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) and 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

Benchmark models were sampled using python 3.6, numpy 1.19.5 and deeptime 0.2.9, the custom python code to generate the data is made available in a public GitHub repository under https://github.com/markovmodel/ivampnets. Synaptotagmin-C2A molecular dynamics data was generated with openMM 7.1.1 as detailed in the cited publication [10.1021/acs.jctc.0c00043] and is publicly available under https://zenodo.org/record/6908073.

Data analysis

python 3.6, numpy 1.19.5, pyTorch 1.8.0, matplotlib 3.1.3, VMD 1.9.4

All data analysis was performed in Jupyter notebooks that are made public in the GitHub repository (https://github.com/markovmodel/ivampnets).

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](#).

All benchmark data can be generated from the Jupyter notebooks published in the GitHub repository (https://github.com/markovmodel/ivampnets). The molecular dynamics data set of synaptotagmin C2A is permanently available at Zenodo, https://doi.org/10.5281/zenodo.6908073. The all-atom crystal structure of synaptotagmin C2A is available under PDB ID 2R83. Restrictions apply to the availability of the villin data set, which were used under license for this study. Data are available from the authors upon request. The villin headpiece folding data is courtesy of D.E. SHAW research [Lindorff-Larsen et al, 10.1126/science.1208351].

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

| Reporting on sex and gender | 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.

- [x] 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](https://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 | Numerical simulations have been carried out to generate maximum possible sample size considering the required computation time and memory limitations for generation and subsequent analyses. No systematic reduction of samples has been performed. |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data exclusions | No data has been excluded.                                                                                                              |
| Replication | No new experimental findings are reported in this study. We ensure reproducibility of our results by publicly sharing the code to generate it.                                                   |
| Randomization | Data was grouped into batches at random.                                                                                               |
| Blinding    | Blinding was not relevant to this manuscript because the reported results are based on a mathematical model described in the study using publicly available code. |
| Materials & experimental systems | Methods          |
|---------------------------------|------------------|
| n/a                             | n/a              |
| ◐ Antibodies                    | ◐ Involved in the study |
| ◐ Eukaryotic cell lines         | ◐ ChIP-seq        |
| ◐ Palaeontology and archaeology | ◐ Flow cytometry  |
| ◐ Animals and other organisms   | ◐ MRI-based neuroimaging |
| ◐ Clinical data                 |                  |
| ◐ Dual use research of concern  |                  |