On CRISPR-controlled Proteases

van Beljouw, S.P.B.

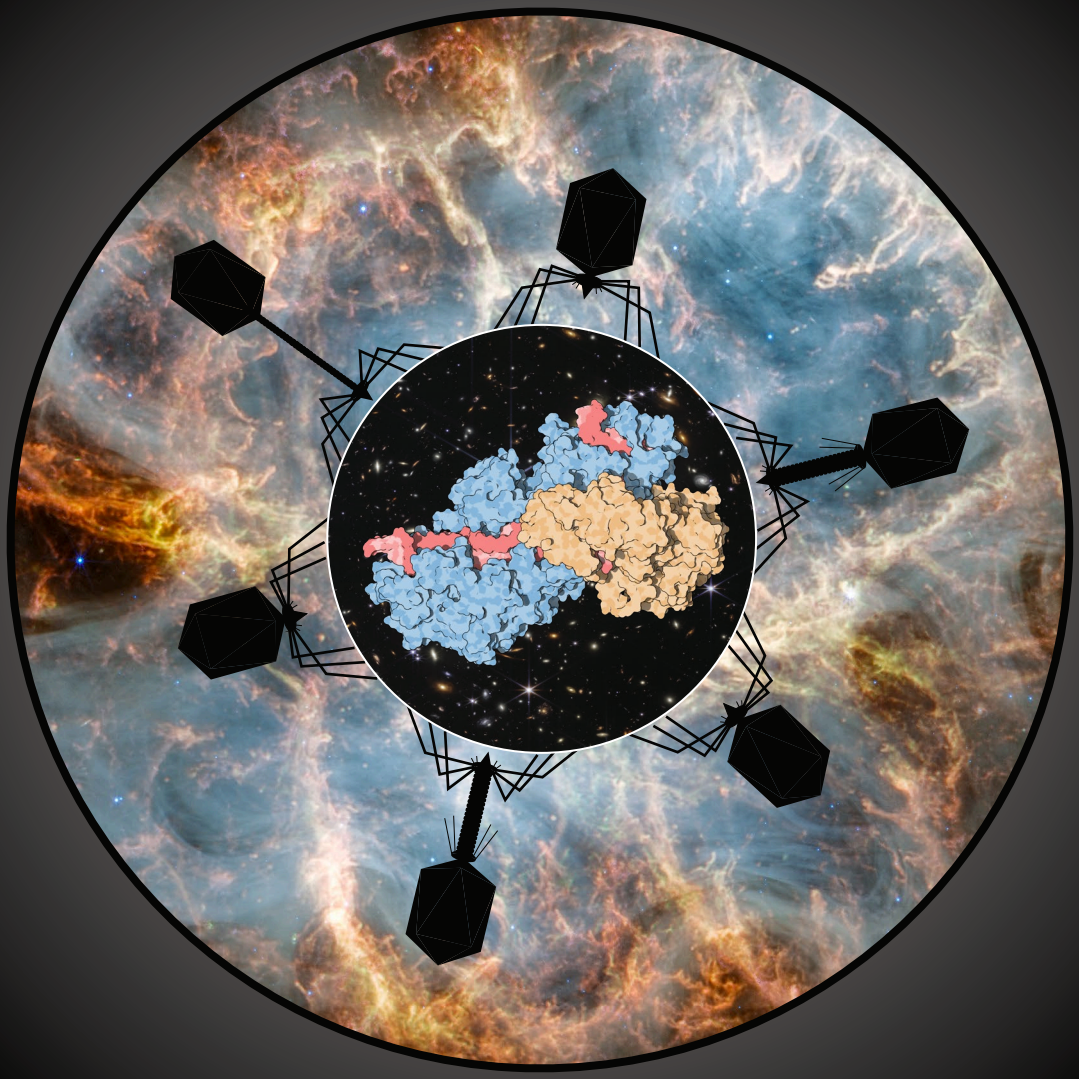
DOI
10.4233/uuid:82a05f19-df6b-463a-b2ab-18e310821262

Publication date
2024

Document Version
Final published version

Citation (APA)
van Beljouw, S. P. B. (2024). *On CRISPR-controlled Proteases*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:82a05f19-df6b-463a-b2ab-18e310821262

ON CRISPR-CONTROLLED PROTEASES



Sam P. B. van Beljouw

Propositions

accompanying the dissertation

ON CRISPR-CONTROLLED PROTEASES

by

Sam Pelle Bertram VAN BELJOUW

1. The single-subunit gRAMP effector evolved from the multi-subunit type III effector instead of the other way around.
   This proposition pertains to Chapter 7 of this thesis.
2. In the greater evolutionary picture, immune and anti-immune systems are better characterized as 'facilitator systems', akin to how the opposing movement of bicycle pedals facilitates a hybrid direction forward.
   This proposition pertains to the Epilogue of this thesis.
3. It makes sense to look for the target substrate of a protease in its immediate genetic surroundings.
   This proposition pertains to Chapter 4 of this thesis.
4. Bacterial cells can distinguish self from non-self by inducing an autoimmune response.
   This proposition pertains to Chapter 2 of this thesis.
5. The domination of metaphysical frameworks by reductionist materialistic thought is a prime suspect in the degradation of Earth's biosphere.
6. In the domain of Craspase research, there is supportive evidence for the colloquial hypothesis suggesting that scientists are more predisposed to sharing each other's toothbrushes than adopting each other's terminology.
7. Current neuroscientific approaches cannot determine whether thoughts are products of the nervous system or autonomous entities that exist independently of it.
8. There is a possible method to selectively kill tumour cells using Craspase.
9. The present scientific endeavour effectively revolves around feeding detailed information to a coming AI that will be leading in the generation of novel hypotheses.
10. Asking whether the mind-altering experience is hallucination or observation is analogous to asking whether mathematics is invented or discovered.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotors Prof. dr. ir. S.J.J. Brouns and Prof. dr. C. Joo.

On CRISPR-controlled Proteases

On CRISPR-controlled Proteases

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on
Friday 7 June 2024 at 12:30 o'clock

by

Sam Pelle Bertram van Beljouw

Master of Science in Biotechnology,
Wageningen University & Research, the Netherlands,
born in Lieshout, the Netherlands.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. S.J.J. Brouns	Delft University of Technology, promotor
Prof. dr. C. Joo	Delft University of Technology, promotor
Independent members:	
Prof. dr. M.E. Tanenbaum	Delft University of Technology & Hubrecht Institute
Prof. dr. M.F. White	University of St Andrews
Dr. R.H.J. Staals	Wageningen University & Research
Dr. H.C. van Leeuwen	Organization for Applied Scientific Research TNO
Other member:	
Dr. M. Pabst	Delft University of Technology
Reserve member:	
Prof. dr. C. Dekker	Delft University of Technology

This work was financially supported by NWO Vici.

Keywords: CRISPR-Cas, type III-E, prokaryotic immunity, RNA, Craspase, CRISPR-controlled proteases, gRAMP, TPR-CHAT, host-parasite evolution.

Front: The pupil and the iris are composed of images taken by NASA's James Webb Space Telescope, showing Webb's first deep field photograph and the stellar remains of the Crab Nebula, respectively. Design by the author.

Printed by: Gildeprint.

Casimir PhD Series, Delft-Leiden 2024.

ISBN 978-94-6496-132-4

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

CONTENTS

SUMMARY

The ancient arms race between bacteria and viruses have driven the evolution of a large repertoire of bacterial immune pathways. Prominent among these are the CRISPR-Cas antiviral systems, well-known for their capacity to recognize and cleave invading nucleic acids. In this dissertation, we discover a unique CRISPR-Cas system that is not only active against RNA, but also protein.

In the **first chapter**, we briefly introduce the principles of bacterial immunity with an emphasis on RNA-targeting CRISPR-Cas biology. Next, we evaluate the bioinformatic prediction of the novel type III-E CRISPR-Cas system that serves as the starting point for our investigations. The type III-E system is characterized by the presence of genes encoding for gRAMP and TPR-CHAT, a putative RNA-targeting CRISPR-Cas protein and caspase-like protease, respectively. This is suggestive of a functional relationship between the CRISPR-Cas and protease families that we seek to uncover.

In the **second chapter**, we sketch out the larger conceptual frameworks of RNA-targeting CRISPR-Cas to situate the type III-E system. We explain the biological rationales of targeting viral RNA, provide a detailed analysis of the RNA-targeting type III and type VI CRISPR-Cas systems, and discuss the common theme of cellular dormancy or death as a bacterial immune outcome.

In the **third chapter**, we transplant the gRAMP gene and CRISPR array from the bacterium *Candidatus* "Scalindua brodae" into *Escherichia coli*, followed by extensive biochemical characterization. We find that gRAMP is a large protein containing a single-stranded guide RNA component derived from the CRISPR array that guides it to complementary RNA molecules. Binding of a target RNA unleashes two RNase enzymatic activities within gRAMP, resulting in cleavage of the bound RNA in two guide-defined locations.

In the **fourth chapter**, we perform pulling assays to reveal a stable interaction between gRAMP and TPR-CHAT. Because the CHAT domain resembles a eukaryotic caspase, we term this complex Craspase, for 'CRISPR-guided caspase'. In the Craspase complex, proteolytic activity of TPR-CHAT is triggered upon binding of RNA to gRAMP. Activated Craspase then cuts a host-encoded protein of unknown function, Csx30, into two fragments. Cleavage of the bound target RNA by gRAMP switches off the proteolytic activity, making Craspase a self-regulating protease that is controlled by specific RNA molecules.

In the **fifth chapter**, we deploy cryo-electron microscopy to capture Craspase in various conformations, allowing structural dissection of its activation and deactivation principles. The guide RNA is threaded through gRAMP, with TPR-CHAT docked onto gRAMP at the 5′-handle of the guide RNA. A gating loop in gRAMP, initially covering the 5′-region of the guide RNA, ensures target RNA pairing initiation at the 3′-side. Target RNA binding sterically displaces the gating loop, allowing pairing across the whole guide RNA. A subsequent clash between the end of the target RNA and a helix in the TPR domain likely relays a conformational change into the CHAT catalytic center to activate the protease.

In the **sixth chapter**, we biochemically characterize the Craspase ortholog from *Candidatus* "Jettenia caeni". This Craspase also cleaves its Csx30 substrate, albeit in a vastly

different amino acid sequence compared to Craspase from *Candidatus* "S. brodae". This demonstrates that Craspase orthologs do not recognize a conserved motif but instead evolved to be specific for their respective Csx30 protein. We exploit this feature by developing a biotechnological tool capable of simultaneous detection of multiple RNA species in a sample, serving purposes for nucleic acids diagnostics.

In the **seventh chapter**, we present a general discussion on Craspase. We provide an overview of important milestones in Craspase research, reflect on the findings of others that detailed and expanded the work contained in this dissertation, ponder upon open questions and unexplained phenomena, and explore the emerging theme of proteases controlled by CRISPR-Cas systems.

In conclusion, this dissertation contains the experimental characterization and conceptual understanding of Craspase, a CRISPR-controlled protease that adds proteins to the repertoire of cleavage targets in CRISPR-Cas immunity. The specific proteolytic action, stringent RNA-dependent activation prerequisites, and inherent self-regulatory ability make Craspase a clear example of the complexity that can emerge out of the endless bacterium-virus arms race.

SAMENVATTING

De eeuwenoude evolutionaire wapenwedloop tussen bacteriën en virussen heeft geleid tot een uitgebreid repertoire van bacteriële immuunsystemen. Prominent hierin zijn de CRISPR-Cas antivirale systemen, bekend om hun vermogen om binnendringende nucleïnezuren te herkennen en te knippen. In deze dissertatie ontdekken we een uniek CRISPR-Cas systeem dat niet alleen actief is tegen RNA, maar ook tegen eiwitten.

In het **eerste hoofdstuk** introduceren we kort de principes van bacteriële immuniteit, met nadruk op de biologie van CRISPR-Cas systemen die RNA knippen. Vervolgens evalueren we de bio-informatische voorspelling van het nieuwe type III-E CRISPR-Cas systeem dat het startpunt vormt voor onze onderzoeken. Type III-E CRISPR-Cas wordt gekenmerkt door de aanwezigheid van genen coderend voor gRAMP en TPR-CHAT, respectievelijk een vermeend RNA-knippend CRISPR-Cas eiwit en een caspase-achtige protease. Dit suggereert een functionele relatie tussen de CRISPR-Cas en protease families die we willen ontrafelen.

In het **tweede hoofdstuk** schetsen we de conceptuele kaders van CRISPR-Cas systemen die RNA knippen. We leggen de biologische redeneringen uit voor het knippen van viraal RNA, geven een gedetailleerde analyse van de RNA-knippende type III en type VI CRISPR-Cas systemen, en bespreken het gemeenschappelijke thema van kiemrust of celdood als een uitkomst van bacteriële immuniteit.

In het **derde hoofdstuk** transplanteren we het gen coderend voor gRAMP samen met de CRISPR-array van de bacterie *Candidatus* "Scalindua brodae" naar *Escherichia coli*, gevolgd door uitgebreide biochemische karakterisering. We ontdekken dat gRAMP een groot eiwit is met een enkelstrengige RNA-component, afkomstig van het CRISPR-array, dat gRAMP naar complementaire RNA-moleculen leidt. Het binden van zo'n RNA activeert twee RNase enzymatische activiteiten binnen gRAMP, resulterend in het knippen van het gebonden RNA op twee gedefinieerde posities.

In het **vierde hoofdstuk** voeren we co-expressie experimenten uit waarmee we een stabiele interactie tussen gRAMP en TPR-CHAT onthullen. Omdat het CHAT-domein lijkt op een eukaryotische caspase noemen we dit complex Craspase, voor 'CRISPR-geleide caspase'. In het Craspase complex wordt de proteolytische activiteit van TPR-CHAT geactiveerd door binding van RNA aan gRAMP. Geactiveerde Craspase knipt vervolgens een cellulair eiwit met onbekende functie, Csx30, in twee fragmenten. Het knippen van het gebonden RNA door gRAMP schakelt de proteolytische activiteit uit, wat Craspase een zelfregulerende protease maakt die wordt gecontroleerd door specifieke RNA-moleculen.

In het **vijfde hoofdstuk** gebruiken we cryo-elektronenmicroscopie om Craspase in verschillende conformaties vast te leggen, waardoor structurele dissectie van zowel de activerings- als deactiveringsprincipes mogelijk is. Het gRAMP eiwit is om het begeleidende RNA gevouwen, met TPR-CHAT aan gRAMP gebonden op het 5′-handvat van de begeleidende RNA. Een poortlus in gRAMP, die aanvankelijk het 5′-gebied van de begeleidende RNA bedekt, zorgt ervoor dat het paren van complementair RNA wordt geïnitieerd aan de 3′-zijde. Het binden van het RNA verplaatst de poortlus, waardoor paren over de hele begeleidende RNA mogelijk is. Een daaropvolgende botsing tussen het

einde van het complementaire RNA en een helix in het TPR-domein geeft waarschijnlijk een conformationele verandering door aan het katalytische centrum van CHAT om de protease te activeren.

In het **zesde hoofdstuk** beschrijven we de biochemische karakterisatie van de Craspase ortholoog van *Candidatus* "Jettenia caeni". Deze Craspase knipt ook het Csx30 substraat, zij het in een aanzienlijk andere aminozuursequentie vergeleken met Craspase van *Candidatus* "S. brodae". Dit toont aan dat Craspase orthologen geen geconserveerd motief herkennen, maar in plaats daarvan zijn geëvolueerd om specifiek te zijn voor hun respectieve Csx30 eiwit. We benutten deze eigenschap door een biotechnologisch instrument te ontwikkelen dat in staat is om gelijktijdige detectie van meerdere RNA soorten in een monster mogelijk te maken, ten behoeve van nucleïnezurendiagnostiek.

In het **zevende hoofdstuk** presenteren we een algemene bespreking van Craspase, waarin we een overzicht geven van belangrijke mijlpalen in het Craspase onderzoek. We reflecteren op de bevindingen van anderen die het werk in deze dissertatie hebben gedetailleerd en uitgebreid, denken na over open vragen en onverklaarde verschijnselen, en verkennen de opkomende trend van protease eiwitten gecontroleerd door CRISPR-Cas systemen.

Ter conclusie bevat deze dissertatie de experimentele karakterisering en het conceptuele begrip van Craspase, een protease gecontroleerd door CRISPR-Cas, waarmee eiwitten worden toegevoegd aan het repertoire van biomoleculen waartegen CRISPR-Cas immuniteit werkzaam is. De specifieke proteolytische werking, strikte RNA-afhankelijke activeringsvereisten en intrinsieke zelfregulerende capaciteit maken van Craspase een duidelijk voorbeeld van de complexiteit voortkomend uit de eindeloos creatieve wapenwedloop tussen bacteriën en virussen.

1

INTRODUCTION AND OVERVIEW

"As the raindrop descends, does it know its body will be absorbed by the root of grass or trees, to be consumed by animals, to flow into milk, one day at last to dance in the blood of a singing child?"

Paul R. Fleischman

A modified version of this chapter has been published as *van Beljouw et al., Nature Reviews Microbiology (2023)* [1] and *van Beljouw & Brouns, Biochemical Society Transactions (2024)* [2].

1

1.1. EVOLUTION IN THE MICROCOSMOS

Since the appearance of primordial microorganisms, life has been in an intricate relation with death. Major contributors to this are parasites: their emergence is thought to be inevitable in even the simplest of replicator systems [3]. Parasites prey on host machinery for their propagation, often at the expense of host viability. However, the simple observation of a rich biosphere around us teeming with life indicates that the first hosts did not simply collapse under parasitic pressure. Instead, defense systems evolved that enabled the host to survive by preventing parasitic propagation. But complete parasite eradication did not happen either: contemporary bioinformatics demonstrate that virtually all life forms harbor diverse evolved parasites, often termed mobile genetic elements (MGEs) [4]. MGEs are simple genetic organisms, such as bacteriophages [5], plasmids [6], conjugative elements [7], transposable elements [8], and phage-inducible chromosomal islands [9], capable of horizontal transference within or between genomes and cells (**Fig. 1**). The wide range of MGEs are thought to have resulted from early parasites that evolved counter-defense to protect against host defense, upon which the host evolved counter-counter-defense, and during this perpetual host-parasite arms race [10], functionalities are often swapped and recruited horizontally [11]. The host-parasite co-evolution is a main driver of increasing biological complexity, with the enormous variation and ingenuity in microbial immune systems as a prime example.

1.2. THE CRISPR-CAS IMMUNE SYSTEM OF BACTERIA

An average bacterium encodes several anti-MGE systems [12, 13]. Although these systems can be exploited by various MGEs to resolve conflict beyond traditionally emphasized host-virus interactions [14], they are often studied as cellular immune systems against viral invaders. In this context, interference of the viral infection can occur at various stages, from blocking the initial penetration at the cell membrane to preventing access to resources by initiating cell death [15]. The only form of adaptive immunity discovered in bacteria and archaea are the CRISPR-Cas (clustered regularly interspaced short palindromic repeats and associated proteins) systems. A hallmark feature of these systems is the presence of a CRISPR array, which consists of repeat sequences that are separated by variable sequences, termed spacers [16]. A spacer corresponds to segment of a virus, known as protospacer [17], which has previously been encountered. The CRISPR array thus provides a genetic memory bank of past invasions. The Cas proteins are often encoded in the neighborhood of CRISPR arrays and are involved in various stages of CRISPR-Cas immunity [18–21].

Transcription of the CRISPR array generates a precursor CRISPR RNA (crRNA), which is further processed to free individual short crRNAs [22, 23]. The crRNAs subsequently assemble with Cas proteins to form ribonucleoproteins, called effector proteins, that surveil the cell in search for nucleic acids that base pair with the loaded crRNA [24]. Successful complementary binding indicates the presence of an intruder, whereupon the effector complex initiates a protective response. Depending on the type of effector complex, two general strategies intended to abrogate infection are discernible: direct dismantling of the invading DNA via crRNA-guided cleavage, and interference with the invader's replication cycle that often involves the activation of downstream immune

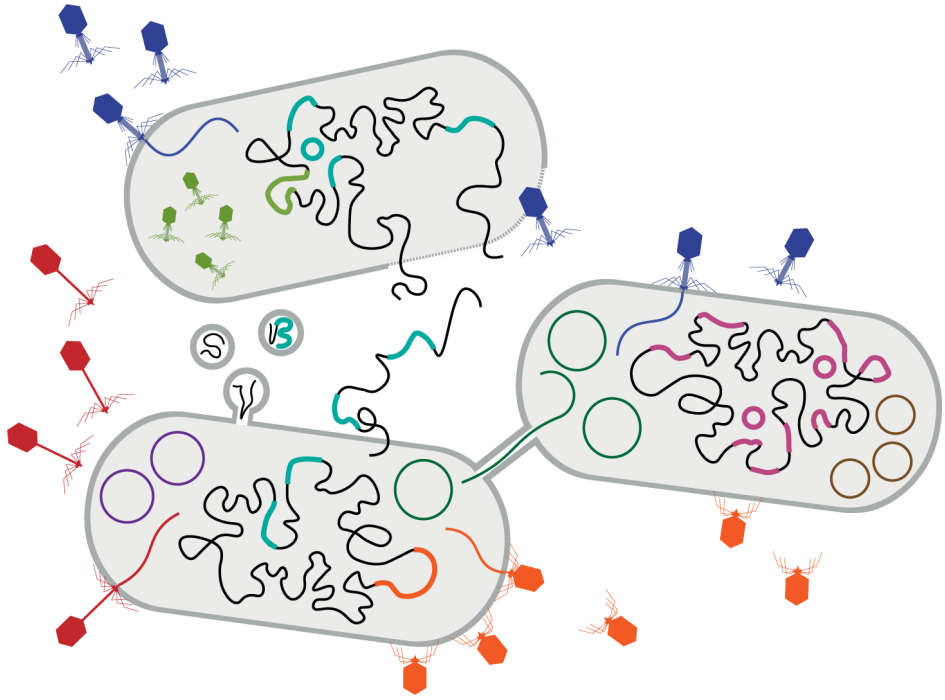


Figure 1. The ecology of mobile genetic elements. Mobile genetic elements (MGEs), such as bacteriophages (in red, orange and dark blue), plasmids (in purple, dark green and brown), transposons (in cyan and pink), and phage-inducible chromosomal islands (in light green), reside within and move between cells and genomes. MGEs can travel via the cytoplasm in cells, the medium separating cells, pores formed between cells, and vesicles budding from cells. Especially the movement of bacteriophages can have detrimental cellular outcomes, such as cellular membrane rupture followed by death (top cell).

proteins. The first strategy is used by the DNA-targeting effector complexes that belong to the CRISPR-Cas subtypes I, II, IV and V. In this case, target recognition is coupled to DNase activity initiated by the effector complex architecture. This leads to rapid degradation of the bound DNA to abolish infection without necessitating further action. The second strategy is used by the RNA-targeting effector complexes from the CRISPR-Cas subtypes III and VI. RNA-activated type III and type VI effectors set in motion a broad, collateral response that often goes far beyond merely cleavage of the bound RNA target; typical immune outcomes include cellular dormancy, a non-replicating cellular state defined by low or inactive metabolism to sustain under stressful conditions, and perhaps even cell death.

1.3. TARGETING VIRAL RNA IN TYPE III CRISPR-CAS

Type III CRISPR-Cas is believed to be the oldest member of the CRISPR-Cas family [25, 26]. Although type III systems can protect against an RNA virus under experimental conditions [27, 28] and spacer matches to RNA viruses have been discovered [29, 30], early

analyses of the CRISPR arrays found spacers matching with DNA viruses [31, 32] and thus viral transcript targeting by type III systems has been primarily studied. The effector complexes are typically composed of multiple subunits with signature subunit Cas10 being the largest component [33–37]. Cas10 is a multi-domain protein harboring a nuclease and a cyclase (also known as polymerase; from here cyclase/polymerase) domain, providing the effector complex with DNase activity as well as the capacity to generate second messengers from adenosine triphosphate (ATP), called cyclic oligoadenylates (cOAs), that can activate downstream immune proteins (**Fig. 2**). Other important enzymatic activities come from the multiple Csm3 (also known as Cas7) subunits in the effector complex, each facilitating generation of a cleavage in the bound target RNA [34]. The target RNA cleavage inactivates the Cas10, stopping cOA production and downstream immune activities [38, 39]. In this way, the type III effector self-regulates to tune its activity proportional to the amount of target RNA present. The other type III subunits are responsible for either complex formation with the Cas6-processed crRNA [22, 40], recruitment of host factors that promote crRNA maturation [41] and nucleic acid clearance [42], or assisting in binding a complementary RNA [34].

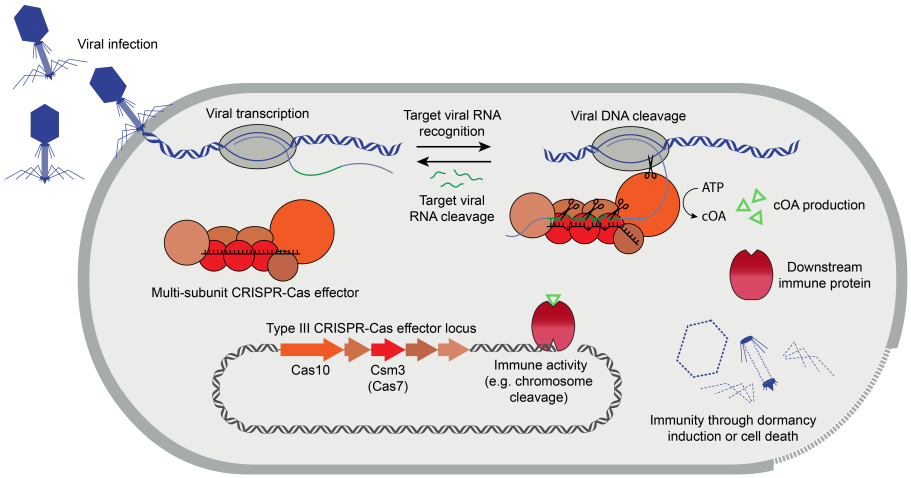


Figure 2. Canonical type III CRISPR-Cas biology. The multi-subunit type III CRISPR-Cas effector is composed of various Csm3 (Cas7) subunits as well as a Cas10 subunit, genetically encoded often near a CRISPR array. Upon transcription of target RNA from an infecting virus, the effector complex is capable of binding the transcript when it is complementary to the guide RNA bound by the effector. Upon pairing between target and guide RNA, the enzymatic activities of the Cas10 and Csm3 subunits are unleashed. Cas10 is capable of both cleaving the viral DNA, presumably to inhibit viral replication, as well as generating cyclic oligoadenylates (cOA) from ATP. The cOA molecules in turn activate downstream immune proteins often encoded in the type III CRISPR-Cas locus. One example of such immune protein is a DNase that induces breakages in the chromosome, leading to the death of the cell to prevent viral replication and its spread to neighboring cells.

1.4. BIOINFORMATIC PREDICTION OF TYPE III-E

In order to keep overview of the fast-moving CRISPR-Cas field, an updated classification of CRISPR-Cas systems is periodically published. In the edition of 2019 [43], a conspicuous new member aroused some consternation, as it was hard to classify along the conventional rules. Traditionally, new CRISPR-Cas members are grouped based on effector composition: class 1 CRISPR-Cas systems contain the multi-subunit effectors (e.g. Cascade in type I and Cmr/Csm in type III), whereas those in class 2 CRISPR-Cas systems comprise a single subunit (e.g. Cas9 in type II and Cas13 in type VI) [44]. The predicted system, found in a handful of primarily marine bacterial species, such as *Candidatus* "Scalindua brodae" and *Candidatus* "Jettenia caeni", was typified by the effector candidate gRAMP (for 'giant repeat associated mysterious protein') (**Fig. 3**). Classifying as a class 2 system seemed logical, as gRAMP is encoded by a single gene. However, domain composition analysis of gRAMP showed high similarity to various subunits of the Cmr/Csm effector complexes, which appear to have fused together to constitute gRAMP. This strongly suggested an evolutionary relationship between gRAMP and the type III systems, and the gRAMP loci were thus added to class 1 as type III-E CRISPR-Cas [43]. The gene encoding gRAMP contains Cas7-like folds —raising the possibility of RNA targeting akin to the canonical type III effector complexes [33–37]—, as well as a Cas11-like fold and a large insertion of unknown origin. A Cas10-like domain is absent from gRAMP, making cOA signaling unlikely; instead, gRAMP co-occurs with a caspase-like protease, TPR-CHAT, suggestive of a relationship between the caspase and CRISPR-Cas families. Furthermore, a transcription factor (RpoE) and two genes of unknown function (Csx30 and Csx31) hinted at potential complex immune pathways that gRAMP could be involved in. In the present dissertation, we sought to uncover these suspected intricacies of type III-E CRISPR-Cas biology.

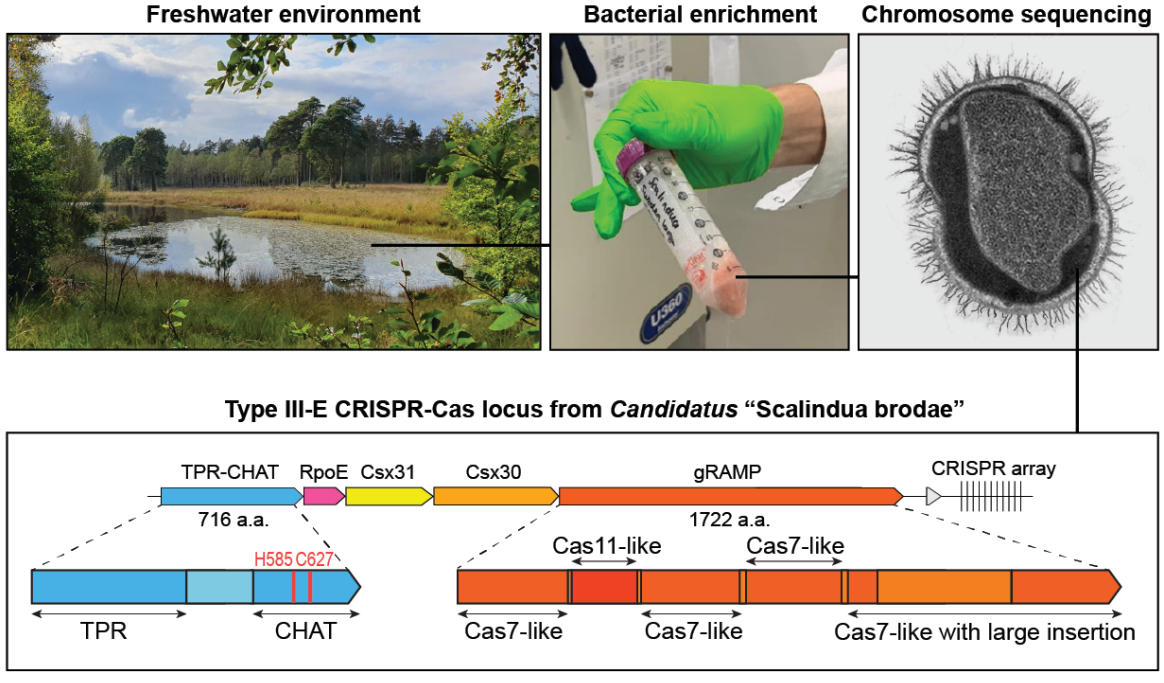


Figure 3. The type III-E CRISPR-Cas system isolated from marine microbes. The type III-E CRISPR-Cas system was bioinformatically predicted to occur in various bacteria, most of them native to freshwater environments. Batch reactor enrichment of *Candidatus* "Scalindua brodae" followed by chromosome sequencing revealed the composition of its type III-E CRISPR-Cas system. TPR-CHAT is a two-domain protein of 716 amino acids (a.a.), with the histidine and cysteine residues expected to form a protease catalytic dyad. The gRAMP protein is 1722 a.a. in size and contains a Cas11-like and multiple Cas7-like domains, with the most C-terminal Cas7-like domain possessing a large insertion of unknown origin. In between TPR-CHAT and gRAMP reside a predicted transcription regulation gene, RpoE, and two genes of unknown function, Csx30 and Csx31. A CRISPR array is located downstream of the gRAMP gene. The image of *Candidatus* "Scalindua" (top right) is adapted from [45], with permission.

1.5. DISSERTATION OVERVIEW

In the **second chapter**, we provide the broader experimental and conceptual understanding of RNA-targeting CRISPR-Cas biology. After exploring the general biological features of RNA-targeting CRISPR-Cas, we conduct an in-depth analysis of the RNA-targeting type III and type VI systems. Lastly, we explain the biological rationale behind immune responses involving dormancy to combat viral infections. In this context, we discover a principle cells employ to passively discriminate between self and non-self, which we term "blindfolded self/non-self discrimination". The bacterium employs this strategy to discern between self and non-self, capitalizing on its ability to temporally slow down metabolism to outlive rapidly replicating viruses.

In the **third chapter**, we describe the molecular composition and function of gRAMP from the anaerobic ammonium oxidizing (anammox) bacterium *Candidatus* "S. brodae". By heterologous expression of gRAMP in *Escherichia coli* and in vitro biochemical characterization, we established that gRAMP is a single effector ribonucleoprotein with several type III domains fused together. The crRNA component of gRAMP binds complementary RNA sequences, whereupon RNase domains in gRAMP cleave bound RNA in two specific positions, six nucleotides apart.

In the **fourth chapter**, we reveal the stable complex formed between gRAMP and TPR-CHAT using pulling assays and uncover its mechanism of action. Sequence-specific binding of RNA to the crRNA in the protein complex is the trigger for peptidase activity of TPR-CHAT, whereas cleavage of the bound RNA switches the peptidase off again. This intricate physical relation between proteins from the CRISPR-Cas and caspase families led us to term the complex Craspase, for 'CRISPR-guided caspase'. We further determine that activated Craspase site-specifically proteolyzes the host-encoded protein Csx30, a type III-E protein of unknown function.

In the **fifth chapter**, we mechanistically dissect Craspase using cryo-electron microscopy and biochemistry approaches. The overall architecture of gRAMP consists of a Cas11-like domain four non-identical Cas7-like domains, one with a big insertion domain. TPR-CHAT is bound at the 5′-handle of the crRNA in gRAMP to form Craspase. Craspase possesses a long linker hovering over the spacer nucleotides adjacent to the repeat, which physically prevents interaction of the crRNA with RNA molecules that are only matching in this region to likely reduce off-target activation. Binding of an actual target RNA leads to steric clashes with a helix in TPR-CHAT, triggering a large conformational change in the protease active center to accommodate the proteolytic activity on Csx30.

In the **sixth chapter**, we experimentally study the Craspase ortholog from anammox bacterium *Candidatus* "J. caeni" and find that it cleaves Csx30 in an amino acid sequence radically different from Craspase in *Candidatus* "S. brodae". This demonstrates that Craspase orthologs do not recognize a conserved motif and instead have an individually specified target sequence in the respective Csx30 substrate. We exploit this feature for the generation of a biotechnological diagnostics tool for multiplexed detection of RNA molecules.

Finally, in the **seventh chapter**, we present a discussion on open questions around Craspase and type III-E immunity, integrated with the plethora of complementary findings that type III-E research gained during the generation of this dissertation.

2

RNA-TARGETING CRISPR-CAS BIOLOGY

"We are drowning in information but starved for knowledge."

John Naisbitt

The CRISPR-Cas family contains adaptive immune systems found in many prokaryotic organisms. These systems provide protection against viral infection by targeting of invading nucleic acid sequences. Whereas some CRISPR-Cas systems recognize and cleave viral DNA, type III and VI CRISPR-Cas systems sense RNA that results from viral transcription and perhaps invasion by RNA viruses. The sequence-specific detection of viral RNA evokes a cell-wide response that typically involves global damage to halt the infection. How to make sense of an immune strategy that encompasses broad, collateral effects rather than specific, targeted destruction? In this chapter, we summarize the current understanding on RNA-targeting CRISPR-Cas systems. We detail the composition and properties of type III and VI systems, outline the cellular defense processes that are instigated upon viral RNA sensing and describe the biological rationale behind the broad RNA activated immune responses as an effective strategy to combat viral infection.

A modified version of this chapter has been published as *van Beljouw et al., Nature Reviews Microbiology (2023)* [1].

2.1. RECOGNITION OF RNA RATHER THAN DNA

Upon DNA genome ejection into the cytoplasm, the infecting virus will not only encounter the transcription and translation machinery required for its propagation, but also immune proteins that aim to stop it. DNA-targeting immune proteins (for example, restriction enzymes, RecBCD, and type I, II, IV and V CRISPR-Cas effectors) can become active immediately upon entry of viral DNA into the cell and given their high abundance in bacterial genomes, they are often dubbed the first intracellular line of immunity [12, 13, 46] (**Fig. 1A**). But viruses have evolved various ways of escaping this first line defense, including deploying anti-immune proteins (for example, regulation disruptors, anti-CRISPR proteins, RecBCD inhibitors and anti-restriction proteins [47–50]), changing recognition sequences (for example, mutating protospacer sequence and the protospacer adjacent motif (PAM) [51]), chemically modifying the DNA (for example, glycosylation and non-canonical nucleotide incorporation [52–54]), and physically protecting the DNA (for example, genome compartmentalization [55, 56]) (**Fig. 1B**). Moreover, even in case of successful targeting, viral genome replication may already have occurred to the point of outrunning the activity of DNA-acting systems. A failed or slow DNA-targeting response results in progression of the infection, beginning with the accumulation of viral transcripts. This is when the RNA-targeting CRISPR-Cas systems can come to the rescue.

The co-occurrence of type I and type III CRISPR-Cas loci [12] and the suggested sharing of spacers between type I and III effectors [57, 58] points at the possibility of cooperation between CRISPR-Cas systems targeting DNA and RNA. Indeed, it was found that viral escape from the DNA-targeting type I-F system was overcome through the targeting by a co-occurring type III-B system that used the type I-F crRNAs [59]. Cross-talk between RNA- and DNA-targeting CRISPR-Cas systems was also observed on the level of CRISPR adaptation, whereby a native type VI-B locus was enriched with functional spacers that were acquired by the machinery of a co-occurring type II-C system [60]. Besides cooperating synergistically, RNA- and DNA-targeting systems can also complement each other. An intriguing example of defense complementation was discovered in *Pseudomonas* and *Serratia* bacteria, in which infecting jumbo viruses build a nucleus-like proteinaceous compartment to shield their genome from various DNA-acting defenses [55, 56, 61] (**Fig. 1B**). RNA-targeting CRISPR-Cas was shown to still provide protection, as viral transcripts are translated outside the protective barrier, where they are exposed to type III and type VI effectors [56, 61]. In other words, CRISPR-Cas systems targeting RNA are not only able to temporally back up DNA-based immunity evasion but also do so spatially. Further studies will undoubtedly uncover more fascinating insights of the biological and ecological interactions between co-occurring CRISPR-Cas systems.

Instead of rapid eradication of the virus, tolerating an invader can be an advantageous cellular strategy. Temperate viruses are capable of integrating into the host chromosome, often bringing genes that potentially benefit host fitness [62]. Whereas systems targeting DNA destroy the virus independent of its potential benefits, RNA-targeting systems can conditionally tolerate viral presence by only suppressing toxic lytic gene transcripts [63, 64] (**Fig. 1C**). This enables the host to use viral gene content, while preventing the virus to become a threat.

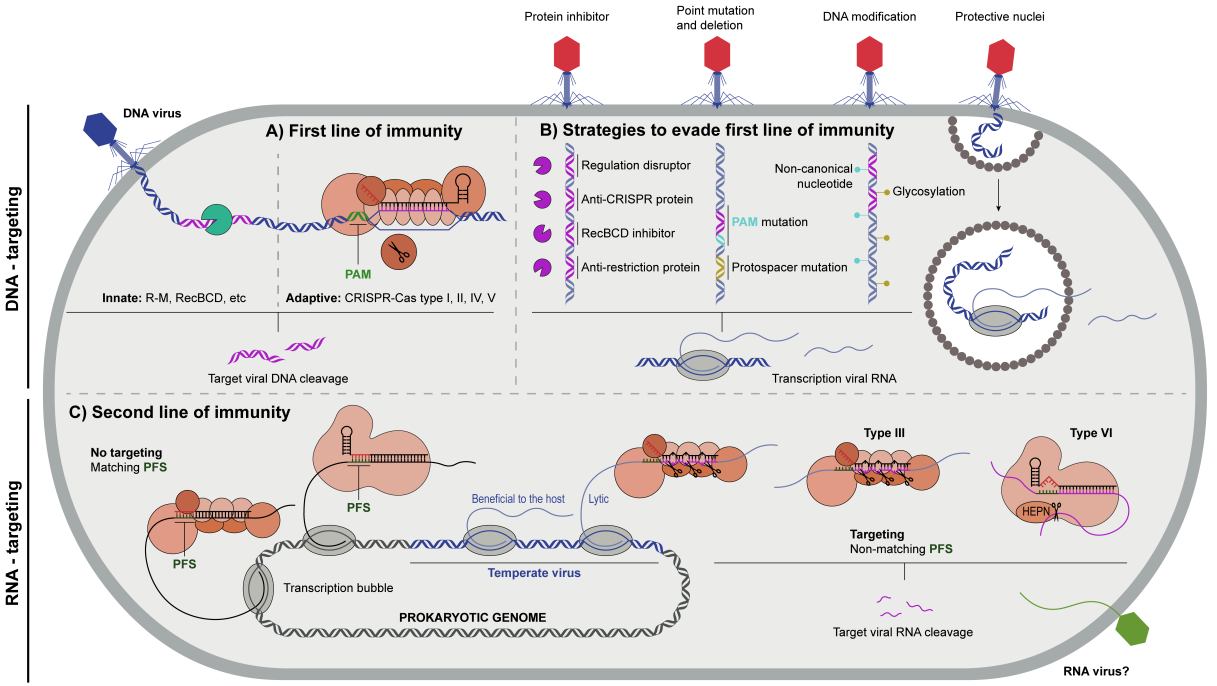


Figure 1. DNA- and RNA-targeting host defense systems. (**A**) Upon entering of the viral genome, DNA-targeting systems are immediately able to attack the target DNA (violet) once identified as non-self (for example, through protospacer adjacent motif (PAM) recognition) and hence form a first line of immunity. (**B**) Viruses have evolved various ways to evade the first line of immunity, including regulation disruptors, anti-CRISPR proteins, RecBCD inhibitors, anti-restriction proteins, and usage of protective nuclei. Additionally, PAM and protospacer mutations as well as DNA modifications can evade efficient DNA-targeting, resulting in progress of the viral lifecycle to transcription. (**C**) The presence of viral transcripts (or, perhaps, viruses with an RNA genome) is sensed by systems of the second line of immunity, including type III and VI CRISPR-Cas effectors. These effectors have lenient target base pairing requirements to enable targeting and cleavage of transcripts from mutated viruses. To prevent self-targeting, the RNA-targeting CRISPR-Cas effectors test pairing in the protospacer flanking sequence (PFS) to distinguish self from non-self. Targeting RNA instead of DNA enables toleration of integrated temperate viruses by only restricting lytic transcripts.

Thus, whereas the DNA-targeting CRISPR-Cas systems display characteristics of a first line of defense strategy (for example, immediate targeting, genome clearance and high specificity), RNA-recognizing CRISPR-Cas systems can be viewed as a second line strategy with various advantages. First, recognition on the RNA level circumvents hindrance by DNA modifications or other genome protecting mechanisms, giving the host an additional route to protect itself against foreign nucleic acids. RNA-recognizing effectors are also generally lenient toward mutations in the target sequence, making the chance of escape mutants less likely. Moreover, due to the transient nature of RNAs, an instance of autoimmunity through self-RNA cleavage is less detrimental to the host compared to self-targeting of DNA, and a full immune response is often not mounted when a self-RNA is recognized. Furthermore, being reliant on transcription enables conditional tolerance of viral presence, which enables the host to potentially benefit from the additional gene content. Finally, different cellular locations of DNA and RNA enables RNA-recognizing CRISPR-Cas systems to back up evaded DNA-targeting systems, as observed in defense against jumbo viruses.

2.2. DISSECTION OF RNA-TARGETING CRISPR-CAS

Although a shared feature of type III and type VI CRISPR-Cas, and some type II and type V systems, is the ability to sense target RNA, they are evolutionarily distant and structurally unrelated, thus differing greatly in terms of immune activity (**Table 1**). In this section, we review the molecular anatomy and mechanistic functioning of type III and type VI CRISPR-Cas systems.

2.2.1. TYPE III CRISPR-CAS

The type III effector complexes surveil the cell to detect RNA molecules complementary to the crRNA (**Fig. 2A**). Initial target RNA pairing occurs at the 3′ end of the crRNA, which induces a conformational change within the effector complex that enables base-pairing between the crRNA and target RNA [68]. To verify the origin of the bound transcript, complementarity is checked between the repeat-derived crRNA portion (also known as 5′ tag) and the PFS (also known as 3′ anti-tag) of the suspected RNA invader [31, 69–71]. Base pairing in this region indicates the binding of a self-RNA. When such false alarm occurs, Cas10 is kept locked in an inactive state to inhibit the immune response and reduce the toxic effects [31, 39, 68, 70, 72]. However, target cleavage does still occur: the RNase domains embedded in the Csm3 (in type III-A and III-D) and Cmr4 (in type III-B and type III-C) subunits of a target RNA-bound effector complex are exposed to the substrate and cleave it with a 6-nucleotide periodicity [27, 73–77]. Cleaved RNA fragments dissociate from the crRNA [38], which is believed to recycle the effector complex for binding of a new target (**Fig. 2A**).

Feature	Type III	Type VI
Class	1	2
Abundance	25% and 34% of total CRISPR-Cas loci in bacteria and archaea, respectively [44]	Rare in bacteria, absent in archaea [65]
Recognized nucleic acid	RNA	RNA
Effector composition	Multi-subunit (type III-A to III-D and III-F) and single-subunit (type III-E)	Single subunit (Cas13)
Pre-crRNA processing	External (Cas6)	Internal (Cas13)
Target RNA cleavage	Csm3 and Cmr4	Internal HEPN
Location of seed region in crRNA	At the 3′ end of the spacer	In the center of the spacer
Location of target RNA cleavage	In the crRNA binding region	Outside the crRNA complementary region
Cleavage specificity	6-nucleotide periodicity	Preference for certain nucleotide or (di)nucleotide motifs (Lsh-Cas13a and Lbu-Cas13a show preferred cleavage at U [66]; Lwa-Cas13a, Cca-Cas13b, Lba-Cas13a and Psm-Cas13b cleave efficiently at AU, UC, AC and GA, respectively [67])
Self/non-self discrimination	crRNA tag–anti-tag pairing	crRNA tag–anti-tag pairing
crRNA tag–anti-tag paired	RNA target cleavage, no Cas10 activation	No RNA target cleavage, no Cas13 collateral activity
crRNA tag–anti-tag unpaired	RNA target cleavage, Cas10 activation	RNA target cleavage, Cas13 collateral activity
Second messenger	cOA (Palm domain in Cas10)	tRNA fragments
DNase activity	Yes (HD domain in Cas10)	No
Experimentally studied secondary effector proteins	NucC, Card1 and Can2, Can1, Csm6, Csx1 and TTHB144, CalpL, Csa3	Csx28
Secondary effector target	Indiscriminate RNA or DNA degradation, protease activity	Indiscriminate RNA degradation and membrane depolarization
Secondary effector regulation	cOA regulation (ring nuclease, target cleavage)	Cas13 regulation (Csx27, WYL1)
Induction of dormancy	Yes	Yes

Table 1. Key aspects of type III and VI CRISPR-Cas. cOA, cyclic oligoadenylate; crRNA, CRISPR RNAs; HD domain, histidine–aspartate domain; HEPN, higher eukaryotes and prokaryotes nucleotide-binding.

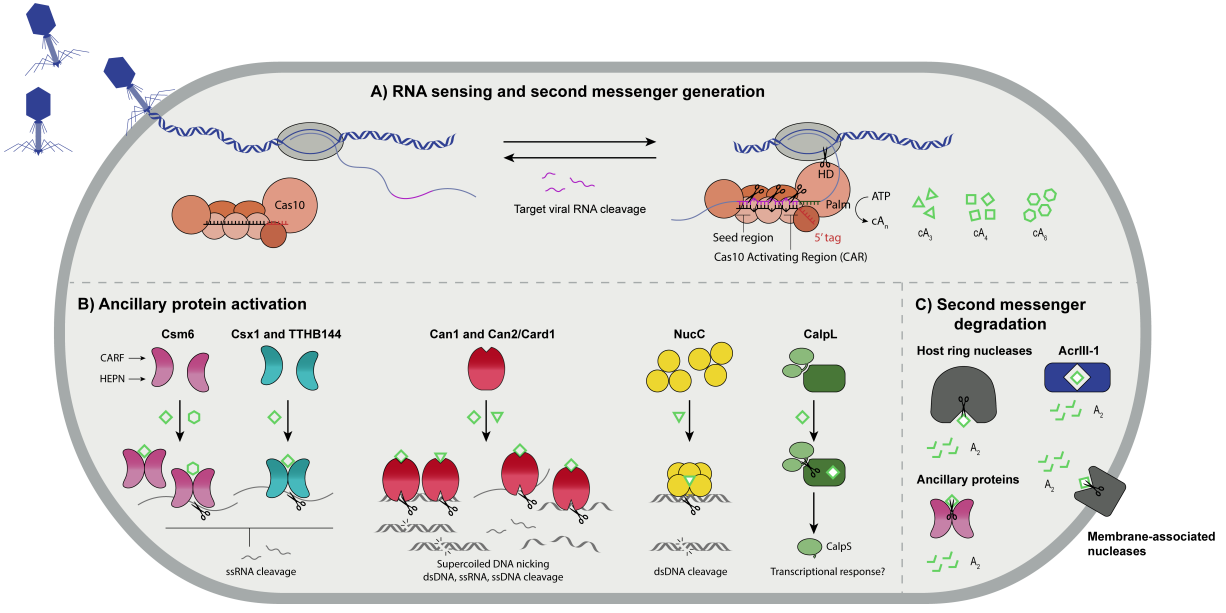


Figure 2. Schematic of type III CRISPR-Cas immunity. (**A**) Transcription of a viral target RNA (violet) is sensed by the multi-subunit type III CRISPR-Cas effector complex. Target recognition is initiated at the seed region in the 3′ end of the CRISPR RNA (crRNA), and sufficient binding in the Cas10-activating region (CAR) with mismatching 5′ tag (red) at the 3′ anti-tag (green) results in Cas10 activation. Type III CRISPR-Cas displays features of a sensor system, whereby recognition and cleavage of a target transcript leads to activation and deactivation of the Cas10 domains, respectively. The histidine–aspartate (HD) domain can cleave single-stranded DNA, perhaps close to the transcription bubble, whereas the Palm domain uses its cyclase functionality to generate cyclic oligoadenylate (cOA) molecules (cA_n, n = 3-6) from adenosine triphosphate (ATP). (**B**) Various ancillary proteins in type III CRISPR-Cas clusters are activated by different cOA species, or perhaps through direct interaction with a target bound type III CRISPR-Cas effector. The enzymatic activities of ancillary proteins are guided towards different forms of nucleic acid moieties and perhaps protein: Csm6, Csx1 and TTHB144 act on single-stranded RNA (ssRNA), Can1 acts on double-stranded DNA (dsDNA) and perhaps ssRNA, Card1 acts on ssRNA and dsDNA, Can2 acts on ssRNA, single-stranded DNA (ssDNA) and perhaps dsDNA, NucC acts on dsDNA, and CalpL cleaves protein. (**C**) Degradation of cOA into linear di-adenylate species (A_2) through ring nuclease activity of host proteins or viral proteins (such as AcrIII-1) can defuse cOA molecules and as such dampen or shut-off the ancillary protein immune response in case of infection alleviation.

When there is both sufficient base pairing in the spacer portion of the crRNA (particularly in the first nucleotides at the 5′ side, also known as Cas10 activating region (CAR) [68] and sufficient mismatches in the PFS region, the bound RNA is most likely of invader origin (**Fig. 2A**). Conformational changes are relayed in the complex [72], unleashing nuclease and cyclase/polymerase catalytic activities within Cas10. The Cas10 nuclease activity is exerted by an histidine–aspartate (HD) domain and can degrade single-stranded DNA (ssDNA) [31, 39, 70, 78]. Whereas the exact role of Cas10 ssDNA activity has to be established, several hypotheses are plausible: ssDNA cleavage might promote immunity via degradation of ssDNA at or near the viral transcription bubble [78, 79] (**Fig. 2A**); the R-loops that arise during viral transcription elongation [80]; or single-stranded replication intermediates of viruses and plasmids [80]. It also has been proposed that Cas10 nuclease activity promotes host mutagenesis through the induction of host chromosomal lesions [81]. Whereas the DNA cleavage can be sufficient for viral protection in case of abundant viral transcription [82], signal amplification is needed when transcription is limited. Here, the Cas10 cyclase/polymerase activity comes into play: two Palm domains facilitate ATP binding [83, 84], whereupon a GGDD motif in one of the Palm domains catalyzes the conversion of bound ATP molecules into cyclic oligoadenylate (cOA) second messengers. This is achieved by 3′–5′ joining of adenosine monophosphates (AMPs) to form rings ranging between two (cyclic di-adenyate (cA_2)) to six (cyclic hexa-adenylate (cA_6)) AMP units [83, 85–87] (**Fig. 2A**). The dispersed cOA molecules in turn bind to CRISPR-associated Rossmann Fold (CARF) domains of proteins often found to be encoded in or near the CRISPR-Cas type III loci [43, 83, 85, 86, 88]. CARF is a nucleotide-binding domain that allosterically activates an attached effector domain upon binding the cOA nucleotide ligand, releasing its immune enzymatic activity [43, 89].

Genomic neighborhood analysis of type III CRISPR-Cas loci have revealed numerous CARF family proteins, amongst others nucleases, transcription factors, proteases, deaminases, nitrilases and membrane-associated proteins [43, 88], suggesting a plethora of potential intriguing cOA-based defense pathways. Various nuclease-type CARF proteins have been characterized: Csm6, Csx1 and TTHB144 [83, 85, 89–93], Can1 [94], and Card1 and Can2 [95, 96] (**Fig. 2B**). Upon cA_4 or cA_6 binding to a Csm6 or Csx1 homodimer, conformational changes bring together the higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domains to constitute a promiscuous RNase pocket that degrades both viral and host transcripts. Similarly, TTHB144 also displayed HEPN-mediated ribonuclease activity after activation by cA_4 [97]. Viral transcript levels during some infections drastically outnumber those of the host [98–100], so in these cases the indiscriminate RNA degradation will primarily affect the virus. However, it is generally believed that the desired outcome of viral suppression is through induction of cellular dormancy due to depleted host transcripts levels. Can1, which contains two CARF-domains, a nuclease domain and a nuclease-like domain, was shown to nick supercoiled DNA upon cA_4 activation. This is thought to destabilize replication forks as they are stalled at the nicked sites, which could in turn interfere with viral replication. Card1 and Can2 were found to possess both ssRNase and ssDNase activities upon activation by cA_4, presumably achieving the protective function through two separate, but perhaps synergistic, processes: cellular dormancy due to transcript depletion and direct invading genome destruction

through cleavage of ssDNA intermediates in DNA replication. As a single type III effector complex is capable of synthesizing multiple cOA species [101, 102], encoding a multi-pronged ancillary nuclease or, alternatively, multiple ancillary proteins that bind different cOAs [94], might enable further fine-tuning an appropriate immune response. Another nuclease that is activated by cOA, albeit in a CARF-independent manner, is NucC [102–104] (**Fig. 2B**). NucC was initially studied as part of the cyclic oligonucleotide-based anti-phage signaling system (CBASS), but some homologs of NucC are associated with type III loci and were demonstrated to be activated by Cas10-generated cA_3. The activation of NucC involves the assembly of two NucC trimers into a homohexamer upon cA_3 binding, leading to complete destruction of the bacterial chromosome to cause cell death.

In some bacteria, proteins in association with type III systems possess a domain called SAVED (SMODS-associated and fused to various effectors domains) instead of CARF [105]. Structural insights suggest that the SAVED domain evolved through fusion of two CARF proteins, broadening the range of cyclic nucleotide molecules that can be detected to activate a fused effector domain [106]. An intriguing example of this is the CRISPR-associated Lon (CalpL) [107] protease, which is encoded in close proximity to a CRISPR-Cas type III-B system. CalpL has an integrated SAVED domain and forms a strong complex with co-occurring CalpT and CalpS (**Fig. 2B**). The protease is activated upon binding cA_4 and cleaves off CalpS, which is thought to bind RNAP to enable a transcriptional response. The fact that also putative families of genes without nucleotide-sensing domains were identified to associate with type III modules, including nucleases, proteases, peptidases and ATPases, suggest that type III signaling goes beyond the usage of cOA second messengers [43, 88]. As an alternative, physical association of the accessory protein with the type III effector complex might function as a means of regulation. Elucidating the mechanism of action of other type III CRISPR-Cas clusters lacking nucleotide-sensing domains is likely to expand our knowledge of type III intermolecular communication routes beyond cOA dependency.

Although most of the type III-associated immune proteins are uncharacterized, domain function inference and experimental data suggest that ensuring an anti-viral response through non-specific, debilitating action on both host and virus is a common theme. These are obviously damaging measures for the host when insufficiently controlled. Therefore, tight regulation of ancillary protein activation has evolved: type III CRISPR-Cas seems to be an intricately regulated signaling system, displaying properties such as signal amplification, self-regulation and tuning of signaling molecule concentrations. Recognition of just a single RNA molecule can potentially generate about a thousand cOA molecules, which in turn enable enzymatic activation of many immune proteins [38, 108]. The cOA concentration increases proportionally to the viral load, for example in case of multiple co-occurring infections or internally replicating viruses, as more viruses generally means more transcripts. This ensures scaling of the immune response to the severity of infection [38, 108]. To limit the detrimental effects to the host after invader clearance, the cOA concentration can be tuned down by shutting off new production as well as destroying excess cOA. New cOA production is stopped through cleavage of the target RNA, as this returns Cas10 to an inactive state [38]. It has been hypothesized that Cas10 inactivation, rather than the protective effect of direct degrada-

tion of viral transcripts, is the primary role of target RNA cleavage by the effector complex [39]. Removal of excess cOA is achieved by dedicated CARF-containing ring nucleases, which cleave cOA rings into inactive linear di-adenylate species [109–111] (**Fig. 2C**). Furthermore, some of the CARF effectors themselves have been shown to degrade cOA via a CARF domain [93, 97], HEPN domain [101, 112, 113] or fusion to a ring nuclease [114], thereby acting as an intrinsic timer to regulate their own activity [115]. Also, a family of membrane-associated nucleases was shown to possess cOA degradation capacity [116], as well as the virus-encoded ring nuclease AcrIII-1 [108, 109]. The potent AcrIII-1 binds cA_4 at a higher affinity than Csx1 and degrades it at a faster rate than host ring nucleases. This effectively reduces the number of activated Csx1 molecules, leading to suppression of the immune response to safeguard cellular integrity.

The wide pool of functions inherent to the type III systems —RNA-targeting, second messenger signaling and the availability of large repertoire of ancillary proteins— raises the intriguing possibility that their functionality extends beyond immune defense [88]. An example of type III cOA signaling transcending direct anti-viral defense is found in the CARF-containing transcription factor Csa3 from a co-occurring type I-A system, which seems to be involved in transcriptional regulation [117–119]. Certain type III CRISPR-Cas proteins secreted by the bacterial pathogen *Mycobacterium tuberculosis* were also shown to function as virulence factors [120]. Deciphering the full range of type III CRISPR-Cas biology, with its many uncharacterized proteins and functionalities, will provide molecular biologists with experimental challenges for years to come as well as potential for new biotechnological applications.

2.2.2. TYPE VI CRISPR-CAS

In type VI CRISPR-Cas, the functionalities for crRNA processing, invader recognition and immune response are contained in a single effector protein: Cas13 [23, 121–123]. Up until now, six different type VI subtypes have been identified: types VI-A to VI-D, Cas13X and Cas13Y [43, 124]. The Cas13 proteins across subtypes are distantly related, only sharing the presence of two HEPN domains [43]. Cas13 adopts a bilobed structure, with one lobe being responsible for RNA target recognition and the other for the RNA nuclease activity. Initially, Cas13 was demonstrated to be capable of targeting RNA viruses in an experimental setting [123]; newer insights showed that type VI spacer sequences match the genome of DNA viruses, indicating binding of viral transcripts [125, 126]. Cas13 has the remarkable capacity of reaching femtomolar sensitivity in finding a target RNA in a population of non-target RNAs [127]. Upon target RNA loading, a conformational shift in the nuclease lobe accommodates the two HEPN domains to form a stable composite RNase pocket that mediates target RNA cleavage as well as hydrolysis of by-stander RNA [121, 128, 129], leading to inhibition of the invading DNA virus [55, 129–133]. Unloaded Cas13 is inactive, indicating the presence of an auto-inhibited conformation that is released upon target recognition.

Initial binding of target RNA to Cas13 occurs at the 'central seed region', a solvent-exposed part in the center segment of the crRNA [122, 123, 128, 134] (**Fig. 3A**). This region is most sensitive to mismatches, as it initiates RNA duplex formation. The nucleotides in the crRNA at the 5′ side of the seed region are known as the 'HEPN-nuclease switch region' (**Fig. 3A**), because imperfect base-pairing in this portion prevents HEPN-nuclease

activation [128]. Extensive base-pairing between the repeat derived crRNA portion and PFS also blocks the formation of the HEPN catalytic pocket, which is likely to prevent autoimmunity [135, 136]. Bound target RNA is cleaved by Cas13 only when there is sufficient base-pairing in the switch region and non-complementarity between 5′ tag and PFS (**Fig. 3A**). Although genomes of viruses that escaped Cas13 targeting were found to contain deletions of tens to hundreds of bases [55, 133], one would imagine that strict matching requirements in the seed and switch regions enables viruses to also escape by point mutations.

In contrast to other CRISPR-Cas effector proteins, in which the catalytic sites are buried deep inside the protein, the HEPN catalytic site is located at the solvent exposed external surface of Cas13 [137]. This leads to RNA cleavage outside of the target RNA-binding region (**Fig. 3A**), with different homologs of Cas13 displaying varying ribonucleotide cleavage preferences [23, 67]. The fact that the crRNA-bound portion of the target RNA is not cleaved seems to suggest that the RNA:RNA duplex stays intact, perhaps preventing target release. This could in turn mean that Cas13 is not able to sequentially bind new targets and that Cas13 collateral cleavage remains activated upon target RNA cleavage. Pioneering work on Cas13 demonstrated that when Cas13 is guided towards early-expressed transcripts, viral DNA does not accumulate, probably because host and viral transcript depletion early in the lytic cycle prevents genome replication [129]. Extensive host transcript depletion interferes with vital cellular pathways, whereupon cells enter dormancy. This state of hibernation is maintained as long as active virus continue to produce target RNA. Inhibition of target transcription (signifying the defeat of the virus through, for example, viral DNA elimination by co-existing restriction-modification systems [138]) was found to reverse dormancy even after 9 hours, implying that cells stay alive during the process [129]. Direct cleavage of the target RNA seems to be less important for the antiviral response, as pre-activation of Cas13 with non-viral RNA is sufficient to clear a virus infection. This indicates that Cas13, once activated, can provide immunity against co-infecting viruses whose transcripts cannot be directly recognized. Indeed, cross-protection was shown to be a feature of Cas13 targeting, establishing broad and nonspecific immunity [129].

