Review

Israel Journal of Chemistry
www.ijc.wiley-vch.de

doi.org/10.1002/ijch.202300169

Engineered RNA-binding Proteins: Studying and Controlling RNA Regulation

Riley W. Sinnott,[a] Yang Cao,[a] and Bryan C. Dickinson*[a]

Abstract: The complexity of eukaryotic organisms is intricately tied to transcriptome-level processes, notably alternative splicing and the precise modulation of gene expression through a sophisticated interplay involving RNA-binding protein (RBP) networks and their RNA targets. Recent advances in our understanding of the molecular pathways responsible for this control have paved the way for the development of tools capable of steering and managing RNA regulation and gene expression. The fusion between a rapidly developing understanding of endogenous RNA regulation and the burgeoning capabilities of CRISPR-Cas and other programmable RBP platforms has given rise to an exciting frontier in engineered RNA regulators. This review offers an overview of the existing toolkit for constructing synthetic RNA regulators using programmable RBPs and effector domains, capable of altering RNA sequence composition or fate, and explores their diverse applications in both basic research and therapeutic contexts.

1. Introduction

The regulation of RNA, a key intermediary in the flow of genetic information in eukaryotic organisms, is foundational to dynamic gene expression and cellular identity.[1] How RNA is regulated within the cell by changes to its stability, chemical composition, sequence, and translation rate has been an intensive research area, and findings over the recent decades have significantly elevated our perception of RNA regulation's importance in the central dogma.[2] Specifically, events that alter RNA composition such as alternative splicing and 3' end processing are foundational to eukaryotic organisms' complexity through their influence on protein composition and localization.[1] We have also come to understand that the regulation of mRNA stability and translation by RNA-binding proteins (RBPs) and noncoding RNA such as miRNA or lncRNA can uniquely tune proteomic homeostasis in distinct tissues.[3] More recently, work by the research group of Chuan He and colleagues into RNA regulatory pathways such as N^6-methyladenosine (m^6A) modification have been foundational to the emergence of the field of "epitranscriptomics" (Figure 1a).[4] Epitranscriptomics focuses on probing how reversible RNA modifications impact RNA metabolism and translation, which can direct cell fate with broad implications from cancer to development.[5]

RBPs that recognize and bind specific RNA modifications or sequence motifs to elicit an effect on the bound RNA are the functional workhorses driving chemical modification-dependent gene expression changes at the RNA level.[6] In fact, we are beginning to understand that the same modification can have diverse downstream effects on a specific RNA, depending on which RBP(s) – each with different specific activities and/or interaction partners – are bound.[7] These RBP-driven effects can include the modulation of: splicing, polyadenylation, stability, localization or translation efficiency. The power and versatility RNA regulation possesses has inspired considerable interest into whether RBPs can be precisely engineered to programmably influence gene expression (Figure 1b). This interest has coalesced into a rapidly expanding field driven by the convergence between advances in our understanding of RNA regulation and also engineering methods informed by the recent breakthroughs in programmable DNA-targeting technologies. In this review, we will discuss the recent progress of engineering proteins to bind specific RNA sequences as well as the functionalization of these proteins to direct diverse processes that modify RNA's composition and/or fate. We also highlight several exciting applications of these systems to probe and study biological systems for basic research, as well as synthetic RBPs with therapeutic functionality for correcting aberrant gene expression. Finally, we will describe further progress into harnessing RNA regulation beyond RBPs.

2. Programmable RNA-Binding Proteins

How native RBPs specifically interact with RNA is an active area of research, with many recently discovered RBPs containing novel domains not previously linked to RNA binding.[8] Well-studied RBPs often consist of multiple weak RNA-binding domains that contact a target RNA simulta-

[a] *R. W. Sinnott, Y. Cao, B. C. Dickinson*
Department of Chemistry, University of Chicago, Chicago IL 60615 USA
E-mail: Dickinson@uchicago.edu

Israel Journal of Chemistry

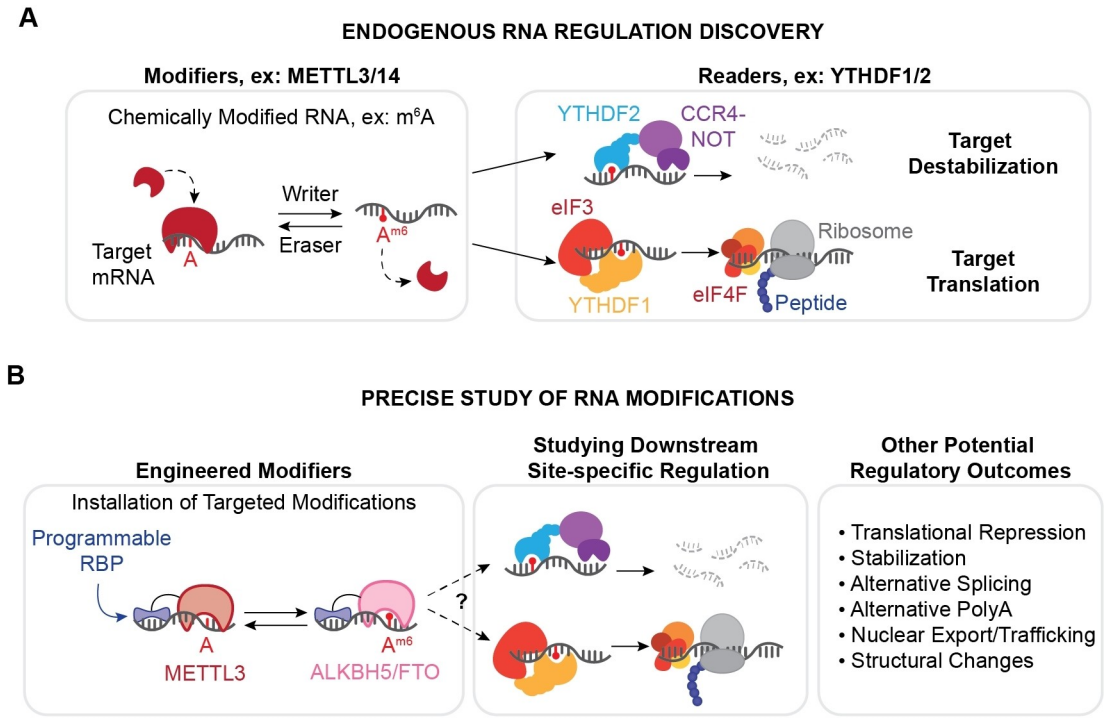
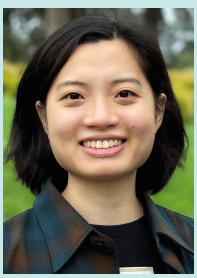


Figure 1. Breakthroughs in epitranscriptomics enable engineered platforms to direct and study RNA regulation. This figure depicts m^{6}A modification as an example RNA regulatory pathway that can be co-opted by engineered RBP platforms. a, m^{6}A modification installation by "writer" enzymes and removal by "eraser" enzymes mediates downstream regulation of the modified RNA by "reader" proteins that can have different effects depending on their interaction partners such as destabilization (YTHDF2) or translational activation (YTHDF1). b, Programmable RBPs can direct fused "writer" or "eraser" enzymes to explicit RNA targets and influence their site-specific modification. The downstream effects of the RNA's modification change can then be studied for basic research or leveraged to cause other known regulatory outcomes depending on the effector domain fused to the RBP.

Riley Sinnott received his B.S. in Biochemistry from Tulane University and gained an interest in engineering biological systems working on transcription factor-based butanol biosensors in the lab of Prof. Nicholas Sandoval. He is currently a fourth-year Chemistry Ph.D. student and NSF Graduate Research Fellow in the lab of Prof. Bryan Dickinson with a research focus on engineering programmable RBPs to address outstanding therapeutic needs for fine-tuning gene expression.



Yang Cao earned her B.S. in Chemistry from Peking University and her Ph.D. in Chemistry from the University of Chicago. She studied enzymes to regulate protein lipidation and developed biotechnologies to control mammalian mRNA translation under the supervision of Prof. Bryan Dickinson during her Ph.D. She is currently a Postdoc at University of California, Berkeley.

Bryan Dickinson earned his B.S. in Biochemistry from the University of Maryland and his Ph.D. in Chemistry from the University of California, Berkeley. After a Jane Coffin Childs Memorial postdoctoral fellowship at Harvard University, he joined the faculty at the University of Chicago in the Department of Chemistry in the Summer of 2014, was promoted to Associate Professor in 2019, and Professor in 2023. The Dickinson Group employs synthetic organic chemistry, molecular evolution, and protein design to develop molecular technologies to study and control chemistry in living systems. The group's current primary research interests include: 1) developing new evolution technologies to reprogram and control biomolecular interactions, 2) engineering RNA-targeting biotechnologies as new therapeutic platforms, and 3) leveraging chemical biology to study biomolecular interactions. The motivating principle of the Dickinson Group is that chemists' ability to create functional molecules through both rational and evolutionary approaches can lead to new breakthroughs in biology and biotechnology.

neously to increase affinity and specificity.[9] However, many RBPs regulate multiple RNA binding partners with similar short sequence motifs and do not possess the specificity necessary for single target RNA-binding. Engineering *de novo* RNA binding domains remains an outstanding challenge, in part due to the inherent flexibility of RNA structures as ligands and our limited understanding of how RBPs achieve sequence-specific binding and how that specificity may be further improved.[10]

Conversely, programmable DNA-targeting with engineered proteins has been achieved through the use of modular, repeat domain proteins, such as TALENs or zinc fingers, which can sequence-specifically bind with 1 or 3 nucleotide(s) per domain, respectively.[11] By fusing several domains together, zinc finger and TALENs repeats are typically designed to bind approximately 18 bases, with additional engineering (e.g., simultaneous binding of a second designer protein to the nearby antisense strand required for DNA cleavage) capable of further increasing target specificity. Zinc fingers and TALENs have thus been functionalized with many different DNA-modifying proteins for a myriad of genome engineering-related applications.

Analogously, Pumilio proteins are RBPs that uniquely derive their RNA-binding capability through modular repeat domains (PUF domains) that each recognize a single RNA base (Figure 2b).[12] Much like DNA-binding TALENs, modular PUF proteins have been engineered for programmable RNA-binding in mammalian cells, however, PUF proteins generally only allow effective targeting of up to 8–9 RNA bases, greatly limiting their specificity and use in many bioengineering applications.[13] Pentatricopeptide repeat (PPR) proteins are another RBP family that recognize single RNA bases through repeated domains and have been pursued for designer RBP engineering. PPR proteins may be amenable to

