Published online 9 December 2022 *Nucleic Acids Research, 2022, Vol. 50, No. 22* **12913–12923**
https://doi.org/10.1093/nar/gkac1151

Target RNA-guided protease activity in type III-E CRISPR–Cas system

Xiaoshen Wang[1,2,†], Guimei Yu[1,†], Yanan Wen[1,†], Qiyin An[3,4,†], Xuzichao Li[1], Fumeng Liao[1], Chengwei Lian[1], Kai Zhang[1], Hang Yin[1], Yong Wei[5], Zengqin Deng[3,*] and Heng Zhang[1,2,*]

[1]The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Haihe Laboratory of Cell Ecosystem, Tianjin Institute of Immunology, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China, [2]State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan, Hubei 430062, China, [3]Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Center for Biosafety Mega-Science, Chinese Academy of Sciences, No.44 Xiao Hong Shan, Wuhan, Hubei 430071, China, [4]University of Chinese Academy of Sciences, Beijing 100049, China and [5]The Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China

Received September 05, 2022; Revised November 11, 2022; Editorial Decision November 16, 2022; Accepted November 17, 2022

ABSTRACT

The type III-E CRISPR–Cas systems are newly identified adaptive immune systems in prokaryotes that use a single Cas7–11 protein to specifically cleave target RNA. Cas7–11 could associate with Csx29, a putative caspase-like protein encoded by the gene frequently found in the type III-E loci, suggesting a functional linkage between the RNase and protease activities in type III-E systems. Here, we demonstrated that target RNA recognition would stimulate the proteolytic activity of Csx29, and protein Csx30 is the endogenous substrate. More interestingly, while the cognate target RNA recognition would activate Csx29, non-cognate target RNA with the complementary 3′ anti-tag sequence inhibits the enzymatic activity. Csx30 could bind to the sigma factor RpoE, which may initiate the stress response after proteolytic cleavage. Combined with biochemical and structural studies, we have elucidated the mechanisms underlying the target RNA-guided proteolytic activity of Csx29. Our work will guide further developments leveraging this simple RNA targeting system for RNA and protein-related applications.

INTRODUCTION

Prokaryotes utilize the CRISPR–Cas (clustered regularly interspaced short palindromic repeat-CRISPR associated genes) systems to defend against invasive nucleic acids (1,2). Two major classes composed of six types of CRISPR–Cas systems have been discovered (3,4). Those in class 1 (types I, III and IV) typically encode multiple Cas effectors, while class 2 systems (types II, V and VI) contain a single Cas effector protein. The Cas effector proteins assemble with crRNA, forming the ribonucleoprotein complex to guide the sequence-specific cleavage of invading DNA or RNA (5). Recently, a new type III-E CRISPR–Cas system has been identified (4,6,7), which, distinct from the traditional class 1 type III systems, employs a single Cas7–11 effector protein with four Cas7-like domains and one Cas11-like domain. The four Cas7-like domains (Cas7.1–7.4) arranged into a helical filament encompass the crRNA. The Cas11 domain resides between Cas7.2 and Cas7.3 domains, facilitating the crRNA-target RNA duplex formation (8,9). The Cas7–11 complex performs site-specific RNA cleavage based on the 6-nt ruler mechanism, akin to the Csm/Cmr complexes in the canonical type III-A/B systems (10–13). Intriguingly, Cas7–11 could target RNA in mammalian cells with minimal off-target effect and cell toxicity, promising an RNA editing tool (7).

The canonical type III systems, such as type III-A/B, feature a Cas10 protein, which would be activated to exert DNase and/or cyclic oligoadenylate (cOA) synthetase activities post cognate target RNA binding (14–20). Cas10 protein, however, is missing in the type III-E systems. By contrast, a gene encoding a caspase-like protein (Csx29, also referred to as TPR-CHAT) is acquired by the type III-E systems (6,7). Csx29 contains a C-terminal caspase-like protease (PS) domain and an N-terminal tetratricopeptide repeat (TPR) domain. Interestingly, Csx29 could directly

*To whom correspondence should be addressed. Tel: +86 22 83336833; Email: zhangheng134@gmail.com
Correspondence may also be addressed to Zengqin Deng. Email: dengzengqin@wh.iov.cn
†The authors wish it to be known that, in their opinion, the first four authors should be regarded as Joint First Authors.

interact with Cas7–11 (6), and our recent study revealed that the cognate target RNA binding would trigger profound conformational changes in the tethered Csx29 (9), indicating a functional linkage between the RNase and the peptidase activities within the complex (also known as CRISPR-guided caspase, Craspase). Caspases belong to a family of cysteine proteases (21), the activation of which is a hallmark of programmed cell death. However, the functional mechanisms of Csx29 in the type III-E systems remain to be characterized.

In this study, we have biochemically and structurally explored the activity and substrate of Csx29. We demonstrated that binding of the cognate target RNA to the Cas7–11 complex would trigger the tethered Csx29 to cleave the endogenous protein Csx30. Detailed mutagenesis studies revealed a preference for Leu and Met residues at the P1 position. Interestingly, besides the recognition and cleavage motif, the substrate structure also appears to affect the cleavage by Csx29. Species-specific substrate specificity is also observed, which is likely to be governed by the divergent cleavable loop and the C-terminal region of the substrate. Further, we found that the 3′ end of crRNA repeat (a 2-nt tag) could function as a determinant site for Csx29 activation, distinguishing self- and non-self-RNA. These findings elucidate the functional relationship between the Cas effector and the accessory protease Csx29, and provide new insights into the mechanism of the type III-E CRISPR–Cas system.

MATERIALS AND METHODS

Protein expression and purification

The genes from *Candidatus 'Scalindua brodae'* encoding Cas7–11, Csx30 and RpoE were subcloned into the pET-derived vector containing an N-terminal His-MBP tag. The gene encoding Csx29 and synthesized crRNA were inserted into the co-transformation 13S-A vector (addgene: 48323). Csx29 was also cloned into the pET-derived vector with an N-terminal His tag. Plasmids were transformed into the BL21 (DE3) cells for induction and expression, alone or in combination. Protein expression was induced by 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cells were lysed with sonication in a lysis buffer containing 25 mM Tris–HCl pH 7.5, 500 mM NaCl, 10 mM imidazole and 2 mM β-mercaptoethanol. Cell lysates were clarified with centrifugation and then incubated with the Ni-NTA resin (QIAGEN). After extensive washing with the buffer supplemented with 20mM imidazole, the target protein was eluted using the elution buffer (25 mM Tris–HCl pH 7.5, 500 mM NaCl, 2 mM β-mercaptoethanol and 300 mM imidazole). Home-made protease Tobacco Etch Virus (TEV) was used to remove the tag if present. The Cas7–11-crRNA and Cas7–11–crRNA–Csx29 complexes and the RpoE protein were further purified using the HiTrap heparin HP column (Cytiva). Further purification of Csx30 was performed using the Ion-exchange chromatography (HiTrap Q) and the Superdex 200 column (Cytiva) in the running buffer (25 mM Tris–HCl pH 7.5, 150 mM NaCl, 2 mM DTT).

In vitro protein cleavage assays by Csx29

The Cas7–11–crRNA–Csx29 complex was mixed with target RNA in a molar ratio of 1:1.2. The candidate substrate proteins (e.g. Csx30) were then added and incubated in the cleavage buffer (25 mM Tris–HCl pH7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT and 5% glycerol) at 37°C for 30 min. The cleaved products were analyzed with 15% SDS-PAGE gel and visualized by the Coomassie blue staining.

N-terminal Edman sequencing of cleaved Csx30 by Csx29

Mixtures of the Csx30 cleavage reactions were separated by SDS-PAGE and transferred to the polyvinylidene difluoride membrane. Bands of cleaved Csx30 were collected for the N-terminal Edman sequencing (PPSQ33A, SHIMADZU). The results were analyzed using the PPSQ-30 Data Processing software.

In vitro target RNA cleavage assay

Synthesized target RNAs (Sangon Biotech) were dissolved in the cleavage buffer containing 25 mM Tris–HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT and 5% glycerol. The cleavage was initiated by mixing the SbCas7–11-crRNA complex with the target in a molar ratio of 1:2 in the cleavage buffer at 37°C. The reaction was stopped by adding 40 mM EDTA and 1 mg/ml proteinase K after 30 min, analyzed using 15% TBE–urea PAGE gel and imaged using the Azure Imaging System (Azure Biosystems).

