Structural features of a phased nucleosome core particle
(chromatin/DNase I/5S rRNA gene)

ROBERT T. SIMPSON* AND DARRELL W. STAFFORD†

*Laboratory of Cellular and Developmental Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20205; and †Department of Zoology, University of North Carolina, Chapel Hill, North Carolina 27594

Communication by Gary Felsenfeld, October 1, 1982

ABSTRACT Chicken erythrocyte inner histones associate with a cloned 260-base-pair (bp) segment of Lytechinus variegatus DNA in a unique location. The fragment contains a 120-bp segment encoding 5S rRNA, a 90-bp flanking sequence to the 5′ side of the transcribed segment, and a 50-bp downstream flanking sequence. Association of DNA, uniquely labeled at one end or the other and at either the 3′ or the 5′ terminus of a given strand, with histones at 0.1 M ionic strength leads to formation of a compact complex which sediments at about 13 S. Analysis of cutting of the complex by DNase I shows that protection from the nuclease is confined to a region beginning 20 bp from the left end of the segment and extending to about 165 bp from the left end. Within the protected region, the two DNA strands differ in their susceptibilities to the nuclease, the precise location of nucleosome cutting sites and the spacing between these sites, and the relative susceptibilities of specific cutting locations. It seems that information present in DNA and the histone octamer is sufficient to create a precisely phased nucleosome in which interactions of the two DNA strands with histones are not the same. The structure of this unique nucleosome is not predicted by the intellectual model based on studies of mixed populations of nucleosome core particles.

Phasing of nucleosomes in vivo has been suggested for a number of gene sequences in various eukaryotic and viral systems (for review, see ref. 1). The mechanism whereby such defined relationships of histone octomers to specific DNA sequences are established has remained obscure; in addition to the possibility of specific histone-DNA interactions per se, phasing could arise from interactions of nonhistone proteins with DNA or as a consequence of the mechanism and specific origin of replication of a chromatin segment. In vitro association of histones with simian virus 40 DNA suggested that certain sites might preferentially interact with histones to form nucleosomes (2); other experiments indicated that sites that are presumed to be free of nucleosomes in vivo (3, 4) can associate with histones in vitro (5). In experiments using short (140 and 200 base pair) segments of DNA containing the lac operator, Chao et al. (6) showed that several defined relationships of histones and DNA might occur; a totally random association of protein and DNA did not occur. These experiments were limited by the short length of the DNA used (so that end effects may have contributed to the results) and by lack of controls in which DNA alone was digested (they anticipated the currently recognized sequence selectivity of DNase I).

We have examined the question of in vitro phasing arising from histone-DNA interactions alone by using a 260-bp segment of DNA containing a 5S rRNA gene of Lytechinus variegatus (7).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

EXPERIMENTAL SECTION

Plasmid pLV405 contains the 248-bp Mbo II fragment of pLV103 (7) subcloned in the EcoRI site of pACYC184. DNA was isolated from cleared lysates by isopycnic banding in CsCl gradients containing ethidium bromide. The cloned 5S segment was excised with EcoRI and purified by gel electrophoresis and electrolution. 5′-End labeling with [γ-32P]ATP (New England Nuclear) and polynucleotide kinase (New England BioLabs) was carried out as described by Maxam and Gilbert (8). 3′-End labeling was performed by filling in the single-stranded EcoRI tails by using [α-32P]dATP (Amersham) and large fragment DNA polymerase I (New England BioLabs), followed by a chase with excess unlabeled dATP and TTP. Secondary restriction of the labeled fragments with Xmn I or Mnl I (New England BioLabs) was followed by fragment separation on small columns of Sephacryl S-300.

Chicken erythrocyte core particles were isolated as described (9) and sedimented through sucrose gradients containing 0.5 M NaCl to remove any residual H1/H5 and nonhistones. Histones and DNA were separated by hydroxyapatite chromatography with 2.5 M NaCl/50 mM sodium phosphate used to elute histones and 0.6 M potassium phosphate at pH 6.0 used to elute the DNA. The four small histones were intact and in the proper stoichiometric ratio. Histones and DNA were dialyzed against 0.1 M NaCl/10 mM Tris-HCl, pH 8.0/1 mM EDTA/1.4 mM 2-mercaptoethanol and stored at 4°C.

