THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Beyond Sequence: Internucleosomal Interactions Dominate Array Assembly

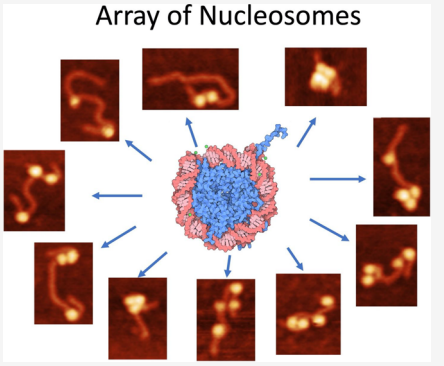
Yaqing Wang, Tommy Stormberg, Mohtadin Hashemi, Anatoly B. Kolomeisky, and Yuri L. Lyubchenko*

Cite This: *J. Phys. Chem. B* 2022, 126, 10813−10821

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The organization of the nucleosome array is a critical component of the chromatin assembly into higher order structure as well as its function. Here, we investigated the contributions of the DNA sequence and internucleosomal interactions on the organization of the nucleosomal arrays in compact structures using atomic force microscopy. We assembled nucleosomes on DNA substrates allowing for the formation of tetranucleosomes. We found that nucleosomes are capable of close positioning with no discernible space between them, even in the case of assembled dinucleosomes. This morphology of the array is in contrast with that observed for arrays assembled with repeats of the nucleosome positioning motifs separated by uniform spacers. Simulated assembly of tetranucleosomes by random placement along the substrates revealed that nucleosome array compaction is promoted by the interaction of the nucleosomes. We developed a theoretical model to account for the role of DNA sequence and internucleosomal interactions in the formation of the nucleosome structures. These findings suggest that, in the chromatin assembly, the affinity of the nucleosomes to the DNA sequence and the strengths of the internucleosomal interactions are the two major factors defining the compactness of the chromatin.



INTRODUCTION

DNA in eukaryotic cells is packaged into chromatin through extensive association with histone proteins.[1−3] The nucleosome is the fundamental unit of chromatin, which regulates the readout and expression of the eukaryotic genome.[4−6] It is a DNA−protein complex with approximately 147 base pairs of DNA wrapped around a protein core complex known as the histone octamer.[7−9] Canonical histone octamers consist of two copies of four core histone proteins, H2A, H2B, H3, and H4.[10] The positively charged histone octamers bind strongly to the negatively charged DNA. X-ray crystallography revealed the atomic structure of the nucleosome and explained how DNA is wrapped around histone octamers in a superhelix of approximately one and three-quarters turns.[11]

The spatial organization of nucleosomes in chromatin continues to be a source of debate. Initially, it was proposed that nucleosomes condense into a 30-nm-diameter chromatin fiber, supported by experiments with the use of electron microscopy (EM) or X-ray scattering analyses of chromatin extracted from various organisms.[12−15] Most recently, however, a study utilizing a combination of EM topography with a developed labeling method (ChromEMT) did not support the assembly of ordered 30 nm fibrils.[16] Instead, they showed the assembly of 10 nm fibers in the cell that are not uniform; rather, they are heterogeneous and vary in diameter between 5 and 24 nm. Potential reasons for this discrepancy are discussed in the recent review article,[17] in which the major role is given to electrostatics, as ionic strength for experiments *in vitro* and *in vivo* are very different. The authors also suggest that the absence of the 30 nm fiber formation can be due to nucleosome loss or irregular nucleosome spacing in native chromatin.

The findings in ref 16 are in line with publications 18−20, in which it has then been proposed that nucleosome fibers exist in a highly disordered, interdigitated state. What is the reason for such irregular spacing of nucleosomes? We have recently shown that the internucleosomal distance within dimers of nucleosomes varies depending on DNA sequence.[21] No such effect is detected if repeats of a highly sequence-specific DNA motif as the Widom 601 sequence are used in similar experiments.[22] Note that many structural studies of chromatin, including papers cited above, utilized repeats of the 601 motif.[23] Furthermore, according to our recent paper,[24] the 601 sequence has a considerably higher affinity for binding nucleosome cores. For a DNA substrate with 372 bp containing the 601 motif, the mononucleosomes are assembled at the position corresponding to the location of the 601 motif

Received: July 27, 2022
Revised: November 28, 2022
Published: December 14, 2022

with a 98% preference. Moreover, this affinity does not affect the DNA wrapping as both 601 and nonspecific sequences wrap the same amount of DNA.[25]

Based on these studies we hypothesize that affinity to histone cores to DNA sequence can be a critical factor for the nucleosome array assembly stabilized by the internucleosomal interactions. To test this hypothesis, here we designed a DNA substrate containing a single instance of the specifically positioning 601 motif and an extended nonspecific sequence from plasmid DNA or a fully nonspecific DNA substrate. The experimental studies utilized single-molecule AFM, which can characterize the nanoscale structure of biological systems.[26−30] We found that nucleosomes on such DNA templates form compact assemblies with close contacts between the nucleosomes. Monte Carlo simulation of random nucleosome placement along these substrates revealed that the compact assemblies observed experimentally occur at a much higher rate than expected with simple nonspecific positioning. We propose a theoretical model according to which the affinity of the nucleosome core to the DNA sequence and the internucleosomal interactions are the two major factors defining the compact assembly of the nucleosome arrays.

METHODS

DNA Substrate. The 601 DNA substrate used in nucleosome assembly contains the 147 bp Widom 601 sequence flanked by plasmid DNA, 113 bp and 738 bp in length (shown in Figure 1A). It is generated from PCR using a plasmid vector pUC57 with the forward primer (5′-GATGTGCTGCAAGGCGATTAAG-3′) and the reverse primer (5′-GGGTTTCGCCACCTCTGAC-3′). The 998 bp nonspecific DNA substrate used is the same length as the 601 substrate, except that the 147 bp Widom 601 sequence is replaced by 147 bp of nonspecific DNA (shown in Figure 1B). The DNA substrates were concentrated from the PCR product and purified using gel electrophoresis. DNA concentration was then determined using a NanoDrop Spectrophotometer (ND-1000, Thermo Fischer) before being used for nucleosome assembly.

