THE UNIVERSITY OF CHICAGO

SYNTHETIC BIOLOGY APPROACHES TO STUDY AND CONTROL THE TRANSCRIPTOME

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOCHEMISTRY AND MOLECULAR BIOPHYSICS

BY
SIMONE RAUCH

CHICAGO, ILLINOIS
MARCH 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Bryan Dickinson for his guidance and support throughout the past several years. Bryan always encouraged and motivated me with his enthusiasm for synthetic and chemical biology. He always had an open door to discuss any and all scientific successes or failures I encountered and provided me with many opportunities to collaborate within or outside our lab and to present my work at conferences. In retrospect, I am especially grateful for the times he 'forced' me to repeatedly practice important public talks, something I could not appreciate at the time. This dissertation would not have been possible without Bryan's thoughtful input and constant support.

I was extremely fortunate to work with such an amazing group of people throughout my PhD. I would like to thank all members of the Dickinson group for creating a collaborative and fun environment to come to work to every day. In particular, I would like to take this opportunity to thank the individuals who played a crucial role during my PhD. When I first joined the lab, it was Dr. Jinyue Pu who patiently introduced me to and familiarized me with many techniques that were vital for my work. Throughout the years in the Dickinson lab, Dr. Huiqing (Jane) Zhou introduced me to biochemical techniques and acted as one of my mentors and role models in the lab. Jane always went above and beyond to give helpful advice and suggestions in response to any scientific or life question I could possibly have. It was often Jane's thoughtful guidance and input that motivated me to keep trying and pushing projects to their completion. I would also like to thank the people who directly worked on these projects with me: Emily He, Michael Srienc, Zijie (Scott) Zhang, and Dr. Krysten Jones. Emily and Michael worked with me on the initial development of CRISPR-Cas-inspired RNA-targeting system (CIRTS) and helped with cloning, protein expression, and luciferase screens. Further, I would like to thank Scott for his help with all bioinformatic analysis and the repeated patient statistics explanations that even a biochemist could understand. Special thanks to Krysten, who joined me on the CIRTS project last year, and

has been a great collaborator. Krysten conducted all the mouse experiments detailed in Chapter 4 and is now working on improving direct CIRTS delivery with lipid nanoparticles. Thanks also goes to Kaitlin Kentala, and Julia Zinkus-Boltz for helpful discussions, and Saara-Anne Azizi and Dr. Somayeh Ahmadiantehrani for invaluable help with editing this thesis.

I would like to sincerely thank my thesis committee, Prof. Chuan He, Prof. Jingyi Fei, and Prof. Joseph Piccirilli, for their invaluable input and support throughout the entirety of my PhD. Their thoughtful comments on my research at all stages of my PhD have benefited my project greatly. I would also like to thank the generous support from the University of Chicago with the Harper Dissertation Fellowship.

Finally, I am profoundly grateful to my friends and family for their continued support and encouragement through the highs and lows of graduate school. I would also like to thank the entire BMB cohort for making this a fun and vibrant experience.

ABSTRACT

Gene expression regulation at the RNA level has emerged as a key point of control of information flow through the central dogma. RNA stability, trafficking, and translation efficiency are heavily regulated by canonical RNA processing mechanisms such as capping, splicing, and polyadenylation as well as chemical RNA modifications. With an ever-increasing number of reported regulatory layers, there is an increased need for new tools to disentangle these complex regulatory networks in a site- and transcript-specific manner within their endogenous environment. Additionally, RNA has become an attractive target for therapeutic intervention due to its rapid turnover and short lifetime allowing for multigenic targeting and higher safety profiles compared to DNA-targeting approaches. In this thesis, we report the development and optimization of programmable RNA effectors to both study and control gene expression regulation at the RNA level. We first developed programmable RNA "reader" proteins, which allow us to interrogate the effects of specific regulatory proteins on single transcripts. Our first-generation programmable epitranscriptomic readers rely on RNA-targeting CRISPR/Cas systems as the programmable delivery vehicle. While powerful, these microbially derived CRISPR/Cas systems are large and present immunogenicity issues when applied in non-native human contexts, limiting their potential future therapeutic applications. Therefore, we next developed a second-generation platform for engineering programmable RNA effector proteins that are significantly smaller than the CRISPR/Cas systems and are built entirely from human parts. Our CRISPR/Cas-inspired RNA targeting system (CIRTS) involves mining the human proteome for functional domains and combining them to create programmable RNA effector proteins. Next, we engineered a small molecule-inducible CIRTS biosensor system for temporal control of CIRTS activity in cells and *in vivo*. Finally, we set out to develop a second-generation CIRTS with improved efficacy, novel delivery methods, and the first steps towards preclinical applications. Taken together, CIRTS, which can induce transcript-specific RNA decay, protein production and RNA edits, will find applications in synthetic biology to control the genetic information flow at the RNA level.

LIST OF PUBLICATIONS BASED ON WORK IN THIS THESIS†

1. **Rauch, S.**, and Dickinson, B. C. Programmable RNA Binding Proteins for Imaging and Therapeutics. *Biochemistry*, **2018**, *57*, 363-364.

2. **Rauch, S.**, He C. and Dickinson, B. C. Targeted m^6A reader proteins to study epitranscriptomic regulation of single RNAs. *J. Am. Chem. Soc.*, **2018**, *140*, 11974-11981.

3. **Rauch, S.**, and Dickinson, B. C. Targeted m^6A reader proteins to study the epitranscriptome. *Methods Enzymol.*, **2019**, *621*, 1.

4. **Rauch, S.**, He, E., Srienc, M., Zhou, H, Zhang, Z., Dickinson, B. C. Programmable RNA-guided RNA effector proteins built from human parts. *Cell*, **2019**, *178*, 122.

5. Zhou, H., **Rauch, S.**, Qing, D., Cui, X., Zhang, Z., Nachtergaele, S., Sepich, C., He, C., Dickinson, B.C. Evolution of a reverse transcriptase to map N^1-methyladenosine in human messenger RNA. *Nat. Methods*, **2019**, *16*, 1281-1288.

6. **Rauch, S.**, and Dickinson, B. C. Expanding the Chemical Scope of RNA Base Editing. *Biochemistry*, **2019**, *58*, 3555-3556.

7. **Rauch, S.***, Jones, K.A.*, and Dickinson, B. C. Small molecule-inducible RNA-targeting systems for temporal control of RNA regulation *in vivo*. *Submitted*, **2020** *on Chemrxiv* DOI: 10.26434/chemrxiv.11673711.v1.

†. The following chapters of this dissertation contain sections and figures adopted from the listed publications with modifications. Chapter 1: publications 1, 4 & 6; Chapter 2: publication 2; Chapter 3: publication 4; Chapter 4: publication 7.

⋆. Denotes equal contribution

CHAPTER 1

INTRODUCTION

1.1 Synthetic biology to control gene expression at the RNA level

Ribonucleic acid (RNA) lies at the center of the central dogma of molecular biology and connects genetic information to protein products. Genes encoded in deoxyribonucleic acid (DNA) are transcribed into RNA before being translated into proteins that carry out a wide range of cellular functions (**Figure 1.1**). In recent years, RNA, which had previously been thought of as a simple intermediate of the central dogma, has emerged as a key regulator of many cellular functions. A transcribed RNA molecule is subject to a complex series of processes that precisely regulate its lifetime, composition, and translation into protein. These processes form a highly interconnected regulatory network, rendering functional studies of individual components in their endogenous context challenging, but providing synthetic biologists with opportunities to create technologies both to study biological functions and even for therapeutic intervention via manipulation of gene expression (**Figure 1.1**).

Synthetic biology approaches to design and construct novel biological components or repurpose existing components for desired purposes have the potential to revolutionize biotechnology and medicine[1]. For instance, researchers have engineered entire organisms for fuel and increased crop production[2]. Genome engineering, a subarea of synthetic biology, involves combining and re-engineering individual molecular building blocks, instead of an entire organism, to serve a particular role[3]. For example, naturally occurring RNA-targeting proteins, such as the Pumilio family (PUF) proteins, have been converted into programmable RNA-targeting moieties by altering their amino acid sequence to bind specific RNA sites of interest[4,5]. Such programmable RNA-targeting proteins have opened up new opportunities to study RNA regulation in its endogenous context by providing a

simple method of transcript- and site-specific delivery of regulatory proteins.

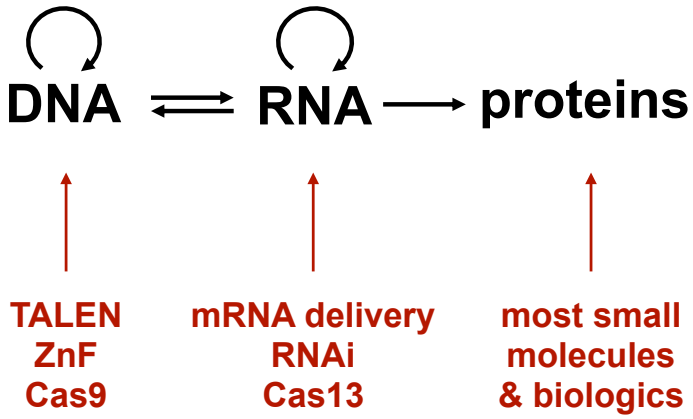

Figure 1.1 Tools to study and control gene expression along the central dogma. Synthetic biology tools have been developed to manipulate gene expression for biological studies or therapeutic intervention at every step along the central dogma. At the protein level, small molecules have been developed and deployed for decades. Only recently has the focus shifted towards gene expression control at the nucleic acid level with the discovery of the RNA interference pathway and programmable DNA/RNA-targeting proteins.

The discovery of the CRISPR/Cas family of proteins has further simplified genome engineering by providing a protein scaffold that is even more readily programmed to target DNA or RNA molecules of interest using base pair complementarity[6–9]. Synthetic biology approaches provide the necessary technology to meet an ever-increasing interest in and demand for tools to study nucleic acids in their endogenous context.

While a multitude of site-specific DNA targeting and effector systems have been developed and have allowed for breakthrough discoveries in DNA regulation, studies in RNA biology have been hindered by a lack of available programmable effector proteins. This chapter will briefly introduce gene expression control at the RNA level and highlight the regulatory pathways relevant to subsequent chapters in this thesis. It will then introduce how RNA regulation is currently studied in live mammalian cells and discuss the challenges and opportunities in the field before detailing the scope of this thesis.

1.1.1 Gene expression and RNA regulation

Post-transcriptional regulation refers to the control of gene expression at the RNA level, after transcription of DNA and before translation into protein[10]. Gene expression regulation at the RNA level has only recently been discovered as both crucial and complex. RNA that has been transcribed in the nucleus undergoes a wide variety of modifications until it reaches its site of translation in the cytoplasm. Canonical regulatory processes such as capping, splicing, polyadenylation, and trafficking occur in the nucleus and can regulate the lifetime and function of RNA[11–14]. More recently, a diverse set of chemical modifications has been found to alter stability, transport, protein binding, and translation of RNA[15–18] (**Figure 1.2**). Together all these regulatory processes, which directly act on messenger RNA (mRNA), influence almost all levels of cellular function. These findings substantiate the role of RNA as a dynamic regulator of cellular processes and demonstrate an increased need for tools to monitor and control gene expression at the RNA level in a site- and transcript-specific manner in the endogenous environment in live cells.

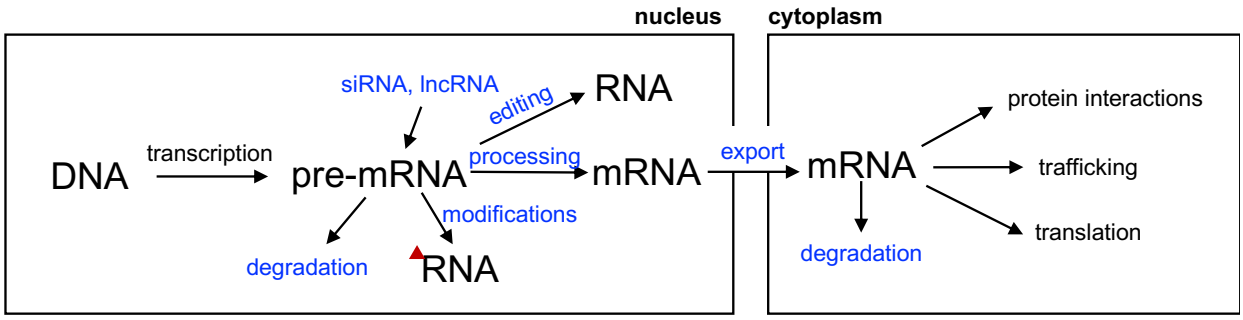

Figure 1.2 Transcriptome regulation.
Mammalian gene expression at the RNA level is heavily regulated by a diverse set of interconnected processes. The disentangled nature of this regulatory layer requires the development of novel tools to study individual regulatory pathways within an endogenous context.

RNA regulation

In addition to well-known capping, splicing, and polyadenylation processing steps that

occur in the nucleus[11–14], individual nucleobases can be altered by addition of chemical modifications or editing via deamination in the nucleus and cytosol[18,19]. Chemical RNA modifications are found in ribosomal, transfer, and cellular RNAs (rRNA, tRNA and mRNA/lncRNA). It is well-known that rRNA is commonly modified by pseudouridines and 2'O-methylation for increased stability[20]. tRNA is the most prevalent modified RNA species and is frequently altered to ensure tRNA discrimination, translation fidelity, and stability[18,21]. In eukaryotic mRNA/lncRNA, the most common modifications of transcribed polymers are N^6-methyladenosine (m^6A), N^1-methyladenosine (m^1A), pseudouridine (ψ) and 2'O-methylations[22,23] (**Figure 1.3A**). In eukaryotes, m^6A is the most prevalent internal mRNA modification. On average, each mRNA transcript contains approximately three m^6A modifications as compared to 0.3-1 other modifications per transcript[18,22,23]. m^6A has also been shown to play a crucial role in cellular development and differentiation[16,18,24]. In addition to the discovery of the 'writers' and 'erasers' that introduce and remove m^6A, respectively, several 'reader' proteins have been discovered, which link site-specific RNA modifications to particular regulatory functions within the cell[25–27] (**Figure 1.3B**). One of the most well-characterized m^6A reader protein families is the YTH521-B homology domain containing family (YTHDF) of proteins[25]. These cytoplasmic reader proteins include YTHDF1/2/3, which have been implicated in regulating translation efficiency through interactions with eukaryotic translation initiation factors (eIFs) and the ribosome, and also in RNA degradation by recruitment of the RNaseP pathway or CCR4-NOT deadenylation machinery[18,24,28,29].

RNA editing in mammalian cells most commonly occurs by enzyme-mediated deamination of ribonucleobases[19]. The two major deamination pathways are adenosine-to-inosine (A-to-I) editing and cytidine-to-uridine (C-to-U) editing[30–33]. In the first pathway, adenosine is hydrolytically deaminated to inosine (which is then translated as guanosine by preferentially base pairing with cytidine) by the Adenosine Deaminase Acting on RNA (ADAR) family of proteins[34,35]. The ADAR family of double-stranded RNA (dsRNA) editing enzymes includes ADAR1/2/3, with ADAR1/2 showing active deamination in cells. A-to-I

edits constitute almost 90% of all editing events in cells[36]. The deamination of cytidine to uracil is catalyzed by a second class of RNA deamination enzymes, the Apolipoprotein B mRNA Editing Catalytic Polypeptide-like (APOBEC) family of proteins[37,38]. RNA editing cannot only alter the protein coding region by introducing codon-changing mutations, it can also affect other cellular processes including splicing and microRNA (miRNA) binding[39,40]. In addition to serving as a crucial regulator of many cellular processes, RNA editing can act as a protective mechanism upon viral infection. Viruses that contain an RNA genome or replicate through an RNA intermediate often produce long double-stranded (ds)RNA fragments. Because dsRNA fragments rarely occur in eukaryotic cells, they are easily recognized as foreign by cellular machinery and hyperedited by ADARs/APOBECs for rapid degradation[41,42]. Finally, aberrant RNA editing has also been linked to cancer development and progression[43]. Taken together, these findings underscore the importance of tightly regulated RNA editing in mammalian cells.

Translation

After undergoing heavy processing in the nucleus and being exported into the cytoplasm, mature mRNAs are translated into protein. Translation is divided into three distinct stages: initiation, elongation, and termination[44,45]. For the purpose of this thesis, we will focus on initiation. During initiation of eukaryotic translation, at least 10 proteins termed eukaryotic initiation factors (eIFs) bind to a transcript's 5' and 3' untranslated regions (UTRs). Binding of the eIF complex stabilizes the formation of the preinitiation complex around the start codon. The 40S ribosomal subunit, the initiator tRNA, and several translation initiation factors form the 43S preinitiation complex (PIC)[46]. Then, eIF4F, which consists of eIF4A, eIF4E, and eIF4G, helps to connect the transcript to be translated to the 5'cap structure to initiate transcript scanning[47]. Once the start codon is reached, release of eIF1 triggers complete ribosome assembly and translation is initiated. Elongation of the polypeptide chain occurs through a cycle of entering tRNAs, peptidyl bond formation, and ribosome translation and continues until the ribosome encounters a stop codon, which

triggers the release of the completed protein product from the ribosome. mRNAs can be simultaneously translated by several ribosomes at a time[48].

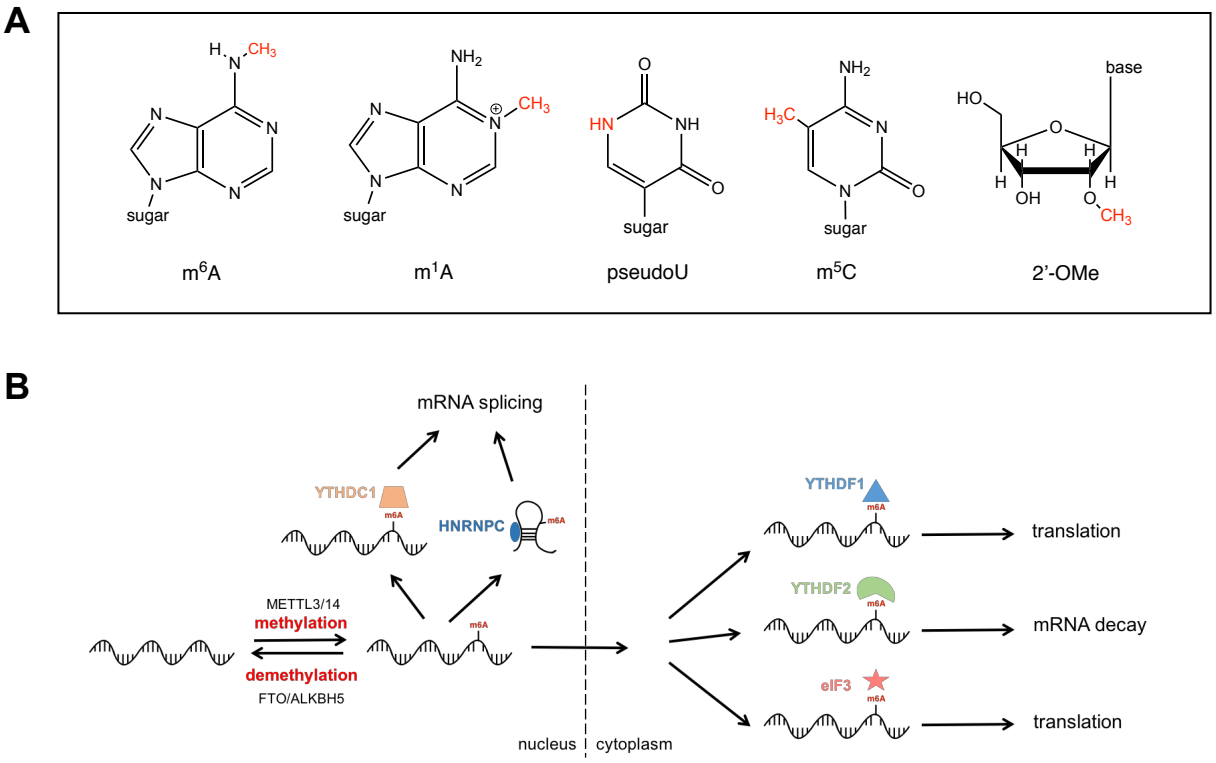

Figure 1.3 Chemical RNA modifications and their regulatory network.
(A) Overview of the most abundant eukaryotic chemical RNA modifications. **(B)** The most prevalent internal mRNA modification N^6-methyladenosine is dynamically installed and erased by 'writer' and 'eraser' proteins and is linked to cellular processes through binding of 'readers'.

In eukaryotic systems, gene expression is heavily regulated at the RNA level. Starting the moment RNA is transcribed, its stability, composition, location, and translation are controlled by a variety of interconnected cellular processes, affecting both the timing and level of protein biosynthesis. Many of these regulatory pathways are mediated by multi-protein complexes and require their endogenous interaction partners for functional studies, rendering study of these mechanisms in traditional transgenic model cell lines challenging. Synthetic biology approaches offer the opportunity to develop a whole new set of tools to accurately interrogate their functions in cells.

1.1.2 Tools to study RNA in cells

As it becomes more and more apparent that RNA provides a critical regulatory layer to the genetic information flow, methods to both study and control the spatial and temporal dynamics of RNAs in physiological conditions are sorely needed. Here, I will introduce the tools and technologies that are currently used to study RNA regulation.

Protein-based technologies

To study the spatiotemporal aspects of a target RNA or the regulatory role of an RNA binding protein (RBP) on a target RNA in live cells, the introduction of an exogenous sequence to the RNA of interest is often required. In general, RNA regulatory proteins possess two main functions: binding of RNA and eliciting a particular cellular response. RNA functional tethering assays are often used to decouple these two functions and to simplify cellular assays[49]. In such assays, an RNA transcript of interest is fused to a well-characterized RNA hairpin sequence (e.g. MS2 hairpin, boxB, etc.) on the 3'UTR of the transcript. The hairpin is then recognized by its native binding partner (MS2 coat protein (MCP) for the MS2 hairpin or lambda peptide for the boxB hairpin), which can be fused to a protein of interest whose function is unknown[50]. Tethering assays have aided the functional annotation of many RNA regulatory proteins including the role of the m^6A reader protein YTHDF1[24]. Eight consecutive MS2 hairpins were fused to luciferase mRNA and co-transfected with a YTHDF1-MCP fusion protein. In the presence of YTHDF1-MCP, an increase in luciferase levels was detected, leading the authors to conclude that YTHDF1 must play a role in translation of proteins[24]. However, although powerful, classic tethering assays rely on the introduction of overexpressed and tagged RNA, which can impact regular cellular functions and have limited abilities to elucidate native functions in cells.

Sequence-specific, programmable RNA binding proteins have expanded our ability to study endogenous RNA transcripts. The native Pumilio homology domain (PumHD) protein family has been engineered to target short RNA sequences of interest by combining

consecutive modules (termed Pumby for Pumilio-based assembly) of varying composition with high specificity[4,5]. Pumbys have been used to monitor translational activity of specific endogenous RNA transcripts in live cells via fusion of two Pumbys to a spit luciferase reporter system. Additionally, specific transcripts can be targeted for increased translation activity by fusing Pumby to the translation initiation factor eIF4e. Engineered Pumilio family-based protein technologies have opened up novel opportunities to study endogenous RNA transcripts[4]. However, Pumbys have to be re-engineered and adjusted for every transcript of interest, which limits the practicality of this method. Overall, advances in this field have been limited by a lack of easily programmable RNA effector proteins.

The discovery of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas system, which evolved as a prokaryotic immune defense mechanism, has transformed our ability to study and manipulate cellular nucleic acids site-specifically[6–8]. A key advantage of CRISPR-Cas systems, as compared to previous methods, such as zinc finger proteins and TALE nucleases for DNA-targeting[5,51–54] or Pumby's for RNA targeting[4], is that they are easily programmed to target virtually any locus of interest. The CRISPR-Cas system is a ribonucleoprotein complex that uses base pair interactions of a displayed guide RNA (gRNA) to interact with a target nucleic acid sequence. The simple nature of base pair-guided targeting creates the possibility to program systems to interact with a defined nucleic acid sequence by simply changing the nucleic acid sequence on the guiding strand. Programmable RNA-targeting tools analogous to the Cas9 DNA-targeting systems hold great promise for studying the mechanisms of RNA regulation and for therapeutic applications. The Cas9 protein family, which natively targets DNA, has been repurposed to act as a programmable nuclease-inactive RNA targeting moiety[55–57]. A protein family related to Cas9, the Cas13 protein family, was then shown to natively target RNA[9]. More recently, an additional class of smaller single-stranded RNA targeting Cas proteins (Cas12g proteins) that exhibit ssRNA and ssDNA collateral cleavage were discovered[58]. Aside from providing a powerful method to degrade target RNAs using the nuclease activity inherent to Cas13 systems, catalytically inactive dead Cas13 proteins (dCas13) have

been used to image RNA by fusions to GFP[59], to edit RNA by delivering an ADAR A-to-I[60] or an evolved C-to-U editing enzyme[61], and to modulate splicing by delivering hnRNPa1 to target transcripts of interest[62]. Additionally, our group recently adopted the dCas13 system to deliver truncated N^6-methyladenosine (m^6A) binding proteins ("readers") to specific sites in the transcriptome, yielding new tools to study and control RNA regulation[63].

Nucleic acid-based technologies

Aside from protein-based technologies, several nucleic acid tools have been developed to study and control gene expression at the RNA level. For example, the native RNA interference (RNAi) pathways have been exploited to knockdown proteins of interest in cellular systems[64]. In its endogenous role, the RNAi pathway lowers mRNA stability and translation through processing of small-interfering RNA (siRNA) or microRNAs (miRNAs) that bind to a target transcript[65]. The resulting short double-stranded RNA fragments recruit the RISC complex and Argonaute proteins to degrade transcripts through cleavage or deadenylation. In the laboratory, these pathways have been used extensively for knockdown of proteins of interest[65,66]. To achieve the desired knockdown, exogenous small hairpin RNAs (shRNAs) or siRNAs targeting specific transcripts are delivered into cells[66]. This method has been particularly useful to unveil protein functions in mammalian systems and has now found its way into therapeutic applications with two FDA-approved RNAi drugs[67,68]. For discovery-driven science applications, proteins can be knocked down to discern their function, whereas in therapeutic applications, diseases that are caused by increased protein production can be treated by shRNA or siRNA administration. Additionally, long non-coding RNA (lncRNA) tools, termed SINEUP as they require an invSINEB2 element (SINE) to UP-regulate translation, have been developed as enhancers of protein production[69–71]. In SINEUP, natural lncRNAs known to induce translation are fused to a guiding region complementary to an RNA transcript of interest, resulting in increased protein production of both reporter and endogenous transcripts.

Besides influencing RNA levels directly, nucleic acids have also been developed as

recruiting moieties for RNA editing in order to overcome off-target effects caused by overexpressing hADAR. For example, chemically modified[72] or non-modified[73] recruiting oligonucleotides can redirect endogenous ADAR to a target site on an RNA of interest. While chemical modifications increase the stability and therefore the performance of these approaches, chemically modified oligonucleotides can also cause immunogenic responses when administered repeatedly, potentially limiting their therapeutic applications. Consequently, Qu *et al.* developed engineered ADAR-recruiting RNAs (arRNAs)[73], which are not chemically modified and can therefore be genetically encoded. Expression of arRNAs resulted in efficient A-to-I editing in human primary cell lines and several disease-relevant contexts, including restoration of transcriptional regulatory activity of mutant *TP53*^{W53X} and recovery of IDUA activity in Hurler syndrome, emphasizing the therapeutic potential of this approach.

1.1.3 Delivery of tools to study and control RNA regulation

One major challenge in the field of RNA technologies is the delivery of such tools to the site of interest. Whether for biological studies or therapeutics, protein- and nucleic acid-based targeting moieties have to be introduced into the native system and reach their target of interest to elicit their desired function. For mammalian cell culture models, technologies can be delivered in plasmid form and conventional lipofection-based transfection reagents can be used to introduce them to cells[74,75]. Lipofection methods rely on cationic liposome formulations that complex with negatively charged nucleic acid molecules to overcome electrostatic repulsions in the cell membrane[75]. More difficult-to-transfect or primary cell types can be electroporated to introduce exogenous DNA[76]. While convenient and efficient for cell culture, DNA-based delivery by lipofection or electroporation are less efficacious for experiments that require high efficiency transfection (e.g. library generation) or for *in vivo* models. An alternative, highly efficient delivery method is viral delivery. The two most commonly used viral delivery systems are lentivirus[77,78]

and adeno-associated virus (AAV)[79]. Both approaches rely on encoding the RNA or protein tool into the viral genome. Lentiviral approaches are commonly applied for library generation and some therapeutic applications in which permanent incorporation into the cellular genome is beneficial[78]. AAV-based techniques, which are more commonly used for therapeutic delivery, are known to be less immunogenic and less frequently integrate their genomic information (or in this case the encoded protein/RNA) into the host's genome when compared to lentiviruses[79,80]. Both forms of viral delivery have the advantage of programmable tissue specificity. For example, there are currently 11 well-characterized AAV serotypes with different tissue specificities, with additional synthetic serotypes constantly being developed[81,82]. All these delivery methods are centered upon delivery of the tool encoded in DNA. While this approach is convenient for many applications, it ultimately relies on up to two amplification steps (transcription and, for protein technologies, translation) that determine the amount produced, making proper dosing of these tools or therapeutics challenging. Furthermore, both aforementioned viral delivery methods rely on replacing the viral genome with the technology to be delivered and have therefore strict packaging size limits (4.7 kb for AAV and 6kb for lentivirus)[83], rendering the delivery of larger protein or RNA technologies challenging. To overcome these challenges, direct RNA or protein delivery methods have been developed[84,85]. Dosing and delivery of RNA tools have been simplified greatly using approaches that limit the amplification steps to transcription or to no amplification at all. RNA and protein tools can be delivered by special lipofectamine reagents or by addition of a charged peptide that neutralizes charge and facilitates cellular uptake analogous to DNA delivery[85,86]. For more complex delivery applications, lipid nanoparticle (LNP) or other nanomaterial capsules have been developed. LNPs are the most clinically advanced non-viral delivery approach. They have been shown to be efficient delivery moieties for nucleic acids as well as peptides in cells and animal models[87,88]. Cargo is encapsulated in an automated flow cell system that combines the cargo solution with a solution containing different lipid compositions mimicking low-density lipoproteins that facilitate cellular uptake via endogenous machinery. Upon entering cells,

the pH-sensitive lipid capsules break apart, releasing the cargo into cells[87]. LNP production and delivery have been applied for therapeutic use with the first FDA-approved RNAi drug Patisiran[89]. Key challenges to address for direct RNA and protein delivery are the stability and the packaging constraints of the cargo. In the absence of one or two amplification steps, the delivered cargo has to be stable enough to survive and function in the cellular environment. Exogenously delivered RNA, which is known to be rapidly turned over, often has to be heavily chemically modified and end-capped to protect it from nuclease degradation and increase its half-life in cells[90,91]. One key challenge for packaging of cargo, regardless of the packaging vesicle, is the size of the cargo. Despite optimization efforts, the packaging limit for direct protein delivery lies around approximately 100 kDa. Although larger cargoes are not impossible to package, their packing efficiency is low and inconsistent[92]. Therefore, technologies for biological studies and therapeutic applications need to be stable enough to survive cellular conditions and small enough in size to be packaged efficiently and reproducibly.

1.1.4 Current challenges and opportunities

With an increasing understanding that RNA regulation provides a crucial regulatory element to gene expression, an increasing number of tools and therapeutics targeting RNA are being developed. A field that started with interrogating RNA and RBP function with exogenously expressed reporter systems has slowly shifted towards using protein-sequence or RNA-guided technologies to study endogenous transcripts directly.

The discovery of the Cas protein family and the tolerance of Cas proteins to fusion proteins has opened up new opportunities for easily programmable interrogations and interventions at the RNA level. However, even though this provides an important basis for synthetic biology tools to study and control RNA functions, many regulatory pathways have been incompletely explored. For example, native nucleases, GFP fusions for imaging

purposes, ADAR fusion proteins for editing, and hnRNP fusions for splicing are the only Cas13-based systems that exist for RNA studies[59,60,62,93]. All these tools rely on Cas13 as the targeting moiety for direct cleavage or delivery of an effector protein. Cas13 itself is a large protein (approx. 100 kDa), which renders efficient delivery beyond DNA delivery methods challenging[62]. Additionally, the proteins of the Cas13 family show vastly different binding efficiency. Recently, a screen of 12 Cas13 family members that all operate with the identical gRNA delivery principle, revealed that only two dCas13-GFP fusions could efficiently target transcripts for imaging purposes, leaving many questions about Cas13 transcript targeting open[93]. Furthermore, Cas13 proteins are of microbial origin, which dramatically limits their therapeutic applications. Since RNA is a high turnover molecule, RNA therapeutics have to be dosed repeatedly. Recent studies have found that up to 85% of the population has pre-existing antibodies to Cas family proteins, which could lead to immunogenicity concerns upon repeated dosing of a therapeutic[94–96].

In the relatively nascent field of RNA technologies, there is a multitude of opportunities for synthetic biologists to develop RNA-targeting tools. Ideally, these tools would be of small size to facilitate delivery and limit disruptions of the cellular environment. Furthermore, it would be beneficial to develop a versatile set of orthogonal RNA-targeting technologies that allow for easy switching of the effector region to study different regulatory proteins and their interactions. In developing a potential RNA-targeting therapeutic, small size and non-microbial origin are key features. Therefore, a majority of this thesis will focus on the development of transcript-specific RNA-targeting tools to study and control gene expression with the challenges and limitations of current therapeutics in mind.

1.1.5 Scope of this thesis

In this thesis, I will focus on the development of RNA-targeting tools for biological studies and potential future therapeutic applications.

Chapter 2 presents the development of dCas13b-guided m^6A reader proteins to study the function of YTHDF1/2 on individual endogenous transcripts.

Chapter 3 presents the development and optimization of programmable RNA-guided RNA effector proteins engineered from human protein domain followed by multiplexed targeting of orthogonal RNA targeting systems and viral delivery.

Chapter 4 discusses the development of a small molecule-inducible RNA targeting system, which focuses on A-to-I editing followed by application of the system in a mouse model.

Chapter 5 introduces a second generation CIRT system with increased effector scope and efficacy followed by preclinical applications.

Chapter 6 summarizes this dissertation, provides a broader perspective and outlook on the field of RNA-targeting technologies and discusses potential future directions.

CHAPTER 2

TARGETED M^{6}A READER PROTEINS TO STUDY EPITRANSCRIPTOMIC REGULATION OF SINGLE RNAS

2.1 Introduction

RNA transcribed from the genome in the nucleus bears little resemblance to the RNA polymer it will ultimately become in the cytoplasm where it is translated into protein. Well-known processes such as capping, splicing and polyadenylation, as well as the recently discovered and ever-expanding list of diverse chemical modifications and editing, significantly alter the properties and fates of a given RNA during the course of its lifetime.[11–14,97,98] These alterations, which regulate critical aspects of RNA function such as stability, transport, protein binding, and translation, are commonly mediated by interaction with RNA binding proteins.[99] Especially in mammalian systems, regulatory processes at the post-transcriptional level are often a key determinant of genetic information flow.[100,101] As the array of known RNA regulatory mechanisms continues to rapidly expand, there is an increasing need for approaches to probe these processes in live cells at the single transcript level. In mammalian cells, *N*^{6}-methyladenosine (m^{6}A) is the most prevalent and well-studied messenger RNA (mRNA) modification.[98] m^{6}A has been implicated in various biological processes, such as cell development and differentiation[22,102–104], viral infections[105–107], and cancer[108–111], by mediating a variety of aspects of RNA processing and regulation such as splicing[15], translation[16,17,111], and stability[99,112]. In addition to the discovery of the 'writers' and 'erasers' that introduce and remove m^{6}A, respectively, several 'reader' proteins have been discovered that are able to link site-specific RNA modifications to particular regulatory functions within the cell.[27] The YT521-B homology domain family (YTHDF) of proteins is a family of cytoplasmic reader proteins that preferentially bind m^{6}A within a DR(m^{6}A)CH consensus site.[25] Early studies mainly focused on two YTHDF

family proteins, YTHDF1 and YTHDF2. YTHDF2 has been linked to the degradation of methylated RNA through recruitment of the CCR4-NOT deadenylation complex to the target transcript. In tethering assays on a reporter construct, the presence of YTHDF2 led to a decrease in RNA stability and the initiation of degradation. In co-immunoprecipitation assays, the subdomain region of YTHDF2 that is directly responsible for binding to its interaction partner CCR4-NOT was uncovered.[28]

YTHDF1 has been linked to the regulation of translation efficiency of a target transcript through interactions with the translation initiation machinery and the ribosome. Tethering assays showed enhanced translation of a reporter gene and knockdown studies showed YTHDF1 target-wide decreases in translation efficiency and ribosomal occupancy.[17] A later study found, however, that YTHDF1 could also induce deadenylation of targeted RNA transcripts in a tethering experiment.[28] Studies of viruses concluded that all YTHDF proteins can induce translation upregulation[113] or RNA degradation[114] and do not have different roles within cells. However, since all these experiments have been performed on a reporter system and with all biological regulation machinery present, it is unclear whether YTHDF1 itself initiates deadenylation or it merely has the capability to recruit other proteins, such as YTHDF2, to proceed with deadenylation, or vice versa. If the YTHDF proteins do not have different functions, there must be cellular machinery to regulate their functions on particular transcripts within the cell. Understanding this regulation is important because the ability of the YTHDF proteins to modulate the expression level of individual methylated transcript has also been linked to viral infection[105,107] and a variety of cancers.[115] While studying cellular readouts as a result of the YTH family proteins with an exogenous reporter plasmid is a critical first step to understanding their biological function, taking these proteins out of their endogenous regulatory network limits these mechanistic studies substantially. More importantly, whole-transcriptome studies using knockdown or overexpression of the reader protein reveals that the same m^6A sites on the same transcript can be regulated by both YTHDF1 and YTHDF2 as well as other reader proteins,[116,117] suggesting either competition or additional undiscovered regulat-

ory mechanisms are critical for controlling the outcome of the target transcript. To address this and related challenges, tools capable of delivering single reader proteins to single transcripts in a controllable, RNA modification-independent manner, would allow for the interrogation of the reader proteins on target RNA stability and translation efficiency.

The recently discovered CRISPR/Cas system has provided an easily programmable way to study nucleic acids in their endogenous environment.[7,8,57] Within the last several years, nuclease inactive DNA-targeting Cas9 (dCas9) has proven to be a versatile platform for the delivery of effector proteins to single sites in the genome. This has led to new methods to modify the epigenetic properties of cells, including histone methylation and histone acetylation status.[118,119] The discovery of RNA-targeting Cas proteins, the Cas13 family of proteins, has opened doors for equivalent studies on endogenous RNA transcripts.[9] For example, the RNA-targeting nuclease LwaCas13a can be programmed to target specific RNAs in mammalian systems and can efficiently degrade target transcripts through nuclease activity. Moreover, a nuclease-inactivated 'dead' version of LwaCas13a (dLwaCas13a) can be fused to green fluorescent protein for endogenous RNA imaging.[59] A nuclease-inactive variant of a Cas13 ortholog, PspCas13b, was fused to the known A-to-I RNA editing enzymes ADAR1 and ADAR2 to show direct RNA editing at endogenous RNA sites.[60] Studies with a 'dead' Cas13d ortholog furthermore demonstrated the ability of Cas13 to deliver a splicing modulator to individual transcripts.[62] Collectively, these studies demonstrate the potential of Cas13-based systems as targeting moieties for tools to study endogenous RNA regulation. In this present work, we developed dCas13b-m^{6}A reader tools to study the molecular basis of m^{6}A-mediated RNA regulation at specific transcripts (**Figure 2.1A**). We engineered dCas13b-YTHDF1 and dCas13b-YTHDF2 effector proteins that can be site-specifically targeted to an RNA of interest. We demonstrate that our constructs maintain the previously reported effects on translation efficiency (YTHDF1, **Figure 2.1B**) and degradation (YTHDF2, **Figure 2.1C**) using a luciferase reporter system, confirming the functions of these targeted readers in live cells. We go on to show that the fusion proteins can degrade endogenous transcripts in a YTHDF2-dependent manner. In-

triguingly, the YTHDF2-mediated decay substantially decreased the expression level of target transcripts, suggesting these tools may find utility for biotechnological applications. Together, this work provides the field with a programmable, versatile set of tools to study endogenous RNA regulation on the molecular level.

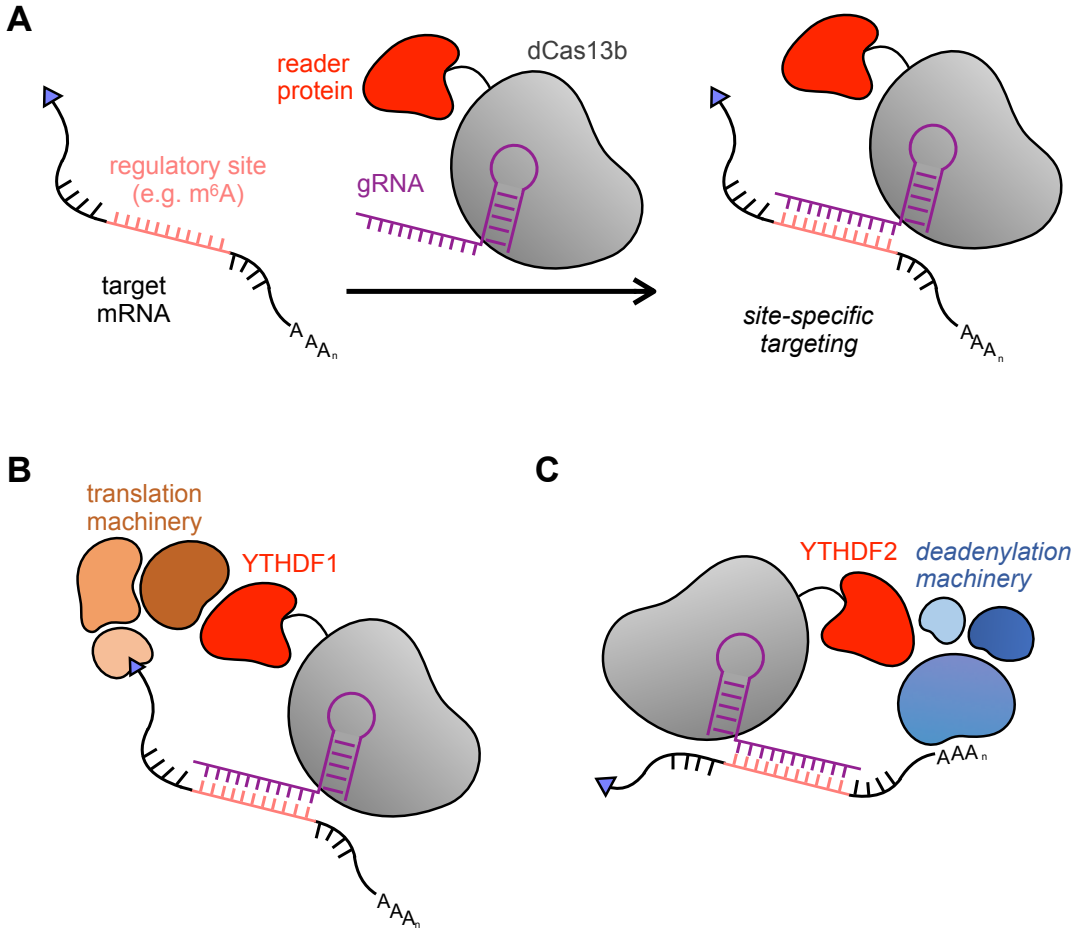

Figure 2.1 dCas13b fusion proteins as site-specific regulatory proteins.
(A) General overview of site-specific RNA targeting using dCas13b-guided fusion proteins. **(B)** dCas13b-YTHDF1 fusion protein can be targeted RNA transcripts to trigger assembly of translation machinery. **(C)** RNA targeted by YTHDF2 fusion proteins undergo degradation mediated by the CCR4-NOT deadenylation complex.

