Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Megan Korn April 14, 2020

Synthesis of a Click-Compatible Phenylacrylamide to Probe A-to-I RNA Editing

by

Megan M. Korn

Jennifer M. Heemstra, Ph.D.
Adviser

Department of Chemistry

Jennifer M. Heemstra, Ph.D.

Adviser

Khalid Salaita, Ph.D.

Committee Member

Carolyn Keogh, Ph.D.

Committee Member

2020

Synthesis of a Click-Compatible Phenylacrylamide to Probe A-to-I RNA Editing

By

Megan M. Korn

Jennifer M. Heemstra, Ph.D.

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Chemistry

2020

Abstract

Synthesis of a Click-Compatible Phenylacrylamide to Probe A-to-I RNA Editing
By Megan M. Korn

Adenosine-to-inosine (A-to-I) editing is one of the most common modifications in the human transcriptome, however full range of its biological function has not been fully explored. A lack of cost-effective, simple, and quick tools to probe A-to-I editing has limited research into functions of A-to-I sites and the therapeutic potential of ADAR. Herein, a click-compatible phenylacrylamide was developed to probe in vitro ADAR interactions with RNA substrates. *N*-(4-ethylnylphenyl)acrylamide (EPhAA) was synthesized in one-step with 42% yield, and carried out 90% percent conversion of inosine ribonucleoside. Using copper click chemistry, a Cy5 moiety was covalently attached to EPhAA-labeled RNA oligonucleotides and percent labeling was quantified with a gel shift assay. In vitro editing of HER1—a known ADAR substrate—was carried out using a mutant E1008Q ADAR1, which edits at a quicker rate than wild type ADAR1. Using EPhAA, Cy5 was click-conjugated to edited HER1 transcripts and editing was quantified with a gel shift assay. EPhAA successfully demonstrated enhanced editing of E1008Q ADAR1 versus wild type ADAR1. This system offers a cheaper, simpler, and quicker way to quantify A-to-I editing, as compared to other chemical labeling methods such as radioactive labeling and ICE-seq. The demonstrated labeling protocol is an ideal system for characterizing relative editing rates of mutated ADARs or for exploring novel RNA substrates of ADAR.

Synthesis of a Click-Compatible Phenylacrylamide to Probe A-to-I RNA Editing

By

Megan M. Korn

Jennifer M. Heemstra, Ph.D.

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Department of Chemistry

2020

Acknowledgements

First and foremost, I would like to thank Dr. Jen Heemstra, Dr. Khalid Salaita, and Dr. Carolyn Keogh for serving on my committee. Your feedback and flexibility in this process has been greatly appreciated and has helped me grow as a researcher.

I'd like to especially thank Jen for giving me the opportunity to work in the Heemstra lab throughout my undergraduate career. My time in the group has defined my Emory experience and I am so grateful for your mentorship, support, and kindness.

I'd like to thank Steve Knutson and Misael Romero-Reyes, both of whom have formally served as my graduate student mentors. Steve has challenged me to keep up with his fast-paced PhD work and has helped me improve my organizational skills, creativity, and chem-bio skillset. Misael introduced me to the ways of organic synthesis and helped integrate me into the lab. I can't thank both of you enough for taking the time to teach me, mentor me, and inspire me to be a better chemist.

While Steve and Misael held the formal title, it's true that every Heemstra lab grad student and post-doc, past and present, has been a mentor to me at one time or another. By name: thank you Colin, T, Erin, Aimee, Rachel, Arventh, Hector, Diane, Brea, Arunava, Satheesh, Meredith, Shayla, and Travis.

I'd like to especially thank Heemstra lab members Steve, Colin Swenson, Tewoderos Ayele, Deanna Dailey, and Ryan Johnson for contributing to this project. I'd also like to acknowledge collaborators: Dr. Peter Beal and Leanna Monteleone, who expressed and purified ADAR1 for use in this project.

Without support from my family, I would not have gotten this far. Thank you to my parents, Lucy and Bob; brother, Nigel; nephew, Gunner; and grandmother, Lucille, for cheering me on.

Table of Contents

Table of Figures

1. Introduction

1.1 Epitranscriptomics

Nucleic acids are biopolymers that encode information necessary for life. RNA, an intermediate molecule that transfers information between DNA and proteins, consists of four canonical bases—adenosine, guanosine, cytosine, and uridine. In messenger RNA (mRNA), the order of these bases directly determines the structure of proteins.[1] However, RNA can be modified extensively after transcription, and these four bases can undergo a variety of chemical transformations that alter their function.[2] Endogenous enzymes catalyze these base-editing reactions to decorate the canonical bases of RNA transcripts, and these subsequent modified bases within a cell comprise the "epitranscriptome".[2] These modifications are found on a range of RNA types, including tRNA, mRNA, and miRNA.[3] There are over 140 different modified bases, and new bases are continually discovered and added to databases that catalog types of post-transcriptional modifications.[4]

1.2 A-to-I Editing

One of the most common editing events in the human epitranscriptome is the deamination of adenosine to form inosine, as shown in Figure 1.[5] Adenosine-to-inosine (A-to-I) editing is carried out by a family of enzymes called ADARs (adenosine deaminases that act on RNA). The reaction proceeds by hydrolytic attack, whereby the oxygen of a water molecule carries out a nucleophilic attack on the C6 of the adenosine base, releasing ammonia as a side product.[6] The active site of ADAR contains a tripeptide sequence—histidine, glutamate, and alanine—which coordinates Zn^{2+} in order to activate water molecules for nucleophilic attack.[6] ADARs also contain lysine-rich

double-stranded RNA-binding domains (dsRBD), that position RNA targets for editing.[6] Recent reports suggest that neighboring base pairs in RNA editing sites serve as ADAR recognition elements to promote ADAR editing at specific locations in the transcript.[7] While some of this recognition arises from the dsRBD of ADAR, these data also suggest that site selectivity is also rooted in the deaminase active site.[8–10]

An important aspect of A-to-I editing is that the process changes the Watson-Crick-Franklin base pairing of the nucleobase.[7] Instead of pairing with a uridine base as adenosine does, the introduced inosine instead base pairs with cytosine, which is illustrated in Figure 1.[7] As such, this edit effectively changes the sequence of the source genome.[7]

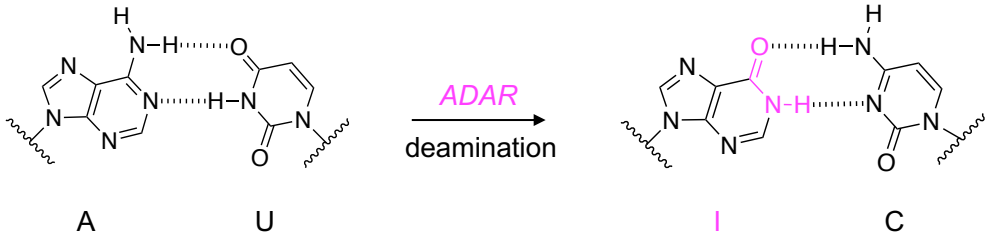


Figure 1. ADAR catalyzes the deamination of adenosine to yield inosine. Adenosine base pairs with uracil, while inosine base pairs with cytosine.

1.3 Biological Function of A-to-I Editing

The biological function of A-to-I editing is of great interest to researchers, particularly because it is a highly pervasive modification in the epitranscriptome.[5] Several functions have been linked to A-to-I editing: it can regulate splicing patterns[11], plays a role in recognition of viral RNA[6], recodes amino acids[7,12,13], contributes to the immune response[14], impacts small RNA function[12], introduces exons in Alu regions[12], prevents circular RNA biogenesis[12], and perhaps drives evolution by diversification of the transcriptome.[12]

A-to-I editing occurs in both coding and noncoding RNA transcripts.[15] As such, researchers hypothesize that the function of inosine is different depending on the type of RNA.[12,13] In the case of coding RNA transcripts, A-to-I editing changes the nucleotide sequence of mRNA and can cause phenotypic changes in the translated protein. Editing sites in protein-coding regions often reside in neurotransmitter receptors, including the glutamate receptor subunit GluR2, a potassium channel called Kv1.1, and the $\alpha 3$ subunit of $GABA_A$ receptor (GABRA3).[13] In these examples, A-to-I editing changes one amino acid to another which changes the overall function of the protein, either by directly inhibiting Ca^{2+} ion flow in the case of GluR2 or by preventing proper protein trafficking and localization in the case of GABRA3.[13] On the other hand, A-to-I editing of noncoding RNAs, such as microRNAs (miRNA), modifies translation of specific mRNA transcripts.[16] miRNA plays an important role in translation regulation by forming a duplex with complementary mRNA to silence that mRNA's expression.[13] Therefore, A-to-I editing of miRNA can change which mRNA region it targets or inhibits its silencing function altogether.[13] For example, editing of miR-376a-5p silences the translation of phosphoribosyl pyrophosphate synthetase-1, a protein that synthesizes uric acid. This suggests A-to-I editing plays a role in regulating waste production in cells.[13,16]

While recent work has improved our understanding of A-to-I editing, there are estimated to be greater than 2 million editing sites, with only a small fraction of those having characterized function.[12] Based on the importance of A-to-I editing in the aforementioned examples, there exists a need to probe further into ADAR and mRNA dynamics.

