CRISPR

The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase

Sam P. B. van Beljouw[1,2], Anna C. Haagsma[1,2], Alicia Rodríguez-Molina[1,2], Daan F. van den Berg[1,2], Jochem N. A. Vink[1,2], Stan J. J. Brouns[1,2]*

Type III CRISPR-Cas immunity is widespread in prokaryotes and is generally mediated by multisubunit effector complexes. These complexes recognize complementary viral transcripts and can activate ancillary immune proteins. Here, we describe a type III-E effector from *Candidatus* "Scalindua brodae" (*Sb*-gRAMP), which is natively encoded by a single gene with several type III domains fused together. This effector uses CRISPR RNA to guide target RNA recognition and cleaves single-stranded RNA at two defined positions six nucleotides apart. *Sb*-gRAMP physically combines with the caspase-like TPR-CHAT peptidase to form the CRISPR-guided caspase (Craspase) complex, suggesting a potential mechanism of target RNA–induced protease activity to gain viral immunity.

Facing constant predation by viruses and other mobile genetic elements (MGEs), prokaryotes have evolved multiple defense systems to protect themselves (*1*). Among these are the CRISPR-Cas systems, which provide adaptive immunity. Invaders are recognized and inactivated by effector ribonucleoprotein complexes, and during this process genetic memory of the invader is generated and stored as spacers in CRISPR arrays (*2*). CRISPR-Cas effector complexes consist of CRISPR-associated (Cas) proteins bound to CRISPR RNA (crRNA) derived from a long transcript of the precursor CRISPR array (pre-crRNA) (*3*). Complementary binding of the crRNA to a target nucleic acid signifies the detection of an invader, setting the immune reaction in motion. Depending on the type of effector protein, immunity is typically reached either through direct cleavage of the invader's genetic material or by the activation of ancillary nucleases. CRISPR-Cas type III proteins encompass both functionalities, making them among the most sophisticated CRISPR-Cas effector proteins known to date (*4*).

Recently, a new CRISPR-Cas subtype that resembles type III systems was bioinformatically predicted and classified as type III-E (*5*). Although type III-E effectors are related to other type III effectors (Fig. 1A), they are quite different in terms of protein architecture. Whereas type III effectors are typified by a multisubunit composition, the type III-E effector seems to have various repeat-associated mysterious protein (RAMP) domains fused together, hence the nickname g(iant)RAMP (*5*). The unusual size of ~1300 to 1900 amino acids makes gRAMP the largest single-unit effector found to date (Fig. 1B). We found that spacers embedded in type III-E CRISPR arrays have targets in MGEs (Fig. 1C and table S1) with a bias toward targeting the coding strand of open reading frames (Fig. 1D), indicating that gRAMP activity likely involves interaction with invader mRNA. Some type III-E loci carry Cas1 fused to a reverse transcriptase (fig. S1), suggesting that the type III-E acquisition machinery actively selects spacers from RNA (*6*). Curiously, the gRAMP gene clusters lack ancillary nuclease genes but often co-occur with a gene encoding a TPR-CHAT protein (fig. S1) (*5*), a caspase-like peptidase found to be involved in regulated bacterial cell death (*7*). This is suggestive of a functional relation between the CRISPR-Cas and caspase families in type III-E antiviral activity that we sought to uncover in the present study.

To investigate the molecular composition and function of the gRAMP protein, we selected gRAMP from the type III-E locus in *Candidatus* "Scalindua brodae" (*Sb*-gRAMP) (Fig. 1E) (*8*). We introduced an *Escherichia coli* codon optimized version of this protein into *E. coli*, together with a plasmid containing five copies of the first spacer from the native CRISPR array (Fig. 2A and table S2). We purified *Sb*-gRAMP to apparent homogeneity in three consecutive chromatography steps and observed distinct 260-nm absorption during size exclusion chromatography (SEC), indicative of copurifying nucleic acids (fig. S2). Multiangle light scattering (MALS) analysis indicated a homogeneous particle of 242.5 ± 2.4 kDa (Fig. 2B), consistent with the expected size for a *Sb*-gRAMP monomer bound to a single-stranded RNA (ssRNA) species in the range of 45 to 60 nucleotides (nt) (see the supplementary text and table S3). Subsequent protein analysis showed a single band at the expected size for *Sb*-gRAMP (Fig. 2C). Nucleic acid extraction revealed the presence of three well-defined RNA populations, of which the population at ~50 nt represents mature crRNA (Fig. 2D). These observations were reproducible when coexpressing *Sb*-gRAMP with a different repeat-spacer pair (fig. S3, A to D), demonstrating that *Sb*-gRAMP can be loaded with a crRNA of choice.

