



HAL
open science

Polymer simulations guide the detection and quantification of chromatin loop extrusion by imaging

Thomas Sabaté, Benoît Lelandais, Edouard Bertrand, Christophe Zimmer

► To cite this version:

Thomas Sabaté, Benoît Lelandais, Edouard Bertrand, Christophe Zimmer. Polymer simulations guide the detection and quantification of chromatin loop extrusion by imaging. *Nucleic Acids Research*, 2023, 10.1093/nar/gkad034 . pasteur-04021372

HAL Id: pasteur-04021372

<https://pasteur.hal.science/pasteur-04021372>

Submitted on 9 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Polymer simulations guide the detection and quantification of chromatin loop extrusion by imaging

Thomas Sabaté^{1,2,3,*}, Benoît Lelandais^{1,4}, Edouard Bertrand^{2,†} and Christophe Zimmer^{1,*}

¹Institut Pasteur, Université Paris Cité, CNRS UMR 3691, Imaging and Modeling Unit, F-75015 Paris, France, ²IGH, University of Montpellier, CNRS UMR 9002, Montpellier, France, ³Sorbonne Université, Collège Doctoral, F-75005 Paris, France and ⁴Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, F-75015 Paris, France

Received June 02, 2022; Revised December 03, 2022; Editorial Decision January 03, 2023; Accepted January 25, 2023

ABSTRACT

Genome-wide chromosome conformation capture (Hi-C) has revealed the organization of chromatin into topologically associating domains (TADs) and loops, which are thought to help regulate genome functions. TADs and loops are understood as the result of DNA extrusion mediated by the cohesin complex. However, despite recent efforts, direct visualization and quantification of this process in single cells remains an open challenge. Here, we use polymer simulations and dedicated analysis methods to explore if, and under which conditions, DNA loop extrusion can be detected and quantitatively characterized by imaging pairs of fluorescently labeled loci located near loop or TAD anchors in fixed or living cells. We find that under realistic conditions, extrusion can be detected and the frequency of loop formation can be quantified from fixed cell images alone, while the lifetime of loops and the speed of extrusion can be estimated from dynamic live-cell data. Our delineation of appropriate imaging conditions and the proposed analytical methods lay the groundwork for a systematic quantitative characterization of loop extrusion in fixed or living cells.

INTRODUCTION

Over the last decade, much progress has been made in understanding the three-dimensional organization of chromatin, thanks to powerful genomic techniques such as Hi-C (1), which provides genome-wide maps of DNA–DNA contact frequencies (2,3). A notable milestone was the discovery of topologically associating domains (TADs) (4,5), sub-megabase scale regions of enhanced chromatin contacts that appear as blocks on the diagonal of Hi-C maps and that are believed to help regulate gene expression by facili-

tating - or on the contrary impeding - enhancer-promoter interactions (6–15). TADs are often associated with off-diagonal peaks (or ‘corner dots’) in the Hi-C maps, reflecting enriched contacts between the two distant loci that define the TAD boundaries (hereafter called ‘anchors’), and which are interpreted as chromatin loops. TAD or loop anchors are typically binding sites for the insulator protein CTCF (1) (with convergent orientation of the CTCF motifs) and TADs with corner dots (also called loop domains) depend on the ring-like cohesin complex, as they disappear upon cohesin removal (16). The formation of TADs and corner dots is now understood as the result of DNA loop extrusion mediated by the cohesin complex (2,3,17–20). In this process, the cohesin ring complex binds to DNA and progressively pulls out a loop of chromatin, until the complex unbinds or stops extruding at obstacles such as CTCF-bound anchor loci. At this point, the loop is temporarily stabilized until the cohesin complex or CTCF dissociate and the anchors detach from each other (21,22). This extrusion mechanism is supported by several lines of evidence, including polymer modeling (19,20,23), Hi-C studies where cohesin (16), CTCF or other regulators of the cohesin complex are experimentally depleted (18,24,25), as well as direct visualization of cohesin-mediated DNA extrusion *in vitro* (17,26). The dynamic nature of TADs and loops is further supported by multiplexed DNA FISH studies, which underlined the high cell-to-cell heterogeneity in chromatin structure within TADs or loops (27–32), and by single-molecule tracking of CTCF and the cohesin subunit RAD21, which showed that these factors have residence times on chromatin orders of magnitude shorter than the cell cycle (1–2 min for CTCF and 22 min for RAD21) (33).

Despite these studies, visualizing and characterizing the dynamic process of loop extrusion directly in single living cells remains a largely unaddressed challenge (34), except for two very recent reports (35,36). A seemingly straightforward experimental approach to visualize loop extrusion in living cells is to track two loop anchors with light

*To whom correspondence should be addressed. Tel: +33 140613170; Email: thomas.sabate@pasteur.fr
Correspondence may also be addressed to Christophe Zimmer. Tel: +33 140613891; Email: czimmer@pasteur.fr
†The authors wish it to be known that, in their opinion, the last two authors should be regarded as Joint Last Authors.

microscopy, using distinct fluorescent reporters located at or near each anchor, and to monitor the progressive decrease of the reporter-reporter distance that is expected to result from extrusion (37). Fluorescent labeling of anchors can be achieved using DNA FISH probes in fixed cells (28), while arrays of repeats (38), dead Cas9 (39–41) or parS-parB (42) can be used to label loci both in fixed and living cells. In practice, however, direct visualization of extrusion by tracking fluorescent loci is complicated by several biological and experimental sources of uncertainty. These include: (i) unavoidable errors in computing spatial coordinates and distances between genomic loci from noisy imaging data, (ii) photobleaching, which limits the number of time points over which loci can be tracked with sufficient signal intensity for accurate localization, (iii) the size of fluorescent labels and the distance of fluorescent reporters to the anchors (hereafter called reporter–anchor separation), (iv) the stochastic movements of chromatin, which can bring together genomically distant loci in space even in absence of any active process such as loop extrusion, (v) the potential rarity of extrusion events and by consequence the potentially large number of cells that must be analyzed to accurately characterize statistical parameters such as the average lifetime of loops. Because of these complicating factors, it is not *a priori* evident whether loop extrusion can be unambiguously visualized by imaging at all, and if it can, under what experimental conditions, and whether key biophysical parameters of loop extrusion can be quantified. Here, we aim to clarify these requirements considering basic expectations from polymer dynamics, taking into account available Hi-C data, and considering various technical limitations of imaging techniques. We approach this by (i) simulating realistic distributions and time series of anchor–anchor distances, (ii) proposing analytical methods to characterize loop extrusion from these data and (iii) quantitatively testing these methods on the simulations. Our results will guide future experimental work aiming to quantify loop extrusion and its dynamics in fixed and living cells.

MATERIALS AND METHODS

Our study involves (i) numerical simulations of polymers undergoing loop extrusion, (ii) analytical models of probability distributions of anchor–anchor vectors and distances, (iii) analyses of simulated data, and (iv) analysis of experimental data from Hi-C, ChIP-seq or imaging. The following provides details on these four methodological parts.

Polymer simulations with loop extrusion

Langevin dynamics simulations. We used polymer simulations to model the dynamics of a chromatin fiber segment subjected or not to loop extrusion. The simulated polymer consisted of 600 beads, and polymer motions were simulated with Langevin dynamics in fixed boundary conditions using LAMMPS (43). Consecutive beads were connected by a harmonic bond with a potential $E_{\text{bond}} = 30(r - 1)^2$, where r is the distance between bead centers. The polymer stiffness was modeled using a harmonic potential $E_{\text{bending}} = K_0(\theta - \theta_0)^2$, where K_0 is the stiffness parameter (set to $K_0 = 0.1$), θ is the angle between three

consecutive beads and $\theta_0 = 180^\circ$ is the equilibrium value (corresponding to three aligned beads). We verified that the contact frequencies did not strongly depend on the exact value of K_0 (see section ‘*Simulated contact frequency maps*’ below). By default, the polymer was confined in a sphere of radius 18 bead diameters using the energy potential $E_{\text{wall}} = 4[(\frac{\sigma}{R})^{12} - (\frac{\sigma}{R})^6]$ for $R < R_c$, where R is the distance between the surface of the confining sphere and the center of a bead, σ is a size factor set to 0.5 bead diameters and R_c is the cutoff distance set to 0.5 bead diameters (for $R \geq R_c$, $E_{\text{wall}} = 0$). This confinement implied a volume occupancy ratio of the polymer of $\rho = 1.3\%$. While this is much lower than estimates of chromatin volume occupancy *in vivo* (~10–15%) (44–46), this discrepancy is not critical given the consistency of our simulation predictions with Hi-C and imaging data (see Results sections ‘*Comparing polymer simulations to Hi-C data*’ and ‘*Comparing polymer simulations to imaging data*’). We considered the polymer to be equilibrated when its radius of gyration and end-to-end distance were both stabilized, which was the case after ~7.5 million time steps (Supplementary Figure S1A, B). After this, we recorded the positions of each bead every 1000 time steps until the end of the simulation (~12 million time steps). Polymer coordinates were imported for further analysis in Python using MDAnalysis (47).

In order to convert simulation time and space dimensions to physical units, we compared Mean Squared Displacement (MSD) from simulations to experimental MSD curves of chromatin loci tracked by live-cell microscopy (40,41,48). This comparison led to the conversion of 1000 simulation time steps to 0.3 s and of 1 bead diameter to $2r_0 = 50$ nm. Assuming a chromatin compaction of $C = 60$ bp/nm (49), as previously estimated by comparing simulation predictions to experimental data in yeast, this implies that 1 bead corresponds to $g = 2r_0C = 3$ kb and the entire 600 bead polymer to 1800 kb of chromatin.

Simulating loop extrusion. To model loop extrusion, we forced the formation of a harmonic bond between non-consecutive beads. We assumed that loop extrusion initially occurs bidirectionally (17,26), i.e. that if beads i and j ($i < j$) are bonded at time t , then beads $i - 1$ and $j + 1$ are bonded at time $t + \Delta t$ (while the bond between beads i and j is deleted), where Δt is the time needed to extrude two beads (Supplementary Figure S2A). The speed of loop extrusion (in base pairs per seconds) is thus defined as $V_0 = 2g/\Delta t$. By default, our simulations assumed that extrusion started at a random location between the beads representing the anchors (extrusion barriers) and proceeded bidirectionally at $V_0 = 1$ kb/s (17,26) until reaching an anchor. Thereafter, loop extrusion proceeded unidirectionally, at the halved speed $V_0/2 = 0.5$ kb/s until reaching the second anchor (18,50–52), whereupon extrusion stopped (Supplementary Figure S2A). By default, we then maintained the bond between the two anchor beads until the end of the simulation. However, for simulations used in the Results sections ‘*Quantifying closed loop lifetimes from live-cell trajectories*’ and ‘*Quantifying the speed of loop extrusion from live-cell trajectories*’, the time spent in the closed state (i.e. when the anchors are maintained in contact by a bond) was drawn from a truncated exponential distribu-

tion. At the end of the closed state, the bond linking the two anchors was deleted and the polymer was allowed to relax without loops until the end of the simulation. We simulated loops ranging from 150 kb to 990 kb in size, and defined the positions of the anchor beads using the following bead indexes: 150 kb loops: {275, 324}, 228 kb loops: {262, 337}, 300 kb loops: {250, 349}, 450 kb loops: {225, 374}, 504 kb loops: {216, 383}, 600 kb loops: {200, 399}, 702 kb loops: {183, 416}, 798 kb loops: {167, 432}, 990 kb loops: {135, 464}.

For Supplementary Figure S11, and unlike elsewhere in the paper, we performed simulations where extrusion always started in the middle of the loop and proceeded bidirectionally at constant speed, until both anchors were reached simultaneously. This was done to ensure that extrusion started and ended at the same time in all cases, which facilitated the comparison of simulations with the theoretical linear model (see Eq. (3) and Supplementary Figure S10B, green dashed line). For each value of these parameters, four hundred synchronized time series were averaged together. We simulated loops of 150 kb, 300 kb, and 600 kb extruding at speeds of 0.2, 1 and 5 kb/s, and loops of 990 kb extruding at speeds of 1 and 5 kb/s. We also repeated these simulations with much weaker confinement (using a sphere of radius 150 rather than 18 bead diameters).

Analytical models

Anchor–anchor vector distributions. In the Results section ‘Estimating the fraction of loop states from static imaging data’, we used an analytical model to estimate the fractions of loops in open vs extruding vs closed states (these states correspond respectively to absence of loops, loops whose size increases with time, and to a stable loop with the two anchors in contact) based on the measured coordinates of anchor–anchor vectors. This model is based on the basic properties of an ideal polymer chain, for which the anchor–anchor vector $\vec{R} = (\delta x, \delta y, \delta z)$ is a random variable that obeys the normal probability density: $P(\vec{R}; \sigma) = P_x(\delta x; \sigma) P_y(\delta y; \sigma) P_z(\delta z; \sigma)$

with: $P_w(\delta w; \sigma) = (\sqrt{2\pi}\sigma)^{-1} \exp(-\frac{\delta w^2}{2\sigma^2})$ for each coordinate $w \in \{x, y, z\}$ and where the variance σ^2 is proportional to the number N_b of Kuhn lengths b separating the anchors and is given by $\sigma^2 = N_b b^2/3$. This implies that: $P(\vec{R}; \sigma) = (2\pi\sigma^2)^{-\frac{3}{2}} \exp(-\frac{R^2}{2\sigma^2})$ and the mean squared anchor–anchor distance (MSAAD) is given by: $\langle R^2 \rangle = N_b b^2$, where the brackets denote statistical averaging. In presence of a loop, we assume that the parts of the polymer outside the loop behave as if the part within the loop was absent, thereby shortening the number of Kuhn lengths between the anchors. Note that in the absence of bending stiffness ($K_0 = 0$), the Kuhn length coincides with the bead diameter ($b = 2r_0 = 50$ nm), which for an anchor–anchor separation of 150 kb, corresponding to $150/g = 50$ beads and 49 Kuhn lengths would imply $\sigma_{\text{open}} = \sqrt{\frac{N_b b^2}{3}} \approx 202$ nm. However, because of the bending potential E_{bending} assumed above, the variance measured on simulated data is slightly larger, namely $\sigma_{\text{open}} = 216$ nm, implying a slightly larger Kuhn length of

$b = 53.5$ nm. Also note that the finite bead radius of $r_0 = 25$ nm used in our simulations affects anchor-to-anchor vectors in the closed state and at the end of the extruding phase, since distances between anchors fluctuate around $2r_0 = 50$ nm. To avoid this bias, which is not accounted for by the above analytical model, we shortened the anchor–anchor distances predicted by the simulations in the closed state by 50 nm, and between 0 and up to 50 nm for the last 60 time points of the extruding state. After this correction, in the closed state, measurement of the MSAAD led to a small value of $\sigma_{\text{closed}} \approx 7$ nm.