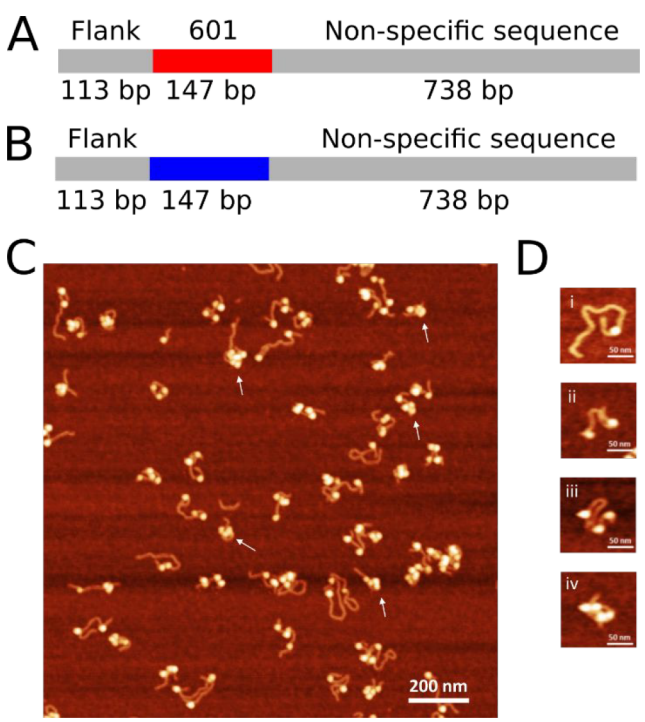


Figure 1. Nucleosome substrates and assembly. (A) Schematic of 601 DNA substrate. (B) Schematic of nonspecific DNA substrate. (C) Representative AFM image of nucleosomes assembled on the 601 DNA substrate. (D) Selected images of different oligonucleosomes assembled on the 601 DNA substrate.

Nucleosome Assembly. Nucleosomes were assembled using a gradient dilution method optimized from our previous research.[21,30−32] Recombinant human histone octamers were purchased from The Histone Source (Fort Collins, CO) for assembly. Before assembly, histones were dialyzed against the initial dialysis buffer (10 mM Tris pH 7.5, 2 M NaCl, 1 mM EDTA, 2 mM DTT) at 4 °C for 1 h. DNA (25 pmol) was then mixed with the histone octamer at a molar ratio of 1:5. The total volume of the mixture was adjusted to 10 μL with 5 M NaCl and DDI H_2O so that the starting concentration of NaCl in the reaction is 2 M. The mixture was diluted with dilution buffer (10 mM Tris pH 7.5) using a syringe pump (0.07 μL/min for 1000 min) to gradually decrease the salt concentration to 0.25 M NaCl, allowing the histone to bind the DNA and form the nucleosome core particle. The nucleosomes were then dialyzed for 1 h against a fresh low salt buffer (10 mM Tris pH 7.5, 2.5 mM NaCl, 1 mM EDTA, 2 mM DTT) before being diluted to 300 nM and stored at 4 °C. The final concentration of the nucleosome was adjusted to 2 nM right before deposition using an imaging buffer (10 mM HEPES pH 7.5, 4 mM $MgCl_2$).

AFM Imaging and Data Analysis. Sample preparation for AFM imaging was performed as described in refs 21, 30, and 33. Briefly, the nucleosome sample was deposited onto the 1-(3-aminopropyl) silatrane (APS) functionalized mica and incubated for 2 min at room temperature. After which, the sample was rinsed with DDI H_2O and dried with a gentle argon flow. The samples were stored under vacuum conditions before being imaged.

Images were acquired using tapping mode under ambient conditions on a MultiMode 8, Nanoscope V system (Bruker, Santa Barbara, CA) using TESPA probes (320 kHz nominal frequency and a 42 N/m spring constant) from the same vendor. The dry sample AFM images were analyzed using the FemtoScan Online software package (Advanced Technologies Center, Moscow, Russia).

DNA contour length analysis was performed by tracing the DNA from one free end to the other. The mean measured contour length of free DNA on an image was divided by the known length of the given substrate, yielding the conversion factor. Flank measurements for the nucleosomes were obtained by measuring from the DNA end to the center of the nucleosome for both arms. Five nanometers was subtracted from each measured flank length to account for the contribution by the histone core. The flank length measurements were divided by the calculated conversion factor to convert from nanometers to base pair values. Schematically, the analysis process is depicted in Supplementary Figure S1.

The initial distinction between compact and well separated nucleosome subpopulations was performed visually, by separately grouping nucleosomes that were observed to be in direct contact with those with clearly distinguishable DNA between nucleosome core complexes. For further analysis, and in simulations and theoretical calculations, the internucleosomal distance cutoff used (30 bp) was determined by measuring the center-to-center distance of dinucleosomes, subtracting 10

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813−10821

nm to account for the contribution of each histone core, and dividing by the calculated conversion factor to convert from nanometers to base pairs. The histograms were generated using OriginPro software (OriginLab Corporation, Northampton, MA, USA).

Monte Carlo Simulations and Data Analysis. Monte Carlo simulations for random placement of 147 bp segments on a 998 bp DNA substrate were performed using bedtools (v2.30.0).[34] Each simulation randomly placed four segments on the DNA substrate, prohibiting overlap of the segments; simulations were repeated 1000 times for each substrate. Simulations were performed for substrates with no sites occupied, one site occupied (113−259 bp) corresponding to the substrate with positioning sequence, and a shifted occupied site (148−294 bp) corresponding to a substrate with positioning sequence allowing nucleosome assembly on both sides of the specific sequence. Data were compiled in Excel and grouped into subpopulations based on their proximity to the next nucleosome along the array using the cutoff above. Nucleosome pairs within 30 bp of one another were considered compact, while those greater than 30 bp from one another were considered well separated.

RESULTS

The 601 DNA substrate consists of a positioning sequence, the 147 bp Widom 601 sequence, flanked by nonspecific DNA sequences that can be wrapped by three or more histone octamers, as shown in Figure 1A. The Widom sequence provides a high affinity binding for the histone octamers,[35] while the nonspecific sequence provides an insight into a more biologically relevant aspect for the assembly of a nucleosome array. The nonspecific DNA substrate is identical in length to the 601 substrate, except that the 147 bp Widom 601 sequence is replaced by 147 bp of nonspecific DNA, as shown in Figure 1B. The nucleosomes were assembled with the DNA substrate and histone octamer at a ratio of 1:5 DNA/octamer. A representative AFM image of the nucleosome assembly is shown in Figure 1C, where the bright features are the nucleosome cores and the string-like strands are the unwrapped DNA in the complexes.

