Corrected 13 June 2023. See full text.

RESEARCH ARTICLE

CRISPR

Craspase is a CRISPR RNA-guided, RNA-activated protease

Chunyi Hu[1]†, Sam P. B. van Beljouw[2,3]†, Ki Hyun Nam[4], Gabriel Schuler[1], Fran Ding[1], Yanru Cui[1], Alicia Rodríguez-Molina[2,3], Anna C. Haagsma[2,3], Menno Valk[2,3], Martin Pabst[5], Stan J. J. Brouns[2,3]*, Ailong Ke[1]*

The CRISPR-Cas type III-E RNA-targeting effector complex gRAMP/Cas7-11 is associated with a caspase-like protein (TPR-CHAT/Csx29) to form Craspase (CRISPR-guided caspase). Here, we use cryo–electron microscopy snapshots of Craspase to explain its target RNA cleavage and protease activation mechanisms. Target-guide pairing extending into the 5′ region of the guide RNA displaces a gating loop in gRAMP, which triggers an extensive conformational relay that allosterically aligns the protease catalytic dyad and opens an amino acid side-chain–binding pocket. We further define Csx30 as the endogenous protein substrate that is site-specifically proteolyzed by RNA-activated Craspase. This protease activity is switched off by target RNA cleavage by gRAMP and is not activated by RNA targets containing a matching protospacer flanking sequence. We thus conclude that Craspase is a target RNA–activated protease with self-regulatory capacity.

It has become clear that RNA-guided DNA/RNA degradation is not the sole mechanism for CRISPR-Cas to confer immunity against foreign genetic elements in prokaryotes (*1*–*5*). Type III CRISPR-Cas systems in particular present a plethora of alternative mechanisms, including RNA-guided secondary messenger production and signaling (*6*, *7*), to activate a range of immune responses (e.g., collateral RNA damage) (*6*–*8*). Type III CRISPR-Cas effectors are typically assembled from multiple protein subunits to enable CRISPR RNA (crRNA) binding and target RNA cleavage, DNA cleavage, and secondary messenger synthesis (*9*, *10*). Type III-E is a recently identified atypical type III system. It encodes a large polypeptide (gRAMP) as a fusion of four Cas7-like domains, one Cas11-like domain, and a big insertion domain (BID), but lacks Cas10, the signature component of a canonical type III system that is required for secondary messenger production (*5*). Subsequent studies demonstrated that the gRAMP ribonucleoprotein (RNP) complex is capable of RNA-guided RNA cleavage at two specific sites (*11*, *12*) that are six nucleotides apart (*11*). Unlike the type VI CRISPR-Cas effector Cas13, gRAMP does not cause collateral RNA cleavage and has no cytotoxicity in eukaryotic cells (*12*).

[1]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA. [2]Department of Bionanoscience, Delft University of Technology, 2629 HZ Delft, Netherlands. [3]Kavli Institute of Nanoscience, 2629 HZ Delft, Netherlands. [4]Department of Life Sciences, Pohang University of Science and Technology, Pohang, Gyeongbuk 37673, Republic of Korea. [5]Department of Environmental Biotechnology, Delft University of Technology, 2629 HZ Delft, Netherlands.
*Corresponding author. Email: ailong.ke@cornell.edu (A.K.); stanbrouns@gmail.com (S.J.J.B.)
†These authors contributed equally to this work.

In type III-E loci, gRAMP frequently associates with TPR-CHAT, a caspase-like protein with N-terminal tetratricopeptide repeats (TPRs) (*5*). Caspases are a family of cysteine proteases controlling programmed cell death pathways in eukaryotes (*13*). Cleavage of gasdermin by caspases, for example, triggers membrane pore formation to cause cell death (*14*, *15*). An equivalent programmed cell death pathway was recently discovered in prokaryotes, where TPR-CHAT was shown to cleave bacterial gasdermin to induce cellular suicide (*2*). In type III-E systems, TPR-CHAT and gRAMP physically interact to form an effector complex named Craspase, for CRISPR-guided caspase (*11*). This observation raised the possibility that Craspase may function as a crRNA-guided protease to prevent the spread of phage infection through a suicide mechanism. However, it remains unknown how Craspase is structurally organized, whether TPR-CHAT in Craspase is a protease, and whether its activity is regulated by RNA (*16*, *17*).

gRAMP structures in resting, RNA-bound, and postcleavage states

To gain insights into the RNA-guided target RNA cleavage mechanisms inside gRAMP, we reconstituted *Candidatus* "Scalindua brodae"-gRAMP (*Sb*-gRAMP) (*11*) and determined its cryo–electron microscopy (cryo-EM) structures in different functional states (Fig. 1A and fig. S1). Consistent with previous results (*11*), *Sb*-gRAMP bound to the complementary RNA target with better than 25 nM affinity and cleaved it at two distinct locations: after the third nucleotide (site 1) and after the ninth nucleotide (site 2) (fig. S1, A to C). Single-particle three-dimensional (3D) reconstruction produced *Sb*-gRAMP RNP in four different functional states: a 3.81-Å structure of the *apo* (resting) state, a 3.65-Å structure of the nonmatching protospacer flanking sequence (PFS) target-bound state, a 3.76-Å structure of the matching PFS target-bound state, and a 3.62-Å structure of the postcleavage state (Fig. 1, B to E, figs. S2 to S4, and table S1).

The overall architecture of *Sb*-gRAMP is similar to that of *Desulfonema ishimotonii* Cas7-11 (*Di*-Cas7-11), recently reported in the target-bound form (*18*). The two structures in the same functional state superimpose with a root mean square deviation of 1.1 Å for Cα atoms, excluding the BID domain, which is less conserved and poorly resolved in the EM density (fig. S5). *Sb*-gRAMP also shares some degree of similarity with the canonical type III-A effector Csm (*10*, *19*–*21*) in overall architecture, guide RNA display, and target RNA–binding mode (fig. S6). The *Sb*-gRAMP backbone consists of four nonidentical Cas7 domains fused together (Cas7.1 - Cas7.4) instead of the three identical Cas7 subunits in Csm (fig. S6). A Zn knuckle is present in each of the four Cas7 domains, which appears to be a shared hallmark among type III effectors (fig. S7A). Csm further contains one copy of Csm4 for 5′-handle recognition, two copies of Csm2 as part of the backbone, and one copy of Csm5 for continued guide-target pairing. By contrast, *Sb*-gRAMP is streamlined: Its Cas7.1 has been repurposed for 5′-handle recognition, the single-copy Cas11 domain has been repurposed for target cleavage, and a structurally distinct BID replaces Csm5 (fig. S6, A to H). On the guide RNA side, the ordered 18-nucleotide (nt) 5′ handle of the crRNA in *Sb*-gRAMP is twice as long as in other class I CRISPR-Cas systems (Fig. 2, A and B). Most of the handle residues are bound by Cas7.1 and shielded on the top by the linker from Cas11 to Cas7.2 and the Zn knuckle in Cas7.2 (fig. S6, S7, B and C). Mutagenesis of the Zn-knuckle structure or sequence-specific contacts to the 5′ handle abolished the in vivo RNA silencing activity of *Sb*-gRAMP, presumably through disruption of RNP assembly (fig. S8). Unexpectedly, *Sb*-gRAMP differs from *Di*-Cas7-11 in crRNA biogenesis. In *Di*-Cas7-11, there is an endoribonuclease center in Cas7.1 for crRNA processing, whereas the equivalent residues in *Sb*-gRAMP are noncatalytic (Fig. 2, C and D) (*18*). This structural difference explains the observation that the crRNA 5′ handle in *Sb*-gRAMP is 3 nt longer. We speculate that *Sb*-gRAMP may rely on certain host nucleases for crRNA biogenesis.