The number of Kuhn lengths N_b between the two anchors depends on the loop state. It is largest ($N_b = N_0$) when the loop is open, equals zero when the loop is closed, and assumes intermediate values during extrusion.

The above model for anchor–anchor vectors applies to the open state with $\sigma_{\text{open}}^2 = N_0 b^2/3$ and to the closed state with $\sigma_{\text{closed}}^2 \ll \sigma_{\text{open}}^2$. Hence:

$$\begin{cases} P_w^{\text{open}}(\delta w) = \frac{1}{(2\pi)^{\frac{1}{2}} \sigma_{\text{open}}} \exp\left(-\frac{\delta w^2}{2\sigma_{\text{open}}^2}\right) \\ P_w^{\text{closed}}(\delta w) = \frac{1}{(2\pi)^{\frac{1}{2}} \sigma_{\text{closed}}} \exp\left(-\frac{\delta w^2}{2\sigma_{\text{closed}}^2}\right) \end{cases}$$

In the extruding state, the number of Kuhn lengths continuously varies from N_0 to 0. As a result, the probability density $P_w^{\text{extruding}}(\delta w)$ is an integral over σ^2 varying from σ_{closed}^2 up to σ_{open}^2 . If the speed of extrusion was constant, then loops of all sizes (from $N_b = 0$ to $N_b = N_0$) would be represented equally in the integral, such that:

$$\begin{aligned} P_w^{\text{extruding}}(\delta w) &= \int_{\sigma_{\text{closed}}^2}^{\sigma_{\text{open}}^2} P_w(\delta w; s) ds^2 \\ &= \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{\sigma_{\text{closed}}^2}^{\sigma_{\text{open}}^2} \frac{1}{s} \exp\left(-\frac{\delta w^2}{2s^2}\right) ds^2. \end{aligned}$$

A complication arises from the fact that the speed of extrusion is halved once the extrusion complex reaches one of the two anchors and extrusion switches from bidirectional to unidirectional. As a consequence, extruding states with unidirectional extrusion are twice more frequent than states with bidirectional extrusion, leading to twice more frequent cases where $N_b = 0$ than where $N_b = N_0$. To take this into account, we introduce a weighting factor $D(\sigma)$ in the integral, such that:

$$\begin{aligned} P_w^{\text{extruding}}(\delta w) &= \int_{\sigma_{\text{closed}}^2}^{\sigma_{\text{open}}^2} D(s) P_w(\delta w; s) ds^2 \\ &= \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{\sigma_{\text{closed}}^2}^{\sigma_{\text{open}}^2} \frac{D(s)}{s} \exp\left(-\frac{\delta w^2}{2s^2}\right) ds^2. \end{aligned}$$

Assuming that extrusion is initiated at uniformly random locations between the two anchors, $D(\sigma)$ increases linearly

with σ^2 between σ_{open}^2 and σ_{closed}^2 , such that:

$$D(\sigma) = \frac{2\sigma_{\text{open}}^2 - \sigma_{\text{closed}}^2 - \sigma^2}{\frac{3}{2}\sigma_{\text{open}}^4 + \frac{3}{2}\sigma_{\text{closed}}^4 - 3\sigma_{\text{closed}}^2\sigma_{\text{open}}^2}$$

This equation can be derived by considering a linear function that equals 2 when $\sigma = \sigma_{\text{close}}$ and 1 when $\sigma = \sigma_{\text{open}}$ and whose integral between σ_{close}^2 and σ_{open}^2 equals 1 (to ensure that it is a probability density).

In practice, the coordinates of anchors (or nearby fluorescent reporters) are not perfectly known, but are computed from noisy images. Because of noise, these coordinates are estimated with a finite precision, which is usually different along the axial and lateral directions of the microscope. To reflect this, we assumed that coordinates are perturbed by random, normally distributed anisotropic errors (hereafter called localization errors), with standard deviations σ_w along each axis $w \in \{x, y, z\}$. This leads to the following modified equations for the probability densities of anchor–anchor coordinate differences in each of the three loop states:

$$\begin{cases} P_w^{\text{open}}(\delta w) = \frac{1}{(2\pi)^{\frac{1}{2}} \sqrt{\sigma_{\text{open}}^2 + 2\sigma_w^2}} \exp\left(-\frac{\delta w^2}{2\sigma_{\text{open}}^2 + 4\sigma_w^2}\right) \\ P_w^{\text{closed}}(\delta w) = \frac{1}{(2\pi)^{\frac{1}{2}} \sqrt{\sigma_{\text{closed}}^2 + 2\sigma_w^2}} \exp\left(-\frac{\delta w^2}{2\sigma_{\text{closed}}^2 + 4\sigma_w^2}\right) \\ P_w^{\text{extruding}}(\delta w) = \frac{1}{(2\pi)^{\frac{1}{2}}} \int_{\sigma_{\text{closed}}^2}^{\sigma_{\text{open}}^2} \frac{D(s)}{\sqrt{s^2 + 2\sigma_w^2}} \exp\left(-\frac{\delta w^2}{2s^2 + 4\sigma_w^2}\right) ds^2 \end{cases} \quad (1)$$

For a combination of the three states with fractions A_{closed} , A_{open} and $A_{\text{extruding}} = 1 - A_{\text{closed}} - A_{\text{open}}$, the full probability density of anchor–anchor coordinate differences reads:

$$P_w(\delta w) = A_{\text{closed}} P_w^{\text{closed}}(\delta w) + A_{\text{open}} P_w^{\text{open}}(\delta w) + A_{\text{extruding}} P_w^{\text{extruding}}(\delta w) \quad (2)$$

Mean anchor–anchor distance as function of time. In the Results section ‘Quantifying the speed of loop extrusion from live-cell trajectories’, we analyzed time series of anchor–anchor distances using an analytical model of the MSAAD $\langle R^2 \rangle(t)$ as function of time. For a given time series, we defined $t = 0$ as the time point when extrusion stops, i.e. the start of the closed state. We further assumed, as above, that the MSAAD obeys the properties of an ideal polymer chain whose length (i.e. its number of monomers) is diminished by the length of the loop. Accordingly, if extrusion proceeded at a constant speed V_0 , then the number of Kuhn lengths between the two anchors would decay linearly with time as $N_b(t) = N_0(V_0/s_0)(-t)$ for $t \in [-\frac{s_0}{V_0}, 0]$, where N_0 is the number of Kuhn lengths in the open state and s_0 is the genomic distance between the anchors, in base pairs.

Under the assumptions above, and ignoring localization errors, we have $\langle R^2 \rangle(t) = N_b(t)b^2 = N_b(t)/N_0 R_0^2$, where R_0^2 is the MSAAD in the open state (before extrusion starts). Because of random localization errors, the measured MSAAD is $\langle R^2 \rangle(t) = N_b(t)/N_0 R_0^2 + R_{\text{loc}}^2$, where $R_{\text{loc}}^2 = 4\sigma_{x,y}^2 + 2\sigma_z^2$ is the contribution of random localization errors.

Thus, for a constant extrusion speed V_0 , the MSAAD obeys the following linear law:

$$\langle R^2 \rangle(t) = R_0^2 (V_0/s_0)(-t) + R_{\text{loc}}^2 \quad (3)$$

However, we assumed that loop extrusion switches from bidirectional to unidirectional, and the speed of extrusion changes from V_0 to $\frac{V_0}{2}$, which leads to a non-linear dependence of the MSAAD with time (Supplementary Figure S10A). For any given set of time series with the same extrusion initiation site, the MSAAD is bounded by R_0^2 and $d_1^2 = R_0^2 \frac{V_0(-t)}{2s_0}$ for times $t \leq -\frac{s_0}{V_0}$, whereas it is bounded by d_1^2 and $d_2^2 = R_0^2 \frac{V_0(-t)}{s_0}$ for times $t \geq -\frac{s_0}{V_0}$ (Supplementary Figure S10B).

We now denote as $\langle\langle R^2 \rangle\rangle(t)$ the ensemble average of the MSAAD (EMSAAD) over many time series, still assuming that $t = 0$ is the start of the closed state for all time series. The theoretical EMSAAD can be derived under the assumption that extrusion is initiated with uniform random probability between the two anchors. With this assumption, for times $t \leq -\frac{s_0}{V_0}$, the MSAAD equals R_0^2 with probability $p_1(t) = -\frac{3V_0 t}{4s_0} - \frac{1}{2}$ and equals d_1^2 with probability $1 - p_1(t)$ (Supplementary Figure S10C). Thus, for $t \leq -\frac{s_0}{V_0}$, we have: $\langle\langle R^2 \rangle\rangle(t) = p_1(t)R_0^2 + (1 - p_1(t))d_1^2$. For times $t \geq -\frac{s_0}{V_0}$, the MSAAD equals d_2^2 with probability $p_2(t) = -\frac{V_0 t}{4s_0}$ and equals d_1^2 with probability $1 - p_2(t)$ (Supplementary Figure S10C), hence $\langle\langle R^2 \rangle\rangle(t) = p_2(t)d_2^2 + (1 - p_2(t))d_1^2$. This leads to a complete, parameter-free model for the EMSAAD at all times:

$$\langle\langle R^2 \rangle\rangle(t) = \begin{cases} -R_0^2 \left(\frac{3V_0 t}{2s_0} \left(\frac{V_0 t}{4s_0} + 1 \right) + \frac{1}{2} \right) + R_{\text{loc}}^2 & \text{if } t < t_0 \\ R_0^2 \frac{V_0 t}{2s_0} \left(\frac{V_0 t}{4s_0} - 1 \right) + R_{\text{loc}}^2 & \text{if } t_0 \leq t \leq 0 \end{cases} \quad (4)$$

where $t_0 = -s_0/V_0$. Note that for $t = t_0$, the EMSAAD is: $\langle\langle R^2 \rangle\rangle(t = t_0) = \frac{5}{8} R_0^2 + R_{\text{loc}}^2$.

Analysis of simulated data

Simulated contact frequency maps. To predict chromatin contact frequency maps for comparison with Hi-C data, we used ensembles of simulated polymer conformations as follows. For simulations of 300 kb loops or larger, we generated 2500 independent simulations and randomly picked 80 polymer conformations per trajectory, yielding 200 000 single conformations in total. For 150 and 228 kb loops, we randomly picked 50 conformations from 4000 independent simulations, also resulting in 200 000 single conformations in total. Predicting a contact map from an ensemble of polymer conformations requires to define a capture radius (i.e. a threshold for the spatial distance between bead centers below which any pair of beads generates a contact event). To determine this radius, we first computed contact maps at 3 kb resolution for various radii, then calculated the average contact frequency along each diagonal of the contact map as function of genomic separation, $P(s)$. We compared simulated $P(s)$ curves to the experimental Hi-C counterpart over the range $s \in [5, 300]$ kb. We achieved the best match for a capture radius of 3 beads (150 nm) (Supplementary Figure S2B). This radius agrees with distance thresholds used in previous studies to reproduce Hi-C maps

from distances between pairs of loci measured by DNA FISH (30,31). As mentioned above (section ‘*Langevin dynamicssimulations*’ above), we also verified that changes in the stiffness parameter K_0 around its assumed default value of 0.1 did not strongly affect simulated $P(s)$ (Supplementary Figure S2D). Contact frequency maps were normalized by the total number of conformations.

Contact peak scores. In order to quantitatively compare the strength of loops in simulated contact maps and Hi-C data, we defined a peak score as the contact frequency between the two anchors divided by a background contact frequency (Figure 2B). For this purpose, the simulated contact maps and the selected Hi-C contact maps (see section ‘*Analysis of ChIP-Seq and Hi-C data*’ below), with genomic resolutions of 3 and 5 kb, respectively, were re-binned to a common resolution of 15 kb and were normalized to a sum of 1. The background frequency was defined as the average contact frequency inside a 30×30 kb (i.e. 2×2 bin) window located at 30 kb from each anchor and inside the loop domain (Figure 2B). Peak scores of simulated or experimental (Hi-C) contact frequency maps were compared using a one-sample t-test.

Quantification of loop extrusion detectability. The ability to detect loop extrusion from anchor–anchor distances measured in fluorescence microscopy images depends on different parameters that include loop size, extrusion speed, localization precision, reporter–anchor separation, fractions of loop states (open, extruding and closed) and the sample size. To assess the ability to detect loop extrusion for a given set of parameters, we randomly picked N anchor–anchor distances from 1000 independent simulation trajectories, each containing 3400 time points, with or without extrusion (Figure 3A). We compared cumulative distance distributions with and without extrusion ($F(x)$ and $G(x)$, respectively) using one-sided Kolmogorov–Smirnov tests (with the alternative hypothesis $F(x) < G(x)$). This test was repeated on 5000 independent random samples of the two distance distributions. We defined detectability of loop extrusion as the percentage of Kolmogorov–Smirnov tests revealing significant differences ($P < 0.05$ after Benjamini–Hochberg correction) (53) (Figure 3B).

To mimic reporter–anchor separation (Figure 3C), we measured the distances between beads that were shifted relative to the beads representing loop anchors. To mimic localization errors (Figure 3D), we shifted the (x, y, z) coordinates of the tracked bead centers using random normally distributed displacements $(\delta x, \delta y, \delta z)$ of mean 0 and standard deviations $\sigma_x = \sigma_y = \sigma_{x,y}$ and $\sigma_z = 2\sigma_{x,y}$, respectively, where the factor 2 reflects the typical anisotropy due to the axial elongation of widefield or confocal point spread functions (54). To model different loop extrusion speeds (Figure 3H), we modified the time interval Δt after which new bonds were created during loop extrusion (see section ‘*Simulating loop extrusion*’ above).

Fitting theoretical distance models. In the Results section ‘*Estimating the fraction of loop states from static imaging data*’, we used the theoretical model from (Eq. 2) above to estimate the three state fractions A_{closed} , A_{open}

and $A_{\text{extruding}} = 1 - A_{\text{closed}} - A_{\text{open}}$ from simulated anchor–anchor vectors. We randomly picked N anchor–anchor vectors from 4000 independent simulation trajectories, each containing 3000 time points and fitted the analytical model to these data. Because the covariance of $P(\vec{R}) = P_x(\delta x)P_y(\delta y)P_z(\delta z)$ is zero, rather than fitting this 3D function to the anchor–anchor vectors \vec{R} , we simultaneously fitted the three 1D probability densities $P_x(\delta x)$, $P_y(\delta y)$ and $P_z(\delta z)$ (Eq. 2) to the three axial projections of anchor–anchor vectors $(\delta x, \delta y$ and $\delta z)$, respectively. We performed these fits using a Python script that employs the `curve_fit` function of the `scipy` package (55), with the three proportions initialized to 1/3 each ($A_{\text{closed}} = A_{\text{open}} = A_{\text{extruding}} = \frac{1}{3}$). Note that to perform this fit we assumed the values of σ_{open}^2 and σ_{closed}^2 defined in section ‘*Anchor–anchor vector distributions*’ above to be known (see Discussion).