Association of histones and DNA was carried out in DNA excess to minimize the possibility of association of more than one octomer with the labeled DNA fragment; a ratio of histone octomer to core particle DNA of 0.7 times the stoichiometric amount typically was used. The labeled sea urchin DNA was added in trace amounts to the association. The histones were incubated at least overnight at 4°C with a 10-fold mass excess of Tris poly(l-glutamate). DNA plus the trace labeled DNA and histones were brought to 37°C, histones were added to the DNA, and the samples were incubated for 2 hr and then cooled to 0°C (10). The histone-DNA complex was separated from unassembled DNA by sedimentation. The peak was localized by scintillation counting of aliquots of fractions and pooled fractions were dialyzed against 0.1 M NaCl/10 mM Tris-HCl, pH 8.0/1 mM EDTA.

DNase I ( Worthington) digestions were carried out at 37°C by using 5 μg of DNA or complex in the above buffer adjusted to contain 10 mM MgCl₂ and 3 mM CaCl₂. Reactions were terminated by addition of sodium dodecyl sulfate and EDTA to concentrations of 0.5% and 25 mM, respectively. DNA was purified, dissolved in 80% (vol/vol) deionized formamide/10 mM NaOH/1 mM EDTA, heated to 90°C for 1 min, and analyzed on 0.4-mm-thick polyacrylamide gels containing 8.3 M urea (8). Size standards were end-labeled DNA cut at guanylic residues by dimethyl sulfate (8).
RESULTS

Fig. 1 presents the sequence (7) of the cloned DNA fragment used for these studies. The start site for transcription of 5S rRNA is at base 91; termination is at base 210. The bottom (C) strand as presented is the coding strand. The fragment is highly appropriate for the study because: (i) its length is sufficient that end effects on possible nucleosome phasing should be small, but, at the same time, short enough that formation of two nucleosomes is unlikely and (ii) a single restriction site for Xmn I (base 10) and closely spaced sites for Mnl I (bases 229 and 249) are located near the ends of the fragment, allowing facile preparation of uniquely end-labeled probes. The only obvious unusual feature of the DNA sequence is the long run of thymidine followed by an even longer run of cytidine in the top (noncoding) (W) strand at positions 208–229.

The fragment associates with histones to form a defined complex, as detected by its sedimentation at about 13 S, compared to the protein-free DNA which sediments at about 6 S or unfolded histone–DNA complexes, exemplified by core particles in high concentrations of urea, at about 6–7 S (11). The breadth of the sedimentation peak is closely similar to that of chicken erythrocyte core particles sedimented under identical conditions.

Fig. 2 shows the results, for both protein-free DNA and the histone–DNA complex, of mapping DNase I cutting sites in the W strand with the labeled site at the 5′ (left) end of the fragment. The "DNA alone" controls show some evidence of sequence selectivity of cutting. Approximately equivalent frequencies of cutting are observed for DNA and the complex up to about 20 bases from the left hand end (data not shown). At this point, obvious differences between the two sets of digestion patterns emerge. Most striking are areas where there is virtually no cutting of the complex, spaced at approximately 10-base intervals. Between these areas are similarly spaced regions where the complex is cut strongly by the nuclease. In some cases these correspond to sites of frequent cutting in the DNA alone (near 40, 70, 80, 90, and 100) but in other cases a weakly cut site in the DNA control is more strongly cut in the complex (near 30, 60, 110, 120, and 130). Clearly, interactions with histones have made the nuclease cut the complex preferentially at a site which is not favored on the basis of DNA sequence. Cutting at 10-base intervals is highly suggestive of a specific interaction of DNA with the histones, but the most telling evidence for such interaction is the occurrence of sites where cutting is strong for DNA

```
AATCCAGAATACCTGAGGATTTATATACGCCGAGCTTAACATACCTCCAG
GGTGGTATTATTGAAAGTCTCCTAAATATTCGGCTACTGGTATTGGGAAG
TTAAATACGTGAACCTCTCTCTCTACCAAGGATATCATTTACTCTGCT
AAATTTATAGGATAGTACGTCTCCTGGATGATCTACGATTTATGGCC
GTCCTGAGTCTAGCTGAGCTCCAGGGATTTATAAGCCGATGACGTCATAACATCCCTGACCC
GGCGGACCCTTATGGCCCACAACATCCGAAAAAAAAGAGGGGGGGGGGGGAGAAACGAAG
GGTTGCTTATTGAAGGTCCCTAAATATTCGGCTACTGGTATTGGGAAG
```

Fig. 1. Sequence of the cloned L. variegatus 5S rRNA gene segment.