The *Candidatus* "Scalindua brodae" type III-E locus contains a CRISPR array comprising 11 spacers interspaced by a 36-nt repeat sequence (table S2). To investigate the characteristics of mature crRNA in more detail, we used polymerase chain reaction to amplify the native CRISPR array from *Candidatus* "Scalindua brodae" genomic DNA and coexpressed it with *Sb*-gRAMP. After RNA extraction (fig. S4) and RNA sequencing (Fig. 2E and fig. S5), we determined that this RNA contains an unusually long 5′ handle of predominantly 27 to 28 nt, including the conserved last 14 nt of the repeat (Fig. 2F and fig. S6). Although a dominant spacer portion of 20 nt was observed, by far most mature crRNAs were truncated within a window of 16 to 25 nt (Fig. 2F and figs. S3E and S6B), suggestive of less-specific processing on the 3′ end of the crRNA. In other type III systems, dedicated proteins (e.g., Cas6) are required for the generation of mature crRNA from pre-crRNA (*9*). Our discovery of spacer-sized RNA bound to *Sb*-gRAMP, combined with the absence of genes known to be involved in pre-crRNA processing (fig. S1), suggests that *Sb*-gRAMP might be capable of processing its own pre-crRNA, similar to Cas12 and Cas13 (*10*, *11*). The larger RNA molecules bound by *Sb*-gRAMP might represent pre-crRNA processing intermediates. We obtained lower *Sb*-gRAMP protein yields and more degraded forms of *Sb*-gRAMP without the provision of pre-crRNA (Fig. 2G), suggesting an essential stabilizing or chaperoning role for the crRNA in *Sb*-gRAMP-crRNA complex formation. These results demonstrate that the *Sb*-gRAMP protein and a crRNA of typically 47 nt form a single ribonucleoprotein complex.

Given the apparent structural domain homology between *Sb*-gRAMP and subunits of other type III effectors, it seems plausible that *Sb*-gRAMP has evolved through a series of domain fusions and insertions. *Sb*-gRAMP contains a small domain with remote resemblance to Cas11 (Csm2) and four domains that have structural homology to Cas7 (Csm3/Cmr4) subunits, one with a large insertion (Fig. 1E) (*5*). Because Csm3/Cmr4 have been demonstrated to be responsible for target RNA cleavage in other type III systems (*12*, *13*), we set out to investigate whether *Sb*-gRAMP also has endoribonuclease activity. We incubated purified *Sb*-gRAMP-crRNA with fluorescently labeled ssRNA substrates complementary to crRNA of spacer 1 (Fig. 3A) or spacer 3 (fig. S7A) and observed two well-defined RNA cleavage products (Fig. 3B and fig. S7B), with the smaller product accumulating over time (fig. S8). Most

[1]Department of Bionanoscience, Delft University of Technology, Delft, Netherlands. [2]Kavli Institute of Nanoscience, Delft, Netherlands.
*Corresponding author. Email: stanbrouns@gmail.com

Downloaded from https://www.science.org at Shanghai Normal University on October 08, 2021

