Structure and mechanism of the RNA dependent RNase Cas13a from *Rhodobacter capsulatus*

Leonhard M. Kick\(^1\), Marie-Kristin von Wrisberg\(^1\), Leander S. Runtsch\(^1\) & Sabine Schneider\(^1\,

Cas13a are single-molecule effectors of the Class II, Type VI family of CRISPR-Cas systems that are part of the bacterial and archaeal defense systems. These RNA-guided and RNA-activated RNA endonucleases are characterized by their ability to cleave target RNAs complementary to the crRNA-spacer sequence, as well as bystander RNAs in a sequence-unspecific manner. Due to cleavage of cellular transcripts they induce dormancy in the host cell and thus protect the bacterial population by aborting the infectious cycle of RNA-phages. Here we report the structural and functional characterization of a Cas13a enzyme from the photo-auxotrophic purple bacteria *Rhodobacter capsulatus*. The X-ray crystal structure of the RcCas13a-crRNA complex reveals its distinct crRNA recognition mode as well as the enzyme in its contracted, pre-activation conformation. Using site-directed mutagenesis in combination with mass spectrometry, we identified key residues responsible for pre-crRNA processing by RcCas13a in its distinct catalytic site, and elucidated the acid-base mediated cleavage reaction mechanism. In addition, RcCas13a cleaves target-RNA as well as bystander-RNAs in *Escherichia coli* which requires its catalytic active HEPN (higher eukaryotes and prokaryotes nucleotide binding) domain nuclease activity. Our data provide further insights into the molecular mechanisms and function of this intriguing family of RNA-dependent RNA endonucleases that are already employed as efficient tools for RNA detection and regulation of gene expression.
Type VI CRISPR-Cas systems are characterized by their large single-protein effector nucleosome with a composite active site consisting of two highly divergent higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domains. The type VI CRISPR-Cas systems most likely evolved from toxin-antitoxin systems through insertion of a HEPN-encoding element/gene next to a CRISPR array, duplication of HEPN domain and multiple insertions that increased specificity of crRNA binding as well as multiple independent acquisitions of the adaptation module. These systems can be found in a number of bacterial species with Cas13a (formerly C2c2) as the hallmark enzyme, but share an overall low sequence identity of 20–30%. All so far characterized Cas13-family members possess two distinctive active sites, one for pre-crRNA processing and one for target RNA cleavage. They become activated by sequence-specific binding of a single-stranded target RNA to the crRNA, which induces a conformational change and alignment of the catalytic tetrad, in the composite HEPN domains, resulting in cleavage of the target ssRNA (cis-cleavage) and sequence-unspecific bystander ssRNAs (trans-cleavage), which has been exploited for the in situ detection of viral RNAs. In Listeria it was shown that activation of Cas13a leads to trans-cleavage of mRNA transcripts, that in turn leads to induction of dormancy in the host, which is sufficient to abort lytic infection and phage propagation in a cellular population. This resembles the strategy of abortive infection systems where dormant enzymes are activated upon phage replication, resulting in inactivation of the cellular translation or transcription machinery, protein phosphorylation or alteration of the membrane integrity, leading to growth arrest or cell death, thus arresting phage infection and spreading. Moreover, phages with target mutations that evade DNA-targeting by CRISPR-Cas systems are neutralized by Cas13a when activated by wild type phages. Therefore, CRISPR-Cas13a systems prevent outbreaks of CRISPR-resistant phages by acting on the host and not directly on virus.

Here we provide data on the structure and mechanisms of the Cas13a-endonuclease from the photo-auxotrophic bacteria Rhodobacter capsulatus (RcCas13a), which shares ~20–25% sequence identity with the to date structurally characterized homologous enzymes from Lachnospiraceae bacterium (Lba), Leptotrichia buccalis (Lbu), and Listeria seeligeri (Lse). The X-ray crystal structure of RcCas13a in complex with its crRNA elucidates how RcCas13a specifically recognizes its crRNA and shows the enzyme in a contracted, pre-activation conformation. In order to determine amino acid residues responsible for pre-crRNA processing by RcCas13a in the not conserved likely processing catalytic site, we employed site-directed mutagenesis in combination with in vitro activity assays and mass spectrometry. This allowed us not only to identify key residues responsible for pre-crRNA processing by RcCas13a, but also to elucidate the reaction mechanism. Moreover, we could show that RcCas13a can be activated in Escherichia coli, where it cleaves target RNA as well as bystander-RNAs, which requires a catalytic active domain. In summary, this study provides further insight into the molecular mechanisms of this intriguing family of RNA-dependent RNA endonucleases.

