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 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) of the 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 whereever 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: All software used for modeling and simulation is named along with version numbers in the Methods section of the main manuscript or the supplementary information and is freely available for academic use or has an open source license, with the exception of AMBER14 for which a license must be purchased, see https://ambermd.org/ and the Schrödinger suite 2015.4 (https://www.schrodinger.com/).

Data analysis: All software used for analysis of the simulation results is named along with version numbers in the Methods section of the main manuscript or the supplementary information and is freely available for academic use or has an open source license.

For manuscripts utilizing custom algorithms/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 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

Fig 5: We provide three associated coordinate files in a zip file in the Supplementary Information.
The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request. There are no restrictions.
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.

- [x] Life sciences
- [ ] Behavioural & social sciences
- [ ] Ecological, evolutionary & environmental sciences

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

Life sciences study design

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

| Sample size | No experiments were performed. Sample size (simulation length and replicas) was considered at all stages of the modeling and simulation as described in the manuscript. |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | No data were excluded |
| Replication | No experiments were performed. Replicas of computations were performed and the results are described in the manuscript. |
| Randomization | No experiments were performed. Randomization was not relevant. |
| Blinding | No experiments were performed. Blinding was not relevant. |

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 |
|-----|-----------------------|
| [x] | Antibodies |
| [x] | Eukaryotic cell lines |
| [x] | Palaeontology and archaeology |
| [x] | Animals and other organisms |
| [x] | Human research participants |
| [x] | Clinical data |
| [x] | Dual use research of concern |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
| [x] | ChIP-seq |
| [x] | Flow cytometry |
| [x] | MRI-based neuroimaging |