Pull-down assay

The MBP-tagged or His-tagged pull-down assay was performed to explore the interaction between SbCas7–11 and co-occurring Csx29, Csx30 and RopE proteins in the type III-E system. The lysates containing MBP or His-tagged bait proteins and tag-free prey proteins were mixed and incubated with the amylose resin (NEB) or Ni-NTA resin (QIAGEN) for 2 h in the binding buffer (25 mM Tris–HCl pH 7.5, 150 mM NaCl, 2 mM DTT, 5% glycerol). The bound proteins were eluted with the binding buffer supplemented with 25 mM maltose (for MBP pull-down) or 250 mM imidazole (for His pull-down) after extensive washing and analyzed by SDS-PAGE.

Isothermal titration calorimetry (ITC) measurements

The ITC measurements were performed using a MicroCal PEAQ-ITC system (Malvern) at 25°C in a buffer containing 25 mM Tris–HCl pH 7.5, 150 mM NaCl. A typical titration involved 19 injections of RpoE protein (500 μM) into the cell containing 25 μM Csx30 protein. Data were analyzed with the 'one set of site' fitting model using the MicroCal PEAQ-ITC analysis software. The experiments were repeated 3 times independently with similar results, and one representative plot with the derived dissociation constant and standard error of fitting was shown.

Cryo-EM sample preparation

Purified catalytic dead SbCas7–11–crRNA binary complex was incubated with the target RNA containing the complementary 3′ anti-tag at a molar ratio of 1:2 on ice for 30 min. The mixture was further purified with size-exclusion chromatography (Superdex 200) and the peak fractions were collected and concentrated to ∼ 1 mg/ml for cryo-EM sample freezing. Purified SbCas7–11–crRNA–Csx29 complex was incubated with the NTR at a molar ratio of 1:3 on ice for 30 min. Aliquots of 3.5 μl of the samples were applied to glow-discharged holey carbon grids (Au 200 mesh R1.2/1.3, Quantifoil) and incubated for 20 s. The grids were then blotted for 2 s, plunge-frozen in liquid ethane using Vitrobot Mark IV (FEI Thermo Fisher) at 100% humidity and 4°C and stored in liquid N2 till data collection.

Cryo-EM data collection, image processing and model building

A CRYO ARM 300 electron microscope (JEOL, Japan) operating at 300 kV equipped with a K3 direct electron detector (Gatan, USA) was used. Data was collected with serialEM (22) at a nominal magnification of 50 000× with a super-resolution pixel size of 0.475 Å/pixel, a frame rate of 40 frames per second, a total accumulated dose of 40 e/Å^2 and a defocus range of −0.5 to −2.5 μm.

Image processing was performed with cryoSPARC v3.3 (23). Patch-based movie alignment and CTF estimation were carried out. Particles were automatically picked using the Topaz picker (24), extracted and then subjected to 2D classification. For the SbCas7–11-crRNA-NTR dataset, 954,132 particles in good 2D classes were selected and passed for ab initio model building and heterogenous refinement requesting eight classes. One good class containing 442,994 particles was selected for homogenous refinement, CTF refinement and the non-uniform refinement (NU-refinement), which yielded a final reconstruction of 2.8 Å.

For the SbCas7–11–crRNA–NTR–Csx29 dataset, 420,156 particles were selected for 3D classification requesting two classes, identifying one class representing NTR RNA-bound state. One good class containing 158 275 particles was selected for further CTF refinement and NU-refinement, yielding a final reconstruction of 3.09 Å.

The final map was first sharpened using Phenix Autosharpen tool (25–27). Our recent SbCas7–11 ternary complex structure was used as the initial searching template (9), which was adjusted and refined in Coot (28). Real-space refinement was then performed in Phenix using *phenix.real_space_refine* (29). The quality of the model was analyzed with MolProbity in Phenix (30). Refinement statistics are summarized in Supplementary Table S1.

RESULTS

Target RNA-dependent proteolytic cleavage of Csx30 by Csx29

The recent studies suggest that *Candidatus 'Scalindua brodae'* Cas7–11 (SbCas7–11) would recruit the ancillary protease Csx29 regardless of the presence or absence of target RNA (6) (Figure 1A). Intriguingly, we found that the target RNA recognition could trigger profound conformational changes of Csx29, thereby opening the potential substrate binding channel of Csx29 (9). To investigate whether the conformational switch is coupled with the protease activity of Csx29, we sought to identify the substrate. Given that several accessory genes, including *rpoE*, *csx30* and *csx31*, co-occur with the type III-E CRISPR–Cas module (4,6,7) (Figure 1A), we reasoned that these gene products might be the possible substrates for Csx29. We expressed and purified these gene products and assessed the proteolytic activity of Csx29 using these proteins. Remarkably, the Csx30 protein would be cleaved in the presence of cognate target RNA (CTR) with the non-complementary 3′ anti-tag sequence (Figure 1B). In contrast, we did not observe the cleaved products when the putative catalytic residues (His585 and Cys627) in Csx29 were mutated (Figure 1B), suggesting that the proteolytic activity is ascribed to Csx29. Furthermore, mutations that disrupt the interactions of Csx29 with Cas7–11 were unable to cut Csx30 (9) (Supplementary Figure S1A). Notably, Csx29 alone had no apparent enzymatic activity (Figure 1B), indicative of an inactivated status. Another purified accessory protein RpoE is resistant to cleavage by Csx29 (Supplementary Figure S1B). Together, these results echo our earlier assertion (9) and further experimentally demonstrate that the proteolytic activity of Csx29 depends on target RNA recognition.

The CTR binding has been shown to initiate the cOA synthesis and/or DNase activities of Cas10 in canonical type III systems, which are switched off after cleavage (11,12,31). Similarly, the Csx29 activity was greatly enhanced in the presence of the catalytically inactive Cas7–11 (D547A/D698A) compared with the wild-type (Figure 1C), suggesting that the dissociation of target RNA from Cas7–11 after cleavage would deactivate Csx29. The target RNA cleavage might function as a timer to regulate the protease activity, which is reminiscent of that in the type III-A system (32). Cas7.4 contains an inserted IPD domain, which could accommodate the crRNA or the crRNA-target RNA duplex (9). The IPD domain, however, is not essential for target RNA cleavage, and no direct interaction is observed between IPD and Csx29 (9). Not surprisingly, the deletion of IPD did not affect the Csx29 activity (Figure 1D). Consistently, the shortened target RNA, without the potential to insert into IPD, could also activate the Csx29 (Supplementary Figure S1C), further supporting the notion that the IPD region is dispensable for Csx29 activation.

Cleavage specificity of Csx29

We next mapped the proteolytic cleavage site on Csx30 by Edman sequencing. The results indicated that Csx29 cut Csx30 between Leu407 (P1) and Asp408 (P1') (Figure 2A, Supplementary S2A-B), within a long flexible loop in the predicted structure (Supplementary Figure S2C). The caspases and caspase-like proteases (such as Separase) usually recognize a favorable tetrapeptide motif in substrates for cleavage (33,34). We, therefore, examined whether there is a favorable recognition motif for Csx29. Alanine substitution mutations were made in the cleavable loop of Csx30. As expected, Csx29 failed to degrade the L407A mutant protein

