

## Open Access Article

The authors, the publisher, and the right holders grant the right to use, reproduce, and disseminate the work in digital form to all users.



Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102  
Volume 27, Issue Number 6, (2010)  
Current Perspectives on Nucleosome Positioning  
©Adenine Press (2010)

# A Structural Perspective on the Where, How, Why, and What of Nucleosome Positioning

<http://www.jbsdonline.com>

Gaurav Arya<sup>1\*</sup>  
Arijit Maitra<sup>1</sup>  
Sergei A. Grigoryev<sup>2\*</sup>

<sup>1</sup>Department of NanoEngineering  
University of California at San Diego,  
MC 0448, 9500 Gilman Drive, La Jolla,  
CA 92093

<sup>2</sup>Department of Biochemistry and  
Molecular Biology, Penn State Univer-  
sity College of Medicine H171, Milton  
S. Hershey Medical Center, P.O. Box  
850, 500 University Drive, Hershey, PA  
17033

## Abstract

The DNA in eukaryotic chromatin is packed by histones into arrays of repeating units called nucleosomes. Each nucleosome contains a nucleosome core, where the DNA is wrapped around a histone octamer, and a stretch of relatively unconstrained DNA called the linker DNA. Since nucleosome cores occlude the DNA from many DNA-binding factors, their positions provide important clues for understanding chromatin packing and gene regulation. Here we review the recent advances in the genome-wide mapping of nucleosome positions, the molecular and structural determinants of nucleosome positioning, and the importance of nucleosome positioning in chromatin higher order folding and transcriptional regulation.

## Introduction

Eukaryotic DNA is present inside cells in the form of a compact DNA-protein fiber called chromatin, which is composed of small repeating units called nucleosomes (1). Each nucleosome contains ~1.7 turns of DNA wrapped around an octamer of the histone proteins H2A, H2B, H3, and H4 (2, 3). Apart from compacting DNA, the chromatin fiber and its nucleosomes also serve an important function of regulating the accessibility of DNA for other macromolecules present in the cell nucleus. Therefore, nucleosomes and chromatin exercise direct control over DNA sequences and DNA-related processes like transcription, replication, recombination, and repair. It is now well known that posttranslational histone modifications, histone variants, remodeling enzymes, and various architectural proteins play an important role in regulating chromatin.

Emerging studies indicate that “nucleosome positioning”—defined as the probability that a nucleosome starts at a given base pair within the genome (4)—plays an equally important role in gene and chromatin regulation, alongside abovementioned mechanisms of regulation. In other words, nucleosomes may not be distributed randomly within the genomic DNA but may be deliberately positioned at specific regions or occluded from specific regions of the genome for some purpose. Nucleosome positioning is by no means independent of the other regulatory mechanisms like histone modifications and chromatin remodeling but is closely entwined with them. Nucleosome positioning may be characterized using four simple descriptors.

- (1) *Nucleosome repeat length* (NRL) is the average length of DNA associated with one nucleosome (linker and core DNA). Since the length of core DNA is fixed (~147 bp), the linker DNA dictates all the variability in NRL (Figure 1a).
- (2) *Nucleosome fuzziness* describes fluctuations in the likelihood of nucleosome positions on a gene locus obtained from an ensemble of measurements (Figure 1b).

\*Phone: 858-822-5542  
Fax: 858-534-9553 (GA)  
E-mail: garya@ucsd.edu  
sag17@psu.edu

Strongly positioned nucleosomes appear as sharp peaks in nucleosome positioning maps in contrast to fuzzy positioning where the peak locations are blurred due to nucleosome delocalization.

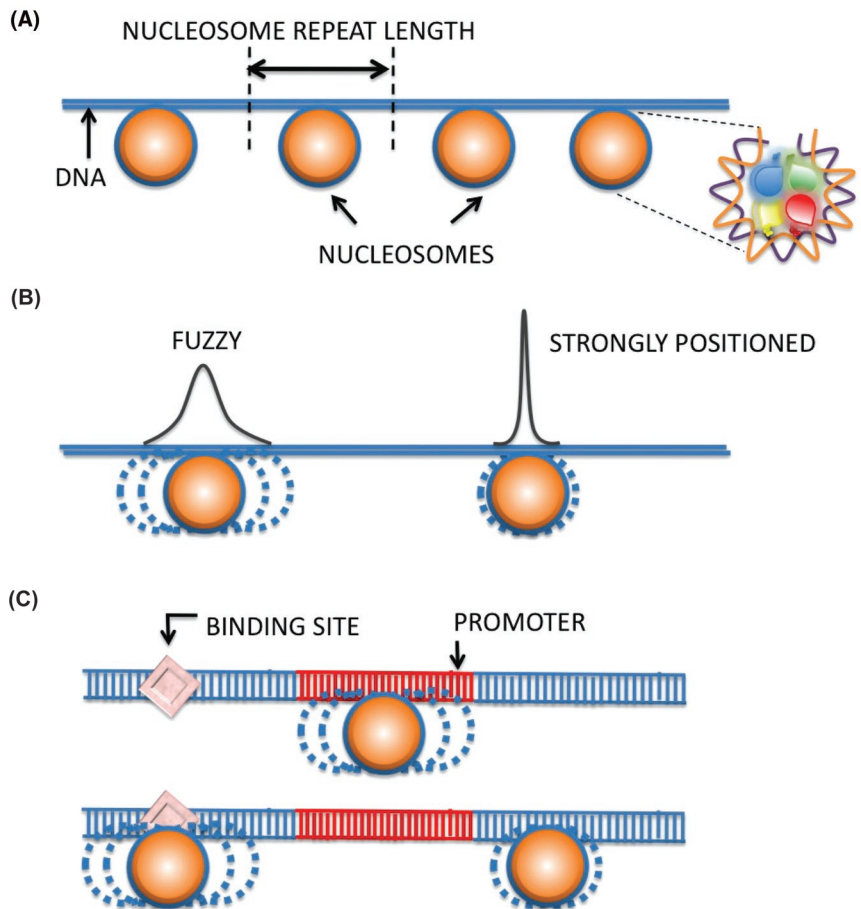
- (3) *Nucleosome occupancy* refers to the presence or absence of nucleosomes over specific DNA sequences in the genome, such as transcription factor binding sites (Figure 1c). Thus, nucleosome occupancy differs from nucleosome positioning in that the former does not care where the nucleosome starts as long as the given base pair is covered by it.
- (4) *Nucleosome phasing* describes an array of repeated nucleosome positions along DNA over certain genes or chromosomal domains (Figure 2c). Nucleosome phasing tends to reflect the NRL but in contrast to the latter is linked to a certain DNA sequence.

