Histone H3 and H4 N-Terminal Tails in Nucleosome Arrays at Cellular Concentrations Probed by Magic Angle Spinning NMR Spectroscopy

Min Gao, † Philippe S. Nadaud, † Morgan W. Bernier, ‡ Justin A. North, ‡ P. Chris Hammel, ‡ Michael G. Poirier, *‡ and Christopher P. Jaroniec*†

†Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States
‡Department of Physics, The Ohio State University, Columbus, Ohio 43210, United States

Supporting Information

ABSTRACT: Chromatin is a supramolecular assembly of DNA and histone proteins, organized into nucleosome repeat units. The dynamics of chromatin organization regulates DNA accessibility to eukaryotic transcription and DNA repair complexes. Yet, the structural and dynamic properties of chromatin at high concentrations characteristic of the cellular environment (≥200 mg/mL) are largely unexplored at the molecular level. Here, we apply MAS NMR to directly probe the dynamic histone protein regions in 13C,15N-enriched recombinant nucleosome arrays at cellular chromatin concentrations and conditions designed to emulate distinct states of DNA condensation, with focus on the flexible H3 and H4 N-terminal tails which mediate chromatin compaction. 2D 1H−13C and 1H−15N spectra reveal numerous correlations for H3 and H4 backbone and side-chain atoms, enabling identification of specific residues making up the dynamically disordered N-terminal tail domains. Remarkably, we find that both the H3 and H4 N-terminal tails are overall dynamic even in a highly condensed state. This significant conformational flexibility of the histone tails suggests that they remain available for protein binding in compact chromatin states to enable regulation of heterochromatin. Furthermore, our study provides a foundation for quantitative structural and dynamic investigations of chromatin at physiological concentrations.

Eukaryotic DNA is dynamically organized into chromatin fibers, which regulate essential functions of the genome including transcription and DNA repair.1,2 The basic building block of chromatin is the nucleosome core particle, which contains ~146 base pairs (bp) of DNA wrapped 1.65 times around a histone protein octamer containing two copies each of histones H2A, H2B, H3, and H4.3 The nucleosome X-ray structure has been determined to near atomic resolution4 and reveals a compact helical core with ~15−30% of the histone sequences protruding from the core as largely unstructured, and presumably flexible, N-terminal tail domains (Figure 1A). The crystal structure of a tetrnucleosome has also been solved,5 but the low (9 Å) resolution of this structure precludes the definition of N-terminal histone tail conformations.

In human cells, single chromatin fibers are chains of ~100 000 nucleosomes located in the nucleus at extremely high concentrations of ≥200 mg/mL,6 with each fiber organized into distinct chromosome territories.7 In vitro, longer (>10−12-mer) nucleosome arrays form a variety of higher order structures in the presence of Mg2+, ranging from an extended beads-on-a-string type “10-nm” fiber in the absence of Mg2+ to a folded “30-nm” fiber at an intermediate (~1 mM) Mg2+ concentration to highly condensed aggregates at high Mg2+ concentrations.8,9 The 30-nm chromatin fiber has been observed in a few distinct cell types.10–11 Interestingly, however, recent small-angle X-ray scattering studies indicate that the 30-nm fiber is not the dominant structural form of chromatin in mitotic chromosomes,12 suggesting that the high cellular concentrations of chromatin may impact its higher order structure.

It is well-established that the positively charged N-terminal tails of histones H3 and H4 mediate the compaction of chromatin into 30 nm diameter fibers and interfiber condensation in vitro, apparently through interactions with DNA and/or acidic regions on the histone octamer surface of neighboring nucleosomes.13–15 Remarkably, peptides with sequences corresponding to the H4 N-terminal tail are also able to mediate self-association of H4 tail-less nucleosome arrays,16 suggesting that charge neutralization plays an important role in chromatin compaction. The mobile N-terminal histone tails have been previously probed by solution nuclear magnetic resonance (NMR) in ~20 mg/mL samples of single nucleosomes17–19 and oligonucleosomes20 in the absence of Mg2+. Of particular note is the recent elegant study of mononucleosomes containing 13C,15N-enriched histones by Bai et al.,19 which found that flexible H3 and H4 domains encompass residues 1−36 and 1−15, respectively, based on sets of backbone amide signals detected in 15N−1H heteronuclear single quantum coherence spectra. On the other hand, the structure and dynamics of H3 and H4 tails in compacted chromatin fibers remain largely unknown, with the analysis of these domains to date being limited to modeling12–23 and relatively indirect biochemical and biophysical approaches such as mutagenesis combined with analytical centrifugation or chemical cross-linking14,24 and hydrogen–deuterium (H/D) exchange coupled with solution NMR.25 Interestingly, the recent H/D exchange NMR study of 12-mer nucleosome arrays25 concluded that the H3 tail forms stable folded structures in highly condensed chromatin fibers. This is in contrast with the observation of a dynamically disordered H3 N-terminus in soluble single
Communication