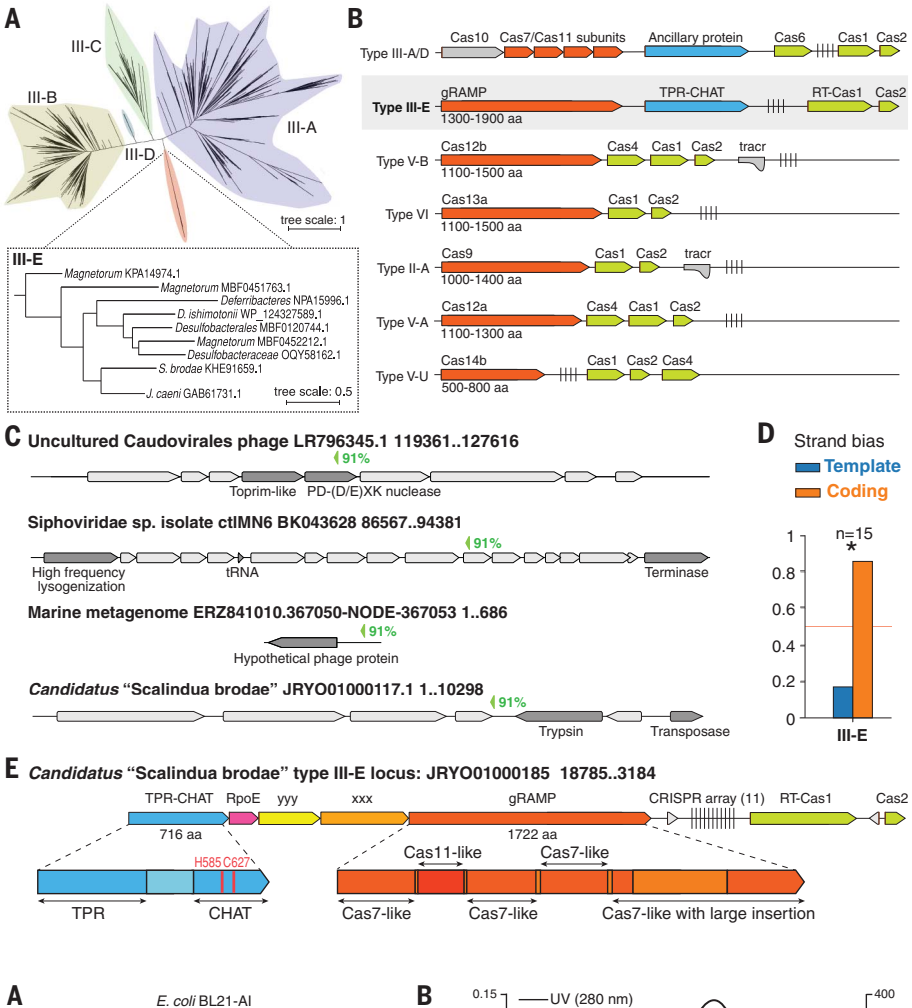


Fig. 1. CRISPR-Cas type III-E composition and characteristics. (**A**) Phylogenetic tree of type III subunit proteins (Csm3 and Cmr4) compiled with type III-E gRAMPs. The tree scale bar represents the number of substitutions per site. (**B**) Illustrations of typical organization of CRISPR-Cas types with sizes of effector proteins indicated in the number of amino acids (*25*). (**C**) Examples of contigs with matches to spacers from type III-E CRISPR arrays. (**D**) Fraction of spacers found in type III-E CRISPR arrays targeting template or coding strand (P = 0.011). (**E**) Detailed outline of the type III-E locus from *Candidatus* "Scalindua brodae."