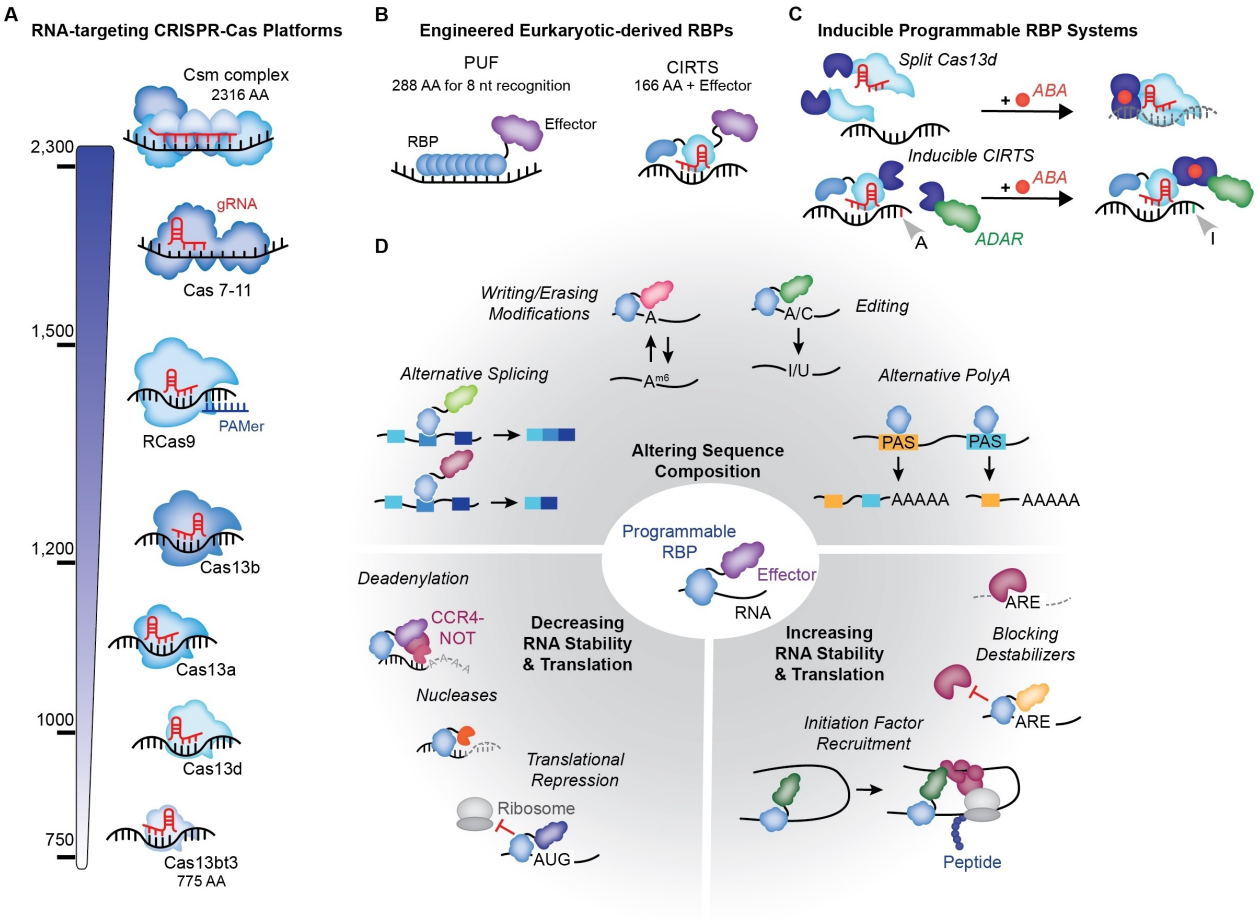


Figure 2. Current programmable RBPs. a, Current RNA-targeting CIRSPR-Cas platforms and their relative size by amino acid (AA) number. Note that the Csm complex AA number includes all subunits (Csm 1–5) required for RNA-targeting and cleavage in their respective stoichiometry. b, Engineered, eukaryotic protein-derived programmable RBPs and their respective AA size without additional necessary effector domain. c, Examples for inducible programmable RBP platforms. Upon ABA addition the two subunits of split Cas13d reconstitute and then can bind and degrade an RNA target. Likewise, inducible CIRTS can bind an RNA target, but only ABA addition can induce the effector (the base editing enzyme ADAR) to interact with the bound RNA and edit the target. d, Range of possible RNA regulatory functionalities that can be programmed either by RBP binding or RBP-mediated direction of an RNA regulatory protein effector. PAS – polyadenylation site

targeting longer RNA sequences than PUF proteins, but their potential for programmable RNA-binding in mammalian cells, though under active investigation, is still unclear.[14]

Overall, development and application of these platforms for programmable RNA binding have been slowed by laborious engineering and screening necessary to uncover domain sequences that bind efficiently to new RNA-targets as well as limited target specificity.[13b,14] Additionally, modular repeat proteins for programmable nucleic acid targeting have largely been supplanted during the last decade by the explosion of CRISPR-Cas technologies and their development into programmable nucleic acid binding platforms.[15]

CRISPR-Cas proteins offer facile programmability through interaction with a guide RNA (gRNA) that directs the Cas protein to a DNA target via antisense complementarity (Figure 2a).[16a] While the CRISPR-Cas systems initially discovered were found to target DNA, successful application to targeting disease relevant RNA in eukaryotic cells was achieved by engineering the DNA-targeting Cas9 variant into a programmable nuclease inactive RNA targeting moiety, Rcas.[16b] These successes inspired the discovery of natively RNA-targeting Cas protein type VI systems with families such as Cas13, which include several variants that have been utilized for programmed RNA manipulation, such as Cas13a, Cas13b, and Cas13d.[17,18] Similar to their DNA-targeting counterparts, Cas13 natively cleaves RNA duplexed with the complementary gRNA.

The nuclease activity of Cas13 proteins can be abolished by mutation to produce catalytically dead Cas13 (dCas13), which yields a gRNA-directed RBP. Furthermore, dCas13 offers straightforward functionalization beyond RNA-cleavage via fusion with different effector protein domains that carry out activity when in close proximity to a target RNA.[19] Additionally, we leveraged the blueprint for programmable RNA-targeting laid out by CRISPR-Cas platforms and recapitulated their functionality using engineered human protein parts. This lead to the development of the CRISPR/Cas-Inspired RNA Targeting System (CIRTS) platform, a protein fusion capable of specific binding to a gRNA and functionalization through addition of further protein domains (Figure 2b).[20] These innovations have driven a rapidly expanding suite of platforms to address fundamental questions about RNA's role in central life processes and offering new therapeutic angles for traditionally "undruggable" diseases.

Recent advances in protein-based RNA-targeting generally fall into two categories: 1) deployment-related improvements to established Cas13 platforms, either through protein engineering to reduce size or gRNA design tools to improve on-target efficacy and reduce gRNA screening efforts; and 2) the discovery and molecular study of novel RNA-targeting CRISPR-Cas platforms from the Cas13 family or other classes. In this section we will discuss both the state-of-the-art in engineered Cas13 variants and emerging platforms with their comparative advantages and disadvantages.

2.1 Engineered Miniature Cas13

Protein size is a key determinant for platform utility and can be a limitation especially in contexts where delivery by viral vectors with strict packaging limits, such as AAV, are a necessity. Consequently, to open up design space for directing RNA modifications in delivery-restrictive tissues, reducing the size of the relatively large Cas13 members has garnered increased attention. These miniaturization efforts have primarily focused on two leading Cas13 platforms: the most potent effector, Cas13d, and the current smallest variant, Cas13bt3 (also known as Cas13X.1; Figure 2a). The expansive and continually deepening Cas13 molecular and structural knowledge base has been an important guide to understanding domain functionality and informing engineered Cas13 variants.[21] Furthermore, recent breakthroughs in protein modeling with AlphaFold2 have bolstered these developments.[22]

For Cas13d, Zhao and Zhang *et al.* described and demonstrated an iterative domain truncation strategy aimed at trimming residues that do not mediate interactions with the gRNA or target RNA, as they are less likely to contribute to either the RNA-binding or cleaving processes.[23] Impressively, the authors demonstrated that this engineering strategy could be effective when informed by Cryo-EM structures or AlphaFold2 structural predictions for several Cas13d variants and a Cas13b variant. The study produced a mini-RfxCas13d (682 AAs) that possessed functionality on par with the wildtype enzyme for both on-target nuclease activity and direction of base editors.

The Cas13bt proteins were discovered via metagenomic mining for more compact Cas13 members of the already established Cas13b family, of which, Cas13bt3 is the smallest at 775 AAs.[24] Comparisons between common Cas13b and Cas13d effectors demonstrated that despite their reduced size, Cas13bt family members are only slightly less efficient for on-target knockdown or base editing, making Cas13bt3 an attractive option for minimal nuclease and base editing applications.

Further domain truncations resulted in a dCas13bt3 mini variant with a total length of 445 AAs with similar or slightly reduced ability to direct base editors as the full-length version, depending on the target context. Further Cas13bt3 mechanistic and structural studies revealed several mutations which can improve its RNA-binding ability, demonstrating that Cas13bt3 can be further engineered for improved efficacy and reduced size (down to a minimal 417 AAs).[25] Overall, the flexibility the minimal Cas13 variants afford for delivering base editors and gRNA arrays, as well as the potential to package larger effector fusions, make these engineered platforms important tools both for therapeutic applications and basic research.

2.2 Cas13 gRNA Design Optimization

Elucidating gRNA spacer design principles to improve on-target and reduce off-target binding carry importance in any application, but especially in areas such as high-throughput RNA-knockdown assays, therapeutic base editing, and RNA sensing, where gRNA efficacy and accuracy are vital. As platforms such as Cas13d gather interest in these areas, several groups have focused on improving RfxCas13d gRNA design via high throughput characterization of tens or even hundreds of thousands of gRNAs in parallel with read outs such as on-target nuclease activity *in cellulo* or *in vitro* RNA-binding.[26]

Data sets generated from these studies have been used to inform both bioinformatic prediction for efficacy for *de novo* designed gRNA and biophysical models for gRNA-directed Cas13d RNA-binding. There now exists several public bioinformatic tools for Cas13d gRNA design for a given target RNA, which has greatly increased accessibility and programmability for this platform.[26b–d,27] Resources to aid with Cas13d gRNA design and prediction can be found at: cas13design.nygenome.org; arcinstitute.org/tools/cas13d; and deepcas13.weililab.org. Furthermore, a molecular understanding of gRNA-directed Cas RNA-binding also informs the design of gRNA with sensitivities to even single mismatches with their target, which may open the door for differential targeting between transcripts from highly similar alleles. For example, targeting splicing-specific transcript isoforms using gRNA landing sites on unique exon-exon junctions may enable specific targeting of almost 90% of human RNA isoforms.[28] Likewise, there has been interest in expanding these screens to other Cas13 platforms such as Cas13a and Cas13b and it is likely that the methods for defining programmability pioneered around Cas13d will be applicable to other emerging programmable RNA-binding platforms.[29]

2.3 Inducible Programmable RBP Systems

All previously described programmable RBP platforms will constitutively bind the RNA-target as long as they are expressed, but there are applications where greater functional control is required. Inducible genome editing systems, activated through various exogenous signals, have previously been engineered, primarily through the fusion of two halves of a genome editor protein to binding partners that only interact when a chemical signal (e.g., a small molecule or peptide ligand) is present.[30] Thus, the activity of the "split" engineered genome editor depends on the inducer molecule causing dimerization of its interaction partners bringing the two halves into close proximity. These chemically inducible platforms possess distinct advantages for temporal control over target manipulation for applications in basic research or precise therapeutics.

This functionality has likewise been integrated into programmable RBP systems by engineering Cas13 into two separate protein subunits that cannot reconstitute into a functional RBP without added small molecules such as abscisic acid (ABA) inducing dimerization domains fused with the Cas13 subunits to interact and bring the Cas13 subunits into close proximity (Figure 2c). Similar inducible systems have been developed for both Cas13b and Cas13d and leveraged to program small molecule-inducible RNA-knockdown in cells and also *in vivo*.[31] Additionally, split Cas13 has been utilized for programmable m^6A writing and other similar split systems comprised of programmable RBPs, such as CIRTS, and effector domains have also be deployed for inducible functions such as base editing, translational activation, and degradation (Figure 2c).[32] Collectively, inducible RBP binding, or effector recruitment allows for additional temporal and reversible control over RNA regulation with applications in basic research in time-sensitive experiments and potentially gene circuits for programmable cellular rewiring in advanced therapeutics.

2.4 Type III RNA-Targeting CRISPR-Cas Platforms

Type-III Cas systems represent the second major group of natively RNA-targeting Cas platforms apart from the aforementioned type VI Cas13 proteins.[33] Unlike type VI effectors, type III CRISPR systems are often composed of multiple proteins which must complex together to recapitulate the gRNA processing, and directable RNA-binding and cleavage processes mediated by single protein Cas13 effectors. Due to this additional complexity, the study and deployment of type III systems has been limited compared to type VI systems, but there still exists interest in developing methods to deploy these multi-subunit effectors or identify type III systems that could be deployed in a single-effector manner due to advantages in specificity and accessory functions over current type III systems.

2.4.1 CRISPR-Csm Complexes

Several molecular and structural studies have focused on elucidating the protein stoichiometry and RNA-targeting capabilities of CRISPR-Csm complexes, a multisubunit type III CRISPR system (Figure 2a).[34] Csm complex gRNA interaction and subsequent RNA-targeting and cleavage requires five Csm proteins in varying amounts and an additional Cas6 protein for gRNA maturation from CRISPR arrays.[35] As with Cas13 effectors, gRNAs can be designed with a specific Csm direct repeat to mediate their interaction with the Csm complex, and a programmable spacer region to obtain directed Csm binding and RNA-cleavage. In a study using eukaryotic cells, it was shown that Csm complexes specifically and potently knock-down target RNA levels, both for nuclear noncoding RNA and mRNA, much like Cas13 effectors.[36] However, unlike Cas13, type III CRISPR-Cas systems do not possess the domains responsible for collateral cleavage of bystander RNA, offering a potential advantage for RNA degradation over native Cas13 in situations where

specificity and cytocompatibility are vital, such as high-throughput RNA knockdown screens.

Another interesting aspect of Csm complexes that may be advantageous in biotechnological applications is that, due to the complex's stoichiometry, Csm complexes can deliver multiple copies of fused proteins to a target RNA as up three protein copies are required of some complex members (Csm3) for RNA-targeting. For example, Colognori *et al.* demonstrated that Csm3-GFP fusions, when incorporated into a catalytically dead Csm complex, can be used to image RNA.[36] While the multisubunit nature of Csm complexes will likely impede their development as therapeutic modalities due to delivery challenges, their unique cleavage specificity may support their adoption as another RNA-knockdown tool to compliment RNAi, especially for targets for which RNAi is less efficient, such as nuclear RNA.