Communication

Figure 1. (A) Nucleosome crystal structure (PDB entry 1KX5). Histones H3 and H4 are colored red and blue, respectively, with selected residues located near the N-terminal tail boundaries highlighted. (B) Representative AFM images of 17-mer nucleosome arrays used for the NMR studies. (C) Amino acid sequences of histones H3 and H4. N-terminal residues that are relatively unstructured in nucleosome core particle crystals are bold. Residues determined in the present study to comprise the flexible N-terminal H3 and H4 tails in highly concentrated 17-mer nucleosome arrays irrespective of the degree of array compaction are underlined with solid lines. Additional residues that possibly belong to the dynamically disordered histone tail domains and cannot be unambiguously identified from the NMR data are underlined with dashed lines. (D,E) 2D $^1H-^{13}C$ (D) and $^1H-^{15}N$ (E) correlation spectra of 17-mer nucleosome arrays containing $^{13}C,^{15}N$-enriched H3 and prepared with 0, 1, and 5 mM Mg$^{2+}$, corresponding to the extended, folded, and aggregated chromatin conformation, respectively, as indicated. The spectra were recorded at 500 MHz $^1H$ frequency, 11.111 kHz MAS rate, and 30 °C using a 2D refocused INEPT pulse scheme described in detail previously. Each spectrum was recorded with acquisition times of 18 and 30 ms in the $^1H$ and $^{13}C/^{15}N$ dimensions, respectively, and a total measurement time of $\sim 48$ h. Resonance assignments based on the average chemical shift values in the BioMagResBank database corresponding to residues located in the flexible N-terminal histone tail are labeled in black font. Indicated in gray font are the approximate locations of signals from unique residues bordering the flexible tail that would be observed if those residues were sufficiently mobile. Note that several narrow signals visible at a $^{13}C$ frequency of $\sim 65-70$ ppm, particularly for the 0 mM Mg$^{2+}$ sample, arise from residual sucrose from the sucrose gradient purification procedure. (F,G) Same as panels (D,E) but for 17-mer nucleosome arrays containing $^{13}C,^{15}N$-labeled H4.

nucleosomes$^{19}$ and suggests possible differences in the conformational flexibilities of histone N-terminal domains between uncompacted and compacted chromatin.

Here, we apply 1D and 2D magic-angle spinning (MAS) NMR to probe histones H3 and H4 within isotopically $^{13}C,^{15}N$-enriched recombinant 17-mer nucleosome arrays at very high densities of $\sim 200-400$ mg/mL typical of the cellular environment. The arrays were incubated with 0, 1, or 5 mM Mg$^{2+}$, conditions that respectively yield extended chromatin, compact 30 nm chromatin fibers, and a highly condensed chromatin state with self-association between multiple fibers. The major advantage of the MAS NMR methodology in the present context is that it enables the amino acid residues comprising the rigid and flexible histone protein segments to be monitored directly under physiologically relevant conditions, in a manner similar to that previously demonstrated for a variety of large macromolecular complexes including membrane and fibrillar protein assemblies.$^{16-28}$

The nucleosome arrays used in this study were reconstituted from a 3046 bp DNA template containing 17 tandem repeats of a Widom 601 nucleosome positioning sequence variant with 30 bp of linker DNA$^{29}$ (Figure S1) and histone octamer containing $^{13}C,^{15}N$-labeled H3 or H4 (see Supporting Information (SI) for details). Sucrose gradient purified 17-mer arrays were analyzed by composite gel electrophoresis, BamHI and AvaI restriction enzyme digestions, and atomic force microscopy (AFM) (Figure 1B) to assess their level of saturation with a histone octamer. These analyses (Figures S2–S3 and accompanying SI text) show that the array samples are highly homogeneous, effectively (>95%) saturated with 17 histone octamers, and stable during the course of the MAS NMR experiments.

As noted above, different levels of chromatin condensation were induced by increasing amounts of Mg$^{2+}$. The extent of self-association of the different array samples was readily evident in their physical appearance and was quantitatively confirmed by using a UV absorbance based sedimentation assay$^{30}$ (Figure S4). For the NMR measurements the array samples were concentrated to $\sim 200-400$ mg/mL by ultracentrifugation, which is typical of chromatin densities found in cells.

The presence of flexible and rigid segments of histones H3 and H4 in the context of the highly concentrated 17-mer nucleosome arrays was first assessed using a set of 1D $^{15}C$ refocused INEPT and cross-polarization MAS NMR spectra recorded as a function of temperature between -20 and 30 °C and described in the SI (Figures S5–S7). The key result from this set of experiments is that, at temperatures of ca. -10 °C and above, both H3 and H4 contain highly mobile domains spanning ~25% of the protein sequences, irrespective of the degree of array compaction. Next, we performed a more detailed analysis of the flexible H3 and H4 domains by 2D MAS NMR methods. Figure 1 shows a series of $^1H—^{13}C$ and $^1H—^{15}N$ correlation spectra recorded at 30 °C using a 2D refocused INEPT pulse scheme$^{6,27}$ as a function of
increasing Mg\(^{2+}\) concentration for arrays containing \(^{15}\)C,\(^{15}\)N–H3 (panels D and E) or H4 (panels F and G). As expected based on the 1D INEPT \(^{13}\)C spectra, the 2D \(^{1}H\)–\(^{13}\)C and \(^{1}H\)–\(^{15}\)N data sets display multiple intense resonances for both histones H3 and H4. A qualitative inspection of the \(^{1}H\)–\(^{15}\)N spectra reveals that all detectable backbone amide protons resonate in a narrow range between 8.0 and 8.5 ppm, characteristic of unstructured proteins. \(^{31}\) Indeed, by assuming that the flexible H3 and H4 tails are located at the N-termini and behave as ensembles of random-coil-like states we have been able to account for the vast majority of the observed protein signals by mapping onto the spectra the average residue- and site-specific chemical shifts obtained from the BioMagResBank database. While the spectral sensitivity and resolution achievable for these array samples were not sufficient to establish sequence specific resonance assignments de novo, analysis of the 2D NMR spectra in conjunction with the H3 and H4 amino acid sequences nevertheless permitted a relatively precise identification of the residues encompassing the flexible H3 and H4 N-terminal domains in the 17-mer arrays as follows. For both histones, resonances characteristic of the various amino acids were accounted for starting from the N-terminus, until signals from one or more amino acids in the sequence that should be clearly discernible in the spectra if sufficiently mobile were no longer detectable. It is important to note that while identification of amino acid types was fairly straightforward, the number of residues contributing to the individual resonances could not be readily determined due to the complicating effect of local protein dynamics on signal intensities.

