# DNA Sequence Directs Placement of Histone Cores on Restriction Fragments during Nucleosome Formation†

Moses V. Chao,* Jay Gralla, and Harold G. Martinson

ABSTRACT: Restriction fragments, 203 and 144 base pairs in length, bearing the *Escherichia coli lac* control region have been reconstituted with the core histones from calf thymus to form nucleosomes. By several criteria the reconstituted nucleosomes are similar to native nucleosomes obtained by micrococcal nuclease digestion of calf thymus nuclei. However, sensitive nuclease digestion studies reveal subtle and important differences between native monosomes and the *lac* reconstitutes. Each reconstitute consists mainly of nucleosomes containing histone cores placed nonrandomly with respect to the DNA sequence. The shorter reconstitute forms asymmetric nucleosomes as evidenced by the DNase I digestion pattern. Exonuclease III digestion followed by 5′-end analysis of the larger reconstitute suggests that, of the many possible arrangements of histone core with DNA sequence, only two are highly favored.

An important question regarding the role of chromatin structure in gene expression concerns the accessibility of specific DNA sequences for transcription and replication. Although overt sequence specificity of histones for DNA has been ruled out as a plausible transcriptional control mechanism (Axel et al., 1974; Polisky & McCarthy, 1975; Steinmetz et al., 1975; Lacy & Axel, 1975; Gadski & Chae, 1978), recent more sensitive studies have suggested that some degree of sequence *preference* may nevertheless be involved in the association of histone cores with DNA (Ponder & Crawford, 1977; Musich et al., 1977). Such arrangements might serve to modulate recognition by regulatory factors; they might also be physiologically irrelevant.

The basic structural subunit of chromatin has been shown to consist of nucleosomes containing approximately 200 base pairs of DNA, 140 of which are associated more tightly than the rest with the histone core; the core itself contains two each of the histones 2A, 2B, 3, and 4 (for a review, see Felsenfeld, 1978). Several groups (Tatchell & Van Holde, 1977; Woodcock, 1977; Steinmetz et al., 1978) have characterized reconstituted nucleosomes prepared from "random-sequence" DNA. The sedimentation properties and accessibility to nuclease attack are very similar to native nucleosomes.

We have begun a series of experiments requiring the reconstitution of DNA restriction fragments containing the *Escherichia coli lac* operator with histones. The DNA fragments used here for reconstitution differ from those used in previous studies in two important ways. First, each has a known unique sequence. Second, each has a discrete length. The shorter, 144 base pairs long, approximates nucleosome core DNA and the longer, 203 base pairs long, approximates the core plus spacer. During the course of this work we noticed that some properties of these reconstituted nucleosomes were very unusual. Here we report that both fragments direct formation of nucleosomes that have preferred arrangements of the histone core with respect to DNA sequence.

## Experimental Procedures

*I. Materials.* Enzymes were obtained from the following sources: exonuclease III and *Hha*I[1] from BioLabs, T4 polynucleotide kinase from Boehringer Mannheim, *Eco*RI (a gift) from Mike Komaromy, bacterial alkaline phosphatase and pancreatic DNase I from Worthington, RNase and micrococcal nuclease from Sigma, and Pronase from Calbiochem. $\gamma$-$^{32}$P-labeled ATP was prepared as described by Maxam & Gilbert (1977) following the procedure of Glynn & Chappell (1964).

*II. Preparation of DNA.* *Lac*-containing plasmid constructed by F. Fuller (PMB9-UV5) was prepared by the method of Clewel (1972) with amplification of plasmid replication by treatment with 150 $\mu$g/mL of chloramphenicol. Briefly, cells were harvested and lysates prepared according to the method of Guerry et al. (1973). The lysates were stored overnight at 4 °C and centrifuged for 30 min at 17 000 rpm in an SS-34 rotor. The resulting supernatant was incubated for 1 h at 37 °C in the presence of 50 $\mu$g/mL of RNase and subsequently with 100 $\mu$g/mL of Pronase for 30 min at 37 °C. Protein was removed by chloroform–isoamyl alcohol (24:1) extraction. The plasmid DNA was ethanol-precipitated and further purified by preparative CsCl isopycnic centrifugation for 48 h at 38 000 rpm in a 50 Ti fixed angle rotor.

PMB9-UV5 contains 203 base pair *E. coli lac* control region DNA inserted in the *Eco*RI cleavage site in vector PMB9. To excise the *lac* insert DNA the *lac*-containing plasmid was digested with *Eco*RI for 1 h at 37 °C in 10 mM Tris, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, pH 7.5, and loaded on an 8% polyacrylamide gel containing 0.27% bisacrylamide, 50 mM Tris–borate, pH 8.3, 1 mM EDTA. The gel was stained with 1 $\mu$g/mL of ethidium bromide and the *lac* fragment band excised, crushed, and soaked in gel elution buffer containing 0.5 M $NH_4OAc$, 10 mM $Mg(OAc)_2$, 0.1% $NaDodSO_4$, 0.1 mM EDTA overnight at 37 °C. After ethanol precipitation, the fragment was treated with bacterial alkaline phosphatase for 1 h at 37 °C in 100 mM Tris, pH 8.0, phenol-extracted three times, and precipitated with 0.3 M sodium acetate and 2.5 volumes of ethanol. The dephosphorylated fragment was labeled at both 5′ ends using [$\gamma$-$^{32}$P]ATP and polynucleotide kinase as described by Maxam & Gilbert (1977). The labeled fragment was then ethanol-precipitated and stored in 10 mM Tris, 1 mM EDTA, pH 7.5.

