

# Sequence and Position-dependence of the Equilibrium Accessibility of Nucleosomal DNA Target Sites

J. D. Anderson<sup>1</sup> and J. Widom<sup>1,2\*</sup>

<sup>1</sup>*Department of Biochemistry  
Molecular Biology and Cell  
Biology and*

<sup>2</sup>*Department of Chemistry  
Northwestern University  
Evanston, IL 60208-3500, USA*

We have previously shown that nucleosomes are conformationally dynamic: DNA sequences that in the time-average are buried inside nucleosomes are nevertheless transiently accessible, even to large proteins (or any other macromolecule). We refer to this dynamic behavior as “site exposure”. Here we show that: (i) the equilibrium constants describing this dynamic site exposure decrease progressively from either end of the nucleosomal DNA in toward the middle; and (ii) these position-dependent equilibrium constants are strongly dependent on the nucleosomal DNA sequence. The progressive decrease in equilibrium constant with distance inside the nucleosome supports the hypothesis that access to sites internal to a nucleosome is provided by progressive (transient) release of DNA from the octamer surface, starting from one end of the nucleosomal DNA. The dependence on genomic DNA sequence implies that a specific genomic DNA sequence could be a major determinant of target site occupancies achieved by regulatory proteins *in vivo*, by either governing the time-averaged accessibility for a given nucleosome position, or biasing the time-averaged positioning (of mobile nucleosomes), which in turn is a major determinant of site accessibility.

© 2000 Academic Press

\*Corresponding author

**Keywords:** nucleosome; chromatin; gene regulation; site exposure

## Introduction

The mechanisms whereby proteins gain access to DNA target sites in chromatin are not known. A problem that arises even at the lowest level of chromatin organization, the nucleosome, is that most of the DNA is inaccessible, yet essential processes such as transcription, replication, recombination, and repair must occur. Studies *in vivo* point to a need for ATP-powered molecular machines to “open chromatin”. Strikingly, however, studies from many laboratories reveal that nucleosomal DNA is accessible to apparently all DNA-binding proteins, even in the absence of exogenous factors, despite the fact that the crystallographic structure of the nucleosome shows that the binding sites would be occluded (Luger *et al.*, 1997).

In earlier work (Polach & Widom, 1995, 1996) we proposed a “site exposure model” that explains this behavior of nucleosomes *in vitro* and successfully predicts the results of widely disparate experiments, such as the equilibrium binding of a protein to a nucleosomal DNA target site, the kinetics of digestion of nucleosomal DNA at internal

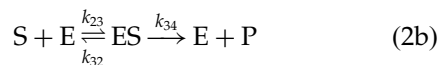
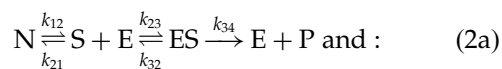
sites by restriction enzymes or at the ends of nucleosomal DNA by exonucleases, and the surprising cooperative binding of pairs of proteins to target sites contained within the same nucleosome. Let N represent a nucleosome in its native (most prevalent) conformation. The site exposure model supposes the existence of a transient state, S, of the nucleosome, at equilibrium with N, in which the nucleosome has transiently uncoiled some of its DNA, such that a DNA target site that was previously inaccessible is now freely accessible to a binding protein, R, or nuclease, E. These can bind to make a complex, RS or ES, the latter of which can go on to catalyze to yield products E and P. We suppose that sufficient DNA is uncoiled so as to make the subsequent steps on the site-exposed nucleosome equivalent to those on naked DNA. Thus:



for equilibrium binding to a nucleosomal target or

E-mail address of the corresponding author:  
j-widom@nwu.edu

naked DNA, respectively, and:



for nuclease digestion of nucleosomal or naked DNA, respectively.

Real nucleosomes *in vitro* behave as though such uncoiling processes are occurring constantly yet transiently, in a rapid pre-equilibrium. As a consequence of this behavior, even sites which are buried within the middle of the nucleosome are constantly but transiently accessible, with equilibrium constant for accessibility (site exposure)  $K_{eq}^{conf} = k_{12}/k_{21}$ . The apparent equilibrium binding affinity for proteins to nucleosomal target sites, and the observed rate constants for digestion of nucleosomal DNA, are reduced from their values on naked DNA by a factor equal to  $K_{eq}^{conf}$ . For further discussion of the site exposure model, see Polach & Widom, (1998) and Widom (1998).

In the site exposure model, equilibrium constants for site exposure are related to sums of free energies for all of the histone-DNA contacts that need to be broken in order to expose the site. This implies that the DNA sequence could be a major determinant of site exposure equilibrium constants. Sequences having especially high affinity for the histone octamer (or equivalently, strong nucleosome positioning sequences (Lowary & Widom, 1997)) should exhibit significantly reduced equilibrium site accessibility. DNA sequence motifs implicated in nucleosome positioning are found with statistically significant enhanced probability in eukaryotic genomes (Widom, 1996; Lowary & Widom, 1998) and are known to actually confer increased nucleosome-positioning power on the DNA segments that contain these motifs (Lowary & Widom, 1998; Thåström *et al.*, 1999). If such high-affinity sequences could strongly suppress site exposure equilibria, this could contribute importantly to gene regulation in two novel ways (see Discussion). Thus, one aim here is to test whether the choice of specific DNA sequence actually can have any significant effect on site exposure equilibrium constants.

A second aim of the study is to examine the position-dependence of equilibrium constants for site exposure. The site exposure model supposes that access to sites internal to a nucleosome is provided by progressive (transient) release of DNA from the octamer surface starting from an end of the nucleosomal DNA. Our earlier study analyzed sites within one side of the nucleosomal DNA, and obtained results consistent with this hypothesis:  $K_{eq}^{conf}$  decreased progressively with distance from the end of the nucleosomal DNA in toward the middle. However, the DNA constructs analyzed in that study had several limitations. They lacked

appropriate sites in the other half, hence they did not allow measurements to be made along the full nucleosomal DNA length. Moreover, each DNA construct used contained only a small number of sites; results from several related, but different, DNA constructs were combined to yield the final picture. Hence, a second goal here is to allow measurement of  $K_{eq}^{conf}$  over the full length of the nucleosomal DNA using a single nucleosome population, to facilitate a clear analysis of the nature of position-dependences to  $K_{eq}^{conf}$ .