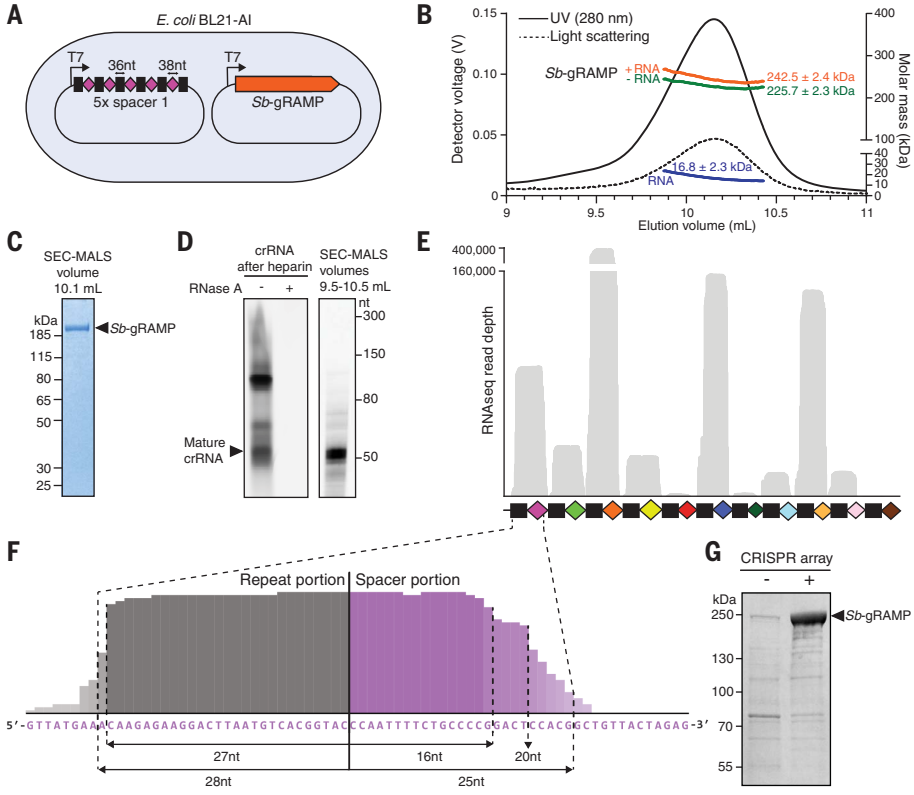


Fig. 2. *Sb*-gRAMP-crRNA is a ribonucleoprotein encoded by a single open reading frame. (**A**) Schematic of the heterologous expression system used for *Sb*-gRAMP-crRNA1 overexpression. (**B**) SEC-MALS peaks corresponding to light-scattering and ultraviolet absorbance at 280 nm (UV_{280}) profiles of purified *Sb*-gRAMP-crRNA1. The molar masses are derived from the peak half-height region (9.9 to 10.4 ml). (**C**) SDS–polyacrylamide gel electrophoresis (PAGE) analysis after SEC-MALS. (**D**) Denaturing urea PAGE analysis of the nucleic acid content in *Sb*-gRAMP-crRNA1. (**E**) RNA-sequencing reads of native crRNA with length 40 to 60 nucleotides mapped onto the native CRISPR array from *Candidatus* "Scalindua brodae." (**F**) RNA-sequencing read depth of the first repeat spacer with predominant repeat and spacer lengths indicated based on consensus processing. (**G**) SDS-PAGE analysis of *Sb*-gRAMP protein expressed in the absence and presence of a CRISPR array with five copies of spacer 1 (crRNA1).

Downloaded from https://www.science.org at Shanghai Normal University on October 08, 2021

target RNA remained uncleaved even at twofold molar excess of gRAMP (fig. S9), suggesting that endoribonuclease activity only proceeds with a limited substrate turnover. Counting from the 5′ side of the crRNA into the target, the cleavage sites were located after nucleotides 3 (site 1) and 9 (site 2) relative to the crRNA, thus 6 nt apart (Fig. 3B and figs. S7C and S10). When we tested target RNA substrates with the fluorescent label on the 3′ end instead of the 5′ end, two smaller cleavage products accumulated with a 6-nt interval (Fig. 3B and fig. S11). This nucleotide spacing between cleavage sites is reminiscent of the 6-nt periodicity found for other type III effector complexes (*14*, *15*), suggesting that the *Sb*-gRAMP retained the ancestral architectural positioning after domain fusion and insertion events. Also analogous to other type III effectors (*16*), complementarity between the protospacer flanking sequence (PFS) and the repeat portion of the crRNA retained the cleavage activity and positioning of *Sb*-gRAMP-crRNA within the complementary region of target RNA (Fig. 3C). Incubation of *Sb*-gRAMP-crRNA with an ssDNA target or noncomplementary ssRNA substrate did not yield cleavage products (Fig. 3D). We also did not observe collateral RNase activity in the presence of a target RNA (fig. S12), demonstrating that *Sb*-gRAMP is a sequence- and position-specific crRNA-guided RNA endonuclease.

The CRISPR-Cas type III Csm3 and Cmr4 domains harbor acidic residues essential for metal-ion coordination and subsequent ribonuclease catalysis (*14*, *15*, *17*). The removal of divalent cations from the cleavage buffer abolished *Sb*-gRAMP-crRNA ability to cleave target RNA (Fig. 3E), showing that RNase activity was dependent on metal ions. This raised the possibility that *Sb*-gRAMP has acidic amino acid residues essential for a functional endoribonuclease active site, akin to ancestral Csm3/Cmr4. To test this, we identified seven conserved aspartic acids and one serine residue that might be active in the ribonuclease catalytic core, based on multiple sequence alignment of various gRAMPs (fig. S13) and structural similarity to the *Streptococcus thermophilus* Csm3 (*12*) (fig. S14). Mutational analysis of these positions indeed revealed that *Sb*-gRAMP-crRNA lost its target RNA cleavage ability at site 2 upon changing a single aspartic acid residue to an alanine (D698A) (Fig. 3F). The single amino acid change renders a *Sb*-gRAMP variant able to sequence specifically cleave target RNA only at site 1, which may provide distinct benefits for biotechnological applications in which specific single RNA cleavage is desired. Despite trying the eight positions (fig. S15), the low sequence similarity of gRAMPs to Csm3/Cmr4 did not permit identifying essential active site residues responsible for the cleavage at site 1 and may require structural analysis. Currently described type III effectors cleave RNA once per Csm3/Cmr4 subunit; *Sb*-gRAMP, composed of four Csm3-like domains, only cleaved twice under our experimental conditions (Fig. 3G). This suggests that the other ribonuclease domains were either rendered dysfunctional over time or evolved into other functionalities.