Segmentation of time series into closed states. In the Results section ‘*Quantifying closed loop lifetimes from live-cell trajectories*’, we analyzed time series of anchor–anchor distances to estimate the duration of closed states. To do this, we segmented these time series into intervals of (inferred) closed states based on a spatial and a temporal threshold (Supplementary Figure S8A). Our procedure to detect closed states in simulated time series of anchor–anchor distances as function of time is as follows. First, we defined the spatial threshold as the 99.9% quantile of the anchor–anchor distances in the closed state (assumed to be known independently, see Discussion). The anchor–anchor distance in the closed state fluctuates around the diameter of 1 bead (50 nm). This is similar to the ~ 40 nm diameter of the cohesin ring (56,57), which may enclose the two anchors in the closed state (17,26,58,59). Second, from time series obtained in simulations without extrusion, we measured the duration of time intervals during which the anchor–anchor distance was always below the spatial threshold (ignoring intervals reduced to a single time point). We then defined the temporal threshold as the 99.9% quantile of these time intervals. Within a time series, all time intervals with anchor–anchor distances below the spatial threshold and with durations exceeding this temporal threshold were segmented as closed state intervals. This resulted in binary time series with values of 1 for inferred closed states and 0 otherwise (i.e. for open or extruding states). In order to reduce the number of false negatives in closed state detections due to brief fluctuations of the distance above the spatial threshold, we then applied a rolling average to this binary time series (with a temporal window equal to the temporal threshold), and re-labelled as closed states all timepoints with values above 0.5.

Estimation of the mean closed state lifetime. In the Results section ‘*Quantifying closed loop lifetimes from live-cell trajectories*’, we analyzed the segmented time series to estimate the mean lifetime of closed states assuming an exponential distribution of closed states (Figure 5C). This was done by fitting a 2-parameter linear function to the logarithm of the histogram of the durations of segmented closed state intervals. In order to reduce the influence of spurious closed states, we used a robust least-squares fit combined with a ‘`soft_L1`’ loss function from `scipy` (55), and the `f_scale`

parameter set to 0.002. We assessed the quality of the fit with 4000 bootstrapped samples of 1000 simulations each, drawn from a total of 10 000 independent simulations (Supplementary Figure S9A).

Estimation of extrusion speed. In the Results section ‘Quantifying the speed of loop extrusion from live-cell trajectories’, we estimated the extrusion speed by fitting an analytical function to $\langle\langle R^2 \rangle\rangle(t)$, the ensemble average of the MSAAD (EMSAAD) over many time series, where $t = 0$ is the start of the closed state for all aligned time series. As described above (section ‘Mean anchor–anchor distance as function of time’), the theoretical EMSAAD follows a quadratic dependence with time, given by (Eq. 4). Nevertheless, for simplicity, and because the theoretical EMSAAD is close to a linear function for $t > -s_0/V_0$ (Supplementary Figure S10), we fitted a linear curve as function of time, whose slope defines an effective extrusion speed V_{eff} :

$$\langle\langle R^2 \rangle\rangle(t) \approx R_0^2 (V_{\text{eff}}/s_0)(-t) + R_{\text{loc}}^2 \quad (5)$$

After segmentation of closed state intervals (see section ‘Segmentation of time series into closed states’ above), we defined a window size of 150 s, which is the minimal duration of extrusion for a 150 kb loop with $V_0 = 1$ kb/s. In general, the minimal duration of extrusion is unknown, but it can be estimated in experimental data by choosing the window size with the highest estimated extrusion speed (see Results section ‘Quantifying the speed of loop extrusion from live-cell trajectories’; Supplementary Figure S9B). We removed all intervals shorter than this time window. We then computed the EMSAAD as function of time and fitted a linear function using least_square from *scipy* (55) (Figure 5D, Supplementary Figure S9B–F).

Analysis of experimental data

Analysis of ChIP-Seq and Hi-C data. For all genomic analyses, we used the hg19 reference genome. For ChIP-Seq data of CTCF, SMC1 and RAD21, we used publicly available data from Rao *et al.* (16). Raw reads were quality-checked with FastQC (60). Reads from different replicates were first mapped independently using bowtie2 v2.2.6.2 (61) with default parameters, and the correlation between replicates was computed using wigCorrelate (62). Replicates with correlations larger than 0.9 were pooled together and mapped again. We removed blacklisted regions (63) and called peaks using MACS2 v2.1 (64) with default parameters. CTCF motifs were identified genome-wide using FIMO (65) with the flags -max-stored-scores 50 000 000 and -thresh 0.001. We then mapped CTCF sites identified with a P -value $< 1 \times 10^{-5}$ onto CTCF ChIP-Seq peaks.

For Hi-C data of HCT-116 cells, we used publicly available contact maps at 5 kb resolution from Rao *et al.* (16). To call contact frequency peaks genome-wide, we used HICCUPS from Juicer 1.19.02 (66) with the flags -r 5000, 10 000 -k KR -f 0.1 -p 4,2 -I 7,5 -t 0.02,1.5,1.75,2 and -d 20 000, 20 000. ChIP-Seq peaks of CTCF, SMC1 and RAD21 were intersected with 20 kb regions centered around the loop anchors using pgltools (67) intersect1D. For comparison with simulation predictions, we selected experimental loops with at least one pair of CTCF-

bound convergent CTCF sites, SMC1 and RAD21 ChIP-Seq peaks at both anchors and genomic lengths $\pm 10\%$ of the simulated loop size.

From the experimental Hi-C maps at 5 kb resolution, we extracted the count matrices of selected loop domains using the dump function of Juicer 1.19.02 (66) (with the flags ‘Observed’ and ‘NONE’). We then converted them to full matrices using the sparse2full function of HiCcompare (68). To identify contact frequency peaks in experimental Hi-C maps re-binned at 15 kb, we searched for the coordinates of the maximum intensity value in a 75×75 kb box centered on peak locations identified by Juicer (66). We removed experimental Hi-C maps that had a peak score above 5 since they were found to originate from artifacts (either bins with unusually low signal in the $30 \text{ kb} \times 30 \text{ kb}$ background window or extremely high peak signal) and were outliers of the peak score distribution.

To build aggregated maps (Figure 2C), we extracted a 150×150 kb window around the loop anchors in each single Hi-C map. We summed these cropped Hi-C maps together and normalized by the number of maps used. We then binned the aggregated map to 15 kb resolution.

Analysis of imaging data. In the Results section ‘Comparing polymer simulations to imaging data’, we considered data from two experimental studies (Supplementary Figure S4). First, we used multiplexed DNA FISH data in mouse embryonic stem cells (mESC) from replicates 1 and 2 of Takei *et al.* (31). This dataset covered one genomic region on each of the 20 chromosomes at 25 kb resolution, each region comprising 60 consecutive loci and spanning at least 1.5 Mb, from which we extracted pairs of genomic loci. We removed from the analysis all cells where more than two identical chromosomes were detected. The pair-wise distances from all chromosomes were pooled together. Second, we considered live-cell imaging data from Gabriele *et al.* (35). From this study, we used data from the C36 clone without auxin treatment or with 4 h auxin treatment (which leads to RAD21 depletion).

From the simulations used for comparisons with these data, and for each considered genomic distance between pairs of loci, we computed the distributions of anchor–anchor distances assuming mean fractions of loop states consistent with experimental Hi-C data (as in Figure 2). We subsequently added random localization errors ($\sigma_{x,y} = 41$ nm, $\sigma_z = 65$ nm) consistent with Takei *et al.* (31).

RESULTS

Overview of simulation and analysis workflow

Our goal is to examine, using simulations, if, how and to what extent loop extrusion can be detected and quantitatively characterized by imaging two fluorescently labeled loop anchors. We considered two types of experiments: (i) imaging fixed or living cells at a single time point (hereafter called static imaging), which provides a snapshot of anchor–anchor distances in a number of single cells and (ii) dynamic imaging of living cells, where the two anchors can be tracked and the anchor–anchor distance can be measured and followed as function of time (Figure 1A). Crucially, we assumed that extrusion can be experimentally abolished,

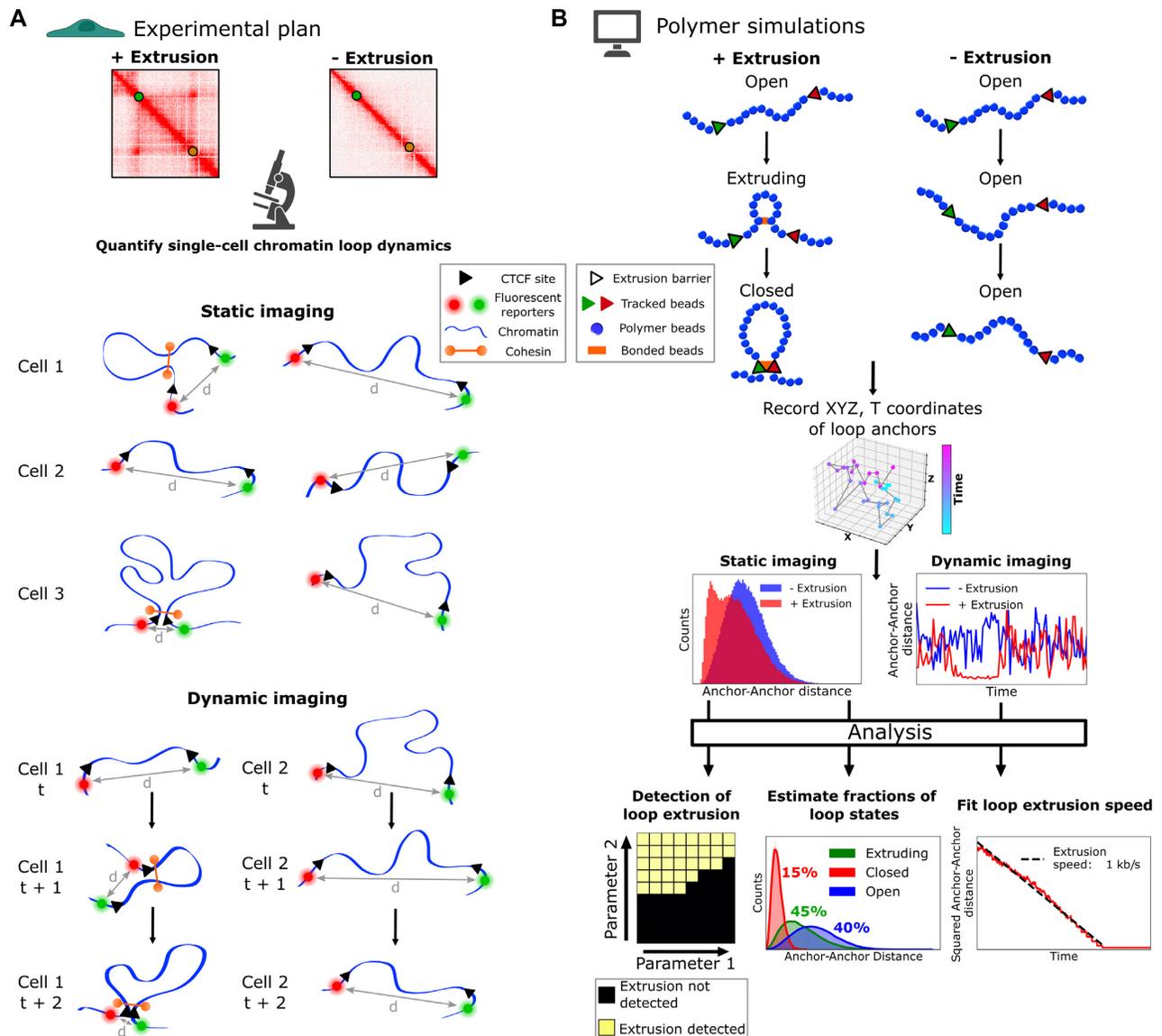


Figure 1. Overview of simulation and data analysis framework to characterize loop extrusion from imaging data. (A) Schematic of the assumed experimental plan. We assumed that 3D imaging data are obtained from cells in which extrusion can be abolished, e.g. by depleting cohesin. The two anchors of a loop are fluorescently labeled in two different colors (here, red and green) and imaged either at a fixed time point (static imaging) or tracked using time-lapse microscopy (dynamic imaging). This allows to determine the approximate 3D coordinates of the two anchors and the anchor–anchor distances (d) in a population of cells, either as a single snapshot or as function of time in single cells. Hi-C data shown on top are from Rao *et al.* (16). (B) Polymer simulations were used to model the dynamics of a generic chromatin segment with or without loop extrusion. The chromatin loop was assumed to transition between three different states, in the following order: open (the polymer is free of loops), extruding (the loop is actively extended) and closed (the loop is fully formed and the two anchors are kept close to each other). We used these simulations to generate 3D coordinates of each anchor as function of time (trajectories in the middle). From these trajectories, we obtained simulated distributions of anchor–anchor distances in cell population snapshots (static imaging) or simulated time series of anchor–anchor distances (dynamic imaging), either in absence (blue) or presence (red) of loop extrusion. We then used dedicated analysis methods to detect the presence of extrusion (bottom left), to estimate the fractions of cells in each of the three loop states (bottom center), or to determine the closed state lifetime and the speed of extrusion (bottom right). Based on the simulations, we determined the ranges of biological and experimental parameters (here noted generically parameters 1 and 2) for which extrusion can or cannot be detected (yellow and black, respectively) or for which loop state fractions or extrusion speed can be determined.

e.g. using an auxin-dependent degron (69,70) fused to RAD21, thus providing an essential negative control without loops in the genome (16) (Figure 1A).

Our analysis comprises the following steps, sketched in Figure 1B. First, we simulated the dynamics of polymers in presence or absence of loop extrusion and assuming various biological parameters, such as the loop size or the speed

of loop extrusion. We supposed that the loop can be found in three different states hereafter called ‘open’, ‘extruding’ and ‘closed’. In the open state, the chromatin segment delimited by the two anchors is free of any loop, e.g. because no extrusion complex is bound to it; in the extruding state, a loop located in between the two anchors is progressively enlarged by the action of an extrusion complex; in the closed

state, the two anchors are brought in close proximity and the loop is temporarily stabilized (the extrusion complex maintains the loop anchors together). Second, we compared simulated contact maps to available Hi-C data in order to determine plausible ranges for the fractions of loops in each state (open, extruding or closed) within a population of cells. Third, we used the same simulations to predict distributions of anchor–anchor distances for different biological and experimental parameters. We then used a statistical test to determine (for each set of parameters), whether and to what extent the presence of loop extrusion can be detected from distance distributions. Fourth, we proposed and tested a method to estimate the fractions of the open, closed, or extruding states based on distance distributions from static imaging data. Fifth, we analyzed how live-cell anchor trajectories can be used to estimate the duration (lifetime) of closed states and the speed of loop extrusion. In the following, we detail each of these steps and determine the conditions under which extrusion can be detected and quantified.