2.4.2 CRISPR-Cas7–11

Interest in identifying other type III CRISPR systems without collateral RNA-cleavage activity led to metagenomic mining efforts focused on identifying whether single effector-like variants of the protein family existed. These efforts led to the discovery of CRISPR-Cas7–11 which, through the native fusion of several domains typically found in separate type III Cas protein subunits, is functionally akin to the single effector Cas13 family (Figure 2a).[37] Indeed, Özcan *et al.* demonstrated that Cas7–11 itself was capable of gRNA array processing and specific RNA cleavage without the cytotoxicity sometimes observed in mammalian cells with RNA-knockdown with some Cas13 variants, as well as dCas7–11 variants capable of directing base editors. The comparable, albeit slightly weaker, on-target RNA degradation compared to the more widely used Cas13b and Cas13d variants made it an immediately viable choice for improved specific RNA-knockdown like Csm complexes, but in a single effector format. However, the smallest Cas7–11 variant identified in the study is almost twice the size (1,367 AAs) of smaller Cas13 variants, which hinders deliverability, even as a nuclease.

Follow up studies leveraged Cryo-EM to elucidate the Cas7–11 domains important for RNA binding and cleavage as well as their molecular mechanism.[38] These insights provided the basis for structure-guided Cas7–11 truncation to produce a minimal Cas7–11S with similar activity to the wildtype protein. While Cas7–11S was still on the larger end of RNA-targeting Cas platforms at 1,290 AAs, it can be delivered as a nuclease in single vector AAV. Even so, the packaging limits of the AAV vector precludes the fusion of any additional effectors. Nonetheless, Cas7–11 and its engineered variants are viable platforms for programmable RNA-targeting and have advantages for RNA-knockdown compared to Cas13 family, as long as delivery is not a constraint, similar to Csm complexes. While beyond the scope of this review, Cas7–11 has recently begun to occupy a unique role in RNA-sensing applications due to the identification of native accessory proteins that allow for specific proteolytic cleavage upon Cas7–11 binding its target RNA.[37b,39] Several groups have leveraged this functionality to create engineered protease targets which generate a measurable signal upon cleavage to create an exciting Cas7–11 RNA-sensing platform.[40]

3. Effectors for Engineering Programmable RBP with Diverse Functions

3.1 Engineering Intrinsic RNA-Targeting CRISPR-Cas Nuclease Activity

While most Cas13 proteins have clearly been demonstrated to possess the ability to not only cleave a target RNA upon duplexing with its gRNA in *cis*, but also bystander ssRNA in *trans* in a gRNA-independent manner (also referred to as collateral cleavage) *in vitro*, there have been conflicting reports about whether this activity occurs in mammalian cells and to what extent.[17a,c,37a,41] Several recent studies have contributed to the understanding that the extent which Cas13-dependent collateral RNA cleavage can be measured in eukaryotic cells is positively correlated with the abundance of the target transcript.[42] Since target recognition is a prerequisite for Cas13 nuclease domain exposure to the cytoplasm, the greater the number of "activated" Cas13 increases bystander RNA cleavage.[21b,42b] When the target is relatively abundant, usually in the case of reporters or housekeeping transcripts, the collateral cleavage can be significant enough to reduce cellular fitness and has even been used to deplete distinct cell populations in complex mixtures.[42b,43]

Though collateral activity has some utility, for most purposes, precise knockdown and no dependance between target abundance and cellular toxicity are preferable. Furthermore, a recent study demonstrated that the collateral activity of many Cas13 variants complicated their delivery by lentivirus, which is an important method used in many library screens.[42c] Some strategies for reducing collateral activity from native Cas13d have been demonstrated, such as limiting Cas13d expression in situations where multiple Cas13d proteins simultaneously bind the target, like in expansion repeat diseases.[44] However, this design may not reduce collateral activity in all contexts and does not ablate it completely.

In order to investigate methods to mitigate collateral cleavage, Tong *et al.* utilized a dual-fluorescent reporter strategy to simultaneously measure both on-target and collateral cleavage from Cas13 targeting in HEK293T cells and confirmed pervasive cleavage of the off-target reporter transcript from multiple native Cas13 variants including Cas13d and Cas13X.[45] The authors then implement a random mutagenesis strategy to alter amino acids in an RNA-binding cleft near the catalytic domains, creating over 100 Cas13d variants and screening them for collateral activity by FACS. A high fidelity Cas13d variant (hfCas13d) was identified by markedly

reduced collateral activity in the reporter screen and was then validated by RNA-seq across multiple endogenous targets. Critically, hfCas13d was also confirmed to possess greatly reduced cellular and *in vivo* toxicity compared to native Cas13d when targeting an endogenous transcript, but with comparible on-target knockdown.

The generalizability of this strategy was then demonstrated by a similar mutagenesis campaign on Cas13X, which also identified a hfCas13X variant with minimal sacrifice to on-target activity, but the lowest measured collateral cleavage across several tested native Cas13b, Cas13d, and Cas13X variants and also hfCas13d for some targets. The mutational hotspot originating the favorable properties of hfCas13X were also independently deduced and verified by a structure-guided engineering approach by Deng and Osikpa *et al.*[46] Their study solved the Cyro-EM structure of target-bound, activated Cas13bt3 and identified positively charged patches on the enzyme's surface near the catalytic domains, which could be mutated to ablate nonspecific cleavage. The authors also found several tyrosine to alanine mutations in this site that were beneficial, which were comparible to the mutations made in the previously engineered hfCas13X, and verified a similar decrease in collateral activity. These engineered enzymes are promising candidates for precise RNA-degradation in sensitive contexts for any but especially abundant targets.

3.2 Engineering and Directing RNA-Effectors for Diverse RNA-Modifications

3.2.1 Programmable RNA-Modification Editing

A wide repertoire of programmable RNA-binding platforms such as dCas9, dCas13a, dCas13b, and dCas13d have been fused to m^6A writers (most commonly METTLE3) and erasers, such as ALKBH5 and FTO, to precisely install or remove m^6A modifications on specific targets in a gRNA-dependent manner (Figure 2d).[47] Temporal control over writing and erasing m^6A has also been implemented by employing the ABA-based inducible system to coordinate dCas13b and a writer or eraser upon addition of ABA.[48] Importantly, the modifications are reversible in the absence of ABA providing a more precise tool to investigate the biological consequences of specific modification events. Further tool development, such as photocaged ABA for light-inducible RNA-modification and ligand stabilized Cas13, continues to provide greater measures of control.[48–49] As more RNA-modifications and their associated writing/erasing pathways are described, such as m^5C, m^1A, and pseudouridine, the tools currently developing around programmable m^6A modification could likely inform programmable writing/erasing platforms to further their study as well.[50] Indeed, there have already been some descriptions for programmable m^1A erasers applied to demethylate endogenous targets.[51]

3.2.2 Altering RNA Sequence and Composition

Processes that rewrite RNA sequences such as editing, alternative splicing, and alternative polyadenylation have the ability to dramatically modify encoded proteins or rewire how an RNA is regulated in the cell.[52] Naturally, eukaryotic cells utilize these pathways to derive much of its transcriptomic diversity which often is reflected in cell differentiation, and when dysregulated, is a causal factor in many diseases. The ability to programmably direct such impactful changes has been an active area of research even before the advent of CRISPR-Cas technologies when PUF proteins were utilized to deliver fused splicing inhibitors or activators to manipulate exon exclusion or inclusion, respectively, into mature transcripts (Figure 2d).[53] Both functionalities were also later replicated with Cas13 but interestingly just Cas13d itself binding to an exon could induce its skipping, with splicing inhibitor effectors only marginally increasing this effect.[17c,19c]

These initial successes have continued to stoke interest in developing more potent splicing activator RBPs. In 2024, Schmok *et al.* utilized large-scale tethering assays to screen over 700 RBPs to identify new proteins capable of mediating alternative exon inclusion on a luciferase splicing reporter.[54] They identified several splice-enhancing proteins, which were then truncated to yield an expanded toolbox of smaller and more effective protein domains to drive exon inclusion when fused to Cas13d. The future application and engineering of RBPs for alternative splicing has exciting therapeutic applications (e.g., correcting aberrant splicing events) as well as supporting our understanding of endogenous splicing networks.

Similar to blocking exon inclusion by Cas13 binding, Tian and Zhang *et al.* demonstrated an analogous blocking mechanism with Cas13 could also be used to disrupt polyadenylation from nearby bound signal sequences to reduce transcript isoform expression derived from specific alternative polyadenylation signals.[55] Additionally, it may also be possible to deliver alternative cleavage and polyadenylation (APA)-inducing effectors to boost APA at a specific signal in an opposing mechanism.[56] Altogether, these programmable platforms are amassing an exciting toolbox to manipulate transcript isoform expression ratios with opportunities to aid study for isoform-specific biology and correcting diseases originating from isoform ratio imbalances. While still on the horizon, the development of methods to use programmable RNA-binding proteins to direct *trans* splicing of user-defined exons, which would have exciting applications in bioengineering and therapeutics alike, is also being explored.[57]

There has also been rapidly growing preclinical interest in developing programmable RBPs with precise RNA base editing activities (Figure 2d). Endogenous RNA base editing is largely mediated by adenosine deaminase acting on RNA (ADAR) enzymes which catalyze directed A-to-I edits (recognized as G in base pairing and translation) conversion.[58] The Zhang lab developed the first programmable RBP-based RNA editing platform, REPAIR, by fusing the human ADAR2

adenosine deaminase domain with Cas13, allowing A-to-I editing of both reporter and endogenous RNAs.[19a] Furthermore, ADAR2 natively recognizes dsRNA substrates, enabling the direction of its activity with Cas13 since the gRNA can be designed to have the target adenosine unpaired in the gRNA: target duplex. ADAR2's specificity was improved by rational mutagenesis to decrease its native RNA-binding ability, combined with a previously described mutation to increase ADAR2's activity, to yield the final REPAIR platform. Overall, the authors demonstrated that REPAIR can edit A in any neighboring base context with a typical editing rate of approximately 10–30% for endogenous, disease relevant transcripts with very limited off-target editing. Since this study was published in 2017, REPAIR has been improved by incorporating later generation Cas13 modalities such as Cas13d and Cas13X (Cas13bt3) to boost its on-target editing efficiency upwards of 50% in some contexts, thus greatly increasing its preclinical potential.[24]

The Zhang lab was also the first to expand RNA editing beyond A-to-I by using an engineered ADAR2 capable of deaminating cytosines, yielding a C-to-U editor named RESCUE.[59] Like REPAIR, RESCUE typically produced editing rates between 10–30%, but due to editing of bystander A sites as well as C, RESCUE had a less favorable specificity profile. Since expanded base editing functionality is valuable for addressing a greater variety of disease relevant mutations, other groups have worked to improve targeted C-to-U editing. Huang *et al.* engineered another C-to-U editing platform by utilizing Cas13 to direct APOBEC3A, a native cytosine deaminase, to unique gRNA which isolate the target C on a bulge that mimics native APOBEC3A substrates.[60] This platform, CURE, was shown to edit multiple reporters and endogenous substrates with relatively high rates (40–50%), but this was largely restricted to APOBEC3A-preffered motifs, where a U proceeded the target C. Future use of CURE or RESCUE would largely depend on the sequence context for the target C, as CURE demonstrated higher editing rates for UC sites, but is likely outperformed or matched by RESCUE in most other contexts.

3.3 Effectors for Repressing or Amplifying Translation

Beyond modifying RNA composition or endonuclease-mediated degradation there are still numerous methods that RBPs employ to ultimately influence a transcript's output. In the case of negative regulation, a storied body of work has shown that mammalian transcripts are often destabilized through association with deadenylase complexes, which simultaneously remove a mRNA's polyA tail and repress its translation.[61] RBP domains that mediate the interaction with deadenylase complexes have also been engineered as effector domains for programmable RNA destabilization and translational repression under multiple platforms such as PUF proteins, dCas13, and CIRTS (Figure 2d).[19b,20,62] Furthermore, direct translational repression without altering transcript abundance has been demonstrated in some contexts, either through direct binding to important translational regulatory elements, such as the start codon, with PUF proteins or dCas13.[63] Repression may be further enhanced by fusion of Cas13 with known translation repressing RBPs such as 4EHP (Figure 2d).[63b]

Platforms to programmably increase protein expression from RNA targets in eukaryotic contexts still represent a major area of need. Several eukaryotic initiation factors or their recruiters such as eIF4e or the N-terminal domain of YTHDF1, which may bind eIF3, have been shown to increase reporter protein expression without affecting transcript levels when directed by PUF proteins or Cas13 and CIRTS respectively, but their application to endogenous targets has been limited (Figure 2d).[19b,20, 63c, 64] These roadblocks have led to the investigation of other methods, beyond directing translation-related proteins, to increase transcript protein expression from programmable RBPs.