**Results and discussion**

**Overall structure of the RcCas13a-crRNA binary complex.** The reference genome of *R. capsulatus* contains three CRISPR-Cas systems: Two Class I systems (Type I and Type III) as well as one Class II, Type VI system (Supplementary Fig. S1), with the Cas13a-signature sequence encoding a 145 kDa protein sharing the family-typical HEPN-endonuclease domains, with the conserved Arg-His-Arg-His catalytic tetrad (Fig. 1a and Supplementary Fig. S2). We expressed RcCas13a alone and together with a CRISPR-locus in *E. coli* and purified apo-RcCas13a as well as the enzyme in complex with the ~58 nt crRNA (RcCas13a-crRNA binary complex). The RcCas13a-crRNA binary complex exhibits a markedly increased stability of about 10 °C (Supplementary Figs. S3 and S4). We determined the X-ray crystal structure of the binary complex by experimental phasing, using crystals of SeMet-labeled RcCas13a protein and refined the final coordinates against data collected on native RcCas13a-crRNA complex crystals to 2.1 Å resolution (for X-ray data processing, structure solution and refinement statistics see Table 1). Protein residues as well as nucleobases of the repeat-region and the first 16nt’s of the spacer in the crRNA are well defined in the electron density. However, some nucleobases in the middle of the spacer sequence and solvent-exposed loop regions in the protein are flexible and partly disordered (Supplementary Fig. S5). RcCas13a shares the common domain architecture and predominantly α-helical-folding topology of the Cas13a-family. Using HHpred we recognized that the C-terminal domain (CTD) has similarity to members of the cold shock domain-protein family, which are multifunctional RNA/DNA binding proteins, exhibiting a β-barrel fold (Supplementary Fig. S7). Albeit the sequence identity is low (~15%), it is likely that the RcCas13a-CTD adopts a similar overall fold. Nevertheless, all our attempts to express and purify solely the RcCas13a-CTD domain for structural characterization failed. The binary complex structure of RcCas13a shows the enzyme in a contracted, inactive conformation. Here, in the composite

| Table 1 Data collection, processing, structure solution, and refinement statistics. |
|---------------------------------|-------------------|-------------------|
| **Native**                      | **SeMet**         |
| **Data collection**             | **Cell dimensions** | **Resolution (Å)** |
|                                  | a, b, c (Å)       | 47.1–2.2          |
|                                  | α, β, γ (°)       | 90.0, 103.7, 97.6 |
|                                  | R_{merge}         | 0.085 (0.913)     |
|                                  | rmerge/Refl       | 9.6 (1.2)         |
|                                  | Completeness (%)  | 97.6 (96.3)       |
|                                  | Redundancy        | 3.6 (3.4)         |
|                                  | CC1/2             | 0.997 (0.726)     |
|                                  | CC2               | 0.999 (0.917)     |
| **Refinement**                  | **Resolution (Å)** | 47.1–2.2          |
|                                  | No. reflections   | 504,148 (48,268)  |
|                                  | R_{merge}/Refl    | 0.218/0.258       |
|                                  | No. atoms         | Protein: 17,795   |
|                                  |                   | Ligand/ion: 1986  |
|                                  |                   | Water: 630        |
|                                  | Bond lengths (Å)  | 0.005             |
|                                  | Bond angles (°)   | 0.88              |

*Values in parentheses are for highest-resolution shell. Data were collected on one single crystal each.*
active site made up by the HEPN1 and HEPN2 domains, His 469 of the conserved catalytic tetrad (Arg 464 + His 469/Arg 1052 + His 1057) is shifted about 18 Å away, preventing endonuclease activity prior to target RNA binding (Fig. 1b, c). Albeit HEPN-domains do not share a significant sequence conservation, they are characterized by an α-helical-folding topology and a conserved Arg and His sequence motif (R(x)4–6H (where x can be any residue; Supplementary Fig. 8). Two HEPN-domains have to align the Arg and His residues, either through dimerization of the two domains contained within one protein, like in Cas13 endonucleases, or of two HEPN-domain proteins, two form the catalytic tetrade required for metal-independent endonuclease activity21,22. As common for type VI endonucleases, RcCas13a possesses two distinct active sites for crRNA processing and target RNA cleavage (Fig. 1c, d). This is in agreement with functional and structural studies on the homologs with ~23% sequence identity to RcCas13a which demonstrated that target RNA binding to the binary complex induces a conformational change and rotation of the helical 2 (Hel 2) domain relative to the HEPN1 domain, resulting in rearrangement of the catalytic tetrad thereby switching-on the RNase activity (Supplementary Fig. S9)9,10,12. Thus, the stable contracted conformation of the binary complex as observed in RcCas13a, likely prevents premature activation of the RNase function.