In this review, we will use the above definitions of nucleosome positioning to discuss where nucleosomes are positioned in eukaryotic genomes, how exactly are nucleosomes positioned, why are nucleosomes positioned from the point of view transcriptional regulation, and what is the role of nucleosome positioning in chromatin higher order folding.

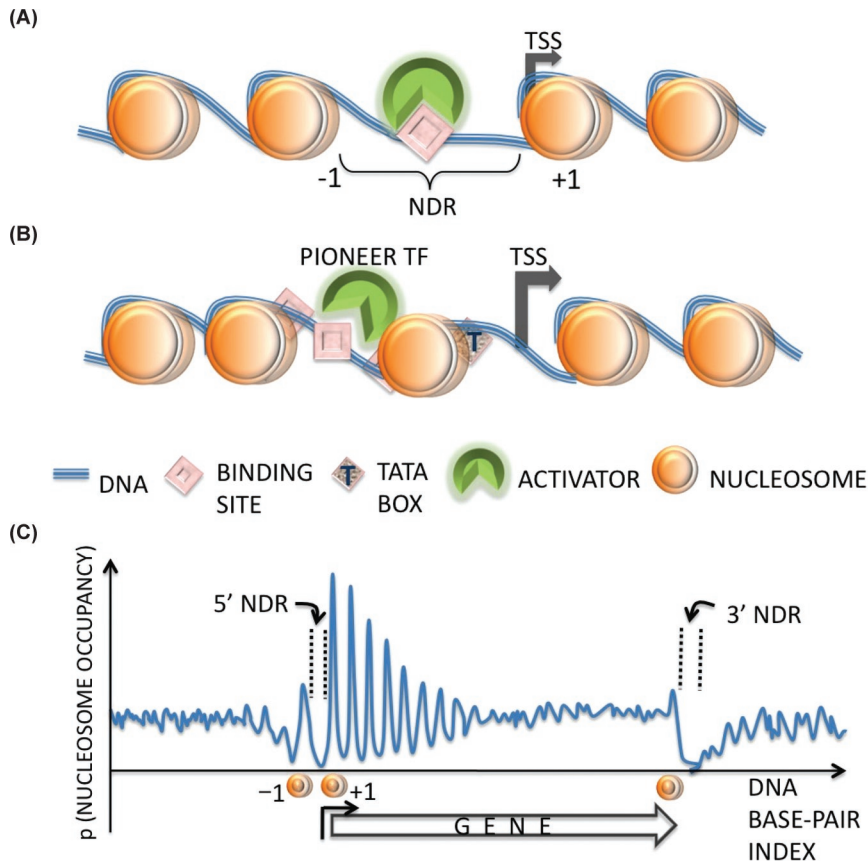
### Where are Nucleosomes Positioned?

#### Global Positioning of Nucleosomes

The nucleosome positioning landscape, especially the linker DNA length, varies strongly from one organism to another and between different tissues. On average,



**Figure 1:** Schematic of nucleosome locations on DNA depicting the nucleosome repeat length (A), nucleosome fuzziness (B), and nucleosome occupancy over or near binding sites (C).



**Figure 2:** Schematic depiction of open (A) and closed or occupied (B) promoters. Also shown is a typical nucleosomal position map of an open promoter showing clear demarcation of the gene locus, identifiable by the presence of the 5' and 3' nucleosome depleted regions (NDR). The transcription start site (TSS) is shown with a right-angled arrow. The 5' NDR is flanked by tightly bound nucleosomes. (C) A typical nucleosomal occupancy map showing clear demarcation of the gene locus, identifiable by the presence of the 5' and 3' nucleosome depleted regions (NDR). The transcription start site (TSS) is shown with a right-angled arrow. The 5' NDR is flanked by rather tightly bound nucleosomes.

yeast (5) and fly (6) chromatin has relatively short linkers (~20 and ~30 bp, respectively), human chromatin (7) has medium-sized linkers (~40 bp), chicken erythrocyte (8) chromatin has longer linkers (~60 bp), and echinoid sperm (9) chromatin has the longest known linkers (up to ~90 bp). *In vitro* the linker length depends on charge neutralization of DNA by counterions so that increased counterion concentration allows the nucleosomes to be more sparse (10). However, this mechanism does not explain tissue-related variations in the NRL, as the salt concentrations are more or less maintained constant *in vivo* across different tissues. The observed variability in linker lengths is more likely related to the transcriptional activity of the genome. For example, the highly transcriptionally active yeast genome has short linkers while the inactive echinoid sperm genome has long linkers (9). In multicellular organisms, the NRL may reflect tissue-specific and cell-specific changes in chromatin structure and gene activity. For example, during erythrocyte differentiation and maturation, the NRL increases from 190 to 212 bp (8).

Global positioning of nucleosomes may also depend on their chromosomal location. Nucleosomes are predicted to be depleted in regions such as the telomeres (5, 11) and strongly occupied within centromeres (11, 12). Nucleosomes are more populated within gene's coding regions as compared to intergenic or noncoding regions such as gene promoters (13). In general, nucleosome occupancy is found to correlate positively with transcriptional activity (13). This could either mean that the act of transcription leads to more crowded nucleosomes, perhaps being promoted by transcription elongation-associated nucleosome reassembly factors (14),

or that transcription requires formation of ordered nucleosome structure, maybe to increase the residence time of the polymerase at each site for attaining better fidelity. However, some highly transcribed coding genes like the ribosomal RNA and transfer RNA are less occupied by nucleosomes, as predicted by a computational model (11). There is clearly a lot of room for further studies to correlate nucleosome occupancy with gene types.