For H3 a number of unique resonances were observed corresponding to residues in the range 1–38, including for example Ga, A\(\gamma\), S\(\beta\), T\(\beta\), V\(\gamma\), L\(\delta\), and K\(\varepsilon\) signals in the \(^{1}H\)–\(^{13}\)C spectra as well as Q\(\gamma\), R\(\varepsilon\), and R\(\eta\) signals in the \(^{1}H\)–\(^{15}\)N spectra. Particularly remarkable is the observation of the V\(\gamma\) signals. Combined with the fact that V35 is the first valine encountered in the sequence, this finding indicates that the flexible H3 tail extends to at least position 35. Equally crucial is the fact that no side-chain \(^{1}H\)–\(^{13}\)C signals characteristic of His or Tyr were observed, which indicates that H39 and Y41 are relatively immobile and defines the H3 tail boundary. In the absence of sequential resonance assignments it is impossible to unambiguously establish from these data whether residues K36–P38 are sufficiently mobile to be detectable, because multiple other residues of the same type that yield observable signals are located between positions 1 and 35. It is also unlikely that any of the H3 signals originate from protein regions other than the flexible N-terminal tail—in particular, it is improbable that the V\(\gamma\) signal could be attributed to any of the other five valines located in the core domain of H3. The reason is that, in addition to the absence of His and Tyr resonances, we also do not observe any characteristic Ile side-chain signals (there are 7 isoleucines in H3 located between positions S1 and 130). Altogether, we conclude that flexible N-terminal H3 tails in concentrated 17-mer nucleosome arrays span amino acids 1–35 and possibly several additional residues in the 36–38 regime (Figure 1C), while the remaining H3 residues (certainly 39–130) are largely immobile.

Analogous analysis of the H4 spectra identified unique signals arising from multiple residues in the range 1–21, including for example Ga, A\(\gamma\), V\(\gamma\), L\(\delta\), H\(\delta\), H\(\varepsilon\), and K\(\varepsilon\) signals in the \(^{1}H\)–\(^{13}\)C spectra as well as R\(\varepsilon\) and R\(\eta\) signals in the \(^{1}H\)–\(^{15}\)N spectra. Key is the observation of the A\(\gamma\), V\(\gamma\), H\(\delta\), and H\(\varepsilon\) signals because A15, H18, and V21 are the first Ala, His, and Val residues in the H4 sequence. Also critical is the finding that characteristic side-chain signals from Asp, Asn, Ile, Glu, and Thr, located in the range 24–30, are missing, which clearly defines the flexible tail boundary. Given that no Tyr and Phe resonances are detected either (there are two Phe and four Tyr residues in H4, with Y98 and F100 located near the C-terminus), we conclude that flexible N-terminal H4 tails in the 17-mer arrays span residues 1–21 while most of the remaining residues are relatively rigid (Figure 1C) (note: H4 residues L22 and R23 may also be somewhat flexible, but this cannot be unambiguously established due to the other leucines and arginines in the 1–21 segment).

In addition to the identification of specific amino acids making up the flexible N-terminal domains of histones H3 and H4 in concentrated chromatin, a central finding of our study is that by and large the same residues remain dynamic in uncompacted, folded, and highly condensed nucleosome arrays as judged by the similarity of the 2D \(^{1}H\)–\(^{13}\)C and \(^{1}H\)–\(^{15}\)N NMR spectra in Figure 1. To compare the extent of protein dynamics between the different array samples, we performed \(^{13}\)N INEPT based NMR experiments to monitor the transverse relaxation rates of the amide protons exclusively for residues encompassing the flexible histone tails (Figure S8). Although rather qualitative, these studies reveal that collectively the amide \(^{1}H\) coherences relax somewhat more rapidly for arrays incubated with 1 mM and 5 mM Mg\(^{2+}\) relative to those prepared in the absence of Mg\(^{2+}\), particularly for H3. This finding is suggestive of a subtle overall reduction in the flexibility of the histone tail domains in the most highly condensed and concentrated arrays, which would lead to less efficient motional averaging of the \(^{1}H\)–\(^{1}H\) dipolar couplings and consequently faster \(^{1}H\) coherence decay. We also recorded complementary electron paramagnetic resonance (EPR) spectra of the 17-mer arrays containing nitroxide spin labeled analogs of histone H3 at two locations including the N-terminus and residue 35 near the flexible tail boundary. The EPR spectra (Figure S9) display nearly identical, relatively narrow line shapes for all Mg\(^{2+}\) concentrations indicating significant flexibility of the H3 N-terminal tail irrespective of the degree of compaction, and, importantly, no detectable broad components characteristic of immobilized protein segments. The moderate line broadening observed for position 35 relative to the N-terminus is consistent with the expected reduction in mobility of the nitroxide spin probe in the vicinity of the nucleosomal DNA.

In summary, we have demonstrated that MAS NMR spectroscopy permits the direct, unequivocal analysis of flexible histone tails in large nucleosome arrays at cellular concentrations. A strong correlation exists between the most dynamic histone residues identified in the current study, amino acids 1–35 in H3 and 1–21 in H4, and the crystal structure of the nucleosome core particle which reveals the same residues as being largely unstructured. Our data are also generally in agreement with the NMR study of single nucleosomes in solution,\(^ {19}\) which reports that flexible H3 and H4 tails encompass residues 1–36 and 1–15, respectively. The finding that H4 residues 16–21 display significant mobility in the array samples could be indicative of somewhat different properties of these domains in the context of single nucleosomes at relatively low concentration versus highly concentrated arrays. Nevertheless, we also cannot discount the possibility that this difference stems from the fact that the solution NMR study focused solely on the protein backbone, while MAS NMR techniques monitor both the backbone and side-chain signals. Conversely, our results do not support the main conclusion of a recent H/D exchange NMR study that the H3 N-terminal tail forms stable folded structures in highly condensed nucleosome arrays.\(^ {20}\) The origin of the discrepancy between the H/D exchange study and MAS NMR
data is unclear. Yet, given that the former method does not visualize mobile residues directly but rather infers information about protein dynamics from amide proton occupancies, it is plausible that the histone tails in the highly condensed arrays could remain flexible yet concurrently somewhat protected from solvent exchange on the relatively short exchange time scales investigated. It is also noteworthy that the significant flexibility of histone tails in condensed nucleosome arrays revealed by MAS NMR is compatible with the EPR data for H3 presented in this study as well as the previous observation of mobile histone protein domains in oligonucleosomes by natural abundance $^{13}$C solution NMR at low and moderate ionic strengths, with the caveat that those experiments were done on relatively dilute nucleosome array samples.