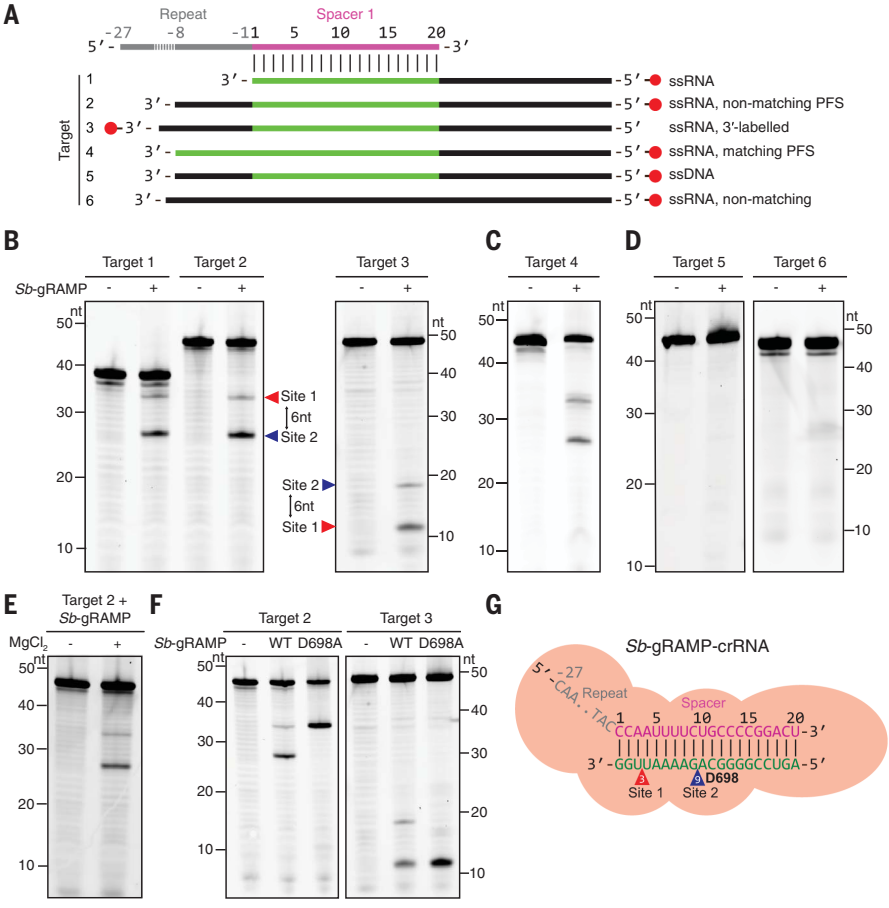


Fig. 3. *Sb*-gRAMP-crRNA drives dual ssRNA cleavage with 6-nt spacing. (**A**) Outline of the Cy5-labeled nucleic acid targets tested for *Sb*-gRAMP-crRNA1 activity. (**B** to **F**) Denaturing urea PAGE gels of *Sb*-gRAMP-crRNA1 incubated with complementary ssRNA [(B), cleavage products are indicated with arrowheads], ssRNA with a matching PFS (C), complementary ssDNA and noncomplementary ssRNA (D), and complementary ssRNA in the presence or absence of $MgCl_2$ (E) and with *Sb*-gRAMP-crRNA1 containing the D698A-inactivating mutation (F). (**G**) Schematic of *Sb*-gRAMP-crRNA1 hybridized to an ssRNA target with the exact cleavage sites indicated.

In canonical type III systems, RNA target recognition with mismatching PFS leads to the activation of the cyclase domain in the Cas10 subunit, which in turn synthesizes cyclic oligo adenylates (cOAs) (*18*, *19*). The cOAs act as allosteric activators of ancillary proteins often found in or near the CRISPR-Cas loci by binding in the CRISPR-associated Rossmann fold (CARF) domain (*20*). It appears that the Cas10 subunit as well as CARF-containing proteins were lost during the evolutionary genesis of gRAMP (*5*). Because this makes cOA signaling in type III-E unlikely, we wondered what the functional relation could be with the co-occurring caspase-like protease TPR-CHAT (fig. S1) (*5*). The tetratricopeptide repeat (TPR) domain is often involved in protein-protein interaction and the formation of protein complexes (*21*), which prompted us to test physical association between TPR-CHAT and *Sb*-gRAMP-crRNA.