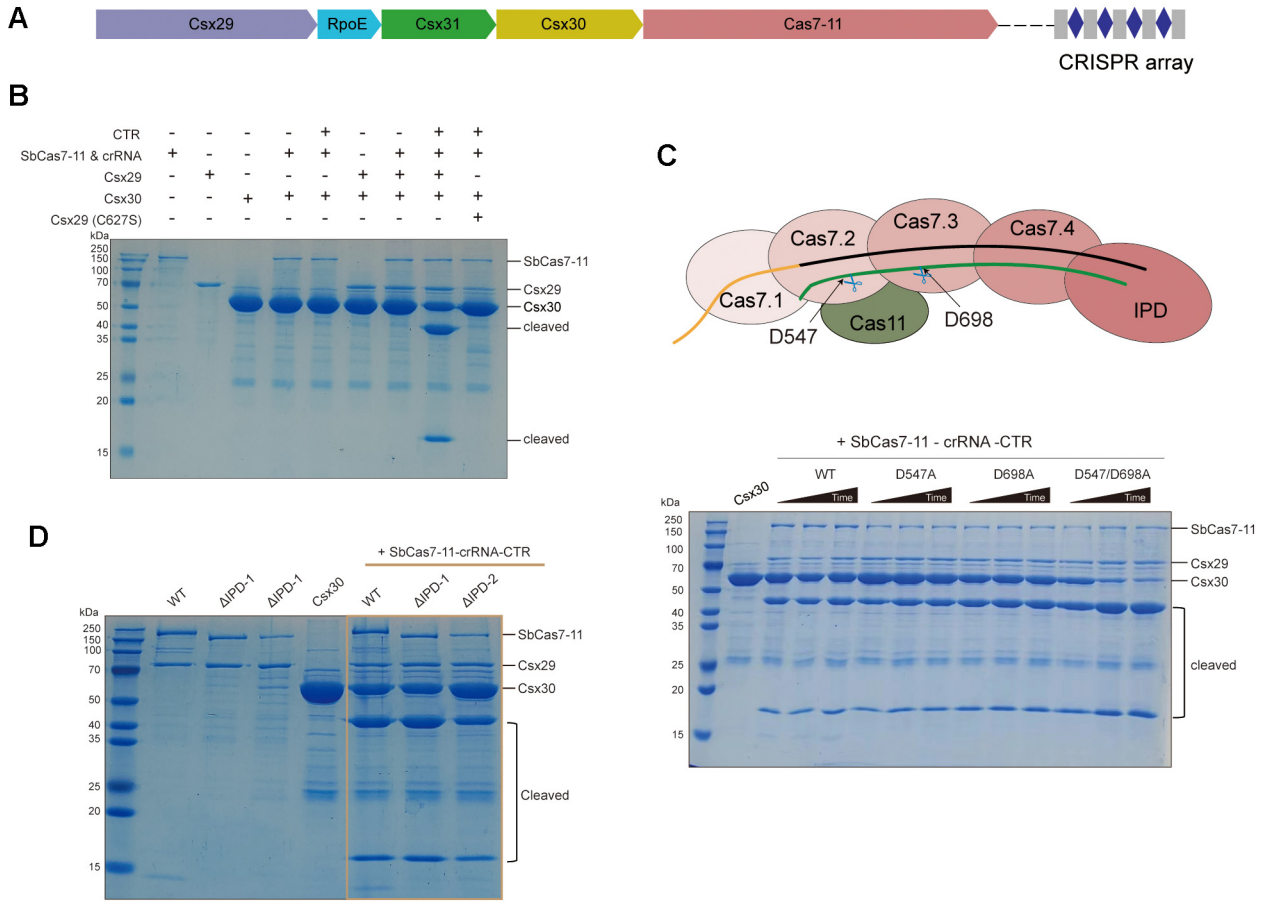

Figure 1. (**A**) Schematic representation of the type III-E CRISPR–Cas locus in *Candidatus 'Scalindua brodae'*. (**B**) *In vitro* cleavage of Csx30 by Csx29 in the presence or absence of functional Cas7–11 complex. Purified Csx30 protein was incubated and treated with indicated SbCas7–11 and Csx29. CTR, the cognate target RNA. Three replicate cleavage experiments were conducted. (**C**) In vitro cleavage of Csx30 by Csx29 in the presence of WT or catalytically inactive Cas7–11. Schematic diagram of Cas7–11-crRNA-target RNA complex (upper panel). Time courses represent 5, 10 and 15 min. This gel is representative of three replicate experiments. (**D**) The IPD domain is dispensable for Csx29 activation. ΔIPD-1: The IPD segment (aa 1032–1387) was deleted in SbCas7–11. ΔIPD-2: The IPD fragment (aa 1038–1383) was deleted in SbCas7–11. This gel is representative of three replicate experiments.

(Figure 2B). Moreover, the P2 (Asp406) and P3 (Ile405) positions of Csx30 are also important for cleavage (Figure 2B). Interestingly, differing from caspases, the residues at the P6 (Leu402) and P7 (Ser401) positions are also likely involved in the substrate recognition and degradation by Csx29 (Figure 2B). To further investigate the preference at P1 site, Leu407 in Csx30 was mutated to various amino acids. Substitutions of Leu407 with polar and charged amino acids were resistant to truncation by Csx29 (Figure 2C). Similarly, substitution to hydrophobic amino acids, including L407I and L407V, also rendered Csx30 insusceptible to Csx29-mediated cleavage. However, Csx29 could efficiently cut the L407M variant comparable to WT (Figure 2C), suggesting that both the hydrophobicity and chemical structure determine the P1 preference.

Target RNA-guided cleavage of a specific protein is a compelling concept for applications. We next explored whether this type III-E system could be extended to target other proteins with the same cleavage motif. The chimeric proteins were generated by fusion of the truncated Csx30 containing the cleavable linker with an anti-CRISPR protein AcrIIA17 (35,36) (Figure 2D). However, we did not observe any proteolytic fragments truncated by Csx29 (Figure 2D). Together, these results indicate that Csx29 may cleave the substrate with both sequence and structure specificity.

Next, we also evaluated the cross-species activity of Csx29 using Cas7–11 from *Desulfonema ishimotonii* (DiCas7–11) (7,8). The assembled DiCas7–11-crRNA-CTR-DiCsx29 complex was incapable of proteolytic degradation of SbCsx30 (Figure 2E). Similarly, the DiCsx30 was also resistant to degradation by SbCsx29. During the preparation of this manuscript, the cleavage site on DiCsx30 was also identified (37). Interestingly, the SbCsx30-chimera with the cleavable fragment replaced with the counterpart in DiCsx30 (aa 424–429) was resistant to cleavage by either DiCsx29 or SbCsx29 (Figure 2E), implying a species-specific recognition mechanism. The sequence alignment revealed that the $Csx30^{NTD}$ is well-conserved across orthologs (Supplementary Figure S3A), whereas the cleavable loop region and $Csx30^{CTD}$ vary remarkably in length and sequence. Consistently, the exposed residues around the catalytic residue in Csx29 are also not well-conserved (Supplementary Figure S3B).