**Recognition of the crRNA by RcCas13a.** The overall common crRNA structure of the binary Cas13a-crRNA complexes consists of a stem-loop, a 2 nt bulge at the 3′ stem and an A-U base pair at the bottom of the stem, implying a shared shape readout mechanism of crRNA recognition by this endonuclease family, combined with some sequence-specific base-readout. RcCas13a recognizes its ~58 nt long crRNA predominantly via shape-recognition, but makes nucleobase-specific interactions with its stem-loop and 3′ flank of the crRNA. The peptide backbone in the helical I (Hel 1) domain (resi 276–279) contacts the exocyclic carbonyl group in C-17 and the Hoogsteen-face of G-16 in the loop (Fig. 2c). The ε-amino group of Lys 179 interacts with A-15, and Asp 293 contacts the exocyclic amino group in A-21 (Fig. 2a, b). In addition, a base-specific readout of the bulge formed by A-8 and C-7 at the 3′ base of the stem-loop occurs via bi-dental hydrogen-bonding of Arg 23 and the peptide backbone of the β-hairpin in the NTD to the Watson–Crick–Franklin face of C-7 (Fig. 2d). The exocyclic amine in A-8 is interacting with the carbonyl oxygen of Ser 684 located in α-helix of the Hel 2 domain. Moreover, G-5 and C-1 in the 3′ flank of the crRNA...
form interactions with Arg 695 (Hel 2) and His 788 (HEPN1-2) (Fig. 2d). All other observed interactions of RcCas13a and its crRNA occur in a nucleobase-unspecific manner and the phosphodiester backbone of the crRNA (Supplementary Figs. S10 and S11). Distinct to the crRNAs from Lba, Lbu, Lse and LshCas13a, the stem-loop in the crRNA of RcCas13a is extended by one C-G base pair and an additional 3.4 Å interaction between the twisted C-28 (3N) and G-5 (1N) at the base of the stem, next to the conserved U-A base-pair can be observed (Supplementary Fig. S11). Moreover, RcCas13a does not form any sequence-specific contacts with the 5′ flank of the crRNA. In contrast, LbuCas13a and LseCas13a interact base-specific with an A and a G, respectively, in the 5′ flank. It was shown for LbuCas13a that it is sensitive to mutation in the 5′ flank of the crRNA as well as to mutations that impact on the overall shape of the stem-loop3,12. Nucleobases at positions 10–14 in the crRNA-spacer region are solvent exposed and are presumably acting as the seed region for target RNA binding (Fig. 1b), as shown for the homolog LbuCas13a23. The 2-nt bulge at the base of the 3′ stem-loop, a common characteristic of so far investigated Cas13a endonucleases (Supplementary Fig. S11), was shown to be important to regulate LseCas13a and LshCas13a activity. Here the extended sequence complementarity and base pairing between the crRNA and target RNA by 8 nucleotides (= anti-tag) blocks activation of the endonuclease. This so called tag:anti-tag interaction appears to allow discrimination between self and non-self, preventing erroneous activation of the endonuclease24,25. A consequence of the extended A-form dsRNA formed between crRNA and target RNA parts of the hairpin and the bulge of the crRNA disintegrate, leading to a misalignment of HEPN 1 + 2-domain catalytic residues and inactive enzyme25.

Mechanism of pre-crRNA processing. For some Cas endonucleases from the Cas13a and Cas12a/Cpf1-families, it was shown that they process their own pre-crRNA in a distinct active site9,10,26, while Cas9-homologs rely on RNaseIII, a general RNA processing factor in bacteria, for crRNA biogenesis27. We investigated the crRNA processing activity of RcCas13 in vitro, using synthetic 3′ 6-Carboxyfluorescein (6-FAM) labeled pre-crRNA. Our experiments show that RcCas13a does not rely on the catalytic HEPN-domains for pre-crRNA processing since RcCas13a mutants, where the active-site residues in one or both of the HEPN1 or HEPN2 domains were mutated to Ala were not affected in their pre-crRNA processing activity (Fig. 3a). Moreover, once bound to RcCas13, the crRNA is not released from the protein again since the binary complex does not cleave 6-FAM-labeled pre-crRNA (Fig. 3a). Furthermore, pre-crRNA processing by RcCas13a is independent of metal ions and the enzyme is still active in the presence of the chelator EDTA in the reaction buffer (Fig. 3b). This indicates an acid-base-catalyzed processing mechanism, in agreement with previous studies on some Cas13a homologs9,10 and Cas12a/Cpf126. In order to gain deeper insights into the pre-crRNA processing mechanism of RcCas13a, we mutated all likely candidate amino acid residues located in the vicinity of the 5′ end of the crRNA in the binary complex structure, since sequence conservation in this area is limited. Following expression and purification, the integrity of all purified

---

**Fig. 2 RcCas13a-crRNA interactions.** a Schematic depiction of the crRNA, with base-specific contacts indicated by red arrows. b Overall RcCas13a complex with the crRNA overlaid with its semi-transparent surface. RcCas13a is colored according to its domain as in Fig. 1. cRNA: stem-loop = golden; spacer = light yellow; seed region = orange; sites with base-specific interaction are highlighted in pink and violet. c RcCas13a base-specific interaction with the tip of the crRNA hairpin, d with the crRNA bulge at the base of stem and C-1 in the 3′ flank.
proteins was analyzed by dynamic light scattering (DLS) to exclude that the introduced mutations have an impact on protein stability and hence catalytic activity, rather than the chemical properties of the respective amino acid side chains (Supplementary Fig. S12).