2.2 Results

2.2.1 Engineering dCas13b fusion proteins to study the epitranscriptome

To provide a simple tool to study RNA modifications, we sought to develop targeted fusion proteins that act in a chemical modification-independent manner. For both YTHDF proteins, the N-terminal domain contains the 'effector' region that mediates cellular responses, while the C-terminal domain contains the m^6A-binding YTH domain. We therefore chose to omit the C-terminal m^6A-binding domain in our fusion proteins to obtain a system that acts independent of RNA methylation. As a result, our constructs can be used to study RNA transcripts in their endogenous context by guiding an effector protein to a known methylation site or to study m^6A-initiated downstream effects decoupled from their native regulatory context. Conversely, eliminating the m^6A binding component of these proteins abolishes an additional regulatory layer of this system and adds the need to carefully control biological experiments to only capture true biological interactions.

To generate a targeted YTHDF1 reader, we cloned a mammalian expression vector with a catalytically-inactive PspCas13b (dPspCas13b) fused to the N-terminal domain of YTHDF1 (NYTHDF1). Truncated versions of the N-terminal domain of YTHDF2 (NYTHDF2) have previously been identified to bind the CCR4-NOT deadenylation machinery,27 maintaining the key functional output of the YTHDF2 reader protein. Therefore, to generate a targeted YTHDF2 reader, we cloned a mammalian expression vector with a dPspCas13b fused to the truncated N-terminal domain of YTHDF2 (YTHDF2(100-200)). We validated the function of the fusion constructs in live HEK293T cells using a dual luciferase reporter, which permits the analysis of relative changes in protein expression by luciferase enzymatic activity, and by RT-qPCR to assess changes in mRNA levels (**Figure 2.2A**). First, we validated our assay setup using the active Cas13b nuclease, which should be capable of degrading a target RNA through active nuclease activity. We transfected

cells with an expression vector for active Cas13b along with a vector that produces either a control off-target gRNA or a gRNA targeting firefly luciferase (**Figure 2.2A**).

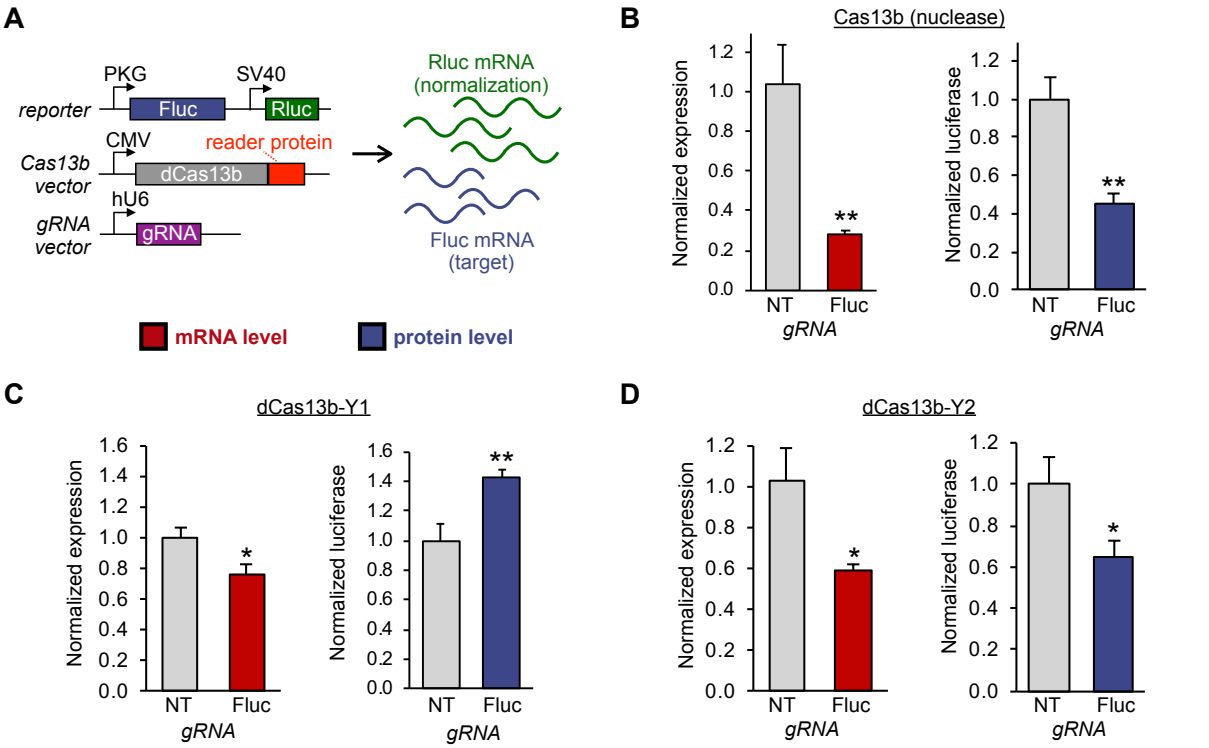

Figure 2.2 dCas13b-YTHDF fusion proteins can trigger enhanced protein production or RNA knockdown on a reporter transcript.
(A) Vector system used to test targeted reader protein tools in a dual luciferase assay. Firefly luciferase (Fluc) is the target of the experiment, while Renilla luciferase (Rluc) is an internal control. Gene expression is monitored by RT-qPCR and protein production is monitored by luciferase luminescence. An off-target gRNA (NT) and a guide RNA targeting Fluc (Fluc) delivery the fusion protein to the transcript. **(B)** HEK293T cells were cotransfected with the vectors shown in (A), including the active Cas13b nuclease, and 48h after transfection subjected to RT-qPCR (red bar) and luciferase assay (blue bar). **(C)** HEK293T cells were cotransfected with the vectors shown in (A), with the dCas13b-YTHDF1 fusion, and assayed by RT-qPCR (red bar) and luciferase assay (blue bar) (n = 9). **(D)** HEK293T cells were cotransfected with the vectors shown in (A), with the dCas13b-YTHDF2 fusion, and assayed by RT-qPCR (red bar) and luciferase assay (blue bar). Student's t-test; *P <0.05, **P <0.01 (n = 3 for RT-qPCR, n = 3 biological x 2 technical for Cas13b and dCas13b-Y2 luciferase assays, n = 3 biological x 2 technical + 3 biological replicates for dCas13b-Y1 luciferase assay).

As expected, we observed a substantial decrease of both luciferase RNA as measured

by RT-qPCR (**Figure 2.2B**, red bar) and protein levels (**Figure 2.2B**, blue bar). This control experiment validates that the dual luciferase reporter is capable of measuring both protein and RNA levels of the model transcript.

Next, we subjected the new YTHDF-fusion proteins to the same analysis as the active Cas13b nuclease. We transfected HEK293T cells with the dCas13b-NYTHDF1 construct along with either a control gRNA or a gRNA targeting the fusion protein to the firefly luciferase mRNA. Interestingly, delivery of the YTHDF1 protein resulted in a slight decrease in mRNA levels as measured by RT-qPCR (**Figure 2.2C**, red bar). Even though the mRNA levels go down, the protein levels consistently increase (**Figure 2.2C**, blue bar) as has been observed in previous tethering assays.[17] These findings indicate that the targeted YTHDF1 protein maintains function as a fusion construct on a model target gene.

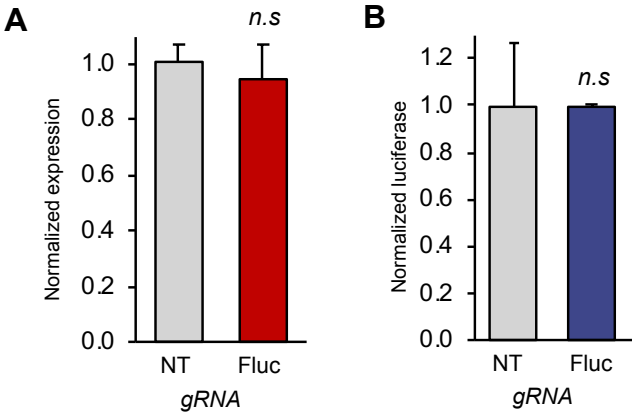

Figure 2.3 Unfused dCas13b has no effect on RNA and protein levels.
(A) Firefly luciferase mRNA levels as measured by qPCR and **(B)** luciferase protein levels (n = 2) as measured by dual luciferase assay do not changes when HEK293T cells are transfected with a dCas13b fusion protein without an effector. Student's t-test; *P <0.05.

We then assayed whether the YTHDF2 fusion construct also retained its function mediating RNA decay as a targeted protein. We transfected HEK293T cells with an expression vector producing the dCas13b-NYTHDF2 fusion protein, again along with either a control gRNA expression vector or a vector that produces a gRNA targeting firefly luciferase. As

expected, targeted YTHDF2 led to a decrease in both mRNA and protein levels (**Figure 2.2D**). The level of luciferase decrease observed agrees well with previously obtained data with the full-length N-terminal domain of YTHDF2 in tethering assays.[17] As a control, dCas13b lacking a fusion protein showed no measurable change in targeted gene expression or protein production (**Figure 2.3**). As some of the previous studies have been conducted in HeLa cells, we also verified the constructs using HeLa cells, observing the same trends as seen for HEK293T cells (**Figure 2.4**). Taken together, these findings indicate that both YTHDF proteins retain previously reported functions as part of a targeted fusion construct and can be used to study m^{6}A regulation dynamics in cells.

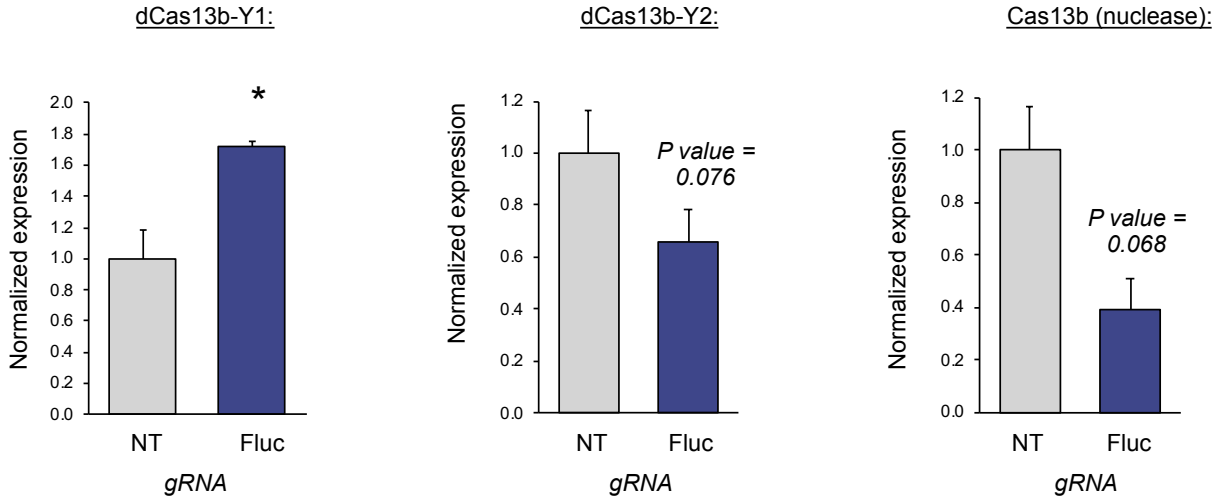

Figure 2.4 dCas13-YTHDF proteins also show activity in HeLa cells.
dCas13b-YTHDF fusion proteins induce an increased protein production or RNA knockdown on a dual luciferase reporter construct in HeLa cells dCas13b-Y1 (n = 2), dCas13b-Y2 (n = 4), Cas13b (n = 5). Student's t-test; *P <0.05.

2.2.2 Determining YTHDF1 fragments responsible for translation activation.

As previously described, YTHDF2 truncations that maintained downstream function independent of the m^{6}A recognition domain had previously been identified. To study

endogenous RNA regulation, we wanted to ensure that the fusion tools are as small and unobtrusive as possible. Unlike YTHDF2, the active subdomain of YTHDF1, which is responsible for binding to the translational machinery, has not been identified. We reasoned that our reporter assay would allow us to identify the subdomain responsible for RNA translation activation. We therefore cloned three different truncations of the N-terminal domain of YTHDF1 (aa1-100, aa100-200, aa200-364) fused with dCas13b (**Figure 2.5A**). We used the same qPCR and dual luciferase assays as described above to compare the differences in mRNA and protein levels of each truncations with previously obtained data for the full-length N-terminal domain. At the mRNA level, none of the fusions caused dramatic changes, except for the unexpected observation that the YTHDF1(aa1-100) fusion seemed to cause some RNA stabilization (**Figure 2.5B**) while inducing less protein production (**Figure 2.5C**). More importantly, we found that only the truncation YTHDF1(aa100-200) retains its translation activation activity (**Figure 2.5C**). However, we observed saw some sample variability in the translational activation effect, resulting in data below the statistically significant threshold. We therefore repeated the YTHDF1(aa100-200) experiment with additional replicates, which and confirmed that the observed translational activation effect (**Figure 2.5D**).

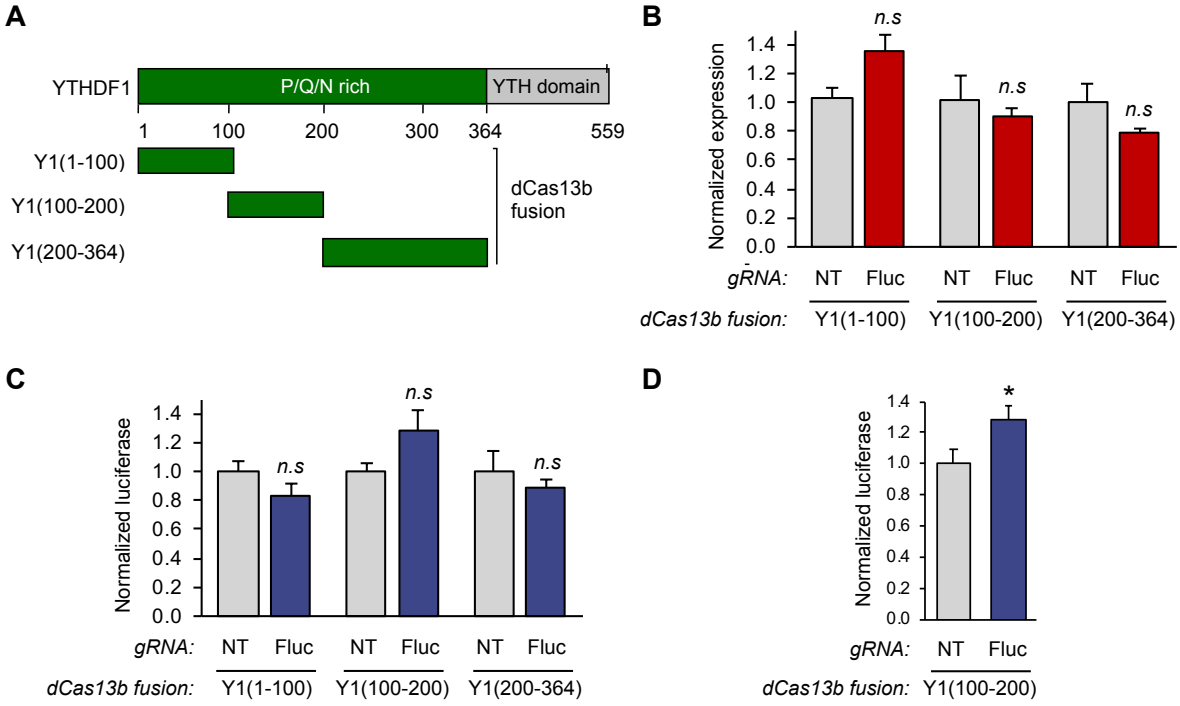

Figure 2.5 A fragment of the N-terminal domain of YTHDF1 is sufficient to bind to the translation initiation machinery.
(A) Schematic diagram of YTHDF1 fragments. **(B)** qPCR analysis of HEK293T cells transfected with Y1(aa1-100), Y1(aa100-200), or Y1(200-364)-dCas13b fusion proteins and either the NT guide or the Fluc guide using the assay shown in Figure 2.2A (n = 3). **(C)** Protein readout analysis of the same Y1(aa1-100), Y1(aa100-200), and Y1(aa200-364) fusion proteins using the dual luciferase reporter system. The Y1(100-200) truncation shows activation of protein production without effecting RNA levels, defining this portion of the protein as the active reader domain (n = 3 biological x 2 technical). **(D)** Luciferase assay analysis of further replicates of the identified dCas13bY1(100-200) protein. Student's t-test; *P <0.03 (n = 9).

Interestingly, unlike with the full-length N-terminal YTHDF1 fusion, we did not observe a decrease in mRNA level for the dCas13b-Y1 (100-200) truncation. The YTHDF1 (aa200-364), however, seemed to promote a RNA decrease similar to the level observed with the whole domain. This could be an indication that aa100-200 is responsible for binding to the translation machinery, while aa200-364 bind to other regulatory proteins, such as YTHDF2, in cells. This observation warrants further investigation and further expands the YTHDF-mediated regulatory repertoire. This complexity in reader subdomains could also

explain some of different properties observed for YTHDF1 for specific transcripts under different conditions in the literature, and further confirms the need for the tools described here.

2.2.3 Endogenous RNA targeting with dCas13b-YTHDF2.

Finally, we tested whether endogenous transcripts could also be targeted by the fusion proteins, using dCas13b-YTHDF2 fusion construct as an exemplar. We selected two transcripts: the low-abundant transcript KRAS and the highly abundant PPIB, selected because each are known to be m^{6}A modified,[120,121] and have also been validated for targeting by Cas13b.[59] To find an appropriate targeting site, we conducted a small gRNA screen for both targets (**Figure 2.6**). To validate our system, we chose to use the gRNAs showing the largest response with our dCas13b-Y2 construct in all following experiments.

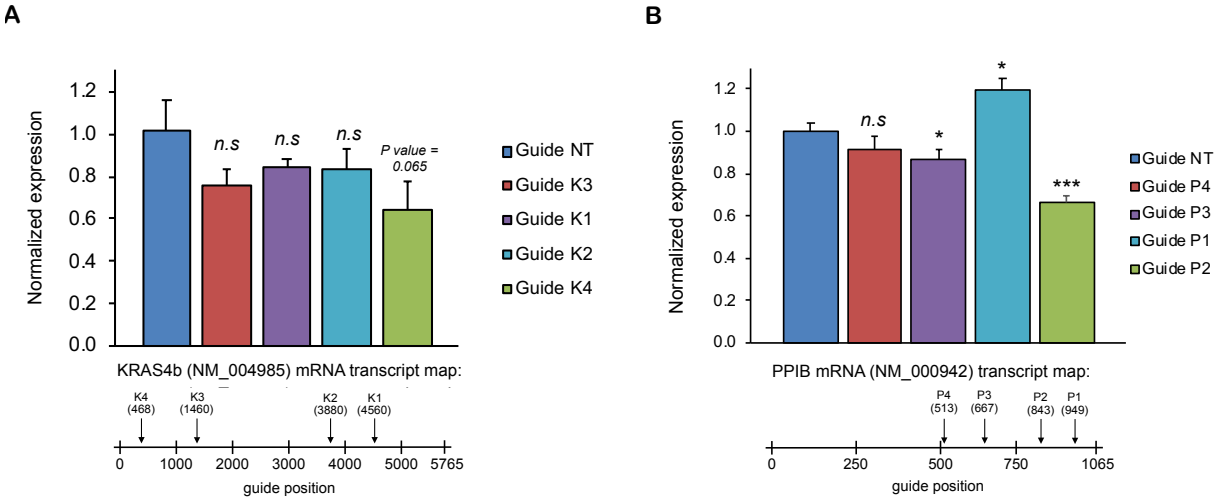

Figure 2.6 gRNA screen with dCas13b-YTHDF2 by RTqPCR.
(A) For KRAS and **(B)** PPIB. We used RT-qPCR to quantify knockdown by dCas13b-YTHDF2 with gRNAs that target different sites of the transcript to identify the regions of most robust response. Any further experiments were conducted with the gRNAs showing the largest response (K4 and P2). Student's t-test; *P <0.05 ***P < 0.001.

After selecting gRNAs targeting either KRAS or PPIB, we transfected HEK293T cells with the dCas13b-YTHDF2 expression vector and a gRNA expression vector. We furthermore transfected cells with the active Cas13b or dCas13b and identical on- and off-target gRNA as positive and negative controls, respectively. For KRAS, we observed a small but reproducible decrease in mRNA levels mediated by YTHDF2 delivery (**Figure 2.7A**). Surprisingly, the knockdown efficiency KRAS was similar to what we observed when using the active Cas13b nuclease to actively degrade the transcript (**Figure 2.7B**).[59] However, an unfused control dCas13b protein had no effect on the transcript (**Figure 2.7C**). PPIB was substantially more efficiently knocked down by both dCas13b-YTHDF2 (**Figure 2.7D**) and nuclease-active Cas13b (**Figure 2.7E**), again with much less effect from a dCas13b control (**Figure 2.7F**). These results suggest that the dCas13b-YTHDF2 approach is a new way to trigger a degradation response on endogenous targets. The observed differences in mRNA decay in response to YTHDF2 targeting again showcases the complexity of RNA regulation and the importance of tools to probe transcript-specific responses. In addition to RT-qPCR, we next verified changes in the expression level of the protein product of the gene PPIB (cyclophilin B: CypB) by Western blot. As anticipated from the qPCR results, we saw the protein level decreased when we transfected cells with our dCas13b-Y2 construct and a PPIB-targeting gRNA (**Figure 2.8**).

Off-target effects are one of the main concern when using Cas-based systems. To assess how much we perturb the overall transcriptome in cells, we conducted a whole-transcriptome analysis by RNA sequencing. We transfected cells with dCas13b-Y2 or active Cas13b and off-target gRNA or PPIB gRNA, prepared mRNAs libraries, and then analyzed each transcriptome by high-throughput sequencing (HTS). We determined that for dCas13b-Y2 there were 26 differentially expressed transcripts between the NT gRNA and PPIB gRNA triplicates (**Figure 2.9B**). We observed changes indicating higher as well as lower expressed transcripts, which could be due to biological on-target effects or mistargeting of dCas13b.

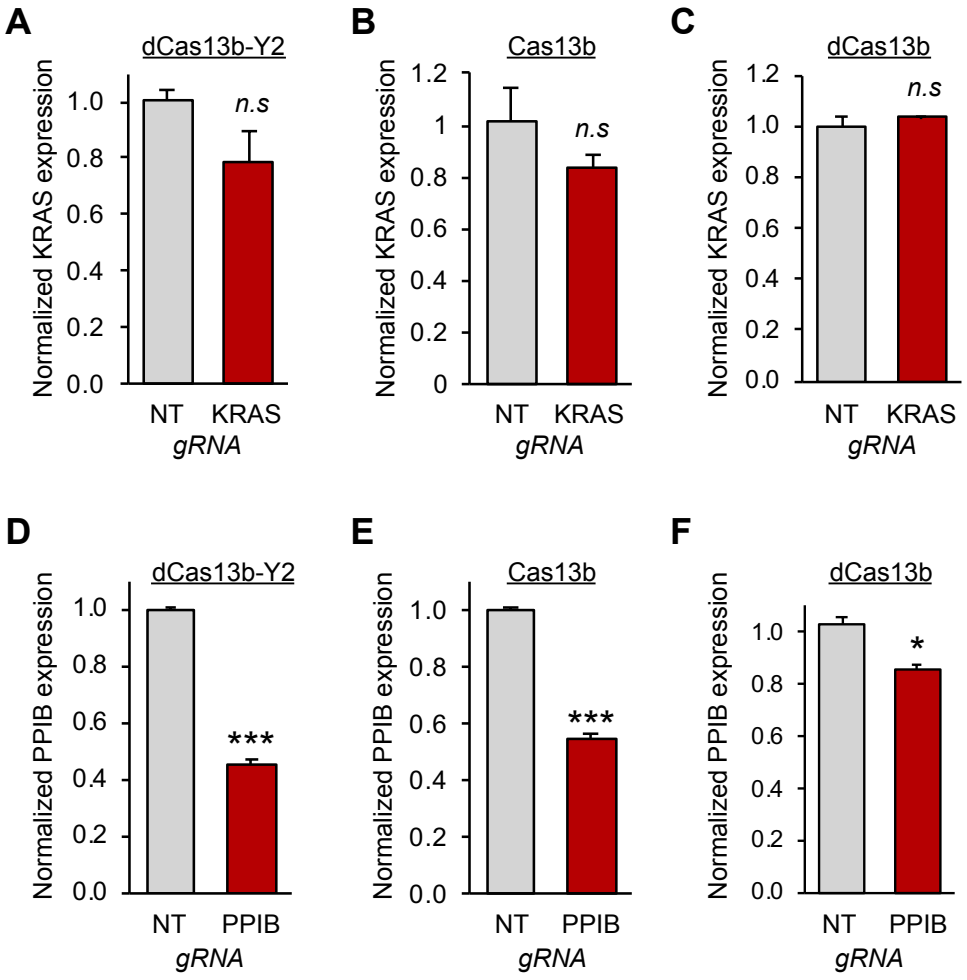

Figure 2.7 Targeted YTHDF2 readers can be used to study the regulation of endogenous methylated transcripts in live cells.
(A) HEK293T cells were transfected with dCas13b-YTHDF2 and either NT gRNA (grey bar) or KRAS gRNA (red bar). 48h later, mRNA levels were determined by RT-qPCR. **(B)** The same analysis was performed in A except the nuclease active Cas13b was delivered. **(C)** The same analysis was performed in A except a dCas13b protein without a fusion was delivered. **(D)** HEK293T cells were transfected with dCas13b-YTHDF2 and either NT gRNA (grey bar) or PPIB gRNA (red bar). 48h later, mRNA levels were determined by RT-qPCR. **(E)** The same analysis was performed in D except the nuclease active Cas13b was delivered. **(F)** The same analysis was performed in D except a dCas13b protein without a fusion was delivered. Student's t-test; *P <0.05, ***P <0.001 (n = 3).

YTHDF2 is known to recruit cellular machinery that can lead to changes in transcript expression and we cannot rule out that this happens when we overexpress it. Nonethe-

less, the off-target effects were comparable to those observed we observed for nuclease-mediated Cas13b targeting (23 differentially expressed transcripts, **Figure 2.9D**) and are substantially less than those observed from other technologies like shRNA.[62]

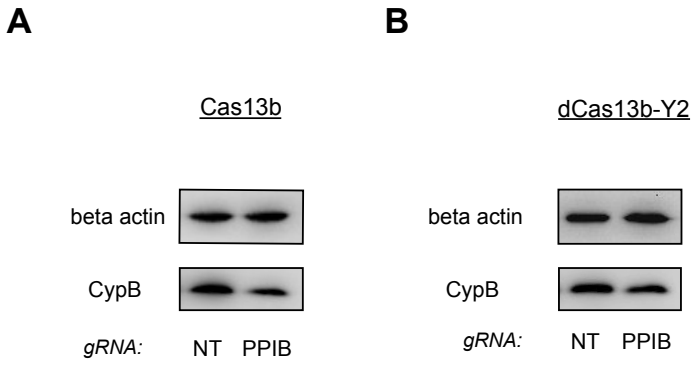

Figure 2.8 Targeted dCas13b-Y2 proteins induce protein level changes.
(A) HEK293T cells were transfected with active Cas13b nuclease and either NT gRNA or PPIB gRNA and protein levels for the protein product of PPIB (cyclophilin B – CypB) were assayed. **(B)** The same analysis as in A was performed but with dCas13b-Y2.

The transcript we targeted with our dCas13b-Y2 construct was not found to be differentially expressed by HTS (**Figure 2.9**). To examine this discrepancy, we conducted further RT-qPCR analyses on the extracted total RNA that was used for HTS. We analyzed the samples prior to mRNA purification and library preparation to rule out biases and artifacts that may emerge during these steps, and also performed comparative analysis with other housekeeping genes to confirm the validity of the on-target effect. Indeed, these additional RT-qPCR experiments reproduced our previous observations, revealing similar decreases in PPIB levels when referencing to either GAPDH or ACTB (**Figure 2.9A and C**). Taken together with our Western blot analysis, we conclude that the relatively modest but significant extent of transcript knockdown is lost during the purification and library preparation steps and was not substantial enough to observe by HTS. To confirm our HTS experimental and analysis pipeline was functioning properly, we also performed transcriptome analysis with PPIB targeted by the active Cas13b nuclease. We found that while

the targeted PPIB transcript is significantly decreased as measured by RT-qPCR, it falls barely into our significance cutoff by HTS. As nuclease-mediated Cas13b outperforms the dCas13b-Y2 knockdown efficiency, we suspect that the overall difference in RNA level must be more substantial to be seen after library preparation and sequencing. It has been previously reported that library preparation and handling can introduce large differences in the outcome.[122] We believe that if we optimized our dCas13b-Y2 conditions further to yield a larger decrease in RNA level, we could observe these by RNA-seq. At this point, we can conclude from our RNA-seq results that expressing dCas13b-Y2 in cells does not cause gross changes to the transcriptome. Any potential Cas-mediated off-target effects will only be uncovered once the system performs well enough for more in-depth sequencing analysis.

2.3 Conclusion and Discussion

We developed and validated dCas13-targeted RNA reader proteins as a new platform to study RNA regulation, focusing here on YTHDF reader proteins of m^6A-modified RNA transcripts. Our system is easily programmable by simply changing the gRNA for Cas13b and can therefore be used to study the regulation of any methylated RNA of interest. Furthermore, due to Cas13's tolerance to fusion proteins, it can also be adapted to study the regulation of any known effector protein of RNA by switching the fusion protein. To our knowledge, this presents the first tool to study RNA regulation dynamics and effects of RNA modifications on single transcripts in an endogenous context. We noticed a high degree of variability in our cell-based assays that we attribute to differences in transfection and cell state. Further optimization is required to ensure more consistent cellular assays. The difference in response between different endogenous KRAS and PPIB transcripts could indicate that there are regulatory differences between genes.

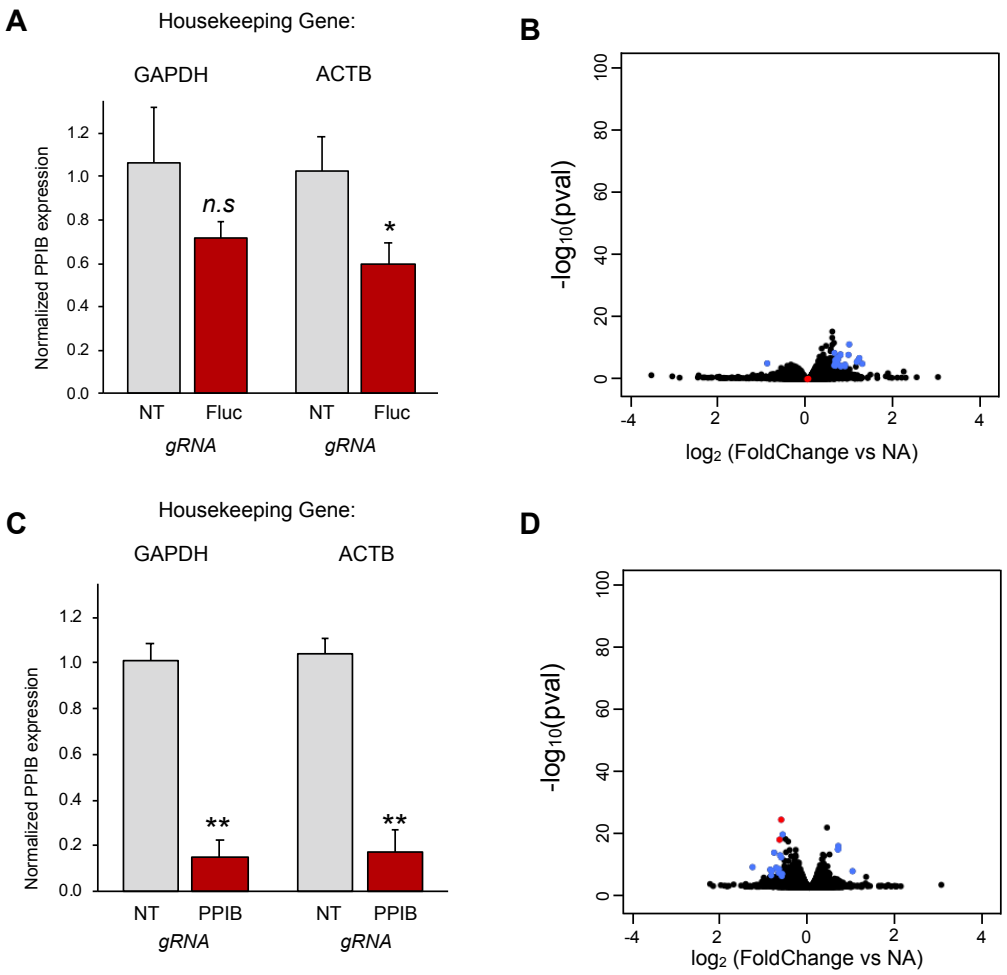

Figure 2.9 On- versus off-target activity of dCas13b-Y2 and Cas13b.
(A) RT-qPCR analysis of dCas13b-Y2 targeting PPIB RNA using two different validated housekeeping genes. This sample was subsequently sent for Illumina sequencing. **(B)** Volcano plot of differentially expressed transcripts between dCas13b-Y2 and NT gRNA versus PPIB-targeting gRNA as determined by RNA sequencing (n = 3). **(C)** RT-qPCR analysis of Cas13b nuclease targeting PPIB RNA using two different validated housekeeping genes. **(D)** Same analysis for dCas13b targeting the endogenous PPIB transcript as shown in B. We find that both dCas13b-Y2 and the active Cas13b nuclease show few differentially expressed transcripts (26 vs 23 significant transcripts – shown in blue). The two PPIB isoforms (shown in red) knocked down by Cas13b barely fall into our confidence cutoff for significant transcripts. Transcripts knocked down with dCas13b-Y2 do not show statistically significant changes. After we verified our dCas13b-Y2 fusion by RT-qPCR with different housekeeping genes and by Western blot, we determined the extend of changes in transcript levels are likely responsible for this discrepancy. Further optimization to maximize the cellular response to our dCas13b-Y2 construct could likely remedy this discrepancy in the future. Student's t-test; *P <0.05 **P < 0.01.

These data also suggest Cas accessibility of a target transcript may be a determinant of overall effectiveness on the transcript, suggesting better RNA-targeting Cas systems are still needed. Overall, these tools provide the ability to investigate such differences on endogenous transcripts and allow for new mechanistic studies regarding RNA regulation of individual transcripts. However, the relatively large size of the currently available RNA-targeting Cas systems could also provide additional complexity to interpreting the results when delivered to transcripts, which could be alleviated by smaller delivery systems.

In this work, we chose to develop and validate our new targeted reader protein tools in the context of transcripts with annotated m^6A sites, because these RNAs are already prone to chemical modification-mediated post-transcriptional gene regulation. However, we found that the system also works on the firefly luciferase reporter RNA. In their work, Liu et al.[123] show that in a similar reporter construct with the same firefly luciferase coding region as in our luciferase reporter construct, there was no substantial methylation of Fluc by m^6A-pulldown followed by RT-qPCR. They furthermore saw the same methylation level with wild type (GGACU) and a 'dead' mutated m^6A (GGAUU) consensus site. Since m^6A is known to exist predominantly at DRACH consensus sites, this suggests there is no substantial methylation of the firefly luciferase reporter system. Therefore, it is reasonable to assume that our dCas13b-Y2 system is acting independent of the methylation level at its targeting site. However, future work, possibly deploying targeted methyltransferases and demethylases, will further clarify these mechanistic questions.

Surprisingly, targeting YTHDF2-mediated decay resulted in dramatic knockdown of certain target genes in our hands, not too different from the knockdown efficiency of nuclease active Cas13b. For example, on the reporter construct, active nuclease-mediated degradation resulted in a 75% decrease in RNA levels and 50% decrease in protein levels, while YTHDF2-mediated decay lead to a 40% decrease in RNA levels and a 40% decrease in protein levels (**Figure 2.2B and C**). The YTHDF2 knockdown of the endogenous genes tested were also quite comparable to active nuclease-mediated degradation.

For example, YTHDF2-mediated decay knocked down PPIB by 60% (**Figure 2.7D**), while active Cas13b-mediated decay knocked the gene down by 55% (**Figure 2.7E**). This efficiency in gene expression control suggests that reader protein-mediated decay may be a useful strategy for both synthetic biology and bioengineering applications.[57,124] Collectively, these experiments indicated that dCas13b-fusion proteins can be used to study single transcripts in a native biological context, and demonstrates there is still much to be learned about RNA regulation.

2.4 Materials and Methods

Cloning.

All plasmids were generated by Gibson Assembly cloning using PCR products amplified with Q5 DNA Polymerase (NEB). The plasmids were sequenced by the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility. All plasmids used in this study are listed in Table 2.1 and 2.2. Vector maps are available online as indicated and are physically available upon request. The original dPspCas13b plasmid (pC0050-CMV-dPspCas13b-longlinker-ADAR2DD(wt)) was a gift from Feng Zhang (Addgene plasmid 103866).

Mammalian cell culture and plasmid transfection.

HEK293T (ATCC) and HeLa (ATCC) cells were maintained using DMEM (L-glutamine, high glucose, sodium pyruvate, phenol red; obtained from Corning) supplemented with 10% fetal bovine serum (FBS, Gemini Benchmark), and 1% penicillin/streptomycin (P/S, Gibco/Life Technologies). For transfections, cells were cultured in in DMEM (L-glutamine, high glucose, sodium pyruvate, phenol red; obtained from Corning) supplemented with 10% FBS. Plasmid transfections were achieved using lipofectamine 2000 (Invitrogen) following the manufacturer's protocol.

Mammalian luciferase assay.

To test changes in protein levels, HEK293T cells were transfected with 150 ng Cas13b/dCas13b-YTHDF1/dCas13b-YTHDF2, 100 ng gRNA, 10 ng DualGlo luciferase reporter (Promega) per well, unless otherwise noted. Due to their lower tolerance of DNA transfection amounts, HeLa cells were transfected with 75 ng Cas13b/dCas13bYTHDF1/dCas13b-YTHDF2, 50 ng gRNA, and 10 ng DualGlo luciferase reporter (Promega). For Figure 2.3B, cells transfected with 80 ng dCas13b unfused construct, 80 ng scFv construct in addition to the luciferase reporter. Cells were plated on a 96-well plate (CellVis) 16 h before transfection and transfected at 80% confluency. A total of 20 µl Opti-MEM I Reduced Serum Medium (ThermoFisher Scientific) per well was used after combining 10 µl Opti-MEM containing 0.5 µl Lipofectamine 2000 and 10 µl Opti-MEM containing the plasmids. Combined solutions were incubated for 15 min before they were slowly pipetted onto cells. After 48 h, luciferase activity was assayed using the DualGlo Luciferase Assay System (Promega) on a Biotek Synergy plate reader according to the manufacturer's instructions. All experiments were performed in at least three biological and two technical replicates. Firefly luciferase readouts were normalized to the corresponding Renilla luciferase readout to account for differences in transfection efficiency.

Total RNA isolation and quantitative PCR.

To determine gene expression levels in HEK293T cells, RNA was isolated and changes in RNA levels were quantified using RT-qPCR. Cells were plated on a 48-well plate (Corning) and transfected with 500 ng Cas13b/dCas13b-fusion, and 300 ng gRNA plasmid. Where applicable, 30 ng DualGlo luciferase reporter (Promega) was added. Total RNA was harvested 48 h after transfection using the RNeasy Mini Kit (Qiagen). Following RNA isolation, RNA was reverse transcribed to cDNA using the PrimeScript RT Reagent Kit (TaKaRa). All qPCR reactions were performed as 20 µL reactions using FastStart Essential DNA Green Master (Roche) and amplified on a LightCycler 96 Instrument (Roche). Expression levels were obtained by subtraction the housekeeping gene (GAPDH or ACTB)

C_t value from target C_t value and normalizing to the non-targeting (NT) gRNA. Relative abundance was determined using $2^{-\Delta C_t}$. All assays were performed with three biological replicates. The qPCR primers used in this study are listed in Table 2.3.

Western blotting.

For Western blotting, HEK293T cells were plated in 12-well plates (Corning) and transfected with 2 µg Cas13b/dCas13b-YTHDF2 and 1.2 µg gRNA plasmids. After 48 h, cells were washed with ice cold PBS and lysed in 50 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% deoxycholate, 2% SDS, pH 7.4). After a 30 min incubation at room temperature, the protein concentration was measured by BCA assay (Thermo Scientific). The appropriate amount of protein (35 µg protein when detecting CypB) was boiled with loading dye for 5 min at 95°C and loaded onto a 15% SDS PAGE gel. The proteins were transferred onto a PVDF membrane (Millipore) and blocked in 5% milk in TBST. Proteins were detected using a 1:1000 dilution CypB antibody (Santa Cruz), and 1:1000 anti-mouse HRP-conjugated antibody (Santa Cruz). Membranes were imaged on Fluor Chem R (Protein Simple) after incubation with Super Signal West Pico Plus (Thermo Scientific).

RNA sequencing and analysis.

To determine the specificity of this dCas13b-based system, we performed RNA-seq analysis. HEK293T cells were plated in 12-well plates (Corning) and transfected with 1.8 µg Cas13b/dCas13b-YTHDF2 and 1.2 µg gRNA plasmid. After 48 h, total RNA was extracted using the RNeasy Mini Kit (Qiagen) followed by a 30 min DNaseI (Fisher) treatment and clean up using the RNA clean up and concentrator kit (Zymo). The resulting total RNA was used as the input for library preparation and for qPCR analysis of the same samples (as described earlier). mRNA extraction and RNA-seq libraries were prepared using the mRNA HyperPrep Kit (KAPA biosystems). Libraries were sequenced on an Illumina HiSeq instrument at the University of Chicago Genomic Facility with at least 14 million reads per library. Reads were mapped to the RefSeqGRCh38 transcriptome, quantified,

and pseudoaligned using kallisto.[125] To find differentially expressed transcripts, we used sleuth.[126] Only genes that had a $\log_2$FoldChange of at least 0.6 and a FDR < 0.05 were considered to be differentially expressed.

2.5 Supplemental information

Table 2.1 List of mammalian plasmids used. All dCas13b plasmids are Kan resistant.

	Number:	Description:	Benchling Link:
1	20-23	Cmv d0 active PspCas13b	https://benchling.com/s/seq-X6scGnmdaSzWQvbHfBwC
2	19-17	Cmv d0 dPspCas13b-GGS-NYTHDF1	https://benchling.com/s/seq-EhcdVC5ME0OYyA9UJfVA
3	20-16	Cmv d0 dPspCas13b-GGS-NYTHDF2(100-200)	https://benchling.com/s/seq-WwYoBoYpesVm1GctvqzO
4	20-17	Cmv d0 dPspCas13b-GGS-NYTHDF1(1-100)	https://benchling.com/s/seq-4IuLfdmD2rxyKOuUoD2E
5	20-18	Cmv d0 dPspCas13b-GGS-NYTHDF1(100-200)	https://benchling.com/s/seq-kVAKoGrrOMfWFufTDo67
6	20-19	Cmv d0 dPspCas13b-GGS-NYTHDF1(200-364)	https://benchling.com/s/seq-XZHJgFuhWuuN861cbKRU
7	20-22	Cmv d0 dPspCas13b-GGS-Sun	https://benchling.com/s/seq-z8IHxkkO37vqX1G0uMn8

Table 2.2 List of gRNA plasmids used. All gRNA plasmids are carb resistant.

