Article

Cell

Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors

Graphical Abstract



Authors

Silvana Konermann, Peter Lotfy, Nicholas J. Brideau, Jennifer Oki, Maxim N. Shokhirev, Patrick D. Hsu

Correspondence

patrick@salk.edu

In Brief

A new family of RNA-targeting CRISPR-Cas systems allows specific modulation of splicing events to reduce pathological tau isoforms in a neuronal model of frontotemporal dementia.

Highlights

- Computational pipeline identifies the RNA-targeting type VI-D CRISPR-Cas family
- Ortholog screen and protein engineering yields the programmable ribonuclease CasRx
- CasRx RNA knockdown exhibits favorable efficiency and specificity relative to RNAi
- Neuronal AAV delivery of dCasRx splice effectors alleviates tau mis-splicing



Article

Silvana Konermann,^{1,2} Peter Lotfy,^{1,2} Nicholas J. Brideau,^{1,2} Jennifer Oki,^{1,2} Maxim N. Shokhirev,³ and Patrick D. Hsu^{1,2,4,*} ¹Laboratory of Molecular and Cell Biology, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA ²Helmsley Center for Genomic Medicine, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA ³Razavi Newman Integrative Genomics and Bioinformatics Core, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA CA 92037, USA

⁴Lead Contact *Correspondence: patrick@salk.edu https://doi.org/10.1016/j.cell.2018.02.033

SUMMARY

Class 2 CRISPR-Cas systems endow microbes with diverse mechanisms for adaptive immunity. Here, we analyzed prokaryotic genome and metagenome sequences to identify an uncharacterized family of RNA-guided, RNA-targeting CRISPR systems that we classify as type VI-D. Biochemical characterization and protein engineering of seven distinct orthologs generated a ribonuclease effector derived from Ruminococcus flavefaciens XPD3002 (CasRx) with robust activity in human cells. CasRx-mediated knockdown exhibits high efficiency and specificity relative to RNA interference across diverse endogenous transcripts. As one of the most compact single-effector Cas enzymes, CasRx can also be flexibly packaged into adeno-associated virus. We target virally encoded, catalytically inactive CasRx to cis elements of pre-mRNA to manipulate alternative splicing, alleviating dysregulated tau isoform ratios in a neuronal model of frontotemporal dementia. Our results present CasRx as a programmable RNA-binding module for efficient targeting of cellular RNA, enabling a general platform for transcriptome engineering and future therapeutic development.

INTRODUCTION

Mapping of transcriptome changes in cellular function and disease has been transformed by technological advances over the last two decades, from microarrays (Schena et al., 1995) to next-generation sequencing and single-cell studies (Shendure et al., 2017). However, interrogating the function of individual transcript dynamics and establishing causal linkages between observed transcriptional changes and cellular phenotype requires the ability to actively control or modulate desired transcripts.

DNA engineering technologies such as CRISPR-Cas9 (Doudna and Charpentier, 2014; Hsu et al., 2014) enable researchers to dissect the function of specific genetic elements or correct disease-causing mutations. However, simple and scalable tools to study and manipulate RNA lag significantly behind their DNA counterparts. Existing RNA interference technologies, which enable cleavage or inhibition of desired transcripts, have significant off-target effects and remain challenging engineering targets due to their key role in endogenous processes (Birmingham et al., 2006; Jackson et al., 2003). As a result, methods for studying the functional role of RNAs directly have remained limited.

One of the key restrictions in RNA engineering has been the lack of RNA-binding domains that can be easily retargeted and introduced into target cells. The MS2 RNA-binding domain, for example, recognizes an invariant 21-nucleotide (nt) RNA sequence (Peabody, 1993), therefore requiring genomic modification to tag a desired transcript. Pumilio homology domains possess modular repeats with each protein module recognizing a separate RNA base, but they can only be targeted to short 8 nt RNA sequences (Cheong and Hall, 2006). While previously characterized type II (Batra et al., 2017; O'Connell et al., 2014) and VI (Abudayyeh et al., 2016; East-Seletsky et al., 2016) CRISPR-Cas systems can be reprogrammed to recognize 20–30 nt RNAs, their large size (~1,200 amino acids [aa]) makes it difficult to package into adeno-associated virus (AAV) for primary cell and *in vivo* delivery.

Reasoning that diverse RNA-targeting CRISPR systems and their associated defense nucleases remain largely unexplored and may harbor advantageous properties, we conducted bioinformatic analysis of prokaryotic genomes to identify sequence signatures of CRISPR-Cas repeat arrays and mine previously uncharacterized, compact Cas ribonucleases that could be developed into RNA targeting tools. We demonstrate that engineered type VI-D CRISPR effectors can be used to efficiently knock down endogenous RNAs in human cells and manipulate alternative splicing, paving the way for RNA targeting applications and further effector domain fusions as part of a transcriptome engineering toolbox.

RESULTS

Computational Identification of a Type VI-like Cas Ribonuclease Family

We first sought to identify previously undetected or uncharacterized RNA-targeting CRISPR-Cas systems by developing a computational pipeline for class 2 CRISPR-Cas loci, which



Figure 1. Type VI CRISPR-Cas13d Is a Family of Single-Effector CRISPR Ribonucleases

(A) Maximum-likelihood phylogenetic tree of Cas13d effectors used in this study, with the full Cas13d CRISPR locus depicted along with conserved HEPN RNase domains. Grey rectangles denote CRISPR direct repeats (DRs) and blue diamonds indicate spacer sequences.

(B) RNA sequencing of a heterologously expressed Cas13d locus from an uncultured *Ruminococcus sp.* sample. Mature gRNAs mapping to the CRISPR array indicate a processed 30 nt DR and a variable spacer length from 14–26 nt. Co-fold analysis of direct repeat truncation indicates a strong hairpin structure.

(C) Purified *E. siraeum* Cas13d and catalytically dead Cas13d (dCas13d) protein are each sufficient to process a guide array into its two component gRNAs. Addition of EDTA does not impair gRNA processing. "d," dCas13d(R295A, H300A, R849A, H854A).

See also Figures S1-S3 and Table S1.

identify new class 2 Cas effectors, we required candidate proteins to be >750 residues in length and within 5 proteincoding genes of the repeat array, as large proteins closely associated with CRISPR repeats are key characteristics of known single effectors. The resulting proteins were classified into 408 putative protein families using single-linkage hierarchical clustering based on homology.

To discard protein clusters that reside in close proximity to CRISPR arrays due to chance or overall abundance in the genome, we next identified additional homologous proteins to each cluster from the NCBI non-redundant protein

require only a single nuclease for CRISPR interference such as Cas9, Cas12a (formerly Cpf1), or Cas13a (formerly C2c2) (Makarova et al., 2015; Shmakov et al., 2015). To improve upon previous strategies for bioinformatic mining of CRISPR systems, which focus on discovering sets of conserved Cas genes involved in spacer acquisition (Shmakov et al., 2015), we defined the minimal requirements for a CRISPR locus to be the presence of a CRISPR array as a search anchor, we first obtained all prokaryotic genome assemblies and scaffolds from the NCBI WGS database and adapted algorithms for *de novo* CRISPR array detection (Bland et al., 2007; Edgar, 2007; Grissa et al., 2007) to identify 21,175 putative CRISPR repeat arrays (Figure S1A).

Up to 20 kb of genomic DNA sequence flanking each CRISPR array was extracted to identify predicted protein-coding genes in the immediate vicinity. Candidate loci containing signature genes of known class 1 and class 2 CRISPR-Cas systems such as Cas3 or Cas9 were excluded from further analysis, except for Cas12a and Cas13a to judge the ability of our pipeline to detect and cluster these known class 2 effector families. To database and determined their proximity to a CRISPR array. Reasoning that true Cas genes would have a high co-occurrence rate with CRISPR repeats, >70% of the proteins for each expanded cluster were required to exist within 20 kb of a CRISPR repeat. These remaining protein families were analyzed for nuclease domains and motifs.

Among the candidates, which include the recently described Cas13b system (Smargon et al., 2017), we identified a family of uncharacterized putative class 2 CRISPR-Cas systems encoding a candidate CRISPR-associated ribonuclease containing 2 predicted HEPN ribonuclease motifs (Anantharaman et al., 2013) (Figure 1A). Importantly, they are among the smallest class 2 CRISPR-Cas13 superfamily is exemplified by sequence-divergent, single-effector signature nucleases and the presence of two HEPN domains. Other than these two *RxxxxH* HEPN motifs (Figure S2A), our candidate effectors have no significant sequence similarity to previously described Cas13 enzymes, so we designated this family of putative CRISPR ribonucleases as type VI Cas13d or type VI-D (Figure S2B).

CRISPR-Cas13d systems are derived from gut-resident microbes, so we sought to expand the Cas13d family via alignment to metagenomic contigs from recent large-scale microbiome sequencing efforts. Comparison of Cas13d proteins against public metagenome sequences without predicted open reading frames (ORFs) identified additional full-length systems as well as multiple effector and array fragments that cluster in several distinct branches (Figure S1B). To generate full-length Cas13d ortholog proteins and loci from the different branches of the Cas13d protein family, we obtained genomic DNA samples from associated assemblies and performed targeted Sanger sequencing to fill in gaps due to incomplete sequencing coverage, such as for the metagenomic ortholog "Anaerobic digester metagenome" (Adm) (Treu et al., 2016).

Cas13d CRISPR loci are largely clustered within benign, Gram-positive gut bacteria of the genus Ruminococcus and exhibit a surprising diversity of CRISPR locus architectures (Figure 1A). With the exception of the metagenomic AdmCas13d system, Cas13d systems lack the key spacer acquisition protein Cas1 (Yosef et al., 2012) within their CRISPR locus, highlighting the utility of a class 2 CRISPR discovery pipeline without Cas1 or Cas2 gene requirements. Cas13d direct repeats (DRs) are highly conserved in length and predicted secondary structure (Figure S2C), with a 36 nt length, an 8-10 nt stem with A/U-rich loop, and a 5'-AAAAC motif at the 3' end of the direct repeat (Figure S2D). This conserved 5'-AAAAC motif has been previously shown to be specifically recognized by a type II Cas1/2 spacer acquisition complex (Wright and Doudna, 2016). In fact, Cas1 can be found in relative proximity to some Cas13d systems (within 10-30 kb for P1E0 and Rfx) while the remaining Cas13d-containing bacteria contain Cas1 elsewhere in their genomes, likely as part of another CRISPR locus.

CRISPR-Cas13d Possesses Dual RNase Activities

To assess if the Cas13d repeat array is transcribed and processed into mature CRISPR guide RNAs (gRNA) as predicted (Deltcheva et al., 2011), we cloned the Cas13d CRISPR locus from an uncultured *Ruminococcus sp.* sample (Ur) into a bacterial expression plasmid. CRISPR systems tend to form self-contained operons with the necessary regulatory sequences for independent expression, facilitating heterologous expression in *E. coli* (Gasiunas et al., 2012). RNA sequencing (Heidrich et al., 2015) revealed processing of the array into ~52 nt mature gRNAs, with a 30 nt 5' direct repeat followed by a variable 3' spacer that ranged from 14–26 nt in length (Figure 1B).

To characterize Cas13d properties *in vitro*, we next purified *Eubacterium siraeum* Cas13d protein (EsCas13d) based on its robust recombinant expression in *E. coli* (Figure S3) and found that EsCas13d was solely sufficient to process its matching CRISPR array into constituent guides without additional helper ribonucleases (Figure 1C; Table S1), a property shared by some class 2 CRISPR-Cas systems (East-Seletsky et al., 2016; Fonfara et al., 2016; Smargon et al., 2017). Furthermore, inactivating the positively charged catalytic residues of the HEPN motifs (Anantharaman et al., 2013) (dCas13d: R295A, H300A, R849A, H854A) did not affect array processing, indicating a distinct RNase activity dictating guide RNA biogenesis analogous to Cas13a (East-Seletsky et al., 2016; Liu et al., 2017).

Cas effector proteins typically form a binary complex with mature gRNA to generate an RNA-guided surveillance ribonucleoprotein capable of cleaving foreign nucleic acids for immune defense (van der Oost et al., 2014). To assess if Cas13d has programmable RNA targeting activity as predicted by the presence of two HEPN motifs, EsCas13d protein was paired with an array or a mature gRNA along with a cognate *in vitro*-transcribed target. Based on the RNA sequencing results, we selected a mature gRNA containing a 30 nt direct repeat and an intermediate spacer length of 22 nt.

Cas13d was able to efficiently cleave the complementary target single-stranded RNA (ssRNA) with both the unprocessed array and mature gRNA in a guide-sequencedependent manner, while non-matching spacer sequences abolished Cas13d activity (Figure 2A). Substitution with dCas13d or the addition of EDTA to the cleavage reaction also abolished guide-dependent RNA targeting, indicating that Cas13d targeting is HEPN- and Mg²⁺-dependent (Figure 2B). To determine the minimal spacer length for efficient Cas13d targeting, we next generated a series of spacer truncations ranging from the unprocessed 30 nt length down to 10 nt (Figure S4A). Cleavage activity dropped significantly below a 21 nt spacer length, confirming the choice of a 22 nt spacer (Figure S4B).

RNA-targeting class 2 CRISPR systems have been proposed to act as sensors of foreign RNAs (Abudayyeh et al., 2016; East-Seletsky et al., 2016), where general RNase activity of the effector nuclease is triggered by a guide-matching target. To assay for a similar property in Cas13d, RNase activity of the binary EsCas13d:gRNA complex was monitored in the presence of a matching RNA target. We observed that EsCas13d can be activated by target RNA to cleave bystander RNA targets (Figure 2C), albeit inefficiently relative to its activity on the complementary ssRNA target. Bystander cleavage is guide sequence- and HEPN-dependent, as the presence of non-matching bystander target alone was insufficient to induce cleavage while substitution of dCas13d or addition of EDTA abolished activity. These results suggest that bystander RNase activity may be a general property of RNA-targeting class 2 systems in CRISPR adaptive bacterial immunity (Figure 2D).

To assess the generalizability of Cas13d reprogramming, we first generated 12 guides tiling a complementary RNA target and observed efficient cleavage in all cases (Figure 3A). Cas13d was unable to cleave a single-stranded DNA (ssDNA) (Figure S4C) or double-stranded DNA (dsDNA) (Figure S4D) version of the ssRNA target, indicating that Cas13d is an RNA-specific nuclease. Further, RNA target cleavage did not appear to depend on the protospacer flanking sequence (PFS) (Figure 3A) in contrast to other RNA-targeting class 2 systems, which require a 3'-H (Abudayyeh et al., 2016) or a double-sided, DR-proximal 5'-D and 3'-NAN or NNA (Smargon et al., 2017). Although we initially observed a slight bias against an adenine PFS (Figure S4E), varying the target PFS base with a constant guide sequence resulted in no significant differences (p = 0.768) in targeting efficiency (Figure S4F).

While DNA-targeting class 2 CRISPR systems (Gasiunas et al., 2012; Jinek et al., 2012; Zetsche et al., 2015) and some



Figure 2. Programmable RNA Targeting by Cas13d *In Vitro*

(A) *E. siraeum* Cas13d requires a matching guide array or mature gRNA to efficiently cleave complementary ssRNA targets. Denaturing gel depicts cleavage reactions incubated at 37°C for 1 hr. NT, non-targeting.

(B) Substitution with dCas13d or addition of EDTA abrogate Cas13d-mediated RNA targeting with both the guide and array. "d," dead Cas13d.

(C) Denaturing gel depicting guide-target matchdependent activation of Cas13d cleavage activity. Scrambled target RNA ("A") is fluorescently labeled, while guide-complementary activator target RNA ("B") is unlabeled. RNA cleavage activity is abolished by the individual removal of guide RNA or complementary target RNA, as well as the addition of EDTA or the catalytic inactivation of Cas13d (indicated as "d").

(D) A model for guide and target-dependent activation of Cas13d RNase activity. The ternary Cas13d:gRNA:target RNA complex is capable of cleaving the complementary target RNA or bystander RNAs.

See also Figure S4 and Table S1.

repeat. EsCas13d exhibited significant preference for uracil bases in both target structures, with lower but detectable activity at all other bases (Figure 3B).

Cas enzymes are found in nearly all archaea and in approximately half of bacteria (Hsu et al., 2014; van der Oost et al., 2014), spanning a wide range of environmental temperatures. To determine the optimal temperature range for Cas13d activity, we next tested a spectrum of cleavage temperature conditions from 16°C-62°C and observed maximal activity in the 24°C-41°C range (Figures S4G and S4H). This temperature range is compatible with a wide range of prokaryotic and eukaryotic hosts, raising the possibility of adapting Cas13d for RNA targeting in different cells and organisms.

Cell-Based Activity Screen of Engineered Orthologs

We next sought to develop the Cas13d nuclease into a flexible tool for program-

RNA-targeting class 1 systems tend to cleave at defined positions relative to the target-guide duplex (Samai et al., 2015; Zhang et al., 2016), the Cas13d cleavage pattern varies for different targets (Figures 2A, 2C, and S4H) and remains remarkably similar despite the guide sequence position (Figure 3A). This suggests that Cas13d may preferentially cleave specific sequences or structurally accessible regions in the target RNA. We tested Cas13d activity on targets containing variable homopolymer repeats in the loop region of a hairpin or as a linear single-stranded mable RNA targeting in mammalian cells. CRISPR orthologs from distinct bacterial species commonly exhibit variable activity (Abudayyeh et al., 2017; East-Seletsky et al., 2017), especially upon heterologous expression in human cells (Ran et al., 2015; Zetsche et al., 2017). We therefore sought to identify highly active Cas13d orthologs in a eukaryotic cell-based mCherry reporter screen.

By synthesizing human codon-optimized versions of 7 orthologs from distinct branches within the Cas13d family



Figure 3. Characterization of Cas13d Target Substrate Preference

(A) Cas13d can be generalizably reprogrammed with multiple guides and does not exhibit a protospacer flanking sequence (PFS) requirement. RNA cleavage by EsCas13d and 12 guides tiling the target RNA is shown. Control lanes are from a separate gel run in parallel.

(B) Cas13d preferentially cleaves uracil bases in the loop of a hairpin or a linear homopolymer repeat, which is interrupted every 5 nt by a transition mutation (X) to enable synthesis. See also Figure S4 and Table S1.

relative to a non-targeting control guide (Figure 4B). In contrast, EsCas13d along with RaCas13d and RffCas13d exhibited limited activity in human cells. Furthermore, none of the HEPN-inactive Rfx-dCas13d constructs significantly affected mCherry fluorescence, suggesting HEPN-dependent knockdown (p > 0.43 for all cases). Robust nuclear translocation of the Rfx and AdmCas13d NLS fusion constructs was observed via immunocytochemistry, while the wildtype effectors remain primarily extra-nuclear (Figure 4C).

Proceeding with RfxCas13d and AdmCas13d as lead candidates, we next compared their ability to knock down endogenous transcripts. To determine the optimal ortholog and guide architecture, we systematically assayed the capability of Rfx and AdmCas13d construct variants to target beta-1,4-N-acetyl-galactosaminyltransferase 1 (*B4GALNT1*) transcripts. In each condition, we again pooled four guides containing distinct spacer sequences tiling the

(Figure S1B), we generated mammalian expression plasmids carrying the catalytically active and HEPN-inactive proteins. Each protein was then optionally fused to N- and C-terminal nuclear localization signals (NLS). These Cas13d effector designs were HA-tagged and paired with two distinct guide RNA architectures, either with a 30 nt spacer flanked by two direct repeat sequences to mimic an unprocessed guide RNA (pre-gRNA) or a 30 nt direct repeat with 22 nt spacer (gRNA) predicted to mimic mature guide RNAs (Figure 4A). For each guide design, four distinct spacer sequences complementary to the mCherry transcript were then pooled to minimize potential spacer-dependent variability in targeting efficiency. We then assessed the ability of Cas13d to knock down mCherry protein levels in a human embryonic kidney (HEK) 293FT cell-based reporter assay.

48 hr post-transfection, flow cytometry indicated that RfxCas13d and AdmCas13d efficiently knocked down mCherry protein levels by up to 92% and 87% (p < 0.0003), respectively,

B4GALNT1 transcript. We found that the RfxCas13d-NLS fusion targeted *B4GALNT1* more efficiently than wild-type RfxCas13d and both variants of AdmCas13d, with both the gRNA and pre-gRNA mediating potent knockdown (~82%, p < 0.0001) (Figure 4D). We therefore chose Cas13d-NLS from *Rumino-coccus flavefaciens* strain XPD3002 (CasRx) for the remaining experiments.