Morphology of Nucleosome Arrays. From the acquired AFM images, we analyzed the morphology of the nucleosome arrays. Nucleosomes appear as bright globular features localized on DNA strands. The nucleosomes on the same DNA template can be separated, but immediately apparent was the compact morphology of nucleosome clusters. Examples of such compact arrangements of the nucleosomes are shown in Figure 1C, as indicated by the white arrows. These arrangements are irregular, resulting in nucleosome clusters of varying size and position. This result indicates that nucleosome arrays assembled on nonspecific DNA adopt heterogeneous structure, as opposed to the ordered nucleosome arrays formed on substrates with repeated positioning DNA sequences imaged with AFM.[36−38]

Along with large nucleosome clusters, mono-, di-, tri-, and tetranucleosomes were observed. Representative snapshots of each type of nucleosome are shown in Figure 1D, with mononucleosome in frame (i), dinucleosome in frame (ii), trinucleosome in frame (iii), and tetranucleosome in frame (iv). We separated each type of nucleosome into groups based on the number of nucleosomes in the array and performed analysis. The yield of each oligonucleosome is shown in Table 1 (n = 515). We found that mononucleosomes are the least commonly observed, while trinucleosomes and tetranucleosomes are the most common, indicating a preference for assembling higher order structures in our experimental setup.

Table 1. 601 Nucleosome Subpopulations

	yield (n = 515)	subpopulation	
free DNA	4.1%	N/A	
mononucleosome	17.5%	N/A	
dinucleosome	21.2%	1−1	78.0%
		2	22.0%
trinucleosome	31.1%	1−1−1	30.6%
		2−1	51.3%
		3	18.1%
tetranucleosome	26.2%	1−1−1−1	6.7%
		2−1−1	18.5%
		2−2	13.3%
		3−1	41.5%
		4	20.0%

We further segregated the groups of oligonucleosomes into subpopulations based on the proximity between nucleosomes on each array. Representative images of each subpopulation are shown in Figure 2. For dinucleosomes (Figure 2A), this resulted in two subpopulations; well separated (1−1) and compact (2) nucleosomes. For trinucleosomes (Figure 2B), three subpopulations exist: well separated (1−1−1), compact (3), and hybrid (2−1) nucleosomes. Tetranucleosomes (Figure 2C) contain five subpopulations: well separated (1−1−1−1), compact (4), and three hybrid structures, depending on the degree of compaction (2−1−1, 2−2, and 3−1). The yields of the subpopulations are shown in Table 1. The well-separated dinucleosome accounts for 78.0% of all the dinucleosome population, while in the trinucleosome group, the well-separated subpopulation is observed in only 30.6% of cases. It decreases substantially in the tetranucleosome group, in which the well-separated tetranucleosomes represent only 6.7% of the population. The decreasing population of the well-separated subgroup in the higher-ordered structures further emphasizes that the compact morphologies are more favorable than the well-separated structures with distant spacing on this DNA substrate. Additionally, the "3−1" subpopulation accounts for 41.5% of all tetranucleosomes observed, compared with 20.0% in the "4" subpopulation. This suggests that the trinucleosome may be the basic unit in the assembly of the nucleosomal array, rather than the tetranucleosome.

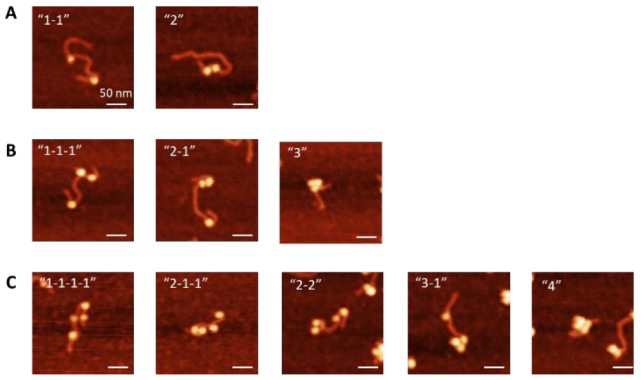


Figure 2. Subpopulations of oligonucleosomes. Scale bars indicate 50 nm.

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813−10821

Internucleosomal Interaction within Oligonucleosomes. In addition to the compact morphology, internucleosomal distances were also characterized. Figure 3 shows the

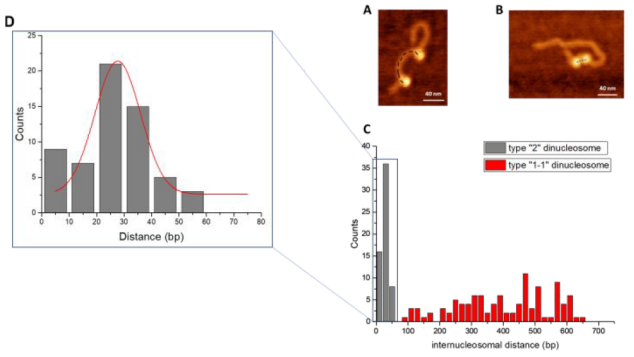


Figure 3. Analysis of internucleosomal distance in the dinucleosome population. (A) Representative image of well separated dinucleosome (type "1−1"). (B) Representative image of compact dinucleosome (type "2"). Scale bars indicate 40 nm. (C) Histogram of internucleosomal distances, n = 167. (D) Distribution of internucleosomal distances in compact dinucleosomes indicate a peak value of 28 ± 2 bp (SEM).

data for both the well separated and compact dinucleosomes. The internucleosomal distances were measured from the center of one nucleosome to the center of the nearest neighboring nucleosome, as shown in Figure 3A,B. A representative image of measured distance for well separated nucleosomes (type 1−1) is shown in Figure 3A, while an image of measured distance for compact nucleosomes (type 2) is shown in Figure 3B. Ten nanometers was subtracted from each measured center-to-center distance to account for the size contributed by the histone core in each nucleosome. The results of these measurements were plotted and are shown in Figure 3C. We see that the separation between the well separated nucleosome and the compacted nucleosomes starts from 50 bp. This distance varies, with distances between nucleosomes as far as 650 bp apart, with no clear preference for positioning. In contrast, the distance of the compact nucleosomes is in the range of 0−40 bp, with a peak population centered at 28 ± 2 bp (SEM), as seen in Figure 3D. This indicates that nucleosomes within 30 bp of one another are most favorable for the internucleosomal interactions to compact the nucleosomes into proximity.

The same measurements were performed on the "2−1" subpopulation of trinucleosomes, and the results can be seen in Figure S2. The distance between the two compacted nucleosomes is termed distance 1, while the distance between the well separated nucleosomes is termed distance 2. The combined data for internucleosomal distance, shown in Figure S2B, indicates the same trend seen in the dinucleosome subpopulation. Compact nucleosomes are found in a very narrow region below 40 bp from one another, while well separated nucleosomes are found anywhere from 90 to more than 500 bp from one another. This result is in agreement with the results for dinucleosomes and further suggests that close proximity of nucleosomes is most favorable for the internucleosomal interactions to compact the nucleosomes into uniform structures.

