and hyperpigmented macules and patches on the face, and dorsal side of hands and feet. Similar to AGS, DSH is associated with mutations in the ADAR1 gene, with approximately 180 mutations identified to date (Ma et al., 2023; Miyamura et al., 2003). Recently, a case was reported involving a 6-year-old female patient with a known history of AGS who also presented with symptoms of DSH. The diagnosis of AGS can be challenging due to its diverse symptoms and the absence of a systematic description of the disease. However, the more apparent effects of DSH symptoms may help in facilitating an earlier diagnosis of AGS (Ahmed et al., 2024).

Systemic lupus erythematosus (SLE) is characterized by an abnormal immune response to autoantigens within the body, leading to symptoms such as joint pain, skin rashes, and fatigue. ADAR1 may contribute to the pathogenesis of SLE through its catalytic activity. Patients with SLE show elevated levels of A-to-I RNA editing, which correlates with increased expression of ADAR1, particularly in those with high levels of ISGs. RNA editing can result in the formation of edited peptides that may function as MHC class I epitopes, potentially triggering an immune response (Roth et al., 2018).

1.5.2 ADAR1 role in cancers

ADAR1 may have pro-oncogenic effects in various cancers due to its role in preventing dsRNA sensing pathways. While immune checkpoint therapies have emerged as an effective cancer treatment approach, many patients do not respond favorably. Interestingly, in a screening study to identify factors that could sensitize cancers to be more vulnerable to killing by cytotoxic T cells, ADAR1 was found to be a top hit candidate across multiple murine cancer models (Lawson et al., 2020).

Loss of ADAR1 function, achieved through knockdown of ADAR1-p150 in tumor cells, improves responses to PD-1 blockade and overcomes common mechanisms of resistance to immunotherapy. This effect is associated with the activation of the MDA5-MAVS sensing pathway, leading to the induction of interferon and the inflammation of the tumor microenvironment, resulting in reduced tumor growth. Alternatively, tumor growth can also be inhibited through the induction of apoptosis via the PKR pathway (Ishizuka et al., 2019; Liu et al., 2019). Increased expression of ADAR1 has been found to negatively impact ZBP1-mediated PANoptosis, which can inhibit tumor growth and influence cancer development and progression in a beneficial manner (Karki et al., 2021; Zhang et al., 2022).

1.6 ADAR1 evolution

The evolution of ADAR has been analyzed through sequencing methods and the identification of A-to-I editing events. By profiling the RNA editomes across the phylogeny of Holozoa (a group that includes animals and their closest unicellular relatives), it was observed that RNA editing was present in the last common ancestor (LCA) of extant metazoans, with the presence of ADAR1 or ADAR2 homologs (Grice and Degnan, 2015; Porath et al., 2017; Zhang

et al., 2023). The first ADAR gene likely evolved from ADAT (tRNA-specific adenosine deaminase) through the acquisition of dsRBDs, which allowed it to work on dsRNA substrates (Grice and Degnan, 2015; Keegan et al., 2004).

Across the vast majority of studied metazoans, repetitive elements were found to be the primary targets of A-to-I RNA editing. The hyper-editing of repeat-derived dsRNA appears to be an ancient phenomenon, present in the LCA of extant metazoans. Furthermore, evolutionary analysis revealed conserved substrate preferences for ADAR-mediated editing, such as the 5' neighboring nucleotide, which was shown to have greater impact on editing selectivity (Zhang et al., 2023).

ADAR-mediated A-to-I RNA editing functions as a widespread regulatory mechanism involved in both transposon safeguarding and antisense-mediated gene regulation across metazoans. Notably, the evolution of RLRs has also originated in metazoans, highlighting the connection between RNA editing and innate immune responses. Additionally, RNA editing enables host organisms to better tolerate the integration of new mobile elements, which can replicate rapidly over relatively short evolutionary timescales (Mukherjee et al., 2014; Porath et al., 2017; Zhang et al., 2023).

A compelling example that illustrates the conserved functions of ADAR-mediated editing can be observed in the comparison between mouse and humans. Despite the absence of the immunostimulatory *irAlu* repeats present in the human genome, the immunosuppressive role of ADAR1 remains conserved in mice.

1.6.1 ADAR in Pacific oysters

ADAR homologs have also been identified in Pacific oysters (*Crassostrea gigas*, also known as *Magallana gigas*). Specifically, two homologs of ADAR1 and one homolog of ADAR2 have been characterized. Notably, these ADAR homologs primarily target non-coding elements within the genome, particularly repetitive sequences, which is consistent with the conservation of ADAR functions observed across different species (Rosani et al., 2019; Zhang et al., 2023).

Ostreid herpesvirus-1 (OsHV-1) is a highly infectious virus that significantly impacts bivalve species, particularly Pacific oysters. The mortality rate associated with OsHV-1 infection can be very high, ranging from 50% to 100% in both juvenile and adult oysters, with younger and smaller individuals exhibiting greater vulnerability to the disease. This virus poses a serious threat to oyster populations worldwide, resulting in substantial losses in aquaculture. It spreads primarily through direct contact with infected shellfish (King et al., 2019; Mandas and Salati, 2017).

Elevated levels of A-to-I RNA editing were observed during OsHV-1 infection, with correlated upregulation of ADAR. This suggests that the ADAR enzyme may play a role in the oysters response to OsHV-1 infection. The role of ADAR in this context remains unclear (Rosani et al., 2022, 2019).

2 Aims and objectives

2.1 Biochemical and structural insights into ADAR1 substrate selectivity properties

ADAR1 is broadly expressed across various tissues and plays a crucial role in regulating pathways associated with innate immune responses. Its A-to-I editing activity is widely distributed throughout the human genome, exhibiting diverse editing efficiencies. Dysregulation of this editing process has been linked to various diseases. Despite its importance, the molecular mechanisms guiding ADAR1's substrate selectivity remain largely unexplored, with much of the current understanding based on comparisons to the more extensively studied ADAR2.

While some functional differences between these two enzymes can be attributed to their distinct cellular localization or expression patterns, previous research has highlighted significant differences in their substrate selectivity properties. For instance, they exhibit different editing offsets due to structural disruptions, which arise from the unique architecture of their RNA binding domains (Zambrano-Mila et al., 2023). Moreover, evidence suggests that the factors influencing ADAR1's substrate selectivity are complex and often reliant on the sequence and structure of the RNA substrate, making it impossible to reduce them to simple determinants. This complexity underscores the need for more thorough investigations into the specific mechanisms that drive ADAR1's substrate recognition and editing capabilities.

By combining structural and biochemical studies, we can gain a deeper understanding of the substrate selectivity properties of ADAR1. This can be accomplished by examining the structures of ADAR1 isoforms in complex with various RNA substrates, which will ultimately improve our understanding of RNA recognition at the molecular level. Structural studies are crucial for understanding how the full-length ADAR1 enzyme aligns its domains with specific RNA features and how these domains interact to facilitate substrate selectivity. Furthermore, the role of ADAR1 dimerization is not well understood, and structural data could illuminate how the formation of dimers influences substrate selectivity.

Gaining structural insights into ADAR1 could also aid in the development of potent inhibitors targeting its catalytic or binding functions, an area of considerable interest within the pharmaceutical industry. Research has shown that inhibiting ADAR1 can enhance the efficacy of immune checkpoint inhibitors in certain cancer therapies. Furthermore, structural information could assist in designing more effective ADAR-recruiting therapeutics aimed at correcting disease-causing mutations at the RNA level through ADAR's recoding capabilities.

Objectives:

Optimization of ADAR1 purification protocol

Current literature has established a purification protocol for ADAR2, which can be successfully overexpressed in yeast or insect expression systems, producing substantial amounts of the active enzyme (Cho et al., 2003; Keegan et al., 2007; Macbeth and Bass, 2007). In contrast, ADAR1 presents greater challenges, and purification protocols for it are quite limited. Although ADAR1 is stable when purified from tissues, the yields are significantly low, only

ranging from 3 to 10 µg per kilogram of calf thymus (O'Connell and Keller, 1994). Additionally, while obtaining sufficient overexpression of ADAR1 is inherently complex, it becomes even more difficult with the longer isoform, ADAR1 p150. So far, only selected isolated individual domains of ADAR1 have been successfully purified, rather than the full-length protein. Therefore, the initial steps of this project will focus on optimizing the purification protocol for ADAR1. The activity of the purified protein will be assessed using an editing assay.

Reconstitution and biochemical characterization of ADAR1-dsRNA complex

In the initial phase, multiple substrates will be designed and synthesized using in vitro transcription. Once this is accomplished, the optimization of complex formation will be carried out to identify the optimal conditions that yield a stable complex suitable for structural analysis. The complex will be examined by employing binding techniques like Electrophoretic Mobility Shift Assay (EMSA) or by assessing binding affinities through fluorescence anisotropy with fluorescently labeled substrates. Reconstituting ADAR1/dsRNA complexes will provide insights into whether there are variations in the binding affinities of specific RNA substrates for the ADAR1 p110 and p150 isoforms.

Structural characterization of ADAR1 bound to dsRNA

The structure of the ADAR-dsRNA complex will be determined using CryoEM, which offers several advantages compared to traditional structure determination methods. It enables the acquisition of structural information for dynamic and flexible complexes, making it particularly useful for investigating the ADAR1/dsRNA complex, which can undergo various conformational changes. Furthermore, ADAR1 contains several flexible regions between its domains, which would likely pose significant challenges for crystallization. Unlike conventional methods, CryoEM does not require large sample sizes or crystalline forms. Additionally, this technique provides the capability to visualize different protein conformations, which might be crucial for a comprehensive understanding of ADAR1 substrate recognition and editing.

Initially, the ADAR1/dsRNA sample will be evaluated using negative staining and room-temperature electron microscopy to confirm its homogeneity and to check for contaminants that might interfere with CryoEM data processing. Following this, CryoEM samples will be prepared, testing various conditions. Screening will be performed with a 200 keV electron microscope, during which the complex will be examined for aggregates, particle distribution, sample concentration, and uniformity of ice thickness across the grid. Once an appropriate sample is identified, data collection will proceed with a 300 keV electron microscope to enable high-resolution data acquisition. The resulting data will then be subjected to computational analysis to elucidate the three-dimensional structure of the complex. Achieving near-atomic resolution will facilitate a deeper understanding of the molecular mechanisms involved in ADAR1 substrate recognition and editing.

2.2 Investigating the in vitro editing activity of ADAR in Pacific oysters

ADAR is found in all metazoan species and has been identified in bivalves, such as Pacific oysters, where it exhibits elevated expression levels in response to OsHV-1 viral infection. ADAR expression and A-to-I editing functions still need to be fully understood in this context. While there is evidence of editing activity in vivo, it remains uncertain whether the protein can also function effectively in vitro. The activity of ADAR can be evaluated through in vitro editing reactions, which can then be analyzed using either Sanger or nanopore sequencing. Sanger sequencing is a well-established technique for detecting modified inosines, whereas nanopore sequencing, capable of monitoring editing levels in individual RNA molecules, is still under active development. The implementation of nanopore sequencing protocols requires careful optimization, and the method is still evolving in terms of equipment improvements and the advancement of bioinformatics pipelines, including the use of machine learning approaches (Chen et al., 2023; Nguyen et al., 2022). As a result, nanopore sequencing is not yet as widely adopted by research groups, as it requires specialized tools that may not be readily available.

Objectives:

To this end, oyster ADAR will be purified using established protocols adapted from human ADAR. The protein's in vitro activity will be evaluated using Sanger sequencing, comparing its performance to that of human ADARs. Following this, the samples will undergo analysis via nanopore sequencing, which will require careful optimization based on existing published research.

2.3 MDA5 interaction with its endogenous target - irAlu repeats

Structural characterization of MDA5 - irAlu filaments

irAlu has been identified as the most abundant endogenous target for MDA5 activation (Ahmad et al., 2018). However, the mechanism by which MDA5 interacts with dsRNA containing structural irregularities remains elusive. To date, structural information has only been provided for perfectly complementary dsRNA, leaving a gap in our understanding of how MDA5 recognizes and binds to irregular RNA structures (Wu et al., 2013a; Yu et al., 2018a, 2021). Currently, the available data on studying the MDA5-irAlu interaction are limited to the analysis of filament length and frequency (Ahmad et al., 2018). Structural analysis of the MDA5-irAlu complex could provide valuable insights into their interaction.

Objectives:

To structurally characterize the MDA5-irAlu filaments, the experimental approach will involve several key steps. The first step will focus on optimizing in vitro transcription of various irAlu repeats. Following this, filaments formation will be optimized using negative staining techniques together with binding methods. The experimental setup will incorporate ATP to simulate cellular conditions. Once stable and homogeneous filament formation is achieved, structural analysis will be conducted using Cryo-EM.

Impact of A-to-I editing on MDA5 interactions with irAlu

Previous studies have demonstrated that ADAR1-dependent editing of endogenous irAlu repeats suppresses MDA5 filament formation and subsequent activation of immune sensing signaling (Ahmad et al., 2018; Mannion et al., 2014). However, the precise mechanism by which inosines within these imperfect RNA duplex structures interfere with MDA5 association remains unclear. One hypothesis is that RNA editing destabilizes the dsRNA structure by modifying A:U pairs, potentially resulting in less stable bonding post-editing. Interestingly, ADAR1 has been shown to preferentially target A:C mismatches, which can actually enhance the stability of the RNA duplex (Wong et al., 2001; Wright et al., 2018, 2007).

Moreover, high-throughput sequencing studies have shown that ADAR1 editing within repetitive elements, including irAlu sites, is generally low, often below 1% (Bazak et al., 2014). In light of these findings, it would be beneficial to investigate the effects of isolated, in vitro edited irAlu on MDA5 binding properties, as this could provide valuable insights into the functional implications of ADAR1-mediated editing.

Objectives:

To explore the effects of A-to-I editing on the binding of MDA5 to irAlu repeats, the experimental approach will involve conducting an editing assay using purified ADAR enzymes on the synthesized irAlu repeats. Following the editing process, binding interactions between MDA5 and the modified irAlu repeats will be examined in comparison to non-edited RNA.

3 Materials and methods

3.1 Molecular cloning

3.1.1 DNA agarose gel

Agarose (PeqGold Universal agarose, VWR) was dissolved to a final concentration of 1% (w/v) in 1x TAE (Tris-acetate-EDTA) buffer by heating it in a microwave. SYBR Safe (Thermo Fisher Scientific) was added (1.5 µl for every 50 ml of gel), and the gel was poured into a cassette with the appropriate comb and left to polymerize fully. DNA samples were mixed with a 6x DNA loading dye (NEB), loaded into the gel, and run for 30 minutes (min) at 130 V in 1x TAE. A 1 Kb Plus DNA ladder (Thermo Fisher) was loaded alongside the samples. Gels were imaged using a Quantum UV imager (Vilber).

3.1.2 InFusion cloning

Primers to amplify the insert were designed with a 15-20 nucleotide flanking sequence that matches the vector insertion site. Inserts were amplified using high-fidelity Q5 DNA polymerase according to the manufacturer's instructions (NEB). Vectors were linearized by PCR or treatment with restriction enzymes. Annealing temperatures were calculated using the NEB™ Calculator. DNA fragments were run on a 1% (w/v) agarose-TAE gel, followed by gel extraction using the QIAquick Gel Extraction Kit (Qiagen). DNA concentration was measured with a Nanodrop 2000 (Thermo Fisher Scientific) using a Nucleic Acids module. A total of 50 ng of insert and 100 ng of vector were mixed with 0.5 µl of In-Fusion Cloning Kit (Takara Bio) in a final volume of 10 µl and incubated for 30 min at 50 °C. 2 µl of the ligated DNA was transformed into XL10 Gold or DH5α *E. coli* chemically competent cells by heat shock for 45 s at 42 °C. 1 ml of LB (Lysogeny Broth) media was added, and the cells were recovered for 45 min at 37 °C with shaking at 300 RPM. Bacteria were plated on LB-agarose plates with the appropriate antibiotic for plasmid resistance. Bacterial colonies were cultured overnight (ON) at 37 °C. Plasmid DNA from positive colonies was grown in LB media with the appropriate antibiotic ON at 37 °C with shaking at 200 RPM. It was then extracted using the Plasmid DNA Mini Kit I (Omega Biotek). Regions containing inserts were sequenced using Sanger sequencing (Microsynth). The list of vectors and competent cells used for cloning can be found in Tables 1 and 2.

3.1.3 Site-directed mutagenesis

The mutations in the plasmid were introduced using an around-the-horn approach. 5' end-phosphorylated primers containing the desired mutations were used to amplify plasmid DNA. The primers were positioned adjacent to each other at the 5' end. The non-mutated plasmid template was removed by treatment with DpnI (NEB) for 1 hour at 37 °C. The DNA product was purified through gel extraction and ligated using T4 DNA ligase (NEB) by incubating DNA for 2 hours at room temperature (RT). The DNA was transformed into XL10 *E.Coli* competent cells. Subsequent steps were performed in the same manner as described in InFusion cloning.

Bacteria strain	Genotype	Purpose
DH5a competent <i>E. Coli</i>	fhuA2Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	For cloning and plasmid amplification
XL10 Gold ultracompetent <i>E.coli</i>	endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR- mrr)173 tetR F'[proAB lacIqZΔM15 Tn10(TetR Amy CmR)]	For cloning
DH10aEM BacY <i>E.coli</i>	F- <i>mcrA</i> Δ(<i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i>) Φ80/ <i>lacZ</i> ΔM15 Δ <i>lacX</i> 74 <i>recA1 endA1</i> <i>araD</i> 139 Δ(<i>ara, leu</i>)7697 <i>galU galk</i> λ- <i>rpsL nupG</i> /pMON14272/pMON7124	For producing bacmid DNA for insect cell expression

Table 1 List of used bacteria strains.

The competent cells were prepared in-house by Anita Testa Salmazo

Vector name	Details	Vector resistance	Purpose
438-A	pFastBac subcloning vector featuring a Histidine ₆ tag for the expression of proteins in insect cells utilizing the Bac-to-Bac baculovirus system.	Ampicillin	For insect cell expression
438-C	pFastBac subcloning vector featuring a Histidine ₆ tag and MBP (maltose-binding protein) for the expression of proteins in insect cells utilizing the Bac-to-Bac baculovirus system.	Ampicillin	For insect cell expression
438-His ₆ -CDS-TwinStrep	pFastBac subcloning vector featuring a Histidine ₆ and TwinStrep tags for the expression of proteins in insect cells utilizing the Bac-to-Bac baculovirus system.	Ampicillin	For insect cell expression
pHDV	Modified pSP64 vector (Promega) for cloning RNA templates for in vitro transcription; includes T7 promoter and hepatitis delta virus (HDV) ribozyme.	Ampicillin	For in vitro transcription of RNAs containing homogenous 3' ends
pCAGGS	Eukaryotic expression vector containing a cytomegalovirus (CMV) enhancer, the strong chicken β-actin promoter followed by a chicken β-actin intron sequence.	Ampicillin	For mammalian cell expression

pcDNA3.1	Eukaryotic expression vector containing a cytomegalovirus (CMV) enhancer-promoter.	Ampicillin	For mammalian cell expression
----------	--	------------	-------------------------------

Table 2 List of used vectors.

3.1.4 Produced plasmids

In the presented thesis, several plasmids were created and produced. The protein-coding sequences used in this work are as follows:

Protein	Sequence accession number
ADAR1 p110	Uniprot P55265-5
ADAR1 p150	Uniprot P55265-1
Δno. ADAR1 p150	Uniprot P55265-1, with deletion of the initial no. of amino acids
ADAR2	Uniprot P78563-2
MDA5 ΔCARDs	Uniprot Q9BYX4-1 with deletion of the initial 2-286 amino acids
CgADAR1v	NCBI XP_065945010.1

Table 3 List of the accession number for protein-coding sequences.

Plasmids for mammalian cell expression	
pCAGGS His ₆ -ADAR1 p110-TwinStrep	pCAGGS His ₆ -Δ33 ADAR1 p150-TwinStrep
pCAGGS His ₆ -ADAR1 p150-TwinStrep	pCAGGS His ₆ -Δ78 ADAR1 p150-TwinStrep
pcDNA3.1 His ₆ -ADAR1 p110-TwinStrep	pCAGGS His ₆ -Δ133 ADAR1 p150-TwinStrep
pcDNA3.1 His ₆ -ADAR1 p150-TwinStrep	pCAGGS His ₆ -Δ133 ADAR1 p150-TwinStrep
	pCAGGS His ₆ -Δ133 ADAR1 p150-dimer mutant (V747A, D748Q, W768V, C773S) -TwinStrep
Plasmids for insect cell expression	
438-A ADAR1 p110	438 His ₆ -TwinStrep-ADAR1 p150
438-C ADAR1 p110	438 His ₆ -TwinStrep-ADAR1 p150
438 His ₆ -TwinStrep-ADAR1 p110	438 His ₆ -ADAR1 p150-TwinStrep
438 His ₆ -ADAR1 p110-TwinStrep	438-A ADAR2
438-A ADAR1 p150	438-A CgADAR1v
438-C ADAR1 p150	438-C MDA5 ΔCARDs
Plasmids for in vitro transcription templates production	
pHDV 5HT2C	pHDV irAlu NICN1 FL
pHDV 5HT2C editing	pHDV irAlu NICN1 A sense
pHDV Gria2 R/G	pHDV irAlu NICN1 A antisense
pHDV Gria2 R/G editing	pHDV irAlu NICN1 C
pHDV NEIL1	pHDV irAlu NICN1 D
pHDV NEIL1 editing	pHDH irAlu BPNT1 FL
pHDV AJUBA	pHDH irAlu BPNT1 A
pHDV TTYH2	pHDH irAlu BPNT1 B
pHDV TTYH2 editing	

Table 4 List of produced plasmids.
RNA sequences are provided in the Appendix data.

3.2 Protein purification methods

3.2.1 Protein expression

3.2.1.1 Insect cell expression system

Protein-coding sequences were cloned into modified pFastBac 438 vectors. Plasmid DNA was electrotransformed into DH10 α EMBacY *E. coli* cells and plated on LB agar containing 10 μ g/ml gentamicin, 30 μ g/ml kanamycin, 10 μ g/ml tetracycline, 150 μ g/ml X-Gal, and 1 mM IPTG. Colonies were grown for 48 hours at 37 °C. Bacmid DNA was isolated from white colonies, and the presence of the gene insert was verified by a PCR reaction. Positive clones were transfected into Sf9 insect cells cultured in Sf-900 III SFM media at 27 °C using Fugene transfection reagent (Promega). Media containing baculoviruses (V0) was collected after 48-72 hours and further amplified in Sf9 suspension culture in the same media with shaking at 125 RPM (V1). Protein expression was carried out in High Five cells (Thermo Fisher Scientific) cultured in ESF21 media (Oxford Expression Technologies) by infecting them with V1 protein-containing baculoviruses. Cells were grown for 48-72 hours at 27 °C with shaking at 125 RPM, harvested by centrifugation, and stored at -80 °C until needed.

3.2.1.2 Mammalian cell expression system

Protein coding sequences were cloned into pCAGGS or pcDNA3.1 vectors. Proteins were transiently expressed in Expi293F mammalian cells (Thermo Fisher Scientific) maintained at 37 °C, 8% CO₂ in Expi293 media (Thermo Fisher Scientific) with shaking at 125 RPM. Cells were transfected using appropriate plasmid purified with NucleoBond Xtra Maxi kit (Macherey-Nagel) and PEI (25-kDa linear polyethylenimine, Thermo Scientific). Following transfection the cells were cultured for 42-48 hours, harvested by centrifugation, and stored at -80 °C until needed.

3.2.2 Protein analysis

3.2.2.1 SDS-Page gel electrophoresis

Purified proteins or cell lysates were mixed with 5x SDS-PAGE loading buffer to a final concentration of 1x (5x LB: 250 mM Tris-HCl (pH 7.0 at 25 °C), 30% (v/v) glycerol, 5% SDS, 0.025% (w/v) bromophenol blue, 5% (v/v) beta-mercaptoethanol), boiled for 5 min at 95 °C and loaded onto Bolt 4-12% Bis-Tris Plus Gels (Thermo Fisher Scientific). The gel was run in a 1x MOPS buffer (Thermo Fisher Scientific) for 50 min at a constant voltage of 170V. Precision Plus Protein Dual Color Standards or Precision Plus Protein Unstained Protein Standards (Bio-Rad) were loaded alongside the samples. Gels were stained with InstantBlue® Coomassie Protein Stain (Abcam) and imaged using a gel scanner (CANON 9000F M II, Reichelt Elektronik).

3.2.2.2 Measuring protein concentration

Protein concentration was measured with a Nanodrop 2000 (Thermo Fisher Scientific) using a Protein A280 nm module. The instrument was blanked with a protein storage buffer and the concentration was calculated based on the extinction coefficient calculated by the ProtParam tool (Expasy).

3.2.2.3 Western blot against His₆ tag

The protein samples were prepared as described in the SDS-PAGE gel electrophoresis protocol. After the run was completed, proteins were transferred to a PVDF Transfer Membrane (Thermo Fisher Scientific) activated in 100% methanol using a Mini Blot Module (Fisher Scientific) in 1x Bolt Transfer Buffer (Thermo Fisher Scientific). The protein transfer was conducted according to the manufacturer's guidelines. The membrane was blocked with 5% (w/v) milk (Lactan) in the PBS-T buffer which consists of phosphate-buffered saline supplemented with 0.1% (v/v) Tween-20. The blocking step was carried out for 1 hour at RT. The membrane was then incubated with anti-His antibodies conjugated with HRP (Miltényi Biotec) diluted to 1:10 000 in 1% (w/v) milk/PBS-T. This incubation also took place at RT for 1 hour. To remove any unbound antibodies, the membrane was rinsed five times with PBS-T, with each wash lasting approximately 5 min. Proteins were detected by applying SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific) to the membrane, and the signal was visualized using a GelDoc Go Gel Imaging System (Bio-Rad) with careful attention to adjusting the exposure time to prevent oversaturation of the signal, ensuring clear and accurate protein band detection.

3.2.3 Optimization of ADAR1 purification protocol

3.2.3.1 Expression optimization in mammalian cells

To determine the expression vector for mammalian cells, Expi293F cells were transfected with either a pcDNA3.1 or pCAGGS plasmid containing ADAR1 p110 or p150 (pCAGGS His₆-TwinStrep-ADAR1 or pcDNA3.1 His₆-TwinStrep-ADAR1). Cell expression was conducted in a 25 ml small-scale culture as described above. Two cell densities were tested during the transfection - 1×10^6 or 3×10^6 cells/ml. The viable cell count was measured after 48 hours of expression, and 1.5×10^4 cells were collected, mixed with 5x SDS-PAGE loading buffer, boiled for 5 min at 95 °C, and analyzed using SDS-PAGE gel electrophoresis.

ADAR1 p110 or p150 (pCAGGS His₆-ADAR1-TwinStrep) were transiently expressed in Expi293F cells to optimize the expression time. A 100 µl cell suspension was collected after 17, 21, 26, 41, 45, 50, 65, 69, and 74 hours, mixed with 5x SDS-PAGE loading buffer, boiled for 5 min at 95 °C and stored at -20 °C until use. A western blot against the His₆ tag was employed to assess the protein expression level. The expression time was determined based on the intensities of the detected bands, measured using the ImageJ gel analysis tool (ImageJ version 1.52k).

3.2.3.2 Buffer screening for protein solubility

The initial buffer optimization was performed using homemade buffers with pH values ranging from 3 to 10. All buffers included 200 mM NaCl and 1 mM DTT, and some contained 25 U of benzonase (Merck). A list of the tested buffers is provided in the Table 5. Tests were conducted using His₆-MBP-ADAR1 p150 expressed in High Five insect cells. For each condition, a total of 10×10^6 cells were suspended in 800 µl of buffer, sonicated, and incubated on ice for 30 min. A 20 µl sample was saved for SDS-PAGE analysis (total lysate). The cells were then centrifuged to clarify the lysates, and another sample for SDS-PAGE was prepared (supernatant). To evaluate protein solubility, samples were analyzed by SDS-PAGE gel electrophoresis as described above.

Buffer composition

50 mM Glycine pH 3.0, 200 mM NaCl, 1 mM DTT

1x SPG pH 4.0, 200 mM NaCl, 1mM DTT
50 mM MES pH 6.0, 200 mM NaCl, 1 mM DTT
1x SPG pH 4.0, 200 mM NaCl, 1 mM DTT
1x SPG pH 4.0, 200 mM NaCl, 1 mM DTT, 25U benzonase
50 mM PIPES pH 7.0, 200 mM NaCl, 1 mM DTT
50 mM sodium phosphate pH 7.5, 200 mM NaCl, 1 mM DTT
50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT
1x SPG pH 10.0, 200 mM NaCl, 1 mM DTT

Table 5 Buffer list used for the initial protein solubility screen.

The next round of the buffer solubility test was performed using thermal unfolding with nanoDSF Prometheus technology. The buffer screening was feasible after acquiring ADAR1 p110 at high purity from the mammalian expression system. The protein was diluted in various buffers (Fig. 15) to a final concentration of 0.18 mg/ml and loaded into standard capillaries. Protein solubility was assessed by measuring two wavelengths, 330 and 350 nm, and plotting their ratio against temperature.

Buffer - Anions screen (96 wells)

1+4		1	2	3	4	5	6	7	8	9	10	11	12	
Buffer & Salt	A	H2O	pH 2	pH 3	pH 4	pH 5	pH 6	pH 7	pH 8	pH 9	pH 10	pH 11	pH 12	
			maleate	glycine	formate	citrate	cacodylate	HEPES	bicine	CHES	borate	CAPS	phosphate	
		250 mM	500 mM	500 mM	500 mM	500 mM	500 mM	500 mM	500 mM	500 mM	250 mM	500 mM	500 mM	
	B	pH 4.0	pH 4.5	pH 5.0	pH 5.5	pH 6.0	pH 6.5	pH 7.0	pH 7.5	pH 8.0	pH 8.5	pH 9.0	pH 9.5	
		acetate	acetate	acetate	MES	MES	MES	Na-phosphate	K-phosphate	Tris/HCl	Tris/HCl	glycine	glycine	
		500 mM	500 mM	500 mM	500 mM	500 mM	500 mM	500 mM	500 mM					
	C	pH 6.0				pH 7.0				pH 8.0				
		500 mM MES				500 mM Na-phosphate				500mM Tris				
		0.5M NaCl	1M NaCl	2.5 M NaCl	5 M NaCl	0.5M NaCl	1M NaCl	2.5 M NaCl	4.6 M NaCl	0.5M NaCl	1M NaCl	2.5 M NaCl	5 M NaCl	
	D	pH 7.5												
		std buffer 1	std buffer 2	std buffer 3	std buffer 4	imidazole	tween	glycerol	bME	DMSO	Trehalose	Arg + Glu	EDTA	
						1 M	0.25%	25%	25 mM	25.00%	25%	100+100 ml	25 mM	

D

std buffer 1 100 mM Tris, 750 mM NaCl, 0.5 mM CaCl2	7.4
std buffer 2 50 mM Hepes, 750 mM NaCl, 0.25% Tween	7.4
std buffer 3 PBS: 60mM PO4 (72 g/l Na2HPO4.2H2O, 12g/l KH2PO4), 685 mM NaCl, 13.5 mM KCl	7.4
std buffer 4 PBS, 0.25% Tween	7.4

Figure 15 Buffers composition used in the protein solubility by thermal unfolding.

3.2.3.3 Optimized ADAR1 purification protocol

Recombinant ADAR1 was transiently expressed in Expi293F mammalian cells using the pCAAGS His₆-ADAR1-TwinStrep plasmid with a 3C cleavage site between the tags. The purification procedure was conducted at 4 °C using an ÄKTA purification system (Cytiva). Approximately 1.5L of culture was lysed in A1 lysis buffer containing 50 mM HEPES-NaOH (pH 8.0 at 4 °C), 500 mM NaCl, 20% (v/v) glycerol, 1 mM DTT, and a protease inhibitor mix (1 mM PMSF, 2 mM benzamidine, 1 μM leupeptin, 2 μM pepstatin), then sonicated and subjected to ultracentrifugation. The clarified lysate was loaded onto a HisTrap Excel 5 ml column (Cytiva) equilibrated with lysis buffer, and unbound proteins were washed away with A1 buffer 2 (50 mM HEPES-NaOH (pH 8.0 at 4 °C), 500 mM NaCl, 20% (v/v) glycerol, 5 mM EDTA, 1 mM DTT) containing 50 mM imidazole. Contaminating chaperones were removed with an A1 lysis

buffer containing 4 mM ATP-MgCl₂. The protein was eluted with A1 buffer 2 with a linear gradient of 50- 500 mM imidazole, then loaded onto a StrepTrap 1 ml column (Cytiva) equilibrated with A1 buffer 2 and eluted with A1 buffer 3 (50 mM HEPES-NaOH (pH 8.0 at 4 °C), 500 mM NaCl, 20% (v/v) glycerol, 1 mM EDTA, 1 mM DTT, 2.5 mM desthiobiotin). Affinity tags were removed using homemade 3C protease, and the protein was subjected to size exclusion chromatography using a HiLoad 16/600 Superdex200 column (Cytiva) equilibrated with A1 SEC buffer (50 mM HEPES-NaOH (pH 8.0 at 4 °C), 350 mM NaCl, 15% (v/v) glycerol, 0.5 mM EDTA, 1 mM DTT). Fractions containing ADAR1 were concentrated with an Amicon Ultra-4 concentrator (50 kDa cutoff), aliquoted, and stored at -80 °C. The protein yield was approximately 300-500 µg per liter of cell culture.

3.2.4 ADAR2 purification protocol

The purification protocol was optimized by Carrie Bernecky. Recombinant ADAR2 was expressed in High Five insect cells by infecting them with V1 baculoviruses containing His₆-3C-ADAR2. The purification procedure was conducted at 4 °C using an ÄKTA purification system (Cytiva). Cells were lysed in A2 buffer 1 (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 500 mM NaCl, 10% (v/v) glycerol, 1 mM DTT) supplemented with 30 mM imidazole and a protease inhibitor mix (1 mM PMSF, 2 mM benzamidine, 1 µM leupeptin, 2 µM pepstatin), then sonicated and ultracentrifuged. The clarified lysate was loaded onto a HisTrap 5 ml column (Cytiva) equilibrated with the same buffer. Unbound proteins were washed away with A2 buffer 1 containing 50 mM imidazole, followed by a high salt wash (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 1 M NaCl, 10% (v/v) glycerol, 1 mM DTT). Protein was eluted in A2 buffer 2 (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 150 mM NaCl, 10% (v/v) glycerol, 1 mM DTT) containing 250 mM imidazole and loaded onto a HiTrap Heparin 5 ml column (Cytiva) equilibrated with A2 buffer 3 (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 150 mM NaCl, 10% (v/v) glycerol, 1 mM DTT), then eluted with a linear gradient from 150 mM to 1 M NaCl. Eluted fractions were cleaved with homemade 3C protease, loaded for reverse HisTrap chromatography, and eluted with A2 buffer 1 containing 50 mM imidazole. The eluted fractions were further purified via size exclusion chromatography using a HiLoad 16/600 Superdex200 column (Cytiva) equilibrated with A2 SEC buffer (20 mM HEPES-NaOH (pH 7.5 at 4 °C), 300 mM NaCl, 10% (v/v) glycerol, 1 mM DTT). ADAR2-containing fractions were concentrated using an Amicon Ultra-4 concentrator (30 kDa cutoff), aliquoted, and stored at -80 °C. The protein yield was approximately 7 mg per liter of cell culture.

3.2.5 ADAR1 / ADAR2 deaminases purification protocol

Recombinant ADAR1 or ADAR2 deaminases were expressed in High Five insect cells by infecting them with V1 baculoviruses containing His₆-3C-ADAR deaminase. The purification procedure was conducted at 4 °C using an ÄKTA purification system (Cytiva). Cells were lysed under the same condition as described in the ADAR2 purification section. The clarified lysate was loaded onto a HisTrap 5-ml column (Cytiva) equilibrated with the A2 buffer 1 containing 30 mM imidazole. Unbound proteins were washed away with A2 buffer 1 supplemented with 50 mM imidazole, followed by a high salt wash (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 1 M NaCl, 10% (v/v) glycerol, 1 mM DTT). Protein was eluted in A2 buffer 2 (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 150 mM NaCl, 10% (v/v) glycerol, 1 mM DTT) using a linear gradient from 50 to 500 mM imidazole. It was then loaded onto a HiTrap Heparin 5 ml column (Cytiva) equilibrated with A2 buffer 3 (50 mM HEPES-NaOH (pH 7.5 at 4 °C), 10% (v/v) glycerol, 1 mM

DTT), containing 50 or 100 mM NaCl for ADAR1 and ADAR2 deaminases, respectively, and subsequently eluted with a linear gradient up to 1 M NaCl. The eluted fractions were further purified via size exclusion chromatography through multiple sample applications using a Superdex75 Increase 10/300 column (Cytiva) equilibrated with A2 SEC buffer. ADAR deaminase-containing fractions were concentrated using an Amicon Ultra-4 concentrator (30 kDa cutoff), aliquoted, and stored at -80 °C. The protein yield was approximately 2-3 mg for ADAR2 deaminases and 125-750 µg for ADAR1 deaminases per liter of cell culture.