1.4 ADAR as a Therapeutic Agent

The ability to change the sequence of mRNA or to influence its translation is of great interest to researchers, especially utilizing an enzyme that is both endogenous and site selective. The ADAR family consists of three types: ADAR1, ADAR2, and ADAR3.[8] ADAR1 and ADAR2 carry out the catalytic deamination reaction in humans, while ADAR3 has no catalytic activity and is hypothesized to serve a regulatory role.[8] ADAR1 and ADAR2 often target different adenosines—ADAR1 tends to edit repetitive sites, while ADAR2 primarily targets non-repetitive editing sites.[8] However, it is unclear what types of sequence context or secondary structure makes certain adenosines more likely to be edited than others. Furthermore, it is unclear what domains of ADAR1 and ADAR2 afford site-selective editing of certain transcripts.

As described in the previous section, editing of mRNA transcripts leads to downstream effects in protein function. Several researchers and companies have proposed harnessing the selectivity and specificity of ADAR1 and ADAR2 to insert coding changes into mRNA transcripts in an effort to transiently change protein function. In this way, ADAR can be used as a therapeutic to treat genetic disease. For example, ProQR, a biotech therapeutics company based in the Netherlands, has been developing Axiomer technology, which uses antisense oligonucleotides to induce A-to-I editing at specific spots in the transcriptome.[17] The company targets diseases caused by G-to-A mutations, which they hypothesize can be treated with A-to-I editing.[18] A key component of accelerating this research is understanding the site-selectivity of ADAR1 and ADAR2 on specific transcripts. However, current tools for studying RNA substrate-ADAR

interactions rely on a tool called RNA-seq, which can be cost-prohibitive and time consuming, as described in section 1.5.

1.5 Methods for Detecting Inosine in the Epitranscriptome

Since A-to-I editing changes the base pairing properties of the nucleotide, inosine can be detected in an RNA sequence by reverse transcribing the sequence, leading to each inosine being coded as a guanosine in the cDNA. Thus, when comparing the original DNA sequence to cDNA data through either Sanger sequencing or next generation sequencing methods like RNA-seq, inosine bases can be identified. However, single nucleotide polymorphisms present in DNA and RNA can increase the background level of changes between the reference and edited transcript, and sequencing errors inherent to the polymerases used in RT-PCR may incorporate G bases erroneously.[5]

In order to overcome these limitations, Suzuki and coworkers developed a method called ICE-seq which utilizes a combination of chemical labeling and next generation sequencing. ICE stands for inosine chemical erasing, named so because the chemical labeling of inosine stops the first subsequent cDNA transcription, essentially removing all of the cDNA with inosine.[19–21] Sequencing of both treated and untreated sample detects the erased reads caused by inosine positions. ICE-seq relies on acrylonitrile as a labeling reagent, which attaches a covalent cyanoethyl moiety to the Watson-Crick-Franklin face of inosine, and thus hinders Watson-Crick-Franklin base pairing with C during reverse transcription. This prevents the polymerase enzymes from forming a cDNA transcript containing inosine. In combination with this improved RT-PCR, software was developed to analyze the sequencing data. The program analyzes A

to G mismatches, and compiles data to eliminate false-positive G mismatches. This combination of chemical modification and sequencing software provides a useful tool for the future of examining potential A-to-I edited sites.[21] However, depending on the amount of sequencing reads required, this tool can be both laborious, low throughput, and expensive.[22]

A cheaper method, such as radiolabeling nucleotides, can be used to detect ADAR activity and identify A-to-I edited sites. In this protocol, an RNA transcript is radiolabeled with [^{32}P] and subjected to *in vitro* editing.[23] The RNA transcript is extracted, purified, and digested to individual nucleotides, and then resolved on a 2D TLC plate to separate the distinct nucleotides to determine if any editing occurred on the transcript.[23] This method is quite sensitive, however is limited by its dangerous protocol, since it requires radioactive materials and generates radioactive waste.

In 2018, Knutson and coworkers developed a method inspired by ICE-seq—an acrylonitrile derivative was synthesized with a fluorescein handle to selectively label A-to-I edited transcripts with the ability to visualize editing fluorescently. The method can also be used to enrich transcripts by using an antifluorescein antibody against the labeled transcripts. Acrylamidofluorescein was limited by its off-target labeling of canonical bases, uridine and guanosine, as well as poor solubility.[24]

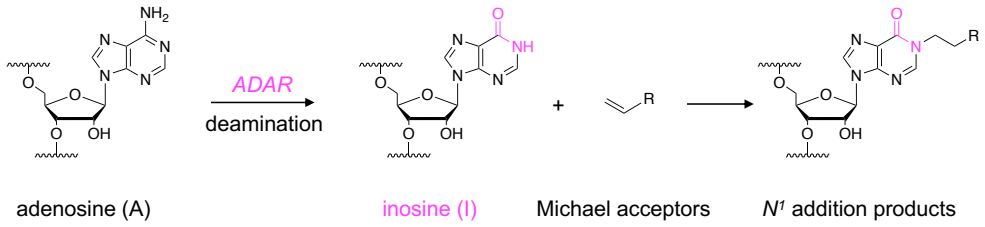


Figure 2. Inosine can be chemically labeled by Michael acceptors, yielding N^1 addition products, which halt reverse transcription enzymes or attach biochemically functional moieties.

In 2019, Li and coworkers developed a similar chemical labeling technique—an acrylonitrile derivative was synthesized with a water-soluble handle and a click handle so as to attach a biotin affinity handle via click chemistry. Using biotin-streptavidin pulldown methodology, A-to-I edited RNA transcripts are enriched prior to sequencing to improve the signal-to-noise ratio and to help discover sites with less frequent editing events.[25] This protocol also helps to lower the cost of sequencing by reducing the amount of sequencing reads necessary to achieve appropriate signal-to-noise ratio. This synthetic method is superior because of its use of an affinity handle and its solubility in aqueous conditions, however, it is limited by a multistep synthesis to create the acrylonitrile derivative and requires synthesis of dangerous organic azide intermediates.[25]

In this work, a phenylacrylamide labeling reagent has been developed to probe A-to-I editing rates on RNA substrates using the aforementioned tools as inspiration. The approaches in the field so far have provided a starting point for identifying and characterizing A-to-I editing, however they are limited to certain applications. High throughput methods like ICE-seq offer a way to identify candidate A-to-I sites and exist to discover new editing sites across the transcriptome. However, there exists a need for a cheaper, quicker alternative to quantitate editing rates of certain RNA substrates for ADAR.

The previously characterized reactivity of acrylonitrile with inosine provides a basis for stable and sufficiently high adduct formation for accurate labeling readouts. Employing a click-functionalized acrylonitrile derivative lowers the cost and amount of time necessary to characterize editing rates of certain RNA substrates. Reducing the

synthetic steps of Li and coworker's acrylonitrile derivative simplifies the protocol and avoids unstable azide intermediates. Combining these qualities into one labeling reagent allows for quick, simple, safe, and efficient A-to-I editing rate analysis, which can catalyze the characterization of ADAR substrates or mutant ADAR activity for therapeutic potential.

2. Results and Discussion

2.1 Acrylamide Derivative Reactivity Panel

Several derivatives of acrylamide were first tested for reactivity with inosine with the intent of finding the most efficient scaffold. As shown in Figure 3, acrylamide derivatives with electron withdrawing groups resulted in the highest activity. This is likely because the reaction undergoes a Michael addition mechanism, whereby the N6 atom on inosine performs a nucleophilic addition onto the C1 of an electron poor alkene. Figure 3 shows the tested derivatives and the corresponding percent inosine labeling. Percent conversion was calculated by comparing the area under the curve of inosine ribonucleoside to the addition product curve on the HPLC trace.

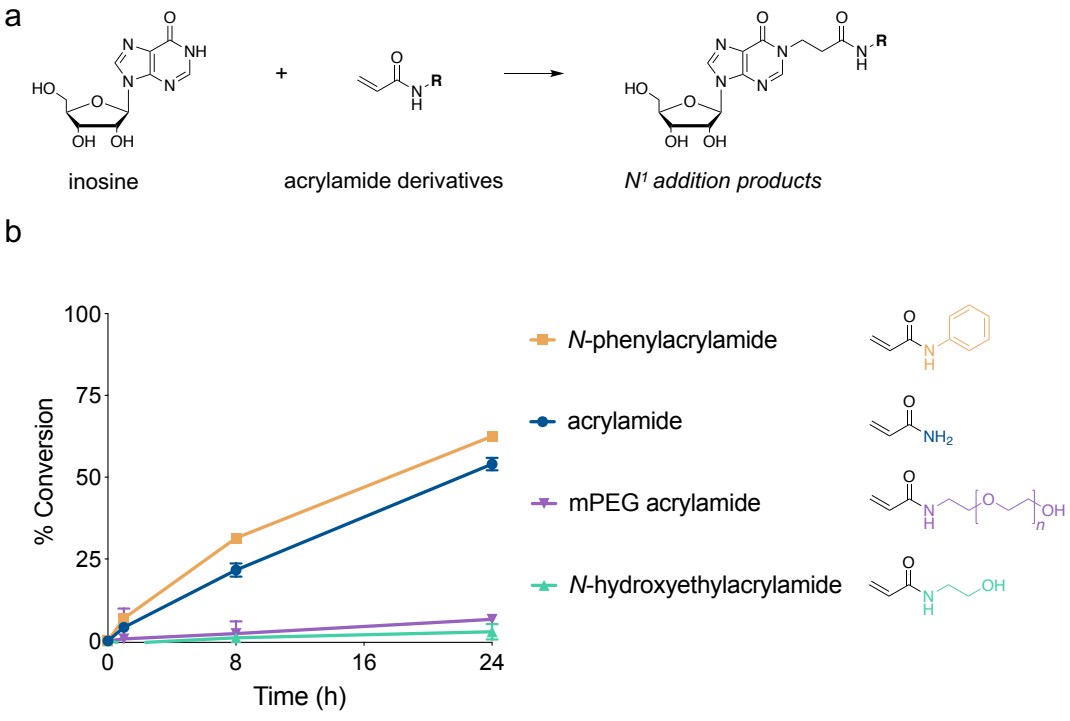


Figure 3. (a) Reaction scheme of inosine ribonucleoside labeling. (b) Percent conversion of selected acrylamide derivatives with inosine ribonucleoside. Calculated from HPLC-analyzed conversion and adduct formation.