Simulating polymer dynamics and loop extrusion

In the following, we considered a generic, 1.8 Mb long chromatin segment centered around two loop anchors separated by 300 kb (unless otherwise stated). We modelled this chromatin segment as a polymer chain made of 600 beads linked by harmonic bonds and subjected to Langevin dynamics. We verified that mean squared displacements (MSD) of an internal bead in a non-extruding polymer (i.e. in the ‘open’ state) obeyed the subdiffusive power law expected from Rouse dynamics (MSD proportional to the square root of time) at small time scales and saturated at long time scales due to confinement (Supplementary Figure S1C).

To model loop extrusion, we assumed that a single loop extrusion complex was loaded on the polymer at a random position between the two anchors, whereafter both sides of the polymer chain were extruded simultaneously (bidirectional extrusion (26,71)). Unless stated otherwise, we assumed that extrusion began at $t = 333$ s after the start of the recorded simulation trajectories and proceeded at a nominal speed of $V_0 = 1$ kb/s (17). As soon as the extrusion complex reached one of the two anchors, it stopped, and the other side of the chain continued to be extruded unidirectionally (18,50,51), at a speed of $V_0/2$ (26), until the extrusion complex reached the second anchor (Supplementary Figure S2A). At this point, extrusion stopped entirely, and the two anchors were maintained in the closed state until the end of the simulation, unless stated otherwise. For any given set of biological or experimental parameters, we ran at least 1000 independent simulation trajectories, with ≈ 12 million time steps each. From these simulations, we directly extracted anchor coordinates as function of time, mimicking live-cell tracking data, or computed the distributions of anchor–anchor distances (and anchor–anchor vectors) in a random ensemble of configurations, mimicking population snapshots taken at a single time point.

Comparing polymer simulations to Hi-C data

Since TADs and loops were primarily defined from Hi-C data (4,5,16), realistic simulations must be able to account

for these features in simulated contact frequency maps. In order to generate contact frequency maps from simulations, we used a capture radius of 150 nm, which provided the best match between simulated and experimental $P(s)$ curves (Supplementary Figure S2B). We first verified that our simulations, in which extrusion was assumed to start at random genomic locations between the anchors, can qualitatively reproduce blocks of enriched contacts with corner peaks on the contact map diagonal, consistent with TADs and loops (Supplementary Figure S2C). Note that if extrusion was instead initiated at one of the two anchors or at the mid-point between them, the contact maps exhibited different features, namely stripes (18,52,72), and hairpins (73) or jets (50), respectively (Supplementary Figure S2C). However, we excluded such structures from our analyses to focus on canonical TADs and loops instead.

The strength of contact blocks and/or corner peaks in Hi-C data is expected to depend on the fractions of cells in which the anchors are in an open, closed, or extruding state. Obviously, these fractions also strongly determine the dynamics of tracked anchors and the expected distribution of anchor–anchor distances, and therefore impact the feasibility of detecting and/or characterizing loop extrusion (37). For example, a high fraction of cells in a closed state could give rise to a bimodal distribution of anchor–anchor distances (one mode at small distances corresponding to closed states and a second mode at larger distances corresponding to open or extruding states), whereas a low fraction might not. These fractions are poorly known - one aim of imaging-based approaches is precisely to estimate them. Nonetheless, to determine a realistic range of state fractions, we proceeded to vary them in our polymer simulations and compared simulated contact maps to a selection of experimental loops in Hi-C data (Figure 2, Supplementary Figures S2E, F and S3). Below, we describe (i) the Hi-C data set considered, (ii) the simulations and the range of parameters explored, (iii) the metric we used to compare data to simulation predictions and (iv) the results of this comparison.

We started from the Hi-C data (16) of RAD21-mAID-GFP tagged HCT-116 cells, in which addition of auxin depletes RAD21 (69). We computationally identified 4470 loops in the 5 kb resolution Hi-C maps of cells without auxin treatment. We then narrowed down this data set to 229 loops that satisfied the following criteria: (i) a size of 300 ± 30 kb, for consistency with the 300 kb size of simulated loops, (ii) the presence of ChIP-Seq peaks of the cohesin subunits SMC1 and RAD21, as well as ChIP-Seq peaks of CTCF at both loop anchors, and (iii) at least one pair of bound convergent CTCF sites at the anchors (Figure 2A). These selection criteria were intended to focus the analysis on loops that depend on cohesin-mediated loop extrusion stabilized by CTCF, rather than on other biological mechanisms (10,74).

Using 200 000 single polymer conformation snapshots from 2500 independent simulations, we built 231 distinct contact maps, each corresponding to a different combination of the open, extruding or closed state fractions (Supplementary Figure S3). The extreme cases, where all conformation snapshots are from the same state, showed the expected behavior: for 100% open states, the contact map displayed no specific structure, and only reflected the uniform decay

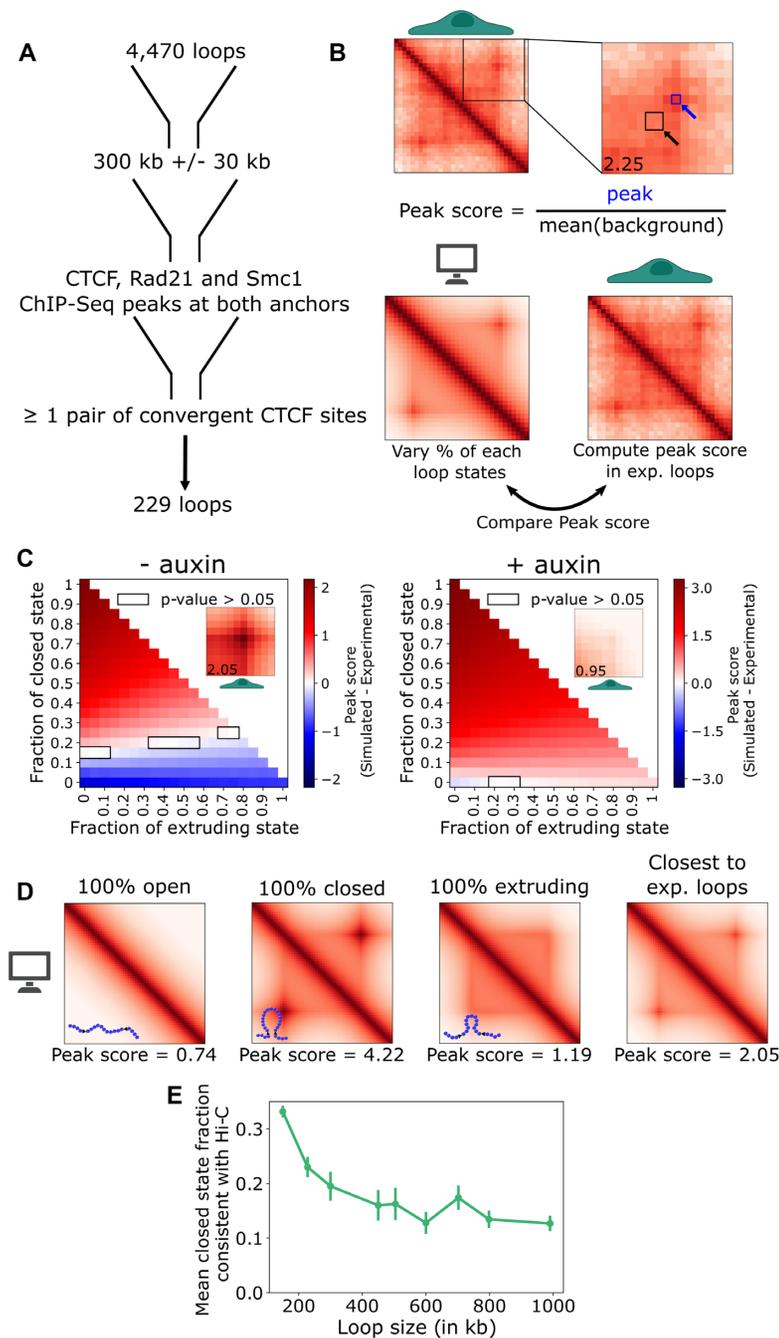


Figure 2. Estimation of loop state fractions consistent with experimental Hi-C data. **(A)** Starting from 4,470 loops detected in Hi-C data of HCT-116 cells from Rao et al (16), we selected a subset of 229 loops with sizes of 300 ± 30 kb, with ChIP-seq peaks of CTCF, RAD21 and SMC1 at both anchors and at least 1 pair of CTCF-bound convergent CTCF sites. **(B)** Top: For each of these loops, a peak score was computed as the ratio of the peak Hi-C contact frequency between the two anchors (blue square) and the background contact frequency (average of the black square). Bottom: Peak scores were used to compare simulated contact frequency maps (for varying fractions of each loop state) to experimental Hi-C maps. **(C)** Heat maps show differences in peak scores between simulated and experimental contact maps for each combination of extruding, closed and open state fractions, for 300 ± 30 kb loops. The cells were subjected or not to a 6 h auxin treatment to deplete the cohesin subunit RAD21 (16). Red and blue colors indicate a positive or negative difference between the simulated and the experimental mean peak scores, respectively, whereas white indicates similar peak scores. The open state fraction is: $A_{\text{open}} = 1 - A_{\text{closed}} - A_{\text{extruding}}$ where A_{closed} and $A_{\text{extruding}}$ are the fractions of closed and extruding states, respectively. The black contour denotes state fractions for which the peak scores of simulated and experimental loops were indistinguishable (P -value > 0.05 for a one sample t-test). Insets show aggregated maps of the experimental Hi-C data centered on the loop anchors and their mean peak score is displayed on the bottom left. The same display range was used for all aggregated maps. **(D)** Simulated contact maps for selected fractions of loop states. The first three maps from the left show extreme cases corresponding to 100% open, 100% closed or 100% extruding states. The fourth map corresponds to the state fractions that achieved the best match with the experimental Hi-C data (40% open, 40% extruding and 20% closed). The peak score for each simulated map is indicated below. All contact maps are shown with the same display range. Simulated and experimental Hi-C maps in panels B–D are shown at 15 kb resolution. **(E)** Mean fraction of closed states in simulations that are consistent with Hi-C data (P -values > 0.05) for loops of different genomic sizes. Error bars indicate the 95% confidence interval.

of contact frequencies as function of genomic distance, $P(s)$; for 100% closed states, the contact map exhibited a strong peak at the corner of a moderately intense contact block; for 100% extruding states, the map showed a block of enhanced contact frequencies delimited by the anchors, without a corner peak (Figure 2D and Supplementary Figure S3).

For a more quantitative comparison of simulation predictions with Hi-C data, we computed a ‘peak score’ in individual contact maps. The peak score was defined as the contact frequency between the two anchors divided by the contact frequency inside a 30×30 kb window located between the two anchors (Figure 2B). A higher fraction of closed states led to a high peak score, while a higher fraction of extruding states led to a higher background and thus a lower peak score (Figure 2D). We then systematically compared simulated and experimental peak scores for fractions of open, extruding and closed states ranging from 0 to 100%. For this purpose, we first verified the consistency of our metric with experimental findings that cohesin depletion leads to elimination of chromatin loops using Hi-C maps of auxin-treated cells (16). As expected, only simulations without closed states and with a large majority of open states (>70–80%) had peak scores statistically consistent with the Hi-C data in these cells (Figure 2C, right; $P > 0.05$).

We then used this peak score to compare simulation predictions to Hi-C maps in cells without auxin treatment. We found that simulation predictions were consistent with the Hi-C data for a range of closed state fractions of 15–25% and a large range of extruding state fractions (Figure 2C, left; $P > 0.05$). Within these ranges, the best match between simulations and Hi-C data was obtained for 20% closed states, 40% extruding states and 40% open loop states (Figure 2C, D). Simulations with 100% closed states were inconsistent with the experiments, leading to peak scores two times higher than in the Hi-C data (Figure 2D, $P < 10^{-150}$), consistent with previous studies arguing against the existence of stable loops in cells (20,33,75,76). However, our metric did not enable us to determine tight constraints on the fractions of open versus extruding states, as a wide range of fractions for each of these two states was compatible with the Hi-C data (Figure 2C and Supplementary Figure S2E).

To account for differences in the genomic sizes of loops, we then repeated the above simulations and analyses for loop sizes ranging from 150 to 990 kb. We found that the mean fraction of closed states consistent with the Hi-C data was larger for small loops than for large loops (33% for 150 kb loops, versus 16% for 504 kb loops and 13% for 990 kb loops) (Figure 2E and Supplementary S2F).

Thus, we used Hi-C data to estimate plausible ranges for the fractions of open, closed and extruding states in our polymer simulations. This allowed us to subsequently simulate loop extrusion dynamics and imaging observations under realistic conditions.

Comparing polymer simulations to imaging data

Having determined the fractions of loop states consistent with experimental Hi-C data, we next compared our simulations to published measurements of distances between chromatin loci from experimental imaging data. We first

compared simulations of 504 kb loops with or without extrusion to recent live-cell imaging data from a study by Gabriele *et al.* (35) that specifically targeted a pair of loop anchors in mouse embryonic stem cells (mESC) with and without cohesin and found good agreement between experiments and simulations in both cases (Supplementary Figure S4). Next, we compared our simulations to a larger experimental data set where DNA FISH was used to measure distances between pairs of loci in mESC for a large range of genomic distances between these loci (31). These experiments did not specifically target loop anchors, but loops were nonetheless present in the targeted genomic regions. As a result, we expected anchor–anchor distances predicted by our simulations with and without extrusion to bracket the distances measured by DNA-FISH. This was indeed the case, and distances predicted by simulations with extrusion were slightly lower than the experimentally measured distances and closely followed the predicted trend as function of genomic distance (Supplementary Figure S4). These comparisons indicated that our simulations were also consistent with imaging data in fixed and living cells.

Detecting chromatin loop extrusion from imaging data

Chromatin loop extrusion is expected to reduce anchor–anchor distances compared to conditions without extrusion. We therefore asked whether loop extrusion can be detected simply by measuring anchor–anchor distances in populations of single cells and comparing them to distances measured when loop extrusion is abolished (Figure 3A). We aimed to assess the ability to reveal loop extrusion (hereafter called ‘detectability’ and defined quantitatively below) depending on several relevant biophysical and experimental parameters. The biological parameters were: the fractions of loop states (open vs. extruding vs. closed), the genomic size of loops, and the speed of loop extrusion. The experimental parameters were: the reporter–anchor separation, which we varied from 0 to 300 kb, the localization precision of fluorescent labels, defined as the standard deviation of errors in estimated coordinates, which we varied from $\sigma_{x,y} = 0$ nm to $\sigma_{x,y} = 200$ nm laterally (with an axial error $\sigma_z = 2\sigma_{x,y}$ due to the anisotropy of the microscope point spread function (54)) and the number of analyzed anchor–anchor distances, which we varied from $N = 10$ to 10 000.