† From the Biochemistry Division of the Chemistry Department and from the Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California 90024. *Received July 17, 1978.* Supported by grants from the National Cancer Institute (CA 19941), the National Science Foundation (PCM 76-00712), the California Institute for Cancer Research, and the Biomedical Research Support Grant (5-S07-RR07009-11).

[1] Abbreviations used: EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; DNase I, pancreatic deoxyribonuclease; *Hha*I, endonuclease *Hha*I from *Haemophilus haemolyticus*; exo III, exonuclease III from *Escherichia coli*; bp, base pair; $PhCH_2SO_2F$, phenylmethanesulfonyl fluoride; P*, phosphorus-32; BSA, bovine serum albumin; $NaDodSO_4$, sodium dodecyl sulfate.

0006-2960/79/0418-1068$01.00/0 

In order to obtain the 144 base pair fragment, 5′-end-labeled, 203 base pair *lac* DNA was restricted with *Hha*I for 1 h at 37 °C in 50 mM NaCl, 6 mM Tris, 6 mM $MgCl_2$, 6 mM 2-mercaptoethanol, pH 7.4, and 100 μg/mL of BSA. The digestion products were run on an 8% polyacrylamide gel and the gel was autoradiographed. The labeled 144 base pair and 57 base pair fragments were excised and eluted as described above for the 203 base pair fragment.

Calf thymus DNA of nucleosome length was isolated by phenol–chloroform–isoamyl alcohol extraction of nucleosomes (see below) and is approximately 140–180 base pairs long.

*III. Preparation of Nucleosomes.* Preparation and characterization of nucleosomes from calf thymus followed the procedure of Shaw et al. (1974), as modified by Martinson et al. (1979a). Briefly, micrococcal nuclease digests of nuclei were fractionated on columns of Sepharose 4B to yield nucleosome monomers, dimers, trimers, and higher multimers. Nucleosome monomers and dimers were labeled with $^{32}P$ at both 5′ ends using [γ-$^{32}P$]ATP and polynucleotide kinase as described (Simpson & Whitlock, 1976; Noll, 1977) except that the reactions were performed at room temperature. [γ-$^{32}P$]ATP and $^{32}P$ were removed by Sephadex G-100 chromatography.

Singly end-labeled mononucleosomes were obtained by mild digestion of dinucleosomes with micrococcal nuclease. Labeled nucleosome dimers were first purified by sucrose gradient centrifugation and digested as described by Woodcock (1977) for 1 min at 20 °C, and the reaction was stopped with a tenfold molar excess of EDTA. The digestion products were centrifuged on a gradient of 5–20% sucrose, 10 mM Tris-HCl, 1 mM EDTA, pH 8, in an SW41 rotor for 12 h at 37 000 rpm at 20 °C. Gradients were fractionated and counted for radioactivity by Cerenkov counts.

*IV. Reconstitution (Martinson et al., 1979a).* Histones from calf thymus nuclei were acid-extracted with 0.3–0.5 N $H_2SO_4$, acetone-precipitated, and dissolved at a concentration of 11–12 mg/mL in 20 mM HCl. Typically, 32 μL of histones at 12 mg/mL is diluted into 128 μL of 9 M of urea/20 mM HCl and allowed to incubate at room temperature for 30 min in order to become fully unfolded. The pH was adjusted to 7.5 with 1 M Tris and phenylmethanesulfonyl fluoride ($PhCH_2SO_2F$) was added to 1 mM. 2-Mercaptoethanol was added to 2% and the histones were incubated at 37 °C under $N_2$ for 1 h. After incubation, the urea was diluted to less than 2 M, yielding a histone concentration of 0.5 mg/mL.

Labeled *lac* DNA fragments isolated as described above were suspended with 100 μg of monosome DNA as carrier in 200 μL of 2 M NaCl, 10 mM Tris, pH 7.5. The final concentration of DNA carrier was 0.5 mg/mL. Two hundred microliters of 0.5 mg/mL calf thymus histones treated as above was added slowly with mixing to the monosome DNA; such addition resulted in a solution containing 1 M NaCl and 0.8 M urea.

The mixture of histones and DNA was then either diluted to 0.6 M NaCl directly with water or dialyzed against 0.6 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5. With respect to the calf thymus carrier, either method yielded a reconstitute which gave normal DNase I digestion (Noll, 1974) as well as UV and tetranitromethane cross-linking (Martinson et al., 1979b) patterns.

The reconstituted material was sedimented in a 5–20% sucrose gradient containing 10 mM Tris, pH 8, 1 mM EDTA, 1 mM $PhCH_2SO_2F$, 0.6 M NaCl for 12 h at 37 000 rpm with an SW41 rotor. Fractions were collected and counted (Cerenkov) for the labeled fragment and monitored at 260 nm for the reconstituted carrier DNA. Monosomes sedimenting at 11 S and free of H1 were pooled and dialyzed against the appropriate nuclease digestion buffer (see below) and stored at 4 °C.

*V. DNA Digestion and Electrophoresis.* DNase I digestions of $^{32}P$-labeled nucleosomes and reconstituted *lac* nucleosomes were carried out at 37 °C in 10 mM $MgCl_2$, 0.1 mM EDTA, 10 mM Tris, pH 8, in the presence of 0.25 mg/mL of unlabeled calf thymus nucleosomes. Digestion was carried out such that 18 to 25% digestion had taken place as measured by perchloric acid soluble material and was stopped with a 50-fold molar excess of EDTA.

Exonuclease digestions were performed in 66 mM Tris, pH 8.0, 0.66 mM $MgCl_2$, 1 mM 2-mercaptoethanol at 37 °C for 30 min in the presence of 0.05 mg/mL of carrier nucleosomes. Ten units of exonuclease III was ordinarily used to digest 25 μg of the carrier. The reaction was stopped with a 50-fold molar excess of EDTA.