Programmable RNA Knockdown in Human Cells with CasRx

Because Cas13d is capable of processing its own CRISPR array, we next leveraged this property for the simultaneous delivery of multiple targeting guides in a simple single-vector system (Figure 5A). Arrays encoding four spacers that each tile the transcripts of mRNAs (*B4GALNT1* and *ANXA4*) or nuclear localized long non-coding RNAs (IncRNAs) (*HOTTIP* and *MALAT1*) consistently facilitated robust (>90%) RNA knockdown by CasRx (p < 0.0001) (Figure 5B).



Figure 4. RNA Knockdown Activity Screen of Engineered Cas13d Orthologs in Human Cells

(A) Schematic for mammalian expression constructs encoding for engineered Cas13d effectors and guides. NLS, nuclear localization signal. pre-gRNA, artificial unprocessed guide RNA containing a single 30 nt spacer sequence flanked by 2 full-length 36 nt DRs. gRNA, predicted mature guide RNA with a single 30 nt processed DR and 22 nt spacer sequence.

(B) Heatmap of mCherry protein knockdown in a Cas13d ortholog activity screen in HEK293FT cells using pools of 4 pre-gRNAs or gRNAs. Normalized MFI, median fluorescent intensity relative to non-targeting condition. Positions in gray were not tested, with n = 3.

(C) Immunocytochemistry of Cas13d showing localization and expression of engineered constructs. Scale bar, 10 μ m. Blue pseudocolor, DAPI staining of nuclei. (D) Comparison of Adm and Rfx Cas13d ortholog constructs for knockdown of endogenous *B4GALNT1* mRNA reveals RfxCas13d-NLS (CasRx) to be most effective for both guide RNA architectures. Pools of 4 guides were used for targeting. NT, non-targeting. Values are mean \pm SEM with n = 3. See also Tables S2, S4, and S5.

We next sought to benchmark CasRx against more established technologies for transcript knockdown or repression, comparing CasRx-mediated RNA interference to dCas9-mediated CRISPR interference (Gilbert et al., 2013, 2014) and spacer sequence-matched small hairpin RNAs (shRNAs) via transient transfection (Figure 5C). For CRISPRi-based repression, we included the most potent dCas9 guide for *B4GALNT1* from previous reports (Gilbert et al., 2014; Zalatan et al., 2015). Across 3 endogenous transcripts, CasRx outperformed shRNAs (11/11) and CRISPRi (4/4) in each case (Figure 5D), exhibiting a median knockdown of 96% compared to 65% for shRNA and 53% for CRISPRi after 48 hr. In addition, we compared knockdown by CasRx to two recently described Cas13a and Cas13b effectors (Abudayyeh et al., 2017; Cox et al., 2017) (Figure S5A). Across three genes and eight guide RNAs, CasRx mediated significantly greater transcript knock-down than both LwaCas13a-msfGFP-NLS and PspCas13b-NES (median: 97% compared to 80% and 66% respectively, p < 0.0001) (Figure S5B).

RNAi has been widely used to disrupt any gene of interest due to a combination of simple re-targeting principles, scalable synthesis, knockdown potency, and ease of reagent delivery. However, widespread off-target transcript silencing has been a consistent concern (Jackson et al., 2003; Sigoillot et al., 2012), possibly due to the entry of RNAi reagents into the endogenous miRNA pathway (Doench et al., 2003; Smith et al., 2017).





(B) Arrays of 4 guides each mediate target knockdown by CasRx in 293FT cells via transient transfection. Knockdown relative to GFP vehicle control was determined by qPCR. Values shown as mean ± SEM with n = 3.

(C) Schematic of CasRx target sequences and spacer position-matched shRNAs.

(D) Relative target RNA knockdown by individual position-matched shRNAs and CasRx gRNAs. NT, non-targeting. CRISPRi, dCas9-mediated transcriptional repression. Values shown as mean \pm SEM with n = 3.

(E) Volcano plot of differential transcript levels between *B4GALNT1* targeting and non-targeting (NT) shRNAs as determined by RNA sequencing (n = 3). 542 non-specific transcript changes were identified.

(F) Volcano plot of differential transcript levels between B4GALNT1-targeting CasRx and non-targeting (NT) guide. Targeting guide position is matched to the shRNA shown in (E). B4GALNT1 was the only transcript exhibiting a significant change, with n = 3.

(G) Summary of significant off-target transcript perturbations by matched shRNAs and CasRx guides.

(H) CasRx targeting of 11 endogenous transcripts, each with 3 guides and a non-targeting (NT) guide in 293FT cells. Transcript levels are relative to GFP vehicle control, mean \pm SEM with n = 3.

See also Figures S5 and S6 and Tables S2, S3, S4, and S5.

Consistent with these reports, upon RNA sequencing of human cells transfected with a *B4GALNT1*-targeting shRNA, we observed widespread off-target transcriptional changes relative to a non-targeting shRNA (>500 significant off-target changes, p < 0.01, Figures 5E and 5G). In contrast, transcriptome profiling of spacer-matched CasRx guide RNAs revealed no significant off-target changes other than the targeted transcript (Figure 5F). This suggests that the moderate bystander cleavage observed *in vitro* (Figure 2C) may not result in observable off-target transcriptome perturbation in mammalian cells. We observed a similar pattern when targeting *ANXA4* (Figure S6), with over 900 significant off-target changes resulting from shRNA targeting compared to zero with CasRx (Figure 5G).

To confirm that CasRx interference is broadly applicable, we selected a panel of 11 additional genes with diverse roles in cancer, cell signaling, and epigenetic regulation and screened 3 guides per gene. CasRx consistently mediated high levels of transcript knockdown across genes with a median reduction of 96% (Figure 5H). Each tested guide mediated at least 80% knockdown, underscoring the consistency of the CasRx system for RNA interference.

Splice Isoform Engineering with dCasRx

Our experiments on RNA targeting with CasRx revealed that target RNA and protein knockdown is dependent on the catalytic activity of the HEPN domains (Figures 2B and 4B). The same guide sequences mediating efficient knockdown with CasRx failed to significantly reduce mCherry levels when paired with catalytically inactive dCasRx (Figure 4B), indicating that targeting of dCasRx to the coding portion of mRNA does not necessarily perturb protein translation. This observation suggested the possibility of utilizing dCasRx for targeting of specific coding and non-coding elements within a transcript to study and manipulate RNA. To validate this concept, we sought to expand the utility of the dCasRx system by creating a splice effector.

Alternative splicing is generally regulated by the interaction of *cis*-acting elements in the pre-mRNA with positive or negative *trans*-acting splicing factors, which can mediate exon inclusion or exclusion (Matera and Wang, 2014; Wang et al., 2015). We reasoned that dCasRx binding to such motifs may be sufficient for targeted isoform perturbation. For proof-of-concept, we identified distinct splice elements in a bichromatic splicing reporter containing DsRed upstream of mTagBFP2 in two different reading frames following an alternatively spliced exon (Orengo et al., 2006) (Figure 6A). Inclusion or exclusion of this second exon toggles the reading frame and resulting fluorescence, facilitating quantitative readout of splicing patterns by flow cytometry. To mediate exon skipping, four guide RNAs were designed to target the intronic branchpoint nucleotide, splice acceptor site, putative exonic splice enhancer, and splice donor of exon 2.

One widespread family of negative splice factors are the highly conserved heterogeneous nuclear ribonucleoproteins (hnRNPs), which typically inhibit exon inclusion via a C-terminal, glycinerich domain (Wang et al., 2015). We targeted the splicing reporter with dCasRx and engineered fusions to the Gly-rich C-terminal domain of hnRNPa1, one of the most abundant hnRNP family members (Figure 6B). Guide position appears to be a major determinant of the efficiency of engineered exon skipping. While each guide position mediated a significant increase in exon exclusion (p < 0.0001 in all cases) relative to the non-targeting guide, targeting the splice acceptor resulted in the most potent exon exclusion (increase from 8% basal skipping to 65% for dCasRx alone and 75% with hnRNPa1 fusion). By comparison, dLwaCas13a-msfGFP-NLS mediated significantly lower levels of exon skipping across all four positions (19% skipping for splice acceptor guide) (Figures S5C and S5D, p < 0.0001).

Targeting all 4 positions simultaneously with a CRISPR array achieved higher levels of exon skipping than individual guides alone (81% for dCasRx and 85% for hnRNPa1 fusion, p < 0.006 compared to splice acceptor site [SA] guide) (Figure 6B). These results indicate that dCasRx allows for tuning of isoform ratios through varying guide placement and suggest that it can be leveraged as an efficient RNA binding module in human cells for targeting and manipulation of specific RNA elements.

Viral Delivery of dCasRx to a Neuronal Model of Frontotemporal Dementia

The Cas13d family averages 930 amino acids in length, in contrast to Cas9 (~1,100 aa to ~1,400 aa depending on subtype, with compact outliers such as CjCas9 or SaCas9), Cas13a (1,250 aa), Cas13b (1,150 aa), and Cas13c (1,120 aa) (Figure S2B) (Chylinski et al., 2013; Cox et al., 2017; Hsu et al., 2014; Kim et al., 2017; Shmakov et al., 2015; Smargon et al., 2017). Although AAV is a versatile vehicle for transgene delivery and gene therapy due to its broad range of capsid serotypes, low levels of insertional mutagenesis, and lack of apparent pathogenicity, its limited packaging capacity (~4.7 kb) makes it challenging to effectively deliver many single effector CRISPR enzymes (Abudayyeh et al., 2017; Ran et al., 2015; Swiech et al., 2015). The remarkably small size of Cas13d effectors render them uniquely suited for all-in-one AAV delivery with a CRISPR array, an optional effector domain, and requisite expression or regulatory elements (Figure 6C).

Frontotemporal dementia with parkinsonism linked to chromosome 17 (FTDP-17) is an autosomal dominant major neurodegenerative disease caused by diverse point mutations in *MAPT*, the gene encoding for tau. Tau exists as two major isoforms in human neurons, 4R and 3R, which are distinguished by the presence or absence of *MAPT* exon 10 and thus contain 4 or 3 microtubule binding domains. The balance of these two isoforms is generally perturbed in FTDP-17 as well as other tauopathies, driving the progression of neurodegeneration (Boeve and Hutton, 2008). Some forms of FTD are caused by mutations in the intron following *MAPT* exon 10 that disrupt an intronic splice silencer and elevate the expression of 4R tau (Kar et al., 2005), thereby inducing pathological changes (Schoch et al., 2016).

We reasoned that dCasRx targeted to *MAPT* exon 10 could induce exon exclusion to alleviate dysregulated 4R/3R tau ratios. Patient-derived human-induced pluripotent stem cells (hiPSCs) were differentiated into cortical neurons via *Neurogenin-2* directed differentiation for 2 weeks (Zhang et al., 2013). Postmitotic neurons were then transduced with AAV1 carrying dCasRx



Figure 6. AAV Delivery of Catalytically Inactive dCasRx Splice Effectors to Manipulate Alternative Splicing

(A) Schematic of bichromatic exon skipping reporter. +1 and +3, reading frame; BP, intronic branch point-targeting guide; SA, splice acceptor siteoverlapping guide; EX, exonic guide; SD, splice donor site-overlapping guide; AUG, start codon; UGA, stop codon. Inclusion of the second exon leads to an out-of-frame (+3), non-fluorescent translation of dsRed followed by inframe mTagBFP2. Exclusion of the targeted exon leads to an in frame translation of dsRed (+1) followed by a stop codon.

(B) Induced exon exclusion by dCasRx and an N-terminal hnRNPa1-dCasRx fusion protein targeted to pre-mRNA. The Gly-rich C-terminal domain of hnRNPa1 is used as the effector domain. Exon skipping efficiency is depicted as a relative percentage of cells carrying primarily the dsRed or BFP isoform, determined through flow cytometry. 'A', CRISPR array carrying all 4 guides. Values are mean \pm SEM with n = 3.

(C) AAV design carrying dCasRx and a three-guide array with total transgene size <4.3 kb, including AAV inverted terminal repeats (ITRs).

(Figure 6D) paired with a repeat array containing 3 spacers that target the exon 10 splice acceptor and two putative exonic splice enhancers (Figure 6E). dCasRx-mediated exon exclusion was able to reduce the relative 4R/3R tau ratio by nearly 50% relative to a BFP vehicle control (Figure 6F) and to a level similar to unaffected control neurons, suggesting that CasRx can be exploited for transcriptional modulation in primary cell types via AAV delivery.

DISCUSSION

Class 2 CRISPR systems are found throughout diverse bacterial and archaeal life. Using a minimal definition of the CRISPR locus for bioinformatic mining of prokaryotic genome and metagenome sequences, which requires only a CRISPR repeat array and a nearby protein, we report the identification of an uncharacterized, remarkably compact family of RNA-targeting class 2 CRISPR systems that we designate Type VI CRISPR-Cas13d.

Because CRISPR systems generally exist as a functional operon within 20 kb of genome sequence, even fragmented metagenome reads may be sufficient to recover useful Cas enzymes for bioengineering purposes. CRISPR genome mining strategies described here and by others (Shmakov et al., 2015), combined with ongoing efforts to profile microbial populations via next-generation sequencing, should be anticipated to contribute mechanistically diverse additions to the genome engineering toolbox.

We biochemically characterized two distinct ribonuclease properties of the Cas13d effector, which processes a CRISPR repeat array into mature guides via a HEPN domain-independent mechanism followed by guide sequence-dependent recognition of a complementary activator RNA. This triggers HEPN-mediated RNase activity, enabling Cas13d to cleave both activator and bystander RNAs, a property shared by other RNA-targeting CRISPR systems. Cas13d additionally exhibits no apparent flanking sequence requirements and was found to be active across crRNAs tiling a target RNA, suggesting the ability to target arbitrary ssRNA sequences.

A comprehensive activity reporter screen in human cells of Cas13d orthologs sampled from distinct branches of the Cas13d family revealed that NLS fusions to Cas13d from

(F) Relative 4R/3R tau transcript ratios in differentiated neurons were assayed via qPCR at 14 days following transduction with AAV. FTD, frontotemporal dementia cells carrying IVS 10 + 16. Values are mean \pm SD with n = 3. ****p < 0.0001.

See also Figure S5 and Tables S2, S4, and S5.

⁽D) Schematic of frontotemporal dementia (FTD) disease modeling. Neurons are generated via *Neurogenin-2* (Ngn2) directed differentiation of patient-derived and control iPSCs followed by transduction with dCasRx or vehicle control AAV (EFS-mTagBFP2).

⁽E) FTD is associated with SNPs in a putative intronic splice enhancer following exon 10 of the *MAPT* transcript encoding for tau. Alternative splicing of *MAPT* exon 10 results in 4R tau (by inclusion) and 3R tau (by exclusion). SNPs in the intronic splice enhancer including the indicated IVS 10 + 16 mutation result in increased exon inclusion and higher levels of 4R tau. gRNAs contained in a dCasRx array were targeted to the exon 10 splice acceptor (g1) as well as two putative exonic splice enhancers indicated in purple (g2, g3).

Ruminococcus flavefaciens strain XPD3002 (CasRx) can be engineered for programmable RNA targeting in a eukaryotic context (Figure 4D). CasRx knocked down a diverse set of 14 endogenous mRNAs and IncRNAs, consistently achieving >90% knockdown with favorable efficiency relative to RNA interference, dCas9-mediated CRISPR interference, and other members of the Cas13 superfamily (Figure S5). Additionally, CasRx interference is markedly more specific than spacer-matching shRNAs, with no detectable off-target changes compared with hundreds for RNA interference.

CasRx is a minimal two-component platform, consisting of an engineered CRISPR-Cas13d effector and an associated guide RNA, and can be fully genetically encoded. Because CasRx is an orthogonally delivered protein, HEPN-inactive dCasRx can be engineered as a flexible RNA-binding module to target specific RNA elements. Importantly, because CasRx uses a distinct ribonuclease activity to process guide RNAs, dCasRx can still be paired with a repeat array for multiplexing applications. We demonstrated the utility of this concept by creating a dCasRx splice effector fusion for tuning alternative splicing and resulting protein isoform ratios, applying it in a neuronal model of frontotemporal dementia.

At an average size of 930 aa, Cas13d is to our knowledge the smallest class 2 CRISPR effector characterized in mammalian cells. This allows CasRx effector domain fusions to be paired with a CRISPR array encoding multiple guide RNAs while remaining under the packaging size limit of the versatile AAV delivery vehicle (Naldini, 2015) for primary cell and in vivo delivery. Further, targeted AAV delivery of CasRx to specific postmitotic cell types such as neurons has the potential to mediate long-term expression of a corrective payload that avoids permanent genetic modifications or frequent re-administration (Chiriboga et al., 2016), complementing other nucleic acid targeting technologies such as DNA nuclease editing or antisense oligonucleotides. RNA mis-splicing diseases have been estimated to account for up to 15% of genetic diseases (Hammond and Wood, 2011), highlighting the potential for engineered splice effectors capable of multiplexed targeting. We envision diverse applications to complement RNA targeting for knockdown and splicing, such as live cell labeling and genetic screens to transcript imaging, trafficking, or regulation. CRISPR-Cas13d and engineered variants such as CasRx collectively enable flexible nucleic acid engineering, transcriptome-related study, and future therapeutic development, expanding the genome editing toolbox beyond DNA to RNA.

STAR***METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Cell culture of Human Embryonic Kidney (HEK) cell line 293FT
 - Cell culture of human bone osteosarcoma epithelial cell line U2OS

 Maintenance of induced pluripotent stem cells and neuronal differentiation

METHOD DETAILS

- Computational pipeline for Cas13d identification
- Protein expression and purification
- Preparation of guide and target RNAs
- Biochemical cleavage reactions
- O Transient transfection of human cell lines
- Flow cytometry
- Gene expression analysis
- Immunohistochemistry
- Bacterial small RNA sequencing and analysis
- Ngn2 lentivirus preparation
- AAV preparation
- RNA-seg library preparation and sequencing
- RNA-seq analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
 O Statistics
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and five tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.02.033.

ACKNOWLEDGMENTS

We thank the entire Hsu laboratory for support and advice; T. Hunter and W. Eckhart for helpful comments; F. Gao for the gift of FTDP-17 patientderived and control iPSC lines; A. Hsu for assistance with data analysis; S. Tyagi for guidance with protein purification; S. Schafer for the gift of Neurogenin-2 plasmid; F. Zhang for the gift of His-MBP-TEV plasmid; J. Karlseder, M. Montminy, J. Ayres, and D. Lyumkis for instrument access; C. Pourcel for CRISPRFinder source code; and K. Reilly, S. Leahy, P. Kougias, P. Weimer, F. Farquharson, J. Dore, and F. Levenez for gifts of genomic DNA samples from Ruminococcus flavefaciens XPD3002, Rum. albus, Rum. bicirculans, Rum. sp. CAG:57, Rum. flavefaciens FD1, and anaerobic digester metagenomes. This work was additionally supported by the Waitt Advanced Biophotonics Core Facility of the Salk Institute with funding from NIH-NCI CCSG P30 014195, NINDS Neuroscience Core Grant NS072031, and the Waitt Foundation; the Stem Cell, NGS, Razavi Newman Integrative Genomics and Bioinformatics, and Flow Cytometry Core Facilities of the Salk Institute with funding from the Helmsley Trust; NIH-NCI CCSG P30 014195; and the Chapman Foundation. S.K. is supported by a Catharina Foundation Fellowship, the Howard Hughes Medical Institute Hanna H. Gray Fellows program, and a Salk Women & Science Special Award. P.D.H. is supported by the NIH through the Office of the Director (5 DP5 OD021369-02), the National Institutes on Aging (5 R21 AG056811-02), and the Helmsley Charitable Trust. Plasmids described in this study will be conveniently distributed to the academic community through the nonprofit repository Addgene.

AUTHOR CONTRIBUTIONS

S.K. and P.D.H. conceived this study, developed the CRISPR identification pipeline, and participated in the design of all experiments. N.J.B., S.K., and P.D.H. led the biochemical characterization of Cas13d enzymes. S.K., P.L., and P.D.H. performed cell-based activity screens and knockdown experiments. J.O. performed microscopy experiments. P.L. cloned most constructs. S.K. and P.D.H. performed RNA-sequencing experiments. M.N.S. and S.K. analyzed the sequencing data. S.K., J.O., P.L., and P.D.H. led the splicing experiments. P.D.H. wrote the manuscript with input from S.K. and P.L. and help from all authors.

Cell

DECLARATION OF INTERESTS

P.D.H. is a founder and scientific advisor for Spotlight Therapeutics. S.K. and P.D.H. are co-inventors on U.S. provisional patent application no. 62/572,963 relating to CRISPR-Cas13 and CasRx, as well as other patents on CRISPR technology.

Received: December 22, 2017 Revised: January 31, 2018 Accepted: February 14, 2018 Published: March 15, 2018

REFERENCES

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*, aaf5573.

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., et al. (2017). RNA targeting with CRISPR-Cas13. Nature *550*, 280–284.

