Effects of Extruder Dynamics and Noise on Simulated Chromatin Contact Probability Curves

Konstantinov V.[1,2*], Artem S.[3,4], Lagunov T.[1,2]

[1] Institute of Cytology and Genetics, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Artificial Intelligence Research Institute, Moscow, Russia
[4] Sirius University of Science and Technology, Sirius, Russia

* Corresponding author

E-mail: v.konstantinov.work@gmail.com

Abstract

Loop extrusion by SMC complexes is a key mechanism underlying chromatin folding during both interphase and mitosis. Despite this shared mechanism, computational models of loop extrusion often rely on fundamentally different assumptions: interphase models typically use dynamic extruders with finite lifetimes, while mitotic models employ static extruders placed according to loop size distributions. In this work, we investigate whether these modeling paradigms are interchangeable or yield intrinsically incompatible results. Using publicly available Hi-C data from mitotic chicken cells, we systematically compare dynamic and static loop extrusion models implemented in the Polychrom framework. We evaluate how key parameters—such as extruder lifetime, extrusion velocity, and spatial noise—affect simulated contact probability curves P(s) and loop size distributions. Our results reveal that while both model types can be tuned to approximate the general shape of P(s), they produce distinct internal structures and divergent relationships between loop size and contact decay. We also show that increased extruder lifetimes lead to excessive nested loop formation, which alters both loop statistics and P(s) derivatives. Introducing spatial exclusion constraints between extruders partially restores consistency with static models. These findings highlight that differences in extruder behavior and polymer noise levels can significantly impact chromatin model outcomes and must be carefully accounted for when interpreting or comparing simulation results across biological conditions.

Author summary

Chromatin organization plays a crucial role in gene regulation and cellular function, yet our understanding of its three-dimensional structure relies heavily on computational modeling and the interpretation of complex experimental data. In this study, we use coarse-grained modeling approaches to simulate chromatin folding and systematically investigate how different analysis metrics and data processing methods influence the conclusions drawn from such models. By comparing widely used metrics and exploring the effects of normalization and noise, we highlight potential pitfalls and biases that can arise in chromatin modeling studies. Our findings provide practical recommendations for researchers in the field, aiming to improve the robustness and reproducibility of computational analyses of chromatin architecture. This work will help guide future studies toward more reliable interpretations of chromatin structure and its biological implications.

Introduction

In recent years, a range of chromosome conformation capture techniques has been developed to study the spatial organization of the genome (1–5)(refs). Among them, Hi-C has emerged as one of the most informative, providing genome-wide maps of chromatin contacts and enabling detailed insights into three-dimensional chromosome architecture (3)(ref). The integration of Hi-C data with computational modeling has led to the proposal of several mechanistic frameworks explaining genome organization (6–9).

One of the most prominent of these is the loop extrusion (LE) model, which posits that proteins of the Structural Maintenance of Chromosomes (SMC) family bind to chromatin and actively translocate in opposite directions along the DNA fiber (7,10). As a result, the loop organisation of chromatin appears. Current evidence suggests that this mechanism plays a critical role in both interphase chromatin organization (7) and mitotic chromosome condensation (8,11).

However, computational models of these two contexts typically rely on different implementations of SMC activity: interphase models often assume dynamically loaded SMC complexes that stochastically dissociate from chromatin (6,7), whereas mitotic models employ static extruders placed randomly at intervals corresponding to the average loop size (8,9). Given that the underlying biochemical mechanisms are likely shared between interphase and mitotic contexts, it is reasonable to pursue a unified modeling framework. Such a framework would enable the estimation of key SMC parameters—such as complex abundance, residence time, extrusion velocity, and the probability of bypassing obstacles. Yet the challenge lies not only in harmonizing the mechanistic assumptions, but also in the diversity of evaluation metrics used to compare models to experimental data.

One commonly analyzed feature is the contact probability as a function of genomic distance, P(s). This dependence is sensitive to both the polymer's fractal properties and the loop extrusion, making it a useful but nontrivial readout for model validation (8,9,12).

In this study, we used publicly available data and baseline parameters from (11) to assess how P(s) is influenced by:

- the choice of modeling framework (we re-implemented the model using the Polychrom toolkit as opposed to the original platform);
- the type of SMC model (dynamic vs static extrusion);
- the level of Gaussian noise added to the simulated contact maps.

Additionally, we compare several metrics for evaluating model-experiment agreement and demonstrate the strong dependence on the normalization point used in P(s) comparisons for most of them.

Results and discussion

Used simulations and experimental data

In this study, we utilized experimental Hi-C data generated using the DpnII restriction enzyme from mitotic chicken DT40 cell cultures, in which a degron system targets subunits of both the condensin II and cohesin complexes (11; see Methods). Under these conditions, the resulting contact probability curves P(s) lack statistically stable genomic contacts beyond 2 Mbp, as more distant interactions are presumably influenced by mechanisms unrelated to direct loop extrusion. As noted in the original study, such experimental perturbation is interpreted to limit the effective range of loop extrusion activity to approximately 2 Mbp.

Simulations were conducted using a modified version of the open2c/polychrom software package (13), which supports modeling chromosome architecture via dynamic loading and unloading of extruders. Extruder parameters were selected

based on those reported in (11). However, in contrast to that study, we did not impose external constraints on the overall geometry of the simulated chromosomes — constraints that were originally applied to recapitulate chromosomal structures at genomic distances beyond 2 Mbp.

Base differences

We began by evaluating the reproducibility of experimental data using a model of a 40 Mbp chromosome with statically placed condensin I motors. The extruders were distributed according to an exponential loop size distribution with a mean of 100 kbp and an inter-extruder distance of 400 bp, parameters previously suggested as optimal for recapitulating contact density profiles in experiments (11). To assess the consistency of modeling results across different implementations, we employed two parameter sets and interaction potentials within the Polychrom framework: the Step potential, used as the default in dynamic extrusion models (14), and the DPD potential, as utilized in (11,15). Further details are provided in the Methods section.

The resulting contact probability curves (Fig. 1) revealed discrepancies between the simulated contact distributions (as a function of genomic distance) and the experimental data for both potentials. These differences can be summarized as follows:

- A poor agreement was observed in the short-range contact region (up to 10 kbp), suggesting a mismatch between the chosen interaction potentials and

the underlying physical parameters (see Methods).

- A generally reduced contact density was noted across the 3 kbp to 2 Mbp range.

- Fluctuations in the derivative of the contact probability are consistent with stochastic noise typical of complex molecular systems under limited sampling.

- The Step potential produced results more consistent with the experimental data than the DPD potential.

