# Thermodynamics of nucleosome breathing and positioning

Kharerin Hungyo, Benjamin Audit, Cédric Vaillant, Alexandre V Morozov

# Thermodynamics of nucleosome breathing and positioning

Kharerin Hungyo,[1,2] Benjamin Audit,[1] Cédric Vaillant,[1] and Alexandre V. Morozov*[,3]
[1)] *CNRS, ENS de Lyon, LPENSL, UMR5672, F-69342 Lyon cedex 07, France*
[2)] *School of Biosciences and Bioengineering, IIT Mandi, Kamand, HP-175005, India*
[3)] *Department of Physics and Astronomy and Center for Quantitative Biology, Rutgers University, Piscataway, New Jersey 08854, USA*

(*Electronic mail: morozov@physics.rutgers.edu)

(Dated: 18 December 2024)

Nucleosomes are fundamental units of chromatin in which a length of genomic DNA is wrapped around a histone octamer spool in a left-handed superhelix. Large-scale nucleosome maps show a wide distribution of DNA wrapping lengths, which in some cases are tens of base pairs (bp) shorter than the 147 bp canonical wrapping length observed in nucleosome crystal structures. Here, we develop a thermodynamic model that assumes a constant free energy cost per unwrapping a nucleosomal bp. Our model also incorporates linker DNA – short DNA segments between neighboring nucleosomes imposed by the folding of nucleosome arrays into chromatin fibers and other higher-order chromatin structures. We use this model to study nucleosome positioning and occupancy in the presence of nucleosome "breathing" – partial unwrapping and re-wrapping of nucleosomal DNA due to interactions with the neighboring particles. We find that, as the unwrapping cost per bp and the chemical potential are varied, the nucleosome arrays are characterized by three distinct states, with low, intermediate, and high densities. The transition between the latter two states proceeds through an equiprobable state in which all nucleosome wrapping lengths are equally likely. We study the equiprobable state theoretically using a mean-field approach, obtaining an excellent agreement with numerical simulations. Finally, we use our model to reproduce *S. cerevisiae* nucleosome occupancy profiles observed in the vicinity of transcription start sites, as well as genome-wide distributions of nucleosome wrapping lengths. Overall, our results highlight the key role of partial nucleosome unwrapping in shaping genome-wide patterns of nucleosome positioning and occupancy.



## I. INTRODUCTION

In eukaryotes, genomic DNA is condensed into a nucleoprotein chromatin complex whose primary unit – the nucleosome – consists of about 147 base pairs (bp) of DNA wrapped around an octamer of histone proteins[1–3]. On the one hand, by packaging DNA, nucleosomes condense the genome within the nuclear space, ensure proper mitotic condensation and segregation, and are a putative driving force for the expansion of eukaryotic genomes[4–7]. On the other hand, as a natural barrier to all DNA-templated processes, nucleosomes contribute to the regulation of transcription, replication, recombination, repair, and DNA insertion or deletion[8–11]. Therefore, nucleosomes, which are present in almost all eukaryotic species, are likely to be a major component of the regulation of genome activity and a key driver of genome evolution and in particular of the emergence of new regulatory pathways[12–14].

One of the key biological questions that remains to be elucidated is to what extent the nucleosomal organization of eukaryotic genomes influences regulation of genome activities: (1) how the nucleosome interferes locally with the regulatory complexes and the components of the machinery that controls DNA accessibility and processing and (2) how the nucleosomal organization along the genome is dynamically self-regulated with respect to various biochemical processes underlying genome functions.

Nucleosome-induced DNA (in)accessibility can be characterized by nucleosome footprints – positions and wrapping lengths of all nucleosomes along the genome[15–17]. These quantities can now be measured experimentally at an unprecedented genomic resolution, at different time points during the cell cycle and in different growth conditions, using wild-type and mutant strains (e.g., in cells where some regulatory complexes have been genetically modified or suppressed), on native but also on *in vitro* reconstituted chromatin[18–20].

One of the main mechanisms that enable transcription factor (TF) accessibility to genomic DNA on the chromatin template is the dynamic unwrapping and re-wrapping of the ends of the nucleosomal DNA[21–24]. The rate and the extent of site exposure/unwrapping may be due to sequence effects, post-translational modifications (PTMs), remodeling activity, TF binding, or some combination of those factors. For instance, it has been shown using fluorescence labeling experiments that DNA sequence affects nucleosomal dynamics[21,25–27]. The intrinsic part of these dynamics, known as "nucleosome breathing", can be attributed to thermal fluctuations[28]. Nucleosome breathing is widely believed to be affected by the action of chromatin remodelers[29]. However, the extent to which chromatin remodeling factors make use of and regulate nucleosome breathing is not well understood.

Due to nucleosome breathing, the nucleosome footprint has variable, time-dependent length. This is consistent

with a range of nucleosome fragment sizes observed in restriction enzyme, site-directed hydroxyl radical, and micrococcal nuclease (MNase) assays[16,20,30,31], even though the observed distributions of nucleosomal lengths might be biased by experimental artifacts such as MNase sequence specificity and the time of exposure to MNase or hydroxyl radicals[20,30,32–34]. For example, a wide range of nucleosome wrapping lengths is observed genome-wide in *S. cerevisiae*, in both genic and intergenic regions (Fig. 1A). Regardless of their wrapping lengths, the nucleosomes are arranged in a stereotypical periodic pattern in the vicinity of genes, with a nucleosome-depleted region immediately to the left of the transcription start site (TSS) (Fig. 1B). Nucleosomes with longer wrapping lengths ($> 140$ bp) occur more frequently than partially unwrapped nucleosomes ($\leq 140$ bp) (cf. rows of Fig. 1B). However, when taken together, the shorter fragments are present prominently in both gene bodies and upstream regulatory regions (Fig. 1C).

Several models have been previously proposed that incorporate nucleosome breathing to explain nucleosome density profiles observed *in vitro* and *in vivo*[23,35–38]. Here, we expand on these works by presenting a systematic study of the effects of the histone-DNA binding affinity and the histone octamer concentration on the genome-wide nucleosome distribution. We focus on the role of breathing in nucleosome organization by describing various thermodynamic properties such as nucleosome density, occupancy, fragment size distribution, and density correlation functions in the presence of nucleosome unwrapping.

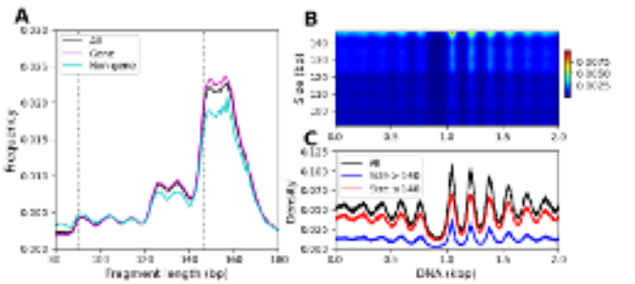


FIG. 1. Large-scale nucleosome maps in *S. cerevisiae*. (A) Genome-wide frequencies of nucleosome fragment lengths with $l \in [91, 147]$ bp are indicated by the region between the vertical dotted lines. (B) Heatmap of the distribution of the fragment length vs. position relative to the transcription start sites (TSS), averaged over all genes[39]. The TSS are aligned at position 1 kbp. Each MNase fragment mapped to the yeast genome was counted if its length $l$ was in the $[91, 147]$ bp range and its center was within 1 kbp of a TSS. The resulting MNase fragment counts for each genomic position and each $l$ were normalized by dividing by the total number of counts and multiplying by 100. (C) The nucleosome density near the TSS, plotted for all fragments and separately for shorter and longer fragments. The figures were generated using MNase data from Ref.[40] (GEO: ID GSE83123, dataset H4_Input_MNase_200U_Replicate_2, medium digestion level).

## II. GENERAL THERMODYNAMICAL PROPERTIES OF A LINEAR "BREATHING" MODEL

Here, we explore the thermodynamic properties of an array of nucleosomes described by a simple linear "breathing" energy model (see Appendix for details). In this model, the energy of a nucleosome located at the position $s$ (we define the genomic position of a nucleosome as the $5'$ end of its DNA) is independent of the genomic position and depends linearly on the total wrapping length $l$:

$$E(s, l) = E(l) = \epsilon_0(l - l_{min}) + E_{core}, \qquad (1)$$

where $l_{min} \leq l \leq l_{max} = 147$ bp, $\epsilon_0$ the incremental wrapping energy ($k_BT$/bp), and $E_{core} = \epsilon_0 l_{min}$ is the energy of a nucleosome with the minimal wrapping length $l_{min}$. The nucleosomes in the array are subject to steric exclusion, which can be viewed as short-range pairwise interactions between nearest-neighbor nucleosomes. The stochastic dynamics of such a nucleosomal array, which include binding and unbinding of histone octamers to the DNA segment, nucleosome sliding, and nucleosome winding and unwinding, are assumed to be in a thermodynamic equilibrium with a thermal bath, which also serves as a reservoir of histone octamers. The statistical properties of the nucleosomal array can then be fully described using the grand-canonical formalism of statistical mechanics[35,38]. Note that we consider both favorable ($\epsilon_0 < 0$) and unfavorable ($\epsilon_0 > 0$) wrapping energies per bp in Eq. (1). In the latter case, smaller-size nucleosomes will be energetically favored. This may happen if fully-wrapped nucleosomes are destabilized by nucleosome-disfavoring sequences such as poly-A/T tracts observed in nucleosome-depleted regions[41], or through the action of chromatin-remodeling enzymes that induce unwrapping followed by nucleosome translocation along the DNA[28].

### A. Arrays of fixed-size nucleosomes

First-generation models of nucleosome positioning and energetics assumed a fixed-size nucleosome wrapping length, typically $l = 147$ bp, on the basis of crystallographic studies[1,2]. In this limit and in the absence of sequence-dependent nucleosome positioning, our model reduces to the well known Tonks gas of non-overalpping one-dimensional particles[42]. The dependence of the equilibrium nucleosome density $\rho$ on the value of the chemical potential $\mu$ (the titration curve) is shown in Fig. 2A for several values of $l$, while the corresponding curves for the nucleosome occupancy $O = \rho l$ are shown in Fig. 2B. Both $\rho$ and $O$ are averaged over the 10 kbp DNA segment which is used in all our calculations.

In the Tonks gas, the equilibrium particle density $\rho$ satisfies the following exact equation[42]:

$$\mu_r = \ln \frac{\rho}{1-\rho l} + \frac{\rho l}{1-\rho l}, \tag{2}$$

where $\mu_r = \mu - E(l)$ is the net chemical potential and $E(l)$ is the energy of formation of a nucleosome with a wrapping length $l$, defined in Eq. (1). Note that both $l$ (bp) and $\rho$ (bp$^{-1}$) are dimensionless quantities. Numerical inversion of this equation leads to the sigmoidal titration curves for $\rho$ or $O$ as a function of $\mu_r$ (Fig. 2A,B). We note that the analytical and the computational results are in excellent agreement (Fig. 2). Furthermore, using Eq. (2) we can derive the chemical potential $\mu_r^*$ at the inflection point, where the susceptibility $\chi = d\rho/d\mu_r$ is at maximum. Even though this titration transition does not correspond to a true thermodynamical phase transition, $\mu_r^*$ still indicates the value of $\mu_r$ for which the nucleosome occupancy is the most sensitive to the changes in the chemical potential and thus represents a "transitional" value separating two distinct nucleosomal states. The inflection in the chemical potential curve occurs when the occupancy is $\rho^* l \approx 0.31$, leading to $\mu_r^* \approx -0.34 + \ln(1/l)$ (Fig. 2C; cf. Appendix F for details). Similarly, if we assume an *in vivo* occupancy value of $O = 0.8$, we obtain $\mu_r^{(0.8)} = 4 + \ln(4/l)$.