#### *Local Positioning Patterns*

More interesting sequence-specific features begin to appear when one examines the local position of nucleosomes across individual genes including the promoter regions. Powerful techniques like ChIP-seq (15, 16) and ChIP-chip (17, 18) can now be used to map nucleosome positions and their modifications across large portions of the genome (13, 19-21). After appropriate clustering of patterns of nucleosome positions at promoters, two general types of nucleosome phasing patterns appear that can be classified into (1) open promoters, and (2) closed or occupied promoters (22). Nucleosome positioning patterns are also found at the transcriptional termination regions and at the boundaries of active and repressed chromatin domains (23). Surprisingly, special nucleosome arrangements are also found at the gene regions corresponding to exon/intron junctions and RNA polyadenylation sites suggesting a link between chromatin organization and RNA processing (24-26).

**Open Promoters:** These promoters, which are generally TATA-less, have a clear and large (~150 bp) nucleosome depleted region (NDR) upstream of transcriptional start site (TSS). In yeast, the NDR is typically flanked by two very well positioned nucleosomes. The nucleosome downstream of NDR is commonly referred to as the “+1” nucleosome and the one upstream is called the “-1” nucleosomes (Figure 2A) (5, 20, 22, 27). The edge of the +1 nucleosome tends to be positioned just over the TSS. Interestingly, fly promoters have a slightly different open promoter: the +1 nucleosome is shifted more downstream (~40-60 bp) of the TSS compared to yeast (6). The nucleosomes downstream of the TSS exhibit strong phasing that decays in a “ripple-like” manner with distance from the TSS; by 5-10 nucleosomes, the phasing becomes quite fuzzy. The pattern of nucleosomes upstream of the TSS generally loses its periodicity much more quickly (within 2-3 nucleosomes). The average occupancy of nucleosomes is often higher downstream of TSS (within the gene) as compared to the intergenic regions upstream. The nucleosomes near the TSS generally contain the H2A.Z variant instead of histone H2A (28).

**Occupied Promoters:** Here the promoters do not contain a clear NDR signature but are rather either fully occupied by nucleosomes (13, 29) or contain a gradient of increasing nucleosome occupancy downstream of the TSS (13) (see Figure 2B). The TSS and most of transcriptional activator binding sites are generally covered, and one site, either in the linker DNA or in the core DNA near its entry/exit site is exposed (30). A classic example of such a promoter is the PHO5 promoter (31). In general, such promoters contain the TATA box, which is placed ~25-125 bp from TSS, and is often buried inside the nucleosome edge (22).

**Transcriptional Terminators:** Recently, it has been demonstrated that a distinct but less striking nucleosome-positioning pattern also exists at the 3' termination site of genes. Specifically, a well-positioned nucleosome occupies the termination site followed by a clear NDR in the intergenic region downstream (5, 32).

**Chromatin Insulators:** Insulators or boundary elements are DNA sequences that functionally segregate neighboring genomic sites. They can, for example, disconnect a promoter from the nearby enhancer or separate repressed chromosomal domains marked by histone H3K27 methylation from active chromatin marked by histone acetylation (23, 33). CTCF is the influential universal insulator-binding

protein whose binding is essential for insulator activity and plays an important role in structural and spatial organization of the genome in the interphase nucleus (33). At vertebrate insulators, CTCF organizes strings of ~10-15 nucleosomes symmetrically positioned at the two flanks of the site occupied by CTCF (23, 34).

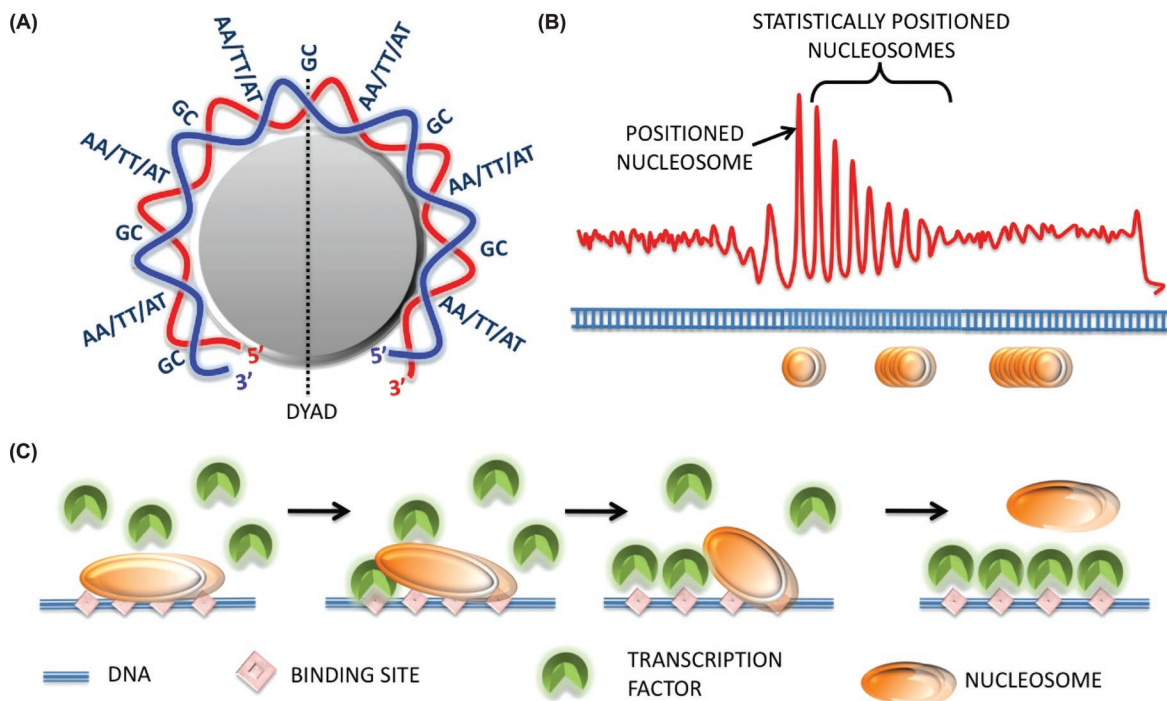
### How are Nucleosomes Positioned?

#### DNA Sequence Effects

The existence of special DNA templates with defined repeating sequence patterns, such as the clone 601 developed by Widom (35) clearly emphasizes that the nucleosomes can be positioned *in vitro* based on the DNA sequence alone. A recent study showing that the yeast genome is nine times more likely to be occupied by nucleosomes *in vitro* compared to that of E-coli also emphasizes the same theme (36). However, the extent to which *in vivo* nucleosome positioning is dictated by the DNA sequence is a continuing debate that is not fully resolved yet.