![Fig. 2. Cutting sites for DNase I in the W strand mapped from the 5′ (left) end. The EcoRI 5S fragment of pLV405 was 5′-end labeled with [32P]ATP and polynucleotide kinase and secondarily restricted with Mnl I and the large fragment either was allowed to associate with histones (nucleosome) or was examined alone as a control (DNA). Digestions with DNase I were performed at the indicated concentrations of nuclease (unit/ml) for 6 sec (DNA) or 2 min (nucleosome). Lanes labeled G are the same labeled DNA after chemical cutting at guanyl residues with dimethyl sulfate.]

alone but does not occur in the histone–DNA complex (near bases 44, 63, 83, 92, and 116, for examples). Such sites must result from a highly selective shielding of the potential cutting site by proteins in the complex. Cutting at 10-base intervals is observed up to 130 bases from the left end. In the sample shown, the next potential site is obscured by chemical cuts occurring during storage in alkaline sample buffer. In other experiments (as in Fig. 4), cutting at 10-base intervals continued to near position 150.

Further indication that the restriction to cutting by DNase I observed for the histone complex reflects a phase core particle derives from the relative susceptibilities of the various cutting sites. The rates of appearance of labeled fragments cut near 20, 30, and 40 bases from the left end are high, suggesting high accessibility of these loci. In contrast, at the 50-base site, the rate of cutting is nearly as low as in the intervening, inaccessible regions. Near 60 and 70 bases from the left end, cutting is again strong. In the succeeding 30-base region, there is clear evidence of cutting at 10-base intervals, near 80, 90, and 100 bases from the left end, but the relative rate of cutting is quite low. Following this region, more active cutting occurs at the sites about 110, 120, and to a lesser extent, 130 bases from the left end of the W strand. This modification of cutting site susceptibility is close to that expected from numerous studies of cutting native and semisynthetic core particles with DNase I (12–16). Thus, if a left-hand border of the nucleosome core particle is at 20 bases from the left end of the DNA, one sees a resistant site 30
bases from the border and a resistant region 60–80 bases from the border, followed by a somewhat resistant site 110 bases from the left 5' end of the core particle DNA segment.

For better examination of the right-hand border of the particle, samples uniquely labeled at the right-hand (5') end of the C strand were analyzed after DNase I digestion of protein-free DNA and the DNA–histone complex. The two samples showed similar cutting site distributions and susceptibilities from the right-hand end of the fragment back to about bases 165–168. Here, an infrequently cut site in the DNA control was cut rather strongly in the complex. Strongly cut sites occurred at about 140 and 150 bases in the C strand, although the results are complicated by the 150-base site being highly susceptible in the DNA alone. Although not as clear as the data in Fig. 2, these results suggested that the right-hand end of the nucleosome core particle is at about 160 bases, as expected from the left-hand border being near base 20 and the known length of the core particle DNA segment.

To analyze cutting sites within the nucleosome for the C strand, we labeled this strand at the 3' end, cut the complex and DNA alone with DNase I, and localized sites as described above (Fig. 3). Again, a striking periodicity of cutting of the complex, which was absent in the controls, is observed; the overall periodicity of both noncut and cut regions is approximately 10 bases, as was the case for the W strand. Cutting at sites less than 20 bases from the left end is similar for DNA and the complex (data not shown). Cuts at bases 29 and 37 are strong relative to nearby sites. A different situation ensues at the next two cutting sites where four (bases 45–48) or three (bases 56–58) bonds are cut with roughly equal frequency. In the central region of the core particle, bases 68–129, cutting appears to occur at one or two bases at each site. Although obfuscated by more nonspecific cleavages, the pattern of cutting at about 10-base intervals persists up to at least 149 bases from the left end. A strongly cut site at 165–168 bases in the complex, absent in the DNA alone, is the last definitive difference between the two species; from here to the right-hand end of the DNA, similar patterns are observed for the two samples when gels are run longer to display this region better.

A striking difference in the behavior of the two strands is apparent on comparison of Figs. 2 and 3. Although the modulation of cutting site susceptibilities in the W strand is roughly that expected from studies of random sequence DNA core particles labeled at both 5' termini, the availability of the various sites for cutting by DNase I in the C strand differs markedly from that in the W strand. Thus, site 4, frequently cut in the W strand, is highly inaccessible in the C strand, being cut even less often than site 3, the first resistant site in the canonical core particle. This may be due in part to sequence selectivity of the nuclease, but this seems less likely on inspection of cutting rates for sites at 78, 98, and 109 bases; these are cut strongly in the complex and infrequently if at all in DNA alone. Most of the cutting in the C strand occurs in the center of the nucleosome, at sites 5–9; this includes region 6–8 which is highly resistant to nuclease in the usual core particle cutting site maps. Similar to site 4, the mirror-image site 10 is also relatively resistant to DNase I, being cut with about the same frequency as the canonical resistant site 11.

