Supplementary information

Methods and Materials

*To*Csm effector complex cloning, expression and purification

The cDNA encoding *To*Cas6-*To*Csm5 (Ton_0892-0897) and crRNA template were chemically synthesized. The crRNA template and *To*Cas6 genes were cloned into the NcoI/SalI and the NdeI/KpnI sites of the pCDFDuet-1 vector, respectively. The *To*Csm1 and *To*Csm4 genes were cloned into the NcoI/SalI and the NdeI/KpnI sites of the pACYCDuet-1 vector respectively. The *To*Csm3 gene was inserted into the NcoI/SalI sites of pETDuet-1, and the *To*Csm2-*To*Csm5 genes were inserted into the HindIII/SalI and the NdeI/XhoI sites of pRSFDuet-1, respectively. *To*Csm2 carries a 6×His-tag at its N-terminus.

Sequence of crRNA template: (from 5' to 3', 5'-handle is marked by shadow, spacer sequence is underlined)

GTTCCAGTAGGACAGAATTGTGTGGAAAGTGGCCCGAAACCCTTCCGGG
AGAGTGAGGACCCTCGAGTTCCAGTAGGACAGAATTGTGTGGAAAGTGG
CCCGAAACCCTTCCGGGAGAGTGAGGACCCTCGAGTTCCAGTAGGACAG
AATTGTGTGGAAAGTGGCCCGAAACCCTTCCGGGAGAGTGAGGACCCTC
GAGTTCCAGTAGGACAGAATTGTGTGGAAAGTGGCCCGAAACCCTTCCG
GGAGAGTGAGGACCCTCGAGTTCCAGTAGGACAGAATTGTGTGGAAAGT
GGCCCGAAACCCTTCCGGGAGAGTGAGGACCCTCGAGTTCCAGTAGGAC
AGAATTGTGTGGAAAG

The *To*Csm effector complex was produced in 6L of *E. coli* strain BL21 (DE3) at 37 °C by induction with 0.5 mM IPTG (isopropyl-β-d-thiogalactoside) for 5 h. The *E. coli* cells were harvested and lysed. The lysate was incubated in a hot water bath at 70 °C for 10 min and on ice for another 10 min to precipitate *E. coli* proteins. Then the lysate was clarified by ultracentrifugation at 32,000× g for 40 min and purified on a nickel-affinity column containing 5 mL Chelating Sepharose Fast Flow matrix (GE Healthcare). Gel filtration was performed as the final step of purification, using a

Superdex 200 10/300 GL column (GE Healthcare) in a buffer containing 50 mM Tris–HCl (pH 8.0), 500 mM NaCl. The best fractions were collected and the *To*Csm effector protein complex was concentrated to 3.6 mg/mL.

_To_Csm effector complex Cryo-EM sample preparation, data acquisition and structure determination

The 3.6 mg/mL *To*Csm complex was diluted to 0.5 mg/mL in buffer containing 25 mM Tris–HCl (pH 8.0), 220 mM NaCl, 4 mM DTT. Cryo-EM grids were frozen using a Vitrobot Mark IV (FEI) as follows: 5 μL of the sample was applied to a glow-discharged Quantifoil R1.2/1.3 holey carbon 300 mesh gold grid, blotted for 5-6 s in 70% humidity at 16 °C, and plunge frozen in liquid ethane cooled by liquid nitrogen.

Cryo-EM data were recorded on a Titan Krios (FEI) operated at 300 kV, equipped with a Gatan K2 Summit camera. SerialEM was used for automated data collection. Movies were collected at a nominal magnification of 22,500× in super-resolution mode resulting in a calibrated pixel size of 1.04 Å/pixel, with a defocus range of approximately -1.8 to -2.5 μm. 32 frames were recorded with a total dose of ~ 60 e^{-} $Å^{-2}$.

1,678 Movie frames were recorded. Beam-induced drift was corrected using MOTIONCORR2[1] . The contrast transfer function parameters of micrographs were estimated by Gctf[2]. Particles were picked by Gautomatch[3]. About 665k particles were collected in total, of which, ~322k are good enough for the 3D classification in Relion[4]. Apart from the rubbish class (~13% particles), two major classes were separated after the 3D classification. The dominant class (~50% particles) is the intact *To*Csm complex. To further improve the reconstruction resolution of the intact complex, a second cycle of 3D classification and various combinations of reconstruction with particles from different classes (the second cycle 3D classification) were performed. Finally, the sole class (group I, ~10% particles) yielded the best resolution (4.33 Å, better than others, using the bin2 data), was used for the final

reconstruction (3.35 Å with the unbinned data). Another class (~37% particles) constitutes an incomplete complex, which lacks the subunit of Csm2 (N-terminal 6×His-tag), possibly due to the degradation or dissociation of Csm2 during the Cryo-EM sample preparation. The resolution was 3.35 Å based on gold-standard Fourier shell correlation (FSC) =0.143 cutoff criterion[5]. Local resolution analysis was performed by ResMap[6].

_To_Csm effector complex Model building

The 3.35 Å density map was of sufficient quality for atomic model building. Briefly, crystal structures of *To*Csm1 and *To*Csm5 were fitted to the Cryo-EM density of *To*Csm effector complex as a rigid body. Atomic models of *To*Csm2, *To*Csm3 and *To*Csm4 were built *de novo* into density with the structures of their homologues as a guide using COOT[7]. The *To*Csm effector complex was finally refined using real-space refinement implemented in PHENIX[8]. All images of the model were created using PyMOL (www.pymol.org).