In order to determine whether extrusion can be detected, we compared the distributions of anchor–anchor distances in simulations with and without extrusion, thereby mimicking the experimental comparison of cells with and without cohesin. For each pair of simulated distance distributions, we performed Kolmogorov–Smirnov tests on thousands of random samples. We then defined the detectability of loop extrusion as the fraction of tests that revealed statistically significant differences between these distributions (Figure 3B).

To explore the effect of the different parameters above on the detectability of extrusion (Figure 3C–I), we varied two parameters at a time, holding the other parameters constant. We used the following state fractions as defaults, based on the above comparison to Hi-C data: 20% closed, 40% extruding, 40% open (Figure 2C). Other

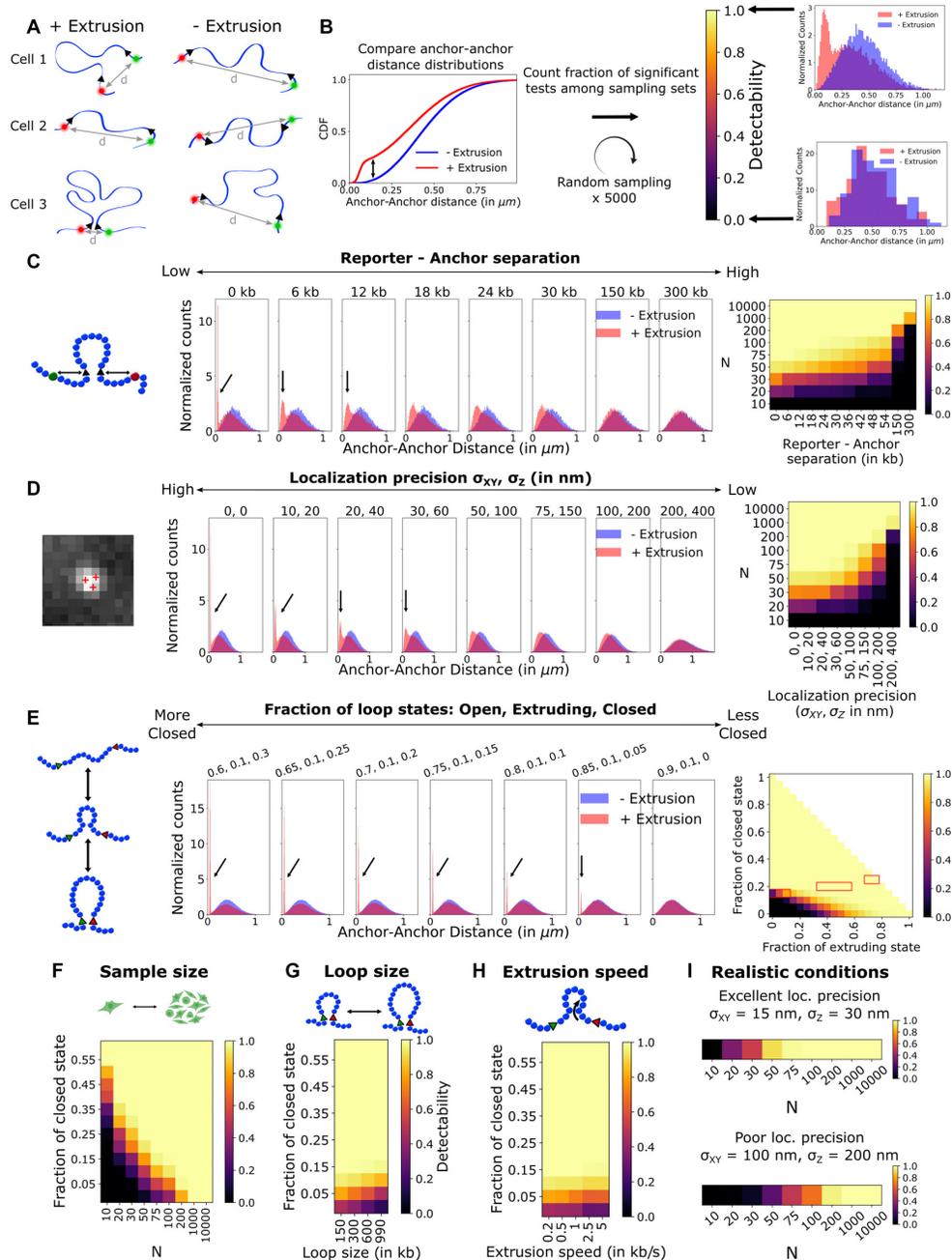


Figure 3. Experimental and biological parameters allowing to detect loop extrusion. (A) Polymer simulations assumed that distances between fluorescently labeled anchors are measured in a population of cells at a given time, both in conditions allowing extrusion, and in a condition where extrusion is experimentally abolished. (B) Cumulative distribution functions (CDFs) of anchor–anchor distances sampled from simulations with (red) or without (blue) extrusion were compared using a Kolmogorov–Smirnov test. Sampling was repeated 5000 times, P -values were corrected for multiple testing with the Benjamini–Hochberg procedure. Detectability of loop extrusion was defined as the fraction of tests revealing significant differences ($P < 0.05$) and is displayed on a color scale from black (no significant difference) to yellow (all tests show significant differences). Histograms show examples of distance distributions for these extreme cases of low and high detectability. (C–H) Effect of experimental or biological parameters on the detectability of loop extrusion. Heat maps show detectability as function of various parameters, including: the reporter–anchor separation (C), the localization precision of fluorescent reporters $\sigma_{x,y} = \sigma_z/2$ (D), the fractions of closed and extruding states (E), the sample size (i.e. the number of distance measurements), N (F), the genomic size of the loop (G) and the speed of loop extrusion (H). Two parameters were varied at a time, with the other parameters set to the following default values: fractions of loop states: 40% open, 40% extruding, 20% closed; loop size: 300 kb; extrusion speed: 1 kb/s; reporter–anchor separation: 0 kb; localization precision: $\sigma_{x,y} = \sigma_z = 0$ nm; sample size: $N = 100$. Distributions of anchor–anchor distances with and without extrusion (red and blue, respectively) and for different parameter values are shown in (C)–(E) (note that we used $N = 100\,000$ to plot these distributions for better legibility, irrespective of the sample size N used for the detectability heat maps). The arrows show the peaks at small distances due to closed states. Note that ‘anchor–anchor distance’ designates the distance between the corresponding reporters (which differ from the actual anchors for panel C). In the heat map of panel (E), the red contour denotes loop state fractions consistent with experimental Hi-C data from Figure 2C (without auxin treatment). (I) Detectability of loop extrusion in realistic experimental conditions with: reporter–anchor separation of $2 \times 3 = 6$ kb and excellent ($\sigma_{x,y} = \sigma_z/2 = 15$ nm, top) or poor ($\sigma_{x,y} = \sigma_z/2 = 100$ nm, bottom) localization precision. Other parameters were set to the default values above.

default parameters were as follows: loop size: 300 kb; extrusion speed: 1 kb/s, based on measurements of DNA loop extrusion *in vitro* (17,26); localization precision $\sigma_{x,y} = \sigma_z = 0$ nm; reporter-anchor separation = 0 kb (corresponding to a direct labeling of the anchors); sample size: $N = 100$ measurements (anchor–anchor distances).

We first analyzed the effect of the reporter-anchor separation (Figure 3C). In the ideal case where this separation is zero (i.e. the anchors themselves are fluorescently labeled), the clearest signature of extrusion was an additional peak in the anchor–anchor distance distribution at very low distances due to closed states (Figure 3C, arrows). In this case, extrusion detectability was high (95%) for a sample size N as low as 50. As expected, increasing the reporter-anchor separation tended to reduce and broaden the low distance peak, and for a total separation of 18 kb or more the distribution became unimodal again (Figure 3C). Consistently, increasing the reporter-anchor separation lowered detectability when the sample size N was kept constant. However, even for reporter-anchor separations of 54 kb, extrusion remained detectable using $N > 100$ measurements (detectability > 92%). Increasing the random localization errors had a qualitatively similar effect as increasing the reporter-anchor separation (Figure 3D). Nevertheless, even for poor localization precisions of $\sigma_{x,y} = \frac{\sigma_z}{2} = 100$ nm, detectability was high (>97%) as long as the sample size N exceeded 200 (Figure 3D).

Obviously, these results strongly depended on the assumed fractions of closed, extruding and open states. For example, for a fixed extruding state fraction of 10%, it became more challenging to detect extrusion from $N = 100$ measurements when closed state fractions were lower than 20% (Figure 3E). Interestingly, increasing the sample size to $N = 1000$ measurements allowed extrusion to be detected for almost all combinations of state fractions, except in the quasi-absence of closed states combined with an extremely high fraction of open states (>80%) (Supplementary Figure S5A).

We then assessed the effects of sample size, loop size and extrusion speed, assuming 40% of extruding states as based on the above analysis of Hi-C data. A larger sample size always helped to detect extrusion in unfavorable conditions. For example, for a small closed state fraction of 5%, increasing the sample size from $N = 100$ to $N = 1000$ improved detectability from 70% to 100% (Figure 3F). The loop size, which we varied between 150 and 900 kb, had only a marginal effect on detectability (Figure 3G). Similarly, varying the speed of extrusion in the range 0.2–5 kb/s, which encompasses extrusion speeds measured *in vitro* (17,26), did not appreciably affect the detectability of loop extrusion (Figure 3H).

Next, we considered the detectability of extrusion as function of localization precision and sample size for realistic values of all other parameters combined (Figure 3I). Specifically, we assumed 20% closed, 40% extruding and 40% open states; a loop size of 300 kb; an extrusion speed of 1 kb/s; and a total reporter-anchor separation of 6 kb. We tested excellent, good and poor localization precisions ($\sigma_{x,y} = \frac{\sigma_z}{2} = 15, 50$ and 100 nm (31,32,77), respectively). The presence of closed loops was identifiable as a low distance peak in the distance histogram for excellent localization pre-

cision, but not for good or poor precision (Supplementary Figure S5B). Nevertheless, even for poor precision, loop extrusion detectability was still high (96%) for a reasonable sample size of $N = 200$ (Figure 3I).

Finally, we performed simulations using experimental and biological parameters approximating those of the above-mentioned Gabriele *et al.* study (35), which reported the identification of loop extrusion from anchor–anchor distances measured using live-cell imaging. Based on these simulations, we found loop extrusion to be detectable at 97% with as little as $N = 100$ measurements, while at least 45 000 distances were measured in (35) (Supplementary Figure S6). Hence, this study's experimental approach indeed matches the conditions defined by our theoretical analysis to detect loop extrusion from imaging data.

In summary, our analysis suggests that under a relatively wide range of realistic biological and experimental conditions, the presence of loop extrusion can be detected in cells by imaging pairs of loop anchors, measuring their distances and comparing them to distances measured in cells where extrusion is abolished. The above results also define the minimal experimental parameters that are required (and satisfied by Gabriele *et al.* (35)) to detect the presence of loop extrusion depending on the fractions of different loop states.

Estimating the fraction of loop states from static imaging data

Next, we aimed to move beyond the mere detection of loop extrusion and assess if static imaging data from a population of single cells allows to quantitatively estimate the fractions of loop states (closed vs. extruding vs. open) (Figure 4). We approached this by fitting to the data a 'three-state' mathematical polymer model that includes these fractions as adjustable parameters. Specifically, we considered the differences in x , y and z coordinates between the two anchors ($\delta x = x_1 - x_2$; $\delta y = y_1 - y_2$; $\delta z = z_1 - z_2$) and fitted a model, $P(\delta x, \delta y, \delta z) = P_x(\delta x)P_y(\delta y)P_z(\delta z)$ to the three distributions of coordinate differences ($\delta x, \delta y, \delta z$) simultaneously. For each coordinate difference $\delta w \in \{\delta x, \delta y, \delta z\}$, $P_w(\delta w)$ is a linear combination of three models, P_w^{open} , P_w^{closed} and $P_w^{\text{extruding}}$, one for each loop state (open, closed and extruding, respectively), weighted by the fractions of each state (A_{open} , A_{closed} , and $A_{\text{extruding}} = 1 - A_{\text{closed}} - A_{\text{open}}$), and is given by (Eq. 2). For the open state, we assumed that the polymer behaves as a freely jointed chain and each coordinate difference $\delta w \in \{\delta x, \delta y, \delta z\}$ follows a normal distribution $P_w^{\text{open}} = \mathcal{N}(0, \frac{N_0 b^2}{3} + 2\sigma_w^2)$, where b is the Kuhn length, N_0 is the number of Kuhn lengths between the anchors, and σ_w is the localization precision along axis $w \in \{x, y, z\}$. For the closed state, where the two anchors are in contact or in close proximity, we assumed that each coordinate difference follows a normal distribution $P_w^{\text{closed}}(\delta w) = \mathcal{N}(0, \sigma_c^2 + 2\sigma_w^2)$, where the added variance σ_c^2 reflects a possible non-zero distance between the anchors in the closed state (e.g. because of the ~40 nm size of the cohesin ring (56,57)). We assumed that the probability distributions P_w^{open} and P_w^{closed} can be determined experimentally (see Discussion).