	Number:	Description:	Benchling Link:
1	18-46	hU6 promoter PspCas13b gRNA dual luciferase reporter 3'UTR	https://benchling.com/s/seq-6PObIxex4v49EEx2nAo4
2	20-24	hU6 promoter PspCas13b gRNA KRAS 1	https://benchling.com/s/seq-9955TvbppwR96Q7QFc7W
3	20-25	hU6 promoter PspCas13b gRNA KRAS 2	https://benchling.com/s/seq-5Kx3WuH-FLrIdov9JCHKM
4	19-20	hU6 promoter PspCas13b gRNA KRAS 3	https://benchling.com/s/seq-STgJGVMHwggItLqEGWGv
5	20-20	hU6 promoter PspCas13b gRNA KRAS 4	https://benchling.com/s/seq-UMRBpa96BoQmgOILyLFk
6	20-28	hU6 promoter PspCas13b gRNA PPIB 1	https://benchling.com/s/seq-Jp6s98d6s1yIEPAyPaKi
7	20-29	hU6 promoter PspCas13b gRNA PPIB 2	https://benchling.com/s/seq-V6h64qrB3cGYAQDa3gVz
8	20-30	hU6 promoter PspCas13b gRNA PPIB 3	https://benchling.com/s/seq-DpK1iX-HAPIIVLdldcKQF
9	20-31	hU6 promoter PspCas13b gRNA PPIB 4	https://benchling.com/s/seq-gJjp9QasQ2RH4VN2BgYY

Table 2.3 List of qPCR primers used.

	Gene:	Forward:	Reverse:
1	GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA
2	ACTB	tcagcaagcaggagtatgac	agccatgccaatctcatct
3	KRAS	tggtggctgatgctttga	cactggatagggttctgtctattc
4	PPIB	AACGCAGGCAAAGACACCAACG	TCTGTCTTGGTGCTCTCCACCT
5	Fluciferase	AGGTTACAACCGCCAAGAAGC	ATGAGAATCTCGCGGATCTTG

CHAPTER 3

PROGRAMMABLE RNA-GUIDED RNA EFFECTOR PROTEINS BUILT FROM HUMAN PARTS

3.1 Introduction

Programmable nucleic acid-binding proteins, including zinc finger proteins, transcription activator-like effector nucleases (TALEN), PUF (Pumilio) proteins, and Cas proteins, have revolutionized genome studies and editing technologies[127–130] and are opening up new therapeutic opportunities to treat human diseases[131,132]. In particular, the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas system, which evolved as a prokaryotic immune defense mechanism, has transformed our ability to study and manipulate cellular DNA site-specifically[6–8]. A key advantage of CRISPR-Cas systems compared to previous methods, such as zinc finger proteins and TALE nucleases[5,51–54], is that they are easily programmable to target virtually any locus of interest. The CRISPR-Cas system is a ribonucleoprotein complex that uses base pair interactions of a displayed guide RNA (gRNA) to interact with a target nucleic acid sequence. The simple nature of base pair-guided targeting opens up the possibility to program systems to interact with a defined nucleic acid sequence by simply changing the nucleic acid sequence on the guiding strand.

While targeting DNA directly will have profound clinical ramifications, diseases that involve subtle alterations to many genes will be challenging to target using DNA editing technologies[133,134]. Additionally, potential side effects or risks of permanent genetic alteration might not be tolerated for some diseases. For example, the genes one may want to target to activate an enhanced wound healing response are likely targets that could pose a risk for cancer development, rendering permanent DNA-based alteration strategies risky[135]. Targeting information flow at the RNA level presents several oppor-

tunities for therapeutic intervention, including but not limited to the ability to halt treatment if side effects emerge, the ability to target genes that would be too risky to alter at the DNA level, and the ability to manipulate gene expression without permanent alterations to the host genome. While inhibiting or enhancing transcription at the genome level provides one possibility for controlling gene expression[136,137], recently discovered RNA epitranscriptomic regulatory mechanisms offer a broad range of RNA regulatory processes to target, including editing, degradation, transport, and translation of RNA transcripts[33,98,138]. Although the mechanisms and consequences of this epitranscriptomic regulatory layer are just beginning to be uncovered, it is apparent that the information flow through RNA is tightly regulated, offering many new opportunities for both basic research discoveries as well as therapeutic development.

Programmable RNA-targeting technologies hold great promise for both biological studies and therapeutic applications. Prior to the discovery of the CRISPR-Cas protein family, programmable PUF (Pumilio) protein-based RNA-targeting tools analogous to TALEN had been developed[4,5,139]. However, the ease with which Cas9-based DNA targeting technologies can be reprogrammed by simply switching a guiding RNA region offers significant advantages over previously developed purely protein-based tools[127]. Cas proteins have been repurposed to act as a programmable RNA targeting moiety to degrade target RNAs using the nuclease activity inherent to Cas13 systems, or catalytically inactive dead Cas13 proteins (dCas13) have been used to image RNA by fusions to GFP[59], to edit RNA by delivering an ADAR A-to-I editing enzyme[60], and to modulate splicing by delivering hnRNPa1 to target transcripts of interest[62]. Additionally, we recently developed a dCas13 system to deliver truncated N^6-methyladenosine (m^6A) binding proteins ("readers") to specific sites in the transcriptome, yielding new tools to study and control RNA regulation[63].

Although the Cas9 and Cas13 systems have revolutionized studies of DNA and RNA, respectively, the large size and bacterial origin of these proteins pose problems for both basic research applications and therapeutic development. Most Cas13 proteins studied to-

date are around 130 kDa in size, and even the smallest member of the family (Cas13d) is approximately 100 kDa[140]. From a translational perspective, the large size presents challenges for viral packaging and direct protein delivery. Moreover, it was recently discovered that a substantial fraction of people already have circulating antibodies to CRISPR-Cas proteins[94–96], suggesting immunogenicity issues may prove problematic in longterm clinical applications. While a one-time DNA editing treatment may not present immunogenicity problems, targeting RNA therapeutically with continuously-delivered microbially derived effector proteins may eventually lead to substantial immunogenicity challenges.

To overcome the large size and microbial-derived nature of current RNA-targeting systems, we present a CRISPR-Cas-inspired RNA targeting system (CIRTS), a general method for engineering programmable RNA effector proteins. We show that the CIRTS strategy permits mining the human proteome for functional parts to build programmable RNA regulatory proteins. Similar to CRISPR-Cas-based systems, CIRTS is a ribonucleoprotein complex that uses Watson-Crick-Franklin base pair interactions with a gRNA to deliver protein cargo site-selectively to the transcriptome. We show that CIRTS can deliver a range of regulatory proteins, including ribonucleases for direct RNA degradation, deadenylation regulatory machinery for transcript degradation, RNA editing proteins for A-to-I editing, and translational activation machinery for enhanced protein production to transcripts in a gRNA-dependent manner. Taken together, this work validates the CIRTS strategy as a viable approach to engineer RNA effector proteins that are small and assembled from human parts.

3.2 Results

3.2.1 Design of a CRISPR-Cas-inspired RNA-targeting system (CIRTS)

While DNA-targeting Cas9-based systems employ complex biophysical mechanisms to unwind DNA and anneal to a target sequence[141–143], mechanistic studies of Cas13 showed that RNA targeting is initiated by a central seed region in the gRNA[9,144–146]. Additionally, Cas13 systems display substantial variability in sequence context targetability on individual transcripts[59,60,62,147]. Together these findings suggest that sequence complementarity between the gRNA and targeted transcript, as well as the accessibility of a given site, are key requirements for RNA targeting. We sought to engineer a Cas13-inspired system that uses a defined protein-RNA interaction to display a gRNA sequence to deliver protein cargoes to a target RNA, similar to previous RNA tethering assays with overexpressed reporter constructs[49]. Indeed, hairpin-binding proteins and covalent RNA fusions have been used to deliver RNA editing machinery to transcripts[148–150].

Based on the current characterization of Cas13[59,60,62,140,144,146,151], we reasoned that a minimal programmable RNA-targeting system will need four components: (1) an RNA hairpin-binding protein that serves as the core of the system and is a selective, high affinity binder to a specific RNA structure displayed on an engineered gRNA, (2) a gRNA that features both the structure that interacts with the engineered hairpin-binding protein and a sequence with complementarity to the target RNA of interest, (3) a charged protein that could bind to the displayed gRNA sequence non-specifically to stabilize and protect the guiding RNA prior to target engagement, and (4) an effector protein, such as a ribonuclease or epitranscriptomic regulator, that acts on the targeted RNA in a proximity-dependent manner (**Figure 3.1**).

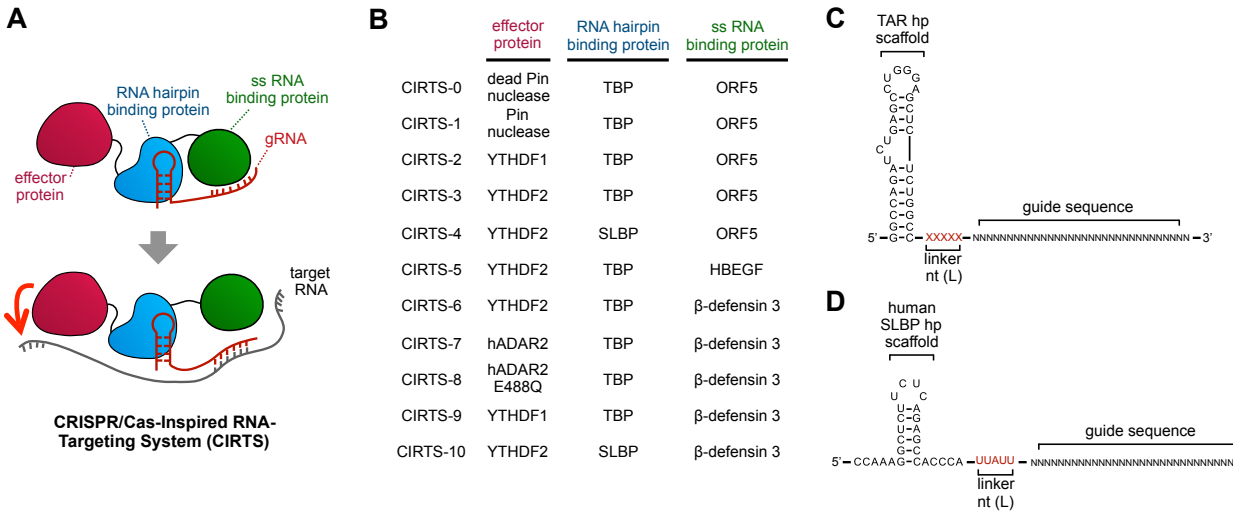

Figure 3.1 Design of a CRISPR-Cas-inspired RNA targeting system (CIRTS). **(A)** Schematic overview of the design strategy. CIRTS is composed of a ssRNA binding protein, an RNA hairpin binding protein, an effector protein, and a guiding RNA. **(B)** List of key CIRTS used in this work. **(C)** Design of the guiding RNA for TBP6.7. The HIV TAR hairpin was fused to a nucleotide linker (L) and a guide sequence. The nucleotide linker was altered during optimization (as described in the supporting figures), but L = UUAUU was used for all work thereafter. **(D)** Design of the guiding RNA for the RNA recognition motif (RRM) of SLBP. The human histone mRNA hairpin was fused to a flexible five nucleotide linker and a guide sequence.

While Cas13 houses all of these functional components in a single protein domain [144,146], we envisioned engineering a system that combines multiple protein domains that each perform one of these functions, which we termed CRISPR-Cas-inspired RNA targeting system (CIRTS). CIRTS vary in their module composition and are uniquely numbered as listed in Figure 3.1A and Figure 3.2.

	effector protein	linker	RNA hairpin binding protein	ss RNA binding protein
CIRTS-11	cytoplasmic Pin nuclease	$(GGS)_6$	TBP	N/A
CIRTS-12	cytoplasmic Pin nuclease	$(GGS)_6$	TBP	ORF5
CIRTS-13	Pin nuclease	$(GGS)_3$	TBP	ORF5
CIRTS-14	Pin nuclease	L8	TBP	ORF5
CIRTS-15	Pin nuclease	helical	TBP	ORF5
CIRTS-16	YTHDF2	L8	TBP	ORF5
CIRTS-17	YTHDF2	helical	TBP	ORF5
CIRTS-18	YTHDF2	$(GGS)_6$	PP7	ORF5

Figure 3.2 CIRTS list continued from Figure 3.1B.
Reference list of all remaining CIRTS used in this work.

3.2.2 Development and in vitro validation of CIRTS-1

For our first-generation system, CIRTS-1, we used an evolved version of the human hairpin-binding protein U1A protein (TBP6.7), which was previously engineered to bind the HIV trans-activation response (TAR) hairpin and has no endogenous human RNA hairpin targets[152,153] (**Figure 3.1B**). We designed a gRNA that includes the TAR hairpin, a nucleotide linker sequence (L), and then a guiding sequence (**Figure 3.1C**). To develop and validate the system, we first engineered a programmable ribonuclease by fusing TBP6.7 to the Pin nuclease domain of human nonsense-mediated mRNA decay factor SMG6, which has been previously used as a non-specific proximity-dependent RNA endonuclease[5,55]. Although this simplest design already displayed promising gRNA-mediated transcript degradation in cell-based luciferase assays (**Figure 3.6A**), the performance was quite poor, which we attributed to the potential degradation of the displayed guiding sequence. The protein surface and hairpin channel of Cas13 systems tend to be highly charged, likely to

non-specifically bind and stabilize the guiding RNA sequence[144]. To engineer this RNA protection function into our system, we added a non-specific, low affinity single-stranded RNA binding protein (ss RNA binding protein) to our construct. However, the human proteome did not readily contain an annotated small, non-specific single-stranded RNA binding protein to the best of our knowledge. Therefore, we developed CIRTS-1 using a small viral ssRNA binding protein, ORF5[154]. Altogether, CIRTS-1 is a protein fusion complex composed of ORF5-TBP6.7-Pin nuclease domain along with a corresponding gRNA (**Figure 3.1B**).

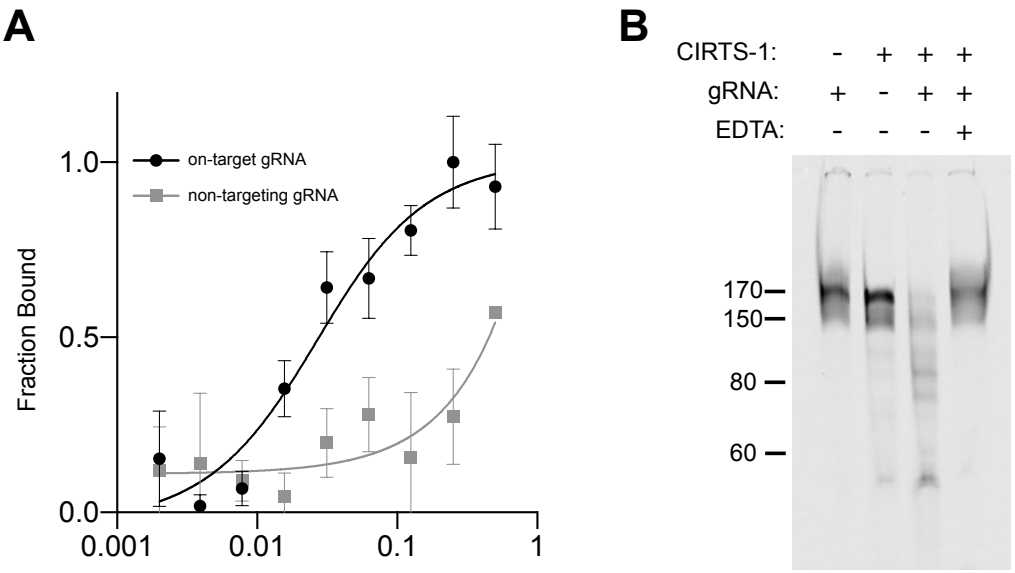

Figure 3.3 CIRTS-1 in vitro binding and RNA cleavage assays.
(A) Filter binding assay evaluating the binding affinity of MBP-CIRTS-1 with on-target gRNA and non-targeting RNA complex to a labeled RNA substrate. Fitting the data to a quadratic binding equation revealed an apparent KD of 22±7 nM for the on-target gRNA:protein complex and an apparent KD around 500 nM for the non-targeting gRNA:protein interaction to the same substrate RNA. **(B)** Cleavage assay run on a 10% denaturing Urea PAGE gel in presence of 0.5 mM $MnCl_2$. An IR800-labeled RNA substrate is cleaved in a gRNA-dependent manner.

We first characterized gRNA-dependent RNA binding and ribonuclease activity of CIRTS-1 *in vitro* on model RNA target substrates. The Pin nuclease domain was previously shown to be active in the presence of Mn^{2+} and activity could be quenched by the addition of EDTA[5]. Directly overexpressing CIRTS-1 led to insoluble protein, which we

resolved by fusing an N-terminal MBP tag to CIRTS-1 (MBP-CIRTS-1). Using purified MBP-CIRTS-1 protein and gRNAs in filter binding assays, we found MBP-CIRTS-1 binds a target RNA (STAR Methods) in a gRNA-dependent manner with an apparent binding dissociation constant (K_D) of 22 nM (**Figure 3.3A**). Critically, if we provide the system with a non-targeting gRNA, we see about 50-fold weaker binding (**Figure 3.3A**). Moreover, in a cleavage assay, we found that MBP-CIRTS-1 cleaves an RNA substrate in a gRNA- and Mn^{2+}-dependent manner, confirming the activity of the Pin ribonuclease domain in the fusion context (**Figure 3.3B**). Collectively, these in vitro results validate the design principles behind CIRTS-1 and motivated us to optimize the system for use in live cells.

3.2.3 Optimization of CIRTS-1

To test the target nuclease activity of CIRTS-1 in live mammalian cells, we established a dual luciferase reporter assay that reports on gRNA-dependent transcriptional changes on a target firefly luciferase (Fluc) RNA (**Figure 3.4A**). Using this system, we optimized the deployment of the Pin nuclease CIRTS by assaying different protein linker types (**Figure 3.5A**), gRNA structures and lengths (**Figure 3.5C**), and CIRTS nuclease cellular localization (**Figure 3.6B**) using a gRNA targeting site we previously found to be effective for Cas13-based knockdown[63]. For this first-generation design, we assayed three linker types with different rigidities between the hairpin-binding protein and the effector domain to assess which one positioned the effector protein best on the target strand while keeping the linker between the ssRNA binding protein and the hairpin binding protein constant. Additionally, we tested different numbers of linking nucleotides (L) on our gRNA (denoted 'XXXXX' in **Figure 3.1C**), ultimately settling on a linker composed of -UUAUU- between the hairpin structure and guiding sequence. We found that a long flexible linker between the hairpin-binding protein and 40 nucleotide long gRNA resulted in the best knockdown efficiency.

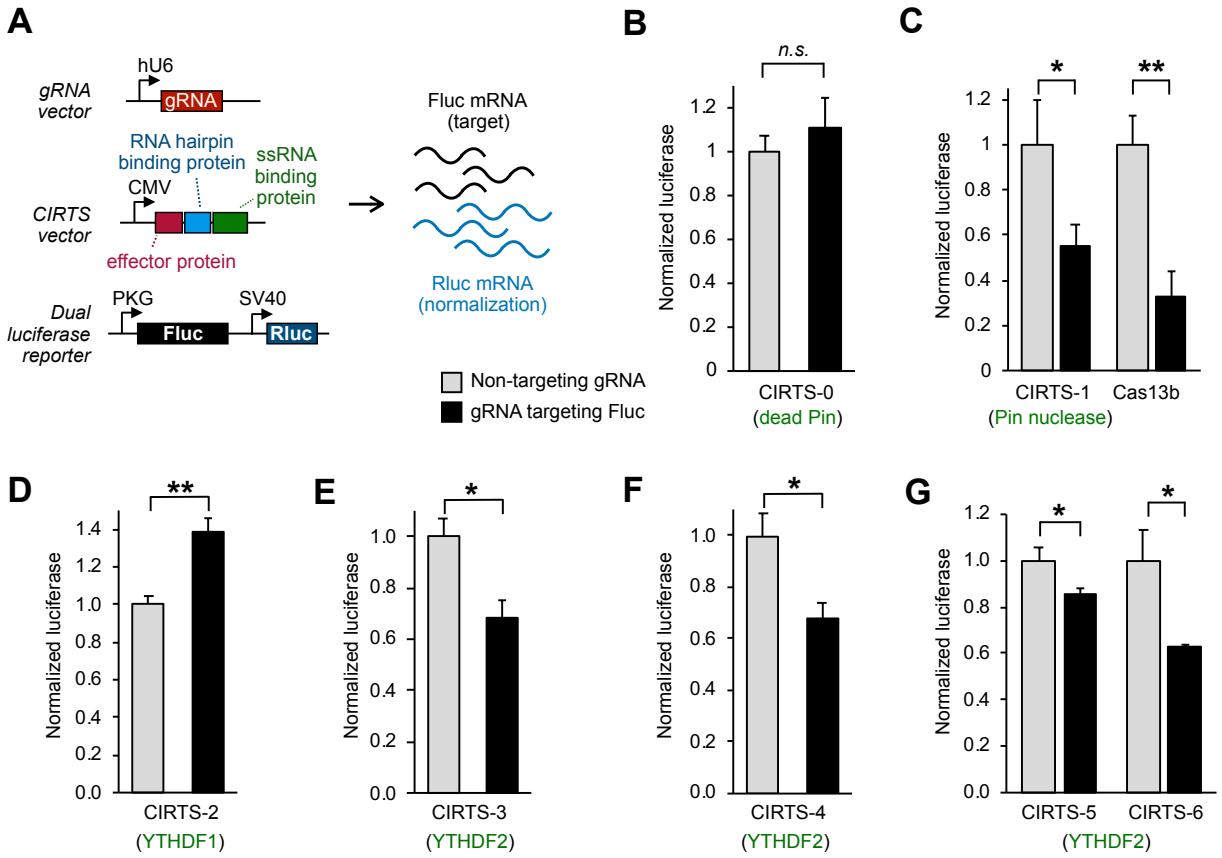

Figure 3.4 CIRTS mammalian cell reporter assays.
(A) General overview of the dual luciferase assay. A reporter construct that contains both firefly luciferase and Renilla luciferase is used in all assays. We targeted CIRTS to the firefly luciferase transcript, while using Renilla luciferase as an internal control. For all subsequent assays, HEK293T cells were transfected with the reporter vector, a CIRTS vector, and a gRNA vector. **(B)** Catalytically inactive CIRTS-0 was used as a control. After 48 h of incubation, we observed no decrease in protein readout. Values shown as mean ± SEM with n = 3 biological replicates. **(C)** Comparison of CIRTS-1 with Cas13b nuclease. Cells transfected with either CIRTS-1 or Cas13 and the corresponding gRNA targeting Fluc show reduced protein levels. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05, **P < 0.01. **(D)** HEK293T cells transfected with CIRTS-2 show an increase in protein level after transfection. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: **P < 0.01. **(E)** HEK293T cells transfected with CIRTS-3 show the anticipated decrease in protein level. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05. **(F)** Switching the hairpin-binding protein to SLBP still results in decrease protein levels after 48 h of transfection. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05.

Figure 3.4 *(continued)* CIRTS mammalian cell reporter assays.
(G) Cells transfected with a fully humanized CIRTS (CIRTS-5 and CIRTS-6) system and an on-target gRNA for firefly luciferase result again in decreased protein levels. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05.

We designed gRNAs that target the firefly luciferase mRNA in the dual-luciferase reporter assay for both CIRTS-1 and Cas13b, as well as control, non-targeting gRNAs (targeting a lambda phage sequence) for each programmable nuclease. To test whether binding to the transcript alters protein levels, we engineered CIRTS-0, which contains a previously reported deactivating mutation in the Pin nuclease domain of CIRTS-1[155], serving as a negative control. We found that CIRTS-0 has no significant effect on the expression level of the target transcript (**Figure 3.4B** and **Figure 3.6C**), indicating binding by the CIRTS ribonucleoprotein to a target RNA is minimally perturbative to the targeted transcript. Additionally, we verified that gRNA binding alone does not introduce detectable target RNA degradation (**Figure 3.6A**). Next, we tested whether CIRTS-1, containing an active nuclease, could mediate degradation of the target. Indeed, we found gRNA-dependent degradation of the target Fluc mRNA, measured at both the protein level as monitored by luciferase activity (**Figure 3.4C**) and the mRNA levels as monitored by RT-qPCR (**Figure 3.6D**). Both results suggest the CIRTS strategy is a viable method in live cells After optimization, we compared the ability of optimized CIRTS-1 (Pin nuclease) to degrade the target reporter RNA to the Cas13b system[60] (**Figure 3.4C**). Although CIRTS-1 is less efficient at targeting the reporter gene as compared to Cas13b, the performance was not dramatically different, especially considering that Cas13b systems have evolved to perform this knockdown function. Encouraged by the performance of CIRTS-1 (Pin nuclease), we next sought to assess the versatility of the design by testing whether each component of the system, including the gRNA, hairpin binding domain, non-specific RNA binding domain, and effector protein, could be swapped for other parts to achieve CIRTS with diverse functions.

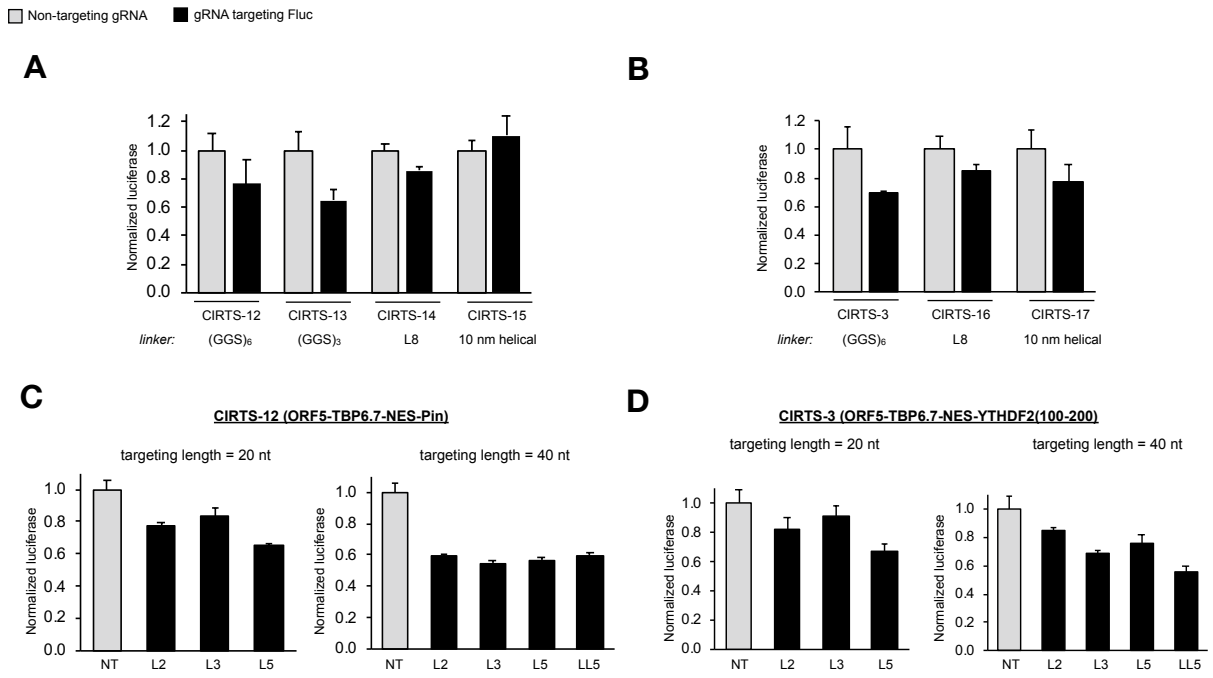

Figure 3.5 CIRTS linker and gRNA optimization. Related to Figure 3.4.
(A) Luciferase assay with the CIRTS nuclease system using different linkers between the hairpin-binding protein and the effector protein. Previously published L8 = SGSETPGTSESATPES[156], 10 nm helical linker = EEEEKKKQQEEEAERLRRIQEEMEKERKRREEDEKRRRKEEEERRMKLEMEAKRKQEEEERKKREDDEKRKKK. **(B)** Luciferase assay with CIRTS-YTHDF2-mediated decay using different linkers between the hairpin-binding protein and the effector protein. **(C)** Different engineered gRNA for TBP6.7 based on the design shown in Figure 3.1C. Two different targeting lengths of 20 and 40 nucleotides were used in combination with different numbers of linking nucleotides (L) between the hairpin and the guiding sequence. The dual luciferase assay was used to assess nuclease-mediated decay. NT = non-targeting, Fluc gRNA containing different linker nuclease (Figure 3.1), L2 = UU, L3 = UUU, L5 = UUAUU. **(D)** The same engineered gRNAs as in Figure 3.5C were used with CIRTS-3 to induce epitranscriptome-induced RNA decay. NT = non-targeting, Fluc gRNA containing different linker nuclease (Figure 3.1), L2 = UU, L3 = UUU, L5 = UUAUU. n = 3 biological replicates.

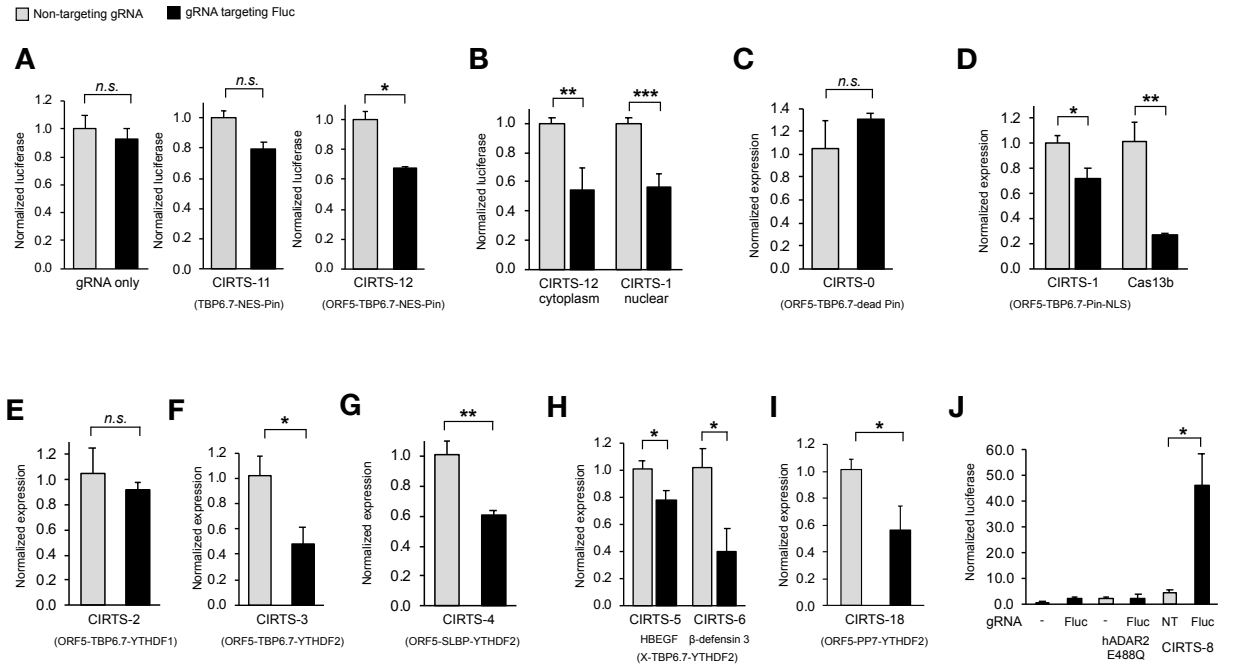

Figure 3.6 Control luciferase assays and RT-qPCRs. Related to Figure 3.4.
(A) Luciferase assay comparing transfection of gRNA only to nuclease-mediated decay of TBP6.7-Pin nuclease domain without (CIRTS-11) and with (CIRTS-12) the additional ssRNA binding protein ORF5. **(B)** CIRTS nuclease can mediate decreases in RNA and therefore protein level in both the nucleus and the cytoplasm (n=6). **(C)** RT-qPCR analysis of RNA levels with the 'dead' Pin nuclease domain CIRTS (CIRTS-0). **(D)** Comparison of RNA levels when cells were transfected with CIRTS-1 and active Cas13b nuclease. CIRTS-1-Pin mediated RNA cleavage showed substantially less RNA degradation compared to the Cas13b system. **(E-H)** All engineered CIRTS system we tested in the dual luciferase assay were also subjected to RT-qPCR analysis to assess changes in RNA levels. CIRTS-2, which contain the YTHDF1 effector domain inducing translation activation showed no significant changes in RNA level while all YTHDF2-containing CIRTS show the expected decrease in RNA levels. **(I)** Engineered CIRTS-18 containing the PP7 dimer as the hairpin binding protein. Knockdown of PPIB after transfection with CIRTS-18 as measured by qPCR. **(J)** Comparison of reporter only, gRNA only, non-TBP6.7 fused hADAR2(E488Q) in the presence of NT or Fluc gRNA, and reporter with CIRTS-8 (hADAR E488Q) with non-targeting or targeting gRNA (3.6F and 3.6H: n = 2 or 3). n = 3 biological replicates unless otherwise noted. Student t-test: *P < 0.05, **P <0.01, ***P <0.001.

3.2.4 Modularity of CIRTS

To explore the versatility of the CIRTS design, we first assayed whether CIRTS could deliver RNA epitranscriptomic regulatory "reader" proteins, which we previously delivered using the dCas13b system[63]. For our study, we chose to focus on regulatory proteins of N^6-methyladenosine, the most prevalent mRNA modification. On average each transcript contains three modifications sites with high m^{6}A abundance detected in the 3'UTR, and m^{6}A has been shown to have regulatory roles in splicing[15], translation[16,17], and stability[28,112]. We exchanged the Pin nuclease effector protein of CIRTS-1 for the N-terminal domain of the YT521-B homology domain family protein 1 (YTHDF1), a cytoplasmic m^{6}A reader protein that recruits the translation machinery[17], to generate CIRTS-2. Note that CIRTS-2 does not include the C-terminal YTH domain of YTHDF1 that recognizes m^{6}A. When CIRTS-2 (YTHDF1) is delivered to the same target sequence as the CIRTS-1 experiments, the RNA levels are relatively unchanged (**Figure 3.6E**), but a significant increase in protein levels from the RNA is generated (**Figure 3.4D**), consistent with the previously reported YTHDF1 activity[17]. We then exchanged the YTHDF1 fragment for a fragment of YTHDF2, an m^{6}A reader protein that recruits the RNA deadenylation machinery and induces RNA degradation[28,112], to generate CIRTS-3. Delivery of CIRTS-3 (YTHDF2) to the reporter mRNA induces degradation of the target transcript as measured by both RNA (**Figure 3.6F**) and protein levels (**Figure 3.4E**). CIRTS-1 through -3 demonstrate the versatility of the design strategy to deliver a range of effector protein cargoes to target RNA in live cells.

After demonstrating the modularity of the effector domain, we set out to assess if other human parts could also be used for the RNA hairpin binding domain and non-specific ssRNA binding protein. We replaced TBP6.7 in CIRTS-3 (YTHDF2) with the RNA hairpin binding domain of the human histone stem loop binding protein (SLBP) to generate CIRTS-4 (YTHDF2). Concurrently, we designed a gRNA based on the histone mRNA stem loop structure (**Figure 3.1D**). Assaying CIRTS-4 (YTHDF2) in the reporter assay (**Figure 3.4F**)

and by RT-qPCR to assess RNA levels (**Figure 3.6G**) revealed similar performance as CIRTS-3, confirming other hairpin binding domains can be used as the core of the CIRTS.

Next, we sought to engineer entirely humanized versions of the CIRTS system. As stated earlier, we designed the initial proof-of-concept systems based on the viral non-specific, single-stranded RNA binding protein, ORF5. Although there are no annotated human single-stranded, non-specific RNA binding proteins, we reasoned highly charged, cationic human proteins could fulfill the role of ORF5 in the CIRTS system[157]. We therefore engineered HBEGF and β-defensin 3, two cationic human proteins, in the place of ORF5 in CIRTS-3 to generate CIRTS-5 and CIRTS-6, respectively. Again, deploying these programmable effectors in the luciferase reporter assay revealed gRNA-dependent degradation of the target gene (**Figure 3.4G and 3.6H**) mediated by the YTHDF2 epitranscriptomic regulation.

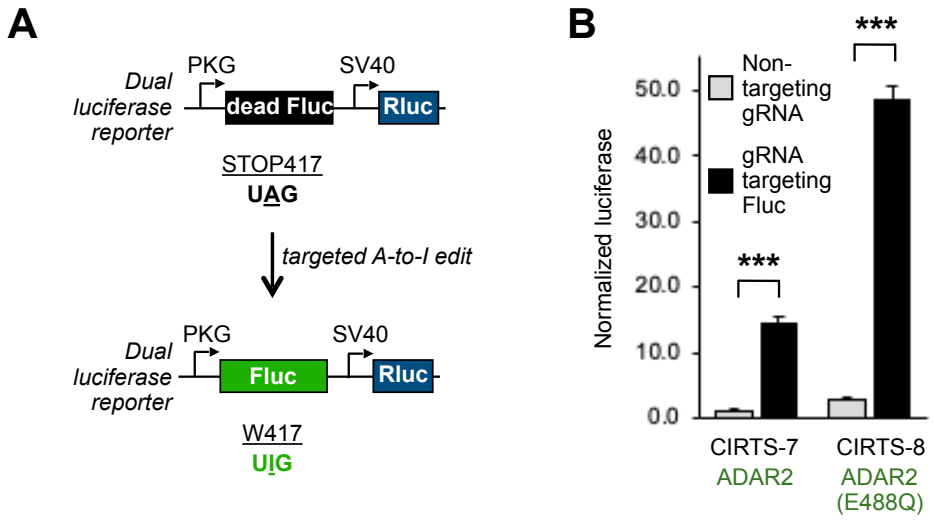

Figure 3.7 CIRTS for RNA editing.
(A) Schematic overview of the RNA editing reporter assay used. A single G-to-A mutation was introduced in the coding sequence of firefly luciferase resulting in a W417X (X = STOP) codon switch and no measurable firefly luciferase signal (Figure 3.6J). **(B)** Delivery of CIRTS-7 (hADAR2 wt) and CIRTS-8 (hADAR E488Q) with an on-target gRNA shows significant RNA editing that results in measurable firefly luciferase signal. Both the background and the editing efficiency of CIRTS-8, the hyperactive hADAR2 mutant, are found to be higher compared to wildtype hADAR2. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: ***P < 0.001.

Finally, we used CIRTS to deliver the catalytic domain of human ADAR2 (hADAR2) to RNA transcripts to confirm CIRTS' versatility in scope of functions with an additional effector protein. We designed a dual luciferase reporter that contains a G-to-A mutation in the coding region of firefly luciferase resulting in a premature stop of translation and no measurable firefly luciferase activity (**Figure 3.7A and 3.6J**). We then deployed CIRTS to deliver wt hADAR2 (CIRTS-7) or hADAR2 E488Q (CIRTS-8), a known hyperactive mutant of hADAR2[158], to the mutated position, which resulted in gRNA-dependent rescue of luciferase activity in both cases (**Figure 3.7B**). The hyperactive hADAR2 mutant showed higher editing efficiency and a higher background in the absence of an on-target gRNA based on luciferase assay. However, using the hyperactive mutant could be beneficial to allow targeting of a wider substrate scope as it has relaxed sequence constraints[158]. To verify that the observed editing signal was indeed gRNA and hairpin-binding protein-dependent, we transfected cells with the gRNA alone or with gRNA and TBP6.7-lacking hADAR2 construct. In our assay, we only observed substantial editing in the presence of both gRNA and our CIRTS editor (**Figure 3.6J**). Collectively, the performance of these various CIRTS in the reporter assays demonstrates the modularity of the CIRTS design, including the hairpin-binding domain and corresponding gRNA, the single-stranded RNA binding protein, and the effector protein.

3.2.5 Targeting endogenous mRNAs with CIRTS

We next sought to assess whether the CIRTS could deliver an effector protein to a target endogenous transcript, using the CIRTS-1 programmable nuclease and CIRTS-3 programmable YTHDF2-mediated decay systems as exemplars. We selected five RNA transcripts that have been previously validated as Cas13 targets, reasoning that these are accessible for RNA targeting by programmable RNA-binding systems. We then designed gRNAs for each target, using the same binding sites on the targets that were previously used in Cas13 experiments[59,62]. We assayed the effects of the CIRTS on RNA levels of

each target by RT-qPCR. When cells were transfected with either CIRTS-1 or CIRTS-3, along with a specific gRNA expressing vector, we observed a significant decrease in RNA level by RT-qPCR for each of the five endogenous mRNA transcripts (**Figure 3.8A and B**). In addition to targeting mRNA, we also verified that we can target other RNA species such as lncRNA by targeting CIRTS-1 (Pin nuclease) to MALAT1 (**Figure 3.9A**). The relative knockdown efficiency varied for each gene, which is also observed in other RNA-targeting systems and is potentially mediated by accessibility, differences in gRNA expression and composition, or other regulatory pathways specific to each gene. Nonetheless, these results confirm that CIRTS can target endogenous transcripts and mediate decay through either active nuclease activity on the target or by triggering endogenous epitranscriptomic regulatory pathways.

Next, we set out to assess whether CIRTS-2 could trigger protein production of an endogenous transcript through a YTHDF1-mediated epitranscriptomic pathways. We selected an abundant transcript PPIB with a reported, reliable antibody for analysis of CypB (the protein product of PPIB) protein production by western blotting. Indeed, cells transfected with CIRTS-2 and an on-target gRNA showed an increase in protein level (**Figure 3.8C and 3.9C**) without a change in RNA levels (**Figure 3.9B**), consistent with prior reported YTHDF1 effects on the transcript[17]. As a control, the same experiment performed with CIRTS-3, which delivers YTHDF2, results in slight decrease in protein levels, which correlates with the decrease in mRNA levels (**Figure 3.8B and 3.9C**).

Finally, as a first test of transcript position-specific effects, we tiled gRNAs along the SMARCA4 mRNA and tested YTHDF2-mediated decay by CIRTS-3. We found dramatically different performance of the system depending on where the gRNA lands on the targeted mRNA (**Figure 3.8D**), which is likely the result of both CIRTS binding accessibility and the regulatory protein sequence requirements. Taken together, these experiments show that the CIRTS platform is functional on endogenous transcripts in a gRNA-dependent manner, and can actively degrade a target transcript, trigger degradation ma-

chinery to act on the target transcript, or activate translation and increase protein production from the target transcript.

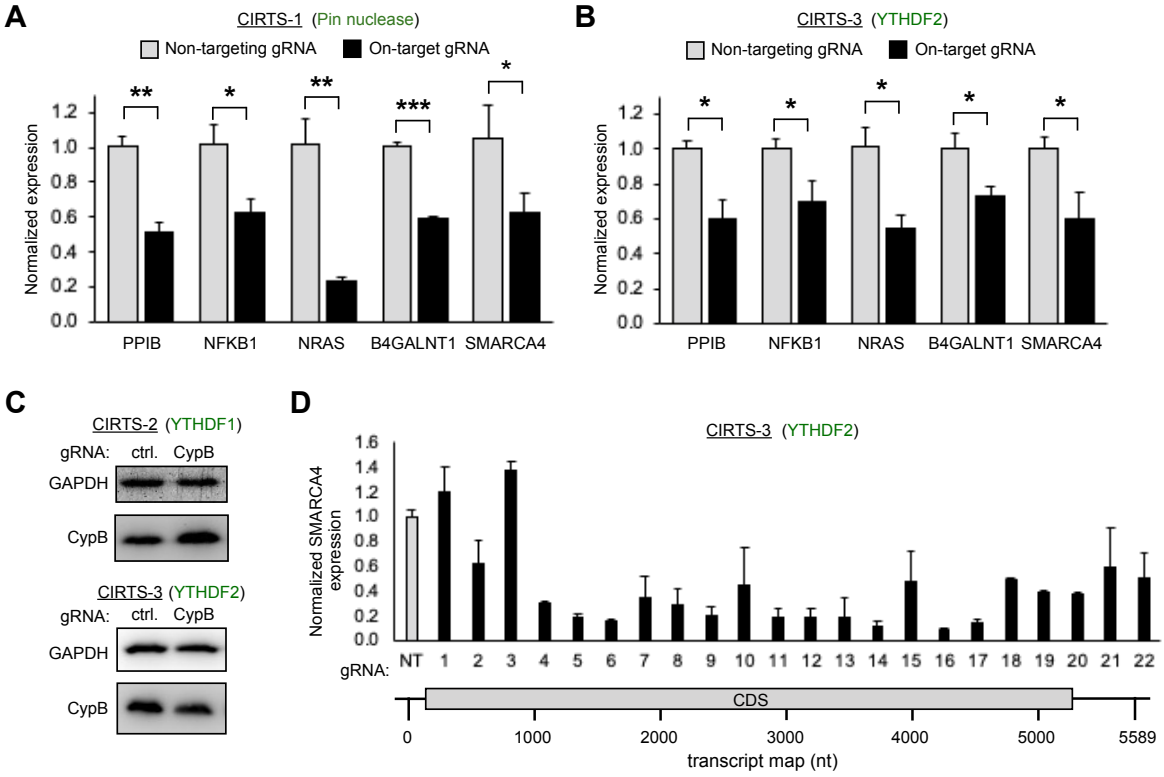

Figure 3.8 Targeting endogenous transcripts with CIRTS.
(A) Nuclease-mediated knockdown of five endogenous transcripts upon transfection of cells with CIRTS-1 as assayed using qPCR. CIRTS-1 can be used to target endogenous transcripts of interest by co-transfecting a gRNA with the corresponding on-target guiding sequence. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05, **P <0.01, ***P <0.001. **(B)** RT-qPCR analysis of YTHDF2-mediated knockdown of endogenous transcripts with CIRTS-3. Cells transfected with CIRTS-3 show gRNA-dependent decreases in RNA level for all five transcripts tested. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05. **(C)** Analysis of protein levels after transfection with CIRTS-2 or CIRTS-3 by Western blot. CIRTS-2 induces an increase in protein levels, whereas CIRTS-3 shows the expected decrease in protein levels, both in a gRNA-dependent manner. **(D**) gRNA screen along SMARCA4 using CIRTS-3 to induce gRNA-dependent RNA decay. We see significant changes in the amount of induced decay dependent on where the transcript is targeted (n = 2 or 3).

