Eukaryotic DNA does not form nucleosomes as readily as some prokaryotic DNA

Robert C. Getts and Michael J. Behe*

Department of Chemistry, Lehigh University, Bethlehem, PA 18015, USA

Received July 30, 1991; Revised and Accepted October 11, 1991

ABSTRACT

Nucleosomal-length DNA was prepared from the genomic DNA of various prokaryotic and eukaryotic organisms by limited nuclease digestion after reconstitution with core histones. The DNAs ranged in base composition from 26.5% to 72% guanosine-plus-cytosine (%GC). The nucleosomal-length DNAs were then used in a competitive reconstitution assay in order to quantitatively determine their relative abilities to form nucleosomes. The results of the assay indicate a linear dependence of the free energy of nucleosome formation on base composition and, surprisingly, show that several prokaryotic DNAs form nucleosomes as well as or better than eukaryotic DNAs.

INTRODUCTION

There are numerous differences between prokaryotes and eukaryotes, but perhaps the most fundamental difference is in the organization of genomic DNA. Eukaryotic DNA occurs as a discrete, multi-level nucleoprotein complex, with a basic repeating structure consisting of nucleosomes (1). The nucleosome core consists of 2 copies each of the four histone proteins H2A, H2B, H3, and H4, around which is wrapped 146 base pairs of DNA. A single copy of a fifth histone, H1, is present in higher eukaryotes and appears to be associated with the linker region between nucleosomes. Prokaryotic DNA, on the other hand, appears to be only loosely associated with organizing proteins (2); repetitive higher-order nucleoprotein organization is difficult to discern and may be nonexistent (3).

Although a number of studies have shown the ability of DNAs from both eukaryotic and prokaryotic sources, as well as synthetic DNAs, to form nucleosomes *in vitro* (1,3), typically these asked only if nucleosomes formed or not; quantitative information was lacking. In the past several years, however, sensitive, quantitative assays have been developed for the determination of the difference in free energies of nucleosome formation when different DNA components are used (4–6). In this study we use such an assay to address a basic question in a quantitative manner: what is the ability of eukaryotic vs. prokaryotic DNA to form nucleosomes? The surprising results show that, although the ability of

prokaryotic DNAs to reconstitute is dependent on their base composition, some prokaryotic DNAs are substantially better at forming nucleosomes than eukaryotic DNAs.

MATERIALS AND METHODS

Materials

Ultrapure grade genomic DNAs from calf thymus, human placenta, *Escherichia coli*, *Clostridium perfringens*, and *Micrococcus luteus* were purchased from Sigma Chemical Company. Salmon testes DNA was also purchased from Sigma and was purified before use by phenol/chloroform extraction followed by cesium chloride density centrifugation. Micrococcal nuclease was obtained from Pharmacia. T4 polynucleotide kinase was purchased from United States Biochemicals. Chicken blood was obtained in Alsever's solution from Lampere Biologicals and arrived in the laboratory the same day as the bleeding was performed.

Preparation of Chicken Nucleosome Monomers and Histones

Nucleosome monomers were prepared from chicken chromatin by limited digestion of erythrocyte nuclei with micrococcal nuclease (7). Monomers were purified by ultracentrifugation at 22,000 rpm in a Beckman SW-27 rotor at 4°C for 40 hours on 5% to 20% linear sucrose gradients containing 10 mM NaCl, 10 mM Tris·HCl, pH 8.0, and 0.1 mM EDTA. Core histones were prepared by the method of Simon & Felsenfeld (8), which involves adsorbing chromatin to a hydroxyapatite column, washing with 0.6 M NaCl to remove H1, and eluting the core histones with 2.0 M NaCl.