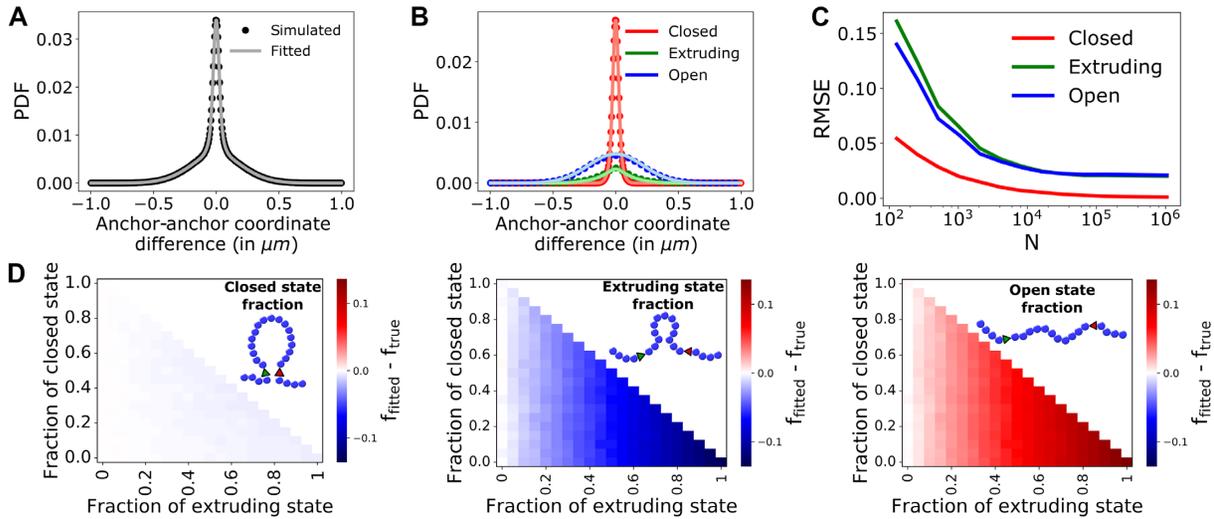


Figure 4. Quantifying fractions of loop states from static images. (A) Total distribution of the differences δw between coordinates of the two anchors ($w \in \{x, y, z\}$) as obtained from simulated data (black dots) and the fitted three-state model (grey curve, Eq. 2). (B) Same as A, but with the three fitted components of the three-state model (corresponding to the closed, extruding and open states) shown separately. Dots show the simulated data and colored curves show the fitted model components. (C, D) Quantification of errors in estimation of loop state fractions. (C) Root mean square error (RMSE) of estimated fractions for the three states as function of the sample size (number of measured anchor–anchor vectors) N . The simulations used for panels (A), (B) and (C) assumed the following loop state fractions: 35% closed, 15% extruding and 50% open. (D) Heat maps show the signed error (bias) in the estimated fractions of closed, extruding or open states, for each combination of loop state fractions, estimated for $N = 500\,000$. Overestimates are shown in red, underestimates in blue. The absolute bias did not exceed 13.1%. All simulations used in this Figure assume a localization precision $\sigma_{x,y} = \sigma_z/2 = 15$ nm.

For the extruding state, we assumed that the coordinate differences follow the probability density $P_w^{\text{extruding}}(\delta w; t) = \mathcal{N}(0, \frac{N_b(t)b^2}{3} + 2\sigma_w^2)$, with $N_b(t) = N_0 - N_{\text{loop}}(t)$, where $t = 0$ is the end of loop extrusion (during extrusion $t < 0$), and $N_{\text{loop}}(t)$ is the number of Kuhn lengths extruded at time t . This assumption effectively derives from the hypotheses that (i) the anchors in the extruding state behave as if part of a shorter polymer from which the extruded loop is absent and (ii) this polymer has time to equilibrate at each step of the extrusion process. In the Materials and Methods section, we derived an explicit model for $P_w^{\text{extruding}}(\delta w; t)$ (Eq. 1) based on the additional assumptions that extrusion initiates with uniformly random probability between the two anchors and switches from bidirectional at constant speed V_0 to unidirectional at constant speed $\frac{V_0}{2}$ upon reaching an anchor. Having specified these models for P_w^{open} , P_w^{closed} and $P_w^{\text{extruding}}$, we could estimate the fractions of each loop state (open, closed and extruding) by fitting the three-state model of (Eq. 2) to the measured distributions of coordinate differences δx , δy and δz .

To test this method, we first simulated coordinates of anchor pairs from 4000 independent simulations assuming 35% closed, 15% extruding and 50% open states (based on the analysis of Hi-C data in Supplementary Figure S2F) for a 150 kb loop, with localization errors $\sigma_{x,y} = \sigma_z/2 = 15$ nm (29,31,32), drew $N = 500\,000$ samples (anchor–anchor vectors) and fitted our three-state model (Eq. 2) to these data (Figure 4A–C). As before, these simulations assumed that extrusion initiates with uniformly random probability between the two anchors, and that the extrusion speed is halved once the first anchor is reached. As shown in Figure 4A and B, the fitted model was in very good

agreement with the simulated data. We computed a root mean square error (RMSE) of estimated fractions A_{closed} , A_{open} and $A_{\text{extruding}}$ for varying sample size N (Figure 4C). For $N = 100$, the RMSE for open and extruding states remained substantial (14% and 16%, respectively), implying that larger sample sizes are needed for accurate estimation of loop state fractions. For $N > 5000$, the RMSE dropped to $<3\%$ for all three states, and decreased further for larger N . Open and extruding states still had RMSE $<2.5\%$ even for large sample sizes (Figure 4C), a residual bias that presumably reflects a minor mismatch between the analytical model and the simulations. We then systematically varied the fractions of open, extruding and closed states and quantified the absolute biases of the estimated fractions for a sample size of $N = 500\,000$ (Figure 4D). The closed state fraction was always very accurately estimated, with absolute bias $<0.5\%$. The extruding state tended to be slightly underestimated, while the open state was slightly overestimated, but the absolute bias always remained at an acceptable level ($<13.1\%$) (Figure 4D).

We repeated this analysis but changed the localization errors to either $\sigma_{x,y} = \sigma_z = 0$ or $\sigma_{x,y} = \sigma_z/2 = 50$ nm (Supplementary Figure S7). Paradoxically, the state fractions were estimated more accurately for larger localization errors (maximum absolute bias of 7.4%) than without localization errors (maximum absolute bias of 17.7%) (Supplementary Figure S7D and H), likely because the Gaussian localization errors reduce the mismatch between the Gaussian model and the simulations in the closed state. Nevertheless, for sample sizes $N > 8000$, the RMSE were $<5\%$ both without localization errors and for large localization errors, assuming 35% closed, 15% extruding and 50% open states (Supplementary Figure S7C and G).

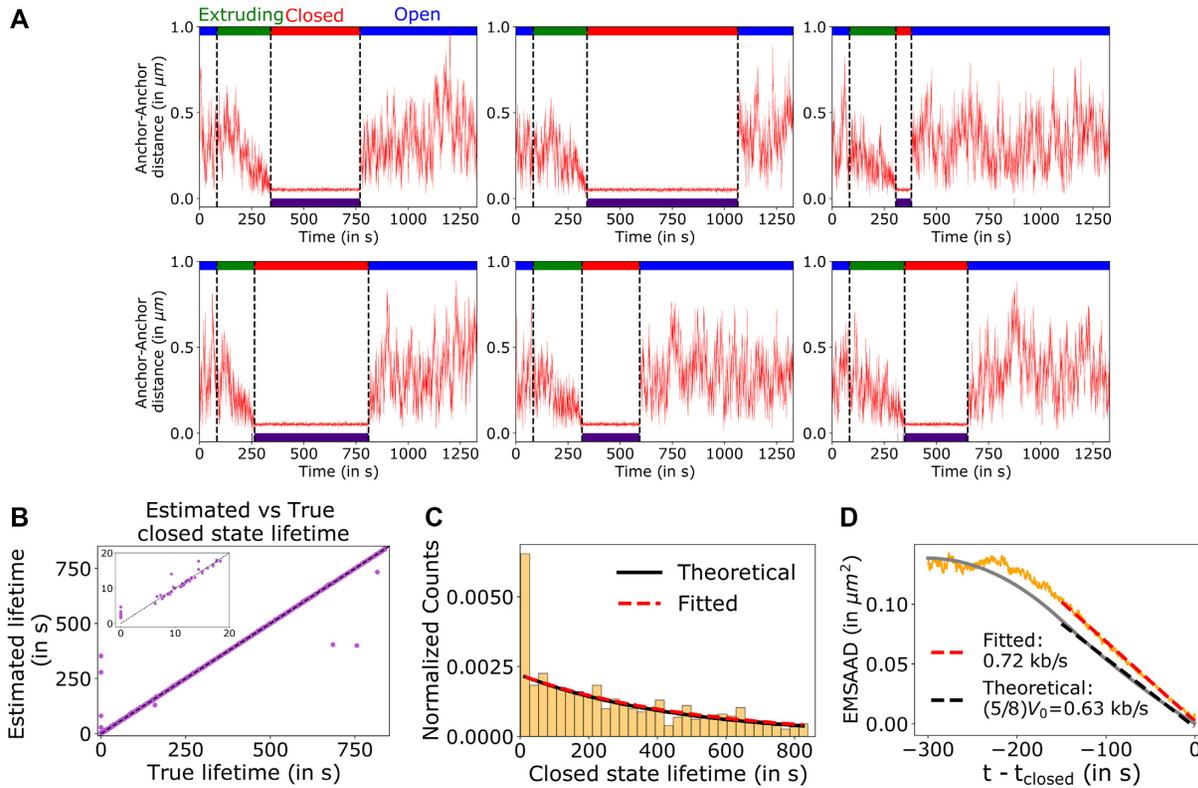


Figure 5. Quantifying the lifetime of closed states and the speed of loop extrusion by dynamic tracking of loop anchors. (A) Six example time series of anchor–anchor distances from simulations with loop extrusion. The true loop state is indicated by color bars on top (blue: open, green: extruding, red: closed). Black dotted lines indicate state changes, i.e. when loop extrusion starts or ends or when closed loops are released. The segmentation of the time series in intervals of closed states is shown below (indigo bar indicates inferred closed state intervals). Each indigo dot corresponds to a distinct segmented closed state interval. Black dotted line denotes $y = x$. The inset shows a close-up for short lifetimes (0–20 s). (B) Estimated vs. true lifetimes of closed loops. (C) Histogram of 1000 estimated lifetimes of closed states. The dotted red line shows a robust exponential fit to the segmented closed state lifetimes and the solid black line shows the theoretical distribution. (D) Ensemble mean squared anchor–anchor distance (EMSAAD) as function of time, averaged over > 1000 time series, after aligning the estimated start time of the closed state to $t = 0$ (solid orange curve). An effective loop extrusion speed $\bar{V}_{\text{eff}} = 0.72$ kb/s was estimated by fitting a linear function for $-150 \text{ s} < t < 0$ (red dotted line). The theoretically expected EMSAAD (Eq. 4) is a non-linear function of time and is shown in solid grey. A linear fit to the theoretical EMSAAD (Eq. 5) is shown as a dashed black line and corresponds to a theoretical speed of 0.63 kb/s. Localization errors were ignored in the simulations used here.

To sum up, our analysis indicates that the fractions of loop states can be estimated accurately by measuring anchor–anchor distances from static imaging data of a sufficiently large number of cells.

Quantifying closed loop lifetimes from live-cell trajectories

While the analysis of static data can provide important quantitative information on loop states, as shown above, a full characterization of the dynamic aspects of loop extrusion, in particular extrusion speed and the lifetime of closed loops, requires dynamic data. Here and in the next section, we assess to what extent these parameters can be estimated from live-cell tracking of loop anchors.

For this purpose, we generated 1000 independent Langevin dynamics simulations of a polymer undergoing extrusion and recorded the trajectories of the two anchors as function of time. We considered a polymer with two loop anchors separated by 150 kb and simulated trajectories of 1330 s each. We assumed that extrusion started 83 s after the beginning of each simulation at a random position between the two anchors, and that extrusion proceeded until the loop was closed. The two anchors were then maintained in this

closed state during a time interval drawn from an exponential distribution, in order to mimic a memory-less (Poisson) process of loop dissociation, with a mean closed state lifetime of 466.6 s (and a maximum of 833 s). After dissociation, the two anchors were released from each other, and the polymer relaxed rapidly to the open state. Typical time series of the anchor–anchor distance during a simulation trajectory are shown in Figure 5A.

In order to quantify closed state lifetimes, we aimed to segment these time series into disjoint time intervals corresponding to closed, open or extruding states. As apparent from the examples in Figure 5A, it is difficult to distinguish the open state from the extruding state, especially in the early phases of the extrusion process where the anchor–anchor distances fluctuate widely. We therefore focused on segmenting the closed state only, without attempting to distinguish open from extruding states. Nonetheless, the distance distribution in the closed state can still significantly overlap that of the extruding and open states, since in these two states anchors can still come into close vicinity due solely to random polymer movements (Supplementary Figure S8A). For this reason, a distance threshold alone is ill suited to segmenting closed states (Supplementary Figure

S8A). We therefore decided to use a combination of spatial and temporal thresholds and identified closed states as long-lived small distance states that were not present in time series without extrusion (as measured in an experimental condition that abolishes extrusion). We defined small distances as those below the 99.9% quantile of distances in the closed state. We defined the temporal threshold as the 99.9% quantile of time intervals during which the anchor–anchor distance was always below the spatial threshold in time series of extrusion-free simulations (Supplementary Figure S8A).

Next, we evaluated this segmentation method on simulated time series with extrusion by quantifying the number of time points correctly or incorrectly classified as closed states. Using the above spatial and temporal thresholds, closed states were identified with a precision (ratio of true positives over all positives) of 99.75% and a recall (ratio of true positives over the sum of true positives and false negatives) of 99.87%, proving high classification quality for individual time points (Supplementary Figure S8B). Precision and recall could be traded-off against each other by varying the spatial and temporal thresholds (Supplementary Figure S8B). We then analyzed the performance of our segmentation method for a total reporter-anchor separation of 6 kb and different localization errors. Recall remained above 99% for localization errors $\sigma_{x,y} = \sigma_z/2 < 50$ nm, while segmentation precision decreased when increasing localization errors (from 96% for $\sigma_{x,y} = \sigma_z = 0$ nm to 78% for $\sigma_{x,y} = \sigma_z/2 = 50$ nm) (Supplementary Figure S8C). Localization errors above $\sigma_{x,y} = \sigma_z/2 = 75$ nm drastically reduced both recall (from 93% for $\sigma_{x,y} = \sigma_z/2 = 75$ nm to 7% for $\sigma_{x,y} = \sigma_z/2 = 200$ nm) and precision (73% for $\sigma_{x,y} = \sigma_z/2 = 75$ nm vs. 49% for $\sigma_{x,y} = \sigma_z/2 = 200$ nm) (Supplementary Figure S8C).

We then proceeded to estimate closed loop lifetimes. This was much more challenging and error-prone, because any time point with a false negative closed state detection led to a fragmentation of segmented closed state intervals and therefore to an underestimation of the true lifetime, while false positive time points generally led to spurious short-lived closed states. Nevertheless, using the thresholds above, the estimated lifetimes of closed states were in excellent overall agreement with the true lifetimes (Figure 5B). The exception to this were very short lifetimes (Figure 5B, C), which originated from the residual false positives in the detection of closed states. 95% of these spurious closed states were detected during extruding states, typically at the end of the extrusion phase (Supplementary Figure S8D).

To estimate the mean lifetime of closed states, we fitted an exponential function to the histogram of estimated lifetimes using a robust approach that rejects outliers. We found that with this robust fit, we could accurately estimate the mean closed state lifetime (Figure 5C and Supplementary Figure S9A). The estimated mean closed lifetime across 4000 bootstrapped samples (each containing 1000 time series) was 465.9 s, in excellent agreement with the expected value of 466.6 s (Supplementary Figure S9A).