**DNA Sequence Bendability:** The nucleosome's preference for specific DNA sequences does not arise from its favorable interactions with particular DNA bases but rather arises from the sequence-dependent mechanics of DNA. The most favorable DNA sequences for positioning nucleosomes contain AA/TT/AT and GC dinucleotides occurring in 10 bp intervals, which are offset with respect to each other by 5 bp (see Figure 3a) (4, 11, 27, 37-40). It is believed that the former pattern allows expansion of the major groove of DNA while the latter pattern allows its contraction to facilitate the overall strong bending of DNA in nucleosomes. Additionally, there has been noted a gradient in AA/TT across the nucleosome binding region (39). The role of such a gradient remains a mystery but we speculate that it may be involved in dictating the polarity of transcription (41).

The Widom and Pugh groups developed the first set of computational models (11, 27) to predict the genome-wide location of nucleosomes, based on correlating genomic



**Figure 3:** (A) Nucleosomal DNA shows repeated alternating motifs of AA/TT/AT dinucleotides and GC dinucleotides. (B) One strongly positioned nucleosome could act as a physical barrier leading to a gradient in occupancy of nearby nucleosomes. (C) Competition between transcription factors (TF) and nucleosomes vying for the same regulatory sites of the DNA.

sequence to experimentally measured nucleosome positioning sequence patterns. The models correctly predicted the positions of more than half of the nucleosomes in the genomic regions examined, and also reproduced experimentally obtained nucleosome-positioning patterns at gene promoter regions (27). In fact, the results led Widom and coworkers (11) to propose the existence of a “nucleosome-positioning” code in the DNA. More recent efforts are directed towards combining different structural and energetic parameters related to DNA sequence to improve the accuracy of the predictions (13, 42-45). Molecular modeling and simulations also hold great promise for discovering new sequence-related parameters that could improve the current methods of predicting nucleosome positions. In general, the models cannot predict the position of all nucleosomes, indicating that the DNA sequence likely codes only for the positions of a limited number of nucleosomes, and that factors other than DNA sequence may also be involved in the positioning of nucleosomes.

Indeed, it is increasingly becoming clear that only a fraction of the nucleosomes need to be positioned to achieve phasing. Several years ago, Trifonov and coworkers found that there exists a strong nucleosome positioning sequence every four nucleosomes in parts of the human genome (46). Though this may not be universally applicable, the authors raised an interesting proposition that periodic occupancy of nucleosomes may not require a positioning code for each and every nucleosome but may be achieved through selective positioning of a few nucleosomes. They provided an analogy with a parking lot where a few well-parked cars promote the proper parking of the rest of the cars. This allows nucleosomes to position almost equally well without adding too much constraint on the genomic sequence. Recently, Peckham *et al.*, found that indeed only a small subset of nucleosomes is positioned by intrinsic DNA sequences using a computational model trained on known nucleosome binding affinities (44).

Several studies have shown that the DNA sequence strongly codes for both the presence of nucleosomes at the  $-1$  and  $+1$  positions of the promoter (5, 45) and the depletion of nucleosomes in the NDR region of the promoter (43, 47). In yeast, the latter is often composed of poly(dA:dT) tracts, which are suggested to exclude nucleosomes (30). Some studies have also indicated absence of a positioning code within the coding regions (5), and also suggest that the 3' terminating nucleosome is also positioned according to the DNA sequence. A recent study however raises some concerns about the importance of DNA sequence in positioning nucleosomes, especially for the  $+1$  H2A.Z nucleosome at the promoters (36). It has been proposed that factors like binding of polymerase enzymes could instead contribute to the positioning.

**Linker Histone Binding:** Another way the DNA sequence could bias the positioning of nucleosome is by constraining the position of linker histone binding. Recently, Cui and Zhurkin (48, 49) have found that there is an extension of the AA/TT/AT pattern beyond the nucleosomal DNA (147 bp) in chicken and flies that have a high linker histone content while yeast genomes, which possess low linker histone content, lack it. This suggests that linker DNA sequence could mediate linker histone binding and determine the nucleosome core/linker boundaries. This discovery implies that nucleosome-positioning predictions need to account for such additional patterns at the nucleosome boundaries for proper positioning of the nucleosome in the genome.

#### *Physical Forces of Positioning*

Below we discuss four purely physical effects that could contribute to the positioning of nucleosomes *in vivo*.

**Statistical Positioning:** Mavrich *et al.*, have recently suggested that the nucleosome phasing observed in genes is primarily dictated by statistical positioning principles (5).

In other words, the presence of a physical barrier restricts the position of the neighboring nucleosome, which then acts as a barrier for the next nucleosome, thus starting a chain of positioning restrictions that decreases with the distance from the original barrier (50-52). Hence, the strong positioning of the +1 and -1 nucleosomes at the TSS could respectively drive the phasing of downstream and upstream nucleosomes through the statistical positioning effect. Interestingly, this effect is predicted to get stronger with increasing nucleosome density (occupancy). Indeed, a longer-ranged nucleosome phasing is observed on the +1 side compared to the -1 side, consistent with the higher nucleosomal density on the +1 side. Another prediction is the increase in nucleosome fuzziness with the lower density of nucleosomes in highly transcribed regions, which is also observed across the genome (5).