Below the transitional chemical potential $\mu_r^*$, the system is in a low-density state (LDS) which is characterized by a simple exponential dependence of the nucleosome density on its binding energy: $\rho \sim \exp(\mu - E_l)$. As the chemical potential increases, the nucleosome density enters the high-density state (HDS) and eventually starts to saturate as the result of hard-core interactions, which introduce an effective repulsion cost between neighboring nucleosomes. The asymptotic value of the nucleosome density is given by $\rho_{max} = 1/l$. Note that in Eq. (2), $l$ can stand for the effective interaction range $l_{\text{eff}}$, the minimum allowed distance between the centers of two neighboring particles ($l_{\text{eff}} = l$ in the case of pure steric exclusion; $l_{\text{eff}} = l + l_k$ when linker lengths $l_k$ are added to the model).

### B. Arrays of nucleosomes with variable DNA wrapping lengths

In this section, we explore the behavior of nucleosome arrays in which nucleosome particles can have an effective size between $l_{min}$ and $l_{max} + l_k$, where $l_{max} = 147$ bp is the crystallographic DNA wrapping length and $l_k$ is the DNA linker length. Note that the average nucleosome spacing may exceed the crystallographic DNA wrapping length if the effects of higher-order chromatin structure and in particular chromatin fiber formation are taken into account. In Figures 3 and 4 we explore the nucleosome density $\rho$, the occupancy $O$, and the mean wrapping length $\langle l \rangle$ as functions of the nucleosome wrapping energy per bp $\epsilon_0$ and the chemical potential $\mu$ using our numerical approach (see Appendix for details). Note that both $\rho$ and $O$ are averaged over the 10 kbp DNA segment. We consider three scenarios: (i) $l_{min} = 91$ bp, $l_k = 0$ bp, which we use to study the effects of partial DNA unwrapping in the absence of linker effects; (ii) $l_{min} = 91$ bp, $l_k = 20$ bp to account for higher-order chromatin structure; (iii) $l_{min} = 1$ bp, $l_k = 20$ bp to reproduce the parameters used in a previous study[35].

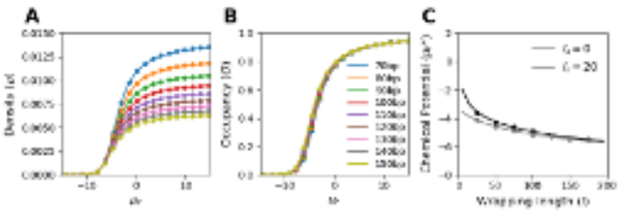


FIG. 2. One-dimensional liquid of fixed-size hard rods. (A) Density titration curves $\rho$ vs $\mu_r$ for several values of $l = 70$ bp (blue) to 150 bp (olive), in 10 bp steps. (B) Same as (A), but for the occupancy $O = \rho l$. (C) $\mu_r^*$ vs nucleosome wrapping length $l$. The lines are plotted using Eq. (2), while the filled circles are obtained from a computational model with fixed-size particles and $l_k = 0$ bp (A,B) and $l_k = 0$ or 20 bp (C), where $l_k$ is the length of a linker between adjacent nucleosomes. All energy units are in $k_BT$.

The histone octamer is likely to disassemble into tetrasome and hexasome subunits well before the nucleosomal DNA is fully unwrapped[43–47]. This motivated the choice of $l_{min} = 91$ bp, consistent with the shortest length of nucleosomal DNA that can still be accommodated by a histone octamer, and by the observation of a "plateau" in the histogram of nucleosome fragment lengths in Fig. 1A. We note that, in this regard, the $l_{min} = 1$ bp setting in the third scenario is less realistic. However, this regime allows us to verify that our numerical results are consistent with previously published observations (cf. Fig. 2C,D in Ref.[35]), validating our overall computational approach.

#### 1. A three-state phase diagram

When the nucleosome density is examined as a function of $\mu$ and $\epsilon_0$, we observe a three-state phase diagram with a low-density state (LDS), an intermediate-density state (IDS), and a high-density state (HDS) (Figs. 3A,D,G and 4A,D,G). For each value of $\epsilon_0$, the system eventually transitions from the LDS to the HDS as $\mu$ increases.

However, for positive $\epsilon_0$ the transition is immediately from the LDS to the HDS, similar to what is observed in fixed-size nucleosome arrays (Fig. 2A). In contrast, for negative $\epsilon_0$ the system goes first through the LDS→IDS and then through the IDS→HDS sigmoidal transitions. Interestingly, the IDS density is almost independent of $\mu$ and $\epsilon_0$ (Fig. 4A,D,G). At high values of $\mu$, the aver-

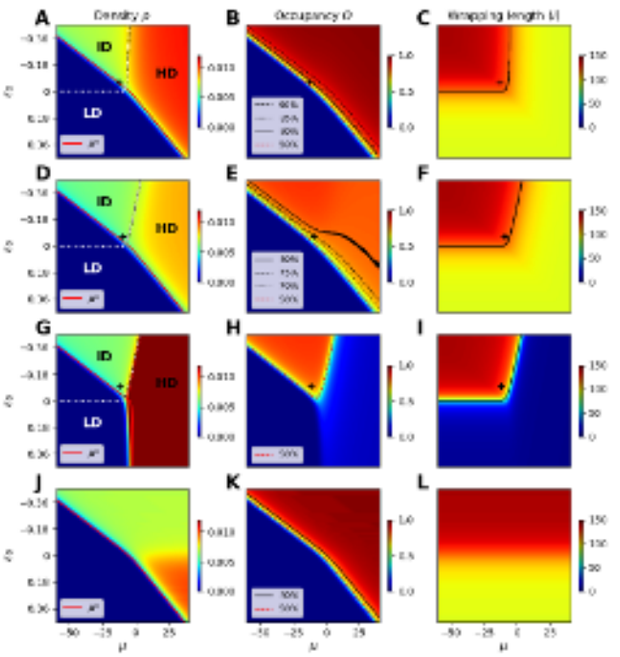


FIG. 3. Heatmaps of the nucleosome density $\rho$ (panels A,D,G,J), the occupancy $O$ (panels B,E,H,K), and the mean wrapping length $\langle l \rangle$ (panels C,F,I,L). The color code for each heatmap ranges from blue (low values) to red (high values). Panels A-C: $l_k = 0$ bp, $l_{min} = 91$ bp, and $l_{max} = 147$ bp (model 1). Panels D-F: $l_k = 20$ bp, $l_{min} = 91$ bp, and $l_{max} = 147$ bp (model 2). Panels G-I: $l_k = 20$ bp, $l_{min} = 1$ bp, and $l_{max} = 147$ bp (model 3; parameters used in Ref.[35]). Panels J-L: The mean-field approximation for the system in panels A-C (see main text for details). The red line in panel J is the locus of the transition chemical potential $\mu^*$ in the mean-field model, also reproduced in panel A. Similarly, the red lines in panels D and G are the $\mu^*$ curves from the corresponding mean-field models. Panels A,D,G: HDS, IDS, and LDS are separated by the equiprobable states shown as black curves. Panels B,E,H,K: The iso-lines of constant occupancy are shown as colored lines, as indicated in the panel legends. Panels C,F,I: the iso-lines of the mean wrapping length $\langle l \rangle = 119$ bp (C), $\langle l \rangle = 119$ bp (F), and $\langle l \rangle = 74$ bp (I) corresponding to the equiprobable states are traced in black. The iso-lines from the panels C,F,I are also replotted in panels A,D,G, respectively (white dotted curves). The black cross marks are the best-fit parameters that reproduce the *in vivo* nucleosome distributions in *S. cerevisiae* for models 1-3 (cf. Table I and Fig. 10).

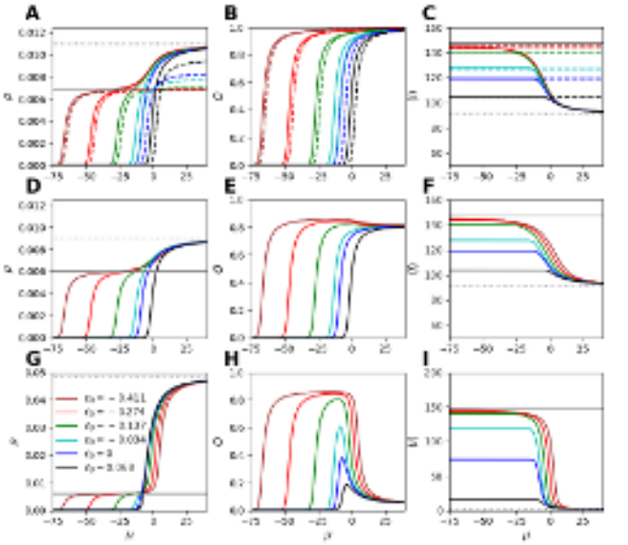


FIG. 4. Nucleosome density $\rho$, occupancy $O$, and the mean wrapping length $\langle l \rangle$ as functions of the chemical potential $\mu$. Panels A-C, D-F, and G-I are the cross sections of the corresponding heatmaps in Fig. 3, for a set of $\epsilon_0$ values shown in the legend of panel G. In panels A,B,C, the computational results are shown as solid lines and the corresponding analytical results are shown as dashed lines of the same colors (brown, red, green, cyan, blue, and black). The analytical results were obtained using Eq. (2) and the $\epsilon_0$-dependent $\langle l \rangle$ (cf. Fig. 3J,K,L and the main text). The maximum-density $\rho_{max} = 1/(l_{min} + l_k)$ is indicated as a dotted grey line; the minimum density $\rho = 1/(l_{max} + l_k)$ is indicated as a solid grey line.

age occupancy $O$ tends to the limiting value which corresponds to fully unwrapped particles with the length $l_{min}$ separated by linkers (Fig. 4B,E,H). Unlike the density diagrams, the occupancy diagrams present only two states with respect to $\epsilon_0$ and $\mu$; the same occupancy can correspond to a range of densities and nucleosome lengths. Finally, the transition chemical potential curves $\mu_r^*(\epsilon_0)$ are nearly straight lines, with distinct slopes for the LDS→IDS, IDS→HDS, and LDS→HDS transitions.

Next, we consider the behavior of the mean wrapping length $\langle l \rangle$ (Figs. 3C,F,I and 4C,F,I). We observe that both the LDS and the IDS are characterized by approximately the same value of $\langle l \rangle$, which is larger than that observed in the HDS. Thus, for a fixed value of $\epsilon_0$ the change of $\langle l \rangle$ with $\mu$ can be viewed as a two-state transition brought about by partial nucleosome unwrapping. The value of $\mu$ for the unwrapping transition coincides with the IDS→HDS boundary in the density plots, for the negative values of $\epsilon_0$. The asymptotic value of $\langle l \rangle$ in the high-$\mu$ limit is $l_{min}$, independent of $\epsilon_0$. Finally, in the LDS/IDS regime $\langle l \rangle$ only weakly depends on $\mu$ and gradually becomes smaller as the negative $\epsilon_0$ decreases in magnitude. At the $\epsilon_0 = 0$ threshold, there is an abrupt transition to fully unwrapped nucleosomes, with $\langle l \rangle$ being bounded by $l_{min}$ from below.