Preparation of Nucleosomal Length DNA

Nucleosomal-length DNAs were made by limited micrococcal nuclease digestion of genomic DNAs that had been reconstituted with chicken histones. In order to obtain representative DNA populations, and to avoid selecting a subpopulation of DNA sequences, as much of the genomic DNA as possible had to be reconstituted. This was done as follows. The genomic DNAs were associated with histones by a stepwise dilution procedure similar to that of Daban and Cantor (9). Briefly, ~90 µg of heterogenous-length genomic DNA was mixed with ~72 µg of

* To whom correspondence should be addressed

core chicken histones (protein/DNA ratio of 0.8) in a total volume of 30 μ l of [2.0 M NaCl, 15 mM Tris·HCl, pH 8.0]. The mixture was agitated gently at room temperature for at least three hours. The salt concentration was then diluted to 1.0 M by the addition of 15 mM Tris·HCl (pH 8.0). Once again the solution was gently agitated for at least 3 hours. The salt concentration was further diluted to 0.5 M and finally to 0.2 M by repeated additions of the same buffer. The sample was then agitated for an additional 10 hours.

Following reconstitution the sample was made 2.5 mM in CaCl_2 and digested with 50 units/ml of micrococcal nuclease at 37°C for 5 to 10 minutes. The digested mixture was layered onto a 5%-20% linear sucrose gradient and spun at 22,000 rpm for 40 hrs at 4°C in an SW27 rotor. The gradients were fractionated and the fractions that contained nucleosome monomers were pooled and dialyzed against 5 mM Tris·HCl (pH 8.0), 0.1 mM EDTA. In all cases except chicken erythrocyte (in which the genomic DNA was digested directly as chromatin in the erythrocyte nucleus) at least 40% of the input DNA was found in the monomer fractions.

The monomer solution was extracted three times with one volume of phenol/chloroform/isoamyl alcohol (25:24:1), and precipitated with ethanol. The nucleosomal-length DNA was resuspended in a minimum volume of 5 mM Tris·HCl (pH 8.0),

0.1 mM EDTA and was loaded onto a 6% polyacrylamide gel. A narrow band of DNA corresponding to approximately 150 ± 10 bp was excised from the gel, eluted into 1.5 M ammonium acetate, extracted once with one volume of n-butanol, and precipitated with two volumes of absolute ethanol. The DNA was then labeled with $\gamma\text{-}^{32}\text{P}\text{-ATP}$ using T4 polynucleotide kinase.

Competitive Reconstitution Assay

The ability of each of the nucleosomal-length genomic DNAs ('test DNAs') to form nucleosomes was quantitated by a competitive reconstitution assay (2). Briefly, 0.5 μ g of labeled test DNA is incubated with 18 μ g of nucleosome monomers and 9 μ g of bulk, unlabeled, nucleosomal-length, chicken DNA (this is a 1:1 mole ratio of nucleosomes to free DNA) in 20 μ l of [800 mM NaCl, 20 mM Tris·HCl, pH 8.0, 0.1 mM EDTA] for 20 minutes at 37°C. The mixture is then diluted over the course of four hours by the repeated addition of [5 mM Tris·HCl, pH 8.0, 0.1% Nonidet P-40] down to a sodium chloride concentration of 80 mM. Each dilution step decreased the salt concentration by a factor of 1.2. The reconstituted mixture is then placed on a 6% polyacrylamide gel and electrophoresed at low voltage. The gel is subsequently stained with ethidium bromide, the bands corresponding to free DNA and nucleosomal DNA excised, and the radioactivity contained in the bands quantitated by scintillation counting.

The differential free energy of nucleosome formation can then be calculated (4) from the equation

$$\Delta\Delta G = -RT \ln K_t/K_s$$

where $\Delta\Delta G$ is the difference in free energy of reconstitution between the test DNA and bulk chicken DNA, K_t is the ratio of counts in the nucleosome band to counts in the free DNA band for the test DNA, and K_s is the same ratio for the standard (labeled chicken DNA that has been put through the competitive reconstitution procedure.)

Genomic and Nucleosomal DNA Melting Temperatures

Melting temperatures of genomic and nucleosomal DNAs were determined using a Gilford System 2600 UV-Vis spectrophotometer with a Gilford 2527 Thermoprogrammer. In a typical measurement 20 μ g of DNA was placed in a buffer of [0.1 M NaCl, 10 mM sodium phosphate, pH 7.0, 0.1 mM EDTA] and the absorbance of the sample at 260 nm was monitored as the temperature was increased at a rate of 1°C/min. The precision of the melting temperature measurement is $\sim \pm 0.5^\circ\text{C}$.