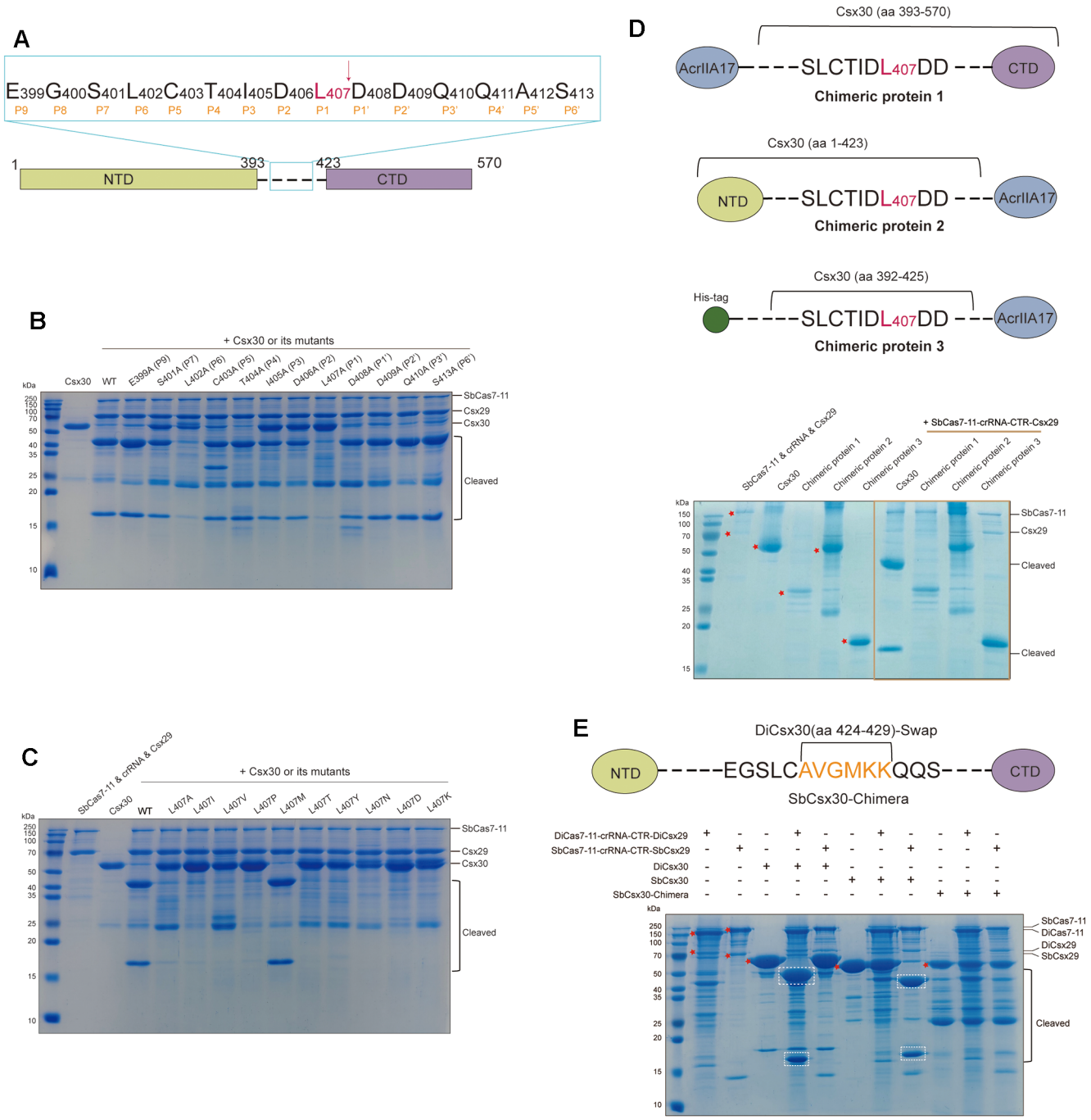

Figure 2. (**A**) Diagram of the Csx29 cleavage site on Csx30 (upper). The lower illustrates fragments of Csx30 used in Figure 2D. (**B, C**) *In vitro* Csx29-mediated substrate cleavage assays using WT and mutant Csx30 proteins. These gels are representative of three replicate experiments. (**D**) *In vitro* Csx29-mediated cleavage assay using different substrates. Schematic diagram of the chimeric substrate containing a cleavable loop of Csx30 (upper). This gel is representative of three replicate experiments. (**E**) Cross-species protease activity analysis of Csx29. This gel is representative of three replicate experiments. SbCsx30-Chimera: the cleavable fragment of SbCsx30 is replaced with the counterpart in DiCsx30 (upper panel).

Collectively, it is plausible that the divergent cleavable loop and CTD region of Csx30 may dictate the species-specific recognition and cleavage. During the revision of the manuscript, another group reported that the divergent DiCsx30^{CTD} directly associates with Csx29, explaining the species-specific observation in our study (38). Kato et al. reported that the DiCsx29 cleaves DiCsx30 after Met427, and other sites of DiCsx30 were found not to be crucial for the cleavage by DiCsx29 (37), highlighting the diversity of the type III-E systems.

Cryo-EM structure of the non-cognate target RNA-bound ternary complex

In canonical type III-A/B systems, both the cognate and non-cognate target RNA could be degraded by Csm/Cmr

complexes (39,40). Likewise, Cas7–11 showed comparable cleavage activities towards the non-cognate target RNA (NTR) with the complementary 3′ anti-tag sequence and the CTR (Figure 3A). In canonical type III-A/B systems, the exposed crRNA repeat (5′ tag) could base pair with NTR to distinguish self- from non-self RNA. Distinct from the type III-A/B systems, most of the crRNA repeat region is wrapped by the Cas7.1 and Cas7.2 domains in the type III-E system, sterically inaccessible for base pairing with target RNA (8,9). To determine if a similar 5′ tag is present in the type III-E system, we reconstituted the SbCas7–11-crRNA-NTR ternary complex and determined its cryo-EM structure at a 2.8 Å resolution (Figure 3B, Supplementary Figure S4A–D) (Supplementary Table S1). The overall structure is nearly identical to that of the CTR-bound complex, with an overall RMSD of 0.967 Å (Supplementary Figure S4E). The most structural variability between these two complexes was found in the bridging loop region between Cas7.1 and Cas11 domains, which is more disordered in the NTR-bound structure (Supplementary Figure S4E). Importantly, as revealed by the structure, the first two nucleotides of the repeat at the 3′-end (named as 2-nt tag hereinafter) base pairs with NTR (Figure 3B), presumably contributing to the self- and non-self discrimination in the type III-E systems.

The non-cognate target RNA deactivates the Csx29 activity

The DNase and cOA synthesis activity ascribed to Cas10 protein could be switched on by CTR but deactivated by NTR in type III-A/B systems (14–18). It is tempting to assume that a similar scenario might be extrapolated to the type III-E system. Indeed, the enzymatic activity of Csx29 was abrogated in the presence of NTR (Figure 3C). In contrast, a single mismatch with the 2-nt tag was well-tolerated (Figure 3D), as evidenced by the proteolytic cleavage of Csx30, indicating that the full base pairing of target RNA with the 2-nt tag is required to suppress Csx29.

Our recent Cas7–11–crRNA–CTR–Csx29 structure reveals that the four nucleotides in the non-complementary 3′ anti-tag sequence protrude into a binding pocket created mainly by the TPR domain of Csx29, possibly triggering the Csx29 activation (9) (Supplementary Figure S5A). As expected, the target RNA without the 3′ anti-tag sequence failed to trigger the hydrolysis of Csx30 by Csx29 (Figure 3D), suggesting that the non-complementary 3′ anti-tag is required for activation. The α16 helix from the binding pocket of TPR domain would undergo substantial conformational rearrangement when binding the CTR, indicating a target RNA sensor role of this helix in activating Csx29. To validate our hypothesis, we mutated the residues located in the binding pocket. As anticipated, alanine mutations of these residues, such as the Y360A in the α16 helix, rendered Csx29 ineffective in the hydrolysis of Csx30 (Supplementary Figure S5B), demonstrating the importance of the non-complementary 3′ anti-tag sequence in Csx29 activation.

Cryo-EM structure of SbCas7–11–crRNA–NTR–Csx29 complex

We next reconstituted the SbCas7–11–crRNA–NTR–Csx29 complex and determined its cryo-EM structure at a 3.1 Å resolution (Supplementary Table S1) (Figure 4A, Supplementary Figure S6A–D). In the present structure, the 2-nt tag of repeat base pairs with the complementary 3′ anti-tag (Figure 4A), whereas we could not observe cryo-EM density for the succeeding nucleotides in the 3′ anti-tag. The CTR binding would induce substantial conformational movements of Csx29 and stimulate the protease activity. However, Csx29 in the NTR-bound quaternary structure is similar to that in the target-free complex structure (Figure 4B, C), indicative of an inactivated state. The bridging loop is mainly responsible for the interaction with the TPR domain in both the target RNA-free and CTR-bound structures. Notably, the bridging loop is disordered in the present structure (Figure 4A), reminiscent of that in the SbCas7–11–crRNA–NTR complex structure (Supplementary Figure S4E). Thus, it seems that the base pairing with the 2-nt tag would trigger the conformational dynamics of the bridging loop. As a result, Cas7–11 makes fewer contacts with TPR domain relative to those in both the target RNA-free and CTR-bound structures (Figure 4A), probably accounting for the TPR movement away from the Cas7–11 domain in the present structure (Figure 4B). The –2'C in the CTR stacks against Tyr360 of TPR domain (Supplementary Figure S5A), which is essential for activation of TPR domain (Supplementary Figure S5B). Nonetheless, the –2'A in the NTR forms a Watson–Crick base pair with the 2-nt tag and fails to engage with Tyr360 (Figure 4A). It appears that the base pairing between the 3′ anti-tag and the 2-nt tag of crRNA would constrain the positioning of the 3′ anti-tag, thereby affecting the conformational movement of Csx29, which is ultimately coupled to the enzymatic activity. Consistently, while our manuscript was in preparation, two independent structural studies of the Cas7–11–crRNA–NTR–Csx29 complex have been posted or published (41,42), revealing that the Csx29 displays a configuration distinct from that in CTR-bound complex, but similar to the inactivated one in Cas7–11–crRNA–Csx29 complex.

