Evidence for Nonrandom Behavior in 208-12 Subsaturated Nucleosomal Array Populations Analyzed by AFM†

Jaya G. Yodh,‡ Yuri L. Lyubchenko,§ Luda S. Shlyakhtenko,|| Neal Woodbury,⊥ and D. Lohr*,⊥

Division of Basic Sciences, Midwestern University, Arizona College of Osteopathic Medicine, Glendale, Arizona 85308, Department of Microbiology and Department of Biology, Arizona State University, Tempe, Arizona 85287, BioForce Laboratory Inc., 2501 North Loop Drive, Suite 614B, Ames, Iowa 50010-8277, and Department of Chemistry and Biochemistry, Arizona State University, P.O. Box 871604, Tempe, Arizona 85287-1604

Received May 6, 1999; Revised Manuscript Received August 23, 1999

ABSTRACT: Atomic force microscopy was used to determine the population distributions in reconstituted, subsaturated 208-12 nucleosomal arrays. The features found in these distributions vary with the average nucleosome loading per template molecule (n_{av}): at $n_{av} < 4$, the distributions show a single peak whose breadth is equal to that expected for a random loading process; at $n_{av} = 4-8$, the distributions are broader than random distributions and are complex; i.e., they contain multiple peaks and/or shoulders. Moreover, the peaks/shoulders typically occur at two nucleosome intervals, i.e., 2, 4, 6 or 3, 5, 7 nucleosomes. This two-nucleosome periodicity is statistically significant. The precise cause for such discrete features within the distributions is unknown, but at least these features would seem to indicate some pairwise preference in nucleosome occupation at these loading levels. In these intermediate-level ($n_{av} = 4-8$) distributions, the major peak contains a larger fraction of the total templates than a random nucleosome loading process would produce. This feature indicates that at these intermediate population levels there is some tendency for correlated nucleosome loading among the templates. Hyperacetylated nucleosomal arrays show only subtle differences in their population distributions compared to nonacetylated arrays and demonstrate the above features. AFM allows one to study unfixed chromatin arrays; we find that nucleosomes on the 208-12 template demonstrate significant lability when they are not glutaraldehyde-fixed.

The eukaryotic chromatin fiber is generally assumed to be densely packed with nucleosomes, with approximately one every 200 bp[1] (*1*). However, functionally active chromatin often contains nucleosome-free regions interspersed among the more densely packed stretches. Locations where such regions are found include replication origins (cf. *2, 3*) and transcriptional regulatory regions of various genes (cf. *4, 5*). Such a "subsaturated" chromatin fiber will have properties that differ from those of a fully saturated, more densely packed one. As a model for such a mixed fiber, we have utilized a polynucleosomal template, the 208-12 array, reconstituted with subsaturating levels of nucleosomes.

The 208-12 DNA template has been previously described (*6*). It is a linear fragment containing 12 tandem repeats of the 208 bp 5S rRNA gene from *L. variegatus*. This template has proven to be extremely useful for it allows the in vitro assembly of polynucleosomal arrays that are homogeneous enough for biophysical and biochemical studies. This is probably due to the ability of each 208 bp unit to bind a nucleosome in discrete positions on the DNA (*7, 8*). Studies utilizing the 208-12 array have yielded valuable information on determinants of chromatin folding, nucleosome assembly and positioning, chromatin remodeling, and transcription at both saturating and subsaturating nucleosome densities (reviewed in *9; 10*−*13*).

We are especially interested in determining how nucleosomes populate the 12-mer template at subsaturating molar ratios (r) of histone octamer to DNA, i.e., when the number of nucleosome binding sites in a template exceeds the number of nucleosomes available to bind. Sedimentation analyses have demonstrated that nucleosomal arrays reconstituted at various subsaturating r values yield distinct sedimentation coefficient distributions (*14*). This indicates that discrete populations of subsaturated nucleosomal arrays can be reconstituted. However, the sedimentation coefficient distributions for the midrange subsaturating arrays ($r = 0.5-0.9$) were broader than those for fully saturated arrays or naked DNA templates, indicating that populations of templates at these intermediate loading levels have some heterogeneity. It is important to know the exact form of the distributions in subsaturated populations for these data can answer important questions: For example, does nucleosome loading follow random behavior? In order to obtain precise population distributions and thus answer these types of

† This study was supported in part by American Cancer Society Grant PF4133 to J.G.Y. and NIH Grant GM54991 to Y.L.L. and L.S.S.

* To whom correspondence should be addressed. Phone: 480-965-5020. Fax: 480-965-2747. E-mail: DLohr@ASU.Edu.

‡ Division of Basic Sciences, Midwestern University, Arizona College of Osteopathic Medicine.

§ Department of Microbiology and Department of Biology, Arizona State University.

|| Department of Microbiology, Arizona State University, and BioForce Laboratory Inc.

⊥ Department of Chemistry and Biochemistry, Arizona State University.

[1] Abbreviations: EDTA, ethylenediaminetetraacetic acid; bp, base pair(s); AP, aminopropyl.

10.1021/bi991034q CCC: $18.00
Published on Web 11/09/1999

questions, we have turned to atomic force microscopy (AFM) to measure the population distributions in 208-12 nucleosomal arrays reconstituted at subsaturating levels.

AFM is a powerful imaging technique that has been used to visualize a number of nucleic acid and nucleoprotein complexes (cf. *15*−*17*). The technique offers many advantages for the type of investigation we will carry out here. One can directly visualize individual templates and measure both the number of nucleosomes present and their positions on the template. By counting many molecules, it is possible to determine the exact population distribution. Thus, AFM offers a way to determine experimental nucleosome population distributions for comparison to theoretical models. AFM has already been used to visualize the 208-18 (an 18-site template) reconstituted nucleosomal array (*18*) and used extensively in studies of native chromatin structure (cf. *19*).

Histones are acetylated in vivo in a nonrandom fashion at specific lysines on the N-terminal tails of all four core histones (*20*). These tails are thought to be involved in interactions with nucleosomal and linker DNA and to play a role in stabilizing internucleosomal contacts in higher-order folded chromatin structures (*21*−*25*). Histone acetylation has been strongly correlated with in vivo transcriptional activity of chromatin in numerous studies, with the general finding that active chromatin is hyperacetylated while more silent regions are hypoacetylated (cf. *26*−*29*). Histone acetylation also plays a role in chromosome replication; newly deposited histones are specifically acetylated (cf. *30*). For these reasons, we have determined nucleosome population distributions for hyperacetylated chromatin. We have also examined the effect of fixation on the nucleosome population distributions obtained.