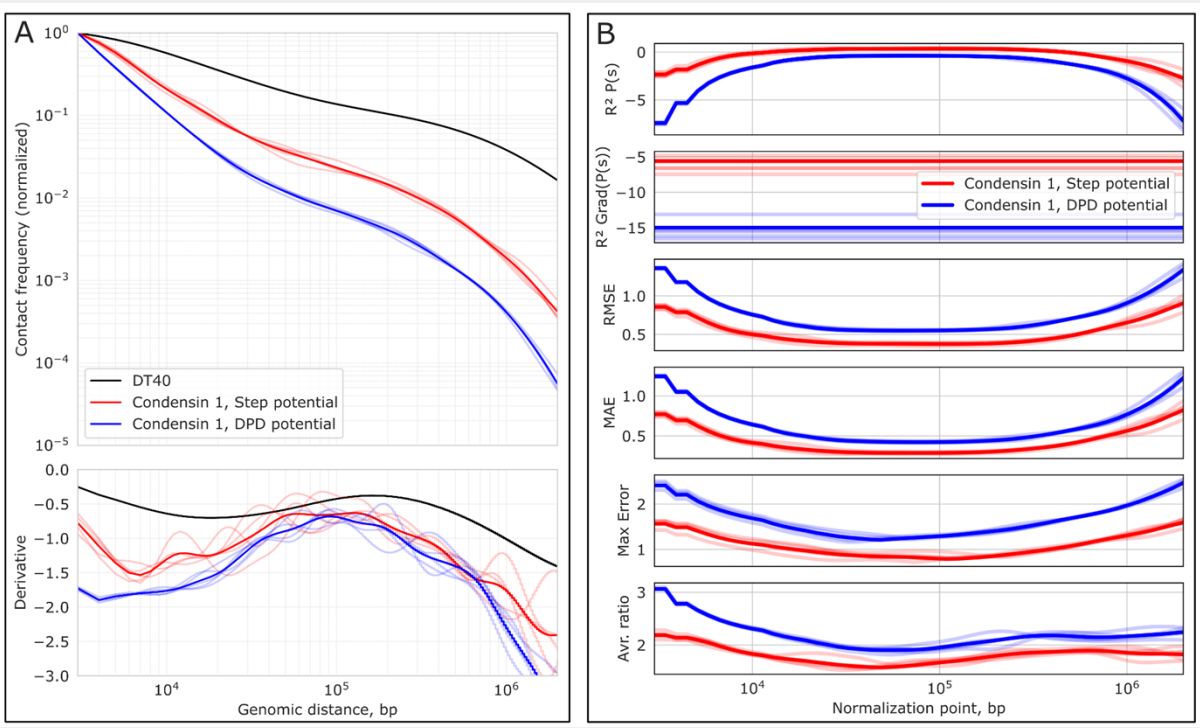

Figure 1. A) P(s) curves in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 [(11)] (black line), raw model single cells (transparent lines),

aggregated raw model data (solid lines). Static extruders with exponential distribution of loop sizes: with Step potential (red line), DPD potential (blue line); B) Results of comparing segments experiment P(s) vs model data with shifting normalisation point on different metrics. The normalization point shifts within the range of 3 kbp to 2 Mbp. The comparable segment of the plot is 3 kbp-2 Mbp. Sigma-noise is 200 nm.

Quantitative metrics

In addition to direct visual comparison, we employed several quantitative metrics to assess the similarity between simulated and experimental contact probability curves. However, a key methodological issue arises: P(s) curves are relative by nature and require normalization at a chosen genomic distance. For example, in (11), normalization is performed at 3 kbp, and comparisons are made over the range of 3 kbp to 2 Mbp. In contrast, (12) applies normalization at 10 kbp or, in some cases, 100 kbp to emphasize longer-range contacts.

These differences in normalization introduce significant variability in commonly used metrics, including R^2 for P(s), root mean square error (RMSE), mean absolute error (MAE), maximum error (ME), and average ratio. In contrast, R^2 calculated for the derivative of P(s) remains unaffected by normalization, since derivatives are independent of the absolute scaling of P(s) (Fig. 1B).

We observed that R^2, RMSE, and MAE values generally increase with normalization points approaching ~100 kbp, reaching a maximum near 1 Mbp. The ME metric exhibits a similar trend, although its maximum occurs around 100 kbp. Beyond these points, all metrics based on P(s) begin to decline in quality, especially as the analysis

approaches the 2 Mbp upper boundary. In contrast, the average ratio metric peaks at 20–30 kbp and shows a pronounced decrease beyond 1 Mbp.

This analysis was repeated across all simulation conditions presented in this study (see Appendix). For consistency, we selected 3 kbp as the normalization point for determining best-fit parameters throughout the paper, aligning with the normalization strategy used in (11).

These observations highlight the importance of carefully selecting both the normalization point and the comparison metric. To ensure comparability between modeling studies, we recommend reporting the R^2 of the P(s) derivative in addition to any primary metric.

Artificial artifacts

Figure 1 illustrates that contact-based analyses under conditions of sparse interaction sampling are inherently sensitive to stochastic fluctuations, necessitating cautious interpretation of contact density variations. One commonly used strategy to address this limitation is the introduction of positional noise, which can reconcile discrepancies between simulated and experimental data by increasing the effective number of contacts. In (11), this is achieved by applying Gaussian displacements to the 3D coordinates of polymer conformations along all spatial axes. This approach effectively generates ‘additional’ conformations, thereby increasing the number of distant, probabilistic contacts—albeit at the cost of uniform short-range contact frequency. It is important to note that while increasing the magnitude of noise (σ)

180 smooths the contact probability curve, excessive noise can obscure or eliminate the
181 underlying structural features of chromatin.

182 The rationale behind this noise addition is rooted in the observed noisiness of Hi-C
183 data, which often displays an almost uniform distribution of short-range contacts. In
184 contrast, simulations generally lack such uniformity, leading to an underestimation of
185 short-distance contact frequencies. Notably, switching the restriction enzyme from
186 DpnII (used in standard Hi-C) to MNase (used in Micro-C) introduces more frequent
187 cutting (~200 bp for MNase vs. ~1–5 kbp for DpnII), yielding a less uniform
188 distribution of short-range contacts and better resolution of chromatin structure (16).
189 In human H1-hESC cells, the most pronounced changes in the derivative of the
190 contact probability occur within the ~1 kb to ~100 kb range, which corresponds
191 closely to the characteristic size of interphase chromatin loops (~100 kb).

192 When σ was varied over a range of [50, 100, 200, 400] nm (Fig. 2B), a predictable
193 decrease in contact density was observed, primarily due to the loss of short-range
194 contacts below 10 kbp. The normalization point for these analyses was fixed at 3
195 kbp. Notably, normalization had no impact on the derivative of P(s), which revealed
196 clear divergence between high-noise model curves and experimental data.

197 We also observed that increasing σ shifted characteristic local maxima and minima
198 in the P(s) derivative toward larger genomic distances. This is consistent with the
199 idea that smaller loops are more susceptible to noise-induced perturbation. A
200 comparable analysis performed using the DPD potential (Fig. A1) yielded
201 qualitatively similar results.

In (11), the optimal σ was reported as 49 nm, producing an average ratio metric of approximately 1.05 using the DPD potential. In our Step potential implementation, we observed a similar value (~1.16) at σ = 200 nm, while the DPD-based implementation yielded a best value of ~1.23 at σ = 100 nm (Appendix). These σ values are consistent with the DpnII digestion scale, where 49 nm corresponds to roughly 5 nucleosomes (~1,000 bp), and 200 nm to ~20 nucleosomes (~4,000 bp).

Based on this analysis, we selected the Step potential with σ = 200 nm as the default configuration for all subsequent simulations, due to its favorable performance across multiple fit metrics.