3.2.6 Pacific oyster ADAR purification protocol

Recombinant CgADAR1v was expressed in High Five insect cells by infecting them with V1 baculoviruses containing His₆-3C-CgADAR1v and subsequent purification steps were followed as described in the ADAR2 purification protocol section.

3.2.7 MDA5 ΔCARDs purification protocol

The purification protocol was optimized by David Michalik, after which minor adjustments were implemented. Recombinant MDA5 ΔCARDs was expressed in High Five insect cells by infecting them with V1 baculoviruses containing His₆-MBP-3C-MDA5 ΔCARDs. The purification procedure was conducted at 4 °C using an ÄKTA purification system (Cytiva). Cells were lysed in M5 buffer 1 (50 mM sodium-phosphate pH 6.0, 400 mM NaCl, 20% glycerol (v/v), 1 mM DTT) supplemented with 20 mM imidazole and a protease inhibitor mix (1 mM PMSF, 2 mM benzamidine, 1 µM leupeptin, 2 µM pepstatin), followed by sonication and ultracentrifugation. The clarified lysate was loaded onto a HisTrap 5-ml column (Cytiva) equilibrated with the same buffer. Unbound proteins were washed away with M5 buffer 1 and a high salt wash (50 mM sodium-phosphate pH 6.0, 2 M NaCl, 20% glycerol (v/v), 30 mM imidazole, 1 mM DTT). Contaminating chaperones were removed using M5 ATP buffer (50 mM sodium-phosphate pH 6.0, 100 mM NaCl, 20% glycerol (v/v), 30 mM imidazole, 3 mM ATP-MgCl₂, 1 mM DTT). Protein was eluted using M5 buffer 2 (50 mM BisTris (pH 6.5 at 4 °C), 400 mM NaCl, 20% glycerol (v/v), 1 mM DTT) with a linear gradient from 20 to 500 mM imidazole. Fractions containing MDA5 were pooled and dialyzed ON in M5 dialysis buffer (50 mM BisTris (pH 6.5 at 4 °C), 200 mM NaCl, 10% glycerol (v/v), 5 mM MgCl₂, 1 mM DTT) with the addition of a homemade 3C protease and protein phosphatase 1 α (PP1α). The treated protein was subjected to a HiTrap Heparin 5-ml column (Cytiva) equilibrated with M5 buffer 3 (50 mM sodium-phosphate pH 6.0, 100 mM NaCl, 10% glycerol (v/v), 1 mM DTT) and eluted with a linear gradient from 100 mM to 2 M NaCl. Subsequently, the protein was loaded for reverse HisTrap chromatography and eluted with M5 buffer 2 containing up to 40 mM imidazole. The MDA5-containing fractions were concentrated using an Amicon Ultra-4 concentrator (30 kDa cutoff) and further purified via size exclusion chromatography using a Superdex75 Increase 10/300 column (Cytiva) equilibrated with M5 SEC buffer (50 mM BisTris (pH 6.5 at 4 °C), 150 mM NaCl, 10% glycerol (v/v), 1 mM DTT). Protein was aliquoted, and stored at -80 °C. The protein yield was approximately 4 mg per liter of cell culture.

3.2.8 Testing the activity of various ADAR proteins

RNA substrates used in the editing reactions had flanking sequences GGUCUUGUCUCUGUGGUCUG/ CUCUCGUCGCGUUGUCCUU that were used for generating the cDNA and amplifying the PCR product.

For initial testing of various ADAR protein activities, RNA was refolded in 10 mM Tris-HCl (pH 8.0 at 25 °C), 50 mM NaCl, and 1 mM EDTA by heating for 5 min at 95 °C, followed by snap cooling on ice. The reaction was carried out in a final volume of 20 µl, consisting of 25 nM RNA and 500 nM protein in 20 mM Tris-HCl (pH 8.0 at 25 °C), 30 mM NaCl, 50 mM KCl, 0.5 mM EDTA, 15% (v/v) glycerol, and 0.5 mM DTT. RNA was edited for 1 hour at 30 or 37 °C.

To prepare the samples for the nanopore sequencing, RNA was refolded in 10 mM HEPES-NaOH (pH 7.5 at 25 °C), and 50 mM NaCl by heating for 5 min at 95 °C, followed by snap cooling on ice. The reaction was carried out in a final volume of 50 µl, consisting of 100 nM RNA and 1 µM protein in 20 mM HEPES-NaOH (pH 7.5 at 25 °C), 30 mM NaCl, 50 mM KCl, 0.5 mM EDTA, 10% (v/v) glycerol, and 1 mM DTT. RNA was edited for 1 hour at 30 °C.

For testing the activity of various deaminases RNA was refolded in 10 mM HEPES-NaOH (pH 7.5 at 25 °C), and 50 mM NaCl by heating for 5 minutes at 95 °C, followed by snap cooling on ice. The reaction was carried out in a final volume of 20 µl, consisting of 250 nM RNA and 2.5 µM protein in 20 mM HEPES-NaOH (pH 7.5 at 25 °C), 100 mM NaCl, 5% (v/v) glycerol, 0.05% NP-40, and 0.5 mM DTT. RNA was edited for 1 hour at 30 °C.

Further steps were carried out in the same manner for all reactions. After editing was completed, RNA was purified using a Monarch RNA clean-up kit (NEB). cDNA was generated with MaximaH Minus reverse transcriptase (Thermo Fisher Scientific) according to the manufacturer's guidelines. cDNA was then amplified using high-fidelity Q5 DNA polymerase following the manufacturer's instructions (NEB). The PCR product was purified and sent for Sanger sequencing (Microsynth). The sequence traces were analyzed using 4Peaks (Nucleobytes).

3.2.9 Checking proteins homogeneity using mass photometry

The homogeneity of different purified ADAR proteins was checked with mass photometry using a Refeyn TwoMP mass photometer (Refeyn Ltd.). Proteins were diluted to 100 nM in a buffer containing 20 mM HEPES-NaOH (pH 7.5 at 25 °C), 150 mM NaCl, 2% (v/v) glycerol, and 1 mM DTT. Calibration was done with the same buffer, into which 2-5 µl of the diluted protein was added and the recording lasted 2 min. A BSA standard, diluted in the same buffer, was used for the mass calibration. Data was processed using DiscoverMP software.

3.3 RNA substrates preparation

3.3.1 In vitro transcription, RNA purification and folding

RNA coding sequences were cloned into the pHDV vector. The plasmids were utilized to generate a PCR template containing the T7 promoter and RNA sequence. RNAs were in vitro transcribed using a homemade T7 RNAP for 4 hours at 37 °C, after which the PCR template was removed by treatment with DNaseI (Promega). RNA was loaded onto an 8% urea-acrylamide gel, visualized by UV shadowing; the bands were crushed, and RNA was extracted ON with 0.3 M NaOAc and 1 mM EDTA. The following day, RNA was filtered to remove gel pieces, ethanol precipitated, dissolved in water, and stored at -80 °C. irAlu NICN1 A strands, both sense and antisense, were annealed by mixing them in equal molar ratios, heating for 5 min at 95 °C, followed by controlled cooling to 4 °C (1 °C/30s). irAlu NICN1 of BPNT1 full length was not subjected to the folding procedure; the shorter variants of irAlu NICN1 or BPNT1 were

folded by heating for 5 min at 95 °C with controlled cooling to 4 °C (1 °C/30s). All other RNAs were refolded by heating for 5 min at 95 °C with snap cooling on ice. A list of the produced RNAs and their sequences is provided in Appendix data.

3.3.2 Urea-acrylamide denaturing gel

RNA was mixed with a 2x urea loading buffer (2x LB: 100 mM Tris-HCl (pH 8.0 at 25 °C), 100 mM Boric acid, 2 mM EDTA, 8 M urea, 0.04% (w/v) Bromophenol blue, 0.04% (w/v) Xylene-cyanol), boiled for 5 min at 95 °C, and loaded onto an 8% acrylamide-urea-TBE (Tris-boric acid-EDTA) gel. The gel was run in a 1x TBE buffer (130 mM Tris, 45 mM boric acid, 2.5 mM EDTA) for 30-60 min at a constant 300 V (time was adjusted to RNA length). A low-range ssRNA ladder (NEB) was loaded alongside the samples. To visualize the RNA, it was stained with SYBR Gold (Thermo Fisher Scientific) and imaged using a Quantum UV imager (Vilber).

3.3.3 RNA fluorescent labeling

RNA was fluorescently labeled through 3' end oxidation. One nanomole of RNA was incubated with a 10-fold excess of sodium periodate (NaIO₄) in 100 mM NaOAc. After 90 min of incubation at 25 °C, the oxidation reaction was stopped by adding a 2-fold excess of sodium thiosulfate (relative to NaIO₄). ATTO-488 hydrazide was added to the reaction in a 30-fold excess over RNA, and the labeling reaction proceeded for 4 hours at 25 °C. RNA was phenol-chloroform extracted, ethanol precipitated, and further purified using the Monarch RNA clean-up kit (NEB) to eliminate unreacted dye. To assess RNA integrity, it was loaded onto an 8% acrylamide urea gel, and after the run was completed, it was imaged with an Amersham Typhoon RGB 9400 scanner (Cytiva). After scanning, the gel was stained with SYBR Gold and imaged using a Quantum UV imager (Vilber).

3.4 ADAR1-RNA complex preparation and characterization

3.4.1 Checking complex formation with Native-PAGE

Refolded RNA was mixed with ADAR1 under various conditions (details of the specific experimental conditions are in the results section). The complex was formed by incubating for 20 min at 25 °C. A 10% (v/v) glycerol solution was added to the samples, which were then loaded onto a NativePAGE 3 to 12% Bis-Tris gel (Invitrogen). Gels were run in 1x NativePAGE running buffer (Invitrogen) at RT (50 min, constant 170V). If fluorescent RNA was used, it was imaged with an Amersham Typhoon RGB 9400 scanner (Cytiva); otherwise, it was stained with SYBR Gold and imaged using a Quantum UV imager (Vilber).

3.4.2 Measuring binding affinity using fluorescence anisotropy

NEIL1 RNA was refolded in 10 mM Tris-HCl (pH 7.5 at 25 °C), 50 mM NaCl, and 1 mM EDTA by heating for 5 min at 95 °C followed by snap cooling on ice. A total of 5 nM RNA was mixed with increasing concentrations of ADAR1 p110 (purified from High Five insect cells) ranging from 0 to 1500 nM. The complex was prepared in a final volume of 20 µl containing 20 mM Tris-HCl (pH 7.5 at 25 °C), 150 mM NaCl, 2% (v/v) glycerol, 0.5 mM EDTA, and 1 mM DTT, and incubated for 25 min at 25 °C. The experiment was conducted in triplicate in a 384-well plate. Fluorescence anisotropy was measured using a Synergy H1-MF Plate Reader (Bio-TEK). Data were analyzed in GraphPad Prism 9 using one site-specific binding equation.

3.4.3 Complex crosslinking

3.4.3.1 BS3

RNA was refolded in 10 mM HEPES-NaOH (pH 7.5 at 25 °C) and 50 mM NaCl by heating for 5 min at 95 °C, followed by snap cooling on ice. A total of 0.25 μM RNA was mixed with 0.5 μM Δ133 ADAR1 p150 in 25 mM HEPES-NaOH (pH 7.5 at 25 °C), 150 mM NaCl, and 2% (v/v) glycerol; the complex was incubated for 5 min at 25 °C. Subsequently, BS3 (bis(sulfosuccinimidyl)suberate; Thermo Fisher Scientific) was added to the reaction (ranging from 0.025 to 2 mM) and incubated for 5 min at RT with shaking at 300 RPM. The crosslinking reaction was terminated by adding 60 mM Tris-HCl (pH 8.0 at 25 °C) with 15 min of incubation at RT. The efficiency of complex crosslinking was assessed using SDS-PAGE electrophoresis (the gel was run for 90 min at 170V) and Native-PAGE as described in previous sections.

3.4.3.2 Glutaraldehyde

RNA was refolded in 10 mM HEPES-NaOH (pH 7.5 at 25 °C), 50 mM NaCl, and 1 mM EDTA by heating for 5 min at 95 °C, followed by snap cooling on ice. A total of 0.41 μM RNA was combined with 0.82 μM Δ133 ADAR1 p150 in 25 mM HEPES-NaOH (pH 7.5 at 25 °C), 150 mM NaCl, 2% (v/v) glycerol, and 0.5 mM EDTA; the complex was incubated for 10 min at 25 °C. Subsequently, 0.05% (v/v) glutaraldehyde (Sigma Aldrich) was added to the reaction and incubated for 5 min on ice. The crosslinking reaction was terminated by adding 50 mM Tris-HCl (pH 8.0 at 25 °C), followed by 10 min of incubation at RT. The efficiency of complex crosslinking was assessed using SDS-PAGE electrophoresis (the gel was run for 90 min at 170V) and Native-PAGE as described in previous sections.

3.4.3.3 SPB

RNA was refolded in 10 mM HEPES-NaOH (pH 7.5 at 25 °C), 50 mM NaCl, and 1 mM EDTA by heating for 5 min at 95 °C, followed by snap cooling on ice. RNA and Δ133 ADAR1 p150 complexes were prepared in 20 mM HEPES-NaOH (pH 7.5 at 25 °C), 150 mM NaCl, 2% (v/v) glycerol, and 0.5 mM EDTA. Different protein concentrations (0.41 and 0.82 μM, each at 2× molar excess over RNA) were incubated for 10 min at 25 °C. 25 or 50 μg/ml SPB (succinimidyl-[4-(psoralen-8-yloxy)]-butyrate; Thermo Fisher Scientific) were added to the reaction to initiate crosslinking, which was conducted under a UV lamp at 365 nm, with a 3 mm glass plate blocking UV <300 nm. Crosslinking lasted 30 min, after which the mixture was spun to remove aggregates. The crosslinking efficiency was assessed by SDS-PAGE (gel run for 90 min at 170V) and Native-PAGE, with validation from mass photometry.

Mass photometry measurements of the complex were conducted using a Refeyn TwoMP mass photometer (Refeyn Ltd.). The instrument was calibrated with a reaction buffer, into which 1-2 μl of the complex was added. The recording lasted for 2 min. Before measurements, the instrument was calibrated with BSA diluted in the reaction buffer. Data processing was performed using DiscoverMP software.

3.5 CryoEM sample preparation

3.5.1 Preparing continuously coated carbon grids

Mica sheets were coated with varying carbon thicknesses (~5 nm for negative staining and ~1.5 nm for cryoEM experiments). Mica sheets were prepared by Anita Testa Salmazo. A carbon film was floated off the mica onto a clean water surface, with grids placed underneath.

The water level was then lowered to allow the carbon to settle onto the grids. Grids were left for drying and stored in the dark until needed. 400 mesh copper grids were used for negative staining, while QUANTIFOIL or C-flat holey carbon grids with varying hole diameters were employed for cryoEM experiments (Science Services).

3.5.2 Preparing graphene-oxide-coated grids

Grids were prepared following the protocol established by (Palovcak et al., 2018). Graphene oxide (GO, Sigma Aldrich) was diluted to 0.2 mg/ml in 83% (v/v) methanol and sonicated to disperse aggregates. The mixture was centrifuged to remove the supernatant with smaller fragments, followed by another round of sonication and centrifugation before diluting the GO again to approximately 0.2 mg/ml in the same solution. The GO was then floated on the water's surface in a Pyrex petri dish with QUANTIFOIL grids placed underneath on a copper mesh platform. Water was gradually removed using a peristaltic pump to allow the GO to settle onto the grids. Before this, the grids were glow-discharged for 30 seconds at ~25 mA (ELMO device), and all utensils were thoroughly washed with chloroform, ethanol, and water to eliminate dust particles that could interfere with the integrity of the GO.

3.5.3 Negative staining

Carbon-coated 400 mesh copper grids were glow discharged for 20 seconds at ~25 mA. A 4 μ l sample (protein concentration 50-100 μ g/ μ l) was applied to the grid and incubated for 2 min. The grid was rinsed with a drop of water, followed by three 20-second washes in 2% (w/v) uranyl formate or 2% (w/v) uranyl acetate. Excess stain was removed with filter paper, and the grid was left to dry. The grids were then imaged using a 120kV Tecnai T12 transmission electron microscope equipped with a bottom-mounted TEM CMOS camera.

3.5.4 Sample vitrification

The CryoEM QUANTIFOIL or C-flat grids without a support layer were glow-discharged for 1 minute at 25 mA, while the carbon or graphene oxide-coated grids were glow-discharged for 10 seconds at 25 mA (details about grid specifications can be found in the respective results sections). A 4 μ l sample was applied to a grid, the excess sample was blotted away, and the sample was plunge-frozen in liquid ethane. For samples prepared with the Vitrobot Mark IV, the device was set to 100% humidity and 4 °C; the grids were blotted with a force of 25 for 2 to 4 seconds. For samples prepared with the Leica GP2 device, the instrument was set to 85% humidity at 5 °C, and the grids were front-blotted for 4 to 6 seconds. Before inserting the grids into the CryoEM microscope, they were clipped.

3.5.5 CryoEM grid screening and data collection

CryoEM grids were screened using a 200kV Glacios Cryo-Electron Microscope equipped with a Falcon III direct detector camera (Thermo Fisher Scientific). Data collection was performed with the FEI EPU package. During screening, the grids were evaluated for overall quality, ice thickness, particle distribution, homogeneity, and density. High-quality images were captured at a nominal magnification of 150kx, corresponding to a pixel size of 0.98Å. Counting mode was employed with a total dose of 50e-/Å² and an exposure time of ~50 s distributed over 36 frames, utilizing -3.5 μ m defocus. Δ 133 ADAR1 p150-Gria2 R/G data set was collected using

the same settings as described above with defocus values ranging from -0.75 to -3.5 μm . In total, 310 micrographs were collected.

To obtain a high-quality dataset from the $\Delta 133$ ADAR1 p150-TTYH2 complex crosslinked with SPB, the data was recorded using a 300 kV Titan Krios Cryo-Electron Microscope equipped with a Gatan K3 BioQuantum direct electron detector (Thermo Fisher Scientific). Data collection was performed using the FEI EPU package. Micrographs were recorded at a nominal magnification of 105kx, corresponding to a pixel size of 0.835 \AA . Counting mode was employed with a total dose of 60.4 $\text{e}/\text{\AA}^2$ and an exposure time of 2.4 s distributed over 40 frames, utilizing a defocus range from -1 to -2.2 μm . In total, 21 963 micrographs were collected.

3.6 CryoEM data processing and analysis

3.6.1 Data processing

For the $\Delta 133$ ADAR1 p150-TTYH2 data from the Titan Krios Initial micrographs data processing was performed using Warp, which included motion correction, estimation of the contrast transfer function, and automated particle picking (Tegunov and Cramer, 2019). The exported particles were then processed through a combination of 2D and 3D classification in RELION 3.1.0 (Zivanov et al., 2018). Further data processing was not feasible as the 3D classification did not yield satisfactory model quality due to particle heterogeneity.

Initial micrograph data processing of the $\Delta 133$ ADAR1 p150-Gria2 R/G data set from the Glacios microscope was conducted using Warp as described above. For particle picking, the software was trained to identify particles approximately 15 nm in size before they were picked (approximate size of ADAR1 dimer with RNA substrate). The exported particles were then processed through a combination of 2D and 3D classification followed by model refinement using CryoSparc (Punjani et al., 2017). A deaminase domain was superimposed into the model using UCSF Chimera software.

3.6.2 Structure prediction using AlphaFold

The first ADAR1 structure was predicted using the AlphaFold2 web server (<https://alphafold.ebi.ac.uk>), which already had a structure for human ADAR (UniProt accession P55265, AF-P55265-F1). This structure was downloaded and analyzed using ChimeraX software.

Following the release of AlphaFold3 (<https://alphafoldserver.com>), predictions for ADAR-RNA complexes became feasible. To predict the structure of the ADAR1-RNA complex, the server received two copies of the coding sequence for human ADAR1 p150 (UniProt P55265-1) with a deletion of the first 133 amino acids ($\Delta 133$ ADAR1 p150), along with the TTYH2 RNA sequence and four Zn^{2+} ions. For the ADAR2 - RNA complex prediction the server received two copies of the human ADAR2 sequence (Uniprot P78563-2) along with the TTYH2 RNA sequence and two Zn^{2+} ions. The predicted structures were subsequently analyzed using ChimeraX software (version 1.6.1).

The structure of CgADAR1v (NCBI XP_065945010.1) was also predicted using AlphaFold3. The server received a single copy of the cgADAR1v protein coding sequence and one Zn^{2+} ion. The obtained structure was analyzed using the ChimeraX software (version 1.6.1).

4 Results

4.1 ADAR1 purification protocol optimization

The initial phase of the PhD work primarily focused on developing a purification protocol for ADAR1. Existing protocols were inadequately described or likely to yield low results (Cho et al., 2003). Since ADARs require IP6 for activity (Macbeth et al., 2005) and undergo post-translational modifications (Keegan et al., 2023), a eukaryotic expression system, such as insect or mammalian cells, was used for optimization.

4.1.1 Choosing the optimal construct for ADAR1 p110 purification

The optimization of the purification protocol has been initiated with the shorter isoform ADAR1 p110, which lacks the Z α domain, thereby facilitating easier handling. Once the technique was established, the rationale was to apply a similar strategy to the longer isoform.

The initial purification step involved affinity chromatography, which separates proteins based on the affinity between the tagged protein and its ligand on the chromatography matrix. In the optimization process, various tags were tested, including a His₆ tag in combination with Ni²⁺ ions on the matrix, an MBP (Maltose Binding Protein) tag utilizing amylose resin, and a TwinStrep tag that interacts with Strep-Tactin, a modified form of streptavidin.

Attempts to purify the ADAR with an N-terminal His₆-MBP tag resulted in a few issues, primarily the protein degradation and the presence of a chaperone protein, indicating improper folding (Fig. 16A). The significant size of MBP (42.5 kDa) may hinder proper folding, contributing to these complications. Further experimentation with the His₆ tag or His₆-TwinStrep tag attached to the N-terminus of ADAR1 similarly yielded a degraded protein (see Fig. 16B), indicating that this configuration was also inefficient for maintaining protein integrity. However, a successful strategy was found by placing the His₆ tag at the N-terminus and the TwinStrep tag at the C-terminus of the ADAR protein. This approach, utilizing two subsequent affinity chromatography steps, allowed for the successful production of intact ADAR protein, effectively eliminating the degradation observed in the previous attempts (Fig. 16C).

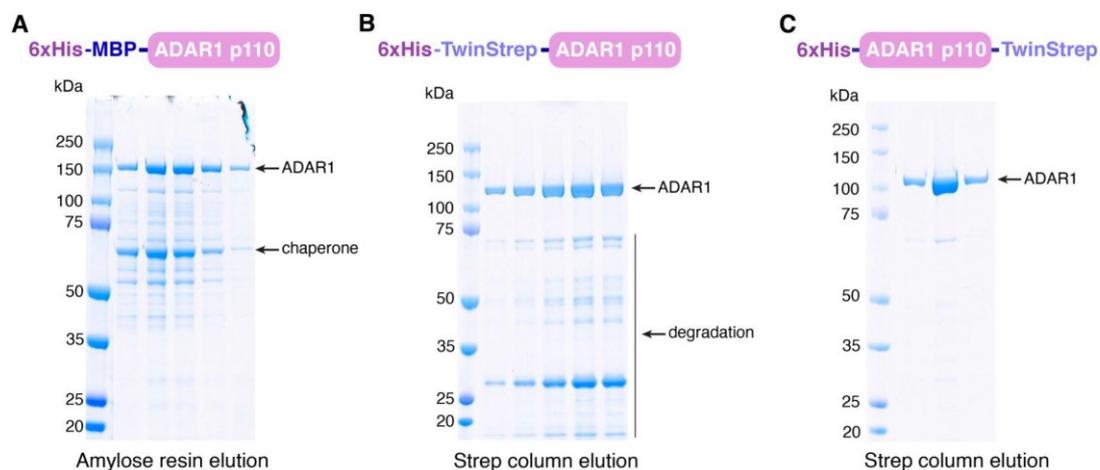


Figure 16 SDS-PAGE analysis following various affinity purification strategies.

- A. SDS-PAGE of purified His₆-MBP-ADAR1, first purified using an HisTrap HP column followed by amylose resin.

- B. SDS-PAGE of purified His₆-TwinStrep-ADAR1, first purified using an HisTrap HP column and subsequently a StrepTrap column.
- C. SDS-PAGE of purified His₆-ADAR1-TwinStrep, first purified using an HisTrap HP column followed by a StrepTrap column.

4.1.2 Evaluating protein solubility through buffer screening

The solubility of ADAR1 was initially tested using various buffers with a wide pH range, from 3.0 to 10.0. This evaluation was conducted on ADAR1 expressed in insect cells, comparing protein levels on SDS-PAGE gels before and after centrifugation to assess the overall solubility. Higher protein levels detected in the supernatant indicated better solubility. The results showed that increasing the pH improved the solubility of ADAR1. At the lowest pH tested, Glycine buffer at pH 3.0, the protein exhibited the poorest solubility. In contrast, the 1x SPG buffer at pH 10.0 yielded the best solubility performance (Fig. 17A). Although the SPG buffer provided the optimal solubility, it was found to be unsuitable for the subsequent purification steps. The high pH of the SPG buffer was too harsh for the HisTrap column resin, which was intended to be used as the first purification step. Additionally benzonase was initially used to remove nucleic acids. Benzonase requires Mg²⁺ ions for its activity, but this ion precipitates in the 1x SPG buffer at pH 10.0. Hence, despite being optimal for solubility, this buffer was unsuitable for the purification process. Ultimately, the ADAR1 p110 purification was effectively achieved using a 50 mM Tris-HCl buffer at pH 8.0. This buffer composition allowed for high-purity protein recovery from both the insect and mammalian cell expression systems, with a higher overall yield obtained from the mammalian cell lines.

In contrast, the optimized protocol proved to be less effective for the longer isoform, ADAR1 p150, which posed significant challenges in achieving high purity. Further optimization efforts included exploring more suitable buffers through thermal unfolding assessments using nanoDSF Prometheus technology on purified ADAR1 p110, alongside a buffer kit from the Vienna Biocenter Protein Technologies Facility (Fig. 15). Unfortunately, despite these extensive efforts, none of the tested buffer conditions were able to surpass the thermal unfolding temperature of the original 50 mM Tris-HCl buffer at pH 8.0, supplemented with 350 mM NaCl, 20% (v/v) glycerol, 0.5 mM EDTA, and 1 mM DTT, which had been employed for the successful purification of ADAR1 p110 (Fig. 17B). This suggested that the longer ADAR1 p150 isoform presented inherent challenges in terms of stability and solubility that could not be readily overcome through simple buffer optimization.

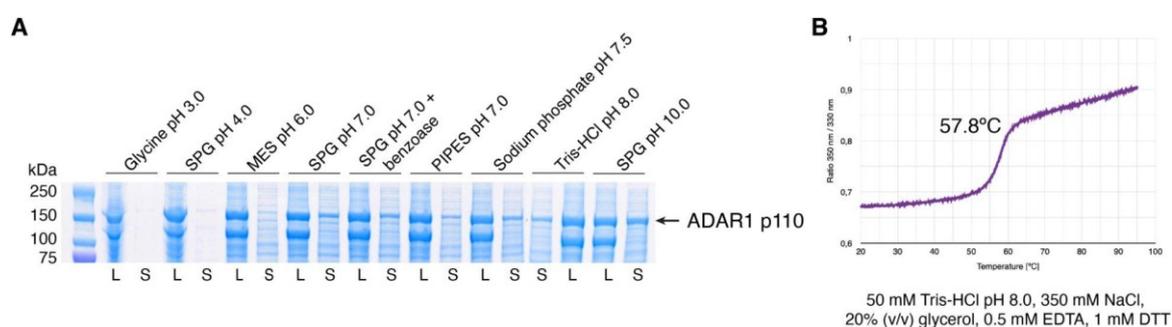


Figure 17 Buffer solubility test.

- A. SDS-PAGE analysis showing buffer solubility at pH ranging from pH 3.0 to 10.0 (L- total lysate before centrifugation, S - supernatant with clarified lysate).
- B. Thermal unfolding of ADAR1 p110 measured using nanoDSF Prometheus Technology.

4.1.3 Optimizing protein expression in mammalian cells

Optimization of protein expression in mammalian cells involved testing various expression vectors, cell densities during transfection, and expression time. To assess the expression levels of ADAR1 isoforms, SDS-PAGE analysis was performed, with equal amounts of lysed cells loaded onto the gel. Among the different vectors examined, the pCAGGS vector proved to be better; it incorporates a cytomegalovirus (CMV) enhancer and a robust chicken β -actin promoter containing an intron sequence, resulting in significantly higher expression levels compared to the pcDNA 3.1 vector, which utilizes a simpler CMV enhancer-promoter system. It was observed that lower cell density during transfection slightly enhanced protein expression levels (Fig. 18A).

Further optimization efforts focused on determining the optimal expression time for ADAR1 isoforms using Western blot analysis with an anti-His₆ tag. The results indicated that ADAR1 p150 reached peak expression levels between 41 and 45 hours post-transfection, while ADAR1 p110 exhibited peak expression between 45 and 50 hours (Fig. 18B). Based on these findings, a refined expression time of 42 hours was established for ADAR1 p150, ensuring maximal yield, while 48 hours was selected for ADAR1 p110.

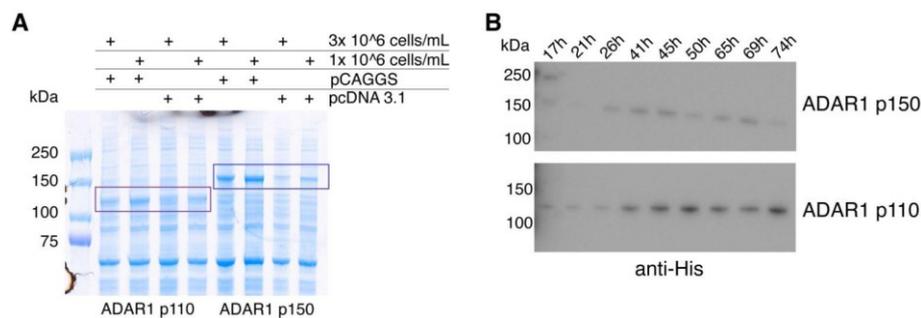


Figure 18 Optimization of protein expression in mammalian cells.

- An SDS-PAGE gel assessing various expression vectors and cell densities employed for protein expression.
- Western Blot analysis targeting the His₆-tag to evaluate expression over time.

4.1.4 Designing the optimal ADAR1 p150 construct

Purification of the ADAR1 p150 isoform proved to be a significant challenge. The protein tended to aggregate and displayed reduced binding to affinity columns, suggesting that the protein tags were inaccessible due to improper folding. Notably, the optimized purification protocol that had been successful for the shorter ADAR1 isoform did not yield high purity or sufficient quantities of the ADAR1 p150 variant.

The AlphaFold analysis showed that the first linker region preceding the ordered Z-alpha domain (comprising amino acids 2-133) was predicted to be intrinsically disordered (Fig. 19A, B). In an attempt to overcome this obstacle, several mutants were designed with various linker truncations (Fig. 19A) and tested for purification from the mammalian expression system. Interestingly, successful purification was achieved for the Δ 77 and Δ 133 ADAR1 p150 mutants, while the Δ 33 variant exhibited similar behavior to the full-length protein.

The purified mutants were then tested for their enzymatic activity by assessing their ability to edit the R/G site in Gria2 RNA (Fig. 19C). Among the constructs tested, only the $\Delta 133$ ADAR1 p150 mutant demonstrated active editing capabilities and was subsequently selected for further experiments throughout the thesis.

Cloning, expression, and purification processes were carried out by intern Nora Rier, whose efforts helped in obtaining the longer ADAR1 p150 isoform.

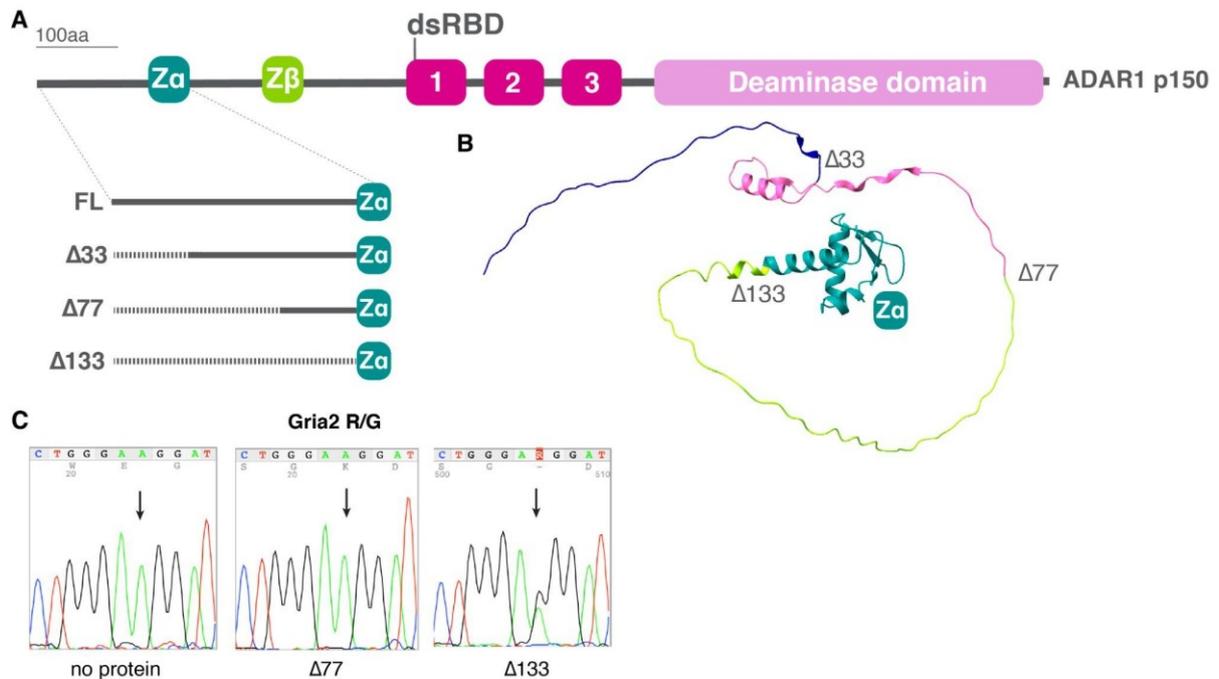


Figure 19 Designing the optimal construct for ADAR1 p150 expression and purification.

- Schematic representation of the various ADAR1 p150 mutants designed.
- Structure prediction for residues 1-197 of AF-P55265-F1, with the regions 1-133 representing the first linker and 134-197 corresponding to the Z-alpha domain (indicated in dar cyan).
- Enzymatic activity of the different mutants through the editing of Gria2 R/G RNA. The R/G editing site is indicated with an arrow. Nucleotide abbreviations: A - adenine, T - thymine, C - cytosine, G - guanine.

4.1.5 Issued faced during purification protocol optimization

In addition to the results presented, the purification of ADAR1 had more challenges during protocol optimization, and while I explored a wider range of conditions, only the most critical findings are included in the thesis.

ADAR1 is an RNA-binding protein that contains multiple domains, each of which interacts with different nucleic acids. During the initial purification attempts, substantial nucleic acid contamination was encountered, necessitating the exploration of various techniques for its removal. The extent of nucleic acid contamination was evaluated by monitoring the absorbance ratio at 260/280 nm, which should be lower than 0.6 in the eluting fractions. Different nucleases, such as benzonase and high-salt active nuclease, were applied to the lysate, but the protein's elution during the first affinity step was inconsistent, sometimes containing nucleic acids and other times appearing free of contaminants. Another strategy involved the use of polyethyleneimine (PEI), a basic cationic polymer, to precipitate nucleic

acids; unfortunately, this approach resulted in considerable protein degradation. Including high NaCl washes provided some improvement but did not eliminate the contamination. The most effective method was adding EDTA to the buffers during the first affinity chromatography, which minimized nucleic acid presence in subsequent purification steps.