We start our quantitative analysis of the computational results shown in Figs. 3 and 4 by considering a one-dimensional (1D) liquid of fixed-size particles – the Tonks gas model. In this model, increasing the chemical potential leads to an increase of the nucleosome density (Fig. 2A). In the low-$\mu$ limit, neighboring nucleosomes do not interact and the density follows the Boltzmann dis-

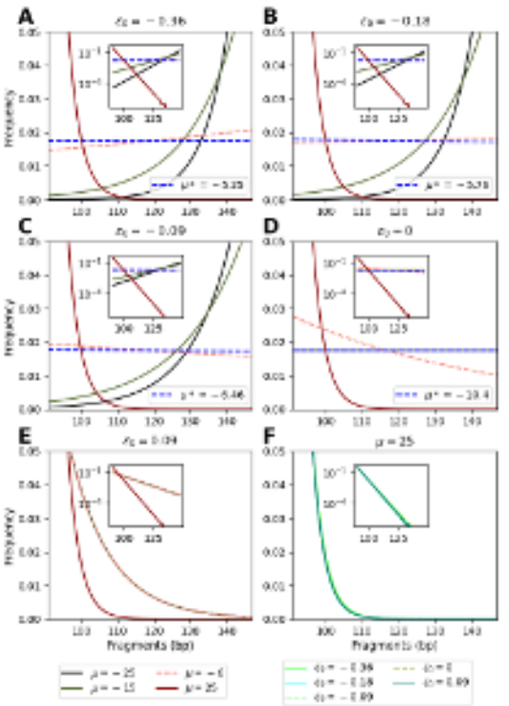


FIG. 5. The distribution of the particle sizes $w(l)$ as a function of the chemical potential $\mu$ and the interaction energy per bp $\epsilon_0$. Panels A-E: particle size distribution for a given $\epsilon_0$ and several values of $\mu$, including $\mu^*$ – the chemical potential at which the distribution of particle sizes is uniform. Panel F: the particle size distribution for a given $\mu$ and several values of $\epsilon_0$. All the insets show the same plots on the log-scale.

tribution: $\rho \sim \exp(\mu - E)$ (Eq. (2)). Around $O \simeq 0.5$, the excluded volume interactions become significant and the adsorption gain in free energy per particle is counterbalanced by an effective repulsion penalty due to steric exclusion. This leads to a slower increase of the nucleosome density with the chemical potential toward an asymptotic maximum value imposed by the hard-core interactions: $\rho_{max} = 1/(l + l_k)$, where $l$ is the fixed particle length. We note that partially unwrapped nucleosomes in contact with each other may have an additional repulsive contribution due to the steric exclusion of unwrapped DNA segments in 3D space[37].

In the extended Tonks gas model with variable particle lengths, the size $l$ of the DNA-bound particles is an additional degree of freedom whose statistical distribution $\rho(l)$ is controlled by both steric exclusion and the internal energy $E(l)$. For example, in the low-$\mu$ limit we expect that $\rho(l) \sim \exp(\mu - E(l))$, similarly to the fixed-size case. Since $E(l)$ is given by Eq. (1), $\rho(l) \sim \exp(\mu)\exp(-\epsilon_0(l - l_{min}))$ between $l_{min}$ and $l_{max}$, consistent with the numerical distributions of $l$ at low $\mu$ (Fig. 5). The most probable value of $l$ is given by $l_{min}$ in the repulsion case ($\epsilon_0 > 0$) and $l_{max}$ in the attraction case ($\epsilon_0 < 0$). These two values of $l$ correspond to the zero-temperature ground states when histone-DNA adsorption is unfavorable or favorable, respectively.

Furthermore, the mean value of $l$ at equilibrium is given by

$$\langle l \rangle = l_{min} + \Delta l \frac{\exp(-\epsilon_0 \Delta l)}{\exp(-\epsilon_0 \Delta l) - 1} + \epsilon_0^{-1}, \qquad (3)$$

where $\Delta l = l_{max} - l_{min}$. This is consistent with the numerical results shown in Fig. 4C,F,I. In the strong-repulsion limit ($\epsilon_0 \to \infty$), $\langle l \rangle \to l_{min}$; in the strong-attraction limit ($\epsilon_0 \to -\infty$), $\langle l \rangle \to l_{max}$. Finally, the density $\rho = \Delta l^{-1} \int_{l_{min}}^{l_{max}} \rho(l) dl$ averaged over all nucleosome lengths can be obtained as

$$\rho = \frac{\exp(\mu)}{\epsilon_0} \left[1 - \exp(-\epsilon_0 \Delta l)\right]. \qquad (4)$$

As long as the wrapping length $l$ is primarily controlled by the internal energy $E(l)$, we expect the shape and the mean of $\rho(l)$ to remain relatively unchanged in the low-density limit as $\mu$ increases. This is indeed observed up to the transition value $\mu^*$, where the system switches to the HDS and the unwrapping transition occurs (Figs. 3C,F,I and 4C,F,I; Fig. 5).

For $\epsilon_0 < 0$ and $\mu < \mu^*$, the LDS→IDS sigmoidal transition occurs that resembles the LDS→HDS transition in fixed-size particle systems. This suggests that the system might be approximated as a single-size system with a $\mu$-independent effective size $\langle l \rangle$. For $\epsilon_0 > 0$, the LDS→HDS transition can be described in the same way. Thus, the system switches from an LDS characterized by $\rho = \exp(\mu - \epsilon_0(\langle l \rangle - l_{min}))$ to a higher-density state (IDS or HDS) characterized by $\rho \sim 1/(\langle l \rangle + l_k)$. This transition occurs at $\mu^* = -0.34 - \ln(\langle l \rangle + l_k)$.

We observe a close correspondence between the titration properties of the mean-field effective system (Figs. 3J,K,L and 4A,B,C) and the full system (Figs. 3A,B,C and 4A,B,C). The largest discrepancy between the two models occurs when $\epsilon_0 \approx 0$, where the length fluctuations are more significant (in fact, $\rho(l)$ is approximately uniform) and the simple mean-field picture is likely to break down. Interestingly, the mean-field $\mu^*$ is slightly larger than the computationally observed values (cf. solid and dashed lines in Fig. 4A,B,C). This suggests, according to how $\mu^*$ changes with respect to $l$ (Fig. 2C), an effective particle size that is slightly larger than $\langle l \rangle$.

The $\epsilon_0 < 0$ case is characterized by a second sigmoidal transition, IDS→HDS[35]. As $\mu$ continues to increase, the system increases its density by reducing the nucleosome wrapping lengths, eventually approaching $l_{min}$ from above. Thus, in contrast to single-size systems, the effective repulsion cost associated with binding of new particles is lowered by the ability of the bound particles to reduce their size. Consequently, the unwrapping transition occurs when, in the IDS, the mean gain in adsorbing a new particle ($\mu$) becomes larger than the mean unwrapping cost needed to accommodate this new particle

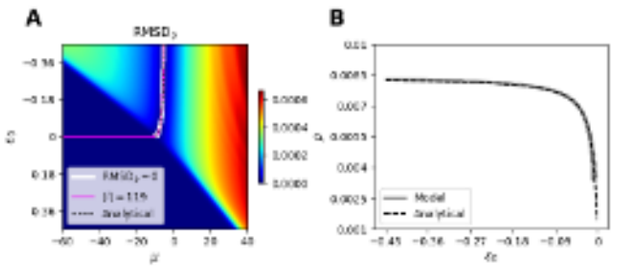


FIG. 6. Equiprobable states as a function of $\epsilon_0$ and $\mu$. (A) Heatmap of the root mean square deviation (RMSD) of the density, $\text{RMSD}_\rho = \sqrt{\langle(\rho(l) - \langle\rho\rangle_l)^2\rangle_l}$, where $\rho(l)$ is the density of nucleosomes of length $l$. The mean wrapping length $\langle l \rangle = 119$ (magenta) and $\text{RMSD}_\rho = 0$ (white) are superimposed with the analytical curve (dashed black line). (B) Density $\rho$ as a function of $\epsilon_0$ along the transition line.

and satisfy the steric exclusion constraints. The stronger the attraction, the higher is the $\mu$ required for the unwrapping transition. After the IDS→HDS transition, the particle density and the mean wrapping length converge asymptotically to $\rho \sim 1/(l_{min} + l_k)$ and $\langle l \rangle \sim l_{min}$ (Fig. 4).

Conceptually, the unwrapping transition can be understood by mapping the system to a Tonks gas with a single size $l_{max}$ and an effective soft-core repulsion pair potential. The neighboring particles can overlap and this overlap is controlled by a pair potential that can be derived directly from the adsorption energy $\epsilon_0$[48]. For $\epsilon_0 > 0$, a single LDS→HDS transition is observed. Since the binding energy $\epsilon_0$ is repulsive, nucleosome wrapping is suppressed and the nucleosomes are formed predominantly in the $l \simeq l_{min}$ state, making a single-size mean-field description accurate in this regime.

In Fig. 5, we show the full distributions of the particle wrapping lengths as a function of $\mu$ and $\epsilon_0$. Overall, the nucleosome wrapping lengths are distributed exponentially in the $\epsilon_0 < 0$ case, with the maximum of the distribution centered on $l_{max}$ for the low values of $\mu$. As $\mu$ increases, the distribution becomes uniform at $\mu^*$ and again exponential but centered on $l_{min}$ at $\mu > \mu^*$ (Fig. 5A-C). In the $\epsilon_0 > 0$ case (Fig. 5E), the distributions are always centered on $l_{min}$ because additional wrapping is energetically unfavorable. In Fig. 5F, we show the wrapping length distributions for $\mu = 25\ k_BT$ and the 5 values of $\epsilon_0$ shown in Fig. 5A-E. Regardless of the value of $\epsilon_0$, the distributions are self-similar and are characterized by an exponential decrease from the most probable value $l_{min}$ toward $l_{max}$.

### 2. The intermediate state with equally probable wrapping lengths

Here, we focus on the $l_{min} = 91$ bp, $l_k = 0$ bp case shown in Fig. 3A,B,C (model 1); however, the arguments can be generalized to the other two cases. As discussed above, in the $\epsilon_0 < 0$ case there is an IDS→HDS unwrapping transition from a distribution peaked at $l_{max}$ to a distribution peaked at $l_{min}$. This observation suggests that the transition occurs through a uniform distribution of wrapping lengths with a mean value $\langle l \rangle = (l_{min} + l_{max})/2 = 119$ bp.

To check this hypothesis, we computed the RMSD of the nucleosome density, $\text{RMSD}_\rho$, as a function of $\epsilon_0$ and $\mu$ (Fig. 6). We observe that the transition is indeed characterized by a thin region with $\text{RMSD}_\rho \simeq 0$ (Fig. 6A), which indicates that the transition occurs through a state where all possible wrapping lengths are equally likely. The uniform distribution of wrapping lengths leads to $\text{RMSD}_\rho = 0$ since the density of DNA fragments is the same for all wrapping lengths $l$. This is further illustrated in Fig. 5A-E, where we report the uniform distribution of $l$ observed for different $\epsilon_0$ values at the corresponding chemical potential $\mu^*(\epsilon_0)$.

The transition curve $\mu^*(\epsilon_0)$ can also be obtained by extracting the $\langle l \rangle = 119$ bp iso-curve from the heatmap in Fig. 3C. As expected, this iso-line closely follows the $\text{RMSD}_\rho = 0$ $\mu*$ curve (Fig. 6A). Finally, we confirm our numerical results by theoretical arguments, where the existence of the equiprobable wrapping-length state was examined analytically, yielding the transition values of $\mu^*$ and $\rho^*$ that define the IDS→HDS unwrapping transition (Eq. (E5) in the Appendix).

As shown in Fig. 6A, the theoretical curve indeed coincides with our numerical derivation of the transition line (either the $\text{RMSD}_\rho = 0$ curve or the $\langle l \rangle = 119$ bp iso-line). The predicted density $\rho^*$ along this theoretical line is also in a very good agreement with the numerical one (Fig. 6B).