*VI. DNA Isolation and Electrophoresis on Denaturing Gels.* DNA was extracted from nucleosomes and the various digestion reactions by the following procedure. The sample was suspended in 0.1% $NaDodSO_4$ and extracted twice with phenol–chloroform (1:1) and twice with chloroform–isoamyl alcohol (24:1). Three volumes of cold ethanol was added after NaCl was brought to 0.1 M. Ethanol precipitation occurred at −78 °C in a dry ice–acetone bath, and then the samples were spun at 10 000 rpm for 20 min. The precipitate was washed with ethanol, respun, and dried. For denaturing gels, each sample was suspended in 25 μL of 7 M urea, boiled for 3 min, and quick-chilled on ice before being loaded on either a 6 or 8% polyacrylamide denaturating gel containing 7 M urea in 100 mM Tris–borate, 2 mM EDTA, 3 mM ammonium persulfate (Maniatis et al., 1975). Electrophoresis was conducted at between 250 and 300 V until the bromophenol blue tracking dye had migrated to within 2 cm of the bottom of the gel. Sequenced *lac* DNA strands of 21, 31, 63, 84, and 124 bases long were used to calibrate the gels. All gels were autoradiographed on Cronex 4 X-ray film with the aid of an Intensifier (Du Pont Quanta II) at −70 °C.

*VII. Partial Sequence Analysis.* Exonuclease generated fragments were eluted from a 6% polyacrylamide–urea denaturing gel after autoradiography and ethanol precipitated as described. These fragments are singly labeled and ready for sequencing analysis.

Partial sequencing of end-labeled DNA fragments from the exo III digest was performed using the guanine-only cleavage reaction described by Maxam & Gilbert (1977), except that methylation was enhanced by adding 1 μL of dimethyl sulfate at 5-min intervals for 20 min at room temperature. This additional methylation resulted in more cleavage closer to the end label and thus, for our purpose, required less DNA for analysis. The cleavage products were analyzed on 20% polyacrylamide gels (in 7 M urea, 50 mM Tris–borate, pH 8.3, 1 mM EDTA) and autoradiographed.

## Results

*Isolation of Lac Reconstitutes.* Figure 1 shows the results of digestion of the $^{32}P$ end-labeled 203 base pair *lac* DNA fragment with restriction endonuclease *Hha*I to obtain the 144 base pair fragment. In addition to the stated number of base pairs, each fragment contains unpaired bases at the ends as a result of restriction digestion. Thus, the larger fragment contains two 207 base strands, each end-labeled with $^{32}P$, while the shorter fragment is end-labeled only on its 150 base strand, and not at the 5′ end of the 144 base strand (see Figure 1B). These fragments contain the entire *lac* control region which

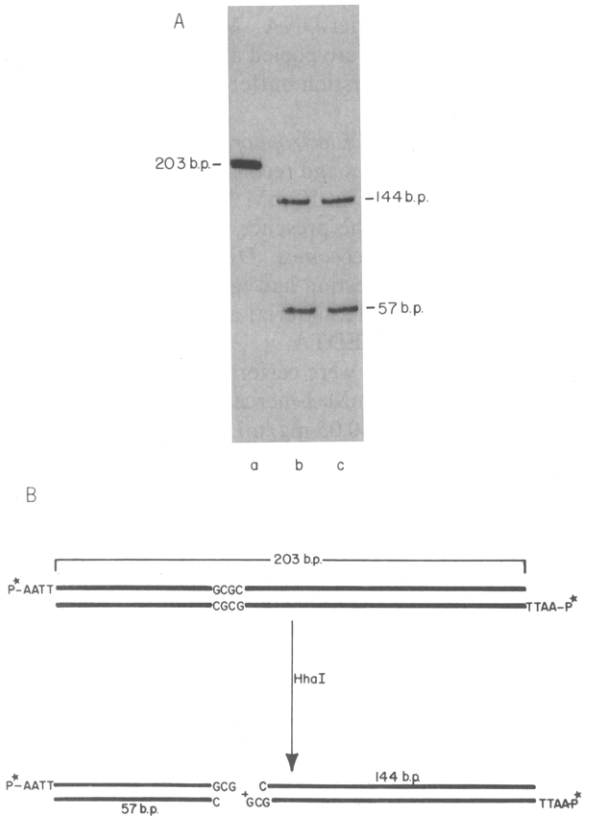


FIGURE 1: *Hha*I digest of $^{32}$P-labeled 203 base pair *lac* DNA. (A) Eight percent native polyacrylamide gel analysis of an *Hha*I digest of kinased 203 base pair *lac* DNA. *Hha*I digestion gives two singly end labeled fragments, measuring 63 and 150 bases long on the labeled strand. Electrophoresis was for 2.5 h at 250 V. The gel was autoradiographed on Kodak No-Screen film for 15 min, and the bands were eluted as described in the text. Lane a shows undigested 203 base pair *lac* DNA, and lanes b and c show *Hha*I digestion of 203 base pair DNA. (B) A schematic drawing showing the lengths and labeling pattern of the fragments.

