nature research

Corresponding author(s): Lennart Randau

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☒	☐ An indication of 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 statistics 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
☒	☐ Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	RNA-seq data was collected at the Max-Planck Genome Centre Cologne. The HiSeq2000 sequencing platform generated the sequencing images and pixel-level raw data collection, image analysis, and base calling were performed by Illumina's Real Time Analysis (RTA V.1.7) software. The corresponding base call files (*.BCL) were converted to qseq files by the Illumina's BCL Converter, and the qseq files were subsequently converted to FASTQ files for downstream analysis. Sequence runs were demultiplexed using CASAVA 1.7. For the Csf5 crystal structures, diffraction data were collected at beamlines ID23-1 and ID29 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France.
Data analysis	RNA-seq: Illumina reads were trimmed (quality score limit of 0.05) and mapped to the reference genome of A. aromatoleum using CLC Genomics Workbench 5.0 (CLC Bio, Aarhus, Denmark). The following mapping parameters were used: mismatch cost: 2, insertion cost: 3, deletion cost: 3, length fraction: 0.5, similarity: 0.8. Structure analysis: Data were processed with the XDS program package for data reduction merging and scaling was performed using the AIMLESS program as implemented in the CCP4 package. Crank2 was used for experimental phasing of the VIL iodine derivative (CCP4 package). Coot in combination with Refmac5 (CCP4 package) and phenix.refine (PHENIX package) was used for iterative model building and refinement. The crystal structure of the 4 °C preparations was subsequently solved by molecular replacement using via the CCP4 implemented program Phaser. Figures were prepared in Pymol (www.pymol.org).

Mass-spectrometry:
Raw data of RNA-protein heteroconjugates were analysed and manually validated with the OpenMS pipeline RNPxl . For protein-protein crosslinks, Raw files were converted to mgf format with ProteomeDiscoverer 1.4 (Thermo Scientifc) and analysed with the software pLink (v. 1.23, pFind group) for identification of crosslinked peptides. Here, default settings were applied with carbamidomethylation of cysteines as fixed and oxidation of methionines as variable modification, FDR was set to 0.01. Results were filtered by excluding crosslinks supported by only one spectrum and additionally by applying a score cut-off value of 3. Interaction networks were visualised by xiNET.

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/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Availability of data and materials:
Crystallographic data and models have been deposited at the protein data bank (PDB) under accessions 6H9H and 6H9I.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	Sample size is not relevant for this study. We report the crystal structure of the Csf5 enzyme and the electron micrsoscopy structures of Csf2 and crRNP complexes.
Data exclusions	No data was excluded from the analyses.
Replication	In vivo Csf5 cleavage assays and the production and TEM visualization of recombinant crRNPs and Csf2 filaments were repeated three times with similar results. The cross-linking experiments were conducted once for each cross-linker concentration and technical duplicates of every sample were measured on the mass spectrometer. RNA-seq analysis of crRNA termini are from a single experiment. For each crystal structure, the data was derived from a single crystal.
Randomization	Randomization was not relevant for this study. Samples were not allocated into experimental groups.
Blinding	Blinding was not relevant for this study and group allocation during data collection was not performed.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

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