Recently, Cao and Li *et al.* showed that incorporation of a SINEB2 RNA domain, which can itself increase translation when directed by a *cis* gRNA sequence, into the gRNA of dCas13d could increase translation of a reporter to a greater degree than the SINEB2 domain alone (an RNA-based technology known as SINEUPs).[65] The authors then showed that this strategy could also be applied to several endogenous targets, demonstrating that amplification of tumor suppressor genes in this manner can successfully decrease tumor size and metastasis in an *in vivo* model. Additionally, Li *et al.* demonstrated a method to stabilize transcripts that have known negative regulator binding sites by directing dCas13 fused to the negative regulator's RNA-binding domain (TTP in their study), competing the native negative regulator off of the transcript of interest thereby increasing its abundance (Figure 2d).[66]

While these are innovative strategies, the further development of direct protein-based methods for both translational repression and activation in eukaryotic cells will likely depend on the continued study and characterization of RBP function. RNA networks are often regulated by many RBPs in concert which can complicate the determination of their individual effects. Additionally, some RBPs' function depends on cellular or bound RNA-sequence context.[8] Overall, a greater understanding of how and when individual RBPs direct their endogenous RNA-targets is necessary for their engineering and deployment as programmable translational effectors. Ongoing research into characterizing individual RBP function is rapidly expanding our knowledge of how mammalian cells dynamically control RNA fate and how to design and build programmable systems.[62c,67]

4. Programmable RBPs in Preclinical Applications

Cas13 has found recent success in mediating the degradation of mRNAs encoding proteins whose "undruggable" aggregation play central roles in devastating neurodegenerative

diseases (Figure 3a). Huntington's disease (HD) is a prime example with longstanding interest in targeting the transcript of the causal autosomal dominant huntingtin (HTT) protein mutant. HD linked HTT mutants harbor CAG trinucleotide repeats that encode toxic polyglutamine expansions in disease carriers.[68] In pioneering proof-of-concept studies, the Yeo lab leveraged RNA-targeting Cas9 to selectively degrade pathogenic HTT mRNA and other toxic, repetitive transcripts demonstrating efficient RNA targeting and clearance both *in vitro* and *in vivo*.[69] Now, the recent description of smaller and more potent RNA-targeting nucleases such as Cas13d has opened the door to further preclinical development for these strategies.

In 2022, Morelli and Wu *et al.* demonstrated Cas13d can efficiently degrade toxic HTT mRNA and reduce pathogenic HTT protein expression in patient cells, as well as in a mouse model of HD, following AAV-mediated delivery to the striatum.[70] Importantly, the authors demonstrated effective degradation in a toxic allele-specific manner without altering levels of wild-type HTT mRNA, which is important for maintaining neuronal homeostasis even in HD patients. This allele-specific targeting represented a major advance over previous efforts with RNAi and antisense oligonucleotides (ASOs), which were either toxic due to off-target degradation of the healthy transcript or did not efficaciously degrade the toxic transcript.[71] Cas13d AAV treatment led to phenotypic improvements in the HD mice for at least eight months without major adverse effects resulting from off-target transcriptomic perturbation. This study showcased that RNA-targeting nucleases could be uniquely positioned to address an important and difficult to manage disease.

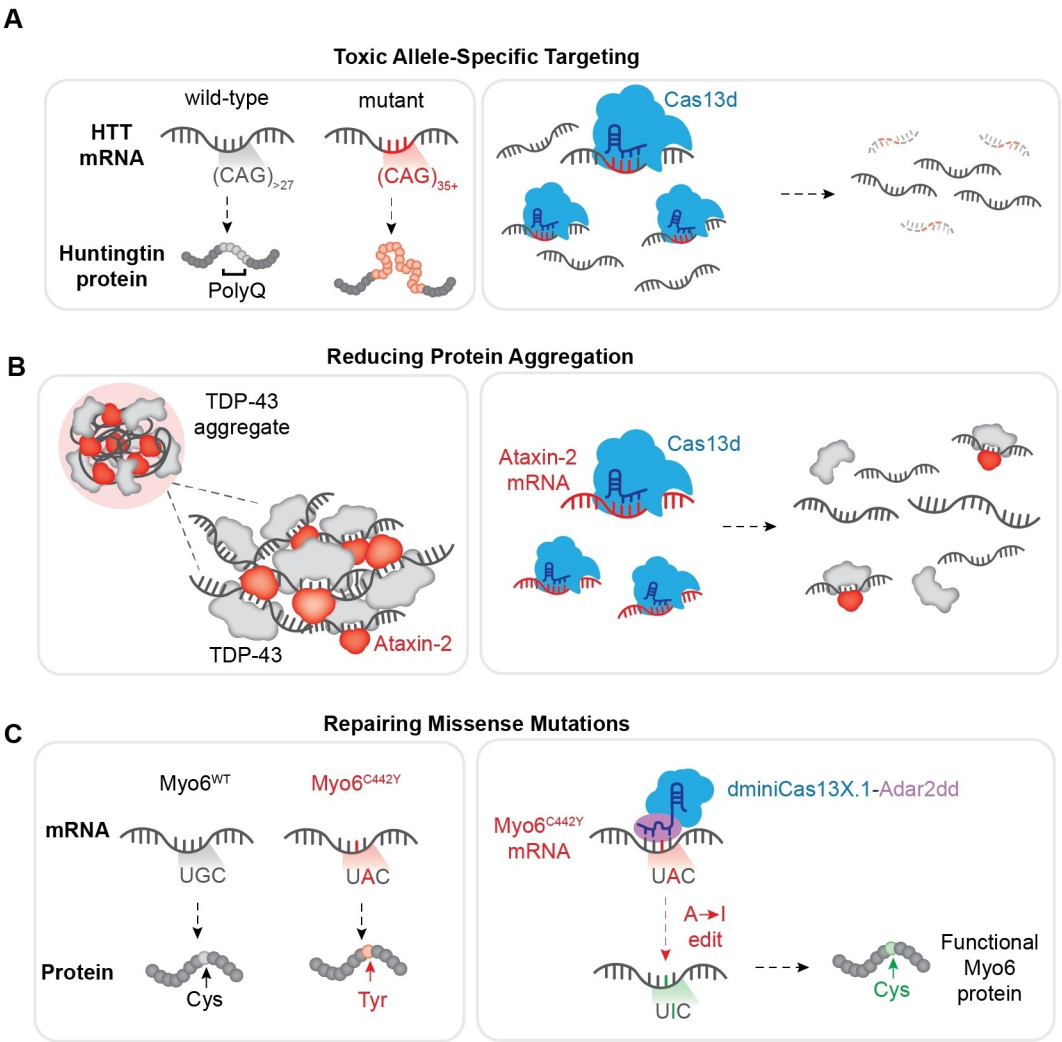


Figure 3. Highlighted preclinical applications of therapeutic programmable RBPs a, Allele-specific targeting of the mutant CAG expansion in the HTT mRNA by Cas13d allows its precise degradation and reduction in toxic Huntingtin protein levels. b, Ataxin-2 mRNA degradation by Cas13d reduces TDP-43-related aggregates associated with several forms of ALS and frontotemporal dementia. c, A missense mutation in Myosin VI leading to deafness can be corrected by Cas13 base editors to restore functional protein expression.

Simultaneously, the Gaj lab also demonstrated the potential for Cas13d–mediated gene silencing in the central nervous system to address neurodegenerative diseases. They first showed that Cas13d–mediated SOD1 mRNA knockdown, which encodes a metalloenzyme whose mutation is linked to some forms of ALS, improves outcomes in mouse models of this disease when delivered via AAV to the spinal cord.[72] Then, in 2023, they applied Cas13d to degrade the ataxin-2 transcript. Ataxin-2 is a positive modifier of neurodegenerative diseases caused by TDP-43 aggregation, including many types of ALS and frontotemporal dementia (Figure 3b).[73] Zeballos *et al.* demonstrated that knockdown of ataxin-2 protein levels by Cas13d effectively reduced wild type TDP-43 aggregation and some TDP-43 proteinopathy-related mutants in cell-based assays. These results then translated to a severe TDP-43-related ALS mouse model characterized by an average lifespan of only 24 days. Intracranial administration of AAV encoding ataxin-2 targeting Cas13d reduced TDP-43 aggregates, improved phenotypic ALS markers, and extended the lifespan of the juvenile mice up to 35 days on average.

Interestingly, the authors also compared off-target transcriptomic effects between native RfxCas13d (used in the mouse model), a high fidelity Cas13d mutant (hifiCas13d), and Cas7–11, and confirmed that hifiCas13d significantly improved the measured knockdown of off-target transcripts over the other two systems in cells. Overall, this study contributes to the strong case for the development of Cas13d–based therapies for "undruggable" neurodegenerative proteinopathies and also provides evidence that hifiCas13d variants may be able to address legitimate concerns related to non-specific cleavage, though more preclinical testing around these variants *in vivo* is necessary.[74]

Cas13X (also called Cas13bt) and its high-fidelity variants (hfCas13X) have also been deployed to demonstrate proof-of-concept targeting in a number of disease contexts by Hui Yang's group. In two studies published in 2023, Yang's group demonstrated both target and contextual adaptability of hfCas13X by degrading disease-relevant mRNA and lncRNA in mouse models of autosomal dominant pigmentosa and Angelman's syndrome respectively.[75] In addition, a collaboration between the Li and Shu groups published in 2022, utilized miniature dCas13X fused to an engineered ADAR2 deaminase domain to edit a mutation in the myosin VI mRNA ($Myo6^{C422Y}$) associated with dominant-inherited deafness (Figure 3c).[76] AAV-delivery of the Cas13X-derived base editor to the inner ear of heterozygous $Myo6^{C422Y/+}$ mice demonstrated an approximately 1.5-fold increase in wildtype Myo6 transcript compared to untreated controls. The authors were then able to show these editing levels correlate to phenotypic improvements to hearing function, which were supported by findings that editing improved hair cell survival and function, both negatively affected by $Myo6^{C422Y}$.

While improvements to editing efficacy would strengthen this method's viability to correct deafness in $Myo6^{C42Y2/+}$ disease models and would likely be necessary in $Myo6^{C422Y/C422Y}$ models, this study is an exciting exhibition of the preclinical potential for programmable RNA-editors. Additionally, applying mini Cas-proteins to deliver fused effectors by AAV and effectively targeting endogenous RNA *in vivo* underscores the potential these platforms possess to direct diverse functions in complex systems. Moreover, as the Li and Shu labs have also previously demonstrated Cas13d–mediated RNA knockdown in the context of other forms of hearing loss, this remains a promising emerging area of application for preclinical programmable RNA-targeting development.[77]

5. Applications for Programmable RBPs in Basic Research

5.1 High-Throughput RNA-Targeting CRISPR Screens

5.1.1 Screening by RNA-Knockdown

The advent of high-throughput CRISPR screens have enabled the parallel study between genotype and phenotype on an unprecedented scale.[78] Typically, these screens rely on DNA-targeting CRISPR-Cas proteins to disrupt a gene specificized by a member of a gRNA library. The library is introduced to a bulk cell population such that each cell receives one gRNA or predefined gRNA mixture, depending on the screen design. A phenotypic output is then induced by some selection (i.e., growth-based after a challenge) which is then tied to gRNA-programmed disruption of a specific gene(s) by high-throughput sequencing methods such as single-cell RNA-seq. These powerful screens have been used to elucidate genetic networks in a myriad of contexts such as cancer biology, immunology, development and more.[79]

The recent description of highly specific RNA-targeting CRISPR platforms has also sparked interest into whether these platforms could also be applied to high-throughput genetic screens. Even before DNA-targeting CRISPR was described RNAi had found success when used in similar forward genetics applications, but suffered from key drawbacks, including murky specificity, and inconsistent knockdown across targets.[80] Due to the advantages platforms like Cas13 have over RNAi in these areas, it stands to reason that RNA-targeting CRISPR-Cas platforms would be viable in these screens.

In order to determine the feasibility and outcomes for Cas13-based CRISPR screens, Wessels and Méndez-Mancilla *et al.* developed a workflow for Cas13 RNA Pertub-seq (named CaRPool-seq) and benchmarked its performance against Cas9-mediated gene knockout and CRISPR interference-based screens.[81] The authors found that CaRPool-seq was more than capable of robust knockdown on the transcript and protein level (65% and 75% average knockdown respectively) from a diverse gRNA library targeting cell surface receptors. By leveraging Cas13d's ability to self-process gRNA arrays, they could also encode multiple gRNA in a single, barcoded RNA molecule, allowing for simulta-

neous multi-target perturbation. In fact, CaRPool-seq outperformed equivalent Cas9-driven gene knockout and CRISPR interference-based methods used in high-throughput screens in identifying cells in a complex population where three genes or transcripts were targeted simultaneously. This performance discrepancy was due to improved detection for the single gRNA array used by Cas13 compared to the three separately expressed gRNA necessary for parallel targeting by Cas9, as each method achieved similar protein-level knockdown for each target.