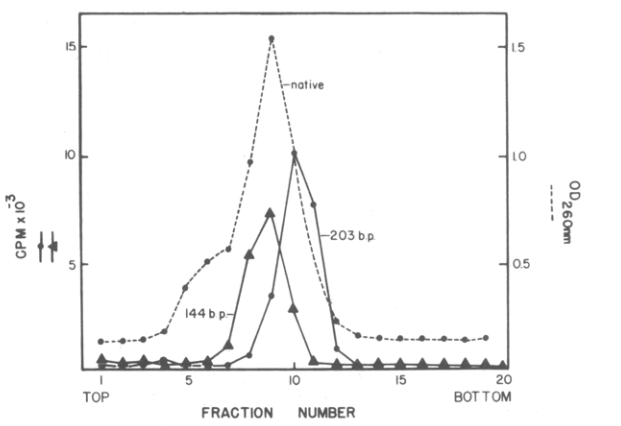


FIGURE 2: Sucrose gradient analysis of reconstituted *lac* DNA nucleosomes. Reconstituted end-labeled 144 and 203 base pair *lac* DNA were run on 5–20% sucrose gradients containing 0.6 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5, fractionated as described in the text. A parallel sucrose gradient containing isolated native calf thymus nucleosomes was run and fractionated, and the $A_{260}$ was measured (dashed line). These native nucleosomes cosedimented with reconstituted calf thymus nucleosomes. Dinucleosomes sediment at fraction 14 in this gradient.

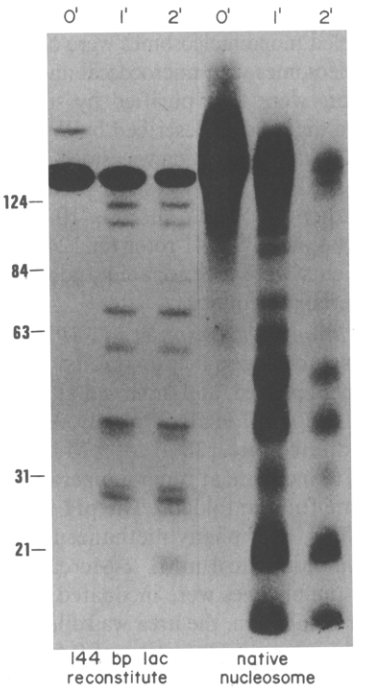


FIGURE 3: Autoradiograph of DNase I digest of reconstituted 144 base pair *lac* DNA and singly labeled calf thymus nucleosomes. Single-stranded DNA fragments resulting from DNase I digestion were run on an 8% polyacrylamide gel containing 7 M urea. Electrophoresis was for 2.5–3 h at 300 V. Markers of 21, 31, 63, 84, and 124 nucleotides were electrophoresed under the same conditions and their positions are designated by identifying numbers on the left side of the gel. The time course of digestion is shown for the 144 base pair *lac* reconstitute and the native nucleosome. Digestion of 144 base pair *lac* DNA under similar conditions (not shown) gave no pattern of bands.

has been sequenced (Gilbert & Maxam, 1973; Dickson et al., 1975).

Each DNA fragment was reconstituted with core histones as described (Martinson et al., 1979a). This procedure has been shown to yield calf thymus reconstitutes of similar sedimentation coefficient, DNase I digestion accessibility, and histone–histone cross-linking pattern to nucleosomes isolated from nuclei (unpublished observations).

Sucrose gradient velocity sedimentation displayed in Figure 2 showed that the 144 base pair reconstitute cosediments with native calf thymus mononucleosomes. The 203 base pair reconstitute sediments slightly faster, as would be expected of a particle with similar shape but larger mass. This sedimentation is a preparative step in our isolation protocol and is run in 0.6 M sodium chloride to remove H1.

*Deoxyribonuclease I (DNase I) Studies.* Digestions of doubly end-labeled nucleosomes (Noll, 1977; Simpson & Whitlock, 1976) using DNase I have shown preferential sites of cleavage by DNase I. We performed the same experiment on our singly labeled 144 base pair *lac* DNA reconstitute and compared the distribution of the end label in the single-stranded fragments produced with those resulting from digestion of native nucleosomes. For strict comparability, native singly end-labeled nucleosomes were prepared by mild micrococcal nuclease digestion of doubly labeled nucleosome dimers followed by isolation in a sucrose gradient.

Figure 3 displays the time course of DNase I digestion of the singly labeled 144 base pair *lac* reconstitute and the native nucleosome control. Both clearly show regions of intensity at 10-base intervals. Additionally, both show the pattern of unequal band intensities revealed by end labeling and characteristic of the differential accessibility of DNA sites in nucleosomes to DNase I (Noll, 1977; Simpson & Whitlock, 1976). This verifies that the DNA folding around the histone core in the *lac* reconstitute is similar to that in native nucleosomes.

There are, however, curious differences between the DNase I digestion patterns of native and *lac* reconstituted nucleo-

somes. The native pattern of relative intensities is similar to those obtained previously (Noll, 1977; Simpson & Whitlock, 1976). Qualitative analysis of these patterns supports the concept of a symmetric nucleosome in which the DNA at sites near 30, 60, 70, 80, and 110 bases from a 5′ end are relatively more resistant to DNase I attack than the other sites (Noll, 1977; Simpson & Whitlock, 1976; Finch et al., 1977).

Although the *lac* reconstitute pattern is very similar, the symmetry is lost; for example, a low-intensity region, flanked by two high-intensity ones, now appears at 50 bases rather than 30. The DNA bands are also very narrow in the *lac* reconstitute and in the low molecular weight portion of the gel; each ten-base region shows clearly resolved doublets or triplets in the autoradiograph. By contrast, the native digest bands are very broad, probably reflecting the heterogeneous lengths of DNA obtained in the original micrococcal nuclease cleavage of nuclei.

Taken together these data suggest that the 144-bp reconstitute has assumed a nonrandom asymmetrical arrangement of the histone core with respect to the fragment of *lac* DNA. The pattern may be viewed as the superimposed digestion of two or three predominant conformations, each having a different distance between the DNA end-label and the DNase I sensitive sites. Unfortunately it is not known how many such arrangements are theoretically possible.