The last two handle nucleotides (5′-A-2 and C-1-3′) are base-pairing competent because they are displayed like a guide (Fig. 2A and fig. S6, H to G). Type I, III, and IV effectors display the crRNA spacer (guide region) in 6-nt segments, with the sixth nucleotide pinned down by the thumb loop of Cas7; the target is

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

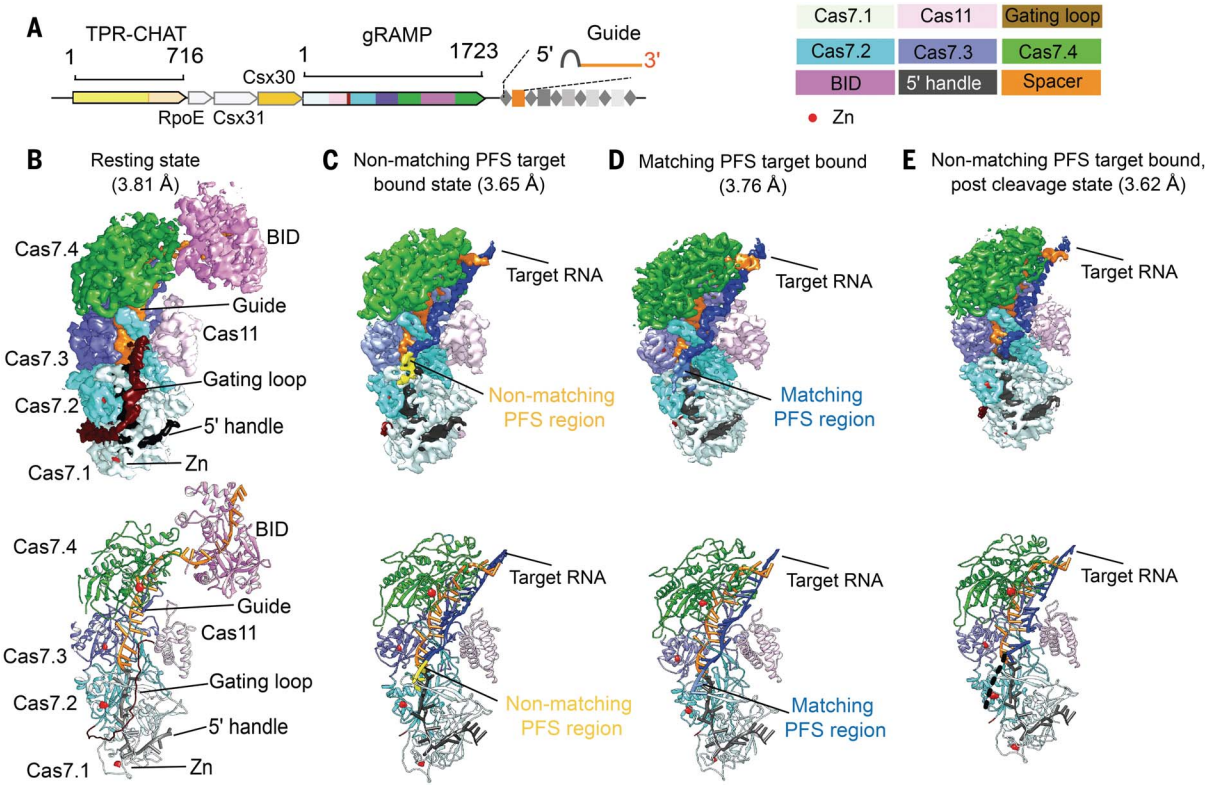


Fig. 1. Structural snapshots of *Sb*-gRAMP RNP in different functional states. (**A**) Type III-E operon in *Ca.* S. brodae. (**B** to **E**) Snapshots of *Sb*-gRAMP at resting state (B), nonmatching PFS RNA–bound state (C), matching PFS RNA–bound state (D), and nonmatching PFS RNA postcleavage state with $MgCl_2$ (E). Top images are cryo-EM densities and bottom images are structural models.

thus recognized in 5-nt segments with the sixth nucleotide unspecified. *Sb*-gRAMP contains major exceptions. The first 5-nt segment contains the last two nucleotides of the 5′ handle and the first three nucleotides of the spacer, a scenario only observed in type III-E (*18*) (Fig. 2, A and E, and fig. S6, H to G). The third segment deviates from the normal again, because an unconventional knotted protein loop from Cas7.4 divides the displayed bases to a 3-nt block and a 6-nt block. The two blocks are divided by a single-peptide crossover rather than by a β-hairpin thumb, so no nucleotide is pinned underneath and the base pairing in the third segment is not interrupted. The following crRNA nucleotides are displayed by the dynamic BID domain (amino acids 1031 to 1385), which is only resolved to low resolution and therefore docked with an AlphaFold-predicted model (*22*) (Fig. 1B).

Off-targeting prevention and RNA cleavage mechanisms in *Sb*-gRAMP

By capturing three additional functional states, we achieved the temporal resolution to interpret the target recognition and cleavage mechanisms by *Sb*-gRAMP. We found that the long linker from Cas11 to Cas7.2 (G375 to E412, here named the gating loop) has acquired important functions for RNase regulation (Fig. 3, A and B, and fig. S9). Its N-terminal portion (G375 to G397) senses RNA substrate binding and controls RNase activities. In the resting state, the gating loop blocks the first segment of the guide RNA and the nearby site 1 cleavage center. This conformation is incompatible with target-guide pairing at the first segment, and the gating loop has to be displaced to enable cleavage at site 1 (Fig. 3A). We therefore envision that the target-guide pairing initiates from the third and second segments and propagates into the first segment (fig. S9), as observed for other type III systems (*23*). In follow-up experiments, we found that *Sb*-gRAMP's RNase activity was optimal against a target with 18-nt complementarity from the 5′ end of the spacer portion: 12-nt or shorter complementarity abolished cleavage and 24-nt or longer complementarity attenuated cleavage (fig. S10). This suggests that at least some base pairing along all three segments of the guide RNA, displayed by Cas7.2 to Cas7.4, is required for efficient RNA cleavage. By contrast, additional base paring with crRNA at the BID is not required or may even be counterproductive (fig. S10). This is consistent with the previous observation that the 3′ end of the crRNA in the endogenous *Sb*-gRAMP is often as short as 20 nt (*11*), and that the BID is dispensable for Cas7-11 activity in human cells (*18*).