Figure 1. Isolation of reconstituted nucleosomes from genomic DNAs using sucrose density gradients. A) *E. coli*; B) *C. perfringens*; C) *M. luteus*; D) human placenta; E) calf thymus; F) salmon sperm. Pooled monomer fractions are indicated in each panel by a bar.

Table 1. Reconstitution of prokaryotic and eukaryotic DNAs.

DNA Type	%GC	%NM	Bound/ Free	$\Delta\Delta G$ (kcal/mol)
Chicken erythrocyte	43	—	0.386	0.000 ± 0.025
Calf thymus	43	48.6	0.435	-0.073 ± 0.028
Human placenta	41	41.1	0.342	0.076 ± 0.051
Salmon sperm	42	44.6	0.412	-0.040 ± 0.042
<i>M. luteus</i>	72	42.4	0.560	-0.225 ± 0.030
<i>C. perfringens</i>	26.5	46.5	0.370	0.027 ± 0.027
<i>E. coli</i>	50	50.4	0.447	-0.088 ± 0.052
100% GC ¹	100	—	—	-0.293 ± 0.085

For each DNA source are listed the base composition (%GC), the percentage of genomic DNA recovered in the nucleosome monomer fraction (%NM), the ratio of bound to free DNA in the reconstitution assay, and the differential free energy of reconstitution ($\Delta\Delta G$).

¹Value calculated from data in reference 4.

RESULTS

Efficiency of Reconstitution of Long Genomic DNAs

In order to perform the competitive nucleosome reconstitution assay, we required nucleosomal-length DNA fragments from the genomic DNAs that were to be tested. Fragments of the proper length could be generated simply by nuclease digestion of the genomic DNAs followed by fractionation with gel electrophoresis, but the possibility that nuclease sequence preferences would bias the sample made this approach unattractive. We therefore determined to generate the fragments by reconstitution of the longer DNAs with histones, followed by nuclease digestion and isolation of the nucleosomal DNA. This approach, however, also has the possibility of yielding a biased DNA sample if the reconstitution efficiency is low. In that instance nucleosomes may form predominantly on the strongest binding sites and subsequent competitive reconstitution would give a misleading value for the genomic DNA as a whole. Thus in our approach it is necessary to obtain efficient reconstitution of genomic DNAs.

As shown in Figure 1 and Table 1 the reconstitution procedure we followed (9) is quite efficient, so that in all cases >40% of the starting DNA is isolated from the nucleosome monomer peak after digestion of the nucleoprotein complex. The fact that all of the genomic DNAs we examined, both prokaryotic and eukaryotic, GC-rich and AT-rich, had reconstitution efficiencies between ~40%–50% suggests that that is the upper limit of bulk reconstitution of long, heterogeneous-sequence DNA *in vitro*. Thus we conclude that the nucleosomal-length DNA we isolated by the procedure described above consists of a random sampling of each genomic DNA.

This conclusion is confirmed by melting temperature experiments. The results from several such experiments are shown in Figure 2. Genomic and nucleosomal DNAs of *C. perfringens* melt over the same temperature range ($T_m = \sim 83 \pm 2^\circ\text{C}$), as do the genomic and nucleosomal DNAs of *E. coli* ($T_m = \sim 92 \pm 2^\circ\text{C}$). The genomic and nucleosomal DNAs from the other organisms listed in Table 1 also melt within $\sim 1^\circ\text{C}$ of each other (data not shown) showing that the formation of nucleosomal-length DNA by the method described above has not biased the base composition of the samples noticeably.