### 3. Statistical nucleosome ordering

A key feature of finite-size particle systems with excluded volume interactions is the spontaneous emergence of particle ordering at sufficiently high densities[15,35,49,50]. The ordering is seen in the appearance of periodic oscillations in the pair correlation function $g^{(2)}(s, \Delta s)$, as shown in Fig. 7A,B for fixed-size particles and Fig. 8 for variable-size particles. The radial pair-correlation function quantifies the probability that, given a particle at a position $s$, another particle is found at a separation distance $\Delta s$, relatively to the non-interacting system with the same particle density $\rho(s)$: $g^{(2)}(s, \Delta s) = \rho_2(s, s + \Delta s)/\rho^2(s)$ (note that $\lim_{\Delta s \to \infty} g^{(2)}(s, \Delta s) = 1$). In the uniform case (position-independent energies and pair interactions), the pair correlation is also position-independent: $g^{(2)}(s, \Delta s) \to g^{(2)}(\Delta s)$.

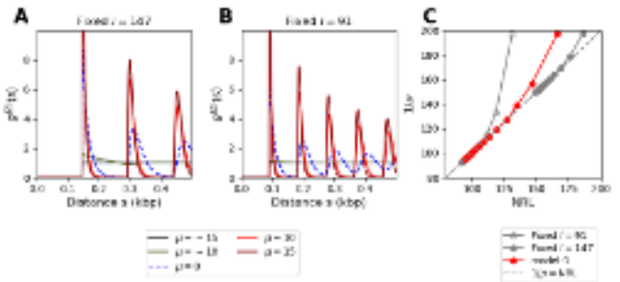


FIG. 7. Radial pair-correlation functions for the arrays of fixed-size nucleosome particles. Each curve in panels A and B corresponds to a value of $\mu$ indicated in the panel B legend. A: $l = 147$ bp, B: $l = 91$ bp. C: A scatter plot of the inverse nucleosome density $1/\rho$ vs NRL. We assume $\epsilon_0 = -0.034\ k_BT$/bp.

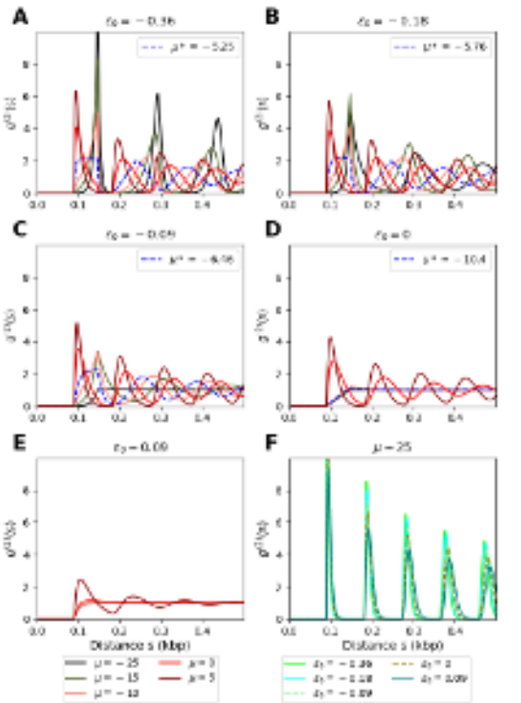


FIG. 8. Radial pair-correlation functions for the arrays of variable-size nucleosome particles (model 1). A-E: Radial pair-correlation functions for several values of the chemical potential $\mu$ and a given binding energy per bp $\epsilon_0$: $\epsilon_0 = -0.36\ k_BT$ (A), $\epsilon_0 = -0.18\ k_BT$ (B), $\epsilon_0 = -0.09\ k_BT$ (C), $\epsilon_0 = 0\ k_BT$ (D), and $\epsilon_0 = 0.09\ k_BT$ (E). F: Radial pair-correlation functions for several values of $\epsilon_0$ and a fixed $\mu = 25\ k_BT$.

At low particle density (the LD state), the pair correlation function can be approximated by the Boltzmann factor: $g^{(2)}(\Delta s) = e^{-u(\Delta s)}$ and thus only depends on the pairwise interaction energy. When the pairwise interactions are reduced to steric exclusion with the range $l_{hc}$, the pair correlation function is given by $g^{(2)} = 0$ for $\Delta s \leq l_{hc}$ and $g^{(2)} = 1$ for $\Delta s > l_{hc}$ (cf. the $\mu = -15, -10\ k_BT$ curves in Fig. 7A,B for $l_{hc} = 147$ bp and 91 bp, respectively). In the HD state, particles start to be confined by the neighboring particles as a result of excluded volume interactions. This leads to the long-range ordering and thus to the periodic modulation of the pair-correlation function around the $g^{(2)} = 1$ value. As shown in Fig. 7A,B, the strength of such ordering increases with the chemical potential, manifesting itself in a concomitant increase of the amplitude and decrease of the period (hereafter called the Nucleosome Repeat Length, or NRL) as well as a reduction of the width of the oscillatory peaks. The NRL is computed as the average peak-to-peak distance of the first 2–4 peaks in the radial pair correlation function $g^{(2)}$. As expected, the value of NRL is strongly correlated with the mean separation distance $\rho^{-1}$ (Fig. 7C).

Next, we investigate how fluctuations in the nucleosome wrapping length influence statistical ordering. In Fig. 8, we report the pair correlation function averaged over $l$ for several values of $\epsilon_0$ and $\mu$. In Fig. 9, we report the NRL heat maps vs. $\epsilon_0$ and $\mu$ (Fig. 9A,C,E), as well as NRL as a function of $\mu$ for several $\epsilon_0$ values (Fig. 9B,D,F). We observe that, as a natural consequence of the correlation between NRL and the nucleosome density, the system is described by a three-state NRL phase diagram that coincides with the three-state density phase diagram (Fig. 3): the LDS is characterized by limited ordering and large values of NRL; the HDS is characterized by strong ordering (Fig. 8F), with the NRL that asymptotically tends to $l_{min}+l_k$. Finally, the IDS exhibits moderate ordering with the NRL $\simeq l_{max} + l_k$ (Fig. 9B,D,F).

Whereas the pair correlation function in the HDS (Fig. 8F) resembles the pair correlation function of the fixed-size system with $l_{hc} = 91$ bp (Fig. 7B), the IDS ($\epsilon_0 < 0$) and the LDS pair correlation functions exhibit slight differences. For $\epsilon_0 < 0$ in the IDS, the pair correlation function shows periodic peaks with the asymptotic NRL $\sim 147$ bp, but the peaks have different shapes. As confirmed by the LDS pair correlation function (Fig. 8C), the fluctuating-size particles behave as effective fixed-size particles with soft-core interaction – a pairwise interaction that allows particles to interpenetrate down to the minimal separation distance $\Delta s_{min} = l_{min} = 91$ bp. This effective soft-core pair potential $u^{eff}(\Delta s)$ can be derived from the LDS pair correlation function, $u^{eff} = -\ln g^{(2)}$.

For the representative case shown in Fig. 8C for $\epsilon_0 = -0.09\ k_BT$, the effective potential is 0 for $\Delta s > 147$ bp, $\infty$ for $\Delta s < 91$ bp, and finite for 91 bp $\leq \Delta s \leq 147$ bp. In the latter case, the shape of the soft-core repulsive potential depends on $\epsilon_0$[48] in the LDS regime: the larger $|\epsilon_0|$, the larger the penalty cost for unwrapping, leading to stronger effective repulsion. For $\epsilon_0 = 0$, the en-

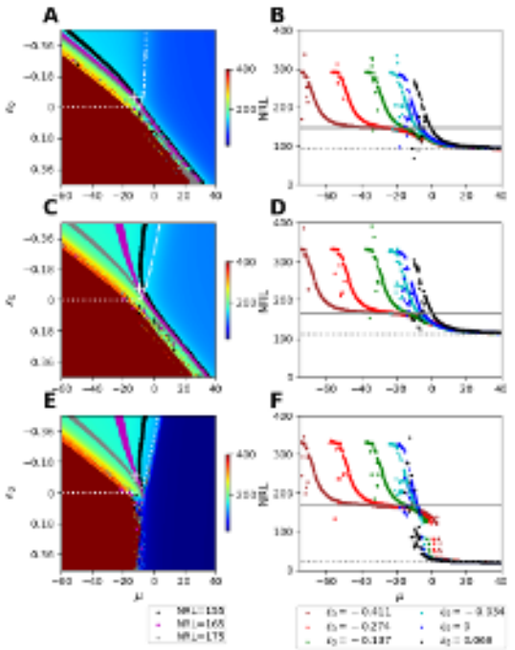


FIG. 9. The nucleosome repeat length (NRL) as a function of $\mu$ and $\epsilon_0$. A,C,E: NRL heatmaps, with NRL iso-lines shown as dotted curves: 155 bp (black), 165 bp (magenta), and 175 bp (grey). The mean wrapping length $\langle l \rangle$ is indicated in each panel as a white dotted line. B,D,F: NRL as a function of $\mu$ for 6 $\epsilon_0$ values. $\mathrm{NRL}_{min} = l_{min} + l_k$ and $\mathrm{NRL}_{max} = l_{max} + l_k$ are shown as dotted and solid grey lines, respectively. The nucleosome length parameters are as in Fig. 3: $l_{min} = 91$ bp, $l_{max} = 147$ bp, $l_k = 0$ bp (A,B); $l_{min} = 91$ bp, $l_{max} = 147$ bp, $l_k = 20$ bp (C,D); $l_{min} = 1$ bp, $l_{max} = 147$ bp, $l_k = 20$ bp (E,F). The cross marks (white) are best-fit parameters that correspond to the *in vivo* conditions for the different models (see Table I and Fig. 10).

thalpic cost of unwrapping is zero and only the entropic penalty remains, leading to the linear LDS $g^{(2)}$ shape for $\Delta s \in [91, 147]$ bp (Fig. 8D). Finally, for the $\epsilon_0 > 0$ case, as illustrated in Fig. 8E, we observe an effective attractive soft-core pair potential that favors the shortest separation distances between neighboring nucleosomes (i.e., $\Delta s \simeq l_{min} + l_k$). These pairwise separation distances promote unwrapping, reducing the overall enthalpic cost.

Note that the quasi-periodic ordering described above is density-dependent and thus has purely entropic origins[15]. It is not imposed by a pairwise potential which could be used to impose a pre-specified periodicity at any particle density. We also emphasize that this entropic periodic ordering only concerns the relative positions of the particles and not their absolute genomic positions. For example, in the HD state the density $\rho(s)$ remain nearly uniform, while the pair-correlation function is strongly periodic. This internal ordering can emerge in the density profile only in the presence of an inhomogeneous energy potential, such as an energy barrier or infinite walls at the boundaries of the DNA segment.

| | $\epsilon_0$, $k_BT$/bp | $\mu$, $k_BT$ | $\langle l \rangle$, bp | NRL, bp | $\langle O \rangle$ |
|---|---|---|---|---|---|
| **Model-1** | -0.057 | -12.8 | 129 | 165 | 0.71 |
| **Model-2** | -0.064 | -9.2 | 125 | 165 | 0.74 |
| **Model-3** | -0.096 | -11.6 | 125 | 165 | 0.76 |
| **Fixed-147** | — | $\mu_r = 5.05$ | 147 | 165 | 0.89 |

TABLE I. Parameter values of the three variable-size models and a fixed-size model with $l = 147$ bp used to predict the experimental data in Fig. 1.