Sb-gRAMP was further incubated with two kinds of RNA targets with PFS that was either matching (complementary) or nonmatching with the 5′ handle in the crRNA, because complementarity in this region may be indicative of a self-target (i.e., antisense transcript from the CRISPR locus) and thus may lead to alternative structural configurations in *Sb*-gRAMP. However, our structures reveal that regardless of the PFS status, RNA binding induces the same set of conformational changes in *Sb*-gRAMP. Where the guide nucleotides are pinned down by the Cas7 thumbs, the corresponding target nucleotides (fourth and tenth) flip outward. Rotation of the backbone orients their 2′-OH toward the previous phosphate, forming the so-called "in-line" conformation that is necessary for RNA cleavage. For target RNA with a matching PFS, the first segment consists of five base pairs, starting from the last two nucleotides of the 5′ handle and ending with the third nucleotide in the spacer portion (Fig. 2E). The rest of the PFS is not traceable in the EM map. For target RNA with a nonmatching PFS, only three base pairs are found between the target RNA and the spacer portion of the guide. Although the first two nucleotides of the PFS do not form hydrogen bonds with the two 5′-handle residues on the opposite side, they remain stacked to complete

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

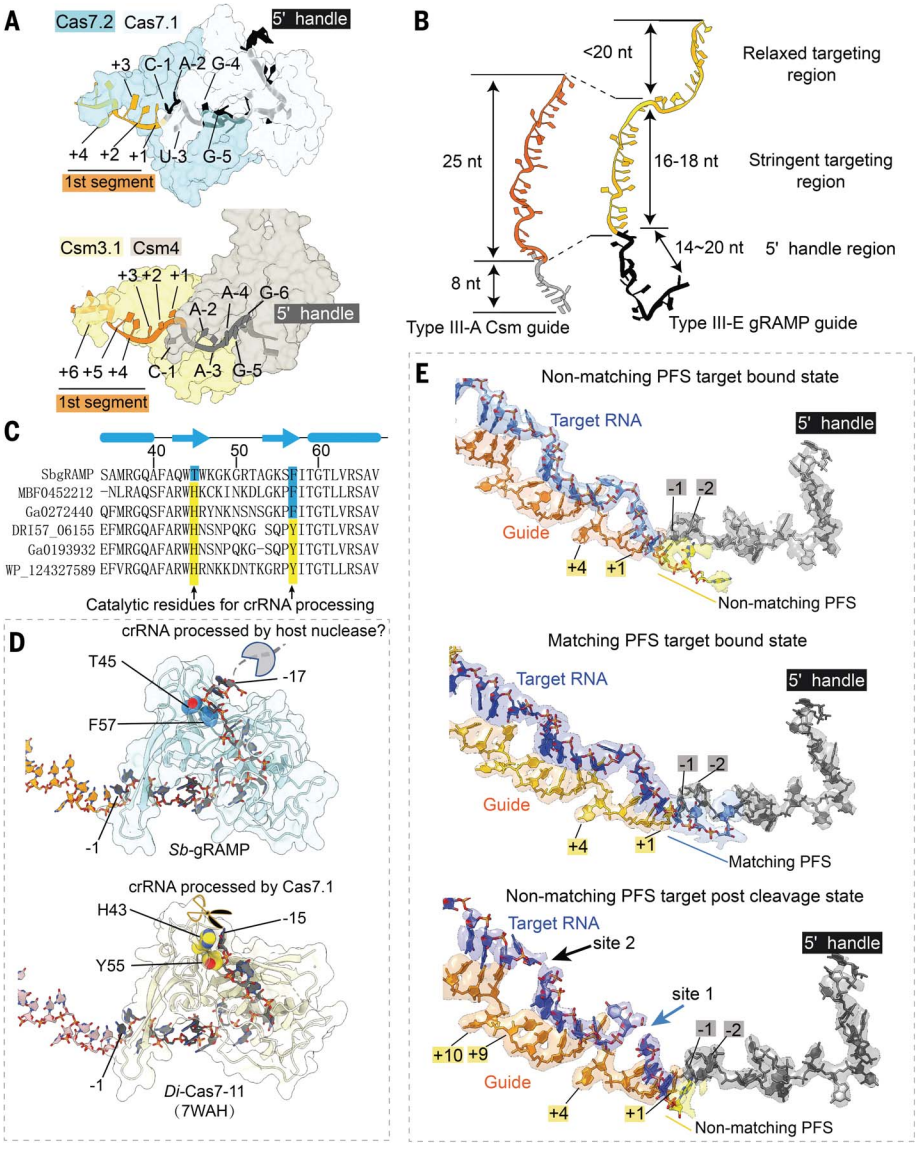


Fig. 2. crRNA accommodation and target RNA recognition mechanisms by *Sb*-gRAMP. (**A** and **B**) Accommodation of the crRNA 5′ handle (A) and spacer region (B) in type III-E *Sb*-gRAMP and type III-A Csm. (**C** and **D**) Primary sequence (C) and 3D structural alignment (D) at the pre-crRNA cleavage center. Catalytic residues in *Di*-Cas7-11 are shown in yellow; equivalent residues in *Sb*-gRAMP are shown in blue. (**E**) Extracted cryo-EM density from nonmatching PFS RNA (top), matching PFS RNA (middle), and nonmatching PFS RNA postcleavage state (bottom).

the first target-guide segment (Fig. 2E). In both PFS-matched and -nonmatched conditions, the impinging gating loop in *Sb*-gRAMP is pushed away from the first segment and becomes entirely disordered (Fig. 3A). Concurrently, the cleavage center at site 1 is exposed and further enhanced by a hinge motion in Cas11 (Fig. 3C and fig. S11A), which aligns catalytic residues among Cas11 and Cas7.2. It should be noted that stacking from the additional 2-nt PFS is not a prerequisite to activate *Sb*-gRAMP, because RNA substrates lacking nucleotides in the PFS region were found to be cleaved efficiently (*11*, *18*). To validate these structural findings, we replaced the tip of the gating loop with a flexible linker to evaluate its importance in target RNA recognition (fig. S9, D and E). Wild-type *Sb*-gRAMP did not bind or cleave RNA that only base-pairs with the first 9 nt of the crRNA spacer. By contrast, the gating loop mutant bound this target RNA efficiently and subsequently cleaved it (Fig. 3D). These experiments suggest that the gating loop plays a pivotal role in preventing off-targeting. Overall, our RNA-bound *Sb*-gRAMP structures support a mechanistic model in which the resting *Sb*-gRAMP exists in an autoinhibited state to avoid sequence-nonspecific RNA binding and cleavage. Target RNA is validated through crRNA pairing in a directional fashion from the 3′ to the 5′ region of the guide. Upon completion of target binding, movement of the gating loop initiates a chain of allosteric events to switch on the RNase centers in gRAMP (Fig. 3E and movie S1).

We further attempted to interpret the cleavage mechanism by comparing the pre- and postcleavage states (Figs. 1 and 3, F and G). EM densities suggest that the RNA substrate was cleaved after the third and ninth nucleotides (site 1 and site 2, respectively) (Fig. 2E), which is consistent with previous reports (*11*, *18*). Because cleavage is metal dependent, we identified multiple candidate residues around the cleavage sites that may contribute to metal coordination (generally acidic residues), proton shuttling (generally polar residues), and transition state stabilization (generally positively charged residues) (Fig. 3, F and G). In subsequent mutagenesis testing (fig. S11, B to D), RNA cleavage at site 1 was abolished by alanine substitutions of D547 in Cas7.2 and of R294, D298, Y367, and K371 in Cas11 (Fig. 3H). Because site 1 is assembled from residues in both Cas11 and Cas7.2, it may only become active after target binding–induced hinge motion in Cas11. Cleavage at site 2 was abolished by Cas7.3 mutations D698A (*11*) and D806A, but not by Cas11 mutations R323A and H328A (Fig. 3H). An allosteric effect was noticed: Site 1–disruptive mutations D547A and D298A impaired site 2 cleavage as well, and site 2 mutation H328A impaired site 1 cleavage instead. These mutants appeared to weaken or alter the RNA-binding mode of *Sb*-gRAMP, as revealed by electrophoretic mobility shift assay (Fig. 3H and fig. S11C). *Sb*-gRAMP containing the double mutations R294A/D698A or D547/D806A was efficient in RNA binding but completely inactive in RNA cleavage (fig. S11C). Such dead-gRAMP variants could be useful in RNA editing, tagging, or tracing applications.