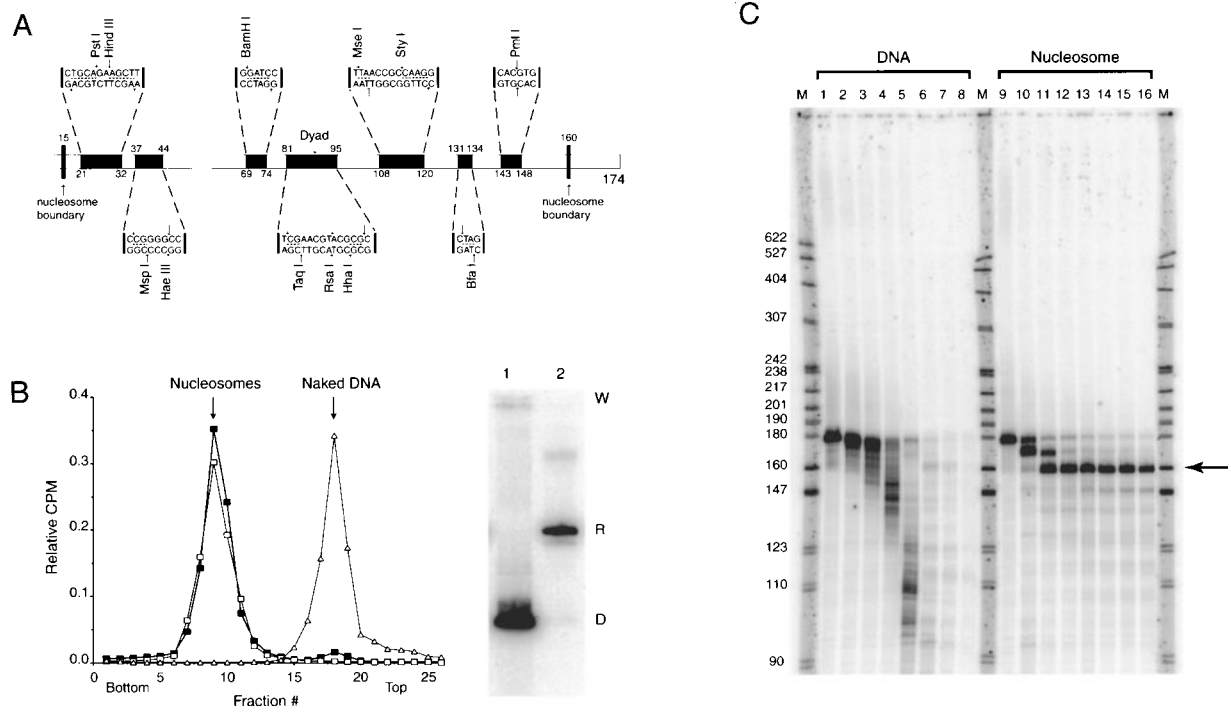
## Results

### Purification and characterization of reconstituted nucleosomes

A new DNA sequence (601.2, Figure 1(a)) was developed for these studies, based on the following considerations. (i) It is derived from a non-natural sequence that was selected for especially high-affinity binding to histone octamer in nucleosome reconstitution (Lowary & Widom, 1998). Such a sequence is also expected to be a strong nucleosome-positioning sequence (useful for creating homogeneous samples) and to exhibit dramatically any DNA-sequence-dependent effects on equilibrium constants for site exposure that might be obtained (Lowary & Widom, 1997). (ii) The DNA has short extensions at each end beyond the expected nucleosome position, used for the mapping studies that establish the actual nucleosome positioning. (iii) We make specific base changes to the DNA sequence to introduce sites for restriction enzymes at locations throughout the nucleosome. These restriction enzymes can then be used to measure the corresponding position-dependent equilibrium constants for site exposure, as described in our earlier studies.

Nucleosomes are prepared by gradual salt dialysis from 2.0 M NaCl and purified by sucrose gradient centrifugation. A typical sucrose gradient profile is shown in Figure 1(b). Reconstituted nucleosomes comigrate with native nucleosomes and are well resolved from residual naked DNA. Peak nucleosome-containing fractions were pooled and further characterized by native gel electrophoresis (Figure 1(b)). A single shifted band was observed, consistent with a strong statistical bias for occupancy by the histone octamer at a single preferred position along the DNA, as would be expected for a DNA sequence with particularly high affinity (Lowary & Widom, 1997). Quantitative analysis revealed negligible contamination of the purified nucleosomes by free DNA ( $\leq 0.5\%$ ).

Since we wished to quantify  $K_{eq}^{conf}$  as a function of position within the nucleosome, it was important to map where the nucleosome is. We anticipated that the high affinity sequence might facilitate mapping using exonuclease III by creating an especially strong "pause" at the protected core particle boundary (Protacio *et al.*, 1997). This proved to be the case. Whereas naked DNA was



