

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a

Confirmed
- ☐

☒
- The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐

☒
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐

☒
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☒

☐
- A description of all covariates tested
- ☐

☒
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐

☒
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐

☒
- For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted
Give P values as exact values whenever suitable.
- ☒

☐
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒

☐
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒

☐
- Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Next-generation sequencing data was collected using the Element Biosciences AVITI platform. Mass spectrometry data was collected using a Waters M-Class UPLC using Xcalibur software (v4.5.474.0). Cryo-EM data were collected using SerialEM (4.0) and Leginon software (3.5).
Data analysis	Structural figures were generated using ChimeraX (v1.10). Protein structures were predicted using AlphaFold (v3). For DRT10 ncRNA searches, covariance models were built with LocaRNA (v2.0.1) and searches were executed with Infernal (v1.1.4). For analysis of DRT10 loci, MAFFT (v7.511) was used for alignments, FastTree (v2.1.11) was used to build phylogenies, and iTOL (online) was used to visualize trees. DNA and RNA sequencing data was processed with cutadapt (v4.2), aligned with bwa-mem2 (v2.2.1) , filtered and converted to BAM with SAMtools (v1.17), converted to bigwig files with bamCoverage (v3.5.1), and visualized in IGV (v2.17.0). Unmapped reads were then extracted for downstream analysis using SAMtools fasta. Motif analysis was performed using MEME (v5.5.7). Mass spectrometry data were acquired in data-dependent mode using Xcalibur software. Cryo-EM image analyses were carried out in cryoSPARC (v4.7.1) and Relion (v3.0). Modeling of the RT-ncRNA complex was performed using ModelAngelo (v1.0.13) and COOT (v0.9.8.1), and model refinement was carried out in Phenix (v1.21.2-5419).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Next-generation sequencing data are available in the National Center for Biotechnology Information Sequence Read Archive (SRA) under the accession PRJNA1535902. Cryo-EM maps were deposited to the appropriate public database (i.e. EMDB, PDB) under the accessions EMD-57959, PDB:30QM (Eco1DRT10 RT-ncRNA complex); EMD-57969, PDB:30QU (Eco3DRT10 RT-ncRNA complex); and EMD-57972, PDB:30QX (dNTP-bound Eco3DRT10 RT-ncRNA complex). Previously determined protein structures used for structural comparisons were obtained from the PDB (6AR1; 9VMA; 9Z6Z; 9C0I). Mass spectrometry data are available in the MassIVE database under the accession MSV000103182. Source data is available for Figures 1b, 1f, 1n, 2a, 3h, 4f, 4j, 5a, Extended Data Figures 1g, 2a, 2b, 2d, 2j, 2k, 3g, 4b, 4c, 5a, 6a, 7e, 7f, 7i, 9a, 9b, 9d, 9g, 9h, 9i, and Supplementary Figures 2b, 2c, 2d, 2e, 2f, and 6c. All other data are available in the manuscript or in Supplementary Information.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were chosen based on previous studies in the phage defense field and are reported in the figure legends. Standard sample sizes were followed for next generation sequencing (Wiegand et al., 2024), biochemistry (Hobbs et al., 2022), mass spectrometry (Tal, Morehouse et al., 2021, Parnas, Jovanovic, Eisenhaure et al., 2015), and bacterial and phage growth assays (Stokar-Avihail et al., 2023, Gao et al., 2020).
Data exclusions	In cryo-EM, particles were excluded during 2D classification and 3D reconstruction as this is standard in the field.
Replication	All data could be reproduced. Experiments presented in the manuscript were repeated at least two times with similar results.
Randomization	Randomization was not applicable for this study design because the study did not involve allocation of human participants or animals to experimental groups. Experiments instead compared predefined bacterial strains, phages, genetic variants, and in vitro biochemical conditions, for which group assignment was determined by the experimental design rather than by random allocation. Control samples were used to establish reference activities.
Blinding	Investigators were not blinded as it was not applicable for this study design. There were no subjective analyses performed that could be affected by the interpretation of the data.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Monoclonal ANTI-FLAG M2 antibody (mouse), Sigma Aldrich, catalog number: F1804, clone M2.
Validation	According to the manufacturer, the "monoclonal ANTI-FLAG® M2 may be used in IP [immunoprecipitation] procedures when used in conjunction with an insoluble carrier matrix, such as a Protein G resin" (https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/175/747/f1804bul-ms.pdf). As suggested, the ANTI-FLAG M2 antibody was used together with Dynabeads Protein G resin (Thermo Fisher) in this study.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HEK293T cells were purchased from the American Type Culture Collection (ATCC CRL-3216)
Authentication	HEK293T cells were authenticated by the supplier using STR profiling.
Mycoplasma contamination	Cells were tested for mycoplasma contamination every 3 months or after thawing fresh passages and consistently tested negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A