EXPERIMENTAL PROCEDURES

Materials. Restriction enzyme *Hin*PI was purchased from New England Biolabs. Glutaraldehyde was from EM Sciences. HeLa cells were obtained from Cellex BioSciences, Inc. The linear 208-12 DNA template for polynucleosome reconstitutions corresponded to a 2530 bp *Hin*PI restriction fragment derived from the pPo1 I 208-12 plasmid (*31*) which was generously provided by Dr. Jeff Hansen at The University of Texas Health Science Center, San Antonio, TX. The template consists of an array of 12 tandem 208 bp repeats of the 5S rRNA gene from *L. variegatus*. Each 5S repeat contains an internal nucleosome positioning site which rotationally positions one nucleosome. The translational positioning of nucleosomes within the 5S repeat has been mapped previously (*7*, *8*).

Purification of Native Histone Octamers. For purification of hyperacetylated histones, HeLa cells were treated with 10 mM sodium butyrate, a deacetylase inhibitor (*32*−*34*), 24 h prior to harvesting. Native nonacetylated and native hyperacetylated histone octamers were prepared from 15 L of untreated or butyrate-treated HeLa cells, respectively, according to the method of Workman et al. (*35*). This procedure consists of a series of extractions of nuclear pellets with buffers containing increasing NaC1 concentrations followed by chromatography on Bio-Gel HTP (BioRad, DNA-grade) to isolate histone octamers free of histone H1 (verified by gel electrophoresis). The histone octamers were concentrated using an Amicon ultrafiltration unit and YM10 membranes, and stored in High Salt Buffer, HSB [50 mM $NaPO_4$ (pH 6.8), 0.5 mM PMSF, and 2.5 M NaC1]. Octamer concentrations were measured by the method of Bradford et al. (*36*) and by absorbance at A_{230} and based on a histone standard.

The histone octamer purity and stoichiometry were analyzed on Coomassie-stained 15% SDS−PAGE gels. All four core histones, H2A, H2B, H3, and H4, are present in equal stoichiometry in our octamer preparation; H1 is completely absent (not shown). Acetylation levels were examined on Triton−acid−urea (TAU) gels (*37*, *38*). The TAU gels contained 14% acrylamide, 1% bisacrylamide, 6 M urea, 0.375% Triton X-100, and 5% acetic acid and were electrophoresed in 5% acetic acid for 24 h at 150 V (anode to cathode). Histone octamers were precipitated with TCA (25% final concentration) before being loaded onto both SDS−PAGE and TAU gels. TAU gels were scanned using SigmaGel software (Jandel Scientific) to determine the extent of histone acetylation in the octamer preparations. Our hyperacetylated octamers contain acetylated forms of all four core histones, with the extent of acetylation being greatest for H4 and H2B. For example, the hyperacetylated histone H4 pool consists of mainly di- and triacetylated species. The nonacetylated H4 pool consists mainly of unacetylated H4.

Reconstitution of Nucleosomal Arrays. Nucleosomal arrays were reconstituted at various saturation levels following the method of Hansen and Lohr (*14*) with some modifications. Briefly, HeLa histone octamers were mixed with the linear 208-12 DNA template at varying molar ratios of histone octamers to DNA. DNA and histone octamers were mixed together on ice at a final concentration of 2 M NaCl, and the final volume was adjusted to 50 μL with TE (10 mM Tris−1 mM EDTA, pH 8.0). Histone octamers were diluted in HSB buffer and added last to the reaction mix. Reaction mixtures were dialyzed (Spectra/Por 6−8 kDa) at 4 °C against 1 L of TE containing 1 M NaCl for >12 h followed by dialysis versus 1 L of TE containing 0.75 M NaCl for 5−6 h. The final dialysis step was against 1 L of 1 mM EDTA (pH 8.0) for >12 h. Chromatin was fixed by dialyzing against fresh 0.1% gluteraldelyde in 1 mM EDTA (pH 8.0) for 24 h. Excess glutaraldehyde was removed by a final dialysis for 24 h against 1 L of 1 mM EDTA (pH 8.0). Fixed samples were recovered and stored on ice until further use.

To confirm that the reconstitutions were successful and to estimate their concentrations, an aliquot of each reconstitute was electrophoresed on a native 3.5% polyacrylamide gel overnight at 4 °C. Oligonucleosomes migrate as a smear between free 208-12 DNA template and the gel well. The more nucleosomes the template contains, the slower it migrates. Those templates carrying ≥8 nucleosomes only minimally enter the gel.

For the experiments on the effects of fixation, reconstituted samples were recovered from dialysis and divided into two aliquots. Half of the sample was not fixed and was stored directly on ice. The other half of the sample was fixed as usual (see above).

Deposition of the Samples onto AP-Mica. The samples for AFM were prepared as described (*15*, *17*, *39*, *40*). Briefly, immediately prior to loading, chromatin samples were diluted approximately 10-fold. Ten microliters of the sample (DNA concentration 0.2−0.4 μg/mL) in 1 mM EDTA (pH 8.0) was pipetted onto pieces of AP-mica for 2 min, rinsed with

deionized water (ModuPure Plus, Continental Water System Corp., San Antonio, TX), and argon-dried.

Atomic Force Microscopy (AFM) Imaging. The AFM images were taken in the air with a MultiMode SPM instrument equipped with D-scanner (Digital Instruments, Inc., Santa Barbara, CA) operating in TappingMode. NanoProbe TESP probes (Digital Instruments, Inc.) and conical sharp silicon tips (K-tek International, Portland, OR) were used for imaging. The typical tapping frequency was 240−280 kHz for TESP tips and 340−380 kHz for the K-tek probes; the scanning rate was 2−3 Hz.

For each reconstitution sample, greater than 100 molecules were analyzed. "Countable" molecules were those which had discernible ends. They could be either extended or more compacted, but in all cases, the nucleosomes had to be easily measurable. Nucleosome loading distributions were obtained by plotting the fraction of molecules with a given number of nucleosomes versus the number of nucleosomes. Distributions were obtained for nonacetylated and hyperacetylated arrays with $n_{av} > 2$ to $n_{av} \simeq 9$. n_{av} is the average number of nucleosomes present on the templates in a distribution (see below).