Csx30 form a complex with RpoE

Besides Csx30, the gene encoding the sigma factor RpoE also frequently occurs within the type III-E CRISPR–Cas loci (6,7) (Figure 1A). We performed the pull-down assay to examine the potential interactions between RpoE and other components of the type III-E system. RpoE could only associate efficiently with Csx30 protein in our experimental condition (Figure 5A, Supplementary Figure S7A). Our ITC data further demonstrated the tight binding between RpoE and Csx30 with an affinity of ∼0.2 μM and a 1:1 stoichiometry (Figure 5B). RpoE makes extensive interactions with the $Csx30^{NTD}$ region in the complex structure predicted by AlphaFold2 with high confidence (43,44), while no direct interaction could be seen between RpoE and the cleavable loop or $Csx30^{CTD}$ (Supplementary Figure S7B). The two separated helical subdomains of RpoE, together with a linker loop, wrap around $Csx30^{NTD}$. Mutations at the predicted binding interfaces compromised the interaction between Csx30 and RpoE (Supplementary Figure S7C, D), validating the structural model of the Csx30-RpoE complex. In line with the structural model, Csx29 was also able to cleave Csx30 in the presence of RpoE (Figure 5C). Fur-

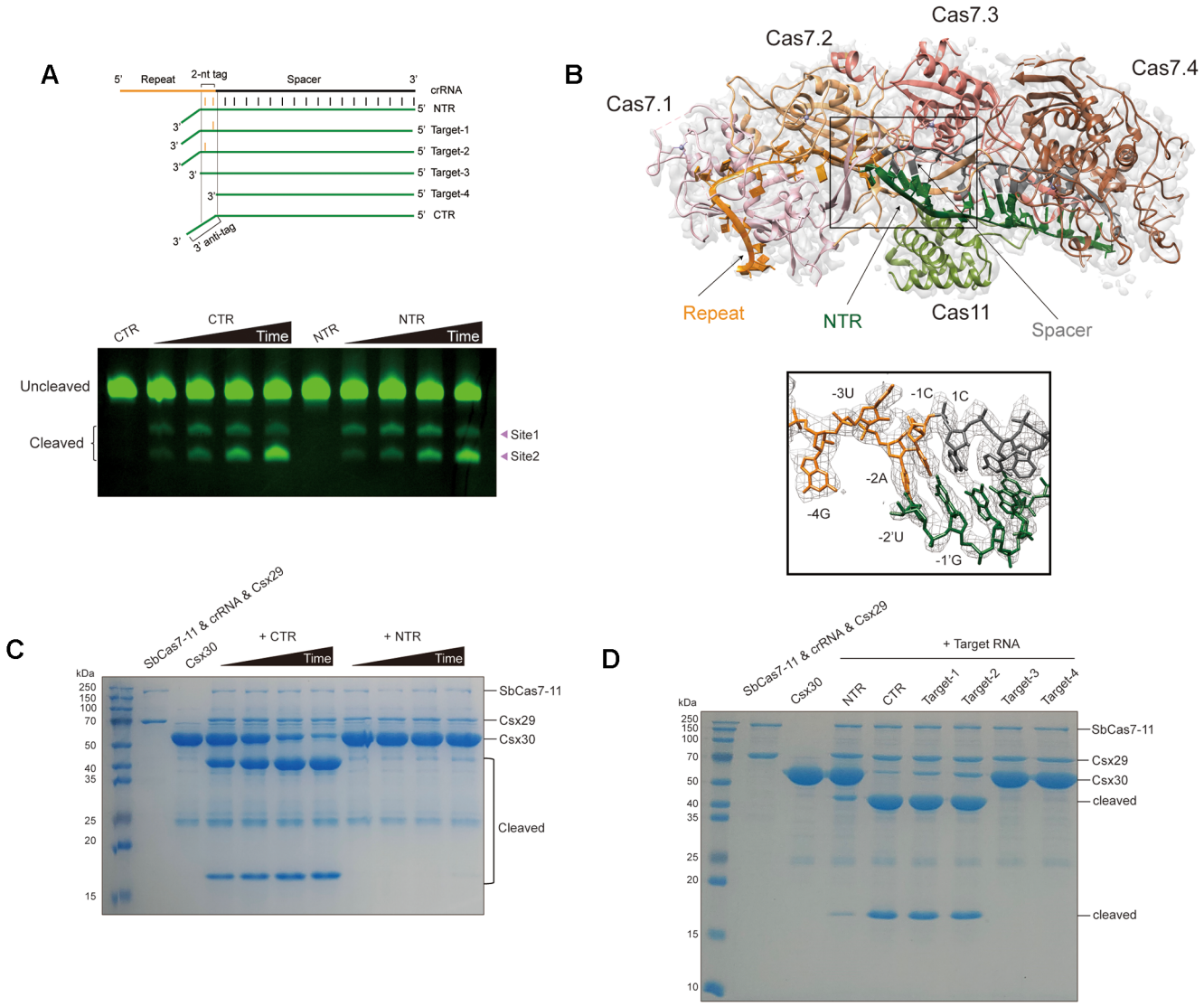

Figure 3. (**A**) (Top) Schematic representation of crRNA and target RNA used in the study. (Bottom) *In vitro* target RNA cleavage assay using CTR and NTR. Time courses represent 5, 15, 30 and 60 min. This gel is representative of three replicate experiments. (**B**) The Cryo-EM structure of the SbCas7–11–crRNA–NTR ternary complex (upper). The four Cas7-like domains, Cas7.1–Cas7.4, are colored in light pink, wheat, warm pink and brown, respectively. Cas11 domain is colored in green. Close-up view of the base pairing between the NTR and the repeat of the crRNA (lower). (**C**) *In vitro* cleavage assay of Csx30 by Csx29 in the presence of NTR or CTR. Time courses represent 5, 15, 30 and 60 min. This gel is representative of three replicate experiments. (**D**) *In vitro* proteolytic cleavage assay of Csx30 by Csx29 with different target RNA. This gel is representative of three replicate experiments.

thermore, the Csx30^{NTD} rather than Csx30^{CTD} could associate with RpoE (Figure 5D). The Csx30^{NTD} and Csx30^{CTD} would be readily separated after cleavage (Figure 5D), and no direct physical association between the Csx30^{NTD}-RpoE binary complex and Csx30^{CTD} was observed (Figure 5D).

DISCUSSION

The recently identified type III-E system is particularly interesting because it combines the features of both class 1 and class 2 CRISPR systems (4,6,7). Here, we biochemically and structurally characterized another attractive property of the type III-E CRISPR system: the RNA-guided protease activity. The cognate target RNA binding would initiate the conformational changes of Csx29, thereby turning on its proteolytic activity. These enzymatic cascade reactions are reminiscent of the target RNA-activated Cas10 in type III-A/B systems. During the preparation of this manuscript, three independent studies about the type III-E systems have been posted or published, and our results are in line with their observations (37,41,42). Notably, such RNA-guided protease activity could potentially be repurposed for targeted protein degradation and protein-based diagnostic applications. We found that the flexible IPD region protruding from Cas7.4 is dispensable for both target RNA cleavage and Csx29 activation (Figure 1D), which would facilitate the rational design of a smaller version of Cas7–11 for delivery and applications. Interestingly, Liu *et al.*, reported that the IPD is indispensable for the RNase activity of SbCas7–11 (42), which