Reconstitution and DNase I digestion of 145 ± 10 base pair "random" sequence end-labeled DNA failed to reveal a clean ten base pair pattern in radioautographs though the unlabeled DNA revealed the usual pattern upon staining (data not shown). Since the unlabeled pattern reflects all size classes produced by internal cleavages, whereas the end-labeling reflects the position of the 5′ end with respect to the internal 10-bp repeat, this simply shows that in the reconstitute the DNA 5′ termini are randomly arranged with respect to the histone cores.

In order to confirm and extend these observations concerning arrangements, we turned to another nuclease digestion probe, exonuclease III.

*Exonuclease III Probe.* Exonuclease III (exo III) digests rapidly (3′ → 5′) DNA tails not associated with the nucleosome core (Prunell & Kornberg, 1977; Riley & Weintraub, 1978). Although excessive digestion gives rise to invasion of the core itself, moderate exo III digestion of 5′-end-labeled nucleosomal DNA should leave a shortened labeled DNA strand, the 3′ nucleotide of which specifies the point at which DNA enters the nucleosome core. If both DNA strands of a restriction fragment are 5′-end labeled, the two 3′ nucleotides should specify the precise location of the histone core with respect to DNA sequence.

In order to detect and locate the histone core of nonrandom reconstitutes, we have performed exonuclease III digestions of the 203 base pair *lac* reconstitute, labeled at both 5′ ends. Since the core consists of only 140 base pairs, many arrangements are possible along the DNA. For each conformation, exo III digestion followed by strand separation should yield a pair of labeled strands which should add up to a length of approximately 347 bases (see Figure 4). If the reconstitution is random, one expects an unresolvable smear of pairs of bands, ranging in length between 140 and 207 bases.

Figure 5A shows that exonuclease III digestion of the 203-bp reconstitute yields specific bands, superimposed on a light background smear. These include a broad band below 207 bases and species with length of 180, 168, and 150 bases. On the other hand, exonuclease digestion of 200 ± 10 base pair "random" sequence DNA reconstitute (Figure 5B) reveals

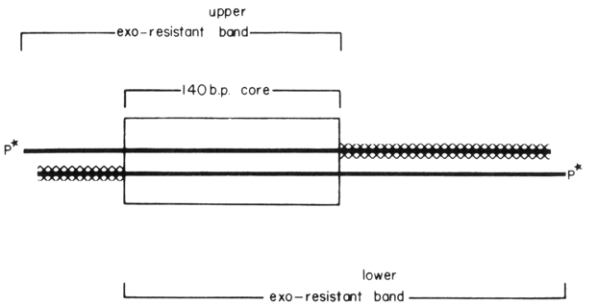


FIGURE 4: The exonuclease experiment. Two bands should arise from exo III digestion (shown by cross-hatching) of a specific conformation of the *lac* reconstitute. Each 207 base strand should be trimmed outside the core, which contains approximately 140 base pairs. Before digestion there are a total of 2 × 207 = 414 bases present, 134 of which are outside the core. Since exo III is 3′ specific, it will eliminate only half (67) of the bases outside the core. Therefore, the total number of resistant bases is 414 − 67 = 347. The two bands, representing the upper and lower DNA strands, which result from digestion should add up in length to 347 bases. A random collection of conformations should give an unresolvable smear of bands, representing many pairs which add up to 347 bases.

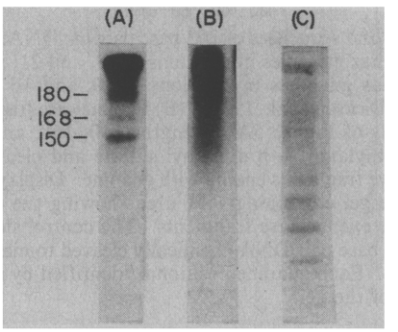


FIGURE 5: Exonuclease III digestion of reconstituted *lac* and "random" sequence DNA. Exonuclease-resistant fragments were purified and electrophoresed on a 6% gel containing 7 M urea for 4–5 h at 250 V. All digestions were carried out at the same enzyme/DNA ratio in the presence of identical concentrations of carrier monosomes for 30 min at 37 °C. (A) Digestion pattern of fragments from an exonuclease digest of reconstituted 203 base pair DNA. The pattern is invariant over a time course of digestion. The three predominant bands are 150, 168, and 180 bases long. There is also a moderate amount of undigested material which is digested at higher ratios of enzyme to DNA (see Discussion). (B) Exo III digestion of reconstituted "random" sequence DNA. $^{32}$P-labeled DNA of 200 ± 10 base pairs length from calf thymus was extracted from a polyacrylamide gel and reconstituted and digested with Exo III. (C) Exo III digestion of $^{32}$P-labeled 203 base pair *lac* DNA. Labeled *lac* DNA was digested under the same conditions as above, and a 30 min digestion point is shown. A time course of digestion showed a continuously changing pattern of bands over a heavy background smear.

the predicted smear, reflecting the many arrangements within this heterogeneous mixture. Naked *lac* DNA was digested as a control. Surprisingly, at early times of digestion, a pattern of discrete bands rather than a smear resulted (Chao et al., unpublished results). However, in contrast to the stable pattern obtained with the *lac* reconstitute (Figure 5A) the pattern of bands for the naked DNA continuously shifted to lower molecular weights as digestion proceeded. A sample time point is shown in Figure 5C. In no instance did the digestion pattern for naked DNA duplicate that for the reconstitute. In particular, for the reconstitute, the dark bands shown in Figure 5A appear fairly quickly (see legend to Figure 5) and bands which appear below 150 bases at longer times of digestion are always faint. These data indicate that the dominant pattern shown in Figure 5A is not a manifestation of digestion to pauses at 10-base intervals within the nucleosome as observed