Initially, optimizations were performed on His₆-tagged ADAR1. Despite using HisTrap purification and testing various ion exchange chromatographies (such as Heparin, MonoQ, or MonoS resins), the protein remained impure and prone to degradation. Attempts to reduce degradation through various proteases and cell disruption techniques were ultimately ineffective. A successful strategy eventually emerged by tagging ADAR1 with a His₆ tag at the N-terminus and a TwinStrep tag at the C-terminus, followed by two affinity chromatography steps. This approach resulted in the purification of a stable, non-degraded protein.

Purifying the longer isoform, ADAR1 p150, posed additional challenges related to aggregation and folding issues. Several expression methods were optimized, including lower temperature expression in insect and mammalian cells, varying viral titers in insect cells, and employing different cell lines such as Expi293F and HEK Freestyle for mammalian cells, as well as High Five and Sf9 for insect cells. While these optimizations provided some improvements, obtaining a pure full-length protein was ultimately achieved only after removing the first intrinsically disordered linker, resulting in high-yield, purified ADAR1 p150.

Moreover, ADAR1 proved sensitive to the materials utilized during affinity tag removal with 3C protease. Cleavage was tested in various types of tubes, including glass vials, standard and low protein-binding 1.5 ml tubes, and 15 ml Falcon tubes. It was observed that ADAR1 precipitated in the Falcon tubes, which had been commonly used, leading to inconsistent and confusing results (Fig. 20A).

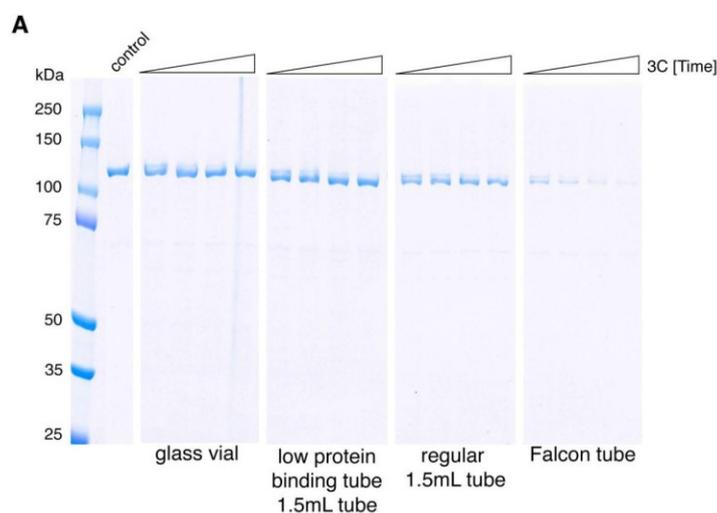


Figure 20 Assessment of various tubes for tag removal using 3C protease.

- A. SDS-PAGE analysis of cleaved ADAR1 p110 across four tube types: glass vial, 1.5 ml low protein-binding tube, standard 1.5 ml tube, and Falcon tube. Cleavage reactions were conducted for 1 hour and 30 minutes, with the initial sample collected after 45 minutes and subsequent samples taken every 15 minutes thereafter.

Due to the numerous challenges encountered during the purification process, achieving a pure, nucleic acid-free, and active form of ADAR1 p110 required approximately 1.5 years of

optimization. This was followed by an additional four months to successfully isolate the longer isoform, ADAR1 p150.

4.1.6 Optimized ADAR1 purification protocol

The ADAR1 p110 (Fig. 21A) and Δ 133 ADAR1 p150 (Fig. 21B) isoforms were successfully expressed and purified using a mammalian expression system, as the insect cell system resulted in lower protein yields and was unable to produce the longer p150 isoform. Following cell lysis, the clarified lysate underwent an initial affinity purification step to capture the His₆-tagged protein. During this step, contaminating chaperones were removed, and the buffers contained 5 mM EDTA to facilitate the removal of contaminating nucleic acids. A HisTrap Excel column, which is resistant to nickel stripping with 5 mM EDTA, was employed for this purpose.

The protein then proceeded to a second affinity chromatography step to capture the TwinStrep tag, facilitating the removal of any degradation products. Afterward, the affinity tags were cleaved, and the protein underwent a final purification step by size exclusion chromatography, effectively removing any remaining aggregates and the 3C protease (Fig. 21A, B). The entire purification process takes approximately 2.5 days.

Mass photometry analysis confirmed the protein homogeneity within the purified samples. For the ADAR1 p110 isoform, the sample predominantly contained monomers (104 kDa) and dimers (218 kDa), with minor contamination from chaperones and higher-order oligomers (174, 328, 434 kDa). In the case of the Δ 133 ADAR1 p150 isoform, the sample was predominantly composed of monomers (121 kDa), with a small population of dimers (246 kDa) and a minor presence of contaminating chaperones (189 kDa).

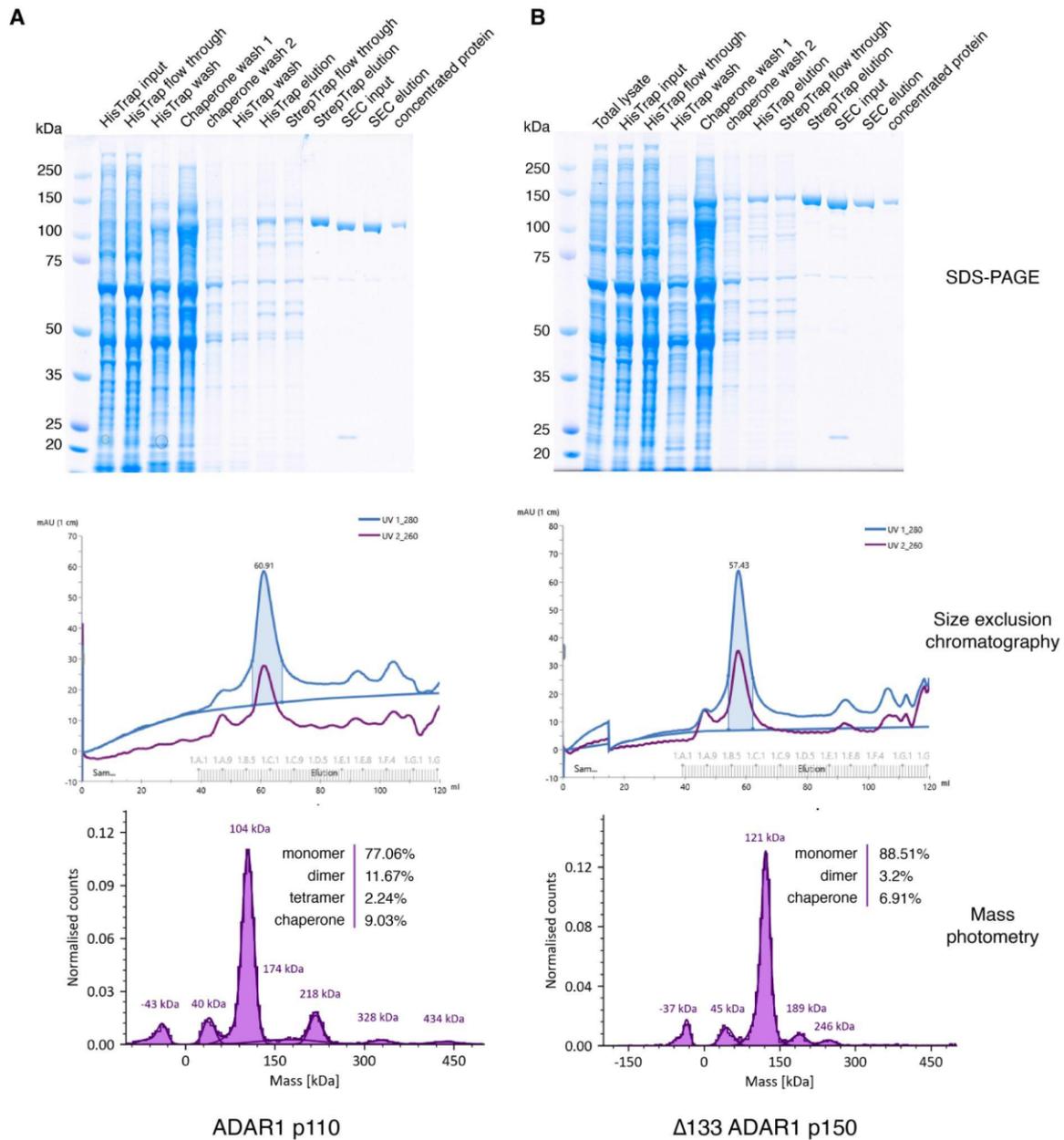


Figure 21 Purification of ADAR1 p110 or 150 from Expi293F mammalian cells.

A. ADAR1 p110,

B. ADAR1 p150.

The analysis includes an SDS-PAGE gel displaying loaded fractions from various purification steps, the elution pattern obtained from size exclusion chromatography, and results from mass photometry analysis.

The successful purification of both isoforms opened the way for further studies aimed at characterizing the ADAR1-RNA complexes. These studies will include a range of functional and structural analyses, including enzymatic activity assays, binding experiments, and efforts to solve the structure using CryoEM techniques.

4.2 ADAR1/2 deaminases purification

To compare the enzymatic activity of the isolated deaminase domain with the full-length protein, various ADAR1 and ADAR2 deaminase (ADAR1 dd or ADAR2 dd) domains were produced using an insect cell expression system. For the ADAR2 dd were generated: the wild-type ADAR2 deaminase, the E488Q mutant with enhanced activity (Matthews et al., 2016), a novel hyperactive mutant, ADAR2 E488Q N496F, to improve activity on 5'-GAN-3' motifs (Katrekar et al., 2022) and ADAR2 RESCUE mutant was designed to convert C to U (cytosine to uracil) (Abudayyeh et al., 2019). For ADAR1, both the wild-type and the E1008Q mutant with enhanced activity (Wang et al., 2015) were produced.

Wild-type and mutant deaminase domains were produced with high purity following the ADAR2 purification protocol. The NaCl binding concentration during heparin chromatography was optimized to improve binding efficiency compared to the full-length ADAR2 protein. The proteins retained their His₆ tags, and the purified deaminases were evaluated for homogeneity using mass photometry, which confirmed that the samples were homogeneous (Fig. 22).

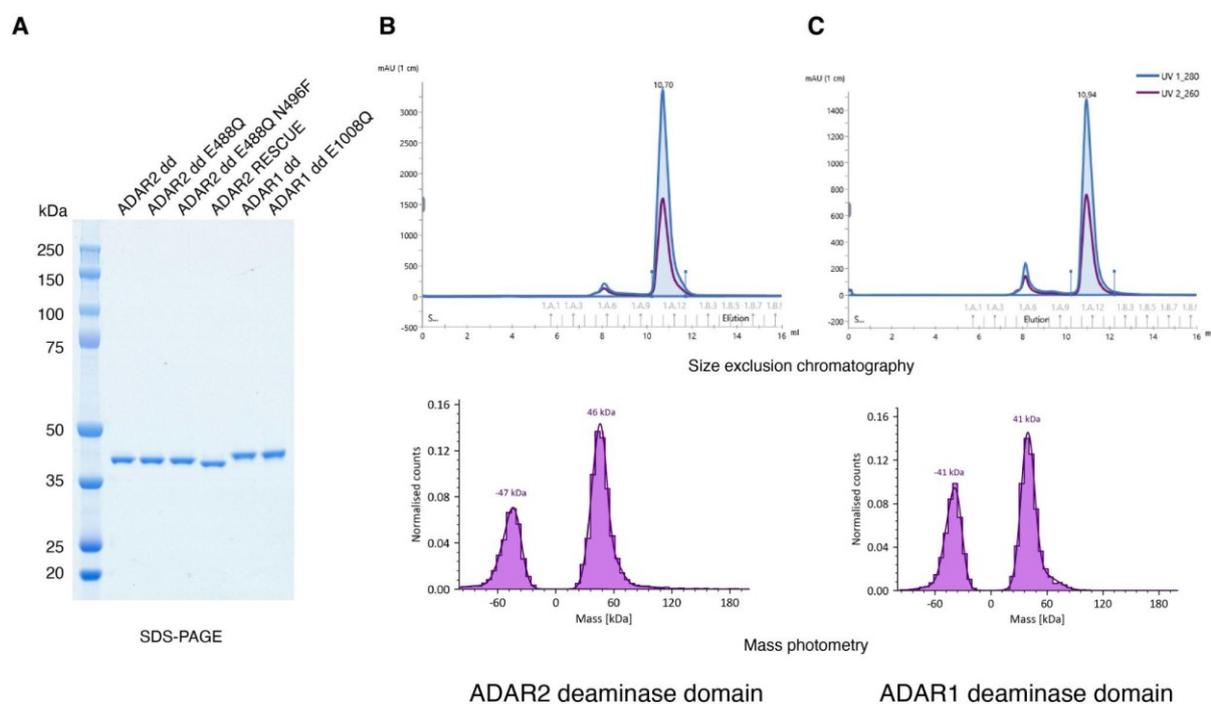


Figure 22 Purification of the ADAR1 or ADAR2 deaminase domains.

The analysis includes an SDS-PAGE gel illustrating the final purified products of various mutants (A), the elution patterns obtained from size exclusion chromatography for wild-type versions of ADAR1 dd (B) and ADAR2 dd (C), and results from mass photometry analysis.

4.3 RNA substrates employed in complex characterization

ADAR1 primarily acts on dsRNA, often editing sites within hairpin structures formed by complementary regions distant from each other in the primary sequence. While high-throughput sequencing can identify editing sites, it has limitations in determining the full dsRNA structure necessary for in vitro studies. For this thesis, previously published and

validated substrates were selected to ensure compatibility with the purified ADAR1 protein, thereby providing a reliable foundation for the experimental work.

The substrates designed for the study varied in length from 80 to 117 nucleotides. This length was carefully chosen to fit the ADAR1 protein domains while being short enough to maintain sample uniformity for CryoEM analysis. The substrates included the glutamate receptor with R/G recoding site (Gria2 R/G), which is derived from the GRIA2 mRNA region. This region is recognized for its efficient editing by the ADAR1 and ADAR2 enzymes (Wong et al., 2001). Another substrate was NEIL1 RNA, derived from the pre-mRNA of the DNA repair enzyme NEIL1, where editing leads to a lysine-to-arginine alteration in the protein's lesion recognition loop (Yeo et al., 2010). The RNA sequence was sourced from a study identifying cis-regulatory elements affecting ADAR editing activity (Liu et al., 2021). Additionally, within the same research, two other RNA sequences derived from mRNAs encoding AJUBA and TTYH2 were also assessed and their reported sequences were used to create these substrates (Fig. 23A).

RNA substrates were produced through T7 in vitro transcription using PCR-generated templates. Following transcription, the RNA substrates were purified via urea gel extraction. The resulting RNA was assessed for purity and homogeneity through denaturing urea gel electrophoresis. This analysis confirmed that all RNA substrates achieved a high level of homogeneity. Notably, the TTYH2 substrate migrated slightly lower than anticipated. This was likely due to strong RNA folding, which was not fully resolved under denaturing conditions rather than a difference in actual size or purity (Fig. 23B).

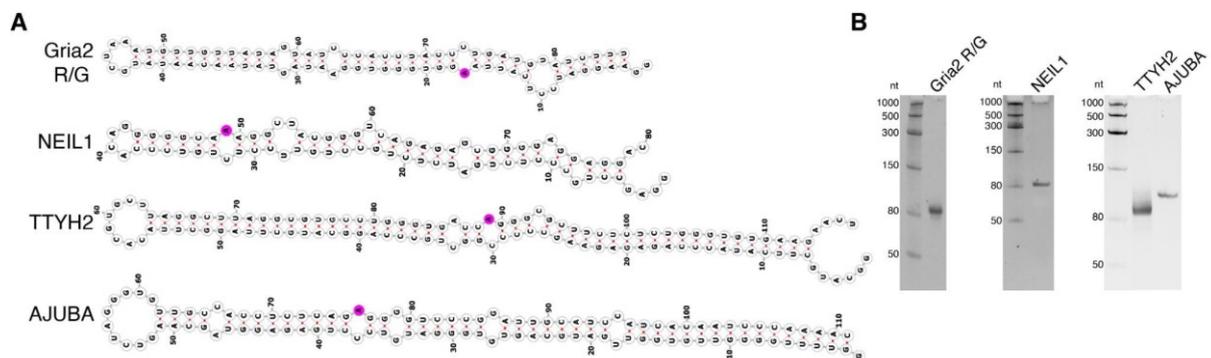


Figure 23 RNA substrates generated for CryoEM studies.

- Predicted secondary structures of Gria2 R/G, NEIL1, TTYH2, and AJUBA, generated using the RNAfold web server. The adenosine targeted for conversion is depicted in magenta, which was highlighted using the Forna visualization tool.
- Urea gel analysis of the synthesized substrates.

4.4 ADAR - RNA complex characterization

4.4.1 Testing ADAR1 enzymatic activity

To assess the enzymatic activity of purified ADAR proteins, in vitro editing assays were performed using synthesized RNA substrates with the ADAR enzymes. The ADAR enzymes were incubated with the RNA substrates for one hour at 37 °C in a reaction buffer containing 20 mM Tris-HCl (pH 8.0 at 25 °C), 30 mM NaCl, 50 mM KCl, 0.5 mM EDTA, 15% (v/v) glycerol, and 0.5 mM DTT. To evaluate the editing activity, a PCR product was generated through reverse transcription of the RNA substrates. The resulting reactions were analyzed using

Sanger sequencing, where editing in the sense strand is represented as an A to G conversion, while in the antisense strand, it appears as a T to C conversion. It was shown previously that the peaks obtained from the opposing strands demonstrated greater consistency (Eggington et al., 2011). All results presented below are shown as the conversion of A (indicated by a green peak) to G (represented by a black peak); however, it should be noted that these sequences were derived from either the sense or antisense strand. Whenever possible, a reverse primer was utilized in the Sanger sequencing; however, in cases where the adenosine was too close to the sequencing coverage limit, the forward primer was used instead. Details regarding the sequenced strands are provided in the figure legend. This methodology effectively illustrates the editing events facilitated by the ADAR enzymes and showcases their enzymatic activity.

One of the most prominent targets for RNA editing by ADAR1 or ADAR2 is the serotonin 2C receptor (5HT2C). Aberrant editing can result in the generation of 24 distinct isoforms of the 5HT2C receptor, which can profoundly influence neurotransmission, contribute to psychiatric disorders such as depression and schizophrenia, and alter the receptor's signaling pathways (Werry et al., 2008). Sites A and B were predominantly edited by ADAR1, exhibiting slightly higher activity for the longer isoform, while site C showed low editing levels exclusively with ADAR2. Notably, site D was completely edited by ADAR2 (Fig. 24A). Importantly, these findings align with the existing literature from both *in vitro* (Eggington et al., 2011) and *in vivo* studies (Yang et al., 2004), confirming that the activity of purified ADAR is indeed effective.

The next target tested was the Gria2 R/G RNA. Editing at R/G site alters the genomic encoding, resulting in a change in translation from arginine to glycine, which facilitates a quicker recovery from desensitization (Lomeli et al., 1994). Both ADAR1 and ADAR2 effectively modified the R/G site. Additionally, editing at the preceding site (site -1) was also noted, consistent with published reports (Kallman, 2003). Furthermore, editing of the site -13 relative to the R/G site was observed, likely due to the extended double-stranded RNA structure arising from the flanking sequences used in the deamination assay (Fig. 24B).

The RNA sequence of NEIL1 is derived from the pre-mRNA of the base excision DNA repair enzyme NEIL1. The editing process modifies the encoding, leading to the conversion of lysine (K) to arginine (R) in the lesion recognition loop, which in turn affects NEIL1's substrate specificity. This site was effectively edited by both ADAR1 and ADAR2, in line with existing literature (Yeo et al., 2010)(Fig. 24C).

The TTYH2 sequence is derived from an intronic region, and the significance of the editing site remains unknown. This substrate has been investigated through mutagenesis and high-throughput screening, revealing important RNA features that influence RNA editing (Liu et al., 2021). Consistent with existing literature, site 0 (marked in magenta) was successfully edited by both ADAR1 and ADAR2. Additionally, low levels of editing were observed at sites -3 and -40 specifically for ADAR2, both enzymes were capable of converting sites -18 and -24 (Fig. 24D).

Similarly, the AJUBA sequence was examined in the same study as TTYH2, with its editing site located within the 3' UTR. Like TTYH2, the impact of this editing site is also currently unknown. Notably, site 0 (marked in magenta) was edited by both ADAR1 isoforms with the longer isoform demonstrating higher efficiency, suggesting that the presence of the Za domain may enhance the editing process. ADAR2 was unable to convert this site (Fig. 24E).

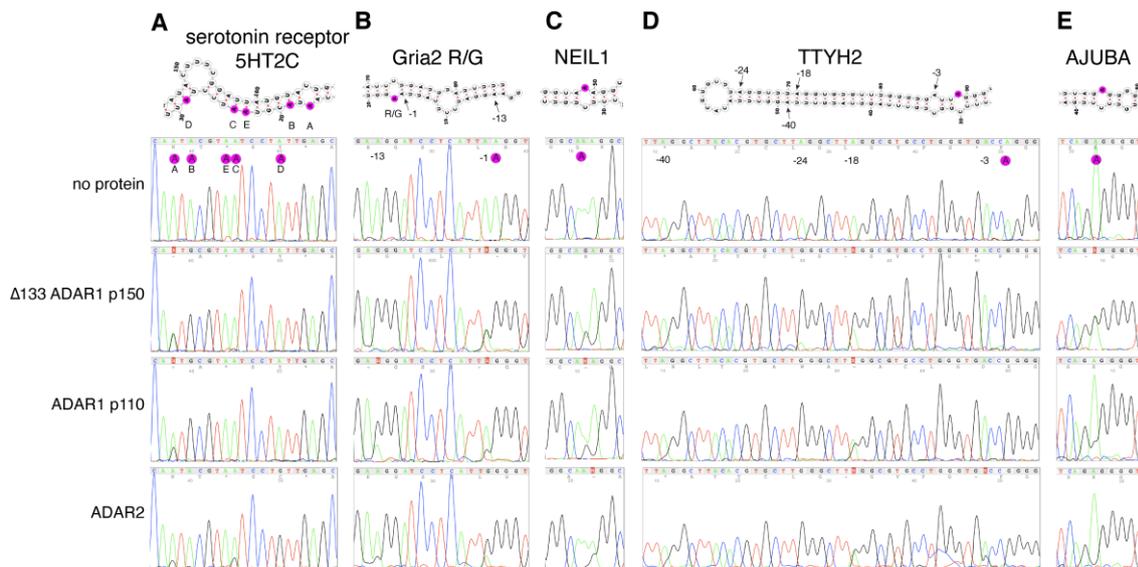


Figure 24 Editing patterns for ADAR1 and ADAR2.

- Editing of 5HT2C. The results were obtained from sequencing the antisense strand, and they are presented as a reverse complement conversion.
- Editing of Gria2 R/G. The results were obtained from sequencing the antisense strand, and they are presented as a reverse complement conversion.
- Editing of NEIL1. The results were obtained from sequencing the sense strand.
- Editing of TTYH2. The sense strand was sequenced. The results were obtained from sequencing the sense strand.
- Editing of AJUBA. The sense strand was sequenced. The results were obtained from sequencing the sense strand.

The adenosine at position 0 is highlighted in magenta, while other converted sites are indicated with arrows. Nucleotide abbreviations: A - adenine, T - thymine, C - cytosine, G - guanine.

4.4.2 Testing ADAR deaminases enzymatic activity

To investigate the editing patterns of isolated deaminase domains, two substrates were tested: Gria2 R/G and TTYH2. The Gria2 R/G site was efficiently edited by all enzymes evaluated; however, editing at site -1 occurred only with the full-length enzymes, indicating that the dsRBD domains are essential for this specific editing process. Site -13 was converted by both full-length ADAR2 and deaminase mutants, while no editing was detected with the wild-type deaminase domains (Fig. 25A).

For TTYH2 (Fig 25B), site 0 was modified by all enzymes, although wild-type deaminases exhibited reduced editing efficiency. Sites -3 and -40 were edited by full-length ADAR2 and its deaminase mutants, E488Q N496F or RESCUE. Conversely, sites -18 and -24 were only converted by the full-length enzymes, with none of the deaminase variants able to effect this change. This suggests that the dsRBD domains are also essential for editing at these particular sites.

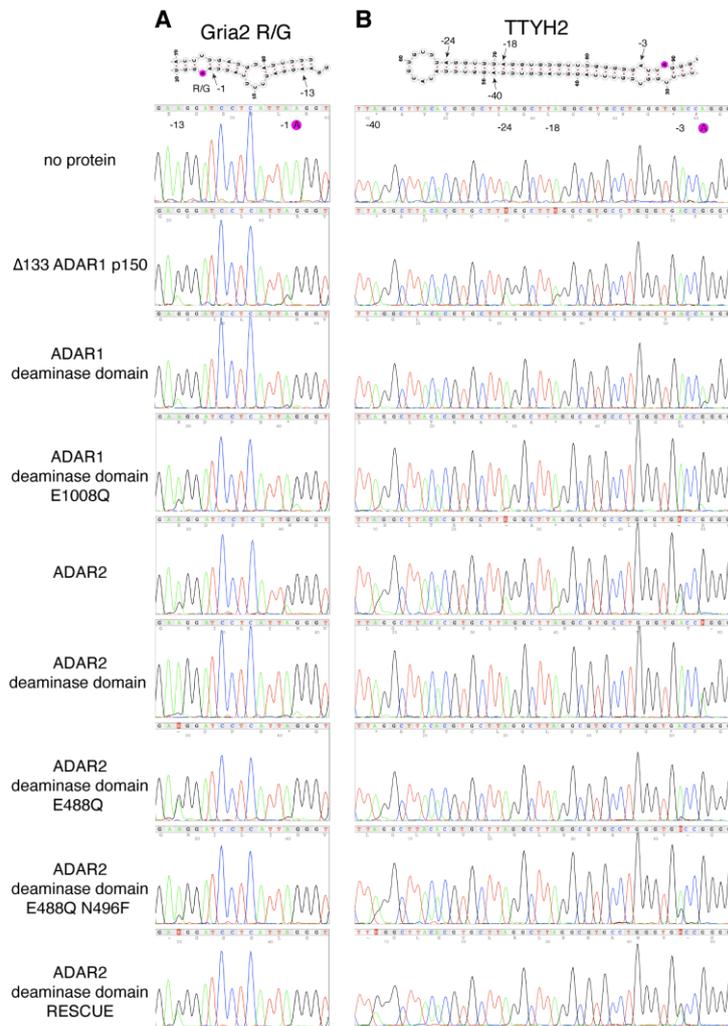


Figure 25 . Comparison of editing patterns for full-length enzymes and isolated deaminase domains.

A. Editing of Gria2 R/G. The results were obtained from sequencing the antisense strand, and they are presented as a reverse complement conversion.

B. Editing of TTYH2. The results were obtained from sequencing the sense strand.

The adenosine at position 0 is highlighted in magenta, while other converted sites are indicated with arrows. Nucleotide abbreviations: A - adenine, T - thymine, C - cytosine, G - guanine.

4.4.3 ADAR1-RNA complex formation

The initial experiments to optimize the complex binding were conducted using ATTO-488 fluorescently labeled Gria2 R/G or NEIL1 substrates at a low RNA concentration of 20 nM. In these experiments, both isoforms of ADAR1 were tested for their binding capabilities. The results showed that both isoforms were able to bind the tested RNAs, however, notable differences were observed in their behavior at varying protein concentrations. The ADAR1 p110 isoform demonstrated a tendency to undergo complex aggregation at higher protein concentrations, (Fig. 26A). In contrast, the ADAR1 p150 isoform exhibited a more stable binding profile, primarily forming monomers with both substrates and as the protein concentration increased, there was a gradual shift from monomers to dimers. The absence of aggregates made the p150 isoform more suitable for cryoEM analysis.

Further binding experiments were carried out using non-labeled substrates, adjusting protein concentrations to levels optimal for freezing samples for cryoEM. All tested substrates were bound to ADAR1 p150, exhibiting distinct binding patterns. Notably, the Gria2 R/G substrate was predominantly enriched for monomeric complexes, while the AJUBA substrate favored the formation of dimers. In contrast, both NEIL1 and TTYH2 substrates displayed a mixture of monomeric and dimeric complexes, reflecting varied binding interactions across the substrates. Additionally, the complexes contained a small fraction of aggregates, which might be attributed to the lower glycerol concentration in the sample preparation. Higher glycerol concentrations could potentially impact the contrast of the samples in cryoEM imaging (Fig. 26B).

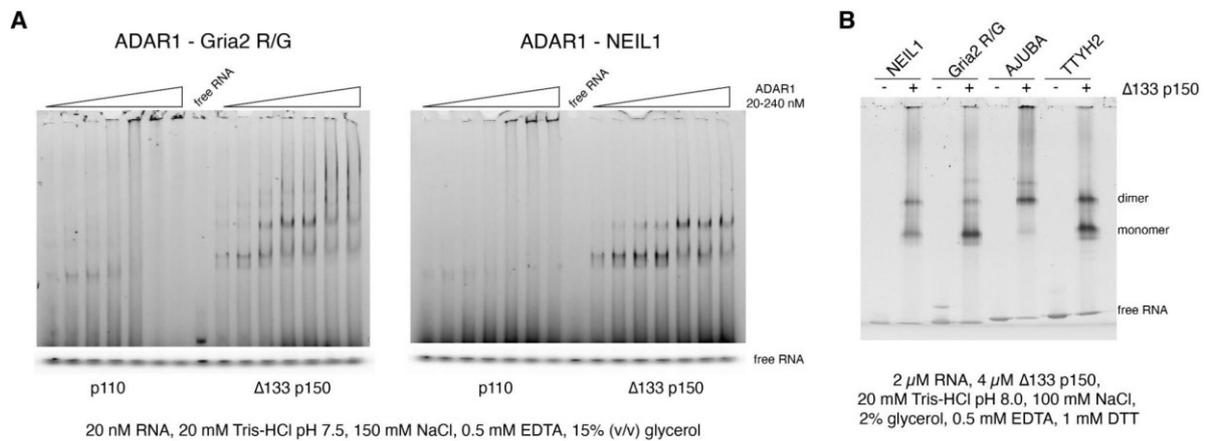


Figure 26 ADAR1 complex formation with different RNA substrates.

- EMSA of ADAR1 isoforms interacting with ATTO-488 Gria2 R/G or NEIL1. The free RNA gel image is cropped from a non-saturated gel scan. The RNA concentration was kept constant at 20 nM, while ADAR1 was used at the following concentrations: 0 (free RNA lane), 20, 40, 80, 120, 160, 200, and 240 nM.
- EMSA of Δ133 ADAR1 p150 with various RNA substrates.

4.4.4 Measuring binding affinities using fluorescence anisotropy

The fluorescence anisotropy data for the interaction between ADAR1 p110 and NEIL1 RNA was analyzed using GraphPad Prism employing the one-site total binding model used for the calculation (<https://rb.gy/4uwj9k>). This model determines the apparent K_d (equilibrium dissociation constant), which reflects the combined effects of both specific and non-specific interactions that may arise at higher protein concentrations which was observed in initial binding experiments (Fig. 26A). The calculated K_d was found to be 337.9 nM, suggesting a rather weak interaction ADAR1 p110 - NEIL1 RNA complex.

While fluorescence anisotropy is a useful technique for assessing binding affinities ((LiCata and Wowor, 2008), it is not without its limitations in this specific context. At elevated protein concentrations, the model may not accurately capture the full nature of the complex formation, particularly the dynamics involved. Furthermore, the existing models may inadequately represent the transient nature of the interactions between ADAR1 and RNA, as well as the potential for non-specific binding events.

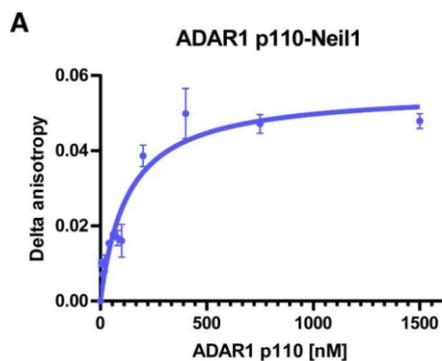


Figure 27 Fluorescence anisotropy data for the ADAR1 p110 - NEIL1 complex. The data points represent the measured fluorescence anisotropy values, with error bars indicating the standard deviation calculated from replicate experiments. The curve illustrates the fitted one-site total binding model.

4.5 ADAR1 - RNA structure determination

4.5.1 Negative staining of ADAR1 - RNA complex

Before the sample preparation for cryoEM analysis, the $\Delta 133$ ADAR1 p150 protein was evaluated using negative staining technique, both in its isolated form and in complex with NEIL1 RNA. The assessment of the isolated $\Delta 133$ ADAR1 p150 protein indicated a homogeneous preparation, confirming its suitability for subsequent analyses (Fig. 28A). In contrast, the complex formed with NEIL1 displayed predominantly homogeneous characteristics; however, a slight presence of aggregates was observed (Fig. 28B). This minimal aggregation suggests that while the complex is largely stable, further optimization may be necessary to enhance its purity for future applications in cryoEM.

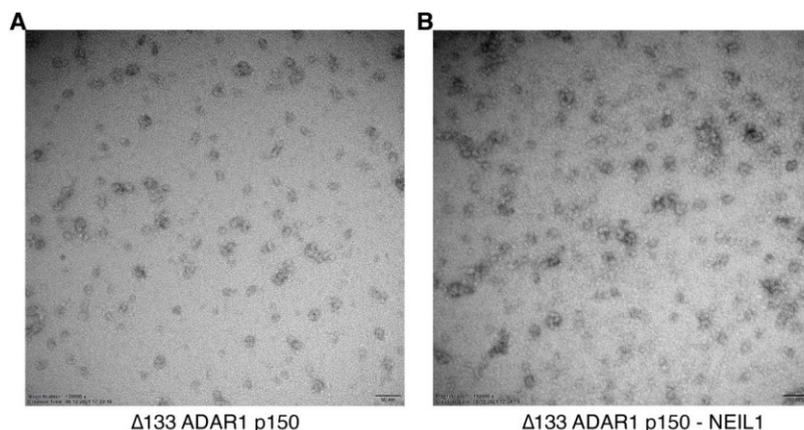


Figure 28 . Negative staining of the $\Delta 133$ ADAR1 p150, both in isolation and in complex with NEIL1 RNA. The images were captured at a nominal magnification of 150,000x

4.5.2 Preparation of cryoEM samples

Specimens for cryoEM are prepared on specialized grids, usually made of a fine metal mesh such as copper or gold. The mesh is coated with a carbon film or gold foil containing a series of small holes, which can vary in both size and arrangement, typically ranging from 0.6 to 3.5 micrometers in diameter. Biological samples are suspended within the holes. Alternatively, a

thin support layer, for example continuous carbon, graphene, or graphene oxide, can be applied to the holey films which provides a stable surface to support the biological complexes. The support film's thickness is crucial as it must allow electrons to pass through the sample while also maintaining the sample support (Fig. 29).

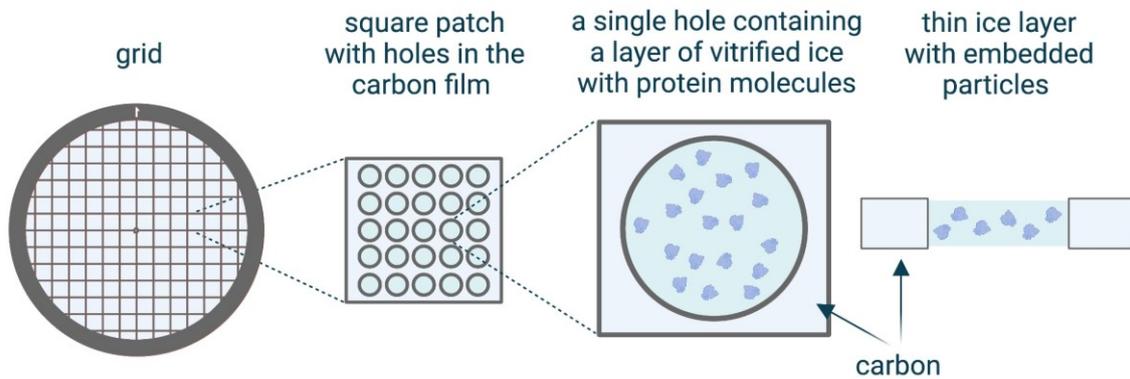


Figure 29 Scheme of the CryoEM grid.
Figure prepared with BioRender based on (Costa et al., 2017)

During the vitrification process, the specimen is embedded in a thin layer of vitrified ice, and it is essential to uniformly distribute the sample within these grid holes to ensure optimal imaging conditions. Achieving this uniform distribution often requires optimization, which can involve experimenting with different types of grids, adjusting the concentration of the biological complexes, and adding specific sample additives (Weissenberger et al., 2021).