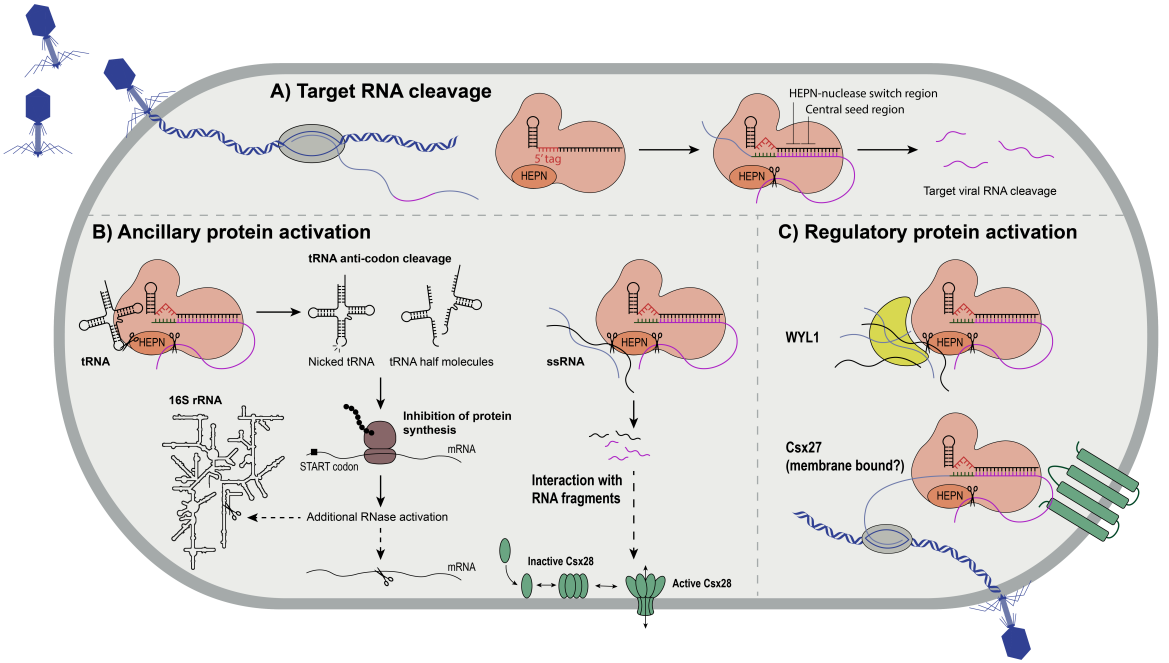


Figure 3. Schematic of type VI CRISPR-Cas immunity. (**A**) Target RNA recognition by Cas13 occurs at the central seed region and leads to the activation of a RNase pocket formed by higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domains when the HEPN-nuclease switch region matches the corresponding target sequence and the 5′ tag (red) mismatches the 3′ anti-tag (green). The activated RNase catalytic site is capable of degrading bound target RNA as well as by-stander RNA. (**B**) Activated Cas13 variants can cleave transfer RNA (tRNA) molecules in the anti-codon loop, which can result in ribosome stalling and subsequent activation of additional RNases to establish a dormant phenotype. Bound Cas13 can activate the ancillary protein Csx28, perhaps by target RNA degradation products or through physical interaction, which form pores in the membrane to instigate membrane depolarization. (**C**) Csx27 and WYL1 are believed to regulate Cas13 activity. Csx27 is thought to anchor Cas13 to the membrane for localized target RNA degradation. Alternatively, physical interaction of Csx27 with Cas13 might downregulate its cleavage activity. WYL1 is believed to upregulate Cas13 activity by confining RNA close to the RNase pocket or by allosterically boosting its cleavage efficiency.

Although the biotechnological potential of Cas13 has gained substantial attention, the study of its biology has received little attention. Recent studies, however, increased the resolution of our understanding of Cas13. The paradigm of strict non-specific RNA-targeting by Cas13 has been challenged as a preliminary study found that Cas13a has a bias towards cleavage of specific RNA molecules, most substantially tRNAs [139] (**Fig. 3B**). Massive tRNA cleavage results in compromised translation and dormant behavior of the cell, limiting the success of the virus in the cell population. Interestingly, in addition to Cas13-mediated tRNA cleavage, certain mRNAs and the 16S rRNA are also cleaved upon Cas13 activation, although not by Cas13 itself. Instead, presumably RNases present in the cell get activated by ribosome stalling as a consequence of tRNA inactivation, leading to the observed additional RNA cleavage (**Fig. 3B**). It is interesting to speculate on the possibility that Cas13 generated tRNA fragments can function as signaling molecules, akin to cOA second messengers in type III, for activation of downstream pathways [139], as broadly observed in eukaryotic systems [140]. Cas13-induced downstream protein activation has recently been shown for the type VI accessory protein Csx28 [126], which forms a membrane pore to enhance anti-viral defense through membrane depolarization [132] (**Fig. 3B**). This indicates that Cas13 has the capacity to also act as a signal relay, besides directly interfering with the viral lifecycle through RNA cleavage. Although the details have to be established, this feature would bring type VI closer to type III in terms of sensor capabilities.

Given the profound cellular consequences of Cas13 activity, the nuclease has to be tightly controlled. Besides low tolerance of mismatches in the central seed region, another layer of Cas13 control is provided by accessory proteins with regulation capacity in some type VI CRISPR loci. The accessory Csx27 is found in type VI-B and functions as an inhibitor, possibly by steric interference of its transmembrane domains with Cas13b, decreasing interference up to 5 orders of magnitude [126, 137]. Another possibility is that Csx27 localizes in the membrane where it keeps Cas13 bound in an inhibited state, to perhaps release active Cas13 for local suppression of transcription during DNA uptake or viral infection [137, 141] (**Fig. 3C**). An additional regulatory protein was identified in type VI-D loci, where WYL1 was shown to interact with Cas13d to stimulate its collateral cleavage capabilities, perhaps through allosteric modulation [125, 142, 143]. Alternatively, because WYL1 possesses affinity for ssRNA, it is hypothesized that WYL1 acts as an RNA sponge that upregulates Cas13d cleavage by confining RNA close to its active pocket (**Fig. 3C**). For both Csx27 and WYL1, the exact mechanistic functioning as well as the biological implications are still unclear. As it is fair to assume that fine-tuned regulation of the toxic Cas13 is important for the cell's viability, future studies may uncover more sophisticated regulatory processes to ensure both cell safety and immune specificity.

2.2.3. RNA-TARGETING CAS9 AND CAS12G

Cas9, the type II CRISPR-Cas effector protein, is well-known for its RNA-guided DNase activity in genome engineering [144], but some Cas9 nucleases are also capable of targeting RNA. Cas9 from *Streptococcus pyogenes* can be programmed to bind and cleave RNA targets in vitro when annealed to a DNA moiety that contains the PAM sequence [145]. Other identified Cas9 variants targeting RNA in *Staphylococcus aureus* [146], *Neisseria*

meningitidis [147] and *Campylobacter jejuni* [146] possess RNase activity even without the addition of PAM oligo's. Although the physiological relevance of Cas9 RNase activity in a natural context is not completely understood, it is speculated to help in clearing invading transcripts during infection [147] and demonstrated to be capable of conferring protection against an RNA virus [146] as well as repressing gene expression in a heterologous host [146]. Accordingly, Cas9 from *Francisella novicida* [148] and *C. jejuni* [149] act as such natural gene regulators, as they were found to target endogenous transcripts to suppress protein expression. The RNA-targeting capacity of Cas9 was exploited to visualize [150] and remove [151] RNAs in human cells, as well as for inhibition of the RNA virus hepatitis C in eukaryotic cells [152]. Also Cas12g, a type V CRISPR-Cas RNA-guided endonuclease, recognizes RNA substrates [153, 154]. Target RNA binding by Cas12g leads to target RNA cleavage as well as collateral RNase and single-stranded DNase activities. No PAM restrictions, small size and high thermal stability make Cas12g an interesting candidate for RNA-specific biotechnological applications [155].

2.2.4. APPLICATIONS OF TYPE III AND VI CRISPR-CAS

RNA-targeting CRISPR-Cas systems are now being exploited for various biotechnological applications [156], including RNA knock-down, RNA editing, RNA imaging, RNA splice modifications, RNA-protein interaction mapping, counter selection in genome engineering and molecular diagnostics in nucleic acid detection. The natural RNA interference functionality of Cas13 can be guided to a desired target by programming it with the complementary crRNA to knockdown transcript levels [23, 123, 157, 158], which allowed for implementation of Cas13 as an antimicrobial agent [159] and a selection tool for virus engineering [130, 133]. Cas13 is used for tagging proteins in the vicinity of specific cellular RNAs to study native protein-RNA interactions in methods called RPL [160] (RNA proximity labelling), CRUIS [161] (CRISPR RNA-unified interaction system) and CAPRID [162] (CRISPR-CasRx-based RNA-targeting and proximity labeling). Genetic fusions of a dead variant of Cas13 (dCas13) yielded various tools: dCas13-GFP is used for fluorescence-based localization of target RNA in a cell [163], dCas13 fused to mRNA splicing repressors can skew protein isoform ratios [157], dCas13 fused to a deaminase domain is used to edit a specific adenosine to an inosine (RNA Editing for Programmable A to I Replacement; REPAIR) [164, 165] or a cytosine to uracil (RNA Editing for Specific C-to-U Exchange; RESCUE) [165, 166], and dCas13 fused to a demethylase was shown to successfully demethylate targeted mRNA in cells [167]. The collateral cleavage capability of Cas13 is exploited for nucleic acid detection in SHERLOCK [67, 127] (Specific High-sensitivity Enzymatic Reporter unLOCKing), SHINE [168] (Streamlined Highlighting of Infections to Navigate Epidemics), SATORI [169] (CRISPR-based amplification-free digital RNA detection), CREST [170] (Cas13-based, rugged, equitable, scalable testing), CARMEN [171] (Combinatorial Arrayed Reactions for Multiplexed Evaluation of Nucleic acids) and droplet microfluidics [172], which are all based on quenched reporter RNAs that become fluorescent upon degradation by target bound Cas13. The same principle been exploited in APC-Cas (Allosteric Probe-initiated Catalysis and CRISPR-Cas13a) for the detection of bacterial pathogens [173]. Also a Cas13 platform using solution turbidity caused by liquid-liquid phase separation as a readout has been employed for RNA detection [174]. Various variations on the theme have been developed for the detection

of microRNA [175–178]. The Cas13-based diagnostics tools vary in terms of sensitivity and specificity and the utility during the SARS-CoV-2 outbreak [179–181].

Also type III systems are engineered for nucleic acid detection: cOA-activated RNases for target-induced cleavage of a reporter are deployed in SCOPE [68] (Screening using CRISPR Oligoadenylate Perceptive Effectors) and MORIARTY [182] (Multipronged, One-pot, target RNA-Induced, Augmentable, Rapid, Test sYstem), and dedicated SARS-CoV-2 detection methods that additionally make use of pyrophosphates and protons generated by Cas10 upon target detection [183]. Moreover, the Can2 [184] and NucC [102] nucleases have been demonstrated to be capable of cOA-activated reporter cleavage for RNA diagnostics. To enhance sensitivity of RNA detection, a pull-down method for capturing and concentrating target RNA from heterogeneous samples was developed using a type III effector complex [184]. The FIND-IT (Fast Integrated Nuclease Detection In Tandem) approach combines Cas13 target recognition with Csm6 RNase activity to enable RNA detection [185].

2.3. DORMANCY AS AN IMMUNE STRATEGY

Bacteria can go into a physiological state of low metabolism known as dormancy, which facilitates survival in unfavorable conditions. Dormancy enables the utilization of energy and resources for processes that attempt to sustain, repair and prevent, rather than those that aim at growth and propagation [186]. When conditions become more favorable, the cell can re-emerge and continue regular metabolism. In scenarios of sustained stress, the cell accumulates so much damage that cell death is inevitable, which can thus be defined as the final stage of cell dormancy. Circumstances that may induce dormancy phenotypes include scarcity of nutrients, extremes of temperature, damage to vital components, oxidative stress, the presence of toxic compounds or parasitic invaders [187].

Many immune systems, including type III and VI CRISPR-Cas, use dormancy as a strategy to halt viral invasion [188]. The broad activities (for example, DNase, RNase, membrane depolarization or proteolysis) responsible for dormancy induction in type III and VI CRISPR-Cas are known or predicted, but the exact cellular pathways involved in generating the dormancy phenotype are often obscure (**Table 2**). Especially the downstream effects of global RNA degradation are to be investigated in detail. Cas13 exhibits collateral RNase activity against primarily tRNAs, resulting in dormancy through protein translation inhibition [139]; for collateral RNase activity in type III, it is not well known whether there is a bias towards degradation of certain transcripts (for example, mRNA, rRNA, tRNA or RNA toxin components of toxin–antitoxin systems). Although the terms ‘dormancy’ and ‘cell death’ are often used interchangeably, reports on cell death caused by the action of RNA-targeting CRISPR-Cas systems are limited. Cas13 was shown to cause cell death when targeted towards an ampicillin resistance gene in cells under ampicillin conditions [163], but there is no data on Cas13 causing cell death during an actual infection. For type III secondary effectors, only NucC has shown a clear cell death phenotype [103, 104].

Effector protein	CRISPR-Cas type	Target	Cellular pathway	Phenotype
NucC	Type III	dsDNA	Host chromosome degradation	Cell death
Card1 and Can2	Type III	ssRNA, ssDNA, dsDNA	Unknown	Dormancy
Can1	Type III	dsDNA, ssRNA	Replication fork destabilization*	Dormancy
Csm6, Csx1 and TTHB144	Type III	RNA	Unknown	Dormancy
Cas13	Type VI	RNA	Ribosome inhibition	Dormancy
Csx28	Type VI	Cell membrane	Perturbed membrane integrity	Dormancy or cell death*

Table 2. Experimentally described type III and type VI CRISPR-Cas proteins able or expected to cause dormancy phenotypes. dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; ssRNA, single-stranded RNA. *Posed as a hypothesis in literature.

An intuitive rationale for why halted cellular activity facilitates immunity is the generation of an inhospitable environment for the infecting virus, with limited access to essential host processes. This results in at least two discernible and probably synergistic scenarios to prevent viral success (**Fig. 4**): interference with the viral lifecycle buys time for the already present DNA-targeting enzymes to destroy the foreign genomes [189, 190]; and viruses are trapped in the cytosol of the dormant cell, preventing it from completing its lifecycle and spreading to neighboring cells, implying a kin-selection strategy [189, 190]. The exact processes required for exiting dormancy after viral clearance are not elucidated and remain an interesting topic of research, but restoration of the damage and replenishment of depleted cellular components are expected to occur. Furthermore, it has been hypothesized that for Cas13, certain RNAs required for restarting cellular processes are protected from degradation through dedicated proteins or tertiary conformations in the RNA [191].

2.3.1. BUYING TIME

The events that happen after viral infection, including viral genome replication and protein synthesis, occur rapidly [192, 193]. Moreover, cells can be infected with multiple viruses at the same time [194] and the DNA target sequence can be in a genomic region that is late-expressed in the infection cycle [32]. This could quickly lead to accumulation of viral DNA and a cellular state in which the viral genome copy number outruns the DNA-acting systems [195]. Induction of dormancy slows down viral replication processes, effectively buying time for DNA-targeting systems that were initially too slow to halt the viral genomes (**Fig. 4**). Additionally, dormant cells can potentially use the extra time to acquire new spacers from the invading genomes, which can then be immediately used against it during a CRISPR-Cas defense. Substantiating this idea, growth inhibited bacteria were shown to have increased spacer acquisition [196]. Kinetic studies on the relations between viral replication and the activity of co-occurring DNA- and RNA-acting systems could shed experimental light on the buying time concept, which remains mostly hypothetical to this day.

2.3.2. KIN PROTECTION

At first glance, unicellular programmed cell death seems paradoxical. Whereas immune systems often confer benefits to the cell carrying them, systems evoking death are clearly not advantageous to the enacting individual. Instead, impeding successful viral development through the abortion of cellular processes releases fewer progeny, effectively lowering the viral epidemic (**Fig. 4**). This decreases the chance for the infection to spread to neighboring cells, which are often closely related kin, and thus likely also susceptible to the infecting virus [197]. Additionally, the sacrificed cell may leak valuable cellular resources into the population [198], perhaps even cOA signaling molecules to prime defense in neighboring cells. So whereas the individual cell does not benefit from suicide, protection of kin makes the maintenance of suicide genes and pathways evolutionary advantageous [199]. Moreover, co-infecting viruses that are resistant to other forms of defense are also taken down in the process, further highlighting why acting on the host and virus simultaneously instead of the invader only can be beneficial [136, 191].

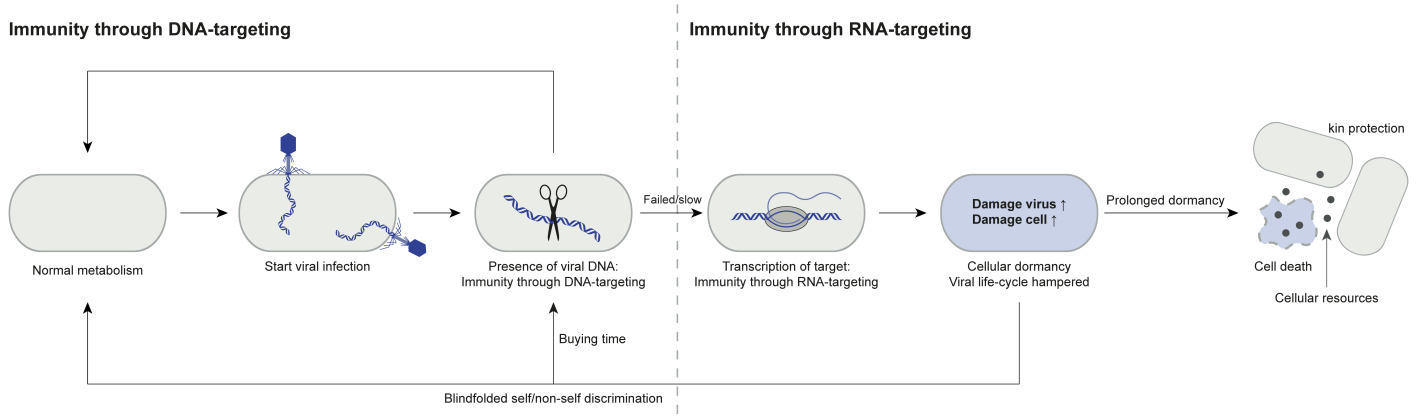


Figure 4. Model of the dormancy strategy used by RNA-targeting type III and VI CRISPR-Cas systems. Infection by a virus is signified by the presence of viral DNA inside the cell. DNA-targeting systems provide the first line of immunity by immediately attacking the invader genomes, usually through genomic DNA cleavage. When the anti-DNA response is unsuccessful (for example, viruses circumventing DNA-targeting mechanisms or viral genome replication outrunning genome cleavage), the viral lifecycle progresses to transcription. Target transcripts are recognized by RNA-targeting type III and VI CRISPR-Cas effector proteins, typically leading to an immune response that involves global damage to both host and virus. The resulting cellular dormancy and hampered progression of the viral lifecycle effectively buys time for initially too slow DNA-targeting immune systems to inactivate remaining viral genomes. The cell can exit from dormancy through restoration of the inflicted self-damage. Alternatively, or synergistically, as the fast viral processes are more severely affected compared to the slow processes of a dormant cell, self is discriminated from non-self through the capacity to recover from the induced damage (blindfolded self/non-self discrimination; see **Figure 5**). In case of sustained dormancy due to prolonged infection, the cell accumulates damage to reach a point of no return: cell death. The invader, requiring a living host for its propagation, is taken down in the process, making cellular suicide an altruistic act to protect kin. Additionally, cellular components leak in the environment to provide nutritional aid to neighboring cells.

2.3.3. BLINDFOLDED SELF/NON-SELF DISCRIMINATION

The use of nucleic acids for information storage is a universal feature of life. Therefore, to prevent accidental targeting of the host genome by DNA-targeting systems, distinguishing features have evolved to discern self and non-self DNA. Two principles for discriminating self and non-self (from here self/non-self discrimination) are well-described: mask self and damage non-self; and recognize non-self and damage non-self. The first principle is used in type II restriction–modification systems, whereby the own genome is masked by methylation sites [200]. This prevents accessibility of the restriction enzymes, such that only the unmethylated viral genomes are detected and restricted. The second principle is found in DNA-targeting CRISPR-Cas systems, which only act on genetic elements that carry a PAM next to the target [201]. The combination of a PAM and target sequence in the own genome is scarce, preventing the recognition of self.

The success of RNA-targeting CRISPR-Cas immunity depends on inflicting damage to both host and virus, without strictly discriminating self from non-self. Due to intrinsic differences between the lifestyles of host and virus, such as the capacity to sustain when important resources are not readily available, the host is often able to survive the global damage, whereas the virus is not. So, during RNA-targeting CRISPR-Cas immunity, another form of self/non-self discrimination seems to apply: damage both self and non-self, outlive non-self. As opposed to pointing directly at non-self prior to the immune response, this principle acts without actively seeking out who is who and effectively differentiates only in hindsight which was self (that is, the biological unit that survived) and which was non-self (that is, the biological unit that perished). We therefore term it 'blindfolded self/non-self discrimination' (**Fig. 5A**). For example, the activity of various RNA-targeting CRISPR-Cas proteins (for example, Cas13, Csm6 and Card1) is guided towards RNA non-specifically. The virus is programmed for fast replication, so depleted RNA levels (for example, tRNA, rRNA, mRNA) and resulting protein scarcity (for example, transcription and translation machinery, structural proteins) can lead to logistical problems in the viral lifecycle. By contrast, the host enters a dormant state, during which slowed metabolic processes do not require fine-tuned RNA and protein concentrations. When the viral processes have been disorganized to the point of virus neutralisation, the host can replenish RNA and protein levels to resume normal life. Another example is presented by the effector protein Can1, which nicks supercoiled DNA that could result in the collapse of replication forks in both the virus and the host [94]. This is detrimental to the virus, where many replication forks are present due to its fast replication cycle. Conversely, the host prevents catastrophic damage by entering a dormant state, ensuring few replication forks. Upon clearance of the virus, the host can repair broken DNA and continue normal metabolism. Thus, presumably autoimmunity — a phenomenon that is generally thought to be avoided — is used for the benefit of the cell during blindfolded self/non-self discrimination.

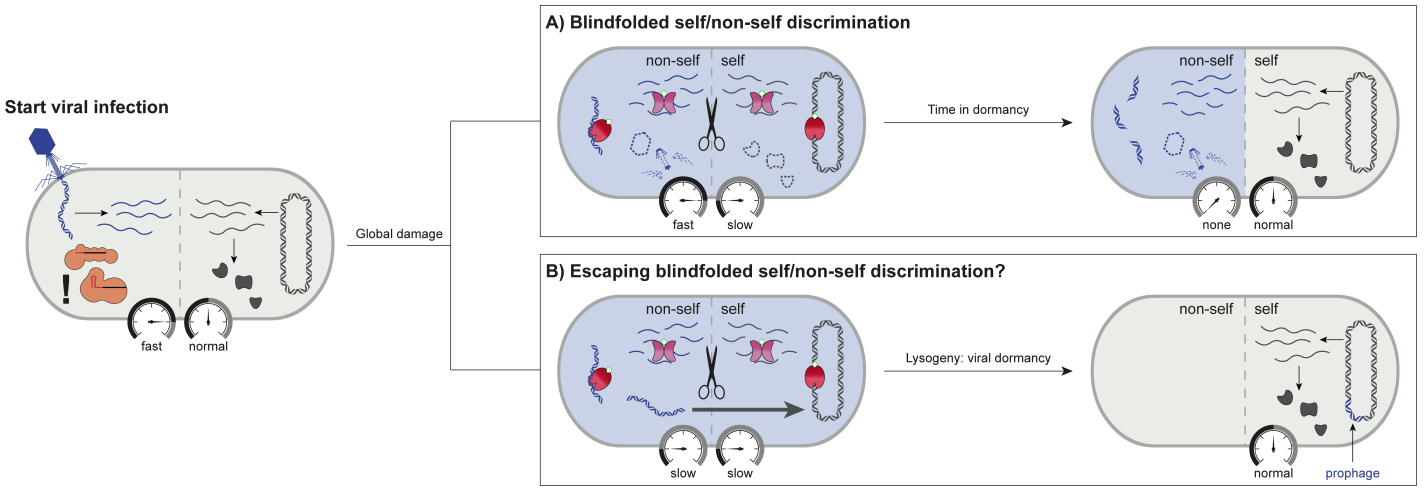


Figure 5. Blindfolded self/non-self discrimination. Upon viral infection, unsuccessful immediate clearance of the virus leads to progression of the viral lifecycle to transcription (rate of metabolism is indicated by the speed-o-meters in the bottom of the cell). Target transcripts are recognized by the type III and VI CRISPR-Cas effector proteins that typically initiate a global immune response. (**A**) Target recognition typically includes damage to both host and virus, as exemplified by the activity of RNase Csm6 (pink) and ssDNA nickase Can1 (red), affecting the processes of replication, transcription and translation. To sustain this stress, the host enters a state of dormancy; by contrast, the virus will pursue their rapid lifecycle and eventually collapse due to the lack of cellular means. In this way, self is discriminated from non-self as if wearing a blindfold; that is, the induced indiscriminate damage can be overcome by self, but not by non-self. (**B**) A way of viral escape from blindfolded self/non-self discrimination could be through mimicking the host in terms of metabolism. Instead of following a fast lifecycle, temperate viruses have the capability of integrating into the host genome as a prophage during the process of lysogeny. This presents a form of viral dormancy, slowing the viral processes down such that they cannot be 'detected' by blindfolded self/non-self discrimination.

It is interesting to speculate that the integration of the viral genome into the host genome, called lysogeny, presents a counter to blindfolded self/non-self discrimination. Instead of rapid replication, the virus goes 'dormant' upon host integration (**Fig. 5B**). By doing so, it effectively mimics the cell in terms of its slow lifestyle, circumventing 'detection' by blindfolded self/non-self discrimination. If viral lysogeny indeed functions as such counter strategy, one might expect lysogenic viruses to be overrepresented in bacteria that use dormancy as an immune response.

Investigating whether blindfolded self/non-self discrimination is also occurring in other forms of prokaryotic immunity adds an intriguing dimension. For example, the burst of newly identified systems in the last years have uncovered cell-wide depletion of the nucleotide pool as a successful antiviral strategy [202, 203]. Viral infection is the trigger for activation of enzymes that hydrolyze deoxynucleotides, thereby reducing the levels of nucleotides in the cell to prevent genome replication. Not only the virus, but also the host is expected to be affected by nucleotide-depletion, although to a lesser extent due to its slower replication rate. The innate difference between host and virus to deal with global damage, in this case nucleotide depletion, allows for blindfolded self/non-self discrimination. So although the principle of blindfolded self/non-self discrimination is discovered in the context of RNA-targeting CRISPR-Cas systems, it may be broader applicable to include other forms of immunity that function through induction of global damage, host dormancy and resurrection upon the disappearance of the invader.

Other global damage strategies are reported, such as depletion of NAD+ and ATP, but are claimed to lead to cell death instead of dormancy [204]. When the host cell dies instead of entering a dormant state upon infection, blindfolded self/non-self discrimination cannot apply. However, it is questionable whether the observation of death phenotypes in laboratory experiments are due to the physiological effect of the immune system or whether there are other contributing factors, such as the artificial over-expression of death effector or an overlooked influence of the invader [205]. In nature, there might be a phase of increasingly worsening dormancy following the global damage response, only leading to cell death when the infection cannot be cleared in time. In the time-window between invader recognition and cell death, blindfolded self/non-self discrimination can apply to rescue the host, even in immune systems that are under laboratory circumstances diagnosed solely as inducers of cell death.

2.4. CONCLUSIONS AND OUTLOOK

CRISPR-Cas immune strategies have far exceeded the 'simple' paradigm of cleaving invader nucleic acids. Type III and type VI CRISPR-Cas systems provide immunity through sensing invader transcripts, whereupon a plethora of broad and often rigorous responses are initiated. Mechanisms include collateral RNA degradation, own chromosome destruction, replication fork collapse, tRNA inactivation and membrane depolarization. But although the general workings of type III and VI CRISPR-Cas are understood, many of the details remain obscure and various open questions are to be answered. For example, under what circumstances does type III and VI CRISPR-Cas immunity lead to cell death? Which cellular pathways generate the dormancy phenotype during type III and VI CRISPR-Cas immunity? To which extent is cellular dormancy induced by viral infection reversible, and which processes are responsible? Do cOA signaling molecules

leak into the environment upon suicide induced by the type III CRISPR-Cas immune response, and are they capable of priming defense in neighboring cells? Are there other systems in the cell that are activated by the cOA produced during the type III CRISPR-Cas immune response? How is the synergy between the DNA- and RNA-targeting immune systems kinetically orchestrated? Can viral lysogeny be considered a counter to 'blindfolded self/non-self discrimination'? How prevalent is Cas13-based downstream activation of ancillary proteins? How, and to which extent, is Cas13 activity regulated by ancillary proteins? How does signal relay occur in type VI CRISPR-Cas? What is the effect of viral RNA modifications (for example, modified nucleosides or secondary structures) on RNA-targeting CRISPR-Cas effectors? Have RNA-targeting CRISPR-Cas systems evolved to protect from RNA viruses? Besides adaptive RNA-targeting immune systems, in what ways do bacteria and archaea use innate immune systems that act on RNA? Such unanswered questions in the field may guide future research priorities. As bacteria and archaea have a long history with virus outbreaks, numerous surprises in RNA-targeting CRISPR-Cas immunity undoubtedly await discovery.

3

THE GRAMP CRISPR-CAS EFFECTOR IS AN RNA ENDONUCLEASE

"He looked at his Soul with a Telescope. What seemed all irregular, he saw and showed to be beautiful Constellations; and he added to the Consciousness hidden worlds within worlds."

Samuel T. Coleridge

Type III CRISPR-Cas immunity is widespread in prokaryotes and is generally mediated by multi-subunit 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", called *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.

A modified version of this chapter has been published as *van Beljouw et al., Science (2021)* [206].

3.1. MAIN TEXT

Facing constant predation by mobile genetic elements (MGEs), prokaryotes evolved multiple defense systems to protect themselves [207]. Among those are the CRISPR-Cas systems, which provide adaptive immunity: invaders are recognized and inactivated by effector ribonucleoprotein complexes, during which genetic memory of the invader is generated and stored as spacers in CRISPR arrays [16]. 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) [24]. 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 the activation of ancillary nucleases. CRISPR-Cas type III encompass both functionalities, making them among the most sophisticated CRISPR-Cas effector proteins known to date [33].

Recently, a new CRISPR-Cas subtype with resemblance to type III systems was bioinformatically predicted and classified as type III-E [43]. Although type III-E effectors are related to other type III effectors (**Fig. 1A**), they are notably different in terms of protein architecture. Whereas type III effectors are typified by a multi-subunit composition, the type III-E effector seems to have various Repeat Associated Mysterious Protein (RAMP) domains fused together; hence the nickname g(iant)RAMP [43]. The unusual size of approximately 1300 to 1900 amino acids makes gRAMP the largest single-unit effector found to date (**Fig. 1B**). 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 [208]. Curiously, the gRAMP gene clusters lack ancillary nuclease genes but often co-occur with a gene encoding a TPR-CHAT protein (**fig. S1**) [43], a caspase-like peptidase found to be involved in regulated bacterial cell death [209].

We found that spacers embedded in type III-E CRISPR arrays have targets in MGEs (**Fig. 1C; table S1**) with a bias towards targeting the coding strand of open reading frames (**Fig. 1D**), indicating that gRAMP activity likely involves interaction with invader mRNA. In order 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) (see **Chapter 1, Fig. 3**) [211]. 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; table S2**). We purified *Sb*-gRAMP to apparent homogeneity via three consecutive chromatography steps and observed distinct 260 nm absorption during size exclusion chromatography (SEC), indicative of co-purifying nucleic acids (**fig. S2**). Multi-angle light scattering (MALS) analysis indicated a homogenous particle of 242.5 ± 2.4 kDa (**Fig. 2B**), consistent with the expected size for a *Sb*-gRAMP monomer bound to an ssRNA species in the range of 45-60 nucleotides (nt) (**Supplementary Text; 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 co-expressing *Sb*-gRAMP with a different repeat-spacer pair (**fig. S3A-D**), demonstrating that *Sb*-gRAMP can be loaded with a crRNA of choice.

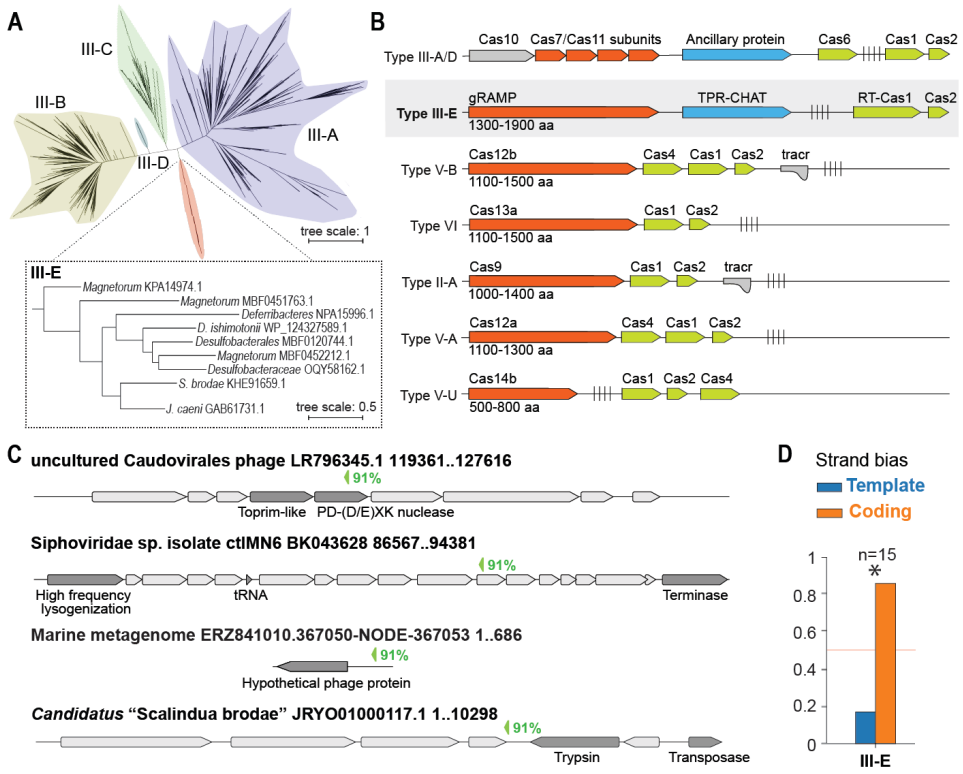


Figure 1. Type III-E CRISPR-Cas composition and characteristics. (**A**) Phylogenetic tree of type III subunit proteins (Csm3 and Cmr4) compiled with type III-E gRAMPs. The tree scale bar represents the number of substitutions per site. (**B**) Illustrations of typical organization of CRISPR-Cas types with sizes of effector proteins indicated in the number of amino acids [210] (**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 ($p = 0.011$).

The *Candidatus* "S. 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 PCR amplified the native CRISPR array from *Candidatus* "S. brodae" genomic DNA and co-expressed it with *Sb*-gRAMP. Following RNA extraction (**fig. S4**) and RNAseq (**Fig. 2E; fig. S5**), we determined that this RNA contains an unusually long 5′-handle of predominantly 27-28 nt, including the conserved last 14 nt of the repeat (**Fig. 2F; fig. S6**). Although a dominant spacer portion of 20 nt was observed, the vast majority of mature crRNAs were truncated within a window of 16-25 nt (**Fig. 2F; fig. S3E; fig. 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 [41]. 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 process-

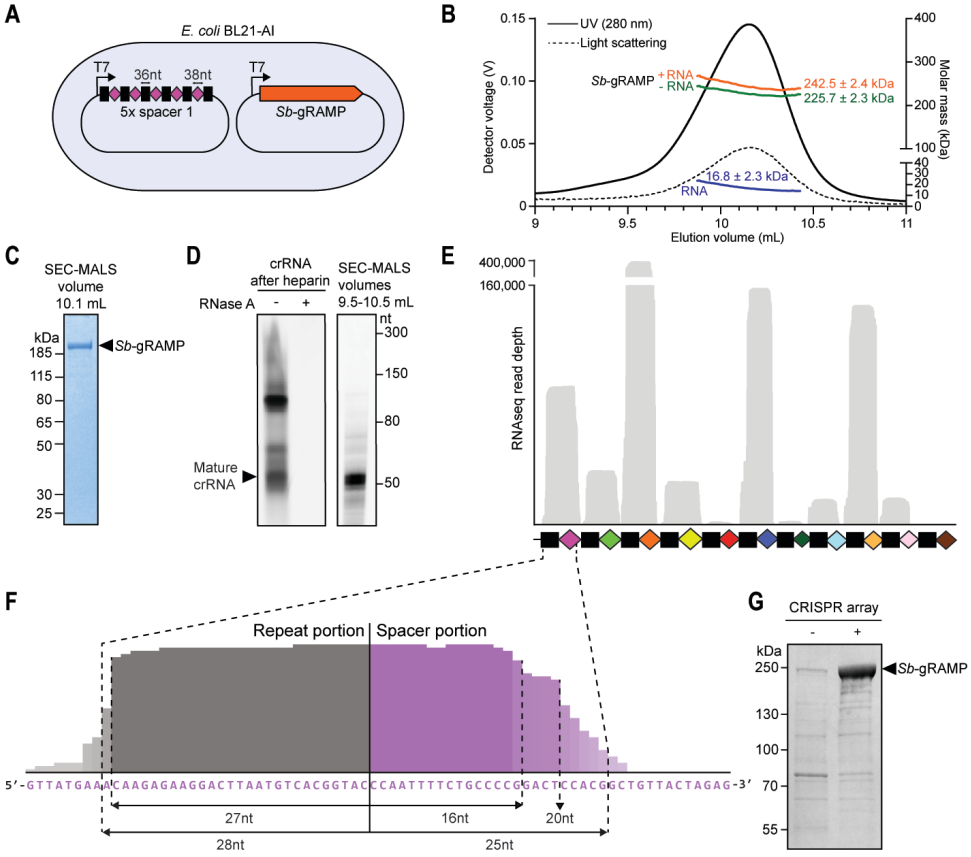


Figure 2. *Sb*-gRAMP-crRNA is a ribonucleoprotein encoded by a single open reading frame. (**A**) Schematic of the heterologous expression system used for *Sb*-gRAMP-crRNA1 overexpression. (**B**) SEC-MALS peaks corresponding to light scattering and UV_{280} absorbance profiles of purified *Sb*-gRAMP-crRNA1. The molar masses are derived from the peak half-height region (9.9-10.4 mL). (**C**) SDS-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. (**F**) RNAseq 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.

ing its own pre-crRNA, similar to Cas12 and Cas13 [23, 212]. The larger RNA molecules bound by *Sb*-gRAMP might represent pre-crRNA processing intermediates. Notably, 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. Hence, 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 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 possess structural homology to Cas7 (Csm3/Cmr4) subunits, one with a large insertion [43]. As Csm3/Cmr4 have been demonstrated to be responsible for target RNA cleavage in other type III systems [213, 214], we set out to investigate whether *Sb*-gRAMP also possesses endoribonuclease activity. We incubated purified *Sb*-gRAMP-crRNA with fluorescently labelled ssRNA substrates complementary to crRNA of spacer 1 (**Fig. 3A**) or spacer 3 (**fig. S7A**) and observed the appearance of two well-defined RNA cleavage products (**Fig. 3B; fig. S7B**), with the smaller product accumulating over time (**fig. S8**). Most target RNA remains uncleaved even at two-fold 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 are located after nucleotides 3 (site 1) and 9 (site 2) relative to the crRNA, thus 6 nt apart (**Fig. 3B; fig. S7C; fig. S10**). When we tested target RNA substrates with the label on the 3′-end instead of 5′-end, two smaller cleavage products accumulated with a 6 nt interval (**Fig. 3B; fig. S11**). This nucleotide spacing between cleavage sites is reminiscent of the 6 nt periodicity found for other type III effector complexes [27, 73], suggesting that the *Sb*-gRAMP retained the ancestral architectural positioning after domain fusion and insertion events. Also analogous to other type III effectors [156], 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 non-complementary 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 type III CRISPR-Cas Csm3/Cmr4 domains harbor acidic residues essential for metal-ion coordination and subsequent ribonuclease catalysis [27, 37, 73]. 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 possesses 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 structure [213] (**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 application where 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. Currently described type III effectors cleave RNA once per Csm3/Cmr4 subunit; *Sb*-gRAMP, composed of four Csm3-like domains, only cleaves twice under our experimental conditions (**Fig. 3G**). This suggests that the other ribonu-

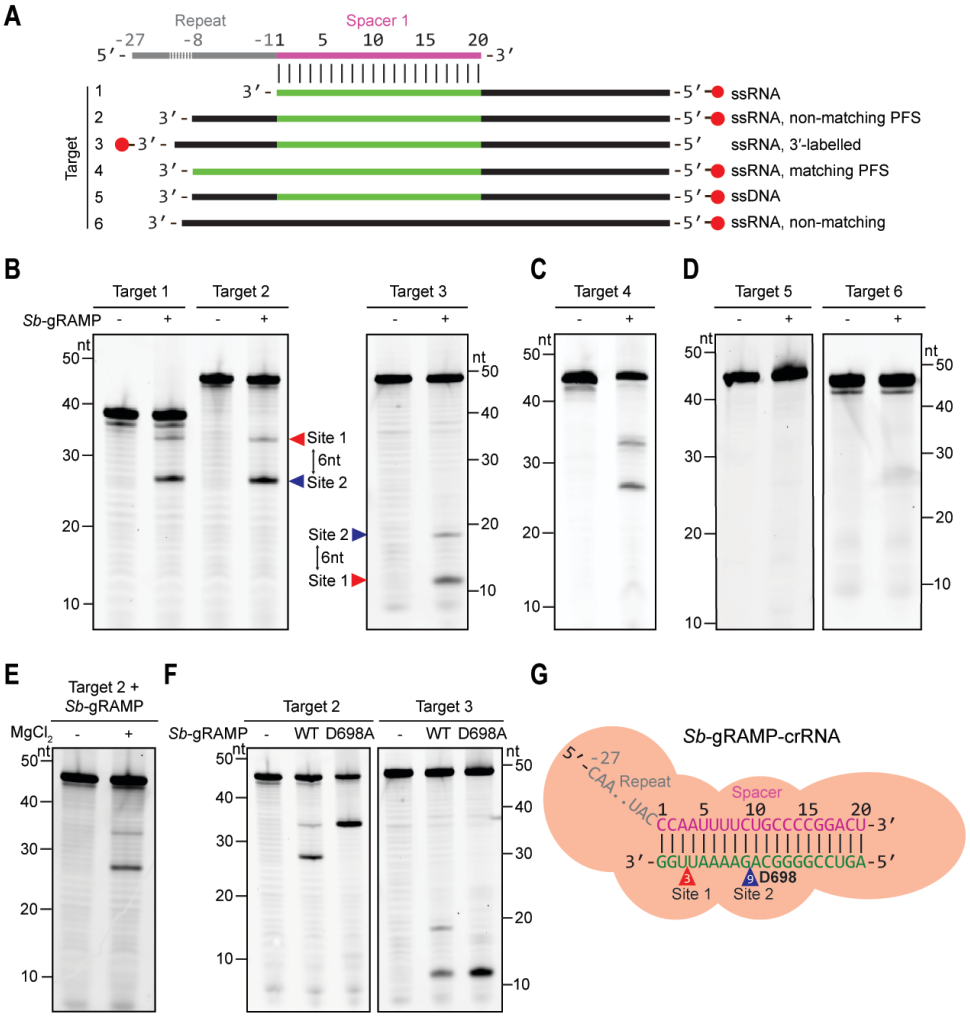


Figure 3. *Sb*-gRAMP-crRNA drives dual ssRNA cleavage with six nt spacing. (**A**) Outline of the Cy5-labelled nucleic acid targets tested for *Sb*-gRAMP-crRNA1 activity. (**B-F**) Denaturing urea PAGE gels of *Sb*-gRAMP-crRNA1 incubated with (**B**) complementary ssRNA (cleavage products indicated with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

clease domains were either rendered dysfunctional over time or evolved into other functionalities. The type III-E gRAMP effector has characteristics of both Class 1 (e.g. type III domains, six nucleotide cleavage spacing, cleavage of RNA independent of matching PFS) and Class 2 (e.g. single protein composition), blurring the boundaries of the traditional CRISPR-Cas classification.

3.2. MATERIALS AND METHODS

PLASMID CLONING

Plasmids and primers used in this study are described in [206] were used for PCR amplification using the Q5 high-fidelity Polymerase (New England Biolabs). Ordered DNA sequences are listed in [206]. All plasmids were verified by Sanger-sequencing (Macrogen Europe, Amsterdam, The Netherlands). Cloning was performed using NEBuilder HiFi DNA Assembly (New England Biolabs) unless stated otherwise. Bacterial transformations for cloning were performed using NEB® 5-alpha Competent E. coli (New England Biolabs) and carried out by electroporation using ECM630 Electrocell Manipulator (2.5 kV, 200 Ω, 25 μF).

For the expression of *Candidatus* "S. brodae" gRAMP (Sb-gRAMP), a two-plasmid expression system was used: a plasmid containing the gRAMP protein (pGRAMP) and a plasmid containing the CRISPR RNA (pCRISPR-1, pCRISPR-3 or pCRISPR-WT). To construct pGRAMP, a coding sequence corresponding to an *E. coli* codon-optimized gRAMP protein variant containing N-terminal Twin-Strep-Tag II and SUMO-tag was designed, ordered (Life Technologies Europe BV) and cloned downstream of the LacI repressed T7 promoter on the plasmid 13S-S (encoding spectinomycin resistance for selection) (Berkeley MacroLab). pGRAMP variants with mutations were generated by amplifying pGRAMP with primers containing the desired mutations.

To construct pCRISPR-1 and pCRISPR-3, a CRISPR array starting with the native *Candidatus* "S. brodae" leader sequence, followed by six native repeats interspaced by five times the first spacer (pCRISPR-1) or the third spacer (pCRISPR-3) in the native CRISPR array was designed, ordered (Life Technologies Europe BV) and cloned downstream of the LacI repressed T7 promoter on pACYC Duet-1 (encoding chloramphenicol resistance for selection) (Novagen EMD Millipore) by restriction with MfeI-HF (New England Biolabs) and NdeI (New England Biolabs) followed by ligation with T4 DNA ligase (New England Biolabs). To construct pCRISPR-WT, the native type III-E CRISPR array was PCR amplified from *Candidatus* "S. brodae" genomic DNA and cloned downstream of the LacI repressed T7 promoter on pACYC Duet-1. In order to get both the Sb-gRAMP protein and the CRISPR-1 array encoded on the same plasmid, the CRISPR array was PCR amplified from pCRISPR-1 and assembled with PCR amplified pGRAMP to yield pGRAMP-CRISPR-1. For the co-expression of an RNA target cognate to crRNA-1, a plasmid (pTarget) encoding the complementary target DNA (5′ - CTCTAGTAACAGCCGTGGAGTCCGGGGCAGAAAATTGG - 3′) downstream of the LacI repressed T7 promoter was generated via "around-the-horn" PCR amplification and self-ligation of p13S-S using phosphorylated primers with flags encoding the target sequence. The region including the LacI repressed T7 promoter, DNA encoding the RNA target cognate to crRNA-1 and downstream T7 terminator was subsequently PCR amplified from pTarget and integrated in pGRAMP-CRISPR-1 to yield pGRAMP-CRISPR-1-target-1.

PROTEIN OVEREXPRESSION FOR PURIFICATION

Electrocompetent cells *E. coli* BL21-AI were transformed with plasmids (for gRAMP-crRNA1, pGRAMP and pCRISPR-1 were used, and for gRAMP-crRNA-WT, pGRAMP and pCRISPR-WT were used; for gRAMP-crRNA3, pGRAMP and pCRISPR-3 were used) and

grown overnight at 37 °C on LB-agar plates containing selection media (50 µg/mL spectinomycin and/or 25 µg/mL chloramphenicol). Colonies were streaked from the plate, resuspended in 100 mL LB containing selection media and grown for ~3 hours before inoculation in 8 L LB medium containing selection media. Cultures were grown at 37 °C and 150 rpm until the cultures reached exponential phase (OD600 0.3-0.5). The grown cultures were incubated on ice for 1 hour and protein expression was induced with a final concentration of 0.2% L-arabinose and 0.5 mM IPTG, followed by overnight incubation at 20 °C and 150 rpm. The overnight grown cultures were harvested by centrifugation at 6000 rpm for 30 minutes. The supernatant was discarded and the pellets were resuspended in PBS (50 mL PBS/initial 1 L culture) and harvested by centrifugation (30 minutes, 3900 rpm 4 °C). The supernatant was discarded, and the pellets stored at -80 °C until further use.

PURIFICATION OF SB-GRAMP-CRRNA AND SB-GRAMP VARIANTS

Cell pellets of each 1 L culture were resuspended in 50 mL of ice-cold lysis buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5). 1 tablet of cOmplete™ EDTA-free Protease Inhibitor Cocktail was added per 50 mL resuspended pellet. The cells were lysed with 3 runs at 1000 bar in a cooled French press and spun down at 16000 rpm at 4 °C for 30 minutes (JA-17 rotor, Avanti). The resulting lysate was filtered through a 0.45 µm syringe filter. To prepare the purification column, a 20 mL gravity column (Bio-Rad) was loaded with 3 mL of a 50% suspension Strep-Tactin ®XT Affinity Resin (IBA Lifesciences GmbH) (corresponding with 1.5 mL column bed volume). The Strep-Tactin ®XT Affinity Resin was washed with 5 column volumes of ice-cold wash buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5). The filtered sample lysate was loaded on the washed Strep-Tactin ®XT Affinity Resin and subsequently washed with 5 column volumes of ice-cold wash buffer. 10 mL of elution buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, 50 mM Biotin, pH 7.5) was used to elute the protein. gRAMP D437A D516A, gRAMP D448A, gRAMP D968A and gRAMP D971A were snap frozen and stored at -80 °C until use in the target RNA cleavage assay.

gRAMP-crRNA, gRAMP D448A D516A, S457A D516A, D698A and D698A D771A were subjected to subsequent purification steps. The pooled fractions were gradually diluted 1.5 times with a wash buffer devoid of NaCl (100mM Tris-HCl, 1mM DTT, 5% glycerol, pH 7.5). The sample was spun down at 13,200 rpm at 4 °C for 10 min whereupon the supernatant was transferred to a new tube. For the Heparin affinity chromatography, a 5 mL HiTrap Heparin HP column (Cytiva) was washed with 2 column volumes of degassed MilliQ at 1 mL/min and equilibrated with degassed low salt buffer (100 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5) at 1 mL/min. The sample was loaded onto the column and washed with 10 column volumes of ice cold degassed low salt buffer. The proteins were eluted by using a NaCl gradient (0-100% in 25 minutes) from low salt buffer to high salt buffer (100 mM Tris-HCl, 1 M NaCl, 1 mM DTT, 5% glycerol, pH 7.5), collecting 1 mL fractions. Pooled fractions were concentrated (Amicon ultra-4 Centrifugal Filter Unit with ultracel-100 membrane) and subjected to size exclusion chromatography using Superdex 200 Increase 10/300 GL (Cytiva) column equilibrated with running buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5) with 0.3 mL/min flow rate using running buffer as mobile phase. For the size exclusion chromatography of

Sb-gRAMP-crRNA prior to SEC-MALS analysis, running buffer without glycerol (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, pH 7.5) was used. Pooled fractions were concentrated, snap frozen in liquid nitrogen and stored at -80 °C until further use.

RNA EXTRACTION

50 µL of purified Sb-gRAMP-crRNA (~1 mg/mL) was thawed and incubated with 3 µL of 800 units/mL proteinase K (New England Biolabs) at 37 °C for 1 hour to digest the protein content, followed by heat inactivation at 95 °C for 5 minutes. To separate the RNA from the digested protein content, acidic phenol (pH 4.5, phenol:chloroform = 5:1, Invitrogen) was added to the sample in a 1:1 ratio, vortexed for 1 minute and centrifuged for 10 minutes at 13200 rpm at room temperature. The aqueous phase was collected and subjected to RNA precipitation (20 µL 3M NaOAc and 500 µL 100% ethanol per 200 µL of sample) for 1 hour at -20 °C. Samples were spun down at 13,200 rpm at 4 °C for 2 hours, washed twice with ice-cold 70% ethanol and spun down at 13,200 rpm at 4 °C for 10 minutes. The pellet was dried in a SpeedVac concentrator (Thermo Fisher Scientific) for 30 minutes at 60 °C and resuspended in RNA grade water. Samples were stored at -80 °C until further use. For RNA digestion, 10 µL of ~50 ng/µL extracted RNA was incubated with 1 µL of 20 mg/mL RNase A (Thermo Fisher Scientific) for 30 minutes at 37 °C. Visualization of RNA was done by mixing 10 µL sample of ~50 ng/µL with 10 µL 2x RNA loading dye (95% formamide, 0.025% SDS and 0.5 mM EDTA) and running the sample on an 8M urea 10% PAGE gel with the Low Range ssRNA Ladder (New England Biolabs). The gel was stained with Sybr Gold and imaged on a Typhoon laser-scanner platform (Cytiva).

RNASEQ OF PURIFIED RNA

RNA sequencing was performed on RNA extracted from Sb-gRAMP-crRNA-WT. 10 µg of phenol-chloroform extracted RNA was incubated with 22 U T4 Polynucleotide Kinase (PNK) (New England Biolabs) and 1x PNK buffer (New England Biolabs) for 4 hours at 37 °C for 2′-3′-dephosphorylation. Subsequently, another 10 U PNK and 1 mM ATP was added for 1.5 hours at 37 °C for 5′-phosphorylation before heat inactivation at 65 °C for 20 minutes. PNK treated samples were subjected to another round of phenol-chloroform purification prior to sending for small RNA sequencing (Macrogen). The small RNA library was prepared using the Illumina TruSeq RNA Library Prep Kit with 20% PhiX and 10% miRNA control added. A HiSeq X lane (2x150bp) sequencing run was performed. The obtained reads (2,649,538 in total) were trimmed using cutadapt [215] to remove the adapters. The single reads of size 40 to 60 (954,096 in total) were filtered and mapped on the CRISPR array sequence using minimap2 [216].

SDS-PAGE ANALYSIS

20 µL of protein sample was supplemented with 5 µL of 5X Laemmli buffer (375 mM Tris-HCl, 9% SDS, 50% glycerol, 0.03% bromophenol blue) and 2.5 µL 1 M DTT. Samples were incubated at 95 °C for 10 minutes before loading on a 4-20% Mini-Protean TGX Precast Gel (Bio-Rad) with PageRuler Prestained Protein Ladder (10 to 180 kDa) or PageRuler Plus Prestained Protein Ladder (10 to 250 kDa). Gels were run in 1X TGS buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS) at 200 V for 30-45 minutes, washed in MilliQ water and stained for 1 hour with BioSafe Coomassie G-250 stain (Bio-Rad) under continuous

shaking. Gels were washed in MilliQ water for 30 minutes before imaging.

TARGET RNA CLEAVAGE

RNA cleavage reactions were performed in 10 µL reaction volume, containing purified 200 nM Sb-gRAMP-crRNA, 100 nM Cy5-labelled RNA oligo, 11 mM Tris-HCl, 67 mM NaCl, 4.5 mM DTT and 1 mM MgCl2 unless stated otherwise. Typical reactions were incubated at 20 °C for 2 hours, after which 0.5 µL of 800 units/mL proteinase K (New England Biolabs) was added for 1 hour at 37 °C and 95 °C for 5 minutes to stop the reactions. 5 µL of the reaction was mixed with 5 µL of 2x RNA loading dye (95% formamide, 0.025% SDS and 0.5 mM EDTA) and loaded on an 8M urea 10% PAGE gel (pre-run at 350 V for 1 hour, sample run at 333V for 2 hours) with 100 nM ladder containing 5′-Cy5 labelled RNA oligo's of sizes 10, 20, 30, 40, 50 and 60 nt (**table S4**). Gels were imaged on a Typhoon laser-scanner platform (Cytiva).

BIOCHEMICAL CHARACTERIZATION OF RNA ENDS

In order to characterize the chemical 5′- and 3′-ends of the cleavage products generated by Sb-gRAMP, cleavage reactions were heat inactivated at 95 °C for 5 minutes and subsequently treated with Antarctic phosphatase (for 5′ and 3′ dephosphorylation) or T4 PNK (for 5′ phosphorylation and 2′-3′-dephosphorylation). For phosphatase treatment, 1.1 µL 10x Antarctic phosphatase reaction buffer (New England Biolabs) and 0.5 µL of 5,000 units/mL Antarctic phosphatase (New England Biolabs) was added and incubated for 30 minutes at 37 °C. For T4 PNK treatment, 1 µL of 10,000 units/mL T4 PNK (New England Biolabs), 1 µL 10 mM ATP and 1 µL 10x T4 PNK buffer (New England Biolabs) was added and incubated for 30 minutes at 37 °C. Reactions were inactivated at 95 °C for 5 minutes, whereupon 0.5 µL of 800 units/mL proteinase K (New England Biolabs) was added and incubated for 30 minutes at 37 °C, followed by inactivation at 95 °C for 5 minutes. For visualization, 10 µL of the reaction was mixed with 2x RNA loading dye (95% formamide, 0.025% SDS and 0.5 mM EDTA) and loaded on an 8M urea 10% PAGE gel (pre-run at 350 V for 1 hour, sample run at 333V for 2 hours). Gels were imaged on a Typhoon laser-scanner platform (Cytiva).

SEC-MALS

The gRAMP-crRNA1 and gRAMP-crRNA3 samples for SEC-MALS were prepared by pooling the SEC elutions corresponding to the protein peaks (10-11 mL). Samples were injected at final concentrations of 1.2 mg/mL (gRAMP-crRNA1) and 1.0 mg/mL (gRAMP-crRNA3) in running buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, pH 7.5) onto a Superdex 200 Increase 10/300 GL (Cytiva). After SEC, the sample passed a DAWN HELEOS 8 light scattering instrument (Wyatt) and a Optilab T-rEX refractive index detector (Wyatt). Light source of the RI detector and wavelength of the laser in the light scattering instrument were 658 and 660 nm, respectively. Molecular mass distribution and concentrations of chromatogram peaks were calculated based on the light scattering signal and the refractive index, respectively, derived from peak half-height regions (9.87-10.43 mL for gRAMP-crRNA1 and 9.76-10.39 mL for gRAMP-crRNA3). For molar mass calculations, ASTRA 7.3.2 software (Wyatt) with the protein conjugate analysis module was used with refractive-index increments (dn/dc) of 0.1850 mL/g (protein)

and 0.1680 mL/g (RNA). Protein UV extinction coefficients of 1.203 mL mg^{-1} cm^{-1}, and RNA UV extinction coefficients of 15.3914 mL mg^{-1} cm^{-1} (gRAMP-crRNA1) and 15.8769 mL mg^{-1} cm^{-1} (gRAMP-crRNA3) were used. UV extinction coefficients were estimated using ExPASy ProtParam [217] and MolBioTools DNA calculator (www.molbiotools.com) using the sequences for gRAMP and TPR-CHAT. For the RNA of gRAMP-crRNA1, the UV extinction coefficient was calculated using 55 nt ssRNA of crRNA1 (AAACAAGAGAAG-GACUUAAUGUCACGGUACCCAAUUUUCUGCCCCGGACUCCACG) with 5′ phosphate. For the RNA of gRAMP-crRNA3, the UV extinction coefficient was calculated using 55 nt ssRNA of crRNA3 (AAACAAGAGAAGGACUUAAUGUCACGGUACAAUUAUCAUUUG-GACAGCUUCCCUC) including
a 5′ phosphate.

RNase T1 ladder generation

For exact cleavage site determination, cleavage products were compared to a ladder consisting of RNase T1 digested target RNA. To generate the ladder, ~20 ng of Target 2 or Target 7 was combined with ~140 ng of *E. coli* RNA that was extracted with the mirVana RNA Isolation Kit (Thermo Fisher Scientific) using the total RNA protocol. Subsequently, 0.1 U of RNase T1 (Thermo Fisher Scientific) was added and incubated for 1 minute at 37 °C, followed by addition of 1.6 U of Proteinase K (New England Biolabs) and incubation for 30 minutes at 37 °C. Heat inactivation was performed for 5 minutes at 95 °C whereupon 2x RNA loading dye (95% formamide, 0.025% SDS and 0.5 mM EDTA) was added. The samples were stored at -20 °C until use.

Type III-E bioinformatics mining

For the identification of assemblies containing gRAMP, earlier reported gRAMPs [43] were used as seed in the non-redundant NCBI database, MGNIFY database and IMG databases on January 13th, 2021. Mmseqs2 easy-cluster was used to find gRAMP homologs based on 60% coverage of the earlier reported gRAMPs with a minimum sequence identity of 25% [218]. Genes of discovered contigs were annotated using hhpred using default settings [219]. Only gRAMPs containing a co-occurring TPR-CHAT were included in further analyses.

Spacer analysis and strand bias determination

The type III-E spacers were collected from previously predicted type III-E systems [220]. Additional spacers from metagenomes were found based on similarity to the repeat sequences from already described arrays and a complete genome of *Desulfonema magnum* also contained additional spacers. Matches in (meta)genomes to these spacers were subsequently found using BLAST as previously described [57]. Hits were removed when found within arrays or with nucleotide identity <90% (no gaps allowed). Two hits from MVRP01000104.1 that were below the 90% were still included based on the low possibility that these hits would occur at random in the same metagenome (e-value $<10^{-6}$). The potential MGE origin of type III-E spacer matches in metagenomes was determined by analyzing and annotating the flanking genes of spacer matches using prokka [221]. If one of the genes were determined to be characteristic of viruses (e.g. viral proteins), transposons (e.g. transposases) or plasmids (e.g. conjugation machinery), the hit with

the spacer match was predicted to be of MGE origin. Strand bias was determined by comparing the orientation of the spacer compared to the orientation of open reading frames predicted on the matching target using Prodigal [57, 222].

PHYLOGENETIC TREE CONSTRUCTION

All Cmr4 and Csm3 protein sequences were downloaded from Interpro on May 19th, 2021, and clustered with the gRAMP proteins. The phylogenetic tree was constructed as previously described [210]. In short, cd-hit v4.8.1 [223] was used to filter redundant proteins. The remaining proteins were aligned using MAFFT v7.475 linsi [224] and Trimal v1.4.rev15 [225] was used to remove columns composed of gaps (-nt 0.95). Tree visualization was done by IQ-TREE v2.1.2 [226] using 1000 bootstraps. To find out to which CRISPR-Cas type the Cmr4 and Csm3 proteins belonged, each protein was blasted onto the NCBI database. In case of an exact hit, the contig in which the hit was found was analyzed using CRISPRCasTyper [220]. The final phylogenetic tree was visualized using figtree v1.4.4 and iTol [227].

3.3. SUPPLEMENTARY INFORMATION

SUPPLEMENTARY TEXT

The molar mass determined for the nucleic acid content of the *Sb*-gRAMP-crRNA1 during SEC-MALS is 16.8 ± 13.9% kDa, or 14.46-19.14 kDa. Approximating the molar mass of an average ribonucleotide monophosphate at 320.5 Da and 5′-phosphate at 79 Da, 14.46 kDa corresponds to (14,460-79.0)/320.5 = 44.9, or 45 nucleotides and 19.14 kDa corresponds to (19,140-79.0)/320.5 = 59.5, or 60 nucleotides. This presents a range for ssRNA (45-60 nt) bound to *Sb*-gRAMP.

SUPPLEMENTARY FIGURES

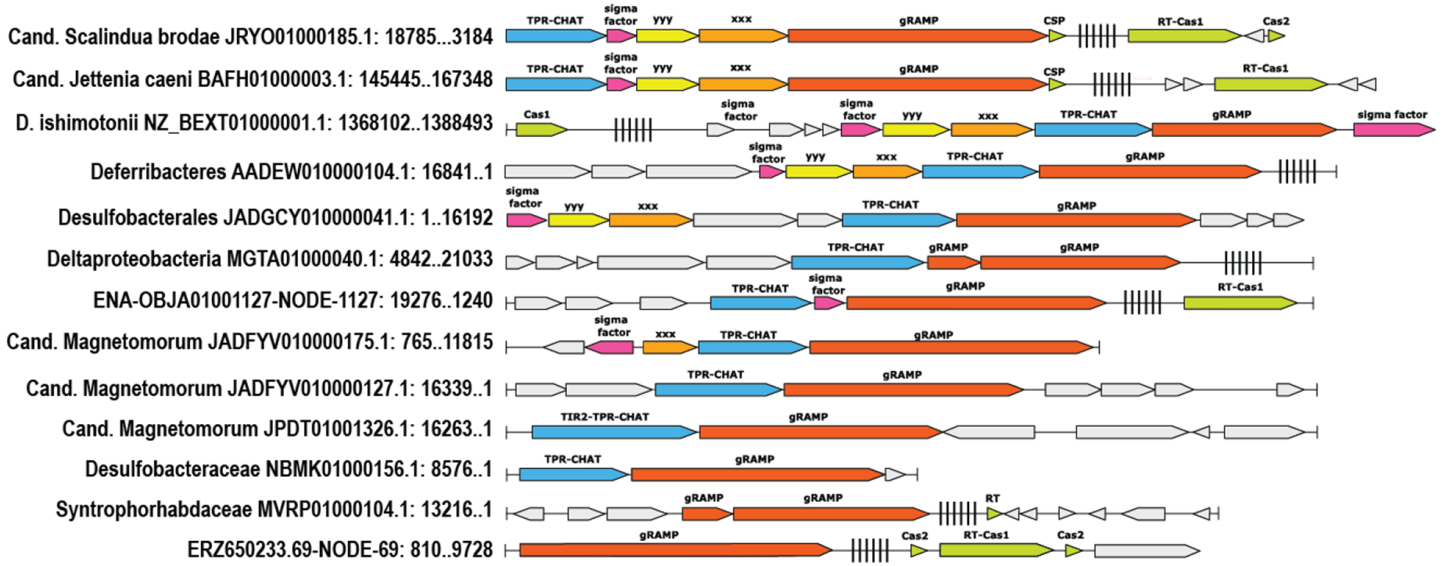


Figure S1. Identified CRISPR-Cas type III-E loci. Typical type III-E components include gRAMP (red), TPR-CHAT (blue), RpoE sigma factor (pink) and two genes of unknown function (yyy, also known as Csx31, in yellow; xxx, also known as Csx30, in light orange). Canonical CRISPR-Cas proteins are annotated in green. CRISPR arrays are indicated with staggered vertical lines. Start and ends of contigs are indicated with single vertical lines at the ends.