N-phenylacrylamide exhibited the best percent conversion and further click-functionalization of this molecule was explored.

2.2 Synthesis of *N*-(4-ethynylphenyl)acrylamide

Attaching an azide handle *para* to the amine was first attempted, however optimization of the synthesis proved to be difficult and the desired product was dangerously unstable, due to a low N:C ratio. Thus, the molecule was click-functionalized with an alkyne handle, which proved to be an easy synthesis and purification, compared to the azide-functionalized version. The synthesis was inspired by Knutson and coworkers' one step synthesis of acrylamidofluorescein, whereby acrylic acid was coupled to a functionalized amine via activation of the acrylic acid and subsequent nucleophilic attack by the amine. The molecule synthesized is *N*-(4-ethynylphenyl)acrylamide (EPhAA).

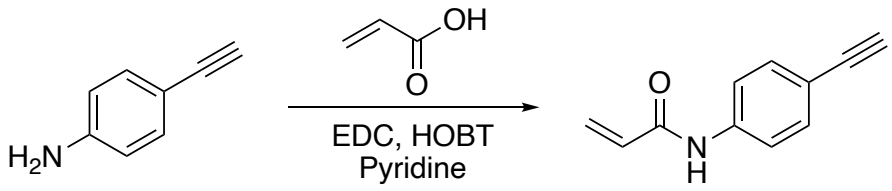


Figure 4. Synthetic scheme for EPhAA. 42% yield was achieved after 16 hours. Detailed conditions are described in Materials and Methods.

2.3 Ribonucleoside Labeling and HPLC Analysis

The reactivity of EPhAA was tested against the canonical nucleosides, inosine, and pseudouridine (Ψ), another type of post-transcriptional modification, to characterize its selectivity. Pseudouridine was included in this experiment because it is also commonly found in mammalian mRNA transcripts.[5] When developing this labeling technology, it was the original intention to utilize EPhAA as a tool for pre-enrichment of inosine-containing mRNA transcripts prior to RNA-seq, and so it was imperative to test the

reagent's selectivity against all RNA bases, both canonical and noncanonical. As shown in Figure 5, inosine ribonucleoside was labeled with EPhAA at various pH values. The reaction proceeds at a much quicker rate under basic conditions, due to deprotonation of inosine, however, RNA oligonucleotides are most stable at pH 4-5 and will degrade via phosphodiester bond cleavage under more alkaline conditions.[26] Consistent with the reaction rates in Figure 5c, the known pK_a of inosine is 8.75, thus a pH of 8.6 was used to balance quick reaction rate with RNA degradation.[27] Nucleosides were incubated with EPhAA for increasing time points and an aliquot of the reaction mixture was taken and analyzed via HPLC. Percent conversion was calculated relative to unreacted ribonucleoside.

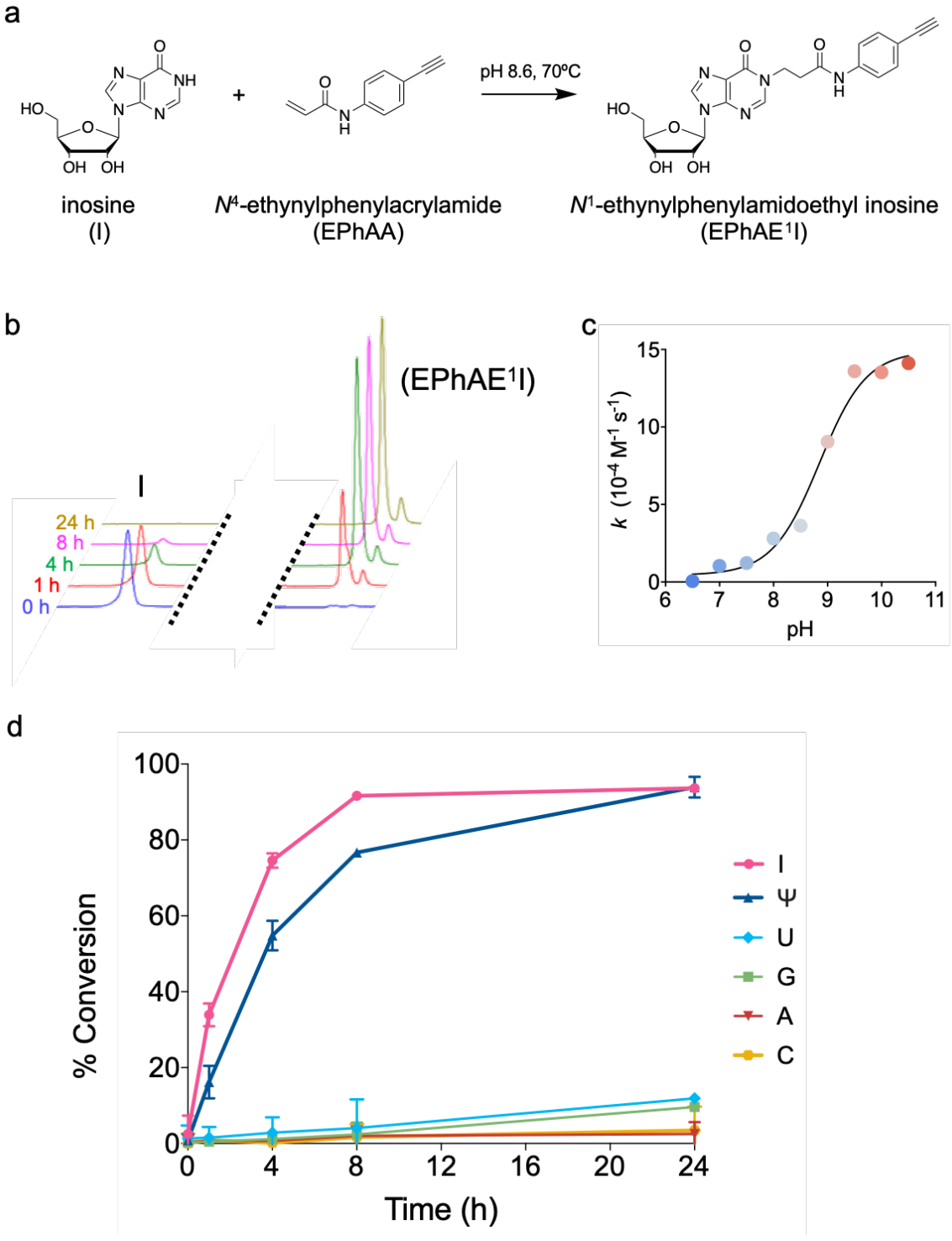
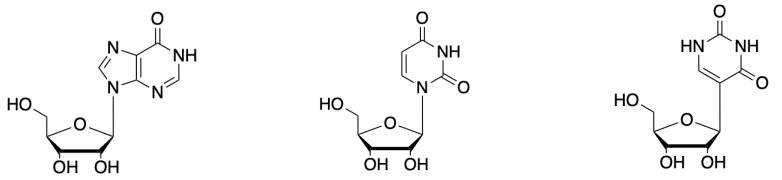


Figure 5. (a) Reaction scheme for the ribonucleoside labeling. (b) A representative HPLC trace used to quantify percent conversion for each ribonucleoside. (c) pH-dependency of the rate of reaction with inosine ribonucleoside. (d) Reactivity of EPhAA with ribonucleosides.

Minimal reactivity was observed with the canonical bases, thus validating EPhAA's specificity on the ribonucleoside level, however there was some labeling of pseudouridine and uridine ribonucleosides, which was probed further on the oligomer scale. However, inosine, uridine, and pseudouridine all have acidic nitrogen atoms, so each can act as a nucleophile to attack the alkene of EPhAA.



inosine (I) uridine (U) pseudouridine (Ψ)

Figure 6. Structure of inosine, uridine, and pseudouridine, which were all observed to react with EPhAA.

2.4 RNA Oligonucleotide Labeling and PAGE Shift Analysis

The reactivity with nucleosides was confirmed on the oligomer scale using 35-mers differing by one nucleotide, either inosine, uridine, or pseudouridine. The sequences used are shown in Materials and Methods. The variable nucleotides were chosen as these exhibited the highest reactivity with EPhAA on the ribonucleoside scale (see Figure 5d). In order to quantify percent labeling, copper click chemistry was used to covalently attach a Cy5 moiety after labeling the oligomers with EPhAA. Percent labeling was quantified via Cy5 intensity, as shown in Figure 7. The oligomers were incubated in separate reactions, with the same conditions as the ribonucleosides, and aliquots were taken at different timepoints over the course of 48 hours. Percent labeling was quantified by a gel shift assay.