Almeida, S., Zhang, Z., Coppola, G., Mao, W., Futai, K., Karydas, A., Geschwind, M.D., Tartaglia, M.C., Gao, F., Gianni, D., et al. (2012). Induced pluripotent stem cell models of progranulin-deficient frontotemporal dementia uncover specific reversible neuronal defects. Cell Rep. *2*, 789–798.

Anantharaman, V., Makarova, K.S., Burroughs, A.M., Koonin, E.V., and Aravind, L. (2013). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol. Direct 8, 15.

Andronescu, M., Condon, A., Hoos, H.H., Mathews, D.H., and Murphy, K.P. (2007). Efficient parameter estimation for RNA secondary structure prediction. Bioinformatics *23*, i19–i28.

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., et al. (2017). Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. Cell *170*, 899–912.

Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J., et al. (2006). 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nat. Methods *3*, 199–204.

Biswas, M.H.U., Almeida, S., Lopez-Gonzalez, R., Mao, W., Zhang, Z., Karydas, A., Geschwind, M.D., Biernat, J., Mandelkow, E.M., Futai, K., et al. (2016). MMP-9 and MMP-2 contribute to neuronal cell death in iPSC models of frontotemporal dementia with MAPT mutations. Stem Cell Reports 7, 316–324.

Bland, C., Ramsey, T.L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N.C., and Hugenholtz, P. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics *8*, 209.

Boeve, B.F., and Hutton, M. (2008). Refining frontotemporal dementia with parkinsonism linked to chromosome 17: introducing FTDP-17 (MAPT) and FTDP-17 (PGRN). Arch. Neurol. *65*, 460–464.

Cheong, C.G., and Hall, T.M. (2006). Engineering RNA sequence specificity of Pumilio repeats. Proc. Natl. Acad. Sci. USA *103*, 13635–13639.

Chiriboga, C.A., Swoboda, K.J., Darras, B.T., Iannaccone, S.T., Montes, J., De Vivo, D.C., Norris, D.A., Bennett, C.F., and Bishop, K.M. (2016). Results from a phase 1 study of nusinersen (ISIS-SMN(Rx)) in children with spinal muscular atrophy. Neurology 86, 890–897.

Chylinski, K., Le Rhun, A., and Charpentier, E. (2013). The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biol. *10*, 726–737.

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

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA matura-

tion by trans-encoded small RNA and host factor RNase III. Nature 471, 602-607.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Doench, J.G., Petersen, C.P., and Sharp, P.A. (2003). siRNAs can function as miRNAs. Genes Dev. *17*, 438–442.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Du, Y.C., Gu, S., Zhou, J., Wang, T., Cai, H., Macinnes, M.A., Bradbury, E.M., and Chen, X. (2006). The dynamic alterations of H2AX complex during DNA repair detected by a proteomic approach reveal the critical roles of Ca(2+)/ calmodulin in the ionizing radiation-induced cell cycle arrest. Mol. Cell. Proteomics 5, 1033–1044.

East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H., Tjian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature 538, 270–273.

East-Seletsky, A., O'Connell, M.R., Burstein, D., Knott, G.J., and Doudna, J.A. (2017). RNA targeting by functionally orthogonal type VI-A CRISPR-Cas enzymes. Mol. Cell *66*, 373–383.

Edgar, R.C. (2007). PILER-CR: fast and accurate identification of CRISPR repeats. BMC Bioinformatics 8, 18.

Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature 532, 517–521.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. USA *109*, E2579–E2586.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. Cell *159*, 647–661.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. *35*, W52-7.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Hammond, S.M., and Wood, M.J. (2011). Genetic therapies for RNA missplicing diseases. Trends Genet. 27, 196–205.

Heidrich, N., Dugar, G., Vogel, J., and Sharma, C.M. (2015). Investigating CRISPR RNA Biogenesis and Function Using RNA-seq. Methods Mol. Biol. *1311*, 1–21.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineagedetermining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262–1278.

Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. (2003). Expression profiling reveals offtarget gene regulation by RNAi. Nat. Biotechnol. *21*, 635–637.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Kar, A., Kuo, D., He, R., Zhou, J., and Wu, J.Y. (2005). Tau alternative splicing and frontotemporal dementia. Alzheimer Dis. Assoc. Disord. *19 (Suppl 1)*, S29–S36.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.

Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K., Cho, H.Y., Song, D.W., Lee, K.J., Jung, M.H., Kim, S., et al. (2017). In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. Nat. Commun. *8*, 14500.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., and Wang, Y. (2017). Two distant catalytic sites are responsible for C2c2 RNase activities. Cell *168*, 121–134.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. Nat. Rev. Microbiol. *13*, 722–736.

Matera, A.G., and Wang, Z. (2014). A day in the life of the spliceosome. Nat. Rev. Mol. Cell Biol. *15*, 108–121.

Naldini, L. (2015). Gene therapy returns to centre stage. Nature 526, 351–360.

O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature *516*, 263–266.

Orengo, J.P., Bundman, D., and Cooper, T.A. (2006). A bichromatic fluorescent reporter for cell-based screens of alternative splicing. Nucleic Acids Res. 34, e148.

Peabody, D.S. (1993). The RNA binding site of bacteriophage MS2 coat protein. EMBO J. *12*, 595–600.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. (2015). Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. Cell *161*, 1164–1174.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467–470.

Schoch, K.M., DeVos, S.L., Miller, R.L., Chun, S.J., Norrbom, M., Wozniak, D.F., Dawson, H.N., Bennett, C.F., Rigo, F., and Miller, T.M. (2016). Increased 4R-tau induces pathological changes in a human-tau mouse model. Neuron *90*, 941–947.

Shendure, J., Balasubramanian, S., Church, G.M., Gilbert, W., Rogers, J., Schloss, J.A., and Waterston, R.H. (2017). DNA sequencing at 40: past, present and future. Nature *550*, 345–353.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. Mol. Cell *60*, 385–397.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.

Sigoillot, F.D., Lyman, S., Huckins, J.F., Adamson, B., Chung, E., Quattrochi, B., and King, R.W. (2012). A bioinformatics method identifies prominent offtargeted transcripts in RNAi screens. Nat. Methods 9, 363–366.

Smargon, A.A., Cox, D.B., Pyzocha, N.K., Zheng, K., Slaymaker, I.M., Gootenberg, J.S., Abudayyeh, O.A., Essletzbichler, P., Shmakov, S., Makarova, K.S., et al. (2017). Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. Mol. Cell 65, 618–630.

Smith, I., Greenside, P.G., Natoli, T., Lahr, D.L., Wadden, D., Tirosh, I., Narayan, R., Root, D.E., Golub, T.R., Subramanian, A., and Doench, J.G. (2017). Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. PLoS Biol. *15*, e2003213.

Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. Nat. Biotechnol. *33*, 102–106.

Treu, L., Kougias, P.G., Campanaro, S., Bassani, I., and Angelidaki, I. (2016). Deeper insight into the structure of the anaerobic digestion microbial community; the biogas microbiome database is expanded with 157 new genomes. Bioresour. Technol. *216*, 260–266.

van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. Nat. Rev. Microbiol. *12*, 479–492.

Wang, Y., Liu, J., Huang, B.O., Xu, Y.M., Li, J., Huang, L.F., Lin, J., Zhang, J., Min, Q.H., Yang, W.M., and Wang, X.Z. (2015). Mechanism of alternative splicing and its regulation. Biomed. Rep. *3*, 152–158.

Wright, A.V., and Doudna, J.A. (2016). Protecting genome integrity during CRISPR immune adaptation. Nat. Struct. Mol. Biol. *23*, 876–883.

Yang, X., Zou, P., Yao, J., Yun, D., Bao, H., Du, R., Long, J., and Chen, X. (2010). Proteomic dissection of cell type-specific H2AX-interacting protein complex associated with hepatocellular carcinoma. J. Proteome Res. *9*, 1402–1415.

Yosef, I., Goren, M.G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic Acids Res. *40*, 5569–5576.

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., and Lim, W.A. (2015). Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell *160*, 339–350.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759–771.

Zetsche, B., Strecker, J., Abudayyeh, O.O., Gootenberg, J.S., Scott, D.A., and Zhang, F. (2017). A survey of genome editing activity for 16 Cpf1 orthologs. bioRxiv. https://doi.org/10.1101/134015.

Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron *78*, 785–798.

Zhang, J., Graham, S., Tello, A., Liu, H., and White, M.F. (2016). Multiple nucleic acid cleavage modes in divergent type III CRISPR systems. Nucleic Acids Res. 44, 1789–1799.

STAR***METHODS**

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-HA 6E2	Cell Signaling Technology	Cat #2367, RRID: AB_10691311
Bacterial and Virus Strains		
E. coli DH5a	Thermo Fisher	Cat# 18258-012
<i>E. coli</i> Rosetta2(DE3)	Novagen	Cat# 71400
pAAV-mTagBFP2	This paper	N/A
pAAV-hU6-DR30-Bsal_EFS-dCasRx-bghpA	This paper	N/A
pLVxTP-Ngn2-2A-Puro_UtOrev	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
Tobacco Etch Virus (TEV) protease	Hsu Laboratory	N/A
Isopropyl-b-D-thiogalactoside (IPTG)	US Biological Life Sciences	Cat# 18500
Pierce Protease-inhibitor tablets, EDTA-free	Thermo Fisher	Cat# A32965
Turbo DNase	Life Technologies	Cat# AM2238
Salt active nuclease	Sigma Aldrich	Cat# SRE-0015
AminoallyI-UTP - ATTO-680	Jena Bioscience	Cat# NU-821-680
Agencourt RNAClean XP Beads	Beckmann Coulter	Cat# A63987
SYBR Gold Nucleic Acid Gel Stain	Life Technologies	Cat# S11494
Lipofectamine 2000	Thermo Fisher	Cat# 11668027
Lipofectamine 3000	Thermo Fisher	Cat# L3000001
TRIzol reagent	Thermo Fisher	Cat# 15596018
T4 Polynucleotide Kinase	NEB	Cat# M0201
RNA 5'-polyphosphatase	Lucigen	Cat# RP8092H
<i>E. coli</i> poly(A) polymerase	NEB	Cat# M0276
T4 RNA Ligase	NEB	Cat# M0204
M-MLV RT/RNase Block	Promega	Cat# M3681
Doxycycline	Sigma Aldrich	Cat# D30-72
ROCK inhibitor Y-27632	Cayman	Cat# 10005583
Bovine Serum Albumin (BSA)	Sigma Aldrich	Cat# A4161
Apo-transferrin	Sigma Aldrich	Cat# T1147
Putrescine	Sigma Aldrich	Cat# P5780
Progesterone	Sigma Aldrich	Cat# P8783
Sodium selenite	Sigma Aldrich	Cat# S5261
Insulin	Roche	Cat# 11376497001
Puromycin	Life Technologies	Cat# A1113803
B27	Thermo Fisher	Cat# 17504044
BDNF	Peprotech	Cat# 450-02
dbcAMP	Sigma Aldrich	Cat# D0627
Cytosine β-D-arabinofuranoside hydrochloride (AraC)	Sigma Aldrich	Cat# C1768
hbEGF	Peprotech	Cat# 100-47
HyClone Fetal Bovine Serum (FBS)	GE Healthcare	Cat# SH30071.03
		L ot# AAD202128

(Continued on next page)

Continued		
REAGENT OF RESOURCE	Behavioran	
Polyethylenimine Max (PEI)	Polysciences	
	Lonza	Cat# 12-725F
Critical Commercial Assays		
NEBNext Ultra II Directional RNA Library Prep Kit	NEB	Cat# E7760
TaqMan Fast Advanced Master Mix	Thermo Fisher	Cat# 4444556
RNeasy Plus Mini Kit	QIAGEN	Cat# 74134
RevertAid First Strand cDNA Synthesis Kit	Thermo Fisher	Cat# K1622
MEGAclear Transcription Clean-Up Kit	Thermo Fisher	Cat# AM1908
Directzol Miniprep Kit	Zymo Research	Cat# R2050
Ribo-Zero rRNA Removal Kit (Bacteria)	Illumina	Cat# MRZMB126
HiScribe T7 High Yield RNA Synthesis Kit	NEB	Cat# E2040
PEG Virus Precipitation Kit	Biovision	Cat# K904
Deposited Data		
GEO (RNA Sequencing of targeting by CasRx and matched shRNAs in 293FT cells)	This paper	GSE108519
Experimental Models: Cell Lines		
Human: HEK293FT (Female)	Thermo Fisher	Cat # R70007
Human: FTD patient iPSC line 3L6 (Male)	Laboratory of Fen-Biao Gao	Biswas et al., 2016
Human: Control iPSC line 2L17 (Male)	Laboratory of Fen-Biao Gao	Almeida et al., 2012
Human: Control iPSC line 2L20 (Male)	Laboratory of Fen-Biao Gao	Almeida et al., 2012
Oligonucleotides		
Guide and target RNA sequences for biochemical assays, see: Table S1	This paper	N/A
Guide RNAs for mammalian RNA targeting, see: Table S2	This paper	N/A
shRNA sequences, see: Table S3	This paper	N/A
qPCR probes and primers, see: Table S4	This paper	N/A
Recombinant DNA		
CRISPR Effector sequences, see: Table S5	This paper	N/A
pAAV1	Penn Vector Core	N/A
pAd DeltaF6	Penn Vector Core	N/A
RG6-mTagBFP2	This paper	Adapted from Addgene #80167 from Thomas Cooper
His-MBP-TEV	Laboratory of Feng Zhang	N/A
psPAX2	Laboratory of Didier Trono	Addgene #12260
pMD2.G	Laboratory of Didier Trono	Addgene #12259
EF1a-PspCas13b-NES(HIV)-2A-EGFP	This paper	Adapted from Addgene #103862 from Feng Zhang
pC0043-PspCas13b crRNA backbone	Cox et al., 2017	Addgene #103854
pC016-LwCas13a guide expression backbone with U6 promoter	Abudayyeh et al., 2017	Addgene #91906
EF1a-LwaCas13a-msfGFP-NLS	This paper	Adapted from Addgene #91924 from Feng Zhang
Software and Algorithms		
Bowtie 2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2
CRISPRfinder	Grissa et al., 2007	Source code was a gift from Christine Pourcel. Online portal: http://crispr.i2bc.paris- saclay.fr/Server
Graphpad Prism 7	GraphPad Software	https://graphpad.com/scientific-software/prism

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
R	R Project	http://www.r-project.org
FlowJo 10.3	FlowJo, LLC	https://flowjo.com
Geneious 10.2	Biomatters	http://www.geneious.com
Image Studio	LI-COR Biosciences	https://www.licor.com/bio/products/software/ image_studio/
DESeq2 v 1.14.1	Love et al., 2014	bioconductor.org/packages/release/bioc/ html/DESeq2.html
2.5.1b STAR aligner	Dobin et al., 2013	https://bioconda.github.io/recipes/star/ README.html
HOMER analysis suite	Heinz et al., 2010	http://homer.ucsd.edu/homer
Clustal Omega	Sievers et al., 2011	http://www.ebi.ac.uk/Tools/msa/clustalo
MAFFT	Katoh and Standley, 2013	https://mafft.cbrc.jp/alignment/server/
PhyML 3.0	Guindon et al., 2010	http://www.atgc-montpellier.fr/phyml/ binaries.php
Other		
Zeiss LSM 880 with Airyscan	Zeiss	N/A
Odyssey Clx Imaging System	LI-COR Biosciences	N/A
LightCycler 480 Instrument II	Roche	N/A
Gel-Doc EZ System	Bio-Rad	N/A
MACSQuant VYB	Miltenyi Biotec	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Patrick D. Hsu (patrick@salk.edu). Key plasmids described in this study will be distributed to the research community via the Addgene plasmid repository under a standard MTA.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture of Human Embryonic Kidney (HEK) cell line 293FT

Human embryonic kidney (HEK) cell line 293FT (female, Thermo Fisher) was maintained in DMEM (4.5 g/L glucose), supplemented with 10% FBS (GE Life Sciences) and 10 mM HEPES at 37°C with 5% CO₂. Upon reaching 80%–90% confluency, cells were dissociated using TrypLE Express (Life Technologies) and passaged at a ratio of 1:2. This cell line was purchased directly from the manufacturer and was not otherwise authenticated.

Cell culture of human bone osteosarcoma epithelial cell line U2OS

Human bone osteosarcoma epithelial U2OS (female) were maintained in DMEM (4.5 g/L glucose) supplemented with 10% FBS and 10 mM HEPES at 37°C with 5% CO₂. Cells were passaged at a 1:3 ratio upon reaching 70% confluence. This cell line was not authenticated.

Maintenance of induced pluripotent stem cells and neuronal differentiation

Stable human iPSC lines containing the FTDP-17 IVS10+16 mutation or age- and sex-matched control lines were obtained from the laboratory of Fen-Biao Gao (Biswas et al., 2016). Briefly, cells obtained from one male patient with the *MAPT* IVS10 + 16 mutation and two separate lines from one male control patient were reprogrammed into hiPSCs (Almeida et al., 2012). iPSCs were transduced with lentivirus containing a doxycycline-inducible *Ngn2* cassette. Lentiviral plasmids were a gift from S. Schafer and F. Gage. iPSCs were then passaged with Accutase and plated into a Matrigel-coated 6-well plate with mTESR media containing ROCK inhibitor Y-27632 (10 µM, Cayman) at 500,000 cells per well. On day 1, media was changed with mTESR. On day 2, media was changed to mTESR containing doxycycline (2 µg/ml, Sigma) to induce *Ngn2* expression. On day 3, culture media was replaced with Neural Induction media (NIM, DMEM/F12 (Life Technologies) containing BSA (0.1 mg/ml, Sigma), apo-transferrin (0.1 mg/ml, Sigma), putrescine (16 µg/ml, Sigma), progesterone (0.0625 µg/ml, Sigma), sodium selenite (0.0104 µg/ml, Sigma), insulin (5 µg/ml, Roche), BDNF (10 ng/ml, Peprotech), SB431542 (10 µM, Cayman), LDN-193189 (0.1 µM, Sigma), laminin (2 µg/ml, Life Technologies), doxycycline (2 µg/ml, Sigma) and puromycin (Life Technologies)). NIM media was changed daily. Following 3 days of puromycin selection,

immature neuronal cells were passaged with Accumax (Innovative Cell Technologies) and plated onto 96-well plates coated with poly-D-lysine and Matrigel in Neural Maturation media (NMM; 1:1 Neurobasal/DMEM (Life Technologies) containing B27 (Life Technologies), BDNF (10 ng/ml, Peprotech), N-Acetyl cysteine (Sigma), laminin (2 µg/ml, Life Technologies), dbcAMP (49 µg/ml, Sigma) and doxycycline (2 µg/ml, Sigma). Media was replaced the next day (day 7) with NMM containing AraC (2 µg/ml, Sigma) to eliminate any remaining non-differentiated cells. On day 8, AraC was removed and astrocytes were plated on top of neurons to support neuron cultures in NMM containing hbEGF (5 ng/ml, Peprotech). Cells were transduced with AAV on day 10 and assayed on day 24.

METHOD DETAILS

Computational pipeline for Cas13d identification

We obtained whole genome, chromosome, and scaffold-level prokaryotic genome assemblies from NCBI Genome in June 2016 and compared CRISPRfinder, PILER-CR, and CRT for identifying CRISPR repeats. The 20 kilobase flanking regions around each putative CRISPR repeat was extracted to identify nearby proteins and predicted proteins using Python. Candidate Cas proteins were required to be > 750 aa in length and within 5 proteins of the repeat array, and extracted CRISPR loci were filtered out if they contained Cas genes associated with known CRISPR systems such as types I-III CRISPR. Putative effectors were clustered into families via all-byall BLASTp analysis followed by single-linkage hierarchical clustering where a bit score of at least 60 was required for cluster assignment. Each cluster of at least 2 proteins was subjected to BLAST search against the NCBI non-redundant (nr) protein database, requiring a bit score > 200 to assign similarity. The co-occurrence of homologous proteins in each expanded cluster to a CRISPR array was analyzed and required to be > 70%. Protein families were sorted by average amino acid length and multiple sequence alignment for each cluster was performed using Clustal Omega and the Geneious aligner with a Blosum62 cost matrix. The RxxxH HEPN motif was identified in the Cas13d family on the basis of this alignment. TBLASTN was performed on all predicted Cas13d effectors against public metagenome whole genome shotgun sequences without predicted open reading frames (ORFs). The Cas13d family was regularly updated via monthly BLAST search on genome and metagenome databases to identify any newly deposited sequences. New full-length homologs and homologous fragments were aligned using Clustal Omega and clustered using PhyML 3.2. CRISPRDetect was used to predict the direction of direct repeats in the Cas13d array and DR fold predictions were performed using the Andronescu 2007 RNA energy model at 37°C (Andronescu et al., 2007). Sequence logos for Cas13d direct repeats were generated using Geneious 10.