_To_Csm1 cloning, expression and purification

*To*Csm1 was cloned into a pET24a vector with a C-terminal 6×His-tag. The recombinant protein was expressed in *E. coli* Rosetta (DE3) strains in LB broth containing 50 μg/mL of kanamycin and 34 μg/mL of chloramphenicol. The *E. coli* cells were grown to an OD600 of 0.8, then protein expression was induced with 0.3 mM IPTG for 16 h at 16 °C. The cells were pelleted, resuspended in buffer containing 50 mM Tris-HCl (pH8.0), 300 mM NaCl and 20 mM imidazole, and lysed in a high-pressure cell disruptor. The supernatant was clarified by centrifugation (16,000 rpm for 40 min at 4 °C). The protein was first purified by a nickel-affinity column. The eluate from the nickel-affinity column was dialyzed against 50 mM Tris-HCl (pH 8.0) and 100 mM NaCl. After this, the dialyzed sample was further purified using Heparin HP, Resource Q, and Superdex200 10/300 GL columns with a FPLC system. The final elution buffer was 50 mM Tris-HCl (pH 8.0), 500 mM NaCl. The protein was

concentrated to 45 mg/mL and was flash-frozen in liquid nitrogen before stored at -80 °C.

*To*Csm1$_{crys}$-2ATP crystallization and data collection

*To*Csm1 protein was diluted to 8 mg/mL, 2mM ATP was added before crystallization. The initial crystals of *To*Csm1 were obtained by vapor diffusion using a reservoir solution of 0.2 M $MgCl_2$, 0.1 M tri-Na citrate pH 5.0 and 10% PEG 20000 at 16 °C. Through the extensive optimization of pH and precipitants, the qualities of crystals were further improved. The crystals were directly frozen in liquid nitrogen without additional cryoprotection. The X-ray diffraction data sets for *To*Csm1$_{crys}$-2ATP were collected at the Shanghai Synchrotron Radiation Facility beamline BL19U. The diffraction data were indexed, integrated and scaled using the HKL2000 program[9]. All images of the model were created using PyMOL (www.pymol.org).

*To*Csm1$_{crys}$-2ATP structural determination

The structure of *To*Csm1$_{crys}$ -2ATP was solved by the Molecular Replacement in the PHENIX program[8] using the crystal structure of apo-Csm1 from *T.onnurineus* as a search model. Further modeling of the structure was performed using the program COOT[7], and then subject to refinement using the PHENIX program[8].

In vitro synthesis of cyclic oligoadenylates (cOAs) and Mass Spectrometry analysis

To identify the product generated by *To*Csm effector complex, 500 nM *To*Csm effector complex, 20 μM noncomplementary target RNA, 500 μM ATP and 5mM $NiSO_4$ were incubated at 55 °C for 40 min in reaction buffer (25 mM Tris-HCl pH 8.0, 50 mM KCl). Reactions were stopped by heating to 95 °C for 10 min. The denatured protein was removed from the products by centrifugation for 15 min at 4 °C. The supernatant was analyzed by Mass Spectrometry at negative mode (MALDI-TOF/TOF Ultraflextreme^{TM}, Brucker, Germany).

In vitro target RNA cleavage assay

To identify the RNA cleavage activity of *To*Csm effector complex, the *To*Csm complex (15 μM), unlabelled 38-nt target RNA substrate (13 μM) and 8 mM $MnCl_2$ were incubated at 55 °C for 0, 5, 10 and 20 min in reaction buffer (25 mM Tris-HCl pH 8.0, 50 mM KCl). The reactions were stopped by the addition of 2× formamide loading buffer (100% formamide, 0.01% bromophenol blue) followed by heating at 95 °C for 3 min. Then the samples were analyzed on denaturing Urea-PAGE.

Reference:

1 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

2 Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12, doi:10.1016/j.jsb.2015.11.003 (2016).

3 Zhang, K. Gautomatch https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/. (2016).

4 Fernandez-Leiro, R. & Scheres, S. H. W. A pipeline approach to single-particle processing in RELION. *Acta Crystallogr D Struct Biol* **73**, 496-502, doi:10.1107/S2059798316019276 (2017).

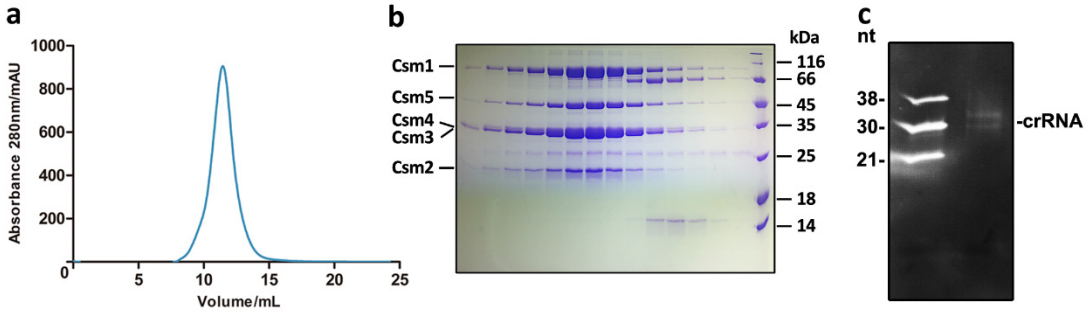
5 Rosenthal, P. B. & Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology* **333**, 721-745, doi:10.1016/j.jmb.2003.07.013 (2003).

6 Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat Methods* **11**, 63-65, doi:10.1038/nmeth.2727 (2014).