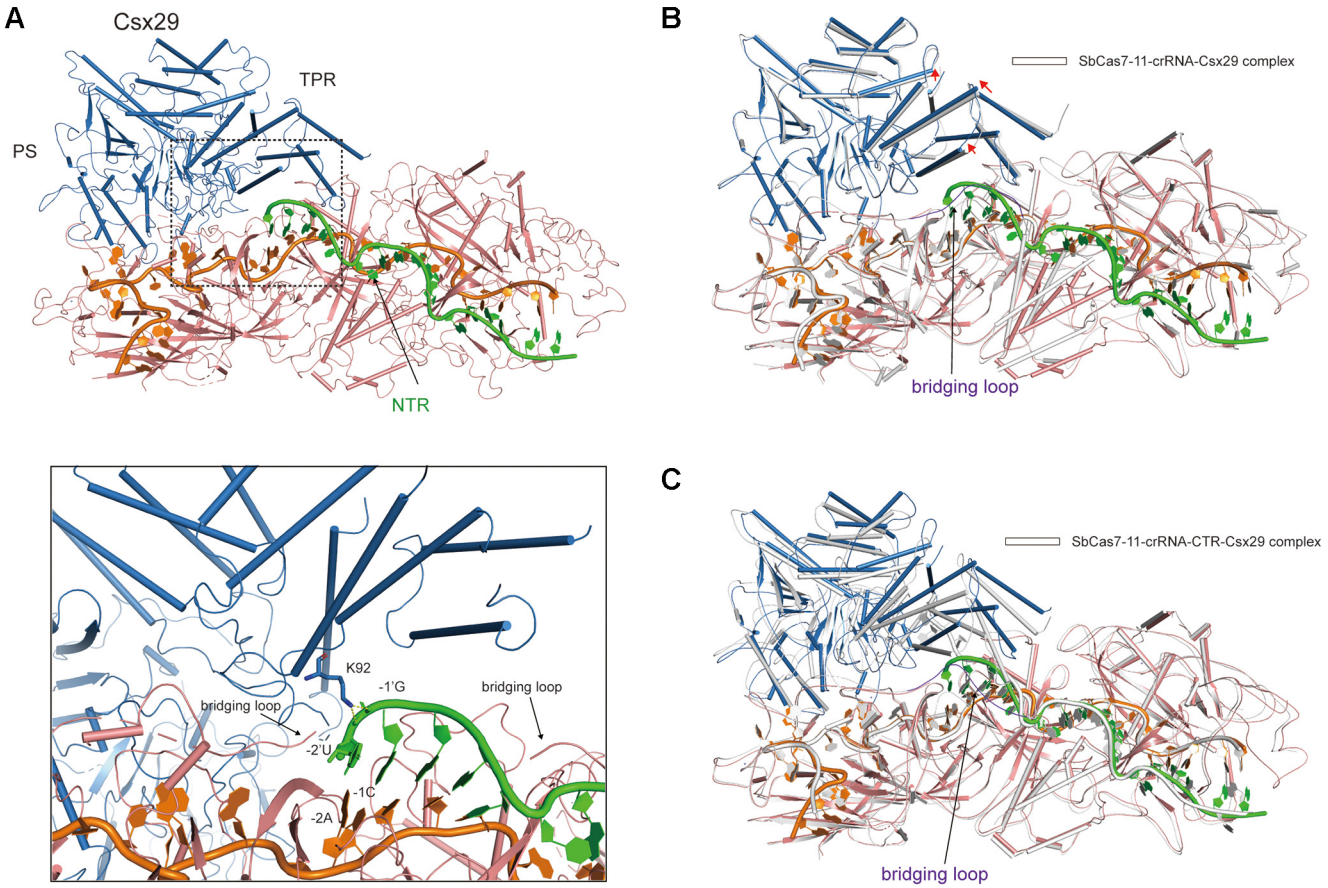

Figure 4. (**A**) The Cryo-EM structure of the SbCas7–11–crRNA–NTR–Csx29 complex (upper). PS: protease domain in Csx29. TPR: tetratricopeptide repeat domain in Csx29. Close-up view of the base pairing involving the 2-nt tag (lower). The Lys92 in TPR domain makes polar contacts with the target RNA. (**B**) Superimposition of SbCas7–11–crRNA–NTR–Csx29 complex and SbCas7–11–crRNA–Csx29 complex (white, PDB: 7X8A). The bridging loop (purple) is ordered in SbCas7–11–crRNA–Csx29 complex, and indicated by an arrow. The slight outward movements of TPR domain in the NTR-bound complex are indicated by red arrows. (**C**) Structural alignment of NTR- and CTR-bound (white, PDB: 7XC7) quaternary complexes. The bridging loop (purple) is ordered in the CTR-bound complex, and indicated by an arrow.

may be related to the linker selection or the truncation length.

The crRNAs of type III-A/B systems contain an 8-nt 5′ tag, the complementarity to which determines whether the target RNA could activate Cas10 (11–13,31). Despite the long repeat (∼15–28 nt) in the type III-E system (6–8), we found that only the 2-nt 3′-end of the repeat is available for base pairing and serves as a determinant site for Csx29 activation. While the CTR could stimulate the enzymatic activity of Csx29, the NTR with complementary nucleotides to the 2-nt tag fails to activate Csx29.

Csx29 seems to be a multiple-turnover enzyme in our system. Caspase and caspase-like proteases are known to have a strong specificity for charged residues at the P1 position (21,34), whereas we found SbCsx29 prefers hydrophobic leucine and methionine residues at the P1 position. This result agrees well with the finding in DiCsx29, which cleaves DiCsx30 after a methionine residue (37). Of note, both sequence and structure specificities of the substrate are potentially required for cleavage by Csx29, as only the full-length Csx30 could be degraded by Csx29 (Figure 2D). In line with our observations, Hu et al. reported the cleavage of an artificial substrate by Csx29, but with low cleavage activity relative to the Csx30 protein (41). However, the structure-dependent recognition would restrict the potential application of the RNA-guided protease activity. Further studies on the exploitation and design of only amino acid sequence-dependent recognition and cleavage would be necessary to repurpose this system for applications.

The sigma factor RpoE is thought to modulate the transcription of specialized regulons in response to the stress signals in *Escherichia coli* and other bacteria (45,46). Anti-sigma factor RseA negatively regulates RpoE through physical association (47,48), which would be released to direct the RNA polymerase core enzyme for transcription after consecutive cleavage of the anti-sigma factor in *E. coli* upon stimulation (46). The working model of RpoE in the type III-E system might be analogous, in some ways, to its ortholog in *E. coli*: Csx30 may play a role similar to the anti-sigma factor. The released $Csx30^{NTD}$-RpoE complex may be further processed to fully liberate RpoE to initiate a cascade of stress responses (Figure 5E). Indeed, Feng Zhang's group experimentally validated that Csx30 could coordinate the transcriptional activity of RpoE, which likely modulates other immune responses, such as spacer acquisition and cell homeostasis (38). The proteases required for sub-