To assess potential interaction, we affinity purified either *Sb*-gRAMP-crRNA or TPR-CHAT from cells that coexpressed both proteins (Fig. 4A) and observed co-elution using either protein as a bait (Fig. 4B and fig. S16A). This indicated the formation of a stable complex between *Sb*-gRAMP-crRNA and TPR-CHAT (table S4), the interaction of which was retained during subsequent heparin chromatography and SEC (Fig. 4B and fig. S17). MALS revealed a single, homogeneous population with a molar mass of 315.4 ± 2.8 kDa (Fig. 4C and fig. S18), corresponding to *Sb*-gRAMP-crRNA and TPR-CHAT complexed in a 1:1 stoichiometry (table S3). The retention time of the complex

Downloaded from https://www.science.org at Shanghai Normal University on October 08, 2021

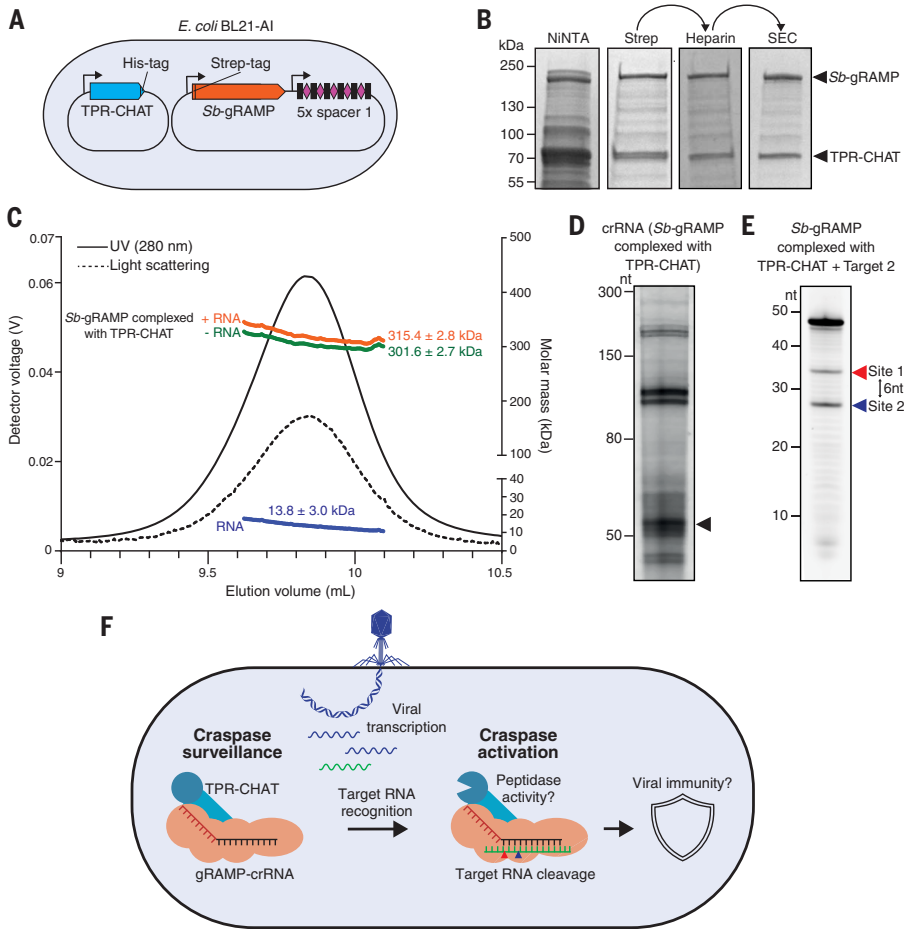


Fig. 4. *Sb*-gRAMP-crRNA and TPR-CHAT form the stable Craspase complex. (**A**) Schematic of the expression system used for overexpression of TPR-CHAT and *Sb*-gRAMP-crRNA1. (**B**) SDS-PAGE gels after the subsequent purification steps of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. (**C**) SEC-MALS light-scattering UV_{280} absorbance profiles of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. Molar masses are derived from the peak half-height region (9.6 to 10.1 ml). (**D**) Denaturing urea PAGE gel of RNA extracted from heparin-purified *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. Mature crRNA size is indicated with a black arrowhead and is based on the mature crRNA fraction found in *Sb*-gRAMP-crRNA1. (**E**) Denaturing urea PAGE gel of target cleavage by heparin-purified *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. Cleavage products are indicated with arrowheads. (**F**) Model of CRISPR-Cas type III-E during phage infection.