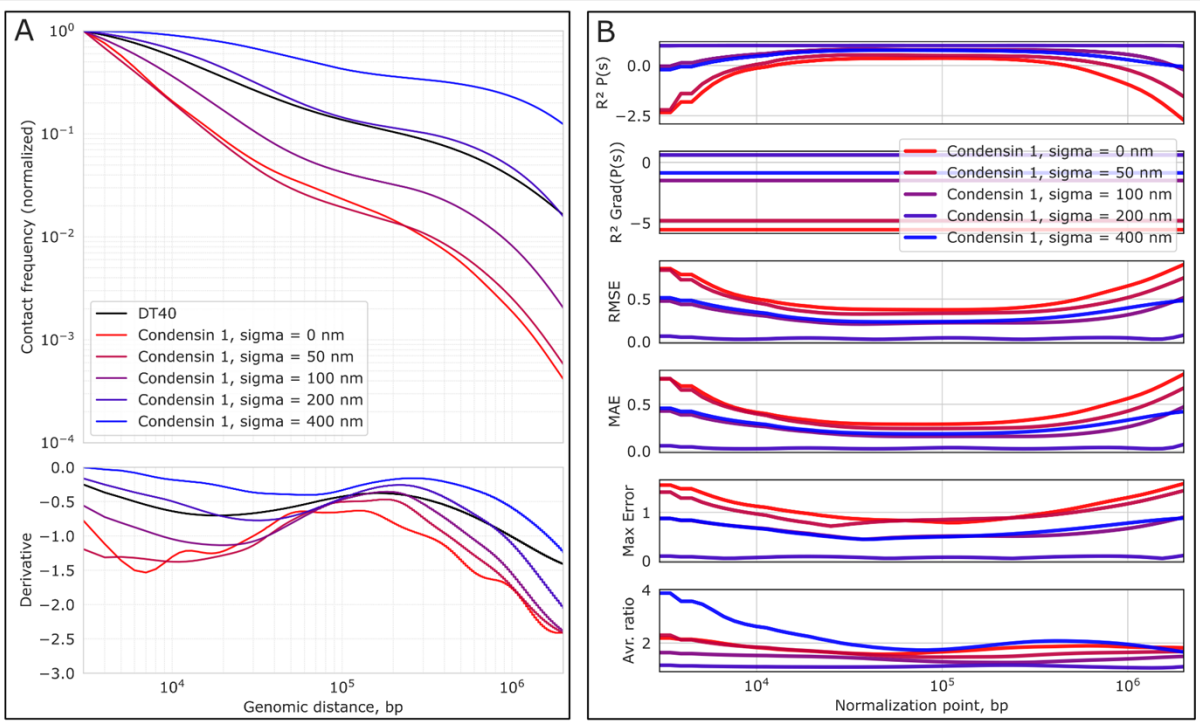

Figure 2. A) Smoothing of P(s) model curves with different sigma-noises in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (11) (black line), aggregated model data for Step potential; B) Results of comparing segments experiment P(s) vs model data with shifting normalisation point on different metrics. The normalization point shifts within the range of 3 kbp to 2 Mbp. The comparable segment of the plot is 3 kbp-2 Mbp. Sigma-noise is 200 nm.

Dynamic extruders

The use of static extruders with an exponential loop size distribution raises important questions regarding their validity for modeling the dynamic behavior of SMC complexes. In (11), static extruders are employed in conjunction with externally imposed large-scale helical constraints, which promote chromosome compaction. In contrast, (12) demonstrates that a similar helical structure can emerge spontaneously in simulations with dynamic extruders, even in the absence of external spatial constraints. In our simulations, however, we did not observe the formation of helical structures under either static or dynamic extrusion regimes, in contrast to the results reported by (12) for dynamic models.

A related issue concerns the selection of optimal parameters for modeling dynamic extruders. The key parameter in this context is the extruder **lifetime**—defined as the average duration of DNA binding—which directly determines the resulting loop size. We estimated the base lifetime using the following relation:

$$\mathbf{lt = (1/6) \times avg_loop_size / velocity} \quad (3)$$

where the average loop size for condensin I is taken to be 100 kbp, the extrusion velocity is set to 1 kbp/sec (17), and the 1/6 factor accounts for the effective scaling of the polymer modeled as a chain of 10 nm beads. This yields a baseline lifetime of approximately 17 seconds, sufficient for ~63% of extruders to reach the next adjacent extruder along the polymer. In the simulations, each condensin complex has a probability of 1/lt to detach from the polymer at each time step and is immediately reloaded at a random location.

To explore the effect of extruder dynamics, we simulated three different lifetimes:

- **17 seconds**, corresponding to the time required for condensin I to extrude its average loop size (~100 kbp);
- **68 seconds**, corresponding to the time required for condensin II to extrude its average loop size (~400 kbp);
- **136 seconds**, used as a control to investigate the effects of prolonged binding.

As shown in Figure 3A, we observe marked differences in contact density profiles between dynamic and static extrusion regimes. For extruders with a 17-second lifetime, contact density is elevated relative to the experimental data at distances in the 10–100 kbp range, followed by a steep decline. This reflects the short binding duration, during which many extruders fail to encounter obstacles before detaching; approximately one-third dissociate within 17 seconds.

For the 68-second condition, the number of short-range contacts (<40 kbp) is reduced relative to the 17-second case, due to fewer short-lived extruders. However,

contact density increases significantly in the 50–200 kbp range, aligning more closely with experimental data in this interval. Beyond this range, the model again overestimates contact decay.

The 136-second extruders exhibit similar behavior up to ~20 kbp, but contact density is slightly lower than in the 68-second case up to ~70 kbp. Between 70 and 600 kbp, however, the contact probability profile closely matches the experimental curve. Beyond 1 Mbp, discrepancies reemerge across all conditions, consistent with the finite influence range of loop extrusion observed in the experimental system.