Nucleosomal Positioning. In order to determine the influence of the positioning 601 motif on the positioning and compaction of nucleosomes, we plotted the lengths of the free DNA flanks observed in our nucleosome arrays. The 601 motif was positioned 113 bp from one end of the DNA substrate; therefore, flank lengths of 113 bp would correspond to nucleosome positioning on the motif. The data for well separated dinucleosome flank length are shown in Figure 4.

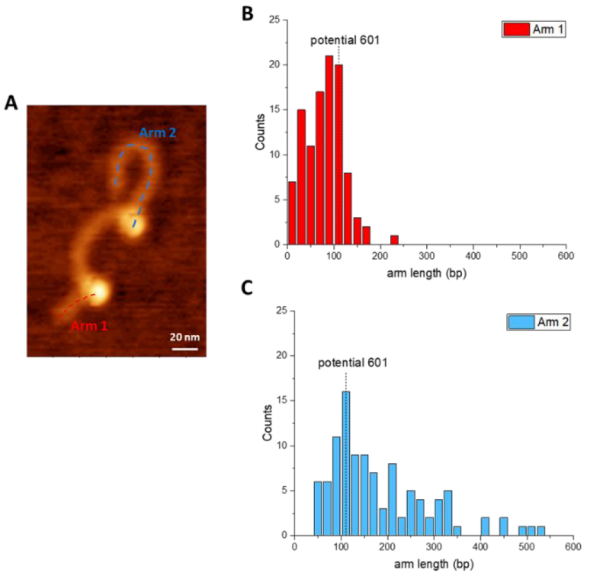


Figure 4. Analysis of flank DNA length in well separated dinucleosomes, n = 106. Scale bar indicates 20 nm.

Interestingly, while a large portion of nucleosomes were positioned at approximately 113 bp, nucleosome position varies, with nucleosomes positioning at or near the end of the DNA. This indicates that the preferential positioning of nucleosomes to the 601 motif is influenced by end binding and the total length of the DNA substrate.

Compact structures of dinucleosomes and trinucleosomes both exhibited a similar positioning. The results are shown in Figure S3. The flank length of compact structures falls within the range of the 601 motif but can vary as far as 250 bp away in the case of dinucleosomes. Compact trinucleosomes exhibit a lower range of positioning according to DNA flank length but contain more structures positioned near the end of the DNA substrate. Overall, the presence of a single positioning nucleosome sequence does not appear to dictate the assembly pattern of other nucleosomes along the substrate, suggesting that array assembly is most dependent on interactions between nucleosomes.

Nucleosome Arrays with No Positioning Sequence. To further investigate the role of a positioning sequence in the assembly of higher order nucleosome structures, we prepared the DNA substrate comprised exclusively of a nonspecific DNA sequence of the same length as our substrate containing the 601 motif. The 147 bp 601 motif was replaced by 147 bp of nonspecific DNA while maintaining the same sequence for the rest of the substrate. A schematic of the substrate is shown in Figure 1B, while a representative image of the assembled nucleosomes acquired through static AFM imaging is shown in Figure S4A. We found that, like nucleosomes assembled on the substrate containing the positioning sequence, the assembly was rather heterogeneous, including compact structures of varying sizes.

We again separated the oligonucleosomes into subpopulations; representative images of the oligonucleosome subgroups

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813−10821

can be seen in Figure S4B–D. A table of the oligonucleosome yields is shown in Table 2 (n = 498). It is notable that well

Table 2. Non-specific Nucleosome Subpopulations

	yield (n = 498)	subpopulation	
free DNA	4.0%	N/A	
mononucleosome	5.4%	N/A	
dinucleosome	17.5%	1–1	65.5%
		2	34.5%
trinucleosome	46.4%	1–1–1	34.6%
		2–1	47.2%
		3	18.2%
tetranucleosome	26.7%	1–1–1–1	6.8%
		2–1–1	28.6%
		2–2	12.0%
		3–1	30.8%
		4	21.8%

separated subpopulation yields decrease in higher order structures, and the "3–1" population is the most common morphology—both observations match those of the substrate containing a single positioning sequence. However, on this substrate, tri- and tetranucleosomes are more common compared with those assembled on the substrate containing a single positioning sequence (73.1% vs 57.3% of total complexes, p = 0.00000013, Fisher's exact test). Moreover, the overall population of well separated nucleosomes is diminished in the absence of a positioning sequence in observed dinucleosomes (65.5% vs 78%, p = 0.0557, Fisher's exact test) and in tetranucleosomes (6.8% vs 6.7%, p = 1.0, Fisher's exact test). Additionally, comparing the overall populations between the two substrates also shows significant differences (p = 0.00596, two-tailed t test). This result indicates that the absence of a positioning sequence may allow nucleosomes to move more freely along the substrate, resulting in more internucleosomal interactions and therefore more compact structures.

Simulations of Nucleosome Positioning. We simulated the random placement of nucleosomes along the substrates to assess the impact of sequence and internucleosomal interactions in nucleosome proximity compared with random placement. For the nonspecific sequence, tetranucleosomes were simulated by randomly placing four particles occupying 147 bp along the substrate with no allowance for overlap. For the 601 sequence, we simulated the substrate with the region containing the 601 motif (from 113 to 259 bp) as occupied to mimic a homogeneously bound nucleosome in that region; three particles were then randomly placed along the remainder of the substrate, reflecting a randomly assembled tetranucleosome array with one nucleosome bound to the 601 region. We also simulated a shifted version of the 601 substrate—the region from 148 to 294 bp was preoccupied, while three particles were randomly placed along the rest of the substrate. This setup allowed for a nucleosome to bind upstream of the excluded region. The results were analyzed and grouped into subpopulations based on their proximity to the next nucleosome along the array.