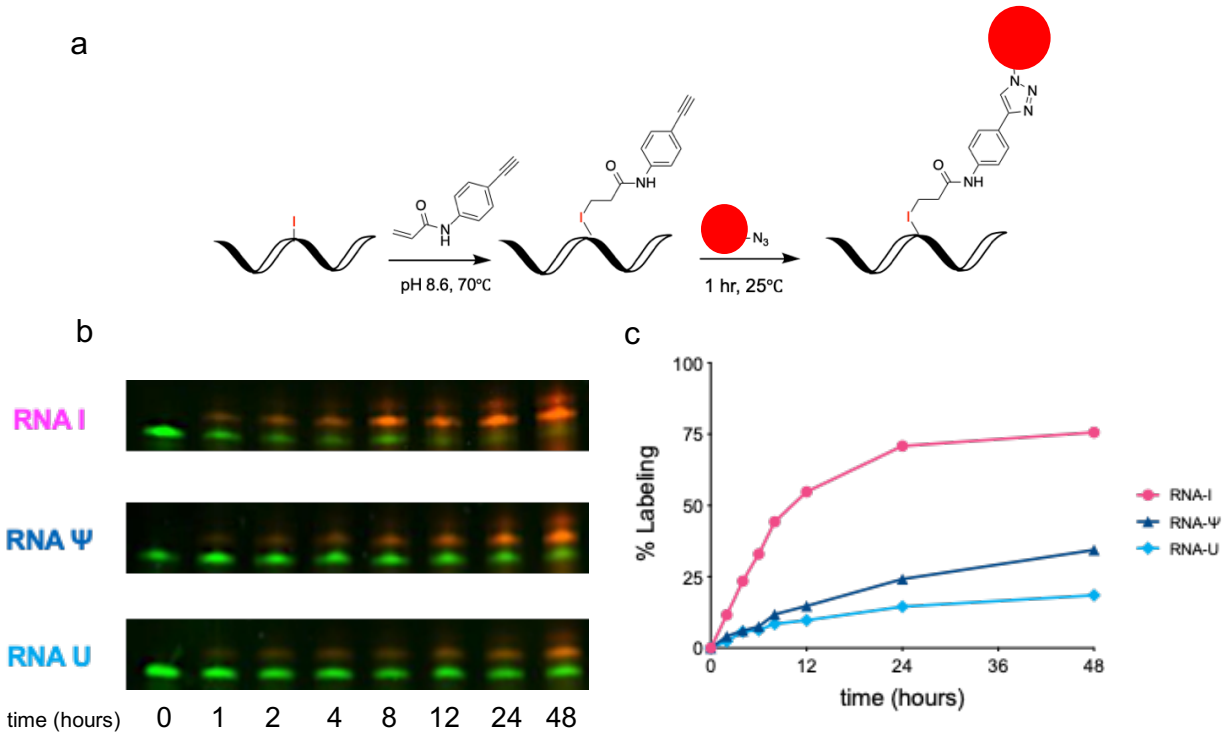


Figure 7. (a) Reaction scheme of EPhAA labeling and subsequent copper click chemistry to attach Cy5 picolyl azide (indicated by red circle). (b) PAGE assay showing EPhAA labeled oligomers (top band, red) versus unlabeled oligomers (bottom band, green) over selected time points. (c) Percent labeling of each oligomer over 48 hours. Values were calculated in ImageJ using densitometry of the labeling oligomer band divided by the total amount of oligomer in each lane.

As shown in Figure 7, RNA I was 80% labeled by EPhAA, with significant labeling of RNA Ψ and RNA U, about 30% and 20%, respectively. In both RNA I and RNA Ψ, a third band began to form around 24 hours, which is most likely labeling of guanosine bases within the sequence. This off-target reactivity with pseudouridine and uridine changed the course of EPhAA's application; the labeling-probe could not be used to enrich inosine-containing RNA from cells because EPhAA's reactivity with uridine and pseudouridine (which is also abundant in mammalian RNA) would introduce false positives. Instead, the focus of this project shifted to labeling of carefully designed synthetic RNA oligomers that minimize uridine and pseudouridine. This method

removes false positive transcripts and allows for quick and cheap *in vitro* characterization of ADAR edited RNA substrates.

2.5 HER1 RNA A Versus HER1 RNA I Labeling and PAGE Shift Analysis

Moving forward, a known RNA substrate of ADAR1 was used to test the efficacy of EPhAA.[8] This mRNA substrate, HER1 shown in Figure 8a, is a 33 nucleotide oligomer hairpin-structured RNA that is modified at a mismatched adenosine ribonucleotide. This mismatch destabilizes the RNA duplex and allows for site-specific editing.[8,28] For our purposes, we used HER1 as a starting point and designed a similar RNA oligomer, however with fewer uridine bases to limit off-target labeling. The newly designed HER1 compared to previously reported HER1 is shown in Figure 8b.

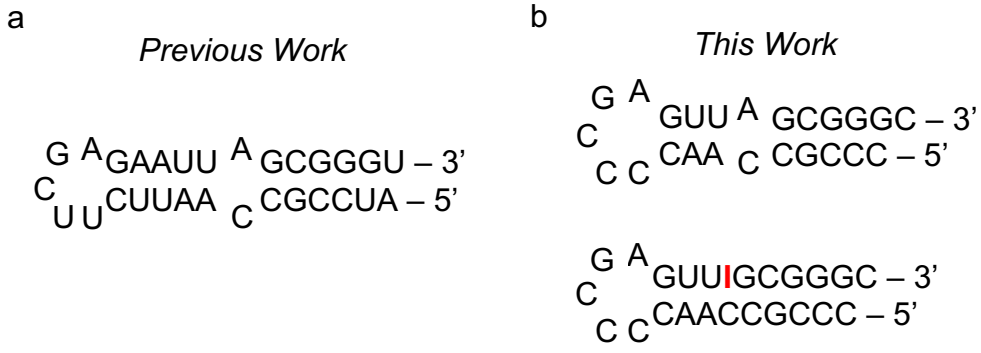


Figure 8. (a) Original HER1 mRNA transcript as described by Wang and coworkers. (b) Redesigned HER1 used in EPhAA labeling experiments in this work. Two versions of HER1 were used: one with an adenosine in the 25th position (HER1 RNA A) and the “edited” version with an inosine in the 25th position (HER1 RNA I).

HER1 RNA A and HER1 RNA I were both incubated with EPhAA over the course of 48 hours to determine the incubation time that yielded the best selectivity of labeling for this specific substrate.

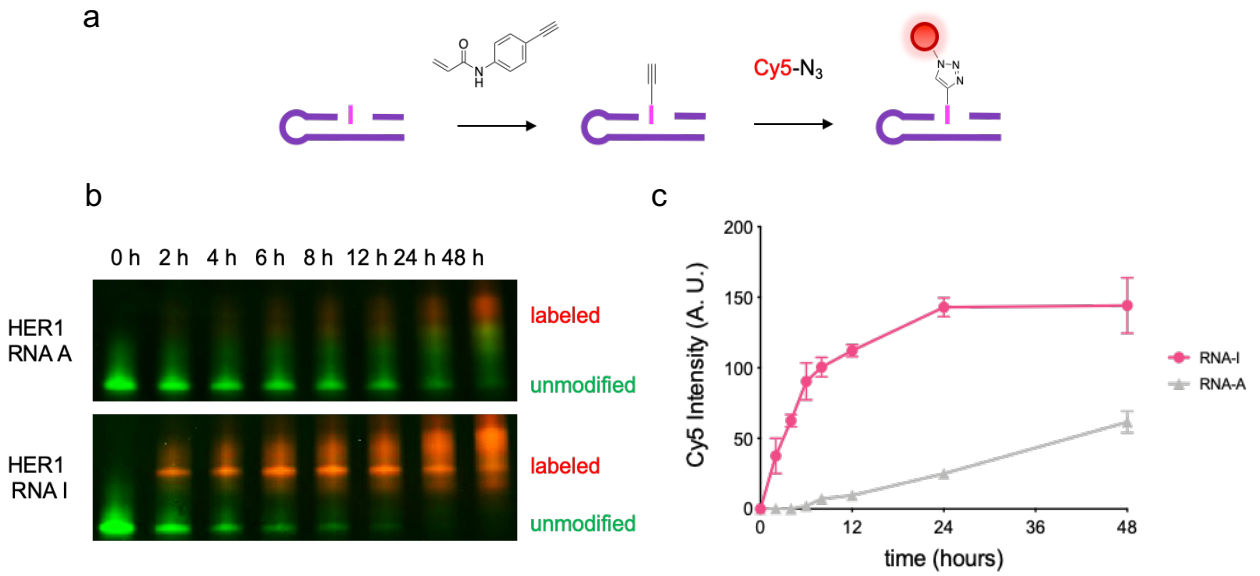


Figure 9. (a) HER1 RNA A and HER1 RNA I were first labeled with EPhAA over selected timepoints, and then using copper click chemistry, were covalently attached to a Cy5 moiety. (b) PAGE assay displaying the labeling over 48 hours of HER1 RNA A (top) versus HER1 RNA I (bottom). (c) Cy5 intensity of HER1 RNA A versus HER1 RNA I over 48 hours.

As shown in Figure 9, the best selectivity of HER1 RNA I over HER1 RNA A occurred at 6 hours of incubation with EPhAA. After 6 hours, other bases within HER1 RNA A, such as cytosine and guanosine, were being labeled by the reagent, causing false positive signal. To reduce this false positive signal, 6 hours was used as the reaction time for labeling inosine in all proceeding experiments.

2.6 HER1 RNA I Labeling Sensitivity and PAGE Shift Analysis

After determining ideal incubation time, the sensitivity of the labeling reagent was probed. Starting at 25 pmol of HER1 RNA I, each reaction was serial diluted 10-fold until 250 amol of HER1 RNA I was reached. The limit of detection by gel shift assay is showed in Figure 10.

