

## MIT Open Access Articles

*Multiscale modeling of genome organization  
with maximum entropy optimization*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Lin, Xingcheng, Qi, Yifeng, Latham, Andrew P and Zhang, Bin. 2021. "Multiscale modeling of genome organization with maximum entropy optimization." The Journal of Chemical Physics, 155 (1).

**As Published:** 10.1063/5.0044150

**Publisher:** AIP Publishing

**Persistent URL:** <https://hdl.handle.net/1721.1/141348>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** Creative Commons Attribution 4.0 International license



# Multiscale modeling of genome organization with maximum entropy optimization F

Cite as: J. Chem. Phys. **155**, 010901 (2021); <https://doi.org/10.1063/5.0044150>  
Submitted: 14 January 2021 • Accepted: 28 April 2021 • Published Online: 01 July 2021

 Xingcheng Lin, Yifeng Qi,  Andrew P. Latham, et al.

## COLLECTIONS

F This paper was selected as Featured



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Perspective on integrating machine learning into computational chemistry and materials science](#)

The Journal of Chemical Physics **154**, 230903 (2021); <https://doi.org/10.1063/5.0047760>

[Two-dimensional electronic-vibrational spectroscopy: Exploring the interplay of electrons and nuclei in excited state molecular dynamics](#)

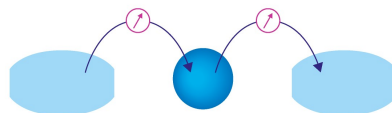
The Journal of Chemical Physics **155**, 020901 (2021); <https://doi.org/10.1063/5.0053042>

[Chemical physics software](#)

The Journal of Chemical Physics **155**, 010401 (2021); <https://doi.org/10.1063/5.0059886>

Webinar

Interfaces: how they make  
or break a nanodevice



March 29th – Register now



Zurich  
Instruments

# Multiscale modeling of genome organization with maximum entropy optimization

Cite as: J. Chem. Phys. 155, 010901 (2021); doi: 10.1063/5.0044150

Submitted: 14 January 2021 • Accepted: 28 April 2021 •

Published Online: 1 July 2021



View Online



Export Citation



CrossMark

Xingcheng Lin,  Yifeng Qi, Andrew P. Latham,  and Bin Zhang<sup>a)</sup> 

## AFFILIATIONS

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

<sup>a)</sup> Author to whom correspondence should be addressed: [binz@mit.edu](mailto:binz@mit.edu)

## ABSTRACT

Three-dimensional (3D) organization of the human genome plays an essential role in all DNA-templated processes, including gene transcription, gene regulation, and DNA replication. Computational modeling can be an effective way of building high-resolution genome structures and improving our understanding of these molecular processes. However, it faces significant challenges as the human genome consists of over  $6 \times 10^9$  base pairs, a system size that exceeds the capacity of traditional modeling approaches. In this perspective, we review the progress that has been made in modeling the human genome. Coarse-grained models parameterized to reproduce experimental data via the maximum entropy optimization algorithm serve as effective means to study genome organization at various length scales. They have provided insight into the principles of whole-genome organization and enabled *de novo* predictions of chromosome structures from epigenetic modifications. Applications of these models at a near-atomistic resolution further revealed physicochemical interactions that drive the phase separation of disordered proteins and dictate chromatin stability *in situ*. We conclude with an outlook on the opportunities and challenges in studying chromosome dynamics.

© 2021 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0044150>

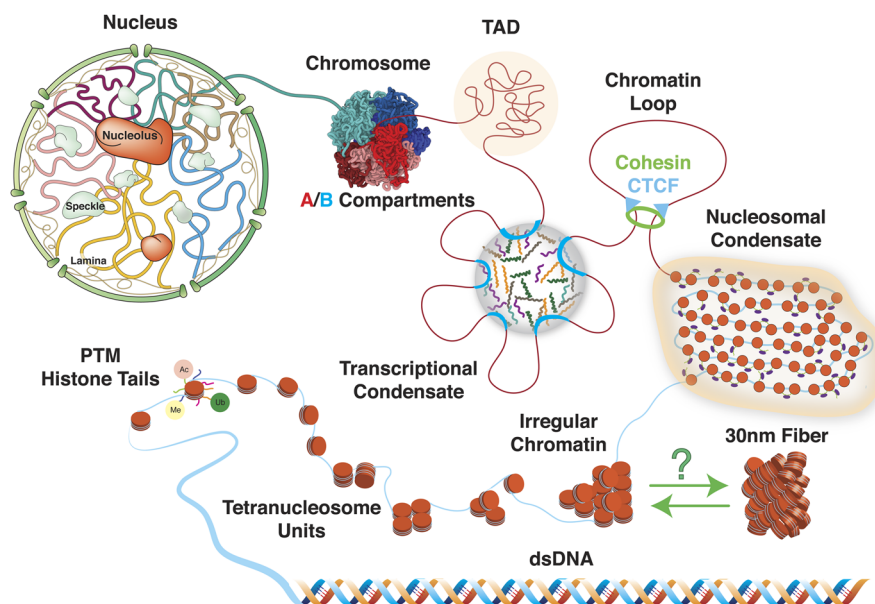
## I. INTRODUCTION

The genome is often hailed as the blueprint of life. It instructs the development and operation of an entire organism. Advancement in sequencing techniques has made possible the determination of nucleic acid sequences that make up the genome for many species, including the completion of the human genome project.<sup>1,2</sup> These sequences provide valuable information about the genome's function and a glimpse into the meaning of life at the atomic level.<sup>3</sup> They allowed discovering and cataloging coding regions or genes that produce protein molecules and define the distinctive features of individual cells. Notably, genes only make up a small fraction of the genome. While the rest of the genome is not transcribed into proteins, it helps encode cellular diversity found in multicellular organisms by regulating the amount of proteins genes produced.<sup>4</sup>

One of the means the non-coding regions regulate gene expression is through three-dimensional (3D) genome organization. The genome, much like enzymes, must fold in 3D to form active sites that catalyze the progression of chemical reactions.<sup>5-7</sup> As shown in Fig. 1, the first layer of eukaryotic genome organization is the

formation of nucleosomes by wrapping DNA around histone proteins. Access to nucleosomal DNA is restricted due to its tight binding with histones.<sup>8</sup> Therefore, regulating nucleosome conformation serves as an effective strategy for controlling genome accessibility and modulating gene expression levels.<sup>9,10</sup> Similar arguments can be made for a string of nucleosomes, i.e., chromatin, as compacting chromatin by stacking nucleosomes close to each other will again limit DNA accessibility and exclude the binding of the transcriptional machinery.<sup>11,12</sup> At even larger scales, genome folding could bring regulatory elements (enhancers and promoters) that are far apart in the linear sequence (10kb to 1Mb) to spatial proximity.<sup>13</sup> Such contacts could achieve targeted control of gene expression by selectively localizing the transcriptional machinery to cell-type-specific genes. A detailed characterization of the genome organization and its dependence on the non-coding regions could improve our understanding of gene regulation in eukaryotes.

Many questions regarding what the genome structure is and how the genome folds remain to be addressed.<sup>14</sup> Globally, the genome organization appears to be poised on the border between order and disorder. While many studies have uncovered



**FIG. 1.** Illustration of the many layers of three-dimensional genome organization. In eukaryotes, the double-stranded DNA first wraps around histone proteins (orange) to form nucleosomes. The N-terminal tails of histone proteins are subject to a wide range of post-translational modifications (PTM), including acetylation (Ac), methylation (Me), and ubiquitination (Ub). A string of nucleosomes, or chromatin, may compact into irregular structures and nucleosomal condensates, although regular structures have also been seen *in vitro*. At larger scales, genomic segments that are far apart in sequence can come in contact due to the formation of chromatin loops, transcriptional condensates, and topologically associating domains (TAD). TADs of similar properties may phase separate, resulting in the compartmentalization of chromosomes into regions enriched with heterochromatin (*B* compartment, blue) or euchromatin (*A* compartment, red). These two chromatin types differ in their compactness, gene density, and nuclear localization. Inside the nucleus, individual chromosomes often occupy non-overlapping regions to form territories.

non-random structural features, the genome is largely amorphous and does not adopt a single unique conformation.<sup>15</sup> How specific interactions emerge from the large pool of non-specific contacts is mostly unknown.<sup>16</sup> It is tempting to assume that the genome adopts a similar sequence to structure relationship as that has been widely accepted for proteins. If so, what exactly is the sequence that dictates the genome organization? The answer to this question is not apparent and cannot be simply the sequence of nucleotides. Epigenetic modifications<sup>17–19</sup> may be the key to forming distinct genome organizations to encode unique gene expression profiles in cells that share identical DNA sequences. Equally important questions arise at smaller scales on the level of single genes (~10kb). While a seemingly more straightforward problem, the structure for a string of nucleosomes remains controversial.<sup>20–22</sup> There is an ongoing debate regarding whether chromatin *in situ* adopts the same set of conformations as those seen *in vitro*. Its interaction with protein molecules inside the nucleus may drive chromatin into phase-separated condensates,<sup>23,24</sup> which renders a high-resolution structural characterization challenging due to its dynamic nature. Addressing these questions will improve our understanding of gene regulatory mechanisms and is promising for uncovering novel genome engineering approaches to alter the genome organization for more targeted and reversible therapeutic approaches. We hope this perspective may introduce the chemical physics community to the fascinating genome folding problem and inspire more theoretical and physical chemistry research efforts.

Much like its success in studying protein folding, computational modeling could be powerful tools for interpreting experimental data, exploring mechanistic hypotheses, and predicting genome organization *de novo*.<sup>25–27</sup> Many simulation tools introduced for protein folding can indeed be transferred to study genome organization. Computational modeling will face unique challenges not seen in protein folding as well, presenting opportunities for the development of new methods and theories. *First*, the human genome consists of over  $6 \times 10^9$  base pairs, a system size that exceeds the capacity of all-atom explicit solvent simulations. Even a small chromatin segment, when fully solvated, can lead to systems of several millions of atoms in size. For whole-genome modeling, coarse-graining will be necessary. *Second*, while the physicochemical interactions that stabilize protein structures are relatively well understood, the same cannot be said for the genome. In particular, many long-range contacts between genomic segments are mediated by protein molecules, the identity of which has yet to be revealed. Since the system's exact composition is lacking, multi-scale approaches<sup>28,29</sup> that aim at deriving coarse-grained force fields from atomistic simulations are inapplicable for parameterizing whole-genome models. Further development of the coarse-grained modeling strategy is needed. *Finally*, the genome is inherently a non-equilibrium system, and its conformational ensemble, in principle, cannot be described with equilibrium statistical mechanics.<sup>30</sup> Efficient algorithms are needed to simulate the impact of ATP-driven enzymes that consume energy to break the detailed balance, so are theoretical tools for interpreting simulation results.

In this perspective, we will review the progress in modeling the human genome and outline the challenges that lie ahead. It is organized as follows: We first discuss modeling efforts at the entire chromosome and the whole-genome level. Particular focus will be devoted to the data-driven mechanistic modeling approach that is effective at unraveling the principles of genome organization and enabling *de novo* structural prediction. In Sec. III, we transition to studies on a much smaller scale for a string of nucleosomes. Arguments will be made to near-atomistic models that can potentially bridge the gap between mesoscopic models and atomistic simulations to characterize chromatin organization *in situ*. While the modeling strategies in Secs. II and III differ significantly, maximum entropy optimization serves as an effective strategy for parameterizing coarse-grained models in both cases. We acknowledge that chromatin organization between the two scales considered here is of critical importance as well, and see the excellent reviews on such topics.<sup>31–33</sup> Finally, we provide an outlook on the opportunities and challenges in modeling chromosome dynamics.

## II. STRUCTURE FOR THE ENTIRE GENOME

The human genome consists of over  $6 \times 10^9$  base pairs, and each one of the 46 chromosomes accounts for tens or hundreds of million bases (MB) in length. When fully extended, the genome accounts for  $\sim 2$  m. Formation of nucleosomes that are 10 nm in size and 200 bp in sequence length reduces the genome length by almost an order of magnitude to 0.3 m, a number that still vastly exceeds the size of a typical cell nucleus ( $\sim 10 \mu\text{m}$  in diameter). Therefore, additional folding and compaction must occur beyond nucleosomes to fit the genome inside the nucleus. We note that most probable polymer configurations are not straight lines but resemble random coils with size  $R_g \propto N^{1/2}a$ ,<sup>34</sup> where  $N$  is the number of nucleosomes and  $a$  is the nucleosome diameter. For a naive estimation with all chromosomes connected together, we have  $N \approx 3 \times 10^7$ ,  $a = 10$  nm, and  $R_g \approx 55 \mu\text{m}$ . Further collapsing coiled configurations into globules that scale as  $N^{1/3}a \approx 3 \mu\text{m}$  does allow the genome to fit inside the nucleus. Great progress has been made toward understanding how the folding proceeds and what the final folded configuration is. Many research groups have contributed to addressing this inherently multi-scale problem using a wide range of tools.

### A. Experimental characterization of genome organization

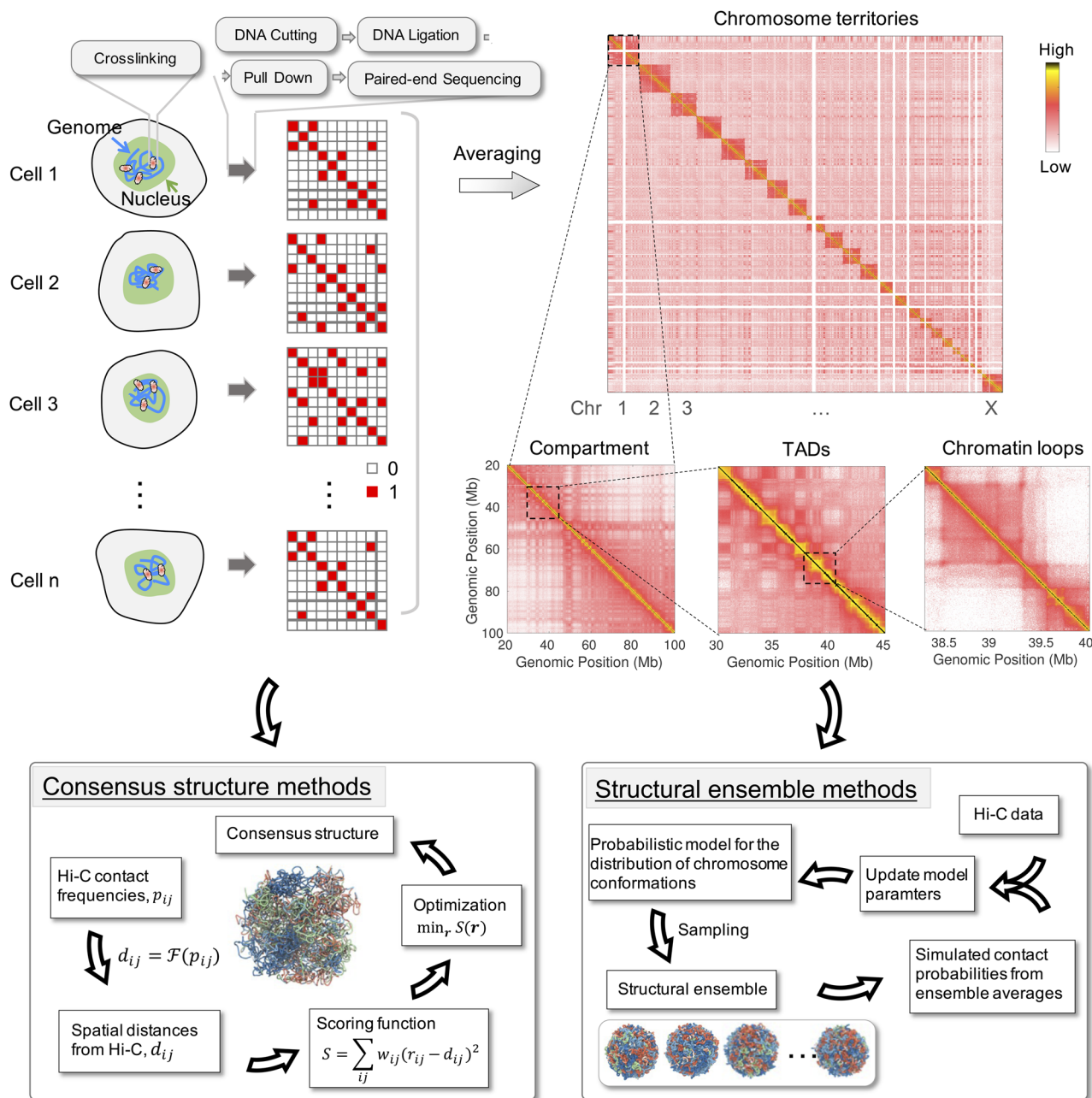
The genome is large and amorphous, and the nucleus environment is heterogeneous and crowded. These features render a high-resolution structural characterization of genome organization challenging. Many well-established techniques that succeed at producing atomic structures for proteins, including x-ray crystallography and Nuclear Magnetic Resonance (NMR), are not directly applicable to the genome. Instead, structural information of the genome was often derived from electron microscopy (EM) or fluorescence *in situ* hybridization (FISH). Applications of EM imaging have led to the discovery of different chromatin types, including heterochromatin and euchromatin that differ in their degree of compactness, gene density, and nuclear localization.<sup>35,36</sup> Electron micrographs of metaphase chromosomes swollen by divalent ions

further revealed the presence of radial loops.<sup>37</sup> Meanwhile, FISH, which uses a probe DNA sequence attached with a fluorophore to hybridize with a corresponding chromosome region via base pairing, allows direct visualization of specific genomic loci with fluorescence microscopy. Experiments that employ multiple probes to measure the spatial distance between genomic loci support the presence of looped structures in interphase as well.<sup>38–40</sup> Labeling whole chromosomes further revealed that they tend to demix and occupy non-overlapping spatial regions termed chromosome territories.<sup>41–43</sup> While imaging-based techniques have successfully characterized chromosome morphology and large scale organization, traditionally, they have not been effective at studying fine-scale structures due to the limited resolution.

High-throughput sequencing-based techniques, including genome-wide chromosome conformation capture (Hi-C) and related methods,<sup>44–46</sup> have served as effective alternatives for studying genome organization. As shown in Fig. 2, Hi-C experiments start by cross-linking DNA segments in spatial proximity inside the nucleus using formaldehyde. The genome is then fragmented with restriction enzymes to result in pairs of contacting segments. Sequencing these contacting pairs reveals their identity, which can be mapped to the reference genome to determine the genomic position of individual loci. For a single structure, this protocol produces a binary contact matrix, each entry of which being either 0 or 1 to indicate whether the corresponding genomic pair is in contact or not. When averaged over a population of millions of cells, as often done in experiments, the protocol produces a contact frequency, or probability, map. The contact map provides rich information on the arrangement of the genome inside the nucleus. Its resolution is bound by the frequency at which restriction enzymes cut the DNA. Since the average spacing between sequences recognized by restriction enzymes is  $\sim 1$  kbp, Hi-C experiments can, in principle, characterize genome organization at high resolution.

Hi-C experiments have confirmed findings from imaging studies and revealed many previously unknown features of genome organization.<sup>47–49</sup> A typical contact probability map from GM12878 cells is shown in Fig. 2, where the probability decreases from yellow to red and to white. Consistent with the formation of chromosome territories, individual chromosomes appear as squares of high contact frequency along the diagonal. Zooming into individual chromosomes revealed block-wise checkerboard patterns that indicate the presence of two chromatin types, often termed A/B compartments<sup>45</sup> and correlated with euchromatin and heterochromatin. The yellow squares appearing along the diagonal line support the formation of topologically associating domains (TADs) with enriched contacts inside the domain than contacts across domains.<sup>50–52</sup> A subset of the TADs exhibits strong contact signals at the corner, supporting the formation of looped structures with pronounced interactions localizing at the ends.<sup>51</sup>

We note that Hi-C experiments are not without limitations. In particular, averaging over a population of cells is often necessary to achieve improved statistics but also casts doubt on the relevance of observed structural motifs in individual cells.<sup>53–55</sup> Hi-C experiments also do not directly measure 3D distances. Sometimes, a higher contact frequency does not translate into a shorter distance,<sup>56,57</sup> rendering a structural interpretation of the contact map less straightforward. The recent development of super-resolution



**FIG. 2.** Computational modeling of genome organization with Hi-C data. Top: illustration of the experimental protocol used in population Hi-C experiments (see the text for details). An example contact probability map for the genome from GM12878 cells is shown on the right, with the probability decreasing from yellow to red and to white. Bottom: illustration of the two popular methods used in building genome structures from Hi-C data. In consensus structure methods (left), pairwise Hi-C contact frequencies ( $p_{ij}$ ) are first transformed to distances ( $d_{ij}$ ) via a mapping function  $\mathcal{F}$ . These distances can be used as constraints to refine computer models and derive consensus structures for the genome. An ensemble of structures can also be used (right) to reproduce Hi-C contact frequencies without converting them to distances. These structural ensemble methods often describe the structures with a probabilistic model and use iterative algorithms to update model parameters.

imaging<sup>58–61</sup> and *in situ* genome sequencing<sup>62</sup> techniques can potentially overcome the shortcomings of traditional imaging and Hi-C to provide high-resolution single-cell structural models for the genome.<sup>43,63</sup>

## B. Computational modeling of genome organization

Computational modeling has been valuable for studying genome organization. In addition to providing intuitive 3D

structural views of the genome, computational approaches are effective at falsifying hypotheses and identifying the ones that are consistent with experimental data.<sup>26,27,64,65</sup> Due to the complexity of the genome organization and our limited knowledge of the folding mechanism, various modeling strategies have been employed in prior studies. We classify them as data-driven approaches that mainly rely on experimental data for structural modeling and hypothesis-driven approaches in which experimental data mostly serve for validation purposes.