**Polymerase Binding:** Recent evidences suggest that nucleosome phasing downstream of the +1 nucleosome at the TSS is related to transcriptional initiation and polymerase binding. First, the +1 nucleosome position is closely linked to the pol II binding site in fly genome (5). In fact, the relative distance between the +1 nucleosome and pol II site is conserved across the yeast and fly genome even though the absolute position of the nucleosome relative to TSS is different, indicating a strong correlation between the two (36). Second, detailed studies indicate strong positioning of the +1 nucleosome and phasing of downstream nucleosomes with stalled and elongating pol II (though the position of the +1 nucleosome is further downstream of the TSS in the latter) and little positioning and/or phasing is observed in the absence of pol II (7). Third, the strong positioning of the +1 nucleosome and phasing downstream is not observed when nucleosomes are assembled on the same genomic template *in vitro* using well-defined procedures, suggesting that *in vivo* factors such as pol II binding and not the DNA sequence may be dictating the positioning of the +1 nucleosome. Fourth, the statistical positioning of nucleosomes is observed only downstream of the TSS, *i.e.*, along the transcription direction, suggesting that it is somehow linked to the polarity of pre-initiation complex. In contrast, nucleosomes interacting with transcribed pol II are randomly positioned indicating that elongating pol II displaces the nucleosomes (53). Thus, the exact mechanism of how pol II binding could induce such positioning of nucleosomes is unclear. It may be related to the statistical positioning phenomenon, where the bound pol II acts like a physical barrier, or to the recruitment of chromatin-remodeling factors by the pre-initiation complex. Clearly, further research is required to shed more light into this effect.

**Nucleosome Packing and Interaction Constraints:** The folding of a one-dimensional (1D) array of nucleosomes into a three-dimensional (3D) chromatin fiber could also impose severe constraints on allowable nucleosome positions, *i.e.*, nucleosome positions that would lead to overlap in the 3D space would not be favored in the 1D space. Moreover, attractive internucleosomal interactions mediated by the histone tails (54, 55) and repulsive interactions between the wound DNA on nucleosomes could also significantly influence the 3D (and therefore 1D) arrangement of nucleosomes. Segal and Widom (4) have proposed that there must exist a competition between favorable DNA sequences promoting nucleosome binding at specific positions and steric, attractive, and repulsive interactions between nucleosomes. They further suggest that this competition becomes strong at high nucleosome packing densities where the latter forces become strong. In fact, a similar concept was used by Widom to explain why the linkers exhibit a ~10 bp quantization in their length (56).

**Interactions with Transcription Factors:** Several researchers have also proposed that binding of DNA-associated proteins such as transcription factors (TF) could compete with nucleosomes for DNA binding sites. Indeed, many transcription factor binding sites are found at the nucleosome-depleted regions such as the linker DNAs (4, 57, 58). Mirny (59) has recently argued through a theoretical model that clusters of closely spaced TF binding sites could act cooperatively

(even when the TFs do not interact with each other) to displace nucleosomes when the TF concentration rises beyond a threshold limit. Specifically, TFs prefer to bind to clusters of closely spaced TF-binding sites over sparsely populated binding sites. This leads to a greater number or choice of binding positions for nucleosomes, and thereby an increase in the translational entropy of nucleosomes. Segal and Widom (4) have suggested that the equilibrium between nucleosome and TF binding could explain why the positions of a few nucleosomes change drastically across different cellular events while those of others remain unchanged. It was proposed that histone modifications or DNA methylation could tilt this equilibrium balance between nucleosome-TF binding for some nucleosomes, leading to drastic changes in their positions but the modifications/methylation may not be strong enough to change the balance for other nucleosomes, which remain unaffected.

Other factors may help promote nucleosome occupancy over certain genomic regions by binding to the nucleosome rather than competing with it. For example, Rap1 protein has been shown to bind *in vivo* to the rotationally exposed major grooves of wound DNA of -1 nucleosome in yeast, causing these nucleosomes to exhibit stronger phasing compared to non-bound nucleosomes (53). Similarly, it has been shown that another transcription factor, Bdf1, binds to the +1 and +2 nucleosomes in active yeast genes (53). Bdf1 is a component of a protein complex (SWR1) responsible for H2A.Z deposition (60) and thus may promote nucleosome phasing by recruiting this histone variant as discussed below.

#### *Enzymatic Action*

Below we discuss the important role of ATP-consuming chromatin remodelers and the putative effects of histone modifications and histone variants on nucleosome positioning.

**Chromatin Remodelers:** The SWI/SNF and ISWI are the two main families of remodelers that have been studied in detail (29, 61, 62). The SWI/SNF remodelers are generally disruptive and can both slide and eject nucleosomes (61). They are often linked to promoter activation because of their involvement in chromatin disorganization. In fact, several SWI/SNF remodelers possess a bromodomain that allows them to target the -1 nucleosome with acetylated histone tails, thus playing an important role in promoter activation (63). These remodelers have been speculated to be involved in the generation of the NDR in open promoters though this has not been proven yet (22). The ISWI remodelers like ACF, on the other hand, are not disruptive to nucleosomes and are involved in positioning of nucleosomes. The exact mechanism by which these remodelers promote uniform linker lengths is not known but it is related to their need to bind to large portions of the entering/exiting linker for efficient functioning (64). Also, the enzymes typically remodel nucleosomes lacking the H4K16 acetylation mark (65) thus linking them to transcriptional repression. Furthermore, they have the capability to move nucleosomes to unfavorable DNA sequence elements (66), *e.g.*, poly(dA:dT) tracts in yeast genome, leading to occlusion of TF-binding sites and gene repression. Recent data show that both the major types of remodelers move nucleosomes through a sequence-independent mechanism though local sequences contribute to local variations between final nucleosome positions (67).

**Histone Variants and Posttranslational Modifications:** It is possible that posttranslational histone modifications could directly alter the interaction between DNA and histones to modify the DNA sequence preference of nucleosomes. For example, acetylation of K115 and K122 residue on the H3 histone could affect histone-DNA interactions, which also affects nucleosome repositioning and their assembly/disassembly (68). The modifications could also indirectly affect nucleosome repeat lengths through their modulation of



internucleosomal interactions (55, 69), interaction with chromatin remodeling complexes (63, 65), and interaction with other architectural proteins (70). Replacement of major histones by histone variants that modulate DNA/histone interactions and alter the internucleosomal interactions mediated by the histone tails could also affect nucleosome positioning in a similar way. For example, the CENP-A variant inserts two amino acids in a loop directly contacting the DNA (71), and the H2A.Z variant could promote internucleosomal interactions and intra-fiber folding through the stronger acidic patch on its surface (72-74). Finally, DNA methylation also has the potential to alter the preference of nucleosomes for certain DNA sequences as the addition of an extra methyl group at CpG regions on DNA could decrease the bendability of DNA through introduction of a bulky group (75).