Statistical Analysis. Experimental nucleosome loading distributions, obtained from atomic force microscopy analysis, were compared to theoretical distributions calculated for the same average number of nucleosomes (n_{av}) and assuming that nucleosomes bind randomly and with equal frequencies to each of the twelve 208 bp sites, regardless of the histone octamer to DNA ratio. For the theoretical (random) distributions, the fraction of DNA fragments with n nucleosomes bound (F_n) is given by

$$F_n = \left(\frac{n_{av}}{12}\right)^n \left(1 - \frac{n_{av}}{12}\right)^{(12-n)} \frac{12!}{(12-n)!n!}$$

Note that $n_{av}/12$ is the random probability that any particular nucleosome site is occupied. Thus, $(n_{av}/12)^n(1 - n_{av}/12)^{(12-n)}$ is simply the total probability that any given arrangement of n nucleosomes will exist, and $12!/(12-n)!n!$ is the total number of arrangements, assuming that nucleosomes are indistinguishable.

RESULTS

Atomic Force Microscopy Images of Reconstituted Nucleosomal Arrays. Figure 1 displays several representative AFM spreads of reconstituted, subsaturated 208-12 nucleosomal arrays. The arrays show the typical "beads-on-a-string" appearance observed at low ionic strength in the absence of linker histones (*18*). These images also dramatically demonstrate the physical effect of nucleosome loading; the more nucleosomes on a DNA template, the generally more compact it becomes (Figure 1A,B). In each field, we counted every molecule in which all the nucleosomes were clearly distinguishable, regardless of its compaction level, to try to avoid bias in the data set. From one spread of roughly 2−4 ng of DNA, one can count as many molecules as needed. AFM spreading on AP-mica offers the advantage of viewing many more molecules per spread than on untreated mica, while maintaining high-quality chromatin spreads (Figure 1).

Nucleosome Population Distributions of Nonacetylated Subsaturated Arrays. To assess the data quantitatively, we determined the population distribution for each reconstitute. This distribution is obtained by plotting the fraction of molecules with a particular number of nucleosomes (F_n) versus the number of nucleosomes (n). Figure 2 shows distributions for nonacetylated nucleosomal arrays reconstituted at various loading levels. We will refer to them by their n_{av}, the average number of nucleosomes per template molecule. Figure 2A,B,C show samples with $n_{av} = 2.6$, 5.4, and 8.5. The following equation is used to calculate n_{av}:

$$n_{av} = \frac{\sum_1^{12} n \times (\text{no. of molecules with } n \text{ nucleosomes})}{\text{total no. of molecules counted}}$$

The error bars in Figure 2 depend on the number of molecules counted at each value of n and are calculated according to the following equation:

$$\text{error} = \frac{\sqrt{\#\text{ of molecules with } n \text{ nucleosomes}}}{\#\text{ of molecules with } n \text{ nucleosomes}} \times F_n$$

The experimental distributions (closed circles, solid lines) are compared to the distributions which would be obtained for a random loading process (dotted lines), calculated for the same experimental n_{av} and assuming random and equivalent binding of nucleosomes to each of the 12 sites, with neither a preference nor a dispreference for nearest-neighbor occupancy. The equation used to calculate these theoretical (random) distributions is given under Experimental Procedures.

First, from the distributions displayed in Figure 2, it is clear that one obtains distinct populations of arrays in these subsaturated reconstitutes. Morever, there is generally a direct relationship between the input histone/DNA ratio and the n_{av}. For example, a plot of input histone/DNA ratio versus n_{av} is linear from $n_{av} = 1.4$ to >9 (not shown). Similar conclusions were obtained in the sedimentation-based analysis of subsaturated arrays (*14*). We note that these AFM results were obtained with HeLa histones while Hansen and Lohr (*14*) used chicken erythrocyte histones.

Comparison of the experimental distributions to the corresponding random distributions allows us to determine whether nucleosome loading on this multisite template follows a random pathway. In general, the results are consistent with a random process. However, there are some variations. These are dependent on the precise level of nucleosome loading.

(i) $n_{av} < 4.0$. At these low loading levels, the distributions closely resemble the single peak, symmetric type of curve that would be expected for a random loading process, and the breadths of the distributions coincide fairly well with the breadths expected for such a process (cf. Figure 2A). A random process is one in which all 208 bp DNA sites bind nucleosomes equally well, with no preference (or dispreference) for nearest-neighbor site occupation.

(ii) $n_{av} > 4.0$. At higher loadings, there is significant complexity in the experimental distributions, i.e., multiple peaks and/or pronounced shoulders in addition to a major peak (cf. Figure 2B). Distributions with $n_{av} < 4.0$ sometimes show suggestions of shoulders, but those distributions are clearly much less complex. The complexity of the distributions at $n_{av} > 4.0$ makes their comparison to random (single