Competitive Reconstitution of Nucleosomal-Length DNAs

The results of the competitive reconstitution assay are shown in Figure 3 and Table 1. After separation by gel electrophoresis and autoradiography labeled test DNA can, as expected, be seen in the bands corresponding to both nucleosomal and free DNA

Figure 2. Percentage change in the absorbance at 260 nm vs. temperature of: genomic and nucleosomal DNAs of *C. perfringens* ($T_m = \sim 83 \pm 2^\circ\text{C}$); genomic and nucleosomal DNAs of *E. coli* ($T_m = \sim 92 \pm 2^\circ\text{C}$).

(Fig. 3). Quantitation of radioactivity shows that the ratio of counts in the nucleosomal band to that in the free DNA band ranges from a low of 0.342 for human placenta to a high of 0.560 for *M. luteus* (Table 1). Each test DNA was put through at least 5 separate reconstitutions and the reproducibility of the measurements was seen to be quite good, with standard deviations ranging from ~ 0.02 – 0.05 kcal/mol (Table 1) for the genomic DNAs. This is comparable to the precision reported by other laboratories for competitive reconstitution experiments (5,6).

The value of the ratio of counts in the nucleosomal band to that in the free DNA band for chicken erythrocyte DNA is 0.386. This number requires some explanation since the bulk unlabeled nucleosomes and DNA against which the test DNAs compete are also from chicken erythrocytes. Since the chicken DNA is competing against itself a ratio of 1.0 might be expected. However, the labeled chicken DNA (and the labeled DNA from other sources as well) was isolated as a narrower band of $\sim 150 \pm 10$ bp cut from a gel electrophoresis lane of nucleosomal DNA which is more heterogeneous in length (see Materials and Methods). When labeled, unfractionated, nucleosomal DNA is used as the test DNA in the competitive reconstitution assay a ratio of 1.0 is obtained. However, the unfractionated nucleosomal DNA also contains longer DNA, up to ~ 200 bp in length. The

Figure 3. Electrophoresis of competitively-reconstituted DNA. A. Lanes a) 123 bp size-standard ladder; lanes b, d, and f are reconstituted *E. coli*, *M. luteus*, and *C. perfringens* DNA, respectively; lanes c, e, and g are the corresponding free DNAs. B. Lanes a) 123 bp size-standard ladder; lanes b, d, f and h are reconstituted human placenta, calf thymus, salmon sperm, and chicken erythrocyte DNA, respectively; lanes c, e, g and i are the corresponding free DNAs.

Figure 4. Differential free energy of nucleosome reconstitution vs. genomic base composition.

longer DNA competes more effectively for nucleosome formation, as expected just from statistical considerations, and thus a decreased ratio of bound/free DNA is obtained with the fractionated DNA. This consideration does not affect the interpretation of our results, however, since only *relative* changes in the free energy of nucleosome formation ($\Delta\Delta G$) are reported, and all test DNAs compete against the same bulk, unlabeled nucleosomes and DNA.

Figure 4 shows the values of the ratios of bound/free test DNA plotted vs. the percentages of G+C found in the respective genomic DNAs. Values for DNA from prokaryotic sources are shown as circles, eukaryotes as squares, and an average value for 9 synthetic polymers composed exclusively of GC base pairs is shown as a triangle. The value for the synthetic DNA is taken from work previously published by our laboratory (4) in an investigation of polymers of the two series poly[d(G_nC_n)]·poly[d(G_nC_n)] and poly[d(G_nC)]·poly[d(GC_n)]. Linear regression of the data, but excluding the values for all eukaryotic DNAs, yielded a correlation coefficient of -0.984 and a slope of $-4.5 \text{ cal/mol} \cdot \%GC$. Eukaryotic DNAs were excluded from the regression analysis since they interact with histones *in vivo* and might be biased to favor or disfavor nucleosome formation. However, since prokaryotic DNA does not interact with histones *in vivo*, we believe it is valid to assume that it is neutral with respect to nucleosome formation and would exhibit the behavior expected for a 'random', heterogeneous sequence. Two of the eukaryotic DNAs, from calf thymus and salmon sperm, are less than one standard deviation from the least-squares line. Including the values for calf thymus and salmon sperm in the regression leaves the results virtually unaltered. However, the other two eukaryotic DNAs, from chicken erythrocyte and human placenta, lie, respectively, 2.2 and 2.6 standard deviations from the line.