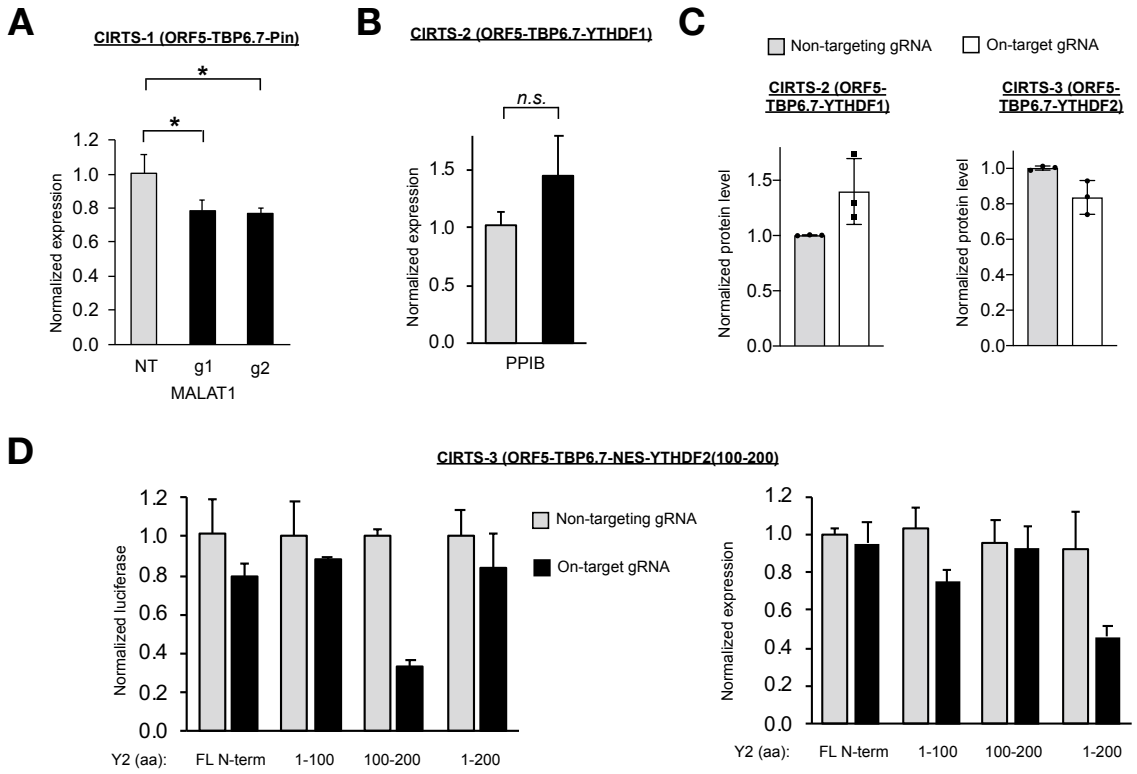

Figure 3.9 Control qPCR, Western Blot, YTHDF2 truncations. Related to Figure 3.8. **(A)** CIRTS-1 can be delivered to RNA species other than mRNA. As a proof-of-principle, we transfected cells with CIRTS-1 (Pin nuclease) and two different gRNAs for the lncRNA MALAT1 and assessed RNA levels by RT-qPCR. **(B)** RT-qPCR analysis of RNA level when cells were transfected with CIRTS-2. As anticipated, no significant changes in RNA level were observed when a YTHDF1-containing protein was used. **(C)** Quantification of protein levels as measured by Western blot when cells were transfected with CIRTS-2 or CIRTS-3 and targeted to PPIB (n=3). **(D**) Different truncations of YTHDF2 were assayed to determine which would be more efficient. We compared luciferase data (left) with qPCR data (right) and concluded to use the Y2(100-200) construct for luciferase analysis and the Y2(1-200) construct for endogenous targeting to enable the best possible quantifications of our tools. n = 3 biological replicates unless otherwise noted. Student t-test: *P < 0.05, **P <0.01.

3.2.6 Targeting specificity of CIRTS

To gain insights into how specific CIRTS is at targeting RNA substrates, we designed a series of experiments that address the sensitivity of the gRNA to mismatches, transcriptome-wide off-targets, and endogenous substrate targeting. To assess mismatch toler-

ance, we designed a luciferase-based mismatch experiment that allows us to assay targeting effects when introducing one, two, or three mismatches into the duplex formed between gRNA and target RNA. We chose to fuse the disease-relevant KRAS4b transcript to our luciferase reporter and asked whether our engineered system can differentiate between the cancer-associated G12D (target 1), the wild type (target 2), the G12C (target 3), and a G12W (target 4) KRAS4b variants (**Figure 3.10A**). We found that CIRTS yields comparable knockdown of the G12D and wild type variants indicating that one mismatch does not cause large changes in targeting specificity (**Figure 3.10B**). However, when we targeted the system to the G12C and G12W reporters, which contain two and three mismatches respectively, CIRTS knockdown efficiency decreased.

We next assessed whether increasing the gRNA length could affect the mismatch tolerance of CIRTS, focusing on mismatches in the center region of the duplex formed between gRNA and target RNA as they showed the largest effect on knockdown efficiency in our assay. As observed with the shorter 20 nt gRNA, we see no difference in knockdown efficiency when we target a reporter with no or one mismatches. However, a longer 40 nt gRNA can rescue some of the effects when the two-mismatch variant was targeted, indicating that the gRNA length contributes to the efficiency of the system (**Figure 3.10D**). As a comparison to existing technology, we subjected Cas13b to the same mismatch assay, which showed that Cas13b is less sensitive to mismatches in general. Targeting Cas13b to reporters with one and two mismatches yielded little change in knockdown efficiency, while three mismatches led to a substantial decrease in knockdown efficiency (**Figure 3.10C**). Both Cas13b and CIRTS are most sensitive to mismatched base-pairing in the center of the duplex formed between gRNA and target RNA, a finding that agrees well with previous studies of Cas13b[59,60].

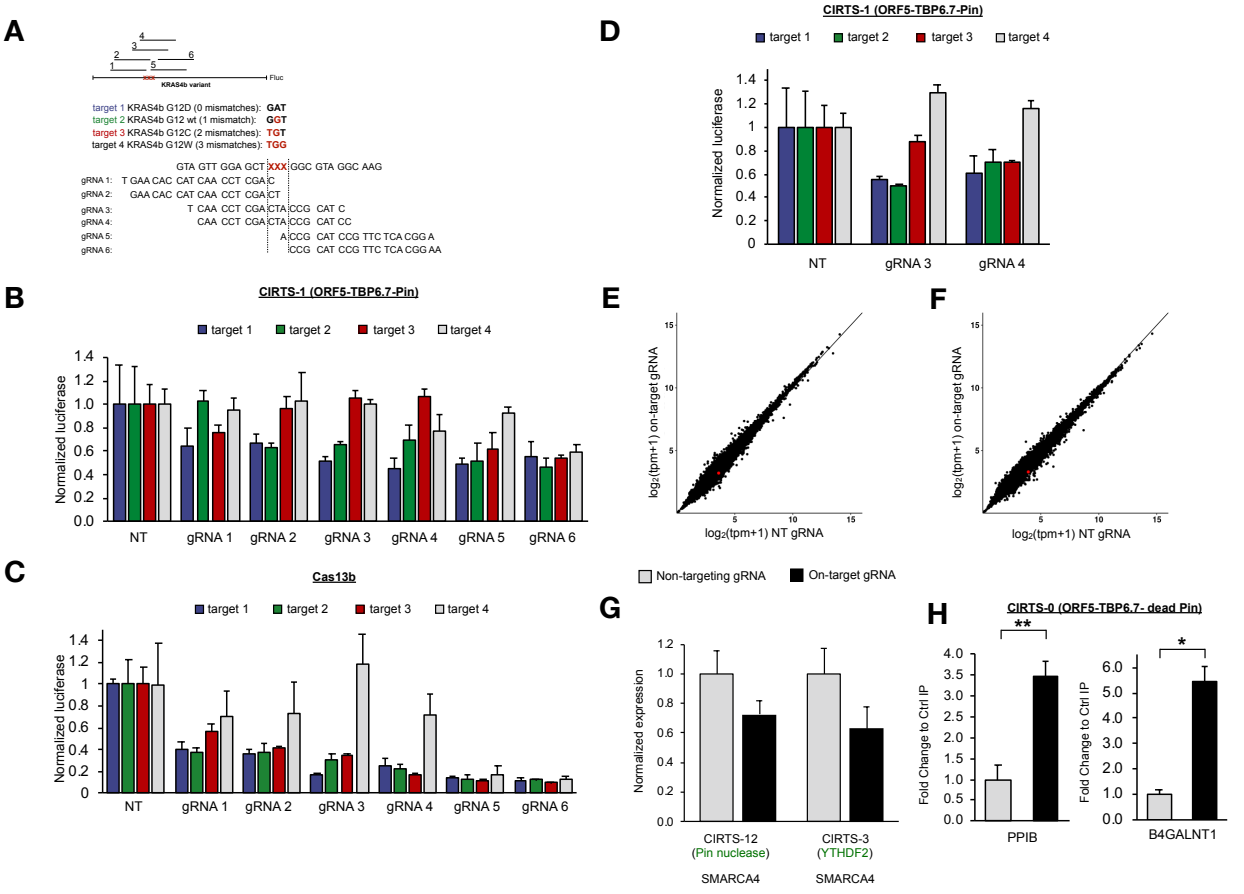

Figure 3.10 Targeting Specificity of CIRTS.

(A) Schematic of the KRAS4b-luciferase mismatch reporter assay. We chose four KRAS4b variants that have an increasing number of mismatches to the designed 20 nt length gRNA and fused it N-terminal to the dual luciferase reporter. **(B)** CIRTS-mediated knockdown of KRAS4b-Fluc with different numbers of mismatches between the gRNA and target RNA as described in Figure 3.10A. CIRTS was found to be most sensitive to mismatches in the middle of its guiding sequence. **(C)** Cas13b-mediated knockdown in the same KRAS4b-Fluc reporter assay as described above. Cas13b shows a higher knockdown efficiency but is also less sensitive to mismatches introduced. Similar to CIRTS, Cas13b is knockdown is most affected by mismatches at the center of the guiding target duplex region. **(D)** Knockdown efficiency of CIRTS on the KRAS4b-luciferase mismatch reporter when using a 40 nt gRNA length. A longer guiding sequence in the gRNA can rescue some of the loss in knockdown efficiency. **(E-F)** Mean expression levels of the transcriptome in log2(transcript per million (TPM) +1) when CIRTS Pin nuclease (E) or CIRTS YTHDF2 (F) are deployed to SMARCA4 (in red) in cells (n=3). Pearson's correlation: 0.990 (E) and 0.991 (F). **(G)** Knockdown levels of SMARCA4 as determined by RNA sequencing. **(H)** Cells were transfected with CIRTS-0-3xFLAG and a gRNA for either PPIB, B4GALNT1, or NT. After crosslinking and FLAG IP, pulled down RNA was quantified using RT-qPCR. Reactions containing on-target gRNA for either transcript showed 3.5 to 5-fold enrichment for these transcripts, indicating guided RNA targeting (n = 2 or 3).

To assess transcriptome-wide off-targets, we subjected our system to RNA sequencing. We assayed effects of the CIRTS Pin nuclease and CIRTS YTHDF2 targeting the endogenous transcript SMARCA4 (**Figure 3.10E and F**). In both cases, we find no statistically significant off-targets. However, while we see knockdown of the targeted transcript and even statistically significant knockdown by CIRTS-3 (YTHDF2) when we look at the target transcript only (pval <0.1), the knockdown levels do not fall into a statistically significant region when evaluated in a transcriptome-wide manner (qval < 0.1) (**Figure 3.10G**). However, together the results of our mismatch assay and the fact that we observed no statistically significant gRNA-dependent off-targets indicate that the selective knockdown efficiency can be further optimized in future studies.

To verify CIRTS bind the transcript of interest, we furthermore performed RNA immunoprecipitation followed by RT-qPCR. We designed gRNAs to target two endogenous transcripts that were previously targeted by Cas13 systems, PPIB, and B4GALNT1[60,62]. We separately delivered each gRNA along with CIRTS-0 (dead nuclease CIRTS) fused to a 3x FLAG-tag. We then subjected lysates to immunoprecipitation with an anti-Flag antibody and quantified the relative amounts of each target RNA bound to the protein. Indeed, both endogenous transcripts were enriched between 2.5- and 5-fold in a gRNA-dependent manner (**Figure 3.10H**), confirming CIRTS function as a programmable RNA-guided RNA binding protein on endogenous transcripts.

3.2.7 Multiplexed targeting of multiple endogenous RNAs with CIRTS

The targeting specificity and the modularity of CIRTS inspired us to extend the application of CIRTS in a multiplexed targeting manner. Rather than delivering a single effector protein and targeting a single transcript at a time, we set out to test whether CIRTS can target more than one transcript or deliver more than one effector protein in the same sample. In principle, CIRTS built from the TBP hairpin binding domain and CIRTS built

from the SLBP hairpin binding domain, which each use separately engineered gRNAs (**Figure 3.1C and D**), should be orthogonal to one another, permitting selective targeting of multiple transcripts with either the same or even different CIRTS.

First, we tested whether a single CIRTS can be used to simultaneously target multiple transcripts. We co-transfected cells with CIRTS-6 (YTHDF2) along with three gRNAs targeting PPIB, SMARCA4, and NRAS and assessed changes in RNA level by RT-qPCR. As expected, we observed a decrease in RNA levels for all three targeted transcripts (**Figure 3.11A and B**). However, we observed a slight decrease in efficiency when we deploy several gRNAs or CIRTS in the same cells, which we attribute to the simultaneous transfection of cells with four plasmids.

To test whether two different types of effectors can be used simultaneously, we next deployed both CIRTS-9, a fully humanized version of the YTHDF1 construct, to target firefly luciferase and CIRTS-10 (YTHDF2) to target SMARCA4 (**Figure 3.11C and D**). We find that both proteins are active and induce the anticipated increase in luciferase protein and decrease in RNA levels respectively. Moreover, to further corroborate the orthogonality of multiple-targeting CIRTS, we deployed two CIRTS, CIRTS-6 (TBP6.7) and CIRTS-10 (SLBP) (**Figure 3.12A and B**), but use different hairpin-binding modules, to deliver YTHDF2 to two different endogenous target mRNAs (**Figure 3.13A**). Each CIRTS degraded the target transcript in an on-target gRNA-dependent manner, with minimal crosstalk between the two systems (**Figure 3.13B**).

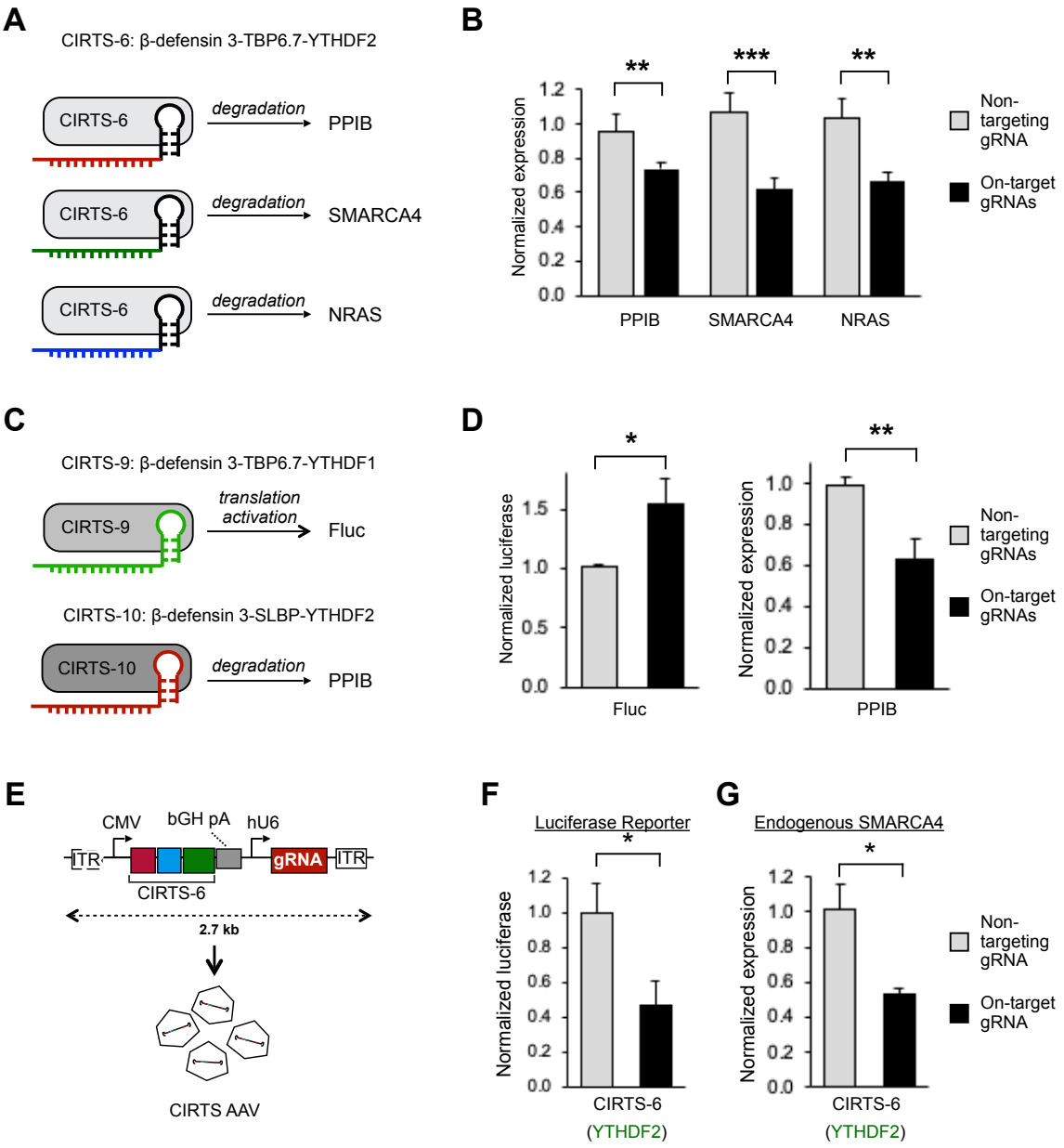

Figure 3.11 Multidimensional targeting and viral delivery of CIRTS.
(A) Schematic of delivery of CIRTS-6 and three gRNAs. **(B)** CIRTS-6 can be delivered with three distinct gRNAs for PPIB, SMARCA4, and NRAS and simultaneously cause knockdown of all three transcripts. n = 5 biological replicates. Student t-test: **P <0.05, ***P <0.001. **(C)** Schematic of simultaneous CIRTS delivery with different effector proteins. **(D)** Changes in luciferase protein levels and PPIB transcript levels when cells were transfected with both CIRTS-9 (YTHDF1) and CIRTS-10 (YTHDF2) and gRNAs for Fluc and PPIB respectively. Both orthogonal CIRTS retain their individual functions and act simultaneously in cells. n = 5 biological replicates. Student t-test: *P < 0.1, **P <0.05.

Figure 3.11 *(continued)* Multidimensional targeting and viral delivery of CIRTS. **(E)** Transfer plasmid for AAV delivery containing both the CIRTS-6 (YTHDF2) as well as the gRNA component of the system. The total insert size between the two inverted terminal repeats (ITR) was 2.7 kb. **(F)** AAV-packaged CIRTS-6 and a gRNA targeting luciferase was delivered to HEK293T cells to knockdown firefly luciferase in the dual luciferase reporter assay. **(G)** AAV-packaged CIRTS-6 and a gRNA targeting SMARCA4 was delivered to HEK293T cells to knockdown the endogenous gene, which revealed efficiency comparable to that achieved by transient transfection. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05.

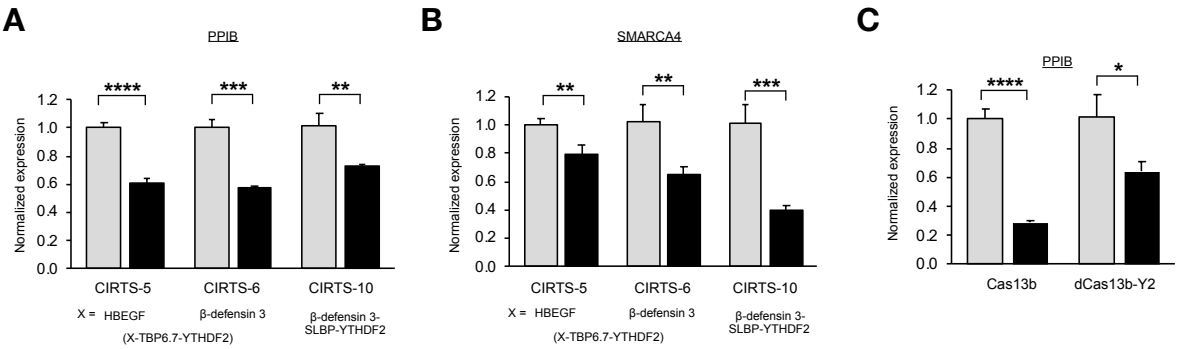

Figure 3.12 Endogenous targeting with CIRTS. Related to Figure 3.11. **(A)** Changes in RNA levels as assessed by RT-qPCR after transfection of CIRTS5-7 alone for PPIB. **(B)** Similar to Figure 3.12A, SMARCA4 levels were assayed when cells were transfected with CIRTS5-7. **(C)** Comparison of knockdown levels of PPIB after delivery of active Cas13b nuclease or an engineered dCas13b-YTHDF2(1-200) construct to PPIB (n = 2 or 3). n = 3 biological replicates unless otherwise noted. Student t-test: *P < 0.01, **P <0.05, ***P <0.01, ****P < 0.001.

At this point, we conclude that the TBP6.7 and SLBP-based CIRTS can each simultaneously target endogenous transcripts in a gRNA-dependent manner. Although not human-derived, we found that other hairpin-binding systems, such as PP7, can also be used to generate CIRTS, suggesting it is possible to generate a range of selective and orthogonal systems (**Figure 3.6I**). CIRTS allow for multiple regulatory proteins to be simultaneously delivered, for example to target one transcript for degradation and another for translational activation, opening up possibilities for cell reprogramming by targeting multiple genes at once in multiple dimensions[159,160].

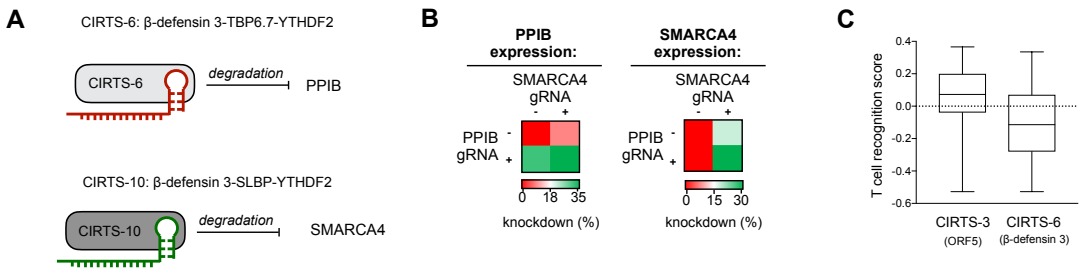

Figure 3.13 Multiplexed targeting with CIRTS. Related to Figure 3.11.
(A) Schematic of vectors used for multiplexed targeting. Cells were transfected with an expression vector for CIRTS-6, and an expression vector for CIRTS-10, an expression vector for a CIRTS-6 gRNA construct targeting PPIB or a non-targeting control, and an expression vector for a CIRTS-9 gRNA targeting SMARCA4 or a non-targeting control. **(B)** Heat map showing knockdown of multiplexed targeting described in (A). When both CIRTS have an on-target gRNA for PPIB or SMARCA4 present, both transcripts can be knocked down in the same samples. Values shown as mean expression level of each target transcript relative to GAPDH, with n = 8 biological replicates. **(C)** Computational prediction of immunogenicity. We first predicted 9-mer peptides that are MHC I binders using the IEDB database and subjected the top one percentile of binders to immunogenicity predictions using the IEDB immunogenicity predictor.

3.2.8 Viral delivery of CIRTS by AAV

Aside from the human-derived nature of CIRTS, another core advantage is the small size of CIRTS, which should permit more efficient viral packaging and delivery. Adenovirus-associated virus (AAV) is a versatile delivery vehicle to deliver transgenes and gene therapies to different cell types due to wide range of serotypes available[79], low immune response stimulation[161], and low risk of genome insertion[79,80]. However, it has been challenging to package and deliver many Cas13 proteins due to a limited packaging capacity of about 4.7 kb[83]. To showcase the possibility of CIRTS to be delivered by AAV, we designed a dual CIRTS-6/gRNA transfer plasmid and packaged it in the AAV delivery vehicle. The total insert, including the CIRTS protein and gRNA, is only 2.7 kb (**Figure 3.11E**). We found that transduction of HEK293T cells with the generated virus recapitulates the knockdown efficiency of CIRTS-6 on both the luciferase reporter as well as an endogenous target

(**Figure 3.11F and G**), confirming viral-delivered CIRTS are still functional and providing a pathway toward clinical deployment. In future applications, one could imagine packing more than one CIRTS into the AAV delivery vehicle to simultaneously target one transcript for upregulation and one transcript for degradation as previously shown by transient transfection.

3.3 Discussion

In summary, here we presented CIRTS, a versatile strategy for engineering programmable RNA effector proteins. CIRTS are small, can be fully humanized, can target endogenous RNAs in live cells, and can work for multidimensional transcriptome control. As research tools, CIRTS should provide advantages to previous methods because of their smaller size. For example, CIRTS-2 and CIRTS-3 are 65 and 36 kDa respectively, while the comparable Cas13b-based programmable YTHDF1 and YTHDF2 systems are 155 and 126 kDa, respectively (**Figure 3.14**). CIRTS-1 is even smaller than the smallest DNA-targeting Cas protein found to date, Cas14a and the smallest Cas12g RNA-targeting protein[58,162].

From a translational perspective, CIRTS should offer several key advantages and opportunities. The humanized nature of the CIRTS will provide a pathway toward avoiding immune responses, opening up the potential for continuously-delivered therapies. While the fusions between the human proteins in the CIRTS present potential limitations in the design where the immune system could respond to[163], this is a problem that can in principle be engineered around. When we computationally predicted the immunogenicity of the highest likelihood MHC I binding peptides in our engineered constructs, we find that the fully humanized CIRTS shows lower propensity to cause immune reactions (**Figure 3.13C**), but further experimental testing is needed to discover where the limitations in the design emerge.

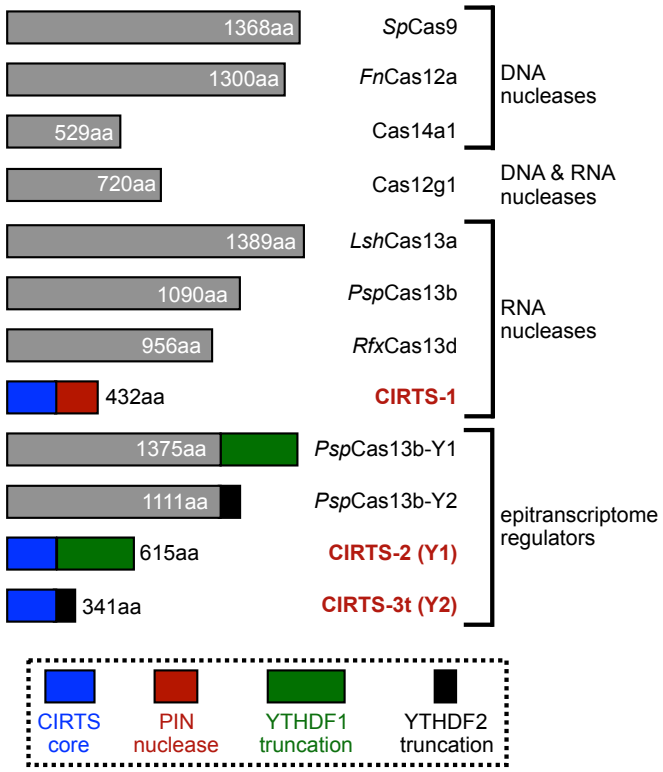

Figure 3.14 Comparing CIRTS to other DNA and RNA-targeting CRISPR-Cas systems.
Schematic size comparison of commonly used Cas9, Cas12, Cas13, and CIRTS.

Several challenges remain with the current CIRTS. First, the alternative hairpin binding protein SLBP in its current form has an endogenous RNA hairpin binding partner, which could influence stem loop RNA trafficking. To minimize endogenous effects of our fusion constructs, we only included the minimal RNA recognition motif (RRM) necessary for hairpin recognition in our system and omitted regions of potential interactions with other proteins or nucleic acids. Likewise, we tried to keep the required RNA hairpin as small as possible to avoid potential endogenous interactions. The stem loop hairpin was already very short and could not be further truncated but we chose to only use the minimally required region necessary for TBP6.7 binding to the TAR hairpin, resulting in a gRNA with less than half the original hairpin length. Second, the cationic peptide, β-defensin 3, in its current form can theoretically still interact with its intracellular binding partners and elicit

unwanted biological responses. However, human β-defensin 3 has been extensively studied[164,165], potentially allowing us to engineer β-defensin 3 mutants that retain the highly charged nature required for CIRTS, but abolish endogenous functions in order to engineer a human part-based, orthogonal RNA targeting system.

From a broader perspective, the CIRTS platform demonstrates the potential of combining parts contained within the human protein toolbox to engineer proteins with new properties. The presented CIRTS were created through minimal protein engineering and optimization efforts, but function nearly as well as the naturally-evolved CRISPR-Cas systems. In particular, when we compared the Cas13b-based knockdown by its endogenous nuclease and by delivering YTHDF2 (**Figure 3.12C**) to CIRTS-mediated knockdown, we find that while the Cas13b nuclease performs substantially better than CIRTS nuclease, we see no difference in knockdown mediated by YTHDF2, suggesting CIRTS-3 (YTHDF2) in its current state can already be used to manipulate the epitranscriptome effectively. Further work optimizing the CIRTS using directed evolution will likely yield variants that have improved performance in mammalian systems[166]. Understanding target site design and effector protein contextual requirements will also improve CIRTS performance, and a better understanding of the epitranscriptomic pathways being exploited will allow us to design better systems. Additionally, there are a range of other regulatory proteins contained within the human proteome that likely house unique RNA control properties[167], which can be coupled with CIRTS to create programmables versions of each protein for both functional characterization and potential translational applications. CIRTS provides an alternative approach for studying and exploiting RNA regulation and will open up future opportunities to intervene in cell regulation for disease treatment.

3.4 Materials and Methods

Escherichia coli strains (10-beta and BL21 (DE3)).

E. coli 10-beta cells were used for all cloning and cultured in Luria-Bertani (LB) broth. *E. coli* BL21 (DE3) cells were used for protein expression for *in vitro* studies. Cells were grown in 2 XYT media to an OD of 0.6 at 37°C before being induced with 0.5 mM IPTG (bioWORLD) and incubated at 16°C for 16 h.

Cell culture of Human Embryonic Kidney (HEK) 293T cells.

Human embryonic kidney (HEK) cell line 293T (female, ATCC) was maintained in DMEM (L-glutamine, high glucose, sodium pyruvate, phenol red; Corning) media supplemented with 10% fetal bovine serum (FBS; Gemini Benchmark), and 1x penicillin/streptomycin (P/S; Gibco/Life Technologies) at 37°C with 5% CO2. When reaching 90%-100% confluency, cells were lifted with Trypsin-EDTA 0.25% (Gibco) and passaged at a ratio of 1:2. This cell line was purchased directly from the manufacturer and no further authenticated was performed.

Cloning.

All plasmids were generated using Gibson Assembly and sequenced by the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility. PCR fragments for Gibson Assembly were amplified using Q5 DNA Polymerase (NEB). All plasmids used in this study are listed in Table 3.1, 3.2, 3.3 with links to fully annotated vector maps and are available upon request. The original pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA was a gift from Feng Zhang (Addgene plasmid # 61591 ; http://n2t.net/addgene:61591 ; RRID:Addgene_61591), the original dPspCas13b plasmid pC0050-CMV-dPspCas13b-longlinker-ADAR2DD(wt) was a gift from Feng Zhang (Addgene plasmid # 103866 ; http://n2t.net/addgene:103866 ; RRID: Addgene_103866), pAAV2/1 and pAdDeltaF6 were gifts from James M. Wilson (Addgene

plasmid # 112862 ; http://n2t.net/addgene:112862 ; RRID:Addgene_112862 and Addgene plasmid # 112867 ; http://n2t.net/addgene:112867 ; RRID:Addgene_112867). Key plasmids will be made available through Addgene.

Protein Expression and Purification.

E. coli-optimized synthetic genes for ORF5, TBP6.7, and the Pin nuclease from SMG6 were purchased as gBlocks (Integrated DNA Technologies) and cloned into a pET vector-derived pMCSG19 protein expression vector (MBP-TVMV-6xHis-TEV) using Gibson Assembly. The ORF5-TBP6.7-Pin construct was transformed into BL21(DE3) cells (NEB). A 10 mL seed culture was used to inoculate 2 L of 2XYT (16 g/L digest peptone, 10 g/L yeast extract, 5 g/L sodium chloride, US Biological) supplemented with 100 µg/mL carbenicillin. Cells were grown to an optical density of 0.6 OD_{600} at 37°C before being chilled to 16°C on ice. Once cooled to 16°C, CIRTS-1 expression was induced with 0.5 mM IPTG (bioWORLD). Induced cultures were grown at 16°C for 12-16 h before the cells were lysed to harvest the protein.

The cells were pelleted by centrifuging at 1,500 g for 15 min at 4°C. The resulting pellet was either directly purified or stored at -80°C for later purification. For purification, cells were lysed with 100 mL lysis buffer (50 mM Tris, 1 M NaCl, 20% glycerol, 10 mM TCEP, pH 7.5) supplemented with protease inhibitors. The resuspended cells were lysed using sonication (Thermo Fisher). Lysates were cleared by centrifuging at 12,000 g for 40 min at 4°C. Cleared lysates were incubated with His60 Ni Superflow Resin (Takara) for 1 h at 4°C with constant gentle agitation. After 1 h, the resin was washed with lysis buffer and eluted with a gradient imidazole elution (10 mM-500 mM). Fractions containing the protein, as assessed by SDS-PAGE, were pooled and concentrated using Ultra-50 Centrifugal Filter Units with 30 kDa cutoff (Amicon, EMD Millipore). After sufficient concentration, the combined fractions were desalted and buffer exchanged into protein storage buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% glycerol, 1 mM DTT) using Sephadex G-25 in PD-10 Desalting Columns (GE Healthcare Life Sciences). The protein was stored at -80°C

or directly used for biochemical assays. The concentration was measured using standard BCA Assay (Thermo Scientific).

***In vitro* RNA preparation.**

DNA oligomer templates containing a T7 RNAP promoter were synthesized by IDT. Templates for all transcribed RNAs were amplified using PCR prior to transcription. For a 250 µL reaction, 12 µg PCR product was incubated with 1x transcription buffer (40 mM Tris-HCl, 2 mM spermidine, 10 mM NaCl), 25 mM MgCl2, 10 mM DTT, 40U SUPERaseIn, 4 mM of each NTP, and 40 µg/mL T7 RNAP at 37°C overnight. The next day, the resulting mixture was DNaseI digested in 1x DNaseI buffer for 30 min at 37°C. RNA was then gel-purified after separation on a 10% 8 M TBE-urea/PAGE gel and purified using the ZR small-RNA PAGE Recovery Kit (Zymo). The target RNA for both the nuclease and the filter binding assays was then 5'-end labeled using the 5' oligonucleotide labeling kit (Vector Labs) with a maleimide-IR800 probe (LI-COR Biosciences) according to the manufacturer's instructions. The labeled RNA was purified using the RNA Clean and Concentrator Kit (Zymo).

Nuclease Cleavage Assay.

The nuclease cleavage assay was performed in Pin domain nuclease buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, 0.5 mM MnCl2, 10% glycerol) with 450 nM labeled substrate (R2), 360 nM gRNA (R4) and 250 nM purified MBP-CIRTS-1. Reactions were incubated for 2 h at 37°C before being quenched with proteinaseK buffer (60 mM EDTA, 4 M urea, proteinaseK). The proteinaseK reaction was incubated for 30 min at 37°C and then denatured further by the addition of 5 M urea and loading dye. Samples were boiled for 7 min at 75°C and analyzed by denaturing gel electrophoresis on a 10% denaturing PAGE gel with 8M urea. Gels were imaged using an Odyssey scanner (LICOR Biosciences).

Filter Binding Assay.

Filter binding assays were performed as previously reporter[168] using two-fold complex dilution of the 1x protein:1x gRNA (0.5 μM – 2 nM). For all reactions, 10 nM labeled substrate were used. The reactions were run in nuclease buffer in the absence of $MnCl_2$ to prevent any cleavage and supplemented with 100 μg/mL heparin and 100 μg/mL tRNA to prevent non-specific interactions. The gRNA and CIRTS were pre-incubated for 5 min at 37°C. Then, the substrate was added and the reaction was incubated for 30 min at 37°C. The reactions were then loaded onto a dot-blot apparatus through Nitrocellulose, Hybond-N+ membranes, filter paper in that order. Membranes were washed with equilibration buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, 10% glycerol) and visualized using an Odyssey scanner (LICOR Biosciences). Data were fit to a quadratic binding equation using Prism (GraphPad Software).

Luciferase Reporter Assay.

To assess changes in protein levels, HEK293T cells were transfected with 12.5 ng dual luciferase reporter plasmid (Promega), 150 ng of the indicated CIRTS expression vector, and 100 ng of the gRNA expression vector. About 16 h before transfection, cells were plated on 96-well plates (Corning) and allowed to grow to 70-80% confluency overnight. The next day, a total of 20 μL Opti-MEM I Reduced Serum Medium (ThermoFisher Scientific) per well was used after combining 10 μL Opti-MEM with 0.5 μL lipofectamine 2000 and 10 μL containing all the transfection DNA. The solutions were combined and incubated for 15 min before slow addition to cells. After 48 h, luciferase readouts of firefly luciferase and Renilla luciferase were sequentially measured using the DualGlo Luciferase Assay System (Promega) on a Biotek Synergy plate reader according to the manufacturer's instructions. All experiments were conducted in at least biological triplicates. Firefly luciferase luminescence levels were normalized to the corresponding Renilla luminescence levels to generate the normalized change in protein levels from the target firefly luciferase gene.

RT-qPCR.

To assess changes in RNA levels after transfection of a CIRTS, total RNA was isolated from HEK293T cells and changes in RNA levels were quantified using RT-qPCR. Cells were plated on 96-well plates (Corning) and transfected at 80% confluency as described above for luciferase assays. Total RNA was harvested 48 h after transfection and isolated using the RNeasy Mini Kit (Qiagen). After RNA isolation, RNA was reverse transcribed to cDNA using the PrimeScript RT Reagent Kit (TaKaRa). All qPCR reactions were run at 20 µL volumes with three biological replicates using FastStart Essential DNA Green Master (Roche) and amplified on a LightCycler 96 Instrument (Roche). Only experiments that showed no amplification of the cDNA control reactions and sharp, single-product melting peaks were used for analysis. All reactions were run with at least three biological replicates. qPCR primers were either identified based on previous publications[62] or verified for specificity using NCBI Primer BLAST. Expression levels were calculated using the housekeeping control gene (GAPDH) cycle threshold (Ct) value and the gene of interest Ct value. The relative expression level of one gene was determined by $2^{-\Delta C_t}$, where $\Delta C_t = C_t$ (gene of interest) - C_t (GAPDH). Relative expression level for targeted gene was obtained upon normalizing the targeted gene expression level of cells experiments treated with the on-target gRNA to those treated by the nontargeting (NT) gRNA. All qPCR primers can be found in Table 3.5.

Western Blotting.

For Western blots, HEK293T cells were plated on 12-well plates (Corning) and transfected with 1.5 µg of a CIRTS expression vector and 1.3 µg of a gRNA expression vector. After 48 h, the cells were washed twice with ice-cold PBS and lysed in 50 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% deoxycholate, 2% SDS, pH 7.4) supplemented with protease inhibitors. After 30 min room temperature incubation, the concentration was measured by BCA assay (Thermo Scientific). 35 µg protein was boiled in protein loading buffer (50 mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.05% bromphenol blue, 100 mM DTT) for 10 min at

95°C and loaded onto a 12% SDS PAGE gel. After stacking at 70 V, the gel was run at 120 V until the dye front reached the bottom, and the proteins were transferred onto a PVDF membrane (Millipore) and blocked in 5% nonfat milk in TBST. Proteins were then detected using 1:500 mouse anti-CypB antibody (Santa Cruz), followed by 1:1000 anti-mouse HRP-conjugated antibody (Santa Cruz). The loading control GAPDH was visualized using 1:5000 HRP-conjugated anti-GAPDH antibody (Proteintech). Membranes were imaged on a Fluor Chem R (Protein Simple) imager after incubation with Super Signal West Pico Plus (Thermo Scientific).

RNA Immunoprecipitation.

For RNA immunoprecipitation experiments, HEK293T were plated on 6-well plates (Corning) and transfected with 1.3 µg of the CIRTS-0 expression vector and 1.7 µg of a gRNA expression vector. After 48 h of incubation, cells were washed twice with ice-cold PBS and then fixed with 1% paraformaldehyde in PBS for 15 min at room temperature. The reaction was then quenched with 250 mM glycine in PBS for 15 min at room temperature. Cells were washed twice with ice-cold PBS before being pelleted at 800 g for 4 min at 4°C. The supernatant was removed, and the pellet was washed once with PBS before lysis. Cells were lysed with 200 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% deoxycholate, 0.1% SDS, pH 7.4) supplemented with protease inhibitors, and SUPERaseIn RNase Inhibitor (Invitrogen). Cells were allowed to lyse on ice for 10 min before being sheared through a 25G needle three times. Insoluble material was pelleted by centrifuging at 13,000 g for 10 min at 4°C and the resulting cleared lysate was used for the FLAG pulldown.