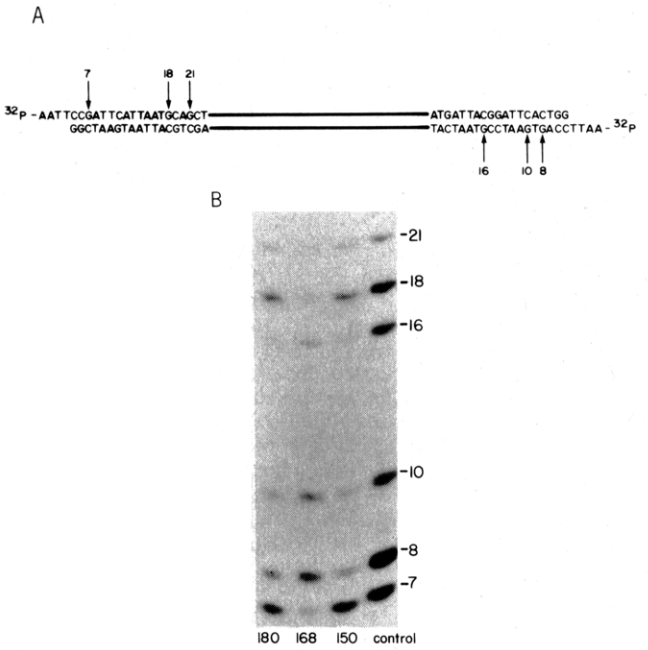


FIGURE 6: (A) End sequences of 203 base pair *lac* DNA. The upper labeled strand has guanines in positions 7, 18, and 21, whereas the lower strand has guanines in positions 8, 10, and 16 (Gilbert & Maxam, 1973; Dickson et al., 1975). (B) Exonuclease fragments from the experiments of Figure 5A of lengths 150, 168, and 180 were chemically methylated with dimethyl sulfate and cleaved in 1 M piperidine to give fragments ending with guanine. Displayed is a 20% polyacrylamide gel containing 7 M urea showing the positions of guanines of the exonuclease fragments. The control slot is doubly labeled *lac* 203 base pair DNA chemically cleaved to mark guanines of both strands. Each guanine position is identified by numbers on the right side of the gel.

by Riley & Weintraub (1978) and Prunell & Kornberg (1977). The sequencing data presented below confirm this interpretation.

The simplest explanation for these results is that this 203 base pair *lac* reconstitute includes two highly favored conformations. One conformation yields the paired lengths of 180 and 168, while the other yields 150 and a species incompletely resolved in the 200-base region of the autoradiograph. This interpretation predicts that each partner of the pair must represent a different DNA strand.

To test this interpretation and to locate precisely the two arrangements with respect to DNA sequence, we performed a partial sequence analysis of the well-resolved species of 180, 168, and 150 bases. Since these strands are 5′-end labeled, their sequence may be determined by the chemical cleavage method (Gilbert & Maxam, 1977). Figure 6A shows that the two DNA strands are easily distinguished by the positions of guanine residues with respect to the 5′-end label. Figure 6B displays the guanine-specific cleavage products of each of the exonuclease-resistant bands. Notice that the 180- and 168-base species clearly represent complementary strands as predicted; the 168-base fragment shows guanines predominantly at positions 8, 10, and 16, while the 180-base fragment shows guanines at 7, 18, and 21 bases from its 5′ end. The light background pattern in all positions results from contaminating strands present in the light background smear of the original exonuclease digest which were excised along with the major band.

We conclude that the 203-bp *lac* reconstitute is highly nonrandom; of the many potential conformations, only the two shown in Figure 7 are highly favored. Notice that these two conformations are not symmetrically related to each other with respect to identical 5′ DNA termini. This suggests that some

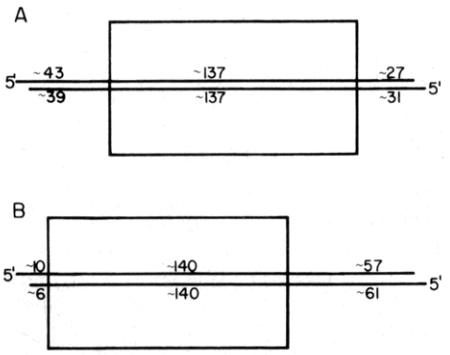


FIGURE 7: Schematic diagram of predominant arrangements of 203 base pair *lac* reconstitute. Arrangement A gives rise upon exonuclease digestion to a 180-base fragment and a 168-base fragment. Exonuclease digestion of arrangement B gives rise to a 150-base fragment and an unresolved fragment (see Figure 5A) of about 200 bases.

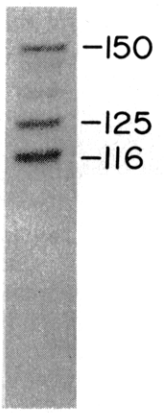


FIGURE 8: Autoradiograph of exonuclease III digest of 144 base pair *lac* reconstitute. The 144 base pair *lac* reconstitute was digested as described and run on a 6% polyacrylamide gel containing 7 M urea. The labeled strand is 150 bases long. The predominant subfragments were sized to be 125 and 116 bases.

aspect of DNA sequence rather than just DNA ends specifies the manner of association of the fragment with the histones.

Next, we applied the exonuclease probe to the 144-bp *lac* reconstitute. Since this fragment is labeled at one end only, a single band rather than a pair should arise from each positional conformer. Figure 8 shows that specific bands of length 150, 125, and 116 bases are dominant. As in the case of the 203 base pair *lac* fragment, digestion of the 144 base pair DNA alone gives a collection of bands mostly smaller than obtained by digestion of the reconstitute. Therefore, it is likely that each band shown in Figure 8 results from exo III digestion halted at a nucleosome core. This suggests that three conformational arrangements may be present in the reconstitute, perhaps those suggested by the doublets and triplets in the DNase I digest. A self-consistent explanation for the DNase I and exonuclease III data is presented under Discussion. However, the conclusion that DNA can reassociate with histones in a sequence specific manner to form asymmetric nucleosome cores does not depend on the details of this model.