The finding that H3 and H4 N-terminal tail domains are flexible even in highly condensed nucleosome arrays strongly suggests that chromatim compaction does not involve specific, high-affinity protein—protein or protein—DNA contacts that would lead to their immobilization. However, given the well-established importance of H3 and H4 histone tails in the folding of chromatin fibers, it is feasible that these flexible domains participate in the chromatin condensation process via multiple weak, transient interactions that shield the electrostatic repulsion between DNA moieties associated with different nucleosome units. The overall conformational flexibility of the H3 and H4 tails indicates that they are accessible to the numerous regulatory proteins that bind histone tails to regulate transcription even in compact heterochromatin at extremely high cellular densities. This provides a mechanism by which transcription activating complexes could directly gain access to histone tails to initiate conversions from heterochromatin to euchromatin. Factors in addition to histone tails that compact chromatin are also dynamic. Heterochromatin Protein 1 (HP1), which binds trimethylated lysine 9 within H3, and linker histone H1, which binds near the nucleosome DNA entry–exit region, facilitate heterochromatin formation. Both HP1 and H1 rapidly exchange in vivo, suggesting that multiple factors, which compact chromatin, function through dynamic mechanisms. Further studies of histone tails and the proteins that bind them are required to determine the role of histone tail dynamics in the conversion between heterochromatin and euchromatin.

**ASSOCIATED CONTENT**

Supporting Information

Materials and methods, detailed descriptions of gel electrophoresis, AFM, UV-based sedimentation and 1D variable temperature NMR studies, and supplementary figures with gel, AFM, NMR, and EPR data. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

Corresponding Authors

mpoirier@mps.ohio-state.edu
jaroniec@chemistry.ohio-state.edu

Notes

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This work was supported by grants from the Center for Emergent Materials at Ohio State (NSF MRSEC, DMR-0820414), the NIH (R01GM094357 to C.P.J. and R01GM083055 to M.G.P.), the Camille & Henry Dreyfus Foundation (Camille Dreyfus Teacher-Scholar Award to C.P.J.), and an OSU-CCC Pelotonia Predoctoral Fellowship (J.A.N.). The authors thank Dr. Swagahta Chakraborty for assistance with the preparation of $^{13}$C,$^1$N-labeled histone protein samples.

**REFERENCES**

Supporting Information

Histone H3 and H4 N-Terminal Tails in Nucleosome Arrays at Cellular Concentrations Probed by Magic Angle Spinning NMR Spectroscopy

Min Gao,† Philippe S. Nadaud,† Morgan W. Bernier,‡ Justin A. North,‡ P. Chris Hammel,† Michael G. Poirier,‡ and Christopher P. Jaroniec†

†Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States
‡Department of Physics, The Ohio State University, Columbus, Ohio 43210, United States
Materials and Methods

Preparation of Nucleosome Arrays

Histones H2A, H2B, H3 and H4 were overexpressed in Escherichia coli BL21(DE3)pLysS, purified as described previously1 using gel filtration and ion-exchange chromatography in 7 M urea followed by dialysis against a solution of 2 mM β-mercaptoethanol (BME) in ultrapure water and lyophilized. Uniformly 13C,15N-labeled H3 and H4 were prepared by using a minimal medium with 13C-glucose (3 g/liter) and 15NH4Cl (1 g/liter) (Isotec/Sigma-Aldrich) as the sole carbon and nitrogen sources, respectively. Histone octamer containing 13C,15N-H3 or H4 was prepared by dissolving the four unfolded histone proteins at concentrations of ≤ 10 mg/ml in 7 M guanidine hydrochloride, 20 mM Tris, 10 mM dithiothreitol, pH 7.5 unfolding buffer in a H2A:H2B:H3:H4 molar ratio of 1.2:1.2:1:1, refolded by double dialysis into 1× TE (10 mM Tris, 1 mM EDTA, pH 8.0), 2 M NaCl, 5 mM BME buffer and purified by gel filtration chromatography in 1× TE, 2 M NaCl buffer as described previously.1

The DNA templates (Figure S1) were prepared by using QIAGEN Giga kits and digested with DdeI (New England Biolabs). The digested DNA mixture contains a 3,046 bp linear dsDNA, corresponding to a 17-mer tandem repeat of a 147 bp variant of the Widom 601 nucleosome positioning sequence with 30 bp of linker DNA.2 The mixture also contains seven shorter DNA fragments of varying length (653, 535, 421, 404, 245, 230, and 161 bp), which serve as buffering DNA in the process of nucleosome array reconstitution and aid in minimizing the non-specific aggregation of the arrays.3

The 17-mer nucleosome arrays were prepared as follows. An aqueous solution was made in 0.5× TE, 2 M NaCl, 1 mM benzamidine (BZA) buffer, containing the digested DNA mixture ([DNA] ≤ 0.2 mg/ml) and a two-fold molar excess of the histone octamer with respect to the DNA, which ensured an effectively complete saturation of all the nucleosome positioning sites. The NaCl was removed by double dialysis at 4 oC against 0.5× TE, 1 mM BZA buffer. The solution was concentrated ~30-fold using Amicon 30 kDa cut-off centrifugal filter devices, followed by purification of the nucleosome arrays using 5-40% sucrose gradient centrifugation in 0.5× TE. The fractions containing pure 17-mer nucleosome arrays were combined and the sucrose removed by exchanging into 0.5× TE buffer using Amicon 30 kDa devices.