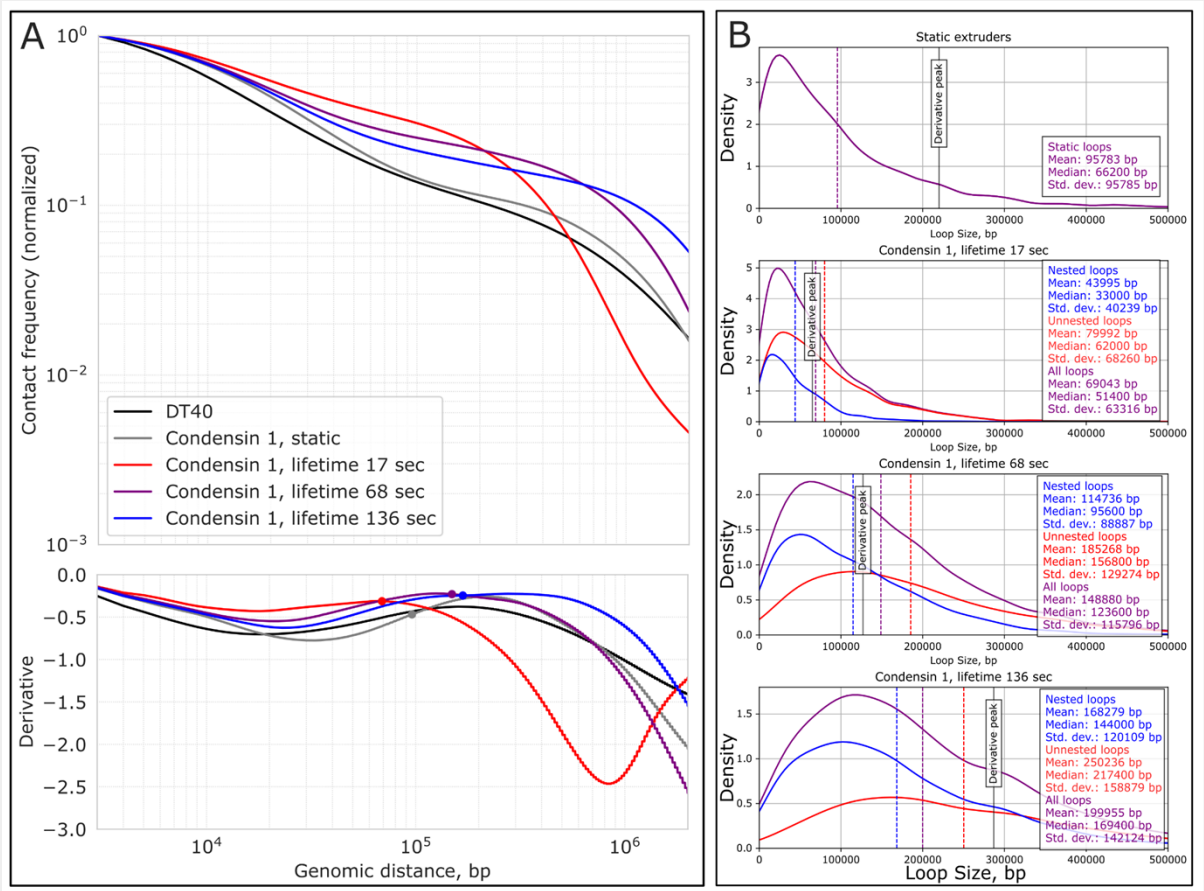

Figure 3. A) P(s) curves in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (black line), aggregated model data. Static (grey line) and dynamic extruders on three lifetimes with base physical params (red, blue, purple lines). Big dots at derivative graphs correspond to mean loop size; B) Loop sizes distribution for different lifetimes. All graphs normalised on area under all-loops graph. Purple line - all extruders, red line - unnested extruders, blue line - nested extruders. 1) Static extruders, 2) Dynamic extruders lt = 17 sec, 3) Dynamic extruders lt - 68 sec, 4) Dynamic extruders 136 sec.

To further investigate the impact of extruder lifetimes on the three-dimensional organization of chromatin, we analyzed the resulting distributions of loop sizes under different dynamic conditions. This analysis explicitly considers the formation of nested loops, which occur when new loops are extruded within preexisting ones. It is important to note that z-loop formation (18) does not occur in our model, as extruders are not permitted to pass through one another.

For extruders with a 17-second lifetime (Fig. 3B), the loop size distribution follows an approximately exponential profile, closely resembling that observed for static extruders. In this regime, non-nested loops significantly outnumber nested ones. As the extruder lifetime increases, the average loop size grows accordingly. Because the total number of extruders on chromatin remains constant, the probability of new extruders binding within existing loops increases, leading to a rise in the number of nested loops. With further increases in lifetime (e.g., 136 seconds), the ratio of nested to non-nested loops continues to rise, indicating progressive structural complexity.

Previous studies have established a correlation between average loop size and the position of the peak in the derivative of the P(s) curve (7,19,20). While our findings support the overall trend — namely, that larger average loop sizes are associated with a rightward shift of the derivative peak — we also observe quantitative discrepancies between these two measures. Specifically, under the 136-second lifetime condition (Fig. 3), the position of the derivative peak substantially exceeds the corresponding average loop size.

A similar discrepancy is observed in the case of static extruders when Gaussian noise exceeds 50 nm (Fig. 3A). Although the loop size is fixed at 100 kbp, the derivative peak shifts from ~90 kbp to nearly 300 kbp as noise increases. These findings suggest that while the derivative of P(s) is a useful qualitative indicator of loop size, it may be quantitatively unreliable under certain conditions, such as high noise levels or extensive nested loop formation. As such, the use of derivative peak position as a proxy for extruder-driven loop size in experimental datasets remains an open question.

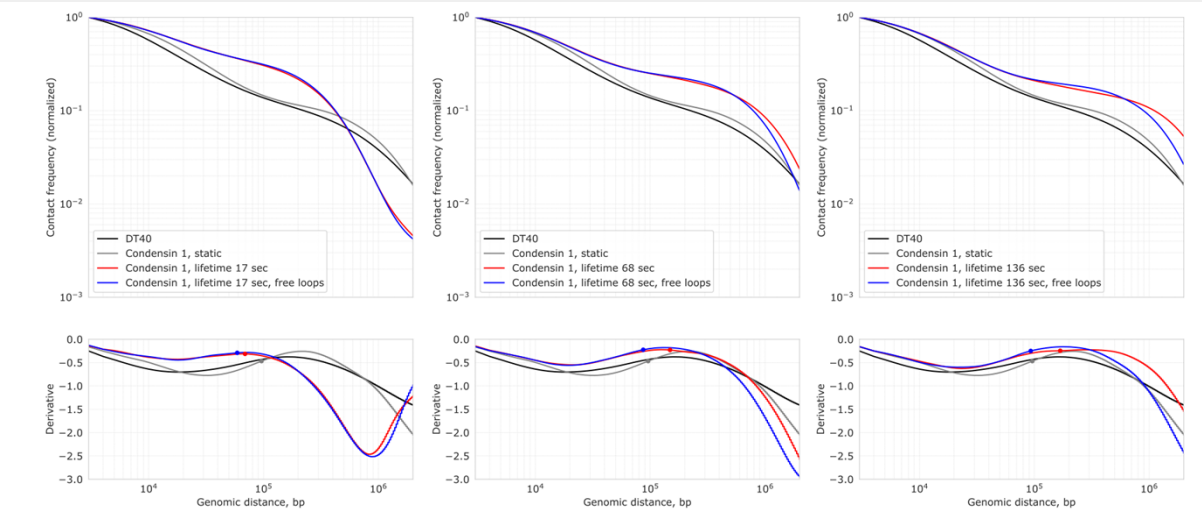

Figure 4. P(s) curves in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (black line) and aggregated model data. Big dots at derivative graphs correspond to mean loop size

As shown in Figure 3, increasing the extruder lifetime leads to a marked rise in the number of nested loops. This effect is associated with a reduction in the effective number of active extruders, resulting from spatial overlap and the hierarchical nature of nested loop organization. To address this limitation, we implemented an alternative strategy introducing spatial exclusion rules, whereby condensin complexes are restricted to loading only within interloop regions, and not within existing loops (Fig. 4).

Under this modified framework, the loop size distribution for extruders with a 17-second lifetime remained unchanged, in agreement with theoretical expectations. However, at longer lifetimes, the distributions shifted substantially, becoming more similar to those observed in the static extruder configuration, as reflected in both the P(s) curves (Fig. 4) and loop size distributions (Fig. A3).

The rationale for this exclusion model stems from experimental observations that condensins can induce chromatin twisting or torsional stress (21), which may affect the physical properties of the substrate and reduce the likelihood of additional extruders binding to already extruded loops. It is therefore theoretically plausible that torsional tension contributes to a reduced probability of condensin loading in regions already occupied by other extruders.

None of the tested models outperformed the static extruder model in terms of quantitative agreement with experimental data. Specifically, the average ratio metric

333 for the static model with optimal Gaussian noise (σ = 200 nm) was 1.16. For the DPD
334 potential model with its best-performing noise level (σ = 100 nm), the value was
335 1.23. The dynamic extruder model with its optimal lifetime (136 sec) yielded a value
336 of 1.37, and the dynamic model with restricted loading (i.e., extruders prohibited from
337 binding within existing loops) achieved a value of 1.35 under the optimal lifetime (136
338 sec).