Protein expression and purification

Recombinant Cas13d proteins were PCR amplified from genomic DNA extractions of cultured isolates or metagenomic samples and cloned into a pET-based vector with an N-terminal His-MBP fusion and TEV protease cleavage site. The resulting plasmids were transformed into Rosetta2(DE3) cells (Novagen), induced with 200 μ M IPTG at OD₆₀₀ 0.5, and grown for 20 hours at 18°C. Cells were then pelleted, freeze-thawed, and resuspended in Lysis Buffer (50 mM HEPES, 500 mM NaCl, 2 mM MgCl2, 20 mM Imidazole, 1% v/v Triton X-100, 1 mM DTT) supplemented with 1X protease inhibitor tablets, 1 mg/mL lysozyme, 2.5U/mL Turbo DNase (Life Technologies), and 2.5U/mL salt active nuclease (Sigma Aldrich). Lysed samples were then sonicated and clarified via centrifugation (18,000 x g for 1 hour at 4°C), filtered with 0.45 μ M PVDF filter and incubated with 50 mL of Ni-NTA Superflow resin (QIAGEN) per 10 L of original bacterial culture for 1 hour. The bead-lysate mixture was applied to a chromatography column, washed with 5 column volumes of Lysis Buffer, and 3 column volumes of Elution Buffer (50 mM HEPES, 500 mM NaCl, 300 mM Imidazole, 0.01% v/v Triton X-100, 10% glycerol, 1 mM DTT). The samples were then dialyzed overnight into TEV Cleavage Buffer (50 mM Tris-HCl, 250 mM KCl, 7.5% v/v glycerol, 0.2 mM TCEP, 0.8 mM DTT, TEV protease) before cation exchange (HiTrap SP, GE Life Sciences) and gel filtration (Superdex 200 16/600, GE Life Sciences). Purified, eluted protein fractions were pooled and frozen at 4 mg/mL in Protein Storage Buffer (50 mM Tris-HCl, 10% glycerol, 2 mM DTT).

Preparation of guide and target RNAs

Oligonucleotides carrying the T7 promoter and appropriate downstream sequence were synthesized (IDT) and annealed with an antisense T7 oligo for crRNAs and PCR-amplified for target and array templates. Homopolymer target RNAs were synthesized by Synthego. The oligo anneal and PCR templates were *in vitro* transcribed with the Hiscribe T7 High Yield RNA Synthesis kit (New England Biolabs) at 31°C for 12 hours. For labeled targets, fluorescently labeled aminoallyI-UTP atto 680 (Jena Biosciences) was additionally added at 2 mM. Guide RNAs were purified with RNA-grade Agencourt AMPure XP beads (Beckman Coulter) and arrays and targets were purified with MEGAclear Transcription Clean-Up Kit (Thermo Fisher) and frozen at -80°C. For ssDNA and dsDNA targets, corresponding oligonucleotide sequences were synthesized (IDT) and either gel purified, or PCR amplified and then subsequently gel purified respectively.

Biochemical cleavage reactions

Purified EsCas13d protein and guide RNA were mixed (unless otherwise indicated) at 2:1 molar ratio in RNA Cleavage Buffer (25mM Tris pH 7.5, 15mM Tris pH 7.0, 1mM DTT, 6mM MgCl2). The reaction was prepared on ice and incubated at 37°C for 15 minutes prior to the addition of target at 1:2 molar ratio relative to EsCas13d. The reaction was subsequently incubated at 37°C for 45 minutes and quenched with 1 μL of enzyme stop solution (10 mg/mL Proteinase K, 4M Urea, 80mM EDTA, 20mM Tris pH 8.0) at 37°C for

15 minutes. The reaction was then denatured with 2X RNA loading buffer (2X: 13mM Ficoll, 8M Urea, 25 mM EDTA), at 85°C for 10 minutes, and separated on a 10% TBE-Urea gel (Life Technologies). Gels containing labeled targets were visualized on the Odyssey Clx Imaging System (Li-Cor); unlabeled array or target cleavage gels were stained with SYBR Gold prior to imaging via Gel Doc EZ system (Bio-Rad).

Transient transfection of human cell lines

Engineered Cas13 coding sequences were cloned into a standardized plasmid expression backbone containing an EF1a promoter and prepared using the Nucleobond Xtra Midi EF Kit (Machery Nagel) according to the manufacturer's protocol. NLS-LwaCas13amsfGFP and PspCas13b-NES-HIV were PCR amplified from Addgene #103854, and #103862, respectively, a gift from Feng Zhang. Cas13d pre-gRNAs and gRNAs were cloned into a minimal backbone containing a U6 promoter. shRNAs and guides for LwaCas13a were cloned into the same backbone and position matched to their corresponding guide RNA at the 3' of the target sequence. Matched gRNAs for PspCas13b were moved to the closest 5'-G nucleotide.

For transient transfection, HEK293FT cells were plated at a density of 20,000 cells per well in a 96-well plate and transfected at > 90% confluence with 200 ng of Cas13 expression plasmid and 200 ng of gRNA expression plasmid using Lipofectamine 2000 (Life Technologies) according to the manufacturer's protocol. Transfected cells were harvested 48-72 hours post-transfection for flow cytometry, gene expression analysis, or other downstream processing.

For reporter assays, HEK293FT cells were transfected in 96-well format with 192ng of Cas13d expression plasmid, 192ng of guide expression plasmid, and 12ng of mCherry expression plasmid with Lipofectamine 2000 (Life Technologies). Cells were harvested after 48 hours and analyzed by flow cytometry.

U2OS cells were plated at a density of 20,000 cells per well in a 96-well plate and transfected at > 90% confluence with 100 ng of Cas13d expression plasmid using Lipofectamine 3000 (Life Technologies) according to the manufacturer's protocol and processed for immunocytochemistry after 48h.

Flow cytometry

Cells were dissociated 48 hours post-transfection with TrypLE Express and resuspended in FACS Buffer (1X DPBS^{-/-}, 0.2% BSA, 2 mM EDTA). Flow cytometry was performed in 96-well plate format using a MACSQuant VYB (Miltenyi Biotec) and analyzed using FlowJo 10. RG6 was a gift from Thomas Cooper (Addgene plasmid # 80167) and modified to replace EGFP with mTagBFP2. All represented samples were assayed with three biological replicates. In the mCherry reporter assay, data is representative of at least 20,000 gated events per condition. In the splicing reporter assay, data is representative of at least 2,500 gated events per condition.

Gene expression analysis

Cells were lysed 48 hours post-transfection with DTT-supplemented RLT buffer and total RNA was extracted using RNeasy Mini Plus columns (QIAGEN). 200 ng of total RNA was then reverse transcribed using random hexamer primers and Revertaid Reverse Transcriptase (Thermo Fisher) at 25°C for 10 min, 37°C for 60 min, and 95°C for 5 min followed by qPCR using 2X Taqman Fast Advanced Master Mix (Life Technologies) and Taqman probes for GAPDH and the target gene as appropriate (Life Technologies and IDT). Taqman probe and primer sets were generally selected to amplify cDNA across the Cas13 or shRNA target site position to prevent detection of cleaved transcript fragments (Table S4). qPCR was carried out in 5 μ L multiplexed reactions and 384-well format using the LightCycler 480 Instrument II (Roche). Fold-change was calculated relative to GFP-transfected vehicle controls using the ddCt method. One-way or two-way ANOVA with multiple comparison correction was used to assess statistical significance of transcript changes using Prism 7.

Immunohistochemistry

For immunohistochemical analysis, U2OS cells were cultured on 96-well optically clear plates (Greiner Bio-One), transfected as previously described, then fixed in 4% PFA (Electron Microscopy Sciences) diluted in PBS (GIBCO) and washed with 0.3M glycine (Sigma) in PBS to quench PFA. Samples were blocked and permeabilized in a PBS solution containing 8% donkey serum (Jackson ImmunoResearch), 8% goat serum (Cell Signaling Technologies), and 0.3% Triton X-100 (Sigma) for one hour, followed by primary antibody incubation in 1% BSA (Fisher Bioreagents), 1% goat serum, and 0.25% Triton-X overnight at 4°C. Samples were washed 3 times with PBS containing 0.1% BSA and 0.1% Triton X-100 before incubating with fluorophore-conjugated secondary antibodies in PBS with 0.05% Triton X-100 and 1% BSA at room temperature for one hour. Cells were washed with PBS with 0.1% Triton-X, stained with DAPI, and then covered with Mounting Media (Ibidi) before imaging. Primary antibody, HA-Tag 6E2 (Cell Signaling, 2367), was used at a 1:100 dilution as per manufacturer's instructions. Secondary antibodies used were goat anti-mouse IgG1-Alexa Fluor 647 (Thermo Fisher, A21240) and Anti-Mouse IgG1 CF 633 (Sigma, SAB4600335). Confocal images were taken using a Zeiss Airyscan LSM 880 followed by image processing in Zen 2.3 (Zeiss).

Bacterial small RNA sequencing and analysis

E. coli DH5α cells were transformed with pACYC184 carrying the CRISPR-Cas13d locus derived from an uncultured *Ruminococcus sp.* strain. Cells were harvested in stationary phase, rinsed in PBS, resuspended in TRIzol (Life Technologies), transferred to Lysing Matrix B tubes containing 0.1 mm silica beads (MP Biomedicals), and homogenized on a Bead Mill 24 (Fisher Scientific) for three 30 s

cycles. Total RNA was isolated by phenol-chloroform extraction, then purified using the DirectZol Miniprep Kit (Zymo Research). RNA quality was assessed on an Agilent 2200 Tapestation followed by Turbo DNase treatment (Ambion). Total RNA was treated with T4 Polynucleotide Kinase (NEB) and rRNA-depleted using the Ribo-Zero rRNA Removal Kit for bacteria (Illumina). RNA was treated with RNA 5' polyphosphatase, poly(A)-tailed with *E. coli* poly(A) polymerase, and ligated with 5' RNA sequencing adapters using T4 RNA ligase 1 (NEB). cDNA was generated via reverse transcription using an oligo-dT primer and M-MLV RT/RNase Block (AffinityScript, Agilent) followed by PCR amplification and barcoding. Resulting libraries were sequenced on Illumina MiSeq, demultiplexed using custom Python scripts, and aligned to the Cas13d CRISPR locus using Bowtie 2. Alignments were visualized with Geneious.

Ngn2 lentivirus preparation

Low passage HEK293FT cells were transfected with Polyethylenimine Max (PEI, Polysciences) and *Ngn2* target plasmid plus pMDG.2 and psPAX2 packaging plasmids (a gift from Didier Trono, Addgene #12259 and #12260) in DMEM + 10% FBS media during plating. The following day, media was changed to serum-free chemically defined minimal medium (Ultraculture supplemented with Glutamax, Lonza). Viral supernatant was harvested 48h later, clarified through a 0.45 micron PVDF filter (Millipore) and concentrated using ultracentrifugation.

AAV preparation

Low passage HEK293FT cells were transfected with Polyethylenimine Max (PEI, Polysciences) and AAV target plasmid plus AAV1 serotype and pAdDeltaF6 helper packaging plasmids (UPenn Vector Core) in DMEM + 10% FBS media during plating. The following day, 60% of the media was changed to chemically defined minimal medium (Ultraculture supplemented with Glutamax, Lonza). 48h later, AAV-containing supernatant was harvested and clarified through a 0.45 µm PVDF filter (Millipore) and concentrated using precipitation by polyethylene glycol (PEG virus precipitation kit #K904, Biovision) following the manufacturer's protocol.

RNA-seq library preparation and sequencing

48h after transfection, total RNA was extracted from 293FT cells using the RNeasy Plus Mini kit from QIAGEN. Stranded mRNA libraries were prepared using the NEBNext II Ultra Directional RNA Library Prep Kit from New England Biolabs (Cat# E7760S) and sequenced on an Illumina NextSeq500 with 42 nt paired end reads. ~15M total reads were demultiplexed per condition.

RNA-seq analysis

Sequenced reads were quality-tested using FASTQC and aligned to the hg19 human genome using the 2.5.1b STAR aligner (Dobin et al., 2013). Mapping was carried out using default parameters (up to 10 mismatches per read, and up to 9 multi-mapping locations per read). The genome index was constructed using the gene annotation supplied with the hg19 Illumina iGenomes collection (Illumina) and sjdbOverhang value of 100. Uniquely mapped reads were quantified across all gene exons using the top-expressed isoform as proxy for gene expression with the HOMER analysis suite (Heinz et al., 2010), and differential gene expression was carried out with DESeq2 v 1.14.1 (Love et al., 2014) using triplicates to compute within-group dispersion and contrasts to compare between targeting and non-targeting conditions. Significant differentially expressed genes were defined as having a false discovery rate (FDR) < 0.01 and a log2 fold change > 0.75. Volcano plots were generated in R 3.3.2 using included plotting libraries and the alpha() color function from the scales 0.5.0 package.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics

All values are reported as mean ± SD or mean ± SEM as indicated in the appropriate figure legends. For comparing two groups, a one-tailed Student's t test was used and statistical significance was determined using the Holm-Sidak method with alpha = 0.05. A one-way ANOVA with Tukey multiple hypothesis correction was used to assess significance between more than two groups. Two-way ANOVA was used when comparing across two factors (i.e., RNA targeting modality and guide position) and adjusted for multiple hypothesis correction by Sidak's multiple comparisons test. For comparing groups that were found to not meet the assumption of a normal distribution by a D'Agostino and Pearson normality test, the non-parametric Friedman test with Dunn's multiple comparison adjustment was performed. PRISM 7.0 was used for all statistical analysis. Sample sizes were not determined *a priori*. At least three biological replicates were used for each experiment, as indicated specifically in each figure.

DATA AND SOFTWARE AVAILABILITY

The accession number for the sequencing data reported in this paper is GEO: GSE108519.

Supplemental Figures



Figure S1. Bioinformatic Pipeline for the Identification of the RNA-Targeting Class 2 CRISPR System Cas13d, Related to Figure 1

(A) Schematic describing a computational pipeline for CRISPR system identification. A minimal definition for a putative class 2 CRISPR locus was used, requiring only a CRISPR repeat array and a nearby protein > 750 aa in length. As per Methods, initial search was performed on prokaryotic genome assemblies derived from NCBI Genome, and later expanded via TBLASTN of predicted Cas13d proteins against public metagenome sequences without predicted open reading frames. DR, direct repeat.

(B) Phylogenetic classification and alignment of full-length Cas13d effectors and metagenomic fragments. Cas13d effectors and metagenomic Cas13d protein fragments cluster into several distinct branches, which are colored for ease of interpretation. Shading indicates residue conservation using the Blosum62 matrix. Full-length Cas13d effectors used in this study were sampled from distinct branches of the Cas13d family. Alignment of Cas13d proteins and protein fragments was performed using ClustalOmega 1.2.4 and maximum-likelihood tree building was performed with PhyML 3.2.





(A) HEPN motif conservation in Cas13d effectors used in this study with conserved residues shaded according to Blosum62. The RxxxH HEPN motif is highlighted.

⁽B) Maximum-likelihood tree of type VI CRISPR-Cas families. Average amino acid lengths of Type VI Cas13 superfamily effectors are indicated in red. Alignment of previously described class 2 CRISPR RNA-targeting proteins (Abudayyeh et al., 2017; Cox et al., 2017; East-Seletsky et al., 2016, 2017; Smargon et al., 2017) and Cas13d effectors was performed using MAFFT 7.38 and maximum-likelihood tree building was performed with PhyML 3.2. Branch labels and scale bar indicate substitutions per site.

⁽C) Predicted Cas13d direct repeat RNA secondary structure.

⁽D) Sequence logo of full length 36 nt Cas13d direct repeats.



Figure S3. Purification of Recombinant Cas13d Protein, Related to Figure 1

EsCas13d was expressed as N-terminal His-MBP fusion and purified by successive affinity, cation exchange, and size exclusion chromatography. The His-tag was removed by TEV protease cleavage.

(A) Chromatogram from Superdex 200 column for EsCas13d.

(B) SDS-PAGE gel of size exclusion chromatography fractions for *E. siraeum* Cas13d.

(C) SDS-PAGE gel of purified E. siraeum Cas13d and dCas13d (R295A, H300A, R849A, H854A mutations of predicted catalytic residues in both HEPN motifs).



(legend on next page)

Figure S4. In Vitro Characterization of Cas13d Properties, Related to Figures 2 and 3

(A) Schematic showing the length and sequence of gRNA spacer truncations and spacer position relative to the complementary ssRNA target.

(B) Denaturing gel depicting EsCas13d cleavage activity of target RNA with different gRNA spacer lengths.

(C) Denaturing gel depicting EsCas13d cleavage reactions paired with 12 guides from Figure 3A tiling a complementary ssDNA version of the ssRNA target. (D) Denaturing gel depicting cleavage reactions using EsCas13d paired with the same 12 guides tiling a dsDNA version of the complementary target.

(F) Cas13d-mediated cleavage of target RNA carrying different PFS bases given an invariant spacer sequence. Quantification of Cas13d cleavage efficiency and a representative denaturing gel depicting EsCas13d cleavage activity are shown. Differences are not significant (one-way ANOVA, p = 0.768). Cleavage percentage is determined as above, and mean is depicted \pm SD with n = 3.

(G and H) Optimal temperature range for Cas13d activity. Denaturing gels depicting EsCas13d cleavage activity at temperatures ranging from 16-62°C for two different target RNAs.

⁽E) Quantification of cleavage efficiency from Figure 3A. Each PFS base is the average of 3 different spacer sequences tiling a complementary target RNA. Cleavage percentage is determined by the ratio of cleaved band intensity divided by total lane intensity. Mean is depicted ± SD with each data point representing an independent replicate.



Figure S5. Comparison of Engineered Cas13 Superfamily Effectors for Targeted Knockdown and Splicing, Related to Figures 5 and 6 (A) Relative target RNA knockdown by individual position-matched gRNAs for CasRx, NLS-LwaCas13a-msfGFP (Abudayyeh et al., 2017) and PspCas13b-NES

(Cox et al., 2017) in HEK293FT cells. NT, non-targeting. Values are mean \pm SEM with n = 3.

(B) Comparison of Cas13 median knockdown efficiencies. n = 3 per guide RNA. **** indicates p < 0.0001 according to Friedman's test.

(C) Exon exclusion by catalytically inactive NLS-dCas13a-msfGFP on the bichromatic splicing reporter. Guides are position-matched to those reported in Figure 6B for CasRx. Values are mean \pm SEM with n = 3.

(D) Comparison of splicing modulation by NLS-dCas13a-msfGFP and CasRx. Fold change in targeted exon exclusion relative to non-targeting guide is shown. Values are mean \pm SEM with n = 3. **** indicates p < 0.0001 according to two-way ANOVA.



Figure S6. RNA Sequencing from CasRx and shRNA Targeting of ANXA4 in Human Cells, Related to Figure 5

(A) Volcano plots of differential transcript levels between ANXA4 targeting and non-targeting (NT) shRNAs as determined by RNA sequencing (n = 3). 915 non-specific transcript changes were identified.

(B) Volcano plot of differential transcript levels for an ANXA4 targeting CasRx array used in Figure 5B containing a guide position matched to the shRNA shown in (A) and a non-targeting (NT) array. ANXA4 was the only transcript exhibiting significant downregulation with n = 3. HIST2HBE was the only transcript identified to exhibit significant upregulation. H2B is a dimer partner of H2AX (Du et al., 2006) which has been shown to interact with ANXA4 (Yang et al., 2010).

Article

Cell

Structural Basis for the RNA-Guided Ribonuclease Activity of CRISPR-Cas13d

Graphical Abstract



Authors

Cheng Zhang, Silvana Konermann, Nicholas J. Brideau, ..., Patrick R. Griffin, Patrick D. Hsu, Dmitry Lyumkis

Correspondence

patrick@salk.edu (P.D.H.), dlyumkis@salk.edu (D.L.)

In Brief

Cryo-EM structures and biochemical analysis of CRISPR-Cas13d in apo, guide-bound, and target-bound states offer insight for engineering this RNAtargeting system.