Characterization of Nucleosome Arrays by Gel Electrophoresis and Atomic Force Microscopy

The purity and level of saturation of the 17-mer nucleosome arrays were confirmed by using composite agarose-polyacrylamide gel electrophoresis for as-prepared arrays and polyacrylamide gel electrophoresis for BamHI-digested and AvaI-digested arrays (see Figure S2 and S3 captions for details).

The sucrose gradient purified arrays were further analyzed by atomic force microscopy (AFM) as follows. Freshly cleaved mica was rinsed with 2 × 200 µl of ultrapure water, followed by treatment with 50 µl of 10 ng/µl aqueous solution of poly-D-lysine (PL; Sigma-Aldrich) and an additional 2 × 200 µl water rinse to remove any unbound PL, and finally dried with a slow stream of dry nitrogen gas. The PL-treated mica was incubated for 5 min with 50 µl of a dilute (~0.2 nM) solution of nucleosome arrays in 0.1× TE buffer, rinsed with 200 µl of ultrapure water and dried as above. The sample was imaged with a Bruker AXS Dimension Icon AFM at a scan rate of 1 Hz.

Additional AFM studies and gel electrophoresis for AvaI-digested arrays were performed in order to assess the stability of the 17-mer nucleosome array samples under typical conditions used for the NMR experiments. These studies were carried out using a dilute nucleosome array solution incubated at a temperature of 30 °C for up to 7 days as well as a small aliquot of previously ultracentrifuged arrays extracted directly from the NMR sample holder following several days of experiments at 30 °C (see Figure S3 caption and NMR Spectroscopy section below for details).
Self-Association of Nucleosome Arrays Monitored by UV Absorbance

The nucleosome array self-association assay was performed as described. A series of array samples with volumes of 25 μl and concentration corresponding to OD<sub>260</sub> ~ 1 (1 mm pathlength) were mixed with 25 μl of aqueous solution containing increasing amounts of MgCl<sub>2</sub> in the 0-5 mM range for the final concentration. After 5 min of incubation at room temperature, the absorbance at 260 nm was measured by using a NanoDrop 2000 spectrophotometer (Thermo Scientific). Each sample was then centrifuged for 5 min at 16,000×g and the absorbance at 260 nm recorded immediately for the supernatant. For each MgCl<sub>2</sub> concentration, the percentage of nucleosome arrays in the supernatant was calculated from the ratio of the absorbance intensities at 260 nm after and before centrifugation.

NMR Spectroscopy

Three 17-mer nucleosome array samples were studied for each type of <sup>13</sup>C,<sup>15</sup>N-labeled histone (H3 or H4), prepared with 0 mM, 1 mM and 5 mM Mg<sup>2+</sup> and corresponding to different degrees of array compaction and self-association. Each nucleosome array sample was pelleted by ultracentrifugation at 4 °C and 80,000 rpm (~320,000×g) using a Beckman-Coulter TLA-100.3 rotor. Due to the distinct physical properties of the array samples prepared with increasing concentrations of Mg<sup>2+</sup> (see Figure S4A), different ultracentrifugation times (7, 3 and 2 h for 0, 1 and 5 mM Mg<sup>2+</sup>, respectively) were required to pellet the samples for the NMR studies. The pellets were transferred to 3.2 mm thin-wall zirconia rotors (36 μl sample volume) and sealed using custom-made spacers to prevent sample dehydration during experiments. The final NMR samples contained between ~8-16 mg of nucleosome arrays with total protein:DNA ratios of approximately 1:1 (w/w) for each sample, corresponding to chromatin concentrations in the ~200-400 mg/ml regime. Note that the 0 mM Mg<sup>2+</sup> samples typically contained smaller quantities of the nucleosome arrays relative to samples prepared with 1 and 5 mM Mg<sup>2+</sup> due to the lower density of the array pellets following ultracentrifugation.

NMR spectra were recorded using a 500 MHz Varian spectrometer equipped with a 3.2 mm T3 magic-angle spinning (MAS) probe in either <sup>1</sup>H-<sup>13</sup>C or <sup>1</sup>H-<sup>15</sup>N configuration. The sample spinning rate was actively regulated to ca. ±3 Hz using a MAS control unit, and sample temperatures were controlled by a stream of dried compressed air delivered to the sample using a variable-temperature (VT) stack. The temperatures listed in the article consistently refer to the VT gas temperature at the sample, with the actual sample temperatures being ~5 °C higher on average caused by the frictional heating due to MAS. The pulse sequences used included standard cross-polarization<sup>5</sup> as well as one- and two-dimensional refocused INEPT<sup>6,7</sup> with parameters described in detail elsewhere.<sup>8</sup>

EPR Spectroscopy

Site directed mutagenesis (Quickchange Lightning Kit, Agilent Technologies) was used to add cysteine residues to the N-terminal tail of the <i>Xenopus laevis</i> histone H3(C110A) mutant. Two mutants were employed for these studies, including V35C and a mutant containing a non-native cysteine residue introduced before the N-terminal alanine found in wild-type H3 (referred to as C0). Recombinant histones carrying the desired mutations for spin labeling were purified from <i>Escherichia coli</i> and refolded into octamer as previously described.<sup>1</sup> Histone octamer was then reduced with 10 mM tris(2-carboxyethyl)phosphine (TCEP) and purified by gel filtration in 5 mM PIPES pH 6.1, 2 M NaCl buffer to remove misfolded aggregates and TCEP. The purified reduced octamer was labeled by increasing the pH to 8.0 using Tris buffer (50 mM final concentration) and rapidly adding 200-fold molar excess of 0.1 M MTSL ((1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate, Toronto Research Chemicals) in acetonitrile. The reaction was allowed to proceed for 1 h at room temperature, then overnight at 4 °C. Unreacted spin label was removed by dialysis into 3 × 500 ml of 5 mM PIPES pH 6.1, 2 M NaCl buffer. The labeling was confirmed by mass spectrometry, and 17-mer arrays containing labeled histone octamer were prepared and purified as described above for the NMR experiments.
Following sucrose gradient purification, arrays were concentrated to \(~1\ \mu\text{M}\), and 10 \(\mu\text{l}\) samples were loaded into quartz capillary tubes for the electron paramagnetic resonance (EPR) measurements.