### 1. Data-driven modeling approaches

*a. Consensus structure refinement approaches.* Hi-C experiments measure the contact probabilities between genomic loci. Assuming a direct mapping between the probabilities and spatial distances exists, one can convert Hi-C data into a distance matrix. Consensus structures can be derived from this matrix by minimizing scoring functions that measure the difference between simulated and experimental distances (Fig. 2). We note that similar ideas have been widely used in predicting protein structures<sup>66</sup> in which the distances between amino acids can be derived from NMR measurements.

The scoring functions are often complex and are of high dimension. Numerous algorithms have been introduced for their efficient optimization to study genome organization in different organisms.<sup>67–78</sup> Duan *et al.* constructed a three-dimensional model of the *Saccharomyces cerevisiae* genome using the interior-point gradient-based methods.<sup>67</sup> The model effectively captures intra- and inter-chromosomal contacts and reveals notable structural features, including individual chromosomal folding and centromere-anchored inter-chromosomal interactions. Molecular dynamics simulations,<sup>73</sup> Monte Carlo sampling,<sup>76</sup> simulated annealing,<sup>69,70,75</sup> Bayesian inference,<sup>77</sup> and manifold embedding<sup>78–80</sup> have also been used for structure building. These techniques allow for the construction of multiple models consistent with experimental data and have provided insight into the connection between 3D structures and 1D genomic features. Notably, Yildirim and Feig<sup>76</sup> derived structural models for *Caulobacter crescentus* at the single base-pair resolution, thanks to its small genome size. The Mozziconacci laboratory introduced a 3D shortest-path reconstruction method based on the multidimensional scaling-based algorithm.<sup>68</sup> This method is computationally efficient and can produce robust results for sparse and noisy contact maps. Algorithms that separate out intra- and inter-chromosomal contacts for high-resolution structural reconstruction at low computational costs have also been developed.<sup>74</sup> Besides the applications on population Hi-C data, distance-based methods have recently been generalized to single-cell Hi-C<sup>81–84</sup> and ChIA-PET data.<sup>71,72</sup>

*b. Structural ensemble refinement approaches.* The pursuit of a consensus structure, while desirable for simplified interpretations, can mask the intrinsic heterogeneity of genome organization within a cell population. Since population Hi-C experiments report the contact frequencies averaged over many cells, there will likely be an ensemble of structures that collectively reproduce the data.<sup>53,81</sup>

To account for large conformational fluctuations around the mean structure, Sasai and co-workers designed a modeling approach using statistical potentials centered at the distances converted from contact probabilities.<sup>85,86</sup> When applying this approach to the yeast

genome, the authors explained experimental chromosome distributions and uncovered a correlation between the transcriptional level of genes and their spatial distribution.

Numerous groups have developed computational techniques to directly fit Hi-C data with an ensemble of structures and avoid converting experimental contact frequencies to spatial distances, the functional form for which remains unknown.<sup>87–95</sup> For example, Alber and co-workers introduced a probabilistic framework based on the maximum likelihood optimization algorithm to reproduce Hi-C contacts with averages from a population of thousands of genome structures.<sup>87–89</sup> Applications of this algorithm revealed the presence of centromere clusters in GM12878 cells and the role of these clusters in positioning chromosomes in the nucleus. Liang and co-workers introduced CHROMATIX, a method that first identifies a minimal set of sufficient interactions and deconvolutes these interactions into a population of single-cell contact states that can then be used for structure reconstruction.<sup>90–93</sup> This method succeeds in characterizing higher-order interactions between multiple genomic regions and structural changes during embryogenesis in *Drosophila*.

Zhang and Wolynes took a more statistical mechanical approach by transforming the problem of structural refinement into parameterization of pairwise interactions between genomic loci, i.e., force field optimization.<sup>94,96</sup> In particular, assuming that the genome organization from a population of cells can be approximated with the Boltzmann distribution of an energy function, finding the set of structures that reproduce experimental data is equivalent to determining the corresponding energy function. Using experimental data to parameterize or improve force field is indeed a topic well studied in the protein folding community. Low-resolution data, including small-angle x-ray scattering (SAXS), Förster resonance energy transfer (FRET), and nuclear magnetic resonance (NMR), are frequently used to improve the determination of protein structures.<sup>97–99</sup> While many functional forms can be defined for the energy function, the one that introduces the least bias, i.e., the optimal model, is derived by maximizing the excess entropy based on the information theory.<sup>100</sup>

The starting point of the Zhang and Wolynes approach is a beads-on-a-string model to mimic the continuity of the DNA molecule, where each bead represents a genomic segment of fixed length. A corresponding energy function,  $U(\mathbf{r})$ , which usually contains terms that account for bonding, bending, and excluding volume effect, can be defined to evaluate the stability of polymer configurations. Additional energy terms that mimic the confinement effect of neighboring chromosomes and the nuclear envelope can also be included. The equilibrium Boltzmann distribution is

$$P(\mathbf{r}) = \frac{e^{-\beta U(\mathbf{r})}}{\int e^{-\beta U(\mathbf{r})} d\mathbf{r}}, \quad (1)$$

where  $\beta = 1/k_B T$ . Defining a threshold function  $f(r_{ij}; r_o)$  that switches from 0 to 1 as the distance  $r_{ij}$  decreases below  $r_o$ , one can compute the contact probability for a pair of genomic loci as  $p_{ij}^{\text{sim}} = \int f(r_{ij}) P(\mathbf{r}) d\mathbf{r}$ .  $r_o$  can be viewed as the minimal contact distance necessary for cross-linking. Since  $U(\mathbf{r})$  does not capture any biological interactions that chromosomes experience inside the nucleus, the simulated contact probabilities are not expected to reproduce Hi-C data.

To improve the model's biological relevance, correction terms can be introduced to the energy function. Maximum entropy optimization has become widely popular in recent years<sup>99,101–103</sup> as an efficient means to improve the agreement between modeling and experiment. Under the maximum entropy framework, one seeks a new model,  $U_{\text{ME}}(\mathbf{r})$ , and a corresponding probability distribution function  $P_{\text{ME}}(\mathbf{r}) = e^{-\beta U_{\text{ME}}(\mathbf{r})} / \int e^{-\beta U_{\text{ME}}(\mathbf{r})} d\mathbf{r}$ . Importantly, for the new model to reproduce experimental contact probabilities, i.e.,  $\int f(r_{ij})P_{\text{ME}}(\mathbf{r})d\mathbf{r} \equiv p_{ij}^{\text{exp}}$ , while maximizing the excess entropy,  $-\int P_{\text{ME}}(\mathbf{r})\ln[P_{\text{ME}}(\mathbf{r})/P(\mathbf{r})]$ , its energy function has a unique solution defined as

$$U_{\text{ME}}(\mathbf{r}) = U(\mathbf{r}) + \sum_{i>j} \alpha_{ij} f(r_{ij}), \quad (2)$$

where  $\alpha_{ij}$  are Lagrangian multipliers, the value of which ensures the agreement between simulation and experiment. They can be determined via an efficient iterative optimization procedure, as detailed in Ref. 94. We refer the interested readers to Refs. 102 and 104–106 for detailed reviews on maximum entropy optimization.

From the energy function, an ensemble of chromosome structures can be constructed via molecular dynamics simulations. Notably, the structural ensemble is guaranteed to reproduce experimental contact probabilities. By inspecting the simulated structures with knot invariants, Zhang and Wolynes found that interphase chromosomes are free of topological entanglements. Their result generalizes beyond using the scaling exponent of contact probability as a function of sequence separation to support knot-free genome organization.<sup>45,107</sup> Since the maximum entropy model matches the complete set of experimental data with minimal assumptions of genome organization, it produces the least biased or “most likely” chromosome structures based on the given knowledge. Conclusions drawn from these structures are thus more reliable than those based on specific features of Hi-C data. When applied to mitotic chromosomes, the maximum entropy approach further revealed cylindrical conformations with a helical twisting along the central axis. This twisting was later confirmed in experiments with higher resolution Hi-C data.<sup>108</sup> Chu and Wang further studied the conformational transition between interphase and metaphase chromosomes and revealed a two-stage pathway for chromosome compaction.<sup>109</sup>

Tiana and co-workers applied a similar maximum entropy optimization approach to derive pairwise contact energies from Hi-C. Monte Carlo sampling of the resulting energy function predicted chromatin conformation heterogeneity within a single TAD and provided insight into asymmetric expression of X-chromosome inactivation.<sup>110,111</sup> Shi and Thirumalai further applied the maximum entropy approach to recreating the structural ensemble by reproducing the pairwise mean distances converted from Hi-C probability.<sup>112</sup> They characterized the heterogeneity of chromosome structures within the same cell type and across different types.

## 2. Hypothesis-driven modeling approaches

The data-driven modeling approaches are useful for reconstructing chromosome structures. However, they face difficulty at addressing questions regarding why chromosomes adopt specific

conformations. Hypothesis-driven approaches are more suited for exploring genome organization mechanisms and have also been widely used in parallel.

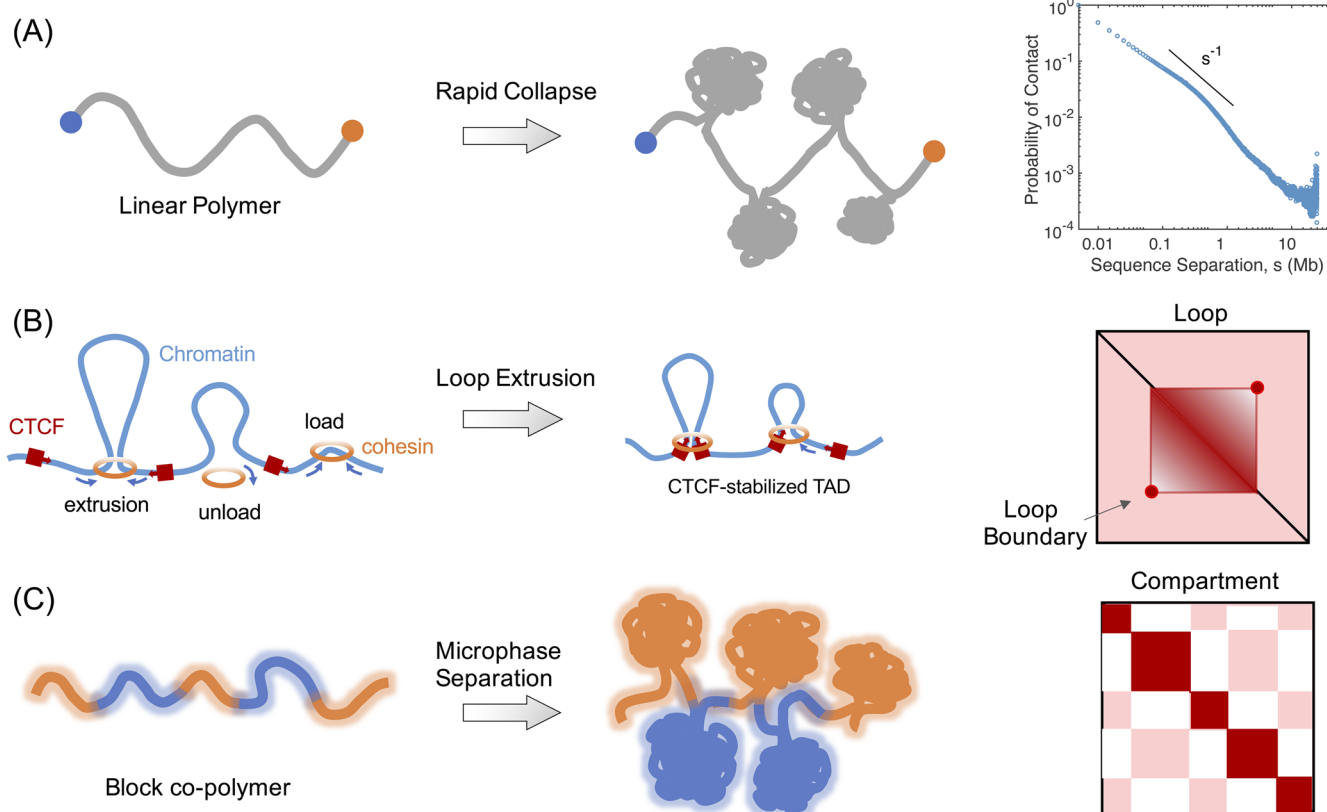
Given that chromosomes are long macromolecules, whose equilibrium conformations will become entangled under confinement, Grosberg *et al.* proposed the fractal globule as a model for chromosome organization<sup>113</sup> [see Fig. 3(a)]. Conformations of the fractal globule are free of entanglements, the presence of which may hinder gene transcription and chromosome separation during mitosis. The fractal globule corresponds to a metastable state that forms when the polymer is rapidly collapsed from an extended conformation. It exhibits a different scaling law from equilibrium conformations regarding the decay of the contact probability as a function of sequence separation. Remarkably, early Hi-C experiments indeed appear to support the scaling exponent predicted by the fractal globule.<sup>45,107</sup> Metastable conformations can persist over the lifetime of a cell, as shown by Rosa and Everaers,<sup>114</sup> due to the slow relaxation time of polymer topology in a crowded environment.

It is worth mentioning that when interpreted strictly, as pointed out by Huang *et al.*,<sup>115</sup> the fractal globule is probably too ideal of a model for genome organization. Its self-similarity property conflicts with the spatial heterogeneity of chromatin packing. Huang *et al.* further introduced a self-returning random walk (SRRW) model that achieves simultaneous high self-interacting frequency and high space-filling heterogeneity. This model makes several predictions consistent with experimental observations, including highly porous chromatin structures, flexible higher-order folding, and irregularly shaped chromosome territory.

In addition to their global topological organization, specific structural features of chromosomes have also been studied extensively. For example, loops have received significant attention in early models of interphase chromosomes to explain the scaling behavior of spatial distances as a function of genomic separation measured by FISH experiments.<sup>38,39,116,117</sup> High-resolution Hi-C data provided direct evidence for the presence of loops in interphase chromosomes.<sup>51</sup> Importantly, connecting the contact map with the underlying DNA sequence further allowed proposing and validating molecular mechanisms of loop formation. For example, the extrusion model<sup>118,119</sup> assumes that these loops form by the processive movement of cohesin molecules that bring close genomic segments far apart in sequence until stopped by CTCF molecules [see Fig. 3(b)]. This model explains the enriched contact within a TAD and the flanking of TADs and loop boundaries with Cohesin and CTCF molecules. Several predictions of the extrusion model have been validated with perturbative Hi-C<sup>120–124</sup> and *in vitro* experiments.<sup>125,126</sup> Other groups have also highlighted the importance of protein binding in chromosome organization.<sup>127,128</sup> In these models, the strong and specific interactions between protein and chromatin drive the contacts within a TAD and the separation between neighboring TADs.

At a larger scale, both imaging experiments and the checkerboard patterns in Hi-C contact maps support the compartmentalization of active and inactive chromatin. Compartmentalization can arise from the microphase separation seen in block copolymers.<sup>129–135</sup> Block copolymers are reasonable models for chromosomes since different chromatin segments can





**FIG. 3.** Illustration of the various mechanisms proposed for genome folding. (a) The metastable fractal globule forms in a process that drives the rapid collapse of a long polymer from an expanded, knotless configuration. The two ends do not have enough time to take part in the collapse, and the polymer remains knotless. (b) In the extrusion model, chromatin loops form as a result of the processive movement of Cohesin molecules along the DNA. CTCF molecules act as blockers to stop Cohesin extrusion, explaining the accumulation of the two proteins at loop boundaries. (c) Microphase separation of block copolymers can lead to contact patterns similar to the compartmentalization seen in Hi-C maps.

exhibit different chemical properties due to their unique histone modification<sup>136</sup> and protein association<sup>137,138</sup> patterns. Additionally, non-equilibrium processes, such as transcription and chromatin remodeling, could contribute to the microphase separation as well. In a simplified model, these processes could give rise to higher effective temperature for active chromatin,<sup>139,140</sup> and the hotter chromatin will separate from the cooler one as a result of their difference in mobility.<sup>141,142</sup> Finally, Wang *et al.* treated the 3D chromosome arrangement as an ellipsoid packing problem to understand the correlation between chromosome positioning and cell geometry.<sup>143</sup>

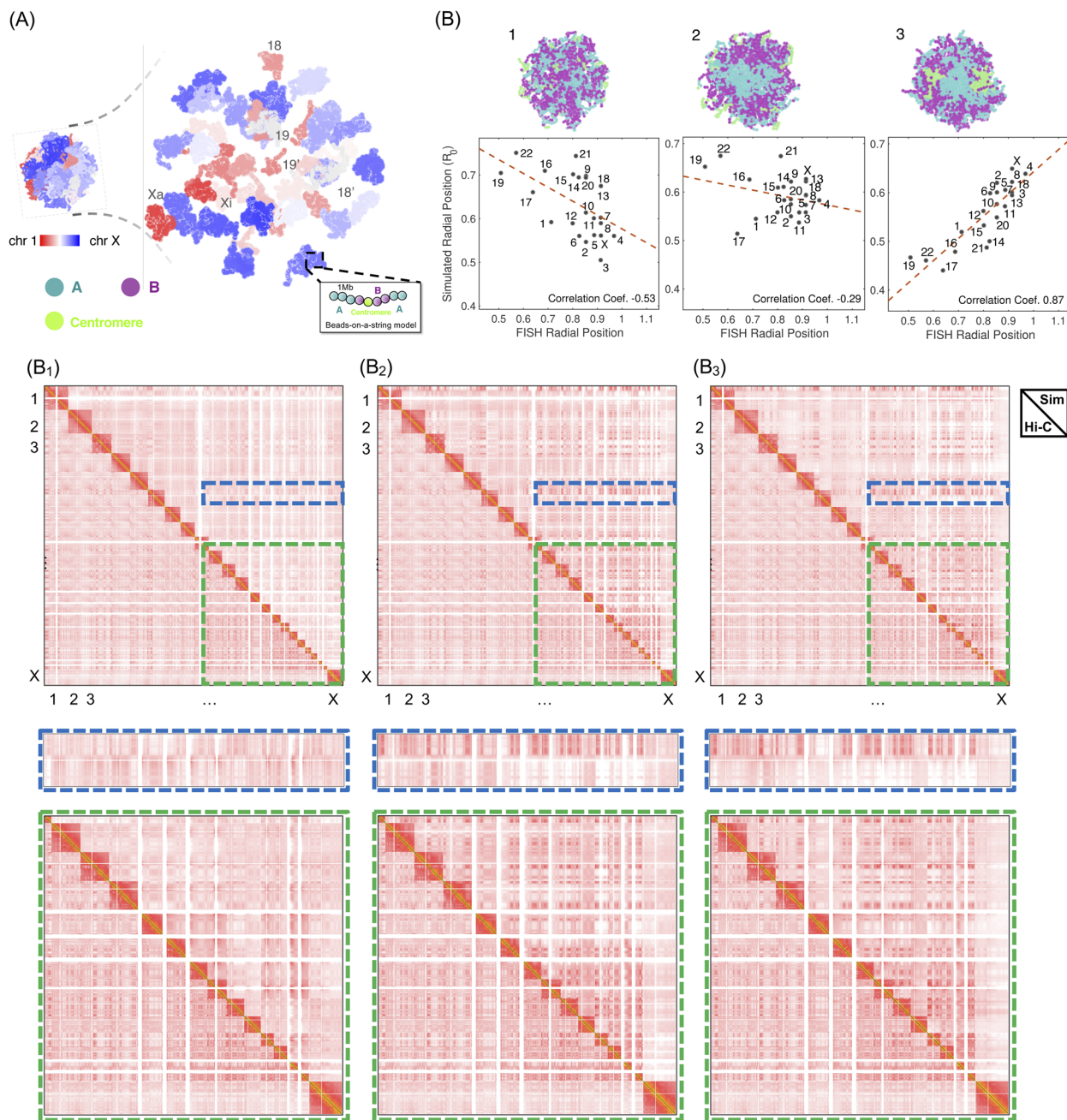
As is evident from the above discussions, multiple models and mechanisms often explain chromosome organization at various scales and reproduce specific features of the experimental data equally well. It is plausible that numerous players co-exist to organize the genome. Quantitatively evaluating the relative significance of various mechanisms is challenging, however, due to the presence of free model parameters whose values are not known *a priori*.