CaRPool-seq was applied in a proof-of-concept effort to screen genes regulating acute myeloid leukemia differentiation. CaRPool-seq was able to recapitulate differentiation drivers identified in single-gRNA Cas9-based screens, but also novel paired interactors in a screen where two gRNA were encoded in each array. These paired interactors also molecularly corroborated observations of increased therapeutic efficacy when the identified pairs were drugged in tandem. This result clearly established the power of multi-target Cas13-based screens to improve current multi-component genetic discovery assays. Similar methods have also been applied to screening understudied classes of lncRNA demonstrating that these screens can be effective beyond protein coding mRNA.[82] Furthermore, the continued development of more specific and potent RNA-targeting platforms and their deployment should only improve their power and robustness.

Identifying the roles of circular RNA (circRNA) in processes like cancer cell proliferation and development has also emerged as another unique application for high-throughput Cas13 RNA-knockdown screens. circRNA are widely produced in eukaryotic cells by alternative splicing events called "back-splicing," where the 5' end of the first exon is spliced to the 3' end of the terminal exon creating a circRNA often identical in sequence to the linear RNA product but with additional covalent linkage.[83] These circRNA products can have differing effects on the cell compared to their linear counterparts due to their increased stability and altered sequence arrangement, but selective circRNA knockdown has been historically difficult to achieve due to partial recognition of the linear RNA by RNAi leading to nonspecific knockdown of both circular and linear transcript isoforms.[84] Recently, the Chen lab demonstrated that Cas13d can be used to specifically knockdown circRNA when the gRNA spans the back-splicing junction without off-target degradation of the linear splicing product.[85] Furthermore, the specificity also included increased knockdown efficacy almost two-fold compared to similar shRNA-based methods and other studies have shown that Cas13 also effectively targets nuclear localized circRNAs which can be difficult to measure on-target knockdown for with shRNA.[86]

Excitingly, these validation studies have enabled high-throughput Cas13-mediated circRNA knockdown screens to identify circRNA central to processes such as cancer cell proliferation, development, and drug resistance.[85a,86–87] For example, Li *et al.* demonstrated these methods were robust enough to screen for circRNA drivers of tumor proliferation in a mouse xenograft model and specifically identified circFAM120 A, whose knockdown reduced tumor mass by over 50%.[85a] Altogether, these results demonstrate both the impact circRNA can have on cellular phenotype and how Cas13 is uniquely positioned to perturb these sequences for further study.

5.1.2 Screening by RNA-Binding

The understanding that Cas13d–binding to an RNA sequence is sufficient to disrupt endogenous RBPs occupying that site has also incited interest in applying dCas13d in high-throughput screens to identify cryptic RBP-associated sequence elements. In splicing, splicing-regulatory elements (SREs) are central to post-transcription regulation but are relatively understudied compared to genomic regulatory elements.[88] Additionally, SREs are hot therapeutic targets for ASO technologies that also have the ability to compete RBPs for SREs to modify splicing patterns in cells and *in vivo*.[89] However, screening for SREs amenable to targeting by ASOs is expensive and thus often limited in the sequence space and targets they interrogate. Therefore, scalable screens for SREs that impact splicing would be impactful both for basic research and drug discovery campaigns.

Indeed, pooled gRNA library screening to tile dCas13d across a target transcript and thereby measure changes in splicing to identify SREs have been recently reported. Recinos *et al.* validated this strategy on the well-studied SMN2 transcript, whose alternative splicing is a promising therapeutic target for spinal muscular atrophy.[90] Their pooled screen not only identified well-known SREs controlling SMN2 splicing, but also a novel distal intronic element far outside previous screens, that impacted exon 7 inclusion. These results were exciting because intronic distal elements have been reported before but are difficult to identify *de novo* due to their undefined reliance on surrounding sequence context and relative distance from the affected exon.[91] Together these results highlight the potential for dCas13-based screens to identify SREs in broad sequence spaces and will likely enable wider SRE discovery campaigns.

5.2 Uncovering Regulatory Networks from Directed Modifications

The ability to site-specifically install or remove RNA modifications such as m^6A represents a major advance in researchers' ability to investigate how modifications on individual RNA affect both their own regulation and the regulatory networks they participate in. In order to study the role of a specific modification, researchers often identify the writer, eraser, or reader responsible for generating the modification or acting on it and then ablate that protein's expression to change the modification level or response to it. However, since writer, eraser, or reader proteins have many RNA targets any effect on their expression level will have

consequences across the transcriptome complicating any phenotypic analysis.

A 2020 publication from the labs of Chuan He, Dali Han, and Yawei Gao studying m^6A modifications on chromosome-associated RNA represents a prime example of the experiments made possible by programmable modifiers.[92] In this study, the authors establish a clear relationship between m^6A levels in chromosome-associated regulatory RNA (carRNA) and chromatin accessibility. Experiments in writer and reader knockout cell lines revealed that m^6A levels on carRNA negatively affect their half-life and, simultaneously, chromatin accessibility that decreases overall transcription. While these global trends were compelling, a clear correlation was drawn when dCas13-FTO was directed to remove m^6A modifications from carRNA which recapitulated the trends observed in the knockout cell lines when compared to non-targeting gRNA or dCas13 fused to a demethylase activity deficient FTO mutant. These results build confidence that programmable RNA modifiers are both complementary tools to reinforce specific findings from whole-cell perturbations and also useful discovery tools themselves going forward.

5.3 Specific RNA Tracking and Proximity Labeling

While many robust technologies and techniques exist to image specific RNAs in fixed cells, localization of RNAs using live cell imaging has historically been challenging.[93] Before programmable RBPs, fluorescent aptamers or hairpins with orthogonal protein binding partners such as MS2 or BoxB would have to be genomically knocked-in to the RNA of interest, but it is difficult to confirm that these additional RNA sequences do not perturb the transcript's localization or function.[94] Utilizing RNA-targeting dCas platforms to direct fused fluorescent proteins or protect fluorophore-conjugated gRNA has now been accepted as a viable method to visualize RNA localization in live cells without the need for genetic manipulation.[16b,95] These advances have allowed researchers to record dynamic cellular processes such as stress granule formation, transcription, nuclear export, and even developmental processes *in vivo* models, such as zebrafish embryos .[95–96]

Beyond localization, studying the protein interaction partners of specific RNAs has also been a challenge. While discovering the inverse, RNA targets of RBPs, is readily accomplished using tandem immunoprecipitation and sequencing methods to isolate a protein of interest and identify its bound RNA, isolating RNA and observing short-lived RNA-protein interactions is more difficult.[97] In order to overcome this shortcoming, several groups have turned to the proximity labeling field, which has developed enzymes capable of uncaging probes that can covalently modify nearby proteins for identification by mass spectrometry.[98] These enzymes, such as the peroxidase APEX2, have been used successfully to map protein members of many diverse interaction networks in live cells, including those with timescales similar to RNA and RBP interactions.[99] Han and Zhao *et al.*, under the direction of Chuan He and Alice Ting, were the first to combine APEX2 and programmable RBPs such as dCas13 to direct proximity labeling at an unmodified, endogenous RNA and identify protein interactors.[100]

Their study sought to detect protein interactors to human telomerase RNA (hTR) by binding a dCas13d-APEX2 fusion to the target locus. Additionally, the authors further modified dCas13 with an additional double stranded RBP to stabilize the duplex formed between the gRNA and target thereby increasing sustained localization. Proteomics from labeling facilitated by dCas13d-APEX2 at hTR identified several proteins not previously known to interact with hTR including m^6A demethylase ALKBH5. This result was then validated via immunoprecipitation of ALBH5 and subsequent hRT detection by RT-qPCR. The elucidation of this previously unknown interaction led to the discovery of novel m^6A sites on hTR which likely play an important role in telomerase regulation, further demonstrating the power and versatility of programmable RBP directed proximity labeling. Similar methods utilizing dCas13-APEX2 or other proximity labeling enzymes, such as ligases, have since been applied to map interaction networks related to processes such as lncRNA regulation, splicing, and mRNA regulation in human cells and model organisms, such as *Drosophila*, will likely continue to be an important tool for RNA-centric RBP interaction discovery.[101]

6. Achieving Selective RNA-binding Beyond Proteins

Well before the first programmable PUF proteins were developed, precise RNA-targeting for inhibiting translation and RNA degradation by ASOs and RNAi were the primary methods for transcriptomic targeting.[102] The storied development and recent applications for these technologies as well as steric blocking oligos to impact splicing have been recently reviewed in great detail.[103] Beyond these relatively mature oligo technologies, basic research into transcriptomic regulation has likewise inspired oligo-based technologies with expanded functionalities. While functionalizing nucleic acids can be less straightforward than proteins for applications other than RNA degradation, there have nevertheless been several recent impactful examples with some even excelling in preclinical applications.

Several aforementioned protein effector modalities have RNA-based counterparts primarily composed of a bifunctional molecule which includes a gRNA domain and a recruiting domain that binds an endogenous protein to elicit a function at the bound RNA. These technologies include ADAR-recruiting oligos, such as LEAPER, which direct site-specific A-to-I editing,[104] and long non-coding RNAs based on SINE domains, called SINEUPs, which can be directed to the start codon of a transcript to promote translation.[65b] Recently, we developed "translation-activating RNAs" (taRNAs), bifunctional engineered RNA molecules that base pair to the 5' or 3'

UTR of a target mRNA and recruit key initiation factors, such as eIF3, to promote target translation.[105] We found that taRNAs were capable of restoring homeostatic protein levels of SYNGAP1 in haploinsufficient iPSC-derived neurons from patients with SYNGAP1 deficiency, pointing to potential therapeutic applications of taRNAs. Finally, RNAs that can induce modifications have been developed, such as RESTART, which site-specifically modifies uridine to pseudouridine, allowing premature stop codon readthrough.[106] Holistically, nucleic acid-based technologies are exciting because they are both a relatively mature field from the standpoint of delivery, stability, and safety aspects, but can be leveraged as an innovative approach to redirect varied aspects of RNA regulation.

Simultaneously, great progress has recently been made in both discovering small molecules that can bind specific RNA structures and functionalizing them to broaden their application scope; this has been thoroughly reviewed.[107] Specifically, the success of RIBOTACs, small molecules that recruit endogenous RNaseL to specifically cleave bound RNA, showcases the potential of using small molecules as a scaffold to also programmably direct regulation within cells and *in vivo*.[108] Additionally, linear or macrocyclic peptides may have reemerged as another option for programmable RNA-binding taking inspiration from improved display techniques such as mRNA display and the recent success of RNA-binding small molecules.[109] Indeed, recently a 12-mer peptide that binds to m^6A sites was identified via phage display and shown to be active *in cellulo* and successfully compete with FTO for m^6A site binding, thus increasing global m^6A abundance and limiting proliferation in cancer cell lines.[110] These findings indicate an exciting future for discovering and designing small molecules and peptides which selectively bind RNA-targets.

7. Summary and Outlook

Over the last decade the fields of RNA regulation and epitranscriptomics have flourished, consequently fueling both demand and inspiration for new tools to deepen their study and leverage their capabilities. Here, we have reviewed in detail the recent rapid discovery and development of gRNA-directed programmable RBPs to meet this need and several impactful applications showcasing their ability to enable precise regulatory interrogation and address therapeutic challenges. Engineering campaigns to minimize and increase binding potency and specificity for programmable RBP platforms complement and support a greater understanding of epitranscriptomic regulatory proteins, ultimately forging new frontiers for directable RNA regulators capable of specific targeting even in complex cell types, tissues, and *in vivo*.

Future directions for the programmable RBP field include the expansion of recent proof-of-concept methods to make basic research discoveries and the iterative improvement to programmable RBPs RNA-binding and on-target effects through further RNA regulation discovery and engineering. For example, the continued application and development of high-throughput screens utilizing RNA-targeting CRISPR-Cas could uniquely uncover new multimodal gene networks and splicing signals. Additionally, the potential functionalities and potency of programmable RBPs can also be expanded via incorporation and testing of novel RNA regulatory proteins and further evolution campaigns to improve their activity. Building on this, the advancement of effector modalities beyond nucleases and base editors into preclinical disease models will greatly expand programmable RBP's therapeutic scope. Finally, investigating new ways to evolve and design programmable RBPs, potentially without the use of a gRNA, would broaden available targeting capabilities and potentially increase our understanding of RBP-RNA interactions.

Acknowledgements

This work was supported by the National Institute of Biomedical Imaging and Bioengineering (R01 EB035016) of the National Institutes of Health (NIH), as well as The G. Harold and Leila Y. Mathers Charitable Foundation. R.S. was supported by a National Science Foundation Graduate Research Fellowship (DGE-2022294368). The authors thank Dr. S. Ahmadiantehrani for assistance with preparing this manuscript.