**Figure 1.** Construction and characterization of reconstituted nucleosomes. (a) Schematic representation of DNA construct 601.2. The 174 bp-long DNA sequence is illustrated as a rectangle. The histone octamer was mapped (see the text) to a single detectable position covering positions 15 to 160 (bp), indicated by the black vertical bars. The corresponding location of the nucleosomal dyad symmetry axis is indicated. Sequences around each nested set of restriction sites are projected above and below the diagram, and specific restriction sites are indicated. The numbers bordering each hatched box indicate the precise locations of those stretches of sequence within the overall 174 bp-long sequence (bp 1 is defined as the left end of the DNA sequence). (b) Sucrose gradient purification and re-analysis by sucrose gradient (left-hand panel) and native gel electrophoresis (right-hand panel). Nucleosomes are reconstituted by gradual salt dialysis and separated from naked DNA and any non-nucleosomal contaminants on 5%-30% (v/w) sucrose gradients. ■, Preparative run of reconstituted core particles; △, marked DNA run in an identical gradient; □, re-analysis of gradient purified particles. Native gel analysis: W, the location of the loading wells; R, the mobility of the reconstituted nucleosomes; D, the mobility of naked DNA. Lane 1, naked DNA; lane 2, purified nucleosomes. The isolated nucleosomes show the appropriate mobility shift (in comparison to other nucleosomes and DNA size standards, not shown). Phosphorimager analysis of the gel reveals contamination by free DNA and other non-nucleosomal aggregates to be  $\leq 0.5\%$ . (c) Exonuclease III mapping. Aliquots from digestion reactions of singly 5' end-labeled naked DNA (left-hand panel) and nucleosome core particles (right-hand panel) were run on 8% denaturing polyacrylamide DNA sequencing gels. Samples were taken prior to initiation (0 minute), and at 1, 2, 4, 8, 16, 32, and 64 minutes after addition of enzyme, corresponding to lanes 1 to 8 for naked DNA and lanes 9 to 16 for nucleosome core particles, respectively. Markers (M) are *MspI*-digested pBR322 fragments (New England Biolabs). Naked DNA is digested rapidly to nucleotides; nucleosome core particles are digested rapidly to the edge of the particles (sharp new band, arrow), whereupon further digestion is greatly slowed. The next digestion pause site can be seen faintly in lanes 13 to 18, corresponding to digestion approximately 10 bp inward from the edge of the particle.

rapidly digested into small oligomers, the nucleosomal DNA was rapidly shortened to a new sharp band corresponding to the end of the core particle (Figure 1(c), lanes 11-16), which persisted for the remainder of the digestion. Similar experiments with the opposite strand labeled (not shown) yielded an independent determination of the other end of the nucleosome. In three separate pairs of such experiments, the nucleosome was mapped to the same stretch of DNA, from position 15-160 on the DNA, as illustrated in Figure 1(a). Evidently, exonuclease III is capable of near base-pair resolution mapping of the strongly preferred nucleosome position on such high affinity DNA molecules.

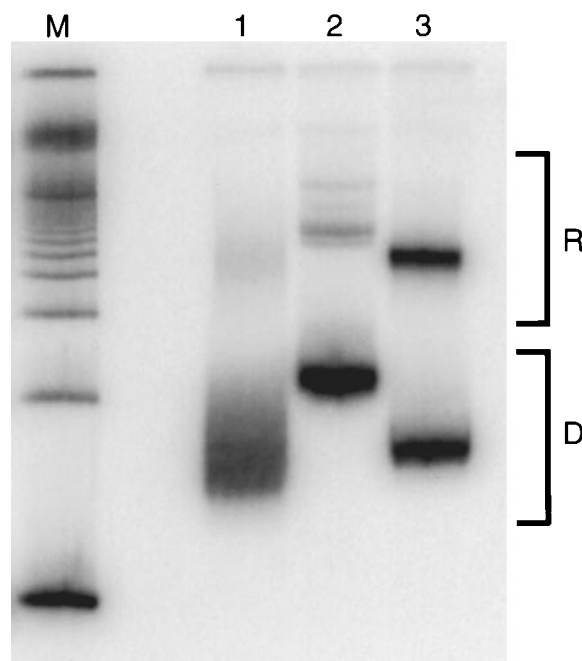
These exonuclease experiments were designed for mapping, hence were not carried out in such a way as to allow a quantitative measurement of  $K_{eq}^{conf}$ . Qualitatively, however, the protection that we observed for this sequence was far stronger than any that we have seen previously, consistent with a substantially decreased  $K_{eq}^{conf}$  at the core particle ends for this high affinity sequence (Protacio *et al.*, 1997).

### Free energy measurements

The DNA sequence designed for this new study (601.2) was based on a high affinity sequence (601) isolated in earlier studies (Lowary & Widom,

1998), but incorporated multiple nucleotide changes so as to introduce restriction sites throughout its length. It is likely that these sequence changes will decrease the affinity, hence it was important to measure the affinity of the new sequence. We also measured the affinities of the various sequences used in our earlier measurements of  $K_{eq}^{conf}$ , since these sequences also incorporated slight modifications of their corresponding natural 5 S RNA gene nucleosome positioning sequence. We used the standard competitive nucleosome reconstitution assay to measure the affinities (free energies) of these various DNA molecules for the histone octamer (Shrader & Crothers, 1989; Lowary & Widom, 1998; Thåström *et al.*, 1999). In this assay, trace amounts of radiolabeled DNA competes with a large excess of unlabeled competitor DNA for limiting amounts of histone octamer. The ratio of nucleosomal to free tracer DNA defines an equilibrium constant (affinity) and a corresponding free energy for the tracer that is valid for that competitive environment. To facilitate comparisons between experiments, we report difference free energies ( $\Delta\Delta G^0$ 's), measuring the free energy of the tracer relative to that of a reference sequence measured at the same time in the identical competitive environment. We used a fragment of the sea urchin 5 S ribosomal RNA gene nucleosome positioning sequence (Simpson & Stafford, 1983) as a reference tracer DNA.

The results of one such experiment are illustrated in Figure 2 and are summarized quantitatively in Table 1. An important control, in which the radiolabeled tracer DNA is a sample of the bulk core particle competitor DNA itself, is shown in lane 1. The ratio of counts in shifted (nucleosomal) products to total counts is in close agreement with the 1:20 molar ratio of octamer supplied in the reaction, confirming that the system is behaving properly. As a second control, we repeated our earlier measurements on the original 601 sequence. Despite differences in details of the competitive environments used in the two studies, our new



**Figure 2.** Native gel analysis of competitive reconstitution assays. Radiolabeled tracer competes with a large excess of unlabeled natural nucleosome core particle DNA for limiting quantities of histone octamer in dialysis from concentrated NaCl. Lane 1, some of the natural nucleosome core particle DNA is labeled and used as tracer. Lane 2, a derivative of the well-characterized natural nucleosome-positioning sequence from sea urchin 5 S rRNA gene (Simpson & Stafford, 1983). Lane 3, high affinity sequence 601.2. M, 100 bp ladder; D, mobility of naked DNA; R, reconstituted nucleosome mobilities. The diverse mobilities represent a range of positionings on the different molecules. Also, the core particle DNA tracer is not homogenous in length. The 5 S derivative yields several distinct nucleosomal positions, whereas 601.2 shows one predominant position. The raw data show that 601.2 competes much more effectively for the limiting histone octamer than does the 5 S sequence (greater ratio of counts in band R *versus* counts in D).