Figure 10. PAGE assay displaying the limit of detection of EPhAA-labeled HER1 RNA I.

Thus, the smallest amount of labeled HER1 RNA I that can be detected via PAGE shift assay is about 2.5 pmol of RNA. This can be extrapolated to determine the amount of RNA substrate to use for a given ADAR-RNA reaction.

2.7 Mock HER1 RNA A Editing and PAGE Shift Analysis

To test the reagent's labeling efficiency in a mock editing environment, several labeling reactions were run, with varying amounts of HER1 RNA I spiked in to simulate different rates of editing. The total amount of RNA was constant for each reaction, with the ratio of HER1 RNA I to HER1 RNA A varying depending on the mock editing rate. As expected, the labeling corresponding to simulated editing rate linearly, verifying the labeling efficiency of EPhAA in a mixed RNA environment.

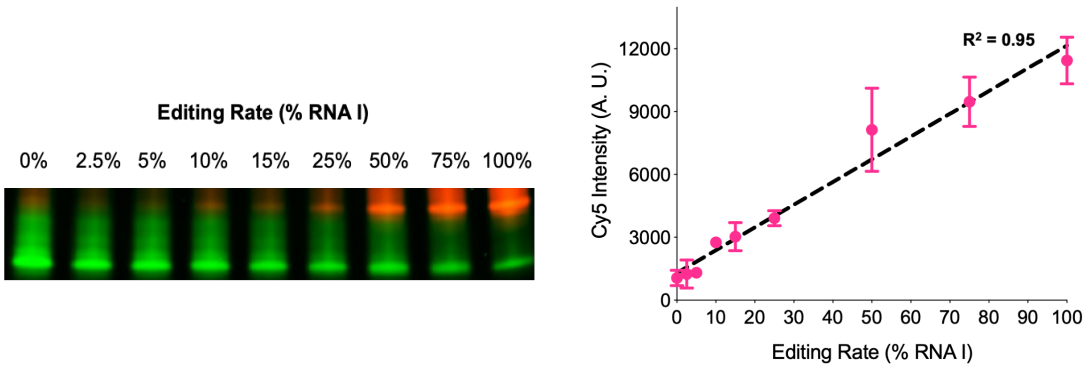


Figure 11. PAGE assay displaying increasing labeling as the ratio of HER1 RNA I : HER1 RNA A increases. A plot of Cy5 intensity versus editing rate shows their linear relationship.

2.8 ADAR1 Editing of HER1 and PAGE Shift Analysis

Validation of EPhAA's labeling efficiency in a mixed RNA environment allowed us to explore the reagent's efficiency under actual ADAR1 editing conditions. Two different ADAR1 enzymes were used: a wild type ADAR1 and a mutant where the glutamate in position 1008 was mutated to glutamine. Glutamate is an amino acid in the deaminase active site that is hypothesized to flip the adenosine base of interest out of the duplex and stabilize the orphan base across from the adenosine.[9] This mutation has been shown to increase editing rates above that of the wild type ADAR1, since glutamine stabilizes orphan cytidine better than glutamate.[9] Using this information, a proof of principle experiment was designed to showcase EPhAA's utility as a quick method to analyze ADAR editing with certain RNA substrates. HER1 RNA-A was incubated with either wild type or mutant ADAR1 to carry out *in vitro* A-to-I editing. Aliquots of RNA were taken at specified timepoints, purified, and EPhAA-labeled using the optimized protocol. Figure 12 validates the expected increased editing activity in the mutant ADAR1, compared to the wild type, and shows the utility of EPhAA as a readout for ADAR activity.

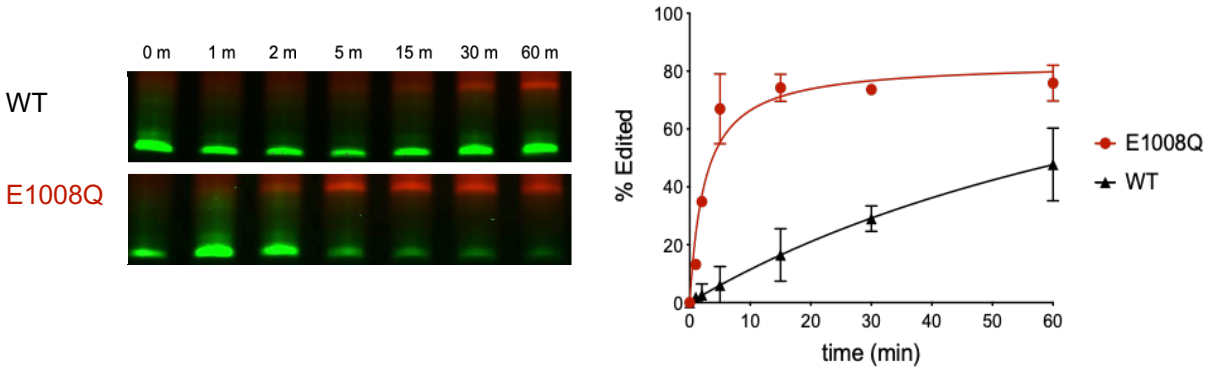


Figure 12. PAGE assay of ADAR1 WT and E1008Q mutant editing of HER1 RNA A over 60 minutes. Percent editing was quantified in ImageJ by densitometry of the labeled oligomer (red band) over total oligomer (sum of red band and green band) in each lane.

3. Conclusions and Limitations

A-to-I editing via site-selective ADAR enzymes is a powerful tool that can be exploited to synthetically modify RNA transcripts for therapeutic purposes. Currently, the editing preferences of ADAR1 and ADAR2 are not well characterized, and as such, it is difficult to predict where A-to-I editing will occur in a given RNA sequence. Currently, there are no chemical biology tools that can quickly, efficiently, and cheaply identify editing rates on specific small RNA substrates. Herein, a click-functionalized phenylacrylamide was developed to selectively label inosine to probe ADAR editing rates upon attaching a fluorescent moiety with copper click chemistry. The labeling system was tested on HER1, a previously characterized dsRNA substrate of ADAR1. The EPhAA-copper click system demonstrated enhanced editing rates of mutated ADAR1. This proof of concept experiment supports the use of this system to quickly and cheaply identify dsRNA substrates of ADAR1 or ADAR2 or to identify enhanced or suppressed editing rates of mutated ADAR enzymes.

While EPhAA demonstrates a useful way to probe ADAR reactivity, there are limitations to this labeling method. First and foremost, the side-reactivity of EPhAA with uridine and pseudouridine limits the scope of RNA substrates that can be probed, especially when probing how sequence context impacts ADAR site selectivity. The similarity in structure of inosine, uridine, and pseudouridine on the Watson-Crick-Franklin face makes difficult the task of finding an acrylonitrile derivative with moieties that make it more inosine selective. Nonetheless, this shortcoming can be avoided through by truncating RNA substrates.

Additionally, the limit of detection of this system can be prohibitive for certain RNA substrates that do not have high editing rates. As it stands, 2.5 pmol of EPhAA-labeled RNA is the lower limit of detection by PAGE shift analysis. For RNA substrates with editing rates of 2.5% or less, the amount of RNA required in an ADAR-activity assay would be at least 100 pmol. While this amount of RNA is achievable through synthesis or *in vitro* transcription, this is an important limitation to the labeling system.

4. Materials and Methods

Acrylamide Derivative Reactivity Panel

Acrylamide, *N*-phenyl acrylamide, and *N*-hydroxyethyl acrylamide were obtained from Sigma Aldrich Corporation (St. Louis, MO) and mPEG-acrylamide (MW 1000 g/mol) was obtained from Creative PEGWorks (Chapel Hill, NC).

In triplicate, 40 μmol of inosine, uridine, pseudouridine, adenosine, cytidine, or guanosine ribonucleoside (Sigma Aldrich Corporation, St. Louis, MO) were incubated in a 250 mM solution of acrylamide, *N*-phenyl acrylamide, *N*-hydroxyethyl acrylamide, or mPEG-acrylamide in 1:1 ethanol: 1M triethylammonium acetate pH 8.6. Reaction was adjusted to pH 8.6 and incubated for 24 hours at 70°C. At 0, 1, 2, 8, and 24 hours, an aliquot of crude reaction mixture was diluted 1:100 in 5% acetonitrile in PBS. Reversed-phase HPLC analysis was performed on an Agilent 1260 Infinity II system using a 4 μm, 150 x 4.6 mm Phenomenex Synergi Fusion-RP 80A C18 column. Acrylamide, mPEG-acrylamide, and *N*-hydroxyethylacrylamide reactions were analyzed using an isocratic mobile phase of 5:95 acetonitrile:water. *N*-phenylacrylamide reactions were analyzed using a linear mobile phase gradient from 5% to 45% acetonitrile in water over 15 minutes. All mobile phases contained 0.1% trifluoracetic acid. Percent conversion in each reaction was defined as the inosine peak area relative to unreacted inosine at the same time point without any reagent.

Synthesis of EPhAA

Unless otherwise noted, all starting materials were obtained from Sigma Aldrich Corporation (St. Louis, MO) and were used without further purification. Column chromatography was carried out using silica gel 60 (230–400 mesh). ^{1}H and ^{13}C NMR

chemical shifts are expressed in parts per million (δ). NMR spectra were obtained on a Varian INOVA 400 spectrometer in the Emory University NMR Research Center. Mass spectrum was obtained on an Agilent 6230 TOF LC/MS.