DISCUSSION

The Competitive Nucleosome Reconstitution Assay

The competitive nucleosome reconstitution assay has been used successfully by ourselves (4) and by other laboratories (5,6) to evaluate the differential free energy of nucleosome binding of specific synthetic DNAs or restriction fragments. The assay has been shown to be sensitive, reproducible, and well-behaved, and to yield results which can be easily rationalized in terms of the known structure of DNA (4–6). Thus it seems to truly reflect changes in binding energy with the DNA component. Recently, however, a brief essay (10) questioned whether results obtained

using the competitive assay, since exchange occurs at $\sim 0.6\text{--}0.8 \text{ M NaCl}$, can be used to infer binding energies at physiological salt concentration. After all, it was argued (10), exchange occurs by definition under near-dissociating conditions.

The answer to this objection is already in the literature. Schrader and Crothers (5) pointed out that both the persistence length of DNA (11) and the sedimentation coefficient of the nucleosome (12) remain constant over the 2–3 fold change in ionic strength between physiological and exchange conditions. Thus the evidence argues that the structure of the intact nucleosome is essentially the same at an ionic strength of 0.2 as it is at an ionic strength of 0.8. Furthermore, if this is so then the free energy differences measured using the exchange assay should be essentially the same as those values at physiological conditions, regardless of the proximity to dissociating conditions. The reason for this is that dissociated histones do not enter into calculation of $\Delta\Delta G$ (see Materials and Methods), only intact nucleosomes and free DNA, both of which appear to be structurally unchanged between physiological conditions and exchange conditions, as mentioned above. The remainder of the Discussion in this paper, then, will assume that the competitive nucleosome reconstitution assay does reflect quantitatively $\Delta\Delta G$ values at physiological ionic strength.

The Free Energy of Nucleosome Formation Varies With Base Composition

As shown in Figure 4 there is clearly a dependence of the $\Delta\Delta G$ of nucleosome reconstitution on DNA base composition. The dependence could arise from any of a number of factors: GC and AT base pairs differ in their structures, physical properties, and hydration (13); short tracts composed predominately of GC or AT base pairs prefer to bend into their major or minor grooves, respectively (14); and the nearest-neighbor compositions of polymers, which account for much (15) of their physical properties, will vary *per force* with varying %GC. The difference in binding energy between a heterogeneous polymer composed entirely of AT base pairs vs. one composed entirely of GC base pairs is calculated to be $\sim 0.45 \text{ kcal/mol}$, based on extrapolation of the line in Fig. 4. The difference in binding between DNAs of 33% GC vs. 66% GC, a range more pertinent to eukaryotic DNA, is $\sim 0.15 \text{ kcal/mol}$. Although such energy differences are not large, they may be important. It has been shown that some sequences which phase nucleosomes *in vivo* form complexes only $\sim 1 \text{ kcal/mol}$ stronger than bulk chicken DNA (5), so large $\Delta\Delta G$ values may not be necessary to have biological effects. It is also interesting to recall that most higher eukaryotic DNAs are significantly AT-rich (3). Although this property may have other effects, it will also cause nucleosomes to be less stable than they would be if the DNA were GC-rich. In fact all of the eukaryotic DNAs examined here reconstitute noticeably less well than the GC-rich *M. luteus* (Fig. 4, Table 1).

Nucleosome formation of prokaryotic vs. eukaryotic DNA

Two of the eukaryotic DNAs, from calf thymus and salmon sperm, are less than one standard deviation from the least-squares line (Fig. 4), and including them in the regression leaves the results virtually unaltered. However, the other two eukaryotic DNAs, from chicken erythrocyte and human placenta, deviate significantly from the linear relationship describing the prokaryotic and synthetic DNA and lie, respectively, 2.2 and 2.6 standard deviations from the line. The deviation in both cases is in the direction of *less* stable nucleosomes. Human DNA forms

nucleosomes that are, on average, ~ 0.2 kcal/mol and ~ 0.3 kcal/mol less stable than particles formed with DNA of the prokaryotes *E. coli* and *M. luteus*, respectively. Since these are average values over the whole genome it is likely that large portions of the human genome bind nucleosomes with even less affinity, perhaps on the order of 1 kcal/mol less than a prokaryotic DNA.