We tested all RccCas13a mutants for their pre-crRNA cleavage activity and could identify Arg 1085, Asn 1089 and His 1113 as key residues essential for pre-crRNA processing (Fig. 3c). While Arg 1085 and Asn 1089 are located at the N-terminus of an α-helix in HEpN2 at the bottom of a cleft between the Hel 1 and HEpN2-domain, His 1113 sits at the tip of a β-hairpin in the HEpN2-domain (Fig. 3d). While in Lbcas13a Arq 1079, Lys 1082 and Asn 1083 residues at structurally equivalent positions at the N-terminus of the α-helix where shown to be crucial in Lbcas13a for pre-crRNA cleavage (Supplementary Fig. S13)4,9, mutation of the corresponding Lys 1088 in Rccas13a did not influence its pre-crRNA processing activity (Fig. 3c) and Supplementary Fig. S15). While in the homolog from L. shahii (LshCas13a) Arq 438 and Lys 441 located in the NTD are required for pre-crRNA processing10. Additionally, in the Lbcas13a complex the 5′ flanl of the pre-crRNA and crRNA are tightly held by the surrounding amino acid side chains. In contrast, the equivalent site in the Rccas13a complex is more open, allowing more flexibility for the 5′ flanl of the crRNA (Supplementary Fig. S14).

Fig. 3 Processing of pre-crRNA by Rccas13a. a Rccas13a cleaves 3′-FAM-labeled pre-crRNA (65 mer) independent of its catalytic HEpN-domains or the C-terminal domain. Rccas13a pre-bound to crRNA does not process 3′-FAM-labeled pre-crRNA. ΔCTD = Δ C-terminal domain; H1* = Rccas13aR464A/H469A/H1057A, H2* = Rccas13aR464A/H469A/R1052A/H1057A. b Pre-crRNA cleavage is not inhibited by EDTA and hence metal-ion-independent. c Identification of residues involved in pre-crRNA processing by mutational analysis. d Close-up view of the 5′ end of the crRNA in the Rccas13a binary structure. residues in the vicinity of the 5′ processing highlighted in pink. Water molecules are shown as spheres. e MALDI-TOF analysis of the pre-crRNA processing reaction shows cleavage of the pre-crRNA 4 or 5 nt 5′ of the stem-loop, resulting in a 2′,3′-cyclic 3′ phosphate (2′,3′cP). f Proposed mechanism of pre-crRNA processing by acid-base catalysis. His 1113 acts as the base, abstracting the proton of the 2′OH next to the scissile phosphate. Arg 1085 acts as a general acid and protonates the 5′-oxygen leaving group.

endonucleases, also requires two residues, Arg 802 and Lys 905 located in its HEpN2 domain for pre-crRNA processing5. These are to our knowledge to date the only examples of catalytic activity HEpN domains that contain a second RNase active site (Supplementary Fig. S8).

Although all Cas13a homologs whose structures have been determined share the same overall domain structure, they vary in the relative orientation of the domains as well as the exact positioning of the crRNA (Supplementary Fig. S6). In general, metal-free acid-base catalysis of nucleases are SN2 reactions. Here a nucleophile attacks the scissile phosphate either from the 3′ side, breaking the 5′-O-P bond, which produces 3′-phosphate and 5′-OH, or the 5′ side, breaking the 3′-O-P bond, resulting in a 5′-phosphate and 3′-OH28. To gain more insights into pre-crRNA processing by Rccas13a, we analyzed the reaction products of the in vitro pre-crRNA processing reaction by matrix-assisted laser desorption ionisation-time of flight (MALDI-TOF) mass spectrometry, which possesses much higher sensitivity and precision in comparison to gel-electrophoresis. This revealed that Rccas13a cleaves either 4 or 5 nt 5′ of the stem-loop, generating a 2′,3′-cyclic 3′ phosphate (2′,3′cP, Fig. 3e and Supplementary Fig. S15). Given that the reaction product is a 2′,3′-cP, the 2′-oxygen next to the scissile phosphate has to carry out the nucleophilic attack and is thus transferred to the 5′ phosphate. Substitution of the ribose by the 2′ deoxyribose (dA or dC) in synthetic pre-crRNA strands abrogates cleavage at this positions (Supplementary Fig. S15), providing further support for the importance of the nucleophilic attack by the 2′OH. Prerequisite for this transfer reaction is deprotonation of the 2′OH group, possibly by the essential His 1113 acting as the base. However, it is also possible that His 1113 abstracts a proton from a water molecule and that the resulting hydroxide ion acts as the effective base, since our structure shows
RcCas13a in complex with the processed, mature crRNA and not the pre-crRNA bound state. Asn 1089 could play a role in the stabilization of the penta-covalent reaction intermediate and/or positioning of the 5′ flank, since mutation of Asn 1089 to Asp does not impact on the cleavage position but mutation to Ala leads to aberrant cleavage 6 nt 5′ to the stem-loop (Supplementary Fig. S15). Arg 1085 could act as the acid and protonates the 5′-oxygen leaving group (Fig. 3f). Due to the high intrinsic pKₐ value of the Arginine’s guanidinium group (pKₐ 13.8 ± 0.1) and delocalization of the positive charge over many atoms involved in a conjugated π-system, the arginine side chain has the unusual ability to retain its charge under all physiological conditions and can thus function as an acid catalyst. As already observed by gel-electrophoresis, the RcCas13aΔCTD mutant does not show any pre-crRNA cleavage activity in the MALDI-analysis of the reaction products (Supplementary Fig. S15). As stated above, in our mutational analysis of residues lining the pre-crRNA processing site in RcCas13a, we could identify Arg 1085 and His 1113 as being clearly essential, and Asn 1089 modulating crRNA cleavage. In contrast, for LbaCas13a, which possesses a narrower pre-crRNA binding site (Supplementary Fig. S15), it was shown that Trp 325 and Asn 1232 are the likely candidates to set up the 2′-OH of the ribose of the scissile base in the pre-crRNA for nucleophilic attack. An activated water molecule, together with Lys 1305, His 328, Lys432, and Lys 1320 participate in the proton transfer reactions.