Nucleosome pairs within 30 bp of one another were considered compact, while those greater than 30 bp from one another were considered well separated, as determined from the data shown in Figure 3. The yield of subpopulations is shown in Table 3 (N = 1000). Nonspecific denotes the

Table 3. Simulated Tetranucleosome Subpopulations

subpopulation	nonspecific yield	601 yield	601 shifted yield
1–1–1–1	33.2%	29.7%	24.8%
2–1–1	51.1%	50.6%	52.4%
2–2	4.5%	4.6%	5.7%
3–1	10.3%	14.8%	16.3%
4	0.9%	0.3%	0.8%

nonspecific DNA substrate simulation; 601 denotes the 601 substrate simulation, and 601 shifted denotes the shifted 601 substrate simulation. There are stark differences between the experimental and simulation data. Simulated data reveal the "2–1–1" subpopulation to be the most common morphology along all three setups, representing approximately half the data set in all three cases. The "4" subpopulation of simulated tetranucleosomes is particularly low, representing less than 1% of the data set in all three cases, dropping as low as 0.3% for 601 tetranucleosomes. Yet, experimental data show that the "2–1–1" subpopulation represents as much as 28.6% of the data set for nonspecific tetranucleosomes. The "4" subpopulation represents slightly less: 21.8% for nonspecific tetranucleosomes. Likewise, for 601 tetranucleosomes, experimental data reveal a mere 18.5% of tetranucleosomes in the "2–1–1" conformation, and 20.0% are observed in the "4" conformation. A full comparison of all three simulations compared with the experimental data sets can be seen in Tables S1 and S2. Overall, the simulation data show a significant preference for well separated nucleosomes compared to the experimental data. In fact, the assembly of the "4" conformation is more than 24 times more likely in experiments (with either substrate) compared with their simulated counterparts, with the "4" conformation in 601 tetranucleosomes being nearly 65 times more likely (Tables S1 and S2). This indicates that nucleosome positioning along the DNA substrate is not random—it is very likely influenced by internucleosomal interactions. These results are also corroborated by statistical analysis that shows significant differences (at $p < 0.005$ or better, χ^2 test) between the simulated and observed population distributions (Table S3).

Theoretical Model for Nucleosome Interaction and Positioning. To understand the dynamics of interactions between nucleosomes and DNA as well as internucleosomal interactions, we develop a simple chemical model that allows us to extract the most relevant properties of these complex processes. Based on the experimental constructs, we consider two DNA molecules of length L = 998 bp, with and without a single 601 segment. Nucleosomes can bind to the DNA with an effective energy E_{ns} (in units of k_BT), if the association happens to the nonspecific region, or with an effective energy E_{601}, if the association happens to the 601 sequence. Each association covers l = 147 bp of the DNA length. In addition, two DNA-bound nucleosomes that are located at distances less than d = 30 bp are assumed to have internucleosomal interactions $E_{int.}$ This cutoff distance is chosen based on the experimental data for dinucleosomes in Figure 3 that show the peak in the distribution of internucleosomal distances as being 28 ± 2 bp (SEM).

Experimental data showed that there are a total of 12 observable chemical states in the system: one state when DNA is totally free, one state with a single bound nucleosome, two states with two bound nucleosomes (with and without internucleosomal interaction), three states with three bound

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813–10821

nucleosomes, and five states with four bound nucleosomes. Because of long times of observation during AFM experiments, it is reasonable to assume that the system reaches an overall chemical equilibrium. This allows us to significantly simplify the analysis to only a few states and transitions between them.

Following this strategy, we first consider the transitions between the free DNA (state 0) and mononucleosome (state 1) for nonspecific DNA. We define the equilibrium probabilities of these states as P_0 and P_1, respectively. The condition of chemical equilibrium then requires that

$$\frac{P_1}{P_0} = e^{-E_{ns}} \tag{1}$$

Using the data from Table 2, one can easily determine that the effective free energy difference is $E_{ns} = 0.3\ k_BT$. Extrapolating based on this, one can obtain the energy of the mononucleosome state, $E_{ns}(L - l)/l = 1.74\ kT$. This means that interactions between nucleosomes and nonspecific segments of DNA are relatively weak, allowing nucleosomes to be quite mobile.

The equilibrium between free DNA and mononucleosomes for the system with the 601 sequence is more involved because the nucleosome can associate with this 601 sequence (we call it a state 1*) or it can associate with any other nonspecific sequence (we call it a state 1). The probabilities, at chemical equilibrium, between these states can be written as

$$\frac{P_1 + P_{1*}}{P_0} = \frac{L - l}{L - l + 1} e^{-E_{ns}} + \frac{1}{L - l + 1} e^{-E_{601}} \tag{2}$$

This equation reflects the fact that only one site of available $L - l + 1$ binding sites is a special one, and $L - l/(L - l + 1)$ of them are nonspecific bindings. Using the already obtained value for $E_{ns} = 0.3\ k_BT$ and the parameters for L and l, we determine that $E_{601} = 8.3\ k_BT$. This shows that the interactions between nucleosomes and the special 601 positioning sequence are very strong, preventing the mobility of the bound nucleosome in this case. Furthermore, the estimate is in good agreement with experimental results published by other groups.[39]

We then explore the interparticle interactions for bound nucleosomes. We consider the equilibrium distributions of two different states for dinucleosomes on nonspecific DNA. To simplify the analysis, let us neglect the end effect of the finite length of the DNA molecule. Then, if one nucleosome is already bound, there are $L - 3l + 2$ sites where the second nucleosome might bind. But $2d$ of them will exhibit an additional internucleosomal interaction energy. This leads to the fraction of associations that lead to state 2 (two bound, interacting nucleosomes) to be

$$f_{int} = \frac{2d}{L - 3l + 2} = \frac{60}{559} \approx 0.11 \tag{3}$$

Then, the fraction of binding events that leads to state 1−1 (two bound, noninteracting nucleosomes) is $1 - f_{int} = 0.89$. The equilibrium distribution of states for dinucleosomes then gives rise to the following:

$$\frac{P_2}{P_2 + P_{1-1}} = \frac{f_{int} e^{-E_{int}}}{f_{int} e^{-E_{int}} + (1 - f_{int})} \tag{4}$$

Using data from Table 2, we estimate that $E_{int} \sim 1.5\ k_BT$, or 0.9 kcal/mol, which is very close to internucleosomal interaction energy obtained in other experimental studies.[40]

The robustness of our analysis can be tested by applying it to evaluate the fraction of states for the dinucleosomes in the case of the special 601 positioning sequence, which was measured independently in our experiments. Because of strong interactions with the 601 sequence, it is reasonable to assume that one nucleosome is always bound to this special site. The second nucleosome can associate to one of the $L - 2l - l_0 + 1 = 592$ sites. $l_0 = 113$ bp is the distance from the 601 segment to one end of DNA, which also imposes that the second nucleosome can only bind to one side of the already bound nucleosome. This gives the following fraction of binding sites that leads to state 2 (two interacting nucleosomes)

$$f_{int} = \frac{d}{L - 2l - l_0 + 1} = \frac{30}{592} \approx 0.05 \tag{5}$$

Substituting this result into eq 4 predicts that ∼18.8% of dinucleosomes formed on the substrate with the special 601 positioning sequence would correspond to state 2, which is comparable to the experimentally measured value of 21.9% for state 2.