To a solution of 4-ethynylaniline (2.0 g, 17.1 mmol), 0.2 eq hydroxybenzotriazole (462.1 mg, 3.42 mmol), and 1.2 eq 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (3.934 g, 20.52 mmol) in anhydrous pyridine (30 mL) under N_2 was added 1.2 eq acrylic acid (1.41 mL, 20.49 mmol). The reaction was stirred at room temperature overnight. Consumption of 4-ethynylaniline was confirmed by TLC in 1:1 hexanes:ethyl acetate. The crude reaction mixture was diluted in 50 mL ethyl acetate and washed sequentially with water and brine. The aqueous layer was back-extracted twice with ethyl acetate, and the collected organic layer was dried with $MgSO_4$ and filtered. The organic layer was concentrated under reduced pressure and purified by column chromatography (1:1 hexanes:ethyl acetate). The purified product was concentrated under reduced pressure and dried *in vacuo* to yield 1.26 g (42%) of a salmon-colored powder. ^{1}H NMR (400 MHz, DMSO-d6) δ 10.29 (s, 1H), 7.65 (d, J = 8.5 Hz, 2H), 7.40 (d, J = 8.3 Hz, 2H), 6.40 (dd, J = 17.6, 10.0 Hz, 1H), 6.29 – 6.19 (m, 1H), 5.78 – 5.70 (m, 1H), 4.06 (s, 1H). ^{13}C NMR (400 MHz, DMSO-d6) δ 163.73, 139.97, 133.23, 132.85, 132.78, 132.04, 127.87,

119.60, 116.81, 83.96, 80.44. HRMS m/z (ESI) calculated for $C_{11}H_{10}NO$ $(M+H)^+$ 172.0762, found 172.0637.

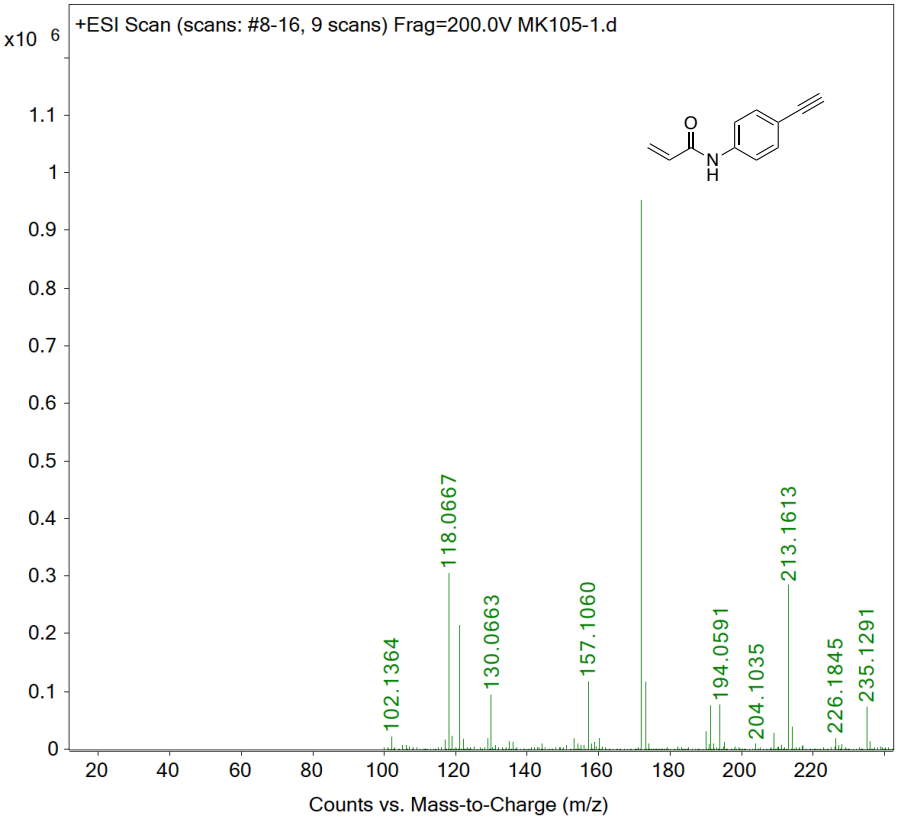


Figure 13. ESI-MS spectrum of *N*-(4-ethynylphenyl)acrylamide. Calculated $(M+H)^+$ for $C_{11}H_{10}NO$ 172.0762.

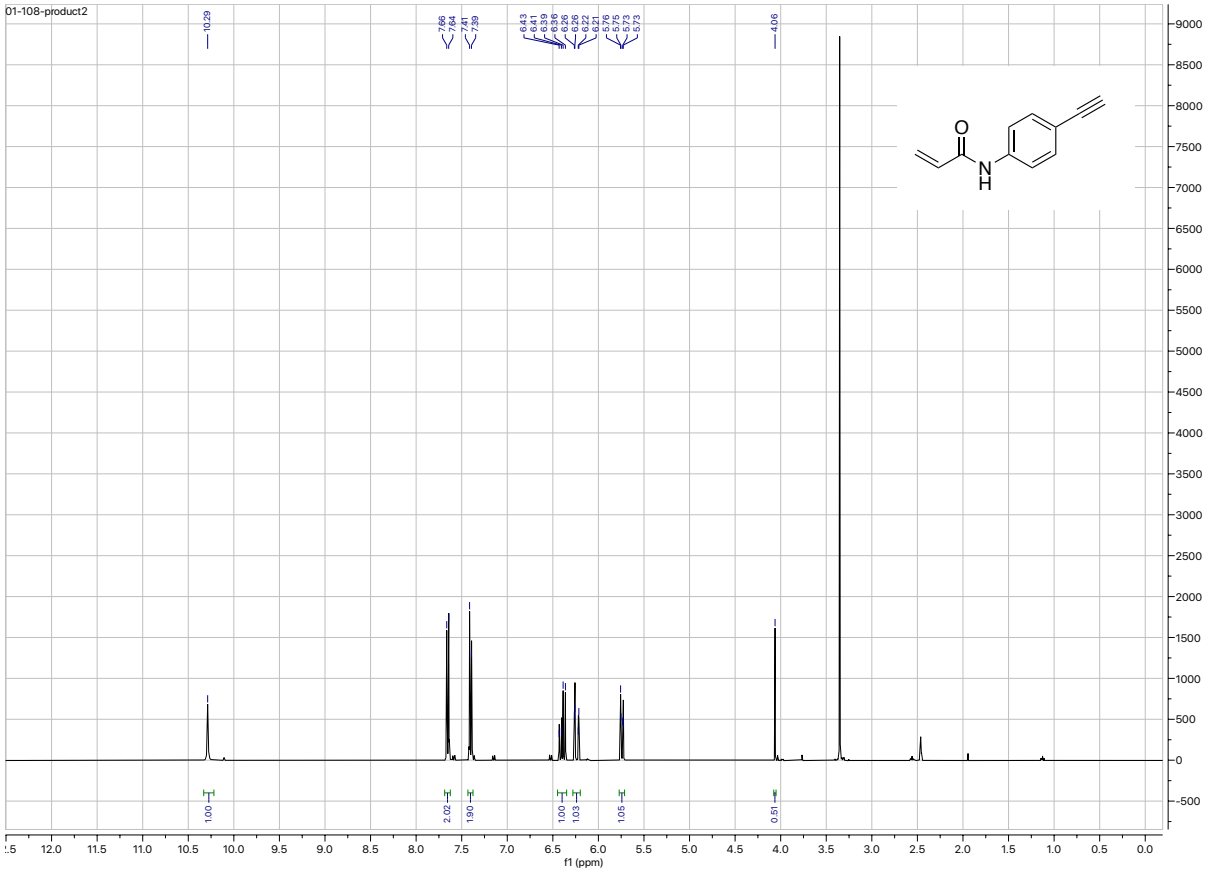


Figure 14. ^{1}H NMR spectrum of *N*-(4-ethynylphenyl)acrylamide (400 MHz, DMSO-d6).