Conflict of Interest

B.C.D. and R.S. holds patents for the CIRTS technology. B.C.D. and Y.C. holds patents for the taRNA technology. B.C.D. is a founder and holds equity in Tornado Bio, Inc. The remaining authors declare no competing interests related to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

[1] J. S. Mattick, *Nat. Rev. Genet.* **2004**, *5*, 316–323.
[2] D. D. Licatalosi, R. B. Darnell, *Nat. Rev. Genet.* **2010**, *11*, 75–87.
[3] a) D. P. Bartel, *Cell* **2009**, *136*, 215–233; b) L. Statello, C. J. Guo, L. L. Chen, M. Huarte, *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118.
[4] I. A. Roundtree, M. E. Evans, T. Pan, C. He, *Cell* **2017**, *169*, 1187–1200.
[5] a) I. Barbieri, T. Kouzarides, *Nat. Rev. Cancer* **2020**, *20*, 303–322; b) M. Frye, B. T. Harada, M. Behm, C. He, *Science* **2018**, *361*, 1346–1349.

[6] S. He, E. Valkov, S. Cheloufi, J. Murn, *Nat. Rev. Genet.* **2023**, *24*, 276–294.
[7] a) X. Wang, Z. Lu, A. Gomez, G. C. Hon, Y. Yue, D. Han, Y. Fu, M. Parisien, Q. Dai, G. Jia, B. Ren, T. Pan, C. He, *Nature* **2014**, *505*, 117–120; b) X. Wang, B. S. Zhao, I. A. Roundtree, Z. Lu, D. Han, H. Ma, X. Weng, K. Chen, H. Shi, C. He, *Cell* **2015**, *161*, 1388–1399.
[8] M. W. Hentze, A. Castello, T. Schwarzl, T. Preiss, *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 327–341.
[9] M. Corley, M. C. Burns, G. W. Yeo, *Mol. Cell* **2020**, *78*, 9–29.
[10] J. P. Mackay, J. Font, D. J. Segal, *Nat. Struct. Mol. Biol.* **2011**, *18*, 256–261.
[11] a) D. Carroll, *Genetics* **2011**, *188*, 773–782; b) J. K. Joung, J. D. Sander, *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 49–55.
[12] X. Wang, J. McLachlan, P. D. Zamore, T. M. Hall, *Cell* **2002**, *110*, 501–512.
[13] a) C. G. Cheong, T. M. Hall, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13635–13639; b) Y. Y. Zhao, M. W. Mao, W. J. Zhang, J. Wang, H. T. Li, Y. Yang, Z. Wang, J. W. Wu, *Nucleic Acids Res.* **2018**, *46*, 4771–4782.
[14] S. Coquille, A. Filipovska, T. Chia, L. Rajappa, J. P. Lingford, M. F. Razif, S. Thore, O. Rackham, *Nat. Commun.* **2014**, *5*, 5729.
[15] G. J. Knott, J. A. Doudna, *Science* **2018**, *361*, 866–869.
[16] a) M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, *337*, 816–821; b) M. R. O'Connell, B. L. Oakes, S. H. Sternberg, A. East-Seletsky, M. Kaplan, J. A. Doudna, *Nature* **2014**, *516*, 263–266.
[17] a) O. O. Abudayyeh, J. S. Gootenberg, S. Konermann, J. Joung, I. M. Slaymaker, D. B. Cox, S. Shmakov, K. S. Makarova, E. Semenova, L. Minakhin, K. Severinov, A. Regev, E. S. Lander, E. V. Koonin, F. Zhang, *Science* **2016**, *353*, aaf5573; b) A. A. Smargon, D. B. T. Cox, N. K. Pyzocha, K. Zheng, I. M. Slaymaker, J. S. Gootenberg, O. A. Abudayyeh, P. Essletzbichler, S. Shmakov, K. S. Makarova, E. V. Koonin, F. Zhang, *Mol Cell* **2017**, *65*, 618–630; c) S. Konermann, P. Lotfy, N. J. Brideau, J. Oki, M. N. Shokhirev, P. D. Hsu, *Cell* **2018**, *173*, 665–676.
[18] O. O. Abudayyeh, J. S. Gootenberg, P. Essletzbichler, S. Han, J. Joung, J. J. Belanto, V. Verdine, D. B. T. Cox, M. J. Kellner, A. Regev, E. S. Lander, D. F. Voytas, A. Y. Ting, F. Zhang, *Nature* **2017**, *550*, 280–284.
[19] a) D. B. T. Cox, J. S. Gootenberg, O. O. Abudayyeh, B. Franklin, M. J. Kellner, J. Joung, F. Zhang, *Science* **2017**, *358*, 1019–1027; b) S. Rauch, C. He, B. C. Dickinson, *J. Am. Chem. Soc.* **2018**, *140*, 11974–11981; c) M. Du, N. Jillette, J. J. Zhu, S. Li, A. W. Cheng, *Nat. Commun.* **2020**, *11*, 2973.
[20] S. Rauch, E. He, M. Srienc, H. Zhou, Z. Zhang, B. C. Dickinson, *Cell* **2019**, *178*, 122–134.
[21] a) G. J. Knott, A. East-Seletsky, J. C. Cofsky, J. M. Holton, E. Charles, M. R. O'Connell, J. A. Doudna, *Nat. Struct. Mol. Biol.* **2017**, *24*, 825–833; b) C. Zhang, S. Konermann, N. J. Brideau, P. Lotfy, X. Wu, S. J. Novick, T. Strutzenberg, P. R. Griffin, P. D. Hsu, D. Lyumkis, *Cell* **2018**, *175*, 212–223; c) I. M. Slaymaker, P. Mesa, M. J. Kellner, S. Kannan, E. Brignole, J. Koob, P. R. Feliciano, S. Stella, O. O. Abudayyeh, J. S. Gootenberg, J. Strecker, G. Montoya, F. Zhang, *Cell Rep* **2019**, *26*, 3741–3751; d) M. R. O'Connell, *J. Mol. Biol.* **2019**, *431*, 66–87.
[22] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583–589.
[23] F. Zhao, T. Zhang, X. Sun, X. Zhang, L. Chen, H. Wang, J. Li, P. Fan, L. Lai, T. Sui, Z. Li, *Nat. Commun.* **2023**, *14*, 5545.
[24] a) C. Xu, Y. Zhou, Q. Xiao, B. He, G. Geng, Z. Wang, B. Cao, X. Dong, W. Bai, Y. Wang, X. Wang, D. Zhou, T. Yuan, X. Huo, J. Lai, H. Yang, *Nat. Methods* **2021**, *18*, 499–506; b) S. Kannan, H. Altae-Tran, X. Jin, V. J. Madigan, R. Oshiro, K. S. Makarova, E. V. Koonin, F. Zhang, *Nat. Biotechnol.* **2022**, *40*, 194–197.
[25] R. Nakagawa, S. Kannan, H. Altae-Tran, S. N. Takeda, A. Tomita, H. Hirano, T. Kusakizako, T. Nishizawa, K. Yamashita, F. Zhang, H. Nishimasu, O. Nureki, *Mol Cell* **2022**, *82*, 3178–3192.
[26] a) H. H. Wessels, A. Mendez-Mancilla, X. Guo, M. Legut, Z. Daniloski, N. E. Sanjana, *Nat. Biotechnol.* **2020**, *38*, 722–727; b) J. Wei, P. Lotfy, K. Faizi, S. Baungaard, E. Gibson, E. Wang, H. Slabodkin, E. Kinnaman, S. Chandrasekaran, H. Kitano, M. G. Durrant, C. V. Duffy, A. Pawluk, P. D. Hsu, S. Konermann, *Cell Syst* **2023**, *14*, 1087–1102; c) X. Cheng, Z. Li, R. Shan, Z. Li, S. Wang, W. Zhao, H. Zhang, L. Chao, J. Peng, T. Fei, W. Li, *Nat. Commun.* **2023**, *14*, 752; d) H. H. Wessels, A. Stirn, A. Mendez-Mancilla, E. J. Kim, S. K. Hart, D. A. Knowles, N. E. Sanjana, *Nat Biotechnol* **2023** doi:10.1038/s41587-023-01830-8; e) H. C. Kuo, J. Prupes, C. W. Chou, I. J. Finkelstein, *Nat. Commun.* **2024**, *15*, 498.
[27] a) X. Lin, Y. Liu, A. Chemparathy, T. Pande, M. La Russa, L. S. Qi, *Cell Rep Med* **2021**, *2*, 100245; b) X. Guo, J. A. Rahman, H. H. Wessels, A. Mendez-Mancilla, D. Haro, X. Chen, N. E. Sanjana, *Cell Genom* **2021**, *1*.
[28] M. D. Schertzer, A. Stirn, K. Isaev, L. Pereira, A. Das, C. Harbison, S. H. Park, H. H. Wessels, N. E. Sanjana, D. A. Knowles, *bioRxiv* **2023**.
[29] a) S. Bandaru, M. H. Tsuji, Y. Shimizu, K. Usami, S. Lee, N. K. Takei, K. Yoshitome, Y. Nishimura, T. Otsuki, T. Ito, *Sci. Rep.* **2020**, *10*, 11610; b) A. Krohannon, M. Srivastava, S. Rauch, R. Srivastava, B. C. Dickinson, S. C. Janga, *BMC Genomics* **2022**, *23*, 172; c) S. Mantena, P. P. Pillai, B. A. Petros, N. L. Welch, C. Myhrvold, P. C. Sabeti, H. C. Metsky, *bioRxiv* **2023**; d) W. Hu, A. Kumar, S. Qi, T. Sadras, J. M. L. Casan, D. Ma, L. M. Brown, M. Haber, I. Voskoboinik, J. A. Trapani, P. G. Ekert, M. Fareh, *bioRxiv* **2022**.
[30] a) B. Zetsche, S. E. Volz, F. Zhang, *Nat. Biotechnol.* **2015**, *33*, 139–142; b) Y. Gao, X. Xiong, S. Wong, E. J. Charles, W. A. Lim, L. S. Qi, *Nat. Methods* **2016**, *13*, 1043–1049.
[31] Y. Ding, C. Tous, J. Choi, J. Chen, W. W. Wong, *bioRxiv* **2023**.
[32] a) Y. Xu, N. Tian, H. Shi, C. Zhou, Y. Wang, F. S. Liang, *J. Am. Chem. Soc.* **2023**, *145*, 5561–5569; b) S. Rauch, K. A. Jones, B. C. Dickinson, *ACS Cent. Sci.* **2020**, *6*, 1987–1996.
[33] M. P. Terns, *Mol. Cell* **2018**, *72*, 404–412.
[34] a) G. Tamulaitis, M. Kazlauskiene, E. Manakova, C. Venclovas, A. O. Nwokeoji, M. J. Dickman, P. Horvath, V. Siksnys, *Mol. Cell* **2014**, *56*, 506–517; b) L. You, J. Ma, J. Wang, D. Artamonova, M. Wang, L. Liu, H. Xiang, K. Severinov, X. Zhang, Y. Wang, *Cell* **2019**, *176*, 239–253 ; c) N. Jia, C. Y. Mo, C. Wang, E. T. Eng, L. A. Marraffini, D. J. Patel, *Mol Cell* **2019**, *75*, 944–956 ; d) M. Guo, K. Zhang, Y. Zhu, G. D. Pintilie, X. Guan, S. Li, M. F. Schmid, Z. Ma, W. Chiu, Z. Huang, *Cell Res* **2019**, *29*, 305–312.
[35] a) R. H. Staals, Y. Zhu, D. W. Taylor, J. E. Kornfeld, K. Sharma, A. Barendregt, J. J. Koehorst, M. Vlot, N. Neupane, K. Varossieau, K. Sakamoto, T. Suzuki, N. Dohmae, S. Yokoyama, P. J. Schaap, H. Urlaub, A. J. Heck, E. Nogales, J. A. Doudna, A. Shinkai, J. van der Oost, *Mol. Cell* **2014**, *56*, 518–530; b) I.

Mogila, M. Kazlauskiene, S. Valinskyte, G. Tamulaitiene, G. Tamulaitis, V. Siksnys, *Cell Rep* **2019**, *26*, 2753–2765 ; c) T. Y. Liu, A. T. Iavarone, J. A. Doudna, *PLoS One* **2017**, *12*, e0170552.

[36] D. Colognori, M. Trinidad, J. A. Doudna, *Nat. Biotechnol.* **2023**, *41*, 1256–1264.

[37] a) A. Ozcan, R. Krajeski, E. Ioannidi, B. Lee, A. Gardner, K. S. Makarova, E. V. Koonin, O. O. Abudayyeh, J. S. Gootenberg, *Nature* **2021**, *597*, 720–725; b) S. P. B. van Beljouw, A. C. Haagsma, A. Rodriguez-Molina, D. F. van den Berg, J. N. A. Vink, S. J. J. Brouns, *Science* **2021**, *373*, 1349–1353.