To prepare the antibody-conjugated magnetic beads, 100 µL Dynabeads Protein G (Thermo Fisher Scientific) were pelleted by magnet and washed twice with wash buffer (PBS + 0.02% Tween 80). Beads were then resuspended in wash buffer and 5 µg of rabbit anti-mouse antibody (Sigma M7023) was added. Samples were incubated on a rotator for 10 min at room temperature. After incubation, beads were washed twice with wash buffer,

resuspended in 100 µL wash buffer and incubated with 5 µg anti-FLAG antibody (Thermo Scientific). Beads were then allowed to conjugate for 10 min at room temperature on the rotator. Then, beads were washed twice with wash buffer and split into two 50 µL fractions. One fraction was allowed to continue to incubate while the other fraction was incubated with 7.5 µg 3xFLAG peptide, which will serve as the control IP. After another 10 min incubation at room temperature, the beads were washed twice and resuspended in 200 µL 1x RIPA buffer with SUPERaseIN RNase inhibitor. The lysates were then incubated with beads overnight at 4°C.

After antibody incubation, the beads were pelleted, washed three times with 1x RIPA, 0.02% Tween 80 and then washed once with DNase buffer (350 mM Tris pH 6.7, 50 mM $MgCl_2$, 5 mM DTT). The beads were then resuspended in DNase buffer and 0.08 U/µL DNaseI (Thermo Fisher) was added. The DNase reaction was incubated for 30 min at 37°C on a rotator. Then, proteins were digested using a final volume 0.1 U/µL proteinaseK (Sigma). The reaction was again incubated for 30 min at 37°C on a rotator before the denaturation step. To denature the sample further, 2.5 M urea was added, and the samples were incubated for 30 min at room temperature. The resulting RNA was purified using the RNA Clean and Concentrator kit (Zymo) and reverse transcribed to cDNA using the PrimeScript RT kit (TaKaRa). RT-qPCRs were run using FastStart Essential DNA Green Master (Roche) and detected on a LightCycler 96 Instrument (Roche). The pulldown data was analyzed using the difference between control-IP and FLAG-IP and then normalized to the non-targeting (NT) gRNA sample. All qPCR reactions were run as 20 µL reactions with at least three biological replicates.

RNA sequencing and analysis.

To determine the targeting specificity of CIRTS, we performed RNA sequencing analysis. HEK293T cells were plated in 96-well plates (Corning) and transfected with 150 ng CIRTS and 100 ng gRNA plasmid. After 48 h, total RNA was extracted using the RNeasy Mini Kit (Qiagen) followed by a 30 min DNaseI (Fisher) treatment. Samples were then

cleaned up using the RNA clean up and concentrator kit (Zymo) and the resulting total RNA was used as the input for library. RNA-seq libraries were prepared using the mRNA Hyper-Prep Kit (KAPA biosystems). Libraries were sequenced on an Illumina HiSeq instrument at the University of Chicago Genomic Facility with at least 10 million reads per library. Reads were mapped to the RefSeqGRCh38 transcriptome, quantified, and pseudoaligned using kallisto[125]. To find differentially expressed transcripts, we used sleuth[126]. Only genes that had a $\log_2$FoldChange > 0.75 and a qval < 0.1 were considered to be differentially expressed.

AAV Preparation.

HEK293T cells were transfected with AAV2/1 serotype, pAdDelta6 helper packing plasmid (Addgene) and the CIRTS-gRNA containing transfer plasmid using polyethylenenimine (Sigma). The serotype and helper plasmids were a gift from James M. Wilson (Addgene #112862 and Addgene #112867). The transfer plasmid was designed based on Addgene plasmid #61591, a gift from Feng Zhang. After 48h of incubation, the AAV-containing supernatant was harvested, clarified through a 0.22 µm PVDF filter (Millipore), and concentrated using PEG-it Virus Precipitation Solution (SBI) according to the manufacturer's protocol.

Immunogenicity Prediction.

To predict how likely our constructs are to induce an immune response, we used previously described prediction strategies to predict T cell epitopes and score the potential immunogenicity of these potential epitopes[169]. We used the Immune Epitope Database (IEDB) prediction tools to assess peptide binding to MHC class I molecules with the default prediction method, which combines artificial neuronal network (ANN)[170], Stabilized matrix method (SMM)[171], and Scoring Matrices derived from Combinatorial Peptide Libraries (Comlib)[172] if any corresponding predictor was available. Otherwise, NetMHCpan[173] was used for the prediction. The top one percentile of the predicted MHC class I restricted

9-mer candidates were used to predict an immunogenicity score using the T cell class I pMHC immunogenicity predictor (IEDB)[174].

Statistics.

All values are reported as the mean ±SD or mean ±SEM as indicated in the figure legend. When comparing two groups, we used a one-tailed Student's t-test and a statistical significance cutoff of at least <0.1 as indicated in the figure legend. For comparison of the two conditions in our RNA-seq data that do not follow a normal distribution, we used sleuth, which assumes a negative binomial distribution. Samples sizes were not determined a priori. At least three biological replicates were used for each experiment unless otherwise noted in the figure legend.

Data and Software Availability.

The accession number for the sequencing data in this thesis is GSE128288.

3.5 Supplemental information

Table 3.1 *E.coli* expression plasmid used in this chapter.

Name:		Description:	Benchling Link:
MBP-CIRTS-1	22-79	MBP-ORF5-TBP6.7-GGS6-dead Pin domain	https://benchling.com/s/seq-7bLaA1P49s8a9N7a1Rkx

Table 3.2 Mammalian expression plasmids for the work in this chapter. All mammalian plasmid have a cmv d0 promoter.

Name:		Description:	Benchling Link:
CIRTS-0	25-60	ORF5-TBP6.7-GGS6-dead Pin domain	https://benchling.com/s/seq-IE6qdJ29AolDiABVG8U9
CIRTS-1	18-48	ORF5-TBP6.7-GGS6-Pin domain-NLS	https://benchling.com/s/seq-9DU44R0ZQvTUrhUX32w8
CIRTS-2	20-15	ORF5-TBP6.7-GGS6-NYTHDF1	https://benchling.com/s/seq-UR8ah7oM49lHQ2mcqM80
CIRTS-3t	19-05	ORF5-TBP6.7-GGS6-YTHDF2 (100-200)	https://benchling.com/s/seq-7VsoN7WImEZmvMvHgQEe
CIRTS-3	26-44	ORF5-TBP6.7-GGS6-YTHDF2 (1-200)	https://benchling.com/s/seq-z4qLYXLPsAxNjgJbavCc
CIRTS-4t	24-47	ORF5-SLBP-GGS6-YTHDF2 (100-200)	https://benchling.com/s/seq-mMJ0o9G0r3yjYHv4QSxD
CIRTS-4	27-07	ORF5-SLBP-GGS6-YTHDF2 (1-200)	https://benchling.com/s/seq-OuS8wRsyHrLweV88tv5k
CIRTS-5t	24-39	HBEGF-TBP6.7-GGS6-YTHDF2 (100-200)	https://benchling.com/s/seq-a0T6GeDe2hYsDzsPwugA
CIRTS-5	27-02	HBEGF-TBP6.7-GGS6-YTHDF2 (1-200)	https://benchling.com/s/seq-aFNu4MYfTPyKjW9TbZQM
CIRTS-6t	25-02	b-defensin 3-TBP6.7-GGS6-YTHDF2 (100-200)	https://benchling.com/s/seq-3ZPqMHgPbgboxWlcx2YT
CIRTS-6	27-03	b-defensin 3-TBP6.7-GGS6-YTHDF2 (1-200)	https://benchling.com/s/seq-wSPzyjo1yHbYTgenr1jQ
CIRTS-7	31-51	b-defensin 3-TBP6.7-GGS6-hADAR(299-701)	https://benchling.com/s/seq-SyiXQTPzSnjadd3x0qh9

Table 3.2 - continued from previous page

Name		Description	Benchling Link
CIRTS-8	31-62	b-defensin 3-TBP6.7-GGS6-hADAR (299-701) E488Q	https://benchling.com/s/seq-mWe7CL35RW9slYljS9cg
CIRTS-9	30-31	b-defensin 3-TBP6.7-GGS6-NYTHDF1	https://benchling.com/s/seq-RTNTinLtTd8hEX5k29Xs
CIRTS-10	26-79	b-defensin 3-SLBP-GGS6-YTHDF2 (1-200)	https://benchling.com/s/seq-8kQOT9WeZ-AnbnpSkTGUo
CIRTS-11	18-17	TBP6.7-GGS6-Pin domain-NES	https://benchling.com/s/seq-ohQaN4BzdY-fWL4BRkm45
CIRTS-12	18-12	ORF5-TBP6.7-GGS6-Pin domain-NES	https://benchling.com/s/seq-oAGtkJLww8cDVB1lEQFN
CIRTS-13	18-63	ORF5-TBP6.7-GGS3-Pin domain	https://benchling.com/s/seq-aPGNWXR4XqcRtSiiKdCU
CIRTS-14	18-65	ORF5-TBP6.7-L8- Pin domain	https://benchling.com/s/seq-EYvRVoETWEGtmnr9WQsd
CIRTS-15	19-63	ORF5-TBP6.7-helical-Pin domain-NES	https://benchling.com/s/seq-m0RRWSGmHas90Nvr9k7J
CIRTS-16	20-35	ORF5-TBP6.7-L8-Y2(100-200)	https://benchling.com/s/seq-9x8CdTanwK1NvT8NWh4Z
CIRTS-17	20-39	ORF5-TBP6.7-helical-Y2(100-200)	https://benchling.com/s/seq-c6WZR6ccaf-dlEThrMUQq
CIRTS-18	25-20	ORF5-PP7-Y2(100-200)	https://benchling.com/s/seq-VzspV3sWZJxL5wWdIUN0

Table 3.3 gRNA vectors used in this chapter.

Name:		Description:	Benchling Link:
TBP-NT	25-61	TBP6.7 non-targeting gRNA I phage	https://benchling.com/s/seq-HulPL4ZTruDzKqvVGRmS
SLBP-NT	24-18	SLBP non-targeting gRNA I phage	https://benchling.com/s/seq-13JVjb-KsEPsI7e0mN9GP
PP7-NT	24-79	PP7 non-targeting gRNA I phage	https://benchling.com/s/seq-eksSYdn2lRWnSQAR9z2X

Table 3.4 gRNA guiding sequences used in this chapter. All used in the same hU6 promoter as TBP-OT.

Name:		Sequence:
Fluciferase-40	18-20	caggtcgactctagactcgaggctagcgagctcgtttaaa
PPIB	23-26	cttggtgctctccaccttccgcaccacctccatgccctct
SMARCA4	23-25	ccgatgcggtgggctcggtcctgcgcttgcaggtcctggt
NFKB1	22-67	gcctccaccagctctctgactgtacccccagagacctcat
NRAS	25-30	aagcatcttcaacaccctgtctggtcttggctgaggtttc
B4GALTN1	24-50	cttcgcaccgcagcgcagcgcggctcagctcccggctcgt
SMARCA4-1	27-67	ggccctggcccttcccctggagccatgctgggccctagcc
SMARCA4-2	27-71	gcccaaccccatttaaccagaaccagctgcaccagctcag
SMARCA4-3	27-74	cctcccaagccctggcctgaaggacccatggcgaatgctg
SMARCA4-4	27-75	cgtcccacccgccgcctcgcccgtgatgccaccgcagacc
SMARCA4-5	27-78	ggaggtggtggtgtgcatgcggagggacacagcgctggag
SMARCA4-6	27-80	tcacaggcaaaatccagaagctgaccaaggcagtggccac
SMARCA4-7	27-81	catggctgaagatgaggaggggtaccgcaagctcatcgac
SMARCA4-8	28-02	ctctggacgagaccagccagatgagcgacctcccggtgaa
SMARCA4-9	28-04	cccaccctgcccgtggaggagaagaagaagattccagatc
SMARCA4-10	28-05	aatatggcgtgtcccaggcccttgcacgtggcctgcagtc

Table 3.4 - continued from previous page

Name		Sequence
SMARCA4-11	28-07	ctcatcacgtacctcatggagcacaaacgcatcaatgggc
SMARCA4-12	28-08	tggtgaaggtgtcttacaagggatccccagcagcaagacg
SMARCA4-13	28-10	accccgccgcctgctgctgacgggcacaccgctgcagaac
SMARCA4-14	28-11	cagtggtttaacgcacccttttgccatgaccggggaaaagg
SMARCA4-15	28-12	acgactcaagaaggaagtcgaggcccagttgcccgaaaag
SMARCA4-16	28-16	ggaaccacgaaggcggaggaccggggcatgctgctgaaaa
SMARCA4-17	28-20	cgccagcgggcgtcaaccccgacttggaggagccacctct
SMARCA4-18	28-26	tgatcaagtacaaggacagcagcagtggacgtcagctcag
SMARCA4-19	28-27	cttcaagaagataaaggagcgcattcgcaaccacaagtac
SMARCA4-20	28-30	agggtcccgagccaagccggtcgtgagtgacgatgacagt
SMARCA4-21	28-32	tatttatacagcagagaagctgtaggactgtttgtgactg
SMARCA4-22	28-33	ggggaacacacgatacctgtttttcttttccgttgctggc
Fluciferase-20		CTAGACTCGAGGCTAGCGAG
Fluciferase-30		CGACTCTAGACTCGAGGCTAGCGAGCTCGT
MALAT1-1	30-36	ATGCTAGCTTGGCCAAGTCTGTTATGTTCACCTGAAAAAG
MALAT1-2	30-37	CACCAGCAAAATGTACTCAGCTTCAATCACAAATACGACT
MALAT1-3	30-38	TTTTGTGGTTATAGCTTGACAAGCAATTAACTTTAAAATG
KRAS-2	22-45	cttgtggtagttggagctga
KRAS-3	22-53	agttggagctgatggcgtag
KRAS-4	22-54	gttggagctgatggcgtagg
KRAS-5	22-65	tggcgtaggcaagagtgcct
KRAS-6	22-66	ggcgtaggcaagagtgcctt

Table 3.5 qPCR primers used in this chapter.

Name:	Forward:	Reverse:
GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA
PPIB	AACGCAGGCAAAGACACCAACG	TCTGTCTTGGTGCTCTCCACCT
Fluciferase	AGGTTACAACCGCCAAGAAGC	ATGAGAATCTCGCGGATCTTG
NFKB1	gcagcactacttcttgaccacc	tctgctcctgagcattgacgtc
NRAS	gaaacctcagccaagaccagac	ggcaatcccatacaaccctgag
B4GALTN1	tgaggctgctttcactatccgc	gaggaaggtcttggtggcaatc
SMARCA4	caaagacaagcacatcctcgcc	gccacatagtgcgtgttgagca
MALAT1	gaagaaggaaggagcgctaa	cctgctaccttcatcaccaa

CHAPTER 4

SMALL MOLECULE-INDUCIBLE RNA-TARGETING SYSTEMS FOR TEMPORAL CONTROL OF RNA REGULATION *IN VIVO*.

Epitranscriptomic regulation has emerged as an important regulatory layer in gene expression, but tools to study and control the dynamics of these regulatory functions in an endogenous environment are lacking. Here, we present a small molecule-inducible RNA-targeting biosensor based on our previously developed CRISPR/Cas-inspired RNA targeting system (CIRTS). We show that the abscisic acid (ABA)-CIRTS biosensor can be used to edit RNA on a luciferase transcript in cells and tolerates swapping of the effector domain to ribonucleases or chemical modification reader proteins to induce RNA degradation or translation, respectively. Using the CIRT editing system, we demonstrate inducible RNA editing on disease-relevant and endogenous transcripts in mammalian cells. Finally, we show the ABA-CIRTS biosensor can edit RNA on a luciferase transcript in a small molecule-dependent manner in mice. This work provides a versatile, inducible RNA-targeting system to both study the dynamics of interconnected RNA regulatory networks and to control gene expression in a temporal manner *in cellulo* and *in vivo*.

4.1 Introduction

Gene expression regulation at the RNA level has recently been uncovered as a crucial and complex regulatory layer. Apart from canonical RNA processing including capping, splicing, polyadenylation, and trafficking[11–14,97], the RNA sequence can be edited or chemically modified[19,98,138]. Adenosine-to-inosine (A-to-I) editing is a common post-transcriptional modification, catalyzed by the ADAR family of proteins[32,33]. The catalytically active ADAR1 and ADAR2 isoforms deaminate A in double-stranded RNA (dsRNA) at the site of a C-A mismatch. The deamination product inosine is thought to mimic guanosine

(G) in cells by base-pairing with cytosine (C), allowing for amino acid changes and changes in splicing[34,35,175]. In addition to RNA editing, RNA bases can be chemically modified. *N*^{6}-methyladenosine (m^{6}A) is the most prevalent internal RNA modification in mammalian systems[138]. m^{6}A is dynamically installed and erased by writer and eraser enzymes, respectively and is linked to cellular functions by recognition and binding of m^{6}A reader proteins[98,176]. Chemical RNA modifications influence RNA stability, splicing, export, and translation efficiency[15,18,28,98] and have been linked to various cancers and viral infections[105,106,111]. With more and more regulatory pathways influencing the entirety of the RNA lifetime being uncovered, precisely controllable technology is required to study the temporal dynamics of these processes in their endogenous environment.

Programmable RNA-targeting tools that allow for site-specific RNA targeting hold great promise for studying the dynamics of biological processes as well as therapeutic applications. Pumilio family (PUF) proteins[4,5] and Cas13 proteins of the CRISPR-Cas system have been developed to deliver a variety of effector proteins, including RNA editors[60,61], splicing modulators[62], or translational activators to RNA transcripts and sites of interest[63]. We recently developed a CRISPR/Cas-inspired RNA targeting system (CIRTS) as a smaller, human-derived delivery moiety for RNA effectors[177]. Analogous to the Cas13 system, CIRTS relies on guide RNA (gRNA) complementarity to bind a target of interest and deliver a tethered cargo protein. However, while these tools provide the means to deliver RNA regulatory proteins to a transcript and site of interest, they offer no temporal control to study effector dynamics. Herein, we describe the development of a small-molecule inducible RNA targeting system based on CIRTS that allows for specific temporal control of RNA editing, degradation, and translation activation.

4.2 Results

4.2.1 Development of a CIRTS biosensor

To engineer a small molecule-inducible RNA targeting system, we coupled CIRTS with the heterodimerization domains of the abscisic acid (ABA) system. We chose the ABA system because it had been successfully applied in Cas9-guided DNA targeting systems for inducible transcription activation of genes of interest or to temporally recruit acetyltransferases to histones[160,178]. The ABA system relies on rapid binding of two heterodimerization domains (ABI and PYL) only in the presence of the small molecule abscisic acid[179]. For our programmable CIRTS biosensor, we fused the targeting component of CIRTS, a single-stranded RNA binding protein and hairpin binding protein, to ABI (CIRTS-ABI) and the RNA effector domain to PYL (**Figure 4.1A**). The targeting CIRTS component binds to a gRNA and engages the target RNA by base pair complementarity. Addition of ABA recruits the effector domain to the targeting moiety of CIRTS to elicit the desired function on the targeted RNA transcript (**Figure 4.1A**).

We validated the ABA biosensor in a luciferase reporter assay, in which a single G-to-A mutation encodes a premature stop codon that prevents luciferase enzyme translation. Targeted editing of this mutation restores full length luciferase expression (**Figure 4.2A**). We transfected HEK293T cells with the targeting CIRTS-ABI, a gRNA targeting the mutated sequence, and PYL fused to a hyperactive version of the catalytic domain of hADAR2(E488Q). Upon induction of the biosensor with 100 μM ABA, we observed the reversion of the stop codon mutation and robust luciferase reporter signal (**Figure 4.2A**). We next constructed a combined CIRTS-ABI and PYL-hADAR(E488Q) vector to reduce the number of transfected plasmid and subjected the ABA-CIRTS-hADAR2 biosensor to a gRNA screen to find the ideal mismatch position along our gRNA (**Figure 4.2B and C**). We determined that a mismatch 15nt into the guiding sequence yielded the best editing

efficiency and we chose this mismatch position for all further editing experiments.

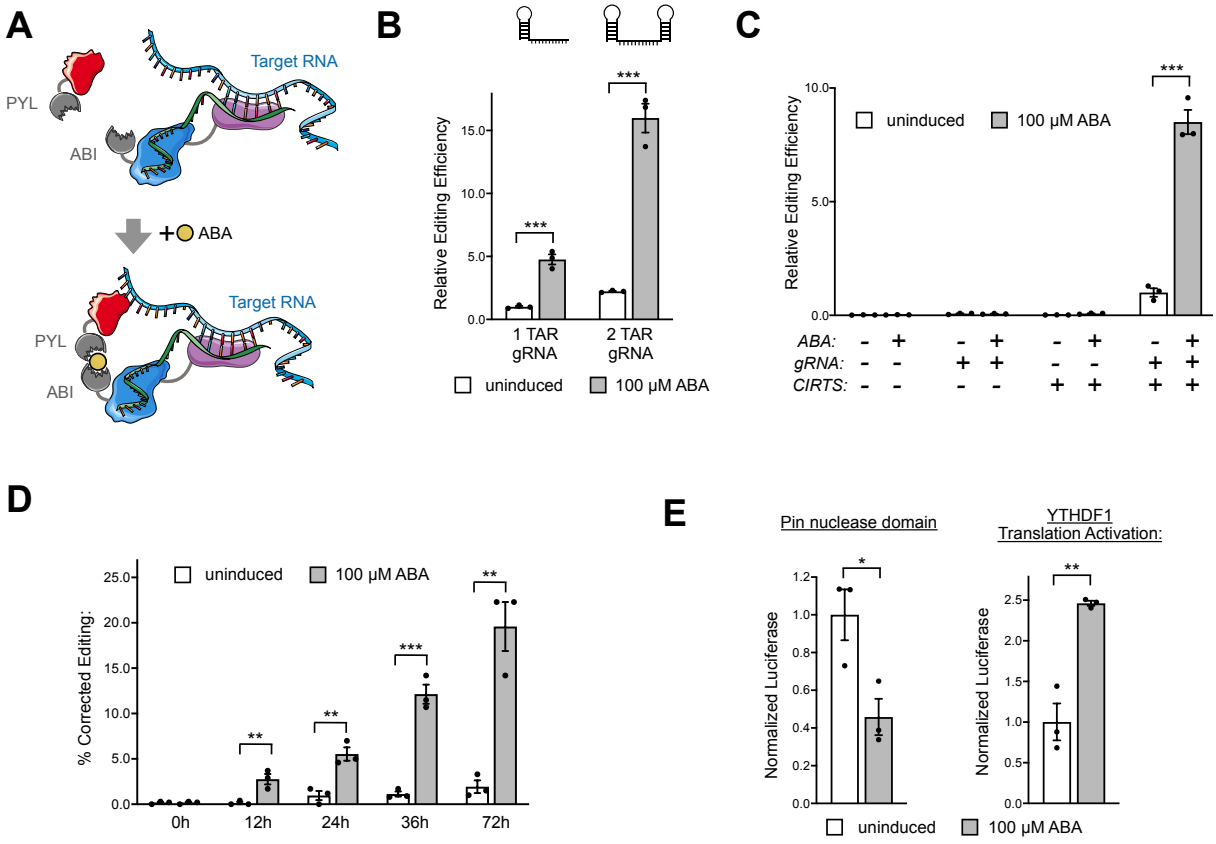

Figure 4.1 Development of a CIRTS biosensor.
(A) Schematic of the small molecule-inducible CIRTS biosensor. **(B)** gRNA screen comparing the original 1 TAR gRNA design with a 2 TAR gRNA. Editing efficiency increased when an additional hairpin was added to the gRNA. **(C**) We transfected cells with each individual component of the biosensor to verify that the observed editing is a result of the targeted, induced CIRTS biosensor. Robust editing was only observed when all CIRTS components were present and ABA was added. **(D**) Quantified editing percentage by RT-Sanger sequencing in a three day time course after addition of ABA. **(E**) The hADAR effector was swapped with previously validated CIRTS effector Pin nuclease and YTHDF1. We observed ABA-dependent RNA degradation via Pin nuclease and translation activation via YTHDF1. All values are mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05, **P <0.01, ***P <0.001.

In previous studies, it had been reported that editing systems relying on hairpin-hairpin binding interactions are more efficient when the gRNA contains two hairpin sequences[148,180]. We therefore tested our inducible CIRTS editor with the original gRNA con-

taining one TAR hairpin in comparison to a gRNA flanked by two TAR hairpins (**Figure 4.1B**). When a second TAR hairpin was added to the gRNA design, we observed a dramatic increase in editing efficiency (**Figure 4.1B**), leading us to adopt this optimized gRNA design for our biosensor. Overexpression or delivery of an RNA strand of gRNA alone has been reported to induce significant levels of editing by endogenous hADAR. We therefore performed a control experiment, delivering each individual component of our system. Delivery of the ABA-CIRTS biosensor or the gRNA alone did not result in dramatic increases of editing over background (**Figure 4.1C**). We did, however, observe low levels of background editing when the gRNA was delivered, which we attributed to editing via endogenous hADARs on target-gRNA duplex (**Figure 4.2D**). Only when we transfected cells with the ABA-CIRTS biosensor, as well as the gRNA, and delivered ABA did we observe robust editing (**Figure 4.1C**). Next, we compared the biosensor CIRTS editor to the full-length CIRTS editor and found that conversion of CIRTS into a biosensor results in a two-fold loss of efficacy (**Figure 4.2E**).

After validating our inducible CIRTS biosensor, we subsequently performed a dynamic characterization of the system. We conducted an ABA time course experiment and measured the resulting RNA editing levels for three days. Editing efficiency was monitored by both a cell-based luciferase stop codon reversion assay and by quantifying RNA edits directly using reverse transcription (RT)-polymerase chain reaction (PCR) followed by Sanger sequencing. Editing levels quantified by RT-PCR-Sanger sequencing are based on the ratio of the peak heights of G and A at the mutation site and are corrected for RT and PCR biases using a standard curve (**Figure 4.2F**). Both luciferase readout and G content at the mutation site increase steadily over the course of 72 h of ABA-induced editing via our CIRTS biosensor (**Figure 4.1D and Figure 4.2G**).

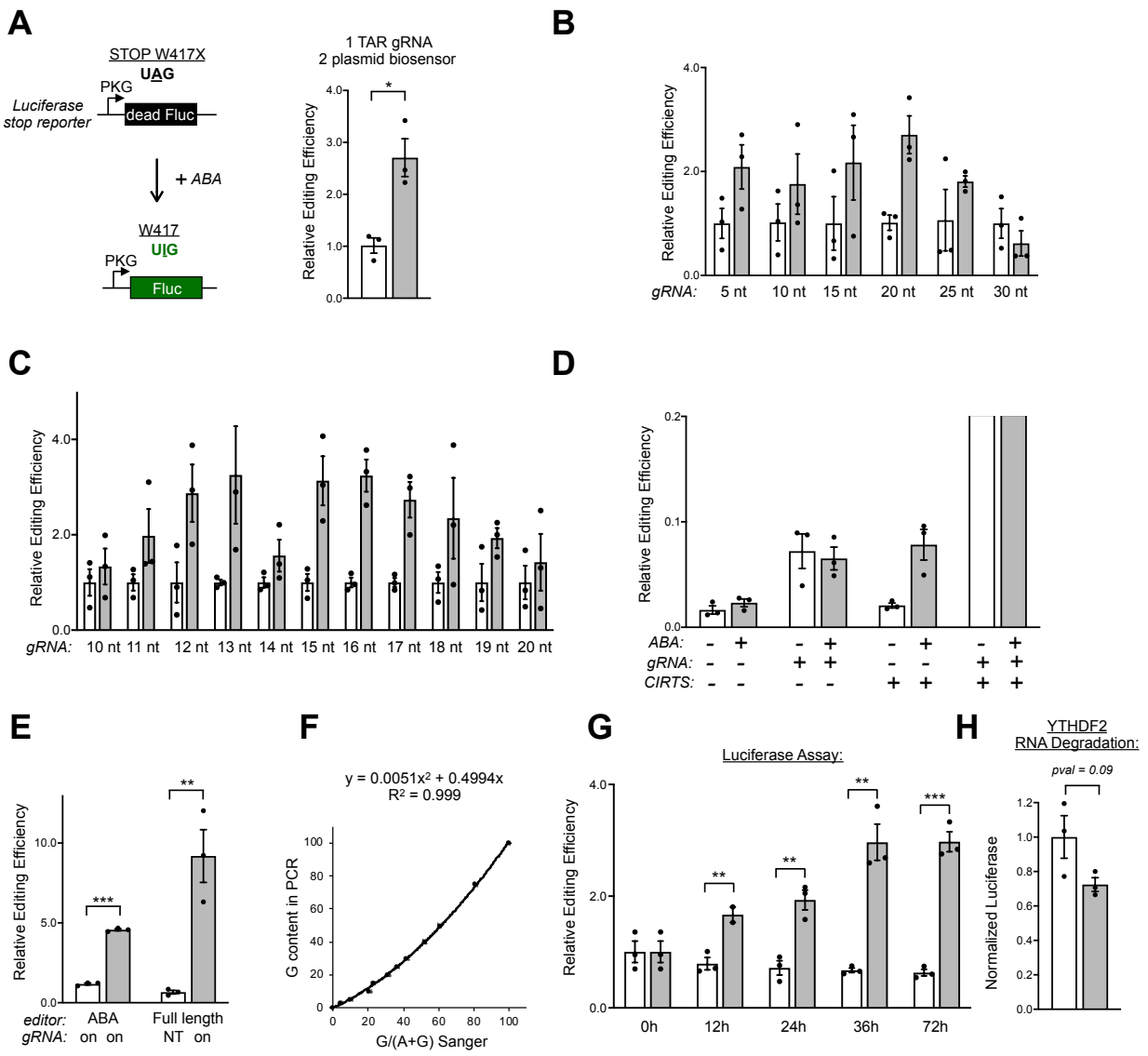

Figure 4.2 CIRTS biosensor optimization.

(A) Schematic of stop codon reversion luciferase assay. Proof-of-principle with 2 plasmid CIRTS-editor system. Stop codon reversion was only observed when 100 μM ABA was added. **(B-C)** gRNA screen with CIRTS-hADAR2 to determine the mismatch position in the gRNA. Nucleotide position label = distance from TAR hairpin. **(D)** Zoomed-in control experiment (Figure 4.1C). Cells transfected with gRNA or induced biosensor without gRNA show low luciferase signal above background. **(E)** Comparison of full length CIRTS-hADAR2 to ABA-inducible CIRTS-hADAR2. **(F)** Standard curve to correct Sanger sequencing quantification of editing for fluorophore and PCR biases. **(G)** Time course with CIRTS editor for 3 days after ABA addition as quantified by stop codon reversion luciferase assay (n=2 or 3). **(H)** CIRTS-YTHDF2 biosensor induces RNA degradation when ABA is added. All values are mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05, **P <0.01, ***P <0.001.

Encouraged by the ability to convert a CIRTS editor into a biosensor, we set out to broaden the effector scope of the biosensor. We therefore swapped the effector domain to three previously validated CIRTS effectors: the Pin nuclease domain, m^{6}A reader protein YTHDF1, which is known to induce translational activation by binding to eukaryotic translation initiation factors and the ribosome[24], and m^{6}A reader YTHDF2, which induces RNA degradation by recruiting the CCR4-NOT deadenylation and RNaseP machineries[18,28,29]. As expected, cells transfected with the Pin nuclease domain and the YTHDF2 reader showed small molecule-inducible RNA degradation (**Figure 4.1E and Figure 4.2H**). ABA induction of the YTHDF1 effector system resulted in the expected increase in luciferase readout (**Figure 4.1E**). Taken together, these findings suggest that a versatile range of effectors can be used with the ABA-inducible CIRTS biosensor to both study RNA binding protein dynamics in cells and control gene expression temporally.

4.2.2 Inducible editing of endogenous and disease-relevant transcripts

We then sought to assess whether our inducible CIRTS editor could be deployed to edit disease-relevant mutations or endogenous transcripts. Nucleic acid editing at the RNA level holds great promise for therapeutic applications as codon changes can be made without risking alterations to the genome. To showcase the ability of the inducible CIRTS editor to reverse a known disease-causing mutation, we chose a stop codon-inducing mutation in adenomatous polyposis coli (APC) as our test case. APC is a tumor suppressor gene, with premature stop codons in APC present in 95% of familial adenomatous polyposis (FAP) patients[181]. We again delivered the CIRTS biosensor, the corresponding gRNA targeting APC, as well as an APC stop codon-simulating reporter, and assessed RNA editing using the RT-qPCR-Sanger sequencing assay. We found that our CIRTS biosensor system can efficiently revert the stop codon mutation in this simulated disease model (**Figure 4.3A**). In addition to disease-relevant targets, we chose two endogenous transcripts, GAPDH and PPIB, to verify the programmable nature of the system. We again

introduced the CIRTS biosensor editor and a gRNA sequence to cells and quantified editing levels 48 h after induction of the system. We observed significant editing for both of the targeted transcripts (**Figure 4.3A**). We acknowledge that the editing efficiency of these endogenous transcripts is relatively moderate and would require further system optimization to improve editing efficiency for biological studies or therapeutic applications.

4.2.3 Inducible editing in vivo

We next aimed to apply our CIRTS biosensor for *in vivo* editing of a luciferase reporter. Initially, we optimized our plasmids for delivery to the mouse liver by constructing a single plasmid encoding the CIRT system and gRNA and a mouse reporter plasmid with the EF1αpromoter for robust, long-term expression of luciferase. We transfected cells with the single CIRTS-gRNA plasmid and the cell or mouse luciferase plasmid and observed comparable editing (**Figure 4.4A-B**). Based on this validation, we delivered the luciferase reporter by hydrodynamic tail vein injection to the mouse liver at 2 or 20 µg to determine optimal reporter concentration. No significant signal was observed with lower DNA concentrations and we selected 2 µg luciferase plasmid for future experiments (**Figure 4.4C**). For *in vivo* editing, we first tested full length CIRTS-hADAR delivered with the luciferase reporter by hydrodynamic tail vein injection. Full length CIRTS-hADAR delivery was optimized so robust luciferase signal at 7 h post injection was only observed in the presence of on-target gRNA (**Figure 4.3B**). No significant editing was observed with a non-targeting gRNA and CIRTS-hADAR as compared to mice receiving reporter alone.

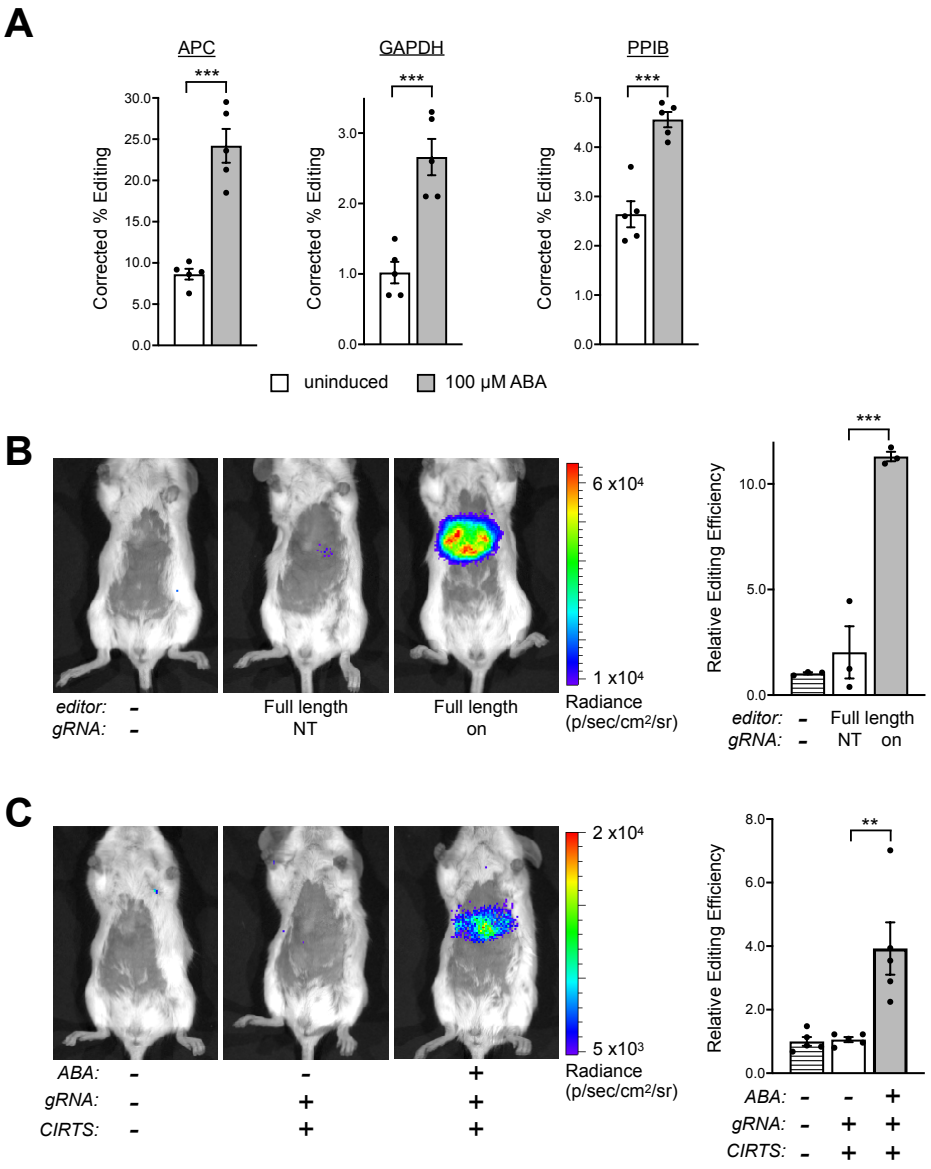

Figure 4.3 Inducible RNA editing on endogenous/disease targets and *in vivo*.
(A) The inducible CIRTS-editor can be deployed to disease-relevant reporter transcripts (APC) or endogenous targets (GADPH and PPIB) and shows ABA-dependent RNA editing. **(B)** In vivo editing of a luciferase reporter with full length CIRTS-hADAR. DNA encoding full length CIRTS-hADAR with on-target or non-targeting gRNA was delivered with luciferase reporter by hydrodynamic tail vein injection. Luciferin was administered i.p. and photon outputs were quantified. **(C**) Small molecule-inducible editing of a luciferase reporter with CIRTS biosensor. DNA encoding ABA-CIRTS-hADAR was delivered with luciferase reporter by hydrodynamic tail vein injection. Mice were imaged as in (B). Representative bioluminescence images are shown for B-C. All values are mean ± SEM with (A-B) n = 3 or (C) n = 5 biological replicates. Student t-test: *P < 0.05, **P <0.01, ***P <0.001.

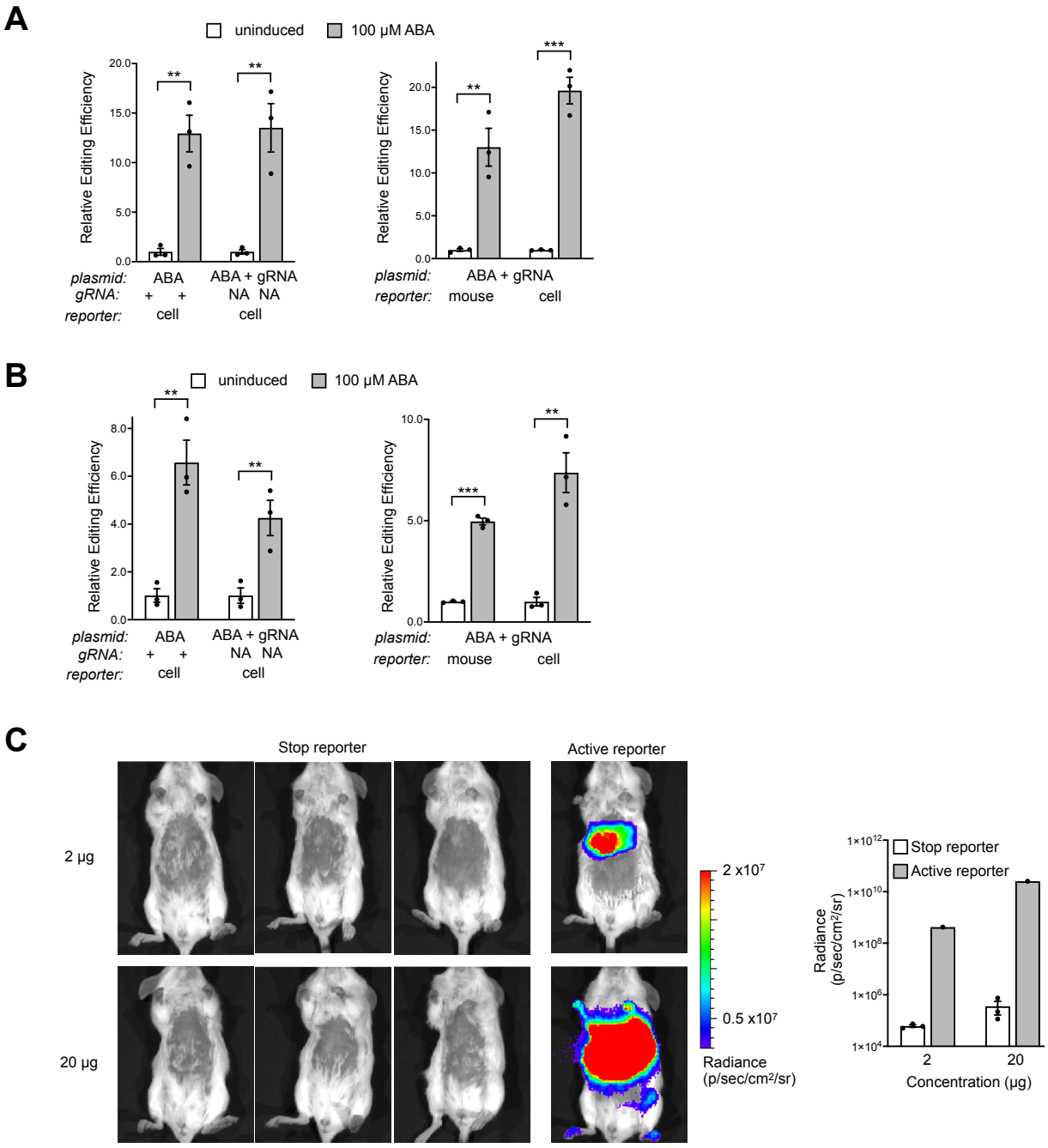

Figure 4.4 Plasmid optimization for in vivo delivery.
(A) Cells transfected with full length CIRTS-hADAR and gRNA or a single plasmid encoding both were screened with the mouse and cell reporter. Comparable editing was observed for all conditions. **(B)** Cells transfected with ABA-CIRTS-hADAR and gRNA or a single plasmid encoding both were screened with the mouse and cell reporter. **(C**) DNA (2 or 20 μg) encoding the stop or active luciferase reporter was delivered to the liver by hydrodynamic tail vein injection. D-luciferin was administered i.p. 24 h after delivery and photon outputs were quantified. Error bars are mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05, **P <0.01, ***P <0.001.

Since the full length CIRTS-ADAR system induced robust editing, we aimed to similarly optimize ABA-CIRTS-hADAR delivery and establish ABA delivery conditions for small molecule-inducible editing. ABA administration post plasmid delivery was based on previously determined gene expression patterns in the liver following hydrodynamic tail vein injection[182,183] and ABA clearance rates[179]. ABA was injected intraperitoneally or intravenously 2 h and 6 h after plasmid expressing ABA-CIRTS-hADAR and gRNA was delivered with the mouse reporter plasmid. Significant editing of the luciferase mutation was only observed in the presence of ABA and could be achieved with either ABA delivery method (**Figure 4.3C**). In the absence of ABA, no significant luminescence was observed over the reporter control conditions. Collectively, these results demonstrate that CIRTS approaches can be utilized for RNA editing of a transcript of interest and small molecule-inducible editing *in vivo*.