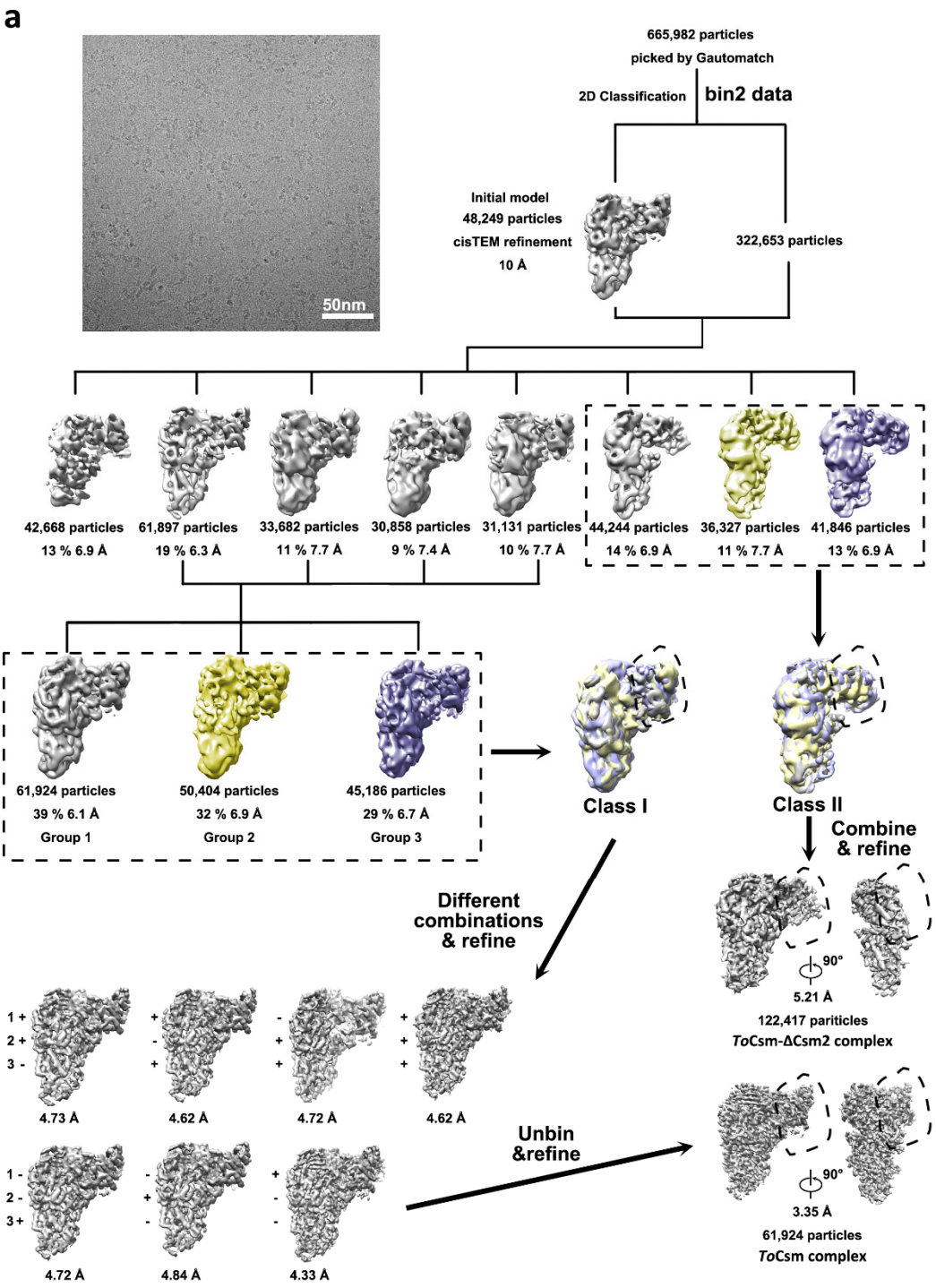
7 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **60**, 2126–2132 (2004).

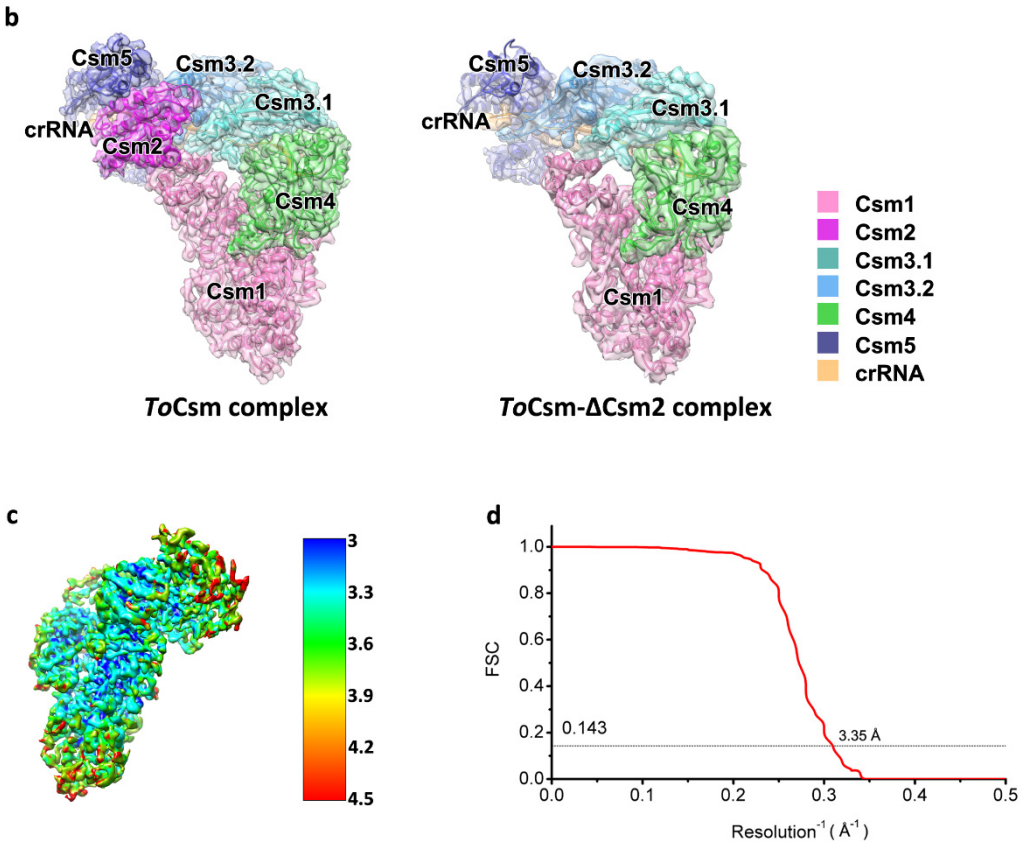
8 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