### ***Why are Nucleosomes Positioned at Gene Promoter Regions?***

As discussed earlier, nucleosomes exhibit specific distributions near gene promoters. That these distributions occur across the genome, and across different organisms, suggests their supreme importance. Moreover, these patterns are most likely involved in transcriptional activation given that they align with respect to the TSS and not the rest of the gene's open reading frame (27). The precise placement of nucleosomes at promoter binding sites could be involved in modulating TF binding, which directly contributes to the assembly of the transcriptional machinery at the promoter, and controlling abortive rounds of short RNA synthesis that are often associated with inactive promoters (76). Here we discuss the function and dynamics of nucleosome patterns at occupied and open promoters from the point of view of transcription initiation.

#### *Nucleosome Positioning in Occupied Promoters*

These promoters are either fully or almost fully occupied with nucleosomes. In particular, the TSS and most of transcriptional activator binding sites are covered, and one site, either in the linker DNA or in the core DNA near its nucleosomal entry or exit region, is exposed for the so-called "pioneer" TFs. Other buried sites possibly require chromatin remodelers and histone modifications for accessibility. Since this reduces the risk of unregulated transcription, covered promoters are often associated with genes that require strict regulation (13), *e.g.*, stress-responsive genes like PHO5 (77, 78). In general, such tightly regulated genes contain TATA-boxes and exhibit strong nucleosome phasing (22, 79). Also, the transcription rate correlates negatively with nucleosome occupancy. In yeast, specifically, the TATA box is located ~40-120 bp upstream of the TSS (22, 80) and is often buried inside the edge of nucleosome providing partial blockage (81). In such cases, the accessibility could be regulated through thermal fluctuations of the entering/exiting linker DNA, as proposed by Widom and coworkers (82).

#### *Nucleosome Positioning in Open Promoters*

Open promoters have a long NDR upstream of TSS that is flanked by +1 and -1 nucleosomes, where the +1 nucleosome is positioned just over the TSS. The phasing of nucleosomes decays downstream and upstream with increasing distance from the TSS. Such promoters are generally associated with constitutive genes (22) that do not require strict regulation, *i.e.*, they can afford to have permanently open TF binding sites. Additionally, these promoters are typically TATA-less and the transcription rate generally is positively correlated with nucleosome occupancy (22). These observations suggest that the primary purpose of nucleosome positioning in open promoters is to produce an NDR that can facilitate recognition and binding of promoter sequences by the TFs for transcription, thus explaining why transcription correlates positively with nucleosome occupancy. In contrast, the purpose of nucleosome positioning in closed promoters of tightly regulated genes is to impede

TF binding; hence, transcription correlates negatively with nucleosome occupancy. It should however be noted that the extent to which nucleosomes inhibit binding of DNA-binding proteins in chromatin varies from protein to protein and likely depends on their DNA-binding strength, sequence specificity, and size.

The 3' transcriptional termination site of many genes also contains a well-positioned nucleosome followed by an NDR (5, 32). It is speculated that the NDR is needed to facilitate disassembly of polymerase machinery or anti-sense pre-initiation complexes (5). Selective enrichment of TFIIB at 3' NDR also supports a gene looping model, an interesting hypothesis that needs to be tested further (5). We envision that similar nucleosome positioning patterns could play key regulatory roles in other genomic processes, which are awaiting discovery.

#### *Patterns in Histone Modifications and Histone Variants at Promoters*

The H2A.Z histone variant is closely associated with positioned nucleosomes. In yeast, H2A.Z is present in the +1 and -1 nucleosomes (6, 20). In flies, H2A.Z is present only in the +1 nucleosome and several downstream nucleosomes (6). In humans, H2A.Z is present in +1 and -1 nucleosomes and several nucleosomes downstream and upstream, respectively (7, 22). In yeast and humans, the H2A.Z variant is preferentially lost from the promoters upon transcriptional induction (7, 83). This might be due to the destabilizing effect of H2A.Z on nucleosome core structure (84) so that nucleosomes containing this histone variant can be removed more easily. Another property of H2A.Z that could bear importance in gene regulation is that, compared to H2A, it promotes fiber compaction across nucleosomes within the same fiber but inhibits nucleosome interactions between fibers thus maintaining the nucleosome array in a dynamically compact state (72-74) in contrast to statically compact self-association.

The promoter nucleosomes are also subjected to activator posttranslational modifications such as methylation of the lysine 4 on the H3 tail and multiple lysine acetylations. Specifically, H3K4 on -2, -1, +1, and +2 nucleosomes are trimethylated; the +3 and +4 nucleosomes are dimethylated, and the +5 and +6 are monomethylated (7). The modification at the +1 nucleosome ensures that these histone modifications provide binding platforms for histone code-reading TFs needed for transcriptional activation (85).

#### *Nucleosome Positioning Dynamics upon Gene Activation or Repression*

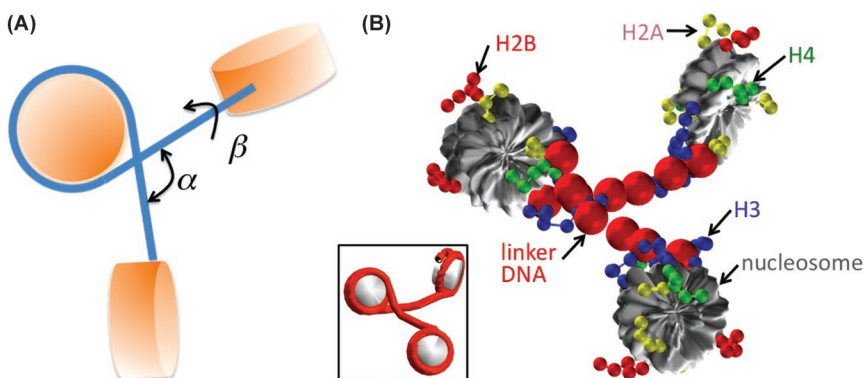
Gene activation or repression triggers noticeable changes in nucleosomal positioning. In general, nucleosomes are evicted from activated genes and nucleosomes are populated in repressed genes (78). However, instead of large scale shuffling of nucleosomes, changes in nucleosome positioning are restricted to few nucleosomes only. For example, in resting and activated genes of CD4+ T-cells (7), the expressed genes show typical nucleosome-positioning patterns near the TSS (see Figure 2) but unexpressed genes show only a strong positioning for the +1 nucleosome and no phasing. Upon activation, the H2A.Z variant comes in to replace canonical H2A at the -3, -2, +1, +2, +3 positions while the -1 H2A.Z is lost. Another change associated with activation of genes is the sliding of a nucleosome from its initial position at the TSS for about a few dozen of nucleotides downstream, both with and without histone acetylation of the +1 nucleosome (86, 87). Often many background changes in nucleosome positions occur even when the transcriptional state of a gene remains unchanged, possibly in preparation for future processes. In sum, there is a higher turnover of promoter region nucleosomes (termed "hot") over those within coding regions (termed "cold"). Remarkably, this high turnover promotes phased nucleosomes, suggesting that nucleosome dynamics at the promoters is governed by a precise and deliberate mechanism rather than a stochastic one that would promote disorder and fuzziness in nucleosome positions.