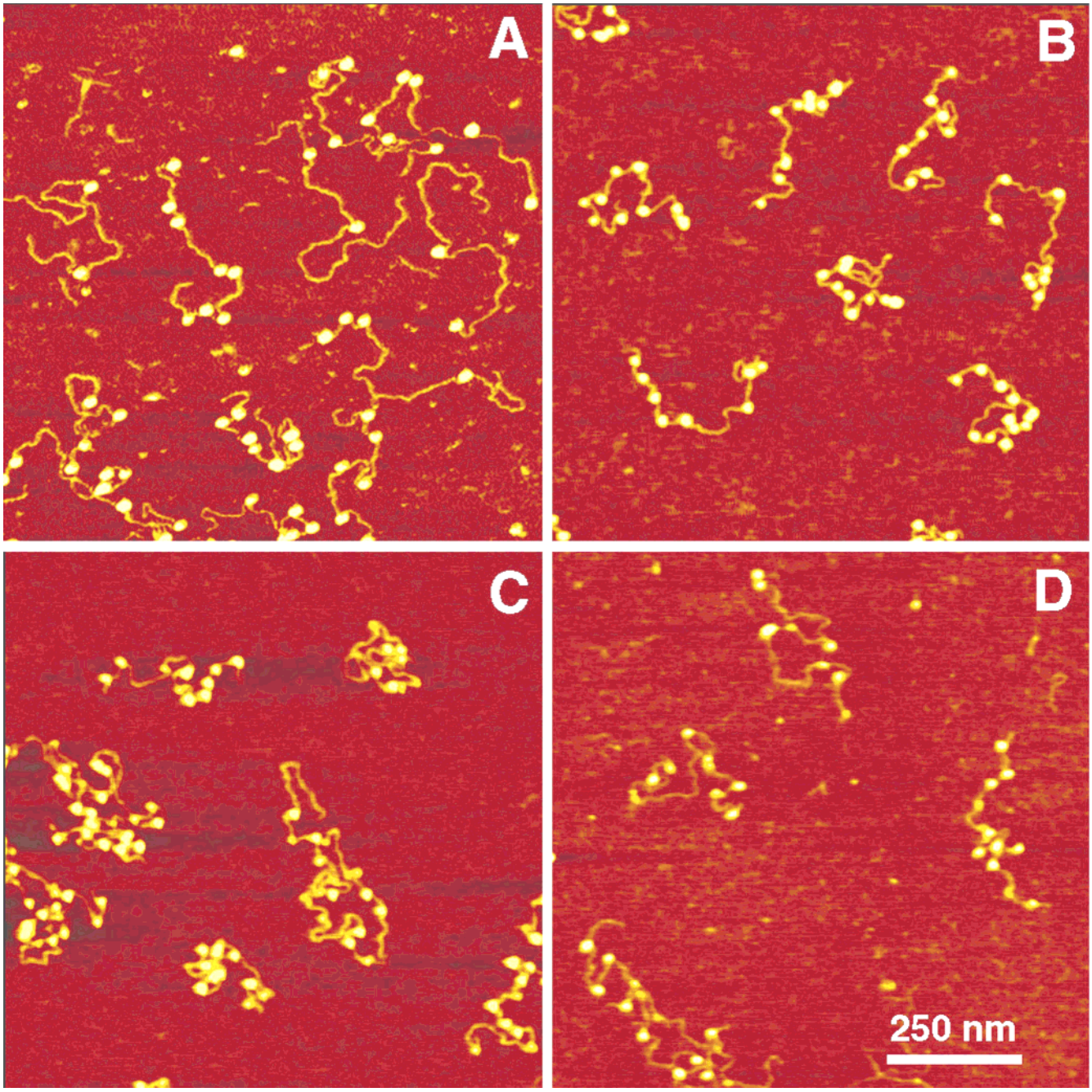

FIGURE 1: Atomic force microscopy images of subsaturated 208-12 nucleosomal arrays. Subsaturated 208-12 nucleosomal arrays were reconstituted at various histone/DNA ratios as described under Experimental Procedures. Representative images of samples from such reconstitutes are shown in fields A−D. Samples are characterized by n_{av}, the average number of nucleosomes per template molecule. (A) Fixed, nonacetylated arrays with $n_{av} = 5.4$. (B) Fixed, nonacetylated arrays with $n_{av} = 8.5$. (C) Unfixed, nonacetylated arrays with $n_{av} = 6.9$. (D) Fixed, hyperacetylated arrays with $n_{av} = 8.5$.

peak) profiles difficult. However, these distributions do seem to be broader than those expected for a simple random process (Figure 2B).

The experimental distributions at $n_{av} > 4.0$ are also broader than those observed at $n_{av} < 4.0$. For example, from sedimentation analysis of subsaturated 208-12 nucleosomal arrays, it was estimated that for any particular subsaturated sample, the effective breadth of the distribution is ±2; i.e., the various templates within the population contain no more than ±2 nucleosomes from the peak value (*14*). The AFM data show that for $n_{av} < 4.0$, more than 90% of the molecules do contain no more than ±2 nucleosomes from the peak value. At $n_{av} > 4$, the percentages fall to around 75−80%, with no consistent variations until significant loading levels are reached, e.g., $n_{av} \simeq 9$ (data not shown). Thus, by this quantitative measure, the distributions at $n_{av} > 4.0$ are broader than those at $n_{av} < 4.0$. This broadening is predicted from the theoretical analysis (cf. Figure 2A versus Figure 2B). The theoretical analysis also predicts that the distributions should again become narrower at high n_{av}, e.g., $n_{av} > 8$. It becomes difficult to count nucleosomes unambiguously as the templates become more highly loaded, so we have not extensively analyzed loadings beyond $n_{av} \simeq 9$.

Nucleosome Population Distributions of Hyperacetylated Arrays. We have also analyzed acetylated histone reconstitutes to test whether this important histone modification