4.3 Discussion

Using the ABA chemical-inducible system, we developed the first inducible RNA targeting system based on our engineered CIRT system. In this study, we achieved ABA-dependent RNA editing, RNA degradation, and translation initiation. Analogous to previously developed inducible DNA-targeting systems based on dCas9 that have been used for targeted transcription or histone targeting, our CIRTS biosensor provides the basic platform for temporal control of protein-based RNA regulation. In this proof-of-concept demonstration, we utilized well-characterized RNA modification inducers or readers, to demonstrate that the system tolerates effector swaps and is likely a generalizable approach for targeted RNA manipulation. In addition to exchanging the effector domain, other orthogonal small molecule- or light-inducible heterodimerization domains, including the FRB/FKBP rapamycin system[184], the GID1/GAI gibberellin system[185], or the blue light-based CRY2/CIBN pairs[186], could open up the possibility of studying temporal dynamics

of orthogonal regulatory pathways simultaneously. We acknowledge that our current biosensor system yields relatively low levels of RNA editing on endogenous transcripts. As we see almost no loss in activity for any of the other RNA effector proteins tested (Pin nuclease, YTHDF1, and YTHDF2), we believe that this could be an editor-specific issue. In future studies, the relative capacity of the CIRTS editing biosensor can likely be improved by protein engineering such as directed evolution of the effector and a linker or orientation screen to improve effector-mutation site alignment. Additionally, there is a range of other RNA regulatory proteins that could be fused to the CIRTS biosensor to study the RNA regulation dynamics in live cells[167]. We further demonstrate that RNA editing with full length CIRTS-hADAR or the ABA editing biosensor can be deployed in mice. While careful tuning of both plasmid and ABA delivery are necessary in this liver model, it ultimately affords temporal control on the hour-timescale. Further optimization of plasmid and ABA delivery could likely afford improvements in editing efficiency and would be necessary if editing on different timescale is needed. We are also currently investigating additional delivery methods of the CIRT system to move towards more clinically relevant *in vivo* targets. Collectively, this validation of the CIRTS biosensor demonstrates the feasibility of an RNA-targeting, inducible biosensor and lays the foundation for temporally regulated studies of RNA regulation in mammalian systems.

4.4 Materials and Methods

Cloning.

All plasmids were cloned using Gibson Assembly and sequenced by the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility. PCR fragments for Gibson Assembly were generated using Q5 DNA Polymerase (NEB). All plasmids used in this study are listed in Table 4.1, 4.2, and 4.3 with links to their vector maps and are available upon request. Key plasmids will be made available through

Addgene.

Mammalian Cell Culture and Transfection.

For cell culture assays, HEK293T cells (ATCC) were used. Cells were maintained in DMEM (L-glutamine, high glucose, sodium pyruvate, phenol red; Corning) media supplemented with 10% fetal bovine serum (FBS; Gemini Benchmark), and 1x penicillin/streptomycin (P/S; Gibco/Life Technologies). For transfections, P/S was omitted from the media. All transfections were conducted using lipofectamine 2000 (Invitrogen) according to the manufacturer's recommendations.

Induction Luciferase Assays.

For ABA induction experiments, HEK293T cells were transfected with 12.5 ng reporter plasmid, 150 ng indicated CIRTS biosensor vector, and 100 ng gRNA expression vector. Approximately 16h before transfection, cells were plated on 96-well plates and grown to 70-80% confluency overnight. The next day, cells were transfected with a total of 20 µL Opti-MEM I Reduced Serum Medium (ThermoFisher Scientific): 10 µL Opti-MEM containing 0.5 µL lipofectamine 2000 per well and 10 µL Opti-MEM containing the plasmid DNA. Diluted lipofectamine 2000 and DNA were combined and incubated for 15 min before slowly adding them to cells. After 24h, the ABA biosensor was induced with 100 µM abscisic acid (ABA, ThermoFisher) and incubated for an additional 24h before luciferase readout on a Biotek Synergy plate reader. Both firefly and Renilla luciferase readouts were measured as previously described (Jove paper Baker & Boyce). First, 40 µL growth media was removed from every well. Then, 40 µL of firefly assay buffer (Triton Lysis Buffer (50 mM Tris, pH 7.0, 75 mM NaCl, 3 mM MgCl2, 0.25% Triton X-100) containing 5 mM DTT, 0.2 mM coenzyme A, 0.15 mM ATP, and 1.4 mg/mL luciferin) was added to lyse the cells and to provide the first substrate for firefly luciferase. After 10 min incubation, the firefly read was taken and 40 µL Renilla assay buffer (45 mM EDTA, 30 mM sodium pyrophosphate, 1.4 M NaCl, 0.02 mM PTC124, 0.003 mM coelentrazine h (CTZ-h)) was added to

stop firefly luciferase activity and provide the substrate for *Renilla* luciferase. The *Renilla* read was taken immediately after addition of the buffer. All experiments were conducted in at least biological triplicates. Firefly luciferase luminescence levels were normalized to the corresponding Renilla luminescence levels to generate the normalized change in protein levels from the target firefly luciferase gene.

Quantification of RNA editing efficiency.

To quantify the editing efficiency of CIRTS in mammalian cells, HEK 293T cells were transfected as described above. Cells were plated to be at 80% confluency and were transfected with 150 ng CIRTS expression vector and 100 ng luciferase reporter-targeting gRNA or 200 ng of gRNA when targeting endogenous transcripts or disease-simulated reporters. After 24h, the ABA biosensor CIRTS was induced with 100 μM ABA and incubated for an additional 24h (or 48h for endogenous and disease-simulating targets) until further processing. Total RNA was harvested using the RNeasy Mini Kit (Qiagen) when isolated from cells or the RNeasy Lipid Tissue Mini Kit (Qiagen) when isolated from mouse liver tissue. After RNA purification, the resulting RNA was reverse transcribed with a target specific primer (Table 4.4) and the PrimeScript RT Reagent Kit (TaKaRa). The resulting cDNA product was PCR-amplified with Q5 DNA polymerase (NEB) using target-specific primers and sent for Sanger sequencing at the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility. Editing efficiency was calculated by taking the ratio of peak heights at the target site (G/(A+G)) as previously reported (Ref in vivo editing ADAR Nat methods). A Sanger sequencing standard curve was used to correct editing efficiencies for sequencing fluorophore and PCR biases **Figure 4.2F**.

***In vivo* delivery of CIRTS and imaging.**

BALB/C mice (Charles River Laboratories) aged 5-6 weeks were injected with plasmids encoding the luciferase reporter (2-20 μg), full length CIRTS-hADAR2 (2-20 μg) or ABA-CIRTS-hADAR2 (2-20 μg) using TransIT-EE Hydrodynamic Delivery Solution (Mirus

Bio) according to the manufacturer's recommendations. Fur on the abdomen of the mice was removed 24-72 h prior to injection using a chemical depilatory cream. For ABA-CIRTS-hADAR2, animals received an i.p. injection of ABA (100 mg/kg, 200 µL per mouse, 2.5% DMSO) or i.v. injection of ABA (5 mM, 100 µL per mouse, 3% ethanol) 2 h and 6 h after plasmid delivery. At 7 h post plasmid delivery, animals received an i.p. injection of D-luciferin (100 mM, 100 µL per mouse, PBS pH 7.4). Mice were anesthetized (2% isoflurane) and placed on the warmed (37°C) stage of an IVIS 200 Imaging System (Xenogen) and imaged with a CCD camera chilled to -90 °C at the University of Chicago Optical Imaging Core Facility. Exposure times were 1 minute with data binning levels set to medium. Regions of interest were selected for quantification and total flux values were analyzed using Living Image software. All animal experiments were performed following the protocols approved by the University of Chicago Institutional Animal Care and Use Committee.

4.5 Supplemental Information

Table 4.1 Mammalian expression plasmids used in this chapter. All mammalian plasmid have a cmv d0 promoter.

Name:		Description:	Benchling Link:
ABA-hADAR(E488Q) sensor	40-31	Cmv b-defensin3-TBP6.7-ABI; cmv PYL-hADAR(E488Q)	https://benchling.com/s/seq-yXH1tpoefWPzACYIMQ4f
ABA-Pin nuclease sensor	35-02	Cmv b-defensin3-TBP6.7-ABI; cmv PYL-Pin nuclease domain	https://benchling.com/s/seq-wtRYVs50hjXhFzyEZEaN
ABA-NYTHDF1 sensor	35-06	Cmv b-defensin3-TBP6.7-ABI; cmv PYL-NYTHDF1	https://benchling.com/s/seq-qhVRzqBWcLuMvLZO4rbR
ABA-NYTHDF2	35-07	Cmv b-defensin3-TBP6.7-ABI; cmv PYL-NYTHDF2	https://benchling.com/s/seq-ClHd0P6NZ5QNHTs4FNf7
FL-hADAR(E488Q)-NT gRNA	40-73	Cmv b-defensin3-TBP6.7-hADAR(E488Q); hU6-2TAR-non-targeting gRNA	https://benchling.com/s/seq-z7m7kz8xeDOYK50E7vTw

Table 4.1 - continued from previous page

Name		Description	Benchling Link
FL-hADAR(E488Q)-Fluc gRNA	40-75	Cmv b-defensin3-TBP6.7-hADAR(E488Q); hU6-2TAR-Fluc gRNA	https://benchling.com/s/seq-MaBOnsmdjH9gFf87uFCr
ABA-hADAR(E488Q) sensor-Fluc gRNA	KJ476	Cmv b-defensin3-TBP6.7-ABI; cmv PYL-hADAR(E488Q); hU6-2TAR-Fluc gRNA	https://benchling.com/s/seq-WaoKjoTdCyO0ZudsfIJv

Table 4.2 gRNA vectors used in this chapter. All gRNA plasmids have a hU6 promoter.

Name:		Description:	Benchling Link:
1 TAR-NT gRNA	25-61	hU6-1 TAR hp non-targeting gRNA	https://benchling.com/s/seq-ZFvJV2QVWlwYbpJeHZTd
2 TAR-NT gRNA	24-69	hU6-2 TAR hp non-targeting gRNA	https://benchling.com/s/seq-fzRpWTpskzLMJ9Nh87gJ
Dead luciferase reporter (cell)	31-50	PGK-Fluc (W417X)-SV40-Rluc	https://benchling.com/s/seq-IP5KD-FXK8NVk9hpcvZdM
Dead luciferase reporter (mouse)	KJ475	EF1α-Fluc (W417X)-SV40-Rluc7(521)	https://benchling.com/s/seq-ITmvpsohG7e7eEEeiWkf

Table 4.3 gRNA guiding sequences used in this chapter. All used in the same hU6 promoter as TBP-OT.

Name:		Sequence:
Fluc-m5		GCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGTAGCCTC
Fluc-m10		GCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGTA
Fluc-m15		TCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGT
Fluc-m20		CGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAG
Fluc-m25		GTAGGCGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCG
Fluc-m35		CCTCGTCCCAGTAGGCGATGTCGCCGCTGTGCAGCCAGCC

Table 4.3 - continued from previous page

Name		Sequence
Fluc-m11		CGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTGT
Fluc-m12		CCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTTG
Fluc-m13		GCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTTT
Fluc-m14		CGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCGTT
Fluc-m16		GTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGCG
Fluc-m17		TGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAGC
Fluc-m18		ATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAG
Fluc-m19		ATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAGAG
Fluc-m20		CGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAG
Fluc-3'UTR		CAGGTCGACTCTAGACTCGAGGCTAGCGAGCTCGTTTAAA
APC-m15		GCTTCCTGCCACTCCCAACAGGTTTCACAGTAAGCGCGTA
GAPDH-m15		CTTGGCCAGGGGTGCCAAGCAGTTGGTGGTGCAGGAGGCA
PPIB-m15		AAAGATCACCCGGCCCACATCTTCATCTCCAATTCGTAGG

Table 4.4 RT, PCR, Sequencing primers used in chapter 4.

Name:		Sequence:
Fluc-RT		GACCCCGGCGTCGAAGATGT
Fluc-PCR-for		GGATGCTCTCCAGTTCGGCT
Fluc-PCR-rev		ACCAGCGCGGCGAGCTGT
Fluc-Seq		GTCCGTGGCCCCATGATCA
APC-RT		TCGTCTTAGTGTAATACTGTAGTGGTCATTAGT
APC-PCR-for		ACTTGCAATAATTCTGCAATGGCCTGT
APC-PCR-rev		CAGTAAAGAGGCTCGGGCCA
APC-Seq		GTGCAGCACTCCACAACATCATT
GAPDH-RT		ATCCACAGTCTTCTGGGTGGCA

Table 4.4 - continued from previous page

Name		Sequence
GAPDH-PCR-for		GCCCCCTCTGCTGATGCCCCCATGTT
GAPDH-PCR-rev		ATGGCATGGACTGTGGTCATGA
GAPDH-Seq		TGGGTGTGAACCATGAGAAGTATGACA
PPIB-RT		TGGTGAAGTCTCCGCCCTGGATCA
PPIB-PCR-1		TTGATTACACGATGGAATTTGCTGTTTTTGTA
PPIB-PCR-2		GCTCCTTGCCGCCGCCCTCAT
PPIB-Seq		CGGGGTCCGTCTTCTTCCTGCT

CHAPTER 5

EXPANDED EFFECTOR SCOPE AND PRECLINICAL APPLICATIONS OF CIRTS

5.1 Introduction

RNA has emerged as an attractive target for therapeutic intervention as it provides the opportunity to control gene expression without permanently altering genetic information. Since all proteins are translated from a messenger RNA (mRNA) intermediate, targeting RNA can potentially be used to either increase protein production by inducing translation activation, to decrease protein output by degrading protein-coding mRNA, or to change the amino acid sequence of the protein product by editing. Key challenges remaining for the development and delivery of protein-based RNA-targeting therapeutics are size and immunogenicity[94,187,188]. In chapter 3, I described the development and optimization of a CRISPR/Cas-inspired RNA-targeting system (CIRTS) that was developed with these challenges in mind. CIRTS are programmable RNA-targeting proteins that are composed of three protein parts (ssRNA binding protein, RNA hairpin binding protein, and effector protein) all of which are derived from human proteins[177]. We carefully chose small proteins or protein domains, keeping the total size of CIRTS as small as possible. The fusion protein is guided to a site of interest by binding via a guide RNA (gRNA) that is composed of an RNA hairpin (bound by CIRTS) and a target-specific guiding sequence (**Figure 5.1A**). While this work laid the foundation for CIRTS-based RNA targeting, the effector scope and overall efficacy of the original system were limited. In this chapter, I will introduce our work to expand the effector scope of CIRTS, to optimize the gRNA structure for maximal cellular half-life and increased overall efficacy, and present our developmental efforts in direct protein delivery of CIRTS and its deployment in preclinical models.

5.2 Results

5.2.1 Expanding the effector scope of CIRTS

In our previous work, we chose the validated Pin nuclease domain as well as a range of RNA-modifying enzymes (YTHDF1, YTHDF2, hADAR2) to conduct the proof-of-principle optimization for RNA degradation, translation activation, and RNA editing[5,55,60,63]. Encouraged that CIRTS readily tolerated changes of the effector domain, we set out to screen new RNA effector proteins for translation activation and RNA degradation to both broaden the effector scope and increase the overall efficacy of CIRTS. First, we chose to fuse CIRTS to five proteins that have been reported to either be part of, or engage with, the translation initiation machinery in eukaryotic systems. Importantly, while there are a multitude of proteins that have annotated functions in translation initiation, only a small subset meet the small size and functional criteria of the CIRTS design. Cap-dependent translation is initiated by binding of the eukaryotic translation initiation factor 4F (eIF4F), which is composed of three proteins: eIF4a, eIF4e, and eIF4G[47]. The scaffold protein eIF4G was too large to be considered as a CIRTS effector[189]. Both of the two remaining components, eIF4a, an RNA helicase, and eIF4e, the RNA cap binding protein, fell within our design limitations and have been shown to induce translation of reporter transcripts[190,191]. We replaced the m^{6}A reader protein YTHDF1, which binds to eIF3 and the ribosome, in the original CIRTS with the eIF proteins. In a cell-based luciferase reporter assay, cells transfected with the new CIRTS effectors and an on-target luciferase gRNA generated significantly increased levels of luciferase, as compared with cells that received a non-targeting (NT) gRNA (**Figure 5.1B**). A second group of proteins known to play a role in translation initiation are pre-mRNA and poly-A binding proteins. We chose to screen three proteins from these categories in the same luciferase reporter system: stem loop binding protein (SLBP)[192], protein boule-like (BOLL)[193], and poly-adenylate binding protein (PAIP)[194]. We observed slightly lower, albeit significant, increases in luciferase produc-

tion using these effectors (**Figure 5.1B**). Taken together, we expanded the effector scope of CIRTS by five proteins capable of inducing translation activation in mammalian cells. The eIF protein family in particular showed great promise for increased overall efficacy of the system.

Aside from translation activators, we also screened several RNases and RNA degradation-inducing proteins to improve the utility of CIRTS as targeted RNA degraders. First, we switched the translation initiation effector for two known human ribonucleases, RNase1 and RNase4[195], repeated our luciferase reporter assay, and observed the anticipated decrease in luciferase levels upon transfection with RNase-CIRTS (**Figure 5.1C**). Next, we swapped the ribonuclease domains for proteins known to either be part of, or to recruit, degradation machinery, such as CNOT7 (a component of the CCR4-NOT deadenylation machinery)[196], ZFP36 and TOB2 (which promote deadenylation)[197,198]. As expected, we observed the anticipated decrease in overall luciferase levels (**Figure 5.1C**), thereby expanding the number of effective CIRTS-degraders by five.

In our proof-of-principle study (Chapter 3), we had demonstrated that CIRTS could also be used for targeted RNA editing. The hADAR protein family is the major known A-to-I editing protein family in mammalian systems, rendering further expansion of the effector scope in this aspect challenging. Nevertheless, we chose to expand our hADAR tools set by engineering CIRTS editors with different localizations within the cell. Our original CIRTS included the hyperactive E488Q hADAR variant with a nuclear localization signal. Here, we show that the CIRTS editor (wild type hADAR and the hyperactive E488Q hADAR variant) could be deployed to either the nucleus with a nuclear localization signal (NLS) or the cytoplasm with a nuclear exclusion signal (NES) while retaining their editing functionality (**Figure 5.1D**), suggesting that the CIRTS editors can be used for studies in either cellular compartment.

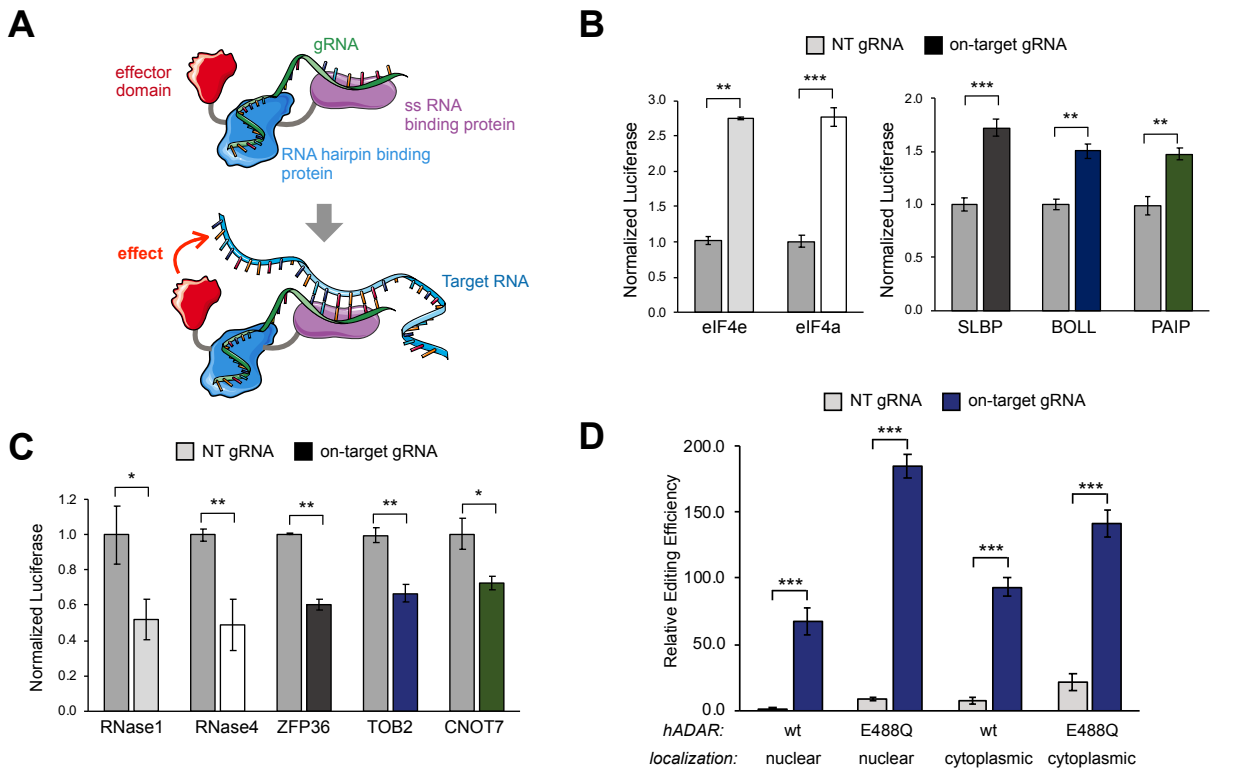

Figure 5.1 Expanded Effector Scope of CIRTS.
(A) Schematic overview of the CIRTS system. **(B)** Effector screen of CIRTS translation activators in a luciferase assay. Cells were transfected with the CIRTS effector and either NT gRNA or on-target gRNA and incubated for 48h prior to taking the measurement. Values shown as mean ± SEM with n = 3 biological replicates. **(C)** Effector screen for CIRTS degraders in HEK293T cells. Values shown as mean ± SEM with n = 2 or 3 biological replicates. **(D)** Localization screen of the CIRTS-hADAR editor. The wild type or hyperactive (E488Q) targeted hADAR editor were tested with nuclear or cytoplasmid localization in a luciferase stop codon- reversion assay. Values shown as mean ± SEM with n = 3 biological replicates. Student t-test: *P < 0.05 (*P <0.1 in C), **P <0.01, ***P <0.001.

Taken together, we conducted screens for different CIRTS effectors to induce translation activation, as well as RNA degradation, and tested if optimization of the CIRTS A-to-I editors was possible. We showed that members of the eIF protein family can be used as programmable RNA effectors to initiate translation with higher efficacy than the first-generation system. Likewise, two of the screened RNases, RNase1 and RNase4, showed increased efficacy over previously developed CIRTS degraders. These findings not only highlight the versatility of CIRTS, namely its ready toleration of changes to the

effector domain, but also open up the possibility of studies in new areas of RNA biology.

5.2.2 gRNA optimization for improved CIRTS efficacy

CIRTS targeting relies on a protein fusion containing an RNA regulatory protein, which is easily programmable by a supplied guiding RNA (gRNA). Overall efficacy of the system is determined by both the chosen effector domain as well as the stability of the gRNA in cells. To overcome effector-based limitations at the target site, we conducted the aforementioned effector screen. A second bottleneck for efficacy of RNA-guided RNA-targeting systems is the stability of the gRNA. In mammalian cells, small RNAs are rapidly degraded by exo- and endonucleases[199], which could result in premature loss of CIRTS target engagement. Previous studies with RNA-guided hADAR editors that rely on an RNA hairpin-guiding sequence to bind to their target, have shown that gRNA designs containing an RNA hairpin and both sides of the guiding sequence (**Figure 5.2A**), increase overall performance of the system[148,180]. We hypothesized that this could be due to either an increased local concentration at the target site or an increase in stability of the gRNA itself. To test this hypothesis, we used the hADAR-CIRTS system in a stop codon-reversion luciferase experiment. We transfected cells with the CIRTS editor and gRNAs containing one, two, or four TAR RNA hairpins (**Figure 5.2A**) and observed a substantial increase in editing efficiency upon addition of the second hairpin (**Figure 5.2B**). Increasing the hairpin number to two on each side, however, did not further increase editing efficiency (**Figure 5.2B**). Next, we engineered a gRNA containing one TAR hairpin to bind CIRTS and one stable hairpin sequence at the other side of the guiding sequence that cannot be bound by CIRTS (1 TAR + 1 stabilizing) to assess whether the increased editing efficiency was a result of increased local concentration of the editor or gRNA stability. We found that editing efficiency with the 2 TAR hairpin and 1 TAR + 1 stabilizing hairpin gRNA designs are comparable, indicating that protection of the gRNA from exonucleases confers an increase in performance (**Figure 5.2C**). We then verified that the 1 TAR + 1 stabilizing gRNA could still work

with our best-performing translation activator (eIF4a) and RNA degrader (RNase1) and tested their efficacy in a luciferase reporter assay. We found that both CIRTS retained their respective functions when transfected with the stabilized gRNA (**Figure 5.2D**).

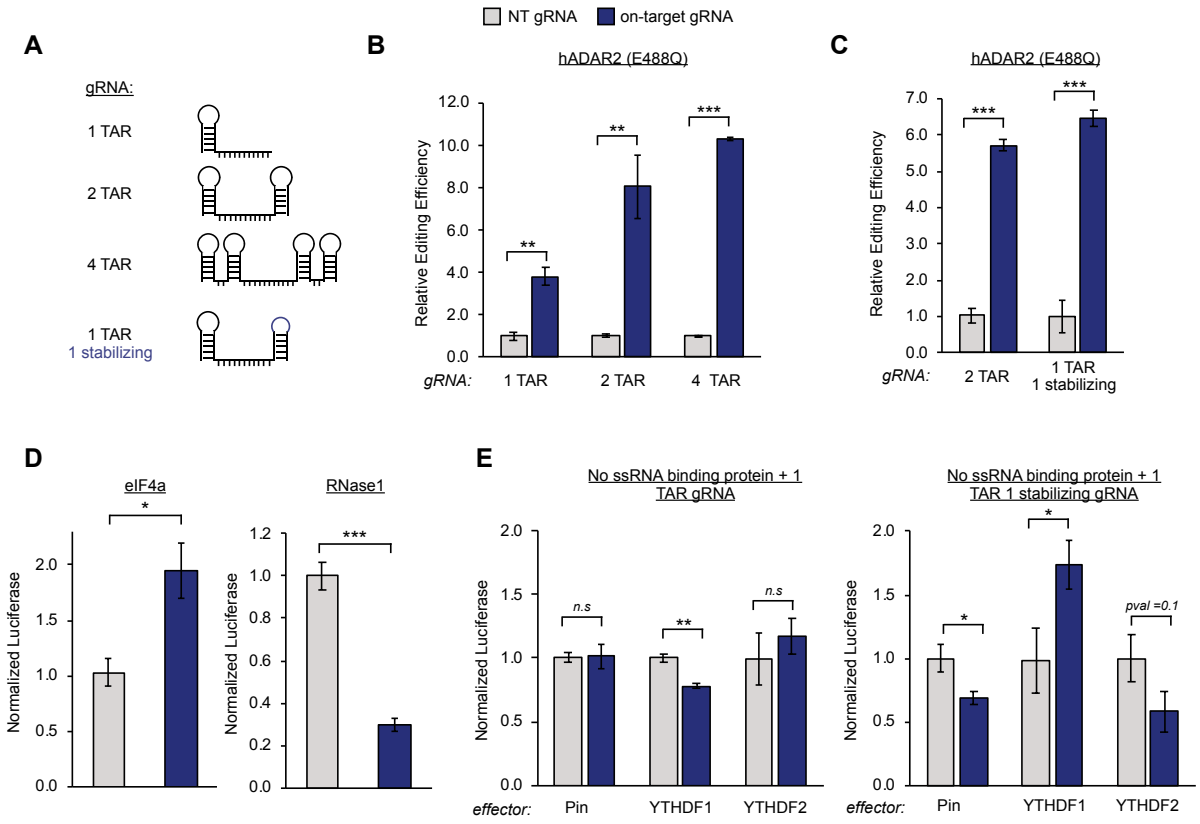

Figure 5.2 CIRTS gRNA optimization.
(A) Schematic overview of the different gRNA designs tested. **(B)** Comparison of 1 TAR, 2 TAR, 4 TAR gRNAs when cells were transfected with a CIRTS-hADAR(E488Q). Values shown as mean ± SEM with n = 3 biological replicates. **(C)** Comparison of editing efficiency of the CIRTS-hADAR editor when co-transfected with either a 2 TAR hp gRNA or with a 1 TAR + 1 stabilizing gRNA design. Values shown as mean ± SEM with n = 2 or 3 biological replicates. **(D)** We verified that non-editing CIRTS effectors retain their function with the new 1 TAR + 1 stabilizing gRNA design. We tested the new gRNA with the two best performing effector from our screen: translation initiator eIF4a, and RNA degrader RNase1. Values shown as mean ± SEM with n = 3 biological replicates. **(E)** We tested a CIRTS design without the ssRNA binding protein with either the 1 TAR or 1 TAR + 1 stabilizing gRNA. The new gRNA could rescue the effects of all tested CIRTS in the absence of the protecting peptide. Values shown as mean ± SEM with n = 2 or 3 biological replicates. Student t-test: *P < 0.05, **P <0.01, ***P <0.001.

In our original CIRTS design, we included a ssRNA binding protein to protect the gRNA from degradation prior to target engagement. We next asked if our new gRNA design, with the inclusion of the stabilizing hairpin, eliminates the need for this additional 'protecting' domain, thereby further shrinking the size of the overall system. We transfected cells with a CIRTS fusion that consists only of the hairpin binding protein and the effector domain and either a 1 TAR gRNA or a 1 TAR + 1 stabilizing gRNA. In the absence of the ssRNA binding protein, no changes in luciferase signal could be observed using the original 1 TAR gRNA design. However, in cells containing the 1 TAR + 1 stabilizing gRNA, the expected changes in luciferase signal were recovered (**Figure 5.2E**). Overall, the improved gRNA design stabilizes the gRNA and can potentially increase the efficacy of CIRTS-mediated RNA degradation and editing. Additionally, stabilizing the gRNA itself eliminates the need to include the ssRNA protein domain originally contained in the CIRTS design, ultimately yielding an even smaller system that can be delivered to cells more efficiently.

5.2.3 Direct protein delivery of CIRTS

One remaining challenge in the field of protein-based therapeutics concerns their delivery. In particular, the large size of the state-of-the-art Cas-based protein therapeutics makes packaging difficult and hinders overall efficacy[200–202]. In our proof-of-concept study (Chapter 3), we demonstrated the feasibility of efficiently packaging CIRTS in adeno-associated virus (AAV) systems. These viral-based delivery methods are efficient but frequently evoke an immune response in patients that renders them impractical, particularly with regard to therapies requiring repeated administration. Additionally, delivering plasmid DNA encoding these protein-based therapeutics relies on two processing steps in cells (DNA to RNA to protein) that cannot be controlled, making dosing challenging. We therefore set out to assess whether we can use direct protein delivery to deliver CIRTS efficiently to cells, while eliminating all dosing-related uncertainties by delivering protein directly.

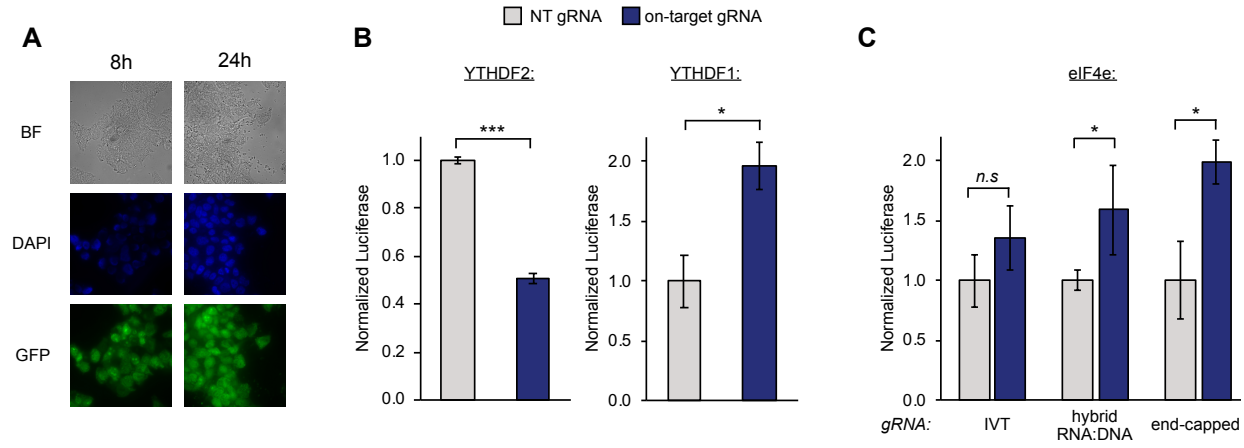

Figure 5.3 Direct protein delivery with CIRTS.
(A) 10 μM CIRTS-GFP was delivered to HEK293T using lipofectamine RNAiMAX. Cells were imaged after 8h and 24h. Representative images of n =3 biological replicates. **(B)** 166 nM CIRTS-YTHDF2/1 were complexed with either non-targeting or firefly luciferase-targeting gRNA and delivered to cells. After 24h incubation, cells were subjected to a luciferase reporter assay to measure luciferase levels. Values shown as mean ± SEM with n = 2 or 3 biological replicates. **(C)** 166 nM CIRTS-eIF4e was co-delivered to cells with either IVT, RNA:DNA, or end-capped gRNA. We found that the protected gRNAs offered an efficacy advantage, most likely due to increased stability and therefore half-life. Values shown as mean ± SEM with n = 2 or 3 biological replicates. Student t-test: *P < 0.05 (*P <0.1 for C), **P <0.01, ***P <0.001.

To visualize protein delivery in cells, we constructed and expressed a CIRTS-GFP variant, incubated it with lipofectamine RNAiMAX (a lipid-based transfection system that has been shown to efficiently deliver RNA and ribonucleoproteins[203,204]) and added the combination to cells. After 8h we observed efficient protein uptake by HEK293T cells that was sustained for up to 24h (**Figure 5.3A**). We next expressed CIRTS-YTHDF2 and CIRTS-YTHDF1 to test if the directly delivered CIRTS retained their function when expressed in *E. coli*. CIRTS-YTHDF2/1 were first complexed with an *in vitro* transcribed gRNA that contained a non-targeting (NT) or a firefly luciferase-targeting guiding sequence and then incubated with RNAiMAX. Approximately 24h after addition of the complex, we measured luciferase levels to assess the efficacy of the effectors. As anticipated, CIRTS-YTHDF2 induced a decrease in luciferase levels, while the YTHDF1-containing CIRTS led to an increased luciferase readout (**Figure 5.3B**). Importantly, the overall efficacy of the effectors

increased when using this direct protein delivery method, possibly due to better overall delivery of the CIRTS/gRNA complex as compared with plasmid delivery of individual components.

We next addressed what we hypothesized to be the remaining weak spot of this system: the small, unprotected *in vitro* transcribed gRNA. To test the effect of protection of the gRNA from exonuclease degradation, we designed and constructed both an RNA hairpin-DNA guiding sequence hybrid gRNA, as well as an end-capped gRNA. As expected, we found that protecting the gRNA by either RNA:DNA hybrid or end-capping improved the overall efficacy of the delivered CIRTS translation activator, as compared with delivery of the usual *in vitro* transcribed gRNA (**Figure 5.3C**). At this point, we conclude that CIRTS retained their function when expressed in *E. coli* and can be efficiently delivered by direct protein delivery. Importantly, we validated our hypothesis that stabilizing the gRNA sequence can improve overall efficacy of the system. In our ongoing work, we are working towards combining a smaller CIRTS, without the additional ssRNA binding protein, with chemically modified gRNAs to further improve ease of delivery and overall performance. One drawback of direct protein delivery is that it relies on the ability to generate sufficient amounts of both modified gRNA as well as protein. We are therefore actively working on improving CIRTS protein expression conditions to scale up protein expression for application in preclinical models.

5.2.4 Preclinical applications

The key advantages of CIRTS over other programmable RNA-targeting systems such as Cas13 are its small size and human-derived nature for potential therapeutic application. RNA has emerged as an exciting therapeutic target, due to its naturally high turnover rate, which provides a wide safety window and the ability to alter multiple targets without risking genomic changes. Both our efficacy and delivery optimization are targeted towards apply-

ing CIRTS in two preclinical models: wound healing and facioscapulohumeral muscular dystrophy (FSHD). For both these applications, we aim to develop orthogonal CIRTS that simultaneously degrade a subset of targets while activating protein production of others.

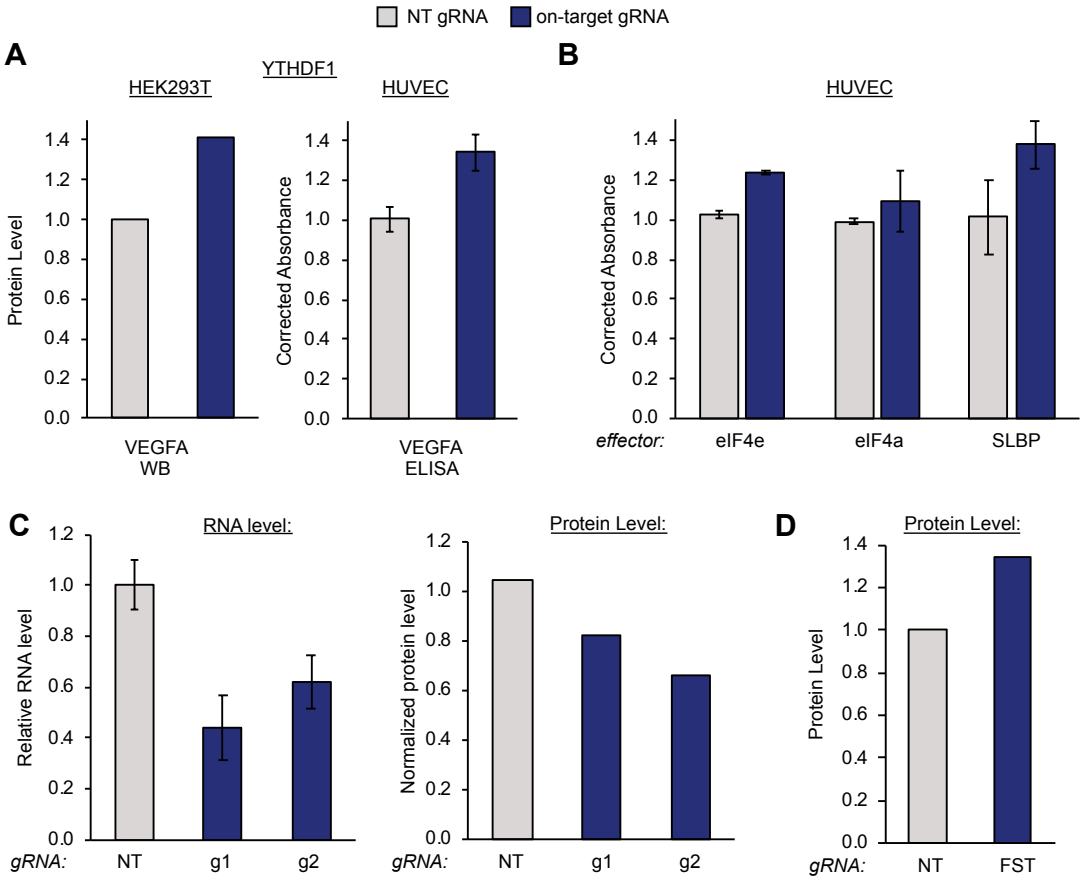

Figure 5.4 Preclinical applications of CIRTS.
(A) To assess CIRTS efficacy in a wound healing preclinical setting, VEGFA protein levels were measured in HEK293T and HUVEC cells after transfection wtih CIRTS-YTHDF1 and NT or on-target gRNA, respectively. WB: 1 biological replicate; ELISA: mean ± SEM with n = 2 biological replicates. **(B)** Protein levels of VEGFA assessed by ELISA for three different CIRTS translation initiation effectors: eIF4e, eIF4a, and SLBP. Values shown as mean ± SEM with n = 2 or 3 biological replicates. **(C)** To test CIRTS efficacy in a FSHD preclinical disease model, RNA and protein levels of DUX4 were assessed by RT-qPCR and Western blot after transfection of cells with CIRTS-RNase1. qPCR: Values shown as mean ± SEM with n = 3 biological replicates. WB: 1 biological replicate. **(D)** Protein levels measured by WB after transfection of U2-OS cells with CIRTS-eIF4e. WB: 1 biological replicate.

Wound healing has been unsuccessfully targeted by pharmaceutical companies for

decades. The intricate and tightly regulated nature of healing local injuries renders therapeutic intervention towards individual targets challenging. However, recent advances in mRNA therapy have successfully demonstrated that delivery of growth factors involved in wound healing, such as vascular endothelial growth factor (VEGF), enhances wound healing[205,206]. We therefore envisioned using a translation initiating CIRTS targeting VEGFA, while deploying a degrader CIRTS to reduce the immune response to synergistically speed up healing of wounds. In our preliminary work, we deployed CIRTS-YTHDF1 in HEK293T and observed an increase protein levels of VEGFA by Western blot (**Figure 5.4A**). Next, we verified that the system remains functional in endothelial cells, a cell type that partakes in wound healing. Both the CIRTS-YTHDF1 effector (**Figure 5.4A**) as well as three additionally tested translational initiators, eIF4e, eIF4a, and SLBP (**Figure 5.4B**), led to increased protein levels as measured by enzyme-linked immunosorbent assay (ELISA). In our ongoing work, we are working to screen additional gRNA sequences targeting VEGFA to identify an optimal targeting region, as well as combining these translation activators with orthogonal RNA degrading CIRTS. After verifying increases in growth factor levels, we will subject cells to a growth factor-stimulated growth experiment to assay phenotypic changes. Furthermore, we are actively working with a collaborator at the University of Chicago Medical School to deploy translation activating CIRTS in their validated mouse model of wound healing.

A second preclinical application that we are actively pursuing is facioscapulohumeral muscular dystrophy. FSHD is a genetic disease of the skeletal muscle caused by a toxic gain-of-function mutation in the transcription factor coding gene *DUX4*. The mutation causes high expression levels of the usually repressed gene, resulting in muscle weakness and atrophy in the face, abdomen, arms, and legs[207–209]. Currently, there is no approved therapy for FSHD. Recently, a gene therapy aimed at improving muscle strength by delivering the natural myostatin antagonist follistatin has shown great promise for improved functional outcomes in a FSHD mouse model[210,211]. We are therefore aiming at developing orthogonal CIRTS that simultaneously degrade the aberrantly produced DUX4

transcripts, while activating protein production of the muscle-protecting protein follistatin. In our preliminary studies, we were able to show that we can use CIRTS-RNase1 to degrade DUX4 RNA, thereby lowering DUX4 protein levels (**Figure 5.4C**). Moreover, we can use CIRTS-eIF4e to activate translation of follistatin, increasing levels of this protein, as measured by Western blot (**Figure 5.4D**). In ongoing work, we aim to consolidate our CIRTS/gRNA vectors to enable studies of orthogonal CIRTS targeting in mammalian cell culture with higher efficiency by decreasing the number of transfected vectors. Once optimized further in tissue culture models, we aim to deploy these orthogonal CIRTS in a FSHD mouse model by direct protein delivery. FSHD is an ideal exemplar disease for RNA-targeting gene therapy because the condition benefits from simultaneous targeting of multiple transcripts. Further developing and deploying CIRTS in these exemplar preclinical applications will not only improve the efficacy of the system but also allow us to learn valuable lessons for improved delivery in cell culture and *in vivo* models.

5.3 Discussion

In summary, we have continued to optimize CIRTS by expanding its effector scope, as well as improving upon our gRNA design. Additionally, we detailed our advances towards using CIRTS in direct protein delivery and in preclinical applications of wound healing and FSHD. These improved and optimized CIRTS open up new opportunities to study components of the translation initiation and degradation machineries in the native cellular context.