Highlights

Check for updates

- Structures of the smallest type VI CRISPR effector in guide and target-bound states
- Mechanistic insights into guide RNA and target RNA recognition
- Insights into apo Cas13d structural dynamics through cryo-EM and HDX-MS
- Rational engineering of Cas13d for minimal coding sequence



Structural Basis for the RNA-Guided Ribonuclease Activity of CRISPR-Cas13d

Cheng Zhang,^{1,3,6} Silvana Konermann,^{2,3,6} Nicholas J. Brideau,^{2,3} Peter Lotfy,^{2,3} Xuebing Wu,⁴ Scott J. Novick,⁵ Timothy Strutzenberg,⁵ Patrick R. Griffin,⁵ Patrick D. Hsu,^{2,3,*} and Dmitry Lyumkis^{1,3,7,*}

¹Laboratory of Genetics, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

²Laboratory of Molecular and Cell Biology, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA ³Helmsley Center for Genomic Medicine, Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

⁴Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

⁵Department of Molecular Medicine, The Scripps Research Institute, 130 Scripps Way, Jupiter, FL 33458, USA

*Correspondence: patrick@salk.edu (P.D.H.), dlyumkis@salk.edu (D.L.) https://doi.org/10.1016/j.cell.2018.09.001

SUMMARY

CRISPR-Cas endonucleases directed against foreign nucleic acids mediate prokaryotic adaptive immunity and have been tailored for broad genetic engineering applications. Type VI-D CRISPR systems contain the smallest known family of single effector Cas enzymes, and their signature Cas13d ribonuclease employs guide RNAs to cleave matching target RNAs. To understand the molecular basis for Cas13d function and explain its compact molecular architecture, we resolved cryoelectron microscopy structures of Cas13d-guide RNA binary complex and Cas13dguide-target RNA ternary complex to 3.4 and 3.3 Å resolution, respectively. Furthermore, a 6.5 Å reconstruction of apo Cas13d combined with hydrogendeuterium exchange revealed conformational dynamics that have implications for RNA scanning. These structures, together with biochemical and cellular characterization, provide insights into its RNA-guided, RNA-targeting mechanism and delineate a blueprint for the rational design of improved transcriptome engineering technologies.

INTRODUCTION

Bacterial life employs diverse CRISPR systems to protect themselves against predatory phage, engaging Cas nucleases with programmable guide RNAs to target invading nucleic acids and endow the host cell with adaptive immunity (Barrangou et al., 2007; Brouns et al., 2008). CRISPR systems are broadly divided into two classes, each with multiple types and subtypes: class 1 systems (types I and III) coordinate multiple proteins that cooperate for target surveillance and defense, while class 2 systems integrate both functions into a single effector enzyme (Koonin et al., 2017).

Class 2 CRISPR-Cas systems include types II, V, and VI, with types II and V shown to target DNA. Adapted over the last half

decade into a remarkably flexible genetic engineering toolbox, class 2 DNA-targeting enzymes such as CRISPR-Cas9 (type II) and CRISPR-Cas12a/Cpf1 (type V) have facilitated many applications, from gene editing to lineage tracing, multi-color chromosomal imaging, and gene drives. Although some class 1 CRISPR systems can target RNA (Hale et al., 2009; Jiang et al., 2016; Kazlauskiene et al., 2017; Niewoehner et al., 2017; Samai et al., 2015), type VI systems have been recently described as the only known single-effector CRISPR nucleases that exclusively target RNA (Abudayyeh et al., 2016; East-Seletsky et al., 2016; Konermann et al., 2018; Shmakov et al., 2015; Smargon et al., 2017; Yan et al., 2018). Cas13, the signature single-effector enzyme family, comprises guide-RNA-directed ribonucleases with four subtypes (Cas13a-d) that each exhibit significant sequence divergence apart from two consensus HEPN (higher eukaryotes and prokaryotes nucleotide-binding domain) RNase motifs, R-X₄₋₆-H. Domains belonging to the HEPN superfamily are frequently found in ribonucleases involved in immune defense (Anantharaman et al., 2013), including in class 1 CRISPR RNases such as Csm6 or the homologous Csx1 (Jiang et al., 2016; Kazlauskiene et al., 2017; Niewoehner and Jinek, 2016) as well as prokaryotic Abi and T-AT defense systems or the anti-viral mammalian RNase L (Han et al., 2014). To defend against viral infection, Cas13 enzymes process pre-crRNA (CRISPR RNA) into mature crRNA guides in a HEPNindependent manner, followed by HEPN-dependent cleavage of a complementary "activator" target RNA in cis. Upon targetdependent activation, Cas13 is also able to cleave bystander RNAs in trans, reflecting a general RNase activity capable of both cis- and trans-cleavage.

Despite functional similarities of crRNA-dependent activation of HEPN-mediated RNA cleavage, Cas13 subtypes characterized to date exhibit key differences beyond their significant divergence at the primary sequence level. Using a computational pipeline for identifying novel class 2 CRISPR-Cas loci from genome and metagenome sequences sourced from large-scale microbiome sequencing efforts, we recently described a Cas13 subtype designated as Cas13d (Konermann et al., 2018). Cas13d enzymes are 20%–30% smaller than other Cas13 subtypes, facilitating flexible packaging into size-constrained

⁶These authors contributed equally

⁷Lead Contact

therapeutic viral vectors such as adeno-associated virus (AAV) (Konermann et al., 2018; Yan et al., 2018).

Cas13 enzymes provide a rich resource for new RNA-targeting technologies and have been recently developed for RNA knockdown (Abudayyeh et al., 2017; Cox et al., 2017; Konermann et al., 2018), editing (Cox et al., 2017), splicing (Konermann et al., 2018), and viral delivery (Konermann et al., 2018). Remarkably, Cas13 subtypes and individual orthologs exhibit highly variable activity in human cells, with Cas13d displaying robust activity for both target cleavage and binding (Cox et al., 2017; Konermann et al., 2018). Here, we sought to understand the molecular and structural basis for Cas13d function, including both guide and target RNA recognition.

RESULTS

Determination of a High-Resolution Cryo-EM Structure of Cas13d in Complex with crRNA

To gain structural insight into Cas13d function, we purified the catalytically active *Es*Cas13d (Konermann et al., 2018) and formed a binary complex containing Cas13d bound to crRNA followed by cryoelectron microscopy (cryo-EM) imaging (Figures S1A and S1B). A large data collection, followed by a computational analysis and refinement of 43,786 particles, led to the derivation of a coulombic potential map of the binary complex bound to crRNA, resolved to a mostly homogeneous resolution of \sim 3.4 Å (Figure S1C; Table S1). Using recently described procedures for characterizing anisotropy in cryo-EM density maps (Tan et al., 2017), we found that the structure maintained approximately even distribution of directional resolution (Figure S1D).

At 954 amino acids (molecular weight ~105 kDa), *Es*Cas13d is considerably smaller than average members of the type VI-A, -B, and -C subtypes. Most residues could be built into the density, with the exception of several flexible loops. All 52 nucleotides spanning the crRNA were observed in the EM density, and 51 of them could be confidently modeled (Figure S1E). The final model is consistent with the cryo-EM map, with good geometry and model statistics (Table S1).

The Structure of crRNA-Bound Cas13d Reveals a Compact Protein Architecture surrounding Solvent-Exposed RNA

The Cas13d binary complex (Figures 1A–1D) maintains a bilobed architecture with five distinct domains organized around the central crRNA guide (Figure 1A). The domains include an N-terminal domain (NTD), a HEPN1 catalytic domain that is split into two distinct regions in sequence space, a first linker domain termed Helical-1, a second HEPN2 catalytic domain, and a second linker domain termed Helical-2. With the exception of the NTD, which is composed of two short α helices flanking a β sandwich region formed by two antiparallel 3-stranded β sheets, the protein is predominantly α helical.

The overall binary ribonucleoprotein architecture is reminiscent of a half-open clam shape surrounding the solvent-exposed crRNA channel. The mature crRNA is divided into a constant direct repeat (DR) region (nucleotides [nt] 1–30), derived from the characteristic repeat of CRISPR arrays, and a spacer region (nt 31–52) complementary to the target protospacers (Figure 1B). In the binary complex, the 5' crRNA handle (also referred to as the DR) is clamped by NTD and HEPN2, with the first two base pairs and its 5 nt loop protruding away from protein density (Figures 1C, 1D, and S1E). Immediately downstream of the DR, the spacer region resides within a cleft and is sandwiched between Helical-1 and Helical-2. HEPN1 provides a structural scaffold connecting the two lobes of Cas13d, reminiscent of a hinge around the largely solvent-exposed RNA density. In this compact configuration, Cas13d forms a "surveillance complex," poised for searching and identifying complementary target sites (Figure 1D).

Determination of a High-Resolution Cryo-EM Structure of Cas13d in Complex with crRNA and Target RNA

Type VI CRISPR-Cas RNases catalyze degradation of singlestranded RNA (ssRNA) through a process that is mediated by the formation of an activated ternary complex containing both spacer and complementary protospacer (Abudayyeh et al., 2016; East-Seletsky et al., 2016; Konermann et al., 2018; Smargon et al., 2017; Yan et al., 2018). To understand the molecular basis for nuclease activation, we sought to determine the ternary structure composed of Cas13d bound to both crRNA and its complementary target RNA. To stabilize the ternary complex in a pre-cleavage state, we mutated all four predicted catalytic HEPN domain residues to alanine (R295A/H300A/R849A/ H854A). We previously reported that this "catalytically dead" Cas13d (dCas13d) retains the ability to bind both crRNA and target RNA but cannot cleave ssRNA (Konermann et al., 2018).

Using similar procedures, we refined 51,885 particles to determine the structure of the Cas13d ternary complex to an average resolution of \sim 3.3 Å (Figures 1E–1G and S1F–S1H), with a more anisotropic distribution of directional resolution (Figure S1I). The quality and overall features of the map were sufficient to derive an atomic model of the ternary complex consistent with the cryo-EM density (Table S1) and resolve most of the polypeptide chain, the entire 52 nt crRNA, and all complementary nucleotides of the target RNA (nt 5–26) (Figures 1G and S1J).

Cas13d Binds Target RNA within a Large Central Cleft Opposite to the Catalytic Site

The structure of the Cas13d-crRNA-target RNA ternary complex shows a similar compact architecture as the binary complex, with the protein subunits of both lobes wrapped around a spacer:protospacer duplex (Figure 1G). All 22 complementary nucleotides of the target RNA (nt 5–26) base pair with the spacer within crRNA (Figures 1E and S1J), and only the terminal two bases of the RNA duplex extend outside of the central cleft (Figure 1G). The 5' handle maintains a solvent-exposed organization as in the binary state, while the guide-target duplex assembles into an A-form RNA helix within the cleft bound by HEPN1, Helical-1, and Helical-2 domains.

On the outside face of the protein, opposite of the central cleft, HEPN1 and HEPN2 form an endoRNase heterodimer. HEPN1 is subdivided in sequence space (residues ~150–344 termed HEPN1-I and ~495–577 termed HEPN1-II) but forms a contiguous tertiary fold. The α 1 of HEPN1-I and the C-terminal portion of HEPN2 form the structural backbone of the bipartite active site and position the four catalytic residues (R295A/H300A/R849A/H854A) of the R-X₄-H motif outward on the external face of



Cas13d. This orientation of the HEPN active site primes the Cas13d ternary complex for cleavage of both target and collateral RNAs.

Bilobed Organization Is Conserved across Class 2 CRISPR Effectors

Class 2 CRISPR-Cas effectors are characterized by their bilobed architectures containing a nucleic acid recognition (REC) lobe that binds crRNA, as well as a nuclease (NUC) lobe that is responsible for cleavage of target nucleic acids (Garcia-Doval and Jinek, 2017) (Figure S2). In contrast to Cas13a, which buries backbone contacts, typically with phosphates but also ribose hydroxyl groups. Individual base interactions occur at lower frequency and are typically constrained to conserved bases within the DR (Figures 2 and 3A). These extensive contacts stabilize the spacer region in a primed pseudo-helical conformation within the solvent-exposed central channel (Figures 3B–3D).

Upon target binding, most of the spacer interactions are shifted toward the 3' end, following formation of the A-form double-stranded RNA (dsRNA) helix. Only four residues (K443, Y447, K376, Y680) interface with RNA in both enzymatic forms, forming backbone contacts with the 5' end of the spacer in binary form.

Figure 1. Structures of Cas13d-crRNA and Cas13d-crRNA-Target RNA

(A) Domain organization of *Es*Cas13d, with boundaries and catalytic sites of the HEPN domains (R295, H300, R849, H854) indicated.

(B) Schematic representation of the crRNA used for binary complex assembly. The DR (5' handle) and spacer of crRNA are colored magenta and cyan, respectively.

(C) Opposing views of the cryo-EM binary reconstruction, resolved to ${\sim}3.4$ Å resolution.

(D) Opposing views of the binary atomic model derived from the cryo-EM density.

(E) Schematic representation of the crRNA used for ternary complex assembly. The DR and spacer of crRNA are colored as in (B); target RNA is red.
(F) Opposing views of the cryo-EM ternary reconstruction, resolved to ~3.3 Å resolution.
(G) Opposing views of the ternary atomic model derived from the cryo-EM density. See also Figures S1 and S2.

the DR within multiple REC lobe domains (Liu et al., 2017a), Cas13d features a more compact REC lobe, and the DR prominently protrudes from the effector. A similar compaction, coupled to solvent-exposure of crRNA and spacer: protospacer duplex, has been observed among other small Cas nucleases such as SaCas9 (Nishimasu et al., 2015).

Cas13d Maintains an Extensive Nucleoprotein Interface throughout the Length of Bound RNA

In both its binary and ternary forms, Cas13d makes extensive nucleoprotein interactions with spacer and complementary target protospacer, with all five protein domains contributing to RNA stabilization. Both complexes maintain a similar configuration of the DR (Figures 1D and 1G). In the binary complex, most of the 22 nt single-stranded spacer interfaces with key residues of Helical-1, Helical-2, HEPN1, and HEPN2 (Figure 2). The majority of interactions are composed of



Figure 2. Cas13d Nucleoprotein Interactions Reconfigure between Its Binary and Ternary Forms

Schematic of protein:DNA interactions within binary and ternary forms of Cas13d. Arrows point to the specific interactions. Standard amino acid and nucleotide abbreviations apply to all residues and bases. Conservation of interacting residues and types of interactions are indicated in the legend.

effectors. In the Cas13d binary structure, we identified multiple residues that interact in either a base- or backbonespecific manner with the DR of crRNA. (Figures 2, 3A, and 3B). The base-specific contacts are concentrated within the unpaired, conserved terminal nucleotides of the DR (nt 22-30) and include G22 and U23 from the 2 nt bulge region, as well as A26, A27, A28, and A29 within the 5' terminal nucleotides. The 2 nt bulge (nt 22-23) appears to be an invariant feature among type VI RNA-guided RNases. Mutagenesis of each of the six crRNA nucleotides forming base-specific contacts abolished Cas13d-mediated ssRNA cleavage, confirming the importance of these interactions for proper crRNA binding and positioning (Figures 3E and 3F).

Given the absence of base-specific contacts along the 5' region of the DR (nt 1-21), we reasoned that EsCas13d can utilize distinct crRNA of other Cas13d orthologs containing conserved 3' terminal nucleotides. As predicted, EsCas13d maintained full target cleavage activity with the UrCas13d cognate crRNA, which contains numerous DR mutations relative to the EsCas13d crRNA but maintains the necessary base-specific contacts. In contrast, crRNAs from RffCas13d and RfxCas13d were predicted to disrupt the critical base-specific G22 and G26 contacts; accordingly, target cleavage activity was abolished (Figure 3G). These data provide a structural basis for defining key base requirements and likely crRNA exchangeability across the Cas13d family, facilitating multiplexed effector applications.

Overall, interactions along the RNA duplex are sparser in ternary and exclusively confined to the RNA backbone (Figure 2).

Cas13d Recognizes the 5' Handle of crRNA

Specific recognition and binding of the constant 5' handle within their cognate crRNAs is a key requirement for all class 2 CRISPR

A prominent feature characterizing the nucleoprotein interface in both the binary and ternary Cas13d forms is the highly ordered, albeit irregularly shaped, ssRNA at the 3' end of the DR (Figures 2, 3A, and 4A). This region (nt 23–29) forms a hairpin loop surrounding a density visible in the cryo-EM map that is likely a centrally located Mg²⁺ ion in both binary (Figure 3A)



and ternary complexes (Figure 4A). While Mg²⁺ is indispensable for target cleavage for all Cas13 subtypes, it is generally not required for pre-crRNA processing by Cas13a and Cas13b. In contrast to a previous report of Mg²⁺-dependent array processing by Cas13d (Yan et al., 2018), we find that Mg²⁺ is not essential for pre-crRNA processing (Figures S3A–S3C). Rather, Mg²⁺ can increase processing efficiency at lower protein:pre-crRNA ratios likely by improving binding affinity of crRNA and Cas13d (Figure S3D), similar to recent observations for type V Cas12a/ Cpf1 (Swarts et al., 2017), highlighting its role across diverse class 2 effectors.

Target Duplex Formation and Lack of PFS Requirement

In the binary Cas13d surveillance complex, most of the singlestranded spacer is solvent exposed and structurally poised for base-pairing with potential targets to enter a ternary state (Figure 4B). For DNA-targeting class 2 CRISPR effectors, target interaction is initiated by protein-PAM (protospacer adjacent motif) interactions. Cas13d lacks the analogous protospacer flanking sequence (PFS) requirements (Konermann et al., 2018; Yan et al., 2018). In contrast, some Cas13a orthologs (including LshCas13a) have been reported to display a single base 3' H (non-G) PFS. This was previously proposed to be caused by base-pairing of the terminal conserved C(30) of crRNA DR with a complementary target G, which would then destabilize critical contacts with the HEPN-1 domain following rotation of the C(30) base away from the protein density (Figure S3E) (Liu et al., 2017b). In EsCas13d, the C(30) base is already rotated toward the target in its ternary form despite a mismatch with the target base (U) (Figure S3F). A complementary G nucleotide would

Figure 3. Recognition of crRNA

(A) Close-up view of the nucleoprotein interactions at the DR-spacer interface in the binary complex.
(B-D) Close-up view of the nucleoprotein interactions at the (B) 5' end, (C) middle region, and (D) 3' end of spacer RNA in the binary complex.
(E) Sequence logo of processed DR from seven Cas13d orthologs (top) with sequences of four selected orthologs shown (bottom). Star, protein:base interactions; dot, protein:backbone interactions. Red bases, mutations used in (F). Blue bases, critical residue differences in (G) between

orthologs. (F) Denaturing gel indicating *Es*Cas13d RNase activity with crRNAs carrying mutant DR. (G) Denaturing gel indicating *Es*Cas13d RNase activity with non-cognate crRNAs.

See also Figure S3.

therefore not be expected to cause any additional rearrangement, suggesting a structural rationale for the absence of a PFS requirement in Cas13d.

Cas13d Binding and Cleavage Are Interlinked

Given the lack of an overt PFS requirement, we sought to understand Cas13d

target binding and cleavage in the context of target complementarity. Class 2 CRISPR-Cas nucleases exhibit distinct binding and cleavage mechanisms. Cas9 from Streptococcus pyogenes exploits a sequential target-binding mechanism with stable binding after ~12 nt of PAM-proximal spacer complementarity (Dahlman et al., 2015; Kiani et al., 2015; Sternberg et al., 2014), while mismatches in this "seed" region are poorly tolerated (Hsu et al., 2013). Cas9 target cleavage, however, requires \sim 4 nt of additional matches to activate the HNH catalytic domain (Sternberg et al., 2015). In contrast to SpCas9, Cas12a/Cpf1 requires extended complementarity of at least 17 nt for stable binding (Singh et al., 2018). Cas13 binding and cleavage requirements are largely unclear, although a central seed region has been proposed for both Cas13a and Cas13b based on the observation that mismatches are least tolerated in the center of the spacer (Abudayyeh et al., 2016; East-Seletsky et al., 2016; Knott et al., 2017; Liu et al., 2017a; Smargon et al., 2017) and can affect HEPN domain activation (Tambe et al., 2018).

To investigate Cas13d target binding and cleavage complementarity, we conducted cleavage assays with a panel of target RNA competitors carrying 4 nt mismatches at different positions along the target sequence (Figure 4C; Table S2). We sought to distinguish between three possible scenarios when a competitor RNA is introduced at high molar excess relative to the target RNA: Cas13d (1) directly cleaves competitor, increasing fluorescent signal relative to target alone; (2) binds competitor without forming a catalytically active ternary complex, thereby sequestering Cas13d from target RNA and decreasing fluorescent signal; or (3) is unable to bind or cut competitor, with no resulting change in fluorescence.



Figure 4. Complementarity Requirements for Cas13d Ribonuclease Activation

(A) Close-up view of the nucleoprotein interactions at the DR-spacer interface in the ternary complex.

(B) Close-up view of the nucleoprotein interactions surrounding the spacer:protospacer duplex.

(C) Schematic of competitor cleavage assay describing three possible scenarios and expected results. MM, mismatched competitor target RNAs; NT, non-targeting competitors.