was shorter than that of *Sb*-gRAMP-crRNA alone, consistent with a particle of higher molecular weight (fig. S19). The heparin-purified complex contained mature crRNA (Fig. 4D) with 6-nt spaced cleavage specificity toward target RNA (Fig. 4E and figs. S20 and S16B). TPR-CHAT with inactivated predicted catalytic residues (H585A and C627A) (*7*) was also able to associate with *Sb*-gRAMP-crRNA (fig. S16), indicating that complex formation occurs independently of TPR-CHAT activity. TPR-CHAT did not dissociate from *Sb*-gRAMP-crRNA in the presence of target RNA (fig. S21), suggesting that *Sb*-gRAMP-crRNA and TPR-CHAT form a stable protein complex that remains assembled upon target RNA recognition. We here name this complex "Craspase" for "CRISPR-guided caspase."

The TPR-CHAT is a predicted protease from the caspase family, the members of which typically catalyze the hydrolysis of specific peptide bonds of target proteins (*22*), many remaining to be identified. Pathways involving caspase-like proteins are often activated during eukaryotic apoptosis but are also abundant in the bacterial kingdom, where they have been shown to function in regulated cell death (*7*). To test whether we could trigger cell death with the Craspase complex, we coexpressed it with an inducible target RNA in *E. coli* and followed the growth kinetics of the bacteria. We did not observe growth defects upon target RNA production (fig. S22), suggesting the absence of the target protein in the transplanted *E. coli* host, the requirement of additional factors for Craspase activation, or a biological role for Craspase in a context other than cell death induction.

The type III-E gRAMP effector has characteristics of both CRISPR-Cas class 1 (e.g., type III domains, 6-nt cleavage spacing, and cleavage of RNA independent of matching PFS) and class 2 (e.g., single protein composition), blurring the boundaries of the traditional CRISPR-Cas classification. Its physical association with TPR-CHAT raises the possibility of crRNA-guided caspase-like activity to reach viral immunity. Although the specific RNA cleavage capability of gRAMP-crRNA could contribute to direct antiviral defense, the primary role of target RNA recognition and cleavage may instead be the on and off switching of Craspase. This switching has been observed in other type III systems where RNA target interaction regulates second-messenger production, which in turn controls ancillary nuclease activity (*23*, *24*). Our findings that *Sb*-gRAMP-crRNA targets RNA and forms a stable complex with TPR-CHAT give rise to a model in which gRAMP-crRNA, instead of using second messengers, allosterically induces peptidase activity upon target RNA recognition to elicit a specific immune response (Fig. 4F).