## Discussion

We have shown that two unique DNA fragments reconstitute with histones in a manner which is nonrandom with respect to DNA sequence. Below we discuss possible implications and uncertainties associated with this finding.

*The 203 Base Pair Reconstitute.* Out of many potential conformations, the 203-bp reconstitute favors the two diagrammed in Figure 7. Although one can find limited sequence

homology by alignment of the two conformations with respect to the 140-bp core, it would be premature to ascribe functional significance to such sequences. It is clear, however, that elements of DNA sequence are responsible for the nonrandom reconstitute. "Random" sequence DNA was shown to reconstitute "randomly". The conformations could not be directed by the restriction termini, since both are TTAA and the two conformations are not located symmetrically with respect to these identical sequences. Unfortunately, for technical reasons (see below and legend to Figure 5), we are unable to quantitate the fraction of the total population in these two highly favored conformations.

The exonuclease digest of the 203 base pair *lac* reconstitute shows that a moderate fraction of strands is completely resistant to digestion at that ratio of enzyme to DNA (see undigested material in Figure 5A). As the ratio is increased, the starting material is digested and a collection of many minor bands appears in the lower molecular weight region of the gel (data not shown). However, the major exo resistant bands do not change in intensity, suggesting that they are in steady-state concentration or that the label is being excised due to the 5′-clipping activity noted by Riley & Weintraub (1978). The relative exo resistance of the intact 203 base pair reconstitute suggests that the DNA termini are in some way associated with the histone octamer despite their position outside the 140 base pair core region. This explanation is consistent with the histone octamer actually containing binding capacity for 200 base pairs of DNA, only 140 base pairs of which constitute the core (Christiansen & Griffith, 1977) which is preferentially exonuclease resistant.

*The 144 Base Pair Reconstitute.* We have derived a self-consistent model which explains the asymmetry of the 144 base pair reconstitute, the presence of multiplets in its DNase I digestion pattern, and the appearance of exo III resistant bands of 150, 125, and 116 bases. Although the full length 150 base band could reflect the same exonuclease resistance of the DNA termini seen in the 203 bp reconstitute digestion, the data are more easily explained if the 150 base band is assumed to represent a specific conformation as discussed below.

The intensity variation within the DNase I digestion pattern of the singly end-labeled 144 base pair reconstitute is shifted approximately 20 bases "up" relative to end-labeled native calf nucleosomes (Figure 3). This shows (see Results) that the reconstitute has distinctly preferred conformation(s) which position the DNA asymmetrically with respect to the histone core. From Figure 3 we can also conclude that any such favored conformation should position the fragment in such a way as to result in a relatively inaccessible DNase I digestion site falling approximately 47 bases in from the uniquely labeled 5′ end (47 bases correspond to the low-intensity position in the reconstitute DNase digest). These DNase I resistant sites could be at any of the tight binding domains of the histone core which occur in native nucleosomes at 30, 60, 70, 80, and 110 base pairs (see Finch et al., 1977). Figure 9 shows three such possible conformations, each of which is asymmetric.

The third conformation alone would yield a DNase I digestion pattern similar to that of Figure 3. However, Figure 3 displays triplets of bands in the low molecular weight region of the gel suggesting the existence of three conformations. Moreover Figure 8 shows the existence of three discrete exo III resistant bands, again suggestive of three preferred arrangements. In particular the lengths of the three exonuclease resistant bands, namely, full length (150 bases), 125 and 116 bases long, suggest directly the existence of each of the three

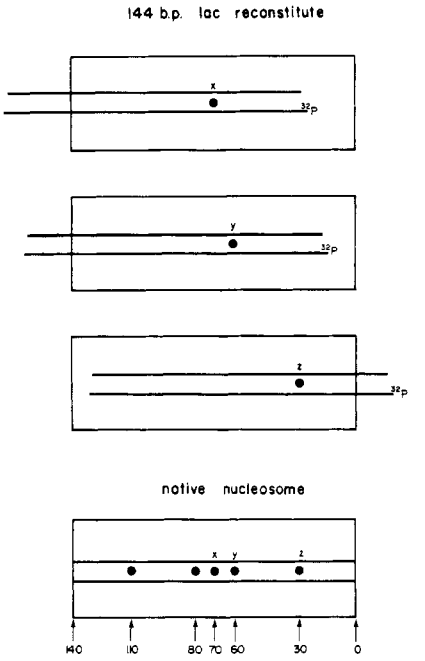


FIGURE 9: A schematic representation of the three proposed conformations of the 144 base pair *lac* reconstitute. A core native nucleosome is shown for reference with dots representing the less accessible DNase I sensitive sites. The distances of these sites from one end of the core DNA are shown. The three conformations of the reconstitute are shown above; each dot is 47 bases removed from the 5′-end label and at the specified position with respect to the native nucleosome core. We should point out that 47 bases is an average, since the triplets of approximately 25, 27, and 29 bases in the DNase I digest of Figure 3 suggest that the less accessible DNase I sites actually occur at 45, 47, and 49 bases for the different conformations of the reconstitute.

preferred arrangements shown in Figure 9. Summation of the DNase I digestion patterns over all three conformations would yield a composite pattern similar to that shown in Figure 3.