Therefore, our analysis suggests that closed state lifetimes can be accurately estimated by tracking loop anchors in live-cell imaging.

Quantifying the speed of loop extrusion from live-cell trajectories

Finally, we assessed whether loop extrusion speed could be quantified from 1000 time series of anchor–anchor distances obtained by live-cell imaging. Because the above approach identified closed loop intervals that depend on the presence of extrusion (thanks to the control condition that abolishes extrusion, Supplementary Figure S8A), we reasoned that timepoints before these intervals should be in the extruding state. Assuming, as before, that the polymer with a loop behaves at each time point as an ideal polymer from which the loop portion is absent, we derived analytical models for the temporal evolution of the mean squared anchor–anchor distance (MSAAD) $\langle R^2 \rangle$ (see Materials and Methods for details). If loop extrusion occurred at a constant speed $V(t) = V_0$, the MSAAD would follow a simple linear function of time, given by (Eq. 3). However, we assumed that the speed of extrusion switches from bidirectional with speed $V(t) = V_0$ to unidirectional with speed $V(t) = V_0/2$ when the extrusion complex reaches one of the two anchors (Supplementary Figure S2A). The time point of switching depends on the genomic site where extrusion is initiated: switching occurs earlier if this site is closer to one of the two anchors. As a result, fitting equation (Eq. 3) to measured MSAAD is expected to lead to an effective speed V_{eff} between $V_0/2$ and V_0 , whose exact value depends on the site of initiation.

Under the additional assumption that the site of extrusion initiation (i.e. the loading site of the extrusion complex) is uniformly random between the two loop anchors, we further derived a model for the ensemble average of the MSAAD (EMSAAD) over many time series, denoted as $\langle \langle R^2 \rangle \rangle(t)$, obtained after aligning the starting times of closed states for all time series at $t = 0$, and given by Eq. (4) (Supplementary Figure S10B).

Equipped with this model, we proceeded to analyze time series simulated with a bidirectional extrusion speed $V_0 = 1$ kb/s that switched to a unidirectional extrusion with speed $V_0/2$ upon reaching an anchor. Although the dependence of $\langle \langle R^2 \rangle \rangle(t)$ on time in Eq. 4 is non-linear, we found that a linear fit of the EMSAAD, assuming a single speed V_{eff} (Eq. 5), led to a reasonably accurate estimation of the mean effective extrusion speed ($\widehat{V}_{\text{eff}} = 0.72$ kb/s estimated speed versus $\frac{5}{8}V_0 = 0.63$ kb/s for the true mean effective speed, Figure 5D). The theoretical EMSAAD was well approximated by a linear fit of slope $-5/8 R_0^2 V_0/s_0$ (Figure 5D). Therefore, multiplying the estimated effective speed \widehat{V}_{eff} by $8/5$ allowed us to recover the bidirectional extrusion speed V_0 . This yielded $V_0 = 1$ kb/s for the theoretical curve, as expected, and $\widehat{V}_0 = 1.15$ kb/s for the estimated speed. The remaining discrepancy between estimated and true bidirectional speeds (\widehat{V}_0 and V_0) can be attributed to two effects: (i) the simulated polymer does not have time to equilibrate given the assumed extrusion speed of $V_0 = 1$ kb/s, and (ii) the polymer is confined. Both effects are ignored by our model and lead to a departure from linearity of the EMSAAD as function of time. Indeed, the EMSAAD is closer to the theoretical (equilibrium) value in simulations with slower extrusion and a relaxed confinement (Supplementary Figure S11). Note that the estimated speed depends on

the size of the time window used for fitting. The optimal window size depends on the actual speed and is therefore not known, but correct time windows and speeds can be determined experimentally from a peak in the estimated speed as function of the time window, as shown in Supplementary Figure S9B. This is explained by the slope of the theoretical EMSAAD, which is maximum at $t = \frac{-s_0}{v_0}$ (Supplementary Figure S10B).

Finally, we assessed how experimental parameters affected the estimation of loop extrusion speed. Unsurprisingly, localization errors negatively impacted the estimation of extrusion speed (by deteriorating the accuracy of closed state segmentation). Localization errors of $\sigma_{x,y} = \frac{\sigma_z}{2} = 30$ nm led to underestimations of the speed by 7%, and $\sigma_{x,y} = \frac{\sigma_z}{2} = 75$ nm led to underestimations by 51% (Supplementary Figure S9C and S9E). Similarly, increasing the reporter-anchor separation from 0 to 30 kb led to an underestimate of the loop extrusion speed by 27% (Supplementary Figure S9D and F).

Nevertheless, this analysis suggests that under realistic conditions it should be possible to extract meaningful estimates of the speed of DNA loop extrusion by analyzing trajectories of loop anchors in time-lapse microscopy of living cells.

DISCUSSION

We used polymer simulations and presented dedicated analysis techniques to determine if, and under which conditions, chromatin loop extrusion can be detected and/or characterized quantitatively by imaging two loop anchors with fluorescence microscopy (Figure 1). We separately considered static imaging experiments, as can be performed in fixed cells, e.g. using DNA-FISH (28), or using single snapshots in live-cells, and dynamic imaging experiments in which two loop anchors are tracked over time (35,78). Our analyses assume that imaging experiments are performed both in cells where loop extrusion takes place and in cells where extrusion is experimentally abolished, e.g. by depletion of the cohesin subunit RAD21 (16), which provides a crucial negative control. We considered the effect of multiple biological and experimental parameters. Biological parameters include the fractions of closed, open and extruding states, the speed of loop extrusion and the genomic size of the loops. Experimental parameters include the reporter-anchor separation, the localization precision of fluorescent reporters, and the sample size (number of measured anchor–anchor distances).

To estimate plausible ranges of loop state fractions, we compared simulations to Hi-C data of ≈ 300 kb large loops (Figure 2), and found that only small (15–25%) fractions of closed states are consistent with the data, in line with previous studies (20,35). Despite this relative rarity of closed loops, we showed that in most cases the presence of loop extrusion can be detected from static images alone for realistic localization precisions, provided that the sample size N (i.e. the number of measured anchor–anchor distances) is large enough (Figure 3). For example, for loops of 300 kb with 20% closed and 40% extruding states, a reporter-anchor separation of 6 kb, and a lateral localization precision of 15 nm (30 nm axially), extrusion can be detected

reliably from as little as $N \approx 75$ measurements (Figure 3I). With a poor lateral localization precision of 100 nm (200 nm axially), reliable detection requires $N \approx 200$ measurements, which is still within reach of standard microscopy experiments (Figure 3I). An important result of our analysis is that loop extrusion can be detected even in absence of a second peak in small anchor–anchor distances, which to our knowledge has never been observed in DNA-FISH (28,79) or live-cell imaging (35–38). Furthermore, we showed that analysis of static imaging data also enables to estimate the fraction of closed loops with high accuracy, and the fractions of loops in extruding and open states with reasonable accuracy (Figure 4). Finally, we demonstrated that under realistic conditions, dynamic imaging data can be analyzed to estimate the lifetime of closed states as well as the speed of loop extrusion (Figure 5). Our study therefore establishes the feasibility of an imaging approach to identify and quantitatively characterize loop extrusion. This is a non-trivial finding considering the many challenges arising from the highly stochastic nature of chromatin motion (80), the potential rarity of closed states, and technical limitations of image acquisition. We have defined the minimum conditions that are required to detect and characterize loop extrusion in terms of localization precision, proximity of reporters to anchors and sample size, and highlighted the critical importance of comparing cells with and without extrusion. We also verified that the recent experimental study of loop extrusion by Gabriele *et al.* in mouse embryonic stem cells indeed matched these conditions (35). The analysis methods proposed here (statistical comparisons of distance distributions and fitting to theoretical polymer models) should be directly applicable to experimental data in follow-up studies.

Nonetheless, we acknowledge that our analyses have caveats. First, our method for detecting loop extrusion (Figure 3) assumes that extrusion can be specifically abolished in experiments such as RAD21 depletion (16). We cannot exclude that potential changes to chromatin structure due to RAD21-depletion that are unrelated to extrusion, such as epigenetic modifications (81,82) or increased compartmentalization (23,24), may be picked up unspecifically by this approach. Second, in our analysis of dynamic imaging data, we identified the closed state as a temporally stable low-distance state absent from conditions without extrusion. This method will miss short-lived closed states whose duration is below the temporal threshold. The segmentation of closed states could potentially be made more specific by also considering conditions that permit loop extrusion but not closed states. Such conditions could be achieved experimentally by removing obstacles to loop extrusion, e.g. by depleting the CTCF protein or mutating CTCF binding sites (19,24). Third, our quantification of closed state fractions from static imaging data and our quantification of closed state lifetimes from dynamic data assume that the distributions of anchor–anchor distances in the closed state are known, which is not straightforward. One possible way to estimate these distances (together with localization errors) is to measure the distance between two fluorescent markers separated by a genomic distance equal to the summed distance between anchors and reporters. Fourth, we did not consider the possibility that multiple cohesin

complexes may simultaneously extrude chromatin loops between the two anchors (76,83,84), which would considerably complicate the analysis of state fractions and extrusion speed. Fifth, our analysis ignores the possibility that loops may be in the closed state at the start or end of dynamic imaging experiments (censoring). Sixth, our assumption of a continuous extrusion activity at constant (bidirectional or unidirectional) speed does not account for the possibility of pausing (85), which may lead to underestimations of the peak extrusion speed. Fully considering these complications will require follow-up work.

In spite of these limitations, our study highlights how imaging in fixed or living cells can be used to rigorously identify and quantitatively characterize the fundamental process of chromatin loop extrusion. With DNA-FISH, fractions of closed, extruding and open states could be estimated in a high-throughput manner by targeting probes against the anchors of TADs and loops for hundreds of genomic loci. Since each of these loci is associated with a unique epigenetic state, gene expression pattern and spatial neighborhood, such a systematic analysis might shed light on the intermingling between these factors and loop extrusion, and its consequences on genome functions (30,31,86,87). The possibility to analyze loop extrusion by live-cell imaging is illustrated by two above-mentioned experimental studies in mESC (35,36), which estimated closed loop fractions ranging from 3% to 27% and median loop lifetimes ranging from 5 to 30 minutes. We expect that subsequent experiments will extend this initial work by analyzing many more TADs and loops and their dynamics in living cells of multiple organisms. By providing clear quantitative guidelines and analytical techniques, we hope that the present study will be instrumental in reaching a better quantitative understanding of how loop extrusion affects chromatin dynamics and function.

DATA AVAILABILITY

The LAMMPS parameter files used for the molecular dynamics simulations and the Python code used to estimate loop state fractions from static imaging data are available at <https://github.com/imodpasteur/LoopExtrusion> and through Zenodo at (88). Numerical data of static and dynamic imaging simulations are available through Zenodo at (89). All other numerical data are available on request to the corresponding authors. Hi-C and ChIP-Seq data used in this work were taken from Rao *et al.* (16) with accession number GSE104334.

SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Jyotsana J. Parmar for stimulating discussions and initial help with the simulations. We thank Davide Normanno for critical reading of the manuscript. We acknowledge the help of the HPC Core Facility of the Institut Pasteur. Parts of some figures were created with BioRender.com.

FUNDING

Investissement d'Avenir [ANR-16-CONV-0005]; Contrat Doctoral Spécifique aux Normaliens; Fondation ARC pour la recherche sur le cancer; Institut Pasteur. Funding for open access charge: Institut Pasteur.

Conflict of interest statement. None declared.

REFERENCES

- Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
- Parmar, J.J., Woringer, M. and Zimmer, C. (2019) How the genome folds: the biophysics of four-dimensional chromatin organization. *Annu. Rev. Biophys.*, **48**, 231–253.
- Davidson, I.F. and Peters, J.-M. (2021) Genome folding through loop extrusion by SMC complexes. *Nat. Rev. Mol. Cell Biol.*, **22**, 445–464.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J. *et al.* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, **485**, 381–385.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376–380.
- Rajpurkar, A.R., Mateo, L.J., Murphy, S.E. and Boettiger, A.N. (2021) Deep learning connects DNA traces to transcription to reveal predictive features beyond enhancer–promoter contact. *Nat. Commun.*, **12**, 3423.
- Espinola, S.M., Götz, M., Bellec, M., Messina, O., Fiche, J.-B., Houbon, C., Dejean, M., Reim, I., Cardozo Gizzi, A.M., Lagha, M. *et al.* (2021) Cis-regulatory chromatin loops arise before tads and gene activation, and are independent of cell fate during early Drosophila development. *Nat. Genet.*, **53**, 477–486.
- Di Stefano, M., Stadhouders, R., Farabella, I., Castillo, D., Serra, F., Graf, T. and Marti-Renom, M.A. (2020) Transcriptional activation during cell reprogramming correlates with the formation of 3D open chromatin hubs. *Nat. Commun.*, **11**, 2564.
- Greenwald, W.W., Li, H., Benaglio, P., Jakubosky, D., Matsui, H., Schmitt, A., Selvaraj, S., D'Antonio, M., D'Antonio-Chronowska, A., Smith, E.N. *et al.* (2019) Subtle changes in chromatin loop contact propensity are associated with differential gene regulation and expression. *Nat. Commun.*, **10**, 1054.
- Hansen, A.S., Hsieh, T.-H.S., Cattoglio, C., Pustova, I., Saldaña-Meyer, R., Reinberg, D., Darzacq, X. and Tjian, R. (2019) Distinct classes of chromatin loops revealed by deletion of an RNA-binding region in CTCF. *Mol. Cell*, **76**, 395–411.
- Zuin, J., Roth, G., Zhan, Y., Cramard, J., Redolfi, J., Piskadlo, E., Mach, P., Kryzhanovska, M., Tihanyi, G., Kohler, H. *et al.* (2022) Nonlinear control of transcription through enhancer–promoter interactions. *Nature*, **604**, 571–577.
- Rodríguez-Carballo, E., Lopez-Delisle, L., Willemin, A., Beccari, L., Gitto, S., Mascrez, B. and Duboule, D. (2020) Chromatin topology and the timing of enhancer function at the HoxD locus. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 31231–31241.
- Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P. and Bickmore, W.A. (2019) Decreased enhancer–promoter proximity accompanying enhancer activation. *Mol. Cell*, **76**, 473–484.
- Xiao, J.Y., Hafner, A. and Boettiger, A.N. (2021) How subtle changes in 3D structure can create large changes in transcription. *eLife*, **10**, e64320.
- Luppino, J.M., Park, D.S., Nguyen, S.C., Lan, Y., Xu, Z., Yunker, R. and Joyce, E.F. (2020) Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes. *Nat. Genet.*, **52**, 840–848.
- Rao, S.S.P., Huang, S.-C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.-R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D. *et al.* (2017) Cohesin loss eliminates all loop domains. *Cell*, **171**, 305–320.

17. Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G. and Peters, J.-M. (2019) DNA loop extrusion by human cohesin. *Science*, **366**, 1338–1345.
18. Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A. and Mirny, L.A. (2017) Emerging evidence of chromosome folding by loop extrusion. *Cold Spring Harb. Symp. Quant. Biol.*, **82**, 45–55.
19. Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J. *et al.* (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6456–E6465.
20. Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. and Mirny, L.A. (2016) Formation of chromosomal domains by loop extrusion. *Cell Rep.*, **15**, 2038–2049.
21. Hansen, A.S., Cattoglio, C., Darzacq, X. and Tjian, R. (2018) Recent evidence that tads and chromatin loops are dynamic structures. *Nucleus*, **9**, 20–32.
22. Rowley, M.J. and Corces, V.G. (2018) Organizational principles of 3D genome architecture. *Nat. Rev. Genet.*, **19**, 789.
23. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. and Mirny, L.A. (2018) Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E6697–E6706.
24. Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J. *et al.* (2017) Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.*, **36**, 3573–3599.
25. Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C., Mirny, L. *et al.* (2017) Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, **551**, 51–56.
26. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J. and Yu, H. (2020) Human cohesin compacts DNA by loop extrusion. *Science*, **366**, 1345–1349.
27. Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N. and Zhuang, X. (2018) Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science*, **362**, eaau1783.
28. Finn, E.H., Pegoraro, G., Brandão, H.B., Valton, A.-L., Oomen, M.E., Dekker, J., Mirny, L. and Misteli, T. (2019) Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell*, **176**, 1502–1515.
29. Mateo, L.J., Murphy, S.E., Hafner, A., Cinquini, I.S., Walker, C.A. and Boettiger, A.N. (2019) Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature*, **568**, 49.
30. Su, J.-H., Zheng, P., Kinrot, S.S., Bintu, B. and Zhuang, X. (2020) Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell*, **182**, 1641–1659.
31. Takei, Y., Yun, J., Zheng, S., Ollikainen, N., Pierson, N., White, J., Shah, S., Thomassie, J., Suo, S., Eng, C.-H.L. *et al.* (2021) Integrated spatial genomics reveals global architecture of single nuclei. *Nature*, **590**, 344–350.
32. Cardozo Gizzi, A.M., Cattoni, D.I., Fiche, J.-B., Espinola, S.M., Gurgo, J., Messina, O., Houbbron, C., Ogiyama, Y., Papadopoulos, G.L., Cavalli, G. *et al.* (2019) Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. *Mol. Cell*, **74**, 212–222.
33. Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R. and Darzacq, X. (2017) CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *Elife*, **6**, e25776.
34. Mirny, L. and Dekker, J. (2021) Mechanisms of chromosome folding and nuclear organization: their interplay and open questions. *Cold Spring Harb. Perspect. Biol.*, **14**, a040147.
35. Gabriele, M., Brandão, H.B., Grosse-Holz, S., Jha, A., Dailey, G.M., Cattoglio, C., Hsieh, T.-H.S., Mirny, L., Zechner, C. and Hansen, A.S. (2022) Dynamics of CTCF- and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science*, **376**, 496–501.
36. Mach, P., Kos, P.I., Zhan, Y., Cramard, J., Gaudin, S., Tünnermann, J., Marchi, E., Eglinger, J., Zuin, J., Kryzhanovska, M. *et al.* (2022) Cohesin and CTCF control the dynamics of chromosome folding. *Nat. Genet.*, **54**, 1907–1918.
37. Brandão, H.B., Gabriele, M. and Hansen, A.S. (2021) Tracking and interpreting long-range chromatin interactions with super-resolution live-cell imaging. *Curr. Opin. Cell Biol.*, **70**, 18–26.
38. Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S. and Weiner, O.D. (2019) Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. *eLife*, **8**, e41769.
39. Fu, Y., Rocha, P.P., Luo, V.M., Raviram, R., Deng, Y., Mazzoni, E.O. and Skok, J.A. (2016) CRISPR-dCas9 and sgRNA scaffolds enable dual-color live imaging of satellite sequences and repeat-enriched individual loci. *Nat. Commun.*, **7**, 11707.
40. Qin, P., Parlak, M., Kusc, C., Bandaria, J., Mir, M., Szlachta, K., Singh, R., Darzacq, X., Yildiz, A. and Adli, M. (2017) Live cell imaging of low- and non-repetitive chromosome loci using CRISPR-Cas9. *Nat. Commun.*, **8**, 14725.
41. Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S. *et al.* (2013) Dynamic imaging of genomic loci in living Human cells by an optimized CRISPR/cas system. *Cell*, **155**, 1479–1491.
42. Germier, T., Audibert, S., Kocanova, S., Lane, D. and Bystrycky, K. (2018) Real-time imaging of specific genomic loci in eukaryotic cells using the ANCHOR DNA labelling system. *Methods*, **142**, 16–23.
43. Thompson, A.P., Aktulga, H.M., Berger, R., Bolintineanu, D.S., Brown, W.M., Crozier, P.S., in ’t Veld, P.J., Kohlmeyer, A., Moore, S.G., Nguyen, T.D. *et al.* (2022) LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.*, **271**, 108171.
44. Woring, M., Darzacq, X. and Izeddin, I. (2014) Geometry of the nucleus: A perspective on gene expression regulation. *Curr. Opin. Chem. Biol.*, **20**, 112–119.
45. Rippe, K. (2007) Dynamic organization of the cell nucleus. *Curr. Opin. Genet. Dev.*, **17**, 373–380.
46. Dekker, J. and Misteli, T. (2015) Long-range chromatin interactions. *Cold Spring Harb. Perspect. Biol.*, **7**, a019356.
47. Michaud-Agrawal, N., Denning, E.J., Woolf, T.B. and Beckstein, O. (2011) MDAAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.*, **32**, 2319–2327.
48. Shao, S., Zhang, W., Hu, H., Xue, B., Qin, J., Sun, C., Sun, Y., Wei, W. and Sun, Y. (2016) Long-term dual-color tracking of genomic loci by modified sgRNAs of the CRISPR/Cas9 system. *Nucleic Acids Res.*, **44**, e86.
49. Arbona, J.-M., Herbert, S., Fabre, E. and Zimmer, C. (2017) Inferring the physical properties of yeast chromatin through bayesian analysis of whole nucleus simulations. *Genome Biol.*, **18**, 81.
50. Guo, Y., Al-Jibury, E., Garcia-Millan, R., Ntagiantas, K., King, J.W.D., Nash, A.J., Galjart, N., Lenhard, B., Rueckert, D., Fisher, A.G. *et al.* (2022) Chromatin jets define the properties of cohesin-driven in vivo loop extrusion. *Mol. Cell*, **82**, 3769–3780.
51. Banigan, E.J. and Mirny, L.A. (2020) The interplay between asymmetric and symmetric DNA loop extrusion. *eLife*, **9**, e63528.
52. Vian, L., Pekowska, A., Rao, S.S.P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., Huang, S.-C., Khattabi, L.E., Dose, M., Pruett, N. *et al.* (2018) The energetics and physiological impact of cohesin extrusion. *Cell*, **173**, 1165–1178.
53. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B (Methodological)*, **57**, 289–300.
54. Zhang, B., Zerubia, J. and Olivo-Marin, J.-C. (2007) Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.*, **46**, 1819–1829.
55. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J. *et al.* (2020) SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods*, **17**, 261–272.
56. Huis in ’t Veld, P.J., Herzog, F., Ladurner, R., Davidson, I.F., Piric, S., Kreidl, E., Bhaskara, V., Aebersold, R. and Peters, J.-M. (2014) Characterization of a DNA exit gate in the human cohesin ring. *Science*, **346**, 968–972.
57. Haering, C.H., Löwe, J., Hochwagen, A. and Nasmyth, K. (2002) Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol. Cell*, **9**, 773–788.
58. Srinivasan, M., Scheinost, J.C., Petela, N.J., Gligoris, T.G., Wissler, M., Ogushi, S., Collier, J.E., Voulgaris, M., Kurze, A., Chan, K.-L. *et al.* (2018) The cohesin ring uses its hinge to organize DNA using non-topological as well as topological mechanisms. *Cell*, **173**, 1508–1519.

59. Higashi, T.L., Eickhoff, P., Sousa, J.S., Locke, J., Nans, A., Flynn, H.R., Snijders, A.P., Papageorgiou, G., O'Reilly, N., Chen, Z.A. *et al.* (2020) A structure-based mechanism for DNA entry into the cohesin ring. *Mol. Cell*, **79**, 917–933.
60. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data.
61. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
62. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. and Karolchik, D. (2010) BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics*, **26**, 2204–2207.
63. Amemiya, H.M., Kundaje, A. and Boyle, A.P. (2019) The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.*, **9**, 9354.
64. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W. *et al.* (2008) Model-based analysis of ChIP-seq (MACS). *Genome Biol.*, **9**, R137.
65. Grant, C.E., Bailey, T.L. and Noble, W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017–1018.
66. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S. and Aiden, E.L. (2016) Juicer provides a one-click system for analyzing loop-resolution hi-C experiments. *Cell Syst.*, **3**, 95–98.
67. Greenwald, W.W., Li, H., Smith, E.N., Benaglio, P., Nariyai, N. and Frazer, K.A. (2017) Pgltools: a genomic arithmetic tool suite for manipulation of hi-C peak and other chromatin interaction data. *BMC Bioinf.*, **18**, 207.
68. Stansfield, J.C., Cresswell, K.G., Vladimirov, V.I. and Dozmorov, M.G. (2018) HiCcompare: an R-package for joint normalization and comparison of HI-C datasets. *BMC Bioinf.*, **19**, 279.
69. Natsume, T., Kiyomitsu, T., Saga, Y. and Kanemaki, M.T. (2016) Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep.*, **15**, 210–218.
70. Li, S., Prasanna, X., Salo, V.T., Vattulainen, I. and Ikonen, E. (2019) An efficient auxin-inducible degron system with low basal degradation in human cells. *Nat. Methods*, **16**, 866–869.
71. Banigan, E.J., van den Berg, A.A., Brandão, H.B., Marko, J.F. and Mirny, L.A. (2020) Chromosome organization by one-sided and two-sided loop extrusion. *eLife*, **9**, e53558.
72. Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R. and Darzacq, X. (2020) Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell*, **78**, 539–553.
73. Matthey-Doret, C., Baudry, L., Breuer, A., Montagne, R., Guiglielmoni, N., Scolari, V., Jean, E., Campeas, A., Chanut, P.H., Oriol, E. *et al.* (2020) Computer vision for pattern detection in chromosome contact maps. *Nat. Commun.*, **11**, 5795.
74. Thiecke, M.J., Wutz, G., Muhar, M., Tang, W., Bevan, S., Malysheva, V., Stocsits, R., Neumann, T., Zuber, J., Fraser, P. *et al.* (2020) Cohesin-dependent and -independent mechanisms mediate chromosomal contacts between promoters and enhancers. *Cell Rep.*, **32**, 107929.
75. Holzmann, J., Politi, A.Z., Nagasaka, K., Hantsche-Grininger, M., Walther, N., Koch, B., Fuchs, J., Dürnberger, G., Tang, W., Ladurner, R. *et al.* (2019) Absolute quantification of cohesin, CTCF and their regulators in human cells. *eLife*, **8**, e46269.
76. Cattoglio, C., Pustova, I., Walther, N., Ho, J.J., Hantsche-Grininger, M., Inouye, C.J., Hossain, M.J., Dailey, G.M., Ellenberg, J., Darzacq, X. *et al.* (2019) Determining cellular CTCF and cohesin abundances to constrain 3D genome models. *eLife*, **8**, e40164.
77. Herbert, S., Brion, A., Arbona, J.-M., Lelek, M., Veillet, A., Lelandais, B., Parmar, J., Fernández, F.G., Almayrac, E., Khalil, Y. *et al.* (2017) Chromatin stiffening underlies enhanced locus mobility after DNA damage in budding yeast. *EMBO J.*, **36**, 2595–2608.
78. Clow, P.A., Du, M., Jillette, N., Taghbalout, A., Zhu, J.J. and Cheng, A.W. (2022) CRISPR-mediated multiplexed live cell imaging of nonrepetitive genomic loci with one guide RNA per locus. *Nat. Commun.*, **13**, 1871.
79. Fudenberg, G. and Imakaev, M. (2017) FISH-ing for captured contacts: towards reconciling FISH and 3C. *Nat. Methods*, **14**, 673–678.
80. Finn, E.H. and Misteli, T. (2019) Molecular basis and biological function of variability in spatial genome organization. *Science*, **365**, eaaw9498.
81. Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Flach, K.D., Teunissen, H., Willems, L., Brummelkamp, T.R., Rowland, B.D. and de Wit, E. (2022) A mediator-cohesin axis controls heterochromatin domain formation. *Nat. Commun.*, **13**, 754.
82. Lee, R., Kang, M.-K., Kim, Y.-J., Yang, B., Shim, H., Kim, S., Kim, K., Yang, C.M., Min, B., Jung, W.-J. *et al.* (2022) CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic Acids Res.*, **50**, 207–226.
83. Brandão, H.B., Ren, Z., Karaboja, X., Mirny, L.A. and Wang, X. (2021) DNA-loop-extruding SMC complexes can traverse one another in vivo. *Nat. Struct. Mol. Biol.*, **28**, 642–651.
84. Kim, E., Kerssemakers, J., Shaltiel, I.A., Haering, C.H. and Dekker, C. (2020) DNA-loop extruding condensin complexes can traverse one another. *Nature*, **579**, 438–442.
85. Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.-H.S., Parsi, K.M., Yang, L., Maehr, R., Mirny, L.A. *et al.* (2020) Ultrastructural details of mammalian chromosome architecture. *Mol. Cell*, **78**, 554–565.
86. Miron, E., Oldenkamp, R., Brown, J.M., Pinto, D.M.S., Xu, C.S., Faria, A.R., Shaban, H.A., Rhodes, J.D.P., Innocent, C., Ornellas, S. de *et al.* (2020) Chromatin arranges in chains of mesoscale domains with nanoscale functional topography independent of cohesin. *Sci. Adv.*, **6**, eaba8811.
87. Gu, B., Comerci, C.J., McCarthy, D.G., Saurabh, S., Moerner, W.E. and Wysocka, J. (2020) Opposing effects of cohesin and transcription on CTCF organization revealed by super-resolution imaging. *Mol. Cell*, **80**, 699–711.
88. Lelandais, B. and Sabaté, T. (2023) imodpasteur/LoopExtrusion: LoopExtrusion. <https://doi.org/10.5281/ZENODO.7521634>.
89. Sabaté, T., Lelandais, B., Bertrand, E. and Zimmer, C. (2023) Polymer simulations guide the detection and quantification of chromatin loop extrusion by imaging. <https://doi.org/10.5281/ZENODO.7525055>.