DISCUSSION

We used AFM to characterize the role of DNA sequence and internucleosomal interactions in the assembly of the nucleosome array. We assembled oligonucleosomes on two DNA substrates, the first comprising 998 bp of DNA with a single nucleosome positioning motif located near the end of the DNA (601 DNA) and the second of the same length, but with no regions of high nucleosome specificity (nonspecific DNA), which mimics natural DNA. The most striking feature of the AFM images of the arrays assembled on these DNA templates is the ability of nucleosomes to assemble in clusters with close proximity of neighboring nucleosomes. Such clusters with two, three, and four nucleosomes comprise more than 90% of tetranucleosome arrays, whereas species with distant locations of nucleosomes comprise only 6.7% on 601 DNA (Table 1) and 6.8% on nonspecific DNA (Table 2). A similar effect of the nucleosome clustering is observed for the trinucleosome assemblies, although a well-separated 1−1−1 arrangement for the trinucleosome is more likely, 30.6% and 34.6% for 601 DNA and nonspecific DNA, respectively. This finding is in a sharp contrast with AFM studies of arrays assembled on repeats of nucleosome-positioning sequences in which arrays of well-separated nucleosomes are the predominant features.[36−38,41]

To quantitatively characterize the internucleosomal interactions, we measured internucleosomal distances for dinucleosome and 2−1 trinucleosome assemblies, for which this parameter can be measured using AFM images. Quantitative analysis for both assemblies shown in Figures 3 and S2 reveals a number of features of the nucleosome assemblies. These data graphically demonstrate that nucleosomes can be tightly clustered in dinucleosome assemblies or separated from each other in a broad range. Importantly, the nucleosomes in the dinucleosome populations, separated at a peak distance as small as 30 bp, showed compact assemblies in which 24% of species display an internucleosomal distance of less than 10 bp. Overall, the distribution is narrow, with a standard deviation of 13 bp. Monte Carlo simulations failed to identify such frequent

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813−10821

and close contacts when nucleosomes were allowed to position freely along the substrates (Tables S1, S2). These data point directly to the role of internucleosomal interactions. According to numerous publications (ref 42 and references therein), histone tails play a critical role in the formation of tight internucleosomal contacts, as post-translational modifications such as acetylation can decrease the nucleosome clustering, including the assembly of dinucleosomes studied with AFM.[37]

According to the molecular modeling of nucleosome arrays,[42] the density of the nucleosome packing depends on the orientation of the nucleosomes in which the tightest distance (less than 1 nm between the nucleosomes) corresponds to stacking of the disk-type particles. The distance can be larger for other orientations of nucleosomes, and the variation of the entry−exit angles between the adjacent nucleosomes, defined by the size of linker DNA between the nucleosome, contributes to the orientation of the nucleosome and hence the internucleosomal distance. We assume that the orientation factor can explain the breadth of the internucleosomal distance distribution in our experiments. Note, however, that histone tails not only bridge the nucleosomes but also can hinder the formation of tight internucleosomal contacts. We have recently shown that truncation of histone H4 can shift the internucleosomal distance for the dinucleosome constructs to smaller distances.[43]

Clustered dinucleosomes were visualized before,[36,38,41] although these were minor species due to the use of DNA templates with positioning sequence repeats, so the nucleosomes primarily occupied the nucleosome-specific motifs. The assembly of dinucleosomes in refs 36 and 38 was explained by a relatively low energetic preference for the assembly of nucleosomes on favored motifs compared with nonfavored ones; however, three-nucleosome clusters have not been observed in those papers, nor four-nucleosome clusters, which in our case were observed with high yield. Our theoretical model and estimates made with a comparison with experiments provide the following explanation for the nucleosome clustering. The affinity of a nucleosome to a nonspecific DNA sequence is as low as 0.3 k_BT. Given that the internucleosomal attraction is relatively high, $E_{int} \sim 1.5\ k_BT$, or 0.9 kcal/mol, as per our calculations above and previous publications,[40] the assembly of clusters is a favorable process. Studies[44−46] have shown that internucleosomal attractions promote the compaction over a large range of scales on the repeated 601 substrates. However, affinity of nucleosomes to positioning sequences can prevent the assembly of nucleosome clusters. It has also been discussed[47] that the relaxed DNA−(H2A−H2B) region may contribute to long-range internucleosomal interactions between DNA and histone tails. Therefore, the decrease in the binding between DNA and histone octamers on the nonspecific substrate compared to that on the 601 sequence benefits the long-range internucleosomal interactions and, ultimately, the compaction of the oligonucleosomes. Importantly, our theoretical analysis explains that differences in distributions of states in dinucleosomes are due to different interaction strengths between the nucleosomes and nonspecific and 601 positioning DNA segments. Thus, the morphology of a nucleosome array is defined by the DNA sequence, so one can expect clusters along with nonclustered segments in the array. However, natural nucleosomal DNA contains only modest sequence periodicity,[38,48] so clustering of nucleosomes should be the most representative morphology of the chromatin. AFM images of trinucleosome and tetranucleosome clusters did not reveal geometric features of a periodic solenoid model of chromatin. The nonregular morphology of tri- and tetra-nucleosome arrays observed in this study better fits the model of irregular chromatin structure.[16] Importantly, the assembly of nucleosomes in clusters can be an important factor in the gene regulation. Promoter regions in eukaryotes do not show a high affinity to nucleosome assembly[49−51] allowing for the nucleosome clustering. Such clusters would create a hurdle for transcription factors to bind regulatory regions as well as for RNA polymerase to transcribe genes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.2c05321.

Scheme for measuring distances in base pairs; dinucleosome and trinucleosome analysis; AFM images of nucleosomes assembled on nonspecific DNA; comparisons of experimental tetranucleosome data to data from simulations of tetranucleosomes; statistical comparison of tetranucleosome populations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Yuri L. Lyubchenko − *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, Nebraska 68198, United States;* orcid.org/0000-0001-9721-8302; Phone: 402-559-1971; Email: ylyubchenko@unmc.edu

Authors

Yaqing Wang − *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, Nebraska 68198, United States; Materials Science Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550, United States*

Tommy Stormberg − *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, Nebraska 68198, United States*

Mohtadin Hashemi − *Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, Nebraska 68198, United States;* orcid.org/0000-0003-2698-9761

Anatoly B. Kolomeisky − *Department of Chemistry and Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States;* orcid.org/0000-0001-5677-6690

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.2c05321

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. Y.W. and T.S. contributed equally to this work.