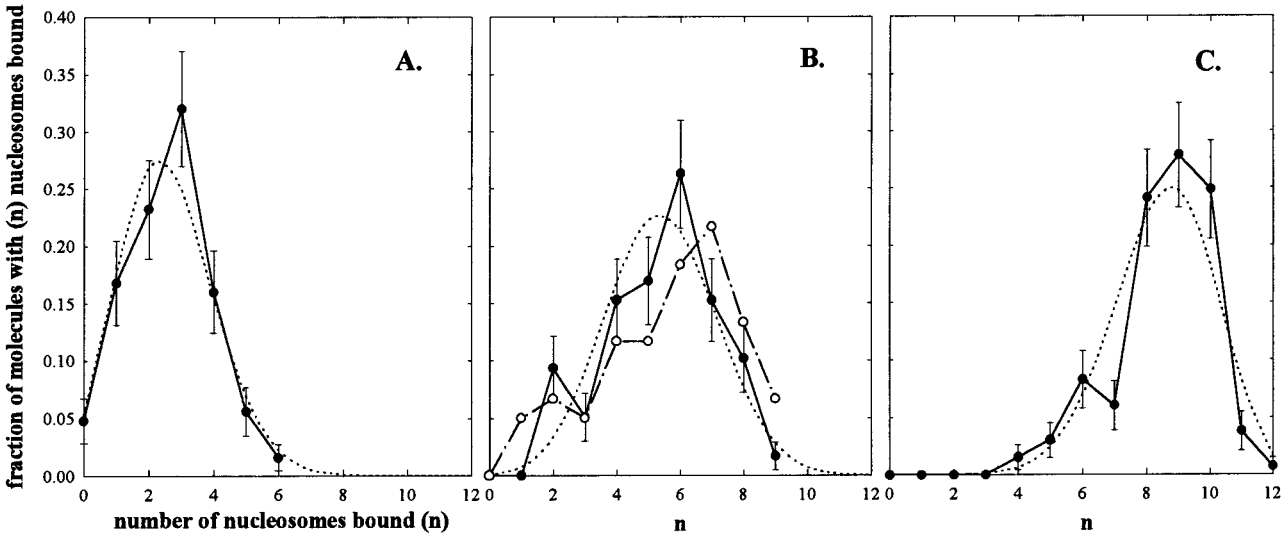

FIGURE 2: Nucleosome population distributions of subsaturated 208-12 nucleosomal arrays. The fraction of molecules with a particular number of nucleosomes bound (F_n) is plotted versus n, the number of nucleosomes, for (A) fixed, nonacetylated arrays with $n_{av} = 2.6$; (B) fixed, nonacetylated arrays with $n_{av} = 5.4$; and (C) fixed, nonacetylated arrays with $n_{av} = 8.5$. Both the experimental results (filled circles, solid lines) and random distributions (dotted lines), calculated as described under Experimental Procedures, are shown. Error bars reflect statistical uncertainties based on the number of molecules counted (see Results). In panel B, we also show a distribution for a fixed, hyperacetylated array with $n_{av} = 5.7$ (open circles, dash−dotted lines).

affects nucleosome loading on these subsaturated templates. Hyperacetylated histones were used in these studies in an attempt to maximize any acetylation-dependent effects. In general appearance, the AFM images of acetylated arrays resemble the nonacetylated ones (cf. Figure 1B versus Figure 1D). A population distribution for an acetylated reconstitute ($n_{av} = 5.7$, open circles) is compared to a comparably loaded nonacetylated reconstitute ($n_{av} = 5.4$, filled circles) in Figure 2B. Again, the two samples give similar results. The distributions for acetylated arrays do appear to be slightly broader than the corresponding (judged by n_{av} value) nonacetylated distributions. For example, in the nonacetylated distribution in Figure 2B, there is ≥5% occupation from $n = 2$ to $n = 8$ while ≥5% occupation extends from $n = 1$ to $n = 9$ in the comparable acetylated distribution. However, the effect is not large, and we are reluctant to try to interpret it at this time. Multiple peaks/shoulders are still observed in the acetylated distributions; in fact, this feature is often more pronounced in these distributions.

It does appear that in the more highly loaded reconstitutes, cf. $n_{av} > 8.0$ (Figure 1B versus Figure 1D), acetylated arrays are on average less compacted than the corresponding nonacetylated arrays. Previous analyses using sedimentation and EM methods (*41*) have shown that saturated, nonacetylated arrays are more compacted than saturated, acetylated arrays. Thus, our results agree with and extend those conclusions to subsaturated arrays. Compaction is thought to depend on internucleosomal contacts arising from electrostatic interactions involving lysines in the N-terminal histone tails. Acetylation reduces these interactions.

Effect of Glutaraldehyde Fixation on Nucleosome Population Distributions. AFM permits sample deposition under a broad variety of solvent conditions (*15, 17, 39, 40*). We took advantage of this feature to analyze the effects of glutaraldehyde fixation on the composition of these 208-12 nucleosomal arrays. The samples compared were simply different aliquots of the same reconstituted chromatin preparation, split just prior to the fixation step (post-reconstitution), and fixed

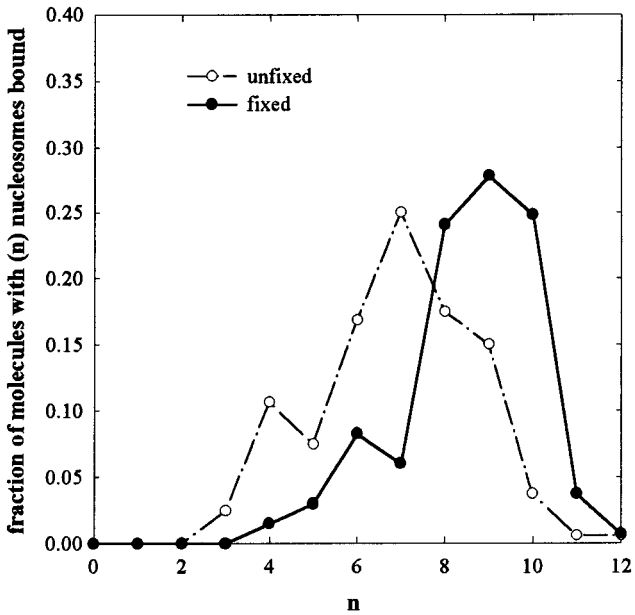

FIGURE 3: Effect of glutaraldehyde fixation on nucleosome population distributions. The fraction of molecules with a particular number of nucleosomes (F_n) is plotted versus n, the number of nucleosomes. The distributions for a glutaraldehyde-fixed sample (filled circles) and an unfixed sample (open circles) from the same (nonacetylated) reconstitute are shown. The fixed sample has an $n_{av} = 8.5$; the unfixed sample has an $n_{av} = 6.9$.

or not fixed. Thus, pipetting errors, histone/DNA ratio differences, and other such artifactual differences between samples can be excluded.

An AFM image of an unfixed sample is shown in Figure 1C. Population distributions comparing the same reconstitute when fixed (filled circles) or unfixed (open circles) are shown in Figure 3. These distributions correspond to the AFM images in Figure 1B,C. Unfixed samples consistently show a lower n_{av} than the corresponding fixed sample. Thus, nucleosomes are apparently somewhat labile when they are not fixed. Since all the AFM operations are carried out at the low ionic strength conditions characteristic of the final

reconstitution step, <1 mM, salt-dependent nucleosome dissociation from DNA cannot explain the lability of the unfixed nucleosomes.

Low-salt nucleosome−DNA dissociation has been previously observed in sedimentation studies of mononucleosomes (*42*) and of 208-12 arrays (*14*). In those cases, the extent of dissociation was ≃5−10%. In this work, the more highly loaded templates (cf. Figure 3) show ≃20% dissociation, based on n_{av} changes. AFM samples are diluted just prior to deposition, and this lowered concentration might enhance the nucleosome dissociation reaction (by mass action). It is also possible that surface forces encountered in the AFM technique could also enhance intrinsic nucleosome lability. However, the same AFM procedure used in this work is gentle enough to permit the visualization of supercoiled DNA, at various extents of supercoiling (*43*), and even to observe highly dynamic local DNA structures such as cruciforms (*44*). Thus, it is not an inherently disruptive procedure.