339 A similar trend was observed for the derivative R^2 metric, which reflects the
340 agreement in slope between modeled and experimental P(s) curves. The static
341 model with σ = 200 nm achieved a derivative R^2 value of 0.98; the DPD model with σ
342 = 100 nm produced 0.97; the unrestricted dynamic extruder model with optimal
343 lifetime (136 sec) gave 0.72; and the restricted dynamic model yielded 0.77 under
344 the optimal lifetime (136 sec).

345 An alternative model assumes that condensins bind DNA once at mitotic entry,
346 extrude loops until they encounter other complexes, and remain stably associated
347 until the end of mitosis. However, this model would require the introduction of
348 additional mechanisms. For instance, the experimentally observed nonzero gaps
349 between loops (11) could be explained by stalling forces arising from chromatin
350 tension exceeding ~2 pN. Similarly, the progressive compaction of chromosomes
351 and the upward shift in the second diagonal of contact maps, despite relatively stable
352 loop sizes for condensin II (11), may reflect the influence of loop extrusion–
353 independent mechanisms, such as activity of topoisomerase II (22).

354

Methods

Polymer Models

We constructed polymer models of chromosomes using polymer dynamics simulations. The simulations were performed using a package based on open2c/polychrom (Imakaev et al., 2019), integrating parameters from both static and dynamic simulations as described in Samejima et al. (2025).

The simulation workflow consists of two main stages:

1. One-dimensional (1D) simulation

   The polymer chain is represented as a 1D lattice composed of beads, each 10 nm in size and corresponding to 200 base pairs (bp).

   1.1. Dynamic extruders, representing condensin I complexes, are modeled as walker objects with two legs that move probabilistically in opposite directions. At each time step, only one leg is active and attempts to move up to approximately 50 nm (5 beads) in its respective direction. The active leg switches with a 50% probability, effectively simulating two-sided loop extrusion. The positions of the legs define the loop boundaries extruded by condensins.

   1.2. Condensins bind randomly to pairs of adjacent free lattice sites and extrude loops until unbinding occurs, after which they immediately rebind at random locations. When two legs of the different extruders collide, extrusion stalls until one leg unbinds.

1.3. Static extruders are also modeled as walker objects, but with legs loaded onto lattice sites following an exponential distribution with a mean loop size of 100 kbp. These static extruders do not perform extrusion or unbinding.

1.4. Condensins identified in the 1D simulation are subsequently implemented as additional harmonic constraints in the 3D polymer model.

1.5. The simulation time step is 0.1 seconds. Condensin velocity is set to 1 kbp/s on naked DNA (17), corresponding to a step size of 50 nm (23); this translates to approximately 6 kbp/s on condensed chromatin. The total simulation duration is 1800 seconds.

1.6. All 1D modelling performed by our scripts (https://github.com/Three-point/1D-polymer-models)

2. Three-dimensional (3D) simulation

This stage employs molecular dynamics simulations using normalized units, where particles have a mass of 1 unit and a bead radius of 0.5 units, with one length unit corresponding to 10 nm.

2.1. Monomers are connected by harmonic springs ((polychrom.forces.harmonic_bonds) with a bond length of 1 unit and a wiggle distance of 0.1 units.

2.2. Polymer stiffness (polychrom.forces.angle_force) is modeled with a persistence length of 150 bp, equivalent to three nucleosome linkers of 50 bp each, implemented via an angular force with k=3.

2.3. Repulsive forces prevent spatial overlap using polynomial repulsion (polychrom.forces.polynomial_repulsive) with a Erep = 3.0, mimicking

real polymer behavior (Formula 1). The collision rate is set to 0.01/picosecond, and the Langevin integrator is used with an error tolerance of 0.01. Each time step comprises 1000 molecular dynamics steps.

2.4. No external spatial constraints are applied to the polymer in this work. This differs from (11), where cylindrical and pulling constraints were used to achieve required densities and patterns.

Each dataset comprises five independent 40 Mbp chains, each initialized with distinct condensin positions and starting conformations generated by random walks. We used the presented parameters for modeling with Step potential.

$$U_{12} = R^{12}*(R^2-1)*E/m \quad (1)$$

$$Udpd = (R-1)^2*E/2 \quad (2)$$

where R - distance between beads in dimensionless units, E - repulsion energy coefficient, m - normalisation constant for this degree of equation.

Experiment types

There are 11 types of model params were used:

- Static extruders. Extruders placed on lattice with exponential distribution of loops once and their positions are not changed with time (no move, binding/unbinding):
    - Step potential
    - DPD potential (Groot and Warren, 1997, Formula 2), where Erep = 3.0 kT, k = 0 (Polymer stiffness), other params equal to Step potential.
- Dynamic extruders. All physical params taken from Step potential. They unbinds from lattice with probability inversely proportional to lifetime, then instantly binds back at random site. Then they make one step by one of legs shoosed randomly (50% chance to switch direction) on 5 beads each time step with probability enough for target extrusion speed:
    - Lifetime = 17 sec of each extruder. This is enough for 60% of extruders to extrude 100 kbp (if encountered no obstacles), forming a distribution closest to exponential. There are low numbers of nested loops (loops on other loops), and a dominating number of unnested loops.
    - Lifetime = 68 sec for each extruder. Nested loops started domination in numbers on short distances.
    - Lifetime = 136 sec for each extruder. Nested loops dominating in numbers on all loop sizes.
- Dynamic extruders which are unable to load at already existing loops, they load only at gaps between loops.
    - Lifetime = 17 sec of each extruder.
    - Lifetime = 68 sec of each extruder.
    - Lifetime = 136 sec of each extruder.

Generation of Hi-C Maps

For each polymer chain in the simulations, we extracted conformations represented by the spatial coordinates of beads at the following time points: 1080 s, 1440 s, and 1800 s. This selection is based on the assumption that a 360 s interval (exceeding three condensin lifetimes) suffices to establish a new loop distribution under dynamic conditions.

Subsequently, each bead position within a conformation was perturbed by adding a random Gaussian displacement with standard deviation σ along each spatial axis. By repeating this perturbation process, we generated 100 new conformations from each original one, resulting in a total of $5 \times 3 \times 100 = 1500$ conformations per dataset.

For each conformation, pairs of beads located within 11 nm were identified as contacts, from which a contact map in .mcool format was constructed (https://github.com/Three-point/1D-polymer-models). Unless otherwise specified, the default value of σ was set to 200 nm, as this value yielded the best agreement with the experimental P(s) curve for static extruder conditions (see figure).

Processed Hi-C data (.mcool) from the DT40 cell line of Gallus gallus (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE262525), sample GSM8171168, were used as experimental data.

The P(s) curves were computed using the cooltools (version 0.5.1) and cooler (version 0.9.1) packages with Gaussian smoothing ($\sigma=0.12$) and excluding the first two diagonals of the Hi-C contact matrix.