EPR measurements were performed on a Bruker EMX X-band EPR spectrometer. Spectra were recorded at room temperature using 20 mW incident power and field modulation of 0.1 G at 100 kHz. A set of ten spectra were acquired for each H3 mutant as a function of Mg\(^{2+}\) concentration, baseline corrected, and normalized to the same number of spins. The resulting data are shown in Figure S9.
Supplementary Results

Characterization of Nucleosome Arrays by Gel Electrophoresis and Atomic Force Microscopy

Gel electrophoresis and atomic force microscopy were employed to assess the purity and level of saturation of the 17-mer nucleosome arrays used for the NMR studies. The gel shift assay in Figure S2A shows a single narrow band for the sucrose gradient purified 17-mer arrays (lane 3), which indicates that arrays saturated with histone octamer are the major species present in the sample. The saturation level of the 17-mer arrays was confirmed by gel electrophoresis analysis of arrays digested with BamHI and AvaI restriction enzymes. The histone octamers were removed following the BamHI digestion to allow for easier quantification of the fraction of cut DNA template. The inability of BamHI (which readily cuts the empty DNA template in the central nucleosome positioning sequence) to cleave the 17-mer arrays (Figure S2B, lane 3) suggests that the arrays are fully saturated with histone octamer making the digestion sites inaccessible to the enzyme. Likewise, AvaI digestion of the 17-mer arrays, which cuts in the linker DNA between nucleosomes (Figure S3D, lane 5), yields primarily single nucleosomes and no free 177 bp DNA fragments, indicating that the arrays are fully saturated. Finally, AFM analysis of dilute decompacted arrays shows a highly homogeneous sample of saturated 17-mer nucleosome arrays (Figure S3A). Quantitative analysis of the AFM data revealed that 98 ± 4% of the nucleosome positioning sites are saturated with histone octamer, corresponding to an average array occupancy of 16.6 ± 0.6 nucleosomes out of 17.

Additional experiments were performed to assess the fidelity of the 17-mer nucleosome array samples when exposed to ambient temperature for an extended period of time, corresponding to typical conditions used for the NMR studies. Specifically, samples of the sucrose gradient purified arrays were incubated at 30 °C for up to one week, digested with AvaI and analyzed by gel electrophoresis (Figure S3D, lanes 3 and 4). These arrays were found to be virtually identical to the control array sample stored at 4 °C, indicating that prolonged incubation of the 17-mer arrays at ambient temperature does not result in appreciable dissociation of histone octamer. Moreover, we have performed the same AvaI-digestion gel electrophoresis assay for a small sample of previously ultracentrifuged 17-mer arrays extracted from the NMR rotor following several days of data acquisition at 30 °C (Figure S3D, lane 6). This array sample was found to contain a detectable amount of free 177 bp DNA, consistent with histone octamer dissociation from a small fraction of the nucleosome positioning sites. However, quantitative analysis of the gel image (see Figure S3 caption for details) suggests that this effect is relatively minor, with 97% or 16.5 out of 17 nucleosome positioning sites per array on average containing histone octamer. Importantly, this finding is further supported by quantitative AFM analysis of the same sample (Figure S3B), which reveals that on average 95 ± 7% of nucleosome positioning sites (i.e., 16.2 ± 1.1 out of 17) are occupied. Altogether, the gel electrophoresis and AFM analyses of the 17-mer nucleosome arrays strongly suggest that the array samples are highly homogeneous, effectively fully saturated with histone octamer, and remain largely intact during the course of NMR experiments.

Self-Association of Nucleosome Arrays

In order to investigate the flexible histone protein domains in 17-mer nucleosome arrays by NMR spectroscopy as a function of the degree of array compaction and self-association, the array samples were prepared with 0 mM, 1 mM and 5 mM Mg²⁺. These 17-mer nucleosome arrays have been demonstrated to compact into 30-nm chromatin fibers by Mg²⁺ under dilute conditions by AFM imaging. To confirm that the 17-mer arrays undergo self-association upon addition of Mg²⁺, we monitored the soluble fraction before and after centrifugation by UV absorbance at 260 nm. Figure S4B shows UV spectra of the soluble fractions for nucleosome array samples that were prepared with 0 and 1 mM Mg²⁺ and briefly centrifuged at 16,000×g. The spectra illustrate that rapid sedimentation of the arrays is effectively complete following the addition of 1 mM Mg²⁺. Figure S4C shows the fraction of nucleosome arrays remaining in the supernatant (calculated from the ratio of absorbance intensities at
260 nm after and before centrifugation) as a function of Mg\(^{2+}\) concentration. These data allowed us to estimate that ~1 mM Mg\(^{2+}\) is required for array self-association, in the range expected based on published reports.\(^9\) In Figure S4D we show a plot of Mg\(^{2+}\) required for 50% self-association of 6-, 7-, 9- and 12-mer nucleosome arrays, determined previously in a systematic study by Hansen and co-workers,\(^10\) along with a fit of these data to a decaying single exponential. These results indicate that, as expected, lower Mg\(^{2+}\) concentrations are required for the compaction of longer nucleosome arrays. Remarkably, the predicted Mg\(^{2+}\) concentration needed for 50% self-association of the 17-mer arrays is in good quantitative agreement with the experimental Mg\(^{2+}\) concentration determined in the present study as seen by superimposing the experimental value for 17-mer arrays onto the plot in Figure S4D. Consistent with these results, the effect of increasing Mg\(^{2+}\) concentration on the physical appearance of the nucleosome array samples was readily evident as shown in Figure S4A, ranging from a fairly clear solution at 0 and 1 mM Mg\(^{2+}\) to a highly opaque one containing a significant amount of precipitate at 5 mM Mg\(^{2+}\).