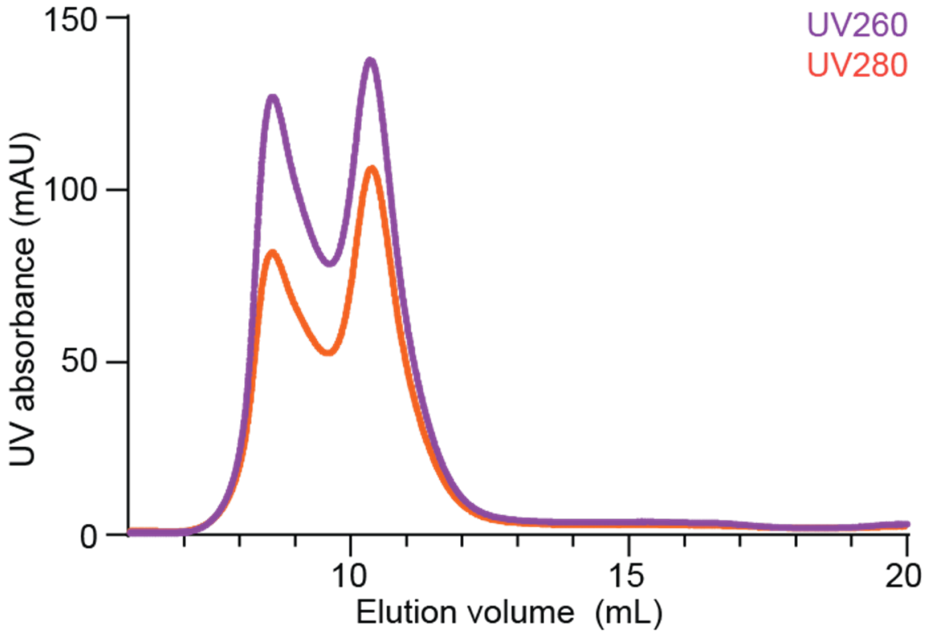


Figure S2. *Sb*-gRAMP size exclusion chromatography. UV absorption peaks at 260 nm (purple) and 280 nm (orange) after size-exclusion chromatography (SEC) (Superdex 200 Increase 10/300 GL) of purified *Sb*-gRAMP that was expressed in the presence of a plasmid containing five copies of the first spacer from the native *Candidatus* "Scalindua brodae" CRISPR array.

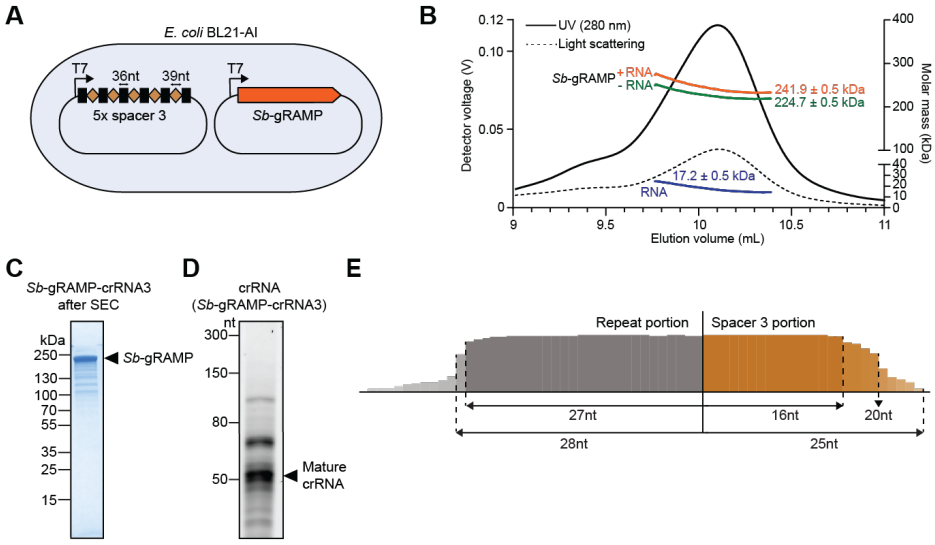


Figure S3. Protein purification and RNA characterization of *Sb*-gRAMP expressed in the presence of a CRISPR array derived from spacer 3 (*Sb*-gRAMP-crRNA3). (**A**) Schematic showing the heterologous expression system used for *Sb*-gRAMP-crRNA3 overexpression. (**B**) SEC-MALS peaks corresponding to light scattering and UV_{280} absorbance profiles of purified *Sb*-gRAMP-crRNA3. The molar masses are derived from the peak half-height region (9.8-10.4 mL). (**C**) SDS-PAGE analysis after SEC. (**D**) Denaturing urea PAGE analysis of the nucleic acid content in *Sb*-gRAMP-crRNA3. (**E**) RNAseq read depth of repeat-spacer 3 with predominant repeat and spacer lengths indicated based on consensus processing (**fig. S6B**).

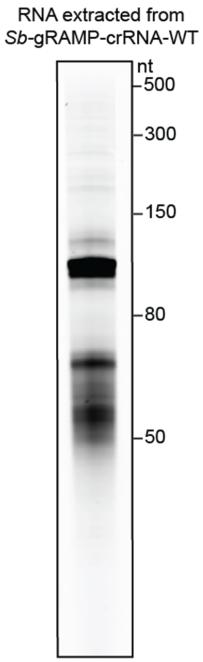


Figure S4. RNA extracted from *Sb*-gRAMP expressed with the native *Candidatus* "Scalindua brodae" CRISPR array (*Sb*-gRAMP-crRNA-WT) after heparin chromatography. The extracted RNA was subsequently used for RNAseq.

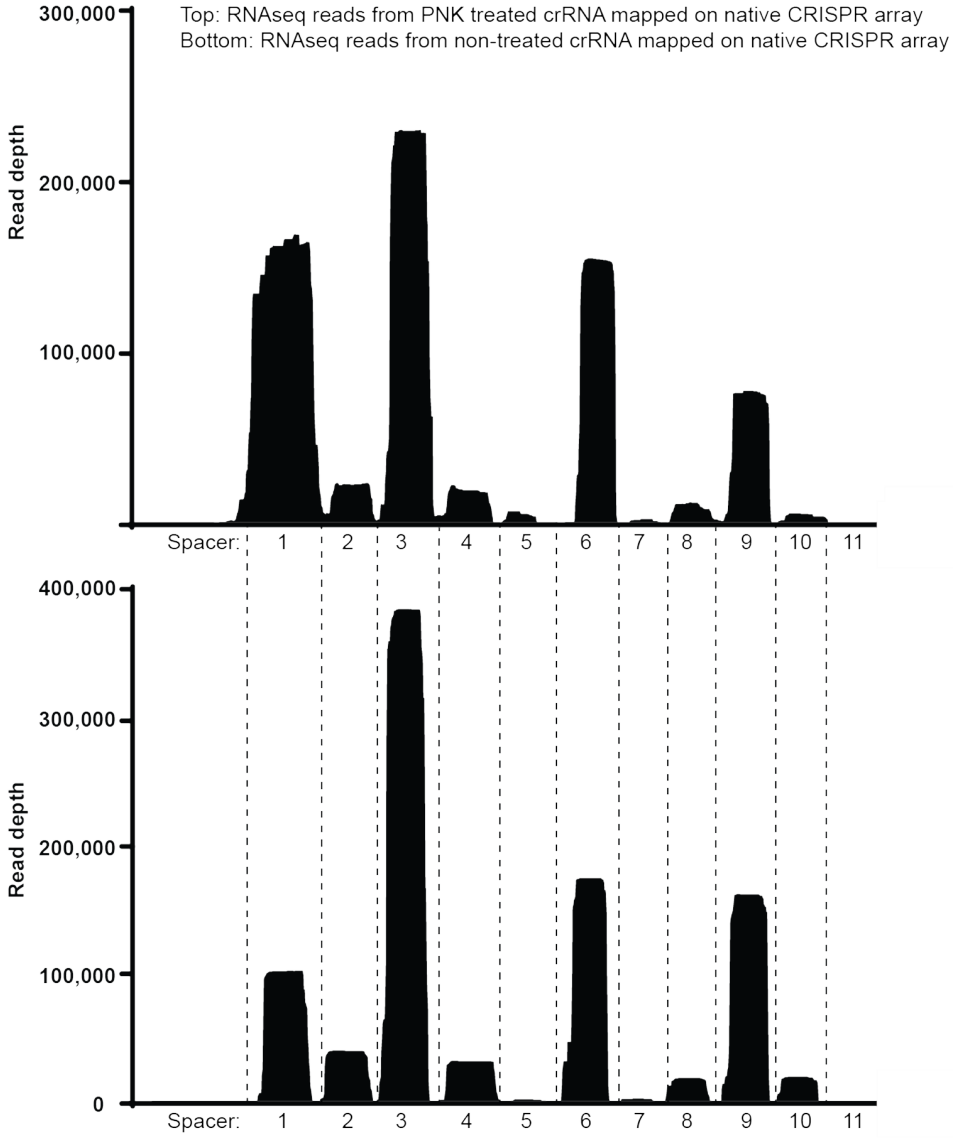


Figure S5. RNAseq of crRNAs in *Sb*-gRAMP. RNAseq read depth of mature crRNA (40 to 60 nt), with (top) or without (bottom) prior T4 Polynucleotide Kinase (PNK) treatment. Both sequence read sets are mapped on the native CRISPR array of *Candidatus* "Scalindua brodae". RNAseq protocols require 5′-P and 3′-OH for appropriate adapter ligation to the RNA. PNK is capable of adding a 5′-phosphate and removing a cyclic 2′-3′-phosphate, thereby ensuring crRNA termini suitable for RNAseq.

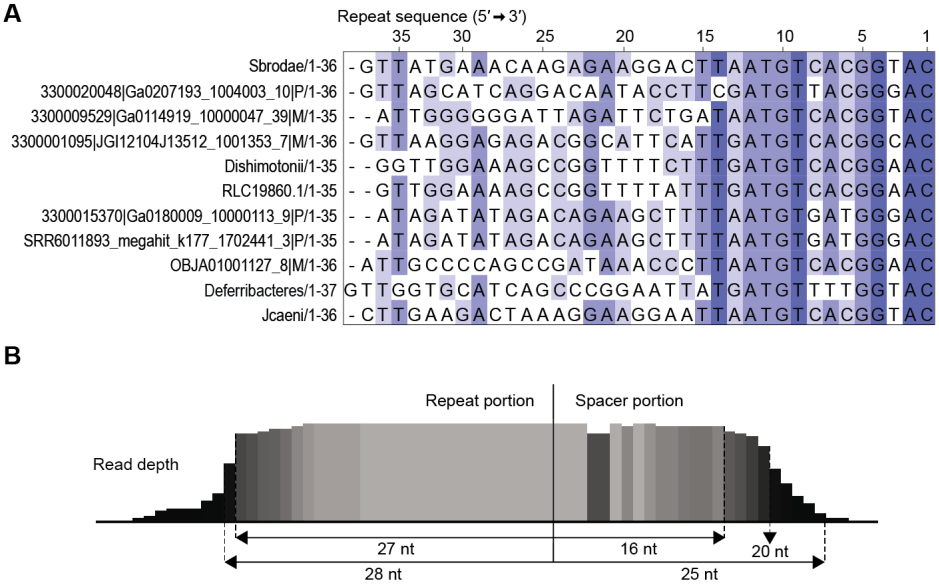


Figure S6. ***Candidatus*** **"Scalindua brodae" native CRISPR array analysis.** (**A**) Alignment of repeats found in various type III-E loci with the consensus nucleotide per position colored on occupancy in the alignment, from white (no consensus nucleotide) to dark (complete or near-complete occupancy of the consensus nucleotide). The alignment was generated using the Cobalt multiple alignment tool [228] with default settings and visualized using Jalview [229]) Read depth of the repeat and spacers with 10,199 reads in which all repeat-spacer pairs are represented. Consensus processing sites are indicated.

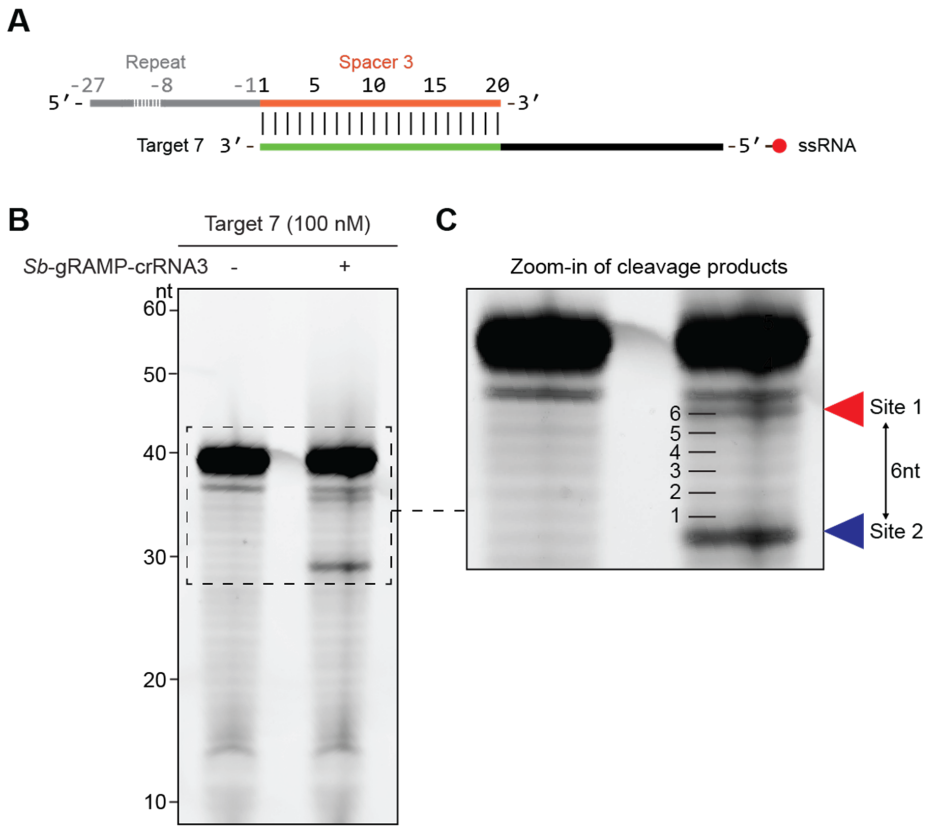


Figure S7. *Sb*-gRAMP-crRNA3 cleaves cognate target ssRNA. (**A**) Outline of the Cy5-labelled RNA target tested for activity of *Sb*-gRAMP loaded with crRNA derived from spacer 3 in the native *Candidatus* "Scalindua brodae" CRISPR array (*Sb*-gRAMP-crRNA3). Substrate details are listed in table S4. (**B**) Denaturing urea PAGE gel of cleavage reactions consisting of 200 nM *Sb*-gRAMP-crRNA3 incubated with cognate Target 7. (**C**) Zoom-in of the cleavage products with single nucleotide steps in the auto-hydrolysis ladder indicated.

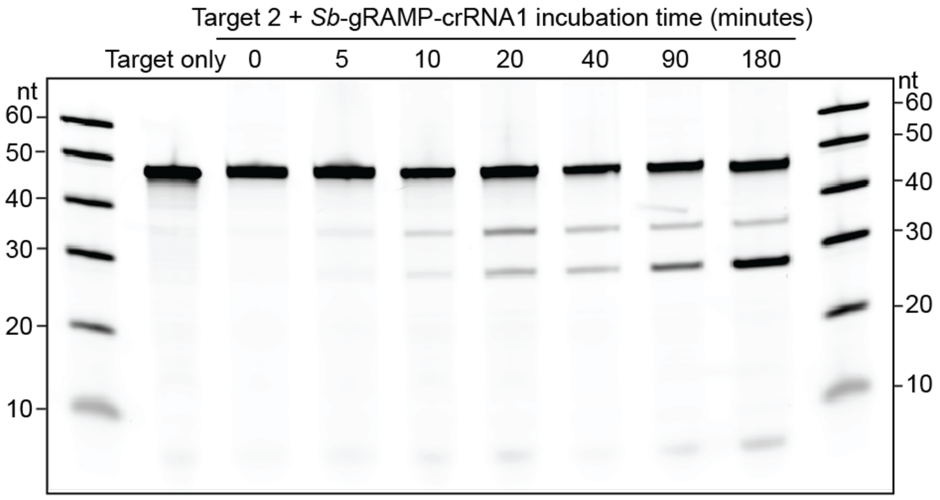


Figure S8. ***Sb*-gRAMP-crRNA1 incubation with complementary ssRNA over time.** Denaturing urea PAGE gel of cleavage reactions consisting of *Sb*-gRAMP-crRNA1 and cognate Target 2 with different incubation times. Substrate details are listed in **table S4**.

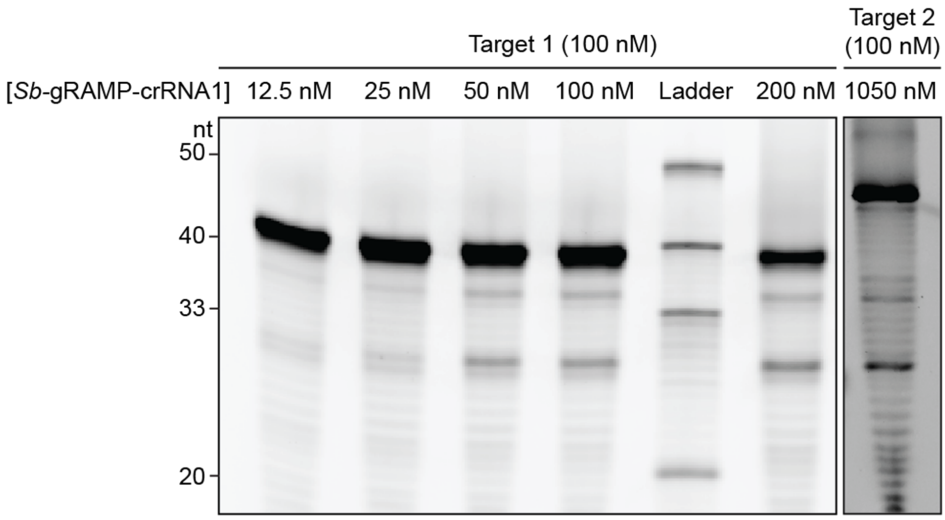
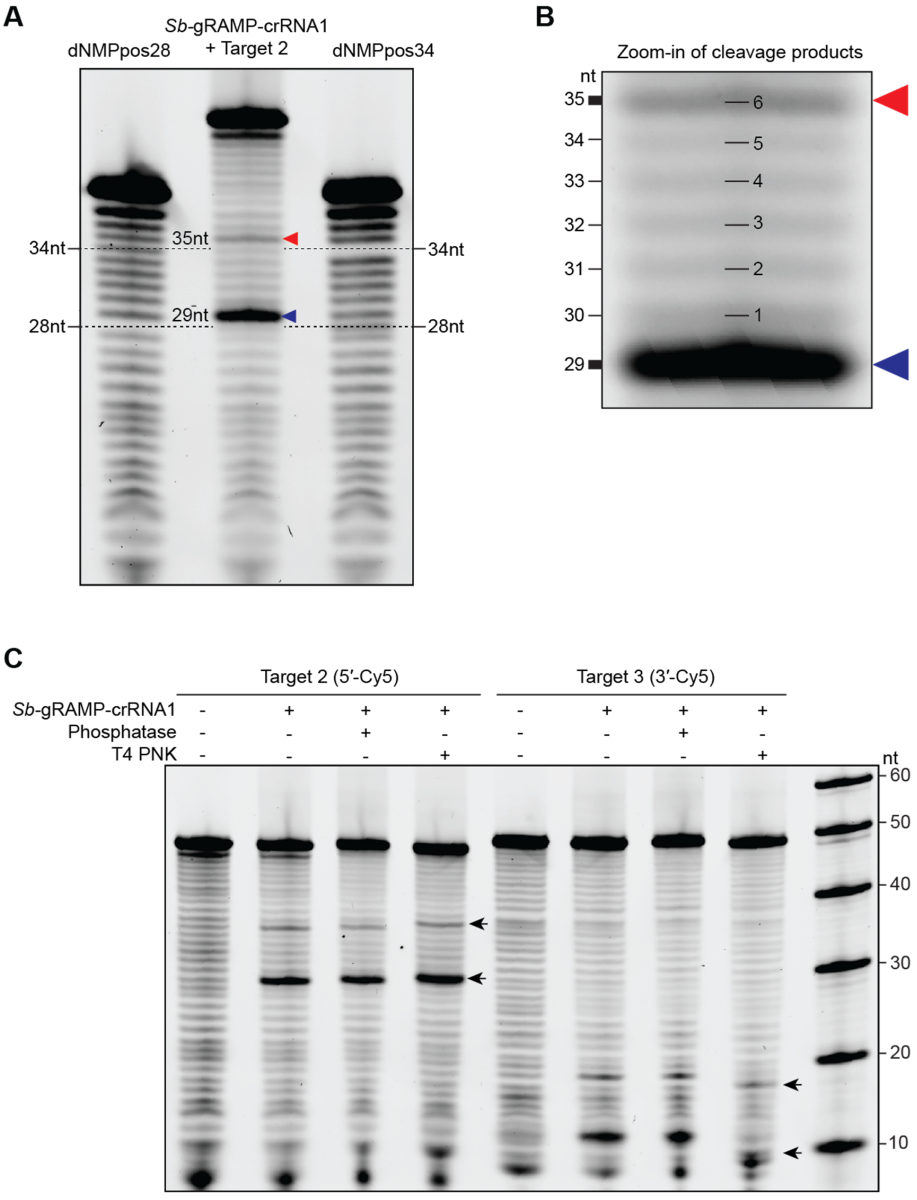
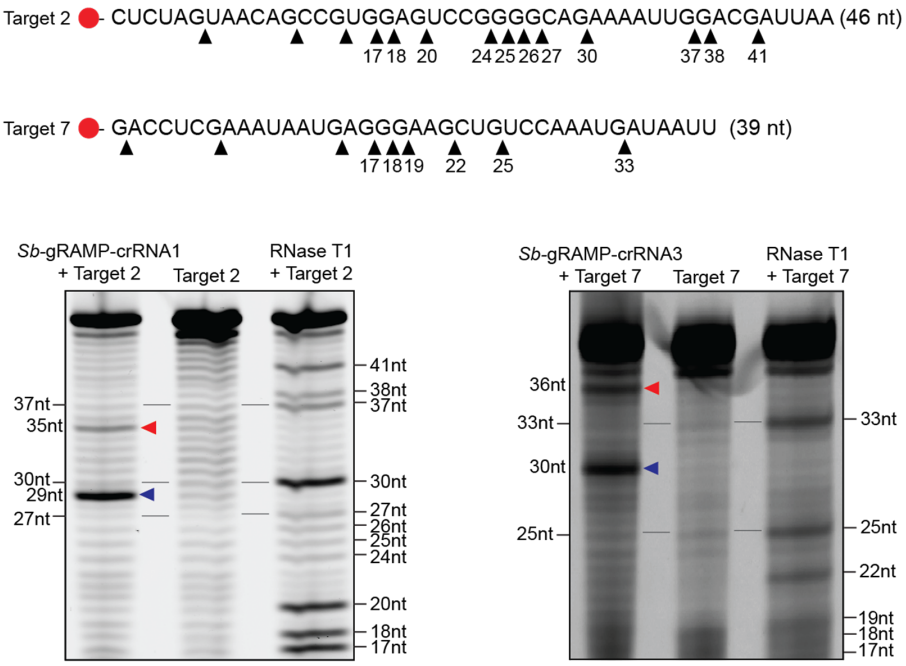


Figure S9. Target RNA cleavage by *Sb*-gRAMP-crRNA1 proceeds with a limited substrate turnover. Denaturing urea PAGE gel of cleavage reactions consisting of different *Sb*-gRAMP-crRNA1 concentrations with 100 nM cognate Target 1 and Target 2. Substrate details are listed in **table S4**.


A
Sb-gRAMP-crRNA1
+ Target 2
dNMPpos28
dNMPpos34
34nt
28nt
35nt
29nt
34nt
28nt
B
Zoom-in of cleavage products
nt
35
34
33
32
31
30
29
6
5
4
3
2
1
C
Target 2 (5′-Cy5)
Target 3 (3′-Cy5)
Sb-gRAMP-crRNA1
Phosphatase
T4 PNK
nt
60
50
40
30
20
10

D


Target 2 CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGGACGAUUAA (46 nt)
17 18 20 24 25 26 27 30 37 38 41
Target 7 GACCUCGAAAUAAUGAGGGAAGCUGUCCAAAUGAUAAUU (39 nt)
17 18 19 22 25 33
Sb-gRAMP-crRNA1 + Target 2
Target 2
RNase T1 + Target 2
37nt
35nt
30nt
29nt
27nt
41nt
38nt
37nt
30nt
27nt
26nt
25nt
24nt
20nt
18nt
17nt
Sb-gRAMP-crRNA3 + Target 7
Target 7
RNase T1 + Target 7
36nt
33nt
30nt
25nt
33nt
25nt
22nt
19nt
18nt
17nt

E

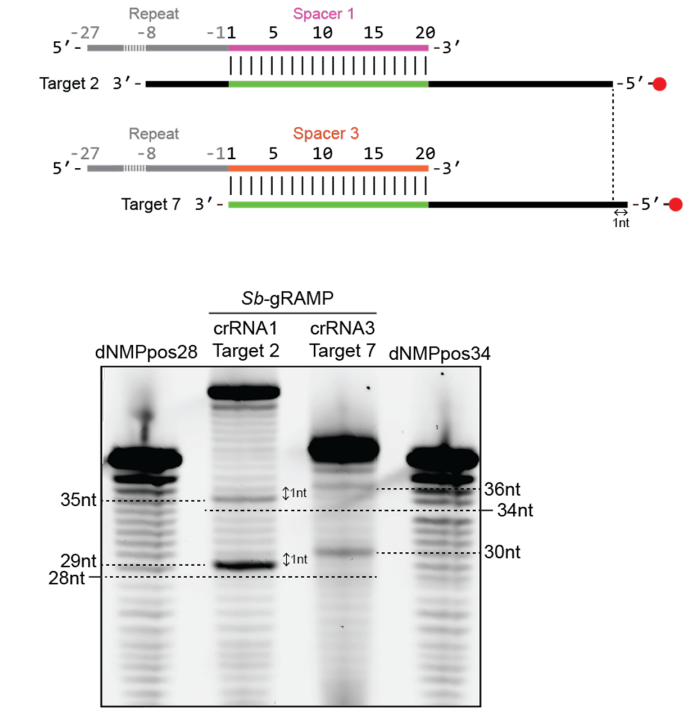

Repeat
Spacer 1
-27 -8 -11 5 10 15 20
5′- -3′
Target 2 3′- -5′
Repeat
Spacer 3
-27 -8 -11 5 10 15 20
5′- -3′
Target 7 3′- -5′
1nt
Sb-gRAMP
crRNA1 Target 2
crRNA3 Target 7
dNMPpos28
dNMPpos34
35nt
29nt
28nt
1nt
1nt
36nt
34nt
30nt

Figure S10. *Sb*-gRAMP-crRNA1 and *Sb*-gRAMP-crRNA3 exact cleavage positions in target RNA. (**A**) Determination of the cleavage sites using the auto-hydrolysis products from the target RNA and the internal absence of an auto-hydrolysis product in ssRNA oligo's due to a deoxyribonucleoside monophosphate (dNMP) at position 28 (dNMPpos28) or position 34 (dNMPpos34). The absence of a 2′-OH at dNMP prevents the nucleophilic attack on the 3′ adjacent phosphorus in the phosphodiester bond and hence prevents the appearance of an auto-hydrolysis product of that size. Cleavage products are indicated with the red and blue arrowheads and are 29 nt and 35 nt. Substrate details are listed in table S4. (**B**) Zoom-in of the cleavage products generated by *Sb*-gRAMP-crRNA1 with counting steps in the auto-hydrolysis ladder. Cleavage products are indicated with arrowheads. (**C**) Cleavage products of cognate target RNA Cy5-labelled 5′ (Target 2) or 3′ (Target 3) after incubation with Antarctic phosphatase and T4 polynucleotide kinase (PNK). Treatment of Target 2 with PNK resulted in an upwards shift (indicated by the arrows) of the cleavage products, suggesting the removal of a 3′-phosphate. Conversely, treatment of Target 3 with PNK resulted in a marked downward shift (indicated by the arrows), indicating the addition of a 5′-phosphate. These results indicate that cleavage products generated by *Sb*-gRAMP-crRNA1 carry 5′-OH and 3′-P, allowing comparison with auto-hydrolysis products in (**A**), as the chemical nature of hydrolysis products is also 5′-OH and 3′-P and thus comparable in electrophoretic mobility. (**D**) Determination of the cleavage sites using RNase T1 digested target RNA as a ladder. RNase T1 cleaves the phosphodiester bond of single-stranded RNA between 3′-guanylic residues and the 5′ of the adjacent nucleotides. RNase T1 cleavage sites of Target 2 (cognate to crRNA1) and Target 7 (cognate to crRNA3) are indicated with black arrowheads and the size is indicated in nt. Cleavage products are indicated with the red and blue arrowheads and are 29 nt and 35 nt for *Sb*-gRAMP-crRNA1, and 30 nt and 36 nt for *Sb*-gRAMP-crRNA3. Target 7 is 1 nt longer (due to a 1 nt longer native spacer) at the 5′ (where the Cy5 label is positioned) compared to Target 2 (**E**), so the observed one nucleotide difference in cleavage products for *Sb*-gRAMP-crRNA1 and *Sb*-gRAMP-crRNA3 indicates that the same positions are cleaved relative to the crRNA. This demonstrates that target RNA cleavage is position specific.

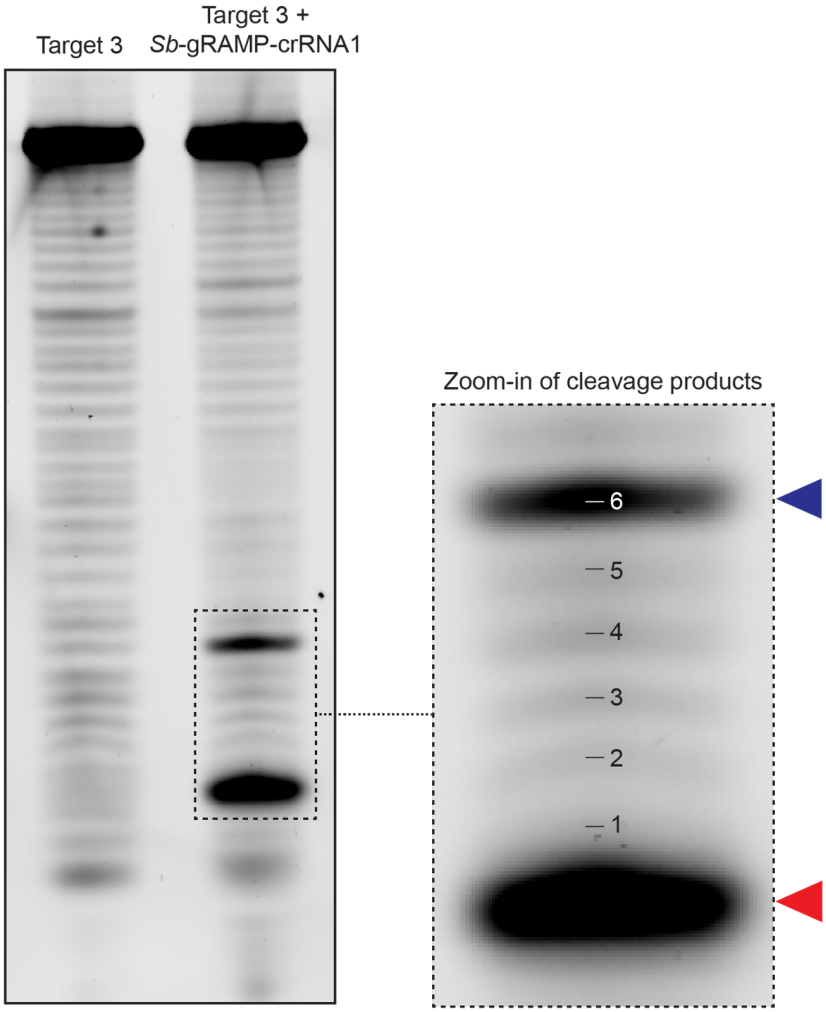


Figure S11. *Sb*-gRAMP-crRNA1 cleavage products of Target 3 are 6 nt apart. Denaturing urea PAGE gel of Target 3 incubated with *Sb*-gRAMP-crRNA1 with a zoom-in of the cleavage products (indicated with arrowheads) in which nucleotide counting steps are indicated.

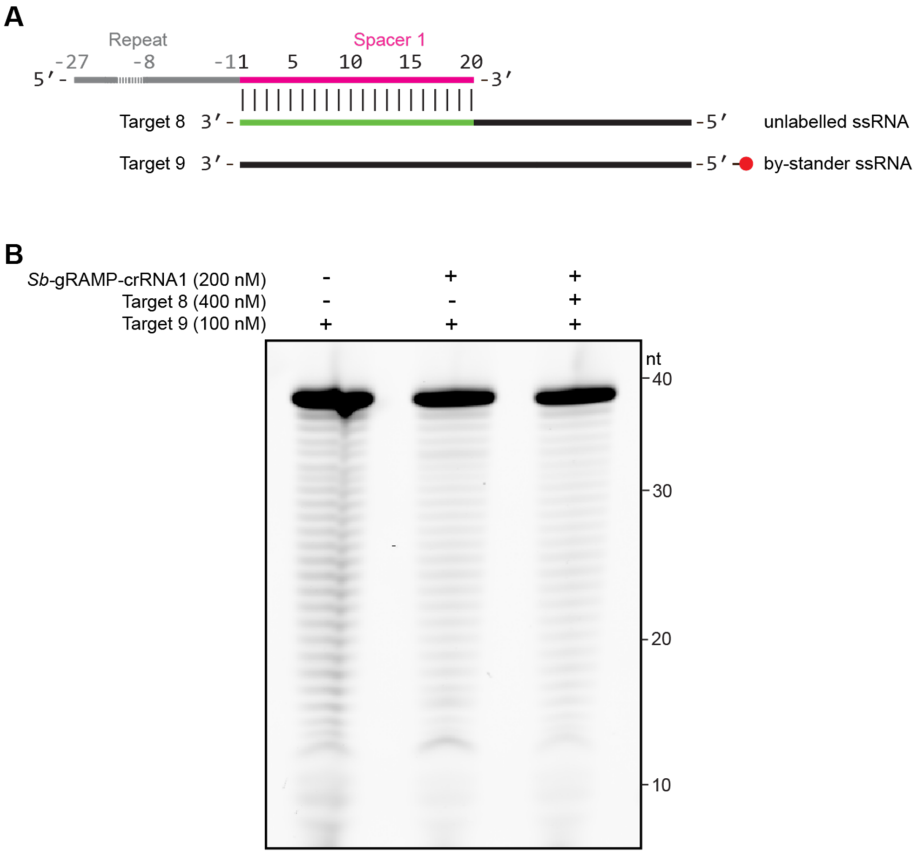
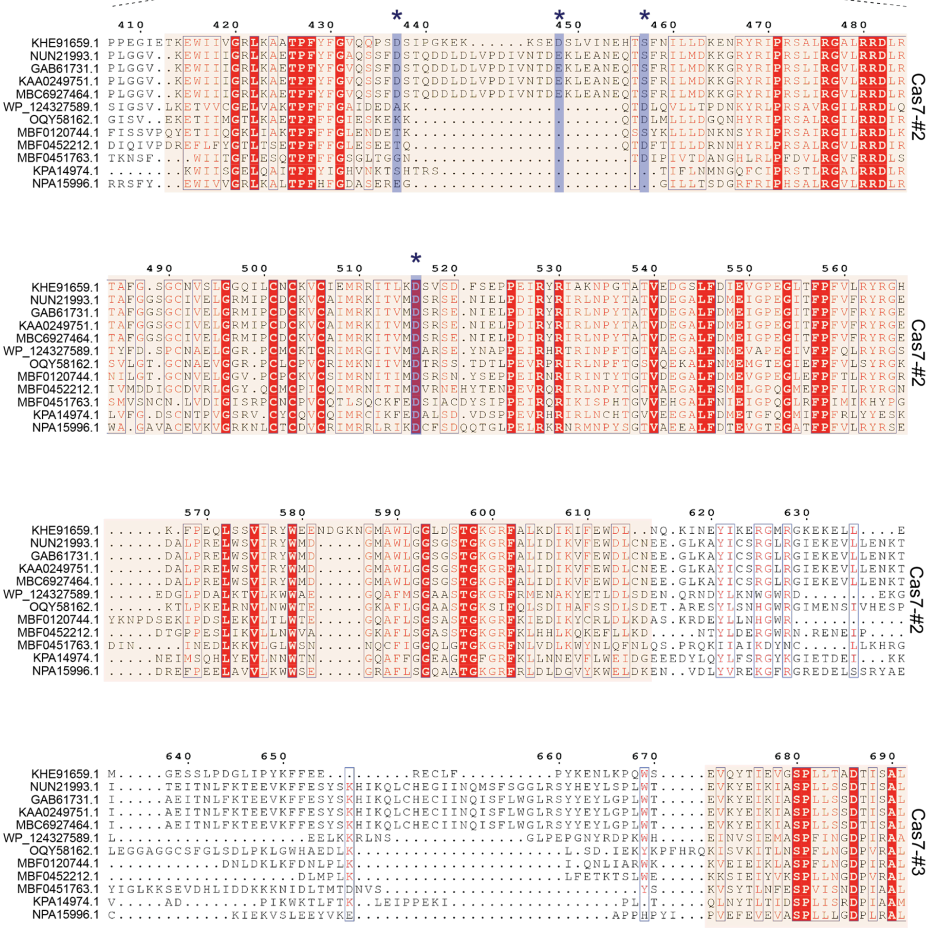


Figure S12. *Sb*-gRAMP-crRNA1 has no collateral activity against non-complementary by-stander RNA. (**A**) Schematic of the collateral cleavage experiment involving unlabeled target RNA (Target 8) complementary to crRNA1 and Cy5-labeled non-complementary bystander RNA (Target 9). Substrate details are listed in table S4. (**B**) Denaturing urea PAGE gel of 200 nM *Sb*-gRAMP-crRNA1 incubated with 2-fold molar excess of Target 8 and 100 nM Target 9.


Cas7-#1
Cas11
Cas7-#2
Cas7-#3
Cas7-#4


KHE91659.1
NUN21993.1
GAB61731.1
KAA0249751.1
MBC6927464.1
WP_124327589.1
OQY58162.1
MBF0120744.1
MBF0452212.1
MBF0451763.1
KPA14974.1
NPA15996.1
Cas7-#2
Cas7-#3

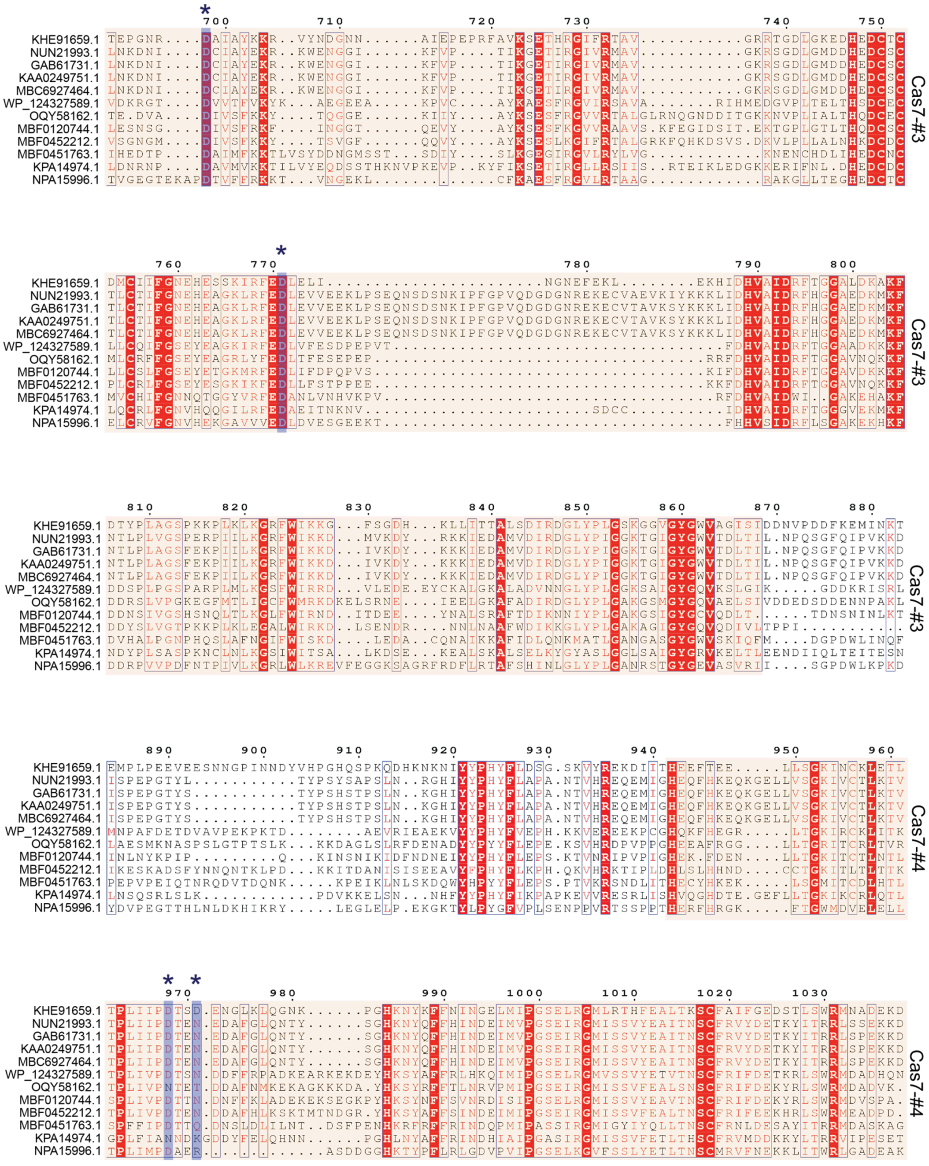


Figure S13. gRAMP multiple sequence alignment of the suspected catalytic Cas7-like domains. Residues with identity are boxed and represented in different shades of red, with full conservation marked in dark red. Generated alanine substitutions are indicated with a dark blue asterisk and the aligning residues are marked with a blue box. Transparent orange blocks demarcate the Cas7-like domains. The alignment was generated using the Cobalt multiple alignment tool [228] default settings and visualized using ESPript 3 [230]

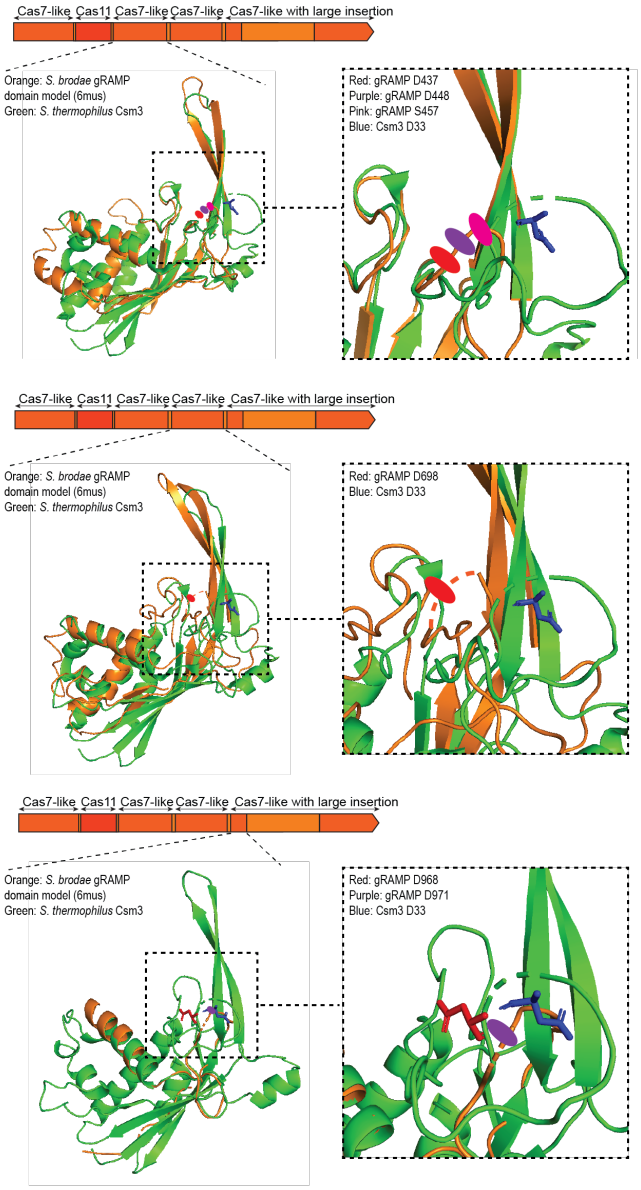


Figure S14. Structural modelling of the Cas7-like domains. Models were generated using Phyre2 intensive mode [231] and aligned with the Csm3 structure of *Streptococcus thermophilus* (PDB: 6IG0) in PyMOL [232] The aspartic acid or serine residues (indicated with sticks or oval representation in red, purple and pink) of the Cas7-like domains in *Sb*-gRAMP that structurally aligned close to the active aspartic acid residue in *S. thermophilus* (indicated with sticks in blue) were subjected to alanine mutational analysis.

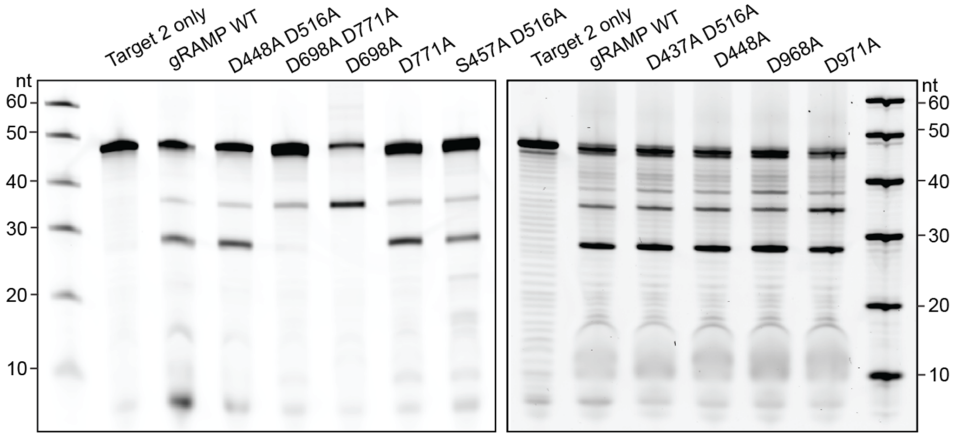


Figure S15. *Sb*-gRAMP mutational analysis. Denaturing urea PAGE gel of target RNA cleavage reactions with crRNA1 loaded *Sb*-gRAMP mutants in which candidate active residues were mutated to alanines.

SUPPLEMENTARY TABLES

Spacer	Target	Identity	Orientation	Source	MGE origin
JAADEW010000104.1	JAADEW010000104.1	100%	Coding	Deferribacteres bacterium	No
NZ_CP061800.1_12:1	SDBT01009396.1	91%	Coding	Marine sediment metagenome	No
NZ_CP061800.1_10:19	Ga0210401_10011668	90%	Coding	Forest soil metagenome	No
JRYO01000185.1_1:6/1-34	OVUK01024005.1	94%	Coding	Aquaculture metagenome	No
NZ_BEXT01000001.1_7:22/1-31	ERZ840537	90%	Coding	Marine metagenome	No
JRYO01000185.1_1:8/1-37	JRYO01000117	92%	Coding	Candidatus Scalindua brodae	Yes
MVRP01000104.1_13:4/1-48	MVRP01000213.1	87%	Coding	Syntrophorhabdaceae bacterium PtaU1	No
MVRP01000104.1_13:1/1-47	MVRP01000160.1	88%	Coding	Syntrophorhabdaceae bacterium PtaU1	No
JPDT01000635.1_5:30	UYTZ01000008.1	91%	Coding	human gut metagenome	No
JPDT01002993.1_36:84	OFGZ01002972.1	94%	Not determined	Mouse gut metagenome	No
JAABRU010000211.1_1:8	LR796420	91%	Template	Phage	Yes
JPDT01002993.1_36:91	LBBO01004998.1	91%	Coding	Hydrothermal vent metagenome	No
JPDT01002663.1_30:10	CEVW01070774.1	94%	Template	Marine metagenome	No
JPDT01002993.1_36:35	CESD01017330.1	91%	Coding	Marine metagenome	No
JPDT01002993.1_36:85	CERB01284573.1	91%	Not determined	Marine metagenome	No
JPDT01000635.1_5:6	BK043628	91%	Coding	Phage	Yes
ERZ650233.69-NODE-69_1:1	APMI01113767.1	100%	Not determined	Wastewater metagenome	No
JPDT01002993.1_36:11	AACY020545844.1	91%	Coding	Marine metagenome	No
JPDT01002663.1_30:10	LR796345	91%	Coding	Phage	Yes
JPDT01000635.1_5:29	ERZ841010	91%	Not determined	Marine metagenome	Yes

Table S1. Type III-E CRISPR-Cas spacer analysis. Listed are spacers from type III-E CRISPR-Cas loci with identified targets, the percentage of nucleotide identity, the target strand orientation, the source of the hit and whether the source of the hit was predicted to be of MGE origin.

Spacer	DNA sequence (5′ to 3′)	Length (nt)
1	CCAATTTTCTGCCCCGGACTCCACGGCTGTTACTAGAG	38
2	AGTTTCCTGTTTTTTTTGCTCCCTAACGCTACTTTGAAT	39
3	AATTATCATTTGGACAGCTTCCCTCATTATTTCGAGGTC	39
4	GAAAAAAAAAAAGTAAAGTTCAGGGGCAAGTGCCAAA	37
5	CCCTTTGCTTCTTCTCTAGTGTTTCTATCCATGTTTGT	38
6	TTACGAAGTATCTCCGTACGAACCTTTTCACTGT	34
7	AGAATTGGTATTATTTTTCCCAGTGTAATAATACC	35
8	ACCATTTTTGTCATTATTTATTGTCATGTTAGAAA	35
9	TCTTCAGCAATTACTTCTTTACGAAGAGATAACTTT	36
10	AAAATCTCAAGCCTCAAGCATATACTCAAAATCATTA	37
11	ATCATTACCATCCATATGTTCTGATGACTGTCTCTGCTGTA	41
Repeat	GTTATGAAACAAGAGAAGGACTTAATGTCACGGTAC	36

Table S2. DNA sequences of the spacers and repeat in the *Candidatus* "Scalindua brodae" CRISPR array.

	Experimentally calculated (*theoretical*) total molar mass (kDa)	Experimentally calculated (*theoretical*) protein molar mass (kDa)	Experimentally calculated (*theoretical*) RNA molar mass (kDa)
***Sb*-gRAMP-crRNA1**	242.5 ± 2.4 (*231.7*)	225.7 ± 2.3 (*214.0*)	16.8 ± 2.3 (*17.7*)
***Sb*-gRAMP-crRNA3**	241.9 ± 0.5 (*231.7*)	224.7 ± 0.5 (*214.0*)	17.2 ± 0.5 (*17.7*)

Table S3. The experimentally calculated versus theoretic molar mass for *Sb*-gRAMP-crRNA. For the theoretical molar mass of the RNA, 55 nt ssRNA of crRNA1 (AAACAAGAGAAGGACUUAAUGUCACGGUACCCAAUUUUCUGCCCCGGACUCCACG) with 5′ phosphate was used and calculated using MolBioTools DNA calculator (www.molbiotools.com).

Name	Sequence 5′ to 3′	nt	Cy5	Description
Target 1	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGG	38	5′	ssRNA complementary to crRNA1
Target 2	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGGACGAUUAA	46	5′	ssRNA complementary to crRNA1 (non-matching PFS)
Target 3	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGGACGAUUAACU	48	3′	ssRNA complementary to crRNA1
Target 4	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGGGUACCGUG	46	5′	ssRNA complementary to crRNA1 (matching PFS)
Target 5	CTCTAGTAACAGCCGTGGAGTCCGGGGCAGAAAATTGGACGATTAA	46	5′	ssDNA complementary to crRNA1
Target 6	GACCUCGAAAUAAUGAGGGAAGCUGUCCAAAUGAUAAUU	39	5′	ssRNA non-complementary to crRNA1
Target 7	GACCUCGAAAUAAUGAGGGAAGCUGUCCAAAUGAUAAUU	39	5′	ssRNA complementary to crRNA3
Target 8	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGG	38	-	Unlabelled ssRNA complementary to crRNA1
Ladder 10 nt	UCGGAUUCUG	10	5′	RNA ladder
Ladder 20 nt	GCGGAUUCUGAAACGGUGGA	20	5′	RNA ladder
Ladder 30 nt	CAAAGUGCUUACAGUGCAGGUAGUGAUAUG	30	5′	RNA ladder
Ladder 40 nt / Target 9	UUUUUUUUUUUCCGCGGCUUUUUUUUUUUUUUUCCGCGGC	40	5′	RNA ladder and by-stander RNA
Ladder 50 nt	GUCAUAGGAGAAGUAUUAACAUCUACAGGGAGUCUAUAUUGAGGUACUAG	50	5′	RNA ladder
Ladder 60 nt	GUCAUAGGAGAAGUAUUAACAUCUACAGGGAGUCUAUAUUGAGGUACUAGUCGGAUUCUG	60	5′	RNA ladder
dNMPpos28 (dNMP bold)	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGG	38	5′	Auto-hydrolysis internal control ladder
dNMPpos34 (dNMP bold)	CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGG	38	5′	Auto-hydrolysis internal control ladder

Table S4. Substrates used for in vitro cleavage experiments.

4

CRASPASE IS A CRISPR RNA-GUIDED, RNA-ACTIVATED PROTEASE

"Now, my own suspicion is that the universe is not only queerer than we suppose, but queerer than we can suppose."

J.B.S. Haldane

In type III-E CRISPR-Cas, the gRAMP effector protein uses a CRISPR RNA guide to bind and cleave viral RNA molecules. Functionality of the other proteins encoded in the type III-E loci are not understood. Here, we find that gRAMP physically combines with the co-localizing caspase-like TPR-CHAT peptidase to form the Craspase (CRISPR-guided Caspase) complex. Craspase is capable of sequence-specific RNA binding, which is the trigger for activation of the protease. We further define Csx30 as the endogenous protein substrate that is site-specifically proteolyzed by activated Craspase, whereupon cleavage of the bound RNA by Craspase turns the protease off again. We thus conclude that Craspase is a target RNA–activated protease with self-regulatory capacity.

A modified version of this chapter has been published as *van Beljouw et al., Science (2021)* [206] and *Hu* & van Beljouw* et al., Science (2022)* [233].

4.1. MAIN TEXT

The gRAMP protein is a single-unit, RNA-targeting effector belonging to subtype III-E CRISPR-Cas. 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) [85, 86]. 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 [89]. It appears that the Cas10 subunit as well as CARF-containing proteins were lost during the evolutionary genesis of gRAMP [43]. Since 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 [43]. The tetratricopeptide repeat (TPR) domain is often involved in protein-protein interaction and the formation of protein complexes [234], 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 co-expressed both proteins (**Fig. 1A**) and observed co-elution using either protein as a bait (**Fig. 1B; fig. S1A**). This indicated the formation of a stable complex between *Sb*-gRAMP-crRNA and TPR-CHAT (**table S1**), whose interaction was retained during subsequent heparin chromatography and size exclusion chromatography (**Fig. 1B; fig. S2**). Multiangle light scattering revealed a single, homogenous population with a molar mass of 315.4 ± 2.8 kDa (**Fig. 1C; fig. S3**), corresponding to *Sb*-gRAMP-crRNA and TPR-CHAT complexed in a 1:1 stoichiometry (**table S2**). The retention time of the complex was shorter than that of *Sb*-gRAMP-crRNA alone, consistent with a particle of higher molecular weight (**fig. S4**). The heparin purified complex contained mature crRNA (**Fig. 1D**) with 6 nt spaced cleavage specificity towards target RNA (**Fig. 1E; fig. S5; fig. S1B**). TPR-CHAT with inactivated predicted catalytic residues (H585A and C627A) [209] was also able to associate with *Sb*-gRAMP-crRNA (**fig. S1**), 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. S6**), suggesting that *Sb*-gRAMP-crRNA and TPR-CHAT form a stable protein complex which remains assembled upon target RNA recognition. We here name this complex Craspase (CRISPR-guided Caspase).

The TPR-CHAT is a predicted protease from the caspase family whose members typically catalyze the hydrolysis of specific peptide bonds of target proteins [235], many of which remain 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 [209]. To test whether we could trigger cell death with the Craspase complex, we co-expressed 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. S7**), 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.

Its physical association with TPR-CHAT raises the possibility of caspase-like activity guided by crRNA to reach viral immunity. Although the specific RNA cleavage capability of gRAMP-crRNA could contribute to direct antiviral defense, the primary role of tar-

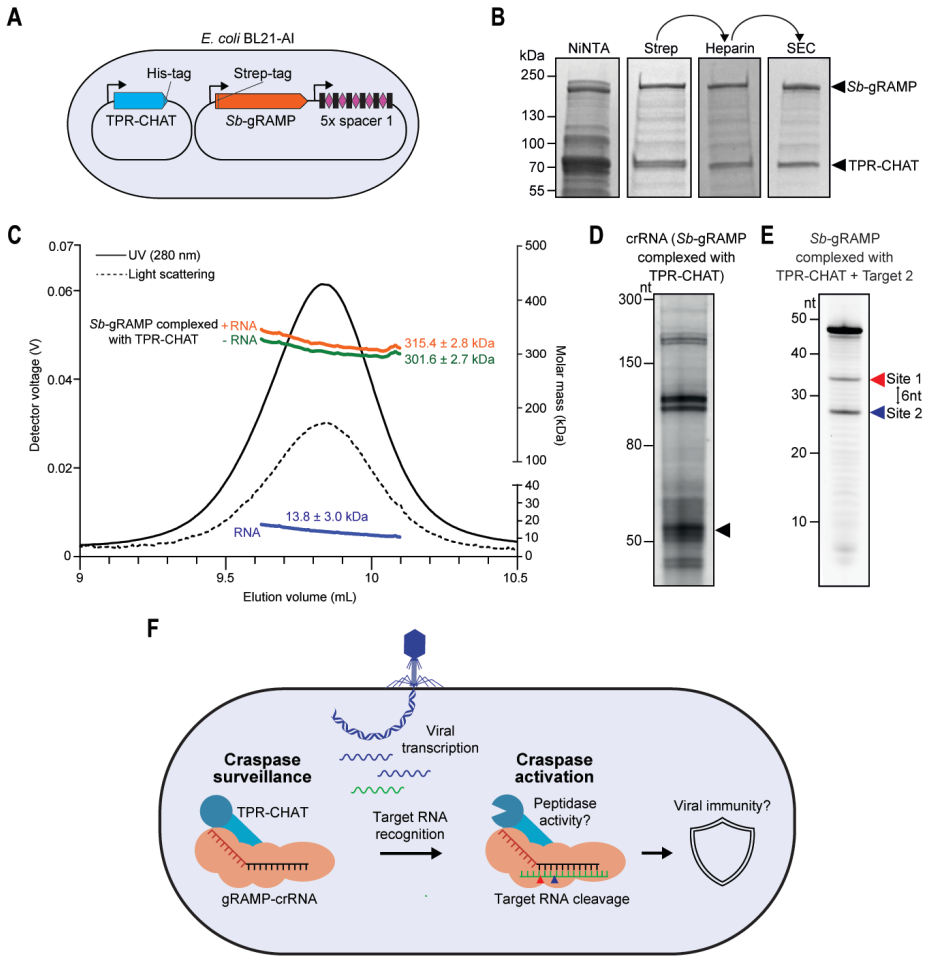


Figure 1. *Sb*-gRAMP-crRNA and TPR-CHAT form the stable Craspase complex. (**A**) Schematic of the expression system used for overexpression of TPR-CHAT and *Sb*-gRAMP-crRNA1. (**B**) SDS-PAGE gels after the subsequent purification steps of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. (**C**) SEC-MALS light scattering UV_{280} absorbance profiles of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. Molar masses are derived from the peak half-height region (9.6-10.1 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 type III-E CRISPR-Cas during phage infection.

get 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 [38, 68]. 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 sec-

ond messengers, allosterically induces peptidase activity upon target RNA recognition to elicit a specific immune response (**Fig. 1F**).

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 [28, 43, 206]. As a protease and its target are often co-localized in the genome [107, 209], we tested Craspase protease activity against these proteins in co-expression experiments (**Fig. 2A**). Expression of full-length Csx30 was strongly reduced in the presence of target bound Craspase, whereas full-length RpoE and Csx31 levels were unaffected (**Fig. 2B**). This effect was alleviated when Craspase carried inactivated cysteine-histidine residues (H585A and C627A) (**Fig. 2B**), suggesting that Craspase possesses proteolytic activity against Csx30. This observation was confirmed in vitro, where purified Craspase processes Csx30 into two distinct fragments (**Fig. 2C; table S3**), 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 (**fig. S8A-B**). Cleavage by Craspase after a leucine residue is consistent with mass spectrometry (**fig. S8A**). Proteolytic digestion could only be observed in the presence of target RNA with non-matching PFS, whereas no cleavage fragments accumulated with non-target RNA or target RNA with matching PFS (**Fig. 2C-D**). As Craspase cleaves bound RNA only under bivalent cation conditions [206], 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 to magnesium rich conditions (**Fig. 2E**), suggesting that target RNA cleavage switches off the peptidase. This is further supported by the finding that the peptidase activity of a nuclease-dead variant of Craspase is not impaired in the presence of magnesium ions (**Fig. 2E**), rendering Craspase R294A D698A a 'stay-on' variant. Binding of a complementary ssDNA, which is not cleaved by Craspase [28, 206], does not activate the peptidase (**Fig. 2E**).

These findings combined support a model (**Fig. 2F**) 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 small fragment. Due to 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. S8C**). However, based on 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.

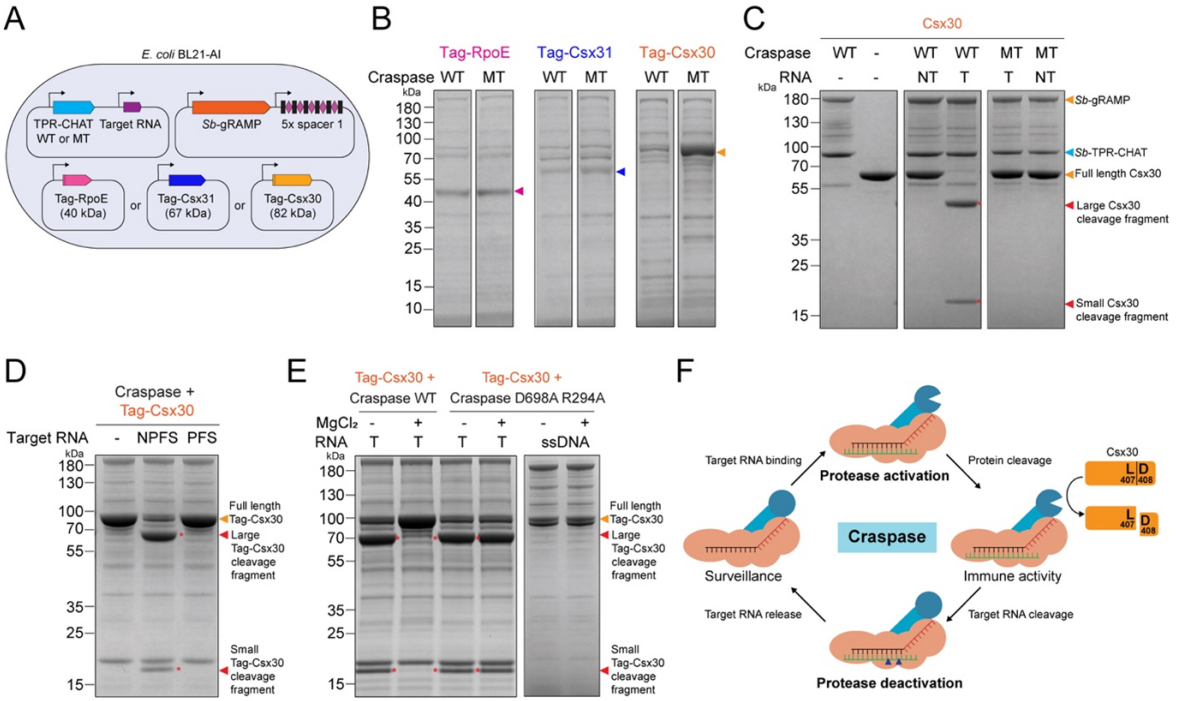


Figure 2. Craspase proteolytically cleaves Csx30 in an RNA-dependent manner. (**A**) Genetic context for RpoE, Csx31 and Csx30 co-expression with Craspase wild-type (WT) or mutant (MT; H585A C627A) and a target RNA in *E. coli* BL21-AI. (**B**) Protein gel showing the eluted protein content from Streptavidin 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 full length protein. (**C**) Protein gels after Craspase WT or Craspase MT (H585A C627A) incubation with Csx30 in the presence of (non-)target RNA. Cleavage products are indicated with a red asterisk. (**D**) Protein gel after Craspase WT incubation with target RNA containing either a non-matching PFS (NPFS) or matching PFS (PFS). (**E**) Left: protein gel after incubation of Tag-Csx30 with target RNA and Craspase WT or Craspase D698A R294A, with or without prior incubation with $MgCl_2$. Right: protein gel after incubation of Tag-Csx30 with target ssDNA and Craspase D698A R294A. (**F**) Model for Craspase functionality. Once unbound Craspase has bound a target RNA, the peptidase is activated. This results in proteolytic cleavage of Csx30 between L407 and D408. Target RNA cleavage by Craspase shuts off the peptidase.