**In vivo activity of RcCas13a.** We investigated the in vivo cleavage activity of RcCas13a in *E. coli* JM109 using plasmid encoded reporter genes (luciferase: luxABCDE; green fluorescence protein: GFP) under control of an anhydrotetracycline Pₜ₆₅ and an IPTG inducible T7-promoter, respectively. RcCas13a and a crRNA with complementary spacer sequence to *luxA* are under transcriptional regulation of the arabinose responsive PBAD-promoter (Fig. 4a). This setup allows the distinct expression of all assay components:
RcCas13α (wt or mutants), crRNA<sub>luxA</sub>, luciferase (reporter for cis-cleavage activity), gfp (reporter for trans- or bystander cleavage activity). Upon expression of wt RcCas13α and crRNA we observe a reduction in bioluminescence due to the sequence-specific degradation of luxABCDE mRNA by RcCas13α (Fig. 4b, left red bar, cis-cleavage), which is independent of expression induction of the second reporter gfp (Supplementary Fig. S16). This observed reduction in bioluminescence is strictly dependent on the presence and expression of the luxA-specific crRNA (Fig. 4b, yellow and pink bars). Deletion of the CTD in Cas13α does not affect RcCas13α activity (Fig. 4b, brown bar), while mutation of any of the catalytic residues in the HEPN-domains abolishes RcCas13α activity (Fig. 4b, blue bars). Sequence independent trans-cleavage was analyzed by GFP-fluorescence. Wt RcCas13α and RcCas13αΔCTD reduce GFP-fluorescence only when co-expressed with crRNA and luxABCDE (Fig. 4c, red and brown hatched bars). Without co-expression of either crRNA or the lux-operon GFP-fluorescence remained unaffected by RcCas13α expression (Fig. 4c, yellow hatched bar and Supplementary Fig. S16). Again RcCas13α-mediated reduction in GFP-fluorescence is dependent on the HEPN-domains and prevented by mutation of any of the active site residues (Fig. 4c, blue hatched bars). This HEPN-dependent cis- and trans-cleavage activity by RcCas13α is common to the Cas13α-family<sup>3,7,9,10,12</sup>. Scrambling of the sequence 10–14 of the spacer (scr-crRNA), which was shown for other Cas13α homologs<sup>24</sup> acting as seed region for interaction between crRNA and target RNA, abolishes reduction of luciferase luminescence upon expression of RcCas13α scr-crRNA (Fig. 4b, right, black bar). Moreover, expression of wt RcCas13α (or RcCas13αΔCTD) in combination with crRNA and target RNA leads to a reduced bacterial growth of the <i>E. coli</i> cells (Supplementary Fig. S17). This is in agreement with previous studies on Type V CRISPR-Cas systems, showing that Cas13α mediates protection against phages by dual function: first, by direct target RNA degradation and secondly, abortion of the infectious cycle by inducing growth arrest (dormancy) in the host cell through trans-cleavage of transcripts. With this indirect target mechanism Type V systems also act on CRISPR-resistant phages, deplete the phage population and thus provide herd immunity to unaffected bacteria<sup>5,31</sup>.