(D) RNA cleavage assay containing Cas13d, crRNA, and competitor RNAs only (no matching target). Signal is normalized to a negative control containing Cas13d and crRNA only. M, fully matched competitor. Each condition represents activity across four distinct target sequences (n = 3, mean ± SD).

(E) RNA cleavage assay containing Cas13d, crRNA, target RNA, and competitor RNA at $25 \times$ molar excess to target. Signal is normalized to the condition with Cas13d, crRNA, and target without competitor. Each condition represents activity across four distinct target sequences (n = 3, mean \pm SD).

(F) Illustration of parallel readout for Cas13d cleavage and binding in human cells using a fluorescent splicing reporter with the RfxCas13d-NLS (CasRx) ortholog. (G) CasRx cleavage and dCasRx splicing (binding) activity as a function of crRNA spacer length. Cleavage and splicing were each normalized to a non-targeting crRNA and represented relative to the activity of the full-length (22 nt) crRNA (n = 3, mean \pm SD). See also Figure S3.

In the absence of target, perfectly matched competitors triggered a robust increase in fluorescence as expected (Figure 4D). Competitors carrying 5'-proximal mismatches activated bystander cleavage in two out of four guides, indicating that complementarity in the DR-distal region of the crRNA spacer is not strictly required for ternary activation (scenario 1). Bystander activity was not activated by other competitors (scenario 3), suggesting a lack of a consistent seed region within the crRNA spacer for target cleavage. In the presence of target, we observed a similar pattern of Cas13d activity across all competitor mismatch positions to the target-free condition (Figure 4E). In particular, none of the mismatch competitors mediated a

decrease in fluorescence (in contrast to scenario 2), suggesting that stable target binding requires at least 18 nt of complementarity and that binding and cleavage are closely coupled in Cas13d.

To further explore the interdependence between nucleotide complementarity and cleavage efficiency, we tested a closely related Cas13d ortholog (*Rfx*Cas13d) in a cell-based reporter to assay for Cas13 knockdown and splicing as a proxy for RNA binding and cleavage activity (Konermann et al., 2018) (Figure 4F). Using a series of crRNA spacer truncations progressing from 22 nt to 10 nt, we observed a simultaneous decrease in knockdown and splicing as a function of decreasing spacer



length (Figure 4G). Half-maximal activity occurred between 18 and 20 nt spacer length.

Taken together, the lack of a clear seed is consistent with the observation of a solvent-exposed spacer stabilized in an accessible configuration along the open channel in the binary complex, positioning crRNA to initiate base-pairing with protospacer at multiple possible locations. Beginning with \sim 18 nt of complementarity, Cas13d likely undergoes a partial reconfiguration of its relevant nucleoprotein interface. Upon reaching complete 22 nt guide-target duplex complementarity, the enzyme transitions to a maximally active and cleavage-competent state.

Target RNA Binding Reconfigures Cas13d into an ssRNA Cleavage Complex and Allosterically Activates the HEPN Domains

The transition from the binary surveillance complex to the ternary complex activates Cas13d for RNA cleavage. Numerous confor-

(A) Schematic representation of Cas13d domain movement. Each vector represents a C α atom translation from its binary to ternary forms. Vector length corresponds to the translational distance. The directions of movement are indicated by arrows.

(B) Quantitative measurement of Cas13d domain movement (mean \pm SD). The residues that undergo largest translations in HEPN2 are highlighted with a dashed box.

(C) Superposition of the crRNA in the ternary complex (DR region in magenta and spacer in cyan) with that in the binary complex (in gray).

(D) Schematic representation of Cas13d crRNA movement. Vector representations as in (A). The position of the phosphate atom in the ternary structure are highlighted in red.

(E and F) Target RNA binding widens the central channel, (E) opens the cleft between NTD and Helical-1, and (F) increases distance between Helical-1 and HEPN2.

(G–I) Stabilization of the HEPN catalytic domain dimer. (G) Surface representation of Cas13d ternary complex, with HEPN catalytic residues highlighted in white. (H) Superposition of HEPN1 and HEPN2 and (I) a close-up view of the catalytic center. The four catalytic residues advance inward and a nearby loop is stabilized upon target binding.

See also Figure S4.

mational rearrangements occur during this transition, stabilizing the activated cleavage-competent state within the catalytic HEPN-domain dimer. HEPN domains function as obligate dimers, with two R-X₄₋₆-H motifs forming a bipartite active site to mediate RNA hydrolysis (Anantharaman et al., 2013). Target RNA binding to Cas13d triggers enzyme acti-

vation. The most dramatic polypeptide rearrangements occur within the Helical-1 domain, which shifts outward by an average C α -C α root-mean-square deviation (RMSD) of \sim 12 Å to accommodate target RNA binding within an expanded cleft (Figures 5A and 5B). To a lesser extent, portions of the HEPN1, Helical-2, and HEPN2 domains, particularly regions proximal to the 3' end of spacer (5' end of protospacer), undergo subtler structural changes. Whereas most of the DR remains unperturbed, the majority of the spacer reorganizes from a single-stranded pseudohelical conformation into a double-stranded A-form RNA helix (Figures 5C and 5D). Collectively, such rearrangements widen the RNA-binding cleft, ranging from \sim 13 Å at the narrowest point between NTD and Helical-1 (Figure 5E) and from 12 Å to 44 Å between Helical-1 and HEPN2 at the 3' end of crRNA (Figure 5F; Video S1). Mutagenesis of key residues forming critical crRNA and target RNA backbone contacts abolished target RNA cleavage but did not affect pre-crRNA processing



(Figures S3G and S3H), confirming their role in stabilizing the spacer:target duplex.

We also observe some intra-domain reorganization within Cas13d (Figure S4A). Most prominently, a kink is introduced into the C-terminal region of HEPN1-1, while the N-terminal loop of HEPN1-II (and, correspondingly, the C-terminal loop residues of Helical-1) reposition by an average Ca-Ca RMSD of ~4 Å. Both of these correspond to the two flexible linkers connecting Helical-1 to HEPN1-I and HEPN1-II, which reposition Helical-1 as it cradles incoming target RNA (Figures S4B and S4C).

The HEPN2 domain, which resides on the "back side" of Cas13d (Figure 5G), undergoes several rearrangements to facilitate target cleavage. A structural alignment of HEPN1-I between binary and ternary forms indicates the catalytic residues of HEPN2 (R849 and H854) reposition by \sim 4 Å, shifting closer to the corresponding catalytic residues of HEPN1-I (R295 and H300) (Figures 5H and 5I). Given that target RNA resides on the opposite side of the HEPN catalytic site, these rearrange-

Figure 6. Apo Cas13d Contains Mobile Domains that Are Stabilized upon RNA Binding

(A) Domain organization of full-length EsCas13d and the visible portion of apo EsCas13d by cryo-EM (see below). Domain boundaries and catalytic sites of the HEPN domains (R295, H300, R849, H854) are indicated. Undetectable regions in the apo Cas13d reconstruction (see [B]-[D]) are represented with dashed lines

(B) Cryo-EM reconstruction of apo EsCas13d, resolved to 6.5 Å resolution.

(C) Rigid-body docked model of EsCas13d based on the cryo-EM density.

(D) Superposition of the structured portion of apo EsCas13d evident within the cryo-EM reconstruction (surface representation) with that of EsCas13d in the binary form.

(E) Differential HDX colored onto the binary protein form of Cas13d. Regions that exhibit statistically significant changes in amide exchange are colored onto the structure and onto a domain organization (top). Insets show four distinct close-up views. White, no statistically significant difference; gray, no coverage in the differential HDX map. See also Figures S5 and S6.

ments suggest allosteric HEPN activation modulated by target RNA binding. Together, the structural changes serve to (1) accommodate target RNA binding and (2) reconfigure the catalytic site into its cleavage-competent form.

Apo Cas13d Utilizes Multiple Dynamic Domains to Form an RNA-Binding Cleft that Is Stabilized upon Guide Binding

To better understand the mechanism of crRNA binding, we examined the apo form of Cas13d. Using similar experi-

mental strategies as before, we obtained 154,889 particles for 3D classification and refinement. We observed variable density within the 2D class averages (Figure S5A), in contrast with the sharp signal present for both the binary and ternary datasets (Figure S1). Despite the apparent heterogeneity, an ab initio 3D reconstruction led to a map that was refined to \sim 6.5 Å resolution (Figures 6A-6B, S5C, and S5D; Table S1). Strikingly, the apo Cas13d reconstruction accounts for only part of the mass in comparison to both binary and ternary forms, with the majority of homogeneous density corresponding to a stable *a*-helical core (Figure 6C). To verify that the remaining density was present in the data, we derived 2D class averages ab initio and found that many of these captured protein density that was otherwise unresolved in the 3D reconstruction (Figure S5E).

We could readily assign the known domain boundaries by docking the binary model into the apo reconstruction. Portions of HEPN1, Helical-2, and HEPN2 (NUC lobe) account for the mass derived by cryo-EM, whereas the entire REC lobe (NTD and Helical-1) was invisible (Figure 6D). The domain organization of NUC lobe in the apo enzyme remains largely unchanged relative to the binary complex at this resolution. These data suggest that the REC lobe (NTD, Helical-1) and portions of HEPN2 may be dynamically arranged in the absence of RNA. Further, because cryo-EM is performed under solution conditions, the observed dynamics are likely an inherent property of the enzyme.

To understand Cas13d conformational dynamics between its apo and binary forms on a residue-by-residue level, we conducted hydrogen-deuterium exchange with mass spectrometry (HDX-MS) to probe for unstructured and flexible regions of Cas13d that undergo deuterium exchange more rapidly than those that are stable and hydrogen bonded (Englander, 2006). Individual exchange profiles of apo and binary Cas13d overall showed similar patterns at the sequence level (Figures S6A and S6B), indicating that the constituent domains, including the REC lobe, exhibit similar overall folds. However, several regions of Cas13d were clearly stabilized upon RNA binding, as indicated by reduced exchange in the differential profiles (Figure 6E). Three of these regions bind distinct segments within crRNA. Residues ~81–131 form part of the NTD, where residues K68, G85, and R139 make critical base-specific interactions with the 3' end of the DR (Figures 2, 3F, and 3G), while residues \sim 371–390 and \sim 594–613 form an RNA-binding interface with the phosphate backbones of the 5' and 3' ends of spacer RNA, respectively. Finally, residues ~325-360 represent the interface between HEPN1-I and Helical-1, with residues 335-340 forming a hinge-like structure connecting Helical-1 to the NUC lobe (Figure 6E).

These data, under solution conditions, indicate that both Cas13d lobes are appropriately folded in the apo configuration and that the REC lobe is likely to be mobile relative to the NUC lobe. 2D class averages indicate the presence of unresolved density in the 3D reconstruction, and HDX analysis indicates increased mobility of the NTD and the linker region (residues \sim 325–360) connecting Helical-1 to the NUC lobe. Upon RNA binding, the Cas13d binary complex is stabilized in multiple regions, and a central, positively charged RNA-binding cleft is formed between the REC and NUC lobes (Figure S6C). Analogous reconfigurations of NUC and REC lobes upon crRNA binding and formation of the central positively charged cleft in the binary complex have been reported for other class 2 CRISPR enzymes, including Cas9 (Jinek et al., 2014; Nishimasu et al., 2015), and may have implications for sampling distinct RNA features for facilitating efficient crRNA recognition and binary transition.

Structure-Guided Cas13d Truncations for Minimal Coding Sequence

The compact size of Cas13d is accompanied by an integration of multiple distinct functions into each individual constituent domain. Each of the protein domains within Cas13d contribute key protein:RNA contacts in addition to their structural and catalytic functions (Figure 2), with many conserved residues among Cas13d orthologs (Figure S7A). As a consequence, regions of high conservation between Cas13d orthologs are dispersed throughout the linear protein sequence and separated by only short stretches of low conservation. We predicted that all five domains of Cas13d would be essential for its RNase activity, in contrast to our previous demonstration that the REC2 domain of *Sp*Cas9 is largely dispensable for target DNA cleavage (Nishimasu et al., 2014). We designed six deletions in the closely related *Rfx*Cas13d ortholog (CasRx) and evaluated the ability of these truncation mutants to knock down a fluorescent reporter in human cells (Figures 7A and 7B). One deletion in Helical-2 (Δ 3), which is located on the external surface of Cas13d and avoids the removal of highly conserved residues, exhibited full activity relative to the wild-type enzyme.

A second round of seven small, surface-localized deletions on top of $\Delta 3$ successfully generated three additional variants with >95% knockdown activity ($\Delta 3.1$, $\Delta 3.3$, and $\Delta 3.7$, Figures 7B and 7C). The most active resulting variant ($\Delta 3.3$) removes 50 amino acids (aa), facilitating AAV-mediated CasRx delivery with increased flexibility (Figures S7B and S7C).

DISCUSSION

The structural, biochemical, and functional analysis of type VI *Es*Cas13d presented here reveals three distinct states of the Cas13d enzyme as it transitions from its inactive apo (Cas13d) to a surveillance (Cas13d-crRNA) and cleavage-competent (Cas13d-crRNA-target RNA) form. These transitions are accompanied by numerous structural rearrangements that accommodate and stabilize these enzymatic forms. Our data suggest a model whereby REC lobe dynamics within Cas13d may facilitate scanning for the crRNA for its proper recognition within a CRISPR array.

Cas13d is among the smallest CRISPR-Cas single effectors, with 20%-30% less mass than other type VI Cas13 endoRNases (Konermann et al., 2018; Yan et al., 2018). Compared to Cas13a (Knott et al., 2017; Liu et al., 2017a, 2017b), Cas13d compacts the analogous NTD and Helical-1 domains of Cas13a into a single, 150 aa NTD (Figure S2) within the REC lobe of Cas13d.

Upon satisfying multiple base-specific contacts within the 3' end of the DR, crRNA binding triggers stabilization of the binary complex and formation of a central positively charged cleft between the REC and NUC lobes of Cas13d. The solvent-exposed, single-stranded spacer region takes on a stabilized pseudo-helical conformation that appears to be poised for target binding at multiple positions. Stable RNA duplex formation requires \sim 18 nt of complementarity and triggers large conformational rearrangements primarily in Helical-1 and HEPN2 to activate the bipartite HEPN domain. Unlike some other type VI RNases, Cas13d does not require any PFS for target recognition, possibly mediated by the flipped C(30) base within crRNA that avoids base-specific pairing with target RNA. Upon Cas13d ternary formation, the HEPN catalytic residues within the HEPN1 and HEPN2 domains migrate toward one another to generate an external-facing active site. This facilitates direct RNA hydrolysis of both guide-complementary activator RNA and non-complementary bystander RNA for antiviral defense.

Although class 2 CRISPR-Cas nucleases share many common traits for nucleic acid sensing, they also vary substantially





Figure 7. Cas13d Truncations and Protein Conservation

(A) Schematic of reporter assay for CasRx-mediated RNA knockdown in human cells.

(B) Knockdown activity of CasRx deletions. Left: Consensus sequence of the multiple sequence alignment of Cas13d orthologs (Figure S7) with domain boundaries and sequence identity as indicated. Pink bars denote conserved protein:crRNA interactions. Deletion junctions are denoted as a linked interval. Knockdown activity is represented relative to non-targeting guide (mean \pm SD, n = 3). Deletions exhibiting >95% reduction of red fluorescent protein (RFP)⁺ cells are bolded, and the corresponding deletion junction is marked in blue.

(C) Surface view of the *Es*Cas13d ternary complex colored by conservation percentage. All truncated regions retaining high knockdown activity (blue) are located on external surface of Cas13d with low (<40%) conservation. Regions with high degree of conservation are concentrated in the RNA-binding cleft (left) as well as the HEPN active site (right). See also Figure S7.

exchangeability across distinct Cas13d orthologs, defining functionally orthogonal subfamilies that could be exploited to facilitate Cas13-based multiplexing applications in both cellular (Abudayyeh et al., 2017; Konermann et al., 2018) and cell-free systems (Gootenberg et al., 2017). Some Cas13d orthologs have accessory proteins (Yan et al., 2018) that enhance activity and could provide clues for improving Cas13d binding or cleavage. In analogy to engineered variants of Cas9 and related nucleases, structure-guided engineering

in size, shape, domain architecture, and organization. Within individual Cas types (e.g., type VI), there is often minimal sequence conservation across subtypes. Cas13d enzymes, for example, do not share sequence homology with Cas13a apart from the minimal 6 aa HEPN catalytic motifs, despite overall similarity of their RNase activities. Independent origins of type V subtypes from mobile genetic elements, as previously suggested for DNA-targeting type II and V effectors and subtypes via distinct TnpB transposase subfamilies (Koonin et al., 2017), may also explain the convergence and divergence of Cas13 superfamily function and structural organization.

Overall, our data elucidate the structural basis of Cas13d RNA-guided RNase activity and its compaction of these properties into a minimal effector size, providing a blueprint for improving Cas13d-based RNA-targeting tools. Further engineering of smaller Cas13d variants, as shown here, will enable flexible packaging into size-limited viral vectors with large regulatory elements for optimal transgene expression and activity. Furthermore, base-specific contacts of Cas13d with the 5' handle of crRNA were sufficient to delineate crRNA

of diverse CRISPR-Cas13d enzymes can be expected to enable improved properties for diverse biomolecular applications of RNA targeting.

STAR***METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Cell culture of Human Embryonic Kidney (HEK) cell line 293FT
- METHOD DETAILS
 - Protein expression and purification
 - Preparation of guide and target RNAs
 - Binary and ternary complex formation
 - Electron microscopy sample preparation and data acquisition

- Image processing of Cas13d binary and Cas13d ternary complex
- Image processing of apo Cas13d
- Model building and refinement
- Biochemical cleavage assays
- Cell-based reporter assays
- Hydrogen-deuterium exchange detected by mass spectrometry
- crRNA filter binding assays
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, two tables, and one video and can be found with this article online at https://doi.org/10.1016/j.cell. 2018.09.001.

ACKNOWLEDGMENTS

We thank Bill Anderson at TSRI and Youngmin Jeon at Salk for help with EM data collection and Bruce D. Pascal for assistance in analysis and rendering of HDX data. We also thank Tony Hunter as well as members of the Hsu and Lyumkis labs for critical reading of the manuscript. This work was supported by NIH-NCI CCSG P30 014195, the Howard Hughes Medical Institute Hannah H. Gray Fellowship (to S.K.), Helen Hay Whitney Foundation Fellowship (to X.W.), NIH DP5 OD021369 (to P.D.H.), NIH DP5 OD021396 and U54GM103368 (to D.L.), and the Helmsley Charitable Trust (to P.D.H. and D.L.).

AUTHOR CONTRIBUTIONS

C.Z. performed all cryo-EM data collections, image analyses, model building, and refinement. S.K. and P.D.H. purified samples, and S.K. generated the complexes. S.K., N.J.B., P.L., and P.D.H. performed the biochemical analyses and genetic analyses. S.J.N. and T.S. performed HDX analyses supervised by P.R.G. X.W. performed crRNA binding analyses. S.K., P.D.H., and D.L. conceived this project, and P.D.H. and D.L. supervised studies. C.Z., S.K., P.D.H., and D.L. interpreted the results and wrote the manuscript.

DECLARATION OF INTERESTS

S.K. and P.D.H. are inventors on patent applications relating to CRISPR-Cas13, as well as other patents on CRISPR technology.

Received: April 25, 2018 Revised: August 1, 2018 Accepted: September 4, 2018 Published: September 20, 2018

REFERENCES

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., et al. (2017). RNA targeting with CRISPR-Cas13. Nature *550*, 280–284.

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*, aaf5573.

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221.

Anantharaman, V., Makarova, K.S., Burroughs, A.M., Koonin, E.V., and Aravind, L. (2013). Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol. Direct 8, 15.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960–964.

Chalmers, M.J., Busby, S.A., Pascal, B.D., He, Y., Hendrickson, C.L., Marshall, A.G., and Griffin, P.R. (2006). Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. Anal. Chem. 78, 1005–1014.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

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

Dahlman, J.E., Abudayyeh, O.O., Joung, J., Gootenberg, J.S., Zhang, F., and Konermann, S. (2015). Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. Nat. Biotechnol. *33*, 1159–1161.

East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H., Tjian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature *538*, 270–273.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501.

Englander, S.W. (2006). Hydrogen exchange and mass spectrometry: A historical perspective. J. Am. Soc. Mass Spectrom. *17*, 1481–1489.

Garcia-Doval, C., and Jinek, M. (2017). Molecular architectures and mechanisms of Class 2 CRISPR-associated nucleases. Curr. Opin. Struct. Biol. 47, 157–166.

Gootenberg, J.S., Abudayyeh, O.O., Lee, J.W., Essletzbichler, P., Dy, A.J., Joung, J., Verdine, V., Donghia, N., Daringer, N.M., Freije, C.A., et al. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. Science *356*, 438–442.