Quantitative Comparison of P(s) Curves

To quantitatively compare the contact probability curves P(s), we focused on the genomic distance range from 3 kbp to 2 Mbp, which encompasses the effective range of condensin I activity (typical loop size ~100 kbp). Prior to metric evaluation, all curves were normalized at a selected genomic coordinate within this range. Given the logarithmic scaling of the genomic distance axis, discrepancies at larger distances exert a disproportionately strong influence on agreement metrics. Therefore, the analyzed range was subdivided into 100 uniformly spaced intervals in log scale.

The following metrics were used to assess the agreement between experimental and simulated P(s) curves in log10–log10 space:

- **Coefficient of determination (R^2):** Values approaching 1 indicate strong agreement; negative values suggest that the model performs worse than a horizontal mean line.

- **Root Mean Square Error (RMSE):** Lower values denote better fit and reduced variance from the experimental curve.

- **Mean Absolute Error (MAE):** Measures the average magnitude of differences; lower values indicate better agreement.

- **Maximum Error (ME):** Captures the largest single deviation between curves; lower values are preferred.

- **Average Ratio:** At each point, the modeled P(s) value is divided by the corresponding experimental value. If the ratio is <1, its reciprocal is taken to ensure symmetry. The final value is the arithmetic mean of these normalized quotients across all distances. A value close to 1 reflects excellent fit, whereas larger values indicate greater overall deviation.

For comparing the derivatives of the P(s) curves, only the **R^2 metric** was employed. This choice is justified by its invariance to normalization, as it captures the similarity in slope and shape of the curves, independent of absolute scaling.

Loop Distribution Analysis

Loop positions were extracted at identical time points (1080 s, 1440 s, and 1800 s) for each simulated cell and categorized into the following groups:

- **Nested loops**: loops fully enclosed within the genomic span of another loop.
- **Unnested loops**: loops that are not nested within any other loop.
- **All loops**: the complete set of loops detected, regardless of nesting status.

For each extruder, the genomic span—defined as the distance between its two extrusion legs in base pairs was calculated. The resulting distributions of loop sizes for each category were visualized using Kernel Density Estimation (KDE) plots, generated with the Seaborn Python library.

509

References

1. Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements [Интернет]. [цитируется по 29 май 2025 г.]. Доступно на: https://genome.cshlp.org/content/16/10/1299

2. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing Chromosome Conformation. Science [Интернет]. 15 февраль 2002 г. [цитируется по 29 май 2025 г.]; Доступно на: https://www.science.org/doi/10.1126/science.1067799

3. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, и др. Comprehensive mapping of long range interactions reveals folding principles of the human genome. Science. 9 октябрь 2009 г.;326(5950):289–93.

4. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, и др. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). Nat Genet. ноябрь 2006 г.;38(11):1348–54.

5. Zhao Z, Tavoosidana G, Sjölinder M, Göndör A, Mariano P, Wang S, и др. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet. ноябрь 2006 г.;38(11):1341–7.

6. Banigan EJ, Tang W, van den Berg AA, Stocsits RR, Wutz G, Brandão HB, и др. Transcription shapes 3D chromatin organization by interacting with loop extrusion. Proceedings of the National Academy of Sciences. 14 март 2023 г.;120(11):e2210480120.

7. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of Chromosomal Domains by Loop Extrusion. Cell Reports. 31 май 2016 г.;15(9):2038–49.

8. Gibcus JH, Samejima K, Goloborodko A, Samejima I, Naumova N, Nuebler J, и др. A pathway for mitotic chromosome formation. Science. 9 февраль 2018 г.;359(6376):eaao6135.

9. Nuebler J, Fudenberg G, Imakaev M, Abdennur N, Mirny LA. Chromatin organization by an interplay of loop extrusion and compartmental segregation. Proceedings of the National Academy of Sciences. 17 июль 2018 г.;115(29):E6697–706.

10. Banigan EJ, Van Den Berg AA, Brandão HB, Marko JF, Mirny LA. Chromosome organization by one-sided and two-sided loop extrusion. eLife. 6 апрель 2020 г.;9:e53558.

11. Samejima K, Gibcus JH, Abraham S, Cisneros-Soberanis F, Samejima I, Beckett

AJ, и др. Rules of engagement for condensins and cohesins guide mitotic chromosome formation. Science [Интернет]. 11 апрель 2025 г. [цитируется по 12 май 2025 г.]; Доступно на: https://www.science.org/doi/10.1126/science.adq1709

12. Dey A, Shi G, Takaki R, Thirumalai D. Structural changes in chromosomes driven by multiple condensin motors during mitosis. Cell Reports. 25 апрель 2023 г.;42(4):112348.

13. Imakaev M, Goloborodko A, hbbrandao. mirnylab/polychrom: v0.1.0 [Интернет]. Zenodo; 2019 [цитируется по 29 ноябрь 2024 г.]. Доступно на: https://zenodo.org/records/3579473

14. Imakaev MV, Fudenberg G, Mirny LA. Modeling chromosomes: beyond pretty pictures. FEBS Lett. 7 октябрь 2015 г.;589(20 0 0):3031–6.

15. Anderson JA, Glaser J, Glotzer SC. HOOMD-blue: A Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. Computational Materials Science. 15 февраль 2020 г.;173:109363.

16. Akgol Oksuz B, Yang L, Abraham S, Venev SV, Krietenstein N, Parsi KM, и др. Systematic evaluation of chromosome conformation capture assays. Nat Methods. сентябрь 2021 г.;18(9):1046–55.

17. Tortora MMC, Fudenberg G. The physical chemistry of interphase loop extrusion [Интернет]. bioRxiv; 2024 [цитируется по 29 май 2025 г.]. с. 2024.08.23.609419. Доступно на: https://www.biorxiv.org/content/10.1101/2024.08.23.609419v1

18. Brandão HB, Ren Z, Karaboja X, Mirny LA, Wang X. DNA-loop extruding SMC complexes can traverse one another in vivo. Nat Struct Mol Biol. август 2021 г.;28(8):642–51.

19. Kim J, Jimenez DS, Ragipani B, Zhang B, Street LA, Kramer M, и др. Condensin DC loads and spreads from recruitment sites to create loop-anchored TADs in C. elegans. Tyler JK, редактор. eLife. 4 ноябрь 2022 г.;11:e68745.

20. Polovnikov KE, Brandão HB, Belan S, Slavov B, Imakaev M, Mirny LA. Crumpled Polymer with Loops Recapitulates Key Features of Chromosome Organization. Phys Rev X. 13 ноябрь 2023 г.;13(4):041029.

21. Janissen R, Barth R, Davidson IF, Peters JM, Dekker C. All eukaryotic SMC proteins induce a twist of −0.6 at each DNA loop extrusion step. Science Advances [Интернет]. 13 декабрь 2024 г. [цитируется по 2 июнь 2025 г.]; Доступно на: https://www.science.org/doi/10.1126/sciadv.adt1832

22. Hildebrand EM, Polovnikov K, Dekker B, Liu Y, Lafontaine DL, Fox AN, и др. Mitotic chromosomes are self-entangled and disentangle through a topoisomerase-II-dependent two-stage exit from mitosis. Molecular Cell. 18 апрель 2024 г.;84(8):1422-1441.e14.

23. Ryu JK, Rah SH, Janissen R, Kerssemakers JWJ, Bonato A, Michieletto D, и др. Condensin extrudes DNA loops in steps up to hundreds of base pairs that are generated by ATP binding events. Nucleic Acids Res. 25 январь 2022 г.;50(2):820–32.

24. Groot RD, Warren PB. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. The Journal of Chemical Physics. 15 сентябрь 1997 г.;107(11):4423–35.