For translational applications, we report improved overall efficacy for targeted protein production or RNA degradation. Additionally, this work expanded the validated delivery options of CIRTS from AAV-based delivery to direct protein delivery. We found that even simple, lipid-based transfection reagents ensure efficient delivery and improved performance in cell culture without the need for major optimization. Together with the ability to

remove the ssRNA binding protein in the presence of a chemically modified gRNA, this demonstration lays the foundation to explore additional delivery methods, including lipid nanoparticle (LNP) delivery. Regardless of the protein delivery method, direct protein delivery relies on the ability to produce large amounts of modified gRNA and protein, which is not without its own practical challenges. Currently, we are working on adjusting the conditions for CIRTS expression to allow for large scale protein production. Should this prove to be more challenging or time consuming than anticipated, we envision testing mRNA delivery of CIRTS. As mRNA has become an attractive therapeutic itself, the LNP and liposome delivery conditions for cell culture and *in vivo* models are established and can readily be tested in our established preclinical applications. This expansion of the CIRTS technology lays the foundation for specific biological applications and paves the way for *in vivo* preclinical validation.

5.4 Materials and Methods

Cloning.

All cloned plasmids were generated using Gibson Assembly (GA) and sequenced by the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility. The PCR fragments used in GA were amplified using Q5 DNA Polymerase (NEB). All plasmids used in this chapter are listed in Table 5.1 with links to their vector maps.

Mammalian Cell Culture and Transfections.

Cell culture experiments were conducted in HEK293T (ATCC), HUVEC (ATCC), and U2-OS (ATCC) cells. HEK293T cells were maintained in DMEM (L-glutamine, high glucose, sodium pyruvate, phenol red; Corning) media supplemented with 10% fetal bovine serum (FBS; Gemini Benchmark), and 1x penicillin/streptomycin (P/S; Gibco/Life Techno-

logies). HUVECs were cultured in Medium 200 (Gibco/Life Technologies) supplemented with low serum growth supplement (LSGS, Gibco/Life Technologies). For transfections, P/S was omitted from the media. U2-OS cells were cultured in McCoy 5a medium (ThermoFisher) supplemented with 10% fetal bovine serum (FBS; Gemini Benchmark), and 1x penicillin/streptomycin (P/S; Gibco/Life Technologies). All HEK293T cell transfections were conducted using lipofectamine 2000 (Invitrogen) according to the manufacturer's recommendations. HUVEC cells were transfected using TransfeX (ATCC). U2-OS cells were transfected with lipofectamine LTX (Invitrogen).

Mammalian Luciferase Assays.

For luciferase assays, HEK293T cells were transfected with 12.5 ng reporter plasmid, 150 ng indicated CIRTS vector, and 100 ng gRNA expression vector. About 16h before transfection, cells were plated on 96-well plates and grown to 70-80% confluency overnight. Cells were transfected with 20 µL Opti-MEM I Reduced Serum Medium (ThermoFisher Scientific). DNA plasmids were combined in 10 µl Opti-MEM and incubated for 15 min before being combined with 0.5 µl lipofectamine 2000 per well. Cells were incubated for 48h after transfection before luciferase reads were taken on a Biotek Synergy plate reader. Both firefly and *Renilla* luciferase readouts were measured as previously reported[212]. 40 µL growth media were removed from every well before addition of 40 µL firefly assay buffer (Triton Lysis Buffer (50 mM Tris, pH 7.0, 75 mM NaCl, 3 mM MgCl2, 0.25% Triton X-100) containing 5 mM DTT, 0.2 mM coenzyme A, 0.15 mM ATP, and 1.4 mg/mL luciferin) to lyse the cells and to provide the first substrate for firefly luciferase. Cells were then incubated for 10 min before the first firefly luciferse read was taken. Immediately thereafter, 40 µL *Renilla* assay buffer (45 mM EDTA, 30 mM sodium pyrophosphate, 1.4 M NaCl, 0.02 mM PTC124, 0.003 mM coelentrazine h (CTZ-h)) was added to stop firefly luciferase activity and provide the substrate for *Renilla* luciferase and the second read was taken. All experiments were conducted in biological triplicates. Firefly luciferase luminescence levels were normalized to the corresponding *Renilla* luminescence levels

to generate the normalized change in protein levels from the target firefly luciferase gene.

Protein expression.

All CIRTS building blocks were purchased as gene blocks from IDT and were E. coli codon optimized. CIRTS were cloned into a pET vector-derived pMCSG19 protein expression vector (MBP-TVMV-6xHis-TEV) using Gibson Assembly and transformed into BL21(DE3) cells (NEB). A 10 mL seed culture was used to inoculated 2L of 2XYT (16 g/L digest peptone, 10 g/L yeast extract, 5 g/L sodium chloride, US Biological) supplemented with 100 µg/mL carbenicillin. When cells reached an optical density of 0.6 OD600 at 37 C, they were chilled to 16 C on ice before being induced with 0.5 mM IPTG (bioWORLD). Induced cultures were grown at 16 C for 12-16 h before the cells were lysed to harvest the protein.

Cells were pelleted by centrifugation at 1,500 g for 15 min at 4°C. The resulting pellet was stored at -80°C or directly purified further. First, cells were lysed with 100 mL lysis buffer (50 mM Tris, 1 M NaCl, 20% glycerol, 10 mM TCEP, pH 7.5) supplemented with protease inhibitors. They were then resuspended and lysed using sonication (Thermo Fisher). Lysates were cleared by centrifuging at 12,000 g for 40 min at 4°C. Cleared lysates were incubated with His60 Ni Superflow Resin (Takara) for 1h at 4°C with constant agitation. Then, the resin was washed extensively using lysis buffer and eluted with a gradient imidazole elution (10 mM-500 mM). Fractions containing the protein were pooled together and concentrated using Ultra-50 Centrifugal Filter Units with 30 kDa cutoff (Amicon, EMD Millipore). After concentration, the solution was desalted and buffer exchanged into lysis buffer (50 mM Tris, 1 M NaCl, 20% glycerol, 10 mM TCEP, pH 7.5) using Sephadex G-25 in PD-10 Desalting Columns (GE Healthcare Life Sciences).

Next, the MBP tag of CIRTS was cleaved off using TEV protease (Genscript) according to the manufacturer's protocol. CIRTS protein (up to 3 µg/1 IU TEV) was incubated in TEV buffer (20 mM Tris-HCl, 50 mM NaCl, pH 7.4) and protease for 4h at 30°C. The cleaved

protein was then directly re-incubated with His60 Ni Superflow Resin (Takara) for 1h at 4°C with constant gentle agitation. The supernatant containing the cleaved protein was collected and immediately concentrated using Ultra-50 Centrifugal Filter Units with 10 kDa cutoff (Amicon, EMD Millipore). The protein was then stored at -80°C or directly used for protein delivery. The concentration was measured using a BCA Assay (Thermo Scientific).

***In vitro* transcription of gRNAs.**

DNA templates for the *in vitro* transcription (IVT) reactions were purchased from IDT (some guiding sequence as the firefly luciferase 3'UTR gRNA) and contained the RNA of interest in addition to a T7 RNA polymerase (RNAP) promoter. IVT reactions were set up at 250 µL volume by combining 12 µg PCR product, 1x transcription buffer (40 mM Tris-HCl, 2 mM spermidine, 10 mM NaCl), 25 mM MgCl2, 10 mM DTT, 40U SUPERaseIn, 4 mM of each NTP, and 40 µg/mL T7 RNAP. The reactions were incubated at 37°C overnight. The next day, the mixture was treated with DNaseI in 1x DNaseI buffer for 30 min at 37°C. RNA was then separated on a 10% 8M TBE-urea PAGE gel and gel-purified using the ZR small-RNA PAGE Recovery Kit (Zymo).

Direct protein delivery.

For direct protein delivery in a 96-well plate (Corning), 166 nM of expressed protein was complexed at a 1:1 ratio with the *in vitro* transcribed gRNA (either non-targeting or firefly luciferase-targeting) in 50 µL Opti-MEM and incubated for 5 min. Then, 1 µL of lipofectamine RNAiMAX per well was added and the mixture was allowed to incubate for an additional 15 min before being slowly added to cells. After 8h or 24h, cells incubated with CIRTS-GFP were imaged on an inverted epifluorescence microscope (Leica DMi8) equipped with a Hamamatsu Orca-Flash 4.0 camera and a 300W Xenon light source (Sutter Lamda XL). Leica LASX software was used to collect images of CIRTS-GFP, Hoechst 33342, and bright field. For any other CIRTS effectors, stable HEK293T-luciferase reporter cells were incubated for 24h after addition of the ribonucleoprotein complex and

used in a luciferase reporter assay as described above.

RT-qPCR.

To verify changes in RNA levels after transfection with CIRTS, we conducted a RT-PCR experiment. Cells were plated on a 96-well plate (Corning) and transfected with CIRTS and gRNA plasmids. Total RNA was harvested 48h after transfection using the RNeasy Mini Kit (Qiagen). The resulting RNA was then reverse transcribed into cDNA using the PrimeScript RT Reagent Kit (TaKaRa). qPCR reactions were run on a LightCycler 96 Instrument (Roche) using FastStart Essential DNA Green Master (Roche). Expression levels were determined using a housekeeping control gene (GAPDH) cycle threshold (C_t) value and the gene of interest C_t value. The relative expression level of one target transcript was calculated by $2^{-\Delta C_t}$, where $\Delta C_t = C_t$. Relative expression level for targeted gene was obtained upon normalizing the targeted gene expression level of cells experiments treated with the on-target gRNA to those treated by the nontargeting (NT) gRNA. All qPCR primers are listed in Table 5.4.

Enzyme-linked immunosorbent assay (ELISA).

A human VEGF ELISA (Biolegend) was used to quantify protein levels in HEK293T and HUVEC cells. Cells were plated on 6-well plates (Corning) and transfected with 1.5 µg of a CIRTS vector and 1.0 µg of a gRNA vector at 80% confluency. After 48h incubation, the cells were lysed with 75 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% deoxycholate, 2% SDS, pH 7.4) supplemented with protease inhibitors. One day prior to running the ELISA, a 96-well ELISA plate (Nunc) was coated with 1x capture antibody in 1x coating buffer A and incubated overnight at 4°C. The day of the experiment, the plate was washed four times with PBST wash buffer (1x PBS + 0.05% Tween-20) and blocked with 1x Assay Diluent A for 1h at room temperature while shaking constantly. The plate was then washed four times with PBST wash buffer before loading of the samples. Each experiment included a full VEGFA standard curve in duplicate to ensure the experimental samples fall within

the detection range of the ELISA kit. Each sample and standard well received 50 µL Assay Diluent D and 50 µL diluted standard or experimental sample and was incubated for 2h at room temperature with constant agitation. After the sample incubation, the plate was washed four times with PBST wash buffer before being incubated for 30 min at room temperature with 1x Avidin-HRP while shaking. Then, the plate was washed five times with PBST and developed using 100 µL substrate solution D and incubating the plate in the dark. After 10 min, the developing reaction was quenched by adding 50 µL 1M H_2SO_4 solution and the absorbance was read out at 450nm and 570nm. The corrected absorbance reading was calculated by subtracting the read at 570nm from the 450nm read. All experiments were conducted in duplicate or triplicate.

Western blot.

For Western blots, U2-OS cells were plated on 6-well plates (Corning) and transfected with 1.5 µg of a CIRTS expression vector and 1.0 µg of a gRNA expression vector using lipofectamine LTX (Invitrogen). After 48 h, the cells were prepared for lysis by washing with PBS twice. Then, 75 µL RIPA buffer (50 mM Tris, 150 mM NaCl, 0.5% deoxycholate, 2% SDS, pH 7.4) supplemented with protease inhibitors was added and the cells were allowed to incubate for 30 min at room temperature. The concentration was measured by BCA assay (Thermo Scientific) and 50 µg protein was boiled in protein loading buffer (50 mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.05% bromphenol blue, 100 mM DTT) for 10 min at 95°C and loaded onto a 10% SDS PAGE gel. Gels were run at 140 V until the dye reached the bottom of the plate, and protein were transferred onto a PVDF membrane (Millipore), before being blocked in 5% nonfat milk in TBST for 1h. Proteins were then detected using 1:100 mouse anti-Dux4 antibody (Santa Cruz) and 1:100 mouse anti-follistatin antibody (Santa Cruz), followed by 1:1000 anti-mouse HRP-conjugated antibody (Santa Cruz). The loading control GAPDH was visualized using 1:5000 HRP-conjugated anti-GAPDH antibody (Proteintech). Membranes were imaged on a Fluor Chem R (Protein Simple) imager after incubation with Super Signal West Pico Plus (Thermo Scientific).

5.5 Supplemental Information

Table 5.1 Mammalian expression plasmids used in this chapter. All mammalian plasmid have a cmv d0 promoter.

Name:		Description:	Benchling Link:
CIRTS-eIF4a	33-02	Cmv b-defensin3-TBP6.7-eIF4a	https://benchling.com/s/seq-jii7Otn-wgWr0srP9bG1q
CIRTS-eIF4e	29-15	Cmv b-defensin3-TBP6.7-eIF4e	https://benchling.com/s/seq-pfj6XdYQsLloCA1v3HhO
CIRTS-SLBP	33-12	Cmv b-defensin3-TBP6.7-SLBP	https://benchling.com/s/seq-4xVlFAN-JDgDP525C19P9
CIRTS-BOLL	33-14	Cmv b-defensin3-TBP6.7-BOLL	https://benchling.com/s/seq-4MKQiywl87xd50IiG40L
CIRTS-PAIP	33-04	Cmv b-defensin3-TBP6.7-PAIP	https://benchling.com/s/seq-2FgA1n9Y22BfJHVplzcb
CIRTS-RNase1	40-30	Cmv b-defensin3-TBP6.7-RNase1(29-156)-NLS	https://benchling.com/s/seq-hpcYndl9VTW8LKP3enhK
CIRTS-RNase4	40-29	Cmv b-defensin3-TBP6.7-RNase4(29-147)-NLS	https://benchling.com/s/seq-8NK5wD8MbCg8l6sKgqJm
CIRTS-ZFP36	32-43	Cmv b-defensin3-TBP6.7-NES-ZFP36	https://benchling.com/s/seq-6Xtuip-GhJBWt7tdtjxGk
CIRTS-TOB2	32-42	Cmv b-defensin3-TBP6.7-NES-TOB2	https://benchling.com/s/seq-epGFTEaPXzPI4Y8lS4o8
CIRTS-CNOT7	39-34	Cmv b-defensin3-TBP6.7-NES-CNOT7	https://benchling.com/s/seq-zNs9rjEYUP4dbKd424Nc

Table 5.2 gRNA vectors used in this chapter. All gRNA plasmids have a hU6 promoter.

Name:		Description:	Benchling Link:
1TAR gRNA-NT	25-61	hU6-1TAR hp-non-targeting gRNA	https://benchling.com/s/seq-RxXTYSa29wJGkr0DUsh7
2TAR gRNA-NT	24-69	hU6-1TAR hp-non-targeting sequence-1TAR hp gRNA	https://benchling.com/s/seq-7Kbgu-pAST3ZLDw8iw1Cj
1TAR + 1stabilizing gRNA-NT	35-44	hU6-1TAR-non-targeting sequence-1 stabilizing hp gRNA	https://benchling.com/s/seq-Rcs4lwzb-SHiCdcswCG6F
4TAR gRNA-NT	32-66	hU6-2TAR hps-non-targeting gRNA-2 TAR hps gRNA	https://benchling.com/s/seq-PfM4blnR-WmkBNlzBL6CA

Table 5.3 gRNA guiding sequences used in this chapter. All used in the same hU6 promoter as TBP-OT.

Name:		Sequence:
Fluc-5'UTR	25-66	GGTGGCTTTACCAACAGTACCGGATTGCCAAGCTTGGGCT
Fluc-3'UTR	18-20	CAGGTCGACTCTAGACTCGAGGCTAGCGAGCTCGTTTAAA
Fluc-mismatch 20	31-47	CGATGTCGCCGCTGTGCAGCCAGCCGTCCTTGTCGATGAG
VEGFA	37-80	TCCTCGGCGCCTCGGCGAGCTACTCTTCCTCCCCGGC-CCGA
DUX4-g1	38-32	TTCGCCGGCCTTCTGGCGGGCCGCGTCTCCCGGGC-CAGGG
DUX4-g2	37-76	TGCCTGCGCGGGCGCCCTGCCACCCTGTCCCGGGT-GCCTG
FST	38-31	CAAAGGCTATGTCAACACTGAACACTTATAGAGAGTTTAC

Table 5.4 qPCR primers used in this chapter.

Name:	Forward:	Reverse:
DUX4	GCGACGGAGACTCGTTTG-GAC	CTGGCGTGACCTCT-CATTCTGA

CHAPTER 6

SUMMARY AND PERSPECTIVES

6.1 Improved tools to control gene expression regulation at the RNA level

RNA regulatory pathways have emerged as key influencers of gene expression in recent years. Novel discoveries and in-depth studies, both enabled by new techniques and tools, have uncovered the importance of these diverse RNA regulatory layers. For example, chemical modifications, initially discovered in the 1970s[213–215], have only recently been shown to play a role in virtually every cellular process[26,138]. The discovery of each novel chemical modification or RNA regulatory protein is frequently functionally validated by genetic manipulation such as knockout/overexpression[18,28,29]. However, detailed mechanistic studies are often hampered by a lack of tools for site-specific manipulation within endogenous cellular contexts. Aside from biological studies, RNA and its regulation have also become an exciting therapeutic target. Both the possibilities to precisely regulate protein production and sequence at the RNA level and the high turnover rate of RNA provide interesting new options for polygenic disease treatment.

Programmable RNA-targeting systems have opened up new opportunities to both study and control individual transcripts in their endogenous environment. The CRISPR/Cas protein families have changed how nucleic acids are studied and manipulated by providing an easily programmable delivery moiety that is customized via simple changes in a short guiding sequence in the gRNA[6–8]. Apart from programmable RNA degradation with the native CRISPR/Cas system, nuclease-inactivated versions of proteins (e.g. dCas) have been fused to GFP for imaging[59], hADAR for RNA editing[60,61], and hnRNPa1 for splicing applications[62]. Additionally, the RNA-targeting RCas9 variant of Cas proteins has been successfully validated in preclinical models of toxic microsatellite repeat expansion

diseases, such as Huntington's disease, C9-ALS, and myotonic dystrophy[55]. Together, these discoveries and early applications of programmable RNA-targeting proteins showcase the need for similarly precise synthetic biology tools, as well as opportunities for both basic and translational research driven by RNA targeting.

In my thesis work, I first summarized the development and validation of programmable m^{6}A reader proteins to study epitranscriptomic regulation of single RNAs (Chapter 2). These programmable m^{6}A reader proteins were guided by Cas13, a large member of the microbial Cas protein family. To address concerns with immunogenicity and size limitations of the Cas13 RNA-targeting system, I developed the CRISPR/Cas-Inspired RNA-targeting System (CIRTS; Chapter 3). CIRTS are composed of all human-derived protein domains and are about half the size of an unfused Cas13 protein. In Chapter 4, I presented a small molecule-inducible biosensor for temporal control of RNA regulation based on CIRTS. Lastly, I showcased our most up-to-date advancements in CIRTS efficacy by conducting an effector protein screen and gRNA optimization, followed by our preliminary work in direct protein delivery and deployment of CIRTS in preclinical models (Chapter 5).

In Chapter 2, I first introduced a new set of programmable m^{6}A reader proteins that rely on dCas13b as the targeting moiety. We chose to fuse a nuclease-inactivated version of Cas13b to two members of the YTHDF m^{6}A reader protein family: YTHDF1, which interacts with the translation machinery to induce protein translation[24], and YTHDF2, which recruits the RNaseP pathway and the CCR4-NOT deadenylation machinery to promote RNA degradation[28,29]. We demonstrated that these dCas13-based fusion systems can be targeted to both a reporter transcript or an endogenous RNA of interest to either induce translation activation (via YTHDF1) or RNA degradation (via YTHDF2). Importantly, we observed that dCas13b-YTHDF2 induced a level of RNA degradation on par with that of active Cas13 nuclease. Taken together, we developed programmable m^{6}A reader proteins that enable biological studies of site-specific m^{6}A reader effects and provide the means to control gene expression at the RNA level by inducing transcript-specific protein

production or nuclease-independent RNA degradation.

As presented in Chapter 3, I developed a CRISPR/Cas-Inspired RNA-targeting system that reduces the overall size of programmable RNA-targeting systems by up to three-fold and is composed of all human protein parts to alleviate the potential for immunogenicity in consideration of future therapeutic deployment. CIRTS is composed of a fusion protein combining a non-specific RNA binding protein, a RNA hairpin binding protein, and an RNA effector protein, as well as a guiding RNA that combines an RNA hairpin and a guiding sequence. The core of the fusion protein can bind to the RNA hairpin on the gRNA and the guiding sequence engages the target RNA using base pair complementarity. We validated this system both *in vitro* and *in cellulo* with a known Pin nuclease domain as the effector and observed gRNA-dependent RNA degradation that was slightly less efficient, albeit comparable, to Cas13. Encouraged by these findings, we next tested the modularity of the system, showcasing its tolerance of switches in each of the individual components (the ssRNA binding protein, the RNA hairpin binding protein, and the effector protein), while retaining effector function on both a reporter, as well as on endogenous targets. Due to their smaller size and modularity, CIRTS is ostensibly capable of simultaneously targeting and delivering orthogonal effectors to different targets. Lastly, we showcased the advantage of smaller size by efficiently packaging CIRTS into an AAV system, a feat that had previously proved difficult with even unfused Cas protein due to AAV packaging limits.

While both CIRTS- and Cas-based RNA-targeting system provide the means to study and control gene expression at the RNA level, they do not allow for studies of temporal dynamics in RNA regulation. To overcome this limitation, we expanded the CIRTS platform into a small molecule-inducible RNA-targeting system based on the abscisic acid (ABA) heterodimerization domains (Chapter 4). The targeting component of CIRTS (ssRNA binding protein and RNA hairpin binding protein) was fused to one heterodimerization protein ABI, while the effector was fused to the other, PYL. Upon addition of ABA, the two domains

dimerize and reassemble a fully functional CIRTS at the target site. We validated this temporally controlled CIRTS biosensor, this time combined with an hADAR editing protein effector domain, in a luciferase stop codon-reversion assay and conducted a gRNA optimization to increase efficacy. Then, we conducted a dynamic characterization by measuring ABA-induced editing levels on the reporter transcript over the course of three days. Encouraged by the inducible RNA editor, we swapped the effector domain to the previously validated Pin nuclease domain, or m^6A readers (YTHDF1 and YTHDF2), and, as expected, observed ABA-dependent target RNA degradation (via Pin and YTHDF2) and translation (via YTDHF1). We further expanded the applications of the small molecule-inducible editor to disease-simulating and endogenous transcripts. Lastly, and importantly, we deployed the inducible CIRTS editor in an *in vivo* luciferase stop codon-reversion assay and measured robust RNA editing with both the full length CIRTS editor and the ABA-inducible editor upon administration of ABA. Addition of the small molecule-inducible components broadens the applicability of CIRTS for biological and translational applications by providing an additional layer of control.

Lastly, I presented our work in expanding the CIRTS toolbox and further improving its efficacy, as well as initial steps towards direct protein delivery in preclinical applications (Chapter 5). In an effort to improve efficacy of the targeted translation activators and RNA degraders, we conducted an effector screen and identified five additional translation activators and five additional RNA degraders that show substantial gRNA-dependent effects. Most noticeably, the eIF proteins and human RNase1/4 stood out as the most efficient effectors. We then optimized our gRNA structure for improved stability by addition of a stabilizing hairpin to the hairpin-unprotected side of the gRNA sequence. Addition of this gRNA feature improved efficacy of the CIRTS editor and degraders and eliminated the need for the ssRNA binding protein, reducing the size of CIRTS even further. Next, we conducted preliminary experiments in direct protein delivery by first delivering an imaging-compatible CIRTS-GFP. Once we visually validated successful delivery, we deployed previously validated effectors and, importantly, recapitulated our previous findings with the

CIRTS-YTHDF1, -YTHDF2, and -eIF4e systems. Notably, the overall performance of the system increased, suggesting the protein delivery is more efficient than transfection of three DNA plasmids. Lastly, we began testing the utility of CIRTS in preclinical applications. We chose to focus on wound healing and FSHD. As a first step, we validated CIRTS-mediated upregulation of important growth factors for wound healing, such as VEGFA, in both HEK293T and HUVEC cells. We plan on expanding this area of research in phenotypic growth assays and ultimately in *in vivo* wound healing models in conjunction with our collaborators. For the FSHD applications, we verified that we CIRTS-dependently degrade the aberrantly expressed DUX4 protein that causes FSHD and upregulate the natural myostatin antagonist follistatin in a model cell line. We are now working on simultaneous targeting of both targets, with hopes of expansion into mouse models of FSHD in the future.

Ultimately, the work presented in this thesis details an expansion of programmable RNA-targeting tools to study and control RNA regulation in an endogenous context. First, by expanding the list of the Cas-based RNA-targeting applications via fusion with m^6A reader proteins, I showcased the amenability of dCas13 to the swapping of the effector domain and therefore its function as a generalizable RNA-targeting platform. Building upon this foundation, I designed and developed a new RNA-targeting system that is equally as versatile yet is significantly smaller and fully derived from human parts. These characteristics make this syste, a promising candidate for future clinical deployment.

6.2 Programmable RNA targeting to study biology

To address basic biology questions, CIRTS provides a versatile platform to study the effects of RNA binding and RNA regulatory proteins in their endogenous context. In our work, we demonstrated that CIRTS readily tolerates switches of the effector domain, allowing for studies of various regulatory pathways. A commonly employed cell-based assay to

study the effects of RNA binding/regulatory proteins on RNA is an RNA tethering assay[49]. In tethering assays, the RNA of interest is fused to multiple RNA hairpins and exogenously introduced by plasmid transfection. The effector protein is then fused to the binding partner of the introduced RNA hairpin (e.g. MS2 hairpin and MS2-coat protein (MCP)) and upon transfection of both components, the fusion protein can bind to the tagged RNA. Often, a reporter RNA, such as luciferase or GFP, is chosen to simplify the measurement of changes in RNA levels. Effects of the protein on the RNA transcript can then be assessed by RT-qPCR (changes in RNA level), luciferase measurement (degradation or translation), or Western blot (protein production). This RNA tethering approach is also frequently used to image individual RNA transcripts by fusing the delivered RNA hairpin-binding protein to fluorescent proteins. Tethering assays have provided invaluable information about the regulation and localization of individual RNA transcripts throughout the years. However, these assays rely on introduction of exogenously overexpressed and tagged RNA, which could alter the RNA's native regulatory environment.

With CIRTS, we hope to provide a small, minimally perturbative yet equally as versatile tool to study location or function of RNA binding proteins on individual RNA transcripts of interest. CIRTS could be used to validate the effects of RNA regulatory proteins identified in standard reporter tethering assays on endogenous transcripts. For example, the functions of m^{6}A reader proteins, YTHDF1 and YTHDF2, have been verified using RNA reporter transcripts fused to multiple RNA hairpins[24,28]. Recent advances in the development of dCas9-guided m^{6}A writers and erasers suggest that m^{6}A positioned in the 5'UTR versus the 3'UTR could have different functional outcomes[216]. To elaborate this finding, known m^{6}A reader proteins, such as YTHDF1 and YTHDF2, could be targeted with CIRTS to determine which reader is responsible for these proposed site-dependent functional consequences. Additionally, CIRTS-guided methyltransferases and erasers could be developed to provide smaller and therefore less perturbative tools to introduce or erase m^{6}A site-specifically and to lay the ground work for a more mechanistic understanding of the roles of individual chemical modification sites. In its essence, CIRTS provides the

means for an 'endogenous tethering assay', eliminating the need to overexpress tagged RNAs for biological studies. Additionally, CIRTS-GFP could be used to image endogenous RNA transcripts by supplying multiple gRNA per transcript. Recent advances in RNA imaging with Cas9 and Cas13[58,217] suggest that an RNA-targeting system approach to image endogenous RNA transcripts is feasible. However, to-date only rather abundant RNA transcripts could be imaged successfully. We imagine that the small size of CIRTS would make delivery of multiple CIRTS-GFP easier and would allow for even more gRNAs on each transcript; this would enable the resolution of even less abundant species. For an even more-improved signal-to-noise ratio at the single transcript level, a Sun-tag approach (CIRTS-24x Sun tag and scFv4-GFP) could be employed for signal amplification for imaging[218].

6.3 Programmable RNA targeting for therapeutic applications

With the development of CIRTS, we hope to contribute to overcoming some of the challenges faced by RNA-targeting therapeutics. Currently, the field of RNA-targeting therapeutics is expanding rapidly, with many biotechnology companies being founded around nucleic acid-, small molecule-, and Cas protein-based RNA-targeting moieties. In fact, with the approval by the FDA of Biogen's Spinraza in December 2016, antisense oligonucleotides (ASOs), which can be programmed to bind an RNA transcript of interest using base pair complementarity to perturb RNA function, have now entered the therapeutic market. Spinraza, marketed for the treatment of spinal muscular atrophy, binds to *SMN2* mRNA, altering its splicing and partially restoring otherwise extremely low SMN2 protein levels[219]. Twenty years after their initial discovery, RNAi drugs finally also entered the therapeutic market in 2018, with the approval of Alnylym's Onpattro[220] and Ionis' Tegsedi[221]. Both of these RNAi therapeutics were approved for treatment of polyneuropathy in hereditary transthyretin-mediated amyloidosis, and function by binding to and

inducing degradation of an abnormal protein coding mRNA of transthyretin. The approval of these initial nucleic acid-based RNA therapeutics lays the foundation for clinically viable RNA-based therapeutics. However, it remains to be seen if protein-based systems, such as CIRTS, can achieve the 20-year timeline from discovery to the clinic experienced by ASOs and RNAi-based therapies. Seven years after their re-discovery, Cas- and other protein-based approaches for DNA-targeting therapies have only now entered clinical trial stages in the United States. Both these DNA-targeting, as well as future RNA-targeting, protein-based therapeutics still face multiple technical hurdles that include delivery, efficacy, and immunogenicity. First and foremost, the large size of many RNA-targeting proteins, such as the Cas protein family, confers several challenges. Virtually all delivery methods for either plasmid, mRNA, or protein delivery are limited in packaging capacity, which often precludes delivery of large fusion proteins of potential therapeutic systems. With increased size, viral packaging becomes inefficient to impossible. The same issues arise with mRNA or direct protein delivery facilitated by cubosomes, lipid nanoparticles, or liposomes. All these delivery methodologies suffer significant losses to their otherwise high efficiency when used to package large cargos. Moreover, once packaged and delivered, proteins with microbial origin can induce immune reactions, especially when administered repeatedly, as is necessary when targeting a high turnover molecule such as RNA. We developed CIRTS in hopes of being able to address these size and immunogenicity problems with bioengineering methodologies.

We note that the CIRTS platform is not without its own challenges. First, the overall efficacy of the system can be relatively low, especially for some translation activators. It will likely require extensive optimization in the context of specific preclinical applications to achieve a relevant level of therapeutic efficacy. However, for some therapeutic applications, as for example wound healing, relatively modest changes in RNA and therefore protein levels are desired to avoid side effects such as cancer. Additionally, CIRTS is a double or triple fusion protein composed of human parts, which renders efficient protein expression for direct protein delivery in lower prokaryotic systems, such as *E. coli*, challen-

ging. Even in the presence of solubility tags, we observe relatively low overall expression yields and substantial inclusion body formation. Furthermore, we have observed significantly different levels of expression for different effector domains, which we believe will either require a switch to a higher order organism, such as yeast, as a universal expression platform, or, alternatively, individual optimization of protein expression conditions for each new CIRTS-effector pair. Nevertheless, similar weaknesses have been previously overcome in the development of other RNA-targeting systems based on the Cas proteins, which leads us to believe that they will likely be addressed with additional optimization and development of CIRTS.

REFERENCES

1. Shapira, P.; Kwon, S.; Youtie, J. Tracking the emergence of synthetic biology. *Scientometrics* **2017**, *112*, 1439–1469.

2. Jagadevan, S.; Banerjee, A.; Banerjee, C.; Guria, C.; Tiwari, R.; Baweja, M.; Shukla, P. Recent developments in synthetic biology and metabolic engineering in microalgae towards biofuel production. *Biotechnol Biofuels* **2018**, *11*, 185.

3. Konig, H.; Frank, D.; Heil, R.; Coenen, C. Synthetic genomics and synthetic biology applications between hopes and concerns. *Curr Genomics* **2013**, *14*, 11–24.

4. Adamala, K. P.; Martin-Alarcon, D. A.; Boyden, E. S. Programmable RNA-binding protein composed of repeats of a single modular unit. *Proc Natl Acad Sci U S A* **2016**, *113*, E2579–88.

5. Choudhury, R.; Tsai, Y. S.; Dominguez, D.; Wang, Y.; Wang, Z. Engineering RNA endonucleases with customized sequence specificities. *Nat Commun* **2012**, *3*, 1147.

6. Cong, L.; Ran, F. A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P. D.; Wu, X.; Jiang, W.; Marraffini, L. A.; Zhang, F. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819–23.

7. Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology* **2013**, *31*, 233.

8. Wiedenheft, B.; Sternberg, S. H.; Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **2012**, *482*, 331.

9. Abudayyeh, O. O.; Gootenberg, J. S.; Konermann, S.; Joung, J.; Slaymaker, I. M.; Cox, D. B. T.; Shmakov, S.; Makarova, K. S.; Semenova, E.; Minakhin, L.; Severinov, K.; Regev, A.; Lander, E. S.; Koonin, E. V.; Zhang, F. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **2016**, *353*, aaf5573.

10. Schaefke, B.; Sun, W.; Li, Y. S.; Fang, L.; Chen, W. The evolution of posttranscriptional regulation. *Wiley Interdiscip Rev RNA* **2018**, e1485.

11. Anderson, P.; Kedersha, N. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 430.

12. Lee, Y.; Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual Review of Biochemistry* **2015**, *84*, 291–323.

13. Chen, M.; Manley, J. L. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 741.

14. Filipowicz, W.; Bhattacharyya, S. N.; Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev. Genet.* **2008**, *9*, 102–14.

15. Xiao, W.; Adhikari, S.; Dahal, U.; Chen, Y. S.; Hao, Y. J.; Sun, B. F.; Sun, H. Y.; Li, A.; Ping, X. L.; Lai, W. Y.; Wang, X.; Ma, H. L.; Huang, C. M.; Yang, Y.; Huang, N.; Jiang, G. B.; Wang, H. L.; Zhou, Q.; Wang, X. J.; Zhao, Y. L.; Yang, Y. G. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell* **2016**, *61*, 507–519.

16. Meyer, K.; Patil, D.; Zhou, J.; Zinoviev, A.; Skabkin, M.; Elemento, O.; Pestova, T.; Qian, S.-B.; Jaffrey, S. 5′ UTR m6A Promotes Cap-Independent Translation. *Cell* **2015**, *163*, 999–1010.

17. Wang, X.; Zhao, B.; Roundtree, I.; Lu, Z.; Han, D.; Ma, H.; Weng, X.; Chen, K.; Shi, H.; He, C. N6-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **2015**, *161*, 1388–1399.

18. Wang, X.; He, C. Dynamic RNA modifications in posttranscriptional regulation. *Mol Cell* **2014**, *56*, 5–12.

19. Kleinman, C. L.; Adoue, V.; Majewski, J. RNA editing of protein sequences: a rare event in human transcriptomes. *RNA* **2012**, *18*, 1586–96.

20. Lafontaine, D. L. Noncoding RNAs in eukaryotic ribosome biogenesis and function. *Nat Struct Mol Biol* **2015**, *22*, 11–9.

21. Kirchner, S.; Ignatova, Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet* **2015**, *16*, 98–112.

22. Zhao, B. S.; Roundtree, I. A.; He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **2017**, *18*, 31–42.

23. du Toit, A. RNA: Expanding the mRNA epitranscriptome. *Nat Rev Mol Cell Biol* **2016**, *17*, 201.

24. Wang, M.; Zuris, J. A.; Meng, F.; Rees, H.; Sun, S.; Deng, P.; Han, Y.; Gao, X.; Pouli, D.; Wu, Q.; Georgakoudi, I.; Liu, D. R.; Xu, Q. Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proc Natl Acad Sci U S A* **2016**, *113*, 2868–73.

25. Dominissini, D.; Moshitch-Moshkovitz, S.; Schwartz, S.; Salmon-Divon, M.; Ungar, L.; Osenberg, S.; Cesarkas, K.; Jacob-Hirsch, J.; Amariglio, N.; Kupiec, M.; Sorek, R.; Rechavi, G. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **2012**, *485*, 201.

26. He, C. Grand challenge commentary: RNA epigenetics? *Nat Chem Biol* **2010**, *6*, 863–5.

27. Edupuganti, R. R.; Geiger, S.; Lindeboom, R. G. H.; Shi, H.; Hsu, P. J.; Lu, Z.; Wang, S. Y.; Baltissen, M. P. A.; Jansen, P.; Rossa, M.; Muller, M.; Stunnenberg, H. G.; He, C.; Carell, T.; Vermeulen, M. N(6)-methyladenosine (m(6)A) recruits and repels proteins to regulate mRNA homeostasis. *Nat Struct Mol Biol* **2017**, *24*, 870–878.

28. Du, H.; Zhao, Y.; He, J.; Zhang, Y.; Xi, H.; Liu, M.; Ma, J.; Wu, L. YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. *Nat Commun.* **2016**, *7*, 12626.

29. Park, O. H.; Ha, H.; Lee, Y.; Boo, S. H.; Kwon, D. H.; Song, H. K.; Kim, Y. K. Endoribonucleolytic Cleavage of m(6)A-Containing RNAs by RNase P/MRP Complex. *Mol Cell* **2019**, *74*, 494–507 e8.

30. Bazak, L.; Haviv, A.; Barak, M.; Jacob-Hirsch, J.; Deng, P.; Zhang, R.; Isaacs, F. J.; Rechavi, G.; Li, J. B.; Eisenberg, E.; Levanon, E. Y. A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res* **2014**, *24*, 365–76.

31. Fossat, N.; Tourle, K.; Radziewic, T.; Barratt, K.; Liebhold, D.; Studdert, J. B.; Power, M.; Jones, V.; Loebel, D. A.; Tam, P. P. C to U RNA editing mediated by APOBEC1 requires RNA-binding protein RBM47. *EMBO Rep* **2014**, *15*, 903–10.

32. 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.; Consortium, G. T.; Laboratory, D. A.; Coordinating Center Analysis Working, G.; Statistical Methods groups Analysis Working, G.; Enhancing, G. g.; Fund, N. I. H. C.; Nih/Nci,; Nih/Nhgri,; Nih/Nimh,; Nih/Nida,; Biospecimen Collection Source Site, N.; Biospecimen Collection Source Site, R.; Biospecimen Core Resource, V.; Brain Bank Repository-University of Miami Brain Endowment, B.; Leidos Biomedical-Project, M.; Study, E.; Genome Browser Data, I.; Visualization, E. B. I.; Genome Browser Data, I.; Visualization-Ucsc Genomics Institute, U. o. C. S. C.; 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. Dynamic landscape and regulation of RNA editing in mammals. *Nature* **2017**, *550*, 249–254.

33. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* **2010**, *79*, 321–49.

34. Bass, B. L.; Weintraub, H. An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* **1988**, *55*, 1089–98.

35. Matthews, M. M.; Thomas, J. M.; Zheng, Y.; Tran, K.; Phelps, K. J.; Scott, A. I.; Havel, J.; Fisher, A. J.; Beal, P. A. Structures of human ADAR2 bound to dsRNA reveal base-flipping mechanism and basis for site selectivity. *Nat Struct Mol Biol* **2016**, *23*, 426–33.

36. Cho, D. S.; Yang, W.; Lee, J. T.; Shiekhattar, R.; Murray, J. M.; Nishikura, K. Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J Biol Chem* **2003**, *278*, 17093–102.

37. Blanc, V.; Davidson, N. O. C-to-U RNA editing: mechanisms leading to genetic diversity. *J Biol Chem* **2003**, *278*, 1395–8.

38. Blanc, V.; Davidson, N. O. APOBEC-1-mediated RNA editing. *Wiley Interdiscip Rev Syst Biol Med* **2010**, *2*, 594–602.

39. Laurencikiene, J.; Kallman, A. M.; Fong, N.; Bentley, D. L.; Ohman, M. RNA editing and alternative splicing: the importance of co-transcriptional coordination. *EMBO Rep* **2006**, *7*, 303–7.

40. Luciano, D. J.; Mirsky, H.; Vendetti, N. J.; Maas, S. RNA editing of a miRNA precursor. *RNA* **2004**, *10*, 1174–7.

41. Gelinas, J. F.; Clerzius, G.; Shaw, E.; Gatignol, A. Enhancement of replication of RNA viruses by ADAR1 via RNA editing and inhibition of RNA-activated protein kinase. *J Virol* **2011**, *85*, 8460–6.

42. Milewska, A.; Kindler, E.; Vkovski, P.; Zeglen, S.; Ochman, M.; Thiel, V.; Rajfur, Z.; Pyrc, K. APOBEC3-mediated restriction of RNA virus replication. *Sci Rep* **2018**, *8*, 5960.

43. Kung, C. P.; Maggi, J., L. B.; Weber, J. D. The Role of RNA Editing in Cancer Development and Metabolic Disorders. *Front Endocrinol (Lausanne)* **2018**, *9*, 762.

44. Pestova, T. V.; Kolupaeva, V. G.; Lomakin, I. B.; Pilipenko, E. V.; Shatsky, I. N.; Agol, V. I.; Hellen, C. U. Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci U S A* **2001**, *98*, 7029–36.

45. Jackson, R. J.; Hellen, C. U.; Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **2010**, *11*, 113–27.

46. Sonenberg, N.; Hinnebusch, A. G. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **2009**, *136*, 731–45.

47. Kumar, P.; Hellen, C. U.; Pestova, T. V. Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev* **2016**, *30*, 1573–88.

48. Kapp, L. D.; Lorsch, J. R. The molecular mechanics of eukaryotic translation. *Annu Rev Biochem* **2004**, *73*, 657–704.

49. Coller, J.; Wickens, M. Tethered function assays: an adaptable approach to study RNA regulatory proteins. *Methods Enzymol* **2007**, *429*, 299–321.

50. Bos, T. J.; Nussbacher, J. K.; Aigner, S.; Yeo, G. W. Tethered Function Assays as Tools to Elucidate the Molecular Roles of RNA-Binding Proteins. *Adv Exp Med Biol* **2016**, *907*, 61–88.

51. Desjarlais, J. R.; Berg, J. M. Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA binding proteins. *Proceedings of the National Academy of Sciences* **1993**, *90*, 2256.

52. Hockemeyer, D.; Wang, H.; Kiani, S.; Lai, C. S.; Gao, Q.; Cassady, J. P.; Cost, G. J.; Zhang, L.; Santiago, Y.; Miller, J. C.; Zeitler, B.; Cherone, J. M.; Meng, X.; Hinkley, S. J.; Rebar, E. J.; Gregory, P. D.; Urnov, F. D.; Jaenisch, R. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **2011**, *29*, 731–4.

53. Joung, J. K.; Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nature Reviews Molecular Cell Biology* **2012**, *14*, 49.

54. Schierling, B.; Dannemann, N.; Gabsalilow, L.; Wende, W.; Cathomen, T.; Pingoud, A. A novel zinc-finger nuclease platform with a sequence-specific cleavage module. *Nucleic Acids Res* **2012**, *40*, 2623–38.

55. Batra, R.; Nelles, D. A.; Pirie, E.; Blue, S. M.; Marina, R. J.; Wang, H.; Chaim, I. A.; Thomas, J. D.; Zhang, N.; Nguyen, V.; Aigner, S.; Markmiller, S.; Xia, G.; Corbett, K. D.; Swanson, M. S.; Yeo, G. W. Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. *Cell* **2017**, *170*, 899–912 e10.

56. Nelles, D. A.; Fang, M. Y.; O'Connell, M. R.; Xu, J. L.; Markmiller, S. J.; Doudna, J. A.; Yeo, G. W. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **2016**, *165*, 488–96.

57. O'Connell, M. R.; Oakes, B. L.; Sternberg, S. H.; East-Seletsky, A.; Kaplan, M.; Doudna, J. A. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **2014**, *516*, 263.