Analogous to Cas10 activation in other type III effectors, Craspase only turns on the protease activity in response to non-self RNA targets, whereas it does not differentiate self and non-self 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? Due to the lack of homology to known proteins, it is difficult to infer the physiological function of Csx30 with confidence. Based on the AlphaFold [236] predicted structure, we speculate that proteolysis may relieve a physical sequestration or trigger a conformational change in Csx30, converting it to the active form (**fig. S8C-D**). An analogous scenario was described for bacterial gasdermin, which only induced its anti-viral effect after site-specific cleavage by TPR-CHAT [209]. Potential involvement of 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 *Candidatus* "Scalindua brodae" [237]. Alternative model organisms may be needed for future functional dissections. 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.

4.2. MATERIALS AND METHODS

PLASMID CLONING

Similar to *Plasmid cloning* in the Materials and Methods section of **Chapter 3**. Additional plasmids and primers used in this study are described in [206, 233] and ordered DNA sequences are listed in [206].

For the expression of *Candidatus* "Scalindua brodae" TPR-CHAT, a coding sequence corresponding to an *E. coli* codon-optimized TPR-CHAT protein variant containing a C-terminal His-tag was designed and ordered as a gBlock (Integrated DNA Technologies) and cloned downstream of the LacI repressed T7 promoter on the plasmid pACYC Duet-1 to yield pTPR-CHAT. pTPR-CHAT H585A C627A was constructed via insertion of a gBlock containing the desired mutations into pTPR-CHAT.

To construct pTag-Csx30, pTag-Csx31 and pTag-RpoE, a sequence corresponding to an *E. coli* codon-optimized Csx30, Csx31 and RpoE proteins were designed, ordered (Life Technologies Europe BV) and cloned into the plasmid pGFPuv (Clontech) using NEBuilder HiFi DNA Assembly (New England Biolabs), combined with a cloned fragment encoding the LacI repressed T7 promoter followed by a N-terminal Twin-Strep-Tag II and a SUMO-tag. For the construction of pTag-Csx30 mutants, primers including the desired mutation were used with pTag-Csx30 as backbone for PCR and subsequent HiFi assembly. For the construction of pTag-Csx30 mutants (pTag-Csx30 T404A, pTag-Csx30 I405A, pTag-Csx30 D406A, pTag-Csx30 L407A, pTag-Csx30 D408A, pTag-Csx30 D409A, pTag-Csx30 Q410A, pTag-Csx30 Q411A), primers including the desired mutation were used with pTag-Csx30 as backbone for PCR and subsequent HiFi assembly. For generation of non-tagged Csx30, non-tagged Csx31 and non-tagged RpoE, the coding sequence corresponding to *E. coli* codon-optimized Csx30, Csx31 and RpoE proteins were cloned into the plasmid pGFPuv using HiFi assembly.

To construct pTPR-CHAT-target1, pTPR-CHAT was enriched with a sequence encoding for target RNA complementary to crRNA1 under control of the IPTG inducible T7 promoter, using NEBuilder HiFi DNA Assembly (New England Biolabs).

PROTEIN OVEREXPRESSION FOR PURIFICATION

Electrocompetent cells *E. coli* BL21-AI were transformed with the desired plasmids (TPR-CHAT: pTPR-CHAT/pTPR-CHAT H585A C627A. Craspase: pGRAMP-CRISPR-1/pGRAMP-CRISPR-1-target-1 and pTPR-CHAT/pTPR-CHAT H585A C627A) followed by overnight growth at 37 °C on LB-agar plates containing selection media (50 μg/mL spectinomycin and/or 25 μg/mL chloramphenicol). Colonies were streaked from the plate, resuspended in 100 mL LB containing selection media and grown for ~3 hours before inoculation in 8 L (for overexpression of gRAMP complexed with TPR-CHAT (Craspase)) or 4 L (for overexpression of TPR-CHAT) LB medium containing selection media. Cultures were grown at 37 °C and 150 rpm until the cultures reached exponential phase (OD600 0.3-0.5). The grown cultures were incubated on ice for 1 hour and protein expression was induced with a final concentration of 0.2% L-arabinose and 1 mM IPTG, followed by overnight incubation at 20 °C and 150 rpm. The overnight grown cultures were harvested by centrifugation at 6000 rpm for 30 minutes. The supernatant was discarded and the pellets were resuspended in PBS (50 mL PBS/initial 1 L culture) and harvested by centrifugation

(30 minutes, 3900 rpm 4 °C). The supernatant was discarded, and the pellets stored at -80°C until further use.

The purification of Craspase was done similar to purification of gRAMP-crRNA in *Purification of Sb-gRAMP-crRNA and Sb-gRAMP variants* in the Materials and Methods section of **Chapter 3**. No cOmplete™ EDTA-free Protease Inhibitor Cocktail was used.

For the purification of TPR-CHAT, cell pellets were unfrozen and resuspended in 200 mL of ice-cooled lysis/wash buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, 25 mM imidazole, pH 7.5). Cell lysis was performed three times in a cooled French press (1 kbar). The lysate was centrifuged at 16,000 rpm for 30 minutes at 4 °C. The supernatant was filtered through a 0.45 μm syringe filter and loaded on a disposable 20 mL column (Bio-Rad) for gravity-flow affinity chromatography containing HIS-Select Nickel Affinity Gel (500 μL/50 mL lysate) (Sigma-aldrich) that was washed with 20 mL of ice-cold lysis/wash buffer. The loaded column was washed with 15 mL of ice-cold lysis/wash buffer. The protein was eluted with ice-cold elution buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, 250 mM imidazole, pH 7.5). Pooled fractions were concentrated, snap frozen in liquid nitrogen and stored at -80 °C until further use.

RNA EXTRACTION

Similar to *RNA extraction* in **Chapter 3**. 50 μL of purified Craspase (~0.6 mg/mL) was used.

SDS-PAGE ANALYSIS

Similar to *SDS-PAGE analysis* in the Materials and Methods section of **Chapter 3**.

TARGET RNA CLEAVAGE

Similar to *Target RNA cleavage* in the Materials and Methods section of **Chapter 3**.

SEC-MALS

Similar to *SEC-MALS* in the Materials and Methods section of **Chapter 3**. The Craspase sample for SEC-MALS was prepared by pooling the SEC elutions corresponding to the protein peaks (9.5-10.5 mL). Samples were injected at final concentrations of 0.6 mg/mL. Molecular mass distribution and concentrations of chromatogram peaks were calculated based on the light scattering signal and the refractive index, respectively, derived from peak half-height regions (9.62-10.10 mL). Protein UV extinction coefficient of 1.225 mL mg^{-1} cm^{-1} and RNA UV extinction coefficient of 15.3914 mL mg^{-1} cm^{-1} were used. The UV extinction coefficient was calculated using 55 nt ssRNA of crRNA1 (AAACAAGA-GAAGGACUUAAUGUCACGGUACCCAAUUUUCUGCCCCGGACUCCACG) with 5′ phosphate.

IN-GEL PROTEOLYTIC DIGESTION AND PROTEIN IDENTIFICATION

To identify the purified proteins, SEC purified Craspase elutions were analyzed using SDS-PAGE (**fig. S4B**, elution 11), whereupon the bands of interest (upper band, ~214 kDa and lower band, ~83 kDa) were cut from the gel and minced into small pieces using a sterile scalpel. Gel pieces were destained using Coomassie destaining solution (100

mM ammonium bicarbonate in 40% acetonitrile) for 15 minutes at 37 °C and shaking at 300 rpm. A reduction was performed using 200 μL of reducing agent (10 mM dithiothreitol) and incubation for 30 minutes at 56 °C. After removal of the reducing agent and cooling down of the gel pieces to room temperature, 200 μL of alkylating solution (55 mM iodoacetamide) was added and incubated in the dark at room temperature for 30 minutes. After removal of the alkylation solution, 200 μL of Coomassie destaining solution was added and incubated for 5 minutes at room temperature with 300 rpm shaking. The solution was discarded and 200 μL of acetonitrile (100%) was added and incubated for approximately 10 minutes at room temperature until the gel pieces were dehydrated. Proteolytic in-gel digestion was initiated by adding 2 μL of a 100 ng/μL trypsin solution (Promega, sequencing grade) dissolved in 100 mM ammonium bicarbonate buffer. After 5 minutes, an additional 50 μL of 100 mM ammonium bicarbonate buffer was added to ensure full coverage of the gel pieces. Samples were incubated overnight at 37 °C with 300 rpm shaking. The next day, samples were spun down and 150 μL of extraction solution (70% acetonitrile, 5% formic acid) was added to the supernatant and incubated for 15 minutes at 37 °C with 300 rpm shaking. The supernatant was collected, 100 μL of acetonitrile (100%) was added and incubated for 15 minutes at 37 °C with 300 rpm shaking. Next, 100 μL of 10:90 (v/v) acetonitrile:water was added to the samples and incubated for 15 minutes at 37 °C with 300 rpm shaking. The supernatant was collected and dried using a SpeedVac concentrator (Thermo Fisher Scientific).

Before spectrometry analysis, the dried samples were resuspended in 15 μL resuspension buffer (3% acetonitrile, 0.1% formic acid) under careful vortexing. An aliquot corresponding to approximately 100 ng protein digest was analyzed using a nanoLC (EASY 1200, Thermo Fisher Scientific) coupled online to a QE plus Orbitrap mass spectrometer (Thermo Fisher Scientific), which was operated in data-dependent acquisition (DDA) mode, acquiring MS2 spectra of the top 10 signals [238]. Mass spectrometric raw data were analyzed using PEAKS Studio X (Bioinformatics Solutions Inc., Canada) by comparing the obtained peptide fragmentation spectra to a constructed database of *E. coli* BL21(DE3) (retrieved from the UniProtKB database, www.uniprot.org) merged with the gRAMP and TPR-CHAT protein sequences, thereby considering 20 ppm parent mass error, 0.2 Da fragment mass error, 2 missed cleavages, carbamidomethylation as fixed, and oxidation and deamidation as variable peptide modifications. Peptide matches and protein identification were filtered for 1% false discovery rate. Protein identification with >1 unique peptide sequences were considered as confident matches.

GROWTH EXPERIMENTS

Electrocompetent *E. coli* BL21-AI cells were transformed by electroporation with the desired plasmids. Individual colonies (of pGRAMP-CRISPR-1, pGRAMP-CRISPR-1-target-1, pTPR-CHAT, pTPR-CHAT H585A C627A, or combinations) were inoculated in 10 mL of LB supplemented with appropriate antibiotic(s) and incubated overnight at 37 °C and 180 rpm. Overnight cultures were diluted to OD600 of 0.05 after which 200 μl was plated in a 96-well plate in duplicate. The cultured plate was incubated at 37 °C in a Synergy™ H1 microplate reader (BioTek) for 1.5 hours. Protein expression was induced with a final concentration of 0.2% L-arabinose and 1 mM IPTG, followed by incubation at 20 °C with shaking. OD600 was measured every 10 minutes.

To verify target RNA expression, two cultures per condition (from three overnight colonies) were grown in LB supplemented with appropriate antibiotic(s) at 37 °C and 180 rpm. Upon reaching OD600 of approximately 0.45, expression was induced for 3 hours using 1 mM IPTG and 0.2% L-arabinose. Total RNA was extracted with the mirVana RNA Isolation Kit (Thermo Fisher Scientific) using the total RNA protocol, with acidic phenol (Invitrogen) according to manufacturer's instructions. Contaminant DNA was removed using TURBO DNA-free kit (Thermo Fisher Scientific) according to manufacturer's instructions. ~100 ng of treated RNA was reverse transcribed using the GoTaq 2-Step RT-qPCR system (Promega) and random primers according to manufacturer's instructions. Control reactions were prepared without reverse transcriptase. The resultant cDNA was used in PCR reactions with GoTaq qPCR Master Mix (Promega) and primers amplifying the target RNA (BN3596 and BN3599; 137 bp), spectinomycin resistance gene (aminoglycoside adenylyltransferase; aadA) RNA (BN3642 and BN3643; 146 bp) as internal plasmid RNA control, or caseinolytic peptidase B (clpB) RNA (BN3649 and BN3650; 157 bp) as internal genomic RNA control, at 95 °C for 2 minutes followed by 25 cycles of denaturation at 95 °C for 15 seconds and annealing and extension at 60 °C for 1 minute. PCR reactions were run on 2% agarose gels and stained with SYBR Safe (Thermo Fisher Scientific) for visualization of nucleic acids.

PURIFICATION OF CSX30, CSX31 AND RPOE CO-EXPRESSED WITH CRASPASE

Plasmid pGRAMP-CRISPR-1, pTPR-CHAT and one of the following three (pTag-RpoE, pTag-Csx31 or pTag-Csx30) were transformed in electrocompetent *E. coli* BL21-AI cells and grown overnight on selection media (100 μg/mL spectinomycin, 25 μg/mL chloramphenicol, 100 μg/mL ampicillin). Colonies were streaked from the plate and grown in 200 mL LB medium containing antibiotics (100 μg/mL spectinomycin, 25 μg/mL chloramphenicol, 100 μg/mL ampicillin) in baffled flasks at 37 °C and 150 rpm until they reached exponential phase (OD600 of 0.3-0.5). The grown cultures were incubated on ice for 1 hour and protein overexpression was induced with a final concentration of 0.2% L-arabinose and 0.5 mM IPTG followed by overnight incubation at 20 °C and 150 rpm. Cells were collected by centrifugation at 16000 rpm at 4 °C. The supernatant was discarded and the pellets were resuspended in PBS (50 mL PBS/initial 1 L culture) and harvested by centrifugation (30 minutes, 3900 rpm 4 °C). The supernatant was discarded, and the pellets stored at -80 °C until further use. Bacterial cell pellets of each 200 mL culture were resuspended in 10 mL of ice-cold lysis buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5). The cells were lysed by sonication (2 minutes, 30 sec on 30 sec off, amplitude 30%). The lysate was centrifuged at 16,000 rpm for 30 minutes at 4 °C, and the supernatant was filtered through a 0.45 μm syringe filter and loaded onto 0.25 mL (column bed volume) of pre-equilibrated Strep-Tactin ®XT Affinity Resin (IBA Lifesciences GmbH).The loaded resin was washed with with 10 column volumes of ice-cold wash buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5) and eluted in 750 μL of elution buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, 50 mM Biotin, pH 7.5).

PURIFICATION OF CSX30 WILD-TYPE AND CSX30 MUTANTS FOR CLEAVAGE EXPERIMENTS

The plasmid pTag-Csx30 was transformed in electrocompetent *E. coli* BL21-AI cells and grown overnight on selection media (100 μg/mL ampicillin). The colonies were streaked from the plate and grown in 4 L or 8 L of LB medium containing antibiotics (100 μg/mL ampicillin) in baffled flasks at 37 °C and 150 rpm until they reached exponential phase (OD600 of 0.3-0.5). The grown cultures were incubated on ice for 1 hour and protein overexpression was induced with a final concentration of 0.2% L-arabinose and 0.5 mM IPTG followed by overnight incubation at 20 °C and 150 rpm. Cells were collected by centrifugation at 16000 rpm at 4 °C. The supernatant was discarded and the pellets were resuspended in PBS (50 mL PBS/initial 1 L culture) and harvested by centrifugation (30 minutes, 3900 rpm 4 °C). The supernatant was discarded, and the pellets stored at -80 °C until further use. Bacterial cell pellets of were resuspended in ice-cold lysis buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5), 50 mL of buffer/1L initial culture. 1 tablet of cOmplete™ EDTA-free Protease Inhibitor Cocktail was added per 50 mL resuspended pellet. The cells were lysed with 3 runs at 1000 bar in a cooled French press. The lysate was centrifuged at 16,000 rpm for 30 minutes at 4 °C, and the supernatant was filtered through a 0.45 μm syringe filter and loaded onto pre-equilibrated Strep-Tactin ®XT Affinity Resin (IBA Lifesciences GmbH), 1.5 mL column bed volume/4L initial volume.The loaded resin was washed with 10 column volumes of ice-cold wash buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5) and eluted in 750 μL of elution buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, 50 mM Biotin, pH 7.5).

For removal of the Csx30 purification Strep-SUMO tag, Tobacco Etch Virus (TEV) protease (Sigma Aldrich; T4455) was added for overnight incubation at 4 °C, followed by affinity chromatography containing HIS-Select Nickel Affinity Gel (Sigma-aldrich) to remove the TEV protease. The collected flow-through was subjected to size exclusion chromatography using Superdex 200 Increase 10/300 GL (Cytiva) column equilibrated with running buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5) with 0.3 mL/min flow rate using running buffer as mobile phase. Pooled fractions were concentrated, flash frozen in liquid nitrogen and stored at -80 °C until further use.

IN VITRO TARGET RNA AND NON-TARGET RNA GENERATION

gBlocks containing the T7 promoter and target RNA (complementary to CRISPR1) or non-target RNA (not complementary to CRISPR1) were synthesized (IDT) and PCR amplified with 5′-TCGATCAGAGCGCTCTTACG and 5′- GGTCCAGTTCAAACACTCCC. ~500 ng of purified PCR fragment was in vitro transcribed overnight using HiScribe^{TM} T7 High Yield RNA Synthesis Kit (NEB) and subsequently treated with DNase I according to the manufacturer's protocol. For RNA extraction, acidic phenol (pH 4.5, phenol:chloroform = 5:1, Invitrogen) was added to the sample in a 1:1 ratio, vortexed for 1 minute and centrifuged for 10 minutes at 13,200 rpm at room temperature. The aqueous phase was collected and subjected to RNA precipitation (20 μL 3M NaAcetate and 500 μL 100% ethanol per 200 μL of sample) for 1 hour at -20 °C. Samples were centrifuged at 13,200 rpm at 4 °C for 2 hours, washed twice with ice-cold 70% ethanol and centrifuged at 13,200 rpm at 4 °C for 10 minutes. The pellet was dried in a SpeedVac concentrator (Thermo Fisher Sci-

entific) for 30 minutes at 60 °C and resuspended in RNA grade water. The four variants of target RNA were mixed in an equimolar ratio before usage in Csx30 protein cleavage reactions.

IN VITRO CSX30 CLEAVAGE REACTIONS

Csx30 cleavage reactions were performed in 10 μL reaction volume, containing purified 2225 nM Craspase (WT, MT (TPR-CHAT H585A C627A) or MT (D698A R294A), without SUMO tag (**Fig. 2C**) or with SUMO tag (**Fig. 2D**, **Fig. 2E**), 5 μM (for incubation with Craspase D698A R294A) or 11 μM (for incubation with target RNA containing matching or non-matching PFS) DualStrep-SUMO-TEV-Csx30 protein (**Fig. 2D**, **Fig. 2E**) or 4 μM tag-less Csx30 protein (**Fig. 2C**), 900 ng in vitro generated target RNA or non-target RNA or 3 μM target RNA containing matching or non-matching PFS, 100 mM Tris, 150 mM NaCl, 10 mM DTT, 0 mM or 2 mM $MgCl_2$ (**Fig. 2C** reactions contain 2 mM $MgCl_2$, **Fig. 2E** was incubated with 90 ng of target RNA and 2 mM $MgCl_2$ for 2 hours at 37 °C prior to addition of 11 μM DualStrep-SUMO-TEV-Csx30 protein). Reactions were run for 1 hour at 37 °C. Afterwards, the reactions were supplemented with 10 μL MilliQ water, 5 μL of 5X Laemmli buffer (375 mM Tris-HCl, 9% SDS, 50% glycerol, 0.03% bromophenol blue) and 2.5 μL 1 M DTT, and incubated at 95 °C for 10 minutes before loading on a 4-20% surePAGE^{TM} Bis-Tris protein gel (GenScript). Gels were run in 1X MOPS buffer (GenScript) at 200 V for 30-45 minutes, washed in MilliQ water and stained for at least 2 hours with BioSafe Coomassie G-250 stain (Bio-Rad) under continue shaking. Gels were washed in MilliQ water for 4 hours before imaging.

MASS SPECTROSCOPY TO CONFIRM THE IDENTITY OF THE CSX30 CLEAVAGE FRAGMENTS AND TO MAP THE CRASPASE CLEAVAGE SITE

To identify the Csx30 protein fragments, bands of interest (large Csx30 cleavage fragment, ~47 kDa and small Csx30 cleavage fragment, ~19 kDa) were cut from the gel and minced into small pieces using a sterile scalpel. Gel pieces were destained using destaining solution (100 mM ammonium bicarbonate in 40% acetonitrile) for 15 minutes at 37 °C and shaking at 300 rpm. A reduction was performed using 200 μL of reducing agent (10 mM dithiothreitol) and incubation for 30 minutes at 56 °C. After removal of the reducing agent and cooling down of the gel pieces to room temperature, 200 μL of alkylating solution (55 mM iodoacetamide) was added and incubated in the dark at room temperature for 30 minutes. After removal of the alkylation solution, 200 μL of Coomassie destaining solution was added and incubated for 5 minutes at room temperature with 300 rpm shaking. The solution was discarded and 200 μL of acetonitrile (100%) was added and incubated for approximately 10 minutes at room temperature until the gel pieces were dehydrated. Proteolytic in-gel digestion was initiated by adding 2 μL of a 100 ng/μL trypsin or chymotrypsin solution (Promega, sequencing grade). Samples were incubated overnight at 37 °C with 300 rpm shaking using an Eppendorf ThermoMixer. The next day, samples were spun down and 150 μL of extraction solution (70% acetonitrile, 5% formic acid) was added to the supernatant and incubated for 15 minutes at 37°C with 300 rpm shaking. The supernatant was collected, 100 μL of acetonitrile (100%) was added and incubated for 15 minutes at 37 °C with 300 rpm shaking. Next, 100 μL of 10:90 (v/v) acetonitrile:water was added to the samples and incubated for 15 minutes at 37 °C

with 300 rpm shaking. The supernatant was collected and dried using a SpeedVac concentrator (Thermo Fisher Scientific). Before spectrometry analysis, the dried samples were resuspended in 15 μL resuspension buffer (3% acetonitrile, 0.1% formic acid) under careful vortexing. An aliquot corresponding to approximately 100 ng protein digest was analyzed using a nano-LC (EASY 1200, Thermo Fisher Scientific, Germany) equipped with an Acclaim PepMap RSLC RP C18 separation column (50 μm x 150 mm, 2 μm), coupled online to a QE plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Germany).

The flow rate was maintained at 350 nL/min over a linear gradient from 5% to 30% solvent B over 38 minutes, and finally to 60% B over 15 minutes, followed by back equilibration to starting conditions. Solvent A consisted of water containing 0.1% formic acid, and solvent B consisted of 80% acetonitrile in water and 0.1% formic acid. The Orbitrap was operated in data depended acquisition mode acquiring peptide signals form 385-1250 m/z at 70 K resolution with a max IT of 100 ms and an AGC target of 3e6. The top 10 signals were isolated at a window of 2.0 m/z and fragmented using a NCE of 28. Fragments were acquired at 17 K resolution with a max IT of 75 ms and an AGC target of 2e5.

Mass spectrometric raw data were analyzed using PEAKS Studio X (Bioinformatics Solutions Inc., Canada) by comparing the obtained peptide fragmentation spectra to a constructed database (retrieved from the UniProtKB) containing the Csx30 protein sequence and relevant background proteins. The database searching was performed analyzing either for trypsin cleaved peptides, or in case of mapping the Craspase cleaved peptides for fully non-specific cleaved variants of the Csx30 protein. The search allowed for 20 ppm parent ion and 0.02 m/z fragment ion mass error, carbamidomethylation as fixed and methionine oxidation and N/Q deamidation as variable modifications. Peptide matches and protein identification were filtered for 1% false discovery rate.

COMPLEX DISSOCIATION TEST

Purified Craspase (0.1 μM) was incubated with ~6.6-fold excess of complementary ssRNA (Target 2) (0.66 μM). The mixture was incubated for 2 hours at 20 °C and then loaded into a Superdex 200 Increase 10/300 GL (Cytiva) previously equilibrated with running buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.5). Fractions were collected and analyzed on an SDS-PAGE gel.

4.3. SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURES

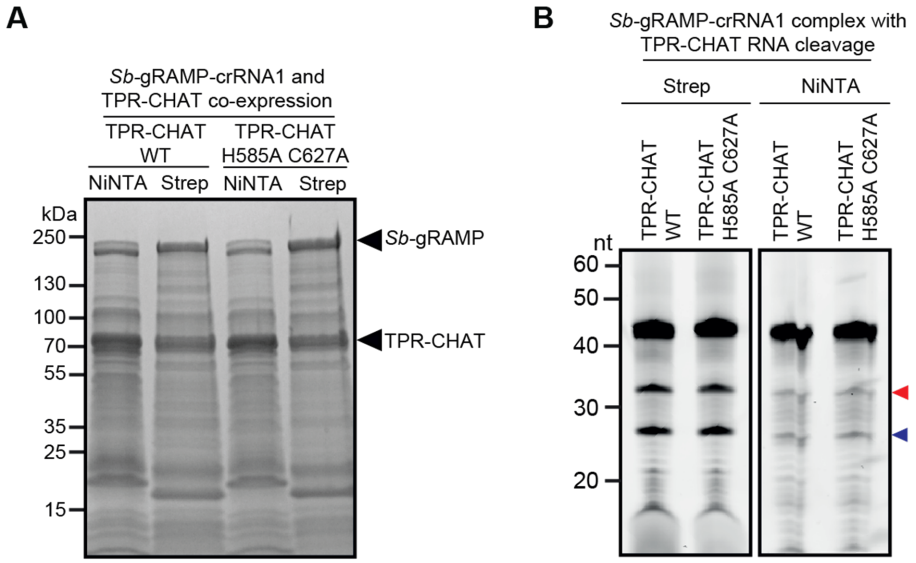


Figure S1. *Sb*-gRAMP-crRNA1 and TPR-CHAT co-expression. (**A**) Pulldown assays were performed using *Sb*-gRAMP-crRNA1 with wild-type TPR-CHAT or TPR-CHAT with inactivating mutations in the predicted protease domain (H585A and C627A), yielding similar band patterns and intensities in each condition. (**B**) The samples after pulldown were incubated with 100 nM target RNA cognate to crRNA1 (Target 2). Cleavage products are indicated with red and blue arrowheads.

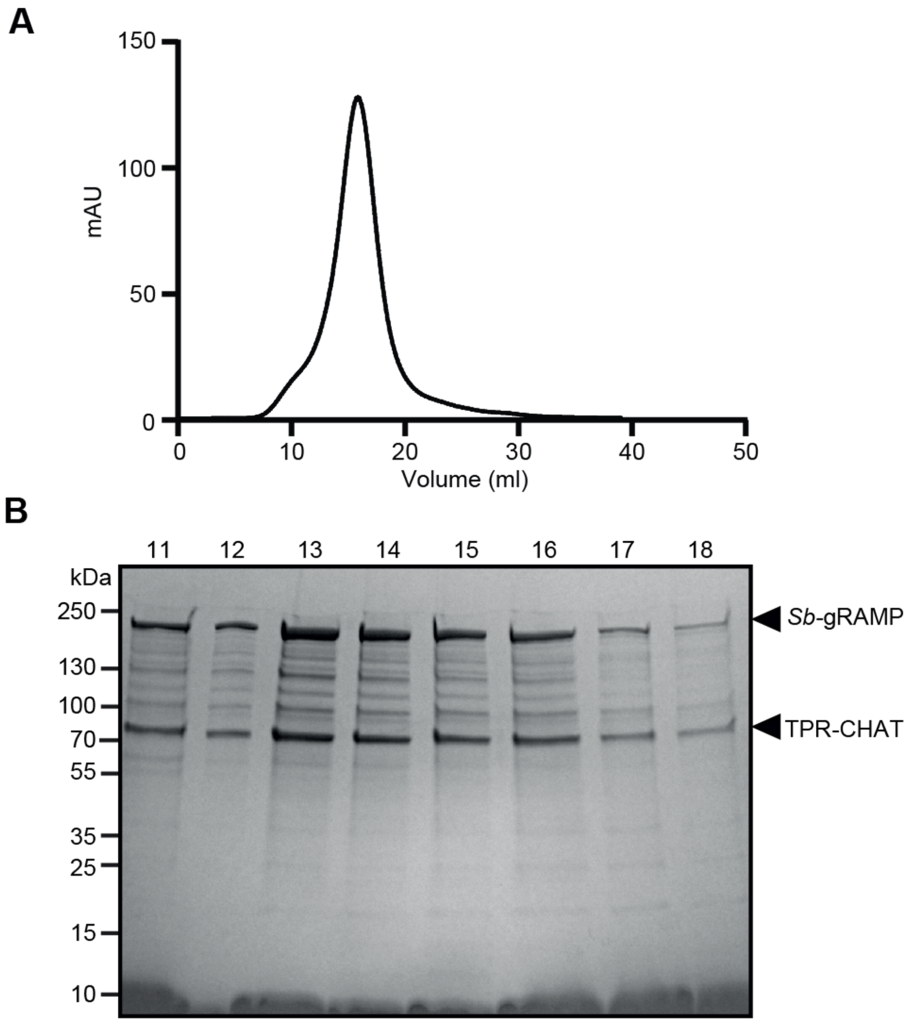


Figure S2. Heparin purification of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. (**A**) The UV_{280} absorbance profile after elution of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT from the heparin column. (**B**) Heparin elution volumes corresponding to the peak analyzed with SDS-PAGE.

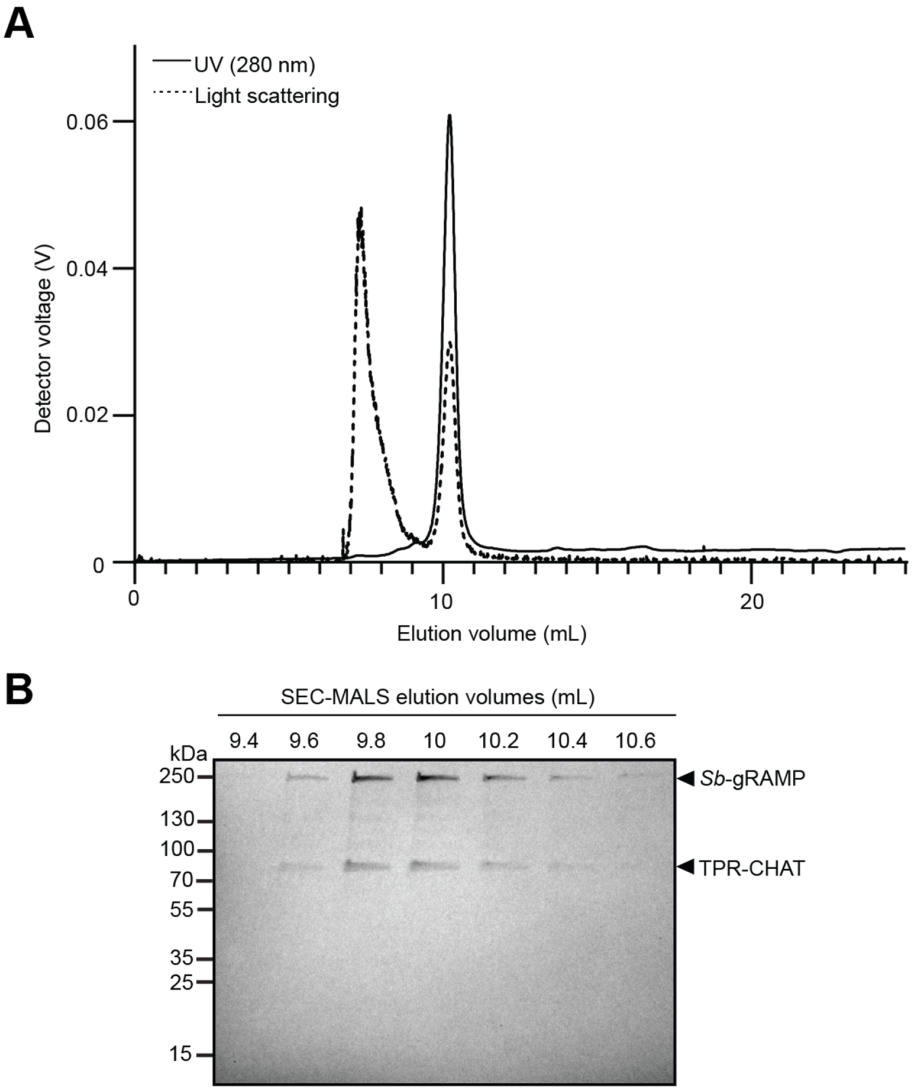


Figure S3. SEC-MALS analysis of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. (**A**) Complete SEC-MALS chromatogram showing the light scattering and UV_{280} absorbance profiles for *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. (**B**) SDS-PAGE analysis of the collected SEC-MALS elution volumes.

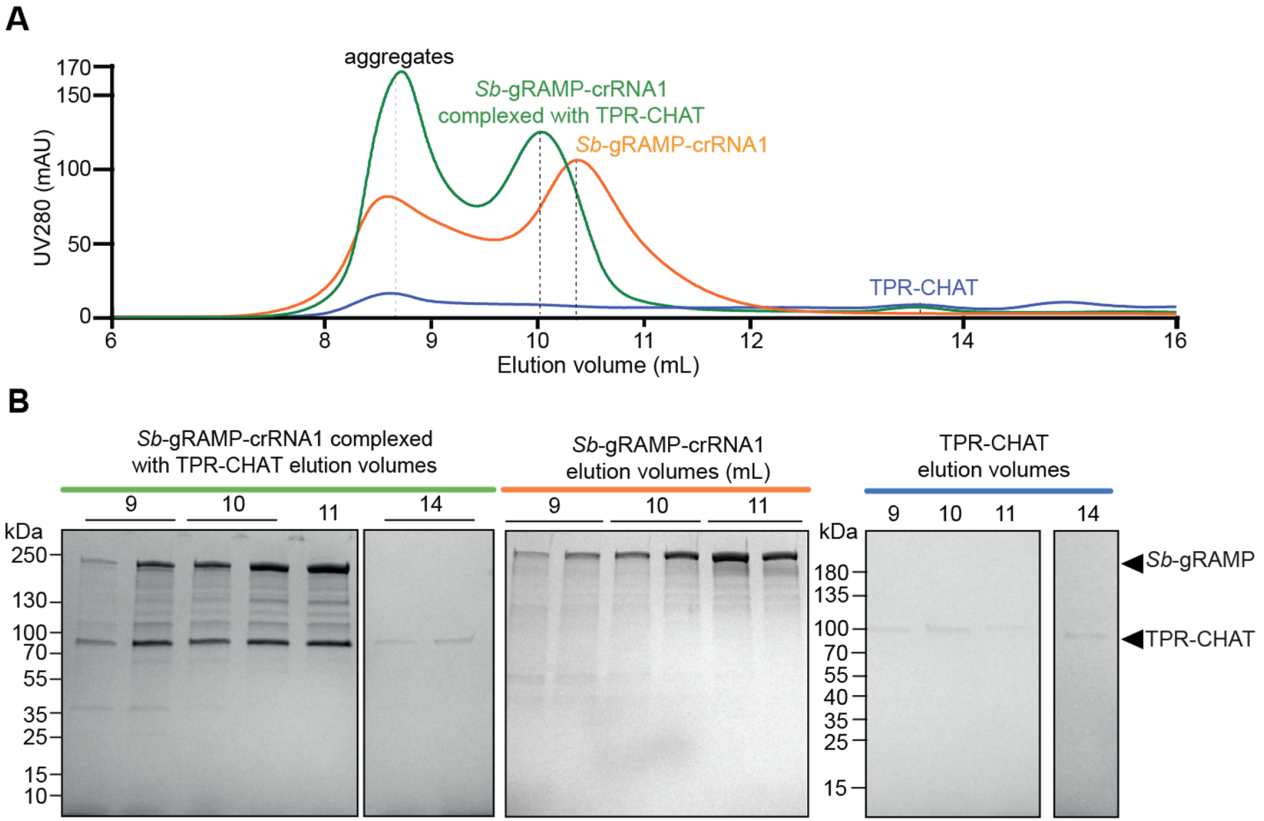


Figure S4. Size-exclusion chromatography (SEC) of *Sb*-gRAMP-crRNA1, *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT and TPR-CHAT alone. (**A**) Chromatograms showing the absorbance at 280 nm. (**B**) SDS-PAGE analysis of the elution volumes corresponding to the peaks.

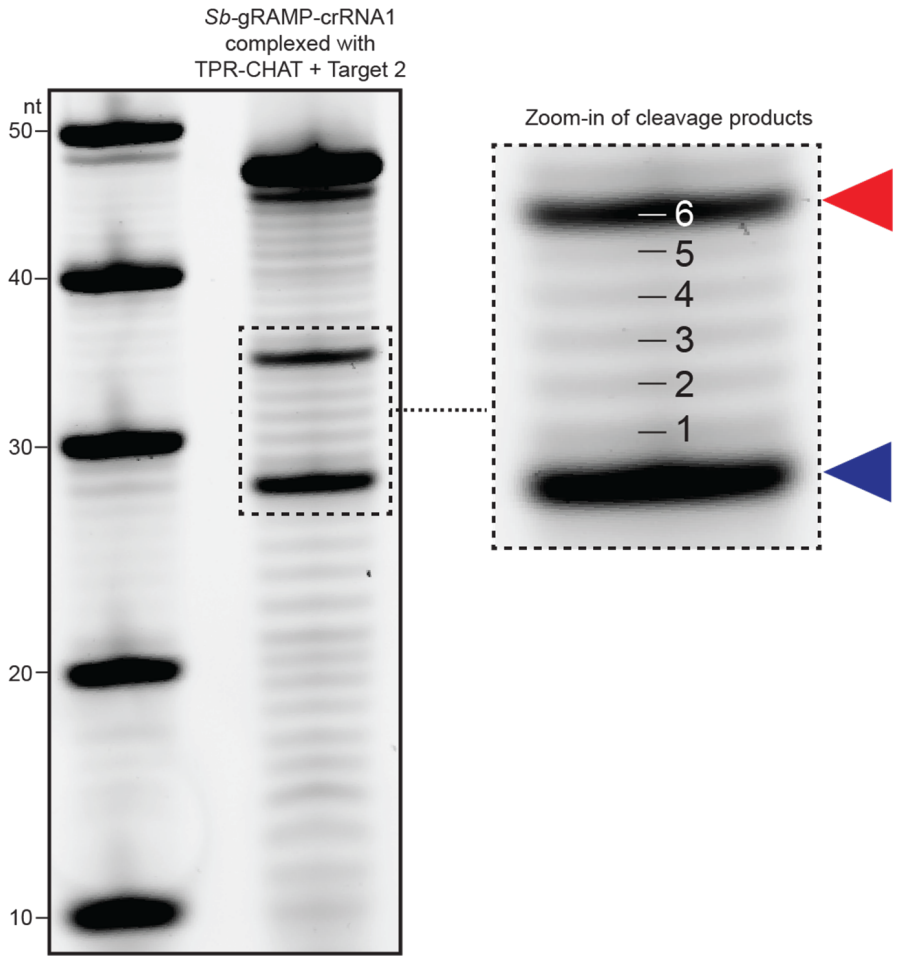


Figure S5. The cleavage products generated by *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT are 6 nt apart. Denaturing urea PAGE gel of a cleavage reaction of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT and cognate Target 2, with a zoom-in of the cleavage products (indicated with arrowheads) in which nucleotide counting steps are indicated.

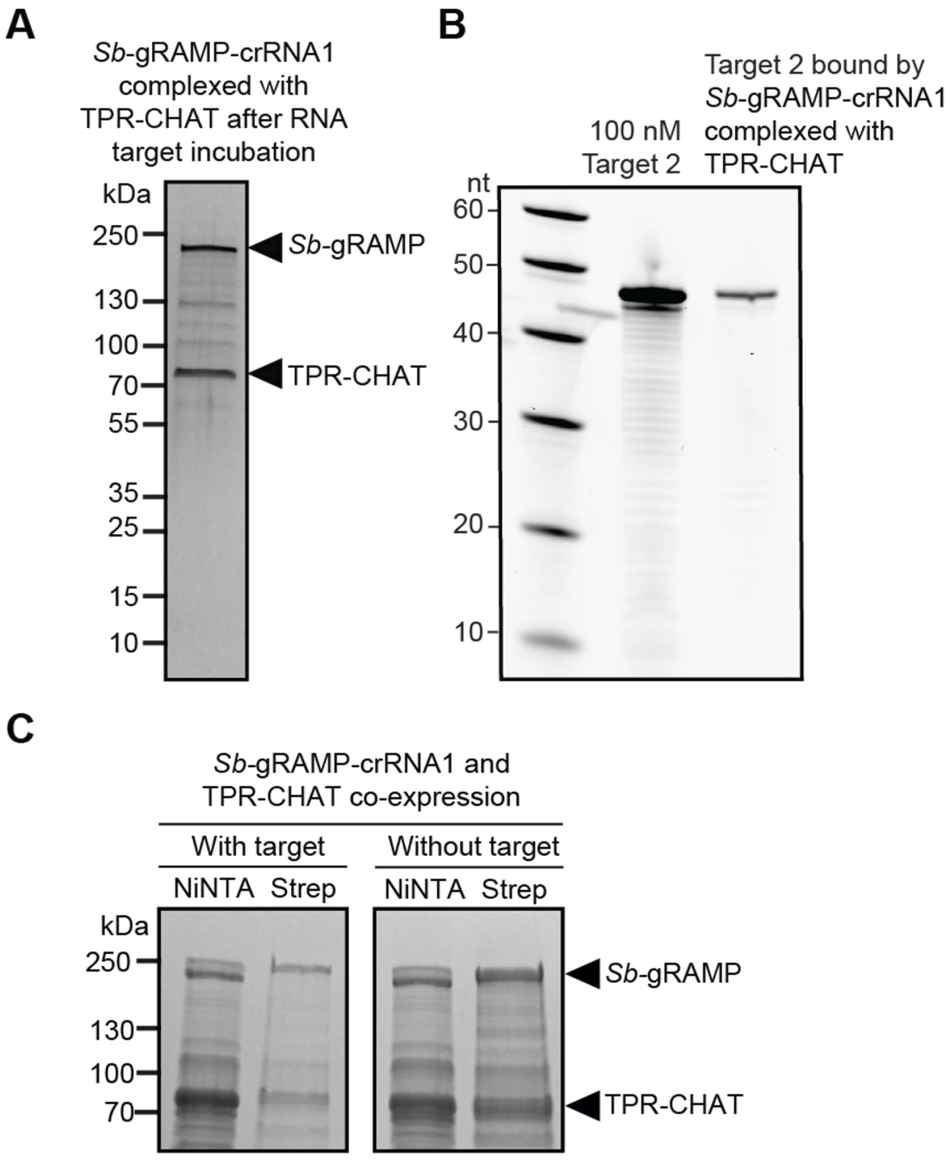


Figure S6. TPR-CHAT remains complexed with *Sb*-gRAMP-crRNA1 upon target recognition. (**A**) SDS-PAGE analysis of SEC purified *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT after incubation with cognate target RNA (Target 2) for 4 hours in the absence of $MgCl_2$. (**B**) RNA bound to SEC purified *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT, indicating successful target binding. (**C**) SDS-PAGE analysis of *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT overexpressed in the presence or absence of a target RNA.

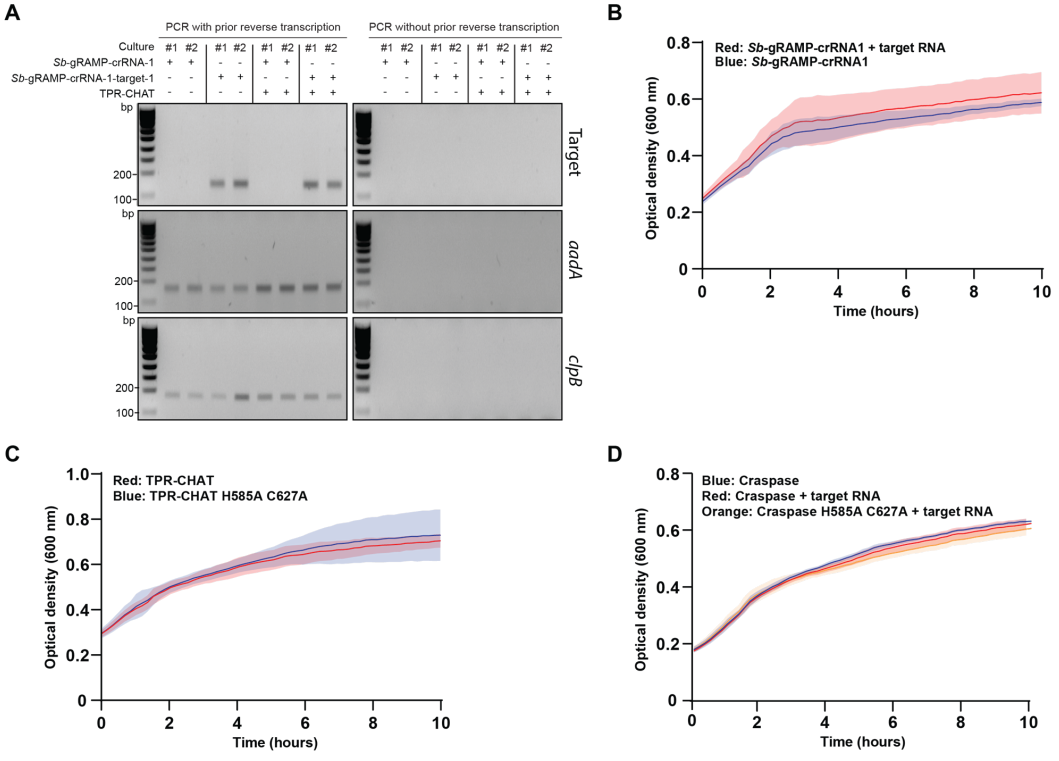


Figure S7. Growth curves of *E. coli* during type III-E expression and target RNA production. (**A**) RT-PCR reactions of RNA extracted from induced cells to verify target RNA and control RNA production. Plasmid encoded spectinomycin resistance (aminoglycoside adenylyltransferase; *aadA*) RNA or genome encoded caseinolytic peptidase B (*clpB*) were used as internal controls. Data represent biological duplicates. (**B**) OD_{600} measurements of cells expressing only *Sb*-gRAMP-crRNA1 (blue) or *Sb*-gRAMP-crRNA1 in the presence of target RNA (red), (**B**) only TPR-CHAT (red) or TPR-CHAT H585A C627A (blue) and (**C**) Craspase or Craspase with TPR-CHAT H585A C627A in the presence (red, orange) and absence (blue) of target RNA. Data represent biological triplicates with standard deviation area bands filled.

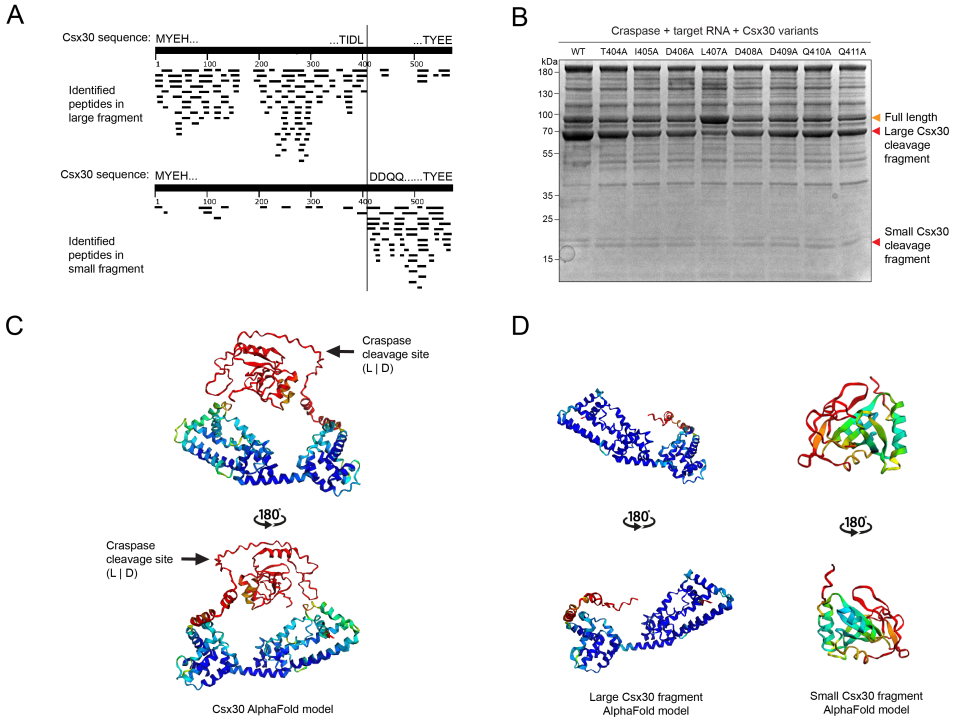


Figure S8. Csx30 processing analysis. (**A**) Mass spectrometry identified peptides in the gel slices of the large and small Csx30 fragments, respectively. Trypsin or chymotrypsin was used for the in-gel digestion of Csx30 fragments. Peptide sequence was defined without assuming cleavage site preference (listed in supplementary table S3 and S4 in [233]), and mapped onto Csx30 using BLAST [239] to reveal a non-trypsin/chymotrypsin cleavage site (TIDL | DDQQ). (**B**) Protein gel after Craspase incubation with target RNA and Csx30 containing alanine mutations of the residues in the cleavage region. (**C**) AlphaFold model of Csx30 with the Craspase cleavage site indicated. The model is colored based on the pLDDT (predicted local distance difference test) score (dark blue: very high confidence, pLDDT>90; light blue: confident, 90>pLDDT>70; yellow: low confidence, 70>pLDDT>50; orange: very low confidence, pLDDT<50). (**D**) AlphaFold models of the Csx30 fragments after processing by Craspase (dark blue: very high confidence, pLDDT>90; light blue: confident, 90>pLDDT>70; yellow: low confidence, 70>pLDDT>50; orange: very low confidence, pLDDT<50).

SUPPLEMENTARY TABLES

Number of unique peptide per protein	Number of spectra	Average mass (Da)	Description	-10 log P	Gel band
251	838	214,018	*Sb*-gRAMP	611.04	Upper
5	5	83,785	TPR-CHAT	106.64	Upper
3	3	88,946	Bifunctional aspartokinase/homoserine dehydrogenase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_4083 PE=3 SV=1	67.33	Upper
89	315	83,785	TPR-CHAT	406.92	Lower
117	175	214,018	*Sb*-gRAMP	366.34	Lower
18	23	69,115	Chaperone protein DnaK OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=dnaK PE=2 SV=1	218.39	Lower
17	22	61,158	30S ribosomal protein S1 OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_2684 PE=3 SV=1	215.08	Lower
8	8	77,101	Polyribonucleotide nucleotidyltransferase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=pnp PE=3 SV=1	140.38	Lower
4	5	28,744	4-hydroxy-tetrahydrodipicolinate reductase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=dapB PE=3 SV=1	121.47	Lower
4	4	48,532	Aspartokinase OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_4013 PE=3 SV=1	102.67	Lower
4	4	66,096	Acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=ECBD_3504 PE=3 SV=1	101.65	Lower
2	2	57,329	60 kDa chaperonin OS=Escherichia coli (strain B / BL21-DE3) OX=469008 GN=groL PE=3 SV=1	53.17	Lower

Table S1. Mass spectrometry analysis of two in-gel digestion samples of *Sb*-gRAMP complexed with TPR-CHAT. Sample derived from in-gel digestion of upper gel band (~214 kDa) or lower gel band (~83 kDa) in **fig. S4B**, elution 11.

	Experimentally calculated (*theoretical*) total molar mass (kDa)	**Experimentally calculated (*theoretical*) protein molar mass (kDa)**	**Experimentally calculated (*theoretical*) RNA molar mass (kDa)**
***Sb*-gRAMP-crRNA1 complexed with TPR-CHAT**	315.4 ± 2.8 (*315.5*)	301.6 ± 2.7 (*297.8*)	13.8 ± 3.0 (*17.7*)

Table S2. The experimentally calculated versus theoretic molar mass for *Sb*-gRAMP-crRNA1 complexed with TPR-CHAT. For the theoretical molar mass of the RNA, 55 nt ssRNA of crRNA1 (AAACAAGAGAAGGACUUAAUGUCACGGUACCCAAUUUUCUGCCCCGGACUCCACG) with 5′ phosphate was used and calculated using MolBioTools DNA calculator (www.molbiotools.com). The molar mass of TPR-CHAT is 83.8 kDa.

Protein	-10 log P	Area Sample	Number of unique peptide per protein	Csx30 cleavage fragment
Csx30	388.80	2.07E10	71	Large
sp|P0CE47|EFTU1_ECOLI	251.48	2.23E8	22	Large
sp|P0CE48|EFTU2_ECOLI	251.48	2.23E8	22	Large
sp|P0ABB4|ATPB_ECOLI	237.83	1.11E8	27	Large
sp|P0A6P9|ENO_ECOLI	207.99	1.95E7	16	Large
sp|P02943|LAMB_ECOLI	189.66	3.24E7	13	Large
sp|P0A6J5|DADA_ECOLI	189.65	2.7E7	14	Large
sp|P0ABC7|HFLK_ECOLI	185.24	1.95E7	11	Large
sp|P0ABH7|CISY_ECOLI	182.45	1.73E7	12	Large
sp|P00393|NDH_ECOLI	180.97	3.01E7	16	Large
sp|P0A6H5|HSLU_ECOLI	176.74	5.95E7	18	Large
sp|P75990|BLUF_ECOLI	169.51	1.64E7	14	Large
sp|P0AG30|RHO_ECOLI	157.13	1,00E+07	13	Large
sp|P0A6Y8|DNAK_ECOLI	155.13	7.01E6	11	Large
sp|P0A855|TOLB_ECOLI	152.39	1.46E7	8	Large
sp|P08622|DNAJ_ECOLI	152.08	8.71E6	10	Large
sp|P08200|IDH_ECOLI	149.68	1.05E7	10	Large
sp|P0A6B7|ISCS_ECOLI	149.11	2.19E7	12	Large
sp|P0AE06|ACRA_ECOLI	146.47	3.01E7	13	Large
sp|P24554|RADA_ECOLI	145.65	8.17E6	11	Large
sp|P27434|RODZ_ECOLI	143.94	9.29E6	8	Large
sp|P31979|NUOF_ECOLI	142.45	1.67E7	11	Large
sp|P68187|MALK_ECOLI	138.47	4.39E6	6	Large
sp|P0AFG6|ODO2_ECOLI	135.44	1.71E7	10	Large
sp|P75876|RLMI_ECOLI	131.00	2.15E6	6	Large
sp|P03004|DNAA_ECOLI	130.73	9.2E6	10	Large
sp|P0C8J8|GATZ_ECOLI	124.37	6.06E6	6	Large
sp|P0A836|SUCC_ECOLI	123.96	2.54E6	7	Large
sp|P30871|3PASE_ECOLI	117.03	3.05E6	7	Large
sp|P0A749|MURA_ECOLI	115.37	2.73E6	6	Large
sp|P0A6H1|CLPX_ECOLI	110.20	4.32E6	6	Large
sp|P0A847|TGT_ECOLI	108.02	2.18E6	5	Large
sp|P00370|DHE4_ECOLI	106.19	2.37E6	7	Large
sp|P00350|6PGD_ECOLI	104.18	2.38E6	6	Large
sp|P02930|TOLC_ECOLI	103.41	3.3E6	8	Large
sp|P69786|PTGCB_ECOLI	102.36	2.44E6	5	Large
sp|P76373|UDG_ECOLI	100.77	1.32E6	5	Large
sp|P37095|PEPB_ECOLI	98.71	3.2E6	5	Large
sp|P31120|GLMM_ECOLI	96.42	2.03E6	5	Large
sp|P46130|YBHC_ECOLI	93.90	4.03E6	6	Large
sp|P0A6F5|CH60_ECOLI	87.89	1.03E6	4	Large
sp|P0AE45|PAEA_ECOLI	84.21	1.37E6	5	Large
Csx30	393.50	6.03E9	49	Small
sp|P0A7W1|RS5_ECOLI	190.95	7.23E7	10	Small
sp|P02359|RS7_ECOLI	187.74	5.76E7	11	Small
sp|P0AA10|RL13_ECOLI	147.43	8.33E7	10	Small
sp|P0A7J3|RL10_ECOLI	133.86	4.7E7	7	Small
sp|P0A7R1|RL9_ECOLI	120.73	4.34E6	7	Small
sp|P02413|RL15_ECOLI	115.54	1.44E7	6	Small
sp|P0ACJ0|LRP_ECOLI	106.74	2.38E6	5	Small
sp|P0ABA0|ATPF_ECOLI	102.89	2.57E6	6	Small
sp|P0A7L3|RL20_ECOLI	101.89	8.4E6	5	Small

Table S3. Mass spectrometry results of two in-gel digestion samples of the large and small Csx30 cleavage fragments. Hits with more than four unique peptides per protein are shown.

5

STRUCTURAL INSIGHTS INTO gRAMP AND CRASPASE

"The particular successions of causes and effects (that delude us with the appearance of mechanism) disappear as infinitely small straight lines in the universal curvature of the organism in which the world itself persists."

Friedrich von Schelling

The type III-E RNA-targeting gRAMP is associated with the caspase-like TPR-CHAT protein to form Craspase. 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 sidechain-binding pocket. This mechanistic analysis defnies how RNA-guided RNA recognition controls the protease activity of Craspase.

A modified version of this chapter has been published as *Hu* & van Beljouw* et al., Science (2022)* [233].

5.1. MAIN TEXT

5.1.1. INTRODUCTION

Most CRISPR-Cas bacterial immune systems function via cleavage of nucleic acids, but there are also certain proteins within the CRISPR-Cas family that cleave proteins. The Craspase protein complex, composed of the RNA-guided gRAMP RNase and the TPR-CHAT protease, is activated by RNA whereupon it cleaves the Csx30 protein. Here we deploy cryo-electron microscopy to dissect how Craspase is structurally and mechanistically orchestrated.

5.1.2. CRYO-ELECTRON MICROSCOPY STRUCTURES OF GRAMP

To gain insights into the RNA-guided target RNA cleavage mechanisms inside gRAMP, we reconstituted *Candidatus* "*Scalindua brodae*" gRAMP (*Sb*-gRAMP) and obtained cryo-electron microscopy (cryo-EM) structures in different functional states (**Figs. 1A, fig. S1**). Consistent with previous results [206], *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 the ninth nucleotide (site 2) (**fig. S1A-C**). Single-particle three-dimensional reconstruction produced *Sb*-gRAMP RNP in four different functional states: a 3.81 Å structure of the *apo* (resting) state, 3.65 Å structure of the non-matching protospacer flanking sequence (PFS) target bound state, 3.76 Å structure of the matching PFS target bound state, and 3.62 Å structure of the post-cleavage state (**Fig. 1B-E, fig. S2-S4**).

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 [240], as a fusion of four Cas7-like domains, one Cas11-like domain, and a big insertion domain (BID), 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 [74, 213, 241, 242] in overall architecture, guide RNA display, and target RNA binding mode (**fig. S6**). The *Sb*-gRAMP backbone consists of four non-identical Cas7 domains fused together, instead of three identical Cas7 subunits in Csm (**figs. S6**). A Zn knuckle is present in each of the four Cas7s, 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. S6A-H**). On the guide RNA side, the ordered 18-nt 5′ handle of the crRNA in *Sb*-gRAMP is twice as long as in other class I CRISPR-Cas systems (**Fig. 2A-B**). The majority 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, fig. S7B-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 [240],

whereas the equivalent residues in *Sb*-gRAMP are non-catalytic (**Fig. 2C-D**). This structural difference rationalizes 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.

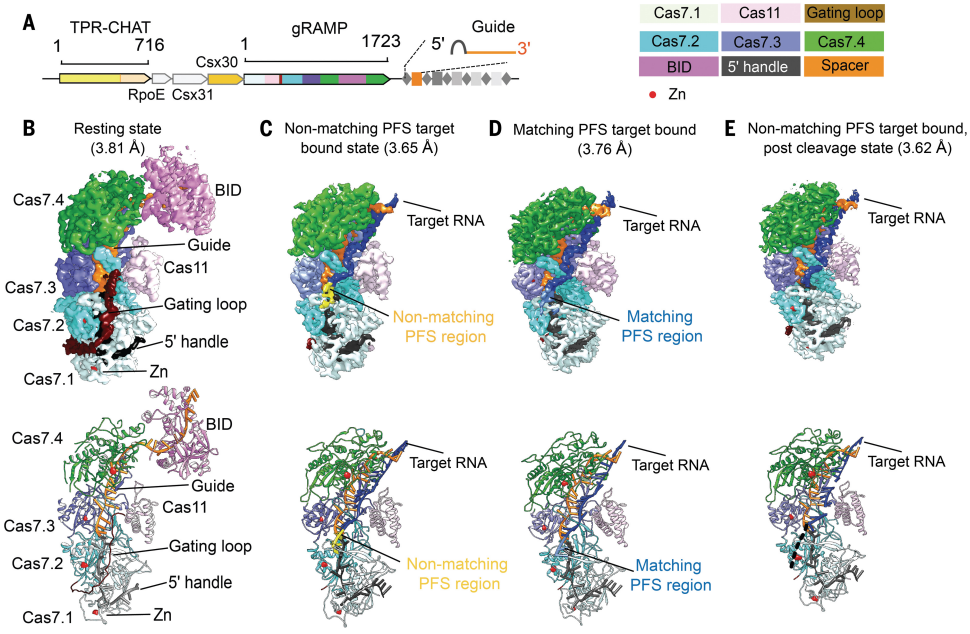


Figure 1. Structural snapshots of *Sb*-gRAMP RNP in different functional states. (**A**) Type III-E operon in *Candidatus* "Scalindua brodae" (*Sb*-gRAMP). (**B-E**) Snapshots of *Sb*-gRAMP at resting state (B), non-matching PFS RNA–bound state (C), matching PFS RNA–bound state (D), and non-matching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

Notably, the last two handle nucleotides (5′-$A_{-2}C_{-1}$-3′) are base-pairing competent because they are displayed like a guide (**Fig. 2A, fig. S6H-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 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 [240] (**Fig. 2A, Fig. 2E, fig. S6H-G**). The third segment deviates from the normal again, as 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 a β-hairpin thumb, therefore 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 [236] predicted model (**Fig. 1B**).