Unfixed samples also appear to be more highly compacted than fixed samples (cf. Figure 1B versus Figure 1C). These results are consistent with the view that chromatin compaction involves the lysines in the tails of histones because glutaraldehyde fixation is thought to act through lysine groups (*45*). These results thus suggest that the nucleosomes in unfixed chromatin probably interact with one another more extensively than the nucleosomes in fixed chromatin and therefore that the full level of nucleosome−nucleosome interactions may not be realized in chromatin folding and compaction studies which use fixed templates.

DISCUSSION

We have carried out an AFM analysis of 208-12 DNA templates reconstituted into chromatin at various subsaturating histone levels, i.e., containing less than a full complement of 12 nucleosomes. Subsaturated nucleosomal arrays have important natural analogues such as eukaryotic origins of replication and gene promoter regions and thus are of great structural interest. To date, 208-12 nucleosomal arrays have been analyzed primarily by sedimentation and gel electrophoresis (ref *9*), methods which yield mainly average properties for the population of molecules present. With AFM, one can directly visualize individual molecules. By counting large numbers of these, it is possible to determine the precise numerical distribution for the population. These distributions will provide insight on the mechanisms by which nucleosomes arrange themselves on these multisite templates and thus, by inference, on eukaryotic chromosomes.

First, the results demonstrate clearly that discrete populations of nucleosomal arrays are produced at subsaturating histone/DNA ratios. As increasing amounts of histone are put into the reconstitution, the average number of nucleosomes on the template, n_{av}, rises smoothly with no evidence of threshold or plateau effects. This type of behavior was observed in, and thus expected from, previous work with 208-12 arrays (*14*).

The numerical distributions we obtained provide insight on how nucleosomes populate these multisite templates. At $n_{av} < 4.0$, the distributions typically contain a single peak whose breadth resembles that predicted for a random process (Figure 2A). Thus, nucleosome loading at these lower levels appears to follow a random process. At $n_{av} > 4.0$, the distributions become broader and have significant complexity, i.e., secondary peaks and/or shoulders (Figure 2B). Thus, the character of the distributions changes around $n_{av} = 4$.

On these 12-site templates, one would intuitively expect a transition in the nucleosome loading process to occur somewhere around $n_{av} = 6$. 208-12 templates with fewer than six nucleosomes will contain more unoccupied 208 bp sites than sites to which nucleosomes are already bound. Thus, an incoming nucleosome may not have to bind in a site that is next to an already occupied site. In templates containing more than six nucleosomes, unoccupied sites are fewer in number than occupied sites, and incoming nucleosomes become increasingly likely to have to bind next to another nucleosome. The changes we note at $n_{av} > 4.0$ probably reflect the onset of this transition from unoccupied binding sites in excess to already occupied sites in excess. That the transition begins well below $n_{av} = 6$ may be due to nonrandom behavior in the nucleosome loading process on these templates (see below). Of course, the binding options available to an incoming nucleosome will depend on precisely where the occupying nucleosomes actually lie within a given template. We are currently analyzing these images to determine this interesting and complementary type of information (Yodh et al., in preparation).

As noted under Results, theoretical analysis predicts that the distributions should again become narrower (and less complex) at high n_{av}, e.g., $n_{av} > 8$, like the distributions at $n_{av} < 4$. We see suggestions of such behavior (cf. Figure 2C). However, nucleosome counting on the more highly populated templates is difficult. Thus, we have not systematically analyzed them and cannot verify this prediction.

The feature of the distributions that we find most surprising is the presence of multiple peaks/shoulders in distributions with $4.0 < n_{av} < 8.0$. These peaks/shoulders usually occur at every other *n*, either as an even series such as 2/4/6 (Figure 2B) or as an odd series such as 3/5/7 (not shown). Such discrete patterns of variation within the population distribution were completely unexpected. We feel that these variations are a real feature of the nucleosomal array population because we always observe them; every experimental distribution obtained in this n_{av} range shows multiple peaks/shoulders. The occurrence of peaks separated by 3 nucleosomes, as in Figure 2B, is much less common.

To test further the validity of the discrete variation pattern within the distributions, we tabulated the number of molecules found at the major peak value, designated as "*p*", and at values $p \pm 1$, $p \pm 2$, $p \pm 3$, i.e., +1, +2, or +3 nucleosomes more and −1, −2, or −3 nucleosomes fewer than the major peak value, in each of our reconstitutes in the range $4 < n_{av} < 8$. We combined these data together in a single plot and compared this experimental curve to the curve expected from a random process (Figure 4). This approach provides a test of whether the discrete variation pattern is a general feature of our results, in a combined data set with a greatly enhanced statistical significance. The results clearly show that templates which contain ±1 nucleosome from the major peak value are present at a lower than random frequency. Templates containing 2 nucleosomes fewer than the major peak value are present at a frequency that is greater than the frequency expected for a random process, while

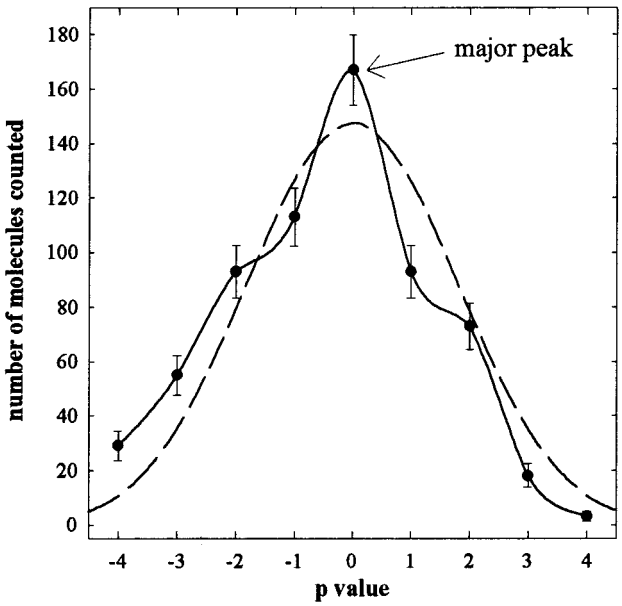

FIGURE 4: Combined analysis of complexity in the nucleosome population distributions. For distributions in the range $n_{av} = 4.0-8.0$, the number of molecules with the major peak value of nucleosomes (called "*p*") and the number containing $p \pm 1$, $p \pm 2$, and $p \pm 3$ nucleosomes were determined, combined into a single data set, and plotted (filled circles, solid line). This experimental curve is compared to the curve expected from a random loading process (dashed line) with the same total number of molecules counted and the same weighted n_{av} values. The experimental and theoretical peak positions are lined up at the major peak value (labeled 0). "*p* value" refers to the number of nucleosomes more (+) or less (−) than the major peak value.

templates with 2 nucleosomes more than the major peak value are present at approximately the random frequency (Figure 4). Thus, this analysis demonstrates that there is a preference for particular numbers of nucleosomes, namely, a strong dispreference for templates that contain ±1 nucleosome from the major peak value and a weaker preference for templates that contain ±2 nucleosomes from the peak, and confirms our consistent observations of discrete variations within individual distributions.