### Conclusion

The photo-auxotrophic bacterium <i>R. capsulatus</i> encodes, next to Class 1, also a Class 2 Type VI CRISPR-Cas system, harboring a 145 kDa Cas13α single-effector endonuclease. This RcCas13α possesses two distinct active sites for crRNA processing and ssRNA endonuclease activity that are typical for this family. The X-ray crystal structure of RcCas13α in complex with its crRNA reveals, that the enzyme combines shape and sequence-specific readout to recognize its crRNA and shows the enzyme in its stable and contracted, inactive conformation, which likely prevents premature activation of the endonuclease activity. Moreover, RcCas13α processes its crRNA, by cleaving either 4 or 5 nt upstream of the stem-loop, which results in a 2′,3′ cyclic phosphate at the 5′ leaving group. Crucial for the crRNA processing activity are three residues located in the processing site: His1113 and Arg1085 likely acting as acid and base. Mutation of Asn1089 to Ala influences the position of the pre-crRNA and resulting in a partial shift in cleavage site preference. Upon expression in <i>E. coli</i>, RcCas13α utilizes its HEPN-domain composite active site to cleave target RNAs with complementary sequence to the crRNA spacer sequence. The presence of crRNA and complementary target RNA also unlocks the HEPN-domain dependent trans-cleavage activity of bystander RNAs and results in a reduced bacterial growth, indicating that RcCas13α protects the bacterial population by aborting lytic infection and phage propagation by inducing dormancy in the host, as reported for homologs. In summary, our data provide further insights into the structure, mechanism and function of this intriguing family of RNA-dependent endonucleases.

### Methods

#### Molecular biology.

Protein and nucleic acid sequences (RcCas13α, RNA, plasmids and primers) are listed in Supplementary Tables S1–S4. Mutants were prepared using standard cloning procedures (restriction enzymes (New England Biolabs), Gibson Assembly, NEBuilder HiFi DNA Assembly (NEB), QuickChange Mutagenesis (Agilent), Q5 site-directed mutagenesis (NEB), SLIM PCR<sup>25</sup>) according to the protocol of manufacturer’s instructions. All used plasmids were verified by Sanger sequencing (GATC Biotech/Geneviz, Germany).

#### Protein expression of RcCas13α and RcCas13α mutants.

For protein expression RcCas13α, annotated as a hypothetical protein (WP_013067728.1), or RcCas13α mutants were cloned into the respect expression plasmid (see Supplementary Table S3) and expressed in <i>E. coli</i> Rosetta2 (DE3). Cells obtained from an overnight culture (LB medium supplemented with respective antibiotics according to Supplementary Table S3, 37 °C, 180 rpm) were diluted 1/100 into OD<sub>600</sub> between 0.2 and 0.4 between 100°C and 20°C. Cells were incubated at 37 °C, 180 rpm until an OD<sub>600</sub> of 0.7 was reached. After shock cooling for 15–30 min on ice, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight at 18–20°C, 180 rpm. Cells were subsequently harvested by centrifugation (5000 g, 10 min; 4 °C).

RcCas13α and RcCas13α-crRNA complex purification. All purification steps were performed at 4–8 °C or on ice unless otherwise noted. The two-step purification procedure of RcCas13α proteins comprised a Ni<sup>2+</sup>-affinity chromatography followed by removal of the His-MBP-tag using Tobacco Etch Virus (TEV) protease and a size-exclusion chromatography (SEC), both carried out on an AKTA-FPLC system (GE Healthcare, now Cytiva). The cell pellet obtained from a 1-liter expression was resuspended in 30 ml His-wash buffer (50 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (Hepes) pH 8.0, 1 M NaCl, 10% (v/v) glycerol, 1 mM diithothreitol (DTT), 10 mM imidazole supplemented with 0.1 mg/ml DNase I (AppliChem), 0.1 mg/ml lysozyme (~20,000 U/mg; Carl Roth) and one tablet Complete<sup>TM</sup> ULTRA EDTA-free protease inhibitor tablet (Roche). Cells were lysed by homogenization using an EmulsiFlex-C5 (Avestin Inc.) or by sonication and cleared lysate was obtained by centrifugation (24,446 g, 30 min; 4 °C) and filtration using a Whatman<sup>TM</sup> folded filter (GE Healthcare, now Cytiva).

Affinity purification used a 1 ml HisTrap FF crude column (GE Healthcare, now Cytiva) at a flow rate of 1 ml/min. After sample application the column was washed with 12 CV His-wash buffer followed by 5 CV His-wash buffer supplemented with imidazole to a final concentration of 30 mM. RcCas13α proteins were eluted with 13 CV His-elution buffer (50 mM Hepes pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, 1 mM DTT) and 1.0 ml fractions were collected. Fractions containing the protein were pooled and concentrated with an Amicon Ultracel Centrifugal filter unit (MWCO 50 kDa, Merck Millipore) to a volume of 500 µl and digested with TEV protease for 2 h at 8 °C. After centrifugation (16,000 g, 10 min; 4 °C) the sample was applied onto a HiLoad Superdex 200 (16/60) column (GE Healthcare, now Cytiva). Separation was performed at a flow rate of 0.8 ml/min in SEC buffer (50 mM Hepes pH 8.0, 500 mM NaCl, 1% (v/v) glycerol, 1 mM DTT) and 1.0 ml fractions were collected. Fractions containing the respective monomeric RcCas13α proteins were pooled, concentrated, and either directly used for crystallization trials or flash-frozen in liquid nitrogen and stored at −80 °C. The RcCas13-crRNA binary complex was obtained by co-expressing RcCas13α containing a C-terminal His<sub>6</sub>-tag with the CRISPR-locus encoded on the same plasmid (pET24-RcCas13α-crRNA). The purification of the RcCas13α-crRNA complex was done analogously to the RcCas13α protein, without removal of the His-tag.