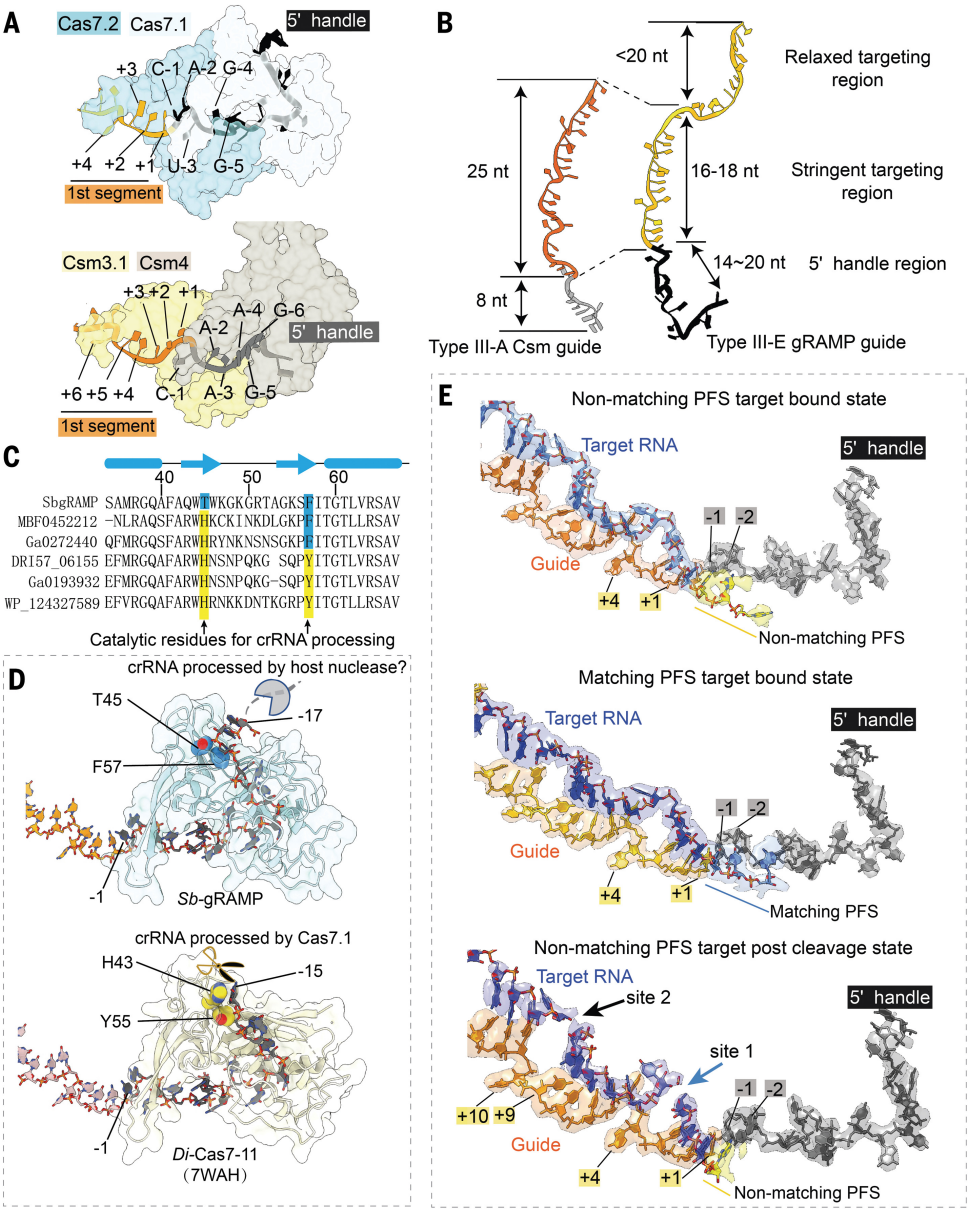


Figure 2. crRNA accommodation and target RNA recognition mechanisms by *Sb*-gRAMP. (**A** and **B**) Accommodation of the crRNA 5′ handle (A) and spacer region (B) in type III-E *Sb*-gRAMP and type III-A Csm. (**C** and **D**) Primary sequence (C) and 3D structural alignment (D) at the pre-crRNA cleavage 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 non-matching PFS RNA (top), matching PFS RNA (middle), and non-matching PFS RNA postcleavage state (bottom).

5.1.3. OFF-TARGETING PREVENTION AND RNA CLEAVAGE MECHANISMS

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-E412, here named the gating loop) has acquired important functions for RNase regulation (**Fig. 3A-B, fig. S9**). Its N-terminal portion (G375 to G397) senses RNA substrate binding and controls RNase activities. In 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 [68]. In follow-up experiments, we found *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-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 [206], and that the BID is dispensable for Cas7-11 activity in human cells [240].

Sb-gRAMP was further incubated with two kinds of RNA targets whose PFS was either matching (complementary) or non-matching with the 5′ handle in the crRNA, as complementarity in this region may be indicative of a self-target (i.e. anti-sense transcript from the CRISPR locus) and thus perhaps leads 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 outwards. Rotation of the backbone orients their 2′-OH towards the previous phosphate, forming the so-called "in-line" conformation, which 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 non-matching PFS, only three base-pairs are found between the target RNA and the spacer portion of the guide. While 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 the first target-guide segment (**Fig. 2E**). In both PFS matched and non-matched 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, 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, as RNA substrates lacking nucleotides in the PFS region were found to be cleaved efficiently [206, 240]. 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. S9D-E**). Wild-type *Sb*-gRAMP did not bind or cleave RNA that

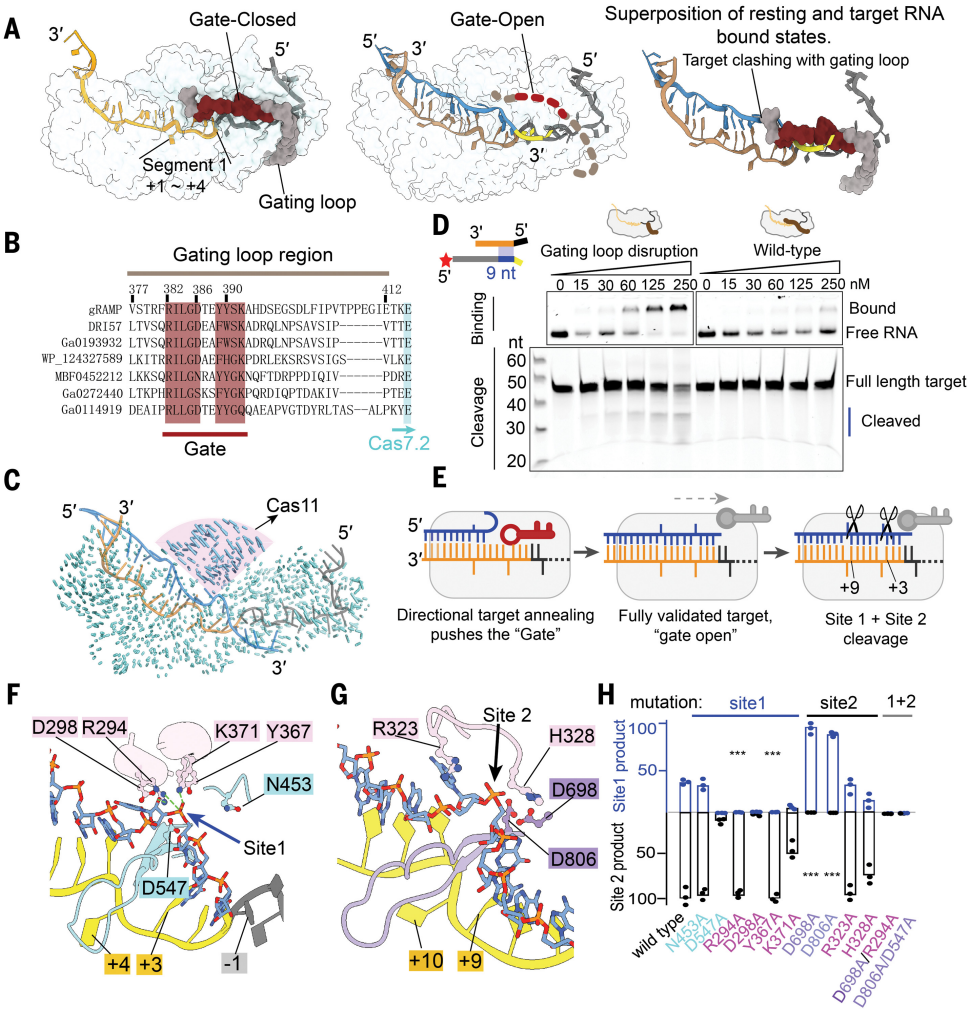


Figure 3. Target validation and cleavage mechanisms by *Sb*-gRAMP. (**A**) Models depicting the gate-closed structure in resting state (left) and the gate-open structure in the target RNA–bound state (middle). Superposition is shown in the right panel. (**B**) Sequence alignment at the gating loop region. Conserved residues are highlighted in burgundy. (**C**) Structural comparison of the resting and non-matching 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.

only base-pairs with the first 9 nt of the crRNA guide. In 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. Over-

all, 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 via crRNA-pairing in a directional fashion from the 3′ to 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**).

We further attempted to interpret the cleavage mechanism by comparing the pre- and postcleavage states (**Fig. 1, Fig. 3F-G**). EM densities suggest 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 [206, 240]. Since 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 residues that are positively charged) (**Fig. 3F-G**). In subsequent mutagenesis testing (**fig. S11B-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**). Since 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 [206] 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, 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.

5.1.4. CRASPASE ARCHITECTURE AND COMPONENT INTERFACES

To gain mechanistic insights into how the RNA-guided protease system works, we reconstituted Craspase in its *apo* state, the matching PFS-containing RNA target bound state, and the non-matching PFS-containing target bound state, and generated their corresponding cryo-EM structures at 3.7 Å, 2.6 Å, and 2.7 Å resolutions, respectively (**fig. S12-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. 4A-B, fig. S15A**). 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 [243, 244], an essential eukaryotic protein that cleaves the cohesin ring to allow chromosome segregation (**fig. S15B-D**). Like separase, the CHAT domain contains a pseudo caspase domain at the N-terminal, a long dimeric coiled-coil mid-insertion, and a C-terminal active protease domain [243, 244]. Although structurally distinct, the two caspase domains pack in a similar fashion as the eukaryotic caspase dimers do [245]. In TPR-CHAT, the β-sheet structure in the pseudo-caspase 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 [246]. 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**).

A ~75 x 35 Å^2 area of the Cas7.1 surface in *Sb*-gRAMP is buried by TPR-CHAT (**Fig. 4C-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, V451) (**Fig. 4C**). A more extensive and mostly hydrophobic interface is found between one of the coiled-coil helix 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, fig. S17**). Given this conformational restriction, we speculated that the energetic barrier for RNase activation may be higher in Craspase compared to *Sb*-gRAMP. Indeed, RNA binding was consistently weaker at different temperatures and the cleavage was slower in Craspase compared to *Sb*-gRAMP (**Fig. 4F, fig. S16B-C**).

5.1.5. 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**). As this exceeds 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 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 guide and target. Constrained by the base-pairing from the first two PFS residues to the guide, the remaining PFS nucleotides point towards the bottom of TPR. Although their densities are difficult to model, possible phosphate densities sug-

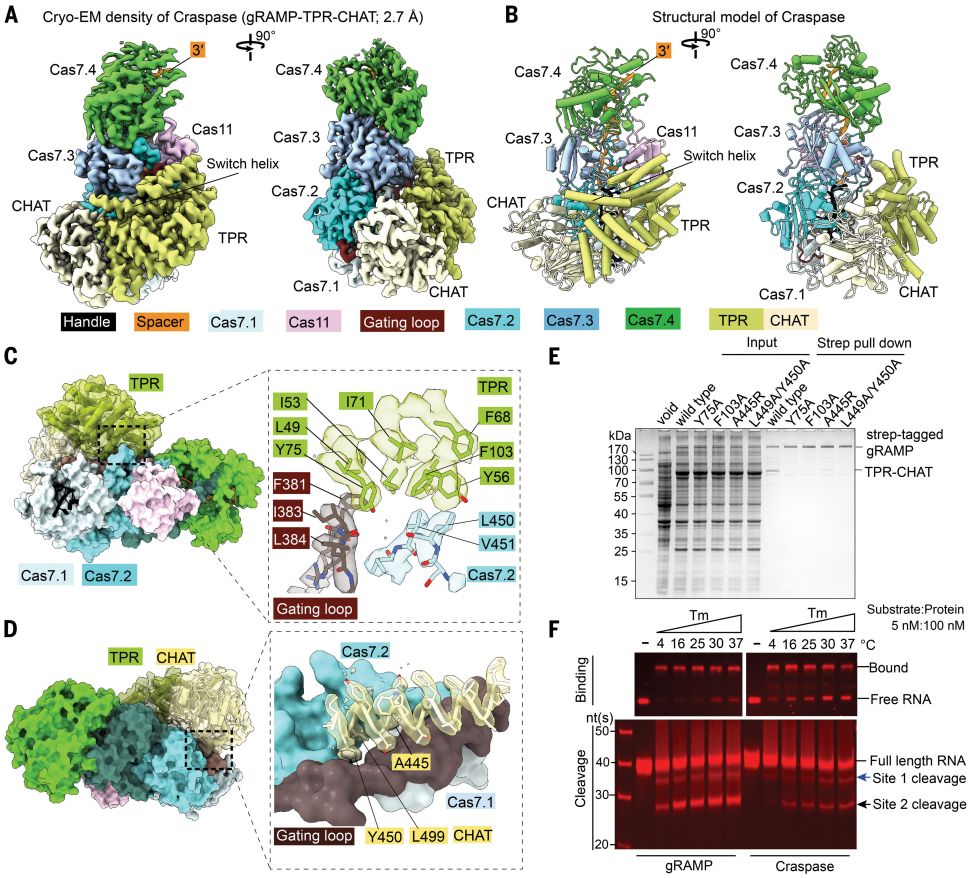


Figure 4. Structural basis for Craspase assembly. (**A**) 2.7-Å cryo-EM density and (**B**) structural model of Craspase. (**C**) Location and enlarged view of the molecular contacts between gRAMP and TPR. Interface residues and corresponding cryo-EM densities are shown. (**D**) Location and enlarged 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.

gest that the PFS travels underneath TPR (**Fig. 5C**). This path may have perturbed the conformation dynamics of the sensing β-hairpin in CHAT, as 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**). Notably, 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 sidechain-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.

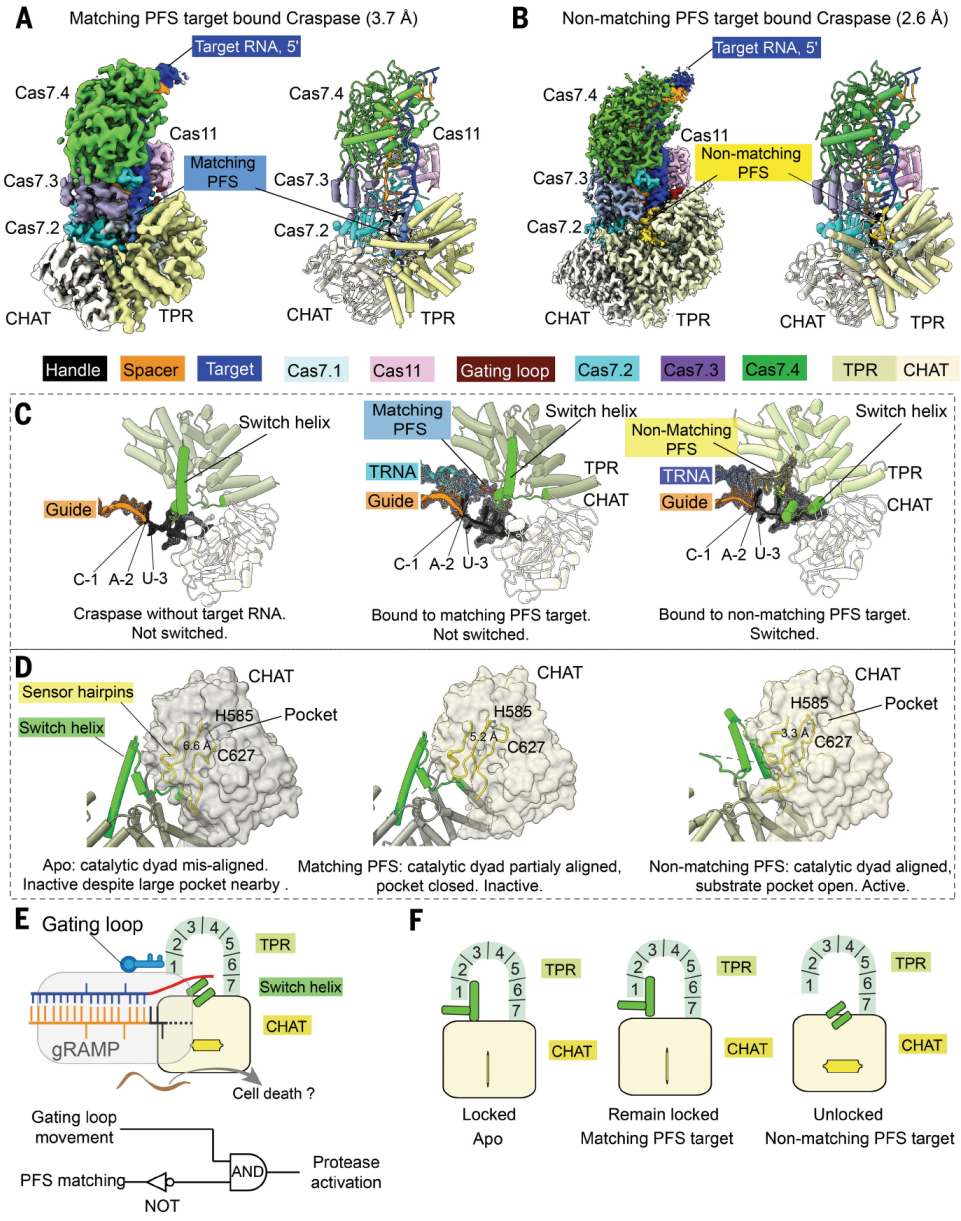


Figure 5. Structural basis for Craspase protease activation. (**A**) 3.7-Å cryo-EM density (left) and structural model (right) of matching PFS RNA–bound Craspase. (**B**) 2.6-Å cryo-EM density (left) and structural model (right) of non-matching PFS RNA–bound Craspase. (**C**) Close-up views of the switch helix in the resting state (left), the matching PFS RNA–bound state (middle), and the non-matching 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**) Top: model depicting non-matching PFS RNA–induced Craspase activation. Bottom: Logic gate diagram illustrating the protease activation mechanism. (**F**) Model depicting TPR-CHAT status in the *apo* state, the matching PFS RNA–bound state, and the non-matching PFS RNA–bound state.

A greater set of conformational changes take place when RNA target containing a non-matching 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**). While the first nucleotide of PFS forms a partially frayed A·C pair, the rest of PFS pivots toward the surface of TPR (**Fig. 5C**). The switch helix is dislodged from the shackle of the padlock, possibly due to clashes with the non-matching PFS. This helix and the preceding loop-helix-loop connection rotates 90 degree and packs against CHAT as a coiled-coil structure (**Fig. 5D, S19**). The sensor hairpin undergoes a larger set of long-range allosteric alterations. Consequently, C627 and H585 become oriented within 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**).

The above mechanistic analysis defines how RNA-guided RNA recognition regulates the protease activity of Craspase (**Fig. 5E**). Sequence complementarity in the target RNA is a prerequisite, which 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. 5F**).

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.

5.2. MATERIALS AND METHODS

PLASMID CLONING

Similar to *Plasmid cloning* in the Materials and Methods section of **Chapter 3**. To construct pCRISPR-GFP guide, the CRISPR array on pCRISPR-1 was replaced with a guide targeting GFP transcripts. To construct pTPR-CHAT-target1, pTPR-CHAT was enriched with a sequence encoding for target RNA complementary to crRNA1 under control of the IPTG inducible T7 promoter, using NEBuilder HiFi DNA Assembly (New England Biolabs).

CRASPASE EXPRESSION AND PURIFICATION

For gRAMP-crRNA and Craspase expression, plasmids pGRAMP, pCRISPR-1, pGRAMP-CRISPR-1, pGRAMP-CRISPR-1-target-1 and pTPR-CHAT were used [206, 233].

For overexpression of gRAMP-crRNA, pGRAMP-CRISPR-1 was transformed in BL21 (DE3) cells. A single colony was picked and grown for overexpression in 4 L of LB media supplemented with streptomycin (50 μg/mL). Expression was induced by adding IPTG to a final concentration of 0.5 mM and incubation at 16 °C overnight. Cells were collected by centrifugation at 5000 rpm on J6 centrifuge and lysed by sonication in buffer A (200 mM NaCl, 50 mM HEPES pH 7.5, 2 mM TCEP, 10% glycerol) with 1 mM phenylmethylsulfonyl fluoride (PMSF). The lysate was centrifuged at 12,000 rpm for 60 minutes at 4 °C, and the supernatant was filtered through a 0.45 μm syringe filter and loaded onto 5 mL of pre-equilibrated Strep-Tactin ®XT Affinity Resin (IBA Lifesciences GmbH). The loaded resin was washed with 25 mL of buffer A and the protein eluted by 20 mL buffer A with 2.5 mM dethio-biotin. The elution was concentrated and further purified by size-exclusion chromatography (Superdex 200 Increase 10/300 GL; Cytiva) equilibrated with buffer B (175 mM NaCl, 25 mM HEPES pH 7.5). The second size-exclusion chromatography peak was collected, concentrated, and flash frozen with liquid nitrogen.

For Craspase expression, the pGRAMP-CRISPR-1 plasmid was co-transformed in Bl21 (DE3) with the pTPR-CHAT plasmid. A single colony was picked and grown for expression in 4 L of LB media. Expression was induced by adding IPTG to a final concentration of 0.5 mM and incubation at 16 °C overnight. Cells were collected by centrifugation and lysed by sonication in buffer A (200 mM NaCl, 50 mM HEPES pH 7.5, 2 mM TCEP, 10% glycerol) with 1 mM phenylmethylsulfonyl fluoride (PMSF). The lysate was centrifuged at 12,000 rpm. for 60 minutes at 4°C, and the supernatant was filtered through a 0.45 μm syringe filter and loaded onto 5 mL of pre-equilibrated Strep-Tactin ®XT Affinity Resin (IBA Lifesciences GmbH). The loaded resin was washed with 20 mL buffer A and eluted by 25 ml buffer A with 2.5 mM dethio-biotin. The elution was then loaded onto 3 mL of pre-equilibrated Ni-NTA resin (Qiagen). After washing with 15 mL of buffer A, the sample was eluted by 15 mL buffer A with 300 mM imidazole. This eluted sample was concentrated and further purified by size-exclusion chromatography (Superdex 200 Increase 10/300 GL; Cytiva) equilibrated with buffer B (175 mM NaCl, 25 mM HEPES pH 7.5). The peak was collected, concentrated, and flash frozen with liquid nitrogen.

RNA CLEAVAGE ASSAYS

20 μL reactions were prepared where 20 nM Cy5 labeled target RNA (**table S1**) was incubated with *Sb*-gRAMP-crRNA in cleavage buffer (100 mM KCl, 50 mM HEPES pH 7.5, 2 mM βME, 5 mM $MgCl_2$) and incubated at room temperature for 30 minutes (unless otherwise specified). The 20 μL reactions were quenched with the addition of EDTA to 150 mM (final concentration) and 1:1 volume of 100% formamide. Samples were heated to 95 °C for 10 minutes and run on 12% urea-PAGE gel. Fluorescent signals were imaged using ChemiDoc (BioRad) and quantified using Image Lab. For the RNA cleavage experiments using varying lengths of complementarity in the target RNA, 10 μL reactions containing 200 nM *Sb*-gRAMP-crRNA with 100 nM target RNA was used in an alternative cleavage buffer (25 mM Tris-HCl, 150 mM NaCl, 10 mM DTT and 2 mM $MgCl_2$). Reactions were incubated at 20 °C for 2 hours, after which 0.5 μL of 800 units/mL proteinase K (New England Biolabs) was added for 1 hour at 37 °C. Reactions were heat inactivated at 95 °C for 5 minutes, followed by mixing with 2x RNA loading dye (95% formamide, 0.025% SDS and 0.5 mM EDTA) and loaded on an 8M urea 10% PAGE gel (pre-run at 350 V for 1 hour, sample run at 333V for 2 hours)

RNA KNOCK DOWN ASSAY IN *E.coli*

pGRAMP-CRISPR-1, pCRISPR-GFP guide wild type or repeat region mutants were co-transformed into BL21(DE3) cells and plated on LB-agar plates with appropriate antibiotics (spectinomycin, kanamycin and ampicillin). Transformation plates were incubated for 20 hours at 37 °C and then scanned via GFP channel using a ChemiDoc (BioRad). For the dot assay, single colonies were picked after transformation and cultured in 3 mL LB media with appropriate antibiotics for 18 hours in 37 °C. All cultures were adjusted to the same OD600 of 0.1. 10 μL from each culture was pipetted on a plastic membrane. Fluorescent signals were imaged using a ChemiDoc (BioRad) via 515 nm to monitor GFP signal and 600 nm to monitor cell density.

ELECTROPHORETIC MOBILITY SHIFT ASSAY

5 nM final concentration of fluorescently labeled target RNA was incubated with specified concentration of *Sb*-gRAMP-crRNA or Craspase in a 20 μL reaction containing binding buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol). After 20 minutes of incubation on ice (unless otherwise specified), 10 μL of each sample was loaded onto 1% agarose gel equilibrated in 0.5x TBE buffer. Electrophoresis was performed at 60 V for 30 minutes at 4 °C. Fluorescent signals were recorded using a ChemiDoc (Bio-Rad).

CRYO-EM SAMPLE PREPARATION

Sb-gRAMP-crRNA was incubated with the specified target RNA in cryo-EM buffer (150 mM NaCl, 25 mM HEPES pH 7.5) for 15 minutes at room temperature. RNA was supplied at a 3-fold molar excess to *Sb*-gRAMP-crRNA (0.5 mg/mL final concentration). 3.5 μL of the incubation was applied to a Quantifoil holey carbon grid (1.2/1.3, 200 mesh) which had been glow-discharged with 20 mA at 0.39 mBar for 30 seconds (PELCO easiGlow). Grids were blotted with Vitrobot blotting paper (Ted Pella, Inc.) for 3 seconds at 4 °C, 100% humidity, and plunge-frozen in liquid ethane using a Mark IV FEI/Thermo Fisher Vitrobot. EM data acquisition Data was collected on a 200 kV Talos Arctica (Thermo

Fisher Scientific) or 300 kV Krios G3i Cryo Transmission Electron Microscope (Thermo Fisher Scientific) with Gatan K3 direct electron detector. The total exposure time of each movie stack led to a total accumulated dose of 50 electrons per $Å^2$ which fractionated into 50 frames without beam tilt. Dose-fractionated super-resolution movie stacks collected from the Gatan K3 direct electron detector were binned to a pixel size of 1.1 Å. The defocus value was set between -1.0 μm to -2.5 μm.

IMAGE PROCESSING AND 3D RECONSTRUCTION

Motion correction, CTF-estimation, blob particle picking, 2D classification, 3D classification and non-uniform 3D refinement were performed in CryoSPARC [247]. High quality 2D classes were hand-picked to do 3D classifications in C1 symmetry. Cryo-SPARC was instructed to generate eight initial models in 3D classification for each reconstruction. After heterologous refinement, the promising 3D models was hand-picked for post 3D refinement. A solvent mask was generated in RELION [248] with 0.1 contour level and was used for all subsequent local refinement steps. CTF post refinement was conducted to refine the beam-induced motion of the particle set, resulting in the final maps. The detailed data processing, refinement and validation statistics are summarized in [233].

ATOMIC MODEL BUILDING

Different modeling procedures were used for *Sb*-gRAMP-crRNA and Craspase. For *Sb*-gRAMP-crRNA with non-matching PFS target, the *Sb*-gRAMP sequence was used to generate an AlphaFold2 prediction, which was split into domains followed by individual docking into the cryo-EM map using UCSF Chimera [249]. Model building was completed through iterative cycles of manual building using Coot [250] and real space and positional refinements using Phenix [251]. Notably, the cryo-EM density for the big insertion domain did not allow de novo model building despite efforts to improve the local resolution. An AlphaFold2 predicted structure was docked into the EM density to model BID. In other structures was built and refined using a similar strategy. The detailed statistics are documented in [233].

MODEL VALIDATION

Map-model Fourier Shell Correlation was computed using Mtriage in Phenix. Map-model FSC resolution of each dataset was estimated from Mtriage FSC curve, using 0.143 cutoff. The masked cross-correlation (CCmask) from Mtriage was reported as representative model-map cross-correlation. Model geometry was validated using the MolProbity [252]. All the deposited models were submitted to MolProbity server to check the clashes between atoms, Ramachandran-plot, bond-angles, bond-lengths, sidechain rotamers, CaBLAM and C-beta outliers. All the model validation stats are summarized in [233].

5.3. SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURES

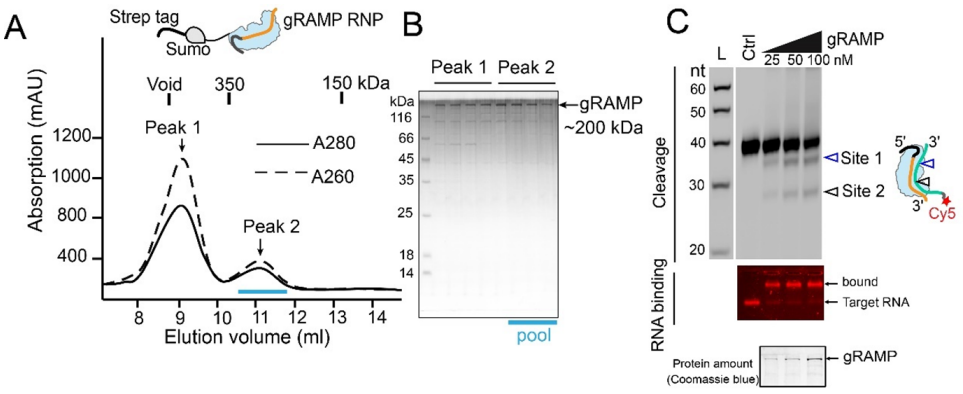


Figure S1. Reconstitution and biochemical characterization of the *Sb*-gRAMP RNP. (**A**) Purification profile on size-exclusion chromatography. (**B**) SDS-PAGE analysis of the peaks on size-exclusion. (**C**) Top: cleavage assay for different concentrations of *Sb*-gRAMP. Middle: EMSA assay using the same titration of *Sb*-gRAMP. Bottom: SDS-PAGE showing protein concentration in cleavage and EMSA assay.

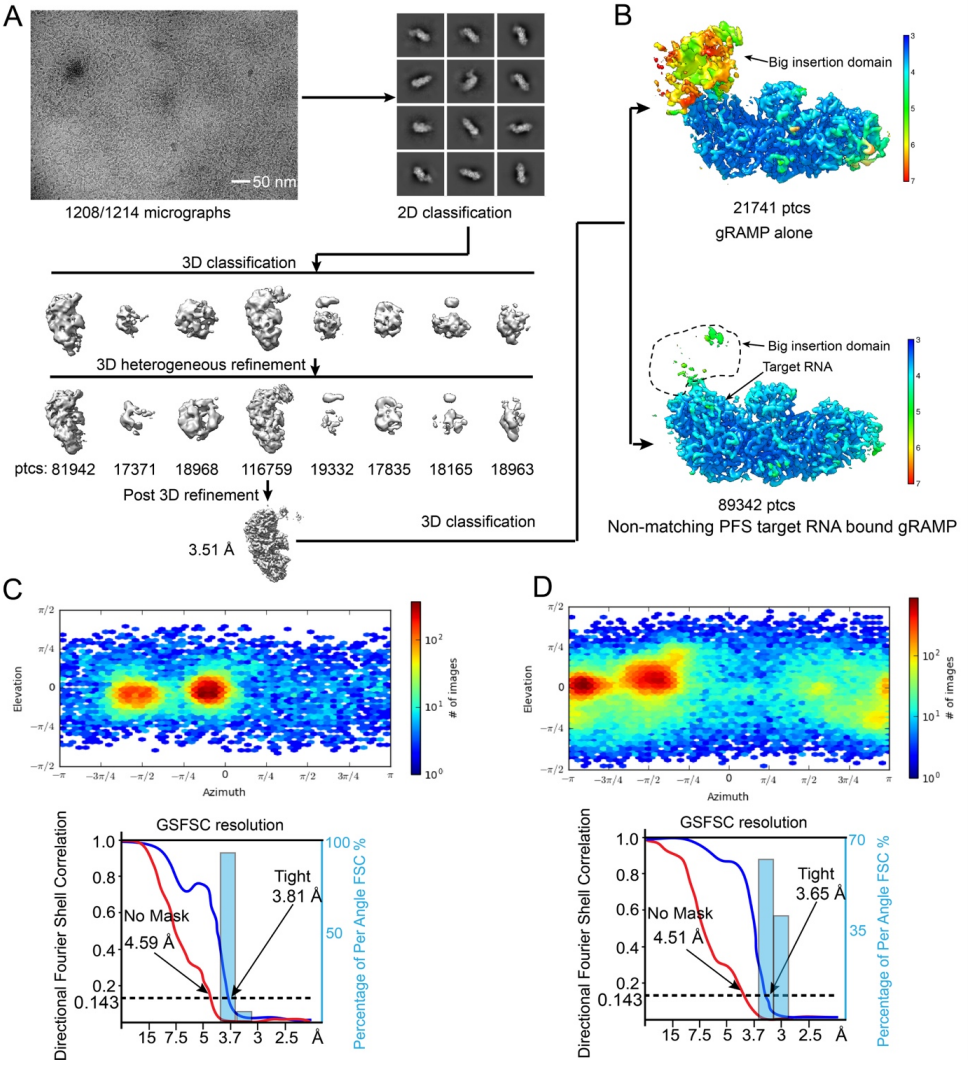


Figure S2. Single particle cryo-EM reconstruction of *Sb*-gRAMP in resting and non-matching PFS RNA bound states. (**A**) Workflow of the cryo-EM image processing and 3D reconstruction for the non-matching PFS RNA bound *Sb*-gRAMP. A portion of the single particles did not bind RNA. They gave rise to the resting state *Sb*-gRAMP reconstruction. (**B**) Final cryo-EM density map showing the local resolution distribution of the two reconstructions. (**C, D**) 3D Euler distribution plot (top panel) and Fourier Shell Correlations (FSC; bottom panel) of *Sb*-gRAMP in the resting and NPFS bound states, respectively. The gold-standard cutoff (FSC = 0.143) was marked with a dotted line.

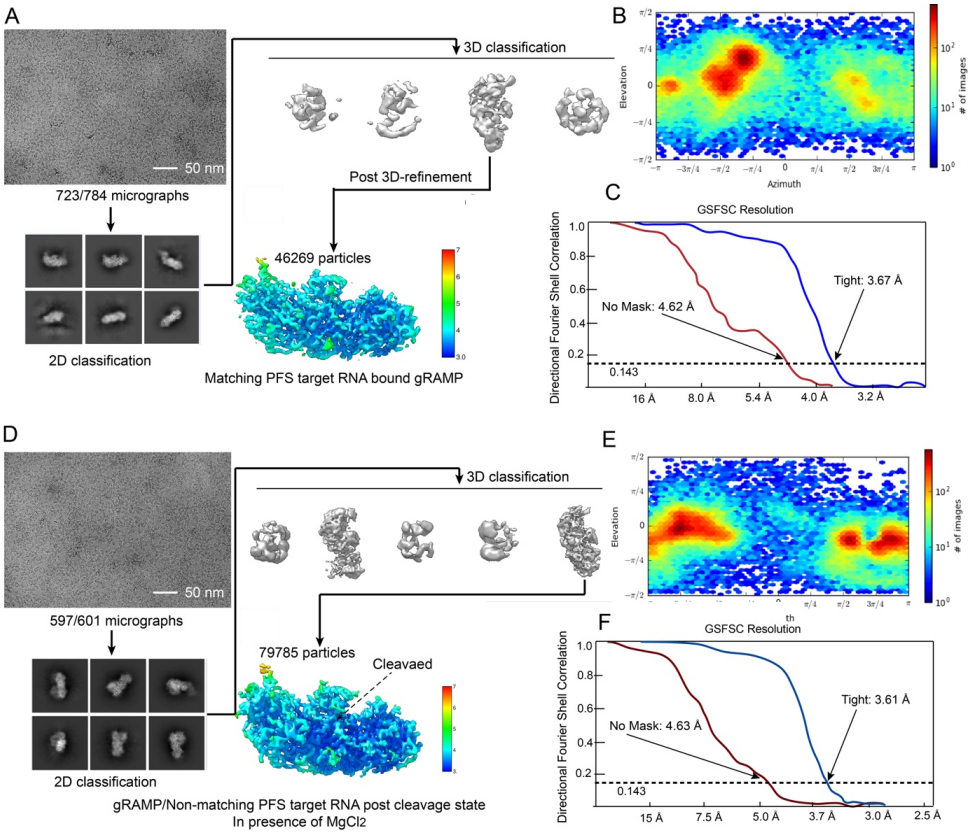


Figure S3. Single particle cryo-EM reconstruction of *Sb*-gRAMP in Matching PFS state and $MgCl_2$-incubation induced post-cleavage state. (**A**) Workflow of the cryo-EM image processing and 3D reconstruction for the *Sb*-gRAMP in matching PFS RNA bound state. The local resolution distribution is colored on the final EM density map. (**B**) 3D Euler distribution plot revealing the representation of the single particle orientation in the final reconstruction. (**C**) Fourier Shell Correlation defined reconstruction resolution (FSC). The dotted line indicates the final resolution defined by the gold-standard FSC cutoff (0.143) (**D**) Workflow of the cryo-EM image processing and 3D reconstruction for the *Sb*-gRAMP in the post-cleavage state, induced by $MgCl_2$ incubation on a Non-matching PFS RNA substrate. The local resolution distribution is colored on the final EM density map. (**E**) 3D Euler distribution plot of the post-cleavage state. (**C**) FSC defined resolution of the post-cleavage state reconstruction.

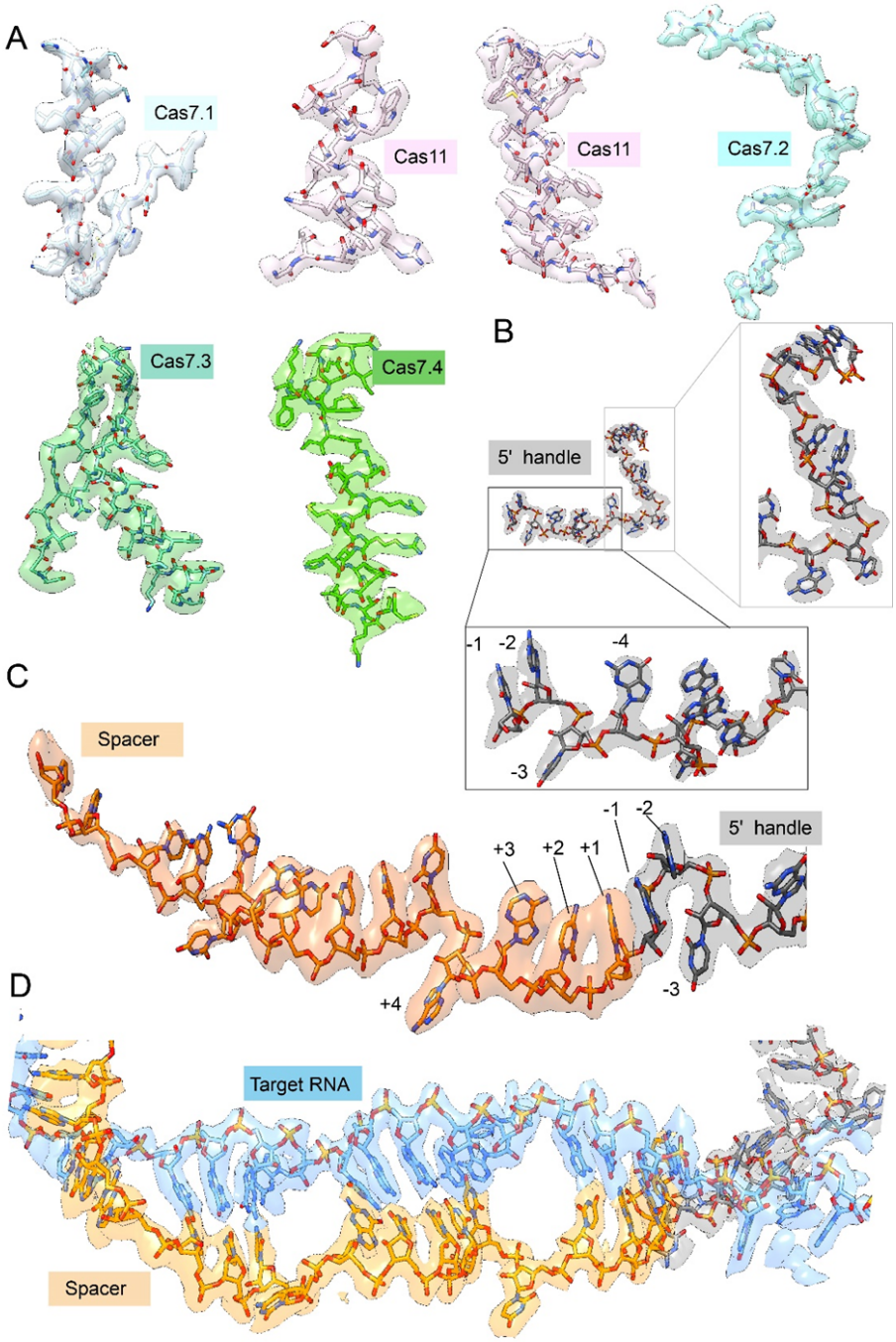


Figure S4. Representative local map density for the different functional states. (**A**) EM densities for representative protein regions inside *Sb*-gRAMP and *Sb*-gRAMP/RNA complex. (**B**) EM densities for the 5′ handle region of the crRNA. The number indicates the base order. (**C**) EM densities for guide RNA region inside *Sb*-gRAMP. (**D**) EM densities for the duplex of target RNA-guide RNA.

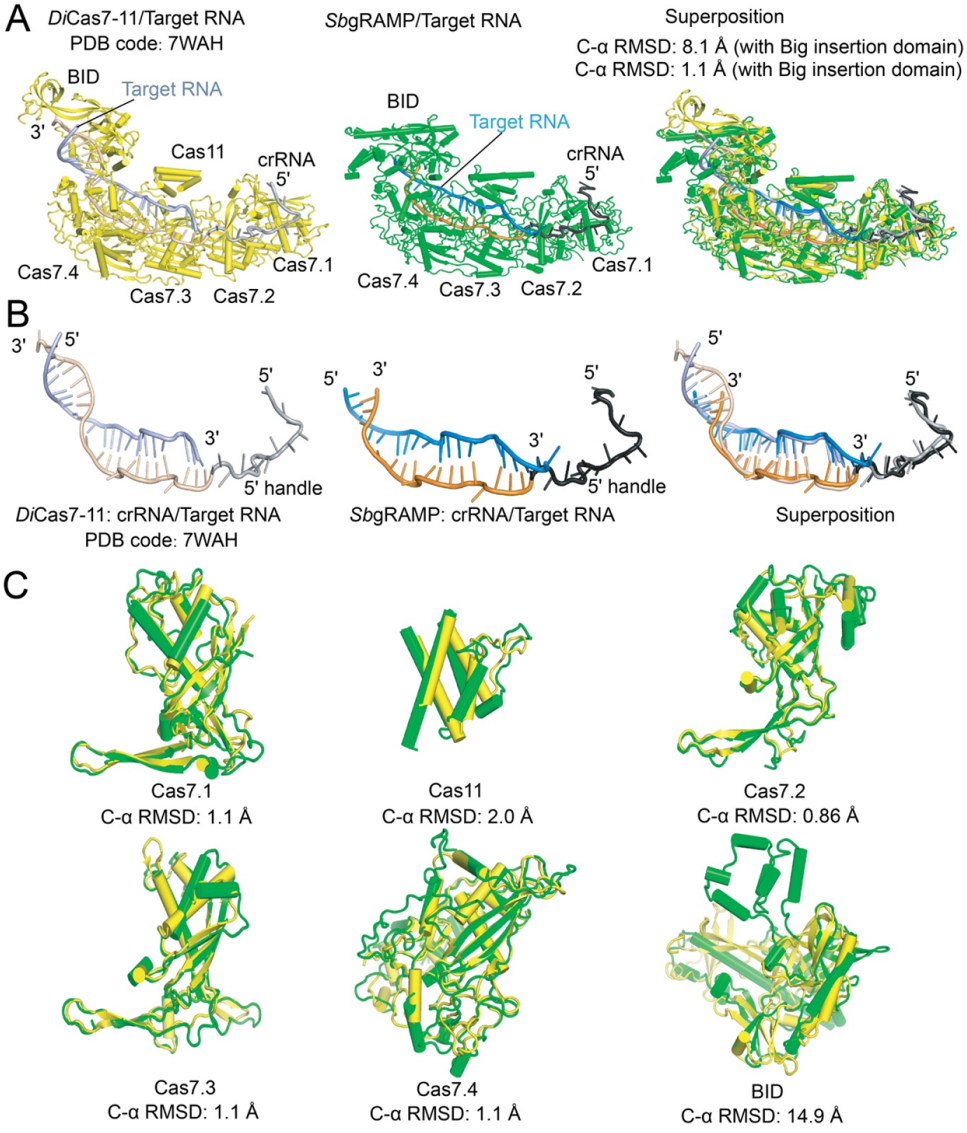


Figure S5. Structural comparison between *Sb*-gRAMP and *Di*-Cas7-11, both in the target RNA bound state. (**A**) Side-by-side and superpositioning of *Sb*-gRAMP and *Di*-Cas7-11 (PDB: 7WAH). The structure and orientation if the big insertion domain are significantly different in the two structures. The rest of the two structures agree quite well. Alignment of all non-BID Cα atoms produced an r.m.s.d. of only 1.1 Å. (**B**) Guide and target RNAs also align well when the two structures are aligned along non-BID Cα atoms. (**C**) Individual domains aligned along Cα atoms superimpose very well, except BID. The BID cryo-EM density in our reconstruction did not allow reliable de novo model building. An AlphaFold predicted model was docked instead.

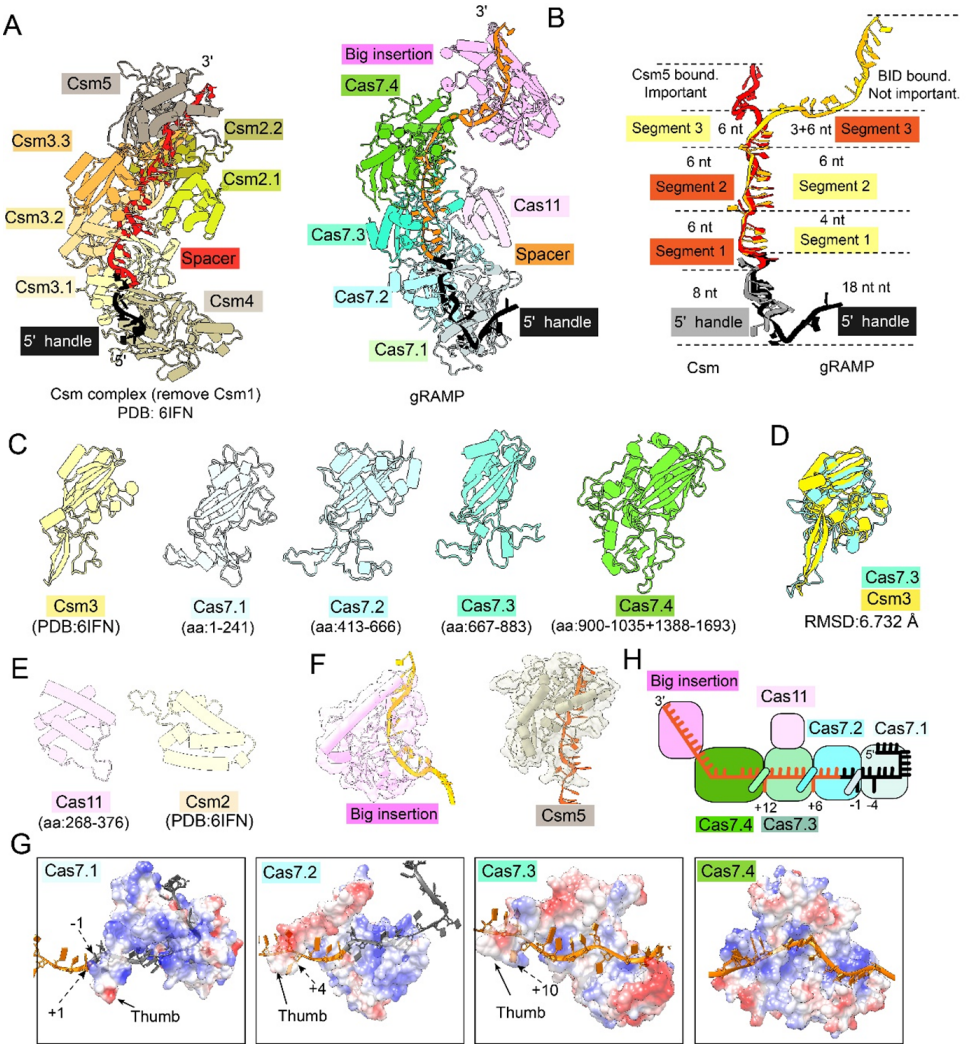


Figure S6. Structural comparison between type III-E *Sb*-gRAMP and type III-A Csm complex. (**A**) Side-by-side comparison between type III-E *Sb*-gRAMP and type III-A Csm (*37*) (PDB: 6IFN) structures. (**B**) Structural alignment and comparison of crRNAs from *Sb*-gRAMP and Csm. (**C**) Side-by-side comparison of Csm-Cas7 and four Cas7 domains in *Sb*-gRAMP. (**D**) Structural alignment showing that Cas7.3 has the highest structure conservation with Csm3. (**E**) Side by side comparison between *Sb*-gRAMP Cas11 and its equivalent Csm2 in the Csm complex. (**F**) Side by side comparison between *Sb*-gRAMP BID and Csm5 in the Csm complex. (**G**) Snapshots of four Cas7 subunits bound to crRNA. Cas7.1, Cas7.2, and Cas7.3 all have an obvious thumb structure to flip out a base of crRNA. (**H**) Cartoon model showing the architecture of the *Sb*-gRAMP RNP complex.

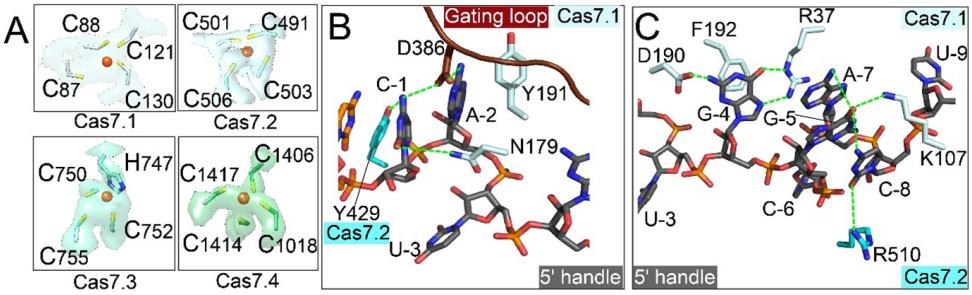


Figure S7. Structural basis for the zinc-finger of *Sb*-gRAMP and 5′ handle recognition. (**A**) Close-up view of the cryo-EM density of zinc-fingers from Cas7.1 to Cas7.4 subunits. (**B**) Close-up view of Cas7.1 and Cas7.2 contacting the -1 to -3 bases of 5′ handle. (**C**) Close-up view of Cas7.1 and Cas7.2 contacting the -4 to -9 bases of 5′ handle.

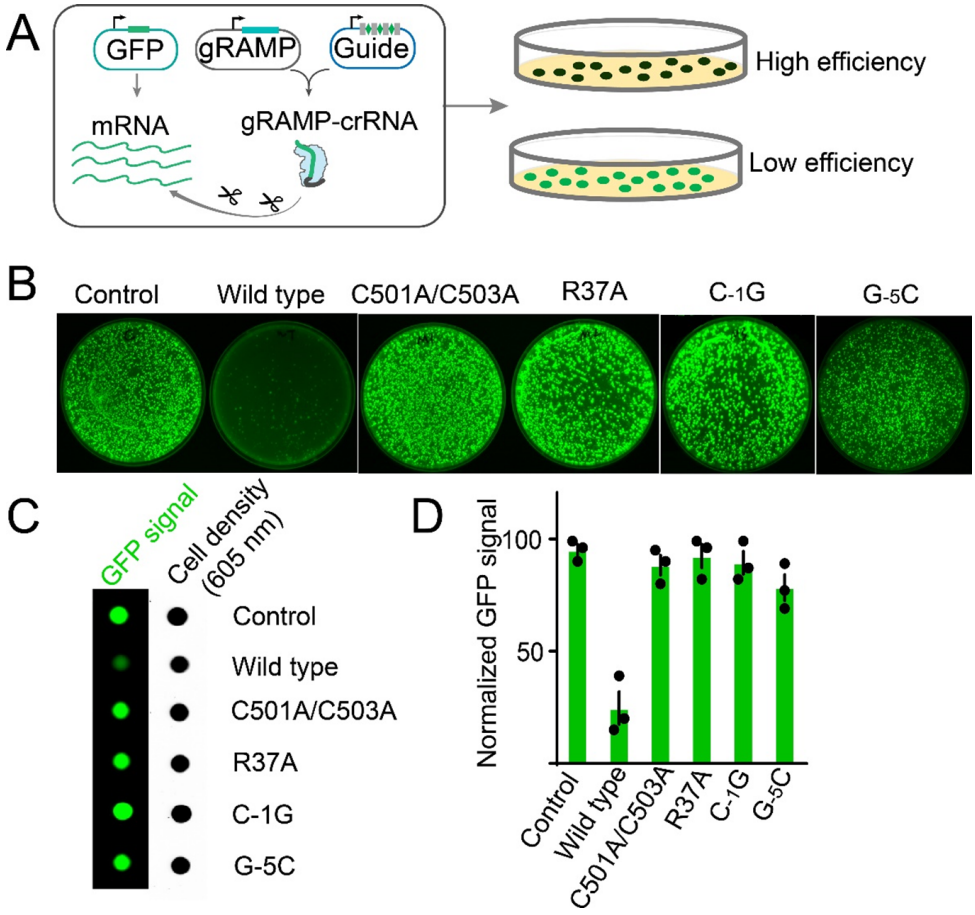


Figure S8. Mutagenesis analysis of *Sb*-gRAMP RNP protein. (**A**) Model depicting the workflow of in vivo RNA knock down assay in *E.coli*. (**B**) Transformation plates of RNA knockdown assay from different mutations (C501A/C503 is the zinc-finger from Cas7.2; R37 from Cas7.1; C-1 and G-5 from the 5′ handle region of crRNA). (**C**) Bacteria solution dot assay to show the GFP signal and cell density. (**D**) Normalized knock down efficiency from different mutagenesis.

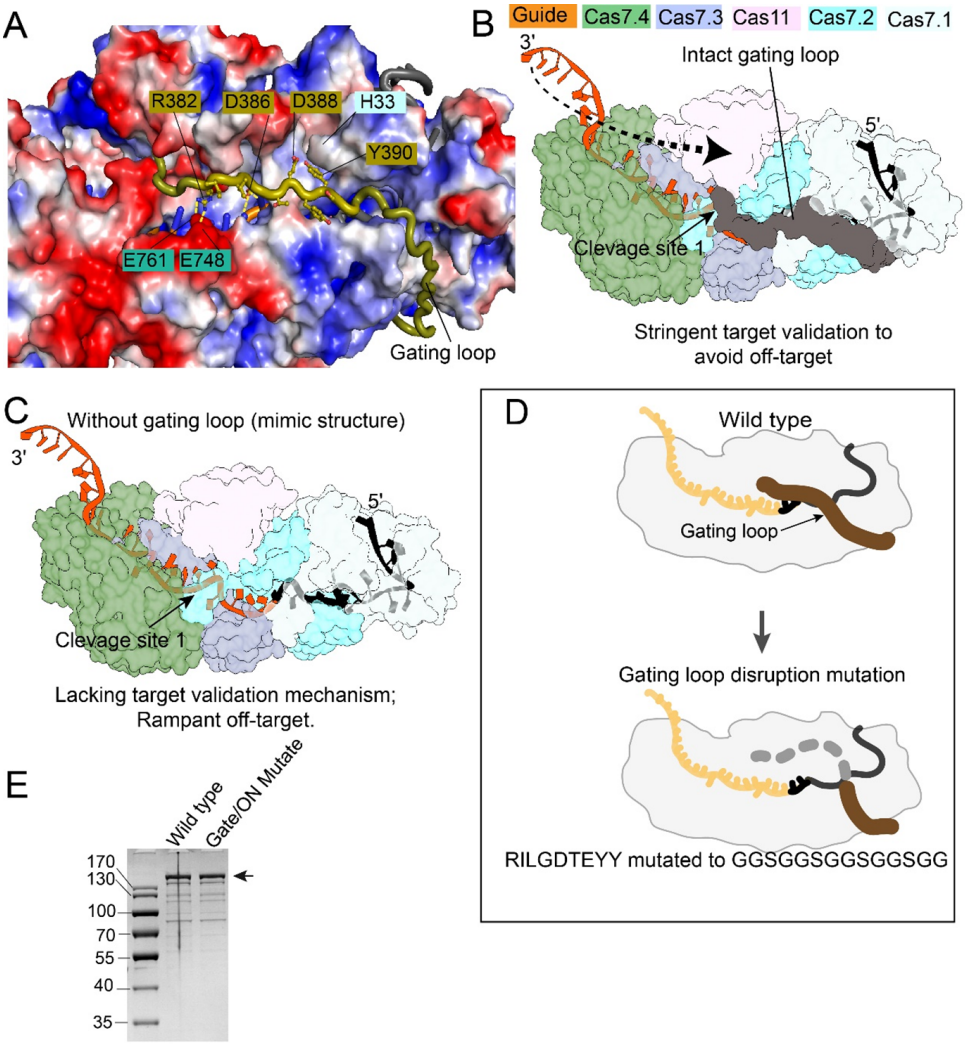


Figure S9. Functional characterization of the gating loop. (**A**) close-up view of the gating loop located in a deep cavity and blocking the "seed region" of crRNA. (**B**) A model showing the direction of base-pairing formation between target RNA and guide RNA. Base-pairings that can only happen after gating loop displacement likely takes place at the end. (**C**) A model showing that without gating loop protection, base-pairing can take place stochastically. Off-targeting may be rampant. (**D**) Cartoon model depicting mutagenesis of the gating loop. (**E**) SDS-PAGE showing the quality of the wild type and gating loop mutant *Sb*-gRAMP samples.

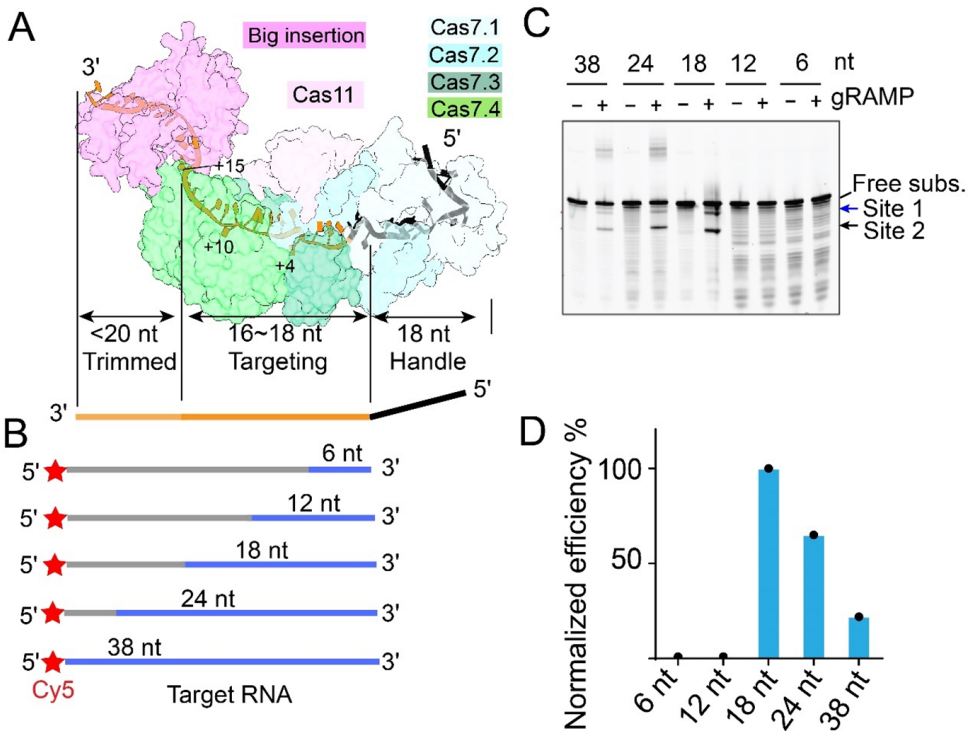


Figure S10. RNA cleavage experiments using varying lengths of complementarity in the target RNA. (**A**) Cartoon model to showing the full-length *Sb*-gRAMP architecture and crRNA region. (**B**) Model of RNA substrates used in assay showing varying lengths of complementarity with crRNA. (**C**) Cleavage assay of RNA substrates and (**D**) the normalized cleavage efficiencies.

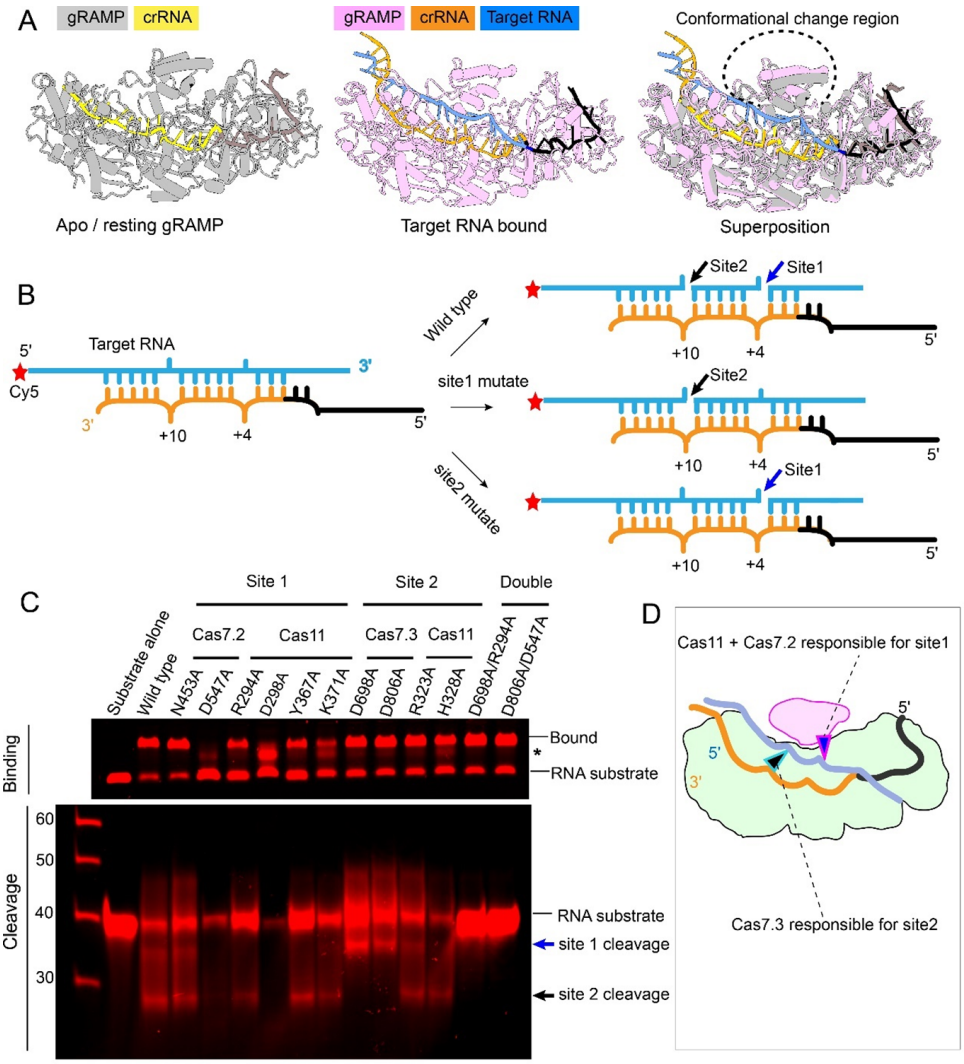


Figure S11. Characterization of *Sb*-gRAMP RNase activity. (**A**) Structural alignment showing the main conformational change is in Cas11 region. (**B**) A model showing the RNA cleavage pattern in wild type, site 1, and site 2 mutations of *Sb*-gRAMP. (**C**) RNA substrate binding and cleavage assay for *Sb*-gRAMP mutations. (**D**) Cartoon model depicting the *Sb*-gRAMP domains involved in each cleavage site.

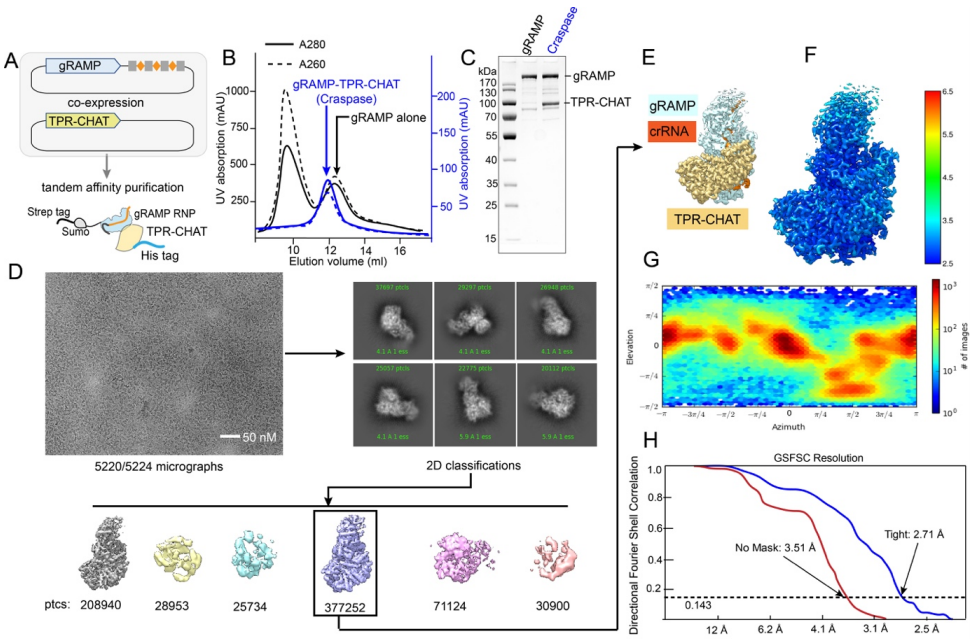


Figure S12. Purification and Cryo-EM single particle reconstruction of *Sb*-gRAMP-TPR-CHAT complex (Craspase). (**A**) Model of method used for co-expression and purification for Craspase. (**B**) Purification profile comparison between *Sb*-gRAMP and Craspase on size-exclusion chromatography. (**C**) Representative SDS-PAGE comparison between *Sb*-gRAMP and Craspase. (**D**) Workflow of the cryo-EM image processing and (**E**) 3D reconstruction for Craspase. (**F**) Final electron density map showing local resolution for Craspase. (**G**) 3D Euler distribution plot revealing the representation of the single particle orientation in the final reconstruction. (**H**) Fourier Shell Correlations (FSC) of the *Sb*-gRAMP-Matching PFS RNA complex reconstruction, with the gold-standard cutoff (FSC = 0.143) marked with a dotted line.