**Observation of Rigid and Flexible Domains of Histones H3 and H4 in Nucleosome Arrays by Variable Temperature 1D \(^{13}\)C MAS NMR**

In order to investigate the rigid and flexible segments of histones H3 and H4 in the context of the highly concentrated 17-mer nucleosome arrays we recorded a set of 1D \(^{13}\)C MAS NMR spectra as a function of temperature between -20 and 30 °C. The \(^{13}\)C magnetization was prepared by using either \(^1\)H-\(^{13}\)C refocused INEPT\(^6\) or cross-polarization,\(^5\) conditions which lead to the suppression of NMR signals arising from the most rigid and most flexible protein residues, respectively. The data sets for the 17-mer arrays incubated with 1 mM MgCl\(_2\) and containing \(^{13}\)C,\(^{15}\)N-labeled H3 and H4 are shown in Figures S5 and S6, respectively. We note here that array samples prepared with 0 and 5 mM Mg\(^{2+}\) showed nearly identical temperature profiles and that repeated temperature cycling had no appreciable effect on the appearance of the NMR spectra, with any spectral changes observed as a function of temperature being largely reversible as shown in Figure S7.

The foremost feature of these experiments is that the INEPT spectra recorded between -10 and 30 °C contain multiple resonances in the aliphatic region, indicative of magnetization transfer through one-bond \(^1\)H-\(^{13}\)C J-couplings. This confirms the presence of flexible segments for both histones H3 and H4 irrespective of the degree of array compaction. Moreover, the spectra recorded at successively higher temperatures show narrower and more intense \(^{13}\)C signals, consistent with enhanced conformational dynamics of the protein backbone and side-chains. In contrast, the INEPT spectra recorded at or below -20 °C are largely devoid of signals. This suggests that around this temperature the large amplitude isotropic-like local protein motions occurring on the sub-millisecond time scale are sufficiently attenuated, leading to the rapid dephasing of \(^1\)H coherences by the residual \(^1\)H-\(^1\)H dipolar couplings that are not efficiently averaged by MAS. The CP spectra, which report on the immobile protein residues, indicate that the majority of amino acids in histones H3 and H4 are relatively rigid in the entire temperature range investigated, highlighted by the fact that the spectra recorded at -20 °C display a similar overall appearance with a ~20-30% increase in signal intensity relative to the 30 °C spectra due to the “freezing out” of protein motions. These findings are consistent with the high-resolution nucleosome X-ray structure, which reveals that unstructured N-terminal tails encompass ~20-25% of the H3 and H4 sequences,\(^11\) as well as previous observations of mobile N-terminal histone tails by solution NMR in mono and oligonucleosomes at relatively low concentrations in the absence of Mg\(^{2+}\).\(^12\^-15\)
**Figure S1.** DNA template for the preparation of nucleosome arrays. The template contains 17 tandem nucleosome binding repeats with 147 bp per repeat (positions 62-208 through 2717-2863) and 8 Ddel restriction sites as indicated. Following Ddel digestion, the DNA mixture contains a 3,046 bp nucleosome binding DNA sequence and seven shorter buffering DNA fragments (653, 535, 421, 404, 245, 230, and 161 bp) that assist in efficient nucleosome array reconstitution as described in Materials and Methods. The template also contains a single BamHI restriction site at position 1526 in the central (9th) nucleosome binding repeat and multiple AvaI digestion sites in the linker regions connecting the nucleosome binding repeats (spaced every 177 bp in positions 221, 398, 575, …, 2876). The BamHI and AvaI enzyme restriction sites are required for the gel electrophoresis assays used to evaluate the levels of nucleosome saturation in the array samples (c.f., Figures S2 and S3).
Figure S2. Gel electrophoresis assays used to assess the saturation level of 17-mer nucleosome arrays. (A) Native agarose-polyacrylamide composite gel (1% agarose-2% polyacrylamide, Tris-borate running buffer) stained with SYBR Gold (Invitrogen). Lane 1: DdeI-digested DNA template containing the 3,046 bp DNA with 17 tandem nucleosome binding repeats and shorter buffering DNAs; Lane 2: reconstituted 17-mer nucleosome arrays prior to sucrose gradient purification; Lane 3: sucrose gradient purified 17-mer nucleosome arrays. (B) Native agarose gel (1% agarose, Tris-acetate EDTA running buffer) stained with ethidium bromide. Lane 1: DdeI-digested DNA template containing the 3,046 bp DNA with 17 tandem nucleosome binding repeats and shorter buffering DNAs; Lane 2: DdeI-digested DNA template further digested with BamHI, which cuts the 3,046 bp DNA into two fragments of roughly equal size; Lane 3: sucrose gradient purified 17-mer nucleosome arrays digested with BamHI (only the 3,046 bp DNA is observed due to the BamHI site being occupied by the histone octamer and not accessible to BamHI), followed by proteinase K treatment to digest the histone proteins; Lane 4: same as lane 3 but for 17-mer nucleosome arrays prior to sucrose gradient purification.