### C. Perspective: Data-driven mechanistic modeling

Coupling the data- and hypothesis-driven approaches together can potentially produce a powerful strategy for modeling genome organization. This strategy ensures the biological relevance of simulated genome structures since all model parameters will be derived from Hi-C data. In the meantime, it will be well suited for mechanistic investigation as the polymer model's energy function will be designed explicitly from biological factors that are known to contribute to genome organization. Pioneering work by Di Pierro *et al.* has shown that a phase separation model among sub-compartments succeeds at reproducing Hi-C data and recapitulating many known aspects of chromosome organization.<sup>134</sup>

#### 1. Exploring the principles of genome organization

We followed the data-driven mechanistic modeling strategy and introduced a polymer model to study the 3D organization of the human diploid genome at 1Mb resolution [Fig. 4(a)].<sup>144</sup> One particular hypothesis that we aimed to evaluate is whether the



**FIG. 4.** Data-driven mechanistic modeling of the whole-genome organization. (a) An example configuration of the diploid human genome colored from red to white and to blue with increasing chromosome ID. Each chromosome is modeled as a string of beads that can either be *A* (cyan) compartments, *B* (purple) compartments, or centromeres (green). (b) Example genome configurations colored by bead types (top), comparison between simulated and experimental chromosome radial positions (bottom), and comparison between simulated (upper triangle) and experimental (lower triangle) genome-wide contact maps for three genome models. In model 1, only one set of parameters was used to model intra- and inter-chromosomal interactions, while two sets of independent parameters were used in model 2. In model 3, in addition to the use of independent parameters for intra- and inter-chromosomal interactions, the centromeric regions were explicitly represented with a new type.

whole-genome organization can be understood from phase separation of different chromatin types. Toward that end, we designed a block copolymer model that includes 46 polymers in spherical confinement. Each polymer represents a chromosome

whose length is determined by the underlying DNA sequence. The spherical confinement mimics the effect of the nuclear envelope. A bead in the polymer was either labeled *A* or *B* to represent the two compartment types seen in Hi-C. The compartment profiles

for individual chromosomes were determined from eigenanalysis on the correlation matrix of intra-chromosomal Hi-C contact maps.<sup>45</sup> For this polymer model, the energy function is defined as follows:

$$U_{\text{Genome}}(\mathbf{r}) = \sum_I U(\mathbf{r}_I) + \sum_I \sum_{i>j \in I} \alpha(|i-j|) f(r_{ij}) + \sum_{i>j} \alpha(T_i, T_j) f(r_{ij}), \quad (3)$$

where  $I$  represents indices over chromosomes and  $i$  and  $j$  represent individual genomic loci, respectively.  $U(\mathbf{r}_I)$  is the generic potential as that in Eq. (2) to account for polymer topology. The second term corresponds to the ideal chromosome potential that depends only on the sequence separation but not on compartment types. It accounts for the interactions within the same chromosome and could arise from protein-mediated contacts<sup>38,116–118,123,127,145,146</sup> and spherical confinement.<sup>93,147</sup> Compartment type ( $T_i \in \{A, B\}$ ) specific interactions are included in the last term. The energy function,  $U_{\text{Genome}}(\mathbf{r})$ , can again be derived by maximizing the excess entropy while enforcing experimental constraints that measure the average contact probabilities at various sequence separations and between different compartment types.<sup>144</sup> Parameters in the energy function,  $\alpha(|i-j|)$  and  $\alpha(T_i, T_j)$ , can be uniquely determined with the iterative optimization algorithm mentioned in the Sec. II B 1 b.

As shown in Fig. 4 B<sub>1</sub>, this model succeeds in recapitulating the microphase separation between the two compartments and the checkerboard pattern seen in Hi-C contact maps. However, a close examination of the simulated genome conformations suggests that heterochromatin ( $B$  compartments) is scattered over the entire nucleus rather than localizing to the nuclear periphery, as shown in EM images.<sup>36</sup> Notably, the unique modeling strategy employed here ensures that the inconsistency between simulation and experiment cannot be resolved by fine-tuning the model parameters, the values of which were inferred from experimental data optimally. The hypotheses introduced in designing the model must be incorrect or incomplete. Therefore, cross-validating simulated structures with imaging experiments offers a feedback mechanism to falsify hypotheses of genome organization mechanisms.

After revisiting the model hypotheses, we found that decoupling the intra- and inter-chromosome interactions is necessary and sufficient for reproducing the peripheral localization of  $B$  compartments (Fig. 4 B<sub>2</sub>). This decoupling indicates the presence of different mechanisms that drive phase separation at various scales. Indeed, it is known that protein-mediated interactions dominate at intra-chromosome contacts.<sup>148</sup> On the other hand, various nuclear landmarks, including lamina and speckles,<sup>149,150</sup> could contribute to the aggregation of various chromosomes. These distinct mechanisms necessitate the use of two sets of independent parameters for intra- and inter-chromosome interactions.

We further found that a third compartment type that corresponds to the central region of chromosomes, or centromeres, is needed to reproduce the radial position of individual chromosomes (Fig. 4 B<sub>3</sub>). Including this type helps capture centromere clustering that has been seen in prior studies.<sup>151</sup> We applied the modeling strategy to study the global rearrangement of genome organization upon tumorigenesis<sup>152</sup> and observed a remarkable change in genome organization upon tumorigenesis. Unlike their interior localization

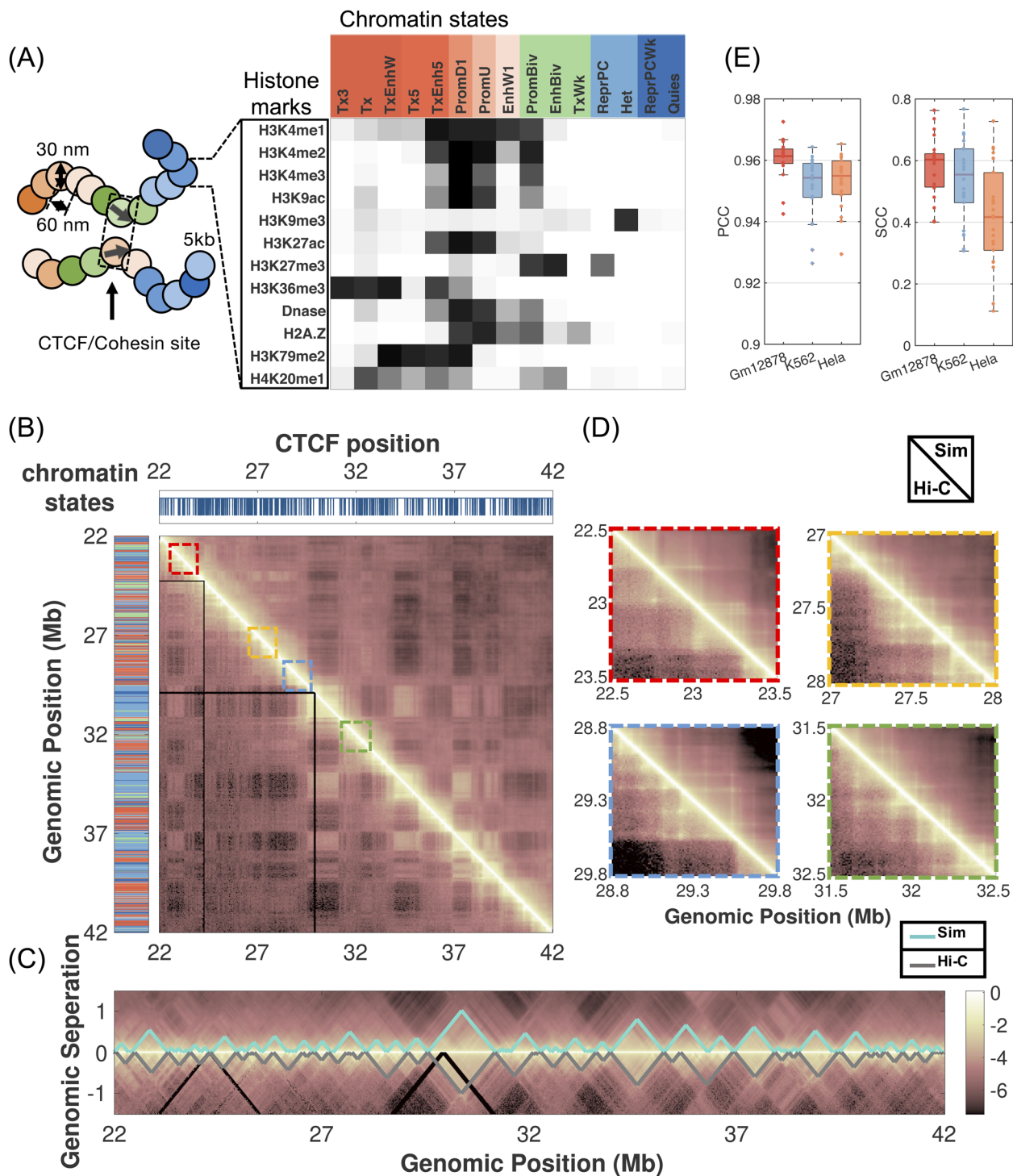
in normal samples,  $A$  compartments in tumors are scattered across the entire nucleus. This observation was later confirmed by electron microscopy.

Most existing computer models do not explicitly include nuclear landmarks such as lamina, nucleoli, and speckles, although studies that account for them are emerging.<sup>85,86,153–156</sup> Modeling the landmarks and the genome together could potentially circumvent the need for an artificial distinction between intra- and inter-chromosome interactions as in the current model. A careful investigation of the coupling between the genome and nuclear lamina may provide insight into the mechanical properties of the nucleus and the impact of mechanical forces on gene transcription.<sup>157,158</sup> Such models could also reveal the role of chromatin network in both the kinetics and thermodynamics of phase separation that drives nucleolus and speckle formation.<sup>156,159,160</sup>

## 2. Predicting chromosome structures from epigenetic marks

In addition to the whole-genome organization, we applied the data-driven mechanistic modeling strategy to study the structure of single chromosomes at 5kb resolution.<sup>161</sup> The model's higher resolution allows detailed characterization of fine-scale structures, including TADs and loops, which play essential roles in gene regulation. Chromosomes can again be viewed as block copolymers, but now the monomer types go beyond the  $A/B$  compartments used in whole-genome modeling. Instead, we represented a chromosome as a sequence of chromatin states that were defined as unique combinatorial patterns of histone modifications [see Fig. 5(a)]. Chromatin states and  $A/B$  compartments are related in many ways. In particular, the states marked with histone modifications H3K4me3 and H3K27ac often correspond to  $A$  compartments, while those marked with H3K9me3 or H3K27me3 overlap with  $B$  compartments. Chromatin states provide a more nuanced view of the variety of chromatin types and a molecular interpretation for the abstract compartments derived from Hi-C data. We note that the use of multiple representations for polymer modeling is not uncommon in the language of protein folding: while two types of amino acids, hydrophobic and hydrophilic, are sufficient to understand the collapse of protein molecules, the chemical specificity of individual amino acids becomes important for predicting high-resolution structures.

We followed the same procedure as whole-genome modeling to define an energy function and derive parameters by enforcing simulated ensemble averages to match average experimental contact probabilities. Details of the model definition and parameterization can be found in Ref. 161. Computer simulations of this model provided a high-resolution structural characterization of chromatin loops, TADs, and compartments. They succeeded in quantitatively reproducing contact probabilities and power-law scaling of 3D contacts, as measured in Hi-C and super-resolution imaging experiments [Figs. 5(b)–5(d)]. In particular, we found that for chromosome 1 of GM12878 cells, the correlation between simulated and experimental contact map exceeds 0.96. The stratum-adjusted correlation coefficient (SCC),<sup>162</sup> which considers the distance-dependence effect of contact maps by stratifying them according to the genomic distance, is 0.7. This value is comparable to that between the experimental replicates. We further examined the



**FIG. 5.** Predicting genome organization with a chromatin-state based polymer model. (a) Overview of the key elements of the computational model. The chromatin is modeled as a string of beads, each assigned with a chromatin state based on the corresponding combinatorial pattern of histone marks. Genomic regions bound by CTCF molecules are also identified to model CTCF mediated loop formation. The polymer model succeeds in quantitatively reproducing compartments (b), TADs (c), and chromatin loops (d) for chromosome 1 from GM12878 cells. (e) The polymer model is transferable across chromosomes and cell types as evidenced by the high correlation between simulated and experimental contact maps measured by Pearson correlation coefficients (PCC—left panel) and stratum-adjusted correlation coefficient (SCC—right panel).

agreement between simulated and experimental contact maps using multiple feature-specific metrics. We found that our model reproduces over 74% of the CTCF-mediated chromatin loops, correctly identifies ~80% of the TAD boundaries,<sup>163</sup> and captures over 57% of the significant enhancer–promoter contacts.

In addition to producing high-resolution chromosome structures, the chromatin-state based model further enables *de novo* prediction. For example, it can be applied to any chromosome, as long as the corresponding sequence of chromatin state is provided as input. Unlike *A/B* compartments that can only be derived from Hi-C data, defining chromatin states only requires histone modifications, and no Hi-C data are needed for structure prediction. We performed additional simulations for chromosomes from GM12878, K562, and HeLa cells, which were not used in model parameterization. We found that the simulated Hi-C data are in good agreement with experimental results as well.

Several other groups have carried out predictive modeling of genome organization using histone modifications with great success as well.<sup>139,164,165</sup> For example, Di Pierro *et al.* simulated chromosome structures at the 50kb resolution using chromatin types defined with histone modification profiles and trained to mimic genomic compartments.<sup>135</sup> MacPherson *et al.* modeled chromosome 16 at the single nucleosome resolution.<sup>164</sup> Using the H3K9me3 pattern derived from ChIP-seq signals, they reproduced the phase separation of euchromatin and heterochromatin seen in Hi-C contact maps. The accuracy of these epigenetic-mark-based models can be further improved. In particular, many protein molecules are known to mediate genomic contacts by directly engaging with the underlying nucleotides. A representation based purely on histone modifications is not sufficient to capture such DNA sequence-specific effects. Encouraging progress is being made at predicting these interactions from the DNA sequence using convolution neural networks.<sup>166–168</sup> Incorporating these studies into the polymer model could provide a more accurate description of genome organization.

### III. STRUCTURE FOR A STRING OF NUCLEOSOMES

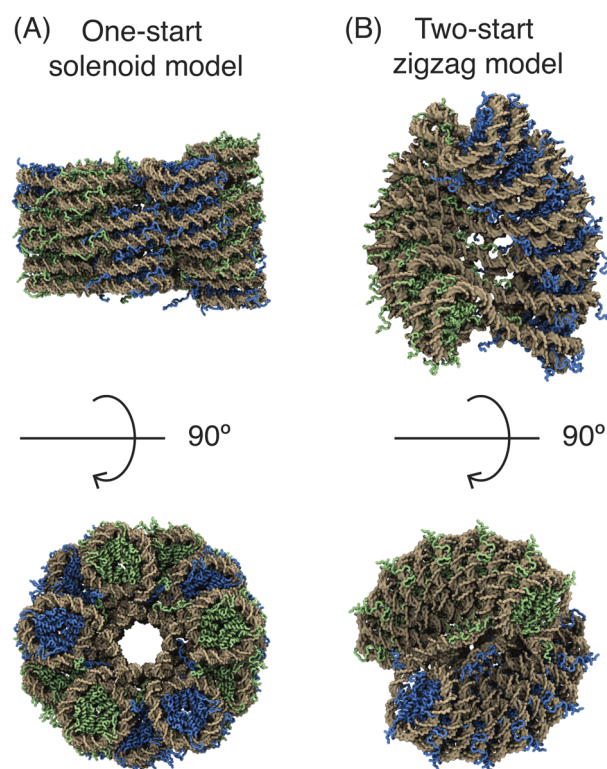
The whole-genome modeling effort and many other studies<sup>169,170</sup> firmly establish the importance of histone modifications in genome organization. Predicting chromosome structures from chromatin states provides further evidence for the sequence–structure relationship of the genome. This relationship suggests that tinkering the post-translational modifications could serve as effective means to alter genome organization and gene expression. Epigenome engineering that aims to modify such epigenetic marks rather than the genetic code has indeed been pursued as a complementary and less invasive approach to genome engineering.<sup>171,172</sup>

To effectively alter chromatin types for epigenome engineering and ensure robust and long-lasting changes in histone modifications, it is essential to study mechanisms that dictate their stability. Histone modifications are subject to constant perturbations from addition and removal enzymes inside the nucleus.<sup>173,174</sup> The chromatin structure has been implicated in mediating the spread and maintenance of histone marks by these enzymes.<sup>136,175–180</sup> Close contacts can facilitate the transfer of enzymes from modified

nucleosomes to unmodified ones, introducing cooperativity among the modifications.<sup>181</sup> Therefore, studying the stability of histone modifications necessitates a detailed characterization of chromatin organization at the resolution of tens of nucleosomes.

#### A. The 30 nm fiber

Packaging the 2-m long human genome inside a nucleus of 10  $\mu\text{m}$  in diameter is a daunting task. Nucleosomes and TADs fold the genome at 200 bp and 1Mb, respectively. Additional compaction occurs to bridge the two rather different length scales. For example, Finch and Klug hypothesized that the 10 nm fiber as a string of nucleosomes could coil into a solenoid to form a 30 nm fiber.<sup>182</sup> As shown in Fig. 6(a), nucleosomes are packed face-to-face around a central axis in the solenoid model, and each helical turn encloses six or seven nucleosomes. Additionally, Woodcock and co-workers introduced the two-start model based on the electron tomography data that exhibited two nucleosome-wide ribbons.<sup>183,184</sup> The two-start model predicts a zigzagging of two nucleosomes that coil into a helical conformation of roughly 30 nm [Fig. 6(b)]. Other fibril models that differ in the nucleosome path have also been proposed.<sup>185</sup> Support of these different structural models mostly come from EM images of the chromatin material extracted from the nucleus under harsh conditions. The use of



**FIG. 6.** Illustration of the two structural models proposed for chromatin fiber. A total of 24 nucleosomes are shown in panel (a) and 12 nucleosomes in panel (b). Histone proteins from the odd and even nucleosomes are shown in blue and green, respectively. The DNA molecule is indicated in gold.

*in vitro* reconstituted nucleosome arrays helped in removing sample heterogeneity in nucleosome spacing and made possible the determination of high-resolution structures with cryo-EM. These studies further support the presence of regular fibril conformations,<sup>186</sup> and the more recent atomic structure favors the two-start zigzag model.<sup>187</sup>

However, the quest for 30 nm fibers inside the nucleus has often led to disappointing conclusions. Many research groups have failed to confirmed its presence employing a wide range of techniques, including cryo-EM,<sup>188</sup> electron spectroscopic imaging,<sup>189</sup> small-angle x-ray scattering,<sup>190,191</sup> super-resolution imaging,<sup>192</sup> Hi-C,<sup>122</sup> and ChromEM tomography.<sup>193</sup> It is becoming increasingly clear that chromatin organization at the kilobase scale is sensitive to a variety of factors,<sup>32,194</sup> including salt concentration, nucleosome spacing, and interaction with non-histone proteins. A systematic characterization of these various factors will be essential for revealing key physicochemical interactions that drive chromatin folding and reconcile the seemingly contradictory experimental observations. However, it has proved challenging due to difficulties in precise chromatin engineering and assembly.

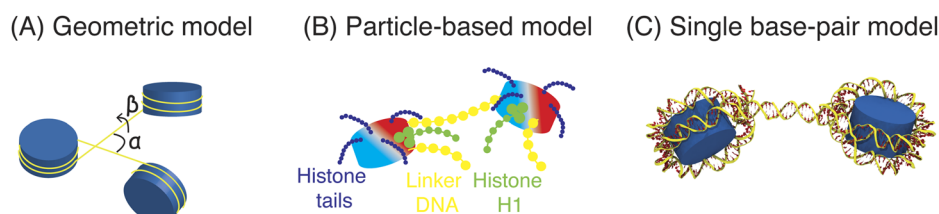
## B. Mesoscopic modeling of chromatin fiber

Computational modeling offers an alternative and promising approach for studying chromatin organization at the kilobase scale. Compared to whole-genome modeling, the problem here is more straightforward, at least conceptually, since the system is well defined. All the components are known, and one can, in principle, carry out all-atom simulations for *de novo* structure prediction.<sup>195,196</sup> Materese *et al.* have shown that such simulations could provide insight into the hydration and electrostatic environments near the nucleosome<sup>197</sup> and the dependence of nucleosome elastic properties on histone variants.<sup>198</sup> Shaytan *et al.* revealed a conformational coupling between histone tails and nucleosomal DNA, with histone tail binding promoting DNA bulging and twisting.<sup>199</sup> The Colleparado-Guevara *et al.* characterized the role of histone modifications on chromatin compaction and the importance of linker histone H1 on chromatin hierarchical looping.<sup>200,201</sup> Winogradoff and Aksimentiev further reported direct observation of spontaneous DNA unwrapping and characterized the role of CG content in such motions.<sup>202</sup> Recently, the Wereszczynski laboratory applied atomistic simulations to study the impact of linker

histones on the structure of an octa-nucleosome array.<sup>203</sup> However, a minimum system for a string of nucleosomes with explicit solvation can easily exceed several millions of atoms,<sup>204</sup> making it challenging to carry out such simulations over long timescales and quantify the stability of various chromatin conformations. Over the years, many coarse-grained models that differ in structural details and energetic terms have been introduced to study chromatin organization.