### 4. Prediction of *in vivo* distributions of nucleosome densities and wrapping lengths

In this section, we focus on finding the model parameters $\epsilon_0$ and $\mu$ that best reproduce the experimental data in Fig. 1. We consider three veriable-size models with $(l_k, l_{min}, l_{max})$ defined in Fig. 3, as well as a fixed-size model with $l = 147$ bp. For each variable-size model, we inspect the data shown in Fig. 3 and Fig. 9A,C,E to find the $(\epsilon_0, \mu)$ parameter combinations that satisfy the following constraints consistent with experimental data: $0.7 \leq O \leq 0.9$, 164.5 bp $\leq$ NRL $\leq$ 165.5 bp, and 120 bp $\leq \langle l \rangle \leq$ 130 bp. Here, $O$ is the average nucleosome occupancy of the 10 kbp DNA segment we used in our calculations, and $\langle l \rangle$ is the mean wrapping length. In the cases where several sets of discretized $(\epsilon_0, \mu)$ parameters satisfy the above constraints, we have chosen the solution with $\langle l \rangle$ closest to 125 bp. The resulting parameter values, along with the values of the constrained variables, are shown in Table I. For the fixed-size model, we tune the relative chemical potential $\mu_r$ to satisfy the constraints on NRL and $O$. In all cases, the TSS was modeled as an infinite barrier at the left edge of the 10 kbp DNA segment. We note that the average nucleosome occupancy is typically between 0.7 and 0.8 in many organisms and chromatin states[51]. Our prediction of $O = 0.89$ for the fixed-size nucleosome model in Table I is on the high side because the nucleosomes are not allowed to unwrap, and is more reasonable in the other three models.

We find that all 3 variable-size models reproduce the distribution of wrapping lengths in *S. cerevisiae* reasonably well, with the exception of the $\simeq 10$ bp oscillations that are likely caused by the histone-DNA interaction patches spaced at roughly 10 bp intervals along the superhelical path of the nuclesomal DNA (Fig. 10A). Note that apart from choosing the values of $\epsilon_0$ and $\mu$ that satisfy the constraints, the models were not fit in any way to the observed distributions of wrapping lengths or nucleosome density profiles.

Interestingly, models 2 and 3 reproduce the observed nucleosome density profiles better than model 1, perhaps due to the presence of the 20 bp linker DNA (Fig. 10B). The only exception is the +1 nucleosome (first peak to the right of the TSS), where models 1 and 2 yield a sharp

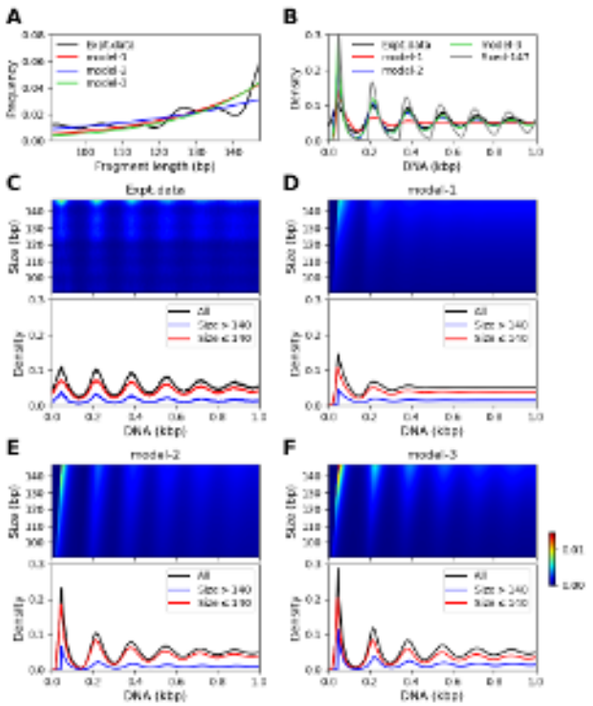


FIG. 10. Comparison of predicted distributions of nucleosome wrapping lengths and nucleosome densities to the right of the TSS with experimental data. The predictions were carried out using three models defined in Fig. 3, with parameters listed in Table I. The experimental data is based on an MNase map of *S. cerevisiae* nucleosomes[40] (Fig. 1). (A) Distributions of nucleosome wrapping lengths. (B) Nucleosome density distributions to the right of the TSS, regardless of the length of the DNA footprint. (C) Heatmap of experimental nucleosome density profiles to the right of the TSS for a range of nucleosome wrapping lengths (upper panel). Experimental nucleosome density profiles for longer and shorter wrapping lengths (lower panel). The data is normalized as in Fig. 1, but using only DNA fragments whose centers map within 1 kbp downstream of the TSS. (D-F) Same as (C), but for the predictions of models 1-3. Note that the black curve in each lower panel of C-F is the same as the corresponding curve in B.

peak. This may be due to the fact that we model the nucleosome distribution over genes using an infinite barrier at the left edge of the 10 kbp DNA fragment to represent the nucleosome-depleted region upstream of the TSS. Figs. 10D-F show how well the variable-size models in Table I reproduce the experimentally observed density profiles of the nucleosomes of various wrapping lengths downstream of the TSS (Fig. 10C). Consistent with our observations above, models 2 and 3 reproduce the data better than model 1, except in the first peak. Overall, nucleosomes with larger wrapping lengths have higher density compared to shorter, partially unwrapped nucleosomes. Compared with experimental data, the fixed-size model results in oscillations that are too prominent (cf. grey line in Fig. 10B).

## III. DISCUSSION AND CONCLUSION

We have presented a comprehensive theoretical and computational investigation of nucleosome positioning in the presence of partial unwrapping and rewrapping of the nucleosomal DNA. Following previous studies[23,35,36,38], we have focused on the dependence of key quantities of biological interest such as nucleosome density, occupancy, and repeat length, as well as the wrapping length distribution with the two control parameters of the model: the chemical potential and the effective binding energy per bp. Specifically, our thermodynamic phase diagrams highlight the existence of three states: (i) the low-density state (LDS), dominated by the largest wrapping length $l_{max}$ in the attractive binding regime and characterized by the absence of ordering; (ii) the intermediate-density state (IDS); and (iii) the high-density state (HDS) characterized by full unwrapping and strong statistical ordering.

Using mean-field arguments, we show that in the low-density state the system behaves similarly to a fixed-size Tonks gas system with the effective particle size $l = \langle l \rangle$, the mean wrapping length. In particular, for the attractive wrapping energies ($\epsilon_0 < 0$), the nucleosome array goes from the LDS to the HDS through an IDS which in fact correspond to the HDS of the mean-field fixed-size system. As the nucleosome density continues to increase, the system undergoes an unwrapping IDS→HDS transition. Interestingly, this unwrapping transition occurs through an equiprobable state, in which all wrapping lengths are equally likely.

The nucleosome repeat length (NRL) is also influenced by the breathing dynamics – the same NRL value can be reached at a lower nucleosome density $\rho$ (which implies a larger mean separation distance) compared with the fixed-size Tonks gas model (Fig. 7C, Table I). This suggests that the wrapping-length fluctuations act as a repulsive "soft-core" pair interaction that, by allowing nucleosomes to overlap, leads (i) to a lower NRL given $\rho$ and (ii) to a higher $\rho$ for the same $\mu$ compared to the hard-core, fixed-size model.

The nucleosome model with wrapping length fluctuations can account for the experimentally observed nucleosome positioning patterns (Fig. 10). Compared to the fixed-size model with $l = 147$ bp, our models better reproduce the statistical ordering patterns, especially after imposing a minimal linker length of 20 bp. This linker length likely reflects both steric constraints due to higher-order chromatin structure and the action of chromatin remodeling factors such as ISWI and CHD1 that are known to control the NRL[52]. The chromatin remodeler action might lead to different linker lengths depending on the genetic background and/or the physiological context, providing a potential explanation for the large range of NRL values observed in different species, cell

types, and functional chromatin states[12,53–55]

The nucleosomal DNA breathing has been previously studied by *in vitro* experiments on mono-nucleosomes or small nucleosome arrays reconstituted on high-affinity positioning sequences[56,57]. In addition, computational analyses of nucleosome unwrapping have been performed using coarse-grained molecular models[58–62]. These experiments and simulations enabled quantitative inference of the wrapping/unwrapping kinetic rates and thus of the unwrapping free energies in a wide range of biochemical and genetic conditions[60,63–65]. Taken together, these studies indicate a rather weak binding association for the entry/exit DNA, in qualitative agreement with our $\epsilon_0$ estimates in Table I. The *in vitro* studies produce a slightly larger value, $|\epsilon_0| \gtrsim 0.1\ k_BT$[61] (Appendix, section 2).

Many factors might contribute to lowering the effective binding energy *in vivo*, including the action of chromatin remodelers that enhance unwrapping by translocating an destabilizing nucleosomes[52]. Interestingly, our models indicate that *in vivo* nucleosomal arrays are in a state that is very close to the unwrapping transition and thus exhibits high plasticity – a small modulation of the effective binding energy or the global density would drive the system toward more unwrapped and thus more accessible states. Finally, since our parameter estimation is based on "bulk" chromatin experimental properties, this plasticity reflects a rather generic feature of chromatin. A more careful analysis of the regional and local modulation of nucleosome plasticity in various species would be required to better understand how genetic and epigenetic factors control the unwrapping transition. For example, one might expect that the boundaries between active and inactive chromatin territories in higher eukaryotes are enriched in more susceptible nucleosomes, whose function is to protect active chromatin from the spreading of repressive factors.

The nucleosome model investigated in this study can be improved in several ways. First, the wrapping length fluctuations could be modeled independently at both ends of the nucleosome core particle. Second, following previous work[36,38], we could consider more complex binding energy rules than the linear model proposed here. In particular, it is possible to take into account the 10 bp periodicity of histone-DNA interactions[2] and the larger binding energies in the 30–40 bp range (cf. Appendix B). Unwrapping may also be promoted by specific DNA sequences such as poly(dA:dT) tracts[41]. Third, future models could explicitly include additional DNA-binding factors such as chromatin remodelers[66], pioneer transcription factors[67,68], and HMG proteins[69] - see e.g. Ref.[70] for a model of this type.

The question of how active (ATP-dependent) and passive (thermodynamic) chromatin remodeling cooperate in regulating genome-wide nucleosome organization and dynamics remains largely unexplored. The development of next-generation models of nucleosome positioning and energetics will enable better understanding of the mechanisms that regulate DNA accessibility in promoters and genic regions, providing biological insights into transcript initiation, elongation, and termination.

## ACKNOWLEDGMENTS

AVM acknowledges the hospitality of the Laboratoire de Physique, CNRS, ENS de Lyon, where this work was initiated due to a CNRS visiting researcher position. BA, CV, and KH acknowledge funding by the Agence Nationale de la Recherche (ANR-20-CE12-013, ANR-21-CE45-0011-01) and by the MITI/CNRS (PIB2023-2024). KH acknowledges HPC IIT Mandi for granting access to their computing facilities. BA and CV thank the French CNRS network GDR "Architecture et Dynamique du Noyau et des Génomes" (ADN&G) for stimulating workshops. All the authors thank R. Blossey and K. Padmanabhan for inspiring discussions.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Kharerin Hungyo**: Conceptualization (equal); Methodology (equal); Software (lead); Validation (equal); Writing – original draft (equal); Writing – review and editing (equal). **Benjamin Audit**: Conceptualization (equal); Funding acquisition (equal); Validation (equal); Writing – review and editing (equal). **Cédric Vaillant**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Validation (equal); Writing – original draft (equal); Writing – review and editing (equal). **Alexandre V. Morozov**: Conceptualization (equal); Formal analysis (equal); Methodology (equal); Validation (equal); Writing – review and editing (equal).