**Table 1.** Quantitative free energy measurements

DNA sequence	$\Delta\Delta G^0$ <sup>a</sup> kcal mol <sup>-1</sup> (n)
601.2	-1.0 ± 0.04 (9)
601 <sup>b</sup>	-2.8 ± 0.03 (6)
Construct a <sup>c</sup>	+0.19 ± 0.17 (10)
Construct b <sup>c</sup>	+0.30 ± 0.06 (8)
Construct c <sup>c</sup>	+0.39 ± 0.38 (6)

Values are expressed as mean ±SD; n, number of independent measurements. Whereas  $\Delta G$  values are specific to particular competitive environments,  $\Delta\Delta G$  values reflect intrinsic properties of the molecules independent of the particular competitive environment.

<sup>a</sup>  $\Delta\Delta G^0$  values measured relative to the 5 S reference sequence (Simpson & Stafford, 1983), and calculated as  $\Delta G_{sample}^0 - \Delta G_{5S}^0$ .  $\Delta G^0 (= -RT \ln K_{eq})$  values obtained using the indicated radiolabeled tracer with chicken erythrocyte core particle DNA as competitor.

<sup>b</sup> See Lowary & Widom (1998).

<sup>c</sup> See Polach & Widom (1995).

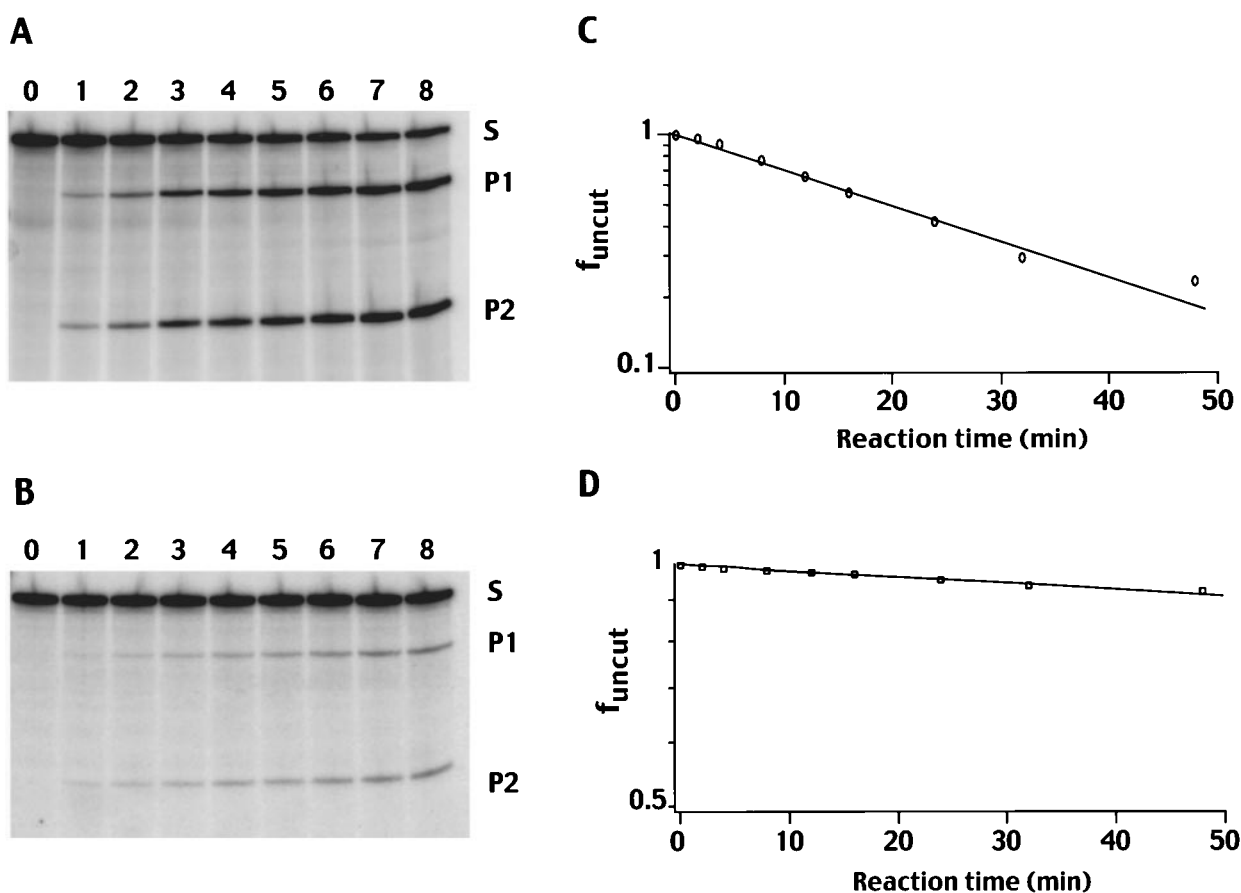
measurements are in close agreement with the earlier results, providing further evidence that the assay is behaving properly. Results for the 601.2 sequence are shown in lane 3, compared to the 5 S reference molecule in lane 2. The 601.2 sequence has a (negative) free energy 1.0(±0.04) kcal mol<sup>-1</sup> greater than that of the 5 S sequence. The three variants of the 5 S sequence used in our earlier studies have free energies that are positive (reduced affinity) relative to that of the unmodified 5 S sequence, with  $\Delta\Delta G$  values of +0.19(±0.17) kcal mol<sup>-1</sup>, +0.30(±0.06) kcal mol<sup>-1</sup>, and +0.39(±0.38) kcal mol<sup>-1</sup>. The corresponding affinity of sequence 601.2 is thus ~10-fold greater than the affinities of these 5 S sequence variants. The site exposure model predicts that  $K_{eq}^{conf}$  values will decrease correspondingly.