The two-angle model introduced by Woodcock *et al.* has been instrumental in providing a conceptual framework for studying the chromatin organization<sup>205</sup> [see Fig. 7(a)]. The angles correspond to the DNA entry–exit angle of individual nucleosomes ( $\alpha$ ) and the relative rotational angle between two connecting nucleosomes ( $\beta$ ). Varying the two angles can give rise to a wide variety of fibril conformations with different arrangements of nucleosomes,<sup>206,207</sup> including the two popular models proposed for the 30-nm fiber: one-start solenoid and two-start zigzag. The dependence of chromatin conformation on histone modifications, histone H1 binding, and linker length can be understood from their impact on the angles.<sup>186,208</sup>

The two angle model's success encouraged the development of more refined models that go beyond geometric arguments. For example, accounting for the elasticity of linker DNAs and interactions between nucleosomes<sup>209</sup> allowed Katritch *et al.* to reproduce the force-extension characteristics of chromatin fiber measured by single-molecule pulling experiments.<sup>210</sup> Wedemann *et al.* further incorporated the electrostatic interactions between linker DNAs to quantify the stability and persistence length of the chromatin fiber and characterize the transition between different fiber structures induced by the binding of linker histone H1.<sup>211,212</sup> Schlick and co-workers introduced a chromatin model that employs particle-based representations for histone tails, linker DNA, and linker histones and accounts for the charge distribution of the nucleosome core particle<sup>213,214</sup> [Fig. 7(b)]. This model was shown to reproduce a variety of experimental observations, including the dependence of chromatin conformation on salt concentration<sup>213</sup> and histone modifications,<sup>200,215</sup> and the compaction of chromatin fiber upon the addition of divalent ions and linker histones.<sup>216</sup> Lequieu *et al.* developed the ICPN model that connects nucleosome core particles with flexible linker DNA.<sup>217</sup> Notably, inter-nucleosome interactions were not computed from the mesoscopic representation but parameterized using free energy calculations



**FIG. 7.** Illustration of the three different types of mesoscopic chromatin models that differ in the representation and energetic contributions. (a) In geometric models, the energetics of the chromatin fiber is fully specified by two angles that correspond to the DNA entry–exit angle of individual nucleosomes ( $\alpha$ ) and the relative rotational angle between two connecting nucleosomes ( $\beta$ ). (b) Particle-based models allow for more accurate treatment of the flexibility of linker DNA and histone tails and histone H1. Inter-nucleosome interactions can be introduced to account for contributions from globular domains of histone proteins. (c) Models with DNA molecules at a single base-pair resolution have been introduced to characterize the bending and twisting of linker DNA with greater details.

carried out with a near-atomistic model. Such a multi-scale strategy could lead to a systematic improvement of model accuracy without sacrificing computational efficiency.

The strong dependence of chromatin organization on the linker DNA length further inspired the development of models with a single-base resolution for the DNA [Fig. 7(c)]. For example, Norouzi and Zhurkin used a knowledge-based potential to evaluate the elastic energy of the linker DNA deformations.<sup>218</sup> Together with electrostatic interactions and specific internucleosomal interactions, the resulting model predicted the presence of different topoisomers that are favored by chromatin with linker DNA of  $10n$  and  $10n + 5\text{ bp}$  in length. The model was further applied to interpret experimental chromatin force-extension curves and dissect the contributions of nucleosome unwrapping and unstacking.<sup>219</sup> de Jong *et al.*<sup>220</sup> applied a rigid-base-pair DNA model to describe un- and re-wrapping of nucleosomal DNA and bending and twisting of linker DNA. Monte Carlo simulations with this model support the stability of different fiber models, two-start and one-start, for chromatin with linker lengths of 20 and 50 bp. A similar model that accounts for the bending and twisting of linker DNA was employed by Koslover *et al.*<sup>221</sup> to study the dependence of the nucleosome path in chromatin fibers on the linker DNA length.<sup>186</sup> The model predicts a number of energetically comparable configurations with different nucleosome–nucleosome interaction patterns, indicating a potential role for kinetic trapping in chromatin fiber formation.

### C. Perspective: Near atomistic chromatin modeling

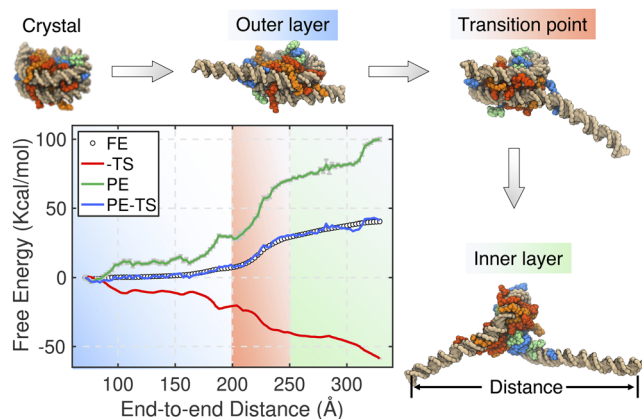
The prior studies have greatly enriched our understanding of the various conformations isolated chromatin can adopt. To further characterize chromatin organization *in situ* and reconcile the difficulty for detecting the 30 nm fiber inside the nucleus, one must study the complex interaction between chromatin and the nucleus environment.<sup>22,200,193</sup> Increasing evidence suggests that intrinsically disordered proteins (IDPs) that interact strongly with chromatin can spontaneously form liquid droplets.<sup>222–228</sup> Collective interactions with these proteins can drive the chromatin into highly dynamic conformations that differ significantly from rigid fibril structures.<sup>229,230</sup> Therefore, accounting for protein chromatin interactions will be essential to model chromatin organization *in situ* but can be challenging with the existing mesoscopic models.

Implicit solvent near-atomistic models offer a promising approach for a detailed characterization of protein–protein and protein–DNA interactions at a single base and residue level. They allow a rigorous treatment of electrostatic, van der Waals, and other interactions between particles based on physical chemistry. Several models have been introduced for the DNA.<sup>231</sup> For example, de Pablo and co-workers developed a DNA model that uses three sites per nucleotide (3SPN).<sup>232–234</sup> Parameters of the model were chosen to reproduce the experimentally measured free energies of base stacking and hybridization. The model was shown to reproduce many properties of DNA molecules accurately, including the persistence length of both single and double-stranded DNA under physiological conditions, the dependence of persistence length on salt concentration and DNA sequence, DNA melting temperatures, and hybridization rate constants.

Integrating 3SPN DNA with similarly coarse-grained protein models has enabled both quantitative and qualitative investigations of chromatin organization. The de Pablo group modeled histone proteins at one bead per amino acid resolution with the atomistic-interaction based coarse-grained protein model (AICG).<sup>235</sup> Interactions among amino acids were parameterized from energies and dynamics of all-atom models via a multiscale protocol. They applied the combined protein–DNA model to reveal that sequence-specific histone binding affinity of the DNA molecule is encoded in their shape.<sup>234</sup> The group further uncovered the coupled role of the DNA-sequence, histone modifications, and chromatin remodelers in positioning nucleosomes<sup>236</sup> and the dependence of nucleosome unwinding barrier on applied tension.<sup>237</sup> Finally, they showed that the model succeeded in reproducing the binding strength between a pair of nucleosomes measured in DNA origami-based force spectrometer experiments.<sup>238,239</sup> Using a closely related protein–DNA model, the Kenzaki and Takada studied the dynamics of DNA unwinding,<sup>240</sup> transcription-factor binding to nucleosomes,<sup>241–243</sup> structures of di- and tri-nucleosomes,<sup>244</sup> and nucleosome sliding/remodeling and twist propagation.<sup>245,246</sup> These studies highlight the usefulness of near-atomistic models in uncovering detailed mechanisms that are otherwise challenging to extract from either experimental studies or mesoscopic models.

To model a nucleosome at a high resolution, we combined 3SPN DNA with the associative memory, water-mediated, structure, and energy model (AWSEM) for protein introduced by Davtyan *et al.*<sup>247</sup> Each amino acid was represented with three beads for the  $C_\alpha$ ,  $C_\beta$ , and  $O$  sites. Interactions among amino acids were optimized following the energy landscape theory prescription to maximize the ratio of folding temperature over glass transition temperature and sculpt a funneled folding landscape for a set of training proteins.<sup>248–250</sup> Similar to the studies of the de Pablo and the Takada groups, protein–DNA interactions were modeled with the Debye–Hückel theory to account for water's dielectric properties and the screened electrostatic interaction. Because of the lack of base-specific hydrogen bonding between histone proteins and the nucleosomal DNA as seen in the crystal structure,<sup>251</sup> electrostatic contributions are expected to dominate the interactions between them.

Using the near-atomistic model for protein and DNA molecules, we determined the free energy landscape for nucleosome unwinding.<sup>252,253</sup> Our study revealed a sizable energetic barrier that decouples the unwinding of the 147 bp long DNA molecule into two separate processes (Fig. 8). This barrier height is in quantitative agreement with the value determined by Brower-Toland *et al.* using single-molecule force spectroscopy<sup>254</sup> and by the Lequieu *et al.* using computer simulations of a closely related near-atomistic model.<sup>237</sup> We note that the mechanistic origin of the energetic barrier has been the focus of several theoretical studies.<sup>255–257</sup> Via a rigorous thermodynamic analysis, we found that the barrier mainly arises from a delayed loss of contacts between disordered histone tails and the DNA. This delay is caused by the dynamical relocation of disordered tails to preserve their DNA contacts. Surprisingly, the energetic penalty is largely offset by an entropic contribution from these newly freed tails. The enthalpy–entropy compensation mechanism provided a fresh perspective on nucleosome stability and highlighted the importance of studying chromatin organization at a high resolution.



**FIG. 8.** Thermodynamics of nucleosome unwinding. The free energy (FE) profile (white dots) as a function of the DNA end-to-end distance supports a three-stage scenario for DNA unwinding. The first stage (blue) corresponds to the unwinding of the outer layer. In the second stage (orange), no significant DNA unwinding occurs, but free energy rises sharply. Finally, the inner layer begins to unwind at a modest free energy cost in the third stage (green). Example nucleosome configurations at different stages are provided on the side, with the DNA indicated in gold and histone proteins indicated in orange, red, blue, and green. The free energy barrier in the transition region is mostly dominated by energetic contributions ( $PE$ , green line), which are compensated by an increase in entropy ( $-TS$ , red line) from the freed histone tails.

We further studied the folding of the basic unit of the chromatin fiber, the tetra-nucleosome.<sup>258</sup> The  $C_{\alpha}$  protein model that represents each amino acid with one bead<sup>259,260</sup> was used to reduce the computational cost. While lacking the sophisticated energy functions used in AWSEM, it suffices to capture the histone octamer's conformational fluctuation around the crystal structure. Direct simulations of tetra-nucleosome folding from extended configurations to collapsed structures are not feasible. Folding will be hindered by slow dynamics that arise from the break and formation of strong, non-specific electrostatic contacts. An advanced sampling technique<sup>261</sup> that combines metadynamics<sup>262</sup>

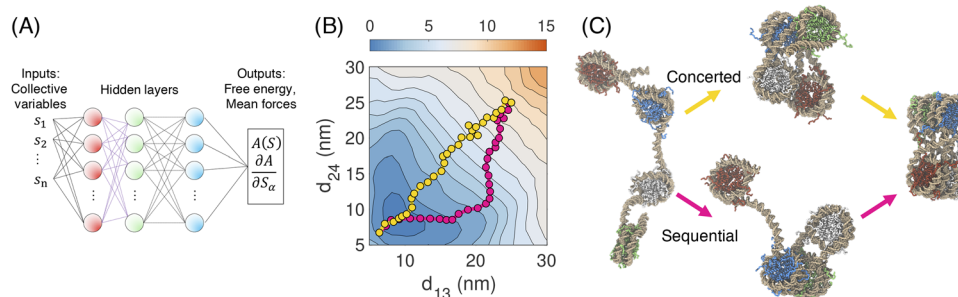
with temperature accelerated molecular dynamics simulations<sup>263,264</sup> enabled efficient conformational sampling and produced various tetra-nucleosome configurations.

To further evaluate the stability of the simulated tetra-nucleosome configurations, we computed the free energy surface as a function of the six inter-nucleosomal distances (Fig. 9). The use of a large number of collective variables is needed to resolve the different conformations but poses challenges to traditional methods such as umbrella sampling. Instead, we determined the surface with a neural network approach by integrating mean forces collected at a series of preselected centers.<sup>265</sup> The free energy surface supports the global stability of the stacked, zigzag configuration resolved in x-ray crystallography<sup>266</sup> and cryo-EM<sup>187</sup> as part of the 30 nm fiber. We determined the most probable folding pathways, or the minimum free energy paths, from open configurations to the zigzag structure using the finite-temperature string method.<sup>267</sup> Notably, the pathways go through intermediate configurations that resemble chromatin configurations observed *in situ*.<sup>268,269</sup> Our study suggests that chromatin configurations observed *in situ* are closely related to the *in vitro* fibril structures and may form as a result of local excitations or unfolding from the global minimum.

The folding intermediates bear comparable stability as the zigzag structure and can be further stabilized by configurational entropy, histone modifications, and variation in the secondary structure of the histone tail. Therefore, chromatin organization is sensitive to both thermal and chemical perturbations. Given the complexity of the nucleus environment, it is perhaps not too surprising to frequently observe the folding intermediates. Chromatin indeed favors more disordered configurations in which proteins mediate contact between non-neighboring nucleosomes,<sup>181</sup> when the binding of the chromatin regulator, Polycomb Repressor Complex 2 (PRC2), was taken into account.

### 1. Study the role of phase separation in chromatin organization with MOFF

The techniques used for tetra-nucleosome folding can be generalized to longer chromatin to evaluate the stability of various fibril structures and the dependence of their stability on the



**FIG. 9.** Stability and folding pathways of the tetra-nucleosome. (a) Illustration of the neural network approach for parameterizing high-dimensional free energy surfaces from mean forces. The neural network takes the six internucleosomal distances ( $S_1, S_2, \dots, S_n$ ) as an input to compute the corresponding free energy  $[A(\mathbf{S})]$  and mean forces ( $\frac{\partial A}{\partial S_i}$ ). (b) Projection of the six-dimensional free energy profile to the distance between 1 and 3 ( $d_{13}$ ) and 2 and 4 ( $d_{24}$ ) nucleosomes. The sequential (pink) and concerted (yellow) pathway for tetra-nucleosome folding are shown on top of the free energy profile with energy unit kcal/mol. (c) Example tetra-nucleosome configurations along the two folding pathways. The DNA molecule is shown in gold, and the histone octamers are shown in green, white, blue, and red.



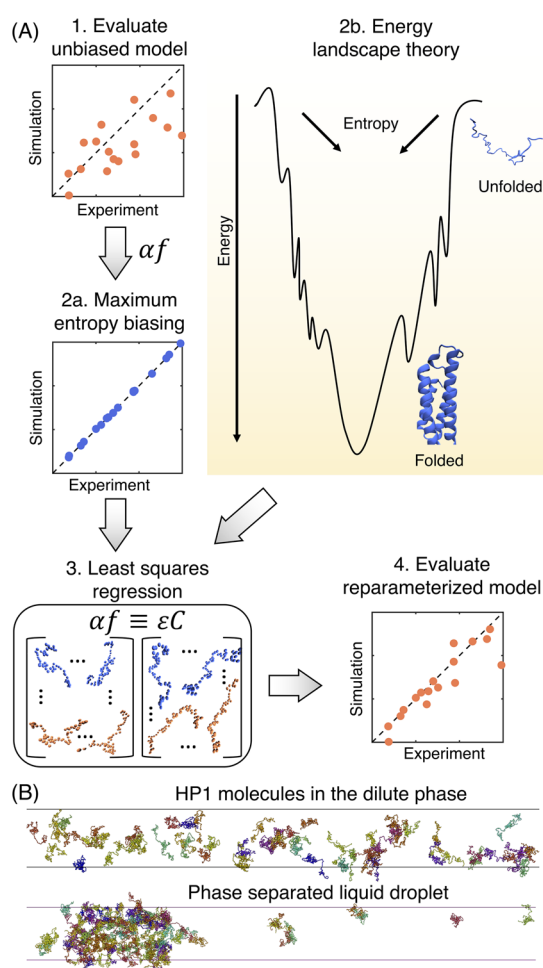
linker DNA length. However, to study the impact of chromatin regulators, a more accurate protein force field is needed. Many chromatin regulators contain disordered regions, and recent evidence suggests that the phase separation of these proteins<sup>270–273</sup> drives changes in the chromatin structure.<sup>222,223,234</sup> We found that the existing coarse-grained force fields, which were often parameterized to fold globular proteins and predict protein structures,<sup>247,274,275</sup> tend to produce overly compact conformations for IDPs. As such, there is widespread interest in developing force fields specifically suited for IDPs.<sup>276–278</sup> Despite the advances, algorithms, which can drive systematic improvements in force field accuracy and ultimately reconcile differences between folded and unfolded models, remain scarce, although improvements have been made in the fully atomistic case.<sup>279–283</sup>

The maximum entropy optimization algorithm introduced in whole-genome modeling is well suited to improve the existing force fields for IDPs. For example, our group and others have incorporated low-resolution experimental data from small-angle x-ray scattering (SAXS) into computer models to refine structure prediction.<sup>284–286</sup> However, an obvious drawback of this approach is that it cannot be applied to proteins for which no experimental data are available.

We developed another algorithm to parameterize a transferable force field for IDPs, which we term maximum entropy optimized force field for IDPs or MOFF-IDP.<sup>287</sup> The algorithm consists of three steps, as outlined in Fig. 10(a). First, simulations are performed for a set of IDPs with an initial force field to *evaluate the unbiased model*. Then, *maximum entropy biasing* is used to determine the least biased correction energies ( $\alpha f$ ) that reproduce protein radii of gyration derived from SAXS experiments. Third, *least squares regression* parameterizes the protein-specific biasing energies with pairwise amino acid potentials ( $\epsilon C$ ) to ensure transferability. These three steps can be repeated to improve the force field accuracy further. The resulting force field, MOFF-IDP, performed well for *de novo* prediction of IDP structural ensembles. For example, it captured the structural rearrangement in the epidermal growth factor receptor C-terminus domain upon phosphorylation.<sup>288</sup>

Just like how the force fields optimized for protein folding and structural prediction cannot be applied to IDPs, most IDP force fields, including MOFF-IDP, are not transferable to globular proteins either. Since both globular proteins and IDPs consist of the same 20 amino acids, it is plausible and desirable to have a consistent force field that describes both types of proteins equally. The maximum entropy optimization algorithm can be readily extended to include folded proteins to improve the force field transferability. Furthermore, it could be combined with other ideas introduced for force field optimization.

Recently, we extended the maximum entropy optimization algorithm with ideas from protein folding studies.<sup>289</sup> For example, energy landscape theory, which succeeded in providing a conceptual framework for studying protein folding kinetics and thermodynamics, requires the protein force field to be funneled toward the native configuration for reliable and efficient folding of globular proteins.<sup>256</sup> Therefore, we added globular proteins to the training set for the first two steps of the algorithm. We further enforced additional constraints for folded proteins when solving the reparameterization equations to require the energy of the native structures to



**FIG. 10.** Coarse-grained protein force field, MOFF, enables large scale simulation of phase separation. (a) Illustration of the maximum entropy optimization algorithm for protein force field parameterization. The algorithm reparameterizes biasing energies ( $\alpha f$ ) determined from maximum entropy optimization with a weighted linear combination of contacts ( $\epsilon C$ ). When solving the reparameterization algorithm with least squares regression, additional constraints can be included for globular proteins to ensure that the native conformations have the lowest energy (step 2b in yellow). (b) Example configurations of HP1 molecules in the dilute and condensed phase.

be lower than that of the misfolded ones for these proteins [see step 2b of Fig. 10(a)]. The reparameterized force field (MOFF) succeeded in predicting the size of both globular and disordered proteins with consistent accuracy.

We have begun to utilize MOFF to study heterochromatin protein 1 (HP1), an essential chromatin regulator known to phase separate. Our results help explain the experimental observations of homolog specific phase separation [Fig. 10(b)].<sup>222,290</sup> When combined with the DNA model mentioned in Sec. III C, MOFF could help address the role of phase-separating proteins in chromatin organization with near-atomistic details.

#### IV. PERSPECTIVE: CHROMOSOME DYNAMICS

Hi-C, imaging, and related methods have provided a comprehensive characterization of genome organization. To connect the structure with function, however, a detailed understanding of chromosome dynamics is required. A prominent example is the eukaryotic gene activation facilitated by the contact formation between enhancers and promoters. Whether these contacts are transient, as in the hit-and-run mechanism,<sup>291</sup> or persistent<sup>292</sup> remains controversial. Distinguishing the two mechanisms have significant implications on genome engineering to control gene expression. Chromosome movement is crucial for the sequence-based homology searches during double-strand break repair,<sup>293</sup> gene rearrangements in antigen receptor repertoire establishment,<sup>294</sup> and telomere length regulation<sup>295</sup> as well. Therefore, a full understanding of the functional role of 3D genome organization can only be achieved by accounting for its dynamical component.

Live-cell imaging has provided great insight into chromatin dynamics.<sup>15,296</sup> When the positions of individual loci were tracked over time, a subdiffusive, visco-elastic motion was observed.<sup>294,297,298</sup> In contrast to regular diffusion, the mean squared displacement of genomic loci scales with time as  $t^\alpha$  with  $\alpha < 1$ . Chromatin dynamics was further found to be highly heterogeneous and is sensitive to a variety of factors, including the nuclear localization,<sup>299,300</sup> the length of the loci,<sup>301,302</sup> and the concentration of lamin A/C protein.<sup>303</sup> Mapping chromatin dynamics across the entire nucleus using Green fluorescent protein (GFP)-tagged histone H2B<sup>304</sup> or fluorescent-labeled DNA<sup>305</sup> uncovered surprisingly long-range correlation among genomic loci over the micrometer scale. This correlated motion is at least partially driven by ATP or transcription.

