nature research

Corresponding author(s): Jennifer A. Doudna, Eva Nogales

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	SerialEM V3.6
Data analysis	Relion2.0, Cryosparc v1, Spider, EMAN2, Coot, Phenix2400, VMD-MDFF v0.4, UCSF-Chimera v1, Pymol v2, Snapgene V4, GraphPad Prism 7,Motioncor2, Gctf V1, Gautomatch V1, pLink1.9, pLabel2.8, Molprobity, MAFFT v7, RAxML v8, iTOL v3, CD-HIT

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

The EM maps have been deposited to EMDB with the following accession codes: EMD-8987, EMD-8988, EMD-8980, EMD-8994, EMD-8996, EMD-8991, EMD-8989, EMD-8890. The refined atomic models have been deposited to PDB with the following accession codes: 6E50, 6E7A, 6E79

All the unctted gels have been included in the source Data. CryoEM raw data and crosslinking data(> 10TB) are available upon email request.

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	Reported in figures
Data exclusions	No data excluded
Replication	All the data presented in this work have been replicated three or more times in biological replicate with consistent results including, E. coli. assays, Mammalian cell gene editing assays, protein purification, DNA cleavage assays, cryo-EM sample preparations, the single particle analysis of CasX with different substrates serve as technical replication for each other.
Randomization	For cryo-EM data collection, the exposure squares are randomly selected across the grid with good ice distribution. In cryoEM analysis, all the particles are randomly assigned to different classes in initial round, then generally converged in later rounds based on the cross correlation value with the class averages.
Blinding	No blinding was undertaken. All the exprements in this story are eitehr in vitro or cell based in vivo. Blindling is not applicable.

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

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Oakes et al. 2016 & The UC Berkeley Cell Culture Facility
Authentication	Cell lines were not re-authenticated for this study
Mycoplasma contamination	Cells were not re-tested for mycoplasma for this study, they were originally tested in the 2016 study, and The UC Berkeley Cell Culture Facility maintains verified mycoplasma free stocks.
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were dissociated from plates with trypsin, quenched with FBS & spun down, the media was decanted and the cells were resuspended in PBS
Instrument	All samples were run on an Attune NXT cytometer with a HTS
Software	FlowJo was used to analyze samples
Cell population abundance	All samples represent at least 10,000 events
Gating strategy	FSC/SSC and then on FSC hight vs area to select for single cells, following this cells were gated on FSC vs GFP for GFP disruption

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.