### Restriction enzyme kinetics assays for $K_{eq}^{conf}$

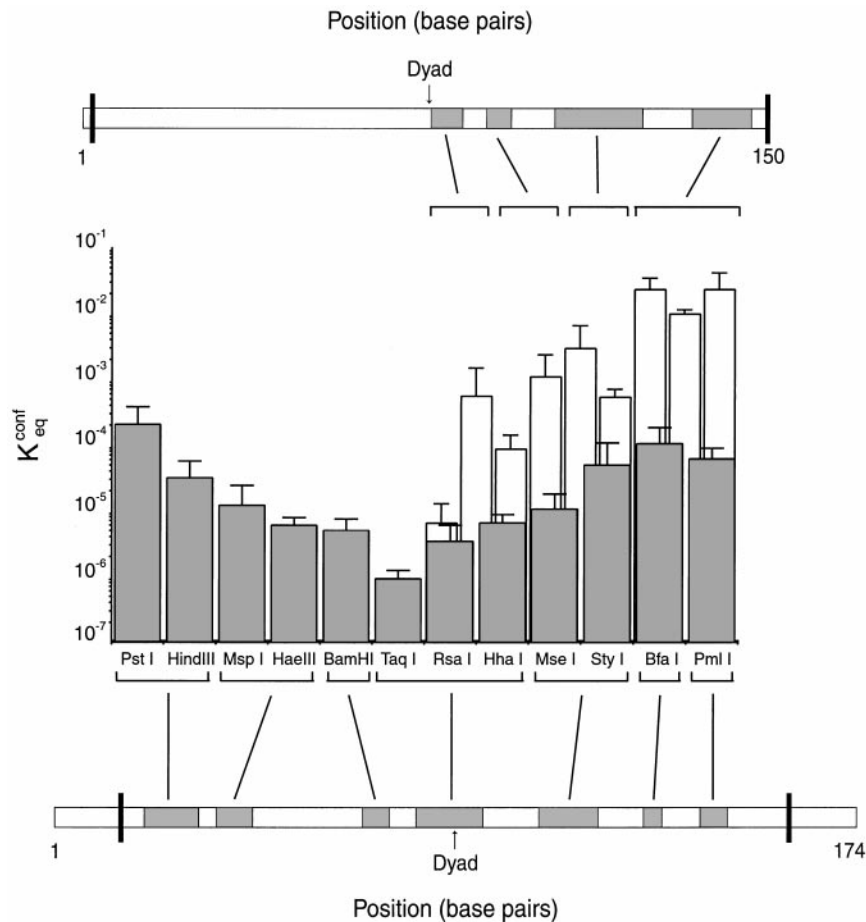
We used the restriction enzyme digestion kinetics assay (Polach & Widom, 1995) to measure  $K_{eq}^{conf}$  at target sites throughout the nucleosome. Parallel digestions of nucleosomal and free DNA were carried out in identical solution conditions. The results of one such experiment, probing a restriction site 40 base-pairs into the nucleosome, are illustrated in Figure 3(a) and (b) for naked DNA and nucleosomes, respectively. The corresponding rate constants for digestion obtained from exponential fits to plots of fraction of uncleaved substrate *versus* time (Figure 3(c) and (d)). In accordance with the prediction of the site exposure model, when compared to our earlier studies on the modified 5 S sequences, nucleosomes containing the high affinity sequence (601.2) are much more slowly cleaved by the restriction enzymes. Because of this increased resistance to

digestion, it is not possible to obtain cleavage of the entire sample over any reasonable period of time. Longer periods of digestion become problematic because of decay of the restriction enzyme activity, not to mention concerns about the integrity of nucleosomes were they to be left for long periods at 37°C. For these reasons, we monitor only the early stages of digestion of the nucleosomes. Possible concerns that these digested nucleosomes might not be representative of the entire sample are refuted by the overall pattern of results (see Discussion).

Values of  $K_{eq}^{conf}$  obtained from many such experiments are shown in Figure 4, tied to their corresponding nucleosomal location on the abscissa (bottom). These results are shown superimposed on our earlier results obtained using the modified 5 S sequences (tied to a separate abscissa, top), to facilitate a comparison.



**Figure 3.** Representative kinetic analysis, probing site-exposure at a site 40 bp from the right-hand end of the core particle. (a), (b) Denaturing polyacrylamide gel analysis of the time-course of digestion using the enzyme *StyI*. (a) Naked DNA, digested with *StyI* at 2 units  $ml^{-1}$ . (b) Nucleosomes, digested with *StyI* at 2000 units  $ml^{-1}$ . Lanes 0 through to 8 are obtained from samples removed 0, 1, 2, 4, 8, 12, 16, 24, 32, and 48 minutes for both naked DNA and nucleosomes. In each case, the substrate (174 nt, S) is converted over time to two products (120 nt, P1, and 54 nt, P2). The sizes of S, P1, and P2 expected from the DNA sequence are confirmed on other gels, not shown, in which their mobilities are compared to the mobilities of size standards. Size standards are routinely omitted on experimental gels simply to increase the number of timepoints that may be quantified under identical conditions. (c), (d) Quantitative analysis of the time-course of digestion from the data in (a) and (b), respectively. The fraction uncut is plotted on a log scale *versus* time. The superimposed lines represent the results of least-squares fits to single exponential decays.



