nature portfolio

Corresponding author(s): Shixin Liu, Yael David

Last updated by author(s): Feb 18, 2022

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

Coarse-grained, implicit-solvent molecular dynamics simulations were performed with one bead per protein residue or DNA base using the maximum entropy optimized force field (MOFF) for proteins and the molecular renormalization group coarse-graining model (MRG-CG) for DNA. All simulations were performed using the GROMACS simulation package.

Data analysis

For each simulated configuration, we first computed a molecular contact matrix. Contacts between two molecules, protein or DNA, were determined if any of their particles are within 1 nm. Using a depth-first search algorithm over the network constructed from the contact matrix, we identified the largest cluster. The high and low densities at a given temperature were determined using molecules whose centers of mass were in regions within 2.5 nm of the largest cluster or 150 nm or more away from it. These densities were used to obtain the critical temperature (TC) by fitting to the analytical expression ρ_H-ρ_L=A(T_c-T)^β, where β = 0.325 is the universality class of a 3D Ising model. Unlike the analytical Ising model, which assumes a homogeneous system, our system is highly heterogeneous, with strong H1:DNA interactions and substantially weaker H1:H1 and DNA:DNA interactions. As such, we note that there is additional uncertainty in TC arising from the sensitivity in density estimation. In particular, while TC was predicted to be less than 300 K for some systems, large clusters with a significant percentage of H1 are prevalent. Therefore, we used simulation data obtained at 300 K to analyze protein-DNA binding and diffusion coefficients. MDAnalysis was used to help with analysis of MD data.

Force and fluorescence data from the .h5 files generated from C-trap experiments were analyzed using tools in the lumicks.pylake Python library supplemented with other Python modules (Numpy, Matplotlib, Pandas) in a custom GUI Python script titled "C-Trap .h5 File Visualization GUI" (https://harbor.lumicks.com/single-script/c5b103a4-0804-4b06-95d3-20a08d65768f). This script was used to extract confocal images and fusion traces from droplet formation, fusion, and FRAP experiments. All specified scripts used to run or analyze C-Trap experiments can be accessed on Lumicks Harbor (harbor.lumicks.com).

The droplet fusion times (τ) were calculated by fitting sigmoidal curves to the force-time data over the time windows of droplet fusion. The fit equation used is F(t)= a/(a+e^(-b(t-t_(1/2)))), where F is the normalized magnitude of high-frequency force data (78 kHz) from the stationary optical trap, t is the time value of each force data point, a and b are the generalized fit parameters, and t_(1/2) is another fit parameter that

approximates the time of half-maximum force. The fits were then used to calculate τ as defined by the time lag between 30% and 80% of the normalized force values.

To analyze the reversibility of H1 foci formation, the lumicks.pylake Python package's greedy line tracking algorithm was applied to define line traces in the regions where the DNA tether was being relaxed.

Fluorescence recovery after photobleaching (FRAP) results were analyzed using a custom script titled "FRAP droplet imaging" (https://harbor.lumicks.com/single-script/3a796fac-dbb3-4fe1-8ce7-8b0cf8c25ad9).

For partition coefficient analysis, confocal image data were extracted from the .h5 files via the lumicks.pylake python library, and droplet masks were generated using the scipy python library. Briefly, Otsu thresholding of the image was followed by scipy.morphology.closing, scipy.morphology.remove_small_objects (min_size=6 pixels), and finally skimage.measure.label to define the mask for droplet regions.

Live-cell fluorescence microscopy still images were processed in FIJI version 2.0.0. Puncta were counted using the 3D Object Counter plugin, and colocalization coefficients were determined using the Coloc2 plugin. Live-cell continuous Z-stack imaging data were deconvolved in Imaris version 9.7.

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

Source data are provided with this paper. Other data are available upon reasonable request.

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	For single-molecule experiments, 15-30 kymographs were collected for each condition. For single-droplet manipulation experiments, at least 12 fusion events were analyzed for each condition based on the difficulty of the experiments and as standard practice in the field. Biochemical measurements were taken in triplicate as standard practice in the field. For live-cell fluorescence microscopy experiments, 10-15 images were collected for each condition. The number of measurements was chosen to ensure sufficient statistical power to distinguish significant differences between our samples.
Data exclusions	For live-cell fluorescence microscopy data analysis, duplicate images of the same cell and dead cells were excluded. For single-molecule experiments, tethers that display abnormal force-extension behaviors were discarded.
Replication	For biochemical measurements, at least three independent experiments were performed and attempts at replication were successful. Live-cell imaging data acquisition was performed across 13 non-consecutive days using fresh transfection and treatments each time. Attempts at replication were successful.
Randomization	Randomization was not applicable as there was no grouping in this study.
Blinding	Blinding was not applicable to this study since the results (single-molecule imaging, biochemical measurements, cell images) are not subjective examinations by the experimenter.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

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)	Human embryonic kidney cells (HEK293T) used in this study were sourced from ATCC.
Authentication	All cell lines used in this manuscript were authenticated by the vendors where we acquired these cell lines. ATCC used morphology, karyotyping, and PCR-based approaches.
Mycoplasma contamination	All cell lines were tested for mycoplasma monthly (MycoFluor Mycoplasma detection kit, Invitrogen) and tested negative for contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.