Notice that each of the arrangements is unusual not only in that it is asymmetric but also in that it necessarily utilizes only partially the capacity of the core to bind 140 base pairs of DNA. This conclusion does not depend on the precise details of the above model. Apparently the contacts directing these conformations provide sufficient energy to compensate for the loss of other available contacts within the core. This would seem reasonable if, as has been suggested (Christiansen & Griffith, 1977), the histone octamer contains capacity for a weaker binding of (perhaps 60) additional base pairs of DNA. In any case, the observation that the 140 base pair capacity of the core is not utilized implies that many potential arrangements are possible even for the 144 bp reconstitute; recall that only three at most predominate. These data further suggest that "random" sequence DNA reconstitutes may be highly heterogeneous even if the DNA utilized is of uniform 140 base pair length. Individual DNA sequences would direct different conformations, of which there are many since not all of the DNA must bind within the core.

*Possible Implications.* Although it generally has been accepted that histone cores exhibit little sequence preference in their association with DNA, these results raise the possibility that a small, but important, subpopulation of preferentially placed nucleosomes may exist in vivo. The DNA fragments we used direct the nonrandom placement of nucleosomes with respect to DNA sequence. One may expect that eukaryotic DNA contains sequences both less and more capable of directing preferential positioning. We cannot assess whether the free energy of interaction of specific DNA sequences with

histone cores is a significant factor among the unknown forces which drive nucleosome formation in vivo. It is clear, however, that the bulk of nucleosomes which form in vitro do not exhibit the same sequence preference, if any, which obtain in vivo. We have already mentioned that, whereas end-labeled native nucleosomes yield the usual DNase I "ladder" pattern, end-labeled reconstitutes derived from calf DNA yield only a smear. This shows that the histone cores are positioned differently on the calf monosome DNA in at least most of the reconstituted nucleosomes compared with the native material. Since analysis of DNase I digested reconstitutes by DNA staining rather than end labeling reveals the normal ladder pattern, this result cannot be ascribed to grossly aberrant reconstitution (see Results). If preferentially situated nucleosomes do exist in vivo, they may function through localized effects in directing the manner of DNA packaging in the vicinity of regulatory sites.

Our results are consistent with several reports in the literature. For example, although restriction enzyme analysis of SV40 histone complexes suggested random nucleosome positioning (Polisky & McCarthy, 1975) a more sensitive assay revealed that the nucleosomes were positioned neither randomly nor uniquely (Ponder & Crawford, 1977). This nonrandom population was also detected in SV40 reconstitutes. Several studies on repetitive DNA have also revealed nonrandom placement of nucleosomes (Musich et al., 1977). Though these are special cases, it does demonstrate that the phenomenon we observe in vitro may have in vivo precedents.

A corollary question is whether the position of histone cores with respect to DNA sequence affects the availability of sequences to be recognized. These *lac* reconstitutes contain sequences which normally specifically bind several *E. coli* regulatory proteins, including *lac* repressor. Our preliminary results indicate that *lac* repressor binds the 203 base pair *lac* reconstitute with high affinity and specificity. We are not yet certain whether the position of the *lac* operator within the histone core is important in determining the binding characteristics.

## References


Axel, R., Melchior, W., Sollner-Webb, B., & Felsenfeld, G. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4101.

Christiansen, G., & Griffith, J. (1977) *Nucleic Acids Res. 4*, 1837.

Clewel, D. B. (1972) *J. Bacteriol. 110*, 66.

Dickson, R. S., Abelson, J., Barnes, W. M., & Reznikoff, W. S. (1975) *Science 187*, 27.

Felsenfeld, G. (1978) *Nature (London) 271*, 115.

Finch, J. T., Lutter, L. C., Rhodes, D., Brown, R. S., Rushton, B., Levitt, M., & Klug, A. (1977) *Nature (London) 269*, 29.

Gadski, R. A., & Chae, C.-B. (1978) *Biochemistry 17*, 869.

Gilbert, W., & Maxam, A. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3581.

Glynn, I. M., & Chappell, J. B. (1964) *Biochem. J. 90*, 147.

Guerry, P., LeBlanc, D. J., & Falkow, S. (1973) *J. Bacteriol. 116*, 1061.

Lacy, E., & Axel, R. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3978.

Maniatis, T., Jeffrey, A., & Van de Sande, H. (1975) *Biochemistry 14*, 3787.

Martinson, H. G., True, R. J., Burch, J. B. E., & Kunkel, G. (1979a) *Proc. Natl. Acad. Sci. U.S.A.* (in press).

Martinson, H. G., True, R. J., Lau, C. K., & Mehrabian, M. (1979b) *Biochemistry 18*, this issue.

Maxam, A., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560.

Musich, P. R., Maio, J. J., & Brown, F. L. (1977) *J. Mol. Biol. 117*, 657.

Noll, M. (1974) *Nucleic Acids Res. 1*, 1573.

Noll, M. (1977) *J. Mol. Biol. 116*, 49.

Polisky, B., & McCarthy, B. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2895.

Ponder, B. A. J., & Crawford, L. V. (1977) *Cell 11*, 35.

Prunell, A., & Kornberg, R. D. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 103.

Riley, C., & Weintraub, H. (1978) *Cell 13*, 281.

Shaw, B. R., Corden, J. L., Sahasrabuddhe, C. G., & Van Holde, K. E. (1974) *Biochem. Biophys. Res. Commun. 61*, 1193.

Simpson, R. T., & Whitlock, J. P. (1976) *Cell 9*, 347.

Steinmetz, M., Streeck, R. E., & Zachau, H. G. (1975) *Nature (London) 258*, 447.

Steinmetz, M., Streeck, R. E., & Zachau, H. G. (1978) *Eur. J. Biochem. 83*, 615.

Tatchell, K., & Van Holde, K. E. (1977) *Biochemistry 16*, 5295.

Woodcock, C. L. F. (1977) *Science 195*, 1350.