**Figure 4.** Summary of measured equilibrium constants for site exposure ( $K_{eq}^{conf}$ ). Means and standard deviations are plotted as a function of location in the nucleosome, for the high affinity sequence 601.2 (dark shading, tied to the corresponding nucleosomal locations below the plot) and for the reference 5 S sequences (light shading, tied to their corresponding nucleosomal locations above the plot). The data for the 5 S reference sequences, including the corresponding nucleosome mapping studies, are from our earlier study (Polach & Widom, 1995).

## Discussion

Two aspects of the results are striking. First, the high affinity sequence exhibits greatly reduced  $K_{eq}^{conf}$  at all sites throughout the nucleosome. Compared to the sequences studied earlier, the high affinity sequence suppresses  $K_{eq}^{conf}$  by ten- to 100-fold throughout the nucleosome. Second, the overall pattern of  $K_{eq}^{conf}$  versus nucleosomal location is roughly symmetric about the mapped location of the nucleosomal center (dyad axis). Values of  $K_{eq}^{conf}$  decrease progressively from either end in toward the middle of the nucleosome. The implications of these two findings are discussed in more detail below.

### Are the digested DNA molecules representative of the sample?

A detailed set of control experiments testing alternative interpretations of the restriction enzyme assay have previously been reported (Polach & Widom, 1995). Those studies demonstrated, among other points, that: the fraction of DNA molecules

participating in the reactions was far greater than the fraction of DNA present as contaminating DNA; nucleosomes remained intact during the digestion; the reactions are first-order in (enzyme), consistent with the assumption of a rapid pre-equilibrium (and implying that the enzymes are not driving the process being monitored); and the rate does not depend on the presence of additional enzymes that lack specific sites (providing additional evidence that in the [enzyme] range used in these studies, neither specific nor non-specific enzyme binding contribute to the site exposure process being studied, but rather the enzymes serve as neutral probes of a naturally occurring dynamic conformational equilibrium).

Here, we focus on three points in particular. First, the accessibilities (actually, rates of cleavage) measured at all sites differ by three to six orders of magnitude from those of naked DNA (Figures 3 and 4). This observation directly rules out the hypothesis that the digestions occur only on pre-existing naked DNA. Also, the extents of digestion measured on these high affinity sequences are still much greater than the fraction of naked DNA

present in the sample, which is typically  $\leq 0.5\%$  (Figure 1(b), right-hand panel), again ruling out that hypothesis.

Second, as an additional check for the present studies, a mock-digestion control was carried out for each reaction to measure the amount of nucleosome dissociation occurring over the course of the experiment. An aliquot was removed from the sample tube prior to addition of restriction enzyme, and was mock-digested through the addition of identical reagents (with the exception of the enzyme itself) as in the actual digestion. After the real digestion reactions were complete, the mock digestion controls were immediately run on a native polyacrylamide gel. Typically, a very small amount of free DNA was detected, in the order of 0.5% or less (data not shown), far less than the total fraction of substrate digested in the real reactions (Figure 3). These results confirm that nucleosome dissociation contributes negligibly to the process monitored in the present study by the restriction enzyme digestion assay.

Finally, the collective results in Figure 4 refute the concern that the nucleosomes that are cleaved, i.e. those molecules actually contributing to the measurement, might not be representative of the entire sample, but might instead represent a small fraction of unstable particles, or perhaps nucleosomes with very different positionings. The results in Figure 4 show that, on the very same samples, there is a greater than 100-fold difference in  $K_{eq}^{conf}$  between differing sites. In order for non-representative species to account for this overall pattern of accessibility, one would have to argue that there are many distinct non-nucleosomal species, one for every different value of  $K_{eq}^{conf}$  obtained; furthermore, one would have to suppose that this ensemble of distinct species had the quite remarkable property of progressively decreasing the accessibility between the DNA ends and middle. We consider these scenarios as unlikely. Rather, in accordance with the presence of a single peak in a sucrose gradient (Figure 1(b), left-hand panel), a single band on native gels (Figure 1(b), right-hand panel), and a single strong pause site at each end during exonuclease III digestion (Figure 1(c)) (delineating, moreover, a  $\sim 146$  bp protected core region; Figure 1(a)), we conclude that the sample is homogeneous, and that the relatively limited extents of digestion followed during the kinetic analysis are nevertheless representative of all of the molecules in the sample.

### Site exposure *via* progressive displacement starting from one end

The progressive decrease in equilibrium constant with distance inside the nucleosome supports the hypothesis that access to sites internal to a nucleosome is provided by progressive (transient) release of DNA from the octamer surface, starting from one end of the nucleosomal DNA. A simple picture based on the structure of the nucleosome (Luger

*et al.*, 1997) that accounts for these results supposes that there are favorable contact-patches ("bonds") between DNA and histone residue every  $\sim 10$  bp, as the DNA backbone (minor groove) wraps around and faces the histone residue surface. Stepwise uncoiling, with each step corresponding to breaking one of these "bonds" (with associated cost in free energy), would lead to the observed position-dependent pattern of equilibrium accessibility.

Two distinct mechanisms could lead to such a progressive displacement of DNA starting from an end (Polach & Widom, 1995). One possibility is that the DNA partially uncoils, with the remaining contacts between histone octamer and DNA left unchanged. One can imagine this as partial uncoiling of a chain that is wrapped (without any slack) around a gear.

A quite different possibility is that the histone octamer could translocate along the DNA (also called sliding), moving along the DNA toward the left in order to progressively expose sites on the right, and *vice versa*.

The distinction between these two models is subtle. At present we have no direct evidence favoring or contradicting either one. We have previously pointed out (Polach & Widom, 1995) that an indirect argument supports the former (uncoiling) mechanism for site exposure: the rate of site exposure appears to be fast in comparison to the rate of sliding (Protacio *et al.*, 1997), suggesting that sliding cannot account for site exposure in those experiments. Indeed, we proposed instead that uncoiling could provide a mechanism for translocation. If DNA that had partially uncoiled from an end were recaptured at a different position along that DNA (most likely displaced by a roughly integral multiple of a DNA helical turn) to create a bulged loop, then propagation of the loop out to either end would result in a net translocation of the histone octamer. These two models could be distinguished in future studies that investigate the effects of appending stretches (e.g. nucleosome or half-nucleosome lengths) of naked DNA on apparent equilibrium constants for site exposure.