Carefully preparing the cryoEM grids is vital for obtaining high-quality data. However, the preparation process can be challenging and often non-reproducible due to the multitude of steps involved. These steps include preparing the biological complex, freezing the sample (a stage where even small device variations can make a difference), clipping the grids before inserting them into the microscope and handling the very fragile grids. Each of these stages requires careful attention and optimization to achieve a uniformly distributed sample while avoiding contamination at each of these steps to ensure the sample is suitable for high-resolution imaging.

The optimization of the cryoEM sample preparation began with the shorter ADAR1 p110 isoform, which was first purified to an adequate concentration for experimentation. The preliminary studies involved complex freezing on ice, primarily utilizing QUANTIFOIL grids, with occasional use of C-flat grids. Early attempts to prepare the ADAR1 p110 and NEIL1 RNA complex using QUANTIFOIL grids did not yield successful results, leading to the decision to switch to C-flat grids featuring a thinner and flatter layer of holey carbon. However, the complex tended to adhere to the holey carbon, making it challenging to visualize within the grid holes, despite various adjustments to the sample preparation and grid selection (Fig. 30A).

Subsequent attempts focused on evaluating the ADAR p110 and Gria2 R/G complex. While the complex appeared promising under negative staining, it exhibited significant aggregation following the vitrification process for cryoEM specimen preparation (Fig. 30B). This

aggregation was likely introduced during the sample freezing step. These experiments highlighted the challenges associated with utilizing the shorter ADAR1 p110 isoform in cryoEM studies, as the initial data indicated a propensity for this construct to aggregate.

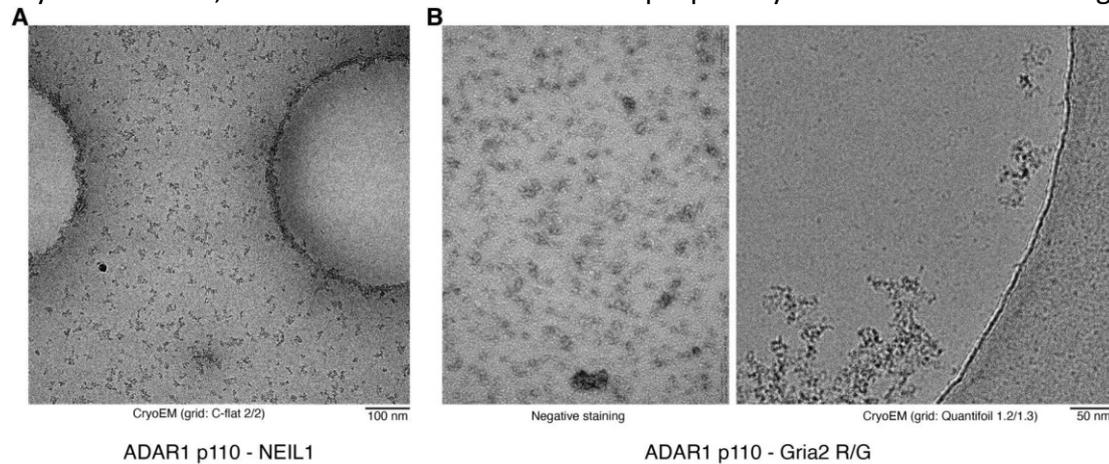


Figure 30 CryoEM analysis of the ADAR1 p110 - RNA complex.

- A. CryoEM image of the ADAR1 p110 - NEIL1 RNA complex.
- B. Negative staining and cryoEM images of the ADAR1 p110 - Gria2 R/G RNA complex.

Subsequent experiments shifted focus to evaluating the longer isoform, $\Delta 133$ ADAR1 p150, in combination with various RNA substrates. Initially, the complex with NEIL1 RNA was tested, but it exhibited behavior similar to that of the shorter isoform, with the complex localizing outside the grid holes. This result indicated that NEIL1 RNA was not suitable for further experiments, despite the change in ADAR1 isoform (Fig. 31A).

Next, complexes with TTYH2 and AJUBA RNA were examined, presenting a new set of challenges. In these samples, a significant amount of free RNA was observed in the acquired grid images (indicated with a white arrow). This observation suggested that the RNA had dissociated from the complex during sample preparation or grid freezing, resulting in a sample that was too heterogeneous for obtaining high-quality data (Fig. 31B, C).

On a more promising note, the complex with Gria2 R/G showed favorable results, as the particles appeared well-distributed within the grid holes. This distribution made it suitable for data acquisition and allowed for an effective assessment of particle quality following their processing (Fig. 31D).

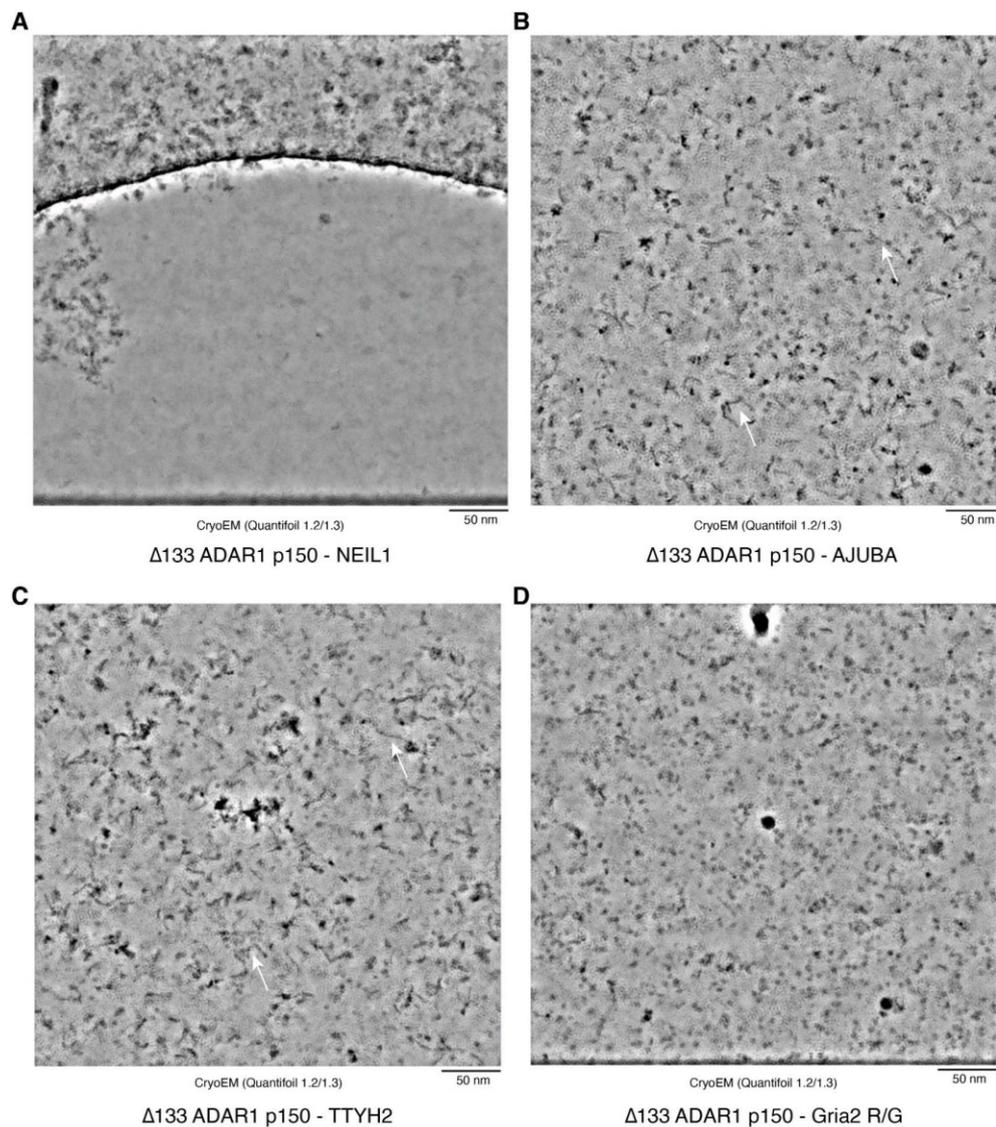


Figure 31 CryoEM analysis of the $\Delta 133$ ADAR1 p150 - RNA complex prepared on ice.

- A. Complex with NEIL1 RNA.
- B. Complex with AJUBA RNA.
- C. Complex with TTYH2 RNA.
- D. Complex with Gria2 R/G RNA.

The contrast of all images was enhanced using Warp software to improve visibility and details of the structural features.

These experiments highlighted the importance of screening multiple RNA substrates and isoforms to identify the most stable and well-behaved complexes for CryoEM studies. The success with the Gria2 R/G complex provided a promising direction for further structural investigations of ADAR1.

To assess the sample quality of the $\Delta 133$ ADAR1 p150-Gria2 R/G complex, a small dataset was collected using a 200 kV Glacios microscope, capturing a total of 310 micrographs during the data collection session. An initial assessment of the particle quality revealed two distinct sizes, approximately 5 nm and 15 nm. When compared to the previously solved ADAR2 dimer-RNA complex (PDB: 6VFF), it became evident that the $\Delta 133$ ADAR1 p150-Gria2 R/G complex should

fall within a larger size range. The results suggest that the smaller particles might represent dissociated complex components (Fig. 32A).

The first stages of data processing were carried out using Warp, which involved particle picking through a machine-learning approach trained to identify particles of varying sizes. After extracting particles around 15 nm, which corresponded to the expected size of the full complex, approximately 13,000 particles were available for further analysis. These particles were subsequently processed in CryoSparc, where both 2D (Fig. 32B) and 3D classifications, including 3D refinement of class 2, were performed (Fig. 32C).

Unfortunately, the resulting model exhibited low quality, with the deaminase domain superimposed (AF-P55265-F1, amino acids: 883-1226), suggesting that RNA likely dissociated from the complex. This observation was also supported by the presence of a significant number of smaller 5 nm particles which probably correspond to the size of the deaminase domain alone. Overall, the data proved too heterogeneous to produce a high-quality model, emphasizing the need for crosslinkers to stabilize the complex for improved results.

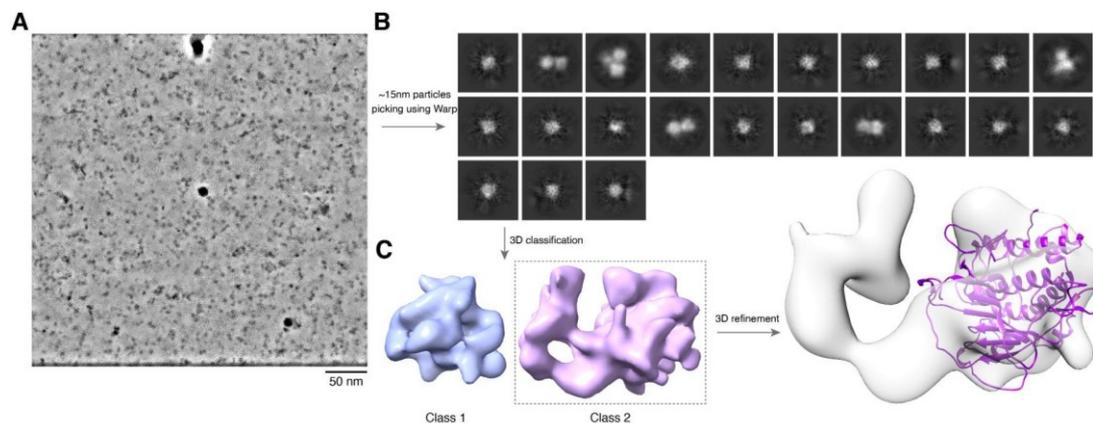


Figure 32 CryoEM analysis of the $\Delta 133$ ADAR1 p150-Gria2 R/G complex.

- Representative micrograph.
- Representative 2D classes derived from the analysis.
- 3D classification and refinement of class 2, with the final structure incorporating the superimposed ADAR1 deaminase (AF-P55265-F1 amino acids: 883-1226).

4.5.3 Stabilizing the complex using crosslinkers

Several crosslinking techniques were explored to stabilize the $\Delta 133$ ADAR1 p150-RNA complex, including BS3, glutaraldehyde, and SPB. The optimization trials focused on fine-tuning parameters such as crosslinking time, concentrations of proteins and reagents, and temperature to prevent excessive crosslinking, which could complicate data analysis and interpretation. The effectiveness of the crosslinking was assessed using SDS-PAGE, comparing the samples before and after applying different crosslinking conditions. Additionally, Native PAGE gel analysis was performed to ensure that crosslinking did not result in the dissociation of the protein and RNA.

4.5.3.1 BS3

BS3 has two reactive N-hydroxysuccinimide (NHS) ester groups reacting with primary amine groups in proteins, specifically targeting lysine residues.

To optimize the reaction conditions, a complex of $\Delta 133$ ADAR1 p150 was formed with various RNA substrates. After the complex was established, BS3 was added to the solution at concentrations ranging from 0.025 to 2 mM, representing a 50 to 4000-fold excess over the protein concentration. In the final experiments, the incubation time was set to 5 minutes at room temperature, with gentle shaking at 300 RPM.

SDS-PAGE analysis demonstrated that crosslinking was indeed effective. However, the results from Native PAGE gel electrophoresis indicated that RNA dissociation occurred at higher concentrations of BS3 (Fig. 33). This finding suggested that BS3 is unsuitable for further experiments. When ADAR1 was treated with BS3, the protein became less available for binding to RNA, compromising the integrity of the protein-RNA complex.

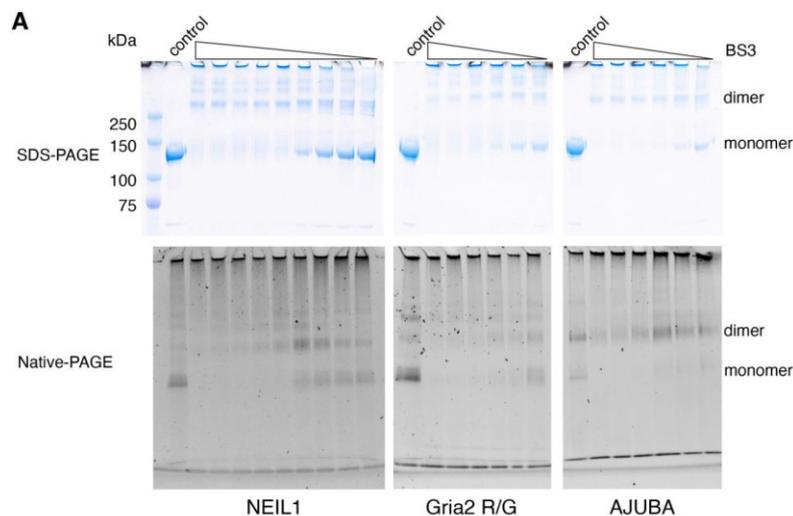


Figure 33 Crosslinking of $\Delta 133$ ADAR1 p150 with RNA substrates using BS3. Crosslinking effectiveness was evaluated through SDS-PAGE and native PAGE analyses.

4.5.3.2 Glutaraldehyde

Glutaraldehyde is a widely used crosslinking agent in structural studies, stabilizing transient protein-protein interactions, which can improve ADAR1-RNA complex stabilization in CryoEM. It has two aldehyde (-CHO) groups, that can interact with the primary amine groups present on proteins, targeting mostly the lysine residues.

The optimized condition for using glutaraldehyde involved a 5-minute incubation on ice with 0.05% (v/v) glutaraldehyde. This treatment efficiently crosslinked the protein, as revealed by SDS-PAGE analysis, and successfully maintained the integrity of the ADAR1-RNA complex without dissociation (Fig. 34A). The successful crosslinking made the sample suitable for subsequent preparation of the cryoEM sample.

The cross-linked complexes of $\Delta 133$ ADAR1 p150 with Gria2 R/G (Fig. 34B), NEIL1 (Fig. 34C) or TTYH2 (Fig. 34D) RNA were prepared using QUANTIFOIL 0.6/1 grids and vitrified using a Vitrobot. The images acquired with the Glacios microscope showed the presence of many free RNA molecules, indicating that the complex had dissociated, despite the effective

crosslinking. This dissociation most likely occurred during the freezing procedure, making the samples unsuitable for further experiments.

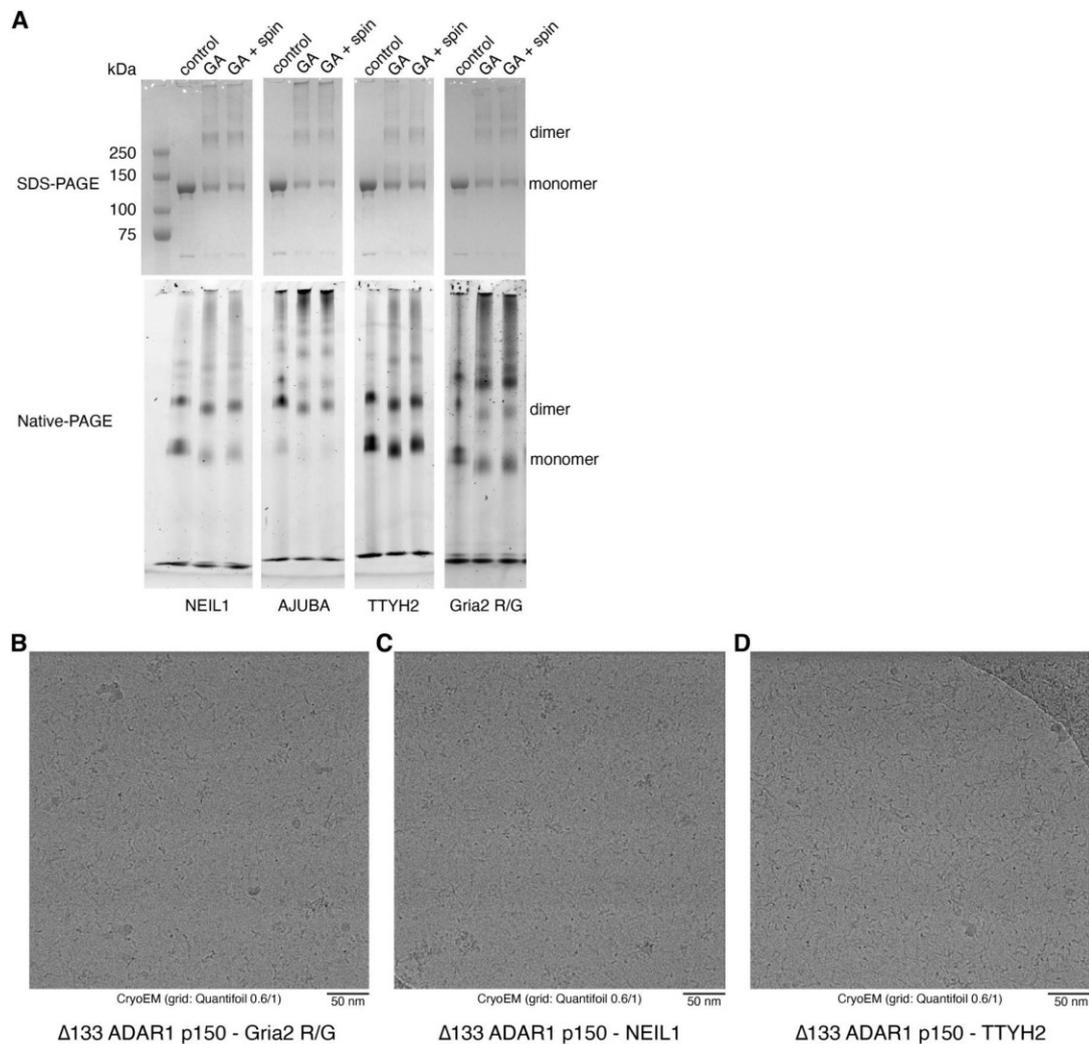


Figure 34 Crosslinking of $\Delta 133$ ADAR1 p150 with RNA substrates using glutaraldehyde.

- Crosslinking effectiveness evaluated through SDS-PAGE and native PAGE analyses. Control - sample before crosslinking, GA - after glutaraldehyde crosslinking, GA + spin - sample after crosslinking followed by a spin to remove the aggregates
- CryoEM analysis of the crosslinked complexes.

4.5.3.3 SPB

Further experiments focused on optimizing crosslinking using SPB (succinimidyl-[4-(psoralen-8-yloxy)]-butyrate), a bifunctional crosslinking agent having two reactive groups: a succinimidyl ester group targeting primary amine groups on proteins and a psoralen moiety interacting with nucleic acids, especially with RNA. The cross-linking is activated by UV light which creates covalent crosslinks. Specifically, upon exposure to UV light at 365 nm, the psoralen group interacts with pyrimidine bases, forming covalent interactions (Fig. 35A). In contrast, wavelengths below 300 nm have the opposite effect and can reverse the crosslinking. The dual-targeting properties of SPB make it suitable for crosslinking ADAR1-RNA complexes, as it can interact with both the protein and the RNA components.

The existing literature provided limited guidance on the practical implementation of the SPB crosslinker. Consequently, the entire process had to be optimized from the ground up. This began with carefully refining the reagent preparation, which was achieved by incubating the dissolved SPB at temperatures slightly above 55 °C for a few minutes. Equally important was the optimization of key parameters, including protein and crosslinker concentrations, sample volume, and crosslinking time. To mitigate any potentially disruptive effects, UV radiation below 300 nm was blocked by placing a 3 mm glass plate between the samples and the UV lamp. The crosslinking optimization was mostly done on $\Delta 133$ ADAR1 p150 - TTYH2 complex.

The optimization of the crosslinking efficiency was validated using mass photometry. This technique helped to demonstrate that the RNA-containing complexes remained intact after the crosslinking process, and it also verified the absence of any unwanted aggregates.

The results showed that the crosslinking was efficient when using a 25 $\mu\text{g/ml}$ concentration of SPB and a 0.05 mg/ml concentration of protein in the reaction mixture. This indicated that no formation of aggregates occurred under these optimized conditions and the RNA remained bound to ADAR1 (Fig. 35B).

These collective optimization efforts including the refinement of reagent preparation and the careful tuning of key parameters established a solid foundation for the successful sample preparation suitable for downstream cryoEM studies.

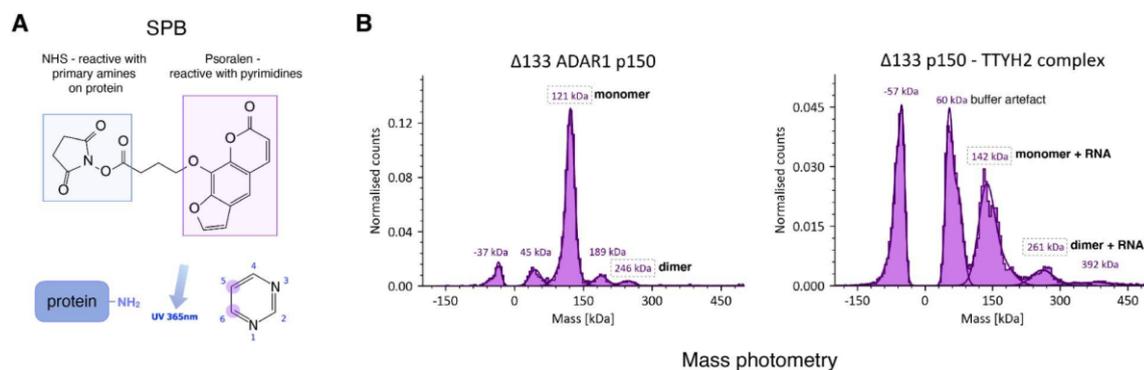


Figure 35 Crosslinking of $\Delta 133$ ADAR1 p150 - TTYH2 RNA complex using SPB.

- Chemical structure of the SPB highlighting the reactive groups.
- Crosslinking effectiveness was assessed through mass photometry analyses.

4.5.4 CryoEM analysis of SPB crosslinked complex

The optimization of crosslinking was found to be most effective at a protein concentration of 0.05 mg/ml, ensuring no formation of aggregates. However, this sample concentration is very low to be able to get enough particles on holey carbon. Therefore the initial attempts involved concentrating the crosslinked complexes. Unfortunately, this approach proved unsuccessful. The concentrating step led to nonspecific RNA binding events, which resulted in the formation of lengthy RNA strands. These unwanted RNA structures were clearly visible in the acquired cryoEM images (Fig. 36A).

The low 0.05 mg/ml protein concentration of the sample, combined with the inability to concentrate it further, necessitated the use of continuous support films on the cryoEM grids. These support films are designed to help concentrate the sample within the grid holes. It is

crucial to minimize their thickness in order to maximize sample contrast during cryoEM data acquisition. Thinner films reduce background noise and improve the achievable resolution. For the experiments, various continuous support film materials were evaluated and optimized, including carbon and graphene oxide. The grid coating with these support layers was prepared carefully according to the described methods.

The initial experiments revealed that the use of graphene oxide as a support layer was causing undesirable sample aggregation, rendering it unsuitable for preparing the complex for cryoEM (Fig. 36B). Consequently, the focus shifted towards optimizing the use of carbon grids instead. However, the initial trials with carbon grids also showed that the complex was prone to aggregation (Fig. 36C), which highlighted the need for refining the sample preparation process. Optimization efforts included adjusting the incubation time on the grid before blotting the sample, testing various blotting devices to improve sample application, and evaluating grids with different hole sizes.

The optimal incubation time before blotting the excess sample was determined to be approximately 10 minutes. Initially, the Vitrobot device was used for blotting. The Vitrobot blots the grids from both sides, leading to increased exposure of the grid to damage, and also increased sample aggregation was observed. To address this, the GP2 device was later introduced. This device blots the excess sample from only one side of the grid, significantly reducing the specimen damage. By gently removing the excess liquid with filter paper, this approach helps preserve the integrity of the carbon layer and maintain the quality of the sample.

A significant aspect of the optimization process involved evaluating grids with varying hole sizes. Grids with smaller holes, specifically those measuring 0.6 μm in diameter, caused the sample to aggregate on one side of the grid, leading to aggregation and uneven distribution (Fig. 36D). In contrast, grids with larger holes measuring 1.2 μm allowed for better sample distribution across the surface (Fig. 36E). The optimal quality was achieved using grids with 3.5 μm diameter holes, which provided the best balance between sample distribution and stability (Fig. 36F). While this setup did not entirely perfect sample enrichment and having more particles would have been ideal, extending the incubation times proved counterproductive, as it increased sample aggregation. The sample was further evaluated through data collection to assess the information about crosslinked complexes.

Following the initial assessment of sample quality using the Glacios microscope, the sample was subsequently imaged on the 300 kV Krios microscope to acquire high-resolution data for image processing. Approximately 22,000 micrographs were collected and processed in Warp, where particles were picked utilizing an automated approach. These exported particles were then analyzed using RELION software. The first step involved an extensive 2D classification, which revealed a highly heterogeneous particle population. During this classification, around 90% of the particles were discarded due to significant structural variability (Fig. 36G). Subsequently, several attempts were made to generate 3D models using different templates, but the data turned out to be highly model-biased (Fig. 36H). Unfortunately, the remaining particle count was insufficient to proceed with data processing, preventing the generation of a reliable 3D model of the complex. This outcome suggests that the sample was indeed highly

heterogeneous, even with the application of crosslinking intended to stabilize the complex.

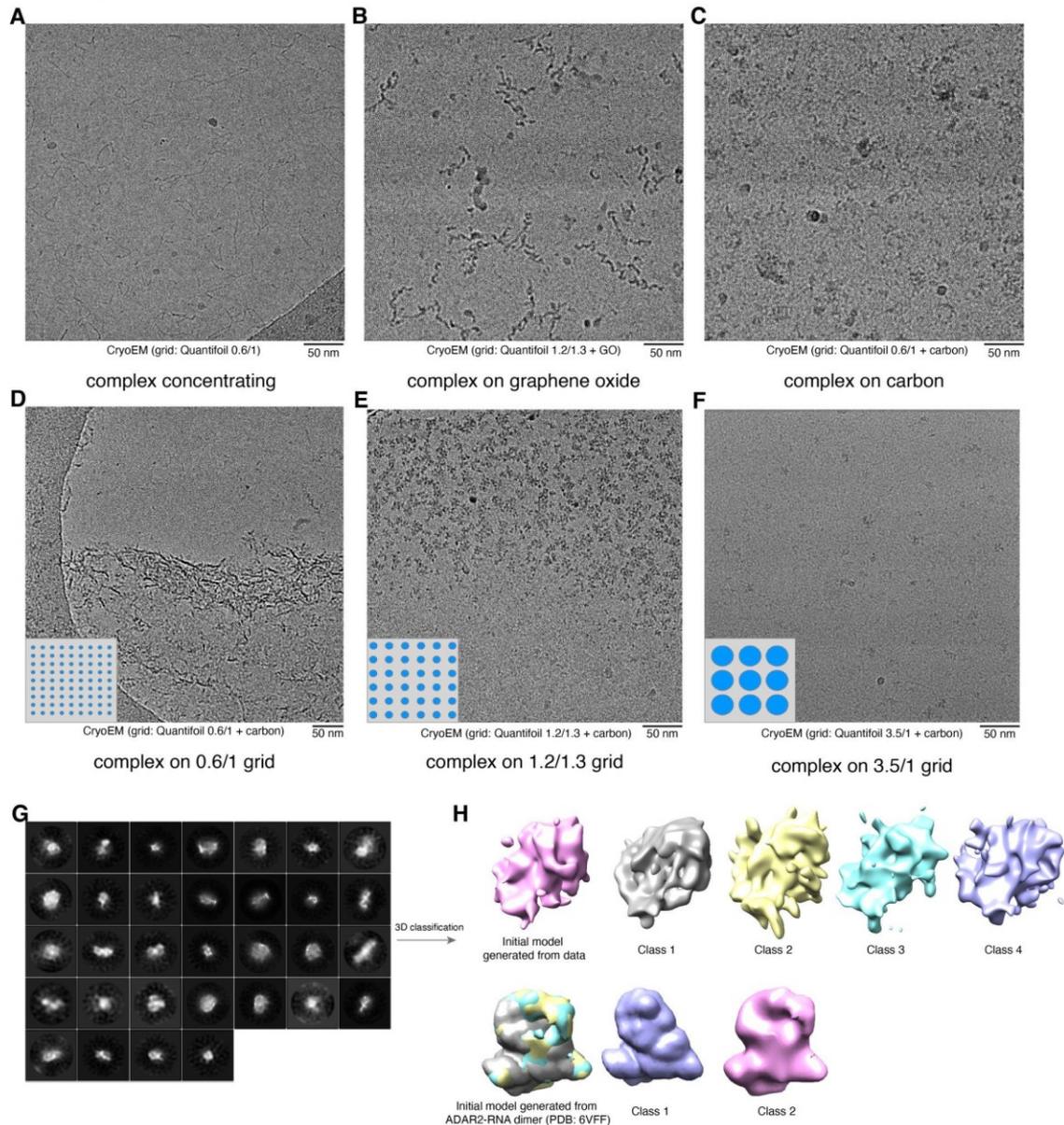


Figure 36 CryoEM analysis of the $\Delta 133$ ADAR1 p150 - TTYH2 RNA complex crosslinked with SPB.

- A. CryoEM analysis of the concentrated complex.
- B. Complex prepared on QUANTIFOIL grids coated with graphene oxide.
- C. Complex prepared on QUANTIFOIL grids coated with continuous carbon, representing initial experimental results.
- D. Analysis of the complex prepared on QUANTIFOIL 0.6/1 grids coated with continuous carbon.
- E. Complex prepared on QUANTIFOIL 1.2/1.3 grids coated with continuous carbon.
- F. Analysis of the complex prepared on QUANTIFOIL 3.5/1 grids also coated with continuous carbon.
- G. 2D classification showing representative classes
- H. 3D classification by using different model templates, models generated from the data, or the ADAR2-RNA complex (PDB: 6VFF)

4.5.5 Reducing complex aggregation by using dimerization mutant

In previous cryo-EM experiments, the ADAR1-RNA complex encountered some issues with aggregation. Recently, a preprint on bioRxiv described a new ADAR1 dimerization mutant that

reveals an interface interaction between the dsRBD3 domains of two monomers (Fig. 37A). This dimerization interface is independent of RNA binding and has been shown to modulate RNA editing activity at specific sites (Mboukou et al., 2023). Disrupting this dimerization interface could potentially reduce protein aggregation while still allowing for effective RNA binding.

The $\Delta 133$ ADAR1 p150 variant, featuring the mutations V747A, D748Q, W768V, and C773S, was successfully cloned and expressed in a mammalian system. Utilizing the established ADAR1 purification protocol, the protein was purified to a high degree of purity. However, analysis of the flow-through from the StrepTrap column indicated that a portion of the protein may have experienced folding issues, as it was not retained by the column (Fig. 37B).

Subsequently, the mutant was evaluated for its binding capabilities with Gria2 R/G, AJUBA, and TTYH2 RNAs, and compared to the non-mutated ADAR1. The mutant was more effective at reducing complex aggregation, as Native PAGE gel analysis showed that aggregates formed at higher protein concentrations compared to the non-mutated version. This suggests that by disrupting the dimerization interface in ADAR1 dsRBD3, it may be possible to reduce aggregation issues encountered in cryoEM experiments.

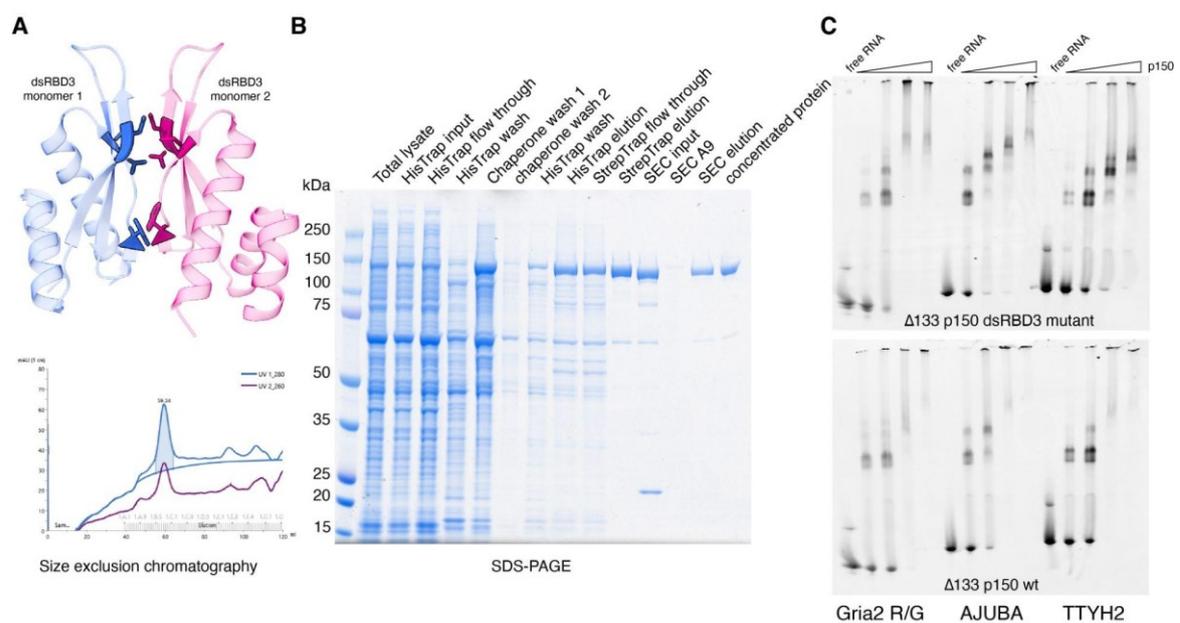


Figure 37 Analysis of the ADAR1 dsRBD3 dimer mutant.

- Crystal structure of dsRBD3, highlighting the dimerization interface, with residues V747, D748, W768, and C773 marked in royal blue or medium violet red (PDB: 7ZJ1).
- Purification of the $\Delta 133$ ADAR1 p150 dsRBD3 dimer mutant, including an SDS-PAGE gel that shows loaded fractions from various purification steps and the elution pattern obtained from size exclusion chromatography.
- EMSA binding assay comparing the mutant and wild-type versions of ADAR1 with various RNA substrates. The RNA concentration was kept constant at 0.25 μ M, while ADAR1 was used at the following concentrations: 0 (free RNA lane), 0.25, 0.5, 1, and 1.5 μ M. The complex was incubated for 20 minutes at 25°C before being loaded onto the gel.

4.5.6 AlphaFold structure prediction

The initial structure prediction was conducted using AlphaFold2 (Jumper et al., 2021). It successfully identified the ordered domains within the ADAR1 protein, including the catalytic deaminase domain, dsRBDs, and Z-domains. However, the model also revealed several low-confidence regions, indicated in grey, which correspond to intrinsically disordered regions (IDRs) located between the structured domains (Fig. 38A).