Temporarily setting aside the role of DNA sequence selectivity of the nuclease in digestion of DNA in a core particle, we
compare directly the digestion patterns of the W and C strands in the complex at four levels of digestion (Fig. 4). In addition to features noted above about the differences in cutting susceptibilities of the two strands, several other features of the digestion patterns are apparent here. (i) There is an apparent polarity in either strand to more frequent cutting toward the 5' end of the strand; thus, in this experiment, the shorter W fragments predominate while, to a lesser extent, the longer C fragments exceed the shorter at comparable extents of digestion. (ii) The rates of cutting of the two strands differ; a good guess is that the rate of cutting of the C strand is about 1/4 to 1/2 that for the W strand. Such marked differences are not observed for the two strands as protein-free DNA. (iii) The stagger of the cutting sites at the different loci in the nucleosome is not constant. Thus, at sites 2, 4, 5, 6, 7, and 8, there is a stagger of about two or three bases to the 3' side for the averaged center of the cutting sites of the C strand compared to the W strand. At sites 9–13, the stagger is about 0.

Fig. 5 presents scans of autoradiograms of DNase I digests of complexes labeled at the 5' end of the W strand or the 3' end of the C strand, digested with DNase I at 1 unit/ml.

DISCUSSION

Phasing of a Nucleosome Assembled in Vitro. The data presented demonstrate that information present in the histone octomer and an unique DNA sequence are sufficient to allow their precise interaction, creating an accurately phased nucleosome in the absence of DNA replication or nonhistone proteins. The sequence features necessary for phasing the nucleosome cannot be determined from a single example. Dot matrix analysis (17) of the sequence has been carried out to ascertain the presence of direct and inverted repeats as well as palindromes. Base pairs 37–46 are a 10-base-pair palindrome, the only striking one in the core particle. Direct repeats (five of seven bases agreeing) are present at DNase I cutting sites 2 and 12, 4 and 10, and 5 and 9. The significance of these is unknown; their occurrence at mirror image sites from the center of the nucleosome makes them of some interest. Finally, the sequence C-A-T occurs in the W strand near nuclease cutting sites 3, 6, and 8—three of the four DNase I-resistant sites in the canonical core particle. Comparison with sequences of other precisely phased core particles and determination of the critical features of this sequence for phasing by in vitro genetic manipulation should allow a gradual dissection of the features necessary and sufficient for unique histone–DNA interactions.

It is not known whether the same phasing of a nucleosome on the L. variegatus 5S gene occurs in vivo. The positioning determined here does agree with one of two phasing arrangements described for Drosophila 55 genes by Louis et al. (18); in contrast, it does not agree with any of four possible phasing arrangements deduced for the Xenopus 5S gene by Gottesfeld and Bloomer (19). If formation of a nucleosome blocks transcription of the involved DNA, this nucleosome is fortunately positioned on the gene sequence. The start site for transcription is the precise center of the nucleosome. The right-hand border of the nucleosome comes very close to the right-hand side of the region bound by the stimulatory factor (TfIIIA) necessary for transcription of the 5S gene in vitro (20, 21); the site of interaction of the factor (in Xenopus) corresponds to bases 140–145 to 170–172 (22). A recent report by Gottesfeld and Bloomer (23) has shown that presence of TfIIIA during in vitro chromatin assembly of a plasmid bearing a Xenopus 5S gene leads to a chromosomal structure on which 5S transcription can occur. It will be of interest to determine whether the phased nucleosome detected in the current study is assembled at the same site in the presence of the transcriptional activator.

Core Particle Structure. In the present study, a nucleosome core particle has been probed by a nuclease by using a fragment labeled at only one end and the cutting of both DNA strands has been examined separately. The results obtained are thus relevant to cutting site maps for various nucleosomes which have been reported in the 6 years since this approach to core particle structure was first used. However, difficulties arise in attempts to make such a comparison. The current data were obtained with a uniquely positioned particle; therefore, cutting site locations and relative susceptibilities may reflect the contributions of DNA sequence selectivity of the nuclease in addition to constraints imposed on nuclease activity due to interactions of DNA with the histones. Conversely, in "averaged" core particles containing (presumably) completely random sequences, (i) experiments are of necessity performed with symmetrically labeled DNA, leading to results including contributions of both strands, and (ii) length heterogeneity in the core particle DNA can obfuscate precise localization of cutting sites.