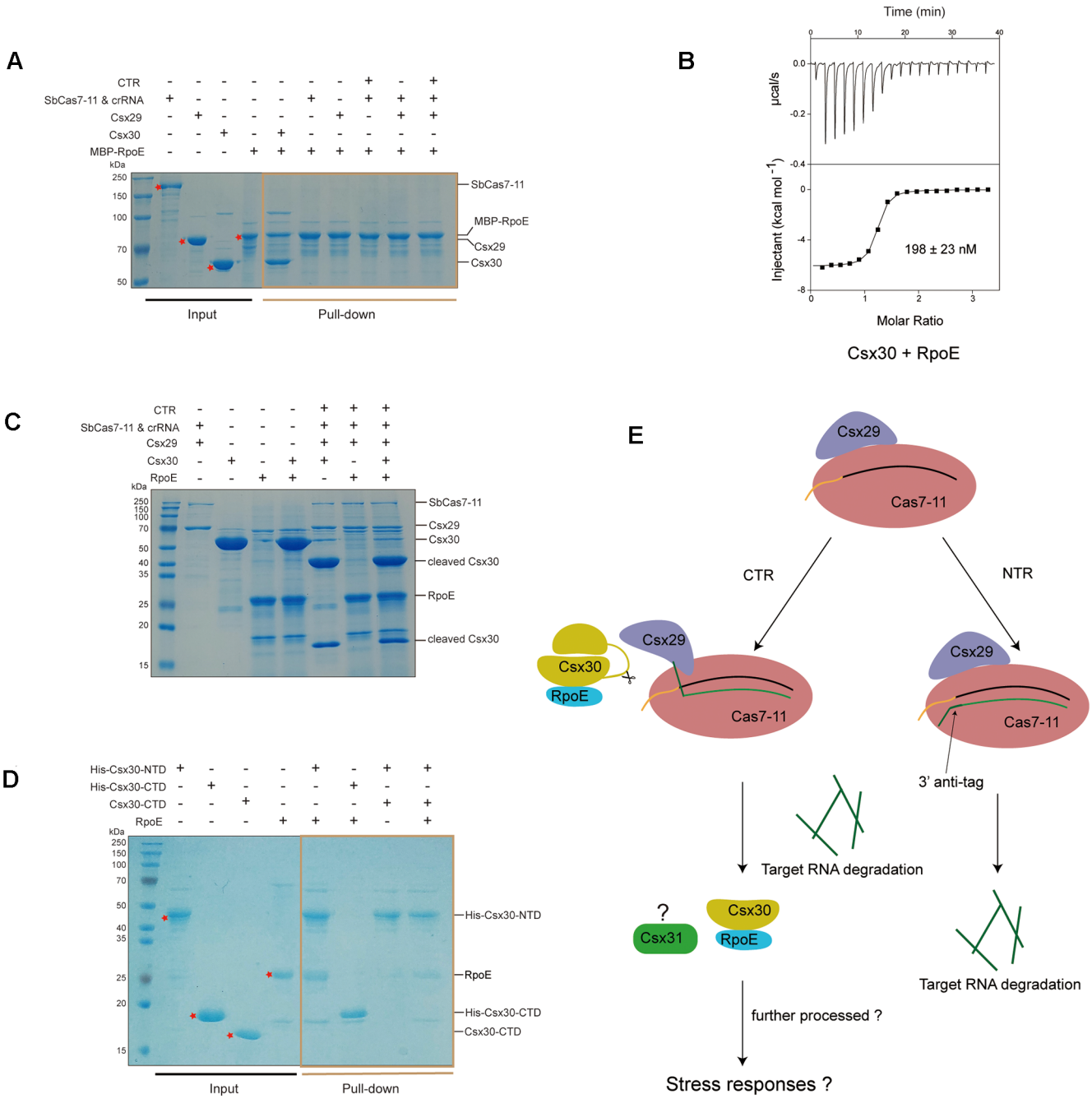

Figure 5. (**A**) *In vitro* pull-down assay of RpoE with SbCas7–11, Csx29 and Csx30. This gel is representative of three replicate experiments. (**B**) Isothermal titration calorimetric (ITC) analysis of the binding between RpoE and Csx30. (**C**) RpoE has no obvious effect on the Csx29-mediated cleavage of Csx30. This gel is representative of three replicate experiments. (**D**) *In vitro* pull-down assay of RpoE by different His-tagged Csx30 proteins. The tag-free $Csx30^{CTD}$ failed to associate with His-tagged $Csx30^{NTD}$. (**E**) The proposed working model of the target RNA-guided proteolytic cleavage of Csx30in the type III-E system.

sequent degradation of $Csx30^{NTD}$ would be identified to elucidate the role of RpoE in the type III system. Alternatively, some interactors may compete with RpoE for binding to $Csx30^{NTD}$, thereby leading to the dissociation of RpoE from $Csx30^{NTD}$. There is another ancillary protein Csx31 found in the type III loci apart from Csx30 and RpoE (6,7) (Figure 1A). However, we could not get the soluble Csx31 protein with homogeneity. Interestingly, two independent studies recently reported that Csx31 could form a ternary complex with $Csx30^{NTD}$ and RpoE, but with controversial points regarding the role of Csx31 (37,42). Therefore, future studies in the natural context may be required to understand the biological function of Csx31.

Interestingly, some signal transduction proteins or systems that sense environmental conditions have been found to colocalize with the type III-E system, such as the two-

component regulatory system (49). For example, the *ntrC* gene neighboring DiCas7–11 encodes an effector of a two-component system, and its ortholog has been demonstrated to interact with the sigma factor to activate the gene transcription (50). These proteins or systems may work in concert with the type III-E system to fight against foreign mobile genetic elements.

DATA AVAILABILITY

The atomic coordinates and the EM map have been deposited to the Protein Data Bank under the accession code 8GNA and 8GU6, and to the Electron Microscopy Data Bank (EMD-34158, EMD-34270).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Center for Instrumental Analysis and Metrology of Wuhan Institute of Virology for supporting cryo-EM data acquisition.

FUNDING

National Key R&D Program of China [2022YFC3400402 to Y.W.]; National Natural Science Foundation of China [32200496 to G.Y., 32071218 to H.Z.]; CAS Pioneer Hundred Talents Program (to Z.D.); Open Project Funding of the State Key Laboratory of Biocatalysis and Enzyme Engineering [SKLBEE2021004]. Funding for open access charge: National Key R&D Program of China [2022YFC3400402 to Y.W.]; National Natural Science Foundation of China [32200496 to G.Y., 32071218 to H.Z.]; CAS Pioneer Hundred Talents Program (to Z.D.); Open Project Funding of the State Key Laboratory of Biocatalysis and Enzyme Engineering [SKLBEE2021004].
Conflict of interest statement. None declared.

REFERENCES

1. Koonin,E.V., Makarova,K.S. and Wolf,Y.I. (2017) Evolutionary genomics of defense systems in archaea and bacteria. *Annu. Rev. Microbiol.*, **71**, 233–261.
2. Mohanraju,P., Makarova,K.S., Zetsche,B., Zhang,F., Koonin,E.V. and Van der Oost,J. (2016) Diverse evolutionary roots and mechanistic variations of the CRISPR–Cas systems. *Science*, **353**, aad5147.
3. Shmakov,S., Smargon,A., Scott,D., Cox,D., Pyzocha,N., Yan,W., Abudayyeh,O.O., Gootenberg,J.S., Makarova,K.S., Wolf,Y.I. *et al.* (2017) Diversity and evolution of class 2 CRISPR–Cas systems. *Nat. Rev. Microbiol.*, **15**, 169–182.
4. Makarova,K.S., Wolf,Y.I., Iranzo,J., Shmakov,S.A., Alkhnbashi,O.S., Brouns,S.J., Charpentier,E., Cheng,D., Haft,D.H., Horvath,P. *et al.* (2020) Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.*, **18**, 67–83.
5. Barrangou,R. and Marraffini,L.A. (2014) CRISPR–Cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell*, **54**, 234–244.
6. van Beljouw,S.P., Haagsma,A.C., Rodríguez-Molina,A., van den Berg,D.F., Vink,J.N. and Brouns,S.J. (2021) The gRAMP CRISPR–Cas effector is an RNA endonuclease complexed with a caspase-like peptidase. *Science*, **373**, 1349–1353.
7. Özcan,A., Krajeski,R., Ioannidi,E., Lee,B., Gardner,A., Makarova,K.S., Koonin,E.V., Abudayyeh,O.O. and Gootenberg,J.S. (2021) Programmable RNA targeting with the single-protein CRISPR effector cas7–11. *Nature*, **597**, 720–725.
8. Kato,K., Zhou,W., Okazaki,S., Isayama,Y., Nishizawa,T., Gootenberg,J.S., Abudayyeh,O.O. and Nishimasu,H. (2022) Structure and engineering of the type III-E CRISPR–Cas7–11 effector complex. *Cell*, **185**, 2324–2337.
9. Yu,G., Wang,X., Zhang,Y., An,Q., Wen,Y., Li,X., Yin,H., Deng,Z. and Zhang,H. (2022) Structure and function of a bacterial type III-E CRISPR–Cas7–11 complex. *Nat. Microbiol.*, **7**, 2078–2088.
10. Osawa,T., Inanaga,H., Sato,C. and Numata,T. (2015) Crystal structure of the CRISPR–Cas RNA silencing cmr complex bound to a target analog. *Mol. Cell*, **58**, 418–430.
11. You,L., Ma,J., Wang,J., Artamonova,D., Wang,M., Liu,L., Xiang,H., Severinov,K., Zhang,X. and Wang,Y. (2019) Structure studies of the CRISPR-Csm complex reveal mechanism of co-transcriptional interference. *Cell*, **176**, 239–253.
12. Jia,N., Mo,C.Y., Wang,C., Eng,E.T., Marraffini,L.A. and Patel,D.J. (2019) Type III-A CRISPR–Cas csm complexes: assembly, periodic RNA cleavage, DNase activity regulation, and autoimmunity. *Mol. Cell*, **73**, 264–277.
13. Guo,M., Zhang,K., Zhu,Y., Pintilie,G.D., Guan,X., Li,S., Schmid,M.F., Ma,Z., Chiu,W. and Huang,Z. (2019) Coupling of ssRNA cleavage with DNase activity in type III-A CRISPR-Csm revealed by cryo-EM and biochemistry. *Cell Res.*, **29**, 305–312.
14. Kazlauskiene,M., Tamulaitis,G., Kostiuk,G., Venclovas,Č. and Siksnys,V. (2016) Spatiotemporal control of type III-A CRISPR–Cas immunity: coupling DNA degradation with the target RNA recognition. *Mol. Cell*, **62**, 295–306.
15. Elmore,J.R., Sheppard,N.F., Ramia,N., Deighan,T., Li,H., Terns,R.M. and Terns,M.P. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the type III-B CRISPR–Cas system. *Genes Dev.*, **30**, 447–459.
16. Estrella,M.A., Kuo,F.-T. and Bailey,S. (2016) RNA-activated DNA cleavage by the type III-B CRISPR–Cas effector complex. *Genes Dev.*, **30**, 460–470.
17. Niewoehner,O., Garcia-Doval,C., Rostøl,J.T., Berk,C., Schwede,F., Bigler,L., Hall,J., Marraffini,L.A. and Jinek,M. (2017) Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. *Nature*, **548**, 543–548.
18. Kazlauskiene,M., Kostiuk,G., Venclovas,Č., Tamulaitis,G. and Siksnys,V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR–Cas systems. *Science*, **357**, 605–609.
19. Grüschow,S., Athukoralage,J.S., Graham,S., Hoogeboom,T. and White,M.F. (2019) Cyclic oligoadenylate signalling mediates mycobacterium tuberculosis CRISPR defence. *Nucleic Acids Res.*, **47**, 9259–9270.
20. Lin,J., Feng,M., Zhang,H. and She,Q. (2020) Characterization of a novel type III CRISPR–Cas effector provides new insights into the allosteric activation and suppression of the cas10 DNase. *Cell Discovery*, **6**, 29.
21. Julien,O. and Wells,J.A. (2017) Caspases and their substrates. *Cell Death Differ.*, **24**, 1380–1389.
22. Mastronarde,D.N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.*, **152**, 36–51.
23. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
24. Bepler,T., Morin,A., Rapp,M., Brasch,J., Shapiro,L., Noble,A.J. and Berger,B. (2019) Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods*, **16**, 1153–1160.
25. Liebschner,D., Afonine,P.V., Baker,M.L., Bunkóczi,G., Chen,V.B., Croll,T.I., Hintze,B., Hung,L.-W., Jain,S., McCoy,A.J. *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix. *Acta Crystallogr. D: Struct. Biol.*, **75**, 861–877.
26. Afonine,P.V., Klaholz,B.P., Moriarty,N.W., Poon,B.K., Sobolev,O.V., Terwilliger,T.C., Adams,P.D. and Urzhumtsev,A. (2018) New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D: Struct. Biol.*, **74**, 814–840.

27. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.-W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 213–221.
28. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.*, **66**, 486–501.
29. Afonine,P.V., Poon,B.K., Read,R.J., Sobolev,O.V., Terwilliger,T.C., Urzhumtsev,A. and Adams,P.D. (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D: Struct. Biol.*, **74**, 531–544.
30. Williams,C.J., Headd,J.J., Moriarty,N.W., Prisant,M.G., Videau,L.L., Deis,L.N., Verma,V., Keedy,D.A., Hintze,B.J., Chen,V.B. *et al.* (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.*, **27**, 293–315.
31. Sofos,N., Feng,M., Stella,S., Pape,T., Fuglsang,A., Lin,J., Huang,Q., Li,Y., She,Q. and Montoya,G. (2020) Structures of the Cmr-β complex reveal the regulation of the immunity mechanism of type III-B CRISPR–Cas. *Mol. Cell*, **79**, 741–757.
32. Rouillon,C., Athukoralage,J.S., Graham,S., Grüschow,S. and White,M.F. (2018) Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife*, **7**, e36734.
33. Poręba,M., Stróżyk,A., Salvesen,G.S. and Drąg,M. (2013) Caspase substrates and inhibitors. *Cold Spring Harb. Perspect. Biol.*, **5**, a008680.
34. Hauf,S., Waizenegger,I.C. and Peters,J.-M. (2001) Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science*, **293**, 1320–1323.
35. Wang,X., Li,X., Ma,Y., He,J., Liu,X., Yu,G., Yin,H. and Zhang,H. (2022) Inhibition mechanisms of CRISPR–Cas9 by acriia17 and acriia18. *Nucleic Acids Res.*, **50**, 512–521.
36. Mahendra,C., Christie,K.A., Osuna,B.A., Pinilla-Redondo,R., Kleinstiver,B.P. and Bondy-Denomy,J. (2020) Broad-spectrum anti-CRISPR proteins facilitate horizontal gene transfer. *Nat. Microbiol.*, **5**, 620–629.
37. Kato,K., Okazaki,S., Schmitt-Ulms,C., Jiang,K., Zhou,W., Ishikawa,J., Isayama,Y., Adachi,S., Nishizawa,T., Makarova,K.S. *et al.* (2022) RNA-triggered protein cleavage and cell growth arrest by the type III-E CRISPR nuclease-protease. *Science*, **378**, 882–889 .
38. Strecker,J., Demircioglu,F.E., Li,D., Faure,G., Wilkinson,M.E., Gootenberg,J.S., Abudayyeh,O.O., Nishimasu,H., Macrae,R.K. and Zhang,F. (2022) RNA-activated protein cleavage with a CRISPR-associated endopeptidase. *Science*, **378**, 874–881.
39. Tamulaitis,G., Kazlauskiene,M., Manakova,E., Venclovas,Č., Nwokeoji,A.O., Dickman,M.J., Horvath,P. and Siksnys,V. (2014) Programmable RNA shredding by the type III-A CRISPR–Cas system of streptococcus thermophilus. *Mol. Cell*, **56**, 506–517.
40. Staals,R.H., Zhu,Y., Taylor,D.W., Kornfeld,J.E., Sharma,K., Barendregt,A., Koehorst,J.J., Vlot,M., Neupane,N., Varossieau,K. *et al.* (2014) RNA targeting by the type III-A CRISPR–Cas csm complex of thermus thermophilus. *Mol. Cell*, **56**, 518–530.
41. Hu,C., van Beljouw,S.P., Nam,K.H., Schuler,G., Ding,F., Cui,Y., Rodríguez-Molina,A., Haagsma,A.C., Valk,M., Pabst,M., Brouns,S.J. and Ke,A. (2022) Craspase is a CRISPR RNA-guided, RNA-activated protease. *Science*, **377**, 1278–1285.
42. Liu,X., Zhang,L., Wang,H., Xiu,Y., Huang,L., Gao,Z., Li,N., Li,F., Xiong,W., Gao,T. *et al.* (2022) Target RNA activates the protease activity of craspase to confer antiviral defense. *Mol. Cell*, **82**, 4503–4518.
43. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with alphafold. *Nature*, **596**, 583–589.
44. Mirdita,M., Schütze,K., Moriwaki,Y., Heo,L., Ovchinnikov,S. and Steinegger,M. (2022) ColabFold: making protein folding accessible to all. *Nat. Methods*, **19**, 679–682.
45. De Las Peñas,A., Connolly,L. and Gross,C.A. (1997) SigmaE is an essential sigma factor in escherichia coli. *J. Bacteriol.*, **179**, 6862–6864.
46. Barchinger,S.E. and Ades,S.E. (2013) Regulated proteolysis: control of the escherichia coli σ E-dependent cell envelope stress response. *Regul. Proteolysis Microorg.*, **66**, 129–160.
47. De Las Peñas,A., Connolly,L. and Gross,C.A. (1997) The σE-mediated response to extracytoplasmic stress in escherichia coli is transduced by RseA and RseB, two negative regulators of σE. *Mol. Microbiol.*, **24**, 373–385.
48. Missiakas,D., Mayer,M.P., Lemaire,M., Georgopoulos,C. and Raina,S. (1997) Modulation of the escherichia coliσE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. *Mol. Microbiol.*, **24**, 355–371.
49. Stock,A.M., Robinson,V.L. and Goudreau,P.N. (2000) Two-component signal transduction. *Annu. Rev. Biochem.*, **69**, 183–215.
50. Buck,M., Gallegos,M.-T., Studholme,D.J., Guo,Y. and Gralla,J.D. (2000) The bacterial enhancer-dependent ς54 (ςN) transcription factor. *J. Bacteriol.*, **182**, 4129–4136.