Craspase architecture and component interfaces

To gain mechanistic insights into how the putative RNA-guided protease system may work, we reconstituted Craspase in its *apo* state, the matching PFS-containing RNA target-bound state, and the nonmatching PFS-containing target-bound state, and generated their corresponding cryo-EM structures at 3.7, 2.6, and 2.7 Å resolutions, respectively (fig. S12 to S14). The TPR-CHAT–binding surface is on top of the buried crRNA 5′ handle in *Sb*-gRAMP, architecturally similar to where the cOA synthetase (Csm1/Cas10) binds in canonical type III-A effector complexes (Fig. 4, A and B; fig. S15A; and movie S2). TPR-CHAT consists of an N-terminal TPR domain (amino acids 1 to 323), a dynamic midregion (amino acids 324 to 399), and a C-terminal cysteine protease from the caspase family (amino acids 400 to 717). The domain arrangement of TPR-CHAT resembles that of separase (*24*, *25*), an essential eukaryotic

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

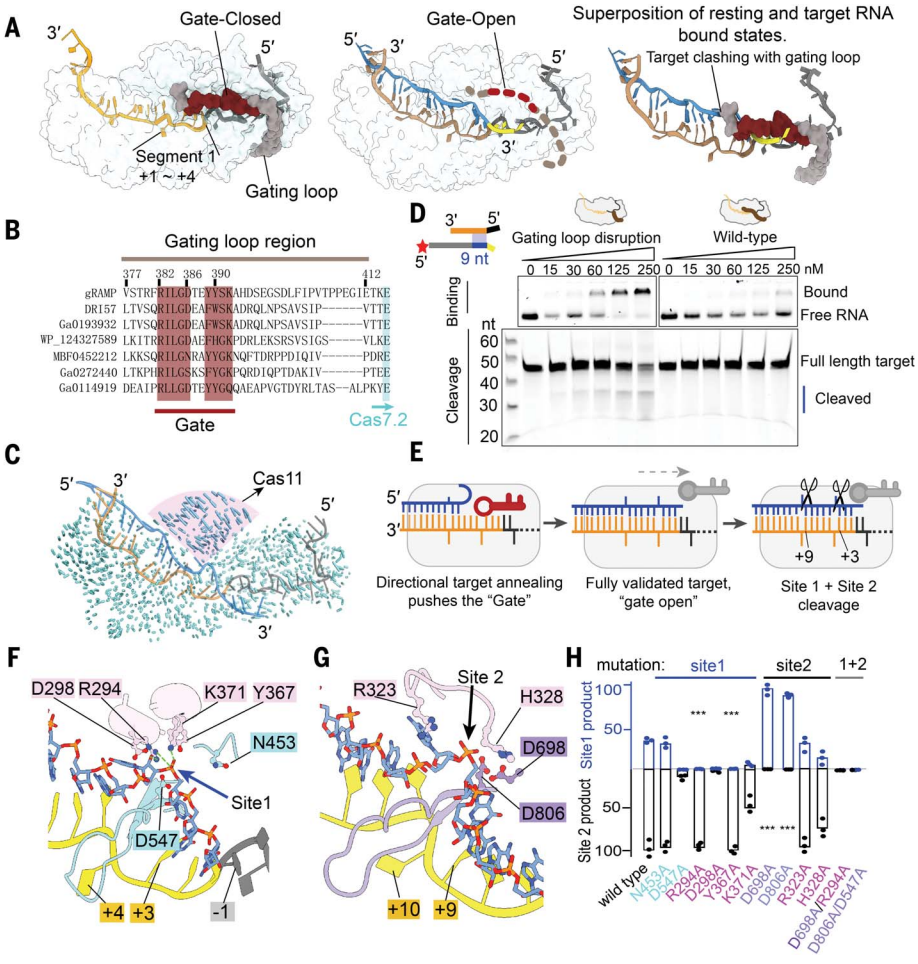


Fig. 3. Target validation and cleavage mechanisms by *Sb*-gRAMP RNP. (**A**) Models depicting the gate-closed structure in resting state (left) and the gate-open structure in the target RNA–bound state (middle). Superposition is shown in the right panel. (**B**) Sequence alignment at the gating loop region. Conserved residues are highlighted in burgundy. (**C**) Structural comparison of the resting and nonmatching PFS RNA–bound states. Vector length is proportional to residue movement distance. Hinge motion in Cas11 is pronounced. (**D**) Electrophoretic mobility shift assay (top) and urea–polyacrylamide gel electrophoresis (bottom) to evaluate the impact of gating loop disruption on the binding and cleavage of partially matching RNA targets. (**E**) Mechanistic model depicting the essential role of the gating loop in target validation. (**F**) Structural basis for site 1 cleavage. (**G**) Structural basis for site 2 cleavage. (**H**) Impact of site 1 (in blue) and site 2 (in black) mutations on target RNA cleavage efficiency.

protein that cleaves cohesin rings to allow chromosome segregation (fig. S15, B to D). Like separase, the CHAT domain contains an N-terminal pseudocaspase domain, a long dimeric coiled-coil midinsertion, and a C-terminal active-protease domain (*24*, *25*). Although structurally distinct, the two caspase domains pack in a similar fashion as the eukaryotic caspase dimers (*26*). In TPR-CHAT, the β-sheet structure in the pseudocaspase domain interacts with the TPR domain, and the midregion serves as the sole anchoring point of CHAT onto *Sb*-gRAMP. The TPR repeats belong to the so-called solenoid domains, which are assembled from repeating structural units and often mediate protein-protein or protein-ligand interactions (*27*). The seven TPR repeats in TPR-CHAT pack side by side to form a C-shaped architecture, with the seventh TPR repeat packing against the β sheet of the globular CHAT domain. TPR-CHAT adopts the rough shape of a padlock, with TPR being the shackle and CHAT the body (fig. S15B). In the Craspase structure without target RNA (*apo*-Craspase), the shackle of the padlock captures a long "switch helix" (amino acids 338 to 362) in the middle. The switch helix is captured by the molecular contacts from the inward-facing loops in the TPR repeats. When wedged in the shackle, the switch helix pins down a loop-helix-loop structure underneath (amino acids 324 to 337). Together, they mediate an extensive set of molecular contacts to multiple regions inside the padlock (fig. S15B), including contacts to the tips of two long β hairpins (sensor hairpins) that further extend all the way to the protease center in CHAT (fig. S15E).

An ~75 × 35 Å^2 area of the Cas7.1 surface in *Sb*-gRAMP is buried by TPR-CHAT (Fig. 4, C and D). However, the actual physical contacts between TPR-CHAT and *Sb*-gRAMP are limited to two surface patches 50 Å apart. On the TPR side, a hydrophobic patch in the first and second TPR repeats makes hydrophobic and main-chain hydrogen bond contacts to a portion of the gating loop (F381, I383, and L384) and a nearby Cas7.2 loop (L450 and V451) (Fig. 4C). A more extensive and mostly hydrophobic interface is found between one of the coiled-coil helixes in the CHAT domain (amino acids 434 to 450) and two regions of *Sb*-gRAMP, namely the C-terminal portion of the gating loop (amino acids 396 to 403) and the Zn knuckle of Cas7.2 (Fig. 4D). In particular, Y450 and L499 of CHAT insert into a hydrophobic pocket on the *Sb*-gRAMP surface, promoting shape complementarity at the interface. The interaction between gRAMP and TPR-CHAT was completely disrupted by Y75A and F103A mutations in the TPR interface and severely impaired by A445R and L449A/Y450A mutations in the CHAT interface (Fig. 4E). An important observation is that the gating loop of *Sb*-gRAMP, which plays a pivotal role in regulating the RNase activity of *Sb*-gRAMP through conformational changes, is sandwiched between *Sb*-gRAMP and TPR-CHAT (fig. S16A). Whereas the entire gating loop becomes unstructured in the RNA-bound *Sb*-gRAMP structure, only the tip of it is rearranged in the RNA-bound Craspase (Fig. 3A and fig. S17). Given this conformational restriction, we speculated that the energetic barrier for RNase activation may be higher in Craspase compared with *Sb*-gRAMP. Indeed, RNA binding was consistently weaker at different temperatures, and the cleavage was slower in Craspase compared with *Sb*-gRAMP (Fig. 4F and fig. S16, B and C).