9 Z, O. & W., M. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).

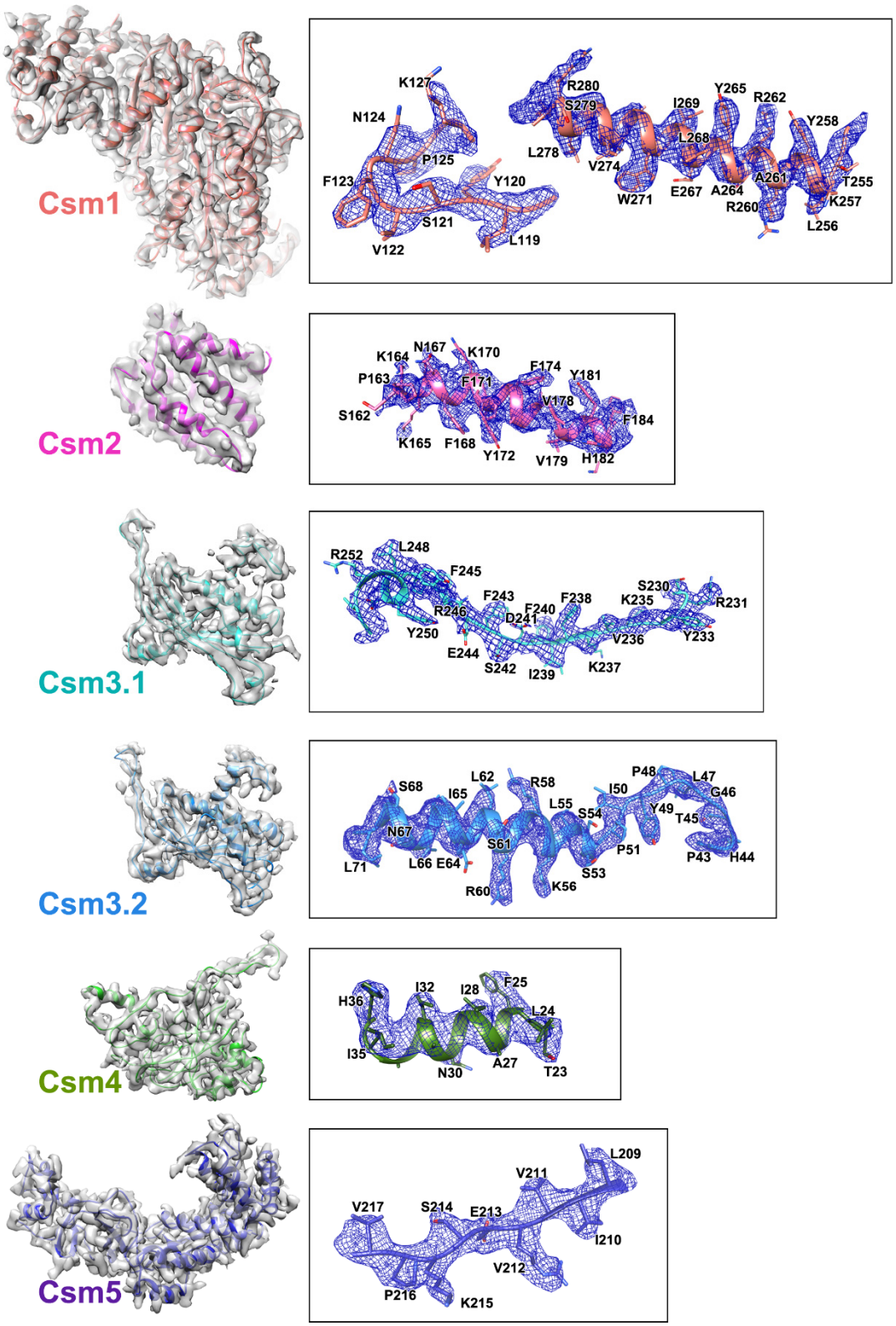


Supplementary information, Fig. S1. Purification of *To*Csm effector complex expressed from *E. coli*. **a** *To*Csm effector complex were fractionated by Superdex 200. **b** SDS-PAGE for protein composition analysis. Gels were stained with Coomassie brilliant blue. **c** Urea-PAGE for crRNA analysis.