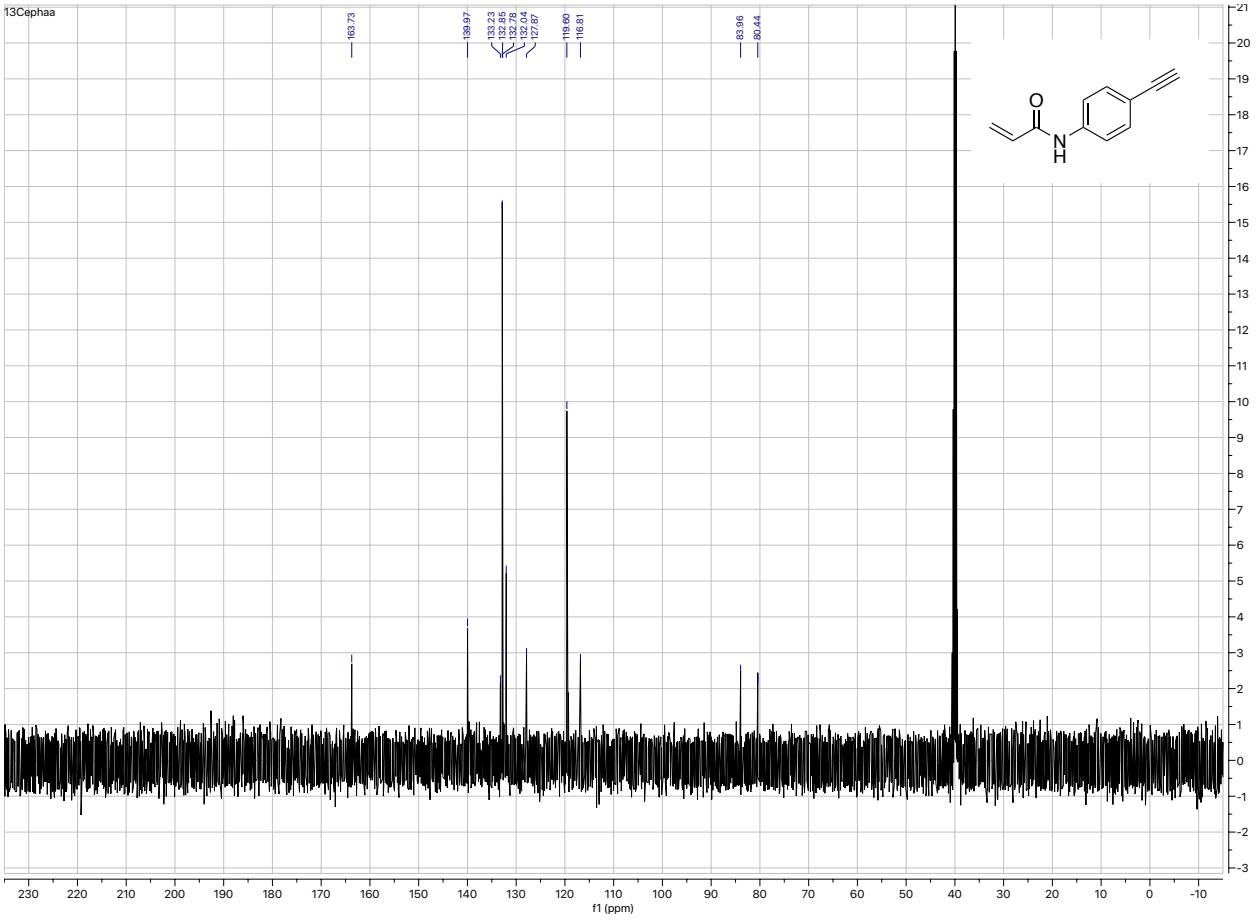


Figure 15. ^{13}C NMR spectrum of *N*-(4-ethynylphenyl)acrylamide (400 MHz, DMSO-d6).

Ribonucleoside Labeling and HPLC Analysis

Ribonucleosides inosine, pseudouridine, guanosine, adenosine, cytidine and uridine were purchased from Sigma Aldrich Corporation (St. Louis, MO). Labeling reaction mixtures were comprised of 50 mM ribonucleoside and 500 mM *N*-(4-ethynylphenyl)acrylamide reagent in 1:1 ethanol:reaction buffer. Phosphate buffered saline (PBS) was used for reactions from pH 6.5-7.5 and 1M triethylammonium acetate (TEAA) for pH 8.0-10.5. Reactions were incubated at 70 °C for the time periods indicated. Reversed-phase HPLC analysis was performed on an Agilent 1260 Infinity II system using a 4µm, 150 x 4.6 mm Phenomenex Synergi Fusion-RP 80A C18 column. Samples were prepared in a stationary phase solution of 5% acetonitrile in PBS. Reactions were analyzed using a linear mobile phase gradient from 5% to 45% acetonitrile in water over 15 minutes. All mobile phases contained 0.1% trifluoracetic acid.

RNA Oligonucleotides

RNA oligonucleotides were custom designed and order from the University of Utah DNA synthesis core facility (Salt Lake City, UT).

RNA-I: GGAAGAAGCAGCAGGAC **(I)** GAGCAGAACAGACCACGGA

RNA-U: GGAAGAAGCAGCAGGAC **(U)** GAGCAGAACAGACCACGGA

RNA-Ψ: GGAAGAAGCAGCAGGAC **(Ψ)** GAGCAGAACAGACCACGGA

HER1 RNA A: CCCGCCAACCCCGAGUU **(A)** GCGGGC

HER1 RNA I: CCCGCCAACCCCGAGUU **(I)** GCGGGC

RNA Oligonucleotide Labeling and PAGE Shift Analysis

In triplicate, 500 pmol of RNA-I, RNA-U, or RNA-Ψ was added to a 100 μL solution of 500 mM EPhAA in 1:1 ethanol : triethylammonium acetate buffer and adjusted to pH 8.6. Reactions were incubated at 70°C. At indicated timepoints, crude reaction mixture was diluted 1:10 with nuclease free water and ethanol precipitated. Purified RNA was reconstituted in nuclease free water and then CuAAC labeled using the Click-&-Go Plus Labeling Kit and 1μL of a 5 mM Cy5 picolyl azide solution in DMSO (Click Chemistry Tools, Scottsdale, AZ). Samples reacted at room temperature for 1 hour. Reactions were ethanol precipitated and purified RNA was diluted 1:1 with RNA loading dye. RNA was resolved on a 12% denaturing polyacrylamide gel then stained in TBE-SYBR Gold solution for 20 minutes before imaging on a GE Amersham Typhoon RBG scanner with Cy5, SYBR Gold, and FITC filters. Densitometric quantification of bands was performed using ImageJ software.

HER1 RNA A Versus HER1 RNA I Labeling and PAGE Shift Analysis

In duplicate, 500 pmol of HER1 RNA A or HER1 RNA I was added to a 100 μL solution of 500 mM EPhAA in 1:1 ethanol : triethylammonium acetate buffer and adjusted to pH 8.6. Reactions were incubated at 70°C. At indicated timepoints, crude reaction mixture was diluted 1:10 with nuclease free water and ethanol precipitated. Purified RNA was reconstituted in nuclease free water and then CuAAC labeled using the Click-&-Go Plus Labeling Kit and 1μL of a 5 mM Cy5 picolyl azide solution in DMSO (Click Chemistry Tools, Scottsdale, AZ). Samples reacted at room temperature for 1 hour. Reactions were ethanol precipitated and purified RNA was diluted 1:1 with RNA loading dye. Mixture was incubated at 70°C for 10 minutes to melt the hairpin structure. RNA was resolved on a 12% denaturing polyacrylamide gel then stained in TBE SYBR

Gold solution for 20 minutes before imaging on a GE Amersham Typhoon RBG scanner with Cy5, SYBR Gold, and FITC filters. Densitometric quantification of bands was performed using ImageJ software.

HER1 RNA I Labeling Sensitivity and PAGE Shift Analysis

In duplicate, 25 pmol, 2.5 pmol, 250 fmol, 25 fmol, 2.5 fmol, or 250 amol of HER1 RNA I was added to a 50 μL solution of 500 mM EPhAA in 1:1 ethanol:triethylammonium acetate buffer and adjusted to pH 8.6. Reactions were incubated at 70°C for 6 hours. Crude reaction mixture was diluted 1:10 with nuclease free water and ethanol precipitated. Purified RNA was reconstituted in nuclease free water and using the Click-&-Go Plus imaging kit (Click Chemistry Tools, Scottsdale, AZ), Cy5 picolyl azide reacted at room temperature for 1 hour with oligomers. Reactions were ethanol precipitated and purified RNA was diluted 1:1 with RNA loading dye. Mixture was incubated at 70°C for 10 minutes to melt the hairpin structure. RNA was resolved on a 12% denaturing polyacrylamide gel then stained in TBE SYBR Gold solution for 20 minutes before imaging on a GE Amersham Typhoon RBG scanner with Cy5, SYBR Gold, and FITC filters. Densitometric quantification of bands was performed using ImageJ software.

Mock HER1 RNA A Editing and PAGE Shift Analysis

In duplicate, varying mixtures of HER1 RNA I and HER1 RNA A were prepared in a 100 μL solution of 500 mM *N*-(4-ethynylphenyl)acrylamide in 1:1 ethanol:triethylammonium acetate buffer, adjusted to pH 8.6, and incubated at 70 °C for 6 hours. Mixtures were defined as follows:

Editing Rate (%)	pmol HER1 RNA I	pmol HER1 RNA A
100	0	100
75	25	75
50	50	50
25	75	25
15	85	15
10	90	10
5	95	5
2.5	97.5	2.5
0	100	0

After labeling, samples were then ethanol precipitated and resuspended in 20 µL nuclease-free water. RNA was then CuAAC labeled using the Click-&-Go Plus Labeling Kit and 1µL of a 5 mM Cy5 picolyl azide solution in DMSO (Click Chemistry Tools, Scottsdale, AZ). Reactions were incubated for 1 hour at room temperature, after which they were ethanol precipitated and resuspended in 10 µL nuclease-free water. 20 pmol of each sample was resolved on a 10% denaturing polyacrylamide gel and imaged with a GE Amersham Typhoon RBG scanner RGB scanner. Densitometric quantification of bands was performed using ImageJ software. Linear regression analysis was performed using GraphPad Prism 8 software.

ADAR1 Editing of HER1 and PAGE Shift Analysis

WT and E1008Q ADAR1 were expressed and purified by Leanna R. Monteleone from the Department of Chemistry, University of California, Davis.