Grant, T., and Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. eLife 4, e06980.

Grant, T., Rohou, A., and Grigorieff, N. (2018). *cis*TEM, user-friendly software for single-particle image processing. eLife 7, e35383.

Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell *139*, 945–956.

Han, Y., Donovan, J., Rath, S., Whitney, G., Chitrakar, A., and Korennykh, A. (2014). Structure of human RNase L reveals the basis for regulated RNA decay in the IFN response. Science *343*, 1244–1248.

Hohn, M., Tang, G., Goodyear, G., Baldwin, P.R., Huang, Z., Penczek, P.A., Yang, C., Glaeser, R.M., Adams, P.D., and Ludtke, S.J. (2007). SPARX, a new environment for Cryo-EM image processing. J Struct Biol. *157*, 47–55.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.

Jiang, W., Samai, P., and Marraffini, L.A. (2016). Degradation of Phage Transcripts by CRISPR-Associated RNases Enables Type III CRISPR-Cas Immunity. Cell *164*, 710–721.

Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science *343*, 1247997.

Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science *357*, 605–609.

Keppel, T.R., and Weis, D.D. (2015). Mapping residual structure in intrinsically disordered proteins at residue resolution using millisecond hydrogen/ deuterium exchange and residue averaging. J. Am. Soc. Mass Spectrom. *26*, 547–554.

Kiani, S., Chavez, A., Tuttle, M., Hall, R.N., Chari, R., Ter-Ovanesyan, D., Qian, J., Pruitt, B.W., Beal, J., Vora, S., et al. (2015). Cas9 gRNA engineering for genome editing, activation and repression. Nat. Methods *12*, 1051–1054.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. eLife 5, e18722.

Knott, G.J., East-Seletsky, A., Cofsky, J.C., Holton, J.M., Charles, E., O'Connell, M.R., and Doudna, J.A. (2017). Guide-bound structures of an RNA-targeting A-cleaving CRISPR-Cas13a enzyme. Nat. Struct. Mol. Biol. *24*, 825–833.

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

Koonin, E.V., Makarova, K.S., and Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. *37*, 67–78.

Lander, G.C., Stagg, S.M., Voss, N.R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P.W., et al. (2009). Appion: an integrated, database-driven pipeline to facilitate EM image processing. J. Struct. Biol. *166*, 95–102.

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

Liu, L., Li, X., Wang, J., Wang, M., Chen, P., Yin, M., Li, J., Sheng, G., and Wang, Y. (2017b). Two Distant Catalytic Sites Are Responsible for C2c2 RNase Activities. Cell *168*, 121–134.e12.

Lu, P., Bai, X.C., Ma, D., Xie, T., Yan, C., Sun, L., Yang, G., Zhao, Y., Zhou, R., Scheres, S.H.W., and Shi, Y. (2014). Three-dimensional structure of human γ -secretase. Nature *512*, 166–170.

Niewoehner, O., Garcia-Doval, C., Rostøl, J.T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L.A., and Jinek, M. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. Nature *548*, 543–548.

Niewoehner, O., and Jinek, M. (2016). Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. RNA *22*, 318–329.

Nishimasu, H., Cong, L., Yan, W.X., Ran, F.A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., and Nureki, O. (2015). Crystal Structure of Staphylococcus aureus Cas9. Cell *162*, 1113–1126.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949.

Pascal, B.D., Willis, S., Lauer, J.L., Landgraf, R.R., West, G.M., Marciano, D., Novick, S., Goswami, D., Chalmers, M.J., and Griffin, P.R. (2012). HDX workbench: software for the analysis of H/D exchange MS data. J. Am. Soc. Mass Spectrom. 23, 1512–1521.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem. *25*, 1605–1612.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Roseman, A.M. (2004). FindEM–a fast, efficient program for automatic selection of particles from electron micrographs. J. Struct. Biol. 145, 91–99.

Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. Cell *161*, 1164–1174.

Scheres, S.H. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. *180*, 519–530.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol. Cell *60*, 385–397.

Singh, D., Mallon, J., Poddar, A., Wang, Y., Tippana, R., Yang, O., Bailey, S., and Ha, T. (2018). Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). Proc. Natl. Acad. Sci. USA *115*, 5444–5449.

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

Snijder, J., Borst, A.J., Dosey, A., Walls, A.C., Burrell, A., Reddy, V.S., Kollman, J.M., and Veesler, D. (2017). Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids. J. Struct. Biol. *198*, 38–42.

Sternberg, S.H., LaFrance, B., Kaplan, M., and Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. Nature *527*, 110–113.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507, 62–67.

Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Leginon system. J. Struct. Biol. *151*, 41–60.

Swarts, D.C., van der Oost, J., and Jinek, M. (2017). Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. Mol Cell *66*, 221–233.e4.

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

Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., and Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nat. Methods *14*, 793–796.

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

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J. Struct. Biol. *193*, 1–12.

Zhang, Z., and Smith, D.L. (1993). Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Protein Sci. 2, 522–531.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

STAR***METHODS**

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli Rosetta2(DE3)	Novagen	Cat# 71400
Chemicals, Peptides, and Recombinant Proteins		
Tobacco Etch Virus (TEV) protease	Laboratory of Patrick Hsu	N/A
Isopropyl-b-D-thiogalactoside (IPTG)	US Biological Life Sciences	Cat# 18500
Salt active nuclease	Sigma Aldrich	Cat# SRE-0015
Pierce Protease-inhibitor tablets, EDTA-free	Thermo Fisher	Cat# A32965
Lipofectamine 2000	Thermo Fisher	Cat# 11668027
HyClone Fetal Bovine Serum (FBS)	GE Healthcare	Cat# SH30071.03 Lot# AAD202128
AminoallyI-UTP - ATTO-680	Jena Bioscience	Cat# NU-821-680
Agencourt RNAClean XP Beads	Beckmann Coulter	Cat# A63987
[α-32P] cordycepin 5'-triphosphate	Perkin Elmer	NEG026250UC
Turbo DNase	Life Technologies	Cat# AM2238
SYBR Gold Nucleic Acid Gel Stain	Life Technologies	Cat# S11494
Novex TBE-Urea Gels 10%	Life Technologies	Cat #EC68755BOX
Poly(A) Polymerase, Yeast	Life Technologies	74225Z25KU
Acid-Phenol:Chloroform, pH 4.5 (with IAA, 125:24:1)	Life Technologies	AM9720
GlycoBlue coprecipitant	Life Technologies	AM9516
Critical Commercial Assays		
MEGAclear Transcription Clean-Up Kit	Thermo Fisher	Cat# AM1908
RNaseAlert Lab Test Kit v2	Thermo Fisher	Cat# 4479768
HiScribe T7 High Yield RNA Synthesis Kit	NEB	Cat# E2040
Deposited Data		
Apo EsCas13d	This paper	EMDB: 9015
Binary EsCas13d (Cas13d-crRNA)	This paper	PDB: 6E9E, EMDB: 9013
Ternary EsCas13d (Cas13d-crRNA-target RNA)	This paper	PDB: 6E9F, EMDB: 9014
Experimental Models: Cell Lines		
Human: HEK293FT (Female)	Thermo Fisher	Cat # R70007
Oligonucleotides		
crRNA for ternary complex, See Table S2.	This paper, Synthego	N/A
Target RNA for ternary complex, See Table S2.	This paper, Synthego	N/A
crRNA and target RNA sequences for <i>in vitro</i> experiments, see Table S2.	This paper	N/A
Recombinant DNA		
EF1a-dCasRx-2A-EGFP plasmid for mammalian expression	Konermann et al., 2018	Addgene, #109050
CasRx gRNA cloning backbone plasmid DNA	Konermann et al., 2018	Addgene, #109053
RG6-mTagBFP2	Konermann et al., 2018	Adapted from Addgene #80167 from Thomas Cooper
His-MBP-TEV	Gift from Feng Zhang	N/A
Software and Algorithms		
Graphpad Prism 7	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
R	R Project	http://www.r-project.org/
FlowJo 10.3	FlowJo, LLC	https://flowjo.com/

Continued				
REAGENT or RESOURCE	SOURCE	IDENTIFIER		
Geneious 10.2	Biomatters	http://www.geneious.com/		
Image Studio	LI-COR Biosciences	https://www.licor.com/bio/products/software/ image_studio/		
Image Lab v 5.2	Bio-Rad Laboratories	http://www.bio-rad.com/en-us/product/image- lab-software?ID=KRE6P5E8Z		
СООТ	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/ pemsley/coot/		
PHENIX	Adams et al., 2010	http://www.phenix-online.org/		
PyMOL	Open-Source PyMOL 1.8.x	http://www.pymol.org/		
UCSF Chimera	Pettersen et al., 2004	https://www.cgl.ucsf.edu/chimera/		
Leginon	Suloway et al., 2005	http://leginon.org/		
Appion	Lander et al., 2009	http://appion.org/		
FindEM	Roseman, 2004	http://emg.nysbc.org/redmine/projects/software/ wiki/FindEM/		
MotionCor2	Zheng et al., 2017	http://msg.ucsf.edu/em/software/motioncor2.html		
CTFFind4	Rohou and Grigorieff, 2015	http://grigoriefflab.janelia.org/ctffind4/		
GCTF	Zhang, 2016	https://www.mrc-lmb.cam.ac.uk/kzhang/		
RELION	Kimanius et al., 2016; Scheres, 2012	https://www2.mrc-lmb.cam.ac.uk/relion/index. php?title=Main_Page/		
cisTEM	Grant et al., 2018	https://cistem.org/		
SPARX	Hohn et al., 2007	http://sparx-em.org/sparxwiki/Introduction/		
MolProbity	Chen et al., 2010	http://molprobity.biochem.duke.edu/		
3DFSC	Tan et al., 2017	https://3dfsc.salk.edu/		
HDX Workbench	Pascal et al., 2012	http://hdxworkbench.com/		
Other				
Plasma System SOLARUS	Gatan	Model 955		
Whatman Qualitative Filter Paper	GE Healthcare	Cat# 1001030		
UltrAuFoil R1.2/1.3 300 mesh Gold grid	Thermo Fisher Scientific	NC1277845		
Amersham Hybond-XL membrane GE Nylon	Thermo Fisher Scientific	45001147		
Amersham Protran NC 0.45um Nitrocellulose	Thermo Fisher Scientific	45004048		
Micro Bio-Spin P-30 Gel Columns, Tris Buffer (RNase-free)	Bio-Rad Laboratories	7326251		
Odyssey Clx Imaging System	LI-COR Biosciences	Clx		
Gel-Doc EZ System	Bio-Rad	1708270		
MACSQuant VYB	Miltenyi Biotec	130-096-116		

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dmitry Lyumkis (dlyumkis@salk.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture of Human Embryonic Kidney (HEK) cell line 293FT

The human embryonic kidney (HEK) cell line 293FT (female, Thermo Fisher) was cultured in DMEM (4.5 g/L glucose), supplemented with 10% FBS (GE Life Sciences) and 10 mM HEPES at 37°C with 5% CO₂. Cells were passaged before reaching 90% confluency using TrypLE Express (Life Technologies) at a ratio of 1:2. The cell line was purchased directly from Thermo Fisher and maintained within the lab for less than 20 passages total following purchase. It was not otherwise authenticated.

METHOD DETAILS

Protein expression and purification

Recombinant *Es*Cas13d proteins were cloned into a pET-based vector with an N-terminal His-MBP fusion and TEV protease cleavage site. The resulting plasmids were transformed into Rosetta2(DE3) cells (Novagen), induced with 200 µM IPTG at OD₆₀₀ 0.5, and grown for 20 hours at 18°C. Cells were then pelleted, freeze-thawed, and resuspended in Lysis Buffer (50 mM HEPES, 500 mM NaCl, 2 mM MgCl2, 20 mM Imidazole, 1% v/v Triton X-100, 1 mM DTT) supplemented with 1X protease inhibitor tablets, 1 mg/mL lysozyme, 2.5U/mL Turbo DNase (Life Technologies), and 2.5U/mL salt active nuclease (Sigma Aldrich). Lysed samples were then sonicated and clarified via centrifugation (18,000 x g for 1 hour at 4°C), filtered with 0.45 µM PVDF filter and incubated with 50 mL of Ni-NTA Superflow resin (QIAGEN) per 10 L of original bacterial culture for 1 hour. The bead-lysate mixture was applied to a chromatography column, washed with 5 column volumes of Lysis Buffer, and 3 column volumes of Elution Buffer (50 mM HEPES, 500 mM NaCl, 300 mM Imidazole, 0.01% v/v Triton X-100, 10% glycerol, 1 mM DTT). The samples were then dialyzed overnight into TEV Cleavage Buffer (50 mM Tris-HCl, 250 mM KCl, 7.5% v/v glycerol, 0.2 mM TCEP, 0.8 mM DTT, TEV protease) before cation exchange (HiTrap SP, GE Life Sciences) and gel filtration (Superdex 200 16/600, GE Life Sciences). Purified, eluted protein fractions were pooled and frozen at 4 mg/mL in Protein Storage Buffer (50 mM Tris-HCl, 1M NaCl, 10% glycerol, 2 mM DTT).

Preparation of guide and target RNAs

In vitro transcription template oligos carrying the T7 promoter were synthesized (IDT) and either annealed with an antisense T7 oligo for crRNAs or PCR amplified for targets and arrays. *In vitro* transcription was performed using the Hiscribe T7 High yield RNA synthesis kit (New England Biolabs) at 31°C for 12 hours. For *in vitro* cleavage reactions, targets were body-labeled through incorporation of AminoallyI-UTP-ATTO-680 (Jena Biosciences) during the *in vitro* transcription. crRNAs and short competitors were purified using RNAclean Agencourt AMPure XP beads (Beckman Coulter) with addition of 50% volume of isopropanol for retention of small RNAs. Longer targets and arrays were purified with the MEGAclear Transcription Clean-Up Kit (Thermo Fisher) and frozen at -80° C. The short 30nt target for ternary complex formation for cryo-EM imaging was synthesized by Synthego.

Binary and ternary complex formation

For cryo-EM binary complex formation, 200 µg *Es*Cas13d was incubated with a 3x molar excess of crRNA in complex formation buffer at 37°C for 1 hr (25mM Tris-HCl, 50mM NaCl, 1mM DTT, 1mM MgCl2, pH 7.5). The resulting binary complex was purified by size exclusion chromatography on a Superdex 200 16/600 column (GE Life Sciences) in S200 complex buffer (25mM Tris HCl, 100mM NaCl, 1mM DTT, 1mM MgCl2, 5% glycerol, pH 7.5). Binary peak fractions were pooled and concentrated to 1.5 mg/mL and processed for cryo-EM sample preparation. Ternary formation was performed sequentially, with 15 min of binary complex formation at 1:2 ratio of *dEs*Cas13d:crRNA followed by 45 min of incubation with target at 1:3 ratio (protein: target). Ternary size exclusion purification and concentration was performed analogous to the binary complex. For HDX binary sample preparation, the purification was scaled up 5X relative to the cryo-EM samples but was otherwise identical. Apo-*Es*Cas13d protein was buffer exchanged into S200 complex buffer and normalized to 1.5 mg/mL prior to Cryo-EM sample preparation.

Electron microscopy sample preparation and data acquisition

All samples, including binary, ternary, and apo, were concentrated to ~1.5 mg/mL prior to vitrification. In all 3 cases, Amphipol A8-35 was added to the sample to a final concentration of 0.1% (w/v) immediately before vitrification on cryo-EM grids, in order to ameliorate preferential specimen orientation, which was established in earlier attempts to collect the data (Lu et al., 2014). Cryo-EM grids were prepared under > 80% humidity at 4°C inside a cold room, and a multi-blotting approach was used to increase particle density (Snijder et al., 2017). Initially, 2ul of sample was applied to an UltrAuFoil R1.2/1.3 300-mesh grid (Quantifoil) after plasma-cleaning (75% argon/25% oxygen atmosphere, 15 W for 7 s using a Gatan Solarus). Next, the grid was side-blotted manually with a filter paper (Whatman No.1) followed by a second round of sample loading and side-blot. Finally, another 2ul sample was added to the grid and blotted immediately before plunging into liquid ethane using a manual plunger. Leginon was used for automated EM image acquisition (Suloway et al., 2005). Micrographs of Cas13d-apo and Cas13d binary complex were collected on a Talos Arctica microscope (FEI) operating at 200kV and equipped with a K2 Summit direct electron detector (Gatan). A nominal magnification of 57,000x was used for data collection, providing a pixel size of 0.73 Å at the specimen level, with a defocus range of -0.5 µm to -2.0 µm. Micrographs of Cas13d ternary complex were acquired on a Titan Krios microscope (FEI) operating at 300kV and equipped with a K2 Summit direct electron detector. A nominal magnification of 37,000x was used for data collection, corresponding to a pixel size of 0.79 Å at the specimen level, with the defocus ranging from $-1.0 \ \mu m$ to $-3.0 \ \mu m$ Movies were recorded in counting mode, with a total dose of ~57e- per Å² for all three samples and under a dose rate of ~2.5 - 3 electrons per pixel per second. All details corresponding to individual datasets are summarized in Table S1.

Image processing of Cas13d binary and Cas13d ternary complex

All pre-processing was performed within the Appion suite (Lander et al., 2009). Motion correction was performed using the program MotionCor2 (Zheng et al., 2017) and exposure-filtered in accordance with the relevant radiation damage curves (Grant and Grigorieff, 2015). For processing of Cas13d binary complex and Cas13d ternary complex datasets, structures of *Lbu*Cas13a-crRNA complex

(PDB:5XWY) and Cas13d-crRNA complex were used as the templates for automatic particle picking in Appion, respectively, using FindEM (Roseman, 2004). The Contrast transfer function (CTF) was estimated using CTFFind4 during data collection on whole micrographs (Rohou and Grigorieff, 2015). After selecting particle coordinates, per-particle CTF estimation was refined using the program GCTF (Zhang, 2016). Stacks containing 400K (Cas13d binary) and 680K (Cas13d ternary) particles were subjected to two rounds of 2D classification, followed by one round of 3D classification in GPU-enabled Relion (Kimanius et al., 2016; Scheres, 2012). The best classes containing 49K (Cas13d binary) and 52K (Cas13d ternary) particles were selected for Relion refinement. Lastly, the parameters were imported into *cis*TEM (Grant et al., 2018), and the last several rounds of orientation and per-particle CTF refinement were performed to improve the resolution by ~0.2 Å for the binary dataset and ~0.3 Å for the ternary data dataset. The spectral amplitudes for each reconstruction were flattened inside *cis*TEM between 8 Å and 3.4 Å or 3.3 Å for the binary and ternary complexes, respectively. The resolutions for both maps were evaluated using conventional Fourier Shell Correlation analysis to evaluate global resolution and directional Fourier Shell Correlation analysis to obtain 3D FSCs and evaluate directional resolution anisotropy (Tan et al., 2017). Due to the manner in which the particles adhered to the air-water interface, the ternary map is characterized by more anisotropic directional resolution.

Image processing of apo Cas13d

Cryo-EM data was processed in a conceptually similar manner as in binary/ternary. The same templates used for particle picking of Cas13d binary complex were also used to select 330,986 particles from the apo Cas13d dataset. After CTF estimation in GCTF (Zhang, 2016)) and 2 rounds of Relion 2D classification to remove bad particles, the extracted stack containing 154,889 particles was imported into cryoSPARC for *ab initio* reconstruction. We used the following parameters in the reconstruction: Number of Ab-initio classes = 1, Initial resolution = 20 and Maximum resolution = 5, which resulted in a map with clearly distinguishable secondary structure elements from 16K particles. Numerous other attempts were performed to obtain an *ab initio* reconstruction, but the particle heterogeneity from the large amount of conformational flexibility precluded our ability to improve the map or use a greater subset of particles at this stage. After a map was generated, the orientations were imported into *cis*TEM (Grant et al., 2018), and the orientations, as well as per-particle CTF parameters were refined for several rounds, resulting in a ~0.2 Å increase in resolution. The final global resolution was estimated at 6.5 Å. Further attempts to classify the data, either through *cis*TEM, or through other processing packages, did not result in visual improvements to the map. We believe that the reasons for the challenges within this dataset have to do with the heterogeneity associated with the apo form of the protein, coupled to its small size. The homogeneous part of the protein only accounts for ~60 kDa of total mass.