### Genomic DNA sequence: a major determinant of the equilibrium accessibility of regulatory target sites in chromatin

A second conclusion of the present study is that the choice of specific genomic DNA sequence is a major determinant of the equilibrium accessibility of target sites within the nucleosome. The DNA sequence influences equilibrium constants for site exposure over several orders of magnitude, and at sites throughout the nucleosome. This conclusion, in turn, implies that the genomic DNA sequence, through influences on site exposure equilibria, can be a major determinant of the actual occupancy of nucleosomal DNA target sites that will be achieved by the regulatory proteins that act at them.

This new conclusion complements our earlier studies which revealed a distinct way in which genomic DNA sequence can regulate the equilibrium accessibility of DNA target sites, namely, through influences on nucleosome positioning. Our earlier results establish that the accessibilities of sites within a nucleosome are strongly dependent on the location of the site within the nucleosome (decreasing by several orders of magnitude as one looks from one end of the nucleosomal DNA in toward the middle (Polach & Widom, 1995)) and that the preferred positioning of nucleosomes is linked to the positioning-dependent free energies of histone-DNA interactions (Lowary & Widom, 1997), with many natural sequences biasing this positioning by 100-fold or more (Thåström *et al.*, 1999). From these facts it follows that the choice of genomic sequence, through biases in the preferred positioning of nucleosomes, can bias by factors of 100-fold or more whether the equilibrium constant for site exposure of a particular target site is potentially  $\sim 1$  or greater, corresponding to preferential positioning of the target site in linker regions, or  $10^{-2}$ - $10^{-3}$ , corresponding to preferential positioning of the target site near the end of the nucleosomal DNA, or as low as  $\sim 10^{-4}$ - $10^{-6}$ , corresponding to preferential positioning of the target site near the nucleosomal dyad.

These two mechanisms by which the choice of genomic DNA sequence can strongly influence the equilibrium accessibility of nucleosomal DNA target sites are actually linked by the site exposure model: a subset of the histone-DNA contacts in the nucleosome that contribute to biasing nucleosome positioning (and thus influence site accessibility indirectly) also contribute to the energy penalty that must be paid to expose the site (thereby influencing site accessibility directly).

If it proves to be the case that nucleosomes are mobile *in vivo* (perhaps owing to the action of macromolecular complexes that catalyze nucleosome mobility (Hamiche *et al.*, 1999; Langst *et al.*, 1999)) and that this mobility occurs rapidly in comparison to time-scales of regulatory decisions, then equilibrium occupancy of regulatory proteins may need to be considered as a set of coupled equilibria. These equilibria may include the forces (free energies) of nucleosome positioning (which depends on DNA sequence (Lowary & Widom, 1998) and many other factors (Yao *et al.*, 1993)), site exposure (which again is sequence-dependent), and regulatory protein binding to DNA.

## Materials and Methods

### Preparation of DNA and histories

DNA sequence 601.2 was derived from sequence 601 (Lowary & Widom, 1998) using PCR to introduce specific base changes. Three oligonucleotides were used, in two successive steps of PCR synthesis (nucleotide changes are capitalized): R601, tacatgcacaggatgtatatatct; L601.1, gccctgCagaaGcTTggtCccgGggccgctcaattggtcgtag-

caagctctGg-ATccgcttGaTcgAacgtacgcgctgtccccc; and L601.2, ggaccctatacgcggGcgccctgCagaaGcTTggtC. A first step of PCR used the primer pair L601.1 and R601 with the original 282 bp-long sequence 601 as template. The product was purified by ion-exchange HPLC (Polach & Widom, 1995) and used as template in the second step of PCR with the primer pair L601.2 and R601, yielding 601.2. This product was also purified *via* ion-exchange HPLC. Histone octamer was purified from chicken erythrocytes as described (Feng *et al.*, 1993).

### Reconstitution and purification of nucleosomes

Nucleosomes incorporating radiolabeled tracer DNA were reconstituted by gradual salt dialysis and purified by sucrose gradient ultracentrifugation as described (Polach & Widom, 1995), using 200 ng specific DNA (including some radiolabeled as tracer), 19.2  $\mu$ g chicken erythrocyte nucleosome core particle DNA (as carrier), and 15.5  $\mu$ g purified chicken erythrocyte histone octamer in a 50  $\mu$ l volume. Peak nucleosome-containing fractions were pooled and exchanged into  $0.5 \times$  TE on Centricon-30 concentrators. The resulting samples were analyzed on an additional sucrose gradient and by native polyacrylamide gel electrophoresis.

### Equilibrium constants for site exposure ( $K_{eq}^{conf}$ )

Construct 601.2 was digested with the following enzymes, all at 37°C: *Pst*I, *Hind*III, *Msp*I, *Hae*III, *Bam*HI, *Rsa*I, *Hha*I, *Mse*I, *Sty*I, *Bfa*I, and *Pml*I. *Taq*I digestions were carried out at 65°C (for discussions of studies at higher temperatures, see Bashkin *et al.*, 1993; Polach & Widom, 1995). All enzymes were obtained in their most concentrated commercially available form from New England Biolabs, and all digestions were carried out with the buffer supplied by NEB. Glycerol was added to all naked DNA digests to the same final concentration present in the corresponding core particle digestion, and never exceeded 5% (v/v) in any reaction. At various timepoints during the reactions, 10  $\mu$ l aliquots were removed and quenched with 40 mM EDTA. The substrate DNA (S) and the two products (P1, P2) were resolved on denaturing polyacrylamide gels and quantified *via* phosphorimager. Background values were determined separately for each band and subtracted. The fraction uncut at each timepoint was defined as (counts in S)/(counts in S + P1 + P2). Apparent rate constants ( $k_{obs}$ ) for each digestion were obtained by non-linear least-squares fitting to a single exponential decay.  $K_{eq}^{conf}$  was obtained from the ratio of  $k_{obs}$  for nucleosomal and naked DNA digests, scaled for enzyme concentration (Polach & Widom, 1995).