Acknowledgments

The degron analysis in this work was supported by the grant of the state program of the «Sirius» Federal Territory «Scientific and technological development of the «Sirius» Federal Territory» (Agreement №26-03, 27/09/2024). Modelling part of the study was funded by the Ministry of Education and Science of the Russian Federation, state project FWNR-2022-0019. The original manuscript text was composed by authors; proofreading was conducted with the assistance of ChatGPT version 4, after that the text was corrected and edited by the authors.

Author contributions

T.L. conceptualized the study. Bioinformatic analyses and polymer simulations were carried out by V.K., with contributions from T.L. and A.S. V.K. drafted the manuscript with contributions by T.L. and A.S. All authors reviewed, revised, and approved the final manuscript.

Code access

The code used in work this study are publicly available at https://github.com/Three-point/1D-polymer-models

Appendix

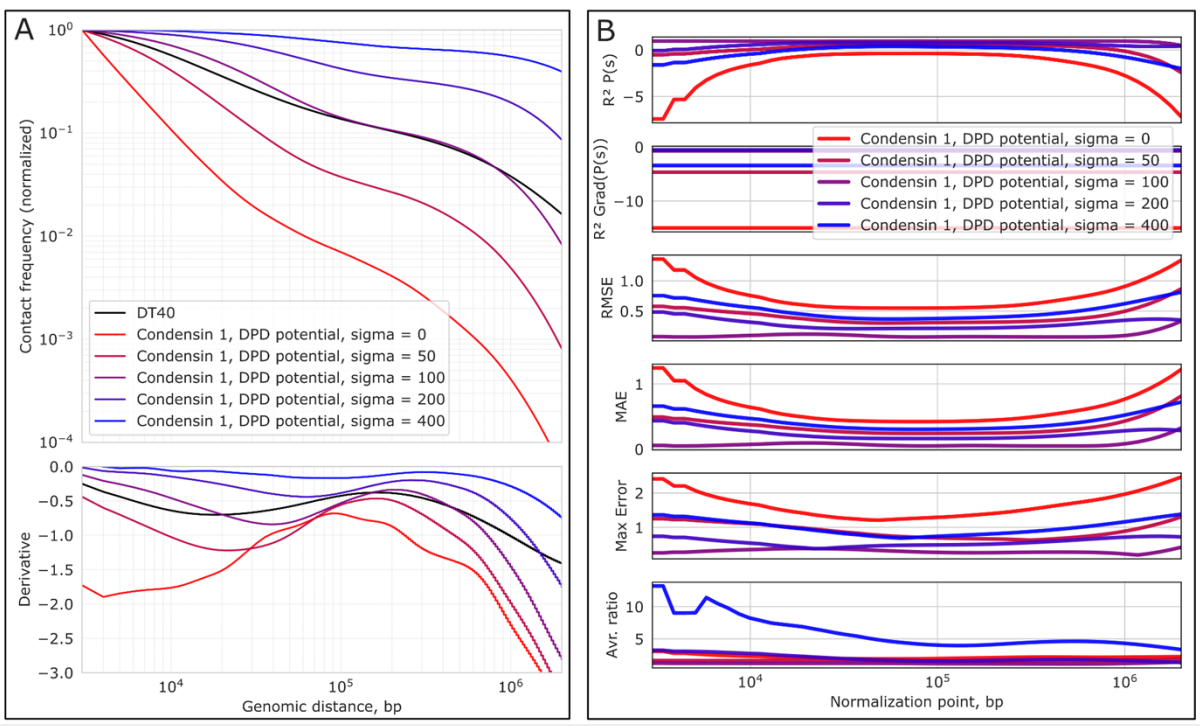

Figure A1. A) Smoothing of P(s) model curves with different sigma-noises in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (black line), aggregated model data for Step potential; B) Results of comparing segments experiment P(s) vs model data with shifting normalisation point on different metrics. The normalization point shifts within the range of 3 kbp to 2 Mbp. The comparable segment of the plot is 3 kbp-2 Mbp. Sigma-noise is 200 nm.

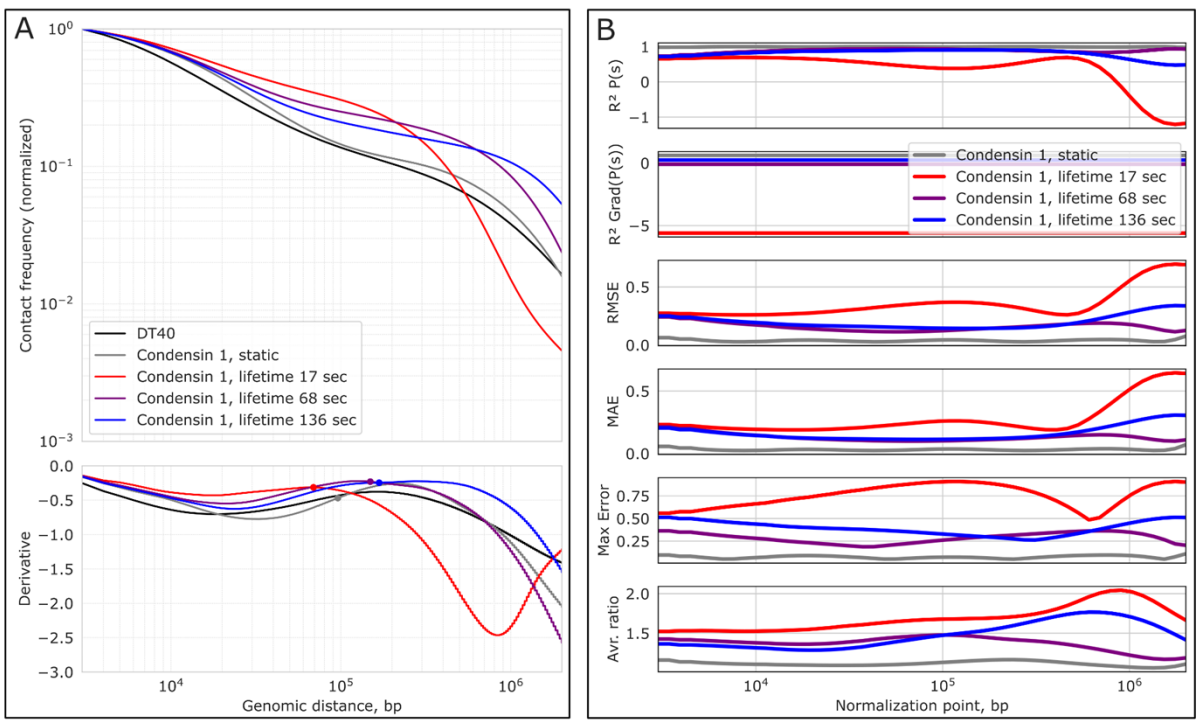

Figure A2. A) P(s) curves in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (black line), aggregated model data. Static (grey line) and dynamic extruders on three lifetimes with base physical params (red, blue, purple lines). Big dots at derivative graphs correspond to mean loop size; B) Results of comparing segments experiment DT40 P(s) vs model data for dynamic extruders on three lifetimes with shifting normalisation point on different metrics. The normalization point shifts within the range of 3 kbp to 2 Mbp. The comparable segment of the plot is 3 kbp-2 Mbp. Sigma-noise is 200 nm.