IDRs are segments of proteins that do not adopt a fixed or ordered three-dimensional structure under physiological conditions (Holehouse and Kragelund, 2024; Struhl, 2024). These regions are commonly found in nucleic acid-binding proteins, where they play crucial roles in facilitating various functions. The positioning and organization of IDRs can vary significantly depending on the cellular context and may change upon binding to RNA substrates (Ottoz and Berchowitz, 2020). This dynamic nature poses challenges for accurately predicting their conformations using AlphaFold technology.

Despite these limitations, the model obtained from AlphaFold2 proved valuable during the initial processing of CryoEM data. Specifically, it allowed for the successful superimposition of the deaminase domain onto the CryoEM model of ADAR Gria2 R/G complex.

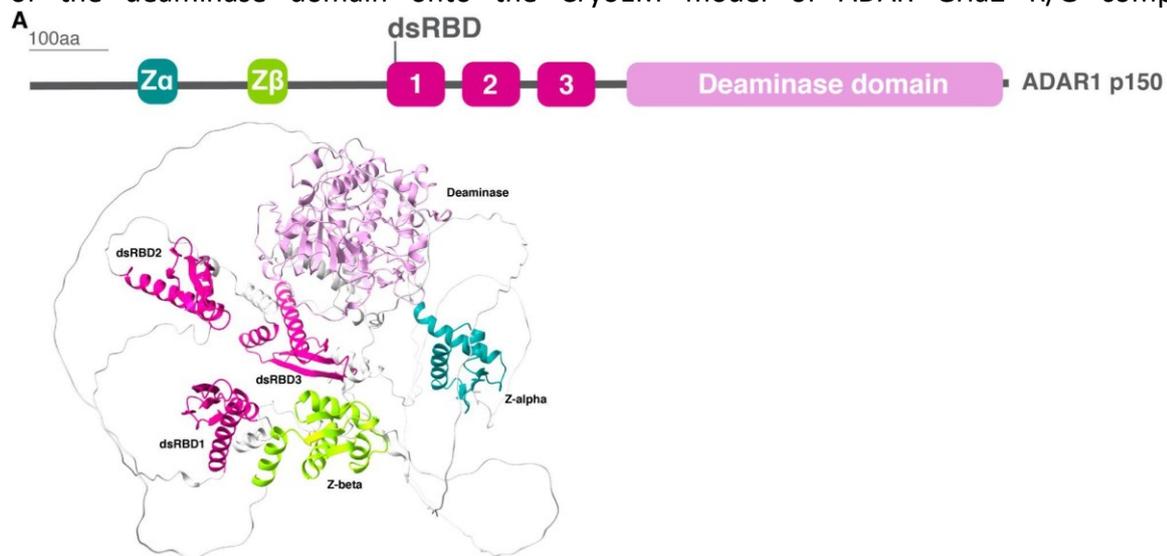


Figure 38 AlphaFold2 ADAR1 structure prediction.

The domains are represented in the following colors: deaminase (plum), dsRBD (medium violet red), Z-alpha (dark cyan), and Z-beta (green yellow). AF-P55265-F1.

ADAR1 structure prediction in a complex with RNA substrate was only feasible when AlphaFold3 was released. It can handle multimer predictions allowing model interactions between multiple proteins and RNA molecules simultaneously (Abramson et al., 2024).

The predicted structure of the Δ 133 ADAR1 p150 - TTYH2 RNA complex was generated using a AlphaFold3 server, which received TTYH2 RNA sequence, two copies of the protein-coding sequence based on previous evidence suggesting that this complex can form dimers (Fig. 26B). Given that the ADAR1 deaminase domain is known to bind to two zinc ions, the model incorporated a total of four zinc ions, with two assigned to each deaminase domain (Park et al., 2020).

IDRs were predicted with low confidence; therefore, these regions were excluded from the structural analysis to enhance the visibility of the ordered domains. A model incorporating the IDRs is provided in the Appendix.

The RNA was predicted to form a double-stranded helix (depicted in dim grey), surrounded by various domains of the ADAR1 protein bound to the RNA (Fig. 39A). The deaminase domain responsible for the RNA modification was predicted to bind at the main editing site of the TTYH2 RNA (Fig. 23A). According to previously published research, the deaminase domains formed the dimerization interface between the two ADAR1 monomers (Fig. 39C) (Thuy-Boun et al., 2020).

The dsRBDs made contact with the TTYH2 RNA. The Z-domains of one monomer were close to the RNA, while the corresponding Z-domains from the second monomer did not directly contact the RNA. This observation suggests that there may not be sufficient spatial accommodation for these domains to bind effectively. Additionally, one region of the RNA remains unoccupied by any domains, indicating a potential need for optimization of the RNA sequence to allow for the effective accommodation of all domains.

To compare the structure with the ADAR2-TTYH2 complex, AlphaFold3 predictions were conducted for this complex. The server utilized the TTYH2 RNA sequence, two copies of the protein-coding sequence, and two zinc ions. It accurately predicted all domains, including the correct positioning of one deaminase domain (monomer 1) at the target adenosine, as well as the dimerization interface between the two deaminases. Additionally, it was observed that the dsRBD2 of monomer 2 interacts with the RNA (Fig. 39B). These findings align with the published crystal structure (Thuy-Boun et al., 2020).

Moreover, in published crystal structure it was speculated that dsRBD2 from monomer 1 could bind to RNA (on the right side relative to deaminase monomer 1) to form asymmetric dimers; however, this interaction was not demonstrated in the predicted model for TTYH2 RNA (Fig. 39B). The predictions also indicated that a significant portion of the RNA remains unoccupied, suggesting it may not be an ideal target for studying the ADAR2-RNA complex.

Interestingly, the orientation of the deaminase domains varies between the two enzymes. In ADAR1, one of the deaminase domains is positioned behind the other, whereas in ADAR2, the predicted orientation is more consistent with the published crystal structure.

Upon examining the catalytic site of ADAR1, the model revealed a base-flipping mechanism consistent with structural studies on the ADAR2-dsRNA complex. The ADAR1 E912 residue, which is essential for the deamination process, is in direct contact with the flipped adenosine. Meanwhile, the ADAR1 E1008 residue occupies the space vacated by the flipped-out base (Fig. 39D).

Structural studies of ADAR2 have demonstrated the presence of a loop that undergoes conformational changes upon binding to the RNA substrate (Matthews et al., 2016). Comparing the sequence alignment of this loop between the two enzymes reveals differences in both sequence and size (Park et al., 2020). By highlighting this loop in yellow, it became evident that only the ADAR2 loop successfully contacts the RNA substrate through positively charged residues according to the previously published crystal structure. The residue His988

from ADAR1, located in the 5' loop (depicted in pink), is involved in coordinating the second zinc ion, which subsequently restricts the loop's binding to RNA (Fig. 39E).

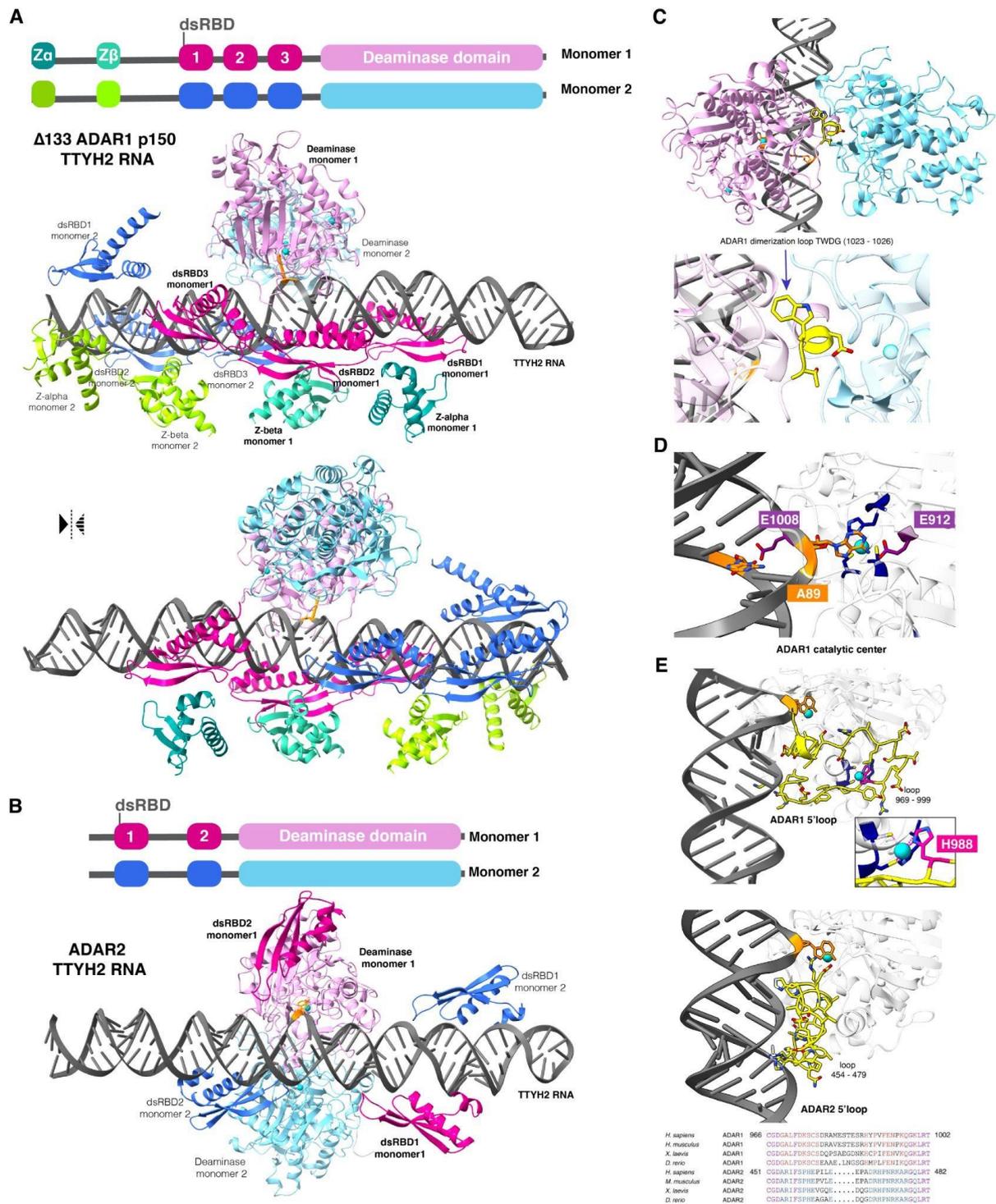


Figure 39 AlphaFold3 Predictions of the ADAR - TTYH2 RNA Complex.

- A. $\Delta 133$ ADAR1 p150 - TTYH2 RNA complex. The structure is colored in accordance to the ADAR1 domains scheme.
- B. ADAR2 - TTYH2 RNA complex. The structure is colored in accordance to the ADAR2 domains scheme.

- C. Dimerization interface of ADAR1 deaminase domain. The interface is illustrated with the TWDG residues from monomer 1 (1023-1026) highlighted in yellow.
- D. The catalytic center of ADAR1. The converted adenosine, which undergoes deamination, is marked in dark orange, and key residues involved in the interaction are highlighted in purple.
- E. Comparison of the 5'loops from ADAR enzymes along with their sequence alignment. The ADAR1 loop also features highlighted residues that coordinate the second zinc ion (dark blue)

4.6 Pacific oyster ADAR

This work was part of a collaborative project with researcher groups from the Department of Biology at the University of Padova, led by Paola Veroni and Umberto Rosani. The majority of the research was conducted by Enrico Rosani, a PhD student who is now a postdoctoral researcher within the same group. The collaborators had identified an ADAR variant that was induced in *C. gigas* infected with Ostreid herpesvirus-1 (OsHV-1). Our contribution to this project was to prepare the recombinant cgADAR1v protein, by using the expertise from purifying human ADARs. We then focused on testing the activity of the cgADAR1v protein and preparing the RNA samples for nanopore sequencing. The nanopore sequencing data was subsequently optimized and analyzed by our collaborators.

4.6.1 *C. gigas* ADAR protein identification and purification

The collaborators identified the ADAR variant that was most strongly induced in the antiviral response by analyzing 95 RNA-seq samples from the oyster *C. gigas* infected with the virus OsHV-1 (Fig. 40a). The analysis revealed that CgADAR1v (LOC105341503) was the most upregulated, with an average expression level (TPM - transcript per kilobase million) of 226 ± 199 , and a peak of 1044 (Fig. 40b). Another ADAR1 candidate, CgADAR1 (LOC105340589), showed intermediate expression levels, with an average TPM of 43.75 (Fig. 40c).

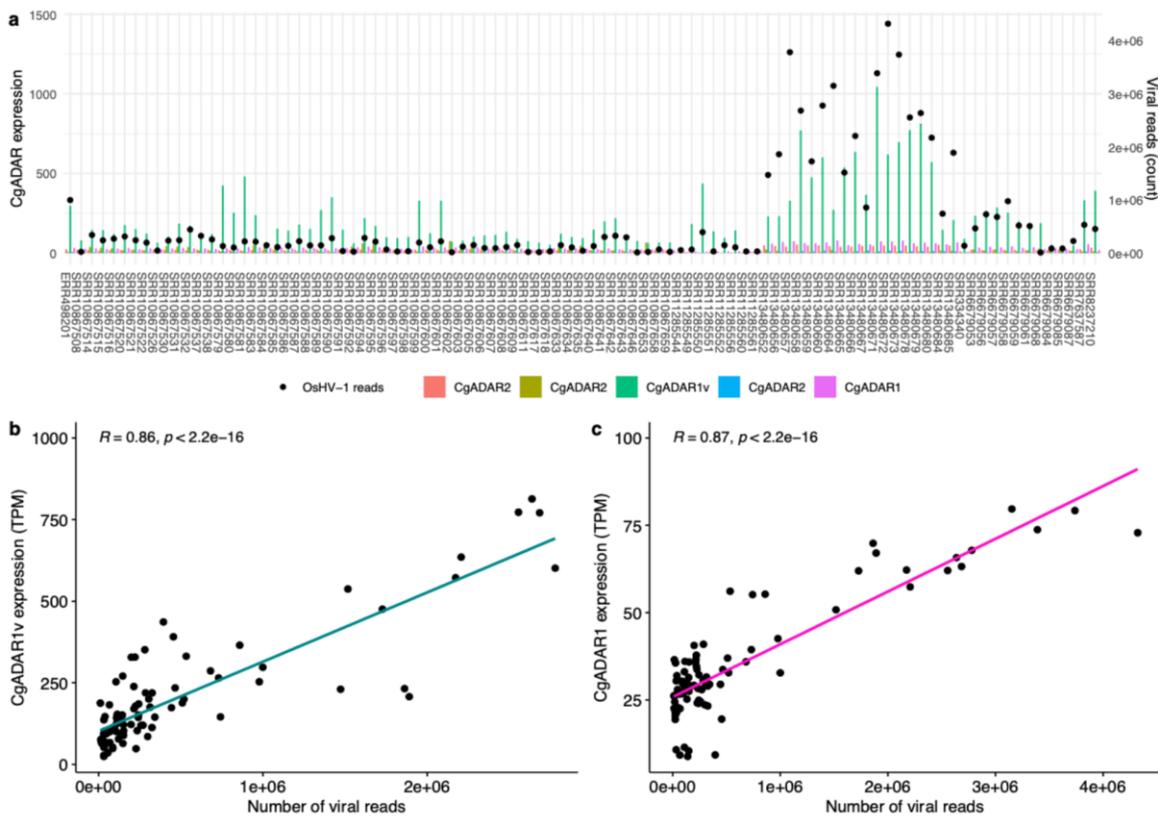


Figure 40 . Identification of ADAR variant in *C. gigas* induced in OsHV-1 antiviral response.

The expression levels of all identified CgADAR genes are represented in distinct colors on the histogram, black dots indicate the number of viral reads displayed on a logarithmic scale (a). The linear regression analyses between the viral read counts and the expression levels of CgADAR1v (b) and CgADAR1 (c).

The AlphaFold3 structure prediction for CgADAR1v indicates that the protein consists of a deaminase domain and three dsRBD3 domains (Fig. 41A). In contrast, earlier analyses conducted with the Phyre2 structure prediction tool, before the availability of the AlphaFold3 model, had only identified two dsRBD domains (data not shown).

For the initial purification trials, an insect cell expression system was chosen due to its cost-effectiveness. Domain architecture of CgADAR1v is more similar to that of ADAR2; specifically, the protein contains only deaminase and dsRBD domains and lacks Z-domains, therefore the protocol for ADAR2 was chosen in initial purification trials which also offers a more straightforward approach for isolating the target protein compared to the ADAR1 purification protocol.

After cell lysis, the CgADAR1v protein was purified using a sequential approach. It was first loaded onto a HisTrap column, followed by a Heparin resin column, tag removal, and a final purification step using size exclusion chromatography. This purification process yielded a high-purity protein, approximately 800 µg from one liter of cell culture. The purity of the protein was confirmed using mass photometry, which revealed a high level of protein purity (Fig. 41 B).

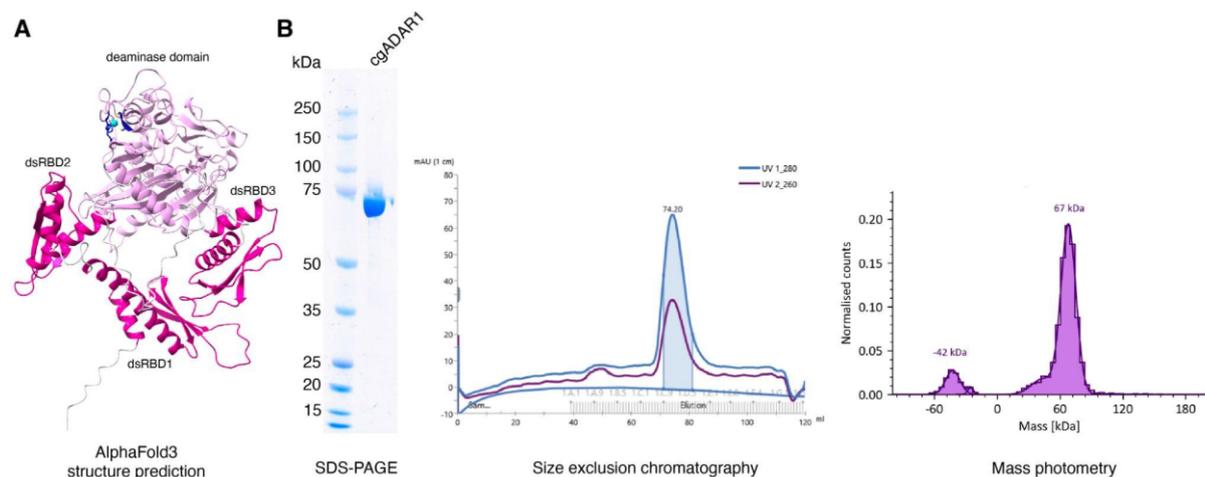


Figure 41 Purification of the CgADAR1v.

- A. Structure prediction was generated using AlphaFold3. Deaminase domain is depicted in plum while the dsRBD domains in medium violet red.
- B. Purification of the CgADAR1v, including an SDS-PAGE gel showing the final purified product, the elution pattern obtained from size exclusion chromatography, and results from mass photometry analyses.

4.6.2 Testing enzymatic activity of CgADAR1v using Sanger sequencing

The activity of CgADAR1v was tested using the same approach as described for the human ADARs. During the initial optimization, different editing temperatures were evaluated to determine the most optimal for CgADAR1v. Among the tested temperatures of 20, 25, and 30 °C, the protein demonstrated the highest activity at both 25 °C and 30 °C, showing comparable levels of activity. Since 30°C is also more optimal for human ADARs, this temperature was chosen for the subsequent experiments. The substrates previously produced for the human ADAR experiments were utilized. Unfortunately, there are no available sequences for CgADAR1v that could be used in the in-vitro experiments.

Analysis of the Sanger sequencing results revealed that the 5HT2C substrate (Fig. 42A) exhibited a significantly lower editing level across all sites compared to human ADARs. Another substrate evaluated was the Gria2 R/G RNA (Fig. 42B), which showed approximately 50% conversion at the R/G site, while the human proteins achieved complete conversion for the same site. Editing activity was also detected at the -1 and -13 sites. In contrast, no editing was observed for NEIL1 RNA (Fig. 42C). The TTYH2 site, marked in magenta, showed lower conversion levels compared to human ADARs, and CgADAR1v demonstrated minimal editing activity at the -40 site (Fig. 42D). These experiments confirmed the activity of the recombinant protein, indicating its potential for further testing using nanopore sequencing.

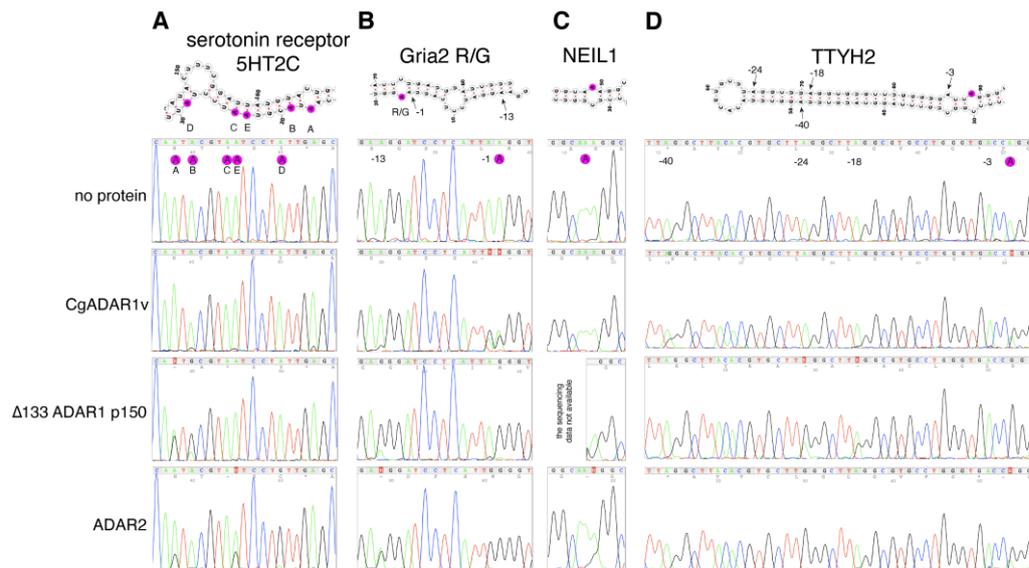


Figure 42 Comparison of editing patterns for cgADAR1v and human ADARs.

- Editing of 5HT2C. The results were obtained from sequencing the antisense strand, and they are presented as a reverse complement conversion.
- Editing of Gria2 R/G. The results were obtained from sequencing the antisense strand, and they are presented as a reverse complement conversion.
- Editing of NEIL1. The results were obtained from sequencing the sense strand.
- Editing of TTYH2. The results were obtained from sequencing the sense strand.

The adenosine at position 0 is highlighted in magenta, while other converted sites are indicated with arrows. Nucleotide abbreviations: A - adenine, T - thymine, C - cytosine, G - guanine.

4.6.3 Testing enzymatic activity of CgADAR1v using nanopore sequencing

The RNA prepared during sample preparation for Sanger sequencing was subsequently analyzed by collaborators using Nanopore sequencing.

The inosines detected in the Nanopore sequencing were in agreement with the previous Sanger sequencing results. However, the editing frequencies differed between the two techniques. For example, the highest editing activity for CgADAR1v on the 5HT2C RNA was detected at site B, which could not be seen in the prior Sanger sequencing. The inosine on the R/G site of Gria2 was detected at a much lower level compared to the previous Sanger sequencing results. The editing frequencies at sites -1 and -13 also did not correlate with the earlier findings. Similar discrepancies were also observed for the NEIL1 and TTYH2 RNAs (Fig. 43).



Figure 43 Nanopore sequencing results.

The frequency of RNA edited sites of the analyzed targets: 5HT2C, Gria2 R/G, NEIL1, TTYH2. Nucleotide abbreviations: A - adenine, T - thymine, C - cytosine, G - guanine.

The obtained result indicates that nanopore sequencing has limitations in studying inosine modifications. The noise generated during the sequencing process can lead to an underestimation of inosine detection, making it challenging to accurately identify and quantify this modification. Furthermore, the current flow cell technology used in nanopore sequencing has constraints in its ability to distinguish properly modified RNA. The design and setup of the flow cell, have not yet achieved the level of precision necessary for the reliable detection of modified nucleotides. Additionally, the results obtained from nanopore sequencing may be influenced by biases present in the bioinformatics analysis.

4.7 MDA5 - irAlu complex characterization

The most abundant endogenous target for MDA5 activation is the irAlu repeats, which are also the RNA sequences most extensively edited by ADAR1. The results presented below focus on investigating the interaction between MDA5 and irAlu repeats, aiming to gain deeper insights into how A-to-I editing influences this interaction.

4.7.1 MDA5 Δ CARDs purification

MDA5 was produced without the CARD domains as the primary goal of this part of the thesis is to understand the binding properties of MDA5 to irAlu RNA, for which the CARD domains are not directly involved. Construct with truncated CARD domains showed little difference in filament morphology when compared to the full length version of the protein (Wu et al., 2013b). Future work involving MDA5 will focus on structural studies with irAlu RNA. The CARD domains are highly dynamic, exhibiting significant flexibility and mobility, making it challenging to capture them using CryoEM. Additionally, working with the full-length protein presents challenges and limitations in preparation, making the truncated version more suitable for this research.

Additionally, during the purification process, the protein was dephosphorylated through treatment with PP1 α . Previous studies have demonstrated that MDA5 is phosphorylated, and dephosphorylation of MDA5 is crucial for regulating innate immune signaling (Wies et al., 2013). While the phosphorylation sites were initially identified in the CARD domains, research conducted in our laboratory has revealed additional phosphorylation sites that are also targeted by PP1 α .

The protein was purified from an insect cell expression system. After cell lysis, it was subjected to affinity purification, capturing the His₆ tag. The protein was then dialyzed in the presence of 3C protease to remove the affinity tags and the PP1 α phosphatase to dephosphorylate the protein. Following the dialysis, the protein was further purified using a Heparin column, HisTrap reverse chromatography, and finally, size exclusion chromatography. The overall protein yield was 4 mg from one liter of the cell culture (Fig. 44). Mass spectrometry analysis showed that after treatment with PP1 α , the protein exhibited reduced phosphorylation levels, and full dephosphorylation could not be achieved.

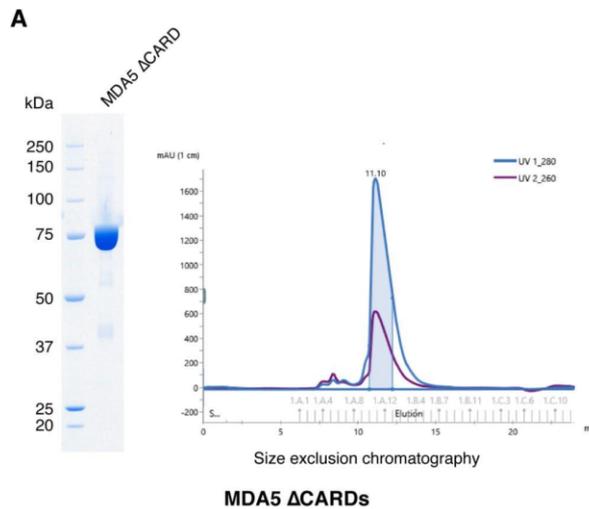


Figure 44 MDA5 Δ CARDs purification.

- A. SDS-PAGE gel showing the final purified product and the elution pattern obtained from size exclusion chromatography. The purified protein has decreased phosphorylation levels.

4.7.2 Preparing irAlu RNA substrates

To study the interaction between MDA5 Δ CARDs and irAlu repeats (irAlus), the different irAlu substrates were prepared, including full-length versions and shorter variants. The substrates used, such as irAlu NICN1 and BPNT1, were designed based on sequences reported in previous studies on MDA5 gain-of-function mutants associated with Aicardi-Goutières syndrome (Ahmad et al., 2018). Initially, the full-length versions were produced, and later, shorter variants were also created (Fig. 45A).

The rationale for preparing the shorter versions was to facilitate structural studies, it was reasoned that these shorter substrates would accommodate fewer MDA5 molecules, making data processing easier. The irAlu NICN1 and BPNT1 substrates were split approximately in half to generate the shorter variants. Additionally, a shorter version of irAlu NICN1 A was produced, with the sequence design based on an ADAR1 editing footprinting assay. Briefly, full-length irAlu NICN1 - MDA5 filaments were formed, followed by the addition of ADAR1 to the reaction to allow for editing. After analyzing the sequencing results, shorter regions lacking editing sites—specifically those covered by MDA5—were selected for designing the shorter substrate.

The RNA substrates were generated through in vitro transcription using PCR templates. Specifically, irAlu NICN1 A was produced by separately synthesizing the sense and antisense strands, which were then annealed. In contrast, all other RNA substrates were synthesized as single pieces and subsequently subjected to folding experiments to produce dsRNA. Various folding techniques were tested, including a fast folding method in which the RNA was heated to 95 °C for 5 minutes and then rapidly cooled on ice. Another approach involved slow folding, where the RNA was heated at 95 °C for 5 minutes and then cooled to 4 °C at a controlled rate of 1 °C every 30 seconds. Control samples were also prepared, including one without any folding treatment and another with RNA denatured in urea loading buffer.

The analyses revealed that the full-length RNA versions were most homogeneous when not subjected to any folding technique, indicating that the RNA purified after in vitro transcription was already in a well-folded conformation. In contrast, the slower folding method proved to be the most effective for the shorter RNA variants, as demonstrated by the Native PAGE analysis (Fig. 45B)

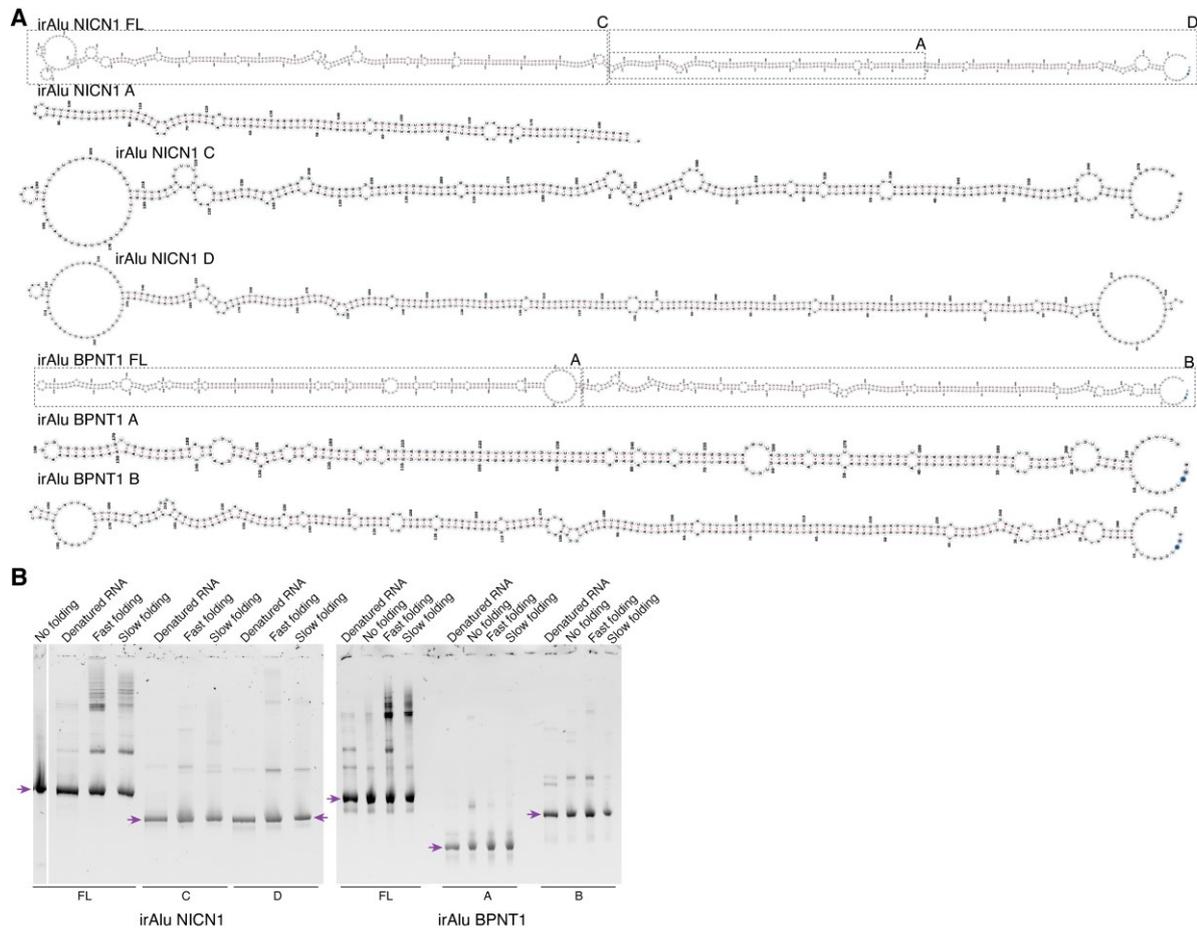


Figure 45 Generating irAlu RNA Substrates.

- A. Predicted secondary structures of the full-length and shorter versions of irAlu NICN1 and BPNT1 substrates. The secondary structures were generated using the RNAfold web server and visualized with the Forna tool.
- B. Native PAGE analysis of the folded irAlu RNA substrates using different techniques, including no folding, fast folding, and slow folding. The band corresponding to the folded RNA is marked with a purple arrow.

4.7.3 Complex binding

Filament formation was investigated in the presence of ATP to simulate the physiological conditions. ATP hydrolysis is essential for MDA5 activation, as it is associated with conformational changes in the protein and the regulation of filament formation (Yu et al., 2018b). Through pilot experiments, a concentration of 5 mM ATP was optimized, revealing that filaments formed at concentrations below 5 mM, while no stable filament formation was observed at higher concentrations, as confirmed by negative staining.

The complex formation between MDA5 Δ CARDs and various irAlu RNAs was tested by running EMSA. Briefly, the folded RNA was incubated with increasing MDA5 concentrations in the

presence of 5 mM ATP for 15 minutes, and the samples were then loaded onto a Native PAGE gel. The analysis revealed that MDA5 Δ CARDs can bind to all the substrates that were produced. Specifically, the full-length irAlu variants exhibited similar binding to their truncated counterparts (Fig. 46). It is important to note that the protein concentrations used in the EMSA experiments with NICN1 C and D were lower than those used for the full-length version of the RNA.

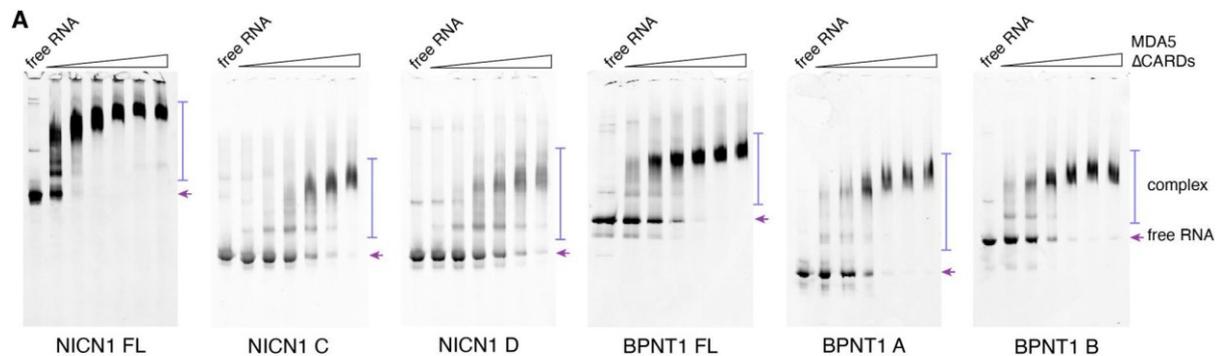


Figure 46 Binding of MDA5 Δ CARDs to various irAlu RNAs analyzed by EMSA.

The free RNA is marked with an arrow while the MDA5-bound (complex) is indicated with a line.

The RNA concentration was kept at a constant 50 nM. In the EMSA experiments with NICN1 FL and BPNT1 FL, A and B, MDA5 was used at the following concentrations: 0 (free RNA lane), 0.25, 0.5, 0.75, 1, 1.25, and 1.5 μ M. For NICN1 C and D, the concentrations of MDA5 were as follows: 0 (free RNA lane), 0.05, 0.125, 0.25, 0.5, 0.75, and 1 μ M.

4.8 Filament formation

The formation of filaments was evaluated using negative staining techniques. The filaments were formed between MDA5 Δ CARDs and various irAlu RNAs in the presence of 5 mM ATP. Efficient filament formation occurred with all tested RNA substrates, in agreement with earlier EMSA results.