## DATA AND SOFTWARE AVAILABILITY

Our software is available at:
`https://github.com/kharerin/NucBreathModel`

## APPENDIX: THERMODYNAMIC MODEL OF NUCLEOSOME POSITIONING AND BREATHING

Here we describe the theoretical and computational foundations of our thermodynamic approach to nucleosome positioning and unwrapping. Nucleosome breathing is the process of transient DNA unwrapping and

rewrapping in the nucleosome core particle due to thermal fluctuations and other factors such as chromatin remodeling activities[21,71,72]. Nucleosomomal DNA can range from 40 bp to 200 bp in MNase digestion experiments[40], although this is partially due to the MNase digestion kinetics—MNase may both under- and over-digest nucleosome-covered DNA depending on its concentration and the times of exposure[20,40,73–75]. To account for these observations, we model nucleosomes with variable-length DNA footprints.

### A. General formulation of the thermodynamic model

We describe arrays of nucleosomes as a 1D fluid of hard rods of variable size in contact with a thermal bath and a particle reservoir[35,38]. The equilibrium properties of this system are fully described by the grand-canonical partition function:

$$\Xi(\mu, L) = \sum_{N=0}^{N_{max}} e^{\mu N} Z(N, L), \tag{A1}$$

where $\mu$ is the chemical potential and $N_{max}$ is the maximum number of particles that can fit on $L$ bp. Note that we set $\beta = 1/k_BT = 1$ ($k_B$ is the Boltzmann constant and $T$ is the temperature), such that $\mu$ and all binding energies are measured in units of $k_BT$. Finally, $Z(N, L)$ is the canonical configuration partition function:

$$Z(N,L) = \int_{s_0}^{s_{N+1}} ds_N \int_{s_0}^{s_N} ds_{N-1} \ldots \int_{s_0}^{s_2} ds_1 \prod_{i=0}^{N-1} \sum_{l_i=1}^{l_i^*} e^{-E(s_i,l_i) - \sum_{j>i} u(s_j,s_i,l_j,l_i)}. \tag{A2}$$

The notation in Eq. (A2) is as follows:

1. $s_i$ is the genomic position of the left (5') edge of the nucleosome $i$; the position of the right edge is given by $s_i + l_i$, where $l_i$ is the DNA wrapping length: $l_i \in [l_{min}, l_{max}]$, with $l_{max} = 147$ bp, the crystallographic size. Note that $s_0 = 0$ and $s_{N+1} = L$ correspond to the ends of the DNA segment; all genomic positions are in dimensionless (bp) units.

2. $l_i^* = l_i^*(s_i, s_{i+1})$ is the maximum possible wrapping length for the particle $i$:

$$l_i^* = \begin{cases} s_{i+1} - s_i - l_k & \text{if } s_{i+1} - s_i - l_k < l_{max} \\ l_{max} & \text{if } s_{i+1} - s_i - l_k > l_{max} \end{cases} \tag{A3}$$

3. $E(s_i, l_i)$ is the total histone-DNA binding energy.

4. $u(s_j, s_i, l_j, l_i)$ is the two-body potential that we will restrict to the nearest-neighbor interactions: $\sum_{j>i} u(s_j, s_i, l_j, l_i) = u(s_{i+1} - s_i, l_i)$. Moreover, we will consider only the steric exclusion component:

$$u(s_{i+1} - s_i, l_i) = \infty \text{ for } s_{i+1} - s_i < l_i + l_k;\ 0 \text{ otherwise}$$

Steric exclusion interactions enforce the non-overlapping constraints: (1) a given nucleosomal DNA bp cannot belong to two neighboring nucleosomes at the same time ($s_{i+1} > s_i + l_i$) and (2) a 3D "steric" constraint that leads to an additional minimal inter-nucleosomal distance $l_k$ ($s_{i+1} > s_i + l_i + l_k$). Here we assume that $l_k$ is constant, i.e. it does not depend on the wrapping configurations of the two neighboring nucleosomes.

### B. Histone-DNA binding energy

We assume that the energy of a nucleosome with a wrapping length $l$ at a genomic position $s$ is given by:

$$E(s, l) = \sum_{n=s}^{s+l-1} \epsilon_0(n), \tag{B1}$$

where $\epsilon_0(n)$ is the local binding free energy per DNA bp at the genomic position $n$. The $\epsilon_0(n)$ free energies account for the local adsorption energy gain counteracted by the mechanical energy cost due to DNA bending into the superhelical conformation and the associated entropy loss. Both contributions to the energy depend on several biochemical factors such as the ionic environment; the mechanical cost of DNA bending is generally assumed to depend on the underlying DNA sequence[76,77]. Note that our simple additive model does not include any higher-order contributions such as dinucleotide-level DNA bending energies.

In-depth studies of the energetics of DNA wrapping are available from the literature[61,62]. At both ends of a fully-wrapped nucleosome, the binding is relatively weak: $\epsilon_0 \approx -0.1\ k_BT$/bp for up to 30 bp. In the $30 - 40$ bp range, the values of $\epsilon_0$ rise in magnitude to $\approx -0.5\ k_BT$/bp, before becoming more moderate again ($\approx -0.1\ k_BT$/bp) in the $40 - 60$ bp range. Overall, unwrapping of 60 bp from either side of the nucleosome costs $8 - 10\ k_BT$. Hence, unwrapping lengths larger than $30 - 40$ bp are unlikely when considering only thermal fluctuations, but might result from the ATP-driven action of chromatin remodellers and RNA polymerases. Furthermore, such extensive unwrapping is often associated with a loss of H2A-H2B dimers and the resulting reduction of the histone octamer to a hexasome or a tetrasome[78–81].

In this paper, we adopt a linear model of the free energy of nucleosome formation with the constant $\epsilon_0$ (Eq. (1)). In this model, the nucleosome formation energy depends only on the DNA wrapping length $l$ (readily

available from nucleosome mapping experiments) and is independent of the genomic coordinates.

### C. Nucleosome density and occupancy

The density of nucleosomes of the wrapping length $l$ is described by the one-point particle distribution $\rho(s,l)$ along the genome, which measures the equilibrium probability of having a nucleosome particle of length $l$ at the genomic position $s$:

$$\rho(s,l) = -\frac{\partial \ln \Xi(\mu, L)}{\partial E(s,l)}, \tag{C1}$$

which can be expressed as:

$$\rho(s,l) = \frac{\Xi(\mu,[0,s))\Xi(\mu,(s+l,L])e^{\mu - E(s,l)}}{\Xi(\mu,L)}. \tag{C2}$$

Here, $\Xi(\mu,[0,s)) \equiv e^{F(s)}$ is the partial partition function for the $[0,s)$ genome segment, $\Xi(\mu,(s+l,L]) \equiv e^{B(s+l)}$ is the partition function for the $(s+l,L]$ genome segment, and $\Xi(\mu,L)$ is the total partition function. Note that Eq. (C2) can be used to compute $\rho(s,l)$ exactly by iteratively summing the partition function[38].

From the density profile one can derive the genomic occupancy $O(s,l)$ – the probability of a bp $s$ to be part of a nucleosome of size $l$:

$$O(s,l) = \sum_{s'=1}^{l} \rho(s-l+s',l). \tag{C3}$$

The total nucleosome density $\rho(s)$, i.e. the probability to find a nucleosome at the genomic position $s$ regardless of its wrapping length $l$, is given by $\rho(s) = \sum_{l_{min}}^{l_{max}} \rho(s,l)$. The probability distribution of the wrapping length $l$ is then given by $w(s,l) = \rho(s,l)/\rho(s)$. The total nucleosome occupancy, i.e. the probability for the base pair $s$ to be part of any nucleosome regardless of its wrapping length $l$, can be obtained as $O(s) = \sum_{l_{min}}^{l_{max}} O(s,l)$. Note that far away from the boundaries of the DNA segment, $\rho(s,l) \to \rho(l)$, $\rho(s) \to \rho$, $w(s,l) \to w(l)$, $O(s,l) \to O(l)$, and $O(s) \to O$.

### D. Pair-distribution function

The two-particle distribution function is given by:

$$\rho_2(s_1,l_1;s_2,l_2) = \frac{1}{\Xi(\mu,L)} \frac{\partial^2 \Xi(\mu,L)}{\partial E(s_1,l_1)\partial E(s_2,l_2)}, \tag{D1}$$

where $s_{1,2}$ and $l_{1,2}$ are the particle genomic positions and lengths, respectively. For hard-core particles without long-range interactions, Eq. (D1) can be written as:

$$\rho_2(s,l_1;s+\Delta s,l_2) = e^{F(s)+M(s,l_1,\Delta s)+B(s,l_2,\Delta s)-F(L)} \times e^{\mu - E(s,l_1))} e^{\mu - E(s+\Delta s,l_2)}, \tag{D2}$$

where $\Xi(\mu,[0,s)) = e^{F(s)}$, $\Xi(\mu,(s+l_1,s+\Delta s)) = e^{M(s,l_1,\Delta s)}$, and $\Xi(\mu,(s+\Delta s+l_2,L]) = e^{B(s,l_2,\Delta s)}$ are the partial partition functions for the genome segments $[0,s)$, $(s+l_1,s+\Delta s)$, and $(s+\Delta s+l_2,L]$, respectively. The total partition function is given by $\Xi(\mu,L) = e^{F(L)}$.

Finally, the radial pair-distribution function is given by:

$$g^{(2)}(\Delta s) = \frac{1}{\rho^2} \sum_{l_1,l_2} \frac{1}{L} \sum_{s} \rho_2(s,l_1;s+\Delta s,l_2), \tag{D3}$$

where $\rho = L^{-1}\sum_s \rho(s)$ is the genome-wide average of the nucleosome density.

### E. Derivation of the equiprobable state

In this section, we demonstrate the existence of the equiprobable wrapping-length state. As described in the main text, the equiprobable state separates the IDS from the HDS for $\epsilon_0 < 0$. It is defined by $\rho(l) = w(l)\rho \equiv \rho^* = \text{const}$. The objective is to show that for every value of $\epsilon_0$, there is a chemical potential $\mu^*$ that leads to the equiprobable state. Numerically, one can extract $\mu^*(\epsilon_0)$ by either computing the RMSD of $\rho(l)$ or the iso-line $\langle l \rangle = (l_{max}+l_{min})/2$, which is the mean wrapping length in the equiprobable state. In the following, we provide an analytical derivation of $\mu^\star(\epsilon_0)$.

For an inhomogeneous mixture of hard rods with sizes $l \in [l_{min}, l_{max}]$, one can obtain (cf. Eq. 2.28 in Ref.[82]):

$$\mu(l) - E(s,l) = \ln\left[\frac{\rho(s,l)}{1-\sum_{l'}\left(\int_{\delta_{l,l'}}^{\Delta_{l,l'}} \rho(s,l')ds\right)}\right] + \sum_{l'} \int_{-\Delta_{l,l'}}^{\delta_{l,l}} \left[\frac{\rho(s,l')}{1-\sum_{l"}\left(\int_{\delta_{l',l"}}^{\Delta_{l',l"}} \rho(s',l")ds'\right)}\right] ds \tag{E1}$$

with:

$$\Delta_{l_1,l_2} = \frac{1}{2}(l_1+l_2), \quad \delta_{l_1,l_2} = \frac{1}{2}(l_1-l_2).$$

If the binding energy is independent of $s$: $E(s,l) = E(l)$, the bulk density $\rho(s,l) \to \rho(l)$, yielding

$$\mu - E(l) = \ln\left[\frac{\rho(l)}{1-\sum_{l'}\rho(l')l'}\right] + \sum_{l'}\left[\frac{l\rho(l')}{1-\sum_{l"}\rho(l")l"}\right]. \tag{E2}$$

The distribution of particle sizes also becomes independent of $s$: $w(s,l) = w(l) = \rho(l)/\rho$. Thus, the equiprobable state is defined by $w^*(l) \equiv w^\star$, $\forall l$. From the normalization condition $\sum_l w(l) = 1$ we obtain $w^* = 1/\Delta l$, where $\Delta l = l_{max} - l_{min} + 1$. This gives for the equiprobable state: $\rho^* = \rho^*_{tot}/\Delta l$, where $\rho^*_{tot} = \sum_l \rho^*$, and

$\sum_l \rho^* l = \rho^*_{tot}\bar{l}$, where $\bar{l} = (\sum_l l)/\Delta l = (l_{max} + l_{min})/2$, the mean value of $l$ in the equiprobable state.