[38] a) K. Kato, W. Zhou, S. Okazaki, Y. Isayama, T. Nishizawa, J. S. Gootenberg, O. O. Abudayyeh, H. Nishimasu, *Cell* **2022**, *185*, 2324–2337 ; b) G. Yu, X. Wang, Y. Zhang, Q. An, Y. Wen, X. Li, H. Yin, Z. Deng, H. Zhang, *Nat Microbiol* **2022**, *7*, 2078–2088; c) Y. Huo, H. Zhao, Q. Dong, T. Jiang, *Nat Microbiol* **2023**, *8*, 522–532; d) H. N. Goswami, J. Rai, A. Das, H. Li, *eLife* **2022**, *11*, e81678.

[39] a) C. Hu, S. P. B. van Beljouw, K. H. Nam, G. Schuler, F. Ding, Y. Cui, A. Rodriguez-Molina, A. C. Haagsma, M. Valk, M. Pabst, S. J. J. Brouns, A. Ke, *Science* **2022**, *377*, 1278–1285; b) X. Wang, G. Yu, Y. Wen, Q. An, X. Li, F. Liao, C. Lian, K. Zhang, H. Yin, Y. Wei, Z. Deng, H. Zhang, *Nucleic Acids Res.* **2022**, *50*, 12913–12923.

[40] a) J. Strecker, F. E. Demircioglu, D. Li, G. Faure, M. E. Wilkinson, J. S. Gootenberg, O. O. Abudayyeh, H. Nishimasu, R. K. Macrae, F. Zhang, *Science* **2022**, *378*, 874–881; b) K. Kato, S. Okazaki, C. Schmitt-Ulms, K. Jiang, W. Zhou, J. Ishikawa, Y. Isayama, S. Adachi, T. Nishizawa, K. S. Makarova, E. V. Koonin, O. O. Abudayyeh, J. S. Gootenberg, H. Nishimasu, *Science* **2022**, *378*, 882–889; c) Q. He, X. Lei, Y. Liu, X. Wang, N. Ji, H. Yin, H. Wang, H. Zhang, G. Yu, *ChemBioChem* **2023**, *24*, e202300401.

[41] L. Liu, X. Li, J. Wang, M. Wang, P. Chen, M. Yin, J. Li, G. Sheng, Y. Wang, *Cell* **2017**, *168*, 121–134 .

[42] a) Y. Ai, D. Liang, J. E. Wilusz, *Nucleic Acids Res.* **2022**, *50*, e65; b) P. Shi, M. R. Murphy, A. O. Aparicio, J. S. Kesner, Z. Fang, Z. Chen, A. Trehan, Y. Guo, X. Wu, *Commun. Biol.* **2023**, *6*, 334; c) Z. Li, Z. Li, X. Cheng, S. Wang, X. Wang, S. Ma, Z. Lu, H. Zhang, W. Zhao, Z. Chen, Y. Yao, C. Zhang, L. Chao, W. Li, T. Fei, *Nat Biomed Eng* **2023**doi:10.1038/s41551-023-01109-y.

[43] J. F. Bot, Z. Zhao, D. Kammeron, P. Shang, N. Geijsen, *bioRxiv* **2023**.

[44] C. P. Kelley, M. C. Haerle, E. T. Wang, *Cell Rep.* **2022**, *40*, 111226.

[45] H. Tong, J. Huang, Q. Xiao, B. He, X. Dong, Y. Liu, X. Yang, D. Han, Z. Wang, X. Wang, W. Ying, R. Zhang, Y. Wei, C. Xu, Y. Zhou, Y. Li, M. Cai, Q. Wang, M. Xue, G. Li, K. Fang, H. Zhang, H. Yang, *Nat. Biotechnol.* **2023**, *41*, 108–119.

[46] X. Deng, E. Osikpa, J. Yang, S. J. Oladeji, J. Smith, X. Gao, Y. Gao, *Nat. Commun.* **2023**, *14*, 5845.

[47] a) C. Wilson, P. J. Chen, Z. Miao, D. R. Liu, *Nat. Biotechnol.* **2020**, *38*, 1431–1440; b) X. M. Liu, J. Zhou, Y. Mao, Q. Ji, S. B. Qian, *Nat. Chem. Biol.* **2019**, *15*, 865–871; c) J. Li, Z. Chen, F. Chen, G. Xie, Y. Ling, Y. Peng, Y. Lin, N. Luo, C. M. Chiang, H. Wang, *Nucleic Acids Res.* **2020**, *48*, 5684–5694; d) Z. Xia, M. Tang, J. Ma, H. Zhang, R. C. Gimple, B. C. Prager, H. Tang, C. Sun, F. Liu, P. Lin, Y. Mei, R. Du, J. N. Rich, Q. Xie, *Nucleic Acids Res.* **2021**, *49*, 7361–7374.

[48] H. Shi, Y. Xu, N. Tian, M. Yang, F. S. Liang, *Nat. Commun.* **2022**, *13*, 1958.

[49] Y. Xu, Y. Wang, F. S. Liang, *Angew. Chem. Int. Ed. Engl.* **2023**, *62*, e202309291.

[50] D. Wiener, S. Schwartz, *Nat. Rev. Genet.* **2021**, *22*, 119–131.

[51] a) S. Xie, H. Jin, F. Yang, H. Zheng, Y. Chang, Y. Liao, Y. Zhang, T. Zhou, Y. Li, *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 19592–19597; b) Y. Wu, Z. Chen, G. Xie, H. Zhang, Z. Wang, J. Zhou, F. Chen, J. Li, L. Chen, H. Niu, H. Wang, *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2119038119.

[52] a) A. R. Kornblihtt, I. E. Schor, M. Allo, G. Dujardin, E. Petrillo, M. J. Munoz, *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 153–165; b) E. Eisenberg, E. Y. Levanon, *Nat. Rev. Genet.* **2018**, *19*, 473–490; c) B. Tian, J. L. Manley, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 18–30.

[53] Y. Wang, C. G. Cheong, T. M. Hall, Z. Wang, *Nat. Methods* **2009**, *6*, 825–830.

[54] J. C. Schmok, M. Jain, L. A. Street, A. T. Tankka, D. Schafer, H. L. Her, S. Elmsaouri, M. L. Gosztyla, E. A. Boyle, P. Jagannatha, E. C. Luo, E. J. Kwon, M. Jovanovic, G. W. Yeo, *Nat Biotechnol* **2024** doi:10.1038/s41587-023-02014-0.

[55] S. Tian, B. Zhang, Y. He, Z. Sun, J. Li, Y. Li, H. Yi, Y. Zhao, X. Zou, Y. Li, H. Cui, L. Fang, X. Gao, Y. Hu, W. Chen, *Nucleic Acids Res.* **2022**, *50*, e26.

[56] K. M. Anderson, P. Poosala, S. R. Lindley, D. M. Anderson, *bioRxiv* **2019**.

[57] J. Borrajo, K. Javanmardi, J. Griffin, S. J. St Martin, D. Yao, K. Hill, P. C. Blainey, B. Al-Shayeb, *bioRxiv* **2023**.

[58] K. Nishikura, *Annu. Rev. Biochem.* **2010**, *79*, 321–349.

[59] O. O. Abudayyeh, J. S. Gootenberg, B. Franklin, J. Koob, M. J. Kellner, A. Ladha, J. Joung, P. Kirchgatterer, D. B. T. Cox, F. Zhang, *Science* **2019**, *365*, 382–386.

[60] X. Huang, J. Lv, Y. Li, S. Mao, Z. Li, Z. Jing, Y. Sun, X. Zhang, S. Shen, X. Wang, M. Di, J. Ge, X. Huang, E. Zuo, T. Chi, *EMBO J.* **2020**, *39*, e104741.

[61] A. C. Goldstrohm, M. Wickens, *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 337–344.

[62] a) A. Cooke, A. Prigge, L. Opperman, M. Wickens, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 15870–15875; b) Z. Abil, C. A. Denard, H. Zhao, *J Biol Eng* **2014**, *8*, 7; c) E. C. Luo, J. L. Nathanson, F. E. Tan, J. L. Schwartz, J. C. Schmok, A. Shankar, S. Markmiller, B. A. Yee, S. Sathe, G. A. Pratt, D. B. Scaletta, Y. Ha, D. E. Hill, S. Aigner, G. W. Yeo, *Nat. Struct. Mol. Biol.* **2020**, *27*, 989–1000.

[63] a) J. Cao, M. Arha, C. Sudrik, A. Mukherjee, X. Wu, R. S. Kane, *Nucleic Acids Res.* **2015**, *43*, 4353–4362; b) A. Apostolopoulos, H. Tsuiji, Y. Shichino, S. Iwasaki, *bioRxiv* **2023**; c) J. Cao, M. Arha, C. Sudrik, D. V. Schaffer, R. S. Kane, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 4900–4904..

[64] a) K. P. Adamala, D. A. Martin-Alarcon, E. S. Boyden, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2579–2588; b) Z. T. Campbell, C. T. Valley, M. Wickens, *Nat. Struct. Mol. Biol.* **2014**, *21*, 732–738

[65] a) C. Cao, A. Li, C. Xu, B. Wu, J. Liu, Y. Liu, *Nucleic Acids Res.* **2023**, *51*, e33; b) S. Zucchelli, F. Fasolo, R. Russo, L. Cimatti, L. Patrucco, H. Takahashi, M. H. Jones, C. Santoro, D. Sblattero, D. Cotella, F. Persichetti, P. Carninci, S. Gustincich, *Front Cell Neurosci* **2015**, *9*, 174.

[66] M. Li, D. Li, L. Lin, P. Wang, W. Zhao, *ACS Synth. Biol.* **2023**, *12*, 2827–2833.

[67] K. Reynaud, A. M. McGeachy, D. Noble, Z. A. Meacham, N. T. Ingolia, *Nat. Struct. Mol. Biol.* **2023**, *30*, 740–752.

[68] a) C. A. Ross, E. H. Aylward, E. J. Wild, D. R. Langbehn, J. D. Long, J. H. Warner, R. I. Scahill, B. R. Leavitt, J. C. Stout, J. S. Paulsen, R. Reilmann, P. G. Unschuld, A. Wexler, R. L. Margolis, S. J. Tabrizi, *Nat. Rev. Neurol.* **2014**, *10*, 204–216; b) P. D. Jacobs, E. Molloy, *Am J Manag Care* **2017**, *23*, 372–377.

[69] a) R. Batra, D. A. Nelles, E. Pirie, S. M. Blue, R. J. Marina, H. Wang, I. A. Chaim, J. D. Thomas, N. Zhang, V. Nguyen, S. Aigner, S. Markmiller, G. Xia, K. D. Corbett, M. S. Swanson, G. W. Yeo, *Cell* **2017**, *170*, 899–912 ; b) R. Batra, D. A. Nelles, D. M. Roth, F. Krach, C. A. Nutter, T. Tadokoro, J. D. Thomas, L. J. Sznajder, S. M. Blue, H. L. Gutierrez, P. Liu, S. Aigner, O. Platoshyn, A. Miyanohara, M. Marsala, M. S. Swanson, G. W. Yeo, *Nat Biomed Eng* **2021**, *5*, 157–168.

[70] K. H. Morelli, Q. Wu, M. L. Gosztyla, H. Liu, M. Yao, C. Zhang, J. Chen, R. J. Marina, K. Lee, K. L. Jones, M. Y. Huang, A. Li, C. Smith-Geater, L. M. Thompson, W. Duan, G. W. Yeo, *Nat. Neurosci.* **2023**, *26*, 27–38.

[71] S. J. Tabrizi, C. Estevez-Fraga, W. M. C. van Roon-Mom, M. D. Flower, R. I. Scahill, E. J. Wild, I. Munoz-Sanjuan, C. Sampaio, A. E. Rosser, B. R. Leavitt, *Lancet Neurol.* **2022**, *21*, 645–658.

[72] J. E. Powell, C. K. W. Lim, R. Krishnan, T. X. McCallister, C. Saporito-Magrina, M. A. Zeballos, G. D. McPheron, T. Gaj, *Sci. Adv.* **2022**, *8*, eabk2485.

[73] C. M. Zeballos, H. J. Moore, T. J. Smith, J. E. Powell, N. S. Ahsan, S. Zhang, T. Gaj, *Nat. Commun.* **2023**, *14*, 6492.

[74] Y. Li, J. Xu, X. Guo, Z. Li, L. Cao, S. Liu, Y. Guo, G. Wang, Y. Luo, Z. Zhang, X. Wei, Y. Zhao, T. Liu, X. Wang, H. Xia, M. Kuang, Q. Guo, J. Li, L. Chen, Y. Wang, Q. Li, F. Wang, Q. Liu, F. You, *Genome Biol.* **2023**, *24*, 20.