In living cells, transcribed chromatin is folded in a higher-order structure (88). It is not fully known how much of an inhibitory role the structure of chromatin plays in genetic processes like transcription. It is likely that the chromatin structure dictated directly by nucleosome positioning could regulate both the transcription rate and fidelity of the enzymes involved in gene functions. There are two main parameters related to nucleosome positioning through which chromatin structure is modulated to regulate gene function. One is the variability in linker lengths of few base pairs range that does not affect the average NRL but strongly affects relative orientation of nucleosomes (internucleosome rotation angle) within chromatin. The other parameter reflects large changes in the absolute length of linkers within the range of 20-90 bp that directly affects the NRL and the distance between neighboring nucleosomes.

*Effect of Local Rotational Variations*

**Partially Folded Chromatin Fibers:** Simplified models often depict chromatin as a regular beads-on-a-string array, though such a structure is observed only in very low ionic strength buffers. At physiological ionic conditions, this linear array of nucleosomes folds into a compact 3D chromatin fiber with a diameter of ~30 nm. In physiological solutions of monovalent cations, the 30 nm fiber is partially folded with mass per unit length of 5-7 nucleosomes/11 nm (89-92). Several experimental approaches including electron microscopy (EM) (91, 93, 94), atomic force microscopy (95), crosslinking studies (96), X-ray crystal structure analysis (97), and DNA cleavage by in-situ irradiation (98) have demonstrated that partially unfolded chromatin such as that observed during interphase has a zigzag conformation with relatively straight linker DNAs crossed in the middle and the nucleosome cores at the periphery of the fiber. Binding of linker histones stabilizes zigzag conformations in chromatin fibers (91, 99, 100). Similar structures have also been observed in computational models of nucleosome arrays (54, 101-103). Thus the zigzag model presents a well-grounded framework for translating nucleosome positioning information into 3D structure of interphase chromatin.

The zigzag model suggests that nucleosomal positions would be major determinants of chromatin structure. Indeed, if no constraints on nucleosome folding are imposed (as in an open zigzag), then the angle between the planes of adjacent nucleosomes ( $\beta$ ) becomes a periodic function of the linker DNA length (Figure 4). For nucleosome arrays with relatively straight linkers, fiber morphology and compactness is directly determined by the internucleosomal rotation angle  $\beta$  and the nucleosome entry/exit angle  $\alpha$ . Such a two-angle model with constant  $\beta$  however does not reproduce existing EM data on native chromatin fibers. A considerable progress in the understanding



**Figure 4:** (A) The two-angle model of chromatin used to model moderately folded chromatin. (B) A more detailed model of nucleosome arrays that treats DNA, nucleosome core, and histone tails at coarse-grained mesoscopic levels and accounts for the energetics, configurational entropy, and dynamics of loosely, moderately folded, and highly compact chromatin. The inset depicts the same configuration in a simpler representation.

of 3D organization of native chromatin fiber was made when the intrinsic irregularities of nucleosome arrays were taken into account. Indeed, linker lengths are not fixed but rather distributed with rough standard deviations of  $\pm 4.5$  bp (104) or  $\pm 2$  bp (105) around a well-defined average that peaks in 10 bp intervals. Even modest  $\pm 2$  bp variations in linker length in native chromatin (105) result in a large  $\pm 72^\circ$  variation in angle  $\beta$  (Figure 4a). Indeed, when variations of  $\pm 2$ -3 bp were introduced into the basic 3D zigzag model, the resulting structures closely resembled 30 nm fibers observed by transmission EM *in vitro* (94) and *in situ* (106), atomic force microscope (95), and cryo-EM (99).