By denoting $\mu^*(\epsilon_0)$ the chemical potential of the equiprobable state, we obtain:

$$\mu^*(\epsilon_0) - E(l) = \ln\left[\frac{\rho^*/\Delta l}{1-\rho^*\bar{l}}\right] + \frac{\rho^* l}{1-\rho^*\bar{l}} \tag{E3}$$

Assuming that $\mu^*(\epsilon_0)$ is independent of $l$ and using the energy model in Eq. (1), we immediately obtain from Eq. (E3):

$$\begin{aligned}\mu^* &= \ln\left[\frac{\rho^*/\Delta l}{1-\rho^*\bar{l}}\right] - \epsilon_0 l_{min}, \\ -\epsilon_0 &= \frac{\rho^*}{1-\rho^*\bar{l}},\end{aligned} \tag{E4}$$

where the chemical potential was redefined to absorb the formation energy of the nucleosome with the minimal wrapping length $l_{min}$: $\mu^* - E_{core} \to \mu^*$.

We thus find that for $\forall \epsilon_0 < 0$, there is a chemical potential $\mu^*$ that leads to the equiprobable density $\rho^*$ state:

$$\begin{aligned}\rho^* &= \frac{|\epsilon_0|}{1+\bar{l}|\epsilon_0|}, \\ \mu^* &= \ln\frac{|\epsilon_0|}{\Delta l} + |\epsilon_0| l_{min}.\end{aligned} \tag{E5}$$

### F. Derivation of the transition point in the fixed-size Tonks gas titration curve

We consider the Tonks gas model with particles of size $l$. For this model, the titration curve is given by Eq. (2):

$$\mu = \ln\left[\frac{O}{1-O}\right] + \frac{O}{1-O}, \tag{F1}$$

where $\mu = \mu_r + \ln(l)$ is the chemical potential and $O = \rho l$ is the particle occupancy.

The transition point $(O^*, \mu^*)$ is given by the inflection point of the titration curve: $d^2O/d\mu^2|_{\mu^*} = 0$. This point is also the inflection point of the $\mu(O)$ curve and thus can be found by solving $d^2\mu/dO^2|_{O^*} = 0$. From Eq. F1 it follows that

$$\frac{d^2\mu}{dO^2} = \frac{-O^3 + O^2 + 3O - 1}{O^2(1-O)^3}. \tag{F2}$$

Thus, the transition value $O^*$ is the solution of the cubic equation $-O^3 + O^2 + 3O - 1 = 0$ in the $[0, 1]$ interval. Numerical computation yields $O^* = \rho^* l \approx 0.31$. According to Eq. (F1), the corresponding value of the chemical potential at the transition is $\mu^* = \mu_r^* + \ln l \approx -0.34$.

### G. Computational implementation

Our nucleosome model has a fixed wrapping energy per bp, $\epsilon_0$ (Eq. (1)) and a wrapping length in the $l \in [l_{min}, l_{max}]$ range. Nucleosomes are modeled as hard-core particles subject to steric exclusion but no other two-body interactions.

To calculate nucleosome densities in the presence of unwrapping, we adopt a dynamic programming approach[38,41,83–85] in which nucleosomes with a given wrapping length are considered a distinct species of DNA-binding particles. Thus, the particle density $\rho_t(i)$ of the nucleosomes of length $l_{min} \le l_t \le l_{max}$ on a DNA segment of length $L$ can be computed as

$$\rho_t(i) = e^{F(i-1)+B(i+l_t)-F(L)} C_t e^{-E_t(i)}, \tag{G1}$$

where the forward partition function $F(i)$ is given by:

$$F(i) = F(i-1) + \ln\left(1 + \sum_{t=1}^{N} e^{F(i-l_t)-F(i-1)} C_t e^{-E_t(i-l_t+1)}\right) \tag{G2}$$

and the backward partition function $B(i)$ is given by:

$$B(i) = B(i+1) + \ln\left(1 + \sum_{t=1}^{N} e^{B(i+l_t)-B(i+1)} C_t e^{-E_t(i)}\right). \tag{G3}$$

Here, $1 \le i \le L$ denotes the genomic position, $C_t = e^{\mu_t}$, and $\mu_t$ is the chemical potential of the particles of length $l_t$. Note that $F(L)$ represents the total partition function. In this study, we use a DNA segment size of $L = 10^4$ bp in all calculations.

[1] C. Davey, D. Sargent, K. Luger, *et al.*, "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution," J. Mol. Biol. **319**, 1097–1113 (2002).
[2] T. J. Richmond and C. A. Davey, "The structure of DNA in the nucleosome core," Nature **423**, 145–150 (2003).
[3] D. Vasudevan, E. Y. D. Chua, and C. Davey, "Crystal structures of nucleosome core particles containing the '601' strong positioning sequence," J. Mol. Biol. **403**, 1–10 (2010).
[4] C. Woodcock, J. Safer, and J. Stanchfield, "Structural repeating units in chromatin. I. evidence for their general occurrence." Exp. Cell Res. **97**, 101–110 (1976).
[5] J. Widom, "A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells," PNAS USA **89**, 1095–1099 (1992).
[6] D. Tolkunov and A. V. Morozov, "Genomic studies and computational predictions of nucleosome positions and formation energies," Adv. Prot. Chem. Struc. Biol. **79**, 1–57 (2010).
[7] R. V. Chereji and A. V. Morozov, "Functional roles of nucleosome stability and dynamics," Brief. Funct. Genom. **14**, 50–60 (2015).
[8] G. C. Yuan, Y. J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, and O. J. Rando, "Genome-scale identification of nucleosome positions in S. cerevisiae," Science **309**, 626–630 (2005).
[9] T. N. Mavrich, C. Z. Jiang, I. P. Ioshikhes, X. Y. Li, B. J. Venters, S. J. Zanton, L. . P. Tomsho, J. Qi, R. L. Glaser, S. C. Schuster, D. S. Gilmour, I. Albert, and B. F. Pugh, "Nucleosome organization in the Drosophila genome," Nature **453**, 358–362 (2008).
[10] A. Weiner, A. Hughes, M. Yassour, *et al.*, "High-resolution nucleosome mapping reveals transcription-dependent promoter packaging," Genome Res. **20**, 90–100 (2010).

[11]M. F. Dion, T. Kaplan, M. Kim, S. Buratowski, N. Friedman, and O. J. Rando, "Dynamics of replication-independent histone turnover in budding yeast," Science **315**, 1405–1408 (2007).
[12]K. E. van Holde, *Chromatin* (Springer, New York, 1989).
[13]K. Rippe, J. Mazurkiewicz, and N. Kepper, "Interactions of histones with DNA: Nucleosome assembly, stability, dynamics, and higher order structure," in *DNA Interactions with Polymers and Surfactants*, edited by R. Dias and B. Lindman (John Wiley & Sons, Inc., 2008) pp. 135–172.
[14]J. Barbier, C. Vaillant, J.-N. Volff, F. G. Brunet, and B. Audit, "Coupling between sequence-mediated nucleosome organization and genome evolution," Genes **12**, 851 (2021).
[15]R. D. Kornberg, "Chromatin structure: a repeating unit of histones and DNA," Science **184**, 868–871 (1974).
[16]A. Flaus, K. Luger, S. Tan, *et al.*, "Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals," PNAS USA **93**, 1370–1375 (1996).
[17]A. Flaus and T. J. Richmond, "Positioning and stability of nucleosomes on MMTV 3'LTR sequences." J. Mol. Biol. **275**, 427–441 (1998).
[18]C. Jiang and B. F. Pugh, "Nucleosome positioning and gene regulation: advances through genomics," Nat. Rev. Genet. **10**, 161–172 (2009).
[19]Z. Zhang, C. J. Wippo, M. Wal, E. Ward, P. Korber, and B. F. Pugh, "A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome," Science **332**, 977–980 (2011).
[20]J. G. Henikoff, J. A. Belskyb, K. Krassovsky, D. MacAlpine, and S. Henikoff, "Epigenome characterization at single base-pair resolution," PNAS USA **108**, 18318–18323 (2011).
[21]G. Li, M. Levitus, C. Bustamante, and J. Widom, "Rapid spontaneous accessibility of nucleosomal DNA," Nat. Struct. Mol. Biol. **12**, 46–53 (2005).
[22]A. Bucceri, K. Kapitza, and F. Thoma, "Rapid accessibility of nucleosomal DNA in yeast on a second time scale," EMBO J **25**, 3123–3132 (2006).
[23]W. Möbius, R. A. Neher, and U. Gerland, "Kinetic accessibility of buried DNA sites in nucleosomes," Phys. Rev. Lett. **97**, 208102 (2006).
[24]M. G. Poirier *et al.*, "Spontaneous access to DNA target sites in folded chromatin fibers," J. Mol. Biol. **379**, 772–786 (2008).
[25]Y.-J. Park, P. N. Dyer, D. J. Tremethick, and K. Luger, "A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome," J. Biol. Chem. **279**, 24274–24282 (2004).
[26]A. Gansen, A. Valeri, F. Hauger, S. Felekyan, S. Kalinin, K. Toth, J. Langowski, and C. A. M. Seidel, "Nucleosome disassembly intermediates characterized by single-molecule FRET," PNAS USA **106**, 15308–15313 (2009).
[27]J. Kim, J. Lee, and T.-H. Lee, "Lysine acetylation facilitates spontaneous DNA dynamics in the nucleosome," J. Phys. Chem. B **119**, 15001–15005 (2015).
[28]H. Schiessel, "The physics of chromatin," J. Phys.: Cond. Matter **15**, R699–R774 (2003).
[29]G. Längst, V. B. Teif, and K. Rippe, "Chromatin remodeling and nucleosome positioning," in *Genome Organization and Function in the Cell Nucleus*, edited by K. Rippe (Wiley-VCH Verlag GmbH & Co. KGaA, 2011) p. 111–138.
[30]K. Brogaard, L. Xi, J. P. Wang, *et al.*, "A map of nucleosome positions in yeast at base-pair resolution," Nature **486**, 496–501 (2012).
[31]H. A. Cole, B. H. Howard, and D. J. Clark, "Activation-induced disruption of nucleosome position clusters on the coding regions of Gcn4-dependent genes extends into neighbouring genes," Nucleic Acids Res. **39**, 9521–9535 (2011).
[32]J. D. McGhee and G. Felsenfeld, "Another potential artifact in the study of nucleosome phasing by chromatin digestion with micrococcal nuclease," Cell **32**, 1205–1215 (1983).
[33]C. Dingwall, G. P. Lomonosoff, and R. A. Laskey, "High sequence specificity of micrococcal nuclease," Nucleic Acids Res. **9**, 2659–2673 (1981).
[34]H.-R. Chung, I. Dunkel, F. Heise, *et al.*, "The effect of micrococcal nuclease digestion on nucleosome positioning data," PLoS ONE **5**, e15754 (2010).
[35]T. Chou, "An exact theory of histone-DNA adsorption and wrapping," Europhys. Lett. **62**, 753–759 (2003).
[36]V. B. Teif, R. Ettig, and K. Rippe, "A lattice model for transcription factor access to nucleosomal DNA," Biophys. J. **99**, 2597–2607 (2010).
[37]D. A. Beshnova, A. G. Cherstvy, Y. Vainshtein, and V. B. Teif, "Regulation of the nucleosome repeat length in vivo by the DNA sequence, protein concentrations and long-range interactions," PLoS Comput. Biol. **10**, e1003698 (2014).
[38]R. V. Chereji and A. V. Morozov, "Ubiquitous nucleosome crowding in the yeast genome," PNAS **111**, 5236–5241 (2014).
[39]R. V. Chereji, S. Ramachandran, T. D. Bryson, and S. Henikoff, "Precise genome-wide mapping of single nucleosomes and linkers in vivo," Genome Biol. **19**, 19 (2018).
[40]R. V. Chereji, J. Ocampo, and D. J. Clark, "MNase-sensitive complexes in yeast: Nucleosomes and non-histone barriers," Mol. Cell **65**, 565–577.e3 (2017).
[41]E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom, "A genomic code for nucleosome positioning," Nature **442**, 772–778 (2006).
[42]L. Tonks, "The complete equation of state of one, two and three-dimensional gases of hard elastic spheres," Phys. Rev. **50**, 955–963 (1936).
[43]J. Zlatanova, T. C. Bishop, J.-M. Victor, V. Jackson, and K. van Holde, "The nucleosome family: dynamic and growing," Structure **17**, 160–71 (2009).
[44]L. J. Hsieh, M. A. Gourdet, C. M. Moore, E. N. Muñoz, N. Gamarra, V. Ramani, and G. J. Narlikar, "A hexasome is the preferred substrate for the INO80 chromatin remodeling complex, allowing versatility of function," Mol. Cell **82**, 2098–2112.e4 (2022).
[45]M. Zhang, A. Jungblut, F. Kunert, T. H. Luis Hauptmann and, O. Kolesnikova, F. Metzner, M. Moldt, F. Weis, F. DiMaio, K.-P. Hopfner, and S. Eustermann, "Hexasome-INO80 complex reveals structural basis of noncanonical nucleosome remodeling," Science **381**, 313–319 (2023).
[46]M. Nishimura, T. Fujii, H. Tanaka, K. Maehara, K. Morishima, M. Shimizu, Y. Kobayashi, K. Nozawa, Y. Takizawa, M. Sugiyama, Y. Ohkawa, and H. Kurumizaka, "Genome-wide mapping and cryo-EM structural analyses of the overlapping tri-nucleosome composed of hexasome-hexasome-octasome moieties," Commun. Biol. **7**, 61 (2024).
[47]B. Onoa, C. Díaz-Celis, C. Cañari-Chumpitaz, A. Lee, and C. Bustamante, "Real-time multistep asymmetrical disassembly of nucleosomes andchromatosomes visualized by high-speed atomic force microscopy," ACS Cent. Sci. **10**, 122–137 (2024).
[48]W. Möbius, B. Osberg, A. M. Tsankov, O. J. Rando, and U. Gerland, "Toward a unified physical model of nucleosome patterns flanking transcription start sites," PNAS **110**, 5719–5724 (2013).
[49]E. Burgos and H. Bonadeo, "Exact solution for the pair distribution function of a onedimensional equilibrium fluid of hard rods with arbitrary size distribution," J. Chem. Phys. **88**, 1163–1169 (1988).
[50]P. V. Giaquinta, "Entropy and ordering of hard rods in one dimension," Entropy **10**, 248–260 (2008).
[51]A. K. Singh and F. Mueller-Planitz, "Nucleosome positioning and spacing: From mechanism to function," J Mol. Biol. **433**, 166847 (2021).
[52]J. Ocampo, R. V. Chereji, P. R. Eriksson, and D. J. Clark, "Contrasting roles of the RSC and ISW1/CHD1 chromatin remodelers in RNA polymerase II elongation and termination," Genome Res. **29**, 407–417 (2019).
[53]H. J. Szerlong and J. C. Hansen, "Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure," Biochem. Cell Biol. **89**, 24–34 (2011).
[54]E. Oberbeckmann, V. Niebauer, S. Watanabe, L. Farnung,