### Exonuclease III mapping of reconstituted core particles

Exonuclease III was used to map the position of the histone octamer in reconstituted core particles. The two ends of the nucleosomal DNA are mapped separately using nucleosomes reconstituted with DNA labeled at a single end. Digestions were carried out on purified nucleosomes and on naked DNA (containing the same chicken erythrocyte DNA carrier) in 100  $\mu$ l reactions containing 75 nM nucleosomes or naked DNA and 100 units  $\text{ml}^{-1}$  exonuclease III in 50 mM Tris-HCl (pH 8.0), 5 mM  $\text{MgCl}_2$ , 5 mM  $\beta$ -mercaptoethanol. Reactions were initiated by mixing two 50  $\mu$ l half reactions, one contain-

ing enzyme, the other substrate, both preheated to 37°C. Volumes of 10 µl aliquots were removed over time, analyzed on a sequencing-size 8% denaturing polyacrylamide gel, and quantified by phosphorimager. Molecular Dynamics Imagequant software was used to determine the relative mobility of digestion products and (in comparison to molecular mass standards) the corresponding locations of the ends of the nucleosome. In three separate mapping experiments, the nucleosome position was mapped to be the same 146 bp stretch of DNA.

### Competitive reconstitutions and free energy measurements

Free energies for histone binding in nucleosome reconstitution were measured using the double-dialysis competitive reconstitution procedure as described (Lowary & Widom, 1998; Thåström *et al.*, 1999). Chicken erythrocyte core particle DNA (30 µg) and tracer amounts of the gel purified, radiolabeled DNA tracer were mixed with 2 µg of histone octamer in a total volume of 50 µl containing 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl, 0.5 mM PMSF, 1 mM benzamidine hydrochloride (BZA). The mixture was loaded into microdialysis buttons, which were then loaded into a dialysis bag containing approximately 200 ml of the same buffer. Samples were dialyzed for ≥ two hours at 4°C in the starting buffer, followed by two dialyses in 0.5 × TE, including PMSF and BZA, for ≥ 12 hours each at 4°C. Aliquots of each competitive reconstitution were run on 5% native polyacrylamide gels containing 1/3 × TBE (33 mM Tris-borate, 0.67 mM EDTA) and quantified by phosphorimager. Equilibrium constants were calculated as the ratios of background subtracted counts in nucleosomal bands (or sets of bands) to counts in free DNA bands, and free energies were calculated from the relationship  $\Delta G^0 = -RT \ln K_{eq}$ .  $\Delta \Delta G^0$  values represent differences between  $\Delta G^0$  values measured for a test sequence and a reference standard sequence, measured at the same time in the identical competitive environment. We used a fragment of the sea urchin 5 S ribosomal RNA gene nucleosome positioning sequence (Simpson & Stafford, 1983) as a reference tracer DNA.

### Acknowledgments

We thank P.T. Lowary and K.J. Polach for valuable assistance and discussions. This work was supported by a grant from the NIH (to J.W.) and by an NIH Cell and Molecular Basis of Disease Traineeship (to J.A.). We acknowledge with gratitude the use of instruments in the Keck Biophysics Facility, which was established with a grant from the W.M. Keck Foundation.

### References

Bashkin, J., Hayes, J. J., Tullius, T. D. & Wolffe, A. P. (1993). The structure of DNA in nucleosome core at

high salt concentration and at high temperature. *Biochemistry*, **32**, 1895-1898.

- Feng, H.-P., Scherl, D. S. & Widom, J. (1993). Lifetime of the histone octamer studied by continuous-flow quasielastic light scattering: test of a model for nucleosome transcription. *Biochemistry*, **32**, 7824-7831.
- Hamiche, A., Sandaltzopoulos, R., Gdula, D. A. & Wu, C. (1999). ATP-dependent nucleosome sliding mediated by the chromatin remodeling complex NURF. *Cell*, **97**, 833-842.
- Langst, G., Bonte, E. J., Corona, D. F. V. & Becker, P. B. (1999). Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell*, **97**, 843-852.
- Lowary, P. T. & Widom, J. (1997). Nucleosome packaging and nucleosome positioning of genomic DNA. *Proc. Natl Acad. Sci. USA*, **94**, 1183-1188.
- Lowary, P. T. & Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19-42.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251-260.
- Polach, K. J. & Widom, J. (1995). Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254**, 130-149.
- Polach, K. J. & Widom, J. (1996). A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites. *J. Mol. Biol.* **258**, 800-812.
- Polach, K. J. & Widom, J. (1998). Restriction enzymes as probes of nucleosome stability. *Methods Enzymol.* **304**, 278-298.
- Protacio, R. U., Polach, K. J. & Widom, J. (1997). Coupled enzymatic assays for the rate and mechanism of DNA site-exposure in a nucleosome. *J. Mol. Biol.* **274**, 708-721.
- Shrader, T. E. & Crothers, D. M. (1989). Artificial nucleosome positioning sequences. *Proc. Natl Acad. Sci. USA*, **86**, 7418-7422.
- Simpson, R. T. & Stafford, D. W. (1983). Structural features of a phased nucleosome core particle. *Proc. Natl Acad. Sci. USA*, **80**, 51-55.
- Thåström, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M. & Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* **288**, 213-229.
- Widom, J. (1996). Short-range order in two eukaryotic genomes: relation to chromosome structure. *J. Mol. Biol.* **259**, 579-588.
- Widom, J. (1998). Structure, dynamics, and function of chromatin *in vitro*. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 285-327.
- Yao, J., Lowary, P. T. & Widom, J. (1993). Twist constraints on linker DNA in the 30 nm chromatin fiber: implications for nucleosome phasing. *Proc. Natl Acad. Sci. USA*, **90**, 9364-9368.

Edited by T. Richmond

(Received 12 October 1999; received in revised form 30 December 1999; accepted 12 January 2000)