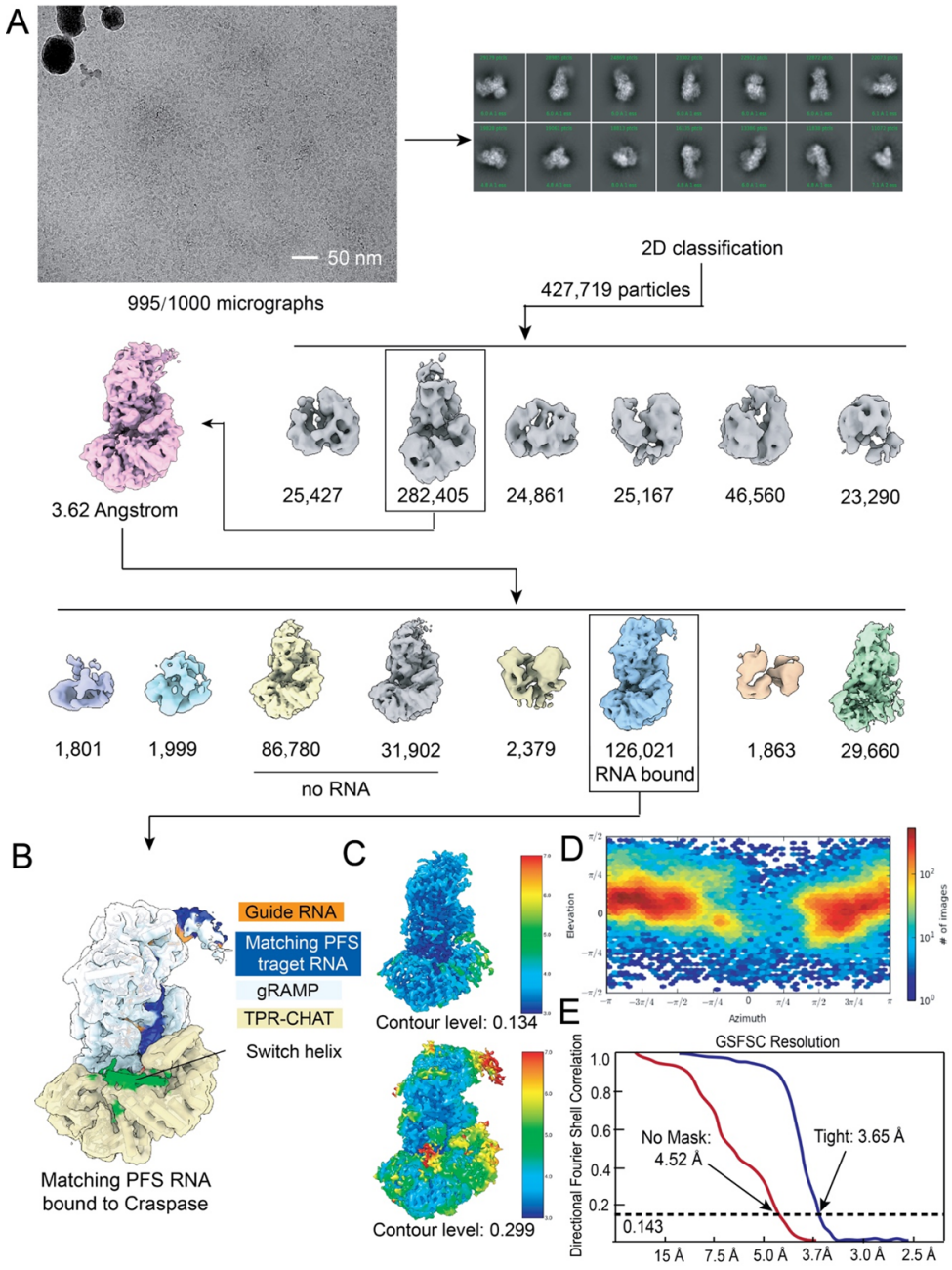


Figure S13. Cryo-EM single particle reconstruction of Craspase bound with a matching PFS RNA target. (**A**) Workflow of the cryo-EM image processing and (**B**) 3D reconstruction for the Craspase bound with a matching PFS RNA. (**C**) Final electron density map showing local resolution at different contour levels. (**D**) 3D Euler distribution plot revealing the representation of the single particle orientation. (**H**) Fourier Shell Correlations defining the reconstruction resolution, with the gold-standard cutoff (0.143) marked with a dotted line.

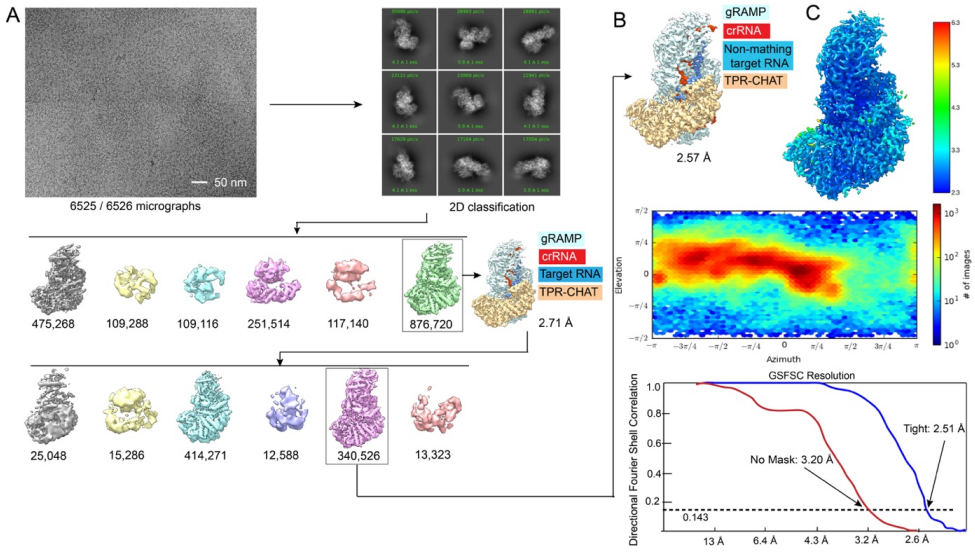


Figure S14. Cryo-EM single particle reconstruction of Craspase bound with a non-matching PFS RNA target. (**A**) Workflow of the cryo-EM image processing and (**B**) 3D reconstruction for the Craspase bound with a non-matching PFS RNA. (**C**) Final electron density map showing local resolution at different contour levels. (**D**) 3D Euler distribution plot revealing the representation of the single particle orientation. (**H**) Fourier Shell Correlations defining the reconstruction resolution, with the gold-standard cutoff (0.143) marked with a dotted line.

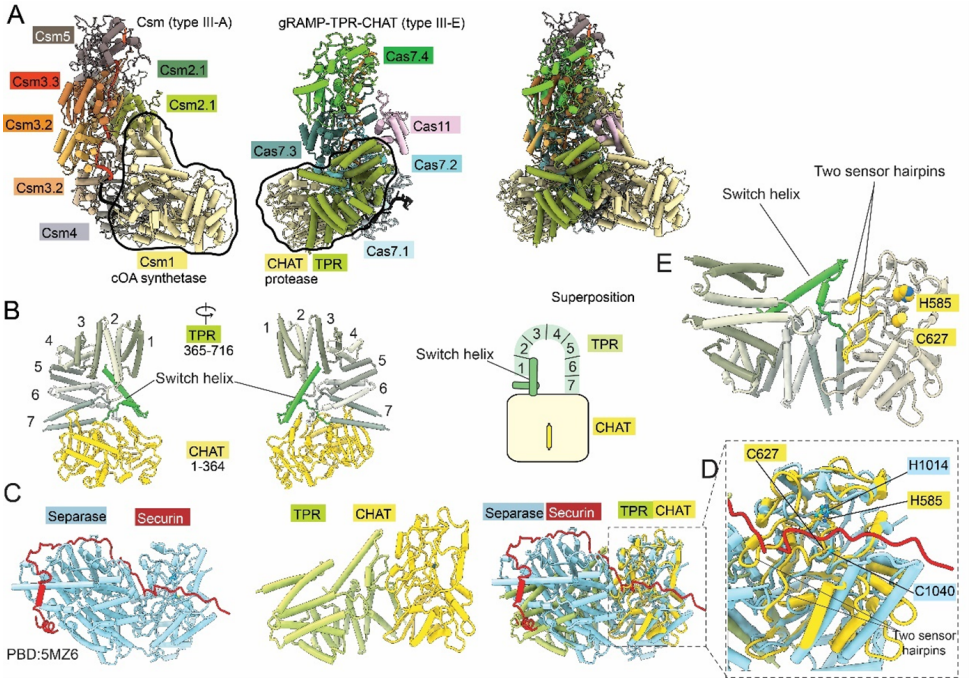


Figure S15. Structural analysis of the TPR-CHAT component of Craspase. (**A**) Left: structural model of Csm complex with cOA synthetase. Middle: structural model of *Sb*-gRAMP with TPR-CHAT protease. Right: structural superposition of the Csm complex and Craspase. (**B**) Right and middle: overall structure of the TPR-CHAT component of Craspase. Right: Cartoon model depicting the architecture of TPR-CHAT. (**C**) Structural comparison between Separase-Securin complex (PDB: 5MZ6) and TPR-CHAT. Left: Separase-Securin complex structure. Middle: TPR-CHAT structure. Right: structural superposition of TPR-CHAT with Separase-Securin. (**D**) Close-up view of the structural alignment of protease domain from Separase and CHAT. (**E**) Close-up view of "switch helix" and "sensor hairpins".

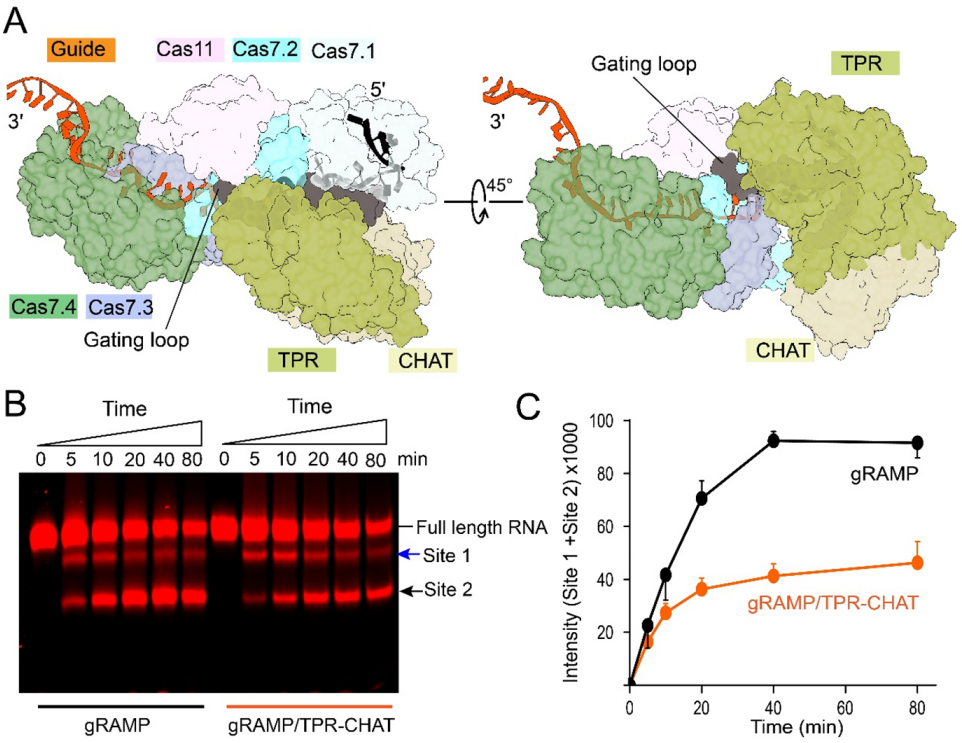


Figure S16. Biochemistry and structure analysis of the Craspase complex. (**A**) Structure model revealing that TPR-CHAT binds to the gating loop and limits gating loop dynamic movement. (**B**) RNA cleavage activity comparison between *Sb*-gRAMP and Craspase with (**C**) quantification of the cleavage products from site 1 and site 2.

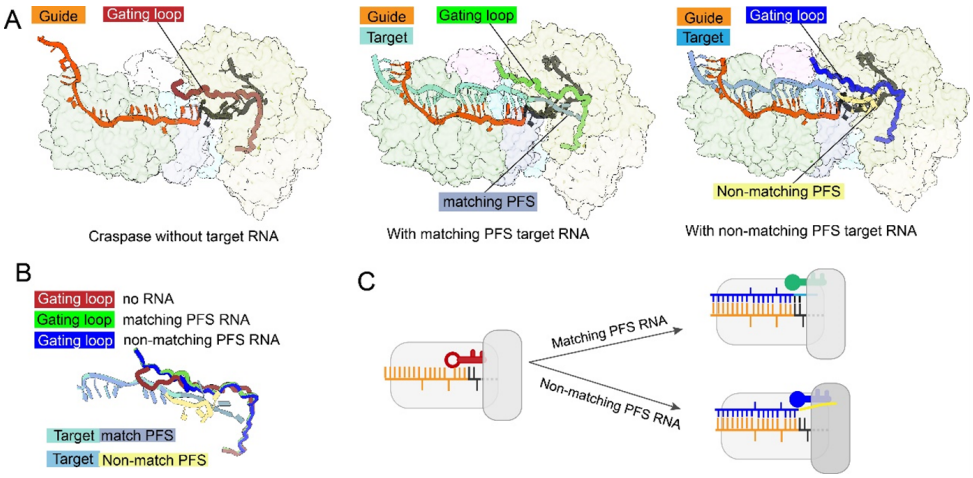


Figure S17. The structural dynamics of the gating loop in Craspase. (**A**) Cartoon model highlighting the gating loop conformation in apo Craspase (left), matching PFS RNA bound (middle) and non-matching PFS RNA bound Craspase complex (right). (**B**) Structural alignment of the gating loop in apo, matching PFS RNA bound and non-matching PFS RNA bound state. (**C**) Model depicting the gating loop dynamics when bound to different substrates.

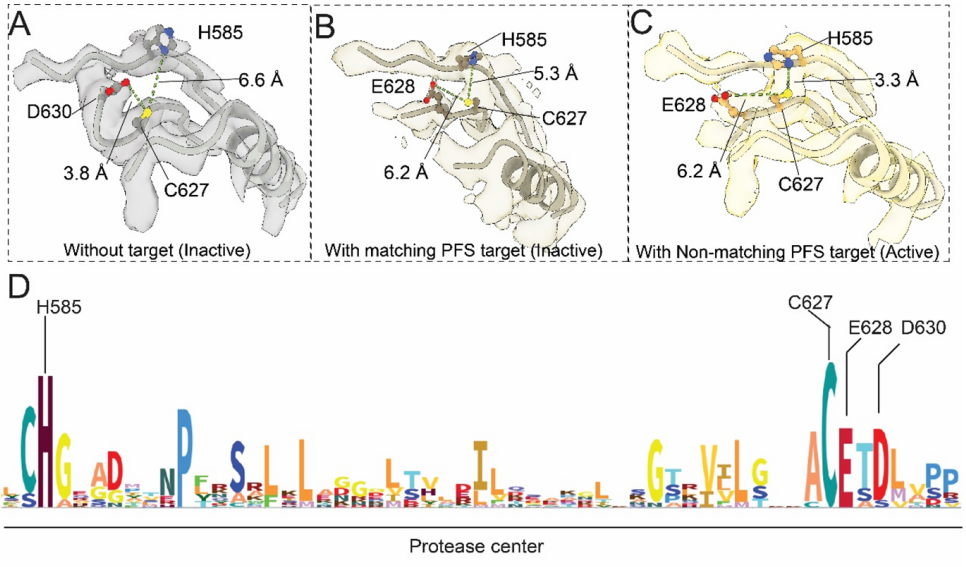


Figure S18. CryoEM density and conservation analyses of the protease center. The extracted local density of protease center from apo state (**A**), matching PFS target RNA bound state (**B**) and non-matching PFS target bound state (**C**). Highlight the protease center residues H585, C627, E628 and D630. The catalytic dyad in the apo state and the matching PFS state are too far apart to allow general acid-base based protease activity in Craspase. (**D**) Amino sequence alignment of protease center from twenty-five TPR-CHAT homologs.

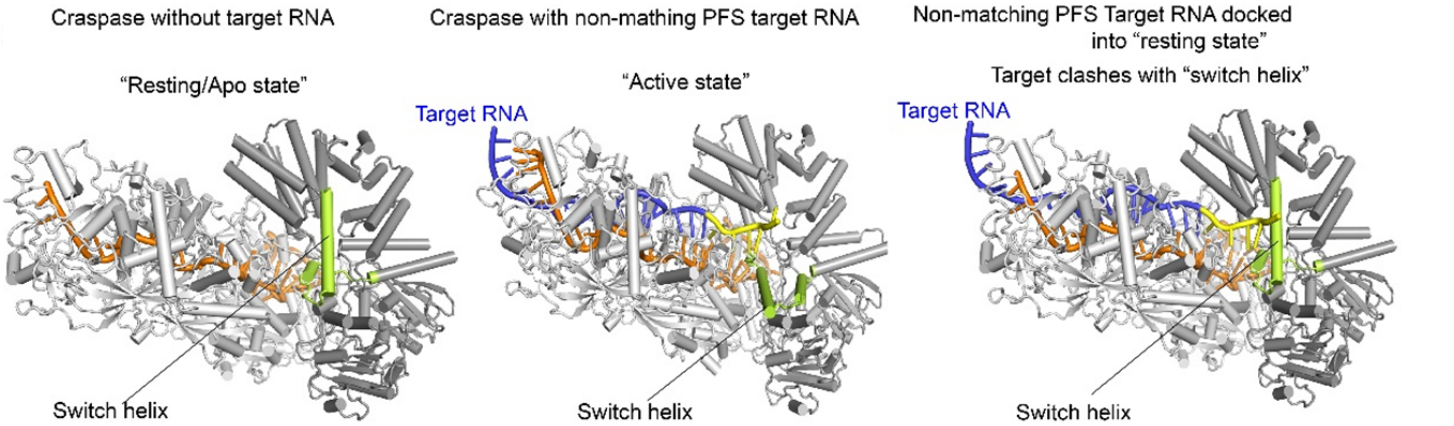


Figure S19. The structural dynamics of the switch helix in TPR-CHAT. Cartoon model highlighting the switch helix conformation in apo Craspase (left), non-matching PFS RNA bound state (middle), and their superposition (right).

SUPPLEMENTARY TABLE

Length of complementarity	Label and RNA sequence (complementarity to crRNA1 in bold)
38nt	Cy5-**CUCUAGUAACAGCCGUGGAGUCCGGGGCAGAAAAUUGG**ACGAUUAA
24nt	Cy5-GAGAUCAUUGUCGG**GUGGAGUCCGGGGCAGAAAAUUGG**ACGAUUAA
18nt	Cy5-GAGAUCAUUGUCGGCACCUC**UCCGGGGCAGAAAAUUGG**ACGAUUAA
12nt	Cy5-GAGAUCAUUGUCGGCACCUCAGGCCC**GCAGAAAAUUGG**ACGAUUAA
9nt	Cy5-GAGAUCAUUGUCGGCACCUCAGGCCCCGU**GAAAAUUGG**ACGAUUAA
6nt	Cy5-GAGAUCAUUGUCGGCACCUCAGGCCCCGUCUU**AAUUGG**ACGAUUAA

Table S1. RNA oligo's used in this study. The RNAs were used for the RNA cleavage experiments using varying lengths of complementarity and to study the impact of the gating loop.

6

CRASPASE ORTHOLOGS CLEAVE A NONCONSERVED SITE IN CSX30

"The question is not what you look at, but what you see."

Henry Thoreau

The Craspase CRISPR-Cas effector consists of the RNA-guided ribonuclease gRAMP and the protease TPR-CHAT, coupling target RNA recognition to protease activation. The natural substrate of Craspase is Csx30, a protein cleaved in two fragments that subsequently activates downstream antiviral pathways. Here, we determined the protease substrate specificity of Craspase from *Candidatus* "Jettenia caeni" (*Jc*-Craspase). We find that *Jc*-Craspase cleaves *Jc*-Csx30 in a target RNA-dependent fashion in A|S, which is different from the sites found in two other studied Craspases (L|D and M|K for *Candidatus* "Scalindua brodae" and Desulfonema ishimotonii, respectively). The fact that Craspase cleaves a nonconserved site across orthologs indicates the evolution of specific protein interactions between Craspase and its respective Csx30 target protein. The Craspase family thus represents a panel of proteases with different substrate specificities, which we exploited for the development of a readout for multiplexed RNA detection.

A modified version of this chapter has been published as *van Beljouw et al., ACS Chemical Biology (2024)* [253].

6.1. MAIN TEXT

The discovery of Craspase revealed a novel feature in the CRISPR-Cas family: coupling of sequence-specific nucleic acid detection to protease activity, all within the same effector complex. Binding of target RNA to Craspase relays a conformational change to activate the protease, which in turn cleaves CRISPR-associated protein Csx30 in a specific position [233, 240, 254, 255]. Csx30 is part of a complex together with Csx31 and RpoE [254], two proteins that are often encoded in type III-E CRISPR-Cas loci [43, 206]. In transplanted *Escherichia coli* cells, cleavage of Csx30 in the Cx30-Csx31-RpoE complex instigates transcriptional pathways via the action of liberated RpoE [255] and induces cellular suicide to protect against phage infection via an unknown process [254]. Despite the bioinformatic prediction of at least ten Craspase orthologs [206], only two –the Craspases from *Candidatus* "Scalindua brodae" (*Sb*-Craspase) [233] and *Desulfonema ishimotonii* (*Di*-Craspase) [240]– have been experimentally described at the protease level.

Here, we set out to characterize the Craspase from *Candidatus* "Jettenia caeni" (*Jc*-Craspase), an anaerobic ammonium-oxidizing bacterium that plays important roles in the global cycling of nitrogen in marine environments [256]. *Jc*-Craspase cleaves *Jc*-Csx30 only in the presence of a target RNA (**Fig. 1A**), in line with observations from the studied Craspase orthologs [233, 240]. Mass spectroscopy on the *Jc*-Csx30 protein fragments revealed the processing site between alanine (A434) and serine (S435) (**Fig. 1B, Fig. 1C**), which is completely different from the other known Craspase cleavage sites (L|D for *Sb*-Craspase [233, 254] and M|K for *Di*-Craspase [240, 255]). This is remarkable given the close phylogenetic relatedness of the gRAMP orthologs, especially those from *Candidatus* "S. brodae" and *Candidatus* "J. caeni" [206]. So contrary to other proteases, which almost always display strong conservation in recognition sites across species [257, 258], Craspase orthologs appear to cleave at a nonconserved site near the C-terminal of target protein Csx30 (**Fig. 1D**). Individual amino acid substitutions of the eight residues surrounding the cleavage site of *Jc*-Csx30 (P4–P4′: AIAA|SSEK) were cleaved with a similar efficiency by *Jc*-Craspase compared to wild-type *Jc*-Csx30 (**Fig. 1E**). These findings strengthen the hypothesis that structural positioning, rather than the identity of the amino acids surrounding the cleavage site, is important for Craspase protein cleavage. While mutational analysis of *Sb*-Csx30 and *Di*-Csx30 similarly did not reveal Csx30 residues that are absolutely essential for cleavage [233, 255], some substitutions (especially in P1) were found to mildly [233, 255] or severely [240, 254, 259] affect the protease activity in the studied timeframes. This implies that the amino acids constituting the Csx30 cleavage site may vary in importance among different orthologs.

Substrate recognition in the physical context of the target protein is reminiscent of gasdermin cleavage by certain eukaryotic caspases [260]. It is believed that the high affinity interaction between caspase and gasdermin provides priming interactions for the insertion of the targeted protein region into the catalytic pocket, bypassing the requirement for a specific motif in the cleavage site. In Craspase, the C-terminal of Csx30 was pinpointed to be the primary protein component for successful Craspase interaction and cleavage [255]. Analysis of the electrostatic surfaces of the TPR-CHAT proteolytic pocket at the Csx30 binding site revealed a large variety in charge distribution between orthologs (**Fig. 2A**), most likely presenting the molecular basis for differential Csx30 recognition. Indeed, we did not observe cleavage of *Jc*-Craspase on *Sb*-Csx30 nor

did we see cleavage of *Sb*-Craspase on *Jc*-Csx30 (**Fig. 2B**), indicating that naturally occurring combinations of Craspase and Csx30 evolved to be specific for each other. The charge complementarity between the CHAT active pocket and the linker of its native Csx30 substrate (**Fig. 2A**) likely contributes to the absence of cross-reactivity. Analysis of Craspase structures in interaction with Csx30 should allow further detailing of the structural recognition code.

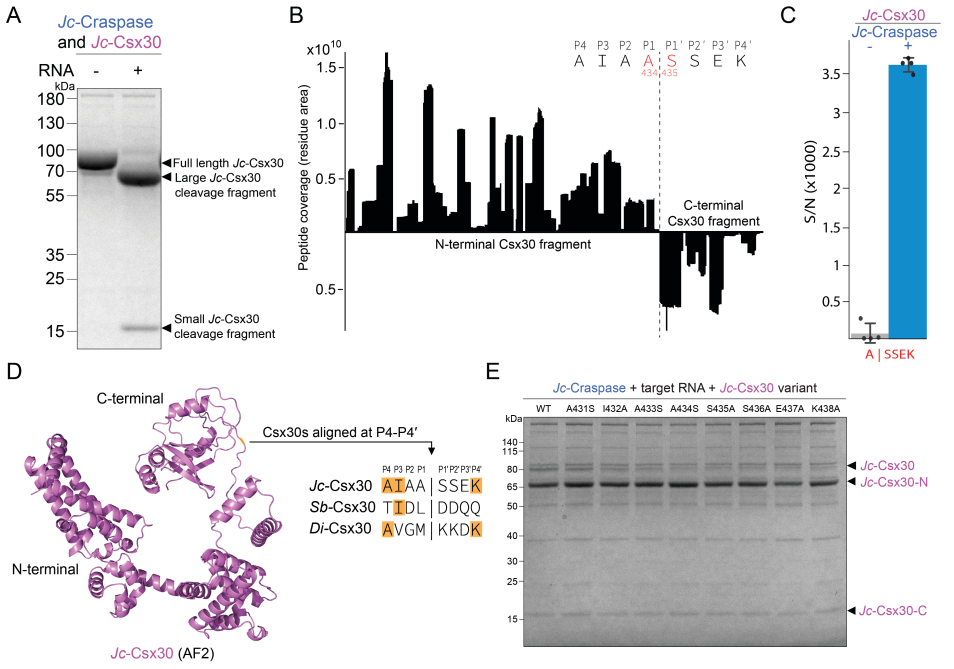


Figure 1. Craspase from *Candidatus* "Jettenia caeni" cleaves Csx30 at A|S. (**A**) Protein gel after *Jc*-Craspase incubation with *Jc*-Csx30 in the presence or absence of a target RNA. (**B**) Peptides identified by MS in the large and small *Jc*-Csx30 cleavage fragments, mapped onto the full *Jc*-Csx30 sequence to reveal the cleavage site at A|S. The eight amino acids residues surrounding the cleavage site (P4–P4′) are shown. (**C**) MS detection of peptides containing the cleavage site A|S in *E. coli* lysates enriched with *Jc*-Csx30, in the absence and presence of *Jc*-Craspase bound to target RNA. S/N indicates signal to noise ratios, as determined by TMTpro reporter ion quantification. (**D**) AlphaFold2 [236] (AF2) model of *Jc*-Csx30 with the position of the cleavage site indicated in orange. Csx30 orthologs are aligned at the cleavage site. Matching residues are indicated in orange. (**E**) Protein gel after incubation of *Jc*-Craspase with target RNA and *Jc*-Csx30 variants containing individual amino acid substitutions around the cleavage site. The large N-terminal and the small C-terminal cleavage fragments are indicated as *Jc*-Csx30-N and *Jc*-Csx30-C, respectively.

We sought to exploit the specificity of Craspase orthologs by creation of a multiplexed Craspase-based assay for detecting and discerning multiple RNA variants in the same reaction. In this assay, the Craspase orthologs are each programed with a specific crRNA to recognize different RNA species and are combined with their native Csx30 proteins. Upon incubation with an RNA sample, the pattern of the resulting Csx30 C-terminal fragments (16 and 18 kDa for cleaved *Jc*-Csx30 and *Sb*-Csx30, respectively) enables de-

duction of which Craspase was activated and consequently which RNAs were present in the sample (**Fig. 2C**). We could successfully detect and discern either one RNA variant or both variants in a single reaction, demonstrating multiplexed RNA detection (**Fig. 2C**). To circumvent the need for a protein gel readout, a recent study fused a fluorescent protein to the C-terminal side of Csx30 [261]. This allows measurement of the time-dependent increase in fluorescence intensity upon Csx30 cleavage, allowing picomolar detection sensitivity.

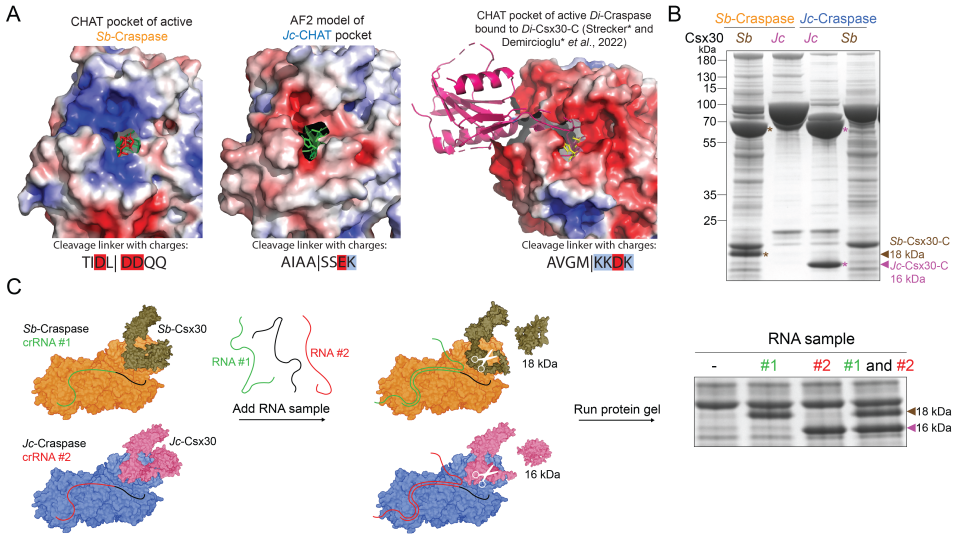


Figure 2. Proteolytic activity of Craspase orthologs is specific toward the corresponding Csx30 target protein. (**A**) Electrostatic surface maps of the CHAT pockets from TPR-CHAT orthologs and the charges in the cleavage sequence of the corresponding Csx30. Red indicates a negative charge, blue indicates a positive charge, and white indicates a neutral charge. Proteolytic residues are represented as sticks in red (*Sb*-TPR-CHAT), green (*Jc*-TPR-CHAT), and yellow (*Di*-TPR-CHAT). *Di*-Csx30-C is in pink. (**B**) Protein gel after incubation of *Jc*-Craspase and *Sb*-Craspase with either *Jc*-Csx30 or *Sb*-Csx30 in the presence of target RNA. Cleavage products of *Sb*-Csx30 and *Jc*-Csx30 are indicated with brown and pink asterisks, respectively. (**C**) Schematic of the multiplexed RNA detection assay in which *Sb*-Craspase and *Jc*-Craspase are programed with specific crRNAs to recognize target RNAs of choice. Upon addition of an RNA sample, each ortholog responds to a different RNA variant to cleave the respective Csx30 protein. This yields C-terminal Csx30 fragments with different sizes (18 kDa for *Sb*-Csx30 and 16 kDa for *Jc*-Csx30). The presence or absence of bands at 18 and 16 kDa on a protein gel reveals whether the target RNAs were in the RNA sample.

In conclusion, we highlight that although Csx30 cleavage by RNA-activated Craspase is conserved across species, Craspase orthologs do not recognize a consensus motif. This provides distinct benefits for biotechnological application that require multiplexed cleavage specificities, for which we provide a proof-of-principle in this study. By programming the Craspases to recognize specific target RNA of pathogens, such as COVID-19 variants [261], one could exploit this assay for point-of-care diagnostics. This is similar to the Cas13- and Cas12-based multiplexing assays, where multiple variants can be detected in a one-pot reactions using effector orthologs and different RNA or DNA reporters [67, 262]. An advantage of the Craspase multiplexing assay is its use of protein-

based reporters, which may provide useful when reporter stability is desired. The assay could in principle be expanded to detection of more RNA variants by adding in additional appropriately programmed Craspase orthologs, provided that all are specific towards their own Csx30. So, besides precise and controllable protease action, Craspase possesses yet another unique feature: a nonconserved cleavage site across protease orthologs. This, in combination with its structural recognition requirements, renders Craspase a protease with high selectivity for Csx30 and potentially a narrow range of additional substrates.

6.2. MATERIALS AND METHODS

PLASMID CLONING

Similar to *Plasmid cloning* in the Materials and Methods section of **Chapter 3**. Used primers and ordered DNA sequences are listed in [2].

To construct pJc-gRAMP-CRISPR, a coding sequence for an *E. coli* codon-optimized *Jc*-gRAMP protein was cloned downstream of a N-terminal Twin-Strep-Tag II, SUMO-tag, and the LacI repressed T7 promoter on the plasmid 13S-S (encoding spectinomycin resistance for selection) (Berkeley MacroLab). For the crRNA part, a CRISPR array starting with a LacI repressed T7 promoter and the native *Candidatus* "Jettenia caeni" leader sequence, followed by six native repeats interspaced by five times the twelfth spacer in the native CRISPR array (5′-CAAGGACGTTGGGAGAAACCAGTATATCAATCGCAAG) was cloned in the same plasmid. To construct pJc-TPR-CHAT, a coding sequence corresponding to *E. coli* codon-optimized *Jc*-TPR-CHAT protein variant was cloned downstream of the LacI repressed T7 promoter on the plasmid pACYC Duet-1. To construct pJc-Csx30, a coding sequence corresponding to an *E. coli* codon-optimized *Jc*-Csx30 protein variant was cloned in the place of *Sb*-Csx30 on pTag-Csx30.

PROTEIN PURIFICATION

Similar to the purification of gRAMP-crRNA in *Purification of Sb-gRAMP-crRNA and Sb-gRAMP variants* in the Materials and Methods section of **Chapter 3**. No cOmplete™ EDTA-free Protease Inhibitor Cocktail was used.

Jc-CSX30 CLEAVAGE ASSAYS

Jc-Csx30 cleavage reactions were performed in 10 μL reaction volume, containing 5 μL of 5 ng/μL *Jc*-Craspase, 5 μL of 6 ng/μL *Jc*-Csx30 protein, 2 μL of 50 μM in vitro generated target RNA, 100 mM Tris, 150 mM NaCl and 10 mM DTT. Reactions were run for 1 hour at 37 °C.

CLEAVAGE POSITION DETERMINATION

In **Fig. 1B**, mass spectrometry was performed similarly to *In-gel proteolytic digestion and protein identification* in the Materials and Methods section of **Chapter 4**. Protein digestions were performed using chymotrypsin. Detected peptides were mapped onto the *Jc*-Csx30 amino acid sequence and are listed in [2].

In **Fig. 1C**, *Jc*-Csx30 (20 pmol) was incubated overnight (18 hours, 30 °C, 350 rpm) with or without *Jc*-Craspase (0.5 pmol) and its target RNA (10 pmol) in four biological replicates. The solution consisted of 25 μl with a final concentration of 100 mM 4-(2-Hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES), 100 mM NaCl, 10 mM Dithiothreitol (DTT), and 5 mM ethylenediaminetetraacetic acid (EDTA) in MilliQ water.

The digestion was stopped with an inactivation mix for a final volume of 100 μl, containing 2.5 M GuHCl, 250 mM HEPES, 10 mM tris(2-carboxyethyl)phosphine (TCEP), 40 mM 2-chroloacetamide (CAA). *E. coli* lysates (cells lysed under native conditions with a hypotonic buffer and probe sonication) were also added to the samples for a total amount of 50 μg. The mix was incubated at 95 °C, 600 rpm for 10 minutes to denature proteins and reduce-alkylate cysteine residues. The individual replicates were labelled

with 200 μg of TMTpro reagents, incubated for 60 minutes at room temperature. The labeling reaction was quenched with 100 mM ammonium bicarbonate (AMBIC), incubated for 30 minutes at room temperature. Individual samples were pooled and SP3 cleanup was performed to remove excess reagents, with the beads finally resuspended in 100 mM HEPES for a final concentration of 1 μg/μl. Protein digestion was performed with 1:50 trypsin: protein ratio overnight (20 hours, 37 °C, 350 rpm). Beads were sonicated and the supernatant was transferred to a new tube, with 10% of the digest removed as the non-enriched sample. Tryptic peptides in the enriched fraction were tagged with undecanal at a ratio of 1:50 protein: undecanal and 50 mM sodium cyanoborohydride ($NaBH_3CN$) in 40% ethanol (EtOH), and incubated for 90 minutes at 50 °C, 450 rpm. Tryptic peptides were depleted by loading the solution to a conditioned SepPak (Waters, 50mg capacity) desalting column, and retrieving the flowthrough. An estimated 500 ng of both non-enriched and N-termini enriched samples were acidified and loaded on EvoTip Pure trap columns with the low input protocol, and queued for mass spectrometry analysis.

Samples were measured with an EvoSep One liquid chromatography (LC) system in line to an Orbitrap Eclipse trybrid mass spectrometer, equipped with a FAIMSpro ion mobility device. Peptide separation was performed with the 20SPD 58 minute gradient method using an Aurora Elite TS Generation 3, 15 cm column (IonOpticks). Peptides were injected with nano spray ionization in positive ion mode with a spray voltage of 2300 V, ion transfer tube temperature of 240 °C, and carrier gas flow of 3.6 L/minute. The instrument was operated in data dependent acquisition (DDA) mode, and two experiments utilizing different compensation voltages (CV) were used during measurement (-45 and -65 V), with otherwise identical settings. MS scans were acquired in the Orbitrap at 60,000 resolution, with a scan range of 375-1500, maximum injection time (IT) at 118 ms, normalized AGC target at 300% and RF lens at 40%. The filters used for precursor selection were: MIPS mode peptide, allowed charged states between 2-7, dynamic exclusion after 2 times for 30 seconds with 10 ppm tolerance, minimum intensity threshold at 20,000, and precursor fit at 70% with a 1.2 m/z fit window. MS/MS scans were recorded with a total cycle time of 1 second for each CV. Precursors were isolated in the quadrupole with 1.2 m/z isolation window and fragmented with HCD at NCE 34%. Scans were acquired in the Orbitrap in centroid mode with a resolution of 30,000, maximum IT of 54 ms, and normalized AGC target at 100%.

Raw data were searched with Proteome Discoverer v2.4. The non-enriched and enriched samples were added as fractions, and TMTpro quantification was selected. The Sequest HT engine was used for PSM detection, and Percolator was used for FDR control (1% strict, 5% relaxed). Craspase and Csx30 sequences were added to the *E. coli* reference proteome (UniprotKB, 4,360 sequences, accessed 17/01/2023). The search was performed with ArgC specificity with semi-specific N-terminal search, with peptide length between 6-46. Methionine oxidation (+15.995 Da), asparagine deamidation (+0.984 Da) N-terminal acetyl (+42.011 Da) and TMTpro (+304.207 Da) were added as variable modifications, while cysteine carbamidomethylation (+57.021 Da) and lysine TMTpro modification (+304.207 Da) were added as fixed modifications. TMT quantification was performed on unique and razor peptides using the Reporter Ions Quantifier, with normalization on median of total peptide amount per channel (N-terminal TMT-

pro modified peptides excluded). Filters for quantification were set to 10 S/N threshold, and 50% co-isolation threshold. Detected peptides are listed in [2].

IN VITRO GENERATION OF RNA

A gBlock containing the T7 promoter and target RNA (complementary to the crRNA in purified *Jc*-Crapase) were synthesized (IDT) and PCR amplified with 5′-TCGATCAGAGCGCTCTTACG. ~500 ng of purified PCR fragment was in vitro transcribed overnight using HiScribeTM T7 High Yield RNA Synthesis Kit (NEB) and subsequently treated with DNase I according to the manufacturer's protocol. For RNA extraction, acidic phenol (pH 4.5, phenol:chloroform = 5:1, Invitrogen) was added to the sample in a 1:1 ratio, vortexed for 1 minute and centrifuged for 10 minutes at 13,200 rpm at room temperature. The aqueous phase was collected and subjected to RNA precipitation (20 μL 3M NaAcetate and 500 μL 100% ethanol per 200 μL of sample) for 1 hour at -20 °C. Samples were centrifuged at 13,200 rpm at 4 °C for 2 hours, washed twice with ice-cold 70% ethanol and centrifuged at 13,200 rpm at 4 °C for 10 minutes. The pellet was dried in a SpeedVac concentrator (Thermo Fisher Scientific) for 30 minutes at 60 °C and resuspended in RNA grade water.

MULTIPLEXED CRASPASE ASSAY

For the multiplexed Craspase assay, reactions were prepared containing purified proteins (5 μL of 5 ng/μL *Jc*-Craspase, 2.5 μL of 2.5 ng/μL *Sb*-Craspase R294A D698A, 5 μL of 5 ng/μL *Sb*-Csx30, and 5 μL of 6 ng/μL *Jc*-Csx30). 2 μL of 50 μM in vitro generated RNA cognate to the crRNA in *Jc*-Craspase and/or *Sb*-Craspase was added and reactions were incubated at 37 °C for 2 hours.

PROTEIN GELS

Similar to *SDS-PAGE analysis* in the Materials and Methods section of **Chapter 3**.

7

DISCUSSION

"Philosophy begins in wonder. And, at the end, when philosophic thought has done its best, the wonder remains."

Alfred North Whitehead

With the discovery of CRISPR-controlled proteases, CRISPR-Cas has moved beyond mere nucleic acid targeting into the territory of targeted protein degradation. Here, we review the understanding of Craspase, the best-studied member of the growing CRISPR RNA-guided protease family. We recollect the original bioinformatic prediction and early experimental characterizations; evaluate some of the mechanistic structural intricacies and emerging biotechnology; discuss open questions and unexplained mysteries; and indicate future directions for the rapidly moving field of the CRISPR proteases.

A modified version of this chapter has been published as *van Beljouw & Brouns, Biochemical Society Transactions (2024)* [2].

7.1. INTRODUCTION

The host-virus arms race serves as obvious facilitator for the emergence of biological complexity, as both the host and invader are challenged to evolve novel traits to outcompete each other for survival. These endless creative impacts resulted in a large repertoire of prokaryotic immune pathways aimed at impeding the propagation of invading viruses [204, 263, 264]. Among the myriad bacterial immune systems identified to date, CRISPR-Cas adaptive immunity stands out as one of the most extensively studied.

The CRISPR-Cas family became well-known for its antiviral activity through guide-defined destruction of DNA [265] and RNA [1]. With the discovery of CRISPR RNA-activated proteases, proteins are added to the repertoire of cleavage targets. The best studied CRISPR protease is Craspase, which upon target RNA activation cleaves a host-encoded protein to instigate a variety of cellular effects. Since the discovery of the predicted genomic loci containing the Craspase module, a number of studies looking into its biology and biotechnology have been published. Here, we distil the obvious and the obscure from these works to present an overview of Craspase. We describe early milestones (e.g. bioinformatic discovery and experimental description of gRAMP and Craspase), ponder upon research findings (e.g. the biological consequences of Craspase proteolytic activity, the biochemistry of Craspase proteolysis, the big insertion domain, the impacts of biotechnology, the fusion nature of gRAMP, the course of TPR-CHAT and gRAMP evolution, and the crRNA dependency of gRAMP), and evaluate the other members of the expanding CRISPR protease family (CalpL and SAVED-CHAT). We end with a subset of the many open questions that remain to be answered, guiding forthcoming progress in this exciting new CRISPR-Cas field.

7.2. CHARACTERIZATION OF GRAMP AND CRASPASE

In the summer of 2021, the single-unit nature of the gRAMP effector was proven [28, 206]. Purified gRAMP —also referred to as Cas7-11 at this point due to the multiple Cas7-like and a Cas11-like domains in its architecture— cleaves target RNA in two guide-defined positions 6 nucleotides apart, reminiscent of the cleavage periodicity in canonical type III effectors. The guide in the gRAMP protein can be re-programmed to cleave an RNA of choice, facilitating usage for RNA knockdown in human cells. Type III-E spacer analysis revealed that some of the crRNAs are complementary to transcripts of MGE genes, suggesting that gRAMP immunity functions through recognizing and cleaving the mRNA of those invaders. But what was the role of the co-localizing predicted protease, TPR-CHAT? Strikingly, co-expression experiments revealed that TPR-CHAT forms a stable complex with gRAMP [206] —technically making the type III-E effector multi-subunit and thus meeting the class 1 classification requirement after all—, without disabling the capacity of gRAMP for binding and cleaving target RNA. The CHAT domain is part of the caspase family, containing proteases that are responsible for programmed cell death [266]. This raised the possibility that binding of RNA to the gRAMP-TPR-CHAT complex, termed Craspase (for 'CRISPR-guided caspase'), unleashes the action of the protease, presumably to instigate cell death. Around this time, a report on a homologous bacterial TPR-CHAT showed precise proteolytic activity on the host-encoded gasdermin [209]. This example hinted that Craspase similarly cuts a specific protein en-

coded by the host, instead of having promiscuous cell-wide activity or activity against a viral protein. However, this hypothesis could not be tested, as target proteins of Craspase were yet to be identified in the vast bacterial proteome. It turned out to be hiding in plain sight.

The hypothesis of RNA-activated protease activity turned out to be correct [233, 240, 254, 255, 259, 267–271]: incubation of Craspase with an RNA complementary to the guide RNA leads to the destruction of Csx30, a conserved protein of unknown function, usually encoded in the type III-E operon. Detailed structural analysis revealed an α-helix in TPR-CHAT, termed 'switch helix', responsible for relaying steric interaction with the target RNA into conformational changes of the peptidase pocket. This facilitates the digestion of Csx30 into two fragments, a large N-terminal and a small C-terminal fragment, mediated by the cysteine-histidine catalytic dyad in the CHAT domain. Intriguingly, Craspase possesses internal self-control, as cleavage of the target RNA presents the off-switch for the protease. This suggested that the main function of RNase activity in Craspase could be protease regulation rather than stopping the invading MGE through interference with its mRNA, a thought borrowed from the RNA-regulated cyclic oligo adenylate (cOA) production in other type III effectors [1].

At this point, the full functional pathway of Craspase —assembly of TPR-CHAT and gRAMP to form Craspase, target RNA binding to activate the protease, Csx30 cleavage, and target RNA cleavage to deactivate the protease— was characterized (**Fig. 1**). But the biological implications of Csx30 cleavage in viral immunity, amongst other mysteries, remained to be elucidated.

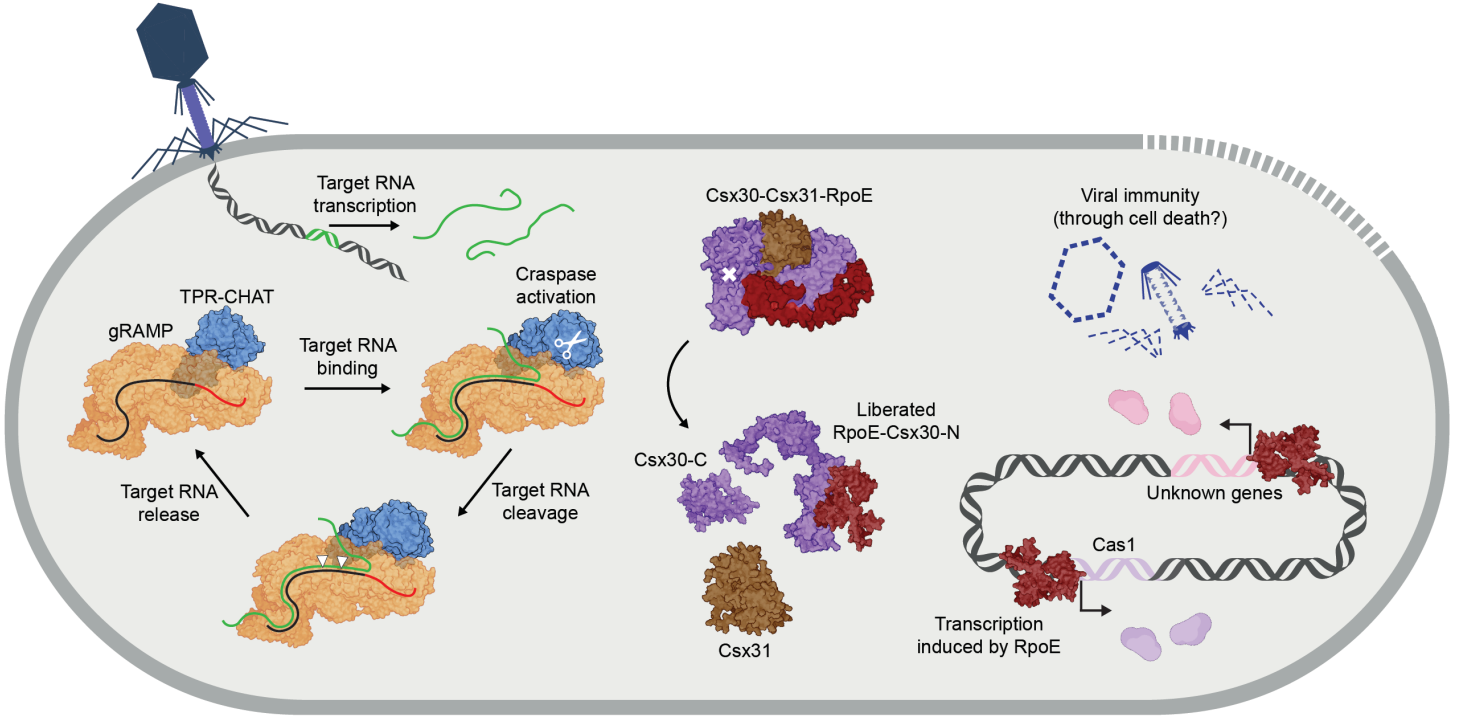


Figure 1. The life-cycle of Craspase. Once an invading virus has progressed to transcription, Craspase will bind a viral transcript complementary to the loaded CRISPR RNA. Target RNA binding will stabilize the protease pocket of Craspase in the active conformation, allowing recognition and cleavage of Csx30 (cleavage site indicated with white cross) in the Csx30-Csx31-RpoE complex. This leads to the liberation of the N-terminal part of Csx30 bound to the transcription factor RpoE, which presumably binds to various sites in the host chromosome, including the promoter region of Cas1 and several unknown genes. Craspase activity is thought to provide viral immunity through an unknown abortive infection pathway.

7.3. THE BIOLOGY, BIOCHEMISTRY, AND BIOTECHNOLOGY OF CRASPASE

7.3.1. BIOLOGICAL CONSEQUENCES OF CSX30 CLEAVAGE

In deciphering the consequences of Csx30 cleavage by Craspase, the type III-E operon turned out to hold more answers. It was found that Csx30 assembles into a complex with Csx31 and RpoE [254], two proteins that are often encoded adjacent to Csx30. In transplanted *Escherichia coli* cells, cleavage of Csx30 in the Cx30-Csx31-RpoE complex induces cellular suicide to protect against invading viruses (**Fig. 1**). This renders Craspase an instrument of altruism, facilitating the sacrifice of the individual for the population, a strategy not uncommon in the bacterial kingdom [1, 209]. But how is the cell death phenotype accomplished, and what happens to the Cx30-Csx31-RpoE upon cleavage? The function of Csx31 is currently unknown, and the limited sequence similarity to other proteins makes its role elusive. The involvement of RpoE, annotated as a sigma factor that facilitates transcription-initiation by recruiting the RNA polymerase to promoter sites in DNA [272], is clearer. It was found that Csx30 keeps RpoE —also termed CASP-σ (for 'CRISPR associated protease sigma factor')— in an inactive state, perhaps through sterically preventing access to the DNA [255]. Cleavage of Csx30 mediated by Craspase —also termed CASP (for 'CRISPR-associated protease')— relieves this inhibition by releasing RpoE-Csx30-N from the complex [255, 268], allowing RpoE/CASP-σ to interact with a DNA binding motif present in various locations in the native organism to potentially drive expression of the upstream genes (**Fig. 1**). One of these genes encodes for Cas1, a key player in the acquisition of novel spacers [273], suggesting that Craspase regulates CRISPR memory formation during phage infection. Also, a putative membrane protein was found near an RpoE/CASP-σ binding site, making membrane pore formation a possible candidate for how the proposed cell death is instigated during type III-E immunity. It is to be noted that cell death has only been observed in transplanted *E. coli* through an unclear molecular pathway, so whether cell suicide is a true phenomenon in the native context remains to be investigated.

As liberated RpoE levels depend on Craspase activity and thus target RNA concentrations, the type III-E pathway effectively links viral load to gene expression levels (**Fig. 2A**). This scenario is analogous to signal amplification in canonical type III CRISPR-Cas effectors, which produce cOA second messenger molecules proportionally to the intracellular target RNA levels [38]. A more speculative layer of scaling in type III-E involves differential promoter affinity for RpoE, in which strong binding of RpoE to some promoters and weak binding to others leads to high or low protein expression, respectively (**Fig. 2B**). In this model, protein expression from strong promoters occurs readily at low doses of free RpoE, whereas significant expression from weak promoters only takes place when a large quantity of liberated RpoE is present. Genes expressed from weak promoters potentially encode the suicide executor proteins, preventing unnecessary activation during mild infection. Proteins with less severe effects, such as Cas1, could be driven by strong promoters for the cell to benefit its enzymatic activity even at low target RNA levels. Exploring the effects of altered expression patterns resulting from RpoE activity is an interesting research direction.

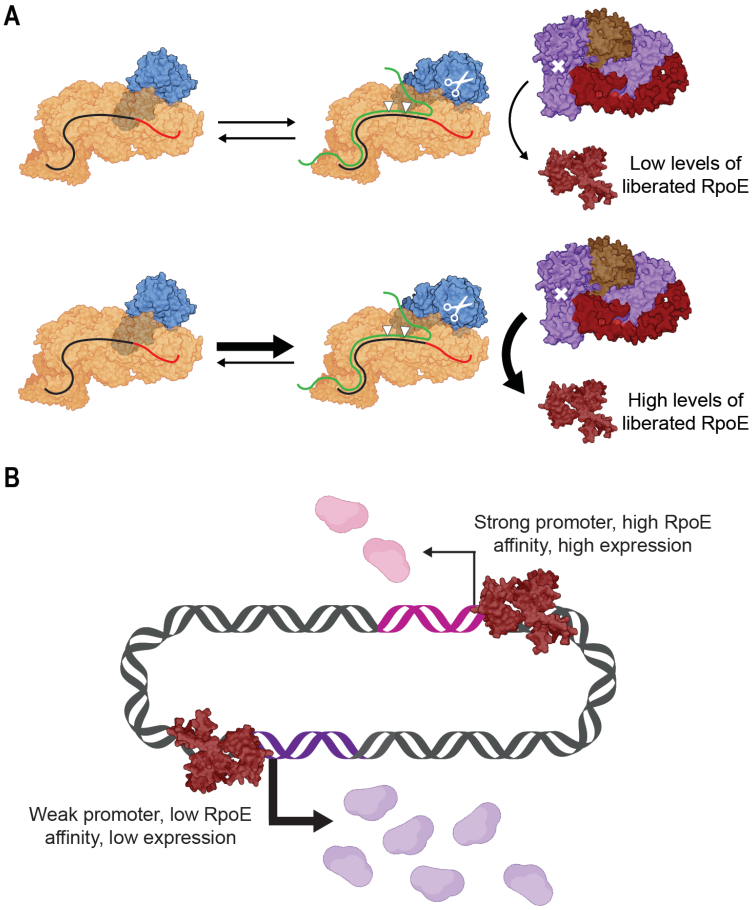


Figure 2. Potential tunability layers in the Craspase pathway. (**A**) The protease activity of Craspase is activated upon binding a target RNA, resulting in cleavage of Csx30 (purple, cleavage site indicated with white cross) to liberate RpoE, whereupon cleavage of the bound target RNA stops proteolytic activity and thus halts RpoE liberation. High levels of target RNA lead to quick replenishment after release of the cleaved RNA fragments from Craspase, effectively increasing the time Craspase spends in the activated state compared to low target RNA levels. This potentially increases the Csx30 cleavage rate and consequently the levels of liberated RpoE. (**B**) Free RpoE binds promoter sites in the host chromosome, driving the expression of various genes. Differential promoter binding strengths allow for variable expression of those genes, adding a layer of intracellular tuning caused by Craspase activity.

7.3.2. CRASPASE PROTEOLYTIC BIOCHEMISTRY

It was hypothesized that the biochemical cause for Craspase protease activation lies in the decreased distance between the catalytic histidine and the cysteine upon target RNA binding, as structural analysis showed a reduction from approximately 7 Å to approximately 3 Å [233, 254]. The reduced distance should allow for deprotonation of the cysteine by the histidine, followed by nucleophilic attack on the substrate and rupture of the scissile bond (**Fig. 3A**).

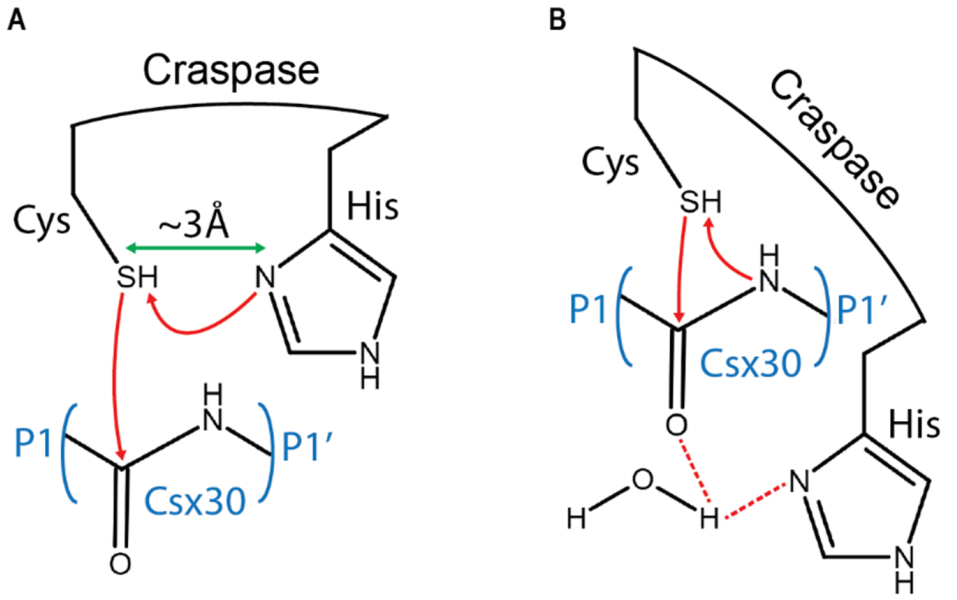


Figure 3. Two models for the proteolytic biochemistry of Craspase. (**A**) Close distance between the cysteine and histidine of the catalytic dyad in Craspase allows for deprotonation of the cysteine by the histidine, facilitating a nucleophilic attack at the scissile bond between P1 and P1′ amino acid positions of the target protein Csx30. This model is adopted from the classical biochemical model for cysteine proteases. (**B**) Alternatively, similar to human legumain and caspases, deprotonation of the cysteine could occur directly by the substrate, followed by the nucleophilic attack.

Although this explanatory model is widespread in the literature on cysteine proteases [274], also an alternative biochemical scenario has been proposed [275]. As fitting of the target protein substrate into the active pocket of caspase requires the cysteine and histidine residues to be separated beyond the hydrogen bonding distance, deprotonation of the cysteine by the histidine was considered unlikely. Instead, molecular density functional calculations suggested that deprotonation of the cysteine directly by the substrate is the most plausible atomistic description (**Fig. 3B**). This mode of proteolytic biochemistry was determined based on human legumain and caspases, so it remains to be investigated whether it is true for Craspase. It is noteworthy that the distance between the catalytic histidine and cysteine in inactive TPR-CHAT is similar to the distance between these residues in inactive human separase [268], hinting at a potential convergence in activation principles. Solving cryo-EM structures of Craspase in interaction with the target protein Csx30 will likely decipher the atomic details at the protease interface, revealing the extent to which Craspase proteolysis works akin to human caspases.

7.3.3. THE BIG INSERTION DOMAIN IN GRAMP

The most conspicuous structural divergence in gRAMP compared to ancestral multi-subunit effector domains took place in the C-terminal Cas7-like domain (Cas7.4), which gained a big insertion domain (BID) protruding from its architecture. The structure of

BID could not be easily resolved [233, 240, 268], suggesting that it is flexibly probing the exterior surrounding gRAMP. The floppy and protruding characteristics make it plausible that the BID acts as a sponge to localize nucleic acids close to the effector for upregulation of target recognition (**Fig. 4**), akin to the strategy suggested for WYL1 in the type VI CRISPR-Cas effector Cas13 [125, 143]. Furthermore, the likely close ancestor of gRAMP, the type III-Dv complex, also contains an insertion in its Cas7 domain and was shown to interact with the 3′ end of the crRNA, where the seed region for target RNA binding was hypothesized to reside [276, 277]. Similarly, the BID in gRAMP is architecturally positioned at the 3′ end of the crRNA where the initial target recognition may take place, suggesting that sponge-like localization of nucleic acids at this region could immediately facilitate target RNA binding. Although a seed region for target RNA binding could be residing in the crRNA portion bound by the BID of gRAMP, the fact that target RNAs lacking complementarity to this crRNA region are being cleaved efficiently [233] shows that this is not an absolute requirement.

An analysis of Craspase from *Desulfonema magnum* dissected the BID in two parts, termed 'insertion-1' and 'insertion-2' [270]. Insertion-1 has direct interactions with the crRNA, whereas insertion-2 is protruding from the protein without crRNA contact. One could speculate that insertion-1 of the BID is involved in crRNA recognition, serving as an anchor for ribonucleoprotein formation. Insertion-2 might possess the proposed sponge functionality, equipping the BID with a double role in both crRNA recognition and RNA localization. A possible additional BID function is slowing down target RNA turnover, as it was observed that target RNA with complementarity extending into the crRNA portion bound by BID are cleaved less efficient compared to shorter ones [233]. Perhaps the BID physically prevents target RNA displacement, lowering target RNA binding kinetics and thus reducing cleavage turnover.

In Craspase from *Desulfonema ishimotonii*, the BID was shown to be dispensable without detectable loss in RNA targeting capacity [28], indicating that it has adopted a role beyond RNase functionality on target RNA. However, removal of the BID in Craspase from *Candidatus* "Scalindua brodae" distinctly reduces cleavage activity, which may be due to involvement in pre-crRNA processing [254]. Functional differences between orthologues Craspases, such as the role of the BID, accommodation of pre-crRNA processing, and regulatory mechanisms of TPR-CHAT (see section **On the origin of gRAMP**), perhaps reflect a need for subcategorization within the type III-E CRISPR-Cas effectors.