Model building and refinement

The model of Cas13d binary complex was built de novo in Coot (Emsley et al., 2010). A poly-ala model with gaps in looped region was first built based on the EM density, and then residues having bulky side chains (Phe, Trp, and Tyr) were registered to facilitate sequence assignment of the remaining protein. The register of crRNA was conducted based on prior knowledge: that the DR region will form a base-paired stem-loop structure and that the spacer is single stranded. This allowed for unambiguous registration of the N-terminal residues 1-57, as well as certain loops scattered throughout the structure and the C-terminal residues 950-954 were poorly ordered, and were thus omitted from the final model. Most of these regions are not conserved among Cas13d orthologs, with ~50% of Cas13d orthologs missing the N-terminal residues (Konermann et al., 2018). For building the model of the Cas13d ternary complex, the binary model was first docked into the ternary cryo-EM map and individual domains were repositioned according to the relevant conformational rearrangements. The NTD, HEPN1, HEPN2 domains and DR of crRNA remain constant, whereas the Helical-1 and Helical-2 domains, as well as the crRNA spacer required repositioning. All connecting loops and any atoms outside of density were rebuilt accordingly. Watson-Crick base pairing between the spacer and target protospacer allowed unambiguous RNA registration. Each model was independently refined in PHENIX (Adams et al., 2010) using phenix.real_space_refine against separate EM half-maps with geometrical, secondary structure, and hydrogen bond restraints. The maps were refined into a working half-map, and improvement of the model was monitored using the free half map. The geometry parameters of the final models were validated in Coot and using MolProbity (Chen et al., 2010). These refinements were performed iteratively until no further improvements were observed. The Cas13d-apo model was generated by rigid-body docking of the Cas13d binary complex structure into the Cas13d-apo cryo-EM map without further refinement or modification, and the parts of the model having poor densities were removed. All the structure figures were prepared in Pymol and UCSF Chimera.

Biochemical cleavage assays

For DR mutant analysis, purified *Es*Cas13d protein and guide RNA were mixed at 2:1 molar ratio in RNA Cleavage Buffer on ice (25mM Tris pH 7.5, 15mM Tris pH 7.0, 1mM DTT, 6mM MgCl2). Protein and and guide RNA were incubated at 37°C for 15 min for binary complex formation prior to the addition of target at 1:2 molar ratio relative to *Es*Cas13d. For pre-crRNA cleavage reactions, purified *Es*Cas13d and *Es*dCas13d proteins were mixed with purified pre-crRNA at 0.5:1, 1:1 and 5:1 molar ratios in RNA cleavage buffer containing 6mM MgCl2 or EDTA. Reactions were prepared on ice and incubated 37°C for 1 hour. Both cleavage reactions were quenched with 1 μ L of enzyme stop solution (10 mg/mL Proteinase K, 4M Urea, 80mM EDTA, 20mM Tris pH 8.0) at 37°C for 15 min. The *in vitro* cleavage reaction was finally denatured in 2X RNA loading buffer (2X: 13mM Ficoll, 8M Urea, 25 mM EDTA), at 85°C for 10 min. Cleavage products were separated on a 10% TBE-Urea gel (Life Technologies). Reactions containing and unlabeled guide

RNA and a fluorescently labeled target (Atto-680, Jena Biosciences) were visualized on the Odyssey Clx Imaging System (Li-Cor); pre-crRNA cleavage gels were stained with SYBR Gold prior to imaging via Gel Doc EZ system (Bio-Rad).

For collateral fluorescent ssRNA reporter assays (Figures 4C–4E), *Es*Cas13d protein was mixed with guide RNA at 1:1 ratio in RNA cleavage buffer on ice and then assembled into protein:guide complexes at 37°C for 15 min. Reactions were put on ice and competitor RNAs containing 4nt mismatches at different positions were mixed in at 25X molar ratio to target RNA. Target RNAs were added at a 1:5 ratio to *Es*Cas13d, and 150nM RNase-Alert substrate (Thermo-Fisher) was added and mixed. Reactions were then incubated in a real-time PCR machine (Bio-Rad, CFX384 Real-Time System) for 180 min at 32°C and measurements were taken every 5 min. Fluorescence values are an average of the last 5 measurements for each condition.

Cell-based reporter assays

Engineered *Rfx*Cas13d mutants were cloned into pXR002: EF1a-dCasRx-2A-EGFP (Addgene #109050) and prepared using the Nucleobond Xtra Midi EF Kit (Machery Nagel) according to the manufacturer's protocol. RfxCas13d gRNAs with variable spacer lengths were cloned into pXR003: CasRx gRNA cloning backbone (Addgene #109053) by golden gate assembly.

HEK293FT cells were transfected in 96-well format with 200 ng of Cas13d expression plasmid, 200 ng of guide expression plasmid, and 20 ng of the bichromatic reporter plasmid with Lipofectamine 2000 (Life Technologies). Cells were harvested in FACS Buffer (1X DPBS^{-/-}, 0.2% BSA, 2 mM EDTA) after 72 hours, then analyzed in 96-well plate format using a MACSQuant VYB (Miltenyi Biotec) followed by analysis using FlowJo 10. RG6 was a gift from Thomas Cooper (Addgene plasmid # 80167) and modified to replace EGFP with mTagBFP2. All represented samples were assayed with three biological replicates.

Hydrogen-deuterium exchange detected by mass spectrometry

Differential hydrogen-deuterium exchange mass sepctrometry (HDX-MS) experiments were conducted as previously described with a few modifications (Chalmers et al., 2006).

Peptide Identification: Peptides were identified using tandem MS (MS/MS) with an Orbitrap mass spectrometer (Fusion Lumos, ThermoFisher). Product ion spectra were acquired in data-dependent mode with the top five most abundant ions selected for the product ion analysis per scan event. The MS/MS data files were submitted to Mascot (Matrix Science) for peptide identification. Peptides included in the HDX analysis peptide set had a MASCOT score greater than 20 and the MS/MS spectra were verified by manual inspection. The MASCOT search was repeated against a decoy (reverse) sequence and ambiguous identifications were ruled out and not included in the HDX peptide set.

HDX-MS analysis: Cas13d (10 μ M) was incubated with or without guide RNA at a 1:5 protein-to-RNA molar ratio for 1 h at room temperature. Next, 5 μ L of sample was diluted into 20 μ L D₂O buffer (50 mM Phosphate, pH 7.5; 150 mM NaCl; 2 mM TCEP) and incubated for various time points (0, 10, 30, 90, 300, 900, and 3600 s) at 4°C. The deuterium exchange was then slowed by mixing with 25 μ L of cold (4°C) 6 M urea, 200 mM TCEP and 1% trifluoroacetic acid. Quenched samples were immediately injected into the HDX platform. Upon injection, samples were passed through an immobilized pepsin column (1mm × 2cm) at 50 μ L min⁻¹ and the digested peptides were captured on a 1mm × 1cm C₈ trap column (Agilent) and desalted. Peptides were separated across a 1mm × 5cm C₁₈ column (1.9 μ L Hypersil Gold, ThermoFisher) with a linear gradient of 4% - 40% CH₃CN and 0.3% formic acid, over 5 min. Sample handling, protein digestion and peptide separation were conducted at 4°C. Mass spectrometric data were acquired using an Orbitrap mass spectrometer (Fusion Lumos, Thermo Fisher). HDX analyses were performed in triplicate, with single preparations of Cas13d and the Cas13d-gRNA complex. The intensity weighted mean m/z centroid value of each peptide envelope was calculated and subsequently converted into a percentage of deuterium incorporation. This is accomplished determining the observed averages of the undeuterated and fully deuterated spectra and using the conventional formula described elsewhere (Zhang and Smith, 1993). Statistical significance for the differential HDX data is determined by an unpaired t test for each time point, a procedure that is integrated into the HDX Workbench software (Pascal et al., 2012). Corrections for back-exchange were made on the basis of an estimated 70% deuterium recovery, and accounting for the known 80% deuterium content of the deuterium exchange buffer.

Data Rendering: The HDX data from all overlapping peptides were consolidated to individual amino acid values using a residue averaging approach. Briefly, for each residue, the deuterium incorporation values and peptide lengths from all overlapping peptides were assembled. A weighting function was applied in which shorter peptides were weighted more heavily and longer peptides were weighted less. Each of the weighted deuterium incorporation values were then averaged to produce a single value for each amino acid. The initial two residues of each peptide, as well as prolines, were omitted from the calculations. This approach is similar to that previously described (Keppel and Weis, 2015).

crRNA filter binding assays

crRNA oligos were synthesized by Synthego, resuspended in RNase-free water at 100 μ M, and 3' end labeled with radioactive cordycepin using yeast poly(A) polymerase in a 10 μ L reaction (1 pmol crRNA oligo, 2 pmol [α -32P] cordycepin 5'-triphosphate, 2 μ L 5X yeast poly(A) polymerase buffer, 1 μ L yeast poly(A) polymerase, 5 μ L RNase-free water). The reaction was incubated at 37°C for 30 min, then 40 μ L RNase-free water was added to increase volume and the 50 μ L sample was passed through a P-30 column to remove free cordycepin. The RNA was extracted with 300 μ L acid-phenol:chloroform and precipitated in 700uL 100% ethanol with 30 μ L 3M Sodium Acetate, pH 5.5, and 1 μ L GlycoBlue (15 mg/mL). The pellet was resuspended in 100 μ L RNase-free water, containing ~10 nM 3' end labeled RNA.

Each binding reaction was performed in 10 μ L containing 1 μ L 3' end labeled crRNA (final concentration ~1 nM), 1 μ L Cas13d at various concentrations, and 1x cleavage buffer (25mM Tris pH 7.5, 15mM Tris pH 7.0, 1mM DTT, 6mM MgCl2 or 6 mM EDTA). Reactions were incubated for 2 hours at 37°C and then filtered through stacked nitrocellulose and nylon membranes. Circular membranes (0.5-inch diameter) were punched from stock, pre-equilibrated with 1x cleavage buffer, and stacked with the nitroceullulose membrane atop the nylon membrane onto the internal pedestal of a Whatman filter holder (Sigma Aldrich #WHA420100) that was inserted into a closed valve of a Visiprep vacuum manifold (Sigma Aldrich #57250-U). For filter binding, 100 μ L of 1x cleavage buffer was applied to the top filter, the valve was opened, the binding reaction was applied, and the membrane stack was immediately washed with 100 μ L ice-cold 1x cleavage buffer. The two membranes were then allowed to air dry and later separated and exposed to a phosphorimaging screen overnight. The detection and quantification was done using Typhoon and ImageQuant, respectively. Curve fitting and K_D calculation was done using Prism 7.

QUANTIFICATION AND STATISTICAL ANALYSIS

All values reported are the average at least three independent replicates from separate cleavage or reporter experiments with the exact number of replicates indicated in the individual figure legends. HDX analyses were performed in triplicate, with single preparations of each purified protein/complex. Error bars represent SEM or SD, also as indicated in the individual legends. For determination of significant differences, one-way ANOVA was used. Statistical significance for the differential HDX data is determined by t test for each time point, and is integrated into the HDX Workbench software. No *a priori* sample size estimation for statistical power was performed.

DATA AND SOFTWARE AVAILABILITY

The cryo-EM structures for corresponding to the binary, and ternary, and apo forms of Cas13d have been deposited into Electron Microscopy Databank under accession codes EMDB: EMD-9013, EMD-9014, and EMD-9015, respectively. The models for the binary and ternary forms have been deposited into the Protein Databank under accession codes PDB: 6E9E and 6E9F, respectively.

Supplemental Figures



Figure S1. Cryo-EM Data Collection, Analysis, and Modeling of Cas13d-crRNA Binary Complex and Cas13d-crRNA-Target RNA Ternary Complex, Related to Figure 1

(A) Representative 2D class averages of Cas13d-crRNA complex calculated using Relion.

(B) Classification strategy and refined maps of Cas13d-crRNA complex, colored by local resolution, calculated using "sxlocres.py" implemented within Sparx. The central RNA-binding groove is characterized by a higher resolution (closer to \sim 3 Å) than the peripheral regions of the protein (closer to \sim 4 Å).

(C) Fourier Shell Correlation (FSC) curves for cross-validation between the maps and models of Cas13d-crRNA complex. Curves calculated between two half maps (red), between the model and the model and the working half map used for model refinement (blue), between the model and the free half map used for validation (magenta), and between the model and the full map (green).

(D) Euler angle distribution plot (left) showing the relative orientation of the particles used in the final 3D reconstruction of Cas13d-crRNA complex.3D FSC isosurfaces (right), thresholded at a cutoff of 0.75 and displayed in three axial orientations, describe the directional resolution (isotropy) of the refined map. (E) EM map carved around crRNA and Mg2+ of Cas13d-crRNA complex

(F) Representative 2D class averages of Cas13d-crRNA-target RNA complex calculated using Relion.

(G) Classification strategy and refined maps of Cas13d-crRNA-target RNA complex, colored by local resolution, calculated using "sxlocres.py" implemented within Sparx. The central RNA-binding groove is characterized by a higher resolution (closer to ~3 Å) than the peripheral regions of the protein (closer to ~4 Å).

⁽H) Fourier Shell Correlation (FSC) curves for cross-validation between the maps and models of Cas13d-crRNA-target RNA complex. Curves calculated between two half maps (red), between the model and the working half map used for model refinement (blue), between the model and the free half map used for validation (magenta), and between the model and the full map (green).

⁽I) Euler angle distribution plot (left) showing the relative orientation of the particles used in the final 3D reconstruction of Cas13d-crRNA-target RNA complex.3D FSC isosurfaces (right), thresholded at a cutoff of 0.75 and displayed in three axial orientations, describe the directional resolution (isotropy) of the refined map. (J) EM map carved around crRNA, target RNA and Mg2+ of Cas13d-crRNA-target RNA complex



Figure S2. Bilobed Structures of Representative Class 2 CRISPR Effectors, Related to Figure 1

(A) Domain organizations of Class 2 CRISPR effectors: SpCas9, SaCas9, AsCas12a, AacCas12b, LbuCas13a and EsCas13d. The Helical-1 domain of Cas13a has been redefined here to include Cas13 regions The REC and NUC domains are colored wheat and light blue, respectively.

Cell

⁽B–I) Surface representations of crystal structures of (B) SpCas9 (PDB: 4008), (C) SaCas9 (PDB: 5CZZ), (D) AsCas12a (PDB: 5B43), (E) AacCas12b (PDB: 5U31) (F) cryo-EM structure of *Lbu*Cas13a in the binary form (PDB: 5XWY), (G) crystal structure of *Lbu*Cas13a in the ternary form (PDB:5XWP), (H) *Es*Cas13d in binary form, (I) *Es*Cas13d in ternary forms. Nucleic acid components and orthogonal views of each structure are shown. The REC lobe of *Lbu*Cas13a was proposed to include NTD, Helical-1, and Helical-2, and the NUC lobe to include HEPN1, Helical3/Linker and HEPN2 domains. In our representation, to be consistent with grouping on each side of crRNA, we defined the REC lobe of *Lbu*Cas13a as including NTD and Helical-2, whereas the NUC lobe would contain Helical-1, HEPN1, HEPN2, and Helical3/Linker. In *Es*Cas13d, The NTD and Helical1 domains are grouped into the REC lobe, and its NUC lobe contains HEPN1, Helical-2 and HEPN2 domains.



Figure S3. Pre-crRNA and Target RNA Cleavage Properties of Cas13d, Related to Figures 3 and 4

(A) Cas13d pre-crRNA cleavage is enhanced in the presence of Mg^{2+} . Denaturing gel of *Es*Cas13d pre-crRNA array cleavage in the presence (M, Mg^{2+}) or absence (E, EDTA) of 6 mM Mg^{2+} given an increasing ratio of protein:pre-crRNA (0.5:1, 1:1 and 5:1 molar ratios of *Es*Cas13d to pre-crRNA) at a 1 hour time point. Red arrowheads indicate mature crRNA 1, which requires two cleavage events to be generated and is abolished by the addition of EDTA at lower protein ratios. Catalytic HEPN residues are not required for pre-crRNA processing as *dEs*Cas13d exhibits equivalent activity as active *Es*Cas13d. (B) Possible cleavage products are indicated as i-v.

(C) Denaturing gel depicting time course of pre-crRNA processing and sensitivity of mature crRNA generation to EDTA. Pre-crRNA processing is conducted at a limiting protein:effector ratio (1:1). In the absence of Mg²⁺, processing of mature crRNA 1 is largely eliminated, while processing of crRNA 2 is only minimally delayed.

(D) Filter binding assays were conducted to determine the binding affinity of Cas13d to mature crRNA 1 in the presence or absence of magnesium. Measured K_D (mean ± SEM, n = 3) were 25.6 ± 5.3 (crRNA-Mg2+) and 262.1 ± 76.3 (crRNA-EDTA).

(E and F) Structural rationale for the 3' PFS requirement in *Lbu*Cas13a and *Es*Cas13d. Close-up views of 3' PFS region of *Lbu*Cas13a (E) and *Es*Cas13d (F). Key interactions are shown as dashed lines. Unlike *Lbu*Cas13a, whose nt-30 of crRNA DR (magenta) faces toward the HEPN-1 domain, nt-30 of crRNA DR in *Es*Cas13d faces toward the solvent and is free to make base-pairing interactions with incoming target RNA. These would be treated as normal spacer:protospacer interactions or, as in this structure, non-Watson-Crick base-pairs.

Cell

⁽G and H) Point mutations of conserved spacer:target RNA interacting residues in *Es*Cas13d impair ssRNA target cleavage but not pre-crRNA processing. Denaturing gel (G) of ssRNA target (arrowhead) cleavage by *Es*Cas13d wild-type (+) and mutant (1-3) proteins, guided by crRNA or pre-cRNA. "1" = N86A/T524A/ N641A, "2" = R386A/R679A/Y680A, "3" = K376A/K443A/Y447A. (H) Denaturing gel of pre-crRNA array cleavage by *Es*Cas13d wild-type and triple mutant proteins. Full length pre-crRNA (i) is cleaved into smaller products (ii-v) as in (B).



Figure S4. Domain Movement of EsCas13d from Binary to Ternary Transition, Related to Figure 5

- (A) Superposition of individual *Es*Cas13d domains between binary and ternary forms.
- (B) Local rearrangement of the hinged region that connects the HEPN1 and Helical-1 domains.
- (C) Superposition of the Helical-1 and HEPN1 domains between binary and ternary forms. The structures are aligned to HEPN1.



Figure S5. Cryo-EM Data Collection, Analysis, and Modeling of Apo EsCas13d, Related to Figure 6 (A) Representative 2D class averages calculated using Relion. Scale bar, 100 nm.

(B) Ab initio cryo-EM reconstruction colored by local resolution, calculated using "sxlocres.py" implemented within Sparx.

(C) Euler angle distribution plot (left) showing the relative orientation of the particles used in the final 3D reconstruction. 3D FSC isosurfaces (right), thresholded at a cutoff of 0.75 and displayed in three axial orientations, describe the directional resolution (isotropy) of the refined map.

(D) Fourier Shell Correlation (FSC) curves calculated between two half maps.

(E) Comparison of *ab initio* 2D class averages calculated using *cis*TEM from the raw data with 2D projections from the apo-*Es*Cas13d model, calculated along optimally assigned Euler angles. Densities that are missing in the 2D projections are often seen within *ab initio* 2D classes averages and are indicated by red arrows.



Figure S6. HDX and Electrostatic Potential Maps for EsCas13d, Related to Figure 6

(A) Percentage of deuterium incorporation was consolidated from peptides and plotted as a function of residue number. The amount of deuterium incorporation for each residue, for apo and binary samples, is colored red and black, respectively. Regions with high Deuterium incorporation are indicated.

(B) The deuterium uptake for individual apo or binary samples is mapped onto the respective structural models. Blue and red colors represent low and high deuterium uptake.

(C) Electrostatic surface maps of EsCas13d in its apo, binary and ternary states. The RNA components are removed for clarity. Red and Blue represent negatively and positively charged regions, respectively.

Cell



Figure S7. Multiple Sequence Alignment and Viral Packaging Schematic of Cas13d, Related to Figure 7

(A) Sequence alignment was conducted for seven Cas13d sequences: Es, Rfx, Adm, P1E0, Ur, Rff and Ra. Residue conservation is indicated by gray-scale shading according to Blosum62. Secondary structural elements observed in *Es*Cas13d binary structure are shown above the sequence. Residues that interact with nucleic acids in the binary and ternary states are labeled with blue and red bars, respectively. The same color scheme as Figure 6 was used to highlight the differential HDX onto *Es*Cas13d protein residues. Magenta triangles, catalytic residues of HEPN domains.

⁽B) Five CasRx deletion variants from Figure 7 exhibiting minimal cleavage activity were tested on the splicing reporter illustrated in Figure 4F to assay for binding activity. All variants displayed reduced or minimal splicing activity compared to full-length dCasRx. This finding is in agreement with close association of binding and cleavage activities in Cas13d (Figure 4) and the participation of HEPN1 and HEPN2 in target binding.

⁽C) AAV construct designs with $\Delta 3.3$ CasRx and a full-length WPRE post-transcriptional element for enhanced transgene expression with payload size < 4.7 kb, the AAV packaging limit.