In duplicate, 100 pmol of HER1 RNA A was added to a 20 µL solution of 1 µM WT ADAR1 or E1008Q ADAR1 in 10 mM Tris HCl pH 7.5, 8.5 mM EDTA, 0.001% Nonidet P-40, 3% glycerol, and 40.5 mM potassium glutamate. Reactions were incubated at

37°C for 60 min. At 0 min, 2 min, 5 min, 15 min, 30 min, and 60 min, an aliquot of the crude reaction mixture was taken and immediately the RNA was extracted using the Monarch RNA Cleanup kit (New England Biolabs, Ipswitch, MA). Samples were eluted in 10 μL of nuclease free water. 10 pmol of purified RNA was added to a 100 μL solution of 500 mM EPhAA in 1:1 ethanol:triethylammonium acetate buffer and adjusted to pH 8.6. Reactions were incubated at 70°C for 6 hours. Crude reaction mixture was diluted 1:2 with nuclease free water and ethanol precipitated. Purified RNA was reconstituted in nuclease free water and CuAAC labeled using the Click-&-Go Plus Labeling Kit and 1μL of a 5 mM Cy5 picolyl azide solution in DMSO (Click Chemistry Tools, Scottsdale, AZ). Samples reacted at room temperature for 1 hour. Reactions were ethanol precipitated and purified RNA was diluted 1:1 with RNA loading dye. Mixture was incubated at 70°C for 10 minutes to melt the hairpin structure. RNA was immediately resolved on a 10% denaturing polyacrylamide gel then stained in TBE SYBR Gold solution before imaging on a GE Amersham Typhoon RBG scanner with Cy5, SYBR Gold, and FITC filters. Densitometric quantification of bands was performed using ImageJ software.

References

(1) Lehninger, A. L.; Cox, M. M.; Nelson, D. L. *Principles of Biochemistry*, 6th ed.; W.H. Freeman: New York, NY, 2013.

(2) Rose, R. E.; Quinn, R.; Sayre, J. L.; Fabris, D. Profiling Ribonucleotide Modifications at Full-Transcriptome Level: A Step toward MS-Based Epitranscriptomics. *Rna* **2015**, *21* (12), 1361–1374. https://doi.org/10.1261/rna.054908.115.

(3) Angelova, M. T.; Dimitrova, D. G.; Dinges, N.; Lence, T.; Worpenberg, L.; Carré, C.; Roignant, J. Y. The Emerging Field of Epitranscriptomics in Neurodevelopmental and Neuronal Disorders. *Front. Bioeng. Biotechnol.* **2018**, *6* (APR), 1–15. https://doi.org/10.3389/fbioe.2018.00046.

(4) Schaefer, M.; Kapoor, U.; Jantsch, M. F. Understanding RNA. **2017**. https://doi.org/10.1098/rsob.170077.

(5) Li, X.; Xiong, X.; Yi, C. Epitranscriptome Sequencing Technologies: Decoding RNA Modifications. *Nat. Methods* **2016**, *14* (1), 23–31. https://doi.org/10.1038/nmeth.4110.

(6) Levin, J. G.; Guo, J.; Rouzina, I.; Musier-Forsyth, K. Nucleic Acid Chaperone Activity of HIV-1 Nucleocapsid Protein: Critical Role in Reverse Transcription and Molecular Mechanism. *Prog. Nucleic Acid Res. Mol. Biol.* **2005**, *80* (05), 217–286. https://doi.org/10.1016/S0079-6603(05)80006-6.

(7) Bass, B. L. RNA Editing by Adenosine Deaminases That Act on RNA. *Annu. Rev. Biochem.* **2002**, *71* (1), 817–846. https://doi.org/10.1146/annurev.biochem.71.110601.135501.

(8) Wang, Y.; Park, S. H.; Beal, P. A. Selective Recognition of RNA Substrates by ADAR Deaminase Domains. *Biochemistry* **2018**, *57* (10), 1640–1651. https://doi.org/10.1021/acs.biochem.7b01100.

(9) Wang, Y.; Havel, J.; Beal, P. A. A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1. *ACS Chem. Biol.* **2015**, *10* (11), 2512–2519. https://doi.org/10.1021/acschembio.5b00711.

(10) Phelps, K. J.; Tran, K.; Eifler, T.; Erickson, A. I.; Fisher, A. J.; Beal, P. A. Recognition of Duplex RNA by the Deaminase Domain of the RNA Editing Enzyme ADAR2. *Nucleic Acids Res.* **2015**, *43* (2), 1123–1132. https://doi.org/10.1093/nar/gku1345.

(11) Netzband, R.; Pager, C. T. Epitranscriptomic Marks: Emerging Modulators of RNA Virus Gene Expression. *Wiley Interdiscip. Rev. RNA* **2019**, No. August, 1–25. https://doi.org/10.1002/wrna.1576.

(12) Tajaddod, M.; Jantsch, M. F.; Licht, K. The Dynamic Epitranscriptome: A to I Editing Modulates Genetic Information. *Chromosoma* **2016**, *125* (1), 51–63. https://doi.org/10.1007/s00412-015-0526-9.

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

(14) Mannion, N. M.; Greenwood, S. M.; Young, R.; Cox, S.; Brindle, J.; Read, D.; Nellåker, C.; Vesely, C.; Ponting, C. P.; McLaughlin, P. J.; et al. The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA. *Cell Rep.* **2014**, *9* (4), 1482–1494. https://doi.org/10.1016/j.celrep.2014.10.041.

(15) Shafik, A.; Schumann, U.; Evers, M.; Sibbritt, T.; Preiss, T. The Emerging

Epitranscriptomics of Long Noncoding RNAs. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2016**, *1859* (1), 59–70. https://doi.org/10.1016/j.bbagrm.2015.10.019.

(16) Helfand, S. L.; Rogina, B.; Helfand, S. L.; Kapahi, P.; Benzer, S.; Goymer, P.; Pletcher, S. D.; Partridge, L.; Tatar, M.; Piper, M. D. W.; et al. Redirection of Silencing Targets by Adenosine-to-Inosine Editing of MiRNAs. **2007**, *315* (February), 1137–1141.

(17) Fiet, L. V. S.; Boudet, J.; Aalto, A.; Redis, R.; Sahin, Ö.; Bruijn, P. De; Wissen, L. Van; Freriks, S.; Turunen, J.; Platenburg, G.; et al. Therapeutic Oligonucleotides for Directing Site-Specific A-to-I Editing by Endogenous ADAR Enzymes : Computational and Empirical Optimization Scan The; 2018.

(18) Janne Juha Turunen , Leiden (NL) ; Antti Aalto , Leiden (NL) ; Bart Klein , Leiden (NL) ; Lenka Van Sint Fiet, L. (N.) ; J. A. G.; Boudet, L. (N.). Application Publication (10) Pub . No . : US 2019 / 0117177 A1, 2019.

(19) Sakurai, M.; Yano, T.; Kawabata, H.; Ueda, H.; Suzuki, T. Inosine Cyanoethylation Identifies A-to-I RNA Editing Sites in the Human Transcriptome. *Nat. Chem. Biol.* **2010**, *6* (10), 733–740. https://doi.org/10.1038/nchembio.434.

(20) Suzuki, T.; Ueda, H.; Okada, S.; Sakurai, M. Transcriptome-Wide Identification of Adenosine-to-Inosine Editing Using the ICE-Seq Method. *Nat. Protoc.* **2015**, *10* (5), 715–732. https://doi.org/10.1038/nprot.2015.037.

(21) Okada, S.; Ueda, H.; Noda, Y.; Suzuki, T. Transcriptome-Wide Identification of A-to-I RNA Editing Sites Using ICE-Seq. *Methods* **2019**, *156* (October 2018), 66–78. https://doi.org/10.1016/j.ymeth.2018.12.007.

(22) Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A Revolutionary Tool for Transcriptomics. **2010**, *10* (1), 57–63. https://doi.org/10.1038/nrg2484.RNA-Seq.

(23) Grosjean, H.; Droogmans, L.; Roovers, M.; Keith, G. Detection of Enzymatic Activity of Transfer RNA Modification Enzymes Using Radiolabeled TRNA Substrates. *Methods Enzymol.* **2007**, *425*. https://doi.org/10.1016/S0076-6879(07)25003-7.

(24) Knutson, S. D.; Ayele, T. M.; Heemstra, J. M. Chemical Labeling and Affinity Capture of Inosine-Containing RNAs Using Acrylamidofluorescein. *Bioconjug. Chem.* **2018**, *29* (9), 2899–2903. https://doi.org/10.1021/acs.bioconjchem.8b00541.

(25) Li, Y.; Göhl, M.; Ke, K.; Vanderwal, C. D.; Spitale, R. C. Identification of Adenosine-to-Inosine RNA Editing with Acrylonitrile Reagents. *Org. Lett.* **2019**, *21* (19), 7948–7951. https://doi.org/10.1021/acs.orglett.9b02929.

(26) Oivanen, M.; Kuusela, S.; Lönnberg, H. Kinetics and Mechanisms for the Cleavage and Isomerization of the Phosphodiester Bonds of RNA by Bronsted Acids and Bases. *Chem. Rev.* **1998**, *98* (3), 961–990. https://doi.org/10.1021/cr960425x.

(27) Fox, J. J.; Wempen, I.; Hampton, A.; Doerr, I. L. Thiation of Nucleosides. I. Synthesis of 2-Amino-6-Mercapto-9-β-D-Ribofuranosylpurine (“Thioguanosine”) and Related Purine Nucleosides. *J. Am. Chem. Soc.* **1958**, *80* (7), 1669–1675. https://doi.org/10.1021/ja01540a041.

(28) Herbert, A.; Rich, A. The Role of Binding Domains for DsRNA and Z-DNA in the in Vivo Editing of Minimal Substrates by ADAR1. *Proc. Natl. Acad. Sci. U. S. A.*

2001, *98* (21), 12132–12137. https://doi.org/10.1073/pnas.211419898.