7.3.4. CRASPASE IN BIOTECHNOLOGY

It was quickly realized that the precise proteolytic action, strict RNA-dependent activation requirements, and self-regulatory capacity make Craspase a strong candidate for biotechnological applications. Craspase self-deactivation is prevented by mutating the RNase pockets in gRAMP, yielding a variant whose protease stays active (i.e. Stay-On Craspase) after target RNA binding [233]. This could provide distinct benefits for usage of Craspase where RNA recognition and protease activity are desired over RNA cleavage, such as in molecular diagnostics. RNA detection tools using Craspase were achieved in multiple ways, all exploiting the principle of coupling RNA recognition and subsequent protease activity to a fluorescence readout. One application involves a protein element recognized by cells for targeted protein degradation, called a degron, linked to fluores-

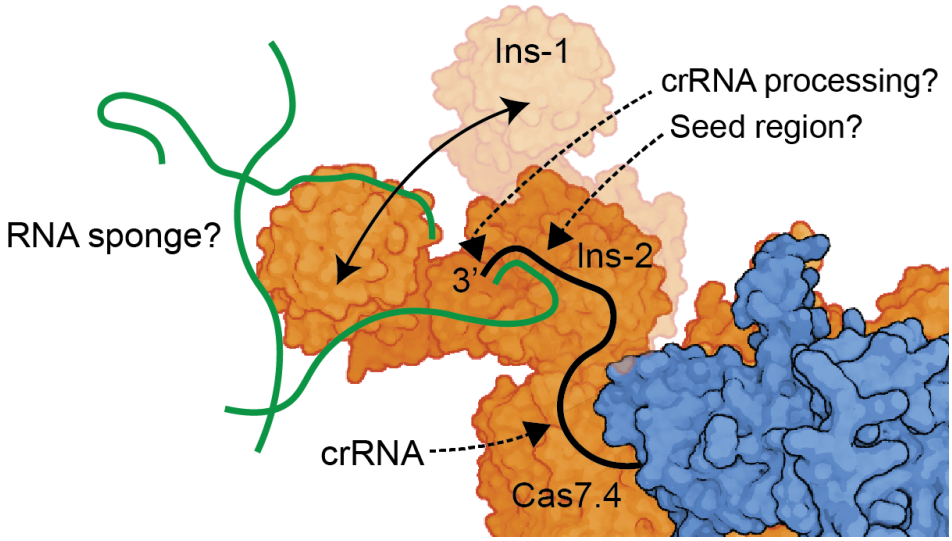


Figure 4. The big insertion domain in gRAMP. The big insertion domain (BID), dissected into insertion-1 (ins-1) and insertion-2 (ins-2) is protruding from Cas7.4 and flexibly probing the environment, perhaps to act as RNA sponge to localize potential target RNAs close to the suspected seed region. The BID in gRAMP from *Candidatus* "Scalindua brodae" has been suggested to also play a role in crRNA processing. TPR-CHAT is visualized in blue.

cent protein via Csx30 [240]. The fluorescent protein remains intact and thus fluorescent only when freed from the degron, which is achieved through Csx30 processing by activated Craspase. In another application, the RpoE promoter sequence was used to drive expression of a fluorescent protein only when Craspase was activated to cleave Csx30 for release of RpoE [255]. A third application involves a recombinase capable of activating an engineered promoter, which was cloned to drive expression of a fluorescent protein [255]. Tethering of the recombinase to the eukaryotic membrane through a part of Csx30 allowed nuclear localization and subsequent fluorescent protein expression only after cleavage by Craspase.

Nucleic acid manipulation and diagnostics using Craspase adds to the existing RNA-targeting CRISPR-Cas toolbox [1], but the unique biotechnological promise lays in its potential for precise protein editing. Currently, available proteases are either promiscuous, uncontrollable, or both, limitations that Craspase overcomes. Craspase is currently specific for the native target protein Csx30 or derivatives thereof, but in the future, Craspase may be engineered to cleave a protein of interest. This opens up a range of new biotechnological avenues, such as cancer therapeutics. Here, Craspase may be evolved to cleave a vital protein in human cells to selectively kill tumour cells, discriminating cancerous from normal cells through precise recognition of oncogenic transcripts. Current targeted therapeutics are based on the inhibition of oncogenic proteins [278], requiring laborious development of novel molecules for each oncogenic variant. Recognition of RNA by Craspase instead provides the benefit of easy reprogramming upon cancer mutagenesis, only requiring loading of Craspase with a guide cognate to the new oncogenic transcript.

7.4. EVOLUTIONARY PERSPECTIVES ON CRASPASE

7.4.1. ON THE ORIGIN OF GRAMP

The clear evolutionary relation between the domains in gRAMP and multi-subunit type III CRISPR-Cas effector complexes strongly suggests that one evolved from the other [43]. The relative abundance of multi-subunit effectors compared to relatively rare gRAMP points at an evolutionary trajectory from multi-subunit to gRAMP as the most parsimonious reconstruction. Likely representing an intermediate step on the evolutionary march from completely multi-subunit to fully fused type III systems is the type III-Dv effector, which is composed of a mix of free subunits (Cas10, Csx19, and Cas7 with an insertion) and fused subunits (Cas7-Cas7 and Cas7-Cas5-Cas11) [276].

What evolutionary benefits could have attributed to the emergence of subunit fusions in type III systems? One rationale is that the use of multi-subunit effectors comes with the risk of producing more subunits than required for the desired stoichiometry of the complex [279]. Fusing the subunits together into a single gene mitigates the unnecessary energy expended during subunit overproduction. Moreover, multi-subunit effectors are restrained in the functional landscape they can explore, as structural morphing of the subunits required for functional change could hinder proper assembly and function of the complex. Multi-subunit type III effectors are usually assembled from multiple Cas7 proteins that are encoded by the same gene and are thus phenotypically identical. Fusion of subunits frees the effector from the restriction of subunit assembly to allow for structural differentiation of the domains, as has become apparent from the idiosyncratic shapes each of the Cas7-like domains in gRAMP display [233] (**Fig. 5**). Such structural diversification allows for change of function, which could be evolutionary advantageous. An example of this may be found in the number of RNA cleavages that are made by the type III effector complexes. In canonical type III effectors, each Cas7 subunit cleaves the target RNA once, constraining the effector to make as much RNA cleavages as there are Cas7 proteins in the complex [1]. In gRAMP, two of the four Cas7-like domains apparently lost their RNA endonuclease activity of the target. As target RNA cleavage serves as the off-switch for protease activity in Craspase, this reduction in number of cleavage sites may have fine-tuned controllability of the protease. Interestingly, there are also multi-subunit effector complexes in which some Cas7 subunits do not cleave the target RNA [280]. It was suggested that the missed cleavage sites are not due to enzymatic incompetency, but rather due to non-cleavable structural architecture of the crRNA:target RNA duplex at the Cas7 interface.

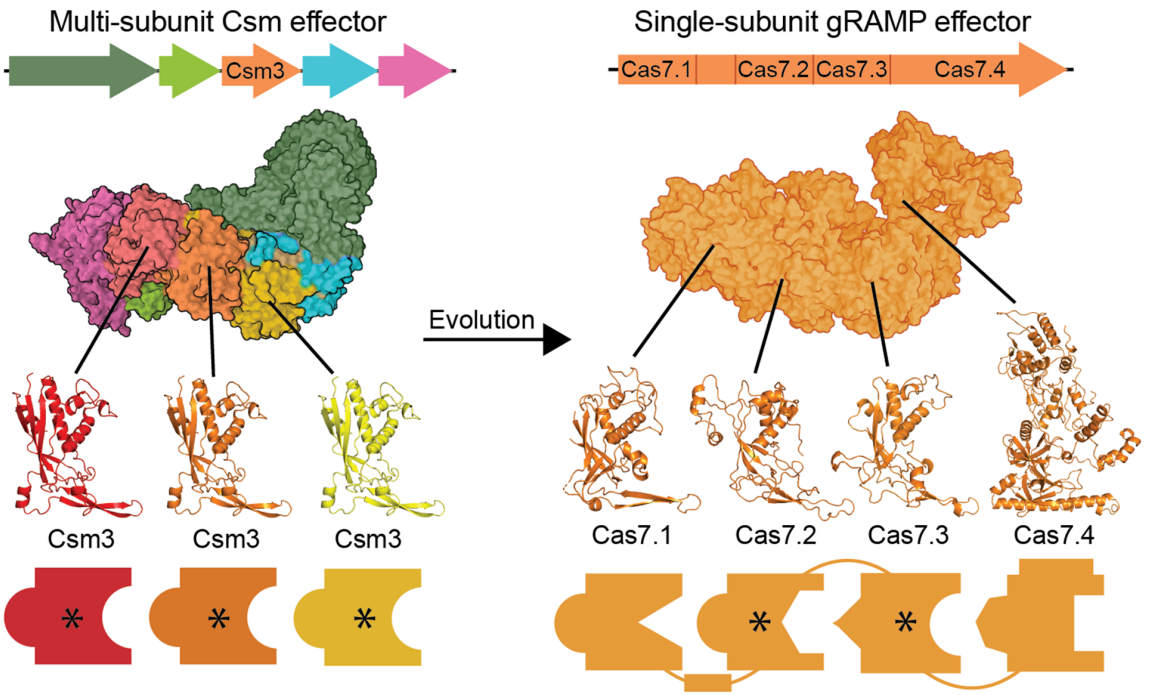


Figure 5. Evolution of gRAMP from the multi-subunit type III effector. The operon for the multi-subunit Csm effector (PDB 6IFN) generally harbours one gene encoding for the RNase Csm3 (Cas7) subunit, which provides multiple copies in the effector complex upon assembly with the other subunits. Structural (indicated with a puzzle piece) or enzymatic (indicated with an asterisk) changes to Csm3 would affect all protein copies in the complex, potentially hampering complex assembly and functional divergence, respectively. This restrains the possible evolutionary explorations that multi-subunit complexes can undertake. In contrast, the Cas7-like domains are fused together in gRAMP, alleviating the need for correct translation levels of the subunits, subunit assembly, and enzymatic homogeneity among the domains. Correspondingly, the Cas7-like domains evolved to take on vastly differing morphologies, most notable Cas7.4 with the big insertion domain. Moreover, contrary to the multi-subunit effectors, not all of them contain an RNase pocket that is active on target RNA.

Another example of how the gained degrees of freedom through subunit fusion resulted in expanded functionality is the observed flexibility of the Cas11-like domain in gRAMP from *D. magnum* [270]. Docking of TPR-CHAT onto gRAMP during formation of Craspase promotes the rigid conformation of Cas11, mediated by a flexible motif in gRAMP termed 'insertion finger'. Stabilized Cas11 results in reduced binding and cleavage of target RNA in Craspase compared to the stand-alone gRAMP, suggesting a role for TPR-CHAT in regulating RNase activity. In Craspase from both *Candidatus* "S. brodae" [233] as well as *D. ishimotonii* [28], a similar reduction in RNA accessibility between gRAMP and Craspase was observed. It was hypothesized that the interaction with TPR-CHAT stabilizes the 'gating loop', a long linker from Cas11 to Cas7.2 that forces target binding to initiate at the 3′ end of the crRNA [233]. Since the gating loop is displaced by incoming target RNA, the stabilized gating loop in Craspase effectively increases the conformational energy barrier for target RNA binding, thus lowering RNA accessibility. So it seems that TPR-CHAT regulates gRAMP in multiple ways, induced by physical interaction with the insertion finger and gating loop (**Fig. 6**).

It makes for interesting speculation whether gRAMP has a cellular role as a stand-alone protein, one that requires increased activity against target RNA compared to when it interacts with TPR-CHAT. Curiously, the insertion finger is not present in *Candidatus* "S. brodae", suggesting that the reduced RNA binding and cleavage activity in Craspase can be ascribed mainly to the rigidity of the gating loop.

7.4.2. THE EVOLUTIONARY RELATIONSHIP BETWEEN TPR-CHAT AND EUKARYOTIC SEPARASE

The CHAT domain in TPR-CHAT exhibits a significant degree of structural similarity to the eukaryotic separases, and both CHAT and separase contain the highly conserved histidine-cysteine catalytic dyad [268, 270]. This suggests a close evolutionary relationship between CHAT and separase. An elegant hypothesis for linking CHAT to the evolution of separase can be traced back to the formation of the early eukaryote. Eukaryogenesis is widely believed to originate from the endosymbiotic fusion of an archaeal and bacterial cell, with the latter eventually evolving into the mitochondrion [281]. During this process, the bacterial endosymbiont might have encoded CHAT-like domains, perhaps to kill the host cell once it became unhospitable [266, 282]. Through the process of gene migration from the DNA of the endosymbiont to the eukaryotic chromosome, the early eukaryote may have adopted these proteins for other purposes [283]. In this scenario, the acquisition of CHAT was followed by functional repurposing for chromosome-related processes, as separases now serve an essential role in eukaryotic chromosome segregation [243]. In the bacterial evolutionary lines, CHAT evolved to have functions in various pathways, such as pyroptosis [209], cell division [284] and, as seems the case for Craspase, viral immunity.

7.4.3. CRRNA DEPENDENCY IN GRAMP

Structural analysis of gRAMP revealed that the crRNA is threaded through the entire protein, engaging in a multitude of RNA-protein interactions. This, combined with the observed decreased stability of gRAMP in the absence of crRNA [206], suggests an essential stabilizing role of crRNA for the conformation of the ribonucleoprotein. Contrary to

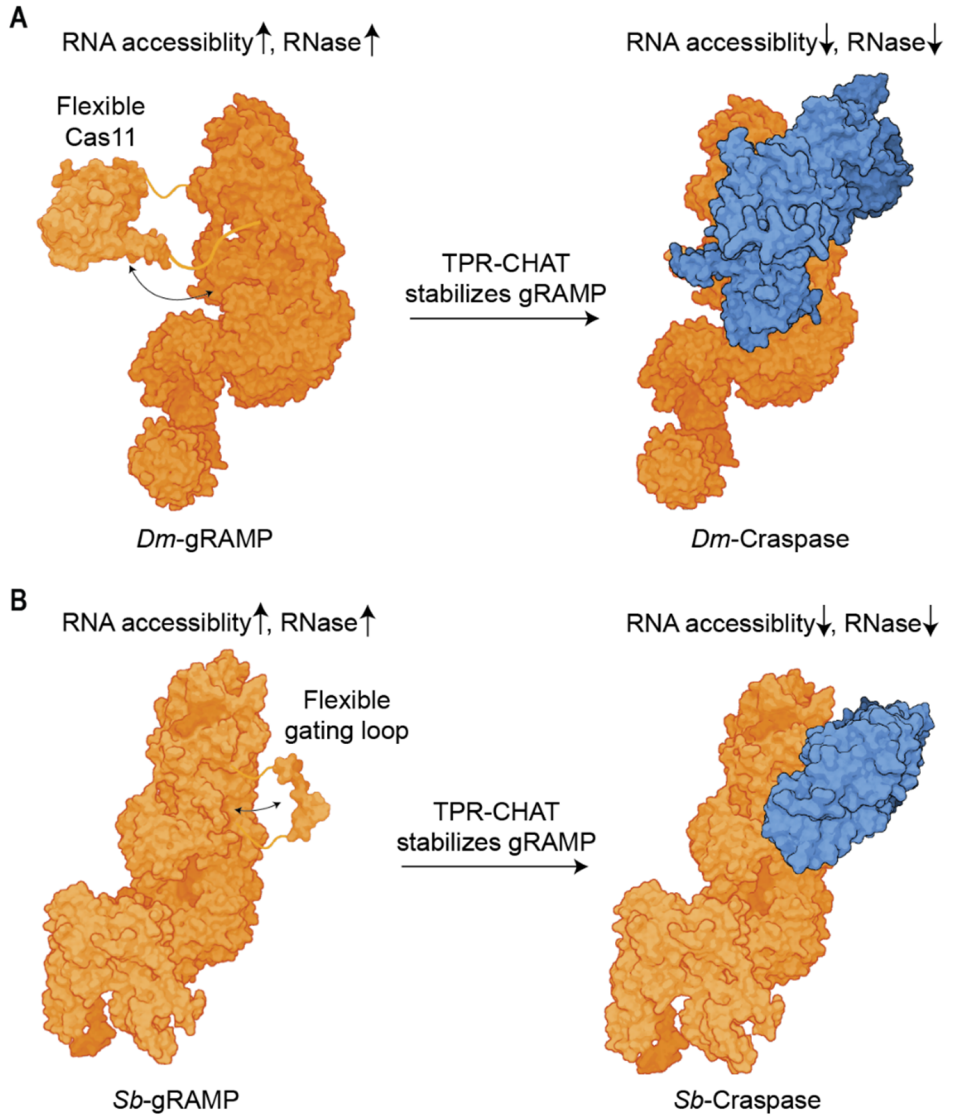


Figure 6. gRAMP regulation by TPR-CHAT. The flexible Cas11-like domain in gRAMP from *Desulfonema magnum* (*Dm*-gRAMP) (**A**) and a flexible long linker, named gating loop, in gRAMP from *Candidatus* "Scalindua brodae" (*Sb*-gRAMP) (**B**) are stabilized through interaction with the respective TPR-CHAT (blue). This increase in conformational rigidity reduces target RNA accessibility and cleavage in Craspase compared to gRAMP.

gRAMP, other single-subunit CRISPR-Cas effectors, such as the type II Cas9, are easily purified in the apo form [285]. How can this discrepancy of crRNA dependence between single-subunit effectors be evolutionarily explained? One speculative hypothesis is that the dependence of gRAMP on crRNA reflects an evolutionary stage where proteins had

not yet acquired the structural features necessary for independent stability. The world preceding protein-dominant cells was likely inhabited by catalysts composed entirely of RNA, often referred to as ribozymes in the 'RNA world' [286]. Ribonucleoproteins may present intermediates on the journey of life from the RNA world to stand-alone proteins, through initial decoration of the RNAs with peptides and subsequent evolutionary complexification into stand-alone proteins. The protein component in early ribonucleoproteins may still mostly be an adaptive trait for the RNA, thus possessing little autonomous stability. In late ribonucleoproteins, the protein component may have evolved to alleviate its RNA dependence and gained complex functionality on its own. Type III effectors are believed to be the earliest of the CRISPR-Cas ribonucleoproteins [43, 287] and perhaps evolved at a time where the protein component was still strictly dependent on the RNA component, explaining the instability of apo gRAMP. Moreover, it was found that RNA targets in type III effectors self-cleave, akin to the mechanism of action in ribozymes [277]. This makes the effectors effectively protein-assisted ribozymes, further supporting a dominant function of RNA in these ribonucleoproteins. The discovery of gRAMP presents a unique possibility to evaluate whether the protein component primarily serves as an adaptive trait for the RNA, as it is the only member of the ancient type III effector family that is single-subunit.

7.5. THE EMERGING THEME OF CRISPR PROTEASES

With the discovery of Craspase, CRISPR-Cas effectors have moved beyond the era of nucleic acids targeting into the domain of protein targeting to instigate anti-viral cascades. Anti-viral defense through deployment of proteases has been described before, such as the protease-mediated cleavage of antitoxin molecules in toxin-antitoxin systems to initiate abortive infection [288]. However, the integration of proteases with CRISPR guidance represents a remarkable advancement: the ability to accurately trigger protease through recognition of a specific nucleic acid specie. Next to Craspase, the CRISPR protease family contains two other experimentally described members (**Table 1**): CalpL (for 'CRISPR-associated Lon') and SAVED-CHAT.

A main difference with Craspase is that CalpL and SAVED-CHAT are not physically associated with the type III effector but instead are triggered by cOA second messenger signaling [107, 289]. CalpL forms a tripartite complex with CalpT and CalpS and, upon binding of cA_4 molecules, cleaves CalpT to liberate CalpS. CalpS is a predicted sigma factor and was found to physically interact with the RNA polymerase, leading to the hypothesis that its transcriptional functioning is unleashed upon release from the complex. SAVED-CHAT, upon binding of cA_3 molecules to the SAVED domain, forms long filaments that lead to protease activation of the CHAT domain. Activated SAVED-CHAT in turn cleaves and activates Prokaryotic Caspase (PCaspase), a protease with specificity towards a range of proteins. One of the target proteins of PCaspase is a putative sigma factor, suggestive of involvement in transcriptional regulation. The prospective transcriptional pathways in both CalpL and SAVED-CHAT provide a convergent evolutionary scenario to the induced transcriptional response in the Craspase system, thus revealing a common theme among CRISPR-associated proteases: the ability to couple nucleic acid detection with transcriptional regulation. Future exploration will unveil the extent of this theme and potential involvement of CRISPR proteases in other molecular pathways.

CRISPR-associated protease	Craspase	CalpL	SAVED-CHAT
CRISPR-Cas subtype	III-E	III-B	III-B
CRISPR-Cas effector	gRAMP	Cmr	Cmr
Protease	TPR-CHAT	SAVED-Lon	SAVED-CHAT
Nucleic acid trigger for activation	RNA	RNA	RNA
Communication between CRISPR-Cas effector and protease	Physical association	Second messengers (cA_4)	Second messengers (cA_3)
Oligomerization of the protease before activation	No	Yes	Yes
Protease catalytic dyad	Cys-His dyad	Ser-Lys dyad	Cys-His dyad
Target protein	Csx30	CalpT	PCaspase (which subsequently cleaves a range of proteins)
Anti-viral effect	Abortive infection	Unknown	Abortive infection
Effect of protease activity on associated transcription factor	Liberation of RpoE/CASP-σ (predicted sigma factor)	Liberation of CalpS (predicted sigma factor)	Cleavage of PCc-σ (predicted sigma factor)

Table 1. Key characteristics of the CRISPR-associated protease family members.

7.6. CONCLUSIONS AND OUTLOOK

Craspase research came a long way from its bioinformatic birth to insights into its physiological role and biotechnological potential. Ongoing efforts will undoubtedly address open questions regarding Craspase biology, biochemistry, and biotechnology in the near future. For example, what are the downstream effect of the genes transcribed by liberated RpoE? What is the role of Csx31? How are spacers acquired in type III-E systems? What are the structural and sequential code details of the interaction between Craspase and Csx30-Csx31-RpoE? How are target RNA binding and cleavage temporally related to protease activity in Craspase? What are the limits and off-targets of RNA and protein knockdown in cells? What can nucleic acid detection using Craspase add to the existing CRISPR-Cas diagnostics toolbox? Are there CRISPR-controlled proteases that are activated by DNA, or can they be created by directed evolution? Do anti-CRISPRs exist for Craspase? Do gRAMP or Craspase variants exist with characteristics different from the currently described ones? Similar questions can be asked for the other CRISPR-controlled proteases, ensuring interesting proteolytic times ahead.

EPILOGUE: TOWARDS COSMIC REVIVAL

"To see a World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And Eternity in an hour"

William Blake

We know more of the natural world than ever before, but at the same time are rapidly destroying it, ourselves included. The main cause of this contradictory behaviour could be traced to our mechanistic worldview, which sees the cosmos as a lifeless machine. Here, I will argue for the fecundity of a dialogue between contemporary molecular biology and process philosophy to develop alternative perspectives, aimed at conceptually bringing the cosmos back to life. I believe that a resulting transformative shift in worldview has the potential to incentivize a direly needed improvement in care for the Earth community.

INTRODUCTION

It is a strange experience to be a laboratory biologist studying the details of life in the midst of runaway climate change [290] and an unfolding human-induced mass extinction [291]. Academic writing desires a distillation of the details from the broader context, resulting in in-depth reports on isolated parts of nature often at the expense of the whole. Consequently, conspicuously absent from this predominantly technical dissertation is an exploration into what its findings could mean for the bigger picture. Yet I feel that in the current seemingly apocalyptic times, we can no longer afford the luxury of providing only dispassionate details when studying the cosmos. Especially given the potential of science-inspired cosmologies to heal or destroy worlds.

Science has revealed that, from the primordial beginning, everything has been creating, growing, progressing, proceeding, and all there is today has been on its way for billions of years to arrive in the present moment. Positioned at the growing tip of this evolutionary journey, humans are in a dubious situation. As far as we know, for the first time in its existence, the cosmos is able to conceive of itself in a deep and profound way through the various faculties that evolution has culminated in us. We possess more knowledge of the world than ever before, but at the same time are rapidly destroying a great deal of the rich manifestations of cosmic development that have ever come into existence.

Ultimately, the commitment of a natural scientist lies in the thorough examination of nature, guided by the directions provided by data. Confronted with accumulation of grim data on the global environmental scale, I find myself compelled to reflect on how my examination of nature could contribute meaningfully, however minutely, to addressing the pressing challenges of the Earth's ongoing crises. Here, I will synthesize personal scientific insight, metaphysical musing, and spiritual intuition on the cosmic whole, aimed at unearthing a root problem of the life-denying tendencies of humans in the 21st century and sowing the seeds for a solution.

THE COSMIC CORPSE

The scientific ideal posits that our theories are informed and constructed by observations. However, it is crucial to recognize that our observations are equally shaped by our theories [292]. Our models of the cosmos determine what we can see of that cosmos, which, in turn, dictates how we interact with it:

> "Beliefs, in short, are rules for action; and the whole function of thinking is but one step in the production of active habits." [293]

Extensively described suspect for the death sentence of the natural world are the beliefs produced by reductionist materialistic thought [294], the dominant framework through which modern humans perceive reality. These dictate that the cosmos is made up of nothing but physical matter that behaves in a purposeless fashion. The embodiment of such thought makes the scientist imagine the cosmos as

> "irreducible brute matter spread throughout space in a flux of configurations, in itself senseless, valueless, purposeless, following a fixed routine imposed by external relations." [295].

In other words, the scientist views its object of study to be inert, passive, a machine, something that is dead. Most paradoxically, this includes the life scientist, testified by his insistence on reducing organisms to dead principles, his 'obsession with mechanism of every sort—"genetic mechanisms", "signalling mechanisms", "regulatory mechanisms", and even "molecular mechanisms of plasticity"' when studying biology [296]. These life-denying beliefs translate into actions revolving around instrumental control, utility, and exploitation, something one would do with a lifeless resource. As the insights propagated by modern science is informing the worldview of the public, it is no wonder that we are killing the planet when we don't realize it is alive in the first place. Currently, detail after detail is added to the vastly growing body of scientific information, but the integration into knowledge is done through the lens of destructive mechanistic frameworks. This cosmology therefore desperately requires updating, not only in the light of the global challenges we face, but also because it is highly contradictory to the insights of modern science [292, 294, 297]. So, besides practical and technological approaches to mitigate damage and loss being essential, acceleration of deep systemic changes in the ways we perceive the cosmos is equally necessary.

COSMIC REVIVAL

An alternative narrative is being developed by the process philosophers. They do not use mechanistic, but organismic and ecological concepts to probe into the nature of reality. Humans have long intuited the cosmos as being animated. Animism and panpsychism, the understanding that all things are alive and possess a form of mentality [298], are in the foundations of ancient and widespread belief structures, and already early Western thinkers like Plato adhered to the idea that

> "this world is indeed a living being endowed with a soul and intelligence ... a single visible living entity containing all other living entities, which by their nature are all related." [299]

When we think of the cosmos as a giant organism with the accompanying organistic principles such as creativity, interrelationality, and evolutionary progression, the world becomes less an instrument that we use and trash, but more something to love, to take care of —something that is more us than we can currently imagine. The cosmos is seen as a living and evolving entity, exhibiting qualities similar to those found in biological organisms such that, informed by process metaphysics,

> "science is taking on a new aspect which is neither purely physical, nor purely biological. It is becoming the study of organisms. Biology is the study of the larger organisms; whereas physics is the study of the smaller organisms." [295]

In my view, MGEs provide tangible evidence for the intuitions of process philosophers. MGEs are processes that are positioned at the edge of dead and alive as we imagine those terms within the current mechanistic paradigm: the most simple MGEs are composed of mere thousands of atoms with no proteins encoded [300], whereas the most complex MGEs are in the range of bacteria in terms of physical dimensions and genomes size, encoding hundreds of proteins, hinting at complex lifestyles [301]. The cell is increasingly considered a playground for MGE interactions [302], allowing

> "consideration of all [mobile] genetic elements, including chromosomes, as parts of a single, interwoven, dynamic community that exploits cell vehicles for propagation and preservation." [303]

MGEs dominate the biosphere in terms of numbers, diversity, and genetic information content [303]. MGEs spread, replicate, disrupt, generate, mix, exchange, recombine, and as such, physically embody the organistic principles (see **Chapter 1, Fig. 1**) [4]. The consequences of these creative impulses become evident across various scales. On the molecular level, integration of insertion sequences [304], CRISPR spacers [273], and prophages [305] morph the nucleic acid molecule to obtain novel genomic features; on the cellular level, acquisition of MGEs results in the sudden availability of novel traits, such as antibiotic resistance factors on plasmids [306]; and on the multicellular level, infiltrating viral proteins are repurposed by the host for crucial functions in complex processes, such as placental development [307]. These are just a couple examples among the endless creative impacts of MGEs that have driven simple early replicators to complex multicellular organisms over the course of 3.5 billion years [11].

Craspase stands among the over one hundred novel immunity pathways that have recently been described in bacteria and archaea, collectively encompassing a wide range of strategies aimed at impeding the propagation of invading MGEs [204, 263, 264]. Contrary to the traditional notion of the immune system as cellular strategy, bioinformatic analyses made apparent that most —or, when acknowledging that the bacterial chromosome itself is a MGE [308], all— immune systems are encoded on MGEs and used for interactions with other MGEs [14, 302, 309]. Intertwined with the evolution of immune systems, MGEs are evolving counter-strategies geared at overcoming the obstacles of immunity [310]. This co-evolutionary entanglement between defense and anti-defense is often portrayed as an arms race marked by conflict, opposition and antagonism, as evidenced by the rich tapestry of war-related metaphors found in MGE literature. This perspective is understandable when viewing immune interactions from the level of the individual MGEs, where flourishing of the one means the demise of the other. However, when transcending the level of the strict individual to consider the broader species perspective, we see that these seemingly antagonistic interactions actually foster evolutionary innovation in the community. This is clearly observed in certain CRISPR-Cas systems: CRISPR-Cas evolved to combat MGE infections by cleaving the invading nucleic acids, which led MGEs to evolve anti-CRISPR proteins to overcome CRISPR-Cas interference [311], which in turn provided the condition for to the evolution of anti-anti-CRISPR proteins [312]. It would, therefore, not be surprising if the evolutionary narrative were to continue with the discovery of anti-anti-anti-CRISPR proteins.

Another intriguing experimental illustration of continuous innovation induced by MGE-MGE interactions are found in the evolution of the phage anti-restriction-induced system (PARIS). As a response to the activity of defense enzymes that cleave its DNA, the phage T7 evolved the counter-defense Ocr (for "overcoming classical restriction") protein. Ocr represses DNA-binding proteins, thereby inhibiting the DNA cleavage ability of cellular restriction enzymes [313]. As an evolutionary reaction, Ocr itself is the trigger for PARIS, which induces an abortive infection phenotype by cleaving the cellular tRNAs [314, 315]. This in turn facilitated a new response to restore infection, namely the evolution of non-cleavable tRNAs [315].

So while immune interactions appear conflictual at the individual level, at the deeper level they exhibit cooperative dynamics in a continuous cycle, geared at facilitating evolutionary trajectories towards ever increasing novelty. Analogously, the two pedals of a bicycle are not competing despite moving in opposite directions, but instead cooperate to facilitate movement of the whole in a hybrid direction forward. In light of this broader context, immune and anti-immune systems can perhaps more aptly be characterized as 'facilitator systems'.

The co-evolutionary dynamics of facilitator systems and horizontal distribution of creative potential in MGE ecologies exemplify a larger cosmic theme, one that now finds support in a plethora of research: we inhabit a cosmos that is an open-ended process into complexity driven by underlying creative principles. The scientific observations on cosmic evolution can be summarized in a simple phrase of mystical proportions:

> "if you let hydrogen gas alone for thirteen billion years it will become giraffes, rose bushes and humans." [316]

This undeniably asserts that there is directionality, unity, and continuity to cosmogenesis that has prevented us from still being simple protonic entities. Similarly, when Darwin reflected on his observations of the natural world, he mused that

> "from so simple a beginning, endless forms most beautiful and most wonderful have been, and are being, evolved." [317]

This is as true for the biological organism as it is for the cosmic organism. We now know that the atoms in our intricate DNA were once churning in the formless mass of stellar cores; that the processes in our multicellular bodies are orchestrated by proteins whose ancestors were already fuelling the march of primordial life in the oceans billions of years ago; that the complex never-ending stream of theory we can mentally conceive are the product of the concepts one can process; that a constellation of protons gained the capacity to write a dissertation about constellations of protons. These findings resonate with the process philosophical axiom that the creative advance into novelty is "the ultimate metaphysical ground" for evolution in the cosmos [318]. Endless form and function from limited form and function — much like a living mystery.

CONCLUSION

Thorough conceptualization of the cosmos in inanimate terms make it unsurprising that we treat it as such. But what if the cosmos as a whole is imagined to be alive? Displaying clear biological principles while only consisting of mere thousands of atoms, MGE are potent agents of innovation at the edge of life and death as we imagine those terms within the current mechanistic paradigm. They shape the advance of the cosmos through the constant horizontal sharing of evolutionary potential and the action of facilitator systems. As such, MGEs are fundamental demonstrations of process, providing physical embodiment to concepts well-couched in process philosophical cosmology, such as interrelationality, cooperation, co-evolutionary dynamics, flux, and becoming. I therefore argue that MGEs serve as a fecund starting point for bringing the euthanized cosmos back to life. A dialogue between the microbiologist and the philosopher of organism is, in the face of the world's urgent problems, eventually all geared towards a solution: conceptual revival of the cosmos.

REFERENCES

[1] S. P. B. van Beljouw, J. Sanders, A. Rodríguez-Molina, and S. J. J. Brouns, "RNA-targeting CRISPR–Cas systems," *Nature Reviews Microbiology*, vol. 21, no. 1, pp. 21–34, 2023.

[2] S. P. B. van Beljouw and S. J. J. Brouns, "CRISPR-controlled proteases," *Biochemical Society Transactions*, p. BST20230962, 2024.

[3] E. V. Koonin, Y. I. Wolf, and M. I. Katsnelson, "Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states," *Biology Direct*, vol. 12, pp. 1–12, 2017.

[4] L. S. Frost, R. Leplae, A. O. Summers, and A. Toussaint, "Mobile genetic elements: The agents of open source evolution," *Nature Reviews Microbiology*, vol. 3, pp. 722–732, 2005.

[5] F. L. Nobrega, M. Vlot, P. A. de Jonge, L. L. Dreesens, H. J. E. Beaumont, R. Lavigne, B. E. Dutilh, and S. J. J. Brouns, "Targeting mechanisms of tailed bacteriophages," *Nature Reviews Microbiology*, vol. 16, pp. 760–773, 2018.

[6] J. Rodríguez-Beltrán, J. DelaFuente, R. Leon-Sampedro, R. C. MacLean, and A. San Millan, "Beyond horizontal gene transfer: the role of plasmids in bacterial evolution," *Nature Reviews Microbiology*, vol. 19, no. 6, pp. 347–359, 2021.

[7] C. M. Johnson and A. D. Grossman, "Integrative and conjugative elements (ICEs): What they do and how they work," *Annual Review of Genetics*, vol. 49, pp. 577–601, 2015.

[8] N. S. Catlin and E. B. Josephs, "The important contribution of transposable elements to phenotypic variation and evolution," *Current Opinion in Plant Biology*, vol. 65, p. 102140, 2022.

[9] R. Martínez-Rubio, N. Quiles-Puchalt, M. Martí, S. Humphrey, G. Ram, D. Smyth, J. Chen, R. P. Novick, and J. R. Penadés, "Phage-inducible islands in the gram-positive cocci," *ISME journal*, vol. 11, pp. 1029–1042, 2017.

[10] L. V. Valen, "A new evolutionary law," *Evolutionary Theory*, vol. 1, pp. 1–30, 1973.

[11] E. V. Koonin, "Viruses and mobile elements as drivers of evolutionary transitions," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 371, no. 1701, p. 20150442, 2016.

[12] A. Bernheim and R. Sorek, "The pan-immune system of bacteria: antiviral defence as a community resource," *Nature Reviews Microbiology*, vol. 18, pp. 113–119, 2020.

[13] F. Tesson, A. Hervé, E. Mordret, M. Touchon, C. d'Humières, J. Cury, and A. Bernheim, "Systematic and quantitative view of the antiviral arsenal of prokaryotes," *Nature communications*, vol. 13, no. 1, p. 2561, 2022.

[14] E. P. C. Rocha and D. Bikard, "Microbial defenses against mobile genetic elements and viruses: Who defends whom from what?" *PLOS Biology*, vol. 20, p. e3001514, 2022.

[15] H. G. Hampton, B. N. Watson, and P. C. Fineran, "The arms race between bacteria and their phage foes," pp. 327–336, 2020.

[16] R. Barrangou, C. Fremaux, H. Deveau, M. Richardss, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath, "CRISPR provides acquired resistance against viruses in prokaryotes," *Science*, vol. 315, pp. 1709–1712, 2007.

[17] H. Deveau, R. Barrangou, J. E. Garneau, J. Labonté, C. Fremaux, P. Boyaval, D. A. Romero, P. Horvath, and S. Moineau, "Phage response to CRISPR-encoded resistance in streptococcus thermophilus," *Journal of Bacteriology*, vol. 190, pp. 1390–1400, 2008.

[18] R. Barrangou and P. Horvath, "A decade of discovery: CRISPR functions and applications," *Nature Microbiology*, vol. 2, 2017.

[19] K. A. Datsenko, K. Pougach, A. Tikhonov, B. L. Wanner, K. Severinov, and E. Semenova, "Molecular memory of prior infections activates the CRISPR-Cas adaptive bacterial immunity system," *Nature Communications*, vol. 3, pp. 1–7, 7 2012.

[20] D. C. Swarts, C. Mosterd, M. W. J. van Passel, and S. J. J. Brouns, "CRISPR interference directs strand specific spacer acquisition," *PLoS ONE*, vol. 7, 2012.

[21] P. M. Nussenzweig, J. McGinn, and L. A. Marraffini, "Cas9 cleavage of viral genomes primes the acquisition of new immunological memories," *Cell host & microbe*, vol. 26, pp. 515–526.e6, 10 2019.

[22] J. Carte, R. Wang, H. Li, R. M. Terns, and M. P. Terns, "Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes," *Genes and Development*, vol. 22, pp. 3489–3496, 12 2008.

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

[24] S. J. J. Brouns, M. M. Jore, M. Lundgren, E. R. Westra, R. J. Slijkhuis, A. P. Snijders, M. J. Dickman, K. S. Makarova, E. V. Koonin, and J. van der Oost, "Small CRISPR RNAs guide antiviral defense in prokaryotes," *Science*, vol. 321, pp. 960–964, 8 2008.

[25] E. V. Koonin and K. S. Makarova, "Discovery of oligonucleotide signaling mediated by CRISPR-associated polymerases solves two puzzles but leaves an enigma," *ACS Chemical Biology*, vol. 13, pp. 309–312, 2018.

[26] G. A. Coleman, A. A. Davín, T. A. Mahendrarajah, L. L. Szánthó, A. Spang, P. Hugenholtz, G. J. Szöllsi, and T. A. Williams, "A rooted phylogeny resolves early bacterial evolution," *Science*, vol. 372, 2021.

[27] G. Tamulaitis, M. Kazlauskiene, E. Manakova, Česlovas Venclovas, A. O. Nwokeoji, M. J. Dickman, P. Horvath, and V. Siksnys, "Programmable RNA shredding by the type III-A CRISPR-Cas system of Streptococcus thermophilus," *Molecular Cell*, vol. 56, pp. 506–517, 2014.

[28] A. Özcan, R. Krajeski, E. Ioannidi, B. Lee, A. Gardner, K. S. Makarova, E. V. Koonin, O. O. Abudayyeh, and J. S. Gootenberg, "Programmable RNA targeting with the single-protein CRISPR effector Cas7-11," *Nature*, vol. 597, pp. 720–725, 9 2021.

[29] Y. I. Wolf, S. Silas, Y. Wang, S. Wu, M. Bocek, D. Kazlauskas, M. Krupovic, A. Fire, V. V. Dolja, and E. V. Koonin, "Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome," *Nature Microbiology*, vol. 5, pp. 1262–1270, 2020.

[30] U. Neri, Y. I. Wolf, S. Roux, A. P. Camargo, B. D. Lee, D. Kazlauskas, I. M. Chen, N. Ivanova, L. Z. Allen, D. Paez-Espino, D. A. Bryant, D. Bhaya, M. Krupovic, R. V. D. Consortium, V. V. Dolja, N. C. Kyrpides, E. V. Koonin, and U. Gophna, "Expansion of the global RNA virome reveals diverse clades of bacteriophages," *Cell*, vol. 185, pp. 4023–4037, 2022.

[31] M. Kazlauskiene, G. Tamulaitis, G. Kostiuk, Česlovas Venclovas, and V. Siksnys, "Spatiotemporal control of type III-A CRISPR-Cas immunity: Coupling DNA degradation with the target RNA recognition," *Molecular Cell*, vol. 62, pp. 295–306, 4 2016.

[32] W. Jiang, P. Samai, and L. A. Marraffini, "Degradation of phage transcripts by CRISPR-associated RNases enables type III CRISPR-Cas immunity," *Cell*, vol. 164, pp. 710–721, 2016.

[33] R. Molina, N. Sofos, and G. Montoya, "Current opinion in structural biology structural basis of CRISPR-Cas type III prokaryotic defence systems," *Current Opinion in Structural Biology*, vol. 65, pp. 119–129, 2020.

[34] G. Tamulaitis, Česlovas Venclovas, and V. Siksnys, "Type III CRISPR-Cas immunity: Major differences brushed aside," *Trends in Microbiology*, vol. 25, pp. 49–61, 2017.

[35] C. Rouillon, M. Zhou, J. Zhang, A. Politis, V. Beilsten-Edmands, G. Cannone, S. Graham, C. V. Robinson, L. Spagnolo, and M. F. White, "Structure of the CRISPR interference complex Csm reveals key similarities with cascade," *Molecular Cell*, vol. 52, pp. 124–134, 10 2013.

[36] J. Zhang, C. Rouillon, M. Kerou, J. Reeks, K. Brugger, S. Graham, J. Reimann, G. Cannone, H. Liu, S. V. Albers, J. H. Naismith, L. Spagnolo, and M. F. White, "Structure and mechanism of the Cmr complex for CRISPR-mediated antiviral immunity," *Molecular Cell*, vol. 45, pp. 303–313, 2 2012.

[37] C. R. Hale, P. Zhao, S. Olson, M. O. Duff, B. R. Graveley, L. Wells, R. M. Terns, and M. P. Terns, "RNA-guided RNA cleavage by a CRISPR RNA-cas protein complex," *Cell*, vol. 139, pp. 945–956, 2009.

[38] C. Rouillon, J. S. Athukoralage, S. Graham, S. Grüschow, and M. F. White, "Control of cyclic oligoadenylate synthesis in a type III CRISPR system," *eLife*, vol. 7, pp. 1–22, 2018.

[39] M. A. Estrella, F. T. Kuo, and S. Bailey, "RNA-activated DNA cleavage by the type III-B CRISPR–Cas effector complex," *Genes and Development*, vol. 30, pp. 460–470, 2 2016.

[40] R. D. Sokolowski, S. Graham, and M. F. White, "Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system," *Nucleic Acids Research*, vol. 42, pp. 6532–6541, 6 2014.

[41] F. C. Walker, L. Chou-Zheng, J. A. Dunkle, and A. Hatoum-Aslan, "Molecular determinants for CRISPR RNA maturation in the Cas10-Csm complex and roles for non-cas nucleases," *Nucleic acids research*, vol. 45, pp. 2112–2123, 2017.

[42] L. Chou-Zheng and A. Hatoum-Aslan, "A type III-A CRISPR-Cas system employs degradosome nucleases to ensure robust immunity," *eLife*, vol. 8, pp. 1–25, 2019.

[43] K. S. Makarova, Y. I. Wolf, J. Iranzo, S. A. Shmakov, O. S. Alkhnbashi, S. J. J. Brouns, E. Charpentier, D. Cheng, D. H. Haft, P. Horvath, S. Moineau, F. J. Mojica, D. Scott, S. A. Shah, V. Siksnys, M. P. Terns, Česlovas Venclovas, M. F. White, A. F. Yakunin, W. Yan, F. Zhang, R. A. Garrett, R. Backofen, J. van der Oost, R. Barrangou, and E. V. Koonin, "Evolutionary classification of CRISPR–Cas systems: a burst of Class 2 and derived variants," *Nature Reviews Microbiology*, vol. 18, pp. 67–83, 2020.

[44] K. S. Makarova, Y. I. Wolf, O. S. Alkhnbashi, F. Costa, S. A. Shah, S. J. Saunders, R. Barrangou, S. J. J. Brouns, E. Charpentier, D. H. Haft, P. Horvath, S. Moineau, F. J. Mojica, R. M. Terns, M. P. Terns, M. F. White, A. F. Yakunin, R. A. Garrett, J. van der Oost, R. Backofen, and E. V. Koonin, "An updated evolutionary

classification of CRISPR-Cas systems," *Nature Reviews Microbiology*, vol. 13, pp. 722–736, 2015.
[45] J. Van De Vossenberg, J. E. Rattray, W. Geerts, B. Kartal, L. Van Niftrik, E. G. Van Donselaar, J. S. Sinninghe Damsté, M. Strous, and M. S. Jetten, "Enrichment and characterization of marine anammox bacteria associated with global nitrogen gas production," *Environmental microbiology*, vol. 10, no. 11, pp. 3120–3129, 2008.
[46] F. Hille, H. Richter, S. P. Wong, M. Bratovič, S. Ressel, and E. Charpentier, "The biology of CRISPR-Cas: Backward and forward," *Cell*, vol. 172, pp. 1239–1259, 2018.
[47] N. Jia and D. J. Patel, "Structure-based functional mechanisms and biotechnology applications of anti-CRISPR proteins," *Nature Reviews Molecular Cell Biology*, vol. 22, pp. 563–579, 2021.
[48] M. Wilkinson, L. Troman, W. A. W. N. Ismah, Y. Chaban, M. B. Avison, M. S. Dillingham, and D. B. Wigley, "Structural basis for the inhibition of RecBCD by gam and its synergistic antibacterial effect with quinolones," *eLife*, vol. 5, pp. 1–12, 2016.
[49] M. Shah, V. L. Taylor, D. Bona, Y. Tsao, S. Y. Stanley, S. M. Pimentel-Elardo, M. McCallum, J. Bondy-Denomy, P. L. Howell, J. R. Nodwell, A. R. Davidson, T. F. Moraes, and K. L. Maxwell, "A phage-encoded anti-activator inhibits quorum sensing in Pseudomonas aeruginosa," *Molecular Cell*, vol. 81, pp. 571–583.e6, 2021.
[50] P. K. Bandyopadhyay, F. W. Studier, D. L. Hamilton, and R. Yuan, "Inhibition of the type I restriction-modification enzymes EcoB and EcoK by the gene 0.3 protein of bacteriophage T7," *journal of Molecular Biology*, vol. 182, pp. 567–578, 1985.
[51] H. Deveau, J. E. Garneau, and S. Moineau, "CRISPR-Cas system and its role in phage-bacteria interactions," *Annual Review of Microbiology*, vol. 64, pp. 475–493, 10 2010.
[52] G. Hutinet, Y. jiun Lee, V. de Crécy-lagard, and P. R. Weigele, "Hypermodified DNA in viruses of E. coli and Salmonella," *EcoSalPlus*, vol. 9, pp. –eESP–0028, 2021.
[53] A. L. Bryson, Y. Hwang, S. Sherrill-Mix, G. D. Wu, J. D. Lewis, L. Black, T. A. Clark, and F. D. Bushman, "Covalent modification of bacteriophage T4 DNA inhibits CRISPR-Cas9," *mBio*, vol. 6, 6 2015.
[54] M. Vlot, J. Houkes, S. J. Lochs, D. C. Swarts, P. Zheng, T. Kunne, P. Mohanraju, C. Anders, M. Jinek, J. van der Oost, M. J. Dickman, and S. J. J. Brouns, "Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR–Cas effector complexes," *Nucleic Acids Research*, vol. 46, pp. 873–885, 1 2018.
[55] S. D. Mendoza, E. S. Nieweglowska, S. Govindarajan, L. M. Leon, J. D. Berry, A. Tiwari, V. Chaikeeratisak, J. Pogliano, D. A. Agard, and J. Bondy-Denomy, "A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases," *Nature*, vol. 577, pp. 244–248, 2020.
[56] L. M. Malone, S. L. Warring, S. A. Jackson, C. Warnecke, P. P. Gardner, L. F. Gumy, and P. C. Fineran, "A jumbo phage that forms a nucleus-like structure evades CRISPR–Cas DNA targeting but is vulnerable to type III RNA-based immunity," *Nature Microbiology*, vol. 5, pp. 48–55, 1 2020.
[57] J. N. A. Vink, J. H. L. Baijens, and S. J. J. Brouns, "Pam-repeat associations and spacer selection preferences in single and co-occurring CRISPR-Cas systems," *Genome Biology*, vol. 22, pp. 1–25, 2021.
[58] S. Majumdar, P. Zhao, N. T. Pfister, M. Compton, S. Olson, C. V. Glover, L. Wells, B. R. Graveley, R. M. Terns, and M. P. Terns, "Three CRISPR-Cas immune effector complexes coexist in Pyrococcus furiosus," *RNA (New York, N.Y.)*, vol. 21, pp. 1147–1158, 2015.
[59] S. Silas, P. Lucas-Elio, S. A. Jackson, A. Aroca-Crevillén, L. L. Hansen, P. C. Fineran, A. Z. Fire, and A. Sánchez-Amat, "Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems," *eLife*, vol. 6, 8 2017.
[60] V. Hoikkala, J. Ravantti, C. Díez-Villaseñor, M. Tiirola, R. A. Conrad, M. J. McBride, S. Moineau, and L. R. Sundberg, "Cooperation between different CRISPR-Cas types enables adaptation in an RNA-targeting system," *mBio*, vol. 12, 2021.
[61] V. Chaikeeratisak, K. Nguyen, K. Khanna, A. F. Brilot, M. L. Erb, J. K. Coker, A. Vavilina, G. L. Newton, R. Buschauer, K. Pogliano, E. Villa, D. A. Agard, and J. Pogliano, "Assembly of a nucleus-like structure during viral replication in bacteria," *Science*, vol. 355, pp. 194–197, 1 2017.
[62] E. Harrison and M. A. Brockhurst, "Ecological and evolutionary benefits of temperate phage: What does or doesn't kill you makes you stronger," *BioEssays*, vol. 39, 2017.
[63] G. W. Goldberg, W. Jiang, D. Bikard, and L. A. Marraffini, "Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting," *Nature*, vol. 514, pp. 633–637, 2014.
[64] G. W. Goldberg, E. A. McMillan, A. Varble, J. W. Modell, P. Samai, W. Jiang, and L. A. Marraffini, "Incomplete prophage tolerance by type III-A CRISPR-Cas systems reduces the fitness of lysogenic hosts," *Nature Communications*, vol. 9, 2018.
[65] S. Shmakov, A. Smargon, D. Scott, D. Cox, N. Pyzocha, W. Yan, O. O. Abudayyeh, J. S. Gootenberg, K. S.

Makarova, Y. I. Wolf, K. Severinov, F. Zhang, and E. V. Koonin, "Diversity and evolution of Class 2 CRISPR-Cas systems," *Nature Reviews Microbiology*, vol. 15, pp. 169–182, 2017.

[66] A. East-Seletsky, M. R. O'Connell, D. Burstein, G. J. Knott, and J. A. Doudna, "RNA targeting by functionally orthogonal type VI-A CRISPR-Cas enzymes," *Molecular Cell*, vol. 66, pp. 373–383.e3, 5 2017.

[67] J. S. Gootenberg, O. O. Abudayyeh, M. J. Kellner, J. Joung, J. J. Collins, and F. Zhang, "Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a and Csm6," *Science*, vol. 360, pp. 439–444, 2018.

[68] J. A. Steens, Y. Zhu, D. W. Taylor, J. P. Bravo, S. H. Prinsen, C. D. Schoen, B. J. Keijser, M. Ossendrijver, L. M. Hofstra, S. J. J. Brouns, A. Shinkai, J. van der Oost, and R. H. Staals, "Scope enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation," *Nature Communications*, vol. 12, pp. 1–12, 2021.

[69] N. C. Pyenson, K. Gayvert, A. Varble, O. Elemento, and L. A. Marraffini, "Broad targeting specificity during bacterial type III CRISPR-Cas immunity constrains viral escape," *Cell host & microbe*, vol. 22, pp. 343–353.e3, 9 2017.

[70] J. R. Elmore, N. F. Sheppard, N. Ramia, T. Deighan, H. Li, R. M. Terns, and M. P. Terns, "Bipartite recognition of target RNAs activates DNA cleavage by the type III-B CRISPR–Cas system," *Genes and Development*, vol. 30, pp. 447–459, 2 2016.

[71] L. A. Marraffini and E. J. Sontheimer, "Self versus non-self discrimination during CRISPR RNA-directed immunity," *Nature*, vol. 463, pp. 568–571, 1 2010.

[72] L. Wang, C. Y. Mo, M. R. Wasserman, J. T. Rostøl, L. A. Marraffini, and S. Liu, "Dynamics of Cas10 govern discrimination between self and non-self in type III CRISPR-Cas immunity," *Molecular Cell*, vol. 73, pp. 278–290.e4, 2019.

[73] R. H. Staals, Y. Zhu, D. W. Taylor, J. E. Kornfeld, K. Sharma, A. Barendregt, J. J. Koehorst, M. Vlot, N. Neupane, K. Varossieau, K. Sakamoto, T. Suzuki, N. Dohmae, S. Yokoyama, P. J. Schaap, H. uraub, A. J. Heck, E. Nogales, J. A. Doudna, A. Shinkai, and J. vanderOost, "RNA targeting by the type III-A CRISPR-Cas Csm complex of Thermus thermophilus," *Molecular Cell*, vol. 56, pp. 518–530, 11 2014.

[74] M. Guo, K. Zhang, Y. Zhu, G. D. Pintilie, X. Guan, S. Li, M. F. Schmid, Z. Ma, W. Chiu, and Z. Huang, "Coupling of ssRNA cleavage with DNase activity in type III-A CRISPR-Csm revealed by cryo-EM and biochemistry," *Cell Research*, vol. 29, pp. 305–312, 2019.

[75] R. H. Staals, Y. Agari, S. Maki-Yonekura, Y. Zhu, D. W. Taylor, E. vanDuijn, A. Barendregt, M. Vlot, J. J. Koehorst, K. Sakamoto, A. Masuda, N. Dohmae, P. J. Schaap, J. A. Doudna, A. J. Heck, K. Yonekura, J. van der Oost, and A. Shinkai, "Structure and activity of the RNA-targeting type III-B CRISPR-Cas complex of Thermus thermophilus," *Molecular Cell*, vol. 52, pp. 135–145, 2013.

[76] C. R. Hale, A. Cocozaki, H. Li, R. M. Terns, and M. P. Terns, "Target RNA capture and cleavage by the Cmr type III-B CRISPR–Cas effector complex," *Genes and Development*, vol. 28, pp. 2432–2443, 2014.

[77] C. R. Hale, S. Majumdar, J. Elmore, N. Pfister, M. Compton, S. Olson, A. M. Resch, C. V. Glover, B. R. Graveley, R. M. Terns, and M. P. Terns, "Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs," *Molecular Cell*, vol. 45, pp. 292–302, 2012.

[78] P. Samai, N. Pyenson, W. Jiang, G. W. Goldberg, A. Hatoum-Aslan, and L. A. Marraffini, "Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity," *Cell*, vol. 161, pp. 1164–1174, 5 2015.

[79] K.-H. Park, Y. An, T.-Y. Jung, I.-Y. Baek, H. Noh, W.-C. Ahn, H. Hebert, J.-J. Song, J.-H. Kim, B.-H. Oh, and E.-J. Woo, "RNA activation-independent DNA targeting of the type III CRISPR-cas system by a Csm complex," *EMBO reports*, vol. 18, pp. 826–840, 2017.

[80] T. Y. Liu, J. J. Liu, A. J. Aditham, E. Nogales, and J. A. Doudna, "Target preference of type III-A CRISPR-Cas complexes at the transcription bubble," *Nature Communications*, vol. 10, 2019.

[81] C. Y. Mo, J. Mathai, J. T. Rostøl, A. Varble, D. V. Banh, and L. A. Marraffini, "Type III-A CRISPR immunity promotes mutagenesis of staphylococci," *Nature*, vol. 592, pp. 611–615, 2021.

[82] J. T. Rostøl and L. A. Marraffini, "Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR–cas immunity," *Nature Microbiology*, vol. 4, pp. 656–662, 2019.

[83] W. Han, S. Stella, Y. Zhang, T. Guo, K. Sulek, L. Peng-Lundgren, G. Montoya, and Q. She, "A type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding," *Nucleic Acids Research*, vol. 46, pp. 10 319–10 330, 2018.

[84] A. I. Cocozaki, N. F. Ramia, Y. Shao, C. R. Hale, R. M. Terns, M. P. Terns, and H. Li, "Structure of the Cmr2 subunit of the CRISPR-Cas RNA silencing complex," *Structure*, vol. 20, pp. 545–553, 2012.

[85] M. Kazlauskiene, G. Kostiuk, Česlovas Venclovas, G. Tamulaitis, and V. Siksnys, "A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems," *Science*, vol. 357, pp. 605–609, 2017.

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

[87] M. Nasef, M. C. Muffly, A. B. Beckman, S. J. Rowe, F. C. Walker, A. Hatoum-Aslan, and J. A. Dunkle, "Regulation of cyclic oligoadenylate synthesis by the Staphylococcus epidermidis Cas10-Csm complex," *RNA*, vol. 25, pp. 948–962, 8 2019.

[88] S. A. Shah, O. S. Alkhnbashi, J. Behler, W. Han, Q. She, W. R. Hess, R. A. Garrett, and R. Backofen, "Comprehensive search for accessory proteins encoded with archaeal and bacterial type III CRISPR-Cas gene cassettes reveals 39 new Cas gene families," *RNA Biology*, vol. 16, pp. 530–542, 2019.

[89] K. S. Makarova, V. Anantharaman, N. V. Grishin, E. V. Koonin, and L. Aravind, "CARF and WYLdomains: Ligand-binding regulators of prokaryotic defense systems," *Frontiers in Genetics*, vol. 5, p. 102, 2014.

[90] O. Niewoehner and M. Jinek, "Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6," *RNA*, vol. 22, pp. 318–329, 2016.

[91] W. Han, S. Pan, B. López-Méndez, G. Montoya, and Q. She, "Allosteric regulation of Csx1, a type IIIB-associated CARF domain ribonuclease by RNAs carrying a tetraadenylate tail," *Nucleic Acids Research*, vol. 45, pp. 10 740–10 750, 2017.

[92] Y. Zhao, J. Wang, Q. Sun, C. Dou, Y. Gu, C. Nie, X. Zhu, Y. Wei, and W. Cheng, "Structural insights into the CRISPR-Cas-associated ribonuclease activity of Staphylococcus epidermidis Csm3 and Csm6," *Science Bulletin*, vol. 63, pp. 691–699, 2018.

[93] C. Garcia-Doval, F. Schwede, C. Berk, J. T. Rostøl, O. Niewoehner, O. Tejero, J. Hall, L. A. Marraffini, and M. Jinek, "Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6," *Nature Communications*, vol. 11, p. 1596, 12 2020.

[94] S. A. McMahon, W. Zhu, S. Graham, R. Rambo, M. F. White, and T. M. Gloster, "Structure and mechanism of a type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate," *Nature Communications*, vol. 11, 2020.

[95] W. Zhu, S. McQuarrie, S. Grüschow, S. A. McMahon, S. Graham, T. M. Gloster, and M. F. White, "The CRISPR ancillary effector Can2 is a dual-specificity nuclease potentiating type III CRISPR defence," *Nucleic Acids Research*, vol. 49, pp. 2777–2789, 2021.

[96] J. T. Rostøl, W. Xie, V. Kuryavyi, P. Maguin, K. Kao, R. Froom, D. J. Patel, and L. A. Marraffini, "The card1 nuclease provides defence during type-III CRISPR immunity," *Nature*, vol. 590, 2021.

[97] J. S. Athukoralage, S. Graham, S. Grüschow, C. Rouillon, and M. F. White, "A type III CRISPR ancillary ribonuclease degrades its cyclic oligoadenylate activator," *journal of Molecular Biology*, vol. 431, pp. 2894–2899, 2019.

[98] D. M. Hinton, "Transcriptional control in the prereplicative phase of T4 development," *Virology journal*, vol. 7, pp. 1–16, 2010.

[99] A. Chevallereau, B. G. Blasdel, J. D. Smet, M. Monot, M. Zimmermann, M. Kogadeeva, U. Sauer, P. Jorth, M. Whiteley, L. Debarbieux, and R. Lavigne, "Next-generation "-omics" approaches reveal a massive alteration of host RNA metabolism during bacteriophage infection of Pseudomonas aeruginosa," *PLoS Genetics*, vol. 12, pp. 1–20, 2016.

[100] B. G. Blasdel, A. Chevallereau, M. Monot, R. Lavigne, and L. Debarbieux, "Comparative transcriptomics analyses reveal the conservation of an ancestral infectious strategy in two bacteriophage genera," *ISME journal*, vol. 11, pp. 1988–1996, 2017.

[101] D. Smalakyte, M. Kazlauskiene, J. F. Havelund, A. Rukšenaite, A. Rimaite, G. Tamulaitiene, N. J. Færgeman, G. Tamulaitis, and V. Siksnys, "Type III-A CRISPR-associated protein Csm6 degrades cyclic hexaadenylate activator using both CARF and HEPN domains," *Nucleic Acids Research*, vol. 48, pp. 9204–9217, 2020.

[102] S. Grüschow, C. S. Adamson, and M. F. White, "Specificity and sensitivity of an RNA targeting type III CRISPR complex coupled with a NucC endonuclease effector," *Nucleic acids research*, vol. 49, pp. 13 122–13 134, 2021.

[103] R. K. Lau, Q. Ye, E. A. Birkholz, K. R. Berg, L. Patel, I. T. Mathews, J. D. Watrous, K. Ego, A. T. Whiteley, B. Lowey, J. J. Mekalanos, P. J. Kranzusch, M. Jain, J. Pogliano, and K. D. Corbett, "Structure and mechanism of a cyclic trinucleotide-activated bacterial endonuclease mediating bacteriophage immunity," *Molecular Cell*, vol. 77, pp. 723–733.e6, 2020.

[104] D. Mayo-Muñoz, L. M. Smith, C. Garcia-Doval, L. M. Malone, K. R. Harding, S. A. Jackson, H. G. Hampton, R. D. Fagerlund, L. F. Gumy, and P. C. Fineran, "type III CRISPR-Cas provides resistance against nucleus-forming jumbo phages via abortive infection," *Molecular cell*, vol. 82, no. 23, pp. 4471–4486, 2022.

[105] A. M. Burroughs, D. Zhang, D. E. Schäffer, L. M. Iyer, and L. Aravind, "Comparative genomic analyses re-

veal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling," *Nucleic Acids Research*, vol. 43, pp. 10 633–10 654, 2015.

[106] B. Lowey, A. T. Whiteley, A. F. Keszei, B. R. Morehouse, I. T. Mathews, S. P. Antine, V. J. Cabrera, D. Kashin, P. Niemann, M. Jain, F. Schwede, J. J. Mekalanos, S. Shao, A. S. Lee, and P. J. Kranzusch, "CBASS immunity uses CARF-related effectors to sense 3′–5′- and 2′–5′-linked cyclic oligonucleotide signals and protect bacteria from phage infection," *Cell*, vol. 182, pp. 38–49.e17, 2020.

[107] C. Rouillon, N. Schneberger, H. Chi, K. Blumenstock, S. D. Vela, K. Ackermann, J. Moecking, M. F. Peter, W. Boenigk, R. Seifert, B. E. Bode, J. L. Schmid-Burgk, D. Svergun, M. Geyer, M. F. White, and G. Hagelueken, "Antiviral signalling by a cyclic nucleotide activated CRISPR protease," *Nature*, vol. 614, pp. 168–174, 2023.

[108] J. S. Athukoralage, S. Graham, C. Rouillon, S. Grüschow, C. M. Czekster, and M. F. White, "The dynamic interplay of host and viral enzymes in type III CRISPR-mediated cyclic nucleotide signalling," *eLife*, vol. 9, pp. 1–16, 2020.

[109] J. S. Athukoralage, S. A. McMahon, C. Zhang, S. Grüschow, S. Graham, M. Krupovic, R. J. Whitaker, T. M. Gloster, and M. F. White, "An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity," *Nature*, vol. 577, pp. 572–575, 2020.

[110] J. S. Athukoralage, C. Rouillon, S. Graham, S. Grüschow, and M. F. White, "Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate," *Nature*, vol. 562, pp. 277–280, 2018.

[111] S. Brown, C. C. Gauvin, A. A. Charbonneau, N. Burman, and C. M. Lawrence, "Csx3 is a cyclic oligonucleotide phosphodiesterase associated with type III CRISPR-Cas that degrades the second messenger ca4," *journal of Biological Chemistry*, vol. 295, pp. 14 963–14 972, 2020.

[112] K. Foster, S. Grüschow, S. Bailey, M. F. White, and M. P. Terns, "Regulation of the RNA and DNA nuclease activities required for Pyrococcus furiosus type III-B CRISPR-Cas immunity," *Nucleic Acids Research*, vol. 48, pp. 4418–4434, 2020.

[113] N. Jia, R. Jones, G. Yang, O. Ouerfelli, and D. J. Patel, "CRISPR-Cas III-A Csm6 CARF domain is a ring nuclease triggering stepwise cA4 cleavage with ApA>p formation terminating RNase activity," *Molecular Cell*, vol. 75, pp. 944–956, 2019.