#### Expression and purification of the SeMet-labeled RcCas13α-crRNA binary complex.

For production of SeMet-labeled protein, was carried out according to the modified protocol by van den Ent and Löwe<sup>25</sup>. In short, a culture of <i>E. coli</i> Rosetta2 (DE3) was grown in M9 minimal medium at 37 °C under agitation until an OD<sub>600</sub> of 0.6 was reached, where an amino acids mix (final concentrations: 100 mg/l-lysine, threonine and phenylalanine, 50 mg/l-leucine, isoleucine, valine and L-selenomethionine) was added as solids and incubated for 15 min. Cells were shock-cooled for 30 min at 4 °C and protein expression was induced with 1 mM IPTG overnight. Purification was performed as for the unlabeled protein, but using 5 mM β-mercaptoethanol (IMAC) and 5 mM DTT (SEC) in the respective buffers.

#### Pre-crRNA cleavage by RcCas13α.

For the in vitro cleavage assay 0.625 µM or 0.208 µM 3′-FAM-labeled pre-crRNA (Ella Biotech; sequence in Supplementary Table S2) was incubated with 2 µM RcCas13α in cleavage buffer (20 mM Hepes pH 80°C).
MALDI-TOF analysis of Cas13a pre-crRNA processing products. Processed pre-crRNA samples for MALDI experiments were obtained by incubation of 3'-FAM-labeled pre-crRNA or respective desoxy-modified pre-crRNA (da-modified/dC-modified) (24 µM; or 15 µM; Ella Biotech; sequences in Supplementary Table S2) with RcCas13a (33 µM) in cleavage buffer for 3 h at 37 °C. The Cas13a protein was removed by heat denaturation (95 °C; 10 min) followed by centrifugation (16,000 x g; 10 min; 4 °C). Unprocessed pre-crRNA (100 µM in H2O) served as control. MALDI-TOF mass spectra were recorded on a Bruker Autoflex II. For the measurements, the samples were desalted against ddH2O using a 0.025 µm VSWP membrane filter (Merck Millipore) and then co-crystallized with a matrix solution (0.7 M 3-hydroxypicolinic acid, 0.07 M diammonium citrate in 1:1 H2O/MeCN).

In vivo activity assay. The sequences encoding the luciferase subunits luxA, luxB, luxC, luxD and luxE from Photorhabdus luminescens (Uniprot IDs P23146, P23113, P23148, P19482) were cloned into pASK-IBA5 (IBA Lifesciences), E. coli with IPTG. Plasmids according to Supplementary Table S4 were transformed into fusion with a C-terminal ssrA degradation tag was cloned into pACYC, inducible order to analyze bystander RNA degradation, the sequence encoding eGFP in RSF1031K, under control of the L-arabinose (ara) inducible pBAD promoter. In