The absence of a clear preference for +2 nucleosomes in the experimental curve of Figure 4 may be due to a "masking" effect arising from proximity to the upper boundary condition, $n_{av} = 12$. The data set contains distributions with n_{av} around 7. In those reconstitutes, templates with 2 nucleosomes more than the peak will contain 9 or more nucleosomes. We have observed clear effects due to proximity to the boundary conditions ($n_{av} = 0$, $n_{av} = 12$) in both low ($n_{av} < 4$) and high ($n_{av} > 8$) distributions (see above). Therefore, the high-*n* side of the experimental curve in Figure 4 may be suppressed due to proximity to the upper boundary condition, thus masking a +2 nucleosome preference. For example, this entire side of the curve appears skewed toward lower frequencies compared to the left (low-*n*) side (Figure 4).

The preference for particular numbers of nucleosomes indicates that there is some nonrandom behavior in the nucleosome occupation of this template. This conclusion does not contradict the apparently random loading observed at $n_{av} < 4.0$ because in the higher levels of loading (1) the frequency of possible neighboring nucleosomes, and thus the potential for internucleosomal contact, is increasing and (2) the electrostatic balance of the templates is changing. The form of this variation, a dispreference for ±1/preference for ±2 nucleosomes, suggests some kind of pairwise feature. These data for a pairwise bias are consistent with in vivo experimental results (*46*, *47*) and in vivo-based model building studies (*48*) which suggest that nucleosome pairs are a basic unit of chromatin organization.

Perhaps the most obvious possibility for the source of this nonrandom behavior is the occurrence of closed-pair nucleosome−nucleosome attractive interactions. If such attractive interactions were strong enough, one would only see templates with even numbers of nucleosomes, i.e., 2/4/6, in the distribution. Thus, if attractive interactions are the explanation for the discrete variation patterns we observe, then the presence of both odd (3/5/7) and even (2/4/6) series suggests that these interactions are not strong enough to restrict the loading to even numbers but only strong enough to bias the reconstitution process against molecules containing ±1 nucleosome and toward molecules containing ±2 nucleosomes from the major peak value, which is itself set by the precise histone input. This explanation seems plausible. However, we emphasize that it is only speculative. For example, the discrete variation pattern is perhaps even more prominent in distributions of acetylated arrays. Acetylation is thought to weaken the degree of internucleosomal contact (*41*, *49*, *50*) and thus would be expected to decrease the peak/shoulder multiplicity in the distributions, if internucleosomal contacts were responsible for this multiplicity. Any feature which establishes a bias toward pairs and away from unpaired singles would explain the data. For example, a tendency for nucleosomes to assemble in particular (pairwise) orientations on chromatin, for DNA torsional constraints or other reasons, could provide such a bias. We are currently carrying out experiments to discover how we can affect these features, in hope of finding the cause of this pairwise preference.

The other intriguing result in Figure 4 is that a greater than random fraction of molecules contain exactly the peak number of nucleosomes; i.e., the experimental distribution is more populated at its major peak than a random process would be. This feature is also observed in individual distributions (cf. Figure 2B). The data in Figure 4 suggest that the true form of these distributions is multimodal, with a rather sharp major peak and secondary peaks at ±2. The increased population at the major peak value suggests some kind of correlation or coherence in nucleosome loading; as templates load up with nucleosomes, the nucleosome numbers on the various molecules stay within a somewhat narrower than random range centered at the major peak value. We do not know what could cause this type of behavior. One possibility could be electrostatic effects of charge neutralization as nucleosomes load on the templates. We are conducting experiments to test this hypothesis further. We note that such coherency in loading might be useful in vivo, to help ensure that both daughter chromosomes become nucleosome-covered at similar rates. In addition to the obvious benefits of keeping daughter chromosome synthesis proceeding synchronously, maintaining coherency might be important for correct maturation of the replicated fiber, for proper chromosome organization and dynamics, and for timely assembly of post-replication nucleoprotein complexes such as transcription factors or structural proteins. In vivo

nucleosome assembly on DNA utilizes protein-dependent mechanisms (*51*) and thus differs from the salt-dependent assembly used in our in vitro work. However, coherence in loading is likely to be an inherent property of the histone/DNA complex itself and thus independent of the precise nucleosome delivery mechanism. For example, the nucleosome assembly process itself follows precisely the same pathway in vitro as it does in vivo (*52*): H3−H4 tetramers on DNA first, followed by H2A−H2B dimers.

ACKNOWLEDGMENT

We thank Jeff Hansen for the 208-12 DNA plasmid, Robert Kingston for support during the histone isolation phase of this work, and Stuart Lindsay and Sanford Leuba for helpful conversations.