Funding

The work was supported by grants to Y.L.L. from NSF (MCB 1515346) and NIH (GM096039, GM100156).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was completed utilizing the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative. The authors thank the Lyubchenko lab members for useful insights in the data interpretation.

■ REFERENCES

(1) Van Holde, K.; Zlatanova, J. Chromatin Higher Order Structure: Chasing a Mirage? *J. Biol. Chem.* **1995**, *270* (15), 8373−8376.
(2) Kornberg, R. D.; Lorch, Y. Twenty-Five Years of the Nucleosome, Fundamental Particle of the Eukaryote Chromosome. *Cell* **1999**, *98* (3), 285−294.
(3) Zhou, K.; Gaullier, G.; Luger, K. Nucleosome Structure and Dynamics Are Coming of Age. *Nature Structural and Molecular Biology* **2019**, *26*, 3−13.
(4) Kornberg, R. D. Chromatin Structure: A Repeating Unit of Histones and DNA. *Science* **1974**, *184* (4139), 868−871.
(5) Armeev, G. A.; Gribkova, A. K.; Pospelova, I.; Komarova, G. A.; Shaytan, A. K. Linking Chromatin Composition and Structural Dynamics at the Nucleosome Level. *Curr. Opin Struct Biol.* **2019**, *56*, 46−55.
(6) Wilson, M. D.; Costa, A. Cryo-Electron Microscopy of Chromatin Biology. *Acta Crystallographica Section D: Structural Biology* **2017**, *73*, 541−548.
(7) McGhee, J. D.; Felsenfeld, G. Nucleosome Structure. *Annu. Rev. Biochem.* **1980**, *49*, 1115−1156.
(8) Kornberg, R. D. Structure of Chromatin. *Annu. Rev. Biochem.* **1977**, *46*, 931−954.
(9) Elgin, S. C.; Weintraub, H. Chromosomal Proteins and Chromatin Structure. *Annu. Rev. Biochem.* **1975**, *44*, 725−774.
(10) Isenberg, I. Histones. *Annu. Rev. Biochem.* **1979**, *48*, 159−191.
(11) Luger, K.; Mäder, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. *Nature* **1997**, *389* (6648), 251−260.
(12) Tremethick, D. J. Higher-Order Structures of Chromatin: The Elusive 30 Nm Fiber. *Cell* **2007**, *128* (4), 651−654.
(13) Schalch, T.; Duda, S.; Sargent, D. F.; Richmond, T. J. X-Ray Structure of a Tetranucleosome and Its Implications for the Chromatin Fibre. *Nature* **2005**, *436* (7047), 138−141.
(14) Chien, F.-T.; van Noort, J. 10 Years of Tension on Chromatin: Results from Single Molecule Force Spectroscopy. *Curr. Pharm. Biotechnol* **2009**, *10* (5), 474−485.
(15) Robinson, P. J. J.; Fairall, L.; Huynh, V. A. T.; Rhodes, D. EM Measurements Define the Dimensions of the "30-Nm" Chromatin Fiber: Evidence for a Compact, Interdigitated Structure. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (17), 6506−6511.
(16) Ou, H. D.; Phan, S.; Deerinck, T. J.; Thor, A.; Ellisman, M. H.; O'Shea, C. C. ChromEMT: Visualizing 3D Chromatin Structure and Compaction in Interphase and Mitotic Cells. *Science* **2017**, *357* (6349), No. eeag0025.
(17) Maeshima, K.; Ide, S.; Babokhov, M. Dynamic Chromatin Organization without the 30-Nm Fiber. *Curr. Opin Cell Biol.* **2019**, *58*, 95−104.
(18) Ohno, M.; Ando, T.; Priest, D. G.; Kumar, V.; Yoshida, Y.; Taniguchi, Y. Sub-Nucleosomal Genome Structure Reveals Distinct Nucleosome Folding Motifs. *Cell* **2019**, *176* (3), 520−534.
(19) Maeshima, K.; Hihara, S.; Eltsov, M. Chromatin Structure: Does the 30-Nm Fibre Exist in Vivo? *Curr. Opin Cell Biol.* **2010**, *22* (3), 291−297.
(20) Luger, K.; Hansen, J. C. Nucleosome and Chromatin Fiber Dynamics. *Curr. Opin Struct Biol.* **2005**, *15* (2), 188−196.
(21) Stormberg, T.; Stumme-Diers, M.; Lyubchenko, Y. L. Sequence-Dependent Nucleosome Nanoscale Structure Characterized by Atomic Force Microscopy. *FASEB J.* **2019**, *33*, 10916−10923.
(22) Filenko, N. A.; Palets, D. B.; Lyubchenko, Y. L. Structure and Dynamics of Dinucleosomes Assessed by Atomic Force Microscopy. *J. Amino Acids* **2012**, *2012*, 1−6.
(23) Boopathi, R.; Dimitrov, S.; Hamiche, A.; Petosa, C.; Bednar, J. Cryo-Electron Microscopy of the Chromatin Fiber. *Curr. Opin Struct Biol.* **2020**, *64*, 97−103.
(24) Sun, Z.; Stormberg, T.; Filliaux, S.; Lyubchenko, Y. L. Three-Way DNA Junction as an End Label for DNA in Atomic Force Microscopy Studies. *Int. J. Mol. Sci.* **2022**, *23* (19), 11404.
(25) Stormberg, T.; Stumme-Diers, M.; Lyubchenko, Y. L. Sequence-dependent Nucleosome Nanoscale Structure Characterized by Atomic Force Microscopy. *FASEB J.* **2019**, *33* (10), 10916−10923.
(26) Miyagi, A.; Ando, T.; Lyubchenko, Y. L. Dynamics of Nucleosomes Assessed with Time-Lapse High-Speed Atomic Force Microscopy. *Biochemistry* **2011**, *50* (37), 7901−7908.
(27) Lyubchenko, Y. L.; Shlyakhtenko, L. S. Imaging of DNA and Protein−DNA Complexes with Atomic Force Microscopy. *Critical Reviews in Eukaryotic Gene Expression.* **2016**, *26* (1), 63−96.
(28) Pan, Y.; Zagorski, K.; Shlyakhtenko, L. S.; Lyubchenko, Y. L. The Enzymatic Activity of APOBE3G Multimers. *Sci. Rep* **2018**, *8* (1), 17953.
(29) Wang, Y.; Sun, Z.; Bianco, P. R.; Lyubchenko, Y. L. Atomic Force Microscopy-Based Characterization of the Interaction of PriA Helicase with Stalled DNA Replication Forks. *J. Biol. Chem.* **2020**, *295* (18), 6043−6052.
(30) Stumme-Diers, M. P.; Stormberg, T.; Sun, Z.; Lyubchenko, Y. L. Probing the Structure and Dynamics of Nucleosomes Using Atomic Force Microscopy Imaging. *J. Visualized Exp.* **2019**, (143),e58820.
(31) Stormberg, T.; Filliaux, S.; Baughman, H. E. R.; Komives, E. A.; Lyubchenko, Y. L. Transcription Factor NF-KB Unravels Nucleosomes. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2021**, *1865* (9), 129934.
(32) Stumme-Diers, M. P.; Banerjee, S.; Hashemi, M.; Sun, Z.; Lyubchenko, Y. L. Nanoscale Dynamics of Centromere Nucleosomes and the Critical Roles of CENP-A. *Nucleic Acids Res.* **2018**, *46* (1), 94−103.
(33) Lyubchenko, Y. L.; Shlyakhtenko, L. S. Imaging of DNA and Protein-DNA Complexes with Atomic Force Microscopy. *Critical Reviews in Eukaryotic Gene Expression* **2016**, *26* (1), 63−96.
(34) Quinlan, A. R.; Hall, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *BIOINFORMATICS APPLICATIONS NOTE* **2010**, *26* (6), 841−842.
(35) Lowary, P. T.; Widom, J. New DNA Sequence Rules for High Affinity Binding to Histone Octamer and Sequence-Directed Nucleosome Positioning. *J. Mol. Biol.* **1998**, *276* (1), 19−42.
(36) Yodh, J. G.; Lyubchenko, Y. L.; Shlyakhtenko, L. S.; Woodbury, N.; Lohr, D. Evidence for Nonrandom Behavior in 208−12 Subsaturated Nucleosomal Array Populations Analyzed by AFM. *Biochemistry* **1999**, *38* (48), 15756−15763.
(37) Bash, R. C.; Yodh, J.; Lyubchenko, Y.; Woodbury, N.; Lohr, D. Population Analysis of Subsaturated 172−12 Nucleosomal Arrays by Atomic Force Microscopy Detects Nonrandom Behavior That Is Favored by Histone Acetylation and Short Repeat Length. *J. Biol. Chem.* **2001**, *276* (51), 48362−48370.
(38) Yodh, J. G.; Woodbury, N.; Shlyakhtenko, L. S.; Lyubchenko, Y. L.; Lohr, D. Mapping Nucleosome Locations on the 208−12 by AFM Provides Clear Evidence for Cooperativity in Array Occupation. *Biochemistry* **2002**, *41*, 3565−3574.
(39) Van Der Heijden, T.; Van Vugt, J. J. F. A.; Logie, C.; Van Noort, J. Sequence-Based Prediction of Single Nucleosome Positioning and Genome-Wide Nucleosome Occupancy. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (38), No. E2514.
(40) Widlund, H. R.; Vitolo, J. M.; Thiriet, C.; Hayes, J. J. DNA Sequence-Dependent Contributions of Core Histone Tails to Nucleosome Stability: Differential Effects of Acetylation and Proteolytic Tail Removal†. *Biochemistry* **2000**, *39* (13), 3835−3841.
(41) Bussiek, M.; Müller, G.; Waldeck, W.; Diekmann, S.; Langowski, J. Organisation of Nucleosomal Arrays Reconstituted with Repetitive African Green Monkey α-Satellite DNA as Analysed by Atomic Force Microscopy. *Eur. Biophys. J.* **2007**, *37* (1), 81−93.