[114] A. Samolygo, J. S. Athukoralage, S. Graham, and M. F. White, "Fuse to defuse: A self-limiting ribonuclease-ring nuclease fusion for type III CRISPR defence," *Nucleic Acids Research*, vol. 48, pp. 6149–6156, 2020.

[115] J. S. Athukoralage and M. F. White, "Cyclic oligoadenylate signaling and regulation by ring nucleases during type III CRISPR defense," *RNA*, vol. 27, pp. 855–867, 2021.

[116] R. Zhao, Y. Yang, F. Zheng, Z. Zeng, W. Feng, X. Jin, J. Wang, K. Yang, Y. X. Liang, Q. She, and W. Han, "A membrane-associated dhh-dhha1 nuclease degrades type III CRISPR second messenger," *Cell Reports*, vol. 32, 2020.

[117] A. A. Charbonneau, D. M. Eckert, C. C. Gauvin, N. G. Lintner, and C. M. Lawrence, "Cyclic tetra-adenylate (cA4) recognition by Csa3; implications for an integrated Class 1 CRISPR-Cas immune response in Saccharolobus solfataricus," *Biomolecules*, vol. 11, p. 1852, 2021.

[118] Q. Ye, X. Zhao, J. Liu, Z. Zeng, Z. Zhang, T. Liu, Y. Li, W. Han, and N. Peng, "CRISPR-associated factor Csa3b regulates CRISPR adaptation and Cmr-mediated RNA interference in Sulfolobus islandicus," *Frontiers in Microbiology*, vol. 11, pp. 1–12, 2020.

[119] P. Xia, A. Dutta, K. Gupta, M. Batish, and V. Parashar, "Structural basis of cyclic oligoadenylate binding to the transcription factor Csa3 outlines cross talk between type III and type I CRISPR systems," *journal of Biological Chemistry*, vol. 298, p. 101591, 2022.

[120] J. Jiao, N. Zheng, W. Wei, J. Fleming, X. Wang, Z. Li, L. Zhang, Y. Liu, Z. Zhang, A. Shen, L. Chuanyou, L. Bi, and H. Zhang, "M. tuberculosis CRISPR-Cas proteins are secreted virulence factors that trigger cellular immune responses," *Virulence*, vol. 12, pp. 3032–3044, 2021.

[121] L. Liu, X. Li, J. Wang, M. Wang, P. Chen, M. Yin, J. Li, G. Sheng, and Y. Wang, "Two distant catalytic sites are responsible for C2c2 RNase activities," *Cell*, vol. 168, pp. 121–134, 1 2017.

[122] L. Liu, X. Li, J. Ma, Z. Li, L. You, J. Wang, M. Wang, X. Zhang, and Y. Wang, "The molecular architecture for RNA-guided RNA cleavage by Cas13a," *Cell*, vol. 170, pp. 714–726.e10, 8 2017.

[123] O. O. Abudayyeh, J. S. Gootenberg, S. Konermann, J. Joung, I. M. Slaymaker, D. B. Cox, S. Shmakov, K. S. Makarova, E. Semenova, L. Minakhin, K. Severinov, A. Regev, E. S. Lander, E. V. Koonin, and F. Zhang, "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," *Science*, vol. 353, p. aaf5573, 8 2016.

[124] C. Xu, Y. Zhou, Q. Xiao, B. He, G. Geng, Z. Wang, B. Cao, X. Dong, W. Bai, Y. Wang, X. Wang, D. Zhou, T. Yuan, X. Huo, J. Lai, and H. Yang, "Programmable RNA editing with compact CRISPR–Cas13 systems

from uncultivated microbes," *Nature Methods*, vol. 18, pp. 499–506, 2021.
[125] W. X. Yan, S. Chong, H. Zhang, K. S. Makarova, E. V. Koonin, D. R. Cheng, and D. A. Scott, "Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a wyl-domain-containing accessory protein," *Molecular Cell*, vol. 70, pp. 327–339.e5, 2018.
[126] A. A. Smargon, D. B. Cox, N. K. Pyzocha, K. Zheng, I. M. Slaymaker, J. S. Gootenberg, O. A. Abudayyeh, P. Essletzbichler, S. Shmakov, K. S. Makarova, E. V. Koonin, and F. Zhang, "Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28," *Molecular Cell*, vol. 65, pp. 618–630.e7, 2 2017.
[127] J. S. Gootenberg, O. O. Abudayyeh, J. W. Lee, P. Essletzbichler, A. J. Dy, J. Joung, V. Verdine, N. Donghia, N. M. Daringer, C. A. Freije, C. Myhrvold, R. P. Bhattacharyya, J. Livny, A. Regev, E. V. Koonin, D. T. Hung, P. C. Sabeti, J. J. Collins, and F. Zhang, "Nucleic acid detection with CRISPR-Cas13a/C2c2," *Science*, vol. 356, pp. 438–442, 4 2017.
[128] A. Tambe, A. East-Seletsky, G. J. Knott, J. A. Doudna, and M. R. O'Connell, "RNA binding and HEPN-nuclease activation are decoupled in CRISPR-Cas13a," *Cell Reports*, vol. 24, pp. 1025–1036, 2018.
[129] A. J. Meeske, S. Nakandakari-Higa, and L. A. Marraffini, "Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage," *Nature*, vol. 570, pp. 241–245, 2019.
[130] B. A. Adler, T. Hessler, B. F. Cress, A. Lahiri, V. K. Mutalik, R. Barrangou, J. Banfield, and J. A. Doudna, "Broad-spectrum CRISPR-Cas13a enables efficient phage genome editing," *Nature microbiology*, vol. 7, no. 12, pp. 1967–1979, 2022.
[131] A. J. Meeske, N. Jia, A. K. Cassel, A. Kozlova, J. Liao, M. Wiedmann, D. J. Patel, and L. A. Marraffini, "A phage-encoded anti-CRISPR enables complete evasion of type VI-A CRISPR-Cas immunity," *Science*, vol. 369, pp. 54–59, 2020.
[132] A. R. VanderWal, J.-U. Park, B. Polevoda, J. K. Nicosia, A. M. M. Vargas, E. H. Kellogg, and M. R. O'Connell, "CRISPR-Csx28 forms a Cas13b-activated large-pore channel required for robust CRISPR-Cas immunity," *Biophysical journal*, vol. 122, no. 3, p. 194a, 2023.
[133] J. Guan, A. Oromí-Bosch, S. D. Mendoza, S. Karambelkar, J. D. Berry, and J. Bondy-Denomy, "Bacteriophage genome engineering with CRISPR–Cas13a," *Nature microbiology*, vol. 7, no. 12, pp. 1956–1966, 2022.
[134] S. Bandaru, M. H. Tsuji, Y. Shimizu, K. Usami, S. Lee, N. K. Takei, K. Yoshitome, Y. Nishimura, T. Otsuki, and T. Ito, "Structure-based design ofgRNA for Cas13," *Scientific Reports*, vol. 10, pp. 1–12, 2020.
[135] B. Wang, T. Zhang, J. Yin, Y. Yu, W. Xu, J. Ding, D. J. Patel, and H. Yang, "Structural basis for self-cleavage prevention by tag:anti-tag pairing complementarity in type VI Cas13 CRISPR systems," *Molecular Cell*, vol. 81, pp. 1100–1115.e5, 2021.
[136] A. J. Meeske and L. A. Marraffini, "RNA guide complementarity prevents self-targeting in type VI CRISPR systems," *Molecular Cell*, vol. 71, pp. 791–801, 2018.
[137] M. R. O'Connell, "Molecular mechanisms of RNA targeting by Cas13-containing type VI CRISPR–cas systems," *journal of Molecular Biology*, vol. 431, pp. 66–87, 2019.
[138] M. C. Williams, A. E. Reker, S. R. Margolis, J. Liao, M. Wiedmann, E. R. Rojas, and A. J. Meeske, "Phage genome cleavage enables resuscitation from Cas13 induced bacterial dormancy," *bioRxiv*, 2022.
[139] I. Jain, M. Kolesnik, K. Kuznedelov, L. Minakhin, N. Morozova, A. Shiriaeva, A. Kirillov, S. Medvedeva, A. Livenskyi, L. Kazieva *et al.*, "trna anticodon cleavage by target-activated crispr-cas13a effector," *Science Advances*, vol. 10, no. 17, p. eadl0164, 2024.
[140] Z. Su, B. Wilson, P. Kumar, and A. Dutta, "Noncanonical roles of tRNAs: tRNA fragments and beyond," *Annual Review of Genetics*, vol. 54, pp. 47–69, 2020.
[141] K. S. Makarova, L. Gao, F. Zhang, and E. V. Koonin, "Unexpected connections between type VI-B CRISPR-Cas systems, bacterial natural competence, ubiquitin signaling network and DNA modification through a distinct family of membrane proteins," *FEMS Microbiology Letters*, vol. 366, pp. 1–6, 2019.
[142] H. Shivram, B. F. Cress, G. J. Knott, and J. A. Doudna, "Controlling and enhancing CRISPR systems," *Nature Chemical Biology*, vol. 17, pp. 10–19, 2021.
[143] H. Zhang, C. Dong, L. Li, G. A. Wasney, and J. Min, "Structural insights into the modulatory role of the accessory protein WYL1 in the type VI-D CRISPR-Cas system," *Nucleic Acids Research*, vol. 47, pp. 5420–5428, 6 2019.
[144] F. Jiang and J. A. Doudna, "CRISPR – Cas9 structures and mechanisms," *Annu. Rev. Biophys.*, vol. 46, pp. 505–531, 2017.
[145] M. R. O'Connell, B. L. Oakes, S. H. Sternberg, A. East-Seletsky, M. Kaplan, and J. A. Doudna, "Programmable RNA recognition and cleavage by CRISPR-Cas9," *Nature*, vol. 516, pp. 263–266, 2014.
[146] S. C. Strutt, R. M. Torrez, E. Kaya, O. A. Negrete, and J. A. Doudna, "RNA-dependent RNA targeting by

CRISPR-Cas9," *eLife*, vol. 7, pp. 1–17, 2018.
[147] B. A. Rousseau, Z. Hou, M. J. Gramelspacher, and Y. Zhang, "Programmable RNA cleavage and recognition by a natural CRISPR-Cas9 system from Neisseria meningitidis," *Molecular Cell*, vol. 69, pp. 906–914.e4, 2018.
[148] T. R. Sampson, S. D. Saroj, A. C. Llewellyn, Y. L. Tzeng, and D. S. Weiss, "A CRISPR-Cas system mediates bacterial innate immune evasion and virulence," *Nature*, vol. 497, pp. 254–257, 2013.
[149] G. Dugar, R. T. Leenay, S. K. Eisenbart, T. Bischler, B. U. Aul, C. L. Beisel, and C. M. Sharma, "CRISPR RNA-dependent binding and cleavage of endogenous RNAs by the Campylobacter jejuni Cas9," *Molecular Cell*, vol. 69, pp. 893–905.e7, 2018.
[150] D. A. Nelles, M. Y. Fang, M. R. O'Connell, J. L. Xu, S. J. Markmiller, J. A. Doudna, and G. W. Yeo, "Programmable RNA tracking in live cells with CRISPR-Cas9," *Cell*, vol. 165, pp. 488–496, 2016.
[151] R. Batra, D. A. Nelles, E. Pirie, S. M. Blue, R. J. Marina, H. Wang, I. A. Chaim, J. D. Thomas, N. Zhang, V. Nguyen, S. Aigner, S. Markmiller, G. Xia, K. D. Corbett, M. S. Swanson, and G. W. Yeo, "Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9," *Cell*, vol. 170, pp. 899–912.e10, 2017.
[152] A. A. Price, T. R. Sampson, H. K. Ratner, A. Grakoui, and D. S. Weiss, "Cas9-mediated targeting of viral RNA in eukaryotic cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, pp. 6164–6169, 2015.
[153] W. X. Yan, P. Hunnewell, L. E. Alfonse, J. M. Carte, E. Keston-Smith, S. Sothiselvam, A. J. Garrity, S. Chong, K. S. Makarova, E. V. Koonin, D. R. Cheng, and D. A. Scott, "Functionally diverse type V CRISPR-Cas systems," *Science*, vol. 363, pp. 88–91, 2019.
[154] Z. Li, H. Zhang, R. Xiao, R. Han, and L. Chang, "cryo-EM structure of the RNA-guided ribonuclease Cas12g," *Nature Chemical Biology*, vol. 17, pp. 387–393, 2021.
[155] B. Tong, H. Dong, Y. Cui, P. Jiang, Z. Jin, and D. Zhang, "The versatile type V CRISPR effectors and their application prospects," *Frontiers in Cell and Developmental Biology*, vol. 8, pp. 1–11, 2021.
[156] M. P. Terns, "CRISPR-based technologies: Impact of RNA-targeting systems," *Molecular Cell*, vol. 72, pp. 404–412, 2018.
[157] S. Konermann, P. Lotfy, N. J. Brideau, J. Oki, M. N. Shokhirev, and P. D. Hsu, "Transcriptome engineering with RNA-targeting type VI-d CRISPR effectors," *Cell*, vol. 173, pp. 665–676.e14, 2018.
[158] J. E. Powell, C. K. Lim, R. Krishnan, T. X. McCallister, C. Saporito-Magriña, M. A. Zeballos, G. D. McPheron, and T. Gaj, "Targeted gene silencing in the nervous system with CRISPR-Cas13," *Science Advances*, vol. 8, pp. 1–12, 2022.
[159] K. Kiga, X. E. Tan, R. Ibarra-Chávez, S. Watanabe, Y. Aiba, Y. Sato'o, F. Y. Li, T. Sasahara, B. Cui, M. Kawauchi, T. Boonsiri, K. Thitiananpakorn, Y. Taki, A. H. Azam, M. Suzuki, J. R. Penadés, and L. Cui, "Development of CRISPR-Cas13a-based antimicrobials capable of sequence-specific killing of target bacteria," *Nature Communications*, vol. 11, pp. 1–11, 2020.
[160] S. Han, B. S. Zhao, S. A. Myers, S. A. Carr, C. He, and A. Y. Ting, "RNA-protein interaction mapping via MS2- or Cas13-based APEX targeting," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, pp. 22 068–22 079, 2020.
[161] Z. Zhang, W. Sun, T. Shi, P. Lu, M. Zhuang, and J. L. Liu, "Capturing RNA-protein interaction via CRUIS," *Nucleic Acids Research*, vol. 48, pp. 1–8, 2020.
[162] W. Yi, J. Li, X. Zhu, X. Wang, L. Fan, W. Sun, L. Liao, J. Zhang, X. Li, J. Ye, F. Chen, J. Taipale, K. M. Chan, L. Zhang, and J. Yan, "CRISPR-assisted detection of RNA–protein interactions in living cells," *Nature Methods*, vol. 17, pp. 685–688, 2020.
[163] O. O. Abudayyeh, J. S. Gootenberg, P. Essletzbichler, S. Han, J. Joung, J. J. Belanto, V. Verdine, D. B. Cox, M. J. Kellner, A. Regev, E. S. Lander, D. F. Voytas, A. Y. Ting, and F. Zhang, "RNA targeting with CRISPR-Cas13," *Nature*, vol. 550, pp. 280–284, 2017.
[164] D. B. Cox, J. S. Gootenberg, O. O. Abudayyeh, B. Franklin, M. J. Kellner, J. Joung, and F. Zhang, "RNA editing with CRISPR-Cas13," *Science*, vol. 358, pp. 1019–1027, 2017.
[165] S. Kannan, H. Altae-Tran, X. Jin, V. J. Madigan, R. Oshiro, K. S. Makarova, E. V. Koonin, and F. Zhang, "Compact RNA editors with small Cas13 proteins," *Nature Biotechnology*, vol. 40, pp. 194–197, 2022.
[166] O. O. Abudayyeh, J. S. Gootenberg, B. Franklin, J. Koob, M. J. Kellner, A. Ladha, J. Joung, P. Kirchgatterer, D. B. Cox, and F. Zhang, "A cytosine deaminase for programmable single-base RNA editing," *Science*, vol. 365, pp. 382–386, 2019.
[167] J. Li, Z. Chen, F. Chen, G. Xie, Y. Ling, Y. Peng, Y. Lin, N. Luo, C. M. Chiang, and H. Wang, "Targeted mRNA demethylation using an engineered dCas13b-ALKBH5 fusion protein," *Nucleic Acids Research*, vol. 48, pp. 5684–5694, 2020.
[168] J. Arizti-Sanz, C. A. Freije, A. C. Stanton, B. A. Petros, C. K. Boehm, S. Siddiqui, B. M. Shaw, G. Adams,

T. S. F. Kosoko-Thoroddsen, M. E. Kemball, J. N. Uwanibe, F. V. Ajogbasile, P. E. Eromon, R. Gross, L. Wronka, K. Caviness, L. E. Hensley, N. H. Bergman, B. L. MacInnis, C. T. Happi, J. E. Lemieux, P. C. Sabeti, and C. Myhrvold, "Streamlined inactivation, amplification, and Cas13-based detection of SARS-CoV-2," *Nature Communications*, vol. 11, 2020.

[169] H. Shinoda, Y. Taguchi, R. Nakagawa, A. Makino, S. Okazaki, M. Nakano, Y. Muramoto, C. Takahashi, I. Takahashi, J. Ando, T. Noda, O. Nureki, H. Nishimasu, and R. Watanabe, "Amplification-free RNA detection with CRISPR–Cas13," *Communications Biology*, vol. 4, 2021.

[170] J. N. Rauch, E. Valois, S. C. Solley, F. Braig, R. S. Lach, and M. Audouard, "A scalable , easy-to-deploy protocol for Cas13-based detection," *journal of Clinical Microbiology*, vol. 59, pp. e02 402–20, 2021.

[171] C. M. Ackerman, C. Myhrvold, S. G. Thakku, C. A. Freije, H. C. Metsky, D. K. Yang, S. H. Ye, C. K. Boehm, T. S. F. Kosoko-Thoroddsen, J. Kehe, T. G. Nguyen, A. Carter, A. Kulesa, J. R. Barnes, V. G. Dugan, D. T. Hung, P. C. Blainey, and P. C. Sabeti, "Massively multiplexed nucleic acid detection with Cas13," *Nature*, vol. 582, pp. 277–282, 2020.

[172] T. Tian, B. Shu, Y. Jiang, M. Ye, L. Liu, Z. Guo, Z. Han, Z. Wang, and X. Zhou, "An ultralocalized Cas13a assay enables universal and nucleic acid amplification-free single-molecule RNA diagnostics," *ACS Nano*, vol. 15, pp. 1167–1178, 2021.

[173] J. Shen, X. Zhou, Y. Shan, H. Yue, R. Huang, J. Hu, and D. Xing, "Sensitive detection of a bacterial pathogen using allosteric probe-initiated catalysis and CRISPR-Cas13a amplification reaction," *Nature Communications*, vol. 11, pp. 1–10, 2020.

[174] W. K. Spoelstra, J. M. Jacques, R. Gonzalez-Linares, F. L. Nobrega, A. C. Haagsma, M. Dogterom, D. H. Meijer, T. Idema, S. J. J. Brouns, and L. Reese, "CRISPR-based DNA and RNA detection with liquid-liquid phase separation," *Biophysical journal*, vol. 120, pp. 1198–1209, 2021.

[175] R. Bruch, J. Baaske, C. Chatelle, M. Meirich, S. Madlener, W. Weber, C. Dincer, and G. A. Urban, "CRISPR-Cas13a-powered electrochemical microfluidic biosensor for nucleic acid amplification-free miRNA diagnostics," *Advanced Materials*, vol. 31, 2019.

[176] T. Zhou, R. Huang, M. Huang, J. Shen, Y. Shan, and D. Xing, "CRISPR-Cas13a powered portable electrochemiluminescence chip for ultrasensitive and specific miRNA detection," *Advanced Science*, vol. 7, pp. 1–10, 2020.

[177] C. Yuan, T. Tian, J. Sun, M. Hu, X. Wang, E. Xiong, M. Cheng, Y. Bao, W. Lin, J. Jiang, C. Yang, Q. Chen, H. Zhang, H. Wang, X. Wang, X. Deng, X. Liao, Y. Liu, Z. Wang, G. Zhang, and X. Zhou, "Universal and naked-eye gene detection platform based on the clustered regularly interspaced short palindromic repeats/Cas12a/13a system," *Analytical Chemistry*, vol. 92, pp. 4029–4037, 2020.

[178] Y. Shan, X. Zhou, R. Huang, and D. Xing, "High-fidelity and rapid quantification of miRNA combining crrna programmability and CRISPR-Cas13a trans-cleavage activity," *Analytical Chemistry*, vol. 91, pp. 5278–5285, 2019.

[179] M. M. Kaminski, O. O. Abudayyeh, J. S. Gootenberg, F. Zhang, and J. J. Collins, "CRISPR-based diagnostics," *Nature Biomedical Engineering*, vol. 5, pp. 643–656, 2021.

[180] C. A. Freije and P. C. Sabeti, "Detect and destroy: CRISPR-based technologies for the response against viruses," *Cell host & microbe*, vol. 29, pp. 689–703, 2021.

[181] P. Fozouni, S. Son, M. D. de León Derby, G. J. Knott, C. N. Gray, M. V. D'Ambrosio, C. Zhao, N. A. Switz, G. R. Kumar, S. I. Stephens, D. Boehm, C. L. Tsou, J. Shu, A. Bhuiya, M. Armstrong, A. R. Harris, P. Y. Chen, J. M. Osterloh, A. Meyer-Franke, B. Joehnk, K. Walcott, A. Sil, C. Langelier, K. S. Pollard, E. D. Crawford, A. S. Puschnik, M. Phelps, A. Kistler, J. L. DeRisi, J. A. Doudna, D. A. Fletcher, and M. Ott, "Amplification-free detection of SARS-CoV-2 with CRISPR-Cas13a and mobile phone microscopy," *Cell*, vol. 184, pp. 323–333.e9, 2021.

[182] S. Sridhara, H. N. Goswami, C. Whyms, J. H. Dennis, and H. Li, "Virus detection via programmable type III-A CRISPR-Cas systems," *Nature Communications*, vol. 12, pp. 1–10, 2021.

[183] A. Santiago-Frangos, L. N. Hall, A. Nemudraia, A. Nemudryi, P. Krishna, T. Wiegand, R. A. Wilkinson, D. T. Snyder, J. F. Hedges, C. Cicha, H. H. Lee, A. Graham, M. A. Jutila, M. P. Taylor, and B. Wiedenheft, "Intrinsic signal amplification by type III CRISPR-Cas systems provides a sequence-specific SARS-CoV-2 diagnostic," *Cell Reports Medicine*, vol. 2, p. 100319, 2021.

[184] A. Nemudraia, A. Nemudryi, M. Buyukyoruk, A. M. Scherffius, T. Zahl, T. Wiegand, S. Pandey, J. E. Nichols, L. N. Hall, A. McVey *et al.*, "Sequence-specific capture and concentration of viral RNA by type III CRISPR system enhances diagnostic," *Nature Communications*, vol. 13, no. 1, p. 7762, 2022.

[185] T. Y. Liu, G. J. Knott, D. C. Smock, J. J. Desmarais, S. Son, A. Bhuiya, S. Jakhanwal, N. Prywes, S. Agrawal, M. D. de León Derby, N. A. Switz, M. Armstrong, A. R. Harris, E. J. Charles, B. W. Thornton, P. Fozouni, J. Shu, S. I. Stephens, G. R. Kumar, C. Zhao, A. Mok, A. T. Iavarone, A. M. Escajeda, R. McIntosh, S. Kim,

E. J. Dugan, J. R. Hamilton, E. Lin-Shiao, E. C. Stahl, C. A. Tsuchida, E. A. Moehle, P. Giannikopoulos, M. McElroy, S. McDevitt, A. Zur, I. Sylvain, A. Ciling, M. Zhu, C. Williams, A. Baldwin, K. S. Pollard, M. X. Tan, M. Ott, D. A. Fletcher, L. F. Lareau, P. D. Hsu, D. F. Savage, and J. A. Doudna, "Accelerated RNA detection using tandem CRISPR nucleases," *Nature Chemical Biology*, vol. 17, pp. 982–988, 2021.

[186] E. S. Rittershaus, S. H. Baek, and C. M. Sassetti, "The normalcy of dormancy: Common themes in microbial quiescence," *Cell host & microbe*, vol. 13, pp. 643–651, 2013.

[187] J. Dworkin and I. M. Shah, "Exit from dormancy in microbial organisms," *Nature Reviews Microbiology*, vol. 8, pp. 890–896, 2010.

[188] A. Lopatina, N. Tal, and R. Sorek, "Abortive infection: Bacterial suicide as an antiviral immune strategy," *Annual Review of Virology*, vol. 7, pp. 1–14, 2020.

[189] E. V. Koonin and F. Zhang, "Coupling immunity and programmed cell suicide in prokaryotes: Life-or-death choices," *BioEssays*, vol. 39, pp. 1–9, 2017.

[190] K. S. Makarova, V. Anantharaman, L. Aravind, and E. V. Koonin, "Live virus-free or die: coupling of antivirus immunity and programmed suicide or dormancy in prokaryotes," *Biology Direct*, vol. 7, pp. 1–10, 2012.

[191] S. D. Mendoza and J. Bondy-Denomy, "Cas13 helps bacteria play dead when the enemy strikes," *Cell host & microbe*, vol. 26, pp. 1–2, 2019.

[192] C. Weigel and H. Seitz, "Bacteriophage replication modules," *FEMS Microbiology Reviews*, vol. 30, pp. 321–381, 2006.

[193] D. Muñoz-Espín, G. Serrano-Heras, and M. Salas, "Role of host factors in bacteriophage ϕ29 DNA replication," *Advances in Virus Research*, vol. 82, pp. 351–383, 2012.

[194] A. Chevallereau, B. J. Pons, S. van Houte, and E. R. Westra, "Interactions between bacterial and phage communities in natural environments," *Nature Reviews Microbiology*, vol. 20, pp. 49–62, 2022.

[195] J. N. A. Vink, K. J. Martens, M. Vlot, R. E. McKenzie, C. Almendros, B. E. Bonilla, D. J. Brocken, J. Hohlbein, and S. J. J. Brouns, "Direct visualization of native CRISPR target search in live bacteria reveals cascade DNA surveillance mechanism," *Molecular Cell*, vol. 77, pp. 39–50.e10, 2020.

[196] T. Dimitriu, E. Kurilovich, U. Łapińska, K. Severinov, S. Pagliara, M. D. Szczelkun, and E. R. Westra, "Bacteriostatic antibiotics promote CRISPR-Cas adaptive immunity by enabling increased spacer acquisition," *Cell host & microbe*, vol. 30, pp. 31–40.e5, 2022.

[197] P. A. de Jonge, F. L. Nobrega, S. J. J. Brouns, and B. E. Dutilh, "Molecular and evolutionary determinants of bacteriophage host range," *Trends in Microbiology*, vol. 27, pp. 51–63, 2019.

[198] A. Ndhlovu, P. M. Durand, and G. Ramsey, "Programmed cell death as a black queen in microbial communities," *Molecular Ecology*, vol. 30, pp. 1110–1119, 2021.

[199] S. H. Peeters and M. I. de Jonge, "For the greater good: Programmed cell death in bacterial communities," *Microbiological Research*, vol. 207, pp. 161–169, 2018.

[200] P. Weigele and E. A. Raleigh, "Biosynthesis and function of modified bases in bacteria and their viruses," *Chemical Reviews*, vol. 116, pp. 12 655–12 687, 2016.

[201] R. T. Leenay and C. L. Beisel, "Deciphering, communicating, and engineering the CRISPR pam," *journal of Molecular Biology*, vol. 429, pp. 177–191, 2017.

[202] N. Tal and R. Sorek, "Snapshot: Bacterial immunity," *Cell*, vol. 185, pp. 578–578.e1, 2022.

[203] B. Y. Hsueh, G. B. Severin, C. A. Elg, E. J. Waldron, A. Kant, A. J. Wessel, J. A. Dover, C. R. Rhoades, B. J. Ridenhour, K. N. Parent, M. B. Neiditch, J. Ravi, E. M. Top, and C. M. Waters, "Phage defence by deaminase-mediated depletion of deoxynucleotides in bacteria," *Nature Microbiology*, vol. 7, pp. 1210–1220, 2022.

[204] F. Rousset and R. Sorek, "The evolutionary success of regulated cell death in bacterial immunity," *Current Opinion in Microbiology*, vol. 74, p. 102312, 2023.

[205] N. Aframian and A. Eldar, "Abortive infection antiphage defense systems: separating mechanism and phenotype," *Trends in Microbiology*, 2023.

[206] S. P. B. van Beljouw, A. C. Haagsma, A. Rodríguez-Molina, D. F. van den Berg, J. N. A. Vink, and S. J. J. Brouns, "The gramp CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase," *Science*, vol. 373, pp. 1349–1353, 2021.

[207] S. Doron, S. Melamed, G. Ofir, A. Leavitt, A. Lopatina, M. Keren, G. Amitai, and R. Sorek, "Systematic discovery of antiphage defense systems in the microbial pangenome," *Science*, vol. 359, p. eaar4120, 2018.

[208] S. Silas, G. Mohr, D. J. Sidote, L. M. Markham, A. Sanchez-Amat, D. Bhaya, A. M. Lambowitz, and A. Z. Fire, "Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein," *Science*, vol. 351, p. aad4234, 2016.

[209] A. G. Johnson, T. Wein, M. L. Mayer, B. Duncan-Lowey, E. Yirmiya, Y. Oppenheimer-Shaanan, G. Amitai,

R. Sorek, and P. J. Kranzusch, "Bacterial gasdermins reveal an ancient mechanism of cell death," *Science*, vol. 375, pp. 221–225, 2022.

[210] P. Pausch, B. Al-Shayeb, E. Bisom-Rapp, C. A. Tsuchida, Z. Li, B. F. Cress, G. J. Knott, S. E. Jacobsen, J. F. Banfield, and J. A. Doudna, "CRISPR-Casphi from huge phages is a hypercompact genome editor," *Science*, vol. 369, pp. 333–337, 2020.

[211] M. Schmid, K. Walsh, R. Webb, W. I. C. Rijpstra, K. V. D. Pas-Schoonen, M. J. Verbruggen, T. Hill, B. Moffett, J. Fuerst, S. Schouten, J. S. Damsté, J. Harris, P. Shaw, M. Jetten, and M. Strous, "Candidatus "Scalindua brodae", sp. nov., Candidatus "Scalindua wagneri", sp. nov., two new species of anaerobic ammonium oxidizing bacteria," *Systematic and Applied Microbiology*, vol. 26, pp. 529–538, 2003.

[212] I. Fonfara, H. Richter, M. Bratoviä, A. L. Rhun, and E. Charpentier, "The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA," *Nature*, vol. 532, pp. 517–521, 2016.

[213] L. You, J. Ma, J. Wang, D. Artamonova, M. Wang, L. Liu, H. Xiang, K. Severinov, X. Zhang, and Y. Wang, "Structure studies of the CRISPR-Csm complex reveal mechanism of co-transcriptional interference," *Cell*, vol. 176, pp. 239–253, 1 2019.

[214] C. Benda, J. Ebert, R. A. Scheltema, H. B. Schiller, M. Baumgärtner, F. Bonneau, M. Mann, and E. Conti, "Structural model of a CRISPR RNA-silencing complex reveals the RNA-target cleavage activity in Cmr4," *Molecular Cell*, vol. 56, pp. 43–54, 2014.

[215] M. Martin, "Cutadapt removes adapter sequences from high-throughput sequencing reads," *EMBnet. journal*, vol. 17, pp. 10–12, 2011.

[216] H. Li, "Minimap2: Pairwise alignment for nucleotide sequences," *Bioinformatics*, vol. 34, pp. 3094–3100, 2018.

[217] E. Gasteiger, C. Hoogland, A. Gattiker, M. R. Wilkins, R. D. Appel, and A. Bairoch, *Protein identification and analysis tools on the ExPASy server.*, 2005.

[218] M. Steinegger and J. Söding, "MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets," *Nature Biotechnology*, vol. 35, pp. 1026–1028, 2017.

[219] J. Söding, A. Biegert, and A. N. Lupas, "The HHpred interactive server for protein homology detection and structure prediction," *Nucleic Acids Research*, vol. 33, pp. 244–248, 2005.

[220] J. Russel, R. Pinilla-Redondo, D. Mayo-Muñoz, S. A. Shah, and S. J. Sørensen, "Crisprcastyper: Automated identification, annotation, and classification of CRISPR-Cas loci," *CRISPR journal*, vol. 3, pp. 462–469, 2020.

[221] T. Seemann, "Prokka: Rapid prokaryotic genome annotation," *Bioinformatics*, vol. 30, pp. 2068–2069, 2014.

[222] L. J. Hauser, D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, and F. W. Larimer, "Prodigal: prokaryotic gene recognition and translation initation site identification," *BMC Bioinformatics*, vol. 11, 2010.

[223] L. Fu, B. Niu, Z. Zhu, S. Wu, and W. Li, "CD-HIT: Accelerated for clustering the next-generation sequencing data," *Bioinformatics*, vol. 28, pp. 3150–3152, 2012.

[224] K. Katoh and H. Toh, "Parallelization of the mafft multiple sequence alignment program," *Bioinformatics*, vol. 26, pp. 1899–1900, 2010.

[225] S. Capella-Gutiérrez, J. M. Silla-Martínez, and T. Gabaldón, "trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses," *Bioinformatics*, vol. 25, pp. 1972–1973, 2009.

[226] L. T. Nguyen, H. A. Schmidt, A. V. Haeseler, and B. Q. Minh, "IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies," *Molecular Biology and Evolution*, vol. 32, pp. 268–274, 2015.

[227] I. Letunic and P. Bork, "Interactive Tree of Life (iTOL) v4: Recent updates and new developments," *Nucleic Acids Research*, vol. 47, pp. 256–259, 2019.

[228] J. S. Papadopoulos and R. Agarwala, "COBALT: Constraint-based alignment tool for multiple protein sequences," *Bioinformatics*, vol. 23, pp. 1073–1079, 2007.

[229] A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp, and G. J. Barton, "Jalview version 2-A multiple sequence alignment editor and analysis workbench," *Bioinformatics*, vol. 25, pp. 1189–1191, 2009.

[230] X. Robert and P. Gouet, "Deciphering key features in protein structures with the new endscript server," *Nucleic Acids Research*, vol. 42, pp. 320–324, 2014.

[231] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. Sternberg, "The Phyre2 web portal for protein modeling, prediction and analysis," *Nature Protocols*, vol. 10, pp. 845–858, 2015.

[232] W. L. DeLano, "Pymol: An open-source molecular graphics tool." *CCP4 Newsletter on protein crystallography*, vol. 40, pp. 82–92, 2002.

[233] C. Hu, S. P. B. van Beljouw, K. H. Nam, G. Schuler, F. Ding, Y. Cui, A. Rodríguez-Molina, A. C. Haagsma, M. Valk, M. Pabst, S. J. J. Brouns, and A. Ke, "Craspase is a CRISPR RNA-guided, RNA-activated protease,"

Science, vol. 377, pp. 1278–1285, 2022.
[234] L. D. D'Andrea and L. Regan, "TPR proteins: The versatile helix," *Trends in Biochemical Sciences*, vol. 28, pp. 655–662, 2003.
[235] L. Aravind and E. V. Koonin, "Classification of the caspase-hemoglobinase fold: Detection of new families and implications for the origin of the eukaryotic separins," *Proteins: Structure, Function and Genetics*, vol. 46, pp. 355–367, 2002.
[236] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, pp. 583–589, 2021.
[237] T. Awata, M. Oshiki, T. Kindaichi, N. Ozaki, A. Ohashi, and S. Okabe, "Physiological characterization of an anaerobic ammonium-oxidizing bacterium belonging to the "Candidatus Scalindua" group," *Applied and Environmental Microbiology*, vol. 79, pp. 4145–4148, 2013.
[238] M. Pabst, D. S. Grouzdev, C. E. Lawson, H. B. Kleikamp, C. de Ram, R. Louwen, Y. M. Lin, S. Lücker, M. C. van Loosdrecht, and M. Laureni, "A general approach to explore prokaryotic protein glycosylation reveals the unique surface layer modulation of an anammox bacterium," *The ISME Journal*, vol. 16, no. 2, pp. 346–357, 2022.
[239] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *journal of Molecular Biology*, vol. 215, pp. 403–410, 1990.
[240] K. Kato, W. Zhou, S. Okazaki, Y. Isayama, T. Nishizawa, J. S. Gootenberg, O. O. Abudayyeh, and H. Nishimasu, "Structure and engineering of the type III-E CRISPR-Cas7-11 effector complex," *Cell*, vol. 185, pp. 2324–2337.e16, 2022.
[241] N. Jia, C. Y. Mo, C. Wang, E. T. Eng, L. A. Marraffini, and D. J. Patel, "Type III-A CRISPR-Cas Csm complexes: Assembly, periodic RNA cleavage, DNase activity regulation, and autoimmunity," *Molecular Cell*, vol. 73, pp. 264–277, 2019.
[242] S. Sridhara, J. Rai, C. Whyms, H. Goswami, H. He, W. Woodside, M. P. Terns, and H. Li, "Structural and biochemical characterization of in vivo assembled Lactococcus lactis CRISPR-Csm complex," *Communications Biology*, vol. 5, pp. 1–12, 2022.
[243] Z. Lin, X. Luo, and H. Yu, "Structural basis of cohesin cleavage by separase," *Nature*, vol. 532, pp. 131–134, 2016.
[244] A. Boland, T. G. Martin, Z. Zhang, J. Yang, X. C. Bai, L. Chang, S. H. Scheres, and D. Barford, "Cryo-EM structure of a metazoan separase-securin complex at near-atomic resolution," *Nature Structural and Molecular Biology*, vol. 24, pp. 414–418, 2017.
[245] Z. Liu, C. Wang, J. Yang, Y. Chen, B. Zhou, D. W. Abbott, and T. S. Xiao, "Caspase-1 engages full-length Gasdermin D through two distinct interfaces that mediate caspase recruitment and substrate cleavage," *Immunity*, vol. 53, pp. 106–114.e5, 2020.
[246] M. L. G. L. Blatch, "The tetratricopeptide repeat: a structural motif mediating protein-protein interactions," *BioEssays*, pp. 932–939, 1999.
[247] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, "CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination," *Nature Methods*, vol. 14, pp. 290–296, 2017.
[248] S. H. Scheres, "RELION: Implementation of a bayesian approach to cryo-EM structure determination," *journal of Structural Biology*, vol. 180, pp. 519–530, 2012.
[249] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, "UCSF chimera - a visualization system for exploratory research and analysis," *journal of Computational Chemistry*, vol. 25, pp. 1605–1612, 2004.
[250] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot," *Acta Crystallographica Section D: Biological Crystallography*, vol. 66, pp. 486–501, 2010.
[251] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, and P. D. Adams, "Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix," *Acta Crystallographica Section D: Structural Biology*, vol. 75, pp. 861–877, 2019.
[252] C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S.

Richardson, and D. C. Richardson, "MolProbity: More and better reference data for improved all-atom structure validation," *Protein Science*, vol. 27, pp. 293–315, 2018.

[253] S. P. B. van Beljouw, A. C. Haagsma, K. Kalogeropoulos, M. Pabst, and S. J. J. Brouns, "Craspase orthologs cleave a nonconserved site in target protein Csx30," *ACS Chemical Biology*, vol. 19, no. 5, pp. 1051–1055, 2024.

[254] X. Liu, L. Zhang, H. Wang, Y. Xiu, L. Huang, Z. Gao, N. Li, F. Li, and W. Xiong, "Target RNA activates the protease activity of Craspase to confer antiviral defense," *Molecular Cell*, vol. 82, no. 23, pp. 4503–4518, 2022.

[255] J. Strecker, F. E. Demircioglu, D. Li, G. Faure, M. E. Wilkinson, J. S. Gootenberg, O. O. Abudayyeh, H. Nishimasu, R. K. Macrae, and F. Zhang, "RNA-activated protein cleavage with a CRISPR-associated endopeptidase," *Science*, vol. 378, pp. 874–881, 2022.

[256] M. Ali, M. Oshiki, T. Awata, K. Isobe, Z. Kimura, H. Yoshikawa, D. Hira, T. Kindaichi, H. Satoh, T. Fujii, and S. Okabe, "Physiological characterization of anaerobic ammonium oxidizing bacterium 'Candidatus Jettenia caeni'," *Environmental Microbiology*, vol. 17, pp. 2172–2189, 2015.

[257] A. Winter, R. Schmid, and R. Bayliss, "Structural insights into separase architecture and substrate recognition through computational modelling of caspase-like and death domains," *PLoS Computational Biology*, vol. 11, 2015.

[258] J. C. Carrington and W. G. Dougherty, "A viral cleavage site cassette: Identification of amino acid sequences required for tobacco etch virus polyprotein processing," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 85, pp. 3391–3395, 1988.

[259] S. Wang, M. Guo, Y. Zhu, Z. Lin, and Z. Huang, "Cryo-EM structure of the type III-E CRISPR-Cas effector gRAMP in complex with TPR-CHAT," *Cell Research*, vol. 32, no. 12, pp. 1128–1131, 2022.

[260] K. Wang, Q. Sun, X. Zhong, M. Zeng, H. Zeng, X. Shi, Z. Li, Y. Wang, Q. Zhao, F. Shao, and J. Ding, "Structural mechanism for GSDMD targeting by autoprocessed caspases in pyroptosis," *Cell*, vol. 180, pp. 941–955.e20, 2020.

[261] Q. He, X. Lei, Y. Liu, X. Wang, N. Ji, H. Yin, H. Wang, H. Zhang, and G. Yu, "Nucleic acid detection through RNA-guided protease activity in type III-E CRISPR-Cas systems," *ChemBioChem*, 9 2023.

[262] J. Chen, N. Quiles-Puchalt, Y. N. Chiang, R. Bacigalupe, A. Fillol-Salom, M. S. J. Chee, J. R. Fitzgerald, and J. R. Penadés, "Genome hypermobility by lateral transduction," *Science*, vol. 362, pp. 207–212, 2018.

[263] D. Mayo-Muñoz, R. Pinilla-Redondo, N. Birkholz, and P. C. Fineran, "A host of armor: Prokaryotic immune strategies against mobile genetic elements," *Cell Reports*, vol. 42, 7 2023.

[264] T. A. Boyle and A. Hatoum-Aslan, "Recurring and emerging themes in prokaryotic innate immunity," *Current Opinion in Microbiology*, vol. 73, p. 102324, 2023.

[265] P. Horvath and R. Barrangou, "CRISPR/Cas, the immune system of bacteria and archaea," *Science*, vol. 327, no. 5962, pp. 167–170, 2010.

[266] E. V. Koonin and L. Aravind, "Origin and evolution of eukaryotic apoptosis: the bacterial connection." *Cell death and differentiation*, vol. 9, pp. 394–404, 2002.

[267] H. N. Goswami, J. Rai, A. Das, and H. Li, "Molecular mechanism of active Cas7-11 in processing CRISPR RNA and interfering target RNA," *eLife*, vol. 11, pp. 1–15, 2022.

[268] N. Cui, J. T. Zhang, Z. Li, X. Y. Liu, C. Wang, H. Huang, and N. Jia, "Structural basis for the non-self RNA-activated protease activity of the type III-E CRISPR nuclease-protease Craspase," *Nature Communications*, vol. 13, 2022.

[269] Y. Huo, H. Zhao, Q. Dong, and T. Jiang, "Cryo-EM structure and protease activity of the type III-E CRISPR-Cas effector," *Nature Microbiology*, 2023.

[270] B. Ekundayo, D. Torre, B. Beckert, S. Nazarov, A. Myasnikov, H. Stahlberg, and D. Ni, "Structural insights into the regulation of Cas7-11 by TPR-CHAT," *Nature Structural and Molecular Biology*, vol. 30, pp. 135–139, 2022.

[271] G. Yu, X. Wang, Y. Zhang, Q. An, Y. Wen, X. Li, H. Yin, Z. Deng, and H. Zhang, "Structure and function of a bacterial type III-E CRISPR – Cas7-11 complex," *Nature Microbiology*, vol. 7, pp. 2078–2088, 2022.

[272] A. Feklístov, B. D. Sharon, S. A. Darst, and C. A. Gross, "Bacterial sigma factors: A historical, structural, and genomic perspective," *Annual Review of Microbiology*, vol. 68, pp. 357–376, 2014.

[273] S. A. Jackson, R. E. McKenzie, R. D. Fagerlund, S. N. Kieper, P. C. Fineran, and S. J. J. Brouns, "CRISPR-Cas: Adapting to change," *Science*, vol. 356, p. 40, 2017.

[274] A. C. Clark, "Caspase allostery and conformational selection," *Chemical Reviews*, vol. 116, pp. 6666–6706, 2016.

[275] B. Elsässer, F. B. Zauner, J. Messner, W. T. Soh, E. Dall, and H. Brandstetter, "Distinct roles of catalytic cysteine and histidine in the protease and ligase mechanisms of human legumain as revealed by DFT-

based QM/MM simulations," *ACS Catalysis*, vol. 7, pp. 5585–5593, 2017.
[276] B. T. U. Schwartz, "Solving the nuclear pore puzzle," *Science*, vol. 376, pp. 1158–1160, 2022.
[277] D. Taylor, E. Schwartz, G. Palermo, J. Bravo, M. Ahsan, L. Macias, C. McCafferty, T. Dangerfield, J. Walker, J. Brodbelt, P. Fineran, and R. Fagerlund, "Type III CRISPR-Cas effectors act as protein-assisted ribozymes during RNA cleavage," *Research Square*, 2023.
[278] A. Ganguly and E. Yoo, "Sotorasib: a KRASG12C inhibitor for non-small cell lung cancer," pp. 536–537, 6 2022.
[279] T. E. Quax, Y. I. Wolf, J. J. Koehorst, O. Wurtzel, R. van der Oost, W. Ran, F. Blombach, K. S. Makarova, S. J. J. Brouns, A. C. Forster, E. G. H. Wagner, R. Sorek, E. V. Koonin, and J. van der Oost, "Differential translation tunes uneven production of operon-encoded proteins," *Cell Reports*, vol. 4, pp. 938–944, 5 2013.
[280] J. Zhang, S. Graham, A. Tello, H. Liu, and M. F. White, "Multiple nucleic acid cleavage modes in divergent type III CRISPR systems," *Nucleic Acids Research*, vol. 44, pp. 1789–1799, 2016.
[281] W. F. Martin, S. Garg, and V. Zimorski, "Endosymbiotic theories for eukaryote origin," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, 2015.
[282] J. M. Frade and T. M. Michaelidis, "Origin of eukaryotic programmed cell death: A consequence of aerobic metabolism?" *BioEssays*, vol. 19, pp. 827–832, 1997.
[283] C. Ku, S. Nelson-Sathi, M. Roettger, F. L. Sousa, P. J. Lockhart, D. Bryant, E. Hazkani-Covo, J. O. McInerney, G. Landan, and W. F. Martin, "Endosymbiotic origin and differential loss of eukaryotic genes," *Nature*, vol. 524, pp. 427–432, 2015.
[284] W. Y. Xing, J. Liu, Z. Q. Wang, J. Y. Zhang, X. Zeng, Y. Yang, and C. C. Zhang, "HetF protein is a new divisome component in a filamentous and developmental cyanobacterium," *mBio*, vol. 12, 2021.
[285] C. Anders and M. Jinek, *In vitro enzymology of Cas9*, 1st ed. Elsevier Inc., 2014, vol. 546.
[286] M. P. Robertson and G. F. Joyce, "The origins of the RNA world," *Cold Spring Harbor Perspectives in Biology*, vol. 4, p. 1, 2012.
[287] E. V. Koonin, K. S. Makarova, and E. V. Koonin, "Origins and evolution of CRISPR-Cas systems," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 374, p. 1772, 2019.
[288] C. F. Schuster and R. Bertram, "Toxin-antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate," *FEMS Microbiology Letters*, vol. 340, pp. 73–85, 2013.
[289] J. A. Steens, J. P. K. Bravo, C. R. P. Salazar, C. Yildiz, A. M. Amieiro, S. Köstlbacher, S. H. P. Prinsen, A. S. Andres, C. Patinios, A. Bardis, A. Barendregt, R. A. Scheltema, T. J. G. Ettema, J. van der Oost, D. W. Taylor, and R. H. J. Staals, "Type III-B CRISPR-Cas cascade of proteolytic cleavages," *Science*, vol. 383, no. 6682, pp. 512–519, 2024.
[290] K. Richardson, W. Steffen, W. Lucht, J. Bendtsen, S. E. Cornell, J. F. Donges, M. Drüke, I. Fetzer, G. Bala, W. von Bloh, G. Feulner, S. Fiedler, D. Gerten, T. Gleeson, M. Hofmann, W. Huiskamp, M. Kummu, C. Mohan, D. Nogués-Bravo, S. Petri, M. Porkka, S. Rahmstorf, S. Schaphoff, K. Thonicke, A. Tobian, V. Virkki, L. Wang-Erlandsson, L. Weber, and J. Rockström, "Earth beyond six of nine planetary boundaries," *Science Advances*, vol. 9, no. 37, p. eadh2458, 2023.
[291] R. H. Cowie, P. Bouchet, and B. Fontaine, "The sixth mass extinction: fact, fiction or speculation?" *Biological Reviews*, vol. 97, no. 2, pp. 640–663, 2022.
[292] I. McGilchrist, *The Matter with Things: Our Brains, Our Delusions, and the Unmaking of the World*. Perspectiva Press, 2021.
[293] C. Peirce, "How to make our ideas clear, by charles s. peirce. popular science monthly 12 (january 1878), 286-302." *Popular Science*, vol. 286, 2016.
[294] M. D. Segall, *Physics of the World-Soul: Whitehead's Adventure in Cosmology*. San Francisco: Lulu Press, 2013.
[295] A. N. Whitehead, *Science and the Modern World*. NY: The Free Press, 1925.
[296] S. Krimsky and J. Gruber, *Genetic explanations: sense and nonsense*. Harvard University Press, 2012.
[297] D. J. Nicholson and J. Dupré, *Everything flows: towards a processual philosophy of biology*. Oxford University Press, 2018.
[298] E. Steinhart, "Scientific animism," in *Animism and Philosophy of Religion*. Springer, 2023, pp. 227–255.
[299] P. Kalkavage, *Timaeus*. Hackett Publishing, 2016.
[300] B. D. Lee, U. Neri, S. Roux, Y. I. Wolf, A. P. Camargo, M. Krupovic, P. Simmonds, N. Kyrpides, U. Gophna, V. V. Dolja, and E. V. Koonin, "Mining metatranscriptomes reveals a vast world of viroid-like circular RNAs," *Cell*, vol. 186, pp. 646–661.e4, 2023.
[301] C. Abergel, M. Legendre, and J. M. Claverie, "The rapidly expanding universe of giant viruses: Mimivirus, pandoravirus, pithovirus and mollivirus," *FEMS Microbiology Reviews*, vol. 39, pp. 779–796, 2015.

[302] E. Pfeifer, J. M. Sousa, M. Touchon, and E. P. Rocha, "When bacteria are phage playgrounds: interactions between viruses, cells, and mobile genetic elements," *Current Opinion in Microbiology*, vol. 70, pp. 1–17, 2022.
[303] M. Jalasvuori and E. V. Koonin, "Classification of prokaryotic genetic replicators: Between selfishness and altruism," *Annals of the New York Academy of Sciences*, vol. 1341, pp. 96–105, 2015.
[304] P. Siguier, E. Gourbeyre, and M. Chandler, "Bacterial insertion sequences: their genomic impact and diversity," *FEMS Microbiology Reviews*, vol. 38, pp. 865–891, 2014.
[305] R. Feiner, T. Argov, L. Rabinovich, N. Sigal, and I. Borovok, "A new perspective on lysogeny: prophages as active regulatory switches of bacteria," *Nature Reviews Microbiology*, vol. 13, 2015.
[306] A. H. V. Hoek, D. Mevius, B. Guerra, P. Mullany, A. P. Roberts, and H. J. Aarts, "Acquired antibiotic resistance genes: An overview," *Frontiers in Microbiology*, vol. 2, pp. 1–27, 2011.
[307] A. Dupressoir, C. Lavialle, and T. Heidmann, "From ancestral infectious retroviruses to bona fide cellular genes: Role of the captured syncytins in placentation," *Placenta*, vol. 33, pp. 663–671, 2012.
[308] S. Humphrey, A. Fillol-salom, N. Quiles-puchalt, A. F. Haag, J. Chen, J. R. Penadés, and R. Ibarra-chávez, "transduction exceeds that of classical mobile genetic elements," *Nature communications*, pp. 1–13, 2021.
[309] D. Hochhauser, A. Millman, and R. Sorek, "The defense island repertoire of the Escherichia coli pangenome," *PLoS genetics*, vol. 19, no. 4, p. e1010694, 2023.
[310] F. Safari, M. Sharifi, S. Farajnia, B. Akbari, M. K. B. Ahmadi, M. Negahdaripour, and Y. Ghasemi, "The interaction of phages and bacteria: the co-evolutionary arms race," *Critical Reviews in Biotechnology*, vol. 40, pp. 119–137, 2020.
[311] A. Pawluk, A. R. Davidson, and K. L. Maxwell, "Anti-CRISPR: Discovery, mechanism and function," *Nature Reviews Microbiology*, vol. 16, pp. 12–17, 2018.
[312] B. A. Osuna, S. Karambelkar, C. Mahendra, A. Sarbach, M. C. Johnson, S. Kilcher, and J. Bondy-Denomy, "Critical anti-CRISPR locus repression by a bi-functional Cas9 inhibitor," *Cell host & microbe*, vol. 28, pp. 23–30.e5, 2020.
[313] A. Isaev, A. Drobiazko, N. Sierro, J. Gordeeva, I. Yosef, U. Qimron, N. V. Ivanov, and K. Severinov, "Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence," *Nucleic acids research*, vol. 48, no. 10, pp. 5397–5406, 2020.
[314] F. Rousset, F. Depardieu, S. Miele, J. Dowding, A.-L. Laval, E. Lieberman, D. Garry, E. P. Rocha, A. Bernheim, and D. Bikard, "Phages and their satellites encode hotspots of antiviral systems," *Cell host & microbe*, vol. 30, no. 5, pp. 740–753, 2022.
[315] N. Burman, S. Belukhina, F. Depardieu, R. A. Wilkinson, M. Skutel, A. Santiago-Frangos, A. B. Graham, A. Livenskyi, A. Chechenina, N. Morozova *et al.*, "Viral proteins activate PARIS-mediated tRNA degradation and viral tRNAs rescue infection," *bioRxiv*, pp. 2024–01, 2024.
[316] B. Swimme, *The universe is a green dragon: A cosmic creation story*. Bear & Company, 1984.
[317] C. Darwin, *On the origin of species: A facsimile of the first edition*. Harvard University Press, 1964.
[318] A. N. Whitehead, *Process and reality: an essay in cosmology*. NY: The Free Press, 1929.

ACKNOWLEDGEMENTS

The law of the universe is such that nothing happens independently of everything else and that nobody exists in isolation. This dissertation and its writer present no exception. Stan, Chirlmin, Alicia, Christos, Anna, Rita, Marina, Jelger, Daan, Kostas, Menno, Jasper, Aswin, Lucia, Reza, Catalina, Tom, Carmen, Andrea, Karlijn, Dani, Tracey, Pippi, Desi, Maartje, Céline, Jasper, Illinka, Baltus, Halewijn, Jill, Thomas, Stefan, Rodrigo, Jan, Eline, Teunke, Becca, Jochem, Cristian, Tobal, Franklin, Boris, Seb, Luc, Ailong, Chunyi, Martin, Johannes, Koen, Joel, Ernie, Matthew, Palmyre, John, Dick, Frank, Tišma, Georgi, Sid, Rachel, Mike, Carolien, Kijun, Cecilia, Theo, Arash, Ramon, Sacha, Biera, Anke, Jan, Joyce, Moritz, Ran, Nicole, Klarinda, Arti, Miyase, Gijs, Sinan, Robert, Simon, Leif, Peter, Gilles, Gitte, Suz, Sil, Ramon, Marouen, Daphne, Stef, Thijs, Khubyb, Pedram, Chaturika, Billy, Didier, Makenna, David, Martijn, Flip, Casey, Julie, Jannemarie, Sue, Kevin, Dirk, Emma, Gertjan, Jelte Jan, Deepika, Roosmarijn, Jeanine, Sandra, Renske, Melvin, Max, Inna, Glenn, Philip, Ruben, Johnick, Hein, Joep, Vincent, Bart, Tim, Willem, Cas, Joey, Doex, Chiem, Janneke, Jesse, Cor, Mom, Dad: you were key in shaping both the pages of this book and the person I am continually becoming. To those not explicitly mentioned here, you know who you are, I know who you are. As we look towards a future that is always upon us, I look forward to our paths crossing again.

Curriculum Vitæ

Sam Pelle Bertram van Beljouw

17-03-1994 Born in Lieshout, the Netherlands.

Education

2016 – 2019 Master of Science in Biotechnology, *Cum Laude*.
Wageningen University, the Netherlands.
Thesis: Following single molecules in *Lactococcus lactis*: Intracellular detection of plasmid DNA.
Supervisor: Dr. rer. nat. J.C. Hohlbein

2014 – 2016 Propaedeutic and premaster in Philosophy.
Tilburg University, the Netherlands.

2012 – 2016 Bachelor of Science in Biomedical Engineering.
Technical University of Eindhoven, the Netherlands.

Work Experience

2024 – present Post doctoral researcher in molecular microbiology, Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, the Netherlands.
Advisor: Prof. dr. ir. S.J.J. Brouns.

2020 – 2024 Doctoral researcher in molecular microbiology, Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, the Netherlands.
Advisor: Prof. dr. ir. S.J.J. Brouns.

2019 – 2020 Intern in microbial engineering, Research & Development, Caribou Biosciences, Inc., Berkeley, USA.
Supervisor: Dr. J.D. Berry.

2018 Research assistant in single-molecule microscopy, Department of Biophysics, Wageningen University, the Netherlands.
Supervisor: Dr. rer. nat. J. Holhbein

AWARDS

2023	International Birnstiel Award, awarded by the Birnstiel Foundation and IMP (Austria).
2023	Young Investigator Award, awarded by the Henner Graeff Foundation (Italy).
2023	Best speaker award, awarded by the Winter School on Proteinases (Italy).
2022	Best speaker award, awarded by the Belgian Society for Viruses of Microbes (Belgium).
2022	Kiem Award, awarded by the Dutch Microbiology Society (the Netherlands).

About the author

Sam Pelle Bertram van Beljouw was born on March 17, 1994, in Lieshout, the Netherlands. He completed his Gymnasium in 2012 at Sint-Joriscollege, Eindhoven, while playing Eredivisie level basketball. In 2016, he earned his Bachelor of Science in Biomedical Engineering from Eindhoven University of Technology. Concurrently, he pursued studies in philosophy, completing a propaedeutic and premaster program at Tilburg University.

After travelling through South-East Asia, Sam undertook a Master of Science in Molecular and Cellular Biotechnology from Wageningen University, graduating Cum Laude in 2019. His thesis focussed on intracellular detection of plasmid DNA in *Lactococcus lactis*, under the supervision of Dr. Johannes Hohlbein. Following a 12-month internship in microbial engineering at Caribou Biosciences in Berkeley, USA, Sam returned to the Netherlands to obtain his PhD in molecular microbiology.

In 2020, he began his doctoral research at the Kavli Institute of Nanoscience, Department of Bionanoscience at the Delft University of Technology, under the guidance of Prof. Stan Brouns. In the following four years, he discovered a unique CRISPR-Cas complex that cleaves proteins, the details of which are described in this dissertation. He received various recognitions for his work, including the International Birnstiel Award, the Kiem Award, and the Young Investigator Award. After his doctoral research, he continued at the same institution as a postdoctoral researcher.

Sam is interested in exploring both the visible and invisible realms of life, especially where they intersect. His methods extend beyond those of scientific and philosophical inquiry to include surrealistic poetry and speculative cosmology. He has authored several poetry collections and is currently working on a novel.

List of Publications

Journal articles related to this dissertation:

6. K. Kalogeropoulos*, **S.P.B. van Beljouw***, S.J.J. Brouns, U. auf dem Keller. *Craspase activity in human and native proteomes.* Manuscript in preparation.

5. **S.P.B. van Beljouw**, A.C. Haagsma, K. Kalogeropoulos, M. Pabst, S.J.J. Brouns. *Craspase orthologs cleave non-conserved motifs in target protein Csx30.* ACS Chemical biology **19**, 5, 1051-1055 (2024).

4. **S.P.B. van Beljouw** & S.J.J. Brouns. *CRISPR-controlled proteases.* Biochemical Society Transactions, BST20230962 (2024).

3. **S.P.B. van Beljouw**, J. Sanders, A. Rodríguez-Molina, S.J.J. Brouns. *RNA-targeting CRISPR-Cas systems.* Nature Reviews Microbiology **21**, 21-34 (2023).

2. C. Hu*, **S.P.B. van Beljouw***, K.H. Nam, G. Schuler, F. Ding, Y. Cui, A. Rodríguez-Molina, A.C. Haagsma, M. Valk, M. Pabst, S.J.J. Brouns, A. Ke. *Craspase is a CRISPR RNA-guided, RNA-activated protease.* Science **377**, 1278-1285 (2022).

1. **S.P.B. van Beljouw**, A.C. Haagsma, A. Rodríguez-Molina, D.F. van den Berg, J.N.A. Vink, S.J.J. Brouns. *The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase.* Science **373**, 1349-1353 (2021).

Other journal articles:

4. K. Eloff*, K. Kalogeropoulos* , O. Morell, A. Mabona, J. Berg Jespersen, W. Williams, **S.P.B. van Beljouw**, M. Skwark, A. Hougaard Laustsen, S.J.J. Brouns, A. Ljungars, E. M. Schoof, J. van Goey, U. auf dem Keller, K. Beguir, N. Lopez Carranza, T. P. Jenkins. *De novo peptide sequencing with InstaNovo: Accurate, database-free peptide identification for large scale proteomics experiments.* BioRxiv (2023).

3. M. Mahler, A.R. Costa, **S.P.B. van Beljouw**, P.C. Fineran, S.J.J. Brouns. *Approaches for bacteriophage genome engineering.* Trends in Biotechnology **41**, 669-685 (2023).

2. K.J.A. Martens*, **S.P.B. van Beljouw***, S. van der Els, J.N.A. Vink, S. Baas, G.A. Vogelaar, S.J.J. Brouns, P. van Baarlen, M. Kleerebezem, J. Hohlbein. *Visualisation of dCas9 target search in vivo using an open-microscopy framework.* Nature Communications **10**, 3552 (2019).

1. **S.P.B. van Beljouw***, S. van der Els*, K.J.A. Martens*, M. Kleerebezem, P. Bron, J. Hohlbein. *Evaluating single-particle tracking by photo-activation localization microscopy in Lactococcus lactis.* Physical biology **16**, 035001 (2019).

Equal contribution indicated with an asterisk (*).

PATENTS:

2. **S.P.B. van Beljouw** & S.J.J. Brouns (2022). *gRAMP protein for modulating a target RNA*. Octrooinummer: 2028346.

1. **S.P.B. van Beljouw** & S.J.J. Brouns (2022). *gRAMP protein and TPR-CHAT protein for modulating a target mRNA or target protein*. Patent No. WO2022255865A1.

DELIVERED (CONFERENCE) PRESENTATIONS:

9. International Society for Viruses of Microbes webinar series, online, *Invited Speaker*, 2024.

8. International Birnstiel Award reception at the Research Institute of Molecular Pathology, Vienna, Austria, *Invited Speaker*, 2023.

7. CRISPR functions beyond defence conference, Berlin, Germany, *Keynote Speaker*, 2023.

6. 40th Winter School on Proteinases and Inhibitors, Tiers, Italy, *Oral Presentation*, 2023.

5. CRISPR-Cas conference, Boston, USA, *Poster Presentation*, 2022.

4. Genome Engineering Seminar Series of Harvard Medical School, online, *Invited Speaker*, 2022.

3. Belgian Society for Viruses of Microbes conference, Leuven, Belgium, *Oral and Poster Presentation*, 2022.

2. Dutch Microbiology Society Spring meeting, Arnhem, the Netherlands, *Invited Speaker*, 2022.

1. Dutch Biophysics Symposium, Veldhoven, the Netherlands, *Poster Presentation*, 2018.

It is believed that the earliest of bacteria were already subject to invasions of viruses. The ongoing interactions between bacteria and viruses have led to a vast diversity and complexity in bacterial immunity, epitomized by the well-known CRISPR-Cas systems. Canonical knowledge states that CRISPR-Cas acts as an immune system through cleavage of invader RNA or DNA. But it now turns out that not only nucleic acids are CRISPR targets.

This dissertation provides an experimental and conceptual characterization of Craspase, a CRISPR-controlled protease. The array of functionalities inherent to Craspase — including precise protein cleavage, guided RNA recognition, and self-regulatory capabilities — highlights the complexity that can emerge from the endless bacterium-virus coevolution.