The negative staining technique facilitated the evaluation of filament morphology and quality. Notably, the filaments formed with the full-length RNA versions were longer and demonstrated a more uniform length distribution compared to their truncated counterparts. However, these full-length filaments exhibited some degree of bending, indicating potential flexibility in their structure. In contrast, the truncated RNA filaments exhibited greater variability in length, suggesting a tendency toward dissociation or incomplete assembly (Fig. 47).

Based on these observations, the full-length RNA versions were selected for further studies utilizing CryoEM.

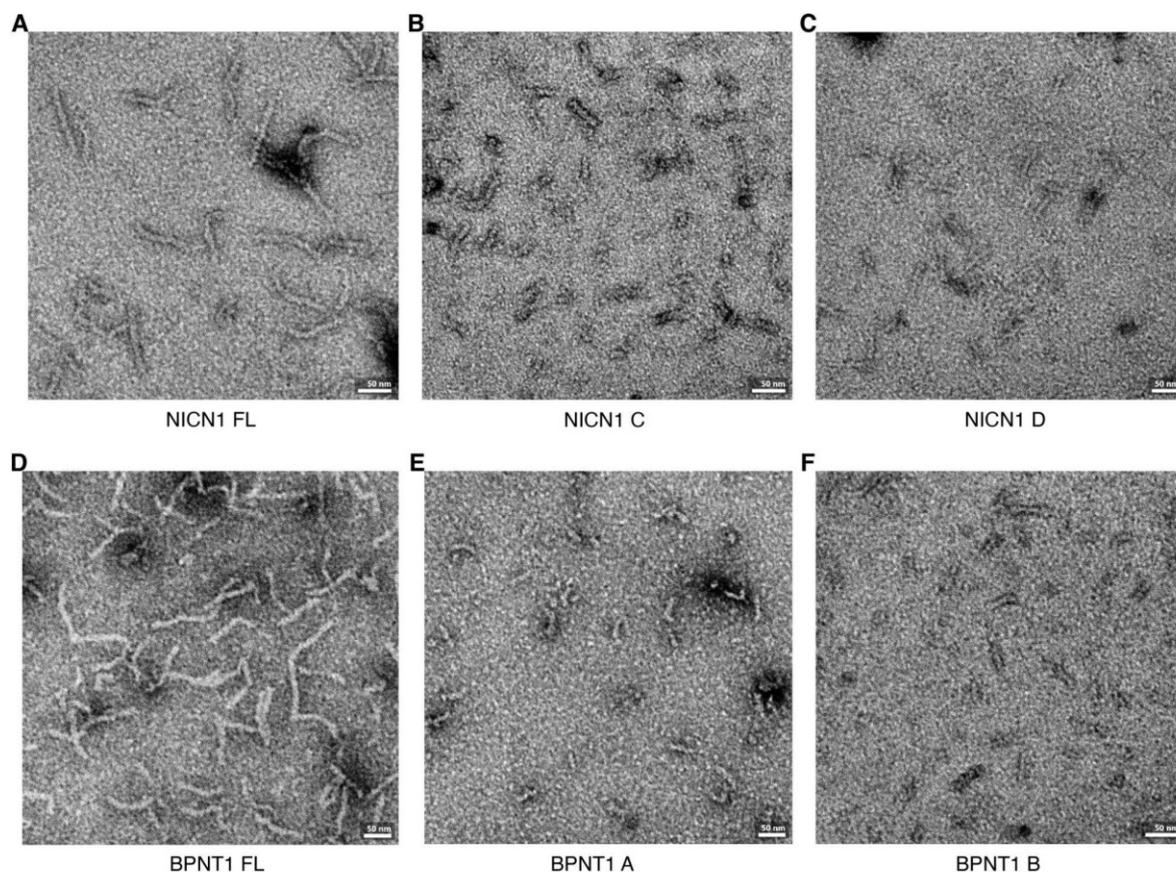


Figure 47 Negative staining analysis of the MDA5 Δ CARDs - irAlu RNAs filaments. irAlu NICN1 FL (A), irAlu NICN1 C (B), irAlu NICN1 D (C), irAlu BPNT1 FL (D) irAlu BPNT1 A (E), irAlu BPNT1 B (F) The images were captured at a nominal magnification of 37 000x

4.8.1 CryoEM analysis

A pilot experiment was conducted to evaluate the sample preparation efficiency for cryoEM analysis. It is crucial that the CryoEM samples contain as little glycerol as possible in the buffer, as glycerol can interfere with image contrast and compromise the quality of the acquired data.

Prior to filament assembly, the protein was concentrated and subjected to size exclusion chromatography to exchange the buffer for a glycerol-free composition. Following this step, filaments were assembled using the full-length RNA variants, and a CryoEM sample was prepared.

The samples were then evaluated using the Glacios microscope to assess filament quality. The analysis revealed that the majority of the RNA was present in a free form, and MDA5 Δ CARDs appeared to be predominantly dissociated (Fig. 48A, B, filaments that still contain MDA5 Δ CARDs are indicated by purple arrows).

The pilot experiment shows a need to stabilize the MDA5 Δ CARDs that probably didn't resist that well in the condition without the glycerol. Additionally, the experiment provided valuable insights into the RNA concentration required to prepare a cryoEM sample that ensures optimal filament distribution and density.

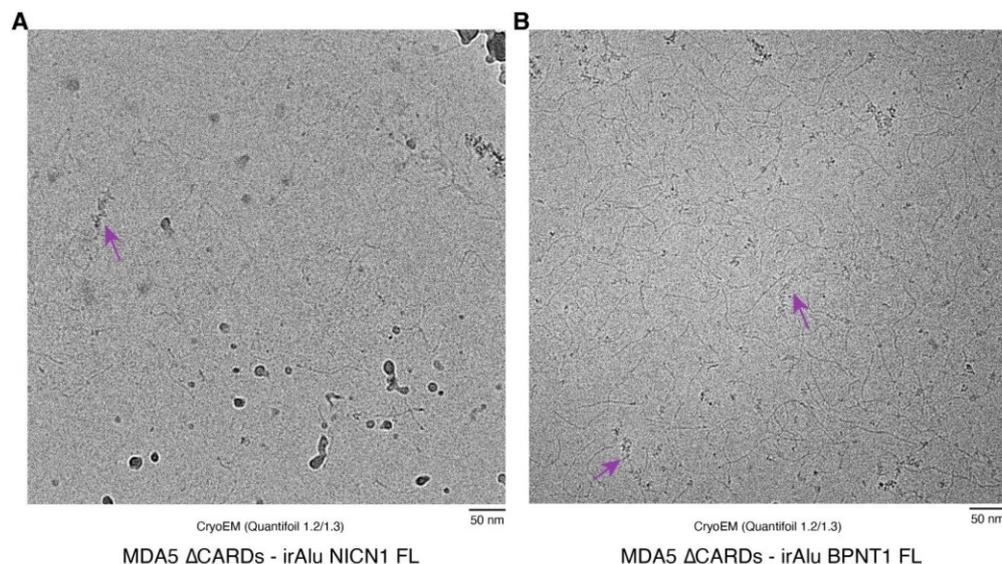


Figure 48 CryoEM analysis of the MDA5 Δ CARDs - irAlu RNAs filaments.

- A. Filaments assembled with irAlu NICN1 FL
- B. Filaments assembled with irAlu BPNT1 FL

The image contrast was adjusted to enhance the visibility of the MDA5 filaments.

4.8.2 irAlu editing by ADAR enzymes

The irAlu elements are among the most frequently edited targets by ADARs, primarily due to their double-stranded RNA properties (Bazak et al., 2014). Studies have shown that editing of these elements effectively inhibits MDA5 binding and activation (Chung et al., 2018).

The editing profiles were assessed using an editing assay, as described earlier in this thesis. The editing reaction was carried out on the full-length irAlu NICN1 RNA, which was subjected to conversion by the human ADAR enzymes, particularly ADAR2 and the p110 isoform of ADAR1. The results for the longer isoform are not shown, as generally lower editing levels were achieved compared to the shorter isoform. The analysis revealed that the irAlu NICN1 RNA could be efficiently edited, as demonstrated by representative regions of the RNA. Editing by ADAR1 at p110 and ADAR2 occurred at multiple sites, with each enzyme demonstrating distinct specificity to a certain degree. Notably, ADAR2 exhibited higher efficiency compared to ADAR1 (Fig. 49).

Furthermore, isolated deaminase domains were tested for their editing efficiency. Surprisingly, no editing activity was observed with the wild-type versions of ADAR1 or the ADAR2 deaminase domains. However, editing was detected using the ADAR2 N496F E488Q mutant. The ADAR2 E488Q mutant was also evaluated, but no editing activity was observed with this variant. These findings suggest that the presence of dsRBDs is essential for efficient editing, which can also be achieved through the tested mutant that enhances editing at 5'-GAN-3' motifs (Fig. 49).

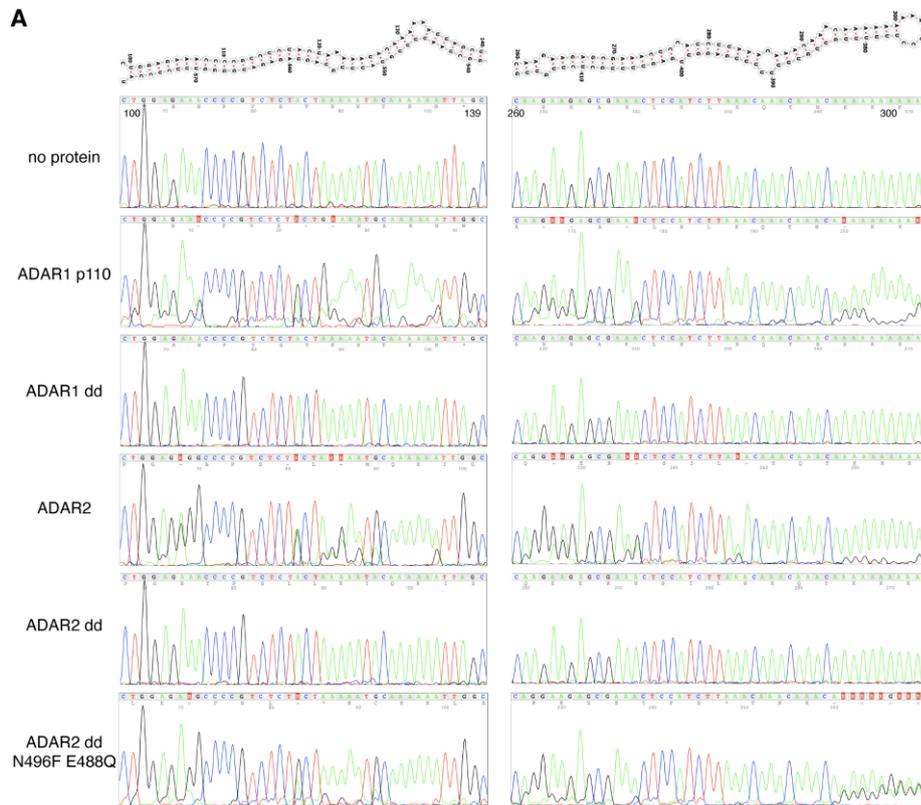


Figure 49 Editing of irAlu NICN1 FL by ADAR enzymes and isolated deaminase domains.

A. The analysis presents two representative regions: nucleotides 100-139 and 260-300. The results were obtained from sequencing the antisense strand, and they are presented as a reverse complement conversion. Nucleotide abbreviations are as follows: A - adenine, T - thymine, C - cytosine, G - guanine

4.8.3 MDA5 Δ CARDs interaction with edited irAlu

The binding ability of MDA5 Δ CARDs to A-to-I modified RNA was evaluated using EMSA. In this experiment, irAlu NICN1 A RNA was edited by ADAR2, while non-modified RNA served as a control. The analysis revealed that the A-to-I modified NICN1 A RNA displayed a lower binding affinity to MDA5 Δ CARDs compared to the non-edited RNA, as the filaments formed more completely at lower concentrations for the non-edited RNA (Fig. 50).

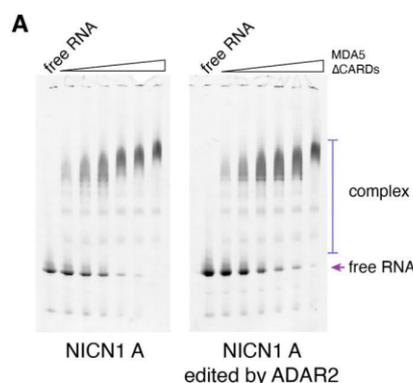


Figure 50 Binding of MDA5 Δ CARDs with irAlu NICN1 A, non-modified or edited with ADAR2.

5 Discussion

5.1 Biochemical and structural insights into ADAR1 substrate selectivity properties

A significant part of the work done during this PhD project focused on understanding the substrate recognition properties of ADAR1, using both biochemical and structural approaches. To achieve this goal, the project involved several steps, culminating in the acquisition of a cryo-EM data of ADAR1 bound to a dsRNA substrate.

Optimization of ADAR1 purification protocol

The project initially focused on optimizing the purification protocol for ADAR1, which presented more challenges than initially expected. The purification protocol that had been previously optimized for ADAR2 was found to be ineffective, necessitating the establishment of a new protocol. The optimization process revealed several factors that complicated the efficient purification of pure and active ADAR1 protein. However, these challenges could be effectively overcome through a comprehensive optimization strategy.

The purification protocol involved several key steps, including minimizing nucleic acid contamination, adding affinity tags to both the N- and C-termini to ensure the retention of intact protein, and removing the intrinsically disordered region from the N-terminus of ADAR1 p150.

ADAR1 has multiple RNA-binding domains that interact with various nucleic acids, which led to significant contamination during the purification process. This contamination could be reduced by adding EDTA to the purification buffer. EDTA acts as a chelator for divalent ions, which are often present as cofactors for nucleic acid-binding proteins. Its inclusion likely reduced the overall binding of nucleic acids within the lysate, thereby enhancing the purity of ADAR1. Although ADAR1 contains two zinc ions in its deaminase domain, these ions are not affected by presence of EDTA. A study investigating binding of the second zinc ion found that zinc binding remained intact even when the protein was stored in a buffer containing 0.5 mM EDTA (Park et al., 2020). Additionally, the presence of EDTA did not impede the enzymatic activity of ADAR1, which relies on zinc in its catalytic center.

Purifying the longer isoform ADAR1 p150 presented additional challenges. This issue could be resolved by removing the first intrinsically disordered region from the N-terminus of the protein. This sequence impaired the protein's binding during the affinity purification steps, which may have also impacted its folding. With the linker removed, the protein maintained its activity, allowing subsequent analyses to focus on structural studies. Moreover the intrinsically disordered regions would not be expected to be resolved in a CryoEM structure due to their dynamic nature (Musselman and Kutateladze, 2021).

By integrating these key factors and optimizing several other steps described in the results section, it became feasible to obtain pure and active protein from the mammalian expression system. Achieving a high degree of protein purity facilitated further progress toward the structural analysis of ADAR1 in complex with its RNA substrate.

Reconstitution and biochemical characterization of ADAR1-dsRNA complex

The biochemical characterization process started with the production of dsRNA substrates known to be compatible with in vitro studies of ADARs. Although sequencing techniques have offered insights into the locations of editing events, they do not reveal the specific regions of the RNA that formed the dsRNA structures targeted by ADARs.

The vast majority of ADAR1 editing sites are found within irAlu repeats, and their secondary structures can be sourced from existing literature. However, these substrates are not suitable due to the excessive number of editing sites, which could complicate the structural analysis. To study the complex substrates, RNAs with a limited number of editing events, such as 5HT2C, Gria2 R/G, NEIL1, TTYH2, or AJUBA RNA, were selected and produced. ADAR1 efficiently edited all tested targets, with the longer isoform exhibiting higher activity. The editing reactions were also evaluated with ADAR2, and the differences in their enzymatic activities could be particularly observed in the editing of 5HT2C, where the enzymes exhibited different target selectivity which is in agreement with the previously published reports (Eggington et al., 2011). Based on the RNA secondary structure prediction it can be speculated that ADAR1 can target adenosines within predicted mismatches, whereas ADAR2 appears to favor RNA stem region. However it's based on isolated case and reliable conclusions could be drawn by analyzing the structure of the complex. The editing of the TTYH2 RNA also exhibited different substrate selectivity, but the underlying reasons for this are unclear.

The optimization of complex formation was tested using EMSA, which indicated that the shorter isoform is more prone to aggregate with the tested substrates. ADAR1 p150 demonstrated better binding, and successful conditions were identified for forming a complex suitable for CryoEM analysis. Binding affinities were measured using fluorescence anisotropy, which revealed a relatively high K_d value (~340nM), compared to the binding experiments conducted with EMSA. Although the ADAR1 p110 and NEIL1 RNA EMSA experiments were not quantified for K_d due to their low reliability and potential biases introduced by the lack of a homogenous complex, it was observed that the complex was present below 120 nM before the aggregates appeared. The K_d value obtained from the fluorescence anisotropy experiments could have been affected by the artifacts introduced by the complex aggregation observed in the EMSA experiments. Additionally, the absence of an appropriate model for accurately calculating binding affinity may have contributed to this elevated K_d value. Given these limitations, further affinity measurements with other RNA substrates were not conducted, as this technique might not be optimal for this system. To better characterize the binding affinities, it would be advisable to shift to different techniques, such as initiating measurements with microscale thermophoresis.

Structural characterization of ADAR1 bound to dsRNA

The initial optimization of Cryo-EM samples concentrated on the shorter isoform; however, the optimization tests revealed that this isoform was unsuitable due to aggregate formation and challenges in localizing it within the grid holes. As a result, the focus shifted to the longer isoform. Of the various RNA substrates tested, only Gria2 R/G RNA was found to be suitable for preparing the complex on holey carbon.

The subsequent data processing revealed that the complex was predominantly dissociated, underscoring the need for implementing a crosslinking agent. Various crosslinkers were

evaluated through a careful assessment of their crosslinking reactions. BS3 was found to be ineffective, as it dissociated the RNA from the complex, while glutaraldehyde did not disrupt the RNA binding but generated nonspecific crosslinks between RNA molecules, which could be observed in the cryo-EM sample.

The most promising crosslinker identified was SPB, which effectively crosslinks both protein and RNA upon exposure to UV light. Through the optimization of various parameters for cryo-EM sample preparation—including support layers, grid types, and blotting devices—a suitable sample was ultimately obtained for data collection. However, data processing revealed that even with crosslinking, the complex remained highly heterogeneous, preventing the successful determination of the complex structure.

While determining the structure using cryo-EM proved to be challenging, AlphaFold predictions of d133 ADAR1 p150 bound to TTYH2 RNA provided valuable insights into its interactions with RNA substrates. As expected, regions critical for catalytic activity were highly conserved, including the Gly-Glu-Gly loop, zinc-coordinating residues, and the glutamate essential for proton shuttling during the catalytic reaction. The model also indicated a flipped-out adenosine in the catalytic center, aligning with the findings from the editing assays (Matthews et al., 2016). Although editing activity was observed at multiple locations within the TTYH2 RNA, this specific adenosine corresponds to one identified in previous studies (Liu et al., 2021). Furthermore, the model successfully predicted dimer formation, featuring a conserved interface similar to that observed in ADAR2 asymmetric dimers (Thuy-Boun et al., 2020).

Significant differences were observed between ADAR1 and ADAR2 in the 5' binding loop region. In alignment with previously published crystal structures, the ADAR2 loop was found to interact with the RNA via positively charged residues (Matthews et al., 2016). In contrast, the ADAR1 loop is situated further away from the RNA binding site, with its interaction being constrained by His988, which is part of the tether that coordinates the second zinc ion (Park et al., 2020). The structural variation implies that ADAR1 may be more effective at targeting adenosines that are followed by loops or bulges in the RNA secondary structure. ADAR1's 5' binding loop is less directly involved in contacting the RNA backbone, which could allow it to better accommodate larger features in the RNA secondary structure.

The AlphaFold predictions provided insightful information on how the various ADAR1 domains can interact with the RNA substrate. The model indicated that one portion of the RNA remains unoccupied, while on the opposite side, there is insufficient space to accommodate all of the domains. It is important to recognize that this is a predictive model; under cellular conditions, the positioning of the dsRBDs or Z-domains may be further influenced by the flexible linkers between these domains. Furthermore, it is unclear if all domains are necessary in this specific context.

Nevertheless, the insights gained from the AlphaFold predictions can help elucidate the heterogeneity seen in the cryo-EM data. The existence of flexible, unoccupied regions in the RNA likely contributes to the aggregation of the complexes, resulting in the observed heterogeneity within the particle populations. These flexible areas also introduce additional noise, which may complicate data processing and analysis. This phenomenon has also been observed with other RNA-binding proteins in our laboratory. While investigating RNA Polymerase II in conjunction with non-coding RNA, reducing the RNA sequence allowed for

the acquisition of higher-resolution structures. In contrast, longer RNA substrates frequently resulted in complex aggregation.

Final conclusions:

Significant efforts were dedicated to understanding the ADAR1 substrate recognition properties from both biochemical and structural perspectives, yet fully addressing this question proved challenging. This emphasizes the complexity of the studied system, which presented several biochemical challenges. The optimization process for protein purification, complex formation, and structure determination via cryo-EM required considerable time and resources.

The thesis outlines an optimized purification protocol for ADAR1 and presents a preliminary low-quality cryo-EM model of the complex. However, due to time limitations, further progress on the sample preparation for the CryoEM was not feasible, and it seems that much more optimization would be required.

The challenges encountered during this study demonstrate the inherent difficulties involved in the structural characterization of RNA-binding proteins, underscoring the need for further improvements in sample preparation and data collection for cryoEM analysis.

Future directions:

1. Optimizing more suitable RNA substrate

The RNA substrates used may not be optimal for cryo-EM analysis, as evidenced by the aggregation observed during sample freezing and the AlphaFold predictions suggesting the presence of unoccupied, flexible regions in the RNA. This underscores the importance of identifying a minimal RNA substrate length that is suitable for binding and editing by ADAR1.

To determine the optimal substrate length, a RNase protection assay could be a valuable approach. Initial attempts were made to optimize the RNA footprinting through RNA hydrolysis, but no cleavage was observed. An alternative strategy may involve testing RNase that is specific to dsRNA regions. Our laboratory is currently exploring the production of the enzyme as it is no longer available for purchase.

Furthermore, the stability of the complex may be insufficient, leading to potential dissociation from the RNA substrate. To enhance stability, RNA substrates incorporating the adenosine analog 8-azanebularine could be utilized. This modification effectively traps the complex in an intermediate state during the catalytic cycle, thereby minimizing dissociation and promoting a more stable interaction with the RNA substrate.

Furthermore, establishing a collaboration with a bioinformatic group could help in the identification of additional dsRNA regions that are suitable for ADAR1 RNA editing. These newly identified regions could then be evaluated using AlphaFold predictions, which have previously shown success in accurately positioning the catalytic domain of ADAR1 at targeted adenosine residues, as evidenced by the observation of flipped adenosine.

2. Testing different ADAR constructs

Simplifying the system by designing shorter constructs of ADAR1, which exclude the Z-domains and concentrate solely on the effects of the catalytic domain and the dsRBDs, could be a beneficial approach. Comparing the RNA editing efficiencies of these shorter constructs with the full-length protein may help assess the contribution and importance of the Z-domains to the overall RNA substrate recognition and editing process on the tested targets.

Furthermore, preliminary findings suggest that disrupting the dimerization interface in the ADAR1 dsRBD3 domain may help in reducing the aggregation. This construct could also be evaluated for further testing.

5.2 Investigating the *in vitro* editing activity of ADAR in Pacific oysters

The identified oyster ADAR variant (CgAdar1v) upregulated during OsHV-1 infection was successfully purified using the established protocols adapted from those developed for human ADARs. This demonstrates the robustness of the protocol, suggesting it may also be applicable to other ADAR homologs.

The purified protein was tested for activity in *in vitro* editing reactions, which were subsequently evaluated using both Sanger and nanopore sequencing techniques. To optimize the conditions for these editing reactions, various temperatures were tested to identify the optimal activity for oyster ADAR. Interestingly, while the typical laboratory growth temperature for oysters is approximately 18°C, the *in vitro* editing activity was significantly lower at 20°C when compared to the higher temperatures of 25°C and 30°C.

By evaluating Sanger sequencing, it was possible to detect enzymatic activity for the oyster ADAR, and in comparison to human ADARs, the protein displayed lower activity. This reduced activity could potentially be attributed to the use of human RNA targets, which may not be the most suitable substrates for the oyster ADAR enzyme.

Subsequently, the edited RNA was analyzed using nanopore sequencing by our collaborators, who began with careful optimization to effectively implement the published protocols. Nanopore sequencing is a direct method that eliminates the need for reverse transcription; in this approach, the RNA strand passes through a solid-state nanopore, and the sequence is determined by changes in electric signals, which are then translated into nucleotide sequences by base caller software. While these signals are well-characterized for canonical nucleotides, modified nucleotides disrupt the current in a way that differs from their unmodified counterparts. Since the software is primarily trained to recognize signals from canonical bases, additional analysis using bioinformatics pipelines is essential for accurately identifying the modified RNA (Bortoletto and Rosani, 2024; Chen et al., 2023; Nguyen et al., 2022).

The collaborators successfully detected inosines, which aligned with the results obtained from Sanger sequencing. However, there were discrepancies in the editing frequencies reported by the two methods. The current base callers do not account for modified nucleotides, which could explain the discrepancy in the detected editing frequencies. The base callers for inosine modification are currently under development, and a better technology may soon be available to more accurately quantify the modified RNA.

5.3 MDA5 interaction with its endogenous target - irAlu repeats

This part of the thesis focused on characterizing the interaction between MDA5 and its endogenous target, the irAlu repeat. Currently, our understanding of MDA5-RNA interactions is primarily derived from structural data involving MDA5 bound to perfectly dsRNA (Wu et al., 2013a; Yu et al., 2021, 2018a). Additionally this research investigated the effects of A-to-I editing on the binding affinity of MDA5 to irAlu repeats.

Structural characterization of MDA5 irAlu filaments

During the research, various irAlu RNAs derived from NICN1 and BPNT1 were successfully synthesized. However, optimizing in vitro transcription posed significant challenges due to the strong secondary structures inherent to irAlu, which often complicated subsequent purification processes. Additionally, the GC-rich nature of these sequences hindered the creation of suitable PCR templates for effective in vitro transcription. To address these challenges, short flanking sequences were added to both ends of the irAlu repeats. This modification facilitated the production of templates for in vitro transcription and enhanced the efficiency of the editing assays during cDNA generation and PCR amplification. Efforts were also made to create independent RNA strands that were subsequently annealed; however, this approach resulted in the formation of higher oligomeric species during RNA folding, rendering these samples unsuitable for further experimentation. Furthermore, the yield of irAlu RNAs was much lower in comparison to the shorter substrates used for characterizing the ADAR1 complex.

The formation of filaments between MDA5 and the synthesized irAlu RNAs was successfully tested in subsequent experiments. It was found that the filaments were only observable at ATP concentrations of 5 mM or lower. Notably, the full-length versions of the RNAs produced more homogeneous filaments in terms of length compared to their truncated counterparts, which exhibited a greater tendency toward dissociation or incomplete assembly. Initially, the shorter regions of irAlus were designed to facilitate downstream processing of the CryoEM samples. Incorporating a non-hydrolyzable analog of ATP could further stabilize these filaments, as previous studies have demonstrated that shorter filaments of approximately 100 bp can be efficiently formed in its presence (Peisley et al., 2012).

The optimization of filament formation in a negative staining setup provided a foundational basis for preparing the Cryo-EM samples enabling the assessment of optimal component concentrations and buffer compositions. However, initial analysis indicated that the samples did not perform well without glycerol in the buffer, necessitating further optimization to achieve more suitable conditions. Additionally, the filaments observed in acquired images exhibited some bending and heterogeneity. However, this was not considered a problem, as the subsequent cryo-EM data processing techniques could account for these variations during the analysis.

Impact of A-to-I editing on MDA5 interactions with irAlu

In the subsequent part of the research, it was aimed to investigate how ADARs A-to-I RNA editing influences the binding properties of MDA5 to irAlus.

Purified irAlu NICN1 RNA was edited using either ADAR1 or ADAR2, along with their respective deaminase domains. Editing was successfully confirmed for both ADAR1 p110 and ADAR2,

revealing distinct differences in their specificity. These discrepancies can be attributed to structural variations between the enzymes, particularly in the 5' binding loop. Furthermore, the presence of an additional zinc ion in ADAR1 may limit the proximity of the ADAR1 loop to the dsRNA, thereby influencing the editing of specific adenosines in distinct ways (Matthews et al., 2016; Park et al., 2020).

Surprisingly, the wild-type deaminase domains were unable to perform the editing activity, indicating that the presence of the dsRBDs is important for ADAR activity on irAlus. This dependence may originate from the dsRBD's role in binding to dsRNA regions or facilitating dimer formation, which cannot be effectively achieved by the isolated deaminase domains alone (Thuy-Boun et al., 2020).

The lack of editing activity by the wild-type versions of the deaminase domains can be effectively compensated by utilizing a mutant that has demonstrated efficacy in targeting the 5'-GAN-3' motifs, where editing is observable. This mutant possesses a catalytically activating mutation, along with the N496F substitution. The N496 residue is located near the adenosine on the unedited strand that base-pairs with the 5' uracil flanking the target adenosine. Notably, the double mutant E488Q, N496F was found to be more effective at editing various 5'-GAN-3' motifs compared to the E488Q single mutant. Additionally, this variant exhibited enhanced editing efficiency at 5'-NAC-3' motifs (Katrekar et al., 2022). However, it is important to note that the editing levels achieved by this mutant did not reach those of the full-length version of ADAR2.

Analysis of the edited RNA revealed a substantial number of editing events, primarily occurring in A-U pairs, thereby supporting the hypothesis that RNA editing can destabilize RNA duplexes. Secondary structure predictions indicate that irAlu NICN1 contains a limited number of A-C mismatches, where RNA editing could enhance stability by introducing more favorable I-C pairs.

Further testing of MDA5's binding ability to ADAR2-pre-edited NICN1 A demonstrated that editing reduced the binding affinity compared to its non-edited counterpart. Initial tests were conducted using a shorter target with a limited number of potential editing sites to make the system simpler. In contrast, the full-length RNA contains multiple sites that can undergo editing, complicating the interpretation of the effects on the resulting RNA properties. Additionally, varying editing efficiencies introduce further complexity to the analysis, as the edited RNA pool may include species with alternative secondary structures.

Overall, while MDA5 binding to edited RNA exhibits reduced affinity, the underlying mechanisms by which inosine influences RNA properties requires further investigation. One promising approach could involve using shorter RNA fragments with inosine modifications at specific sites, allowing for a more controlled setup. This could be complemented with RNA secondary structure prediction tools that specifically account for inosine modifications.

These initial findings establish a foundation for further exploration of MDA5 binding to endogenous irAlu targets and the effects of ADAR-mediated RNA editing on these binding properties. Currently, the research remains in a preliminary phase, with the determination of the structure as a future objective. Moreover, the independent discoveries highlight the need to revise the complex formation conditions, which are discussed below.

Future directions

1. Understanding the functional consequences of MDA5 phosphorylation on irAlu recognition

The experiments aimed to utilize dephosphorylated MDA5, as previous studies have shown that dephosphorylation plays a crucial role in regulating its activity (Wies et al., 2013). Alternative studies in our lab have shown that phosphorylation is crucial for regulating MDA5's activity on dsRNA or poly-IC (RNA that mimics viral infections), with this activity being further influenced by various ATP nucleotide analogs. This finding highlights the importance of understanding the role of phosphorylation in modulating MDA5's activity and its interaction with irAlu repeats as the effects of MDA5 phosphorylation in this context are still unclear.

Before proceeding with the analysis of MDA5 bound to irAlu, future research should focus on clarifying how phosphorylation affects the binding of MDA5 to irAlu repeats. This might involve employing phosphomimetic mutants to identify the specific phosphorylation sites that may modulate this interaction.

2. Revising the optimal MDA5 construct in studying its interaction with irAlu repeats.

Initially, the plan was to work with a truncated version of the MDA5 protein that lacked the CARD domains, as previous studies had indicated that these domains were not essential for MDA5's binding to RNA. This was further supported by the observed similar filament morphology between the truncated construct and the full-length MDA5 protein bound to dsRNA (Wu et al., 2013b).

Recent evidence suggests that the CARD domains are essential for MDA5's ability to translocate along long dsRNA molecules. This ATP-hydrolysis-driven translocation facilitates CARD-CARD interactions between MDA5 proteins, leading to the recruitment of multiple motors near the dsRNA (Han et al., 2024). However, the specific role of the CARD domains in MDA5's interaction with irAlu repeats remains unclear. While the current studies emphasize structural analysis, it is important to acknowledge that the functionality of the CARD domains may significantly influence MDA5's binding properties and cooperativity. These aspects will be further investigated in subsequent experiments by measuring binding affinities.

Given this new information, it is crucial to reevaluate and refine the selection of the MDA5 construct for this study. A comprehensive comparison between the truncated version and the full-length MDA5 protein is necessary. This should begin with examining their binding interactions using EMSA and analyzing filament morphology through negative staining techniques.

Moreover, possible stronger binding interactions between MDA5 and irAlu could also impact the sample preparation for cryo-EM analysis. Tighter, more stable complexes would likely result in less heterogeneity in the sample, potentially leading to higher-quality cryo-EM data and a better structural model.

3. Investigating the effect of LGP2 on MDA5 binding to irAlu

Previous studies have demonstrated that LGP2 can assist MDA5 in binding to shorter dsRNA molecules (Bruns et al., 2014). Considering that irAlus in cells have an average length of approximately 240 bp (Li et al., 2022), LGP2 may enhance MDA5's capacity to bind to these shorter targets.

Moreover, recent reports have further elucidated the regulatory role of LGP2 in the MDA5-dsRNA interaction. It has been demonstrated that LGP2 binding at internal sites on dsRNA promotes the nucleation of MDA5 filament assembly, leading to the formation of shorter filaments (Singh et al., 2024). Additional evidence suggests that LGP2 functions as a "switch," modulating the translocation activity of MDA5 along dsRNA (Han et al., 2024).

Although the original intention was to study the interaction of MDA5 in isolation, this approach may need to be revised. Exploring the interactions of irAlu-MDA5 filaments in the presence of LGP2 could provide valuable insights into the dynamics of this complex system.