RNA-guided protease activation mechanism in Craspase

When Craspase is in the resting state, the catalytic dyad in the TPR-CHAT protease center, Cys627 and His585, are 6.6 Å apart (fig. S18). Because this exceeds the hydrogen-bonding distance by a large margin, C627 could not be deprotonated by H585 and thus could not initiate the nucleophilic attack on the peptide substrate. Our structure therefore suggests that TPR-CHAT in the *apo*-Craspase is an inactive protease. When Craspase is bound to a target RNA with a matching PFS (Fig. 5A), a perfectly base-paired first segment is formed between the guide and target. Constrained by the base pairing from the first two PFS residues to the guide, the remaining PFS nucleotides point toward the bottom of TPR. Although their densities are difficult to model, possible phosphate

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

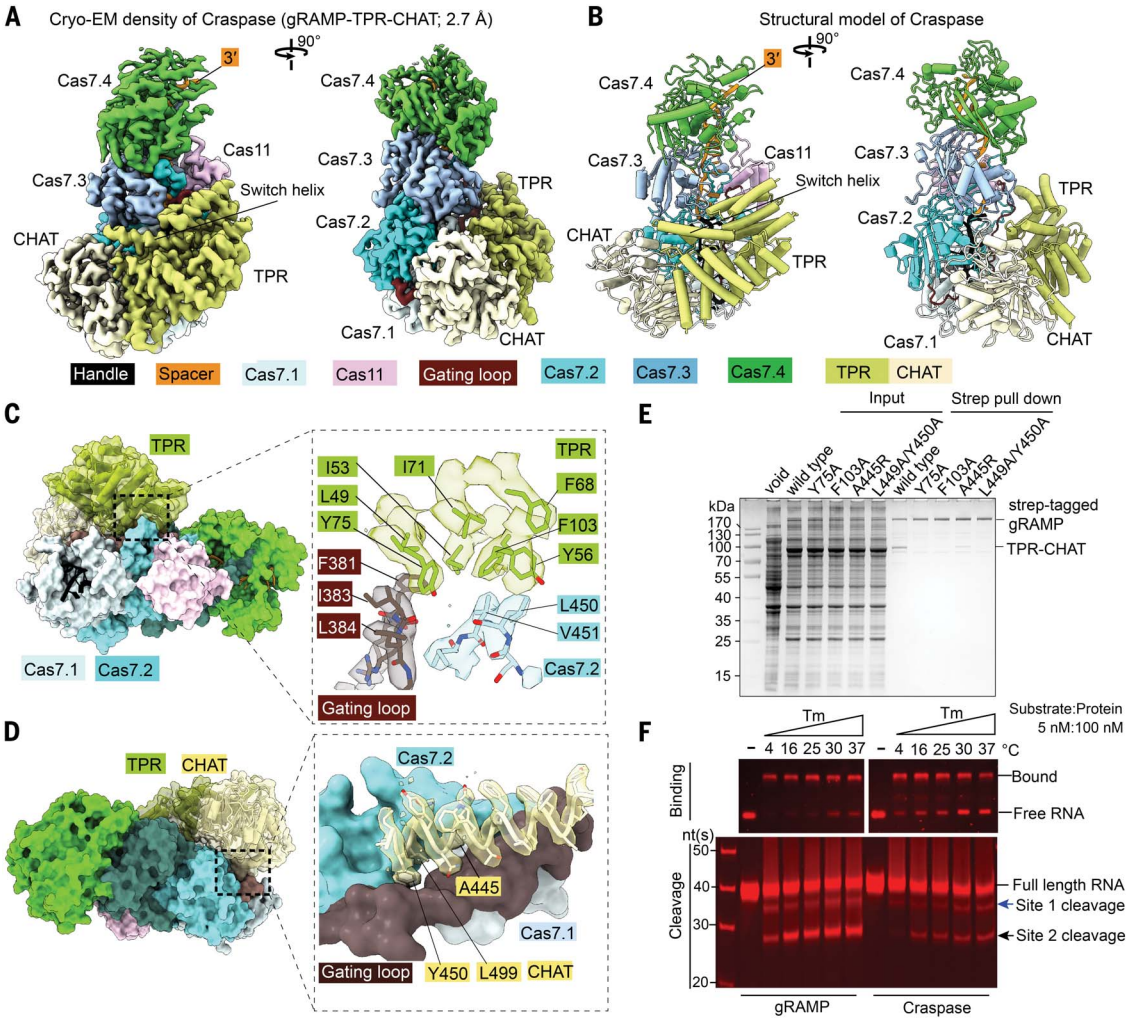


Fig. 4. Structural basis for Craspase assembly. (**A**) 2.7-Å cryo-EM density and (**B**) structural model of Craspase (gRAMP-TPR-CHAT). (**C**) Location and enlarged view of the molecular contacts between gRAMP and TPR. Interface residues and corresponding cryo-EM densities are shown. (**D**) Location and enlarged view of the molecular contacts between gRAMP and CHAT. (**E**) Strep-tag affinity purifications quantifying the impact of interface mutations on Craspase complex formation. (**F**) Electrophoretic mobility shift assay (top) and urea–polyacrylamide gel electrophoresis (bottom) to quantify activity differences in RNA binding and cleavage by gRAMP and Craspase.

densities suggest that the PFS travels underneath TPR (Fig. 5C and movies S3 to S5). This path may have perturbed the conformation dynamics of the sensing β hairpin in CHAT because its tip that may contact PFS becomes disordered. This coincides with a backbone twitch at the protease center, on the opposite end of the sensing hairpin (amino acids 626 to 631) (Fig. 5D). C627 and H585 reside on the two strands of the sensing hairpin. The allosteric change shortens their distance from 6.6 to 5.2 Å (fig. S18B). This distance, however, is still too far to allow H585-mediated C627 deprotonation. Moreover, the nearby side-chain–binding pocket found in the *apo* structure is closed after the structural rearrangement (Fig. 5D). Therefore, the matching PFS RNA–bound Craspase is not expected to be proteolytically active either.

A greater set of conformational changes take place when RNA target containing a nonmatching PFS is bound by Craspase (Fig. 5B). Lacking sequence complementarity to the first 2 nt of PFS, the base pairing in the first guide-target segment is incomplete and the gating loop is only partially dislodged (fig. S17). Although the first nucleotide of PFS forms a partially frayed A•C pair, the rest of PFS pivots toward the surface of the TPR (Fig. 5C). The switch helix is dislodged from the shackle of the padlock, possibly due to clashes with the nonmatching PFS. This helix and the preceding loop-helix-loop connection rotates 90° and packs against CHAT as a coiled-coil structure (Fig. 5D and fig. S19). The sensor hairpin undergoes a larger set of long-range allosteric alterations. Consequently, C627 and H585 become oriented within the hydrogen-bonding distance (3.3 Å) (Fig. 5D), and a hydrophobic pocket opens nearby (fig. S18C). The entire CHAT domain further undergoes a rigid-body movement. As the result, the cleft between *Sb*-gRAMP and TPR-CHAT widens, which may enable the peptide substrate to access binding surfaces (Fig. 5D).