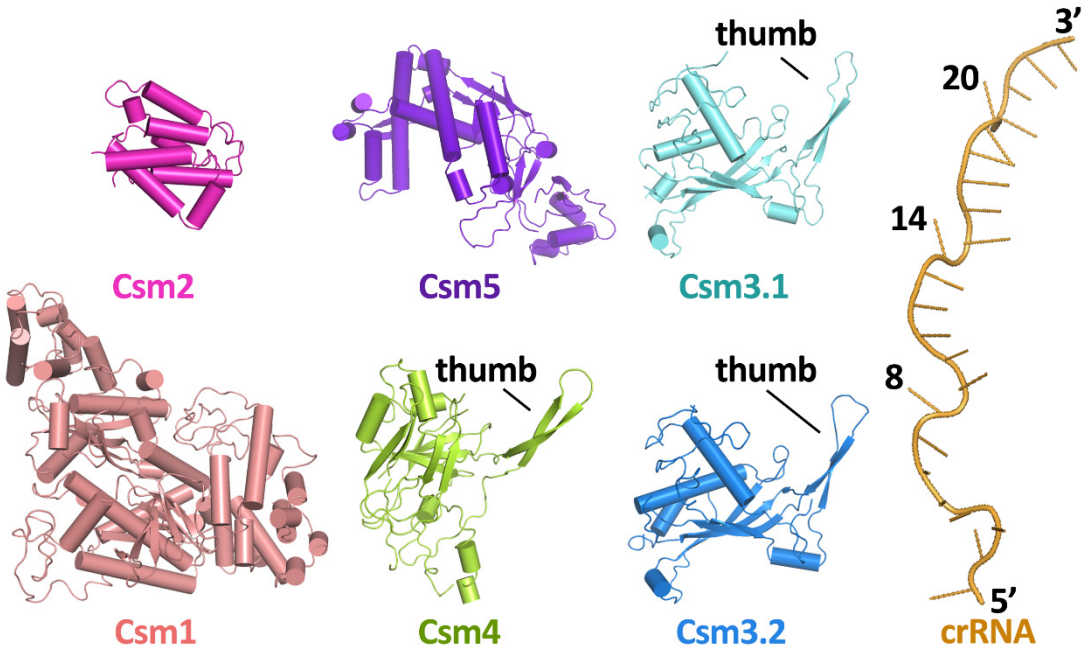

a
50nm
665,982 particles
picked by Gautomatch
2D Classification
bin2 data
Initial model
48,249 particles
cisTEM refinement
10 Å
322,653 particles
42,668 particles
13 % 6.9 Å
61,897 particles
19 % 6.3 Å
33,682 particles
11 % 7.7 Å
30,858 particles
9 % 7.4 Å
31,131 particles
10 % 7.7 Å
44,244 particles
14 % 6.9 Å
36,327 particles
11 % 7.7 Å
41,846 particles
13 % 6.9 Å
61,924 particles
39 % 6.1 Å
Group 1
50,404 particles
32 % 6.9 Å
Group 2
45,186 particles
29 % 6.7 Å
Group 3
Class I
Class II
Combine
& refine
Different
combinations
& refine
90°
5.21 Å
122,417 pariticles
ToCsm-ΔCsm2 complex
1 +
2 +
3 -
4.73 Å
4.62 Å
4.72 Å
4.62 Å
Unbin
&refine
1 -
2 -
3 +
4.72 Å
4.84 Å
4.33 Å
90°
3.35 Å
61,924 particles
ToCsm complex



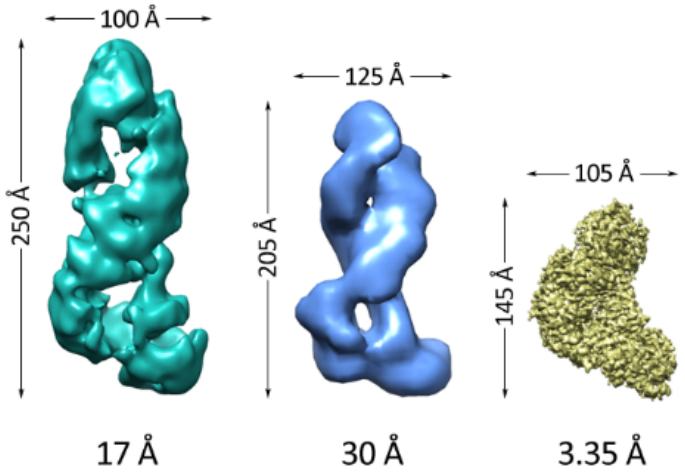
Supplementary information, Fig. S2. Cryo-EM data processing and resolution evaluation. a, b A representative micrograph of *To*Csm effector complex and flow chart of Cryo-EM data processing. Apart from the rubbish class (~13% particles), two major classes were separated after the 3D classification. The dominant class (~50% particles) is the intact *To*Csm complex. Another class (~37% particles) constitutes an incomplete complex, which lacks the subunit of Csm2 (*To*Csm-ΔCsm2 complex). **c** The final map local resolution variation was estimated with ResMap. **d** The gold standard Fourier shell correlation (FSC) curves of the final map. The resolution is 3.35 Å at FSC=0.143.