As mentioned above, since prokaryotic DNAs do not form nucleosomes *in vivo* we assume that their genome is neutral with regard to interactions with histones. On the other hand, because eukaryotic DNAs do form nucleosomes *in vivo* we assume that their DNA sequences are optimized for histone-DNA interactions. Surprisingly, as shown in Fig. 4, this appears to mean a *decreased* stability of the eukaryotic nucleosome compared to complexes formed with prokaryotic DNA. Thus we conclude that the DNA sequence of eukaryotes is biased to interact less strongly with histones than it otherwise would.

At first glance this point of view appears odd in light of studies which have determined sequence preferences for nucleosome positioning (5,6,14) and studies which have shown nucleosome formation on preferred sites (16–19) and their effect *in vivo* (20). However, in those cases the studies were specifically looking for sequences which bound nucleosomes and bound them well. In this paper we consider the genome as a whole. Upon further reflection, then, there is no *a priori* reason for thinking that eukaryotic DNA sequences should be biased to favor or disfavor nucleosome formation since nucleosomes have to both form and dissociate during various genomic functions. Additionally, since the apparent purpose of some covalent modifications of histones, such as acetylation, is to weaken histone-DNA interactions (3), the purpose of nucleosome formation is clearly not to always form the tightest possible histone-DNA complex. Our results show that, for the eukaryotic DNAs considered here as a whole, weakened nucleosomal interactions are apparently preferred.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health grant GM36343 to MJB.

REFERENCES

1. McGhee, J.D. & Felsenfeld, G. (1980) *Ann. Rev. Biochem.* **49**, 1115–1156.
2. Broyles, S.S. & Pettijohn, D.E. (1986) *J. Mol. Biol.* **87**, 47–60.
3. van Holde, K.E. (1989) *Chromatin*. Springer-Verlag, New York.
4. Jayasena, S.D. & Behe, M.J. (1989) *J. Mol. Biol.* **208**, 297–306.
5. Shrader, T.E. & Crothers, D.M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7418–7422.
6. Shrader, T.E. & Crothers, D.M. (1990) *J. Mol. Biol.* **216**, 69–84.
7. Simpson, R.T. & Whitlock, J.P. Jr., (1976) *Cell* **9**, 347–353.
8. Simon, R.H. & Felsenfeld, G. (1979) *Nucleic Acids Res.* **6**, 689–696.
9. Daban, J.R. & Cantor, C.R. (1982) *J. Mol. Biol.* **156**, 749–769.
10. Drew, H.R. (1991) *J. Mol. Biol.* **219**, 391–392.
11. Hagerman, P.J. (1981) *Biopolymers* **20**, 1503–1535.
12. Stacks, P.C. & Schumaker, D.N. (1979) *Nucleic Acids Res.* **7**, 2457–2467.
13. Saenger, W. (1979) *Principles of Nucleic Acid Structure*. Springer-Verlag, New York.
14. Satchwell, S.C., Drew, H.R. & Travers, A.A. (1986) *J. Mol. Biol.* **191**, 659–675.
15. Breslauer, K.J., Frank, R., Blöcker, H. & Marky, L.A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
16. Simpson, R.T. & Stafford, D.W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 51–55.
17. Ramsay, N., Felsenfeld, G., Rushton, B.M. & McGhee, J.D. (1984) *EMBO J.* **3**, 2605–2611.
18. Stein, A. (1987) *J. Biol. Chem.* **262**, 3872–3879.
19. FitzGerald, P.C. & Simpson, R.T. (1985) *J. Biol. Chem.* **260**, 15318–15324.
20. Simpson, R.T. (1990) *Nature* **343**, 387–391.