58. Yan, W. X.; Hunnewell, P.; Alfonse, L. E.; Carte, J. M.; Keston-Smith, E.; Sothiselvam, S.; Garrity, A. J.; Chong, S.; Makarova, K. S.; Koonin, E. V.; Cheng, D. R.; Scott, D. A. Functionally diverse type V CRISPR-Cas systems. *Science* **2019**, *363*, 88–91.

59. Abudayyeh, O. O.; Gootenberg, J. S.; Essletzbichler, P.; Han, S.; Joung, J.; Belanto, J. J.; Verdine, V.; Cox, D. B. T.; Kellner, M. J.; Regev, A.; Lander, E. S.; Voytas, D. F.; Ting, A. Y.; Zhang, F. RNA targeting with CRISPR–Cas13. *Nature* **2017**, *550*, 280.

60. Cox, D. B. T.; Gootenberg, J. S.; Abudayyeh, O. O.; Franklin, B.; Kellner, M. J.; Joung, J.; Zhang, F. RNA editing with CRISPR-Cas13. *Science* **2017**, *358*, 1019–1027.

61. 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. A cytosine deaminase for programmable single-base RNA editing. *Science* **2019**, *365*, 382–386.

62. Konermann, S.; Lotfy, P.; Brideau, N. J.; Oki, J.; Shokhirev, M. N.; Hsu, P. D. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell* **2018**, *173*, 665–676.e14.

63. Rauch, S.; He, C.; Dickinson, B. C. Targeted m(6)A Reader Proteins To Study Epitranscriptomic Regulation of Single RNAs. *J Am Chem Soc* **2018**, *140*, 11974–11981.

64. Kim, D.; Rossi, J. RNAi mechanisms and applications. *Biotechniques* **2008**, *44*, 613–6.

65. Wilson, R. C.; Doudna, J. A. Molecular mechanisms of RNA interference. *Annu Rev Biophys* **2013**, *42*, 217–39.

66. Agrawal, N.; Dasaradhi, P. V.; Mohmmed, A.; Malhotra, P.; Bhatnagar, R. K.; Mukherjee, S. K. RNA interference: biology, mechanism, and applications. *Microbiol Mol Biol Rev* **2003**, *67*, 657–85.

67. Titze-de Almeida, R.; David, C.; Titze-de Almeida, S. S. The Race of 10 Synthetic RNAi-Based Drugs to the Pharmaceutical Market. *Pharm Res* **2017**, *34*, 1339–1363.

68. Aagaard, L.; Rossi, J. J. RNAi therapeutics: principles, prospects and challenges. *Adv Drug Deliv Rev* **2007**, *59*, 75–86.

69. Zucchelli, S.; Cotella, D.; Takahashi, H.; Carrieri, C.; Cimatti, L.; Fasolo, F.; Jones, M. H.; Sblattero, D.; Sanges, R.; Santoro, C.; Persichetti, F.; Carninci, P.; Gustincich, S. SINEUPs: A new class of natural and synthetic antisense long non-coding RNAs that activate translation. *RNA Biol* **2015**, *12*, 771–9.

70. Zucchelli, S.; Fasolo, F.; Russo, R.; Cimatti, L.; Patrucco, L.; Takahashi, H.; Jones, M. H.; Santoro, C.; Sblattero, D.; Cotella, D.; Persichetti, F.; Carninci, P.; Gustincich, S. SINEUPs are modular antisense long non-coding RNAs that increase synthesis of target proteins in cells. *Front Cell Neurosci* **2015**, *9*, 174.

71. Espinoza, S.; Scarpato, M.; Damiani, D.; Manago, F.; Mereu, M.; Contestabile, A.; Peruzzo, O.; Carninci, P.; Santoro, C.; Papaleo, F.; Mingozzi, F.; Ronzitti, G.; Zucchelli, S.; Gustincich, S. SINEUP Non-coding RNA Targeting GDNF Rescues Motor Deficits and Neurodegeneration in a Mouse Model of Parkinson's Disease. *Mol Ther* **2019**,

72. Merkle, T.; Merz, S.; Reautschnig, P.; Blaha, A.; Li, Q.; Vogel, P.; Wettengel, J.; Li, J. B.; Stafforst, T. Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nat Biotechnol* **2019**, *37*, 133–138.

73. Qu, L.; Yi, Z.; Zhu, S.; Wang, C.; Cao, Z.; Zhou, Z.; Yuan, P.; Yu, Y.; Tian, F.; Liu, Z.; Bao, Y.; Zhao, Y.; Wei, W. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nat Biotechnol* **2019**,

74. Dalby, B.; Cates, S.; Harris, A.; Ohki, E. C.; Tilkins, M. L.; Price, P. J.; Ciccarone, V. C. Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods* **2004**, *33*, 95–103.

75. Cardarelli, F.; Digiacomo, L.; Marchini, C.; Amici, A.; Salomone, F.; Fiume, G.; Rossetta, A.; Gratton, E.; Pozzi, D.; Caracciolo, G. The intracellular trafficking mechanism of Lipofectamine-based transfection reagents and its implication for gene delivery. *Sci Rep* **2016**, *6*, 25879.

76. Potter, H. Transfection by electroporation. *Curr Protoc Mol Biol* **2003**, *Chapter 9*, Unit 9 3.

77. Manjunath, N.; Wu, H.; Subramanya, S.; Shankar, P. Lentiviral delivery of short hairpin RNAs. *Adv Drug Deliv Rev* **2009**, *61*, 732–45.

78. Yang, L.; Bailey, L.; Baltimore, D.; Wang, P. Targeting lentiviral vectors to specific cell types in vivo. *Proc Natl Acad Sci U S A* **2006**, *103*, 11479–84.

79. Gao, G.; Vandenberghe, L. H.; Wilson, J. M. New recombinant serotypes of AAV vectors. *Curr Gene Ther* **2005**, *5*, 285–97.

80. Naso, M. F.; Tomkowicz, B.; Perry, r., W. L.; Strohl, W. R. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs* **2017**, *31*, 317–334.

81. Pillay, S.; Zou, W.; Cheng, F.; Puschnik, A. S.; Meyer, N. L.; Ganaie, S. S.; Deng, X.; Wosen, J. E.; Davulcu, O.; Yan, Z.; Engelhardt, J. F.; Brown, K. E.; Chapman, M. S.; Qiu, J.; Carette, J. E. Adeno-associated Virus (AAV) Serotypes Have Distinctive Interactions with Domains of the Cellular AAV Receptor. *J Virol* **2017**, *91*.

82. Zincarelli, C.; Soltys, S.; Rengo, G.; Rabinowitz, J. E. Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol Ther* **2008**, *16*, 1073–80.

83. Wu, Z.; Yang, H.; Colosi, P. Effect of genome size on AAV vector packaging. *Mol Ther* **2010**, *18*, 80–6.

84. Bruno, B. J.; Miller, G. D.; Lim, C. S. Basics and recent advances in peptide and protein drug delivery. *Ther Deliv* **2013**, *4*, 1443–67.

85. Thompson, D. B.; Cronican, J. J.; Liu, D. R. Engineering and identifying supercharged proteins for macromolecule delivery into mammalian cells. *Methods Enzymol* **2012**, *503*, 293–319.

86. Gao, X.; Tao, Y.; Lamas, V.; Huang, M.; Yeh, W. H.; Pan, B.; Hu, Y. J.; Hu, J. H.; Thompson, D. B.; Shu, Y.; Li, Y.; Wang, H.; Yang, S.; Xu, Q.; Polley, D. B.; Liberman, M. C.; Kong, W. J.; Holt, J. R.; Chen, Z. Y.; Liu, D. R. Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* **2018**, *553*, 217–221.

87. Jyotsana, N.; Sharma, A.; Chaturvedi, A.; Budida, R.; Scherr, M.; Kuchenbauer, F.; Lindner, R.; Noyan, F.; Suhs, K. W.; Stangel, M.; Grote-Koska, D.; Brand, K.; Vornlocher, H. P.; Eder, M.; Thol, F.; Ganser, A.; Humphries, R. K.; Ramsay, E.; Cullis, P.; Heuser, M. Lipid nanoparticle-mediated siRNA delivery for safe targeting of human CML in vivo. *Ann Hematol* **2019**, *98*, 1905–1918.

88. Finn, J. D.; Smith, A. R.; Patel, M. C.; Shaw, L.; Youniss, M. R.; van Heteren, J.; Dirstine, T.; Ciullo, C.; Lescarbeau, R.; Seitzer, J.; Shah, R. R.; Shah, A.; Ling, D.; Growe, J.; Pink, M.; Rohde, E.; Wood, K. M.; Salomon, W. E.; Harrington, W. F.; Dombrowski, C.; Strapps, W. R.; Chang, Y.; Morrissey, D. V. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Rep* **2018**, *22*, 2227–2235.

89. Kulkarni, J. A.; Witzigmann, D.; Chen, S.; Cullis, P. R.; van der Meel, R. Lipid Nanoparticle Technology for Clinical Translation of siRNA Therapeutics. *Acc Chem Res* **2019**, *52*, 2435–2444.

90. Corey, D. R. Chemical modification: the key to clinical application of RNA interference? *J Clin Invest* **2007**, *117*, 3615–22.

91. Hassler, M. R.; Turanov, A. A.; Alterman, J. F.; Haraszti, R. A.; Coles, A. H.; Osborn, M. F.; Echeverria, D.; Nikan, M.; Salomon, W. E.; Roux, L.; Godinho, B.; Davis, S. M.; Morrissey, D. V.; Zamore, P. D.; Karumanchi, S. A.; Moore, M. J.; Aronin, N.; Khvorova, A. Comparison of partially and fully chemically-modified siRNA in conjugate-mediated delivery in vivo. *Nucleic Acids Res* **2018**, *46*, 2185–2196.

92. Lee, Y. W.; Luther, D. C.; Kretzmann, J. A.; Burden, A.; Jeon, T.; Zhai, S.; Rotello, V. M. Protein Delivery into the Cell Cytosol using Non-Viral Nanocarriers. *Theranostics* **2019**, *9*, 3280–3292.

93. Yang, L. Z.; Wang, Y.; Li, S. Q.; Yao, R. W.; Luan, P. F.; Wu, H.; Carmichael, G. G.; Chen, L. L. Dynamic Imaging of RNA in Living Cells by CRISPR-Cas13 Systems. *Mol Cell* **2019**,

94. Charlesworth, C. T.; Deshpande, P. S.; Dever, D. P.; Camarena, J.; Lemgart, V. T.; Cromer, M. K.; Vakulskas, C. A.; Collingwood, M. A.; Zhang, L.; Bode, N. M.; Behlke, M. A.; Dejene, B.; Cieniewicz, B.; Romano, R.; Lesch, B. J.; Gomez-Ospina, N.; Mantri, S.; Pavel-Dinu, M.; Weinberg, K. I.; Porteus, M. H. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nature Medicine* **2019**, *25*, 249–254.

95. Simhadri, V. L.; McGill, J.; McMahon, S.; Wang, J.; Jiang, H.; Sauna, Z. E. Prevalence of Pre-existing Antibodies to CRISPR-Associated Nuclease Cas9 in the USA Population. *Mol Ther Methods Clin Dev* **2018**, *10*, 105–112.

96. Wagner, D. L.; Amini, L.; Wendering, D. J.; Burkhardt, L.-M.; Akyüz, L.; Reinke, P.; Volk, H.-D.; Schmueck-Henneresse, M. High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nature Medicine* **2018**,

97. Lim, C.; Allada, R. Emerging roles for post-transcriptional regulation in circadian clocks. *Nature Neuroscience* **2013**, *16*, 1544.

98. Roundtree, I. A.; Evans, M. E.; Pan, T.; He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **2017**, *169*, 1187–1200.

99. Glisovic, T.; Bachorik Jennifer, L.; Yong, J.; Dreyfuss, G. RNAbinding proteins and posttranscriptional gene regulation. *FEBS Letters* **2008**, *582*, 1977–1986.

100. Battle, A.; Khan, Z.; Wang, S. H.; Mitrano, A.; Ford, M. J.; Pritchard, J. K.; Gilad, Y. Impact of regulatory variation from RNA to protein. *Science* **2015**, *347*, 664.

101. Maier, T.; Güell, M.; Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Letters* **2009**, *583*, 3966–3973.

102. Ivanova, I.; Much, C.; Di Giacomo, M.; Azzi, C.; Morgan, M.; Moreira, P. N.; Monahan, J.; Carrieri, C.; Enright, A. J.; O'Carroll, D. The RNA m6A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the Maternal Transcriptome and Oocyte Competence. *Molecular Cell* **2017**, *67*, 1059–1067.e4.

103. Aguilo, F.; Zhang, F.; Sancho, A.; Fidalgo, M.; Di Cecilia, S.; Vashisht, A.; Lee, D.-F.; Chen, C.-H.; Rengasamy, M.; Andino, B.; Jahouh, F.; Roman, A.; Krig, S. R.; Wang, R.; Zhang, W.; Wohlschlegel, J. A.; Wang, J.; Walsh, M. J. Coordination of m6A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming. *Cell Stem Cell* **2015**, *17*, 689–704.

104. Wen, J.; Lv, R.; Ma, H.; Shen, H.; He, C.; Wang, J.; Jiao, F.; Liu, H.; Yang, P.; Tan, L.; Lan, F.; Shi, Y. G.; He, C.; Shi, Y.; Diao, J. Zc3h13 Regulates Nuclear RNA m6A Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Molecular Cell* **2018**, *69*, 1028–1038.e6.

105. Tirumuru, N.; Zhao, B. S.; Lu, W.; Lu, Z.; He, C.; Wu, L. N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *eLife* **2016**, *5*, e15528.

106. Lichinchi, G.; Gao, S.; Saletore, Y.; Gonzalez, G. M.; Bansal, V.; Wang, Y.; Mason, C. E.; Rana, T. M. Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat Microbiol* **2016**, *1*, 16011.

107. Lichinchi, G.; Zhao, B. S.; Wu, Y.; Lu, Z.; Qin, Y.; He, C.; Rana, T. M. Dynamics of Human and Viral RNA Methylation during Zika Virus Infection. *Cell Host & Microbe* **2016**, *20*, 666–673.

108. Li, Z.; Weng, H.; Su, R.; Weng, X.; Zuo, Z.; Li, C.; Huang, H.; Nachtergaele, S.; Dong, L.; Hu, C.; Qin, X.; Tang, L.; Wang, Y.; Hong, G.-M.; Huang, H.; Wang, X.; Chen, P.; Gurbuxani, S.; Arnovitz, S.; Li, Y.; Li, S.; Strong, J.; Neilly, M. B.; Larson, R. A.; Jiang, X.; Zhang, P.; Jin, J.; He, C.; Chen, J. FTO Plays an Oncogenic Role in Acute Myeloid Leukemia as a N6-Methyladenosine RNA Demethylase. *Cancer Cell* **2017**, *31*, 127–141.

109. Cui, Q.; Shi, H.; Ye, P.; Li, L.; Qu, Q.; Sun, G.; Sun, G.; Lu, Z.; Huang, Y.; Yang, C.-G.; Riggs, A. D.; He, C.; Shi, Y. m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Reports* **2017**, *18*, 2622–2634.

110. Zhang, S.; Zhao, B. S.; Zhou, A.; Lin, K.; Zheng, S.; Lu, Z.; Chen, Y.; Sulman, E. P.; Xie, K.; Bögler, O.; Majumder, S.; He, C.; Huang, S. m6A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. *Cancer Cell* **2017**, *31*, 591–606.e6.

111. Lin, S.; Choe, J.; Du, P.; Triboulet, R.; Gregory, R. The m6A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Molecular Cell* **2016**, *62*, 335–345.

112. Wang, X.; Lu, Z.; Gomez, A.; Hon, G. C.; Yue, Y.; Han, D.; Fu, Y.; Parisien, M.; Dai, Q.; Jia, G.; Ren, B.; Pan, T.; He, C. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **2013**, *505*, 117.

113. Kennedy, E.; Bogerd, H.; Kornepati, A. R.; Kang, D.; Ghoshal, D.; Marshall, J.; Poling, B.; Tsai, K.; Gokhale, N.; Horner, S.; Cullen, B. Posttranscriptional m6A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host & Microbe* **2016**, *19*, 675–685.

114. Gokhale, N. S.; McIntyre, A. B. R.; McFadden, M. J.; Roder, A. E.; Kennedy, E. M.; Gandara, J. A.; Hopcraft, S. E.; Quicke, K. M.; Vazquez, C.; Willer, J.; Ilkayeva, O. R.; Law, B. A.; Holley, C. L.; Garcia-Blanco, M. A.; Evans, M. J.; Suthar, M. S.; Bradrick, S. S.; Mason, C. E.; Horner, S. M. N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. *Cell Host Microbe* **2016**, *20*, 654–665.

115. Wang, S.; Sun, C.; Li, J.; Zhang, E.; Ma, Z.; Xu, W.; Li, H.; Qiu, M.; Xu, Y.; Xia, W.; Xu, L.; Yin, R. Roles of RNA methylation by means of N6-methyladenosine (m6A) in human cancers. *Cancer Letters* **2017**, *408*, 112–120.

116. Patil, D. P.; Chen, C.-K.; Pickering, B. F.; Chow, A.; Jackson, C.; Guttman, M.; Jaffrey, S. R. m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **2016**, *537*, 369.

117. Huang, H.; Weng, H.; Sun, W.; Qin, X.; Shi, H.; Wu, H.; Zhao, B. S.; Mesquita, A.; Liu, C.; Yuan, C. L.; Hu, Y.-C.; Hüttelmaier, S.; Skibbe, J. R.; Su, R.; Deng, X.; Dong, L.; Sun, M.; Li, C.; Nachtergaele, S.; Wang, Y.; Hu, C.; Ferchen, K.; Greis, K. D.; Jiang, X.; Wei, M.; Qu, L.; Guan, J.-L.; He, C.; Yang, J.; Chen, J. Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nature Cell Biology* **2018**, *20*, 285–295.

118. Hilton, I. B.; D'Ippolito, A. M.; Vockley, C. M.; Thakore, P. I.; Crawford, G. E.; Reddy, T. E.; Gersbach, C. A. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nature Biotechnology* **2015**, *33*, 510.

119. Kearns, N. A.; Pham, H.; Tabak, B.; Genga, R. M.; Silverstein, N. J.; Garber, M.; Maehr, R. Functional annotation of native enhancers with a Cas9–histone demethylase fusion. *Nature Methods* **2015**, *12*, 401.

120. Linder, B.; Grozhik, A. V.; Olarerin-George, A. O.; Meydan, C.; Mason, C. E.; Jaffrey, S. R. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods* **2015**, *12*, 767.

121. Schwartz, S.; Mumbach, M.; Jovanovic, M.; Wang, T.; Maciag, K.; Bushkin, G. Mertins, P.; Ter-Ovanesyan, D.; Habib, N.; Cacchiarelli, D.; Sanjana, N.; Freinkman, E.; Pacold, M.; Satija, R.; Mikkelsen, T.; Hacohen, N.; Zhang, F.; Carr, S.; Lander, E.; Regev, A. Perturbation of m6A Writers Reveals Two Distinct Classes of mRNA Methylation at Internal and 5′ Sites. *Cell Reports* **2014**, *8*, 284–296.

122. Lahens, N. F.; Kavakli, I. H.; Zhang, R.; Hayer, K.; Black, M. B.; Dueck, H.; Pizarro, A.; Kim, J.; Irizarry, R.; Thomas, R. S.; Grant, G. R.; Hogenesch, J. B. IVT-seq reveals extreme bias in RNA sequencing. *Genome Biology* **2014**, *15*, R86.

123. Yue, Y.; Liu, J.; Cui, X.; Cao, J.; Luo, G.; Zhang, Z.; Cheng, T.; Gao, M.; Shu, X.; Ma, H.; Wang, F.; Wang, X.; Shen, B.; Wang, Y.; Feng, X.; He, C.; Liu, J. VIRMA mediates preferential m6A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. *Cell Discovery* **2018**, *4*, 10.

124. Rauch, S.; Dickinson, B. C. Programmable RNA Binding Proteins for Imaging and Therapeutics. *Biochemistry* **2018**, *57*, 363–364.

125. Bray, N. L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **2016**, *34*, 525.

126. Pimentel, H.; Bray, N. L.; Puente, S.; Melsted, P.; Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods* **2017**, *14*, 687.

127. Chandrasegaran, S.; Carroll, D. Origins of Programmable Nucleases for Genome Engineering. *Journal of Molecular Biology* **2016**, *428*, 963–989.

128. Kim, H.; Kim, J.-S. A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics* **2014**, *15*, 321.

129. O'Connell, M. R. Molecular Mechanisms of RNA Targeting by Cas13-containing Type VI CRISPR-Cas Systems. *J Mol Biol* **2019**, *431*, 66–87.

130. Terns, M. P. CRISPR-Based Technologies: Impact of RNA-Targeting Systems. *Mol Cell* **2018**, *72*, 404–412.

131. Liao, H.-K.; Hatanaka, F.; Araoka, T.; Reddy, P.; Wu, M.-Z.; Sui, Y.; Yamauchi, T.; Sakurai, M.; O'Keefe, D. D.; Núñez-Delicado, E.; Guillen, P.; Campistol, J. M.; Wu, C.-J.; Lu, L.-F.; Esteban, C. R.; Izpisua Belmonte, J. C. In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell* **2017**, *171*, 1495–1507.e15.

132. Monteys, A. M.; Ebanks, S. A.; Keiser, M. S.; Davidson, B. L. CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. *Molecular Therapy* **2017**, *25*, 12–23.

133. Fuxman Bass, J. I.; Sahni, N.; Shrestha, S.; Garcia-Gonzalez, A.; Mori, A.; Bhat, N.; Yi, S.; Hill, D. E.; Vidal, M.; Walhout, A. J. M. Human gene-centered transcription factor networks for enhancers and disease variants. *Cell* **2015**, *161*, 661–673.

134. Granados-Riveron, J. T.; Aquino-Jarquin, G. CRISPR-Cas13 Precision Transcriptome Engineering in Cancer. *Cancer Res* **2018**, *78*, 4107–4113.

135. Sundaram, G. M.; Ismail, H. M.; Bashir, M.; Muhuri, M.; Vaz, C.; Nama, S.; Ow, G. S.; Vladimirovna, I. A.; Ramalingam, R.; Burke, B.; Tanavde, V.; Kuznetsov, V.; Lane, E. B.; Sampath, P. EGF hijacks miR-198/FSTL1 wound-healing switch and steers a two-pronged pathway toward metastasis. *J Exp Med* **2017**, *214*, 2889–2900.

136. Du, D.; Roguev, A.; Gordon, D. E.; Chen, M.; Chen, S.-H.; Shales, M.; Shen, J. P.; Ideker, T.; Mali, P.; Qi, L. S.; Krogan, N. J. Genetic interaction mapping in mammalian cells using CRISPR interference. *Nature Methods* **2017**, *14*, 577.

137. Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **2013**, *152*, 1173–83.

138. Zhao, B. S.; Roundtree, I. A.; He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **2017**, *18*, 31–42.

139. Abil, Z.; Denard, C. A.; Zhao, H. Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. *J Biol Eng* **2014**, *8*, 7.

140. Yan, W. X.; Chong, S.; Zhang, H.; Makarova, K. S.; Koonin, E. V.; Cheng, D. R.; Scott, D. A. Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. *Mol Cell* **2018**, *70*, 327–339 e5.

141. Rutkauskas, M.; Sinkunas, T.; Songailiene, I.; Tikhomirova, M.; Siksnys, V.; Seidel, R. Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Reports* **2015**, *10*, 1534–1543.

142. Sternberg, S. H.; Redding, S.; Jinek, M.; Greene, E. C.; Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **2014**, *507*, 62.

143. Szczelkun, M. D.; Tikhomirova, M. S.; Sinkunas, T.; Gasiunas, G.; Karvelis, T.; Pschera, P.; Siksnys, V.; Seidel, R. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences* **2014**, *111*, 9798.

144. Liu, L.; Li, X.; Ma, J.; Li, Z.; You, L.; Wang, J.; Wang, M.; Zhang, X.; Wang, Y. The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell* **2017**, *170*, 714–726 e10.

145. Slaymaker, I. M.; Mesa, P.; Kellner, M. J.; Kannan, S.; Brignole, E.; Koob, J.; Feliciano, P. R.; Stella, S.; Abudayyeh, O. O.; Gootenberg, J. S.; Strecker, J.; Montoya, G.; Zhang, F. High-Resolution Structure of Cas13b and Biochemical Characterization of RNA Targeting and Cleavage. *Cell Rep* **2019**, *26*, 3741–3751 e5.

146. Tambe, A.; East-Seletsky, A.; Knott, G. J.; Doudna, J. A.; O'Connell, M. R. RNA Binding and HEPN-Nuclease Activation Are Decoupled in CRISPR-Cas13a. *Cell Rep* **2018**, *24*, 1025–1036.

147. Smargon, A. A.; Cox, D. B. T.; Pyzocha, N. K.; Zheng, K.; Slaymaker, I. M.; Gootenberg, J. S.; Abudayyeh, O. A.; Essletzbichler, P.; Shmakov, S.; Makarova, K. S.; Koonin, E. V.; Zhang, F. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol Cell* **2017**, *65*, 618–630 e7.

148. Montiel-Gonzalez, M. F.; Vallecillo-Viejo, I. C.; Rosenthal, J. J. An efficient system for selectively altering genetic information within mRNAs. *Nucleic Acids Res* **2016**, *44*, e157.

149. Sinnamon, J. R.; Kim, S. Y.; Corson, G. M.; Song, Z.; Nakai, H.; Adelman, J. P.; Mandel, G. Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proc Natl Acad Sci U S A* **2017**, *114*, E9395–E9402.

150. Vogel, P.; Moschref, M.; Li, Q.; Merkle, T.; Selvasaravanan, K. D.; Li, J. B.; Stafforst, T. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs. *Nat Methods* **2018**, *15*, 535–538.

151. Gootenberg, J. S.; Abudayyeh, O. O.; Kellner, M. J.; Joung, J.; Collins, J. J.; Zhang, F. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **2018**, *360*, 439.

152. Blakeley, B. D.; McNaughton, B. R. Synthetic RNA recognition motifs that selectively recognize HIV-1 trans-activation response element hairpin RNA. *ACS Chem Biol* **2014**, *9*, 1320–9.

153. Crawford, D. W.; Blakeley, B. D.; Chen, P. H.; Sherpa, C.; Le Grice, S. F.; Laird-Offringa, I. A.; McNaughton, B. R. An Evolved RNA Recognition Motif That Suppresses HIV-1 Tat/TAR-Dependent Transcription. *ACS Chem Biol* **2016**, *11*, 2206–15.

154. Zhou, Z.; Dell'Orco, M.; Saldarelli, P.; Turturo, C.; Minafra, A.; Martelli, G. P. Identification of an RNA-silencing suppressor in the genome of Grapevine virus A. *J Gen Virol* **2006**, *87*, 2387–95.

155. Eberle, A. B.; Lykke-Andersen, S.; Muhlemann, O.; Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol* **2009**, *16*, 49–55.

156. Guilinger, J. P.; Thompson, D. B.; Liu, D. R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol* **2014**, *32*, 577–582.

157. Cronican, J. J.; Beier, K. T.; Davis, T. N.; Tseng, J. C.; Li, W.; Thompson, D. B.; Shih, A. F.; May, E. M.; Cepko, C. L.; Kung, A. L.; Zhou, Q.; Liu, D. R. A class of human proteins that deliver functional proteins into mammalian cells in vitro and in vivo. *Chem Biol* **2011**, *18*, 833–8.

158. Kuttan, A.; Bass, B. L. Mechanistic insights into editing-site specificity of ADARs. *Proc Natl Acad Sci U S A* **2012**, *109*, E3295–304.

159. Bao, Z.; Jain, S.; Jaroenpuntaruk, V.; Zhao, H. Orthogonal Genetic Regulation in Human Cells Using Chemically Induced CRISPR/Cas9 Activators. *ACS Synth Biol* **2017**, *6*, 686–693.

160. Gao, Y.; Xiong, X.; Wong, S.; Charles, E. J.; Lim, W. A.; Qi, L. S. Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat Methods* **2016**, *13*, 1043–1049.

161. Vasileva, A.; Jessberger, R. Precise hit: adeno-associated virus in gene targeting. *Nat Rev Microbiol* **2005**, *3*, 837–47.

162. Harrington, L. B.; Burstein, D.; Chen, J. S.; Paez-Espino, D.; Ma, E.; Witte, I. P.; Cofsky, J. C.; Kyrpides, N. C.; Banfield, J. F.; Doudna, J. A. Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science* **2018**, *362*, 839.

163. Glaesner, W.; Vick, A. M.; Millican, R.; Ellis, B.; Tschang, S. H.; Tian, Y.; Bokvist, K.; Brenner, M.; Koester, A.; Porksen, N.; Etgen, G.; Bumol, T. Engineering and characterization of the long-acting glucagon-like peptide-1 analogue LY2189265, an Fc fusion protein. *Diabetes Metab Res Rev* **2010**, *26*, 287–96.

164. Dhople, V.; Krukemeyer, A.; Ramamoorthy, A. The human beta-defensin-3, an antibacterial peptide with multiple biological functions. *Biochim Biophys Acta* **2006**, *1758*, 1499–512.

165. Kluver, E.; Schulz-Maronde, S.; Scheid, S.; Meyer, B.; Forssmann, W. G.; Adermann, K. Structure-activity relation of human beta-defensin 3: influence of disulfide bonds and cysteine substitution on antimicrobial activity and cytotoxicity. *Biochemistry* **2005**, *44*, 9804–16.

166. Hu, J. H.; Miller, S. M.; Geurts, M. H.; Tang, W.; Chen, L.; Sun, N.; Zeina, C. M.; Gao, X.; Rees, H. A.; Lin, Z.; Liu, D. R. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* **2018**, *556*, 57.

167. Dominguez, D.; Freese, P.; Alexis, M. S.; Su, A.; Hochman, M.; Palden, T.; Bazile, C.; Lambert, N. J.; Van Nostrand, E. L.; Pratt, G. A.; Yeo, G. W.; Graveley, B. R.; Burge, C. B. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Molecular Cell* **2018**, *70*, 854–867.e9.

168. Boyle, E. A.; Andreasson, J. O. L.; Chircus, L. M.; Sternberg, S. H.; Wu, M. J.; Guegler, C. K.; Doudna, J. A.; Greenleaf, W. J. High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proc Natl Acad Sci U S A* **2017**, *114*, 5461–5466.

169. Moutaftsi, M.; Peters, B.; Pasquetto, V.; Tscharke, D. C.; Sidney, J.; Bui, H. H.; Grey, H.; Sette, A. A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat Biotechnol* **2006**, *24*, 817–9.

170. Andreatta, M.; Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **2016**, *32*, 511–7.

171. Peters, B.; Sette, A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* **2005**, *6*, 132.

172. Sidney, J.; Assarsson, E.; Moore, C.; Ngo, S.; Pinilla, C.; Sette, A.; Peters, B. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome Res* **2008**, *4*, 2.

173. Jurtz, V.; Paul, S.; Andreatta, M.; Marcatili, P.; Peters, B.; Nielsen, M. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J Immunol* **2017**, *199*, 3360–3368.

174. Calis, J. J.; Maybeno, M.; Greenbaum, J. A.; Weiskopf, D.; De Silva, A. D.; Sette, A.; Kesmir, C.; Peters, B. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* **2013**, *9*, e1003266.

175. Zheng, Y.; Lorenzo, C.; Beal, P. A. DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res* **2017**, *45*, 3369–3377.

176. Yue, Y.; Liu, J.; He, C. RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev* **2015**, *29*, 1343–55.

177. Rauch, S.; He, E.; Srienc, M.; Zhou, H.; Zhang, Z.; Dickinson, B. C. Programmable RNA-Guided RNA Effector Proteins Built from Human Parts. *Cell* **2019**, *178*, 122–134 e12.

178. Chen, T.; Gao, D.; Zhang, R.; Zeng, G.; Yan, H.; Lim, E.; Liang, F. S. Chemically Controlled Epigenome Editing through an Inducible dCas9 System. *J Am Chem Soc* **2017**, *139*, 11337–11340.

179. Liang, F. S.; Ho, W. Q.; Crabtree, G. R. Engineering the ABA plant stress pathway for regulation of induced proximity. *Sci Signal* **2011**, *4*, rs2.

180. Katrekar, D.; Chen, G.; Meluzzi, D.; Ganesh, A.; Worlikar, A.; Shih, Y. R.; Varghese, S.; Mali, P. In vivo RNA editing of point mutations via RNA-guided adenosine deaminases. *Nat Methods* **2019**, *16*, 239–242.

181. Plawski, A.; Banasiewicz, T.; Borun, P.; Kubaszewski, L.; Krokowicz, P.; Skrzypczak-Zielinska, M.; Lubinski, J. Familial adenomatous polyposis of the colon. *Hered Cancer Clin Pract* **2013**, *11*, 15.

182. Liu, F.; Song, Y.; Liu, D. Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA. *Gene Ther* **1999**, *6*, 1258–66.

183. Yang, J.; Chen, S.; Huang, L.; Michalopoulos, G. K.; Liu, Y. Sustained expression of naked plasmid DNA encoding hepatocyte growth factor in mice promotes liver and overall body growth. *Hepatology* **2001**, *33*, 848–59.

184. Banaszynski, L. A.; Liu, C. W.; Wandless, T. J. Characterization of the FKBP.rapamycin.FRB ternary complex. *J Am Chem Soc* **2005**, *127*, 4715–21.

185. Shani, E.; Weinstain, R.; Zhang, Y.; Castillejo, C.; Kaiserli, E.; Chory, J.; Tsien, R. Y.; Estelle, M. Gibberellins accumulate in the elongating endodermal cells of Arabidopsis root. *Proc Natl Acad Sci U S A* **2013**, *110*, 4834–9.

186. Kennedy, M. J.; Hughes, R. M.; Peteya, L. A.; Schwartz, J. W.; Ehlers, M. D.; Tucker, C. L. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat Methods* **2010**, *7*, 973–5.

187. Rehman, K.; Hamid Akash, M. S.; Akhtar, B.; Tariq, M.; Mahmood, A.; Ibrahim, M. Delivery of Therapeutic Proteins: Challenges and Strategies. *Curr Drug Targets* **2016**, *17*, 1172–88.

188. Weng, Z.; DeLisi, C. Protein therapeutics: promises and challenges for the 21st century. *Trends Biotechnol* **2002**, *20*, 29–35.

189. Imataka, H.; Sonenberg, N. Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol Cell Biol* **1997**, *17*, 6940–7.

190. Feoktistova, K.; Tuvshintogs, E.; Do, A.; Fraser, C. S. Human eIF4E promotes mRNA restructuring by stimulating eIF4A helicase activity. *Proc Natl Acad Sci U S A* **2013**, *110*, 13339–44.

191. De Gregorio, E.; Baron, J.; Preiss, T.; Hentze, M. W. Tethered-function analysis reveals that eIF4E can recruit ribosomes independent of its binding to the cap structure. *RNA* **2001**, *7*, 106–13.

192. Whitfield, M. L.; Zheng, L. X.; Baldwin, A.; Ohta, T.; Hurt, M. M.; Marzluff, W. F. Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. *Mol Cell Biol* **2000**, *20*, 4188–98.

193. Collier, B.; Gorgoni, B.; Loveridge, C.; Cooke, H. J.; Gray, N. K. The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells. *EMBO J* **2005**, *24*, 2656–66.

194. Craig, A. W.; Haghighat, A.; Yu, A. T.; Sonenberg, N. Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **1998**, *392*, 520–3.

195. Lu, L.; Li, J.; Moussaoui, M.; Boix, E. Immune Modulation by Human Secreted RNases at the Extracellular Space. *Front Immunol* **2018**, *9*, 1012.

196. Aslam, A.; Mittal, S.; Koch, F.; Andrau, J. C.; Winkler, G. S. The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation. *Mol Biol Cell* **2009**, *20*, 3840–50.

197. Lai, W. S.; Kennington, E. A.; Blackshear, P. J. Tristetraprolin and its family members can promote the cell-free deadenylation of AU-rich element-containing mRNAs by poly(A) ribonuclease. *Mol Cell Biol* **2003**, *23*, 3798–812.

198. Funakoshi, Y.; Doi, Y.; Hosoda, N.; Uchida, N.; Osawa, M.; Shimada, I.; Tsujimoto, M.; Suzuki, T.; Katada, T.; Hoshino, S. Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. *Genes Dev* **2007**, *21*, 3135–48.

199. Houseley, J.; Tollervey, D. The many pathways of RNA degradation. *Cell* **2009**, *136*, 763–76.

200. Peng, R.; Lin, G.; Li, J. Potential pitfalls of CRISPR/Cas9-mediated genome editing. *FEBS J* **2016**, *283*, 1218–31.

201. Lino, C. A.; Harper, J. C.; Carney, J. P.; Timlin, J. A. Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv* **2018**, *25*, 1234–1257.

202. Cebrian-Serrano, A.; Davies, B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm Genome* **2017**, *28*, 247–261.

203. Zhao, M.; Yang, H.; Jiang, X.; Zhou, W.; Zhu, B.; Zeng, Y.; Yao, K.; Ren, C. Lipofectamine RNAiMAX: an efficient siRNA transfection reagent in human embryonic stem cells. *Mol Biotechnol* **2008**, *40*, 19–26.

204. Yu, X.; Liang, X.; Xie, H.; Kumar, S.; Ravinder, N.; Potter, J.; de Mollerat du Jeu, X.; Chesnut, J. D. Improved delivery of Cas9 protein/gRNA complexes using lipofectamine CRISPRMAX. *Biotechnol Lett* **2016**, *38*, 919–29.

205. Sun, N.; Ning, B.; Hansson, K. M.; Bruce, A. C.; Seaman, S. A.; Zhang, C.; Rikard, M.; DeRosa, C. A.; Fraser, C. L.; Wagberg, M.; Fritsche-Danielson, R.; Wikstrom, J.; Chien, K. R.; Lundahl, A.; Holtta, M.; Carlsson, L. G.; Peirce, S. M.; Hu, S. Modified VEGF-A mRNA induces sustained multifaceted microvascular response and accelerates diabetic wound healing. *Sci Rep* **2018**, *8*, 17509.

206. Gan, L. M.; Lagerstrom-Fermer, M.; Carlsson, L. G.; Arfvidsson, C.; Egnell, A. C.; Rudvik, A.; Kjaer, M.; Collen, A.; Thompson, J. D.; Joyal, J.; Chialda, L.; Koernicke, T.; Fuhr, R.; Chien, K. R.; Fritsche-Danielson, R. Intradermal delivery of modified mRNA encoding VEGF-A in patients with type 2 diabetes. *Nat Commun* **2019**, *10*, 871.

207. Tassin, A.; Laoudj-Chenivesse, D.; Vanderplanck, C.; Barro, M.; Charron, S.; Ansseau, E.; Chen, Y. W.; Mercier, J.; Coppee, F.; Belayew, A. DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J Cell Mol Med* **2013**, *17*, 76–89.

208. Cruz, J. M.; Hupper, N.; Wilson, L. S.; Concannon, J. B.; Wang, Y.; Oberhauser, B.; Patora-Komisarska, K.; Zhang, Y.; Glass, D. J.; Trendelenburg, A. U.; Clarke, B. A. Protein kinase A activation inhibits DUX4 gene expression in myotubes from patients with facioscapulohumeral muscular dystrophy. *J Biol Chem* **2018**, *293*, 11837–11849.

209. Haynes, P.; Bomsztyk, K.; Miller, D. G. Sporadic DUX4 expression in FSHD myocytes is associated with incomplete repression by the PRC2 complex and gain of H3K9 acetylation on the contracted D4Z4 allele. *Epigenetics Chromatin* **2018**, *11*, 47.

210. Giesige, C. R.; Wallace, L. M.; Heller, K. N.; Eidahl, J. O.; Saad, N. Y.; Fowler, A. M.; Pyne, N. K.; Al-Kharsan, M.; Rashnonejad, A.; Chermahini, G. A.; Domire, J. S.; Mukweyi, D.; Garwick-Coppens, S. E.; Guckes, S. M.; McLaughlin, K. J.; Meyer, K.; Rodino-Klapac, L. R.; Harper, S. Q. AAV-mediated follistatin gene therapy improves functional outcomes in the TIC-DUX4 mouse model of FSHD. *JCI Insight* **2018**, *3*.

211. Pearsall, R. S.; Davies, M. V.; Cannell, M.; Li, J.; Widrick, J.; Mulivor, A. W.; Wallner, S.; Troy, M. E.; Spaits, M.; Liharska, K.; Sako, D.; Castonguay, R.; Keates, S.; Grinberg, A. V.; Suragani, R.; Kumar, R. Follistatin-based ligand trap ACE-083 induces localized hypertrophy of skeletal muscle with functional improvement in models of neuromuscular disease. *Sci Rep* **2019**, *9*, 11392.

212. Baker, J. M.; Boyce, F. M. High-throughput functional screening using a homemade dual-glow luciferase assay. *J Vis Exp* **2014**,

213. Desrosiers, R.; Friderici, K.; Rottman, F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci U S A* **1974**, *71*, 3971–5.

214. Perry, R. P.; Kelley, D. E.; Friderici, K.; Rottman, F. The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* **1975**, *4*, 387–94.

215. Dubin, D. T.; Taylor, R. H. The methylation state of poly A-containing messenger RNA from cultured hamster cells. *Nucleic Acids Res* **1975**, *2*, 1653–68.

216. Liu, X. M.; Zhou, J.; Mao, Y.; Ji, Q.; Qian, S. B. Programmable RNA N(6)-methyladenosine editing by CRISPR-Cas9 conjugates. *Nat Chem Biol* **2019**, *15*, 865–871.

217. Wang, H.; Nakamura, M.; Abbott, T. R.; Zhao, D.; Luo, K.; Yu, C.; Nguyen, C. M.; Lo, A.; Daley, T. P.; La Russa, M.; Liu, Y.; Qi, L. S. CRISPR-mediated live imaging of genome editing and transcription. *Science* **2019**, *365*, 1301–1305.

218. Tanenbaum, M. E.; Gilbert, L. A.; Qi, L. S.; Weissman, J. S.; Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **2014**, *159*, 635–46.

219. Mercuri, E.; Darras, B. T.; Chiriboga, C. A.; Day, J. W.; Campbell, C.; Connolly, A. M.; Iannaccone, S. T.; Kirschner, J.; Kuntz, N. L.; Saito, K.; Shieh, P. B.; Tulinius, M.; Mazzone, E. S.; Montes, J.; Bishop, K. M.; Yang, Q.; Foster, R.; Gheuens, S.; Bennett, C. F.; Farwell, W.; Schneider, E.; De Vivo, D. C.; Finkel, R. S.; Group, C. S. Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. *N Engl J Med* **2018**, *378*, 625–635.

220. Kristen, A. V.; Ajroud-Driss, S.; Conceicao, I.; Gorevic, P.; Kyriakides, T.; Obici, L. Patisiran, an RNAi therapeutic for the treatment of hereditary transthyretin-mediated amyloidosis. *Neurodegener Dis Manag* **2019**, *9*, 5–23.

221. Benson, M. D.; Waddington-Cruz, M.; Berk, J. L.; Polydefkis, M.; Dyck, P. J.; Wang, A. K.; Plante-Bordeneuve, V.; Barroso, F. A.; Merlini, G.; Obici, L.; Scheinberg, M.; Brannagan, r., T. H.; Litchy, W. J.; Whelan, C.; Drachman, B. M.; Adams, D.; Heitner, S. B.; Conceicao, I.; Schmidt, H. H.; Vita, G.; Campistol, J. M.; Gamez, J.; Gorevic, P. D.; Gane, E.; Shah, A. M.; Solomon, S. D.; Monia, B. P.; Hughes, S. G.; Kwoh, T. J.; McEvoy, B. W.; Jung, S. W.; Baker, B. F.; Ackermann, E. J.; Gertz, M. A.; Coelho, T. Inotersen Treatment for Patients with Hereditary Transthyretin Amyloidosis. *N Engl J Med* **2018**, *379*, 22–31.