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813−10821

(42) Lyubartsev, A. P.; Korolev, N.; Fan, Y.; Nordenskiöld, L. Multiscale Modelling of Nucleosome Core Particle Aggregation. *J. Phys.: Condens. Matter* **2015**, *27* (6), 064111.
(43) Stormberg, T.; Vemulapalli, S.; Filliaux, S.; Lyubchenko, Y. L. Effect of Histone H4 Tail on Nucleosome Stability and Internucleosomal Interactions. *Sci. Rep.* **2021**, *11* (1), 1−11.
(44) Stehr, R.; Kepper, N.; Rippe, K.; Wedemann, G. The Effect of Internucleosomal Interaction on Folding of the Chromatin Fiber. *Biophys. J.* **2008**, *95* (8), 3677−3691.
(45) Kulaeva, O. I.; Zheng, G.; Polikanov, Y. S.; Colasanti, A. v.; Clauvelin, N.; Mukhopadhyay, S.; Sengupta, A. M.; Studitsky, V. M.; Olson, W. K. Internucleosomal Interactions Mediated by Histone Tails Allow Distant Communication in Chromatin. *J. Biol. Chem.* **2012**, *287* (24), 20248−20257.
(46) Montel, F.; Menoni, H.; Castelnovo, M.; Bednar, J.; Dimitrov, S.; Angelov, D.; Faivre-Moskalenko, C. The Dynamics of Individual Nucleosomes Controls the Chromatin Condensation Pathway: Direct Atomic Force Microscopy Visualization of Variant Chromatin. *Biophys. J.* **2009**, *97* (2), 544−553.
(47) Sengupta, B.; Huynh, M.; Smith, C. B.; McGinty, R. K.; Krajewski, W.; Lee, T.-H. The Effects of Histone H2B Ubiquitylations on the Nucleosome Structure and Internucleosomal Interactions. *Biochemistry* **2022**, *61* (20), 2198−2205.
(48) Shrader, T. E.; Crothers, D. M. Artificial Nucleosome Positioning Sequences. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (19), 7418.
(49) Lee, W.; Tillo, D.; Bray, N.; Morse, R. H.; Davis, R. W.; Hughes, T. R.; Nislow, C. A High-Resolution Atlas of Nucleosome Occupancy in Yeast. *Nat. Genet.* **2007**, *39* (10), 1235−1244.
(50) Mavrich, T. N.; Ioshikhes, I. P.; Venters, B. J.; Jiang, C.; Tomsho, L. P.; Qi, J.; Schuster, S. C.; Albert, I.; Pugh, B. F. A Barrier Nucleosome Model for Statistical Positioning of Nucleosomes throughout the Yeast Genome. *Genome Res.* **2008**, *18* (7), 1073.
(51) Ozsolak, F.; Song, J. S.; Liu, X. S.; Fisher, D. E. High-Throughput Mapping of the Chromatin Structure of Human Promoters. *Nat. Biotechnol.* **2007**, *25* (2), 244−248.

https://doi.org/10.1021/acs.jpcb.2c05321
J. Phys. Chem. B 2022, 126, 10813−10821