References

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D., Jumper, J.M., 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- Abudayyeh, O.O., Gootenberg, J.S., Franklin, B., Koob, J., Kellner, M.J., Ladha, A., Joung, J., Kirchgatterer, P., Cox, D.B.T., Zhang, F., 2019. A cytosine deaminase for programmable single-base RNA editing. *Science* 365, 382–386. <https://doi.org/10.1126/science.aax7063>
- Ahmad, S., Mu, X., Yang, F., Greenwald, E., Park, J.W., Jacob, E., Zhang, C.-Z., Hur, S., 2018. Breaching Self-Tolerance to Alu Duplex RNA Underlies MDA5-Mediated Inflammation. *Cell* 172, 797–810.e13. <https://doi.org/10.1016/j.cell.2017.12.016>
- Ahmed, F., Do, N., Vanderver, A.L., Treat, J.R., 2024. Dyschromatosis symmetrica hereditaria: A clue to early diagnosis of Aicardi–Goutières syndrome. *Pediatr. Dermatol.* 41, 156–157. <https://doi.org/10.1111/pde.15437>
- Athanasiadis, A., Placido, D., Maas, S., Brown, B.A., Lowenhaupt, K., Rich, A., 2005. The Crystal Structure of the Z β Domain of the RNA-editing Enzyme ADAR1 Reveals Distinct Conserved Surfaces Among Z-domains. *J. Mol. Biol.* 351, 496–507. <https://doi.org/10.1016/j.jmb.2005.06.028>
- Athanasiadis, A., Rich, A., Maas, S., 2004. Widespread A-to-I RNA Editing of Alu-Containing mRNAs in the Human Transcriptome. *PLoS Biol.* 2, e391. <https://doi.org/10.1371/journal.pbio.0020391>
- Bajad, P., Ebner, F., Amman, F., Szabó, B., Kapoor, U., Manjali, G., Hildebrandt, A., Janisiw, M.P., Jantsch, M.F., 2020. An internal deletion of ADAR rescued by MAVS deficiency leads to a minute phenotype. *Nucleic Acids Res.* 48, 3286–3303. <https://doi.org/10.1093/nar/gkaa025>
- Bartas, M., Slychko, K., Brázda, V., Červeň, J., Beaudoin, C.A., Blundell, T.L., Pečinka, P., 2022. Searching for New Z-DNA/Z-RNA Binding Proteins Based on Structural Similarity to Experimentally Validated Z α Domain. *Int. J. Mol. Sci.* 23, 768. <https://doi.org/10.3390/ijms23020768>
- Bass, B., 1987. A developmentally regulated activity that unwinds RNA duplexes. *Cell* 48, 607–613. [https://doi.org/10.1016/0092-8674\(87\)90239-X](https://doi.org/10.1016/0092-8674(87)90239-X)
- Bass, B.L., 2002. RNA Editing by Adenosine Deaminases That Act on RNA. *Annu. Rev. Biochem.* 71, 817–846. <https://doi.org/10.1146/annurev.biochem.71.110601.135501>
- Bass, B.L., Nishikura, K., Keller, W., Seeburg, P.H., Emeson, R.B., O'Connell, M.A., Samuel, C.E., Herbert, A., 1997. A standardized nomenclature for adenosine deaminases that act on RNA. *RNA N. Y.* 3, 947–949.
- Bass, B.L., Weintraub, H., 1988. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* 55, 1089–1098. [https://doi.org/10.1016/0092-8674\(88\)90253-X](https://doi.org/10.1016/0092-8674(88)90253-X)
- Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., Isaacs, F.J., Rechavi, G., Li, J.B., Eisenberg, E., Levanon, E.Y., 2014. A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.* 24, 365–376. <https://doi.org/10.1101/gr.164749.113>

- Berke, I.C., Modis, Y., 2012. MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA: MDA5 forms dimers and filaments on binding dsRNA. *EMBO J.* 31, 1714–1726. <https://doi.org/10.1038/emboj.2012.19>
- Berke, I.C., Yu, X., Modis, Y., Egelman, E.H., 2012. MDA5 assembles into a polar helical filament on dsRNA. *Proc. Natl. Acad. Sci.* 109, 18437–18441. <https://doi.org/10.1073/pnas.1212186109>
- Boccaletto, P., Stefaniak, F., Ray, A., Cappannini, A., Mukherjee, S., Purta, E., Kurkowska, M., Shirvanizadeh, N., Destefanis, E., Groza, P., Avşar, G., Romitelli, A., Pir, P., Dassi, E., Conticello, S.G., Aguilo, F., Bujnicki, J.M., 2022. MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* 50, D231–D235. <https://doi.org/10.1093/nar/gkab1083>
- Boo, S.H., Kim, Y.K., 2020. The emerging role of RNA modifications in the regulation of mRNA stability. *Exp. Mol. Med.* 52, 400–408. <https://doi.org/10.1038/s12276-020-0407-z>
- Bortoletto, E., Rosani, U., 2024. Bioinformatics for Inosine: Tools and Approaches to Trace This Elusive RNA Modification. *Genes* 15, 996. <https://doi.org/10.3390/genes15080996>
- Bruns, A.M., Leser, G.P., Lamb, R.A., Horvath, C.M., 2014. The Innate Immune Sensor LGP2 Activates Antiviral Signaling by Regulating MDA5-RNA Interaction and Filament Assembly. *Mol. Cell* 55, 771–781. <https://doi.org/10.1016/j.molcel.2014.07.003>
- Brusa, R., Zimmermann, F., Koh, D.-S., Feldmeyer, D., Gass, P., Seeburg, P.H., Sprengel, R., 1995. Early-Onset Epilepsy and Postnatal Lethality Associated with an Editing-Deficient *GluR-B* Allele in Mice. *Science* 270, 1677–1680. <https://doi.org/10.1126/science.270.5242.1677>
- Carpenter, S., O’Neill, L.A.J., 2024. From periphery to center stage: 50 years of advancements in innate immunity. *Cell* 187, 2030–2051. <https://doi.org/10.1016/j.cell.2024.03.036>
- Cenci, C., Barzotti, R., Galeano, F., Corbelli, S., Rota, R., Massimi, L., Di Rocco, C., O’Connell, M.A., Gallo, A., 2008. Down-regulation of RNA Editing in Pediatric Astrocytomas. *J. Biol. Chem.* 283, 7251–7260. <https://doi.org/10.1074/jbc.M708316200>
- Chen, C.-X., Cho, D.-S.C., Wang, Q., Lai, F., Carter, K.C., Nishikura, K., 2000. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* 6, 755–767. <https://doi.org/10.1017/S1355838200000170>
- Chen, L., Ou, L., Jing, X., Kong, Y., Xie, B., Zhang, N., Shi, H., Qin, H., Li, X., Hao, P., 2023. DeepEdit: single-molecule detection and phasing of A-to-I RNA editing events using nanopore direct RNA sequencing. *Genome Biol.* 24, 75. <https://doi.org/10.1186/s13059-023-02921-0>
- Chilibeck, K.A., Wu, T., Liang, C., Schellenberg, M.J., Gesner, E.M., Lynch, J.M., MacMillan, A.M., 2006. FRET Analysis of in Vivo Dimerization by RNA-editing Enzymes. *J. Biol. Chem.* 281, 16530–16535. <https://doi.org/10.1074/jbc.M511831200>
- Cho, D.-S.C., Yang, W., Lee, J.T., Shiekhattar, R., Murray, J.M., Nishikura, K., 2003. Requirement of Dimerization for RNA Editing Activity of Adenosine Deaminases Acting on RNA. *J. Biol. Chem.* 278, 17093–17102. <https://doi.org/10.1074/jbc.M213127200>
- Chung, H., Calis, J.J.A., Wu, X., Sun, T., Yu, Y., Sarbanes, S.L., Dao Thi, V.L., Shilvock, A.R., Hoffmann, H.-H., Rosenberg, B.R., Rice, C.M., 2018. Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. *Cell* 172, 811–824.e14. <https://doi.org/10.1016/j.cell.2017.12.038>
- Costa, T.R.D., Ignatiou, A., Orlova, E.V., 2017. Structural Analysis of Protein Complexes by Cryo Electron Microscopy, in: Journet, L., Cascales, E. (Eds.), *Bacterial Protein Secretion Systems, Methods in Molecular Biology*. Springer New York, New York, NY, pp. 377–413. https://doi.org/10.1007/978-1-4939-7033-9_28

- Dawson, T.R., Sansam, C.L., Emeson, R.B., 2004. Structure and Sequence Determinants Required for the RNA Editing of ADAR2 Substrates. *J. Biol. Chem.* 279, 4941–4951. <https://doi.org/10.1074/jbc.M310068200>
- De Reuver, R., Maelfait, J., 2024. Novel insights into double-stranded RNA-mediated immunopathology. *Nat. Rev. Immunol.* 24, 235–249. <https://doi.org/10.1038/s41577-023-00940-3>
- De Reuver, R., Verdonck, S., Dierick, E., Nemegeer, J., Hessmann, E., Ahmad, S., Jans, M., Blancke, G., Van Nieuwerburgh, F., Botzki, A., Vereecke, L., Van Loo, G., Declercq, W., Hur, S., Vandenabeele, P., Maelfait, J., 2022. ADAR1 prevents autoinflammation by suppressing spontaneous ZBP1 activation. *Nature* 607, 784–789. <https://doi.org/10.1038/s41586-022-04974-w>
- Deininger, P., 2011. Alu elements: know the SINEs. *Genome Biol.* 12, 236. <https://doi.org/10.1186/gb-2011-12-12-236>
- Del Toro Duany, Y., Wu, B., Hur, S., 2015. MDA5—filament, dynamics and disease. *Curr. Opin. Virol.* 12, 20–25. <https://doi.org/10.1016/j.coviro.2015.01.011>
- Delaunay, S., Helm, M., Frye, M., 2024. RNA modifications in physiology and disease: towards clinical applications. *Nat. Rev. Genet.* 25, 104–122. <https://doi.org/10.1038/s41576-023-00645-2>
- Dhir, A., Dhir, S., Borowski, L.S., Jimenez, L., Teitell, M., Rötig, A., Crow, Y.J., Rice, G.I., Duffy, D., Tamby, C., Nojima, T., Munnich, A., Schiff, M., De Almeida, C.R., Rehwinkel, J., Dziembowski, A., Szczesny, R.J., Proudfoot, N.J., 2018. Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature* 560, 238–242. <https://doi.org/10.1038/s41586-018-0363-0>
- Doherty, E.E., Karki, A., Wilcox, X.E., Mendoza, H.G., Manjunath, A., Matos, V.J., Fisher, A.J., Beal, P.A., 2022. ADAR activation by inducing a *syn* conformation at guanosine adjacent to an editing site. *Nucleic Acids Res.* 50, 10857–10868. <https://doi.org/10.1093/nar/gkac897>
- 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., 2021. Rational Design of RNA Editing Guide Strands: Cytidine Analogs at the Orphan Position. *J. Am. Chem. Soc.* 143, 6865–6876. <https://doi.org/10.1021/jacs.0c13319>
- Durbin, A.F., Wang, C., Marcotrigiano, J., Gehrke, L., 2016. RNAs Containing Modified Nucleotides Fail To Trigger RIG-I Conformational Changes for Innate Immune Signaling. *mBio* 7, e00833-16. <https://doi.org/10.1128/mBio.00833-16>
- Eggington, J.M., Greene, T., Bass, B.L., 2011. Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* 2, 319. <https://doi.org/10.1038/ncomms1324>
- Eifler, T., Pokharel, S., Beal, P.A., 2013. RNA-Seq Analysis Identifies a Novel Set of Editing Substrates for Human ADAR2 Present in *Saccharomyces cerevisiae*. *Biochemistry* 52, 7857–7869. <https://doi.org/10.1021/bi4006539>
- Fisher, A.J., Beal, P.A., 2017. Effects of Aicardi-Goutières syndrome mutations predicted from ADAR-RNA structures. *RNA Biol.* 14, 164–170. <https://doi.org/10.1080/15476286.2016.1267097>
- Gabay, O., Shoshan, Y., Kopel, E., Ben-Zvi, U., Mann, T.D., Bressler, N., Cohen-Fultheim, R., Schaffer, A.A., Roth, S.H., Tzur, Z., Levanon, E.Y., Eisenberg, E., 2022. Landscape of adenosine-to-inosine RNA recoding across human tissues. *Nat. Commun.* 13, 1184. <https://doi.org/10.1038/s41467-022-28841-4>
- Gallo, A., 2003. An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. *EMBO J.* 22, 3421–3430. <https://doi.org/10.1093/emboj/cdg327>
- Gao, Y., Vasic, R., Song, Y., Teng, R., Liu, C., Gbyli, R., Biancon, G., Nelakanti, R., Lobben, K., Kudo, E., Liu, W., Ardasheva, A., Fu, X., Wang, X., Joshi, P., Lee, V., Dura, B., Viero, G., Iwasaki, A., Fan, R., Xiao, A., Flavell, R.A., Li, H.-B., Tebaldi, T., Halene, S., 2020. m6A Modification Prevents Formation of

- Endogenous Double-Stranded RNAs and Deleterious Innate Immune Responses during Hematopoietic Development. *Immunity* 52, 1007-1021.e8. <https://doi.org/10.1016/j.immuni.2020.05.003>
- George, C.X., Samuel, C.E., 1999. Human RNA-specific adenosine deaminase *ADAR1* transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. *Proc. Natl. Acad. Sci.* 96, 4621–4626. <https://doi.org/10.1073/pnas.96.8.4621>
- George, C.X., Wagner, M.V., Samuel, C.E., 2005. Expression of Interferon-inducible RNA Adenosine Deaminase *ADAR1* during Pathogen Infection and Mouse Embryo Development Involves Tissue-selective Promoter Utilization and Alternative Splicing. *J. Biol. Chem.* 280, 15020–15028. <https://doi.org/10.1074/jbc.M500476200>
- Gerber, A., O'Connell, M.A., Keller, W., 1997. Two forms of human double-stranded RNA-specific editase 1 (hRED1) generated by the insertion of an Alu cassette. *RNA N. Y. N* 3, 453–463.
- Gong, C., Maquat, L.E., 2011. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 470, 284–288. <https://doi.org/10.1038/nature09701>
- Grice, L.F., Degnan, B.M., 2015. The origin of the ADAR gene family and animal RNA editing. *BMC Evol. Biol.* 15, 4. <https://doi.org/10.1186/s12862-015-0279-3>
- Grover, D., Mukerji, M., Bhatnagar, P., Kannan, K., Brahmachari, S.K., 2004. Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. *Bioinformatics* 20, 813–817. <https://doi.org/10.1093/bioinformatics/bth005>
- GTEX Consortium, Tan, M.H., Li, Q., Shanmugam, R., Piskol, R., Kohler, J., Young, A.N., Liu, K.I., Zhang, R., Ramaswami, G., Ariyoshi, K., Gupte, A., Keegan, L.P., George, C.X., Ramu, A., Huang, N., Pollina, E.A., Leeman, D.S., Rustighi, A., Goh, Y.P.S., Chawla, A., Del Sal, G., Peltz, G., Brunet, A., Conrad, D.F., Samuel, C.E., O'Connell, M.A., Walkley, C.R., Nishikura, K., Li, J.B., 2017. Dynamic landscape and regulation of RNA editing in mammals. *Nature* 550, 249–254. <https://doi.org/10.1038/nature24041>
- Guo, X., Liu, S., Sheng, Y., Zenati, M., Billiar, T., Herbert, A., Wang, Q., 2023. *ADAR1* Z α domain P195A mutation activates the MDA5-dependent RNA-sensing signaling pathway in brain without decreasing overall RNA editing. *Cell Rep.* 42, 112733. <https://doi.org/10.1016/j.celrep.2023.112733>
- Ha, S.C., Choi, J., Hwang, H.-Y., Rich, A., Kim, Y.-G., Kim, K.K., 2009. The structures of non-CG-repeat Z-DNAs co-crystallized with the Z-DNA-binding domain, hZ α *ADAR1*. *Nucleic Acids Res.* 37, 629–637. <https://doi.org/10.1093/nar/gkn976>
- Ha, S.C., Lowenhaupt, K., Rich, A., Kim, Y.-G., Kim, K.K., 2005. Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature* 437, 1183–1186. <https://doi.org/10.1038/nature04088>
- Han, X.-P., Rao, M., Wu, S.-R., Zhang, Q., Hou, F., Zhang, S.-Q., Chen, L.-L., Liu, J., 2024. Observation of Higher-Order Assemblies Controlled by Protein One-dimensional Movement. <https://doi.org/10.1101/2024.06.06.597732>
- Hartner, J.C., Schmittwolf, C., Kispert, A., Müller, A.M., Higuchi, M., Seeburg, P.H., 2004. Liver Disintegration in the Mouse Embryo Caused by Deficiency in the RNA-editing Enzyme *ADAR1*. *J. Biol. Chem.* 279, 4894–4902. <https://doi.org/10.1074/jbc.M311347200>
- Haudenschild, B.L., Maydanovych, O., Véliz, E.A., Macbeth, M.R., Bass, B.L., Beal, P.A., 2004. A Transition State Analogue for an RNA-Editing Reaction. *J. Am. Chem. Soc.* 126, 11213–11219. <https://doi.org/10.1021/ja0472073>

- Heraud-Farlow, J.E., Chalk, A.M., Linder, S.E., Li, Q., Taylor, S., White, J.M., Pang, L., Liddicoat, B.J., Gupte, A., Li, J.B., Walkley, C.R., 2017. Protein recoding by ADAR1-mediated RNA editing is not essential for normal development and homeostasis. *Genome Biol.* 18, 166. <https://doi.org/10.1186/s13059-017-1301-4>
- Herbert, A., Alfken, J., Kim, Y.-G., Mian, I.S., Nishikura, K., Rich, A., 1997. A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc. Natl. Acad. Sci.* 94, 8421–8426. <https://doi.org/10.1073/pnas.94.16.8421>
- Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., Seeburg, P.H., 2000. Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* 406, 78–81. <https://doi.org/10.1038/35017558>
- Higuchi, M., Single, F.N., Köhler, M., Sommer, B., Sprengel, R., Seeburg, P.H., 1993. RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency. *Cell* 75, 1361–1370. [https://doi.org/10.1016/0092-8674\(93\)90622-W](https://doi.org/10.1016/0092-8674(93)90622-W)
- Holehouse, A.S., Kragelund, B.B., 2024. The molecular basis for cellular function of intrinsically disordered protein regions. *Nat. Rev. Mol. Cell Biol.* 25, 187–211. <https://doi.org/10.1038/s41580-023-00673-0>
- Hou, F., Sun, L., Zheng, H., Skaug, B., Jiang, Q.-X., Chen, Z.J., 2011. MAVS Forms Functional Prion-like Aggregates to Activate and Propagate Antiviral Innate Immune Response. *Cell* 146, 448–461. <https://doi.org/10.1016/j.cell.2011.06.041>
- Hu, S.-B., Heraud-Farlow, J., Sun, T., Liang, Z., Goradia, A., Taylor, S., Walkley, C.R., Li, J.B., 2023. ADAR1p150 prevents MDA5 and PKR activation via distinct mechanisms to avert fatal autoinflammation. *Mol. Cell* 83, 3869–3884.e7. <https://doi.org/10.1016/j.molcel.2023.09.018>
- Hu, S.-B., Li, J.B., 2024. RNA editing and immune control: from mechanism to therapy. *Curr. Opin. Genet. Dev.* 86, 102195. <https://doi.org/10.1016/j.gde.2024.102195>
- Hubbard, N.W., Ames, J.M., Maurano, M., Chu, L.H., Somfleth, K.Y., Gokhale, N.S., Werner, M., Snyder, J.M., Lichauro, K., Savan, R., Stetson, D.B., Oberst, A., 2022. ADAR1 mutation causes ZBP1-dependent immunopathology. *Nature* 607, 769–775. <https://doi.org/10.1038/s41586-022-04896-7>
- Hur, S., 2019. Double-Stranded RNA Sensors and Modulators in Innate Immunity. *Annu. Rev. Immunol.* 37, 349–375. <https://doi.org/10.1146/annurev-immunol-042718-041356>
- 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., 2019. Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. *Nature* 565, 43–48. <https://doi.org/10.1038/s41586-018-0768-9>
- Janeway, C.A., 1989. Approaching the Asymptote? Evolution and Revolution in Immunology. *Cold Spring Harb. Symp. Quant. Biol.* 54, 1–13. <https://doi.org/10.1101/SQB.1989.054.01.003>
- Janeway, C.A., Medzhitov, R., 2002. Innate Immune Recognition. *Annu. Rev. Immunol.* 20, 197–216. <https://doi.org/10.1146/annurev.immunol.20.083001.084359>
- Jiao, H., Wachsmuth, L., Wolf, S., Lohmann, J., Nagata, M., Kaya, G.G., Oikonomou, N., Kondylis, V., Rogg, M., Diebold, M., Tröder, S.E., Zevnik, B., Prinz, M., Schell, C., Young, G.R., Kassiotis, G., Pasparakis, M., 2022. ADAR1 averts fatal type I interferon induction by ZBP1. *Nature* 607, 776–783. <https://doi.org/10.1038/s41586-022-04878-9>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-

- Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstern, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kallman, A.M., 2003. ADAR2 A->I editing: site selectivity and editing efficiency are separate events. *Nucleic Acids Res.* 31, 4874–4881. <https://doi.org/10.1093/nar/gkg681>
- Kang, D., Gopalkrishnan, R.V., Lin, L., Randolph, A., Valerie, K., Pestka, S., Fisher, P.B., 2004. Expression analysis and genomic characterization of human melanoma differentiation associated gene-5, mda-5: a novel type I interferon-responsive apoptosis-inducing gene. *Oncogene* 23, 1789–1800. <https://doi.org/10.1038/sj.onc.1207300>
- Karikó, K., Buckstein, M., Ni, H., Weissman, D., 2005. Suppression of RNA Recognition by Toll-like Receptors: The Impact of Nucleoside Modification and the Evolutionary Origin of RNA. *Immunity* 23, 165–175. <https://doi.org/10.1016/j.immuni.2005.06.008>
- Karki, A., Campbell, K.B., Mozumder, S., Fisher, A.J., Beal, P.A., 2024. Impact of Disease-Associated Mutations on the Deaminase Activity of ADAR1. *Biochemistry* 63, 282–293. <https://doi.org/10.1021/acs.biochem.3c00405>
- Karki, R., Kanneganti, T.-D., 2023. ADAR1 and ZBP1 in innate immunity, cell death, and disease. *Trends Immunol.* 44, 201–216. <https://doi.org/10.1016/j.it.2023.01.001>
- Karki, R., Sundaram, B., Sharma, B.R., Lee, S., Malireddi, R.K.S., Nguyen, L.N., Christgen, S., Zheng, M., Wang, Y., Samir, P., Neale, G., Vogel, P., Kanneganti, T.-D., 2021. ADAR1 restricts ZBP1-mediated immune response and PANoptosis to promote tumorigenesis. *Cell Rep.* 37, 109858. <https://doi.org/10.1016/j.celrep.2021.109858>
- Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., Hiiragi, A., Dermody, T.S., Fujita, T., Akira, S., 2008. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *J. Exp. Med.* 205, 1601–1610. <https://doi.org/10.1084/jem.20080091>
- Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., Uematsu, S., Jung, A., Kawai, T., Ishii, K.J., Yamaguchi, O., Otsu, K., Tsujimura, T., Koh, C.-S., Reis E Sousa, C., Matsuura, Y., Fujita, T., Akira, S., 2006. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 441, 101–105. <https://doi.org/10.1038/nature04734>
- Katrekar, D., Xiang, Y., Palmer, N., Saha, A., Meluzzi, D., Mali, P., 2022. Comprehensive interrogation of the ADAR2 deaminase domain for engineering enhanced RNA editing activity and specificity. *eLife* 11, e75555. <https://doi.org/10.7554/eLife.75555>
- Kawai, T., Takahashi, K., Sato, S., Coban, C., Kumar, H., Kato, H., Ishii, K.J., Takeuchi, O., Akira, S., 2005. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat. Immunol.* 6, 981–988. <https://doi.org/10.1038/ni1243>
- Keegan, L.P., Hajji, K., O’Connell, M.A., 2023. Adenosine Deaminase Acting on RNA (ADAR) Enzymes: A Journey from Weird to Wondrous. *Acc. Chem. Res.* 56, 3165–3174. <https://doi.org/10.1021/acs.accounts.3c00433>
- Keegan, L.P., Leroy, A., Sproul, D., O’Connell, M.A., 2004. Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. *Genome Biol.* 5, 209. <https://doi.org/10.1186/gb-2004-5-2-209>
- Keegan, L.P., Rosenthal, J.J., Roberson, L.M., O’Connell, M.A., 2007. Purification and Assay of ADAR Activity, in: *Methods in Enzymology*. Elsevier, pp. 301–317. [https://doi.org/10.1016/S0076-6879\(07\)24014-5](https://doi.org/10.1016/S0076-6879(07)24014-5)

- Khan, S., Godfrey, V., Zaki, Md.H., 2019. Cytosolic Nucleic Acid Sensors in Inflammatory and Autoimmune Disorders, in: *International Review of Cell and Molecular Biology*. Elsevier, pp. 215–253. <https://doi.org/10.1016/bs.ircmb.2018.10.002>
- Kim, D.D.Y., Kim, T.T.Y., Walsh, T., Kobayashi, Y., Matise, T.C., Buyske, S., Gabriel, A., 2004. Widespread RNA Editing of Embedded *Alu* Elements in the Human Transcriptome. *Genome Res.* 14, 1719–1725. <https://doi.org/10.1101/gr.2855504>
- Kim, Y.-G., Muralinath, M., Brandt, T., Percy, M., Hauns, K., Lowenhaupt, K., Jacobs, B.L., Rich, A., 2003. A role for Z-DNA binding in vaccinia virus pathogenesis. *Proc. Natl. Acad. Sci.* 100, 6974–6979. <https://doi.org/10.1073/pnas.0431131100>
- King, W.L., Jenkins, C., Seymour, J.R., Labbate, M., 2019. Oyster disease in a changing environment: Decrypting the link between pathogen, microbiome and environment. *Mar. Environ. Res.* 143, 124–140. <https://doi.org/10.1016/j.marenvres.2018.11.007>
- Kleinova, R., Rajendra, V., Leuchtenberger, A.F., Lo Giudice, C., Vesely, C., Kapoor, U., Tanzer, A., Derdak, S., Picardi, E., Jantsch, M.F., 2023. The ADAR1 editome reveals drivers of editing-specificity for ADAR1-isoforms. *Nucleic Acids Res.* 51, 4191–4207. <https://doi.org/10.1093/nar/gkad265>
- Kuttan, A., Bass, B.L., 2012. Mechanistic insights into editing-site specificity of ADARs. *Proc. Natl. Acad. Sci.* 109. <https://doi.org/10.1073/pnas.1212548109>
- Lawson, K.A., Sousa, C.M., Zhang, X., Kim, E., Akthar, R., Caumanns, J.J., Yao, Y., Mikolajewicz, N., Ross, C., Brown, K.R., Zid, A.A., Fan, Z.P., Hui, S., Krall, J.A., Simons, D.M., Slater, C.J., De Jesus, V., Tang, L., Singh, R., Goldford, J.E., Martin, S., Huang, Q., Francis, E.A., Habsid, A., Climie, R., Tieu, D., Wei, J., Li, R., Tong, A.H.Y., Aregger, M., Chan, K.S., Han, H., Wang, X., Mero, P., Brumell, J.H., Finelli, A., Ailles, L., Bader, G., Smolen, G.A., Kingsbury, G.A., Hart, T., Kung, C., Moffat, J., 2020. Functional genomic landscape of cancer-intrinsic evasion of killing by T cells. *Nature* 586, 120–126. <https://doi.org/10.1038/s41586-020-2746-2>
- Lehmann, K.A., Bass, B.L., 2000. Double-Stranded RNA Adenosine Deaminases ADAR1 and ADAR2 Have Overlapping Specificities. *Biochemistry* 39, 12875–12884. <https://doi.org/10.1021/bi001383g>
- Lehmann, K.A., Bass, B.L., 1999. The importance of internal loops within RNA substrates of ADAR1. *J. Mol. Biol.* 291, 1–13. <https://doi.org/10.1006/jmbi.1999.2914>
- Li, Q., Gloudemans, M.J., Geisinger, J.M., Fan, B., Aguet, F., Sun, T., Ramaswami, G., Li, Y.I., Ma, J.-B., Pritchard, J.K., Montgomery, S.B., Li, J.B., 2022. RNA editing underlies genetic risk of common inflammatory diseases. *Nature* 608, 569–577. <https://doi.org/10.1038/s41586-022-05052-x>
- Liang, Z., Chalk, A.M., Taylor, S., Goradia, A., Heraud-Farlow, J.E., Walkley, C.R., 2023. The phenotype of the most common human ADAR1P150 α mutation P193A in mice is partially penetrant. *EMBO Rep.* 24, e55835. <https://doi.org/10.15252/embr.202255835>
- LiCata, V.J., Wowor, A.J., 2008. Applications of Fluorescence Anisotropy to the Study of Protein–DNA Interactions, in: *Methods in Cell Biology*. Elsevier, pp. 243–262. [https://doi.org/10.1016/S0091-679X\(07\)84009-X](https://doi.org/10.1016/S0091-679X(07)84009-X)
- Liddicoat, B.J., Piskol, R., Chalk, A.M., Ramaswami, G., Higuchi, M., Hartner, J.C., Li, J.B., Seeburg, P.H., Walkley, C.R., 2015. RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* 349, 1115–1120. <https://doi.org/10.1126/science.aac7049>
- Lind, N.A., Rael, V.E., Pestal, K., Liu, B., Barton, G.M., 2022. Regulation of the nucleic acid-sensing Toll-like receptors. *Nat. Rev. Immunol.* 22, 224–235. <https://doi.org/10.1038/s41577-021-00577-0>

- Liu, A., Wang, X., 2022. The Pivotal Role of Chemical Modifications in mRNA Therapeutics. *Front. Cell Dev. Biol.* 10, 901510. <https://doi.org/10.3389/fcell.2022.901510>
- Liu, A., Ying, S., 2023. AICARDI–GOUTIÈRES syndrome: A monogenic type I interferonopathy. *Scand. J. Immunol.* 98, e13314. <https://doi.org/10.1111/sji.13314>
- Liu, H., Golji, J., Brodeur, L.K., Chung, F.S., Chen, J.T., deBeaumont, R.S., Bullock, C.P., Jones, M.D., Kerr, G., Li, L., Rakiec, D.P., Schlabach, M.R., Sovath, S., Growney, J.D., Pagliarini, R.A., Ruddy, D.A., MacIsaac, K.D., Korn, J.M., McDonald, E.R., 2019. Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. *Nat. Med.* 25, 95–102. <https://doi.org/10.1038/s41591-018-0302-5>
- Liu, X., Sun, T., Shcherbina, A., Li, Q., Jarmoskaite, I., Kappel, K., Ramaswami, G., Das, R., Kundaje, A., Li, J.B., 2021. Learning cis-regulatory principles of ADAR-based RNA editing from CRISPR-mediated mutagenesis. *Nat. Commun.* 12, 2165. <https://doi.org/10.1038/s41467-021-22489-2>
- Lomeli, H., Mosbacher, J., Melcher, T., Höger, T., Geiger, J.R.P., Kuner, T., Monyer, H., Higuchi, M., Bach, A., Seeburg, P.H., 1994. Control of Kinetic Properties of AMPA Receptor Channels by Nuclear RNA Editing. *Science* 266, 1709–1713. <https://doi.org/10.1126/science.7992055>
- Lu, M., Zhang, Z., Xue, M., Zhao, B.S., Harder, O., Li, A., Liang, X., Gao, T.Z., Xu, Y., Zhou, J., Feng, Z., Niewiesk, S., Peeples, M.E., He, C., Li, J., 2020. N6-methyladenosine modification enables viral RNA to escape recognition by RNA sensor RIG-I. *Nat. Microbiol.* 5, 584–598. <https://doi.org/10.1038/s41564-019-0653-9>
- Lv, T., Jiang, S., Wang, X., Hou, Y., 2023. Profiling A-to-I RNA editing during mouse somatic reprogramming at the single-cell level. *Heliyon* 9, e18133. <https://doi.org/10.1016/j.heliyon.2023.e18133>
- Ma, Q., Che, L., Chen, Y., Gu, Z., 2023. Identification of five novel variants of ADAR1 in dyschromatosis symmetrica hereditaria by next-generation sequencing. *Front. Pediatr.* 11, 1161502. <https://doi.org/10.3389/fped.2023.1161502>
- Macbeth, M.R., Bass, B.L., 2007. Large-Scale Overexpression and Purification of ADARs from *Saccharomyces cerevisiae* for Biophysical and Biochemical Studies, in: *Methods in Enzymology*. Elsevier, pp. 319–331. [https://doi.org/10.1016/S0076-6879\(07\)24015-7](https://doi.org/10.1016/S0076-6879(07)24015-7)
- Macbeth, M.R., Schubert, H.L., VanDemark, A.P., Lingam, A.T., Hill, C.P., Bass, B.L., 2005. Inositol Hexakisphosphate Is Bound in the ADAR2 Core and Required for RNA Editing. *Science* 309, 1534–1539. <https://doi.org/10.1126/science.1113150>
- Malik, T.N., Doherty, E.E., Gaded, V.M., Hill, T.M., Beal, P.A., Emeson, R.B., 2021. Regulation of RNA editing by intracellular acidification. *Nucleic Acids Res.* 49, 4020–4036. <https://doi.org/10.1093/nar/gkab157>
- Mandas, D., Salati, F., 2017. Ostreid herpesvirus: A pathogen of oysters. *Viol. Res. Rev.* 1. <https://doi.org/10.15761/VRR.1000112>
- Mannion, N.M., Greenwood, S.M., Young, R., Cox, S., Brindle, J., Read, D., Nellåker, C., Vesely, C., Ponting, C.P., McLaughlin, P.J., Jantsch, M.F., Dorin, J., Adams, I.R., Scadden, A.D.J., Öhman, M., Keegan, L.P., O’Connell, M.A., 2014. The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA. *Cell Rep.* 9, 1482–1494. <https://doi.org/10.1016/j.celrep.2014.10.041>
- Matthews, J.M., Zheng, Y., Tran, K., Phelps, K.J., Scott, A.I., Havel, J., Fisher, A.J., Beal, P.A., 2016. Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. *Nat. Struct. Mol. Biol.* 23, 426–433. <https://doi.org/10.1038/nsmb.3203>
- Maurano, M., Snyder, J.M., Connelly, C., Henao-Mejia, J., Sidrauski, C., Stetson, D.B., 2021. Protein kinase R and the integrated stress response drive immunopathology caused by mutations in the RNA deaminase ADAR1. *Immunity* 54, 1948–1960.e5. <https://doi.org/10.1016/j.immuni.2021.07.001>

- Mboukou, A., Rajendra, V., Messmer, S., Catala, M., Tisné, C., Jantsch, M.F., Barraud, P., 2023. Dimerization of ADAR1 modulates site-specificity of RNA editing.
- Medzhitov, R., 2009. Approaching the Asymptote: 20 Years Later. *Immunity* 30, 766–775. <https://doi.org/10.1016/j.immuni.2009.06.004>
- Miyamura, Y., Suzuki, T., Kono, M., Inagaki, K., Ito, S., Suzuki, N., Tomita, Y., 2003. Mutations of the RNA-Specific Adenosine Deaminase Gene (DSRAD) Are Involved in Dyschromatosis Symmetrica Hereditaria. *Am. J. Hum. Genet.* 73, 693–699. <https://doi.org/10.1086/378209>
- Mukherjee, K., Korithoski, B., Kolaczowski, B., 2014. Ancient Origins of Vertebrate-Specific Innate Antiviral Immunity. *Mol. Biol. Evol.* 31, 140–153. <https://doi.org/10.1093/molbev/mst184>
- Musselman, C.A., Kutateladze, T.G., 2021. Characterization of functional disordered regions within chromatin-associated proteins. *iScience* 24, 102070. <https://doi.org/10.1016/j.isci.2021.102070>
- Nguyen, T.A., Heng, J.W.J., Kaewsapsak, P., Kok, E.P.L., Stanojević, D., Liu, H., Cardilla, A., Praditya, A., Yi, Z., Lin, M., Aw, J.G.A., Ho, Y.Y., Peh, K.L.E., Wang, Y., Zhong, Q., Heraud-Farlow, J., Xue, S., Reversade, B., Walkley, C., Ho, Y.S., Šikić, M., Wan, Y., Tan, M.H., 2022. Direct identification of A-to-I editing sites with nanopore native RNA sequencing. *Nat. Methods* 19, 833–844. <https://doi.org/10.1038/s41592-022-01513-3>
- Nie, Y., Hammond, G.L., Yang, J.-H., 2007. Double-Stranded RNA Deaminase ADAR1 Increases Host Susceptibility to Virus Infection. *J. Virol.* 81, 917–923. <https://doi.org/10.1128/JVI.01527-06>
- Nishikura, K., 2016. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* 17, 83–96. <https://doi.org/10.1038/nrm.2015.4>
- O’Connell, M.A., Keller, W., 1994. Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. *Proc. Natl. Acad. Sci.* 91, 10596–10600. <https://doi.org/10.1073/pnas.91.22.10596>
- Ottoz, D.S.M., Berchowitz, L.E., 2020. The role of disorder in RNA binding affinity and specificity. *Open Biol.* 10, 200328. <https://doi.org/10.1098/rsob.200328>
- Palovcak, E., Wang, F., Zheng, S.Q., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A., Cheng, Y., 2018. A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *J. Struct. Biol.* 204, 80–84. <https://doi.org/10.1016/j.jsb.2018.07.007>
- 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., 2020. High-throughput mutagenesis reveals unique structural features of human ADAR1. *Nat. Commun.* 11, 5130. <https://doi.org/10.1038/s41467-020-18862-2>
- Peisley, A., Jo, M.H., Lin, C., Wu, B., Orme-Johnson, M., Walz, T., Hohng, S., Hur, S., 2012. Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments. *Proc. Natl. Acad. Sci.* 109. <https://doi.org/10.1073/pnas.1208618109>
- Peisley, A., Lin, C., Wu, B., Orme-Johnson, M., Liu, M., Walz, T., Hur, S., 2011. Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition. *Proc. Natl. Acad. Sci.* 108, 21010–21015. <https://doi.org/10.1073/pnas.1113651108>
- Pestal, K., Funk, C.C., Snyder, J.M., Price, N.D., Treuting, P.M., Stetson, D.B., 2015. Isoforms of RNA-Editing Enzyme ADAR1 Independently Control Nucleic Acid Sensor MDA5-Driven Autoimmunity and Multi-organ Development. *Immunity* 43, 933–944. <https://doi.org/10.1016/j.immuni.2015.11.001>