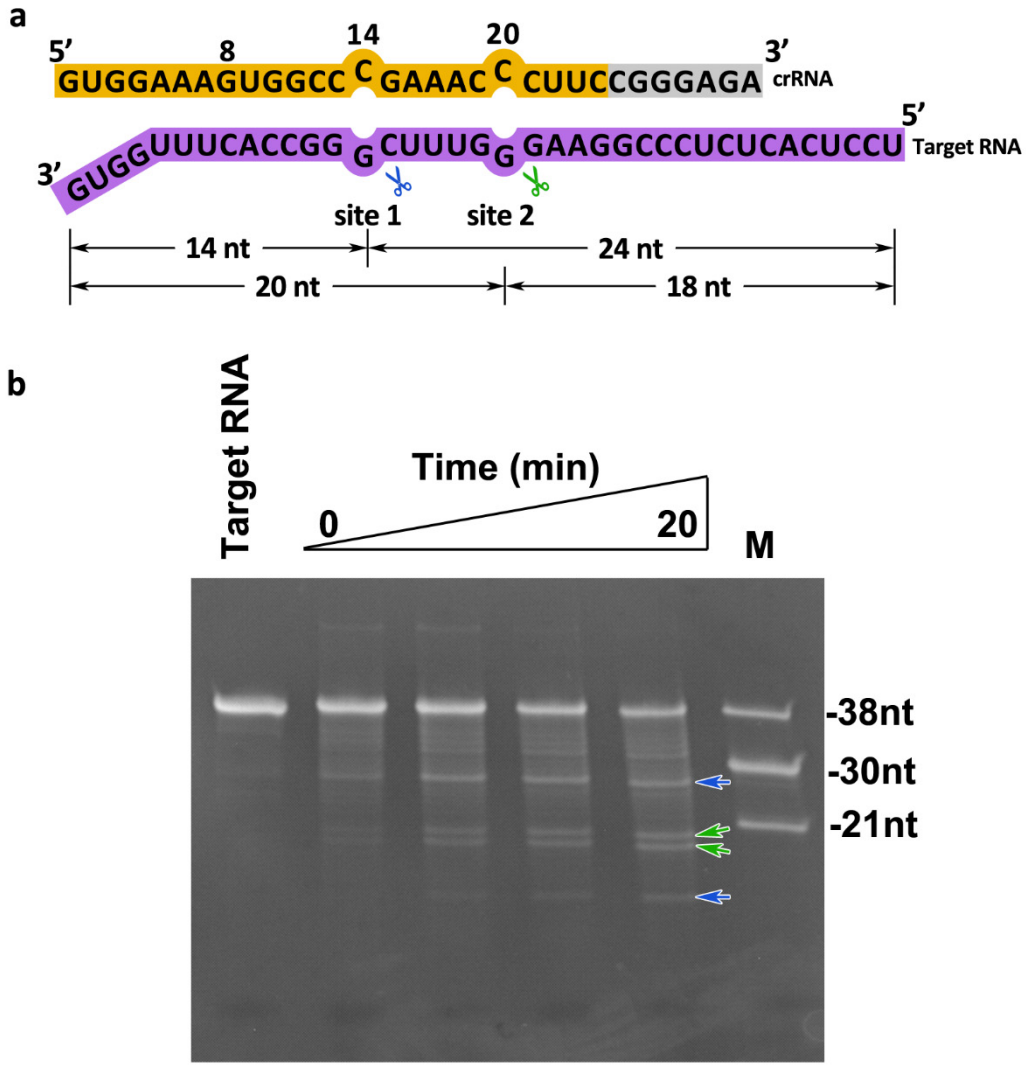
Supplementary information, Fig. S3. Typical electron density maps are shown overlaid with the corresponding model.



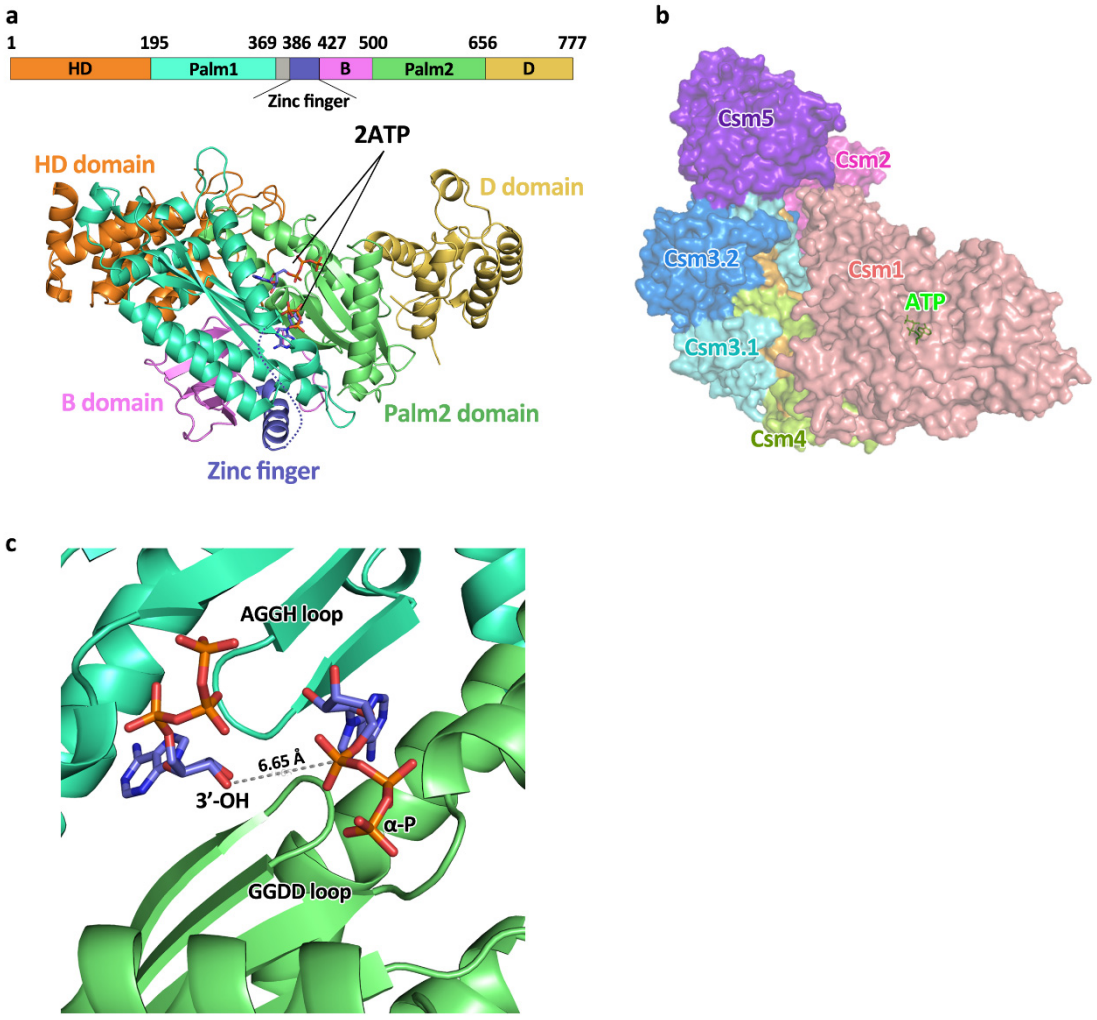
Supplementary information, Fig. S4. Structure of individual subunit in Cryo-EM *To*Csm effector complex structure.