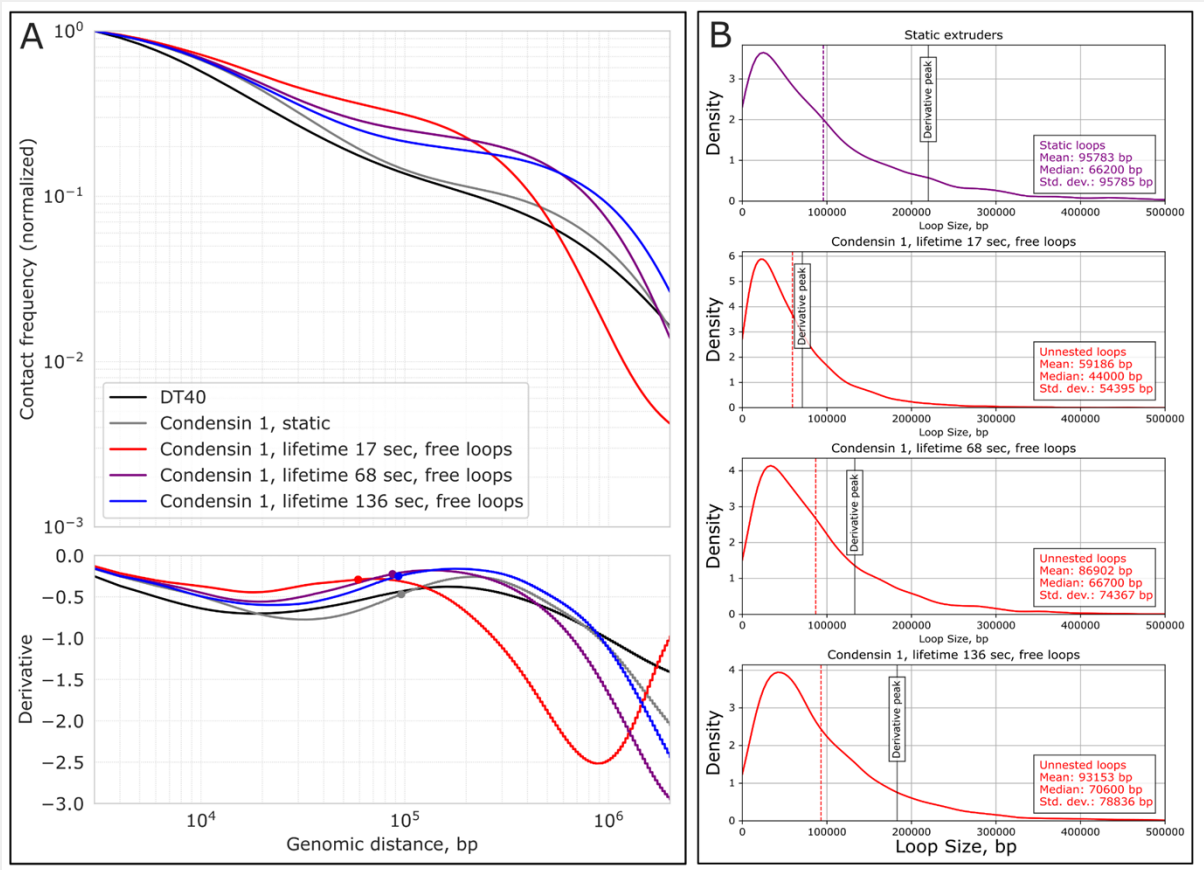

Figure A3. A) P(s) curves in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (black line), aggregated model data. Static (grey line) and dynamic extruders, which are unable to load at existing loops, on three lifetimes for Step potential params (red, blue, purple lines). Big dots at derivative graphs correspond to mean loop size; B) Loop sizes distribution for different lifetimes. All graphs normalised on area under all-loops graph. Purple line - all extruders, red line - unnested extruders, blue line - nested extruders. 1) Static extruders, 2) Dynamic extruders lt = 17 sec, 3) Dynamic extruders lt - 68 sec, 4) Dynamic extruders 136 sec.

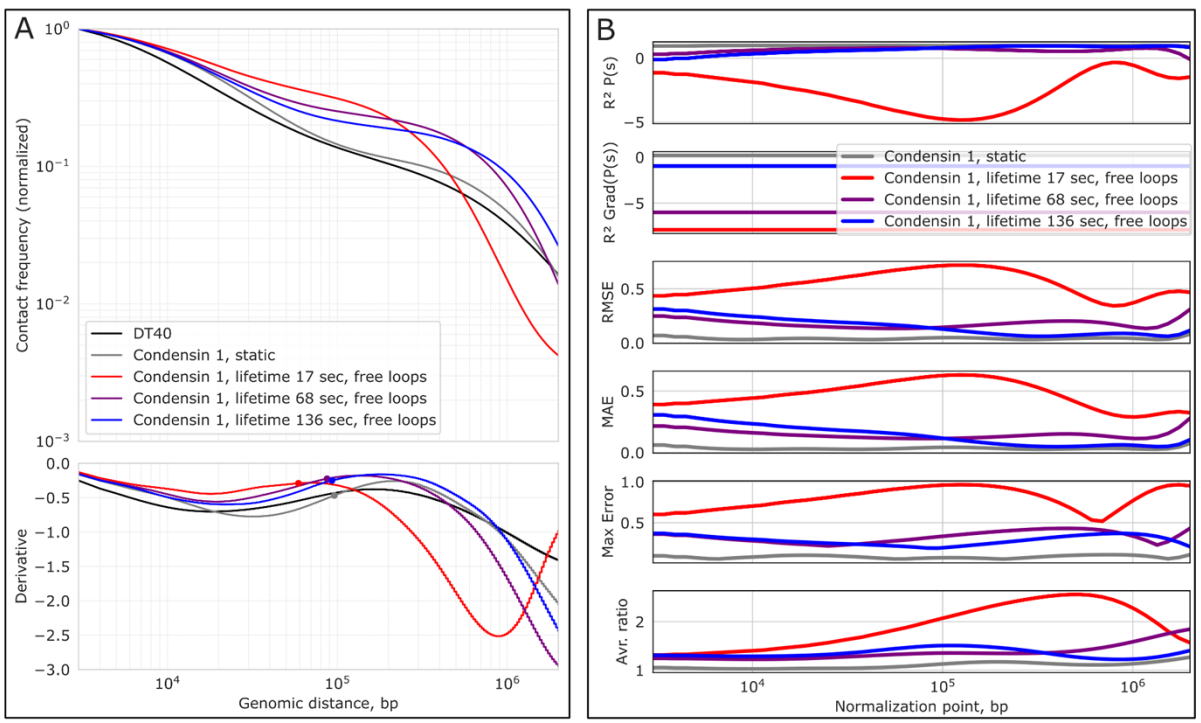

Figure A4. A) P(s) curves in range 3 kbp - 2 Mbp for experimental data for chicken cells line DT40 (black line), aggregated model data. Static (grey line) and dynamic extruders, which are unable to load at existing loops, on three lifetimes with base physical params (red, blue, purple lines). Big dots at derivative graphs correspond to mean loop size; B) Results of comparing segments experiment DT40 P(s) vs model data on different metrics. The normalization point shifts within the range of 3 kbp to 2 Mbp. The comparable segment of the plot is 3 kbp-2 Mbp. Sigma-noise is 200 nm.