Experimental findings have inspired numerous theoretical studies.<sup>306,307</sup> Much progress has been made to understand the origin of anomalous diffusion in a crowded environment. The elastic interaction between a locus and its neighboring consecutive segments because of the polymer nature of the genome alone could give rise to sub-diffusive behavior at intermediate timescales.<sup>308</sup> The folded organization of chromosomes can lead to quantitative changes in the scaling exponent.<sup>309,310</sup> Passive interactions between genomic loci and the nuclear compartments could further restrict their diffusivity via the so-called continuous-time random walk (CTRW).<sup>311</sup> Finally, a particle moving through a viscoelastic environment will undergo subdiffusive motion over a range of time scales due to elastic stresses within the medium.<sup>294,312–314</sup> This environment may, in fact, arise from the chromatin itself, which has been shown to exhibit glassy<sup>131</sup> and gel-like property.<sup>315–317</sup>

While theoretical studies have been very successful at deriving general principles, they often have to introduce significant assumptions regarding the biological complexity of the underlying system to make the problem analytically solvable. Consequently, these studies face challenges in explaining the heterogeneity of single-locus dynamics. In addition, generalizing them to the whole genome is non-trivial. The correlated motion among loci is naturally a more complex problem. Hydrodynamics, motors, nuclear compartments, and phase separation all could potentially contribute to such a correlation. Computer simulations that treat the different factors on equal footing could quantify their contributions to chromosome dynamics. The detailed microscopic mechanism arising from numerical simulations could inspire more quantitative theories on chromatin dynamics.

The whole-genome model introduced in the Sec. II will be valuable for studying chromosome dynamics. For example, it will be straightforward to characterize the subdiffusive behavior of telomeres and understand the mechanistic origin that gives rise to the heterogeneity among different loci. In addition to the behavior of individual loci, the dynamical correlation among loci can also be characterized by computing the spatial-temporal correlation defined as

$$C^{\Delta t}(r) = \left\langle \frac{\sum_{i>j} [\Delta \mathbf{r}_i(t; \Delta t) \cdot \Delta \mathbf{r}_j(t; \Delta t)] \delta(r_{ij}(t) - r)}{\sum_{i>j} \delta(r_{ij}(t) - r)} \right\rangle_t, \quad (4)$$

which quantifies the displacement correlations between loci separated by a distance  $r$  over the time interval  $\Delta t$ . The angular brackets represent averaging over time. Prior simulation studies have shown that correlated motions are present within individual chromosomes due to chain organization.<sup>318,319</sup> The whole-genome model offers a unique opportunity to examine the correlation over scales beyond chromosome territories.<sup>304</sup>

A complete understanding of chromosome dynamics cannot be achieved without accounting for the role of ATP-driven remodeling enzymes. These enzymes can affect the dynamics of single loci<sup>320–322</sup> and the large scale correlated motions. The whole-genome model can be modified straightforwardly to study non-equilibrium dynamics, at least approximately, by introducing colored noise to the equation of motion.<sup>323</sup> In addition to numerical simulations, it is crucial to develop analytical approaches that help conceptualize the impact of non-equilibrium motors on the chromatin structure and dynamics. Wang and Wolynes introduced a perturbation theory to map the non-equilibrium steady state in terms of an effective temperature via a systematic expansion of the many-body master equation.<sup>139,324</sup> This approach is appealing as it could open up the door of applying equilibrium statistical mechanical theories for non-equilibrium systems. We followed similar ideas<sup>140,325</sup> to show that the effect of ATP-driven chromatin remodeling enzymes on nucleosome positioning could be well approximated by effective equilibrium models with rescaled temperatures and interactions. Numerical simulations support the theory's accuracy in predicting both kinetic and steady-state quantities, including the effective temperature and the radial distribution function, in biologically relevant regimes. Generalizing these studies could complement numerical simulations to understand the role of molecular motors in genome organization.

It is worth noting that the force field derived from Hi-C data only provides an effective energy landscape, which is not guaranteed to reproduce dynamical measurements. The maximum caliber method,<sup>326,327</sup> which is a generalization of the maximum entropy principle for dynamical trajectories, could be used to incorporate dynamical information<sup>328,329</sup> and improve the dynamical properties of the whole-genome model.

#### AUTHORS' CONTRIBUTIONS

X.L. and Y.Q. contributed equally to this work.

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (Grant Nos. MCB-1715859 and MCB-2042362) and the National Institutes of Health (Grant No. R35GM133580). A.L. acknowledges support by the National Science Foundation Graduate Research Fellowship Program.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

- 1 International Human Genome Sequencing, "Initial sequencing and analysis of the human genome," *Nature* **409**, 860–921 (2001).
- 2 X. Liu, N. Milshina, K. Glasser, K. Nelson, S. Hannehalli, K. Chaturvedi, K. Wolfe, G. L. Gabor Miklos, J. Carnes-Stine, R. Turner, R. Rodriguez, M. Lewis, W. Rowe, F. Lu, M. Caminha, F. Kalush, R. Brandon, Q. Zhang, Y. Lei, A. Glodek, V. Bafna, D. Busam, P. D. Thomas, C. Vech, M. Flanigan, M. Peterson, A. Wang, A. Gluecksmann, R. Sanders, C. Kraft, R. Wides, R. J. Roberts, W. Zhong, J. Ye, D. Gilbert, G. Wang, C. Mobarry, E. Pratts, X. Zhu, L. Curry, C. Fosler, T. McIntosh, H. Gire, B. Neelam, G. Spier, C. Dahlke, H. Zhang, G. G. Sutton, J. C. Venter, G. Subramanian, E. Stewart, H. An, S. Istrail, N. Nguyen, K. A. Ketchum, D. Wu, C. Sitter, L. Kline, M. Zhan, C. Jordan, R. Lippert, S. Esparham, J. Zhang, R. Charlab, B. Hart, J. Abu-Threideh, M. Gorokhov, C. Evangelista, D. Allen, C. Xiao, R. Scott, D. Ma, A. Muruganujan, A. Kejarawal, F. Zhong, N. N. Tint, R. J. Mural, S. Hladun, N. Garg, P. Amanatides, R.-R. Ji, Z. Ke, J. Kasha, M. D. Adams, P. Guan, S. Pan, Z. Gu, M. Donnelly, Z. Lai, E. Beasley, E. Suh, X. H. Zheng, D. Baldwin, T. J. Heiman, M.-H. Wei, J. Peck, E. Venter, C. Yan, J. Jordan, A. K. Naik, J. Hoover, M. Nodell, N. Guo, J. Wetter, H. Qureshi, A. Awe, C. A. Evans, A. Sprague, M. Simpson, T. Howland, A. D. Mays, D. Nusskern, D. B. Rusch, W. Ge, V. D. Francesco, A. J. Levine, S. C. Zhu, J. D. Gocayne, M. Yandell, A. Basu, V. A. McKusick, R. Schwartz, K. Remington, Y. Liang, H. O. Smith, Y.-H. Rogers, X. Wang, N. Zinder, C. Carter, K. V. Sjolander, L. Moy, W. Majoros, H. M. Moore, R. Thomas, G. V. Merkulov, S. Baumhueter, S. Salzberg, J. Johnson, V. Bonazzi, R. M. Ballew, D. Jennings, T. Smith, Z. Y. Wang, J. Heil, A. Delcher, E. W. Myers, M. Moy, V. A. Narayan, I. Dew, W. Gan, M. E. Higgins, J. Wang, R. Strong, H. Baden, R. Desilets, R. A. Holt, T. Hatton, T. Stockwell, J. Houck, F. Gong, V. Puri, S. Kravitz, K. Dodson, F. Mann, B. Karlak, S. Koduru, W. Shao, S. Tse, J. Lopez, L. Chen, M. Wen, A. G. Clark, R. Bolanos, K. Biddick, A. E. Gabrielian, T. Nguyen, B. Shue, K. Eilbeck, S. Yoosheph, L. Doup, C. Pfannkoch, Q. Zhao, K. Beeson, S. Zhao, A. Halpern, D. Fasulo, I. Chandramouliswaran, L. Davenport, A. Cravchik, S. Sato, C. Heiner, S. McCawley, S. Danaher, Z. Deng, S. Windsor, F. Ali, D. May, K. Zaveri, M. L. Cheng, M. Simon, A. Carver, J. Baxendale, S. Broder, D. H. Huson, D. Hostin, X. Lin, R. Guigó, D. Romblad, S. Levy, M. Hunkapiller, C. Ibegwam, A. Yao, C. Haynes, D. Ely, M. Wang, C. Nelson, Y.-H. Chiang, J. Nadeau, L. Zheng, M. Reardon, A. Levitsky, M. Harris, M. Williams, S. Ferreira, B. Ruhfel, P. W. Li, P. Dunn, J. Li, C. Slayman, B. Murphy, P. Caulk, K. Graham, M. Wu, S. Glanowski, L. Florea, M. Coyne, A. Love, S. Murphy, Z. Li, B. Lazareva, J. Zaveri, A. Xia, M. Newman, J. R. Wortman, J. McDaniel, T. Woodage, I. McMullen, L. Kagan, J. Haynes, J. Sun, A. Center, M. J. Campbell, M. Smallwood, L. Blick, K. Diemer, S. Henderson, C. D. Kodira, E. Winn-Deen, A. Zandieh, W. Zhang, B. Walenz, B. Gropman, M. Barnstead, K. Reinert, S. Williams, H. Mi, I. Barrow, M. Cargill, J. F. Abril, A. Narechania, M. Dombroski, J. Scott, S. Dietz, and M. Skupski, "The sequence of the human genome," *Science* **291**, 1304–1351 (2002).
- 3 E. Schrödinger, *What is Life?*, Canto Classics (Cambridge University Press, 2014).
- 4 Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y.-C. Wu, A. R. Pfening, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shores, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K.-H. Farh, S. Feizi, R. Karlic, A.-R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthal, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. M. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L.-H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, and M. Kellis, "Integrative analysis of 111 reference human epigenomes," *Nature* **518**, 317–329 (2015).
- 5 T. Misteli, "Beyond the sequence: Cellular organization of genome function," *Cell* **128**, 787–800 (2007).
- 6 B. van Steensel and E. E. M. Furlong, "The role of transcription in shaping the spatial organization of the genome," *Nat. Rev. Mol. Cell Biol.* **20**, 327 (2019).
- 7 B. Bonev and G. Cavalli, "Organization and function of the 3D genome," *Nat. Rev. Genet.* **17**, 661–678 (2016).
- 8 K. Luger, A. W. Mäder, D. K. Richmond, D. F. Sargent, and T. J. Richmond, "Crystal structure of the nucleosome core particle at 2.8 Å resolution," *Nature* **389**, 251–260 (2002).
- 9 W. K. M. Lai and B. F. Pugh, "Understanding nucleosome dynamics and their links to gene expression and DNA replication," *Nat. Rev. Mol. Cell Biol.* **18**, 548–562 (2017).
- 10 L. Bai and A. V. Morozov, "Gene regulation by nucleosome positioning," *Trends Genet.* **26**, 476–483 (2010).
- 11 B. Li, M. Carey, and J. L. Workman, "The role of chromatin during transcription," *Cell* **128**, 707–719 (2007).
- 12 G. J. Narlikar, R. Sundaramoorthy, and T. Owen-Hughes, "Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes," *Cell* **154**, 490–503 (2013).
- 13 D. U. Gorkin, D. Leung, and B. Ren, "The 3D genome in transcriptional regulation and pluripotency," *Cell Stem Cell* **14**, 762–775 (2014).
- 14 B. Zhang and P. G. Wolynes, "Genomic energy landscapes," *Biophys. J.* **112**, 427–433 (2017).
- 15 T. Cremer, M. Cremer, B. Hübner, H. Strickfaden, D. Smeets, J. Popken, M. Sterr, Y. Markaki, K. Rippe, and C. Cremer, "The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments," *FEBS Lett.* **589**, 2931–2943 (2015).
- 16 J. van Arensbergen, B. van Steensel, and H. J. Bussemaker, "In search of the determinants of enhancer–promoter interaction specificity," *Trends Cell Biol.* **24**, 695–702 (2014).
- 17 B. D. Strahl and C. D. Allis, "The language of covalent histone modifications," *Nature* **403**, 41–45 (2000).
- 18 O. J. Rando, "Combinatorial complexity in chromatin structure and function: Revisiting the histone code," *Curr. Opin. Genet. Dev.* **22**, 148–155 (2012).
- 19 V. W. Zhou, A. Goren, and B. E. Bernstein, "Charting histone modifications and the functional organization of mammalian genomes," *Nat. Rev. Genet.* **12**, 7–18 (2011).
- 20 K. Maeshima, S. Hihara, and M. Eltsov, "Chromatin structure: Does the 30-nm fibre exist *in vivo*?", *Curr. Opin. Cell Biol.* **22**, 291–297 (2010).
- 21 E. Fussner, R. W. Ching, and D. P. Bazett-Jones, "Living without 30 nm chromatin fibers," *Trends Biochem. Sci.* **36**, 1–6 (2011).
- 22 K. Luger, M. L. Dechassa, and D. J. Tremethick, "New insights into nucleosome and chromatin structure: An ordered state or a disordered affair?", *Nat. Rev. Mol. Cell Biol.* **13**, 436–447 (2012).
- 23 S. Sanulli, M. J. Trnka, V. Dharmarajan, R. W. Tibble, B. D. Pascal, A. L. Burlingame, P. R. Griffin, J. D. Gross, and G. J. Narlikar, "HP1 reshapes nucleosome core to promote phase separation of heterochromatin," *Nature* **575**, 390–394 (2019).
- 24 B. A. Gibson, L. K. Doolittle, M. W. G. Schneider, L. E. Jensen, N. Gamarra, L. Henry, D. W. Gerlich, S. Redding, and M. K. Rosen, "Organization of chromatin by intrinsic and regulated phase separation," *Cell* **179**, 470–484.e21 (2019).
- 25 M. A. Marti-Renom and L. A. Mirny, "Bridging the resolution gap in structural modeling of 3D genome organization," *PLoS Comput. Biol.* **7**, e1002125 (2011).
- 26 J. Dekker, M. A. Marti-Renom, and L. A. Mirny, "Exploring the three-dimensional organization of genomes: Interpreting chromatin interaction data," *Nat. Rev. Genet.* **14**, 390–403 (2013).

- <sup>27</sup>C. A. Brackey, D. Marenduzzo, and N. Gilbert, "Mechanistic modeling of chromatin folding to understand function," *Nat. Methods* **17**, 767–775 (2020).
- <sup>28</sup>M. G. Saunders and G. A. Voth, "Coarse-graining methods for computational biology," *Annu. Rev. Biophys.* **42**, 73–93 (2013).
- <sup>29</sup>W. G. Noid, "Perspective: Coarse-grained models for biomolecular systems," *J. Chem. Phys.* **139**, 090901 (2013).
- <sup>30</sup>A. Zidovska, "The self-stirred genome: Large-scale chromatin dynamics, its biophysical origins and implications," *Curr. Opin. Genet. Dev.* **61**, 83–90 (2020).
- <sup>31</sup>J. Moller and J. J. de Pablo, "Bottom-up meets top-down: The crossroads of multiscale chromatin modeling," *Biophys. J.* **118**, 2057–2065 (2020).
- <sup>32</sup>G. Ozer, A. Luque, and T. Schlick, "The chromatin fiber: Multiscale problems and approaches," *Curr. Opin. Struct. Biol.* **31**, 124–139 (2015).
- <sup>33</sup>A. Bendandi, S. Dante, S. R. Zia, A. Diaspro, and W. Rocchia, "Chromatin compaction multiscale modeling: A complex synergy between theory, simulation, and experiment," *Front. Mol. Biosci.* **7**, 15 (2020).
- <sup>34</sup>P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, 1979).
- <sup>35</sup>E. Heitz, *Das Heterochromatin der Moose* (Bornträger, 1928).
- <sup>36</sup>E. Fedorova and D. Zink, "Nuclear architecture and gene regulation," *Biochim. Biophys. Acta, Mol. Cell Res.* **1783**, 2174–2184 (2008).
- <sup>37</sup>M. P. F. Marsden and U. K. Laemmli, "Metaphase chromosome structure: Evidence for a radial loop model," *Cell* **17**, 849–858 (1979).
- <sup>38</sup>R. K. Sachs, G. van den Engh, B. Trask, H. Yokota, and J. E. Hearst, "A random-walk/giant-loop model for interphase chromosomes," *Proc. Natl. Acad. Sci. U. S. A.* **92**, 2710–2714 (1995).
- <sup>39</sup>C. Mu, E. Nkel, R. Eils, S. Dietzel, D. Zink, C. Mehring, G. Wedemann, T. Cremer, and J. E. Rg Langowski, "Compartmentalization of interphase chromosomes observed in simulation and experiment," *J. Mol. Biol.* **285**, 1053–1065 (1999).
- <sup>40</sup>W. A. Bickmore and K. Oghene, "Visualizing the spatial relationships between defined DNA sequences and the axial region of extracted metaphase chromosomes," *Cell* **84**, 95–104 (1996).
- <sup>41</sup>T. Cremer and C. Cremer, "Chromosome territories, nuclear architecture and gene regulation in mammalian cells," *Nat. Rev. Genet.* **2**, 292–301 (2001).
- <sup>42</sup>A. Bolzer, G. Kreth, I. Solovei, D. Koehler, K. Saracoglu, C. Fauth, S. Müller, R. Eils, C. Cremer, M. R. Speicher, and T. Cremer, "Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes," *PLoS Biol.* **3**, e157 (2005).
- <sup>43</sup>J.-H. Su, P. Zheng, S. S. Kinrot, B. Bintu, and X. Zhuang, "Genome-scale imaging of the 3D organization and transcriptional activity of chromatin," *Cell* **182**, 1641–1659.e26 (2020).
- <sup>44</sup>J. Dekker, K. Rippe, M. Dekker, and N. Kleckner, "Capturing chromosome conformation," *Science* **295**, 1306–1311 (2002).
- <sup>45</sup>E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, "Comprehensive mapping of long-range interactions reveals folding principles of the human genome," *Science* **326**, 289–293 (2009).
- <sup>46</sup>E. D. Wit and W. D. Laatz, "A decade of 3C technologies: Insights into nuclear organization," *Genes Dev.* **26**, 11–24 (2012).
- <sup>47</sup>J. H. Gibcus and J. Dekker, "The hierarchy of the 3D genome," *Mol. Cell* **49**, 773–782 (2013).
- <sup>48</sup>J. Fraser, I. Williamson, W. A. Bickmore, and J. Dostie, "An overview of genome organization and how we got there: From FISH to Hi-C," *Microbiol. Mol. Biol. Rev.* **79**, 347–372 (2015).
- <sup>49</sup>A. D. Schmitt, M. Hu, and B. Ren, "Genome-wide mapping and analysis of chromosome architecture," *Nat. Rev. Mol. Cell Biol.* **17**, 743–755 (2016).
- <sup>50</sup>B. D. Pope, T. Ryba, V. Dileep, F. Yue, W. Wu, O. Denas, D. L. Vera, Y. Wang, R. S. Hansen, T. K. Canfield, R. E. Thurman, Y. Cheng, G. Gülsoy, J. H. Dennis, M. P. Snyder, J. A. Stamatoyannopoulos, J. Taylor, R. C. Hardison, T. Kahveci, B. Ren, and D. M. Gilbert, "Topologically associating domains are stable units of replication-timing regulation," *Nature* **515**, 402–405 (2014).
- <sup>51</sup>S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, "A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping," *Cell* **159**, 1665–1680 (2014).
- <sup>52</sup>J. Dekker and E. Heard, "Structural and functional diversity of topologically associating domains," *FEBS Lett.* **589**, 2877–2884 (2015).
- <sup>53</sup>B. Bintu, L. J. Mateo, J.-H. Su, N. A. Sinnott-Armstrong, M. Parker, S. Kinrot, K. Yamaya, A. N. Boettiger, and X. Zhuang, "Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells," *Science* **362**, eaau1783 (2018).
- <sup>54</sup>J. A. Beagan and J. E. Phillips-Cremins, "On the existence and functionality of topologically associating domains," *Nat. Genet.* **52**, 8–16 (2020).
- <sup>55</sup>W. J. Xie, Y. Qi, and B. Zhang, "Characterizing chromatin folding coordinate and landscape with deep learning," *PLoS Comput. Biol.* **16**, e1008262 (2020).
- <sup>56</sup>G. Fudenberg and M. Imakaev, "FISH-ing for captured contacts: Towards reconciling FISH and 3C," *Nat. Methods* **14**, 673–678 (2017).
- <sup>57</sup>G. Shi and D. Thirumalai, "Conformational heterogeneity in human interphase chromosome organization reconciles the FISH and Hi-C paradox," *Nat. Commun.* **10**, 3894 (2019).
- <sup>58</sup>B. J. Beliveau, A. N. Boettiger, G. Nir, B. Bintu, P. Yin, X. Zhuang, and C.-t. Wu, "In situ super-resolution imaging of genomic DNA with OligoSTORM and OligoDNA-PAINT," in *Super-Resolution Microscopy: Methods and Protocols* (Springer, New York, NY, 2017), pp. 231–252.
- <sup>59</sup>S. Wang, J.-H. Su, B. J. Beliveau, B. Bintu, J. R. Moffitt, C.-t. Wu, and X. Zhuang, "Spatial organization of chromatin domains and compartments in single chromosomes," *Science* **353**, 598–602 (2016).
- <sup>60</sup>G. Nir, I. Farabella, C. Pérez Estrada, C. G. Ebeling, B. J. Beliveau, H. M. Sasaki, S. D. Lee, S. C. Nguyen, R. B. McCole, S. Chatteraj, J. Erceg, J. AlHaj Abed, N. M. C. Martins, H. Q. Nguyen, M. A. Hannan, S. Russell, N. C. Durand, S. S. P. Rao, J. Y. Kishi, P. Soler-Vila, M. Di Pierro, J. N. Onuchic, S. P. Callahan, J. M. Schreiner, J. A. Stuckey, P. Yin, E. L. Aiden, M. A. Marti-Renom, and C.-t. Wu, "Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling," *PLoS Genet.* **14**, e1007872 (2018).
- <sup>61</sup>L. Xie, P. Dong, X. Chen, T. S. Hsieh, S. Banala, M. De Marzio, B. P. English, Y. Qi, S. K. Jung, K. R. Kieffer-Kwon *et al.*, "3D ATAC-PALM: Super-resolution imaging of the accessible genome," *Nat. Methods* **17**, 430–436 (2020).
- <sup>62</sup>A. C. Payne, Z. D. Chiang, P. L. Reginato, S. M. Mangiameli, E. M. Murray, C.-C. Yao, S. Markoulaki, A. S. Earl, A. S. Labade, R. Jaenisch, G. M. Church, E. S. Boyden, J. D. Buenostro, and F. Chen, "In situ genome sequencing resolves DNA sequence and structure in intact biological samples," *Science* **371**, eaay3446 (2020).
- <sup>63</sup>Y. Takei, J. Yun, S. Zheng, N. Ollikainen, N. Pierson, J. White, S. Shah, J. Thomassie, S. Suo, C.-H. L. Eng *et al.*, "Integrated spatial genomics reveals global architecture of single nuclei," *Nature* **590**, 344 (2021).
- <sup>64</sup>D. Meluzzi and G. Arya, "Computational approaches for inferring 3D conformations of chromatin from chromosome conformation capture data," *Methods* **181–182**, 24 (2019).
- <sup>65</sup>R. Zhou and Y. Q. Gao, "Polymer models for the mechanisms of chromatin 3D folding: Review and perspective," *Phys. Chem. Chem. Phys.* **22**, 20189–20201 (2020).
- <sup>66</sup>J. T. Seffernick and S. Lindert, "Hybrid methods for combined experimental and computational determination of protein structure," *J. Chem. Phys.* **153**, 240901 (2020).
- <sup>67</sup>Z. Duan, M. Andronescu, K. Schutz, S. McIlwain, Y. J. Kim, C. Lee, J. Shendure, S. Fields, C. A. Blau, and W. S. Noble, "A three-dimensional model of the yeast genome," *Nature* **465**, 363 (2010).
- <sup>68</sup>A. Lesne, J. Riposo, P. Roger, A. Cournac, and J. Mozziconacci, "3D genome reconstruction from chromosomal contacts," *Nat. Methods* **11**, 1141–1143 (2014).
- <sup>69</sup>D. Baù, A. Sanyal, B. R. Lajoie, E. Capriotti, M. Byron, J. B. Lawrence, J. Dekker, and M. A. Marti-Renom, "The three-dimensional folding of the  $\alpha$ -globin gene domain reveals formation of chromatin globules," *Nat. Struct. Mol. Biol.* **18**, 107–115 (2011).
- <sup>70</sup>D. Baù and M. A. Marti-Renom, "Genome structure determination via 3C-based data integration by the integrative modeling platform," *Methods* **58**, 300–306 (2012).