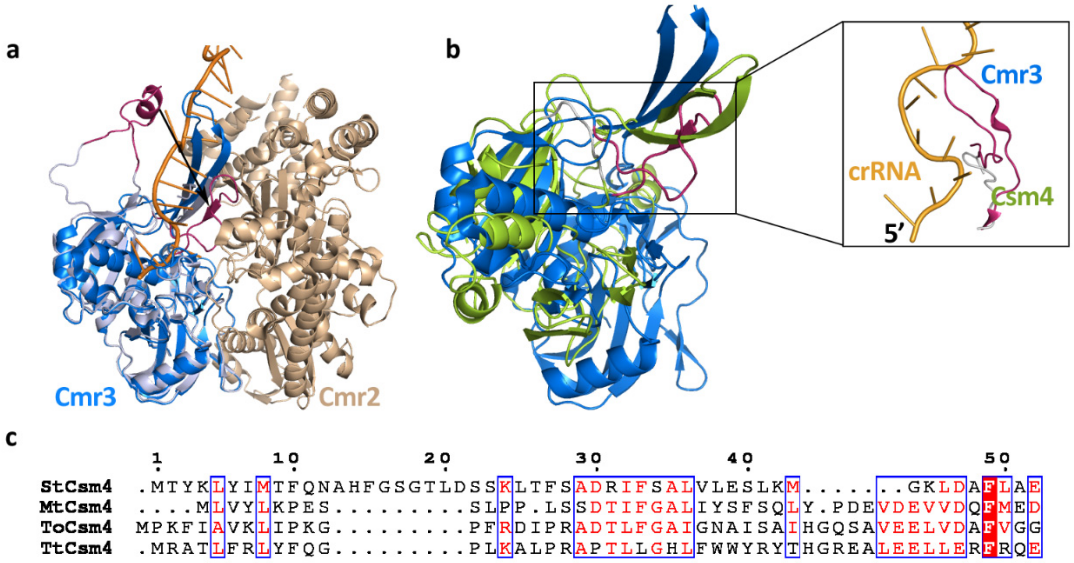
Supplementary information, Fig. S5. Comparison of the EM structure of Csm complexes from different species. *T. Thermophilus* (left), *S. Solfataricus* (middle), *T. onnurineus* (right). Height and width are shown to each Csm complex.



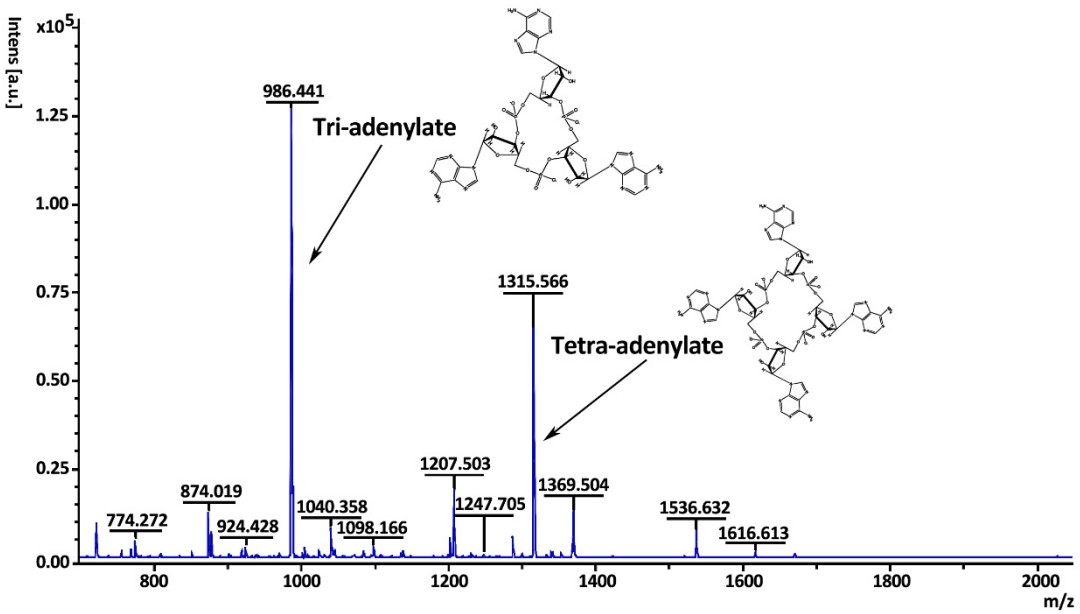
Supplementary information, Fig. S6. Target RNA cleavage activity of *To*Csm effector complex. a Schematic representation of target RNA cleavage sites. The constructed 24-nt crRNA is colored in orange, the unconstructed flexible 3'-end of crRNA is shown in grey. **b** Cleavage of 38-nt unlabelled target RNA by the *To*Csm effector complex in a time course (0, 5, 10, 20 min). Arrows indicate the cleaved fragments on a denaturing Urea-PAGE, corresponding to 24 nt/14 nt (site 1) and 18 nt/20 nt (site 2).



Supplementary information, Fig. S7. ATP-binding site in *To*Csm1$_{crys}$-2ATP and *To*Csm effector complex. a Domain arrangement of *To*Csm1 and crystal structure of *To*Csm1$_{crys}$-2ATP. The unstructured region is shown in grey. **b** Surface presentation of Cryo-EM structure of the *To*Csm effector complex with one ATP molecule. **c** Close-up view of two ATP molecules in the *To*Csm1$_{crys}$-2ATP structure.