REFERENCES

1. Van Holde, K. E. (1989) in *Chromatin*, pp 293−301, Springer-Verlag, Inc., New York.
2. Lohr, D., and Torchia, T. (1988) *Biochemistry 27*, 3961−3965.
3. Sogo, J. M., Stahl, H., Koller, T., and Knippers, R. (1986) *J. Mol. Biol. 189*, 189−206.
4. Lohr, D. (1984) *Nucleic Acids Res. 12*, 8457−8479.
5. Lohr, D. (1997) *J. Biol. Chem. 272*, 26795−26798.
6. Simpson, R. T., Thoma, F., and Brubaker, J. M. (1985) *Cell 42*, 799−808.
7. Dong, F., Hansen, J. C., and van Holde, K. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5724−5728.
8. Meersseman, G., Pennings, S., and Bradbury, E. M. (1991) *J. Mol. Biol. 220*, 89−100.
9. Fletcher, T. M., and Hansen, J. C. (1996) *Crit. Rev. Eukaryotic Gene Expression 6*, 149−188.
10. Bazett-Jones, D. P., Cote, J., Landel, C. C., Peterson, C. L., and Workman, J. L. (1999) *Mol. Cell. Biol. 19*, 1470−1478.
11. Logie, C., and Peterson, C. L. (1997) *EMBO J. 16*, 6772−6782.
12. Logie, C., Tse, C., Hansen, J. C., and Peterson, C. L. (1999) *Biochemistry 38*, 2514−2522.
13. Steger, D. J., Eberharter, A., John, S., Grant, P. A., and Workman, J. L. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 12924−12929.
14. Hansen, J. C., and Lohr, D. E. (1993) *J. Biol. Chem. 268*, 5840−5848.
15. Lyubchenko, Y. L., Gall, A. A., Shlyakhtenko, L. S., Harrington, R. E., Jacobs, B. L., Oden, P. I., and Lindsay, S. M. (1992a) *J. Biomol. Struct. Dyn. 10*, 589−606.
16. Lyubchenko, Y. L., Jacobs, B. L., and Lindsay, S. M. (1992b) *Nucleic Acids Res. 20*, 3983−3986.
17. Lyubchenko, Y. L., Jacobs, B. L., Lindsay S. M., and Stasiak A. (1995) *Scanning Microsc. 9*, 705−727.
18. Allen, M. J., Dong, X. F., O'Neill, T. E., Yau, P., Kowalczykowslci, S. C., Gatewood, J., Balhorn, R., and Bradbury, E. M. (1993) *Biochemistry 32*, 8390−8396.
19. Zlatanova, J., Leuba, S. H., and Van Holde, K. (1998) *Biophys. J. 74*, 2554−2566.
20. Turner, B. M. (1991) *J. Cell Sci. 99*, 12−20.
21. Garcia-Ramirez, M., Dong, F., and Ausio, J. (1992) *J. Biol. Chem. 267*, 19587−19595.
22. Krajewski, W. A., and Ausio, J (1996) *Biochem. J. 316*, 395−400.
23. Fletcher, T. M., and Hansen, J. C. (1995) *J. Biol. Chem. 270*, 25359−25362.
24. Tse, C., and Hansen, J. C. (1997) *Biochemistry 36*, 11381−11388.
25. Schwarz, P. M., Felthauser, A., Fletcher, T. M., and Hansen, J. C. (1996) *Biochemistry 35*, 4009−4015.
26. Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988) *EMBO J. 7*, 1395−1402.
27. Jeppeson, P., and Turner, B. M. (1993) *Cell 74*, 281−289.
28. Braunstein, M., Rose, A. B., Holmes, S. G., Allis, C. D., and Broach, J. R. (1993) *Genes Dev. 7*, 592−604.
29. Howe, L., Ranalli, T. A., Allis, C. D., and Ausio, J. (1998) *J. Biol. Chem. 273*, 20693−20696.
30. Brownell, J., and Allis, D. (1996) *Curr. Opin. Genet. Dev. 16*, 1176−1184.
31. Georgel, P., Demeler, B., Terpening, C., Paule, M. R., and van Holde, K. E. (1993) *J. Biol. Chem. 268*, 1947−1954.
32. Candido, E. P. M., Reeves, R., and Davie, J. R. (1978) *Cell 14*, 105−113.
33. Sealy, L., and Chalkley, R. (1978) *Cell 14*, 115−121.
34. Vidali, G., Boffa, L. C., Bradbury, E. M., and Allfrey, V. G. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2239−2243.
35. Workman, J. L., Taylor, I. C. A., and Kingston, R. E. (1991) *Methods Cell Biol. 35*, 429−430.
36. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
37. Zweilder, A. (1978) *Methods Cell Biol. 17*, 223−233.
38. Hardison, R., and Chalkley, R. (1978) *Methods Cell Biol. 17*, 235−251.
39. Lyubchenko, Y. L., Blankenship, R. E., Lindsay, S. M., Simpson, L., and Shlyakhtenko, L. S. (1996) *Scanning Microsc., Suppl. 10*, 97−109.
40. Lyubchenko, Y. L., and Lindsay, S. M. (1998) in *Procedures in Scanning Probe Microscopy* (Colton, R., Engel, A., Frommer, J., Gaub, H., Gewirth, A., Guckenberger, R., Rabe, J., Heckl, W., and Parkinson, B., Eds.) pp 493−496, J. Wiley & Sons, Ltd., Rochester, New York, Weinheim, Brisbane, Singapore, Toronto.
41. Garcia-Ramirez, M., Rocchini, C., and Ausio, J. (1995) *J. Biol. Chem. 270*, 17923−17928.
42. Yager, T. D., and Van Holde, K. E. (1984) *J. Biol. Chem. 259*, 4212−4222.
43. Lyubchenko, Y. L., and Shlyakhtenko, L. S. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 496−501.
44. Shlyakhtenko, L. S., Potaman, V. N., Sinden, R. R., and Lyubchenko, Y. L. (1998) *J. Mol. Biol. 280*, 61−72.
45. Thoma, F., Koller, T., and Klug, A. (1979) *J. Cell Biol. 83*, 403−427.
46. Burgoyne, L. A., and Skinner, J. D. (1981) *Biochem. Biophys. Res. Commun. 99*, 893−899.
47. Lohr, D., and Van Holde, K. E. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6326−6330.
48. Woodcock, C. L., Frado, L. L., and Rattner, J. B. (1984) *J. Cell Biol 99*, 42−52.
49. Tse, C., Sera, T., Woffei, A. P., and Hansen, J. C. (1998) *Mol. Cell. Biol. 18*, 4629−4638.
50. Walia, H., Chen, H. Y., Sun, J., Holth, L. T., and Davie, J. R. (1998) *J. Biol. Chem. 273*, 14516−14522.
51. Wolffe, A. P. (1994) in *Regulation of Chromatin Structure & Function*, pp 59−67, R. G. Landers Company, Georgetown, TX.
52. Hansen, J. C., van Holde, K. E., and Lohr, D. E. (1991) *J. Biol. Chem. 266*, 4276−4282.

BI991034Q