On the basis of the observed structural features in the protease center, we designed candidate peptides to probe for potential RNA-guided peptidase activity in Craspase. We noticed that one designed peptide showed Craspase-dependent cleavage in thin-layer chromatography assays (Fig. 5E and fig. S20, A and B). Consistent with our mechanistic predictions, the activity was stronger in the presence of a nonmatching PFS RNA substrate than a matching PFS substrate (Fig. 5E). This peptide could also be cleaved by Craspase in the context of an interdomain protein linker, and the cleavage was stimulated by non-PFS target RNA (fig. S20, C to F). Mass spectrometry revealed that the cleavage took place after a leucine residue

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

(fig. S21). Judging by the fact that only one of the two leucine residues in the peptide was selectively cleaved (Fig. 5F) and that the cleavage activity was low and only partially RNA dependent, Craspase clearly specifies additional sequences nearby.

The above mechanistic analysis defines how RNA-guided RNA recognition regulates the protease activity of Craspase (Fig. 5G). Sequence complementarity in the target RNA is a prerequisite that is indirectly read out from the gating loop movement. A NOT logic gate is also in place to avoid activation by a self-RNA. Craspase is only activated when both conditions are true. The structural feature performing the logic calculation is the switch helix: Its movement triggers a stepwise conformational relay that allosterically unlocks the TPR-CHAT padlock and switches on the protease activity (Fig. 5H).

Craspase proteolytically cleaves Csx30 in an RNA-dependent manner

Type III-E *loci* encode three other well-conserved proteins: the putative sigma-factor RpoE and two proteins of unknown function denoted Csx30 and Csx31 (*5*, *11*, *12*). Because a protease and its target are often co-localized in the genome (*2*, *8*), we tested Craspase protease activity against these proteins in co-expression experiments (Fig. 6A). Full-length Csx30 was strongly reduced in the presence of target-bound Craspase, whereas full-length RpoE and Csx31 levels were unaffected (Fig. 6B). This effect was alleviated when Craspase carried inactivated cysteine-histidine residues (H585A and C627A) (Fig. 6B), suggesting that Craspase has proteolytic activity against Csx30. This observation was confirmed in vitro, where purified Craspase processes Csx30 into two distinct fragments (Fig. 6C and table S2), demonstrating that Csx30 is a natural protein target of Craspase. Mutational analysis of the amino acids encompassing the cleavage site showed that L407 in Csx30 is important for Craspase activity (tables S3 and S4 and fig. S22, A and B). Cleavage by Craspase after a leucine residue is consistent with mass spectrometry (fig. S22A) and the peptide cleavage experiments (Fig. 5, E and F). Corroborating the structural insights, proteolytic digestion could only be observed in the presence of target RNA with nonmatching PFS, whereas no cleavage fragments accumulated with nontarget RNA or target RNA with matching PFS (Fig. 6, C and D). Because Craspase cleaves bound RNA only under bivalent cation conditions (*11*), we reasoned that the peptidase in target-bound Craspase would stay active in the absence of magnesium ions. We indeed observed a marked increase in Csx30 processing under magnesium-poor conditions compared with magnesium-rich conditions (Fig. 6E), suggesting that target RNA cleavage switches off the peptidase. This is further supported by the finding that

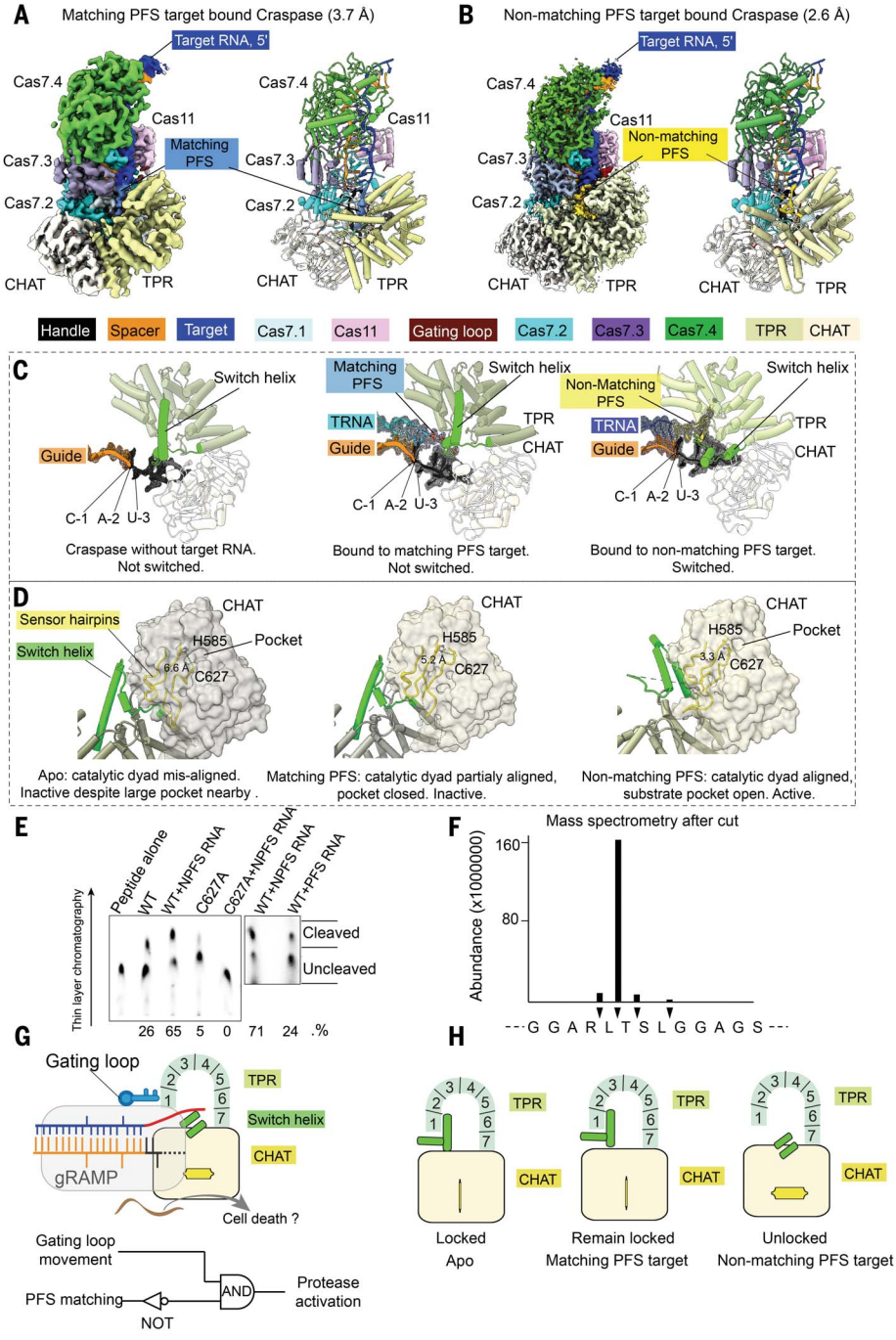


Fig. 5. Structural basis for Craspase protease activation. (**A**) 3.7-Å cryo-EM density (left) and structural model (right) of matching PFS RNA–bound Craspase. (**B**) 2.6-Å cryo-EM density (left) and structural model (right) of nonmatching PFS RNA–bound Craspase. (**C**) Close-up views of the switch helix in the resting state (left), the matching PFS RNA–bound state (middle), and the nonmatching PFS RNA–bound state (right). The switch helix is highlighted in green, and the density of crRNA and target RNA (TRNA) are shown in mesh. (**D**) Conformation of the switch helix and sensor hairpin in three states. Changing status in the catalytic dyad and the nearby side-chain–binding pocket in CHAT (gray surface) are highlighted. (**E**) TLC-based peptide cleavage assay by Craspase. (**F**) Cleavage site mapping by mass spectrometry. Cleavage percentages are quantified. (**G**) Top: model depicting nonmatching PFS RNA–induced Craspase activation. Bottom: Logic gate diagram illustrating the protease activation mechanism. (**H**) Model depicting TPR-CHAT status in the *apo* state, the matching PFS RNA–bound state, and the nonmatching PFS RNA–bound state.