References
1. Makarova, K. S. et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat. Rev. Microbiol. 18, 67–83 (2020).
2. Watanabe, S. et al. Composition and diversity of CRISPR-Cas13a systems in the genus Lettovicia. Front. Microbiol. 10, 2838 (2019).
3. East-Seletsky, A. et al. Two distinct RNAse activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature 538, 270–273 (2016).
4. Liu, L. et al. Two distant catalytic sites are responsible for C2c2 RNA activities. Cell 168, 121–134 e112 (2017).
5. Zhang, B. et al. Two HEPN domains dictate CRISPR RNA maturation and target cleavage in Cas13d. Nat. Commun. 10, 2544 (2019).
6. Zhang, B. et al. Structural insights into Cas13b-guided CRISPR RNA maturation and recognition. Cell 178, 1198–1210 (2018).
7. Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353, aaf5537 (2016).
8. East-Seletsky, A., O’Connell, M. R., Burstein, D., Knott, G. J. & Doudna, J. A. RNA targeting by functionally orthogonal type VI-A CRISPR-Cas enzymes. Mol. Cell. 66, 375–383 e373 (2017).
9. Liu, L. et al. The molecular architecture for RNA-guided RNA cleavage by Cas13a. Cell 170, 714–726 e710 (2017).
10. Knott, G. J. et al. Guide-bound structures of an RNA-targeting A-claving CRISPR-Cas13a enzyme. Nat. Struct. Mol. Biol. 24, 825–833 (2017).
11. Kellner, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O. & Zhang, F. SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat. Protoc. 14, 2986–3012 (2019).
12. Meeske, A. J. et al. A phage-encoded anti-CRISPR enables complete evasion of type VI-A CRISPR-Cas immunity. Science 360, 54–59 (2020).
13. Lopatina, A., Tal, N. & Sorek, R. Absorbive infection: bacterial suicide as an antiviral immune Strategy. Annu. Rev. Virol. 7, 371–384 https://doi.org/10.1146/annurev-virology-011620-040628 (2020).
14. Deveau, H. et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophiles. J. Bacteriol. 190, 1390–1400 (2018).
15. Pyenson, N. C., Gayvert, K., Varble, A., Elemento, O. & Marraffini, L. A. Broad targeting specificity during bacterial Type III CRISPR-Cas immunity constrains viral escape. Cell Host Microbe. 22, 343–353 e437 (2017).
16. van Houte, S. et al. The diversity-generating benefits of a prokaryotic adaptive immune system. Nature 532, 385–388 (2016).
17. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. Nature 577, 327–336 (2020).
18. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. Curr. Opin. Microbiol. 37, 67–78 (2017).
19. Zimmermann, L. et al. A completely reimplented MPI Bioinformatik Tool with a new HHPred server at its core. J. Mol. Biol. 430, 2237–2243 (2018).
20. Amir, M. et al. Sequence, structure and evolutionary analysis of cold shock domain proteins, a member of OB fold family. J. Evol. Biol. 31, 1903–1917 (2018).
21. Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V. & Aravind, L. Comprehensive analysis of the HELEN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol. Direct 8, 15 (2013).
22. Koonin, E. V. & Makarova, K. S. Functions and evolution of CRISPR-Cas systems. Philos Trans R. Soc. Lond. B Biol. Sci. 374, 20180087 (2019).
23. Tambe, A., East-Seletsky, A., Knott, G. J., Doudna, J. A. & O’Connell, M. R. RNA binding and HELEN-nucleic acid activation are decoupled in CRISPR-Cas13a. Cell Rep. 24, 1025–1036 (2018).
24. Meeske, A. J. & Marraffini, L. A. RNA guide complementarity prevents self-targeting in type VI CRISPR systems. *Mol. Cell* 71, 791–801 e793 (2018).
25. Wang, R. et al. Structural basis for self-cleavage prevention by tag:anti-tag pairing complementarity in type VI Cas13 CRISPR systems. *Mol. Cell* 81, 1100–1115 (2021).
26. Swarts, D. C., van der Oost, J. & Jinek, M. Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a. *Mol. Cell* 68, 221–233 e224 (2017).
27. Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNRase III. *Nature* 471, 602–607 (2011).
28. Yang, W. Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* 44, 1–93 (2011).
29. Fitch, C. A., Platzer, G., Okon, M., Garcia-Moreno, B. E. & McIntosh, L. P. Arginine: its pKa value revisited. *Protein Sci.* 24, 752–761 (2015).
30. Silva, P. I., Schulz, C., Jahn, D., Jahn, M. & Ramos, M. J. A tale of two acids: when arginine is a more appropriate acid than H2O\(^+\). *J. Phys. Chem. B* 114, 8994–9001 (2010).
31. Meeske, A. J., Nakandakari-Higa, S. & Marraffini, L. A. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* 570, 241–245 (2019).
32. Chiu, J., March, P. E., Lee, R. & Tillett, D. Site-directed, ligase-independent mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucelic Acids Res.* 32, e174 (2004).
33. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* 229, 105–124 (1993).
34. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132 (2010).
35. Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst.* 34, 843–844 (2001).
36. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* 64, 112–122 (2008).
37. Skubak, P. et al. A new MR-SAD algorithm for the automatic building of protein models from low-resolution X-ray data and a poor starting model. *IUCrJ.* 5, 166–171 (2018).
38. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501 (2010).
39. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* 75, 861–877 (2019).
40. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ.* 1, 213–220 (2014).
41. Schrödinger, L. L. C. The PyMOL Molecular Graphics System, Version 1.8. (2015).

**Acknowledgements**

This work was financially supported by the Deutsche Forschungsgemeinschaft (SCHN 1273–5/SPP 2141, SCHN 1273–6; Center for Integrated Protein Science Munich (CIPSM, EXC114) and SFB 1309). We thank the European Synchrotron Radiation Facility (ESRF) and the Swiss Light Source (SLS) for beamtime, and the staff of beamlines PX 1 (SLS), MASSIF-1 and ID23-2 (ESRF) for setting up the beamlines for data collection. We also thank Christopher Scheidler and Philipp Merz for assistance with protein purification for some activity assays.

**Author contributions**

L.M.K. designed experiments, optimized protein purification and crystallization, determined the structure, performed in vivo activity assays. M.K.W. produced mutant proteins, carried out pre-cRNA processing and in vivo assays. L.S.R. carried out MALDI-ToF measurements, S.S. supervised the project, planned experiments, analyzed data, and wrote first draft of the manuscript. All authors discussed the results and edited the final manuscript version.

**Funding**

Open Access funding enabled and organized by Projekt DEAL.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s42003-022-03025-4.

**Correspondence** and requests for materials should be addressed to Sabine Schneider.

**Peer review information** *Communications Biology* thanks Cheng Zhang and the other anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Joanna Timmins and Anam Akhtar. Peer reviewer reports are available.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.