- <sup>71</sup>P. Szalaj, P. J. Michalski, P. Wróblewski, Z. Tang, M. Kadlof, G. Mazzocco, Y. Ruan, and D. Plewczynski, "3D-GNOME: An integrated web service for structural modeling of the 3D genome," *Nucleic Acids Res.* **44**, W288–W293 (2016).
- <sup>72</sup>P. Szalaj, Z. Tang, P. Michalski, M. J. Pietal, O. J. Luo, M. Sadowski, X. Li, K. Radew, Y. Ruan, and D. Plewczynski, "An integrated 3-dimensional genome modeling engine for data-driven simulation of spatial genome organization," *Genome Res.* **26**, 1697–1709 (2016).
- <sup>73</sup>W. J. Xie, L. Meng, S. Liu, L. Zhang, X. Cai, and Y. Q. Gao, "Structural modeling of chromatin integrates genome features and reveals chromosome folding principle," *Sci. Rep.* **7**, 2818 (2017).
- <sup>74</sup>M. R. Segal and H. L. Bengtsson, "Reconstruction of 3D genome architecture via a two-stage algorithm," *BMC Bioinf.* **16**, 373 (2015).
- <sup>75</sup>B. Adhikari, T. Trieu, and J. Cheng, "Chromosome3D: Reconstructing three-dimensional chromosomal structures from Hi-C interaction frequency data using distance geometry simulated annealing," *BMC Genomics* **17**, 886 (2016).
- <sup>76</sup>A. Yildirim and M. Feig, "High-resolution 3D models of *Caulobacter crescentus* chromosome reveal genome structural variability and organization," *Nucleic Acids Res.* **46**, 3937–3952 (2018).
- <sup>77</sup>M. Hu, K. Deng, Z. Qin, J. Dixon, S. Selvaraj, J. Fang, B. Ren, and J. S. Liu, "Bayesian inference of spatial organizations of chromosomes," *PLoS Comput. Biol.* **9**, e1002893 (2013).
- <sup>78</sup>G. Zhu, W. Deng, H. Hu, R. Ma, S. Zhang, J. Yang, J. Peng, T. Kaplan, and J. Zeng, "Reconstructing spatial organizations of chromosomes through manifold learning," *Nucleic Acids Res.* **46**, e50 (2018).
- <sup>79</sup>J. Paulsen, O. Gramstad, and P. Collas, "Manifold based optimization for single-cell 3D genome reconstruction," *PLoS Comput. Biol.* **11**, e1004396 (2015).
- <sup>80</sup>E. Guarnera, Z. W. Tan, and I. N. Berezovsky, "Three-dimensional chromatin ensemble reconstruction via stochastic embedding," (to be published, 2021).
- <sup>81</sup>T. Nagano, Y. Lubling, T. J. Stevens, S. Schoenfelder, E. Yaffe, W. Dean, E. D. Laue, A. Tanay, and P. Fraser, "Single-cell Hi-C reveals cell-to-cell variability in chromosome structure," *Nature* **502**, 59–64 (2013).
- <sup>82</sup>T. J. Stevens, D. Lando, S. Basu, L. P. Atkinson, Y. Cao, S. F. Lee, M. Leeb, K. J. Wohlfahrt, W. Boucher, A. O'Shaughnessy-Kirwan *et al.*, "3D structures of individual mammalian genomes studied by single-cell Hi-C," *Nature* **544**, 59–64 (2017).
- <sup>83</sup>V. Ramani, X. Deng, R. Qiu, K. L. Gunderson, F. J. Steemers, C. M. Disteche, W. S. Noble, Z. Duan, and J. Shendure, "Massively multiplex single-cell Hi-C," *Nat. Methods* **14**, 263–266 (2017).
- <sup>84</sup>L. Tan, D. Xing, C.-H. Chang, H. Li, and X. S. Xie, "Three-dimensional genome structures of single diploid human cells," *Science* **361**, 924–928 (2018).
- <sup>85</sup>N. Tokuda, T. P. Terada, and M. Sasai, "Dynamical modeling of three-dimensional genome organization in interphase budding yeast," *Biophys. J.* **102**, 296–304 (2012).
- <sup>86</sup>N. Tokuda and M. Sasai, "Heterogeneous spatial distribution of transcriptional activity in budding yeast nuclei," *Biophys. J.* **112**, 491–504 (2017).
- <sup>87</sup>H. Tjong, W. Li, R. Kalhor, C. Dai, S. Hao, K. Gong, Y. Zhou, H. Li, X. J. Zhou, M. A. Le Gros, C. A. Larabell, L. Chen, and F. Alber, "Population-based 3D genome structure analysis reveals driving forces in spatial genome organization," *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1663–E1672 (2016).
- <sup>88</sup>R. Kalhor, H. Tjong, N. Jayathilaka, F. Alber, and L. Chen, "Genome architectures revealed by tethered chromosome conformation capture and population-based modeling," *Nat. Biotechnol.* **30**, 90–98 (2012).
- <sup>89</sup>H. Tjong, K. Gong, L. Chen, and F. Alber, "Physical tethering and volume exclusion determine higher-order genome organization in budding yeast," *Genome Res.* **22**, 1295–1305 (2012).
- <sup>90</sup>A. Perez-Rathke, Q. Sun, B. Wang, V. Boeva, Z. Shao, and J. Liang, "CHROMATIX: Computing the functional landscape of many-body chromatin interactions in transcriptionally active loci from deconvolved single cells," *Genome Biol.* **21**, 13 (2020).
- <sup>91</sup>Q. Sun, A. Perez-Rathke, D. M. Czajkowsky, Z. Shao, and J. Liang, "High-resolution single-cell 3D-models of chromatin ensembles during *Drosophila* embryogenesis," *Nat. Comm.* **12**, 205 (2021).
- <sup>92</sup>G. Gürsoy, Y. Xu, and J. Liang, "Spatial organization of the budding yeast genome in the cell nucleus and identification of specific chromatin interactions from multi-chromosome constrained chromatin model," *PLoS Comput. Biol.* **13**, e1005658 (2017).
- <sup>93</sup>G. Gürsoy, Y. Xu, A. L. Kenter, and J. Liang, "Spatial confinement is a major determinant of the folding landscape of human chromosomes," *Nucleic Acids Res.* **42**, 8223–8230 (2014).
- <sup>94</sup>B. Zhang and P. G. Wolynes, "Topology, structures, and energy landscapes of human chromosomes," *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6062–6067 (2015).
- <sup>95</sup>D. Meluzzi and G. Arya, "Recovering ensembles of chromatin conformations from contact probabilities," *Nucleic Acids Res.* **41**, 63–75 (2013).
- <sup>96</sup>B. Zhang and P. P. Wolynes, "Shape transitions and chiral symmetry breaking in the energy landscape of the mitotic chromosome," *Phys. Rev. Lett.* **116**, 248101 (2016).
- <sup>97</sup>J. W. Pitera and J. D. Chodera, "On the use of experimental observations to bias simulated ensembles," *J. Chem. Theory Comput.* **8**, 3445–3451 (2012).
- <sup>98</sup>E. Boura, B. Rozycki, D. Z. Herrick, H. S. Chung, J. Vecer, W. A. Eaton, D. S. Cafiso, G. Hummer, and J. H. Hurley, "Solution structure of the ESCRT-I complex by small-angle X-ray scattering, EPR, and FRET spectroscopy," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9437–9442 (2011).
- <sup>99</sup>B. Roux and J. Weare, "On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method," *J. Chem. Phys.* **138**, 084107 (2013).
- <sup>100</sup>E. T. Jaynes, "Information theory and statistical mechanics. II," *Phys. Rev.* **108**, 171–190 (1957).
- <sup>101</sup>R. Crehuet, P. J. Buigues, X. Salvatella, and K. Lindorff-Larsen, "Bayesian-maximum-entropy reweighting of IDP ensembles based on NMR chemical shifts," *Entropy* **21**, 898 (2019).
- <sup>102</sup>A. Cesari, S. Reißer, and G. Bussi, "Using the maximum entropy principle to combine simulations and solution experiments," *Computation* **6**, 15 (2018).
- <sup>103</sup>M. Reppert, A. R. Roy, J. O. B. Tempkin, A. R. Dinner, and A. Tokmakoff, "Refining disordered peptide ensembles with computational amide I spectroscopy: Application to elastin-like peptides," *J. Phys. Chem. B* **120**, 11395–11404 (2016).
- <sup>104</sup>D. B. Amirkulova and A. D. White, "Recent advances in maximum entropy biasing techniques for molecular dynamics," *Mol. Simul.* **45**, 1285–1294 (2019).
- <sup>105</sup>R. R. Stein, D. S. Marks, and C. Sander, "Inferring pairwise interactions from biological data using maximum-entropy probability models," *PLoS Comput. Biol.* **11**, e1004182 (2015).
- <sup>106</sup>W. Boomsma, J. Ferkinghoff-Borg, and K. Lindorff-Larsen, "Combining experiments and simulations using the maximum entropy principle," *PLoS Comput. Biol.* **10**, e1003406 (2014).
- <sup>107</sup>L. A. Mirny, "The fractal globule as a model of chromatin architecture in the cell," *Chromosome Res.* **19**, 37–51 (2011).
- <sup>108</sup>J. H. Gibcus, K. Samejima, A. Goloborodko, I. Samejima, N. Naumova, J. Nuebler, M. T. Kanemaki, L. Xie, J. R. Paulson, W. C. Earnshaw, L. A. Mirny, and J. Dekker, "A pathway for mitotic chromosome formation," *Science* **359**, eaao6135 (2018).
- <sup>109</sup>X. Chu and J. Wang, "Conformational state switching and pathways of chromosome dynamics in cell cycle," *Appl. Phys. Rev.* **7**, 031403 (2020).
- <sup>110</sup>L. Giorgetti, R. Galupa, E. P. Nora, T. Piolot, F. Lam, J. Dekker, G. Tiana, and E. Heard, "Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription," *Cell* **157**, 950–963 (2014).
- <sup>111</sup>G. Tiana, A. Amitai, T. Pollex, T. Piolot, D. Holcman, E. Heard, and L. Giorgetti, "Structural fluctuations of the chromatin fiber within topologically associating domains," *Biophys. J.* **110**, 1234–1245 (2016).
- <sup>112</sup>G. Shi and D. Thirumalai, "From Hi-C contact map to three-dimensional organization of interphase human chromosomes," *Phys. Rev. X* **11**, 011051 (2021).
- <sup>113</sup>A. Y. Grosberg, S. K. Nechaev, and E. I. Shakhnovich, "The role of topological constraints in the kinetics of collapse of macromolecules," *J. Phys.* **49**, 2095–2100 (1988).
- <sup>114</sup>A. Rosa and R. Everaers, "Structure and dynamics of interphase chromosomes," *PLoS Comput. Biol.* **4**, e1000153 (2008).

- <sup>115</sup>K. Huang, Y. Li, A. R. Shim, R. K. A. Virk, V. Agrawal, A. Eshein, R. J. Nap, L. M. Almossalha, V. Backman, and I. Szleifer, "Physical and data structure of 3D genome," *Sci. Adv.* **6**, eaay4055 (2020).
- <sup>116</sup>C. Münkler and J. Langowski, "Chromosome structure described by a polymer model," *Phys. Rev. E* **57**, 5888–5896 (1998).
- <sup>117</sup>M. Bohn and D. W. Heermann, "Diffusion-driven looping provides a consistent framework for chromatin organization," *PLoS One* **5**, e12218 (2010).
- <sup>118</sup>A. L. Sanborn, S. S. P. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, K. P. Geeting, A. Gnirke, A. Melnikov, D. McKenna, E. K. Stamenova, E. S. Lander, and E. L. Aiden, "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 (2015).
- <sup>119</sup>G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, "Formation of chromosomal domains by loop extrusion," *Cell Rep.* **15**, 2038–2049 (2016).
- <sup>120</sup>J. H. I. Haarhuis, R. H. van der Weide, V. A. Blomen, J. O. Yáñez-Cuna, M. Amendola, M. S. van Ruiten, P. H. L. Krijger, H. Teunissen, R. H. Medema, B. van Steensel, T. R. Brummelkamp, E. de Wit, and B. D. Rowland, "The cohesin release factor WAPL restricts chromatin loop extension," *Cell* **169**, 693–707 (2017).
- <sup>121</sup>G. Fudenberg, N. Abdennur, M. Imakaev, A. Goloborodko, and L. A. Mirny, "Emerging evidence of chromosome folding by loop extrusion," *Cold Spring Harbor Symp. Quant. Biol.* **82**, 45–55 (2017).
- <sup>122</sup>W. Schwarzer, N. Abdennur, A. Goloborodko, A. Pekowska, G. Fudenberg, Y. Loe-Mie, N. A. Fonseca, W. Huber, C. H. Haering, L. Mirny, and F. Spitz, "Two independent modes of chromatin organization revealed by cohesin removal," *Nature* **551**, 51–56 (2017).
- <sup>123</sup>E. P. Nora, A. Goloborodko, A.-L. Valton, J. H. Gibcus, A. Uebersohn, N. Abdennur, J. Dekker, L. A. Mirny, and B. G. Bruneau, "Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization," *Cell* **169**, 930–944.e22 (2017).
- <sup>124</sup>S. S. P. Rao, S.-C. Huang, B. Glenn St Hilaire, J. M. Engreitz, E. M. Perez, K.-R. Kieffer-Kwon, A. L. Sanborn, S. E. Johnstone, G. D. Bascom, I. D. Bochkov, X. Huang, M. S. Shamim, J. Shin, D. Turner, Z. Ye, A. D. Omer, J. T. Robinson, T. Schlick, B. E. Bernstein, R. Casellas, E. S. Lander, and E. L. Aiden, "Cohesin loss eliminates all loop domains," *Cell* **171**, 305–320.e324 (2017).
- <sup>125</sup>Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, and H. Yu, "Human cohesin compacts DNA by loop extrusion," *Science* **366**, 1345–1349 (2019).
- <sup>126</sup>I. F. Davidson, B. Bauer, D. Goetz, W. Tang, G. Wutz, and J.-M. Peters, "DNA loop extrusion by human cohesin," *Science* **366**, 1338–1345 (2019).
- <sup>127</sup>M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, "Complexity of chromatin folding is captured by the strings and binders switch model," *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16173–16178 (2012).
- <sup>128</sup>C. A. Brackley, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, "Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains," *Nucleic Acids Res.* **44**, 3503–3512 (2016).
- <sup>129</sup>D. Jost, P. Carrivain, G. Cavalli, and C. Vaillant, "Modeling epigenome folding: Formation and dynamics of topologically associated chromatin domains," *Nucleic Acids Res.* **42**, 9553–9561 (2014).
- <sup>130</sup>M. Di Pierro, B. Zhang, E. L. Aiden, P. G. Wolynes, and J. N. Onuchic, "Transferable model for chromosome architecture," *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12168–12173 (2016).
- <sup>131</sup>G. Shi, L. Liu, C. Hyeon, and D. Thirumalai, "Interphase human chromosome exhibits out of equilibrium glassy dynamics," *Nat. Commun.* **9**, 3161 (2018).
- <sup>132</sup>S. Liu, L. Zhang, H. Quan, H. Tian, L. Meng, L. Yang, H. Feng, and Y. Q. Gao, "From 1D sequence to 3D chromatin dynamics and cellular functions: A phase separation perspective," *Nucleic Acids Res.* **46**, 9367–9383 (2018).
- <sup>133</sup>R. Laghmach, M. Di Pierro, and D. A. Potoyan, "Mesoscale liquid model of chromatin recapitulates nuclear order of eukaryotes," *Biophys. J.* **118**, 2130–2140 (2020).
- <sup>134</sup>D. Hnisz, K. Shrinivas, R. A. Young, A. K. Chakraborty, and P. A. Sharp, "A phase separation model for transcriptional control," *Cell* **169**, 13–23 (2017).
- <sup>135</sup>M. Di Pierro, R. R. Cheng, E. Lieberman Aiden, P. G. Wolynes, and J. N. Onuchic, "De novo prediction of human chromosome structures: Epigenetic marking patterns encode genome architecture," *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12126–12131 (2017).
- <sup>136</sup>W. J. Xie and B. Zhang, "Learning the formation mechanism of domain-level chromatin states with epigenomics data," *Biophys. J.* **116**, 2047–2056 (2019).
- <sup>137</sup>G. J. Filion, J. G. van Bommel, U. Braunschweig, W. Talhout, J. Kind, L. D. Ward, W. Brugman, I. J. de Castro, R. M. Kerkhoven, H. J. Bussemaker, and B. van Steensel, "Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells," *Cell* **143**, 212–224 (2010).
- <sup>138</sup>F. Erdel and K. Rippe, "Formation of chromatin subcompartments by phase separation," *Biophys. J.* **114**, 2262–2270 (2018).
- <sup>139</sup>S. Wang and P. G. Wolynes, "Communication: Effective temperature and glassy dynamics of active matter," *J. Chem. Phys.* **135**, 000219–342 (2011).
- <sup>140</sup>Z. Jiang and B. Zhang, "Theory of active chromatin remodeling," *Phys. Rev. Lett.* **123**, 208102 (2019).
- <sup>141</sup>N. Ganai, S. Sengupta, and G. I. Menon, "Chromosome positioning from activity-based segregation," *Nucleic Acids Res.* **42**, 4145–4159 (2014).
- <sup>142</sup>A. Y. Grosberg and J. F. Joanny, "Nonequilibrium statistical mechanics of mixtures of particles in contact with different thermostats," *Phys. Rev. E* **92**, 032118 (2015).
- <sup>143</sup>Y. Wang, M. Nagarajan, C. Uhler, and G. V. Shivashankar, "Orientation and repositioning of chromosomes correlate with cell geometry-dependent gene expression," *Mol. Biol. Cell* **28**, 1997–2009 (2017).
- <sup>144</sup>Y. Qi, A. Reyes, S. E. Johnstone, M. J. Aryee, B. E. Bernstein, and B. Zhang, "Data-driven polymer model for mechanistic exploration of diploid genome organization," *Biophys. J.* **119**, 1905–1916 (2020).
- <sup>145</sup>C. A. Brackley, S. Taylor, A. Papantonis, P. R. Cook, and D. Marenduzzo, "Nonspecific bridging-induced attraction drives clustering of DNA-binding proteins and genome organization," *Proc. Natl. Acad. Sci. U. S. A.* **110**, E3605–E3611 (2013).
- <sup>146</sup>D. Krepel, R. R. Cheng, M. Di Pierro, and J. N. Onuchic, "Deciphering the structure of the condensin protein complex," *Proc. Natl. Acad. Sci. U. S. A.* **115**, 11911–11916 (2018).
- <sup>147</sup>H. Kang, Y. G. Yoon, D. Thirumalai, and C. Hyeon, "Confinement-induced glassy dynamics in a model for chromosome organization," *Phys. Rev. Lett.* **115**, 198102 (2015).
- <sup>148</sup>A. R. Strom and C. P. Brangwynne, "The liquid nucleome—Phase transitions in the nucleus at a glance," *J. Cell Sci.* **132**, jcs235093 (2019).
- <sup>149</sup>Y. Chen and A. S. Belmont, "Genome organization around nuclear speckles," *Curr. Opin. Genet. Dev.* **55**, 91–99 (2019).
- <sup>150</sup>B. van Steensel and A. S. Belmont, "Lamina-associated domains: Links with chromosome architecture, heterochromatin, and gene repression," *Cell* **169**, 780–791 (2017).
- <sup>151</sup>A. E. Wiblin, W. Cui, A. J. Clark, and W. A. Bickmore, "Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells," *J. Cell Sci.* **118**, 3861–3868 (2005).
- <sup>152</sup>S. E. Johnstone, A. Reyes, Y. Qi, C. Adriaens, E. Hegazi, K. Pelka, J. H. Chen, L. S. Zou, Y. Drier, V. Hecht *et al.*, "Large-scale topological changes restrain malignant progression in colorectal cancer," *Cell* **182**, 1474–1489 (2020).
- <sup>153</sup>A. Maji, J. A. Ahmed, S. Roy, B. Chakraborty, and M. K. Mitra, "A lamin-associated chromatin model for chromosome organization," *Biophys. J.* **118**, 3041 (2020).
- <sup>154</sup>M. Chiang, D. Michieletto, C. A. Brackley, N. Rattanaivirotkul, H. Mohammed, D. Marenduzzo, and T. Chandra, "Polymer modeling predicts chromosome reorganization in senescence," *Cell Rep.* **28**, 3212–3223.e6 (2019).
- <sup>155</sup>Q. Li, H. Tjong, X. Li, K. Gong, X. J. Zhou, I. Chiolo, and F. Alber, "The three-dimensional genome organization of *Drosophila melanogaster* through data integration," *Genome Biol.* **18**, 145 (2017).
- <sup>156</sup>Y. Qi and B. Zhang, "Chromatin network retards droplet coalescence," [bioRxiv:433564](https://doi.org/10.1101/2021.03.15.433564) (2021).
- <sup>157</sup>C. Uhler and G. V. Shivashankar, "Regulation of genome organization and gene expression by nuclear mechanotransduction," *Nat. Rev. Mol. Cell Biol.* **18**, 717–727 (2017).