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

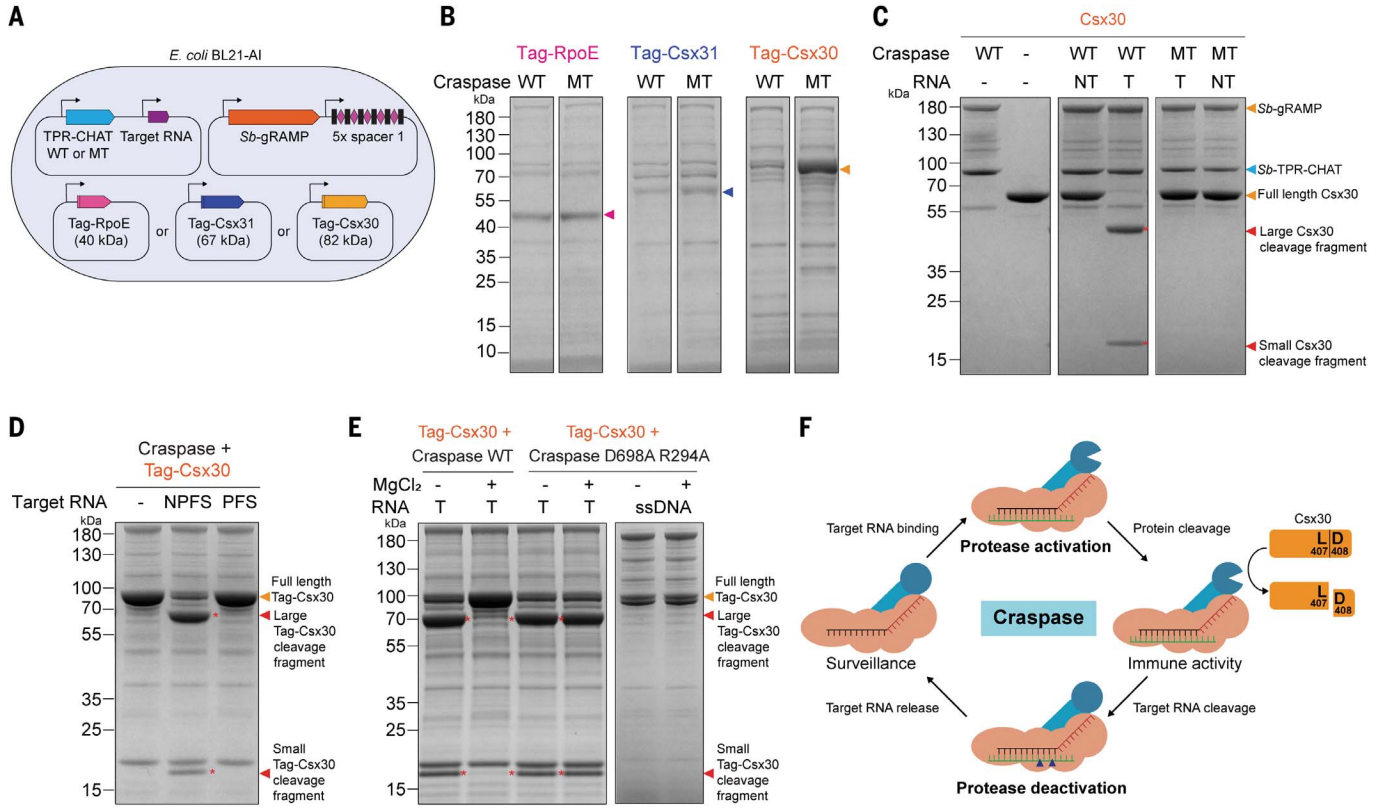


Fig. 6. Craspase proteolytically cleaves Csx30 in an RNA-dependent manner. (**A**) Genetic context for RpoE, Csx31, and Csx30 co-expression with Craspase wild-type (WT) or mutant (MT; H585A C627A) and a target RNA in *E. coli* BL21-AI. (**B**) Protein gel showing the eluted protein content from streptavidin purifications of Tag-RpoE, Tag-Csx31, and Tag-Csx30 after co-expression with either Craspase WT or Craspase MT (H585A C627A). Colored arrows indicate the expected size for the full-length protein. (**C**) Protein gels after Craspase WT or Craspase MT (H585A C627A) incubation with Csx30 in the presence of target RNA or nontarget RNA. Protein cleavage products are indicated with a red asterisk. (**D**) Protein gel after Craspase WT incubation with target RNA containing either a nonmatching PFS (NPFS) or matching PFS. (**E**) Left: protein gel after incubation of Tag-Csx30 with RNA and Craspase target WT or Craspase D698A R294A, with or without prior incubation with $MgCl_2$. Right: protein gel after incubation of Tag-Csx30 with single-stranded target DNA and Craspase D698A R294A. (**F**) Model for Craspase functionality. Once unbound Craspase has bound a target RNA, the peptidase activity is activated. This results in proteolytic cleavage of Csx30 between L407 and D408. Upon target RNA cleavage by Craspase, the peptidase activity is shut off.

the peptidase activity of a nuclease-dead variant of Craspase is not impaired in the presence of magnesium ions (Fig. 6E), rendering Craspase R294A D698A a "stay-on" variant. Binding of a complementary single-stranded DNA that is not cleaved by Craspase (*11*, *12*) does not activate the peptidase (Fig. 6E). These findings combined support a model (Fig. 6F) in which the peptidase activity of Craspase is switched on upon target RNA binding to cleave Csx30 after L407, separating a large N-terminal fragment of ~47 kDa from a small C-terminal fragment of ~19 kDa. Because of the low sequence and structural similarity to known proteins, a prediction of the function of the two protein fragments cannot be made with confidence (fig. S22C). However, on the basis of analogous defense systems, processed Csx30 fragments likely enable an immune response, possibly by eliciting toxicity to the native host cell. Craspase then self-regulates through target RNA cleavage to switch the peptidase off, thereby timing the duration of the immune response and possibly recycling the Craspase complex to bind new target RNAs.

Discussion

A new frontier in CRISPR-Cas biology has emerged, in which the RNA-guided effectors control physiological responses using mechanisms other than nucleic acid degradation. Here, we define how the Craspase protease is allosterically activated by target RNA recognition and inactivated by target RNA cleavage to cleave the native substrate Csx30 in a binary fashion. We tuned its dynamic response range using mechanism-inspired mutants, which will pave the way for biotechnological and therapeutic applications. Our observations suggest the possibility that the cleavage sequence in the native protein substrate is read out in the context of the 3D structure, which is also the case for the molecular recognition of gasdermins by eukaryotic caspases (*26*, *28*). We await follow-up studies revealing the missing recognition codes in substrate recognition and cleavage.