M. Moldt, A. Schmid, P. Cramer, C. L. Peterson, S. Eustermann, K.-P. Hopfner, and P. Korber, "Ruler elements in chromatin remodelers set nucleosome array spacing and phasing," Nat. Commun. **12**, 3232 (2021).

[55]A. K. Singh, T. Schauer, L. Pfaller, T. Straub, and F. Mueller-Planitz, "The biogenesis and function of nucleosome arrays," Nat. Commun. **12**, 7011 (2021).

[56]K. J. Polach and J. Widom, "Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation," J. Mol. Biol. **254**, 130–49 (1995).

[57]B. T. Donovan, Y. Luo, Z. Meng, and M. G. Poirier, "The nucleosome unwrapping free energy landscape defines distinct regions of transcription factor accessibility and kinetics," Nucleic Acids Res. **51**, 1139–1153 (2023).

[58]S. Li, T. Wei, and A. R. Panchenko, "Histone variant H2A.Z modulates nucleosome dynamics to promote DNA accessibility," Nat. Commun. **14**, 769 (2023).

[59]G. N. Rychkov, A. V. Ilatovskiy, I. B. Nazarov, A. V. Shvetsov, D. V. Lebedev, A. Y. Konev, V. V. Isaev-Ivanov, and A. V. Onufriev, "Partially assembled nucleosome structures at atomic detail," Biophys. J. **112**, 460–472 (2017).

[60]J. Culkin, L. de Bruin, M. Tompitak, R. Phillips, and H. Schiessel, "The role of DNA sequence in nucleosome breathing," Eur. Phys. J. E. **40**, 106 (2017).

[61]D. Zhao, J. V. Le, M. A. Darcy, K. Crocker, M. G. Poirier, C. Castro, and R. Bundschuh, "Quantitative modeling of nucleosome unwrapping from both ends," Biophys. J. **117**, 2204–2216 (2019).

[62]R. A. Forties, J. A. North, S. Javaid, O. P. Tabbaa, R. Fishel, M. G. Poirier, and R. Bundschuh, "A quantitative model of nucleosome dynamics," Nucleic Acids Res. **39**, 8306–8313 (2011).

[63]M. Simon, J. A. North, J. C. Shimko, R. A. Forties, M. B. Ferdinand, M. Manohar, M. Zhang, R. Fishel, J. J. Ottesen, and M. G. Poirier, "Histone fold modifications control nucleosome unwrapping and disassembly," PNAS **108**, 12711–12716 (2011).

[64]J. A. North, J. C. Shimko, S. Javaid, A. M. Mooney, M. A. Shoffner, S. D. Rose, R. Bundschuh, R. Fishel, J. J. Ottesen, and M. G. Poirier, "Regulation of the nucleosome unwrapping rate controls DNA accessibility," Nucleic Acids Res. **40**, 10215–10227 (2012).

[65]A. V. Onufriev, "Biologically relevant small variations of intracellular ph can have significant effect on stability of protein–DNA complexes, including the nucleosome," Front Mol. Biosci. **10**, 106778 (2023).

[66]M. F. Poyton, X. A. Feng, A. Ranjan, Q. Lei, F. Wang, J. S. Zarb, R. K. Louder, G. Park, M. H. Jo, J. Ye, S. Liu, T. Ha, and C. Wu, "Coordinated DNA and histone dynamics drive accurate histone H2A.Z exchange," Sci. Adv. **8**, eabj5509 (2022).

[67]Y. Sun, C.-Y. Nien, K. Chen, H.-Y. Liu, J. Johnston, J. Zeitlinger, and C. Rushlow, "Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation," Genome Res. **25**, 1703–14 (2015).

[68]S. R. Joseph, M. Pálfy, L. Hilbert, M. Kumar, J. Karschau, V. Zaburdaev, A. Shevchenko, and N. L. Vastenhouw, "Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos," Elife **6**, e23326 (2017).

[69]R. Hock, T. Furusawa, T. Ueda, and M. Bustin, "HMG chromosomal proteins in development and disease," Trends Cell Biol. **17**, 72–9 (2007).

[70]V. B. Teif and K. Rippe, "Nucleosome mediated crosstalk between transcription factors at eukaryotic enhancers," Phys. Biol. **8**, 044001 (2011).

[71]H. S. Tims, K. Gurunathan, M. Levitus, and J. Widom, "Dynamics of nucleosome invasion by DNA binding proteins," J. Mol. Biol. **411**, 430–448 (2011).

[72]G. A. Armeev, A. S. Kniazeva, G. A. Komarova, M. P. Kirpichnikov, and A. K. Shaytan, "Histone dynamics mediate DNA unwrapping and sliding in nucleosomes," Nat. Commun. **12**, 2387 (2021).

[73]S. Kubik, M. J. Bruzzone, P. Jacquet, J.-L. Falcone, J. Rougemont, and D. Shore, "Nucleosome stability distinguishes two different promoter types at all protein-coding genes in yeast," Mol. Cell **60**, 422–34 (2015).

[74]S. Brahma and S. Henikoff, "RSC-associated subnucleosomes define mnase-sensitive promoters in yeast," Mol. Cell **73**, 238–249.e3 (2019).

[75]S. Wernig-Zorc, F. Kugler, L. Schmutterer, P. Räß, C. Hausmann, S. Holzinger, G. Längst, and U. Schwartz, "nucMACC: An MNase-seq pipeline to identify structurally altered nucleosomes in the genome," Sci. Adv. **10**, eadm9740 (2024).

[76]W. K. Olson, A. A. Gorin, X.-J. Lu, L. M. Hock, and V. B. Zhurkin, "DNA sequence-dependent deformability deduced fromprotein–dna crystal complexes," PNAS **95**, 11163–11168 (1998).

[77]A. V. Morozov, K. Fortney, D. A. Gaykalova, V. M. Studitsky, J. Widom, and E. D. Siggia, "Using DNA mechanics to predict in vitro nucleosome positions and formation energies," Nucleic Acids Res. **37**, 4707–22 (2009).

[78]R. F. Levendosky, A. Sabantsev, S. Deindl, and G. D. Bowman, "The Chd1 chromatin remodeler shifts hexasomes unidirectionally," Elife **5**, 5:e21356 (2016).

[79]E. A. Morrison, L. Baweja, M. G. Poirier, J. Wereszczynski, and C. A. Musselman, "Nucleosome composition regulates the histone H3 tail conformational ensemble and accessibility," Nucleic Acids Res. **49**, 4750–4767 (2021).

[80]Y. Arimura, H. Tachiwana, T. Oda, M. Sato, and H. Kurumizaka, "Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome," Biochemistry **51**, 3302–9 (2012).

[81]M. Brehove, E. Shatoff, B. T. Donovan, C. M. Jipa, R. Bundschuh, and M. G. Poirier, "DNA sequence influences hexasome orientation to regulate dna accessibility," Nucleic Acids Res. **47**, 5617–5633 (2019).

[82]T. K. Vanderlick, H. T. Davis, and J. K. Percus, "The statistical mechanics of inhomogeneous hard rod mixtures," J. Chem. Phys. **91**, 7136–7145 (1989).

[83]N. Kaplan, I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, D. Tillo, Y. Field, E. M. LeProust, T. R. Hughes, J. D. Lieb, J. Widom, and E. Segal, "The DNA-encoded nucleosome organization of a eukaryotic genome," Nature **458**, 362–366 (2009).

[84]E. A. Ozonov and E. van Nimwegen, "Nucleosome free regions in yeast promoters result from competitive binding of transcription factors that interact with chromatin modifiers," PLoS Comput. Biol. **9**, e1003181 (2013).

[85]H. Kharerin and L. Bai, "Thermodynamic modeling of genome-wide nucleosome depleted regions in yeast," PLoS Comput. Biol. **17**, e1008560 (2021).