A critical question thus is how much the sequence selectivity of the nuclease influences the frequency of cutting DNA at the various sites. Analysis of the dinucleotide pairs where cutting occurs in the nucleosome shows that 15 of the 16 possible pairs are cut at least once. Inspection of the data in Figs. 2 and 3 reveals a number of examples in which weakly cut sites in DNA alone are strongly cut sites in the nucleosome. We thus think it possible that many of the features in the cutting site maps for this nucleosome derive from specific features of interactions of the nucleic acid with the histone octomer. The possibility remains that the sequence selectivity of the nuclease is altered when DNA is bound on a nucleosome surface in such a fashion as to lead to the current results. Clearly, a definitive answer to the question of sequence selectivity can only come from studies of several phased nucleosomes that differ in sequence.

The most striking difference between the current study and those with core particles labeled at both ends derives from the ability here to assess cutting in the two DNA strands indivi-
ually. It is quite clear that the susceptibilities of cutting at the various sites is different for the two strands. Susceptibilities in the W strand are reminiscent of those observed for averaged core particles, with moderately resistant sites at 3, 6, and 11 and a highly resistant site at 8, although, in addition, the rate at site 7 is also very low. In contrast, in the C strand, much of the cutting occurs in the central region at sites 5–9, sites 4 and 10 are highly resistant, and sites 1 and 2 appear to be cut with relatively low frequency. In other studies with particles labeled at both ends, the results obviously include (perhaps in nonequal proportions) contributions from both strand patterns.

The two strands also differ in the spacing (and consequently the stagger) of the nuclease cutting sites. A great deal of experimental effort has gone into relating the spacing of cutting sites for DNA in the nucleosome to the helical repeat of DNA in chromatin and in solution (13–16, 24); this has arisen from considerations involving the "linking number paradox" (25–27). In the W strand, the average position of centers of cutting sites from site 2 to site 12 is about 10N, although the actual distances between sites vary from 8 to 12 bases. In the C strand, the position of cutting from sites 2 to 13 starts at 7 + 10N but gradually increases from site 4–10 by 2 base pairs, to 9 + 10N. Others have observed even larger changes in spacing in the central region of the core particle DNA in double end-labeled experiments (14, 16, 24). The stagger of cutting sites changes in the region where the spacing increases in the C strand (Fig. 5). Studies by others have suggested an average stagger of cutting sites for DNase I of two bases, 3' extended (14, 25). This stagger is found here for the major portion of the nucleosome.

Although it is tempting to suggest that the cutting site spacing of near 10 for the large part of this nucleosome does indeed reflect a DNA helical repeat of 10.0 base pairs in the core particle and that other, undetermined, features of nucleosome structure lead to the variations in actual spacing of cutting sites, the differences between the two strands seem to require a more conservative interpretation—that spacing of nuclease cutting sites in chromatin is not solely a result of the helical repeat of DNA when bound in a nucleosome. When considered in the context of a smoothly wound DNA helix, as in the current structural model for the nucleosome (27, 29), arguments that attempt to equate modulation of cutting site susceptibility (30) or spacing of cutting sites for nucleases (25) with the angle of exposure of the susceptible phosphate ester bond are difficult to reconcile with the current data.

The current model for the structure of the nucleosome core particle (27, 29) includes a dyad symmetry axis for the histone octomer and the wrapped DNA. A true dyad is clearly absent in any unique, nonpalindromic DNA. In this uniquely positioned nucleosome, the cutting site susceptibilities for the two strands suggest the absence of a pseudodyad also. Some elements of symmetry are apparent in the susceptibilities of cutting at various sites: 6 > 8 in the W strand and 8 > 6 in the C strand; resistant sites at 3 and 11 in the W strand and at 3, 4, and 10, 11 in the C strand. Other features are more consistent with a strictly polar arrangement of types of nuclease cutting, partic-ularly the occurrence of sites that are cut at multiple loci with about equal intensity (e.g., 3 and 4 C) or at one major locus (e.g., 10 and 11 C). Clearly, at least one other phased nucleosome, or derivatives of this one, must be examined to allow conclusions as to what symmetry elements are consistently present in the distribution of DNase cutting sites of core particles.

Models derived from studies of mixed populations of nucleosomes do not necessarily apply to any unique nucleosome. The data presented here for cutting site locations and susceptibilities in a uniquely positioned nucleosome would not have been predicted by the current conceptualization of core particle structure shared by most workers involved in structural studies of this chromosomal element.

References