Figure S3. AFM and gel electrophoresis analysis of the stability of 17-mer nucleosome arrays under conditions used for NMR studies. (A) AFM images of 17-mer nucleosome arrays recorded prior to pelleting for NMR experiments. (B) AFM images of a small aliquot of a pellet containing 17-mer nucleosome arrays following several days of NMR experiments at 30 °C. The pellet was extracted from the NMR rotor and incubated in 5 mM Tris, 0.5 mM EDTA, pH 8.0 buffer for one week to ensure a complete resuspension of the pelleted arrays prior to AFM image acquisition. Analysis of ~50-60 arrays for each of the images shown in panels (A) and (B) determined, respectively, that 98 ± 4% and 95 ± 7% of the nucleosome positioning sites are saturated with histone octamer, corresponding to an average occupancy of 16.6 ± 0.6 and 16.2 ± 1.1 nucleosomes out of 17 before and after the NMR experiments. (C) Cartoon depicting the Aval-digestion of saturated and unsaturated nucleosome arrays. For the saturated arrays the Aval digestion yields primarily single nucleosome particles, while both nucleosomes and free 177 bp fragments are obtained for arrays that are not fully saturated. (D) Native polyacrylamide gel (5% polyacrylamide, Tris-borate EDTA running buffer) stained with ethidium bromide. Lane 1: DdeI-digested DNA template containing the 3,046 bp DNA with 17 tandem nucleosome binding repeats and shorter buffering DNAs; Lane 2: DdeI-digested DNA template further digested with Aval, which cuts the 3,046 bp DNA into 17 fragments of 177 bp each; Lanes 3-5: sucrose gradient purified 17-mer nucleosome arrays incubated at 30 °C for 48 h (lane 3), 30 °C for 1 week (lane 4), and stored at 4 °C (lane 5), and subsequently digested with Aval; Lane 6: sucrose gradient purified 17-mer nucleosome arrays extracted from the ultracentrifuged pellet inside the NMR rotor following several days of experiments at 30 °C. For the pelleted arrays a weak band corresponding to 177 bp DNA is observed in addition to the single nucleosome band (no 177 bp DNA band is observed for the unpeled arrays incubated at 30 °C). The intensity ratio of the single nucleosome to 177 bp DNA bands was estimated to be 5.7 using ImageJ (http://rsb.info.nih.gov/ij). Given that the intensity in the single nucleosome band arises from 30 bp of linker DNA, the observed band intensity ratio indicates that 97% or 16.5 out of 17 nucleosome positioning sites per array are occupied with histone octamer. (E) Summary of average nucleosome array size before and after the NMR experiments estimated from AFM and Aval-digestion assays.
Figure S4. (A) 17-mer nucleosome arrays incubated with 0, 1 and 5 mM MgCl₂. (B) UV spectra of the supernatant fractions for samples containing 17-mer nucleosome arrays incubated for 5 min with 0 mM MgCl₂ (red) and 1 mM MgCl₂ (blue) and centrifuged for 5 min at 16,000×g. (C) Percentage of nucleosome arrays in the supernatant as a function of MgCl₂ concentration calculated from the ratio of absorbance intensities at 260 nm after and before centrifugation. Representative UV spectra are shown in panel (B). (D) Experimental data (open circles) obtained from the study of Schwarz et al.¹⁰ showing the MgCl₂ concentration required to achieve 50% self-association of nucleosome arrays as a function of the array size in the range of 6 to 12 nucleosomes. Fit of experimental data of Schwarz et al. to a single exponential decay (line). MgCl₂ concentration required for 50% self-association of the 17-mer nucleosome arrays determined in this study (filled circle).
Figure S5. Variable temperature study of $^{13}$C-$^{15}$N-H3 17-mer nucleosome arrays incubated with 1 mM MgCl$_2$. The series of $^{13}$C spectra on the left and right, respectively, were recorded at the indicated temperatures using refocused INEPT and CP pulse schemes, and display resonances corresponding to the flexible and rigid segments of H3, respectively. Each spectrum was recorded with 1024 scans and recycle delay of 2.5 s, and all spectra of the same type were processed in identical fashion and are displayed on the same horizontal and vertical scales.
Figure S6. Variable temperature study of $^{13}$C,$^{15}$N-H4 17-mer nucleosome arrays incubated with 1 mM MgCl$_2$. The series of $^{13}$C spectra on the left and right, respectively, were recorded at the indicated temperatures using refocused INEPT and CP pulse schemes, and display resonances corresponding to the flexible and rigid segments of H4, respectively. Each spectrum was recorded with 1024 scans and recycle delay of 2.5 s, and all spectra of the same type were processed in identical fashion and are displayed on the same horizontal and vertical scales.
Figure S7. Effect of temperature cycling on $^{13}$C refocused INEPT (A,C) and CP (B,D) spectra of 17-mer nucleosome arrays incubated with 1 mM MgCl$_2$. The spectra for arrays containing $^{13}$C,$^{15}$N-H3 (A,B) and $^{13}$C,$^{15}$N-H4 (C,D) were all obtained at 0 °C for a fresh array sample previously handled and stored at 4 °C (blue), after previously heating the sample to 30 °C (green), and after previously heating the sample to 30 °C followed by cooling to -20 °C (red). The acquisition and processing parameters used were identical to those used to obtain the spectra in Figures S5 and S6.
Figure S8. (A) Refocused INEPT based pulse scheme used to monitor transverse amide $^1$H relaxation in 17-mer nucleosome arrays containing $^{13}$C,$^{15}$N-labeled H3 or H4. (B) Series of $^{15}$N spectra recorded at 30 °C as a function of the $^1$H relaxation delay, $2(\Delta+\delta)$, in the 4-12 ms range for arrays containing $^{13}$C,$^{15}$N-H3 and incubated with 0, 1 and 5 mM MgCl$_2$ as indicated. (C) Same as panel (B) but for arrays containing $^{13}$C,$^{15}$N-H4. Note that the spectra within each relaxation series are shown on the same vertical scale. However, given that only the rate of $^1$H signal decay within each series is of interest, the spectra for samples prepared with different Mg$^{2+}$ concentrations were not normalized with respect to each other based on the sample quantities but rather scaled to have a similar overall signal intensity for the first relaxation delay in the series.
**Figure S9.** X-band CW EPR spectra of 17-mer nucleosome arrays containing C0 (left) or V35C (right) H3 variants spin labeled with MTSL and incubated with 0, 1 and 5 mM MgCl₂ as indicated. All spectra were recorded at room temperature and are normalized to represent the same number of spins. The relatively narrow lineshapes resemble those reported for other unstructured proteins\textsuperscript{16,17} and indicate that the H3 N-terminal tail exhibits significant flexibility.
References