- <sup>158</sup>Y. A. Miroshnikova, M. M. Nava, and S. A. Wickström, “Emerging roles of mechanical forces in chromatin regulation,” *J. Cell Sci.* **130**, 2243–2250 (2017).
- <sup>159</sup>D. S. Lee, N. S. Wingreen, and C. P. Brangwynne, “Chromatin mechanics dictates subdiffusion and coarsening dynamics of embedded condensates,” *Nat. Phys.* **17**, 531 (2021).
- <sup>160</sup>Y. Zhang, D. S. Lee, Y. Meir, C. P. Brangwynne, and N. S. Wingreen, “Mechanical frustration of phase separation in the cell nucleus by chromatin,” *Phys. Rev. Lett.* (to be published, 2021).
- <sup>161</sup>Y. Qi and B. Zhang, “Predicting three-dimensional genome organization with chromatin states,” *PLoS Comput. Biol.* **15**, e1007024 (2019).
- <sup>162</sup>T. Yang, F. Zhang, G. G. Yardımcı, F. Song, R. C. Hardison, W. S. Noble, F. Yue, and Q. Li, “HiCRep: Assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient,” *Genome Res.* **27**, 1939–1949 (2017).
- <sup>163</sup>F. Serra, D. Baù, M. Goodstadt, D. Castillo, G. J. Filion, and M. A. Marti-Renom, “Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors,” *PLoS Comput. Biol.* **13**, e1005665 (2017).
- <sup>164</sup>Q. MacPherson, B. Beltran, and A. J. Spakowitz, “Bottom-up modeling of chromatin segregation due to epigenetic modifications,” *Proc. Natl. Acad. Sci. U. S. A.* **115**, 12739–12744 (2018).
- <sup>165</sup>A. Esposito, S. Bianco, A. M. Chiariello, A. Abraham, L. Fiorillo, M. Conte, R. Campanile, and M. Nicodemi, “Polymer physics and machine learning reveal a combinatorial code linking chromatin 3D architecture to 1D epigenetics,” [bioRxiv:433416](https://doi.org/10.1101/2021.04.16.433416) (2021).
- <sup>166</sup>R. Schwessinger, M. Gosden, D. Downes, R. C. Brown, A. M. Oudelaar, J. Telenius, Y. W. Teh, G. Lunter, and J. R. Hughes, “DeepC: Predicting 3D genome folding using megabase-scale transfer learning,” *Nat. Methods* **17**, 1118–1124 (2020).
- <sup>167</sup>G. Fudenberg, D. R. Kelley, and K. S. Pollard, “Predicting 3D genome folding from DNA sequence with Akita,” *Nat. Methods* **17**, 1111–1117 (2020).
- <sup>168</sup>W. Li, W. H. Wong, and R. Jiang, “DeepTACT: Predicting 3D chromatin contacts via bootstrapping deep learning,” *Nucleic Acids Res.* **47**, e60 (2019).
- <sup>169</sup>T. Kouzarides, “Chromatin modifications and their function,” *Cell* **128**, 693–705 (2007).
- <sup>170</sup>T. Suganuma and J. L. Workman, “Signals and combinatorial functions of histone modifications,” *Annu. Rev. Biochem.* **80**, 473–499 (2011).
- <sup>171</sup>P. I. Thakore, J. B. Black, I. B. Hilton, and C. A. Gersbach, “Editing the epigenome: Technologies for programmable transcription and epigenetic modulation,” *Nat. Methods* **13**, 127–137 (2016).
- <sup>172</sup>M. Park, A. J. Keung, and A. S. Khalil, “The epigenome: The next substrate for engineering,” *Genome Biol.* **17**, 183 (2016).
- <sup>173</sup>A. J. Ruthenburg, H. Li, D. J. Patel, and C. David Allis, “Multivalent engagement of chromatin modifications by linked binding modules,” *Nat. Rev. Mol. Cell Biol.* **8**, 983–994 (2007).
- <sup>174</sup>A. Imhof, “Epigenetic regulators and histone modification,” *Briefings Funct. Genomics Proteomics* **5**, 222–227 (2006).
- <sup>175</sup>A. Sood and B. Zhang, “Quantifying epigenetic stability with minimum action paths,” *Phys. Rev. E* **101**, 062409 (2020).
- <sup>176</sup>I. B. Dodd, M. A. Micheelsen, K. Sneppen, and G. Thon, “Theoretical analysis of epigenetic cell memory by nucleosome modification,” *Cell* **129**, 813–822 (2007).
- <sup>177</sup>F. Erdel and E. C. Greene, “Generalized nucleation and looping model for epigenetic memory of histone modifications,” *Proc. Natl. Acad. Sci. U. S. A.* **113**, E4180–E4189 (2016).
- <sup>178</sup>D. Michieletto, E. Orlandini, and D. Marenduzzo, “Polymer model with epigenetic recoloring reveals a pathway for the *de novo* establishment and 3D organization of chromatin domains,” *Phys. Rev. X* **6**, 041047 (2016).
- <sup>179</sup>D. Jost and C. Vaillant, “Epigenomics in 3D: Importance of long-range spreading and specific interactions in epigenomic maintenance,” *Nucleic Acids Res.* **46**, 2252–2264 (2018).
- <sup>180</sup>M. Katava, G. Shi, and D. Thirumalai, “Chromatin dynamics controls epigenetic domain formation,” [bioRxiv:427115](https://doi.org/10.1101/2021.04.16.433416) (2021).
- <sup>181</sup>R. Leicher, E. J. Ge, X. Lin, M. J. Reynolds, W. Xie, T. Walz, B. Zhang, T. W. Muir, and S. Liu, “Single-molecule and in silico dissection of the interaction between Polycomb repressive complex 2 and chromatin,” *Proc. Natl. Acad. Sci. U. S. A.* **117**, 30465–30475 (2020).
- <sup>182</sup>J. T. Finch and A. Klug, “Solenoidal model for superstructure in chromatin,” *Proc. Natl. Acad. Sci. U. S. A.* **73**, 1897–1901 (2006).
- <sup>183</sup>R. A. Horowitz, D. A. Agard, J. W. Sedat, and C. L. Woodcock, “The three-dimensional architecture of chromatin in situ: Electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon,” *J. Cell Biol.* **125**, 1–10 (1994).
- <sup>184</sup>C. L. Woodcock, L. L. Frado, and J. B. Rattner, “The higher-order structure of chromatin: Evidence for a helical ribbon arrangement,” *J. Cell Biol.* **99**, 42–52 (1984).
- <sup>185</sup>S. P. Williams and J. P. Langmore, “Small angle x-ray scattering of chromatin. Radius and mass per unit length depend on linker length,” *Biophys. J.* **59**, 606–618 (1991).
- <sup>186</sup>P. J. J. Robinson, L. Fairall, V. A. T. Huynh, and D. Rhodes, “EM measurements define the dimensions of the ‘30-nm’ chromatin fiber: Evidence for a compact, interdigitated structure,” *Proc. Natl. Acad. Sci. U. S. A.* **103**, 6506–6511 (2006).
- <sup>187</sup>F. Song, P. Chen, D. Sun, M. Wang, L. Dong, D. Liang, R.-M. Xu, P. Zhu, and G. Li, “Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units,” *Science* **344**, 376–380 (2014).
- <sup>188</sup>M. Eltsov, K. M. MacLellan, K. Maeshima, A. S. Frangakis, and J. Dubochet, “Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ,” *Proc. Natl. Acad. Sci. U. S. A.* **105**, 19732–19737 (2008).
- <sup>189</sup>K. Ahmed, H. Dehghani, P. Rugg-Gunn, E. Fussner, J. Rossant, and D. P. Bazett-Jones, “Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo,” *PLoS One* **5**, e10531 (2010).
- <sup>190</sup>Y. Nishino, M. Eltsov, Y. Joti, K. Ito, H. Takata, Y. Takahashi, S. Hihara, A. S. Frangakis, N. Imamoto, T. Ishikawa, and K. Maeshima, “Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure,” *EMBO J.* **31**, 1644–1653 (2012).
- <sup>191</sup>K. Maeshima, R. Rogge, S. Tamura, Y. Joti, T. Hikima, H. Szerlong, C. Krause, J. Herman, E. Seidel, J. DeLuca, T. Ishikawa, and J. C. Hansen, “Nucleosomal arrays self-assemble into supramolecular globular structures lacking 30-nm fibers,” *EMBO J.* **35**, 1115–1132 (2016).
- <sup>192</sup>M. A. Ricci, C. Manzo, M. F. García-Parajo, M. Lakadamyali, and M. P. Cosma, “Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo,” *Cell* **160**, 1145–1158 (2015).
- <sup>193</sup>H. D. Ou, S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, and C. C. O’Shea, “ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells,” *Science* **357**, eaag0025 (2017).
- <sup>194</sup>K. Maeshima, S. Ide, and M. Babokhov, “Dynamic chromatin organization without the 30-nm fiber,” *Curr. Opin. Cell Biol.* **58**, 95–104 (2019).
- <sup>195</sup>T. C. Bishop, “Molecular dynamics simulations of a nucleosome and free DNA,” *J. Biomol. Struct. Dyn.* **22**, 673–685 (2005).
- <sup>196</sup>Z. Li and H. Kono, “Distinct roles of histone H3 and H2A tails in nucleosome stability,” *Sci. Rep.* **6**, 31437 (2016).
- <sup>197</sup>C. K. Materese, A. Savelyev, and G. A. Papoian, “Counterion atmosphere and hydration patterns near a nucleosome core particle,” *J. Am. Chem. Soc.* **131**, 15005–15013 (2009).
- <sup>198</sup>D. P. Melters, M. Pitman, T. Rakshit, E. K. Dimitriadis, M. Bui, G. A. Papoian, and Y. Dalal, “Intrinsic elasticity of nucleosomes is encoded by histone variants and calibrated by their binding partners,” *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24066–24074 (2019).
- <sup>199</sup>A. K. Shaytan, G. A. Armeev, A. Goncarencu, V. B. Zhurkin, D. Landsman, and A. R. Panchenko, “Coupling between histone conformations and DNA geometry in nucleosomes on a microsecond timescale: Atomistic insights into nucleosome functions,” *J. Mol. Biol.* **428**, 221–237 (2016).
- <sup>200</sup>R. Collepardo-Guevara, G. Portella, M. Vendruscolo, D. Frenkel, T. Schlick, and M. Orozco, “Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: A multiscale computational study,” *J. Am. Chem. Soc.* **137**, 10205–10215 (2015).
- <sup>201</sup>A. Sridhar, S. E. Farr, G. Portella, T. Schlick, M. Orozco, and R. Collepardo-Guevara, “Emergence of chromatin hierarchical loops from protein disorder

- and nucleosome asymmetry," *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7216–7224 (2020).
- <sup>202</sup>D. Winogradoff and A. Aksimentiev, "Molecular mechanism of spontaneous nucleosome unraveling," *J. Mol. Biol.* **431**, 323–335 (2019).
- <sup>203</sup>D. C. Woods, F. Rodriguez-Ropero, and J. Wereszczynski, "The dynamic influence of linker histone saturation within the poly-nucleosome array," *J. Mol. Biol.* **433**, 166902 (2021).
- <sup>204</sup>J. Jung, W. Nishima, M. Daniels, G. Bascom, C. Kobayashi, A. Adedoyin, M. Wall, A. Lappala, D. Phillips, W. Fischer, C. S. Tung, T. Schlick, Y. Sugita, and K. Y. Sanbonmatsu, "Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations," *J. Comput. Chem.* **40**, 1919–1930 (2019).
- <sup>205</sup>C. L. Woodcock, S. A. Grigoryev, R. A. Horowitz, and N. Whitaker, "A chromatin folding model that incorporates linker variability generates fibers resembling the native structures," *Proc. Natl. Acad. Sci. U. S. A.* **90**, 9021–9025 (1993).
- <sup>206</sup>H. Schiessel, W. M. Gelbart, and R. Bruinsma, "DNA folding: Structural and mechanical properties of the two-angle model for chromatin," *Biophys. J.* **80**, 1940–1956 (2001).
- <sup>207</sup>P. M. Diesinger and D. W. Heermann, "Two-angle model and phase diagram for chromatin," *Phys. Rev. E* **74**, 031904 (2006).
- <sup>208</sup>H. Wong, J. M. Victor, and J. Mozziconacci, "An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length," *PLoS One* **2**, e877 (2007).
- <sup>209</sup>V. Katritch, C. Bustamante, and W. K. Olson, "Pulling chromatin fibers: Computer simulations of direct physical micromanipulations," *J. Mol. Biol.* **295**, 29–40 (2000).
- <sup>210</sup>Y. Cui and C. Bustamante, "Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure," *Proc. Natl. Acad. Sci. U. S. A.* **97**, 127–132 (2000).
- <sup>211</sup>G. Wedemann and J. Langowski, "Computer simulation of the 30-nanometer chromatin fiber," *Biophys. J.* **82**, 2847–2859 (2002).
- <sup>212</sup>N. Kepper, D. Foethke, R. Stehr, G. Wedemann, and K. Rippe, "Nucleosome geometry and internucleosomal interactions control the chromatin fiber conformation," *Biophys. J.* **95**, 3692–3705 (2008).
- <sup>213</sup>J. Sun, Q. Zhang, and T. Schlick, "Electrostatic mechanism of nucleosomal array folding revealed by computer simulation," *Proc. Natl. Acad. Sci. U. S. A.* **102**, 8180–8185 (2005).
- <sup>214</sup>G. D. Bascom and T. Schlick, "Mesoscale modeling of chromatin fibers," in *Nuclear Architecture and Dynamics* (Elsevier, 2018), pp. 123–147.
- <sup>215</sup>G. Arya, Q. Zhang, and T. Schlick, "Flexible histone tails in a new mesoscopic oligonucleosome model," *Biophys. J.* **91**, 133–150 (2006).
- <sup>216</sup>S. A. Grigoryev, G. Arya, S. Correll, C. L. Woodcock, and T. Schlick, "Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13317–13322 (2009).
- <sup>217</sup>J. Lequeieu, A. Córdoba, J. Moller, and J. J. De Pablo, "ICPN: A coarse-grained multi-scale model of chromatin," *J. Chem. Phys.* **150**, 215102 (2019).
- <sup>218</sup>D. Norouzi and V. B. Zhurkin, "Topological polymorphism of the two-start chromatin fiber," *Biophys. J.* **108**, 2591–2600 (2015).
- <sup>219</sup>D. Norouzi and V. B. Zhurkin, "Dynamics of chromatin fibers: Comparison of Monte Carlo simulations with force spectroscopy," *Biophys. J.* **115**, 1644–1655 (2018).
- <sup>220</sup>B. E. de Jong, T. B. Brouwer, A. Kaczmarczyk, B. Visscher, and J. van Noort, "Rigid basepair Monte Carlo simulations of one-start and two-start chromatin fiber unfolding by force," *Biophys. J.* **115**, 1848–1859 (2018).
- <sup>221</sup>E. F. Koslover, C. J. Fuller, A. F. Straight, and A. J. Spakowitz, "Local geometry and elasticity in compact chromatin structure," *Biophys. J.* **99**, 3941–3950 (2010).
- <sup>222</sup>A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, and G. J. Narlikar, "Liquid droplet formation by HP1 $\alpha$  suggests a role for phase separation in heterochromatin," *Nature* **547**, 236–240 (2017).
- <sup>223</sup>A. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, and G. H. Karpen, "Phase separation drives heterochromatin domain formation," *Nature* **547**, 241–245 (2017).
- <sup>224</sup>A. J. Plys, C. P. Davis, J. Kim, G. Rizki, M. M. Keenen, S. K. Marr, and R. E. Kingston, "Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2," *Genes Dev.* **33**, 799–813 (2019).
- <sup>225</sup>R. Tatavosian, S. Kent, K. Brown, T. Yao, H. N. Duc, T. N. Huynh, C. Y. Zhen, B. Ma, H. Wang, and X. Ren, "Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation," *J. Biol. Chem.* **294**, 1451–1463 (2019).
- <sup>226</sup>B. R. Sabari, A. Dall'Agnese, A. Boija, I. A. Klein, E. L. Coffey, K. Shrinivas, B. J. Abraham, N. M. Hannett, A. V. Zamudio, J. C. Manteiga, C. H. Li, Y. E. Guo, D. S. Day, J. Schuijers, E. Vasile, S. Malik, D. Hnisz, T. I. Lee, I. I. Cisse, R. G. Roeder, P. A. Sharp, A. K. Chakraborty, and R. A. Young, "Coactivator condensation at super-enhancers links phase separation and gene control," *Science* **361**, eaar3958 (2018).
- <sup>227</sup>A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, and R. A. Young, "Transcription factors activate genes through the phase-separation capacity of their activation domains," *Cell* **175**, 1842 (2018).
- <sup>228</sup>S. Chong, C. Dugast-Darzacq, Z. Liu, P. Dong, G. M. Dailey, C. Cattoglio, A. Heckert, S. Banala, L. Lavis, X. Darzacq, and R. Tjian, "Imaging dynamic and selective low-complexity domain interactions that control gene transcription," *Science* **361**, eaar2555 (2018).
- <sup>229</sup>T. Cheutin, A. J. McNairn, T. Jenuwein, D. M. Gilbert, P. B. Singh, and T. Misteli, "Maintenance of stable heterochromatin domains by dynamic HP1 binding," *Science* **299**, 721–725 (2003).
- <sup>230</sup>A. Klosin and A. A. Hyman, "Molecular biology: A liquid reservoir for silent chromatin," *Nature* **547**, 168–170 (2017).
- <sup>231</sup>D. A. Potoyan, A. Savelyev, and G. A. Papoian, "Recent successes in coarse-grained modeling of DNA," *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **3**, 69–83 (2013).
- <sup>232</sup>G. S. Freeman, D. M. Hinckley, and J. J. De Pablo, "A coarse-grain three-site-per-nucleotide model for DNA with explicit ions," *J. Chem. Phys.* **135**, 165104 (2011).
- <sup>233</sup>D. M. Hinckley, G. S. Freeman, J. K. Whitmer, and J. J. De Pablo, "An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization," *J. Chem. Phys.* **139**, 144903 (2013).
- <sup>234</sup>G. S. Freeman, J. P. Lequeieu, D. M. Hinckley, J. K. Whitmer, and J. J. De Pablo, "DNA shape dominates sequence affinity in nucleosome formation," *Phys. Rev. Lett.* **113**, 168101 (2014).
- <sup>235</sup>W. Li, P. G. Wolynes, and S. Takada, "Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3504–3509 (2011).
- <sup>236</sup>J. Lequeieu, D. C. Schwartz, and J. J. de Pablo, "In silico evidence for sequence-dependent nucleosome sliding," *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9197–E9205 (2017).
- <sup>237</sup>J. Lequeieu, A. Córdoba, D. C. Schwartz, and J. J. De Pablo, "Tension-dependent free energies of nucleosome unwrapping," *ACS Cent. Sci.* **2**, 660–666 (2016).
- <sup>238</sup>J. Moller, J. Lequeieu, and J. J. De Pablo, "The free energy landscape of inter-nucleosome interactions and its relation to chromatin fiber structure," *ACS Cent. Sci.* **5**, 341–348 (2019).
- <sup>239</sup>J. J. Funke, P. Ketterer, C. Lieleg, S. Schunter, P. Korber, H. Dietz, C. Lieleg, J. J. Funke, P. Ketterer, H. Dietz, and S. Schunter, "Uncovering the forces between nucleosomes using DNA origami," *Sci. Adv.* **2**, e1600974 (2016).
- <sup>240</sup>H. Kenzaki and S. Takada, "Partial unwrapping and histone tail dynamics in nucleosome revealed by coarse-grained molecular simulations," *PLoS Comput. Biol.* **11**, e1004443 (2015).
- <sup>241</sup>S. Watanabe, Y. Mishima, M. Shimizu, I. Suetake, and S. Takada, "Interactions of HP1 bound to H3K9me3 dinucleosome by molecular simulations and biochemical assays," *Biophys. J.* **114**, 2336–2351 (2018).
- <sup>242</sup>R. Kanada, T. Terakawa, H. Kenzaki, and S. Takada, "Nucleosome crowding in chromatin slows the diffusion but can promote target search of proteins," *Biophys. J.* **116**, 2285–2295 (2019).