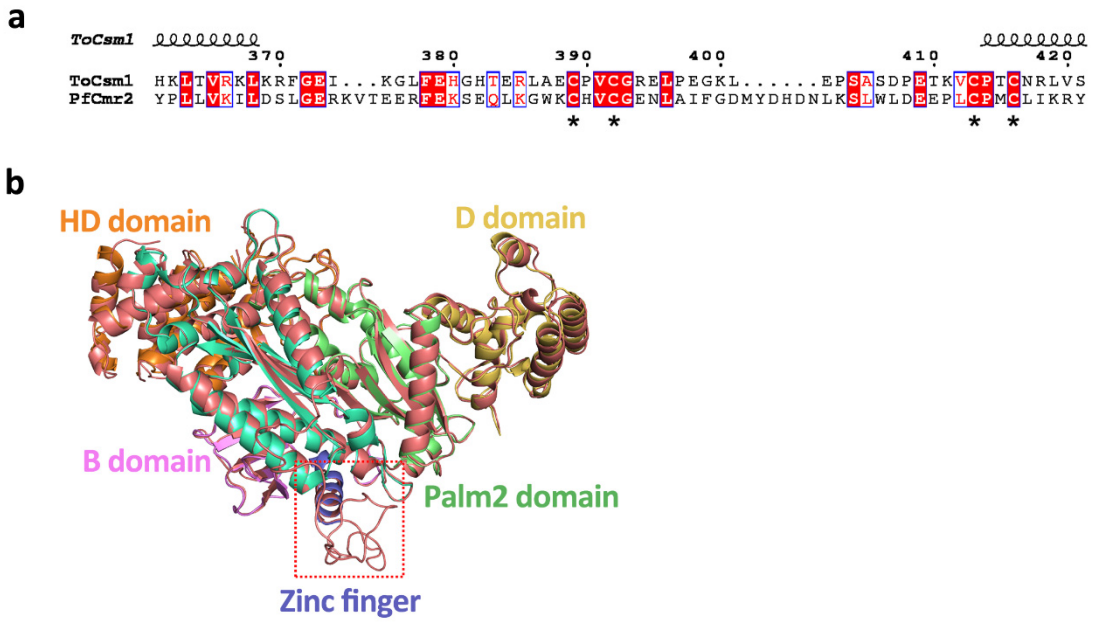


Supplementary information, Fig. S8. Structural comparison and sequence alignment of *To*Csm1 with its orthologous. a Structural comparison of *Pf*Cmr3 with crRNA bound (PDB: 3X1L) and unbound (PDB: 3W2W) shows the loop (warmpink，residues D10-S26) swinging a far distance to coordinate with crRNA. **b** Comparison of the loop in *To*Csm4 (gray) with that in *Pf*Cmr3 (warmpink, residues G13-D17) indicates the loop in *To*Csm4 is shorter. **c** Sequence alignment of *To*Csm4 with Csm4s from different species reported to be able to produce cOA_4/cOA_6.



Supplementary information, Fig. S9. MS analysis of ATP reaction products mediated by the *To*Csm effector complex. Both cOA_3 (Tri-adenylate) and cOA_4 (Tetra-adenylate) are the major products of the *in vitro* reaction, and there are no larger peaks corresponding to the size of cOA_5 or cOA_6. The theoretical molecular weight of cOA_3, cOA_4, cOA_5 and cOA_6 are 987.158, 1316.210, 1645.263 and 1974.315, respectively.

,



Supplementary information, Fig. S10. Sequence alignment of the zinc finger motif and structural comparison between *To*Csm1$_{crys}$-2ATP and *To*Csm1. a Sequence alignment of the zinc finger motif from *ToCsm1* with *Pf*Cmr2 shows the zinc finger sequence is conserved between the two species. The conserved residues are marked with asterisks. **b** Structural superimposition of ToCsm1$_{crys}$-2ATP with *To*Csm1 from the effector complex (salmon) shows that the latter possesses an ordered zinc finger structure.

Supplementary information, Table S1. Cryo-EM Data collection and Refinement Statistics

*To*Csm effector complex	
Data collection and processing	
Magnification	22,500
Voltage (kV)	300
Electron exposure (e-/Å^2)	60
Defocus range (μm)	1.8-2.5
Pixel size (Å)	1.04
Symmetry imposed	C1
Initial particle image (no.)	665,982
Final particle image (no.)	61,924
Map resolution (Å)	
FSC threshold	0.143
Map resolution (Å)	3.35
Refinement	
Initial model used (PDB code)	NA
FSC threshold	0.143
Model resolution (Å)	3.35
Model sharpening B factor (Å^2)	-90
Model composition	
Non-hydrogen atoms	16,246
Protein residues	1,993
Nucleotide	24
Ligands	ATP: 1
R.m.s deviations	
Bonds lengths (Å)	0.008
Bonds angles (º)	1.047
Validation	
Clashscore	7.93
Ramachandran plot	
Favored region (%)	85.65
Allowed region (%)	14.04
Outlier region (%)	0.31

Supplementary information, Table S2. Crystallographic Data collection and Refinement Statistics

	*To*Csm1$_{crys}$-2ATP
Data collection	
Beamline	SSRF BL-19U
Space group	C2
Cell dimensions	
a, b, c (Å)	158.20, 55.24, 92.95
α,β,γ (°)	90, 99.44, 90
Resolution (Å)	46.39-1.69 (1.75-1.69) *
I/σI	15.3 (8.1)
R_{pim}	0.044 (0.102)
Completeness (%)	99.22 (94.02)
Redundancy	6.7 (6.4)
Refinement	
Resolution (Å)	1.69
No. reflections	88086 (8282)
Rwork / Rfree	0.184/0.209
Protein residues	721
No. atoms	
ligands	62
Solvent	973
B-factors (Å^2)	
Protein	15.99
Ligands	12.37
Solvent	27.96
R.m.s. deviations	
Bond length (Å)	0.008
Bond angles (°)	1.26
Ramachandran plot	
Favored region (%)	97.75
Allowed region (%)	2.25
Outlier region (%)	0.00

*Highest resolution shell is shown in parentheses.