- Phelps, K.J., Tran, K., Eifler, T., Erickson, A.I., Fisher, A.J., Beal, P.A., 2015. Recognition of duplex RNA by the deaminase domain of the RNA editing enzyme ADAR2. *Nucleic Acids Res.* 43, 1123–1132. <https://doi.org/10.1093/nar/gku1345>
- Placido, D., Brown, B.A., Lowenhaupt, K., Rich, A., Athanasiadis, A., 2007. A Left-Handed RNA Double Helix Bound by the Z α Domain of the RNA-Editing Enzyme ADAR1. *Structure* 15, 395–404. <https://doi.org/10.1016/j.str.2007.03.001>
- Porath, H.T., Knisbacher, B.A., Eisenberg, E., Levanon, E.Y., 2017. Massive A-to-I RNA editing is common across the Metazoa and correlates with dsRNA abundance. *Genome Biol.* 18, 185. <https://doi.org/10.1186/s13059-017-1315-y>
- Poulsen, H., Jorgensen, R., Heding, A., Nielsen, F.C., Bonven, B., Egebjerg, J., 2006. Dimerization of ADAR2 is mediated by the double-stranded RNA binding domain. *RNA* 12, 1350–1360. <https://doi.org/10.1261/rna.2314406>
- Poulsen, H., Nilsson, J., Damgaard, C.K., Egebjerg, J., Kjems, J., 2001. CRM1 Mediates the Export of ADAR1 through a Nuclear Export Signal within the Z-DNA Binding Domain. *Mol. Cell. Biol.* 21, 7862–7871. <https://doi.org/10.1128/MCB.21.22.7862-7871.2001>
- Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290–296. <https://doi.org/10.1038/nmeth.4169>
- Raghava Kurup, R., Oakes, E.K., Manning, A.C., Mukherjee, P., Vadlamani, P., Hundley, H.A., 2022. RNA binding by ADAR3 inhibits adenosine-to-inosine editing and promotes expression of immune response protein MAVS. *J. Biol. Chem.* 298, 102267. <https://doi.org/10.1016/j.jbc.2022.102267>
- Ramaswami, G., Li, J.B., 2014. RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res.* 42, D109–D113. <https://doi.org/10.1093/nar/gkt996>
- Rebagliati, M.R., Melton, D.A., 1987. Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell* 48, 599–605. [https://doi.org/10.1016/0092-8674\(87\)90238-8](https://doi.org/10.1016/0092-8674(87)90238-8)
- Rehwinkel, J., Gack, M.U., 2020. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nat. Rev. Immunol.* 20, 537–551. <https://doi.org/10.1038/s41577-020-0288-3>
- Rehwinkel, J., Mehdipour, P., 2024. ADAR1: from basic mechanisms to inhibitors. *Trends Cell Biol.* S096289242400120X. <https://doi.org/10.1016/j.tcb.2024.06.006>
- Ren, X., Linehan, M.M., Iwasaki, A., Pyle, A.M., 2019. RIG-I Selectively Discriminates against 5'-Monophosphate RNA. *Cell Rep.* 26, 2019-2027.e4. <https://doi.org/10.1016/j.celrep.2019.01.107>
- Rice, G.I., Del Toro Duany, Y., Jenkinson, E.M., Forte, G.M.A., Anderson, B.H., Ariaudo, G., Bader-Meunier, B., Baildam, E.M., Battini, R., Beresford, M.W., Casarano, M., Chouchane, M., Cimaz, R., Collins, A.E., Cordeiro, N.J.V., Dale, R.C., Davidson, J.E., De Waele, L., Desguerre, I., Faivre, L., Fazzi, E., Isidor, B., Lagae, L., Latchman, A.R., Lebon, P., Li, C., Livingston, J.H., Lourenço, C.M., Mancardi, M.M., Masurel-Paulet, A., McInnes, I.B., Menezes, M.P., Mignot, C., O'Sullivan, J., Orcesi, S., Picco, P.P., Riva, E., Robinson, R.A., Rodriguez, D., Salvatici, E., Scott, C., Szybowska, M., Tolmie, J.L., Vanderver, A., Vanhulle, C., Vieira, J.P., Webb, K., Whitney, R.N., Williams, S.G., Wolfe, L.A., Zuberi, S.M., Hur, S., Crow, Y.J., 2014. Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling. *Nat. Genet.* 46, 503–509. <https://doi.org/10.1038/ng.2933>
- Rice, G.I., Kasher, P.R., Forte, G.M.A., Mannion, N.M., Greenwood, S.M., Szykiewicz, M., Dickerson, J.E., Bhaskar, S.S., Zampini, M., Briggs, T.A., Jenkinson, E.M., Bacino, C.A., Battini, R., Bertini, E., Brogan, P.A., Brueton, L.A., Carpanelli, M., De Laet, C., De Lonlay, P., Del Toro, M., Desguerre, I., Fazzi, E., Garcia-Cazorla, À., Heiberg, A., Kawaguchi, M., Kumar, R., Lin, J.-P.S.-M., Lourenco, C.M., Male, A.M.,

- Marques, W., Mignot, C., Olivieri, I., Orcesi, S., Prabhakar, P., Rasmussen, M., Robinson, R.A., Rozenberg, F., Schmidt, J.L., Steindl, K., Tan, T.Y., Van Der Merwe, W.G., Vanderver, A., Vassallo, G., Wakeling, E.L., Wassmer, E., Whittaker, E., Livingston, J.H., Lebon, P., Suzuki, T., McLaughlin, P.J., Keegan, L.P., O'Connell, M.A., Lovell, S.C., Crow, Y.J., 2012. Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. *Nat. Genet.* 44, 1243–1248. <https://doi.org/10.1038/ng.2414>
- Rosani, U., Bai, C.-M., Maso, L., Shapiro, M., Abbadi, M., Domeneghetti, S., Wang, C.-M., Cendron, L., MacCarthy, T., Venier, P., 2019. A-to-I editing of Malacoherpesviridae RNAs supports the antiviral role of ADAR1 in mollusks. *BMC Evol. Biol.* 19, 149. <https://doi.org/10.1186/s12862-019-1472-6>
- Rosani, U., Bortoletto, E., Montagnani, C., Venier, P., 2022. ADAR-Editing during Ostreid Herpesvirus 1 Infection in *Crassostrea gigas* : Facts and Limitations. *mSphere* 7, e00011-22. <https://doi.org/10.1128/msphere.00011-22>
- Roth, S.H., Danan-Gotthold, M., Ben-Izhak, M., Rechavi, G., Cohen, C.J., Louzoun, Y., Levanon, E.Y., 2018. Increased RNA Editing May Provide a Source for Autoantigens in Systemic Lupus Erythematosus. *Cell Rep.* 23, 50–57. <https://doi.org/10.1016/j.celrep.2018.03.036>
- Sakurai, M., Shiromoto, Y., Ota, H., Song, C., Kossenkov, A.V., Wickramasinghe, J., Showe, L.C., Skordalakes, E., Tang, H.-Y., Speicher, D.W., Nishikura, K., 2017. ADAR1 controls apoptosis of stressed cells by inhibiting Staufen1-mediated mRNA decay. *Nat. Struct. Mol. Biol.* 24, 534–543. <https://doi.org/10.1038/nsmb.3403>
- Sansam, C.L., Wells, K.S., Emeson, R.B., 2003. Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proc. Natl. Acad. Sci.* 100, 14018–14023. <https://doi.org/10.1073/pnas.2336131100>
- Satoh, T., Kato, H., Kumagai, Y., Yoneyama, M., Sato, S., Matsushita, K., Tsujimura, T., Fujita, T., Akira, S., Takeuchi, O., 2010. LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proc. Natl. Acad. Sci.* 107, 1512–1517. <https://doi.org/10.1073/pnas.0912986107>
- Schuberth-Wagner, C., Ludwig, J., Bruder, A.K., Herzner, A.-M., Zillinger, T., Goldeck, M., Schmidt, T., Schmid-Burgk, J.L., Kerber, R., Wolter, S., Stümpel, J.-P., Roth, A., Bartok, E., Drosten, C., Coch, C., Hornung, V., Barchet, W., Kümmerer, B.M., Hartmann, G., Schlee, M., 2015. A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity* 43, 41–51. <https://doi.org/10.1016/j.immuni.2015.06.015>
- Schwartz, T., Rould, M.A., Lowenhaupt, K., Herbert, A., Rich, A., 1999. Crystal Structure of the α Domain of the Human Editing Enzyme ADAR1 Bound to Left-Handed Z-DNA. *Science* 284, 1841–1845. <https://doi.org/10.1126/science.284.5421.1841>
- Shen, S., Zhang, L.-S., 2023. The regulation of antiviral innate immunity through non-m6A RNA modifications. *Front. Immunol.* 14, 1286820. <https://doi.org/10.3389/fimmu.2023.1286820>
- Singh, R., Wu, Y., Herrero Del Valle, A., Leigh, K.E., Mong, S., Cheng, M.T.K., Ferguson, B.J., Modis, Y., 2024. Contrasting functions of ATP hydrolysis by MDA5 and LGP2 in viral RNA sensing. *J. Biol. Chem.* 300, 105711. <https://doi.org/10.1016/j.jbc.2024.105711>
- Sinigaglia, K., Cherian, A., Du, Q., Lacovich, V., Vukić, D., Melicherová, J., Linhartova, P., Zerad, L., Stejskal, S., Malik, R., Prochazka, J., Bondurand, N., Sedláček, R., O'Connell, M.A., Keegan, L.P., 2024. An ADAR1 dsRBD3-PKR kinase domain interaction on dsRNA inhibits PKR activation. *Cell Rep.* 43, 114618. <https://doi.org/10.1016/j.celrep.2024.114618>
- Slavov, D., Gardiner, K., 2002. Phylogenetic comparison of the pre-mRNA adenosine deaminase ADAR2 genes and transcripts: conservation and diversity in editing site sequence and alternative splicing patterns. *Gene* 299, 83–94. [https://doi.org/10.1016/S0378-1119\(02\)01016-8](https://doi.org/10.1016/S0378-1119(02)01016-8)

- Sommer, B., Köhler, M., Sprengel, R., Seeburg, P.H., 1991. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* 67, 11–19. [https://doi.org/10.1016/0092-8674\(91\)90568-J](https://doi.org/10.1016/0092-8674(91)90568-J)
- Song, B., Shiromoto, Y., Minakuchi, M., Nishikura, K., 2022. The role of RNA editing enzyme ADAR1 in human disease. *WIREs RNA* 13, e1665. <https://doi.org/10.1002/wrna.1665>
- Song, Y., Yang, W., Fu, Q., Wu, L., Zhao, X., Zhang, Y., Zhang, R., 2020. irCLASH reveals RNA substrates recognized by human ADARs. *Nat. Struct. Mol. Biol.* 27, 351–362. <https://doi.org/10.1038/s41594-020-0398-4>
- Strehblow, A., Hallegger, M., Jantsch, M.F., 2002. Nucleocytoplasmic Distribution of Human RNA-editing Enzyme ADAR1 Is Modulated by Double-stranded RNA-binding Domains, a Leucine-rich Export Signal, and a Putative Dimerization Domain. *Mol. Biol. Cell* 13, 3822–3835. <https://doi.org/10.1091/mbc.e02-03-0161>
- Struhl, K., 2024. Intrinsically disordered regions (IDRs): A vague and confusing concept for protein function. *Mol. Cell* 84, 1186–1187. <https://doi.org/10.1016/j.molcel.2024.02.023>
- Sun, L., Wu, J., Du, F., Chen, X., Chen, Z.J., 2013. Cyclic GMP-AMP Synthase Is a Cytosolic DNA Sensor That Activates the Type I Interferon Pathway. *Science* 339, 786–791. <https://doi.org/10.1126/science.1232458>
- Sun, T., Yu, Y., Wu, X., Acevedo, A., Luo, J.-D., Wang, J., Schneider, W.M., Hurwitz, B., Rosenberg, B.R., Chung, H., Rice, C.M., 2021. Decoupling expression and editing preferences of ADAR1 p150 and p110 isoforms. *Proc. Natl. Acad. Sci.* 118, e2021757118. <https://doi.org/10.1073/pnas.2021757118>
- Tan, T.Y., Sedmík, J., Fitzgerald, M.P., Halevy, R.S., Keegan, L.P., Helbig, I., Basel-Salmon, L., Cohen, L., Straussberg, R., Chung, W.K., Helal, M., Maroofian, R., Houlden, H., Juusola, J., Sadedin, S., Pais, L., Howell, K.B., White, S.M., Christodoulou, J., O’Connell, M.A., 2020. Bi-allelic ADARB1 Variants Associated with Microcephaly, Intellectual Disability, and Seizures. *Am. J. Hum. Genet.* 106, 467–483. <https://doi.org/10.1016/j.ajhg.2020.02.015>
- Tegunov, D., Cramer, P., 2019. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* 16, 1146–1152. <https://doi.org/10.1038/s41592-019-0580-y>
- Thuy-Boun, A.S., Thomas, J.M., Grajo, H.L., Palumbo, C.M., Park, S., Nguyen, L.T., Fisher, A.J., Beal, P.A., 2020. Asymmetric dimerization of adenosine deaminase acting on RNA facilitates substrate recognition. *Nucleic Acids Res.* 48, 7958–7972. <https://doi.org/10.1093/nar/gkaa532>
- Unterholzner, L., Keating, S.E., Baran, M., Horan, K.A., Jensen, S.B., Sharma, S., Sirois, C.M., Jin, T., Latz, E., Xiao, T.S., Fitzgerald, K.A., Paludan, S.R., Bowie, A.G., 2010. IFI16 is an innate immune sensor for intracellular DNA. *Nat. Immunol.* 11, 997–1004. <https://doi.org/10.1038/ni.1932>
- Uzonyi, A., Nir, R., Shliefer, O., Stern-Ginossar, N., Antebi, Y., Stelzer, Y., Levanon, E.Y., Schwartz, S., 2021. Deciphering the principles of the RNA editing code via large-scale systematic probing. *Mol. Cell* 81, 2374-2387.e3. <https://doi.org/10.1016/j.molcel.2021.03.024>
- Valente, L., Nishikura, K., 2007. RNA Binding-independent Dimerization of Adenosine Deaminases Acting on RNA and Dominant Negative Effects of Nonfunctional Subunits on Dimer Functions. *J. Biol. Chem.* 282, 16054–16061. <https://doi.org/10.1074/jbc.M611392200>
- Wagner, R.W., Smith, J.E., Cooperman, B.S., Nishikura, K., 1989. A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and *Xenopus* eggs. *Proc. Natl. Acad. Sci.* 86, 2647–2651. <https://doi.org/10.1073/pnas.86.8.2647>

- Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J., Nishikura, K., 2004. Stress-induced Apoptosis Associated with Null Mutation of ADAR1 RNA Editing Deaminase Gene. *J. Biol. Chem.* 279, 4952–4961. <https://doi.org/10.1074/jbc.M310162200>
- Wang, Y., Beal, P.A., 2016. Probing RNA recognition by human ADAR2 using a high-throughput mutagenesis method. *Nucleic Acids Res.* 44, 9872–9880. <https://doi.org/10.1093/nar/gkw799>
- Wang, Y., Havel, J., Beal, P.A., 2015. A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1. *ACS Chem. Biol.* 10, 2512–2519. <https://doi.org/10.1021/acscchembio.5b00711>
- Wang, Y., Kanneganti, T.-D., 2021. From pyroptosis, apoptosis and necroptosis to PANoptosis: A mechanistic compendium of programmed cell death pathways. *Comput. Struct. Biotechnol. J.* 19, 4641–4657. <https://doi.org/10.1016/j.csbj.2021.07.038>
- Wang, Y., Park, S., Beal, P.A., 2018. Selective Recognition of RNA Substrates by ADAR Deaminase Domains. *Biochemistry* 57, 1640–1651. <https://doi.org/10.1021/acs.biochem.7b01100>
- Ward, S.V., George, C.X., Welch, M.J., Liou, L.-Y., Hahm, B., Lewicki, H., De La Torre, J.C., Samuel, C.E., Oldstone, M.B., 2011. RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis. *Proc. Natl. Acad. Sci.* 108, 331–336. <https://doi.org/10.1073/pnas.1017241108>
- Weissenberger, G., Henderikx, R.J.M., Peters, P.J., 2021. Understanding the invisible hands of sample preparation for cryo-EM. *Nat. Methods* 18, 463–471. <https://doi.org/10.1038/s41592-021-01130-6>
- Werry, T.D., Loiacono, R., Sexton, P.M., Christopoulos, A., 2008. RNA editing of the serotonin 5HT2C receptor and its effects on cell signalling, pharmacology and brain function. *Pharmacol. Ther.* 119, 7–23. <https://doi.org/10.1016/j.pharmthera.2008.03.012>
- Wies, E., Wang, M.K., Maharaj, N.P., Chen, K., Zhou, S., Finberg, R.W., Gack, M.U., 2013. Dephosphorylation of the RNA sensors RIG-I and MDA5 by the phosphatase PP1 is essential for innate immune signaling. *Immunity* 38, 437–449. <https://doi.org/10.1016/j.immuni.2012.11.018>
- Wong, S.K., Sato, S., Lazinski, D.W., 2001a. Substrate recognition by ADAR1 and ADAR2. *RNA* 7, 846–858. <https://doi.org/10.1017/S135583820101007X>
- Wright, D.J., Force, C.R., Znosko, B.M., 2018. Stability of RNA duplexes containing inosine-cytosine pairs. *Nucleic Acids Res.* 46, 12099–12108. <https://doi.org/10.1093/nar/gky907>
- Wright, D.J., Rice, J.L., Yanker, D.M., Znosko, B.M., 2007. Nearest Neighbor Parameters for Inosine-Uridine Pairs in RNA Duplexes. *Biochemistry* 46, 4625–4634. <https://doi.org/10.1021/bi0616910>
- Wu, B., Peisley, A., Richards, C., Yao, H., Zeng, X., Lin, C., Chu, F., Walz, T., Hur, S., 2013a. Structural Basis for dsRNA Recognition, Filament Formation, and Antiviral Signal Activation by MDA5. *Cell* 152, 276–289. <https://doi.org/10.1016/j.cell.2012.11.048>
- Wu, J., Chen, Z.J., 2014. Innate Immune Sensing and Signaling of Cytosolic Nucleic Acids. *Annu. Rev. Immunol.* 32, 461–488. <https://doi.org/10.1146/annurev-immunol-032713-120156>
- 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., 2014. Adenosine Deaminase Acting on RNA 1 Limits RIG-I RNA Detection and Suppresses IFN Production Responding to Viral and Endogenous RNAs. *J. Immunol.* 193, 3436–3445. <https://doi.org/10.4049/jimmunol.1401136>

- Yang, W., Wang, Q., Kanes, S.J., Murray, J.M., Nishikura, K., 2004. Altered RNA editing of serotonin 5-HT_{2C} receptor induced by interferon: implications for depression associated with cytokine therapy. *Mol. Brain Res.* 124, 70–78. <https://doi.org/10.1016/j.molbrainres.2004.02.010>
- Yeo, J., Goodman, R.A., Schirle, N.T., David, S.S., Beal, P.A., 2010. RNA editing changes the lesion specificity for the DNA repair enzyme NEIL1. *Proc. Natl. Acad. Sci.* 107, 20715–20719. <https://doi.org/10.1073/pnas.1009231107>
- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., Fujita, T., 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.* 5, 730–737. <https://doi.org/10.1038/ni1087>
- Yu, Q., Herrero Del Valle, A., Singh, R., Modis, Y., 2021. MDA5 disease variant M854K prevents ATP-dependent structural discrimination of viral and cellular RNA. *Nat. Commun.* 12. <https://doi.org/10.1038/s41467-021-27062-5>
- Yu, Q., Qu, K., Modis, Y., 2018. Cryo-EM Structures of MDA5-dsRNA Filaments at Different Stages of ATP Hydrolysis. *Mol. Cell* 72, 999-1012.e6. <https://doi.org/10.1016/j.molcel.2018.10.012>
- Zambrano-Mila, M.S., Witzemberger, M., Rosenwasser, Z., Uzonyi, A., Nir, R., Ben-Aroya, S., Levanon, E.Y., Schwartz, S., 2023. Dissecting the basis for differential substrate specificity of ADAR1 and ADAR2. *Nat. Commun.* 14, 8212. <https://doi.org/10.1038/s41467-023-43633-0>
- Zhang, P., Zhu, Y., Guo, Q., Li, J., Zhan, X., Yu, H., Xie, N., Tan, H., Lundholm, N., Garcia-Cuetos, L., Martin, M.D., Subirats, M.A., Su, Y.-H., Ruiz-Trillo, I., Martindale, M.Q., Yu, J.-K., Gilbert, M.T.P., Zhang, G., Li, Q., 2023. On the origin and evolution of RNA editing in metazoans. *Cell Rep.* 42, 112112. <https://doi.org/10.1016/j.celrep.2023.112112>
- Zhang, T., Yin, C., Fedorov, A., Qiao, L., Bao, H., Beknazarov, N., Wang, S., Gautam, A., Williams, R.M., Crawford, J.C., Peri, S., Studitsky, V., Beg, A.A., Thomas, P.G., Walkley, C., Xu, Y., Poptsova, M., Herbert, A., Balachandran, S., 2022. ADAR1 masks the cancer immunotherapeutic promise of ZBP1-driven necroptosis. *Nature* 606, 594–602. <https://doi.org/10.1038/s41586-022-04753-7>
- Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., Scheres, S.H., 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* 7, e42166. <https://doi.org/10.7554/eLife.42166>
- Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., Siddell, S.G., Ludewig, B., Thiel, V., 2011. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* 12, 137–143. <https://doi.org/10.1038/ni.1979>

Appendix data

List of generated RNAs

RNA name	RNA sequence	Description
5HT2C extension	GGUCUUGUCUCUGUGGUCUGGGAUCGGUAUGUAG CAAUACGUAAUCCUAUUGAGCAUAGCCGUUCAAU UCGCGGACUAAGGCCAUCAUGAAGAUUGCUAUUGU UUGGGCAAUUUCUAUAGGUAAAUAAAACUUUUUG GCCAUAGAAUUGCAGCGGCUAUGCUCAAUACUUU CGGAUUUAUGUACUGUGAACAACGUACAGACGUCGA CUGGUAA CUCUCGUCGCGUUGUCCUU	dsRNA substrate for ADAR deamination assay derived from 5HT2C gene (serotonin receptor)
Gria2 R/G	GGAAGGAUCCUCAUUAAGGUGGGUGGAAUAGUAA AACAAUAUGCUAUUAAGUUGUUAUAGUAUCCACCU ACCCUGAUGUAUCUUU	dsRNA substrate for ADAR1 complex formation derived from GRIA2 gene (site R/G)
Gria2 R/G extension	GGUCUUGUCUCUGUGGUCUGGGAAGGAUCCUCAU UAAGGUGGGUGGAAUAGUUAACAAUAUGCUIAA UGUUGUUAUAGUAUCCACCUACCCUGAUGUAUCU UU CUCUCGUCGCGUUGUCCUU	dsRNA substrate for deamination assay derived from GRIA2 gene (site R/G)
NEIL1	GGAGCCUGCCCUCUGAUCUCUGCCUGUUCUCUGU CCCACAGGGGGCAAAGGCUACGGGUCAGAGAGCGG GGAGGAGGAC	dsRNA substrate for ADAR1 complex formation derived from NEIL1 gene
NEIL1 extension	GGUCUUGUCUCUGUGGUCUGGGAGCCUGCCCUCUG AUCUCUGCCUGUUCUCUGUCCACAGGGGGCAA GGCUACGGGUCAGAGAGCGGGGAGGAGGAC CUCUC GUCGCGUUGUCCUU	dsRNA substrate for ADAR deamination assay derived from <i>NEIL1</i> gene
AJUBA	GGUUUUGGGGUUGUGGUUGAUGCAGUGUGGGAU GUCCUGAGAGGUAGCAAGUCUAGGGUGUUGCCAC CUCUCAGAGGGGUCCCGGAUUGCAUCCAUCACAAU CCCAAAC	dsRNA substrate for ADAR deamination assay and ADAR1 complex formation derived from <i>AJUBA</i> gene
TTYH2	GGCAUGCUICAUACCCAGAGAGAAGCCCCGGCUGC CCAGGCAUGCUIUAGGCUUACACGUGCUUAGGCUUA GGCGUGCCUGGGUGACCAGGGCGUUCUCUCUGGG UGUGAAGAACU	dsRNA substrate for ADAR1 complex formation derived from <i>TTYH2</i> gene
TTYH2 extension	GGUCUUGUCUCUGUGGUCUGGGCAUGCUICAUACC CAGAGAGAAGCCCCGGCUGCCCAGGCAUGCUIUAGG CUUACACGUGCUUAGGCUUAGGCGUGCCUGGGUGA CCAGGGCGCUUCUCUCUGGGUGUGAAGAA CUCUCG UCGCGUUGUCCUU	dsRNA substrate for ADAR deamination assay derived from <i>TTYH2</i> gene
irAlu NICN1 full-length extension	GGUCUUGUCUCUGUGGUCUGGGCAGGGUGCAGUG GCUCACGCCUGUAAUCCAGCACUUUGGGAGGCCCA GGCAGGCGGAUCACCUGAGGUCAGGAGUUCAAGAC	Alu: Alu hybrid element derived from Nicn1 gene. Used for the ADAR

	<p>CAGCCUGACCAACCUUGGAGAAACCCCGUCUCUACUA AAAAUACAAAAAUUAGCUGGGCGUGGUGGUAGGC ACCUGUAAUCCCAGCUACUUGGGAGGCUGAGGCAG GAGAAUCUCUUGAACCCGGGAAGUGGAGGUUGCGG ACCUGAGAUCUUGCCAUUGCACUCCAGCCUGGGCAA GAAGAGCGAAACUCCAUCUAAAACAAACAAACAAA AAAAAGAACUGGGACCCUUCUGCCAUCUGACAUA GCCCAAAGCACAUCUCUAUCCUUCUCCCAGUUGCC CCUCUCCUUUUUUGUUGUUUUUUUUGAGGUUGAG UUUUGCUCUUGUUGCCCAGGCUGGAGUGCAAUAG UGCAAUCUUGGCUAACUGCAACCUCCGCCUCCCAGG UUCAAGCAAUUCUCCUGCCUCAGUCUCCCAGUAGC UGGGAUUACAGUCAUGCAUCACCAUGCCUGGCUAA UUUUGUAUUUGUAGUAGAGAUGGGGUUUCUCCA GUUGGUCAGGCUGGUCUAAACACCUAGCCUCAGG UGAUCUGCCUGCCUUGGCCUCCAAAGUGCUGGGA UUACAGGCAUGAGCCACCGCGCCCGCCCUUCUGUC GCGUUGUCCU</p>	deamination assay and complex formation with MDA5.
irAlu NICN1 C extension	<p>GGUCUUGUCUCUGUGGUCUGUACUUGGGAGGCUG AGGCAGGAGAAUCUCUUGAACCCGGGAAGUGGAGG UUGCGGACCUGAGAUCUUGCCAUUGCACUCCAGCC UGGGCAAGAAGAGCGAAACUCCAUCUAAAACAAACA AACAAAAAAGAACUGGGACCCUUCUGCCAUCU GACAUAGCCCAAAGCACAUCUCUAUCCUUCUCCCA GUUGCCCCUCUCCUUUUUUGUUGUUUUUUUGAG GUUGAGUUUUGCUCUUGUUGCCCAGGCUGGAGUG CAAUAGUGCAAUCUUGGCUAACUGCAACCUCCGCCU CCCAGGUUCAAGCAAUUCUCCUGCCUCAGUCUCCCG AGUACUCUCGUCGCGUUGUCCU</p>	A truncated version of Alu: Alu hybrid element derived from the Nicn1 gene.
irAlu NICN1 D extension	<p>GGUCUUGUCUCUGUGGUCUGGGCAGGGUGCAGUG GCUCACGCCUGUAAUCCCAGCACUUUGGGAGGCCCA GGCAGGCGGAUCACCUAGAGGUCAGGAGUUCAAGAC CAGCCUGACCAACCUUGGAGAAACCCCGUCUCUACUA AAAAUACAAAAAUUAGCUGGGCGUGGUGGUAGGC ACCUGUAAUCCCAGCCCUUCUGCCAUCUGACAUAG CCCAAAGCACAUCUCUAUCCUUCGCUGGGAUUACA GUCAUGCAUCACCAUGCCUGGCUAAUUUUGUAUUU GUAGUAGAGAUGGGGUUUCUCCAUGUUGGUCAGG CUGGUCUCAAAACACCUAGCCUCAGGUGAUCUGCCU GCCUUGGCCUUCCAAAGUGCUGGGAUUACAGGCAU GAGCCACCGCGCCCGCCCUUCUGUCGCGUUGUCCU U</p>	A truncated version of Alu: Alu hybrid element derived from the Nicn1 gene.
irAlu NICN1 A sense	<p>GGAUCACCUAGAGGUCAGGAGUUCAAGACCAGCCUG ACCAACCUUGGAGAAACCCCGUCUCUACUAAAAUAC AAAAAUUAGCUGGGCGUGGUG</p>	A truncated version of Alu: Alu hybrid element derived from the Nicn1 gene. Strand sense

irAlu NICN1 A antisense	GGCACCAUGCCUGGCUAUUUUUGUAUUUGUAGUA GAGAUGGGGUUUCUCAUGUUGGUCAGGCUGGUC UCAACACCUGACCUCAGGUGAU	A truncated version of Alu: Alu hybrid element derived from the Nicn1 gene. Strand antisense
irAlu BPNT1 full-length extension	GGUCUUGUCUCUGUGGUCUG AAAGUUUCAUUUGG CCGGGCGCGGUGGCUCAUGCCUGUAAUCCAGCAC UUUUGGGAGGCCGAGGCAGGUGGAUCACUUGAGCUC AGGAGUUUGAGACCAGCCUGGGCAAUAUCGUGAGA CCCCAUCUCUACAAAAUACAAAUUAACUGGGCAUC CUGUCAUGCGCCUGUCAUCCAGCUACUUGAGAGG CUGAAGCAGAAGAAUCUCUUGAGCCCGGAAGGCAG AGGUUGCAGUGAGCUGAGAUCGUGCCACUGCACUC CAGCCUGAGUGACAGGAGUUAAGCCUGUCUCAGA AAAAAAAACUCUUUUUUUUUCCUGAGACGGAGUCU CGCUCUGUAGCCUAGGCUAGAAUGUAGUGGCGUGA UCUCGGCUCACUGCAAGCUCGCCUCCCGGGUUCAU GCCAUUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGA CUGCAGGCACCGCCACCACGCCUGGCUAAUUUUUU GUUUUUUAGUAGAGAAGGUGUUUCACCGUGUUA GCCAGGAUGGUCUUGAUCUCCUGACAUCGUGAUCU GUCUGCCUCGGACUCCCAAAGUGCUGGGAAUACAG GUGUGAGCCACCGCACCUGGCCUAAACCAGAUUUC UCUCGUCGCGUUGUCCU	Alu: Alu hybrid element derived from Bpnt1 gene. Used for the ADAR deamination assay and complex formation with MDA5.
irAlu BPNT1 A extension	GGUCUUGUCUCUGUGGUCUG CCUGUCAUCCAGCU ACUUGAGAGGCUGAAGCAGAAGAAUCUCUUGAGCC CGGAAGGCAGAGGUUGCAGUGAGCUGAGAUCGUGC CACUGCACUCCAGCCUGAGUGACAGGAGUUAAGCCC UGUCUCAGAAAAAAAACUCUUUUUUUUUCCUGAG ACGGAGUCUCGCUCUGUAGCCUAGGCUAGAAUGUA GUGGCGUGAUCUCGGCUCACUGCAAGCUCGCCUC CCGGGUUCAUGCCAUUCUCCUGCCUCAGCCUCCCGA GUAGCUGGGACUGCAGG CUCUCGUCGCGUUGUCCU U	A truncated version of Alu: Alu hybrid element derived from the Bpnt1 gene.
irAlu BPNT1 B extension	GGUCUUGUCUCUGUGGUCUG AAAGUUUCAUUUGG CCGGGCGCGGUGGCUCAUGCCUGUAAUCCAGCAC UUUUGGGAGGCCGAGGCAGGUGGAUCACUUGAGCUC AGGAGUUUGAGACCAGCCUGGGCAAUAUCGUGAGA CCCCAUCUCUACAAAAUACAAAUUAACUGGGCAUC CUGUCAUGCGCACCGCCACCACGCCUGGCUAAUUUU UUGUAUUUUUAGUAGAGAAGGUGUUUCACCGUGU UAGCCAGGAUGGUCUUGAUCUCCUGACAUCGUGAU CUGUCUGCCUCGGACUCCCAAAGUGCUGGGAAUAC AGGUGUGAGCCACCGCACCUGGCCUAAACCAGAUU UCUCGUCGCGUUGUCCU	A truncated version of Alu: Alu hybrid element derived from the Bpnt1 gene.

Table 6 List of generated RNAs.

In blue is highlighted an extension sequence EL15/16 used in editing assay or PCR amplification

AlphaFold3 models

Models of $\Delta 133$ ADAR1 p150 - TTYH2 or ADAR2 - TTYH2 RNA complex including IDRs

Model of ADAR2 - TTYH2 RNA complex including IDRs

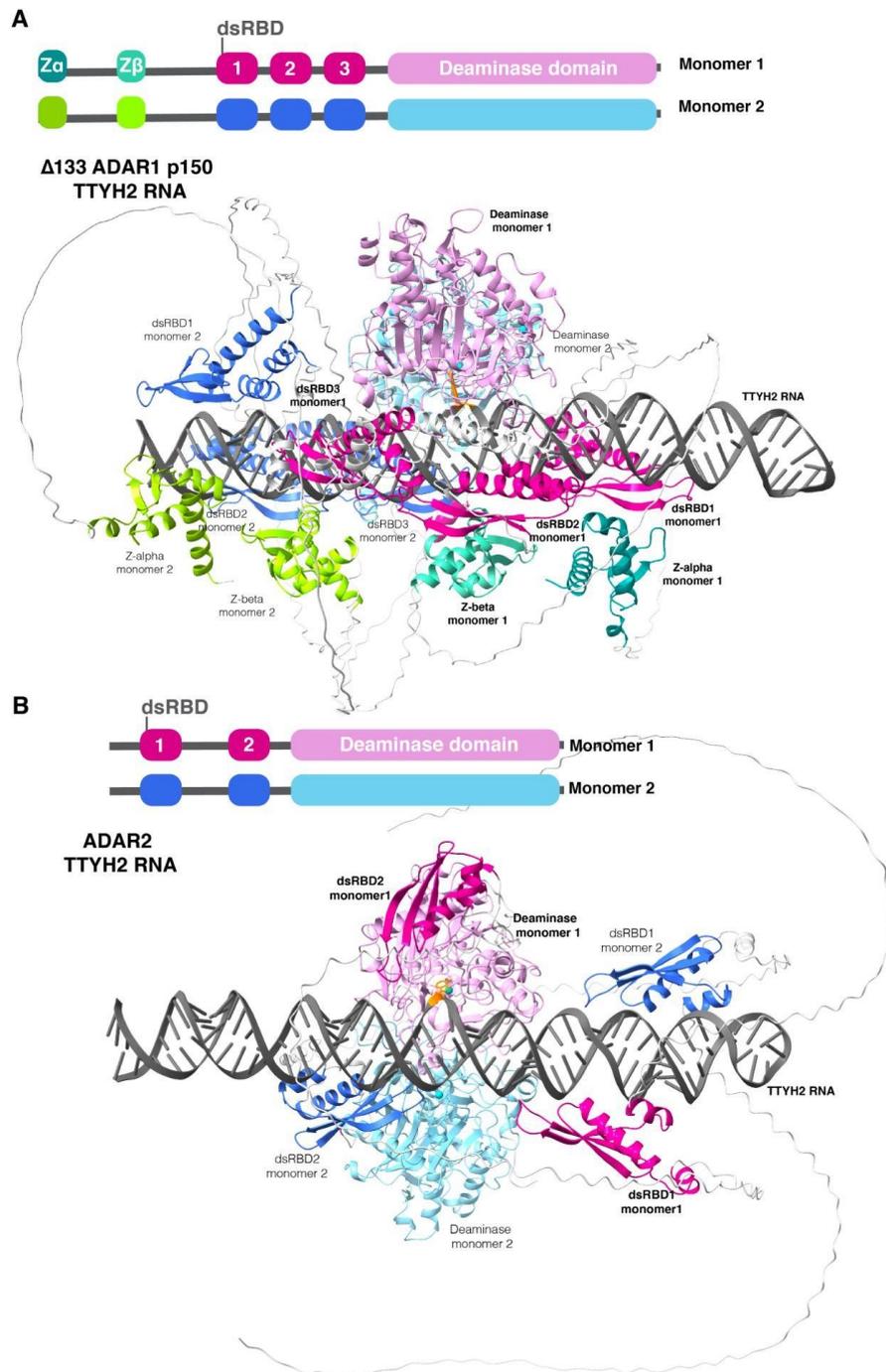


Figure 51 Models of ADAR - TTYH2 RNA complex including IDRs.

A) $\Delta 133$ ADAR1 p150 - TTYH2 RNA

B) ADAR2 - TTYH2 RNA