- <sup>243</sup>C. Tan and S. Takada, "Nucleosome allostery in pioneer transcription factor binding," *Proc. Natl. Acad. Sci. U. S. A.* **117**, 20586–20596 (2020).
- <sup>244</sup>L. Chang and S. Takada, "Histone acetylation dependent energy landscapes in tri-nucleosomes revealed by residue-resolved molecular simulations," *Sci. Rep.* **6**, 34441 (2016).
- <sup>245</sup>G. B. Brandani, T. Niina, C. Tan, and S. Takada, "DNA sliding in nucleosomes via twist defect propagation revealed by molecular simulations," *Nucleic Acids Res.* **46**, 2788–2801 (2018).
- <sup>246</sup>G. B. Brandani and S. Takada, "Chromatin remodelers couple inchworm motion with twist-defect formation to slide nucleosomal DNA," *PLoS Comput. Biol.* **14**, e1006512 (2018).
- <sup>247</sup>A. Davtyan, N. P. Schafer, W. Zheng, C. Clementi, P. G. Wolynes, and G. A. Papoian, "AWSEM-MD: Protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing," *J. Phys. Chem. B* **116**, 8494–8503 (2012).
- <sup>248</sup>J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, "Theory of protein folding: The energy landscape perspective," *Annu. Rev. Phys. Chem.* **48**, 545–600 (1997).
- <sup>249</sup>R. A. Goldstein, Z. A. Luthey-Schulten, and P. G. Wolynes, "Optimal protein-folding codes from spin-glass theory," *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4918–4922 (1992).
- <sup>250</sup>M. P. Eastwood, C. Hardin, Z. Luthey-Schulten, and P. G. Wolynes, "Statistical mechanical refinement of protein structure prediction schemes: Cumulant expansion approach," *J. Chem. Phys.* **117**, 4602–4615 (2002).
- <sup>251</sup>C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond, "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution," *J. Mol. Biol.* **319**, 1097–1113 (2002).
- <sup>252</sup>B. Zhang, W. Zheng, G. A. Papoian, and P. G. Wolynes, "Exploring the free energy landscape of nucleosomes," *J. Am. Chem. Soc.* **138**, 8126–8133 (2016).
- <sup>253</sup>T. Parsons and B. Zhang, "Critical role of histone tail entropy in nucleosome unwinding," *J. Chem. Phys.* **150**, 185103 (2019).
- <sup>254</sup>B. D. Brower-Toland, C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, and M. D. Wang, "Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA," *Proc. Natl. Acad. Sci. U. S. A.* **99**, 1960–1965 (2002).
- <sup>255</sup>B. Sudhanshu, S. Mihardja, E. F. Koslover, S. Mehraeen, C. Bustamante, and A. J. Spakowitz, "Tension-dependent structural deformation alters single-molecule transition kinetics," *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1885–1890 (2011).
- <sup>256</sup>I. M. Kulić and H. Schiessel, "DNA spools under tension," *Phys. Rev. Lett.* **92**, 228101 (2004).
- <sup>257</sup>I. V. Dobrovolskaia and G. Arya, "Dynamics of forced nucleosome unraveling and role of nonuniform histone–DNA interactions," *Biophys. J.* **103**, 989–998 (2012).
- <sup>258</sup>X. Ding, X. Lin, and B. Zhang, "Stability and folding pathways of tetra-nucleosome from six-dimensional free energy surface," *Nat. Commun.* **12**, 1091 (2021).
- <sup>259</sup>P. C. Whitford, J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu, and J. N. Onuchic, "An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields," *Proteins: Struct., Funct., Bioinf.* **75**, 430–441 (2009).
- <sup>260</sup>C. Clementi, H. Nymeyer, and J. N. Onuchic, "Topological and energetic factors: What determines the structural details of the transition state ensemble and 'en-route' intermediates for protein folding? An investigation for small globular proteins," *J. Mol. Biol.* **298**, 937–953 (2000).
- <sup>261</sup>M. Chen, M. A. Cuendet, and M. E. Tuckerman, "Heating and flooding: A unified approach for rapid generation of free energy surfaces," *J. Chem. Phys.* **137**, 024102 (2012).
- <sup>262</sup>A. Laio and M. Parrinello, "Escaping free-energy minima," *Proc. Natl. Acad. Sci. U. S. A.* **99**, 12562–12566 (2002).
- <sup>263</sup>L. Maragliano and E. Vanden-Eijnden, "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations," *Chem. Phys. Lett.* **426**, 168–175 (2006).
- <sup>264</sup>J. B. Abrams and M. E. Tuckerman, "Efficient and direct generation of multidimensional free energy surfaces via adiabatic dynamics without coordinate transformations," *J. Phys. Chem. B* **112**, 15742–15757 (2008).
- <sup>265</sup>E. Schneider, L. Dai, R. Q. Topper, C. Drechsel-Grau, and M. E. Tuckerman, "Stochastic neural network approach for learning high-dimensional free energy surfaces," *Phys. Rev. Lett.* **119**, 150601 (2017).
- <sup>266</sup>T. Schalch, S. Duda, D. F. Sargent, and T. J. Richmond, "X-ray structure of a tetranucleosome and its implications for the chromatin fibre," *Nature* **436**, 138–141 (2005).
- <sup>267</sup>E. Vanden-Eijnden and M. Venturoli, "Revisiting the finite temperature string method for the calculation of reaction tubes and free energies," *J. Chem. Phys.* **130**, 194103 (2009).
- <sup>268</sup>S. Cai, D. Böck, M. Pilhofer, and L. Gan, "The in situ structures of mono-, di-, and trinucleosomes in human heterochromatin," *Mol. Biol. Cell* **29**, 2450–2457 (2018).
- <sup>269</sup>M. Ohno, T. Ando, D. G. Priest, V. Kumar, Y. Yoshida, and Y. Taniguchi, "Sub-nucleosomal genome structure reveals distinct nucleosome folding motifs," *Cell* **176**, 520–534.e25 (2019).
- <sup>270</sup>C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoegge, J. Gharakhani, F. Julicher, and A. A. Hyman, "Germline P granules are liquid droplets that localize by controlled dissolution/condensation," *Science* **324**, 1729–1732 (2009).
- <sup>271</sup>R. van der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gspöner, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, V. N. Uversky, P. E. Wright, and M. M. Babu, "Classification of intrinsically disordered regions and proteins," *Chem. Rev.* **114**, 6589–6631 (2014).
- <sup>272</sup>R. Schneider, M. Blackledge, and M. R. Jensen, "Elucidating binding mechanisms and dynamics of intrinsically disordered protein complexes using NMR spectroscopy," *Curr. Opin. Struct. Biol.* **54**, 10–18 (2019).
- <sup>273</sup>A. A. Hyman, C. A. Weber, and F. Jülicher, "Liquid-liquid phase separation in biology," *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
- <sup>274</sup>D. H. De Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman, and S. J. Marrink, "Improved parameters for the Martini coarse-grained protein force field," *J. Chem. Theory Comput.* **9**, 687–697 (2013).
- <sup>275</sup>A. Liwo, M. Baranowski, C. Czaplewski, E. Gołaś, Y. He, D. Jagieła, P. Krupa, M. Maciejczyk, M. Makowski, M. A. Mozolewska, A. Niadzvedtski, S. Oldziej, H. A. Scheraga, A. K. Sieradzki, R. Ślusarz, T. Wirecki, Y. Yin, and B. Zaborowski, "A unified coarse-grained model of biological macromolecules based on mean-field multipole–multipole interactions," *J. Mol. Model.* **20**, 2306 (2014).
- <sup>276</sup>G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, "Sequence determinants of protein phase behavior from a coarse-grained model," *PLoS Comput. Biol.* **14**, e1005941 (2018).
- <sup>277</sup>H. Wu, P. G. Wolynes, and G. A. Papoian, "AWSEM-IDP: A coarse-grained force field for intrinsically disordered proteins," *J. Phys. Chem. B* **122**, 11115–11125 (2018).
- <sup>278</sup>U. Baul, D. Chakraborty, M. L. Mugnai, J. E. Straub, and D. Thirumalai, "Sequence effects on size, shape, and structural heterogeneity in intrinsically disordered proteins," *J. Phys. Chem. B* **123**, 3462–3474 (2019).
- <sup>279</sup>R. B. Best, W. Zheng, and J. Mittal, "Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association," *J. Chem. Theory Comput.* **10**, 5113 (2014).
- <sup>280</sup>P. Robustelli, S. Piana, and D. E. Shaw, "Developing a molecular dynamics force field for both folded and disordered protein states," *Proc. Natl. Acad. Sci. U. S. A.* **115**, E4758–E4766 (2018).
- <sup>281</sup>A. Vitalis and R. V. Pappu, "ABSINTH: A new continuum solvation model for simulations of polypeptides in aqueous solutions," *J. Comput. Chem.* **30**, 673–699 (2009).
- <sup>282</sup>J.-M. Choi and R. V. Pappu, "Improvements to the ABSINTH force field for proteins based on experimentally derived amino acid specific backbone conformational statistics," *J. Chem. Theory Comput.* **15**, 1367–1382 (2019).
- <sup>283</sup>J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell, "CHARMM36m: An improved force field for folded and intrinsically disordered proteins," *Nat. Methods* **14**, 71–73 (2016).
- <sup>284</sup>A. P. Latham and B. Zhang, "Improving coarse-grained protein force fields with small-angle x-ray scattering data," *J. Phys. Chem. B* **123**, 1026–1034 (2019).

- <sup>285</sup>S. Yang, S. Park, L. Makowski, and B. Roux, "A rapid coarse residue-based computational method for x-ray solution scattering characterization of protein folds and multiple conformational states of large protein complexes," *Biophys. J.* **96**, 4449–4463 (2009).
- <sup>286</sup>B. Rozycki, Y. Kim, and G. Hummer, "SAXS Ensemble Refinement of ESCRT-III CHMP3 Conformational Transitions," *Structure* **19**, 109–116 (2011).
- <sup>287</sup>A. P. Latham and B. Zhang, "Maximum entropy optimized force field for intrinsically disordered proteins," *J. Chem. Theory Comput.* **16**, 773–781 (2020).
- <sup>288</sup>R. Regmi, S. Srinivasan, A. P. Latham, V. Kukshal, W. Cui, B. Zhang, R. Bose, and G. S. Schlau-Cohen, "Phosphorylation-dependent conformations of the disordered carboxyl-terminus domain in the epidermal growth factor receptor," *J. Phys. Chem. Lett.* **11**, 10037–10044 (2020).
- <sup>289</sup>A. P. Latham and B. Zhang, "Consistent force field captures homolog resolved HP1 phase separation," *J. Chem. Theory Comput.* **17**(5), 3134–3144 (2021).
- <sup>290</sup>L. Wang, Y. Gao, X. Zheng, C. Liu, S. Dong, R. Li, G. Zhang, Y. Wei, H. Qu, Y. Li, C. D. Allis, G. Li, H. Li, and P. Li, "Histone modifications regulate chromatin compartmentalization by contributing to a phase separation mechanism," *Mol. Cell* **76**, 646–659.e6 (2019).
- <sup>291</sup>M. Spivakov, "Spurious transcription factor binding: Non-functional or genetically redundant?," *BioEssays* **36**, 798–806 (2014).
- <sup>292</sup>H. Chen, M. Levo, L. Barinov, M. Fujioka, J. B. Jaynes, and T. Gregor, "Dynamic interplay between enhancer–promoter topology and gene activity," *Nat. Genet.* **50**, 1296–1303 (2018).
- <sup>293</sup>A. Seeber, M. H. Hauer, and S. M. Gasser, "Chromosome dynamics in response to DNA damage," *Annu. Rev. Genet.* **52**, 295–319 (2018).
- <sup>294</sup>J. S. Lucas, Y. Zhang, O. K. Dudko, and C. Murre, "3D trajectories adopted by coding and regulatory DNA elements: First-passage times for genomic interactions," *Cell* **158**, 339–352 (2014).
- <sup>295</sup>J. C. Schmidt, A. J. Zaug, and T. R. Cech, "Live cell imaging reveals the dynamics of telomerase recruitment to telomeres," *Cell* **166**, 1188–1197.e9 (2016).
- <sup>296</sup>K. Bystricky, "Chromosome dynamics and folding in eukaryotes: Insights from live cell microscopy," *FEBS Lett.* **589**, 3014–3022 (2015).
- <sup>297</sup>W. F. Marshall, A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, and J. W. Sedat, "Interphase chromosomes undergo constrained diffusional motion in living cells," *Curr. Biol.* **7**, 930–939 (1997).
- <sup>298</sup>I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini, "Transient anomalous diffusion of telomeres in the nucleus of mammalian cells," *Phys. Rev. Lett.* **103**, 018102 (2009).
- <sup>299</sup>H. Hajjoul, J. Mathon, H. Ranchon, I. Goiffon, J. Mozziconacci, B. Albert, P. Carrivain, J.-M. Victor, O. Gadal, K. Bystricky, and A. Bancaud, "High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome," *Genome Res.* **23**, 1829–1838 (2013).
- <sup>300</sup>S. Shinkai, T. Nozaki, K. Maeshima, and Y. Togashi, "Dynamic nucleosome movement provides structural information of topological chromatin domains in living human cells," *PLoS Comput. Biol.* **12**, e1005136 (2016).
- <sup>301</sup>B. Chen, L. A. Gilbert, B. A. Cimini, J. Schnitzbauer, W. Zhang, G.-W. Li, J. Park, E. H. Blackburn, J. S. Weissman, L. S. Qi, and B. Huang, "Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system," *Cell* **155**, 1479–1491 (2013).
- <sup>302</sup>X. Wang, Z. Kam, P. M. Carlton, L. Xu, J. W. Sedat, and E. H. Blackburn, "Rapid telomere motions in live human cells analyzed by highly time-resolved microscopy," *Epigenet. Chromatin* **1**, 4 (2008).
- <sup>303</sup>I. Bronshtein, E. Kepten, I. Kanter, S. Berezin, M. Lindner, A. B. Redwood, S. Mai, S. Gonzalo, R. Foisner, Y. Shav-Tal, and Y. Garini, "Loss of lamin A function increases chromatin dynamics in the nuclear interior," *Nat. Commun.* **6**, 8044 (2015).
- <sup>304</sup>A. Zidovska, D. A. Weitz, and T. J. Mitchison, "Micron-scale coherence in interphase chromatin dynamics," *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15555–15560 (2013).
- <sup>305</sup>H. A. Shaban, R. Barth, and K. Bystricky, "Formation of correlated chromatin domains at nanoscale dynamic resolution during transcription," *Nucleic Acids Res.* **46**, e77 (2018).
- <sup>306</sup>A. Amitai and D. Holcman, "Polymer physics of nuclear organization and function," *Phys. Rep.* **678**, 1–83 (2017).
- <sup>307</sup>M. M. Tortora, H. Salari, and D. Jost, "Chromosome dynamics during interphase: A biophysical perspective," *Curr. Opin. Genet. Dev.* **61**, 37–43 (2020).
- <sup>308</sup>M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).
- <sup>309</sup>M. V. Tamm, L. I. Nazarov, A. A. Gavrilov, and A. V. Chertovich, "Anomalous diffusion in fractal globules," *Phys. Rev. Lett.* **114**, 178102 (2015).
- <sup>310</sup>K. E. Polovnikov, M. Gherardi, M. Cosentino-Lagomarsino, and M. V. Tamm, "Fractal folding and medium viscoelasticity contribute jointly to chromosome dynamics," *Phys. Rev. Lett.* **120**, 088101 (2018).
- <sup>311</sup>E. W. Montroll and G. H. Weiss, "Random walks on lattices. II," *J. Math. Phys.* **6**, 167–181 (1965).
- <sup>312</sup>S. C. Weber, J. A. Theriot, and A. J. Spakowitz, "Subdiffusive motion of a polymer composed of subdiffusive monomers," *Phys. Rev. E* **82**, 011913 (2010).
- <sup>313</sup>E. Kepten, I. Bronshtein, and Y. Garini, "Ergodicity convergence test suggests telomere motion obeys fractional dynamics," *Phys. Rev. E* **83**, 041919 (2011).
- <sup>314</sup>K. Burnecki, E. Kepten, J. Janczura, I. Bronshtein, Y. Garini, and A. Weron, "Universal algorithm for identification of fractional Brownian motion. A case of telomere subdiffusion," *Biophys. J.* **103**, 1839–1847 (2012).
- <sup>315</sup>N. Khanna, Y. Zhang, J. S. Lucas, O. K. Dudko, and C. Murre, "Chromosome dynamics near the sol-gel phase transition dictate the timing of remote genomic interactions," *Nat. Commun.* **10**, 2771 (2019).
- <sup>316</sup>H. Strickfaden, T. O. Tolsma, A. Sharma, D. A. Underhill, J. C. Hansen, and M. J. Hendzel, "Condensed chromatin behaves like a solid on the mesoscale in vitro and in living cells," *Cell* **183**, 1772–1784 (2020).
- <sup>317</sup>R. J. Biggs, N. Liu, Y. Peng, J. F. Marko, and H. Qiao, "Micromanipulation of prophase I chromosomes from mouse spermatocytes reveals high stiffness and gel-like chromatin organization," *Commun. Biol.* **3**, 542 (2020).
- <sup>318</sup>L. Liu, G. Shi, D. Thirumalai, and C. Hyeon, "Chain organization of human interphase chromosome determines the spatiotemporal dynamics of chromatin loci," *PLoS Comput. Biol.* **14**, e1006617 (2018).
- <sup>319</sup>M. D. Pierro, D. A. Potoyan, P. G. Wolynes, and J. N. Onuchic, "Anomalous diffusion, spatial coherence, and viscoelasticity from the energy landscape of human chromosomes," *Proc. Natl. Acad. Sci. U. S. A.* **115**, 7753–7758 (2018).
- <sup>320</sup>S. Put, T. Sakaue, and C. Vanderzande, "Active dynamics and spatially coherent motion in chromosomes subject to enzymatic force dipoles," *Phys. Rev. E* **99**, 032421 (2019).
- <sup>321</sup>S. Joo, X. Durang, O.-c. Lee, and J.-H. Jeon, "Anomalous diffusion for active Brownian particles cross-linked to a networked polymer: Langevin dynamics simulation and theory," *Soft Matter* **16**, 9188 (2020).
- <sup>322</sup>D. Osmanović and Y. Rabin, "Dynamics of active Rouse chains," *Soft Matter* **13**, 963–968 (2017).
- <sup>323</sup>T. F. Farage, P. Krinninger, and J. M. Brader, "Effective interactions in active Brownian suspensions," *Phys. Rev. E* **91**, 042310 (2015).
- <sup>324</sup>S. Wang and P. G. Wolynes, "Tensegrity and motor-driven effective interactions in a model cytoskeleton," *J. Chem. Phys.* **136**, 145102 (2012).
- <sup>325</sup>Z. Jiang and B. Zhang, "On the role of transcription in positioning nucleosomes," *PLoS Comput. Biol.* **17**, e1008556 (2021).
- <sup>326</sup>S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, "Principles of maximum entropy and maximum caliber in statistical physics," *Rev. Mod. Phys.* **85**, 1115–1141 (2013).
- <sup>327</sup>R. Capelli, G. Tiana, and C. Camilloni, "An implementation of the maximum-caliber principle by replica-averaged time-resolved restrained simulations," *J. Chem. Phys.* **148**, 184114 (2018).
- <sup>328</sup>J. Guan, H. Liu, X. Shi, S. Feng, and B. Huang, "Tracking multiple genomic elements using correlative CRISPR imaging and sequential DNA FISH," *Biophys. J.* **112**, 1077–1084 (2017).
- <sup>329</sup>Y. Takei, S. Shah, S. Harvey, L. S. Qi, and L. Cai, "Multiplexed dynamic imaging of genomic loci by combined CRISPR imaging and DNA sequential FISH," *Biophys. J.* **112**, 1773–1776 (2017).