[75] a) Z. Yan, Y. Yao, L. Li, L. Cai, H. Zhang, S. Zhang, Q. Xiao, X. Wang, E. Zuo, C. Xu, J. Wu, H. Yang, *Mol. Ther. Nucleic Acids* **2023**, *33*, 750–761; b) J. Li, Z. Shen, Y. Liu, Z. Yan, Y. Liu, X. Lin, J. Tang, R. Lv, G. Geng, Z. Q. Xiong, C. Zhou, H. Yang, *Mol. Ther.* **2023**, *31*, 2286–2295.

[76] Q. Xiao, Z. Xu, Y. Xue, C. Xu, L. Han, Y. Liu, F. Wang, R. Zhang, S. Han, X. Wang, G. L. Li, H. Li, H. Yang, Y. Shu, *Sci. Transl. Med.* **2022**, *14*, eabn0449.

[77] Y. Guo, L. Han, S. Han, H. Tang, S. Wang, C. Cui, B. Chen, H. Li, Y. Shu, *Mol. Ther. Nucleic Acids* **2022**, *28*, 643–655.

[78] C. Bock, P. Datlinger, F. Chardon, M. A. Coelho, M. B. Dong, K. A. Lawson, T. Lu, L. Maroc, T. M. Norman, B. Song, G. Stanley, S. Chen, M. Garnett, W. Li, J. Moffat, L. S. Qi, R. S. Shapiro, J. Shendure, J. S. Weissman, X. Zhuang, *Nat Rev Methods Primers* **2022**, *2*, 8.

[79] a) F. M. Behan, F. Iorio, G. Picco, E. Goncalves, C. M. Beaver, G. Migliardi, R. Santos, Y. Rao, F. Sassi, M. Pinnelli, R. Ansari, S. Harper, D. A. Jackson, R. McRae, R. Pooley, P. Wilkinson, D. van der Meer, D. Dow, C. Buser-Doepner, A. Bertotti, L. Trusolino, E. A. Stronach, J. Saez-Rodriguez, K. Yusa, M. J. Garnett, *Nature* **2019**, *568*, 511–516; b) A. Dixit, O. Parnas, B. Li, J. Chen, C. P. Fulco, L. Jerby-Arnon, N. D. Marjanovic, D. Dionne, T. Burks, R. Raychowdhury, B. Adamson, T. M. Norman, E. S. Lander, J. S. Weissman, N. Friedman, A. Regev, *Cell* **2016**, *167*, 1853–1866; c) D. A. Jaitin, A. Weiner, I. Yofe, D. Lara-Astiaso, H. Keren-Shaul, E. David, T. M. Salame, A. Tanay, A. van Oudenaarden, I. Amit, *Cell* **2016**, *167*, 1883–1896; d) X. Jin, S. K. Simmons, A. Guo, A. S. Shetty, M. Ko, L. Nguyen, V. Jokhi, E. Robinson, P. Oyler, N. Curry, G. Deangeli, S. Lodato, J. Z. Levin, A. Regev, F. Zhang, P. Arlotta, *Science* **2020**, *370*, 6520.

[80] a) M. Boutros, J. Ahringer, *Nat. Rev. Genet.* **2008**, *9*, 554–566; b) S. E. Mohr, J. A. Smith, C. E. Shamu, R. A. Neumuller, N. Perrimon, *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 591–600.

[81] H. H. Wessels, A. Mendez-Mancilla, Y. Hao, E. Papalexi, W. M. Mauck, 3rd, L. Lu, J. A. Morris, E. P. Mimitou, P. Smibert, N. E. Sanjana, R. Satija, *Nat. Methods* **2023**, *20*, 86–94.

[82] D. Xu, Y. Cai, L. Tang, X. Han, F. Gao, H. Cao, F. Qi, P. Kapranov, *Sci. Rep.* **2020**, *10*, 1794.

[83] L. S. Kristensen, M. S. Andersen, L. V. W. Stagsted, K. K. Ebbesen, T. B. Hansen, J. Kjems, *Nat. Rev. Genet.* **2019**, *20*, 675–691.

[84] A. T. He, J. Liu, F. Li, B. B. Yang, *Signal Transduct Target Ther* **2021**, *6*, 185.

[85] a) S. Li, X. Li, W. Xue, L. Zhang, L. Z. Yang, S. M. Cao, Y. N. Lei, C. X. Liu, S. K. Guo, L. Shan, M. Wu, X. Tao, J. L. Zhang, X. Gao, J. Zhang, J. Wei, J. Li, L. Yang, L. L. Chen, *Nat. Methods* **2021**, *18*, 51–59; b) S. Li, H. Wu, L. L. Chen, *Nat. Protoc.* **2022**, *17*, 2085–2107.

[86] Y. Zhang, T. M. Nguyen, X. O. Zhang, L. Wang, T. Phan, J. G. Clohessy, P. P. Pandolfi, *Genome Biol.* **2021**, *22*, 41.

[87] a) C. Chen, H. Yu, F. Han, X. Lai, K. Ye, S. Lei, M. Mai, M. Lai, H. Zhang, *Mol. Cancer* **2022**, *21*, 46; b) V. Rengganaten, C. J. Huang, M. L. Wang, Y. Chien, P. H. Tsai, Y. T. Lan, H. T. Ong, S. H. Chiou, K. B. Choo, *BMC Cancer* **2023**, *23*, 1088.

[88] J. Ule, B. J. Blencowe, *Mol. Cell* **2019**, *76*, 329–345.

[89] a) M. Matsuo, *JMA J* **2021**, *4*, 232–240; b) Z. Han, C. Chen, A. Christiansen, S. Ji, Q. Lin, C. Anumonwo, C. Liu, S. C. Leiser, Meena, I. Aznarez, G. Liau, L. L. Isom, *Sci Transl Med* **2020**, *12*.

[90] Y. Recinos, D. Ustianenko, Y. T. Yeh, X. Wang, M. Jacko, L. V. Yesantharao, Q. Wu, C. Zhang, *bioRxiv* **2023**.

[91] a) A. B. Rosenberg, R. P. Patwardhan, J. Shendure, G. Seelig, *Cell* **2015**, *163*, 698–711; b) M. T. Lovci, D. Ghanem, H. Marr, J. Arnold, S. Gee, M. Parra, T. Y. Liang, T. J. Stark, L. T. Gehman, S. Hoon, K. B. Massirer, G. A. Pratt, D. L. Black, J. W. Gray, J. G. Conboy, G. W. Yeo, *Nat. Struct. Mol. Biol.* **2013**, *20*, 1434–1442.

[92] J. Liu, X. Dou, C. Chen, C. Chen, C. Liu, M. M. Xu, S. Zhao, B. Shen, Y. Gao, D. Han, C. He, *Science* **2020**, *367*, 580–586.

[93] P. Le, N. Ahmed, G. W. Yeo, *Nat. Cell Biol.* **2022**, *24*, 815–824.

[94] S. Tyagi, *Nat. Methods* **2009**, *6*, 331–338.

[95] D. A. Nelles, M. Y. Fang, M. R. O'Connell, J. L. Xu, S. J. Markmiller, J. A. Doudna, G. W. Yeo, *Cell* **2016**, *165*, 488–496.

[96] a) H. Wang, M. Nakamura, T. R. Abbott, D. Zhao, K. Luo, C. Yu, C. M. Nguyen, A. Lo, T. P. Daley, M. La Russa, Y. Liu, L. S. Qi, *Science* **2019**, *365*, 1301–1305; b) L. Z. Yang, Y. Wang, S. Q. Li, R. W. Yao, P. F. Luan, H. Wu, G. G. Carmichael, L. L. Chen, *Mol Cell* **2019**, *76*, 981–997 ; c) Y. Huang, B. Q. Gao, Q. Meng, L. Z. Yang, X. K. Ma, H. Wu, Y. H. Pan, L. Yang, D. Li, L. L. Chen, *Genome Biol.* **2023**, *24*, 15.

[97] C. Lin, W. O. Miles, *Nucleic Acids Res.* **2019**, *47*, 5490–5501.

[98] W. Qin, K. F. Cho, P. E. Cavanagh, A. Y. Ting, *Nat. Methods* **2021**, *18*, 133–143.

[99] V. Hung, N. D. Udeshi, S. S. Lam, K. H. Loh, K. J. Cox, K. Pedram, S. A. Carr, A. Y. Ting, *Nat. Protoc.* **2016**, *11*, 456–475.

[100] S. Han, B. S. Zhao, S. A. Myers, S. A. Carr, C. He, A. Y. Ting, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 22068–22079.

[101] a) W. Yi, J. Li, X. Zhu, X. Wang, L. Fan, W. Sun, L. Liao, J. Zhang, X. Li, J. Ye, F. Chen, J. Taipale, K. M. Chan, L. Zhang, J. Yan, *Nat. Methods* **2020**, *17*, 685–688; b) Z. Zhang, W. Sun, T. Shi, P. Lu, M. Zhuang, J. L. Liu, *Nucleic Acids Res.* **2020**, *48*, e52; c) X. Lin, M. A. S. Fonseca, J. J. Breunig, R. I. Corona, K. Lawrenson, *RNA Biol.* **2021**, *18*, 2203–2217; d) Z. Zhang, Y. Zhang, J.-L. Liu, *bioRxiv* **2022**.

[102] Y. Zhu, L. Zhu, X. Wang, H. Jin, *Cell Death Dis.* **2022**, *13*, 644.

[103] a) B. Hu, L. Zhong, Y. Weng, L. Peng, Y. Huang, Y. Zhao, X. J. Liang, *Signal Transduct Target Ther* **2020**, *5*, 101; b) S. T. Crooke, B. F. Baker, R. M. Crooke, X. H. Liang, *Nat. Rev. Drug Discovery* **2021**, *20*, 427–453.

[104] L. Qu, Z. Yi, S. Zhu, C. Wang, Z. Cao, Z. Zhou, P. Yuan, Y. Yu, F. Tian, Z. Liu, Y. Bao, Y. Zhao, W. Wei, *Nat. Biotechnol.* **2019**, *37*, 1059–1069.

[105] Y. Cao, H. Liu, S. S. Lu, K. A. Jones, A. P. Govind, O. Jeyifous, C. Q. Simmons, N. Tabatabaei, W. N. Green, J. L. Holder Jr., S. Tahmasebi, A. L. George Jr., B. C. Dickinson, *Nat. Commun.* **2023**, *14*, 6827.

[106] a) Z. Yi, Y. Zhao, Z. Yi, Y. Zhang, G. Tang, X. Zhang, H. Tang, W. Zhang, Y. Zhao, H. Xu, Y. Nie, X. Sun, L. Xing, L. Dai, P. Yuan, W. Wei, *Genome Biol.* **2023**, *24*, 243; b) H. Adachi, Y. Pan, X. He, J. L. Chen, B. Klein, G. Platenburg, P. Morais, P. Boutz, Y. T. Yu, *Mol Cell* **2023**, *83*, 637–651 ; c) J. Song, L. Dong, H. Sun, N. Luo, Q. Huang, K. Li, X. Shen, Z. Jiang, Z. Lv, L. Peng, M. Zhang, K. Wang, K. Liu, J. Hong, C. Yi, *Mol Cell* **2023**, *83*, 139–155 .

[107] J. L. Childs-Disney, X. Yang, Q. M. R. Gibaut, Y. Tong, R. T. Batey, M. D. Disney, *Nat. Rev. Drug Discovery* **2022**, *21*, 736–762.

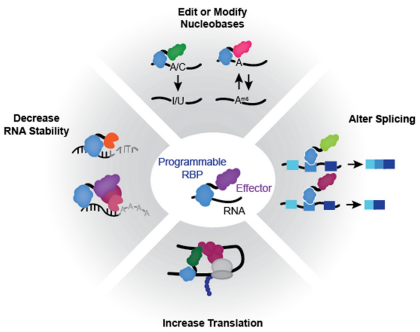
[108] a) M. G. Costales, Y. Matsumoto, S. P. Velagapudi, M. D. Disney, *J. Am. Chem. Soc.* **2018**, *140*, 6741–6744; b) Y. Tong, Y. Lee, X. Liu, J. L. Childs-Disney, B. M. Suresh, R. I. Benhamou, C. Yang, W. Li, M. G. Costales, H. S. Haniff, S. Sievers, D. Abegg, T. Wegner, T. O. Paulisch, E. Lekah, M. Grefe, G. Crynen, M. Van Meter, T. Wang, Q. M. R. Gibaut, J. L. Cleveland, A. Adibekian, F. Glorius, H. Waldmann, M. D. Disney, *Nature* **2023**, *618*, 169–179.

[109] X. Mou, C. K. Kwok, *J. Am. Chem. Soc.* **2023**, *145*, 18693–18697.

[110] R. Rauff, S. M. Abedeera, S. Schmocker, J. Xie, S. C. Abeysirigunawardena, *ChemMedChem* **2023**, *18*, e202200549.

Manuscript received: November 30, 2023
Revised manuscript received: January 15, 2024
Version of record online: ■■, ■■

REVIEW



*R. W. Sinnott, Y. Cao, B. C. Dickinson**