Despite the large structural distinctions, our studies revealed that type III-E systems share fundamental mechanistic similarities with canonical CRISPR-Cas type III systems. Analogous to Cas10 activation in other type III effectors, Craspase only turns on the protease activity in response to nonself RNA targets; it does not differentiate self and nonself RNA targets at the RNA cleavage level. This, combined with the observation that Craspase switches off protease activity upon target RNA cleavage, suggests that the protease activity may only be desired temporarily in the cell, which points to a possible ominous consequence of turning on the Craspase pathway. Does Csx30 proteolysis lead to cell dormancy or possibly programmed cell death? Because of the lack of homology to known proteins, it is difficult to infer the physiological function of Csx30 with confidence. On the basis of the AlphaFold-predicted structure (*22*), we speculate that proteolysis may relieve a physical sequestration or trigger a conformational change in Csx30, converting

Downloaded from https://www.science.org on February 04, 2024

Corrected 13 June 2023. See full text.

it to the active form (fig. S22, C and D). An analogous scenario was described for bacterial gasdermin, which only induced its antiviral effect after site-specific cleavage by TPR-CHAT (*2*). The potential involvement of the other Craspase-associated proteins, RpoE and Csx31, needs to be assessed in future experiments. However, unraveling the biological details is complicated by the difficulty of working with the native host, *Ca.* S. brodae (*29*). Alternative model organisms may be needed for future functional studies. On the application side, the fact that the Craspase peptidase is only active in the presence of a specific RNA species renders it useful for both in vivo (e.g., gene expression profiling) and in vitro (e.g., RNA diagnostics) biotechnological applications. This represents a major expansion of the range of biomolecular engineering possibilities of CRISPR-Cas effectors.

REFERENCES AND NOTES

1. S. Doron *et al*., *Science* **359**, eaar4120 (2018).
2. A. G. Johnson *et al*., *Science* **375**, 221–225 (2022).
3. S. A. Shah *et al*., *RNA Biol.* **16**, 530–542 (2019).
4. K. S. Makarova *et al*., *Nucleic Acids Res.* **48**, 8828–8847 (2020).
5. K. S. Makarova *et al*., *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
6. M. Kazlauskiene, G. Kostiuk, Č. Venclovas, G. Tamulaitis, V. Siksnys, *Science* **357**, 605–609 (2017).
7. O. Niewoehner *et al*., *Nature* **548**, 543–548 (2017).
8. C. Rouillon *et al*., SAVED by a toxin: Structure and function of the CRISPR Lon protease. bioRxiv, 2021.2012.2006.471393 (2021).
9. R. Molina, N. Sofos, G. Montoya, *Curr. Opin. Struct. Biol.* **65**, 119–129 (2020).
10. L. You *et al*., *Cell* **176**, 239–253.e16 (2019).
11. S. P. B. van Beljouw *et al*., *Science* **373**, 1349–1353 (2021).
12. A. Özcan *et al*., *Nature* **597**, 720–725 (2021).
13. Y. Shi, *Cell* **117**, 855–858 (2004).
14. J. Ding *et al*., *Nature* **535**, 111–116 (2016).
15. J. Shi *et al*., *Nature* **526**, 660–665 (2015).
16. M. L. Hochstrasser, J. K. Nuñez, *Nat. Microbiol.* **6**, 1481–1482 (2021).
17. R. J. Catchpole, M. P. Terns, *Mol. Cell* **81**, 4354–4356 (2021).
18. K. Kato *et al*., *Cell* **185**, 2324–2337.e16 (2022).
19. N. Jia *et al*., *Mol. Cell* **73**, 264–277.e5 (2019).
20. M. Guo *et al*., *Cell Res.* **29**, 305–312 (2019).
21. S. Sridhara *et al*., *Commun. Biol.* **5**, 279 (2022).
22. J. Jumper *et al*., *Nature* **596**, 583–589 (2021).
23. J. A. Steens *et al*., *Nat. Commun.* **12**, 5033 (2021).
24. Z. Lin, X. Luo, H. Yu, *Nature* **532**, 131–134 (2016).
25. A. Boland *et al*., *Nat. Struct. Mol. Biol.* **24**, 414–418 (2017).
26. Z. Liu *et al*., *Immunity* **53**, 106–114.e5 (2020).
27. G. L. Blatch, M. Lässle, *BioEssays* **21**, 932–939 (1999).
28. K. Wang *et al*., *Cell* **180**, 941–955.e20 (2020).
29. T. Awata *et al*., *Appl. Environ. Microbiol.* **79**, 4145–4148 (2013).

ACKNOWLEDGMENTS

We thank E. Tan for the critical reading of the manuscript. **Funding:** This work was supported by the National Institutes of Health (grant GM118174 to A.K.); the Department of Defense through the National Defense Science & Engineering Graduate Fellowship Program (G.S.); the National Science Foundation (MRSEC program grant DMR-1719875 to the Cornell Center for Materials Research Shared Facilities); the Department of Energy Office of Biological and Environmental Research (grant KP1607011 to the Laboratory for BioMolecular Structure); the Netherlands Organisation for Scientific Research (NWO VICI grant VI.C.182.027 to S.J.J.B.); the European Research Council (ERC) Consolidator Grants under the European Union's Horizon 2020 Research and Innovation Program (grant 101003229 to S.J.J.B.). **Author contributions:** Conceptualization: A.K., C.H., S.J.J.B., S.P.B.v.B.; Funding acquisition: A.K., S.J.J.B.; Investigation: C.H., G.S., K.H.N., F.D., Y.C., A.K., S.J.J.B., S.P.B.v.B., A.C.H., A.R.-M., M.P., M.V.; Methodology: C.H., G.S., K.H.N., F.D., Y.C., A.K., S.J.J.B., S.P.B.v.B., A.C.H., A.R.-M., M.P.; Project administration: A.K., S.J.J.B.; Supervision: A.K., S.J.J.B.; Visualization: C.H., A.K.; Writing – original draft: A.K., C.H., G.S., S.J.J.B., S.P.B.v.B.; Writing – review and editing: A.K., C.H., G.S., S.J.J.B., S.P.B.v.B. **Competing interests:** S.P.B.v.B. and S.J.J.B. are inventors on patent application N2028346 and PCT/NL2022/050296 submitted by Delft University of Technology that covers uses of gRAMP and Craspase. A provisional patent application related to this research has been filed by Cornell University. The remaining authors declare no competing interests. **Data and materials availability:** The resting-gRAMP coordinates and cryo-EM density map have been deposited in the Protein Data Bank (PDB:8D97) and the Electron Microscopy Data Bank (EMD-27257); gRAMP/nonmatching PFS RNA–bound (PDB:8D8N, EMD-27252); gRAMP/matching PFS RNA–bound (PDB:8D9E, EMD-27259); gRAMP/nonmatching PFS RNA postcleavage state (PDB:8D9I, EMD-27263); Craspase complex (PDB:8D9F, EMD-27260); Craspase/matching PFS RNA complex (PDB:8D9H, EMD-27262); Craspase/nonmatching PFS RNA complex (PDB:8D9G, EMD-27261). Plasmids used in this study are available upon request. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add5064
Materials and Methods
Figs. S1 to S22
Tables S1 to S5
References (*30*–*36*)
Movies S1 to S7
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

Submitted 15 June 2022; accepted 16 August 2022
Published online 25 August 2022
10.1126/science.add5064

Downloaded from https://www.science.org on February 04, 2024