REFERENCES AND NOTES

1. S. Doron *et al.*, *Science* **359**, eaar4120 (2018).
2. R. Barrangou *et al.*, *Science* **315**, 1709–1712 (2007).
3. S. J. J. Brouns *et al.*, *Science* **321**, 960–964 (2008).
4. R. Molina, N. Sofos, G. Montoya, *Curr. Opin. Struct. Biol.* **65**, 119–129 (2020).
5. K. S. Makarova *et al.*, *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
6. S. Silas *et al.*, *Science* **351**, aad4234 (2016).
7. A. G. Johnson *et al.*, bioRxiv 447441 [Preprint]. 7 June 2021.
8. M. Schmid *et al.*, *Syst. Appl. Microbiol.* **26**, 529–538 (2003).
9. F. C. Walker, L. Chou-Zheng, J. A. Dunkle, A. Hatoum-Aslan, *Nucleic Acids Res.* **45**, 2112–2123 (2017).
10. I. Fonfara, H. Richter, M. Bratovič, A. Le Rhun, E. Charpentier, *Nature* **532**, 517–521 (2016).
11. A. East-Seletsky *et al.*, *Nature* **538**, 270–273 (2016).
12. L. You *et al.*, *Cell* **176**, 239–253.e16 (2019).
13. C. Benda *et al.*, *Mol. Cell* **56**, 43–54 (2014).
14. R. H. J. Staals *et al.*, *Mol. Cell* **56**, 518–530 (2014).
15. G. Tamulaitis *et al.*, *Mol. Cell* **56**, 506–517 (2014).
16. M. P. Terns, *Mol. Cell* **72**, 404–412 (2018).
17. C. R. Hale *et al.*, *Cell* **139**, 945–956 (2009).
18. M. Kazlauskiene, G. Kostiuk, Č. Venclovas, G. Tamulaitis, V. Siksnys, *Science* **357**, 605–609 (2017).
19. O. Niewoehner *et al.*, *Nature* **548**, 543–548 (2017).
20. K. S. Makarova, V. Anantharaman, N. V. Grishin, E. V. Koonin, L. Aravind, *Front. Genet.* **5**, 102 (2014).
21. L. D. D'Andrea, L. Regan, *Trends Biochem. Sci.* **28**, 655–662 (2003).
22. L. Aravind, E. V. Koonin, *Proteins* **46**, 355–367 (2002).
23. C. Rouillon, J. S. Athukoralage, S. Graham, S. Grüschow, M. F. White, *eLife* **7**, e36734 (2018).
24. J. A. Steens *et al.*, bioRxiv 429135 [Preprint]. 1 February 2021.
25. P. Pausch *et al.*, *Science* **369**, 333–337 (2020).

ACKNOWLEDGMENTS

We thank M. Jetten and G. Nuijten (Radboud University) for providing *Candidatus* "Scalindua brodae" gDNA; C. de Agrela Pinto (Delft University of Technology) for performing SEC-MALS analysis; M. Pabst and C. de Ram (Delft University of Technology) for performing mass spectrometry analysis; R. Mans, C. Koster, C. Bastiaanssen, and J. van der Torre (Delft University of Technology) for provision of equipment and reagents; Brouns laboratory members for helpful discussions;

Downloaded from https://www.science.org at Shanghai Normal University on October 08, 2021

and A. Rita Costa (Delft University of Technology) for experimental help during revisions. **Funding:** S.J.J.B. is supported by the Netherlands Organisation for Scientific Research (NWO VICI; VI.C.182.027) and has received funding from the European Research Council (ERC) CoG under the European Union's Horizon 2020 research and innovation program (grant agreement no. 101003229).
Author contributions: Conceptualization: S.P.B.v.B., S.J.J.B.; Funding acquisition: S.J.J.B.; Investigation: S.P.B.v.B., A.C.H., A.R.-M., D.F.v.d.B., J.N.A.V.; Methodology: S.P.B.v.B., S.J.J.B., A.C.H.; Supervision: S.P.B.v.B., S.J.J.B.; Writing – original draft: S.P.B.v.B., S.J.J.B.; Writing – review and editing: S.P.B.v.B., S.J.J.B. with input from all other authors.

Competing interests: S.P.B.v.B. and S.J.J.B. are inventors on patent application N2028346 submitted by Delft University of Technology that covers uses of gRAMP and Craspase. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Plasmids and strains are available upon request from S.J.J.B.

SUPPLEMENTARY MATERIALS

https://science.org/doi/10.1126/science.abk2718
Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S8
References (*26*–*45*)
MDAR Reproducibility Checklist

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

2 July 2021; accepted 16 August 2021

Published online 26 August 2021
10.1126/science.abk2718

Downloaded from https://www.science.org at Shanghai Normal University on October 08, 2021

Science

The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase

Sam P. B. van BeljouwAnna C. HaagsmaAlicia Rodríguez-MolinaDaan F. van den BergJochem N. A. VinkStan J. J. Brouns

Science, 373 (6561), • DOI: 10.1126/science.abk2718

CRISPR and Caspase meet

Many prokaryotes use CRISPR RNA–bound proteins to sense viral RNA instead of DNA to set an immune response in motion that protects from virus infection. Although these ribonucleoproteins are typically composed of many protein subunits, van Beljouw *et al*. discovered that CRISPR-Cas type III-E systems are formed by a large, single-component effector protein capable of double RNA cleavage. Distinct from other systems, this effector forms a complex with a peptidase from the caspase family, raising the intriguing possibility that viral RNA activates a protease activity to prevent virus propagation by host suicide. —DJ

View the article online
https://www.science.org/doi/10.1126/science.abk2718
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org at Shanghai Normal University on October 08, 2021

Use of think article is subject to the Terms of service

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.
Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works