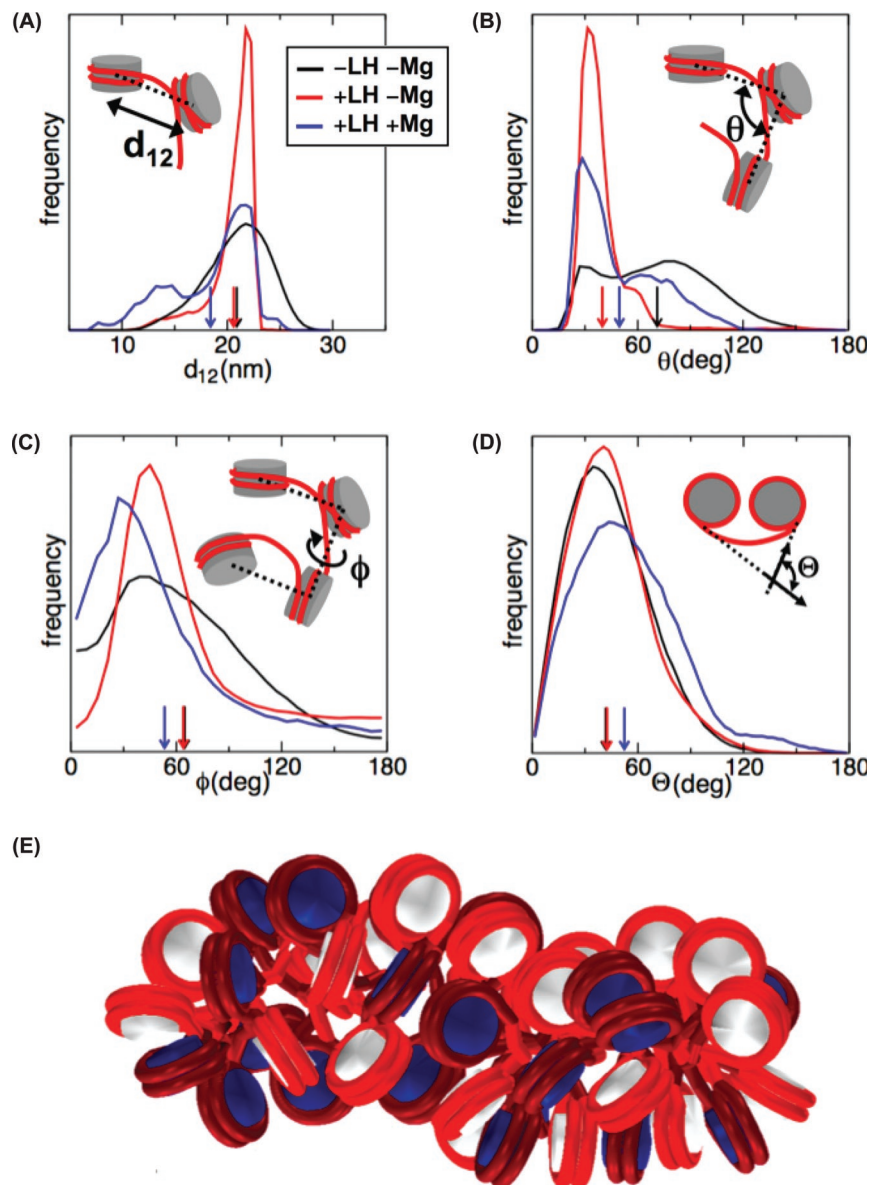
**Completely Folded Chromatin Fibers:** The most compact form of chromatin fibers (11-15 nucleosomes/11 nm) is typically only achieved in the presence of divalent cations and/or linker histones (9, 107). The exact configuration of DNA linkers in the compact fiber is still not well understood. Several models are based on straight linkers in the chromatin fibers (108, 109). Alternative models suggest that linkers are superhelically coiled between the nucleosomes leading to solenoid arrangements (91, 110). All these models invoke regular nucleosome spacing within the fiber, implying that the degree of rotational variations in nucleosome positions is critical for achieving most compact forms of chromatin. It could be argued that irregular spacing between nucleosomes would not allow chromatin to pack to its maximum possible density, leading to structures similar to the irregular zigzag model of chromatin (94, 95) that only lead to modest compaction. In contrast, certain uniform nucleosome orientations in a fiber would allow nucleosomes to stack much better against each other. Such stacking could facilitate strong internucleosomal interactions such as those between the H4 histone tail and acidic patch on adjacent nucleosome (2, 3, 96, 111, 112). In fact, the overall interactions between nucleosomes can be as large as 10-14 kT (where kT is thermal energy), as deduced from single-molecule pulling of chromatin (113).

The assumption that highly compact chromatin requires uniform linker lengths was tested recently using experiments with clone 601-based 12-nucleosome arrays in which linker DNA length was either uniform (207x12) or varied in a repeating pattern (205-207-209)x4 mimicking the natural variation in nucleosome repeat length ( $\pm 2$  bp) (93). Surprisingly, the uniform and variable nucleosome arrays exhibited similar compaction in unfolded, partially folded, and completely folded states. Furthermore, an EM study showed no difference in nucleosome-to-nucleosome contacts between the two fibers. These results suggest that linker DNA has enough flexibility to absorb small rotational variability between the neighboring nucleosomes.

How can a fully compact chromatin fiber accommodate rotational variability of nucleosomes? Recently, we have computationally studied 207 bp NRL nucleosome arrays using detailed mesoscopic modeling and simulations (54, 69, 93, 103). Our model of the nucleosome array integrates coarse-grained structural and energetic features of the nucleosome core, histone tails, linker histone, and linker DNA within a salt medium. An efficient Monte Carlo methodology is then employed to generate the ensemble of thermodynamically favorable configurations of the nucleosome array. Essentially, the model inputs coarse-grained representations and interactions of the various chromatin components and the simulations output the detailed configurations of the arrays, such as the internucleosomal distance  $d_{12}$ , the triplet angles  $\theta$ , and the dihedral angles  $\phi$ . Our simulations reveal a compact fiber with 30 nm outer diameter matching EM (93). Furthermore, computational analysis reveals a much greater conformational variability in linker angle leading to a surprisingly irregular internal linker conformations that combine features of zigzag and solenoid organization in a new heteromorphic chromatin fiber (Figure 5). It thus appears that the chromatin fiber structure has an intrinsic property to incorporate conformational variability in the DNA linkers without a notable effect on the fiber architecture.

NRL differences in tens of nucleotides strongly influence chromatin morphology and its compaction. It is now recognized that there is a visible increase in the chromatin diameter from ~33 nm for chromatin with linkers 30-60 bp long to ~42 nm for linker lengths of 70-90 bp. This sudden increase in the fiber diameter with linker length is also accompanied by a jump in nucleosome packing to 10–15 nucleosomes per 11nm of fiber (110). This crossover in chromatin packing and diameter was explained in terms of a transition from the helical ribbon configuration to a crossed zigzag configuration (114). The observed structural transitions are also consistent with a series of polymorphic chromatin models with straight linkers tangentially oriented in the fiber without crossing the fiber axis (115).

### A Structural Perspective on Nucleosome Positioning



**Figure 5:** Distributions in (A) internucleosomal distance, (B) triplet angle, (C) dihedral angle, and (D) linker bend angle observed from our simulations of 48-unit oligonucleosomes at 0.15 M monovalent salt without LH and  $Mg^{2+}$  (black lines), with LH, without  $Mg^{2+}$  (red lines), and with LH and  $Mg^{2+}$  (blue lines). The black, red, and blue arrows represent the averages. (E) Representative polymorphic structure of a 48-unit oligonucleosome obtained from our simulations. The odd and even-numbered nucleosomes are colored differently to emphasize interactions between every second and third nucleosomes. Adapted with permission from Reference (93).

Indeed, the internal structure of chromatin has been a subject of debate for several decades (116, 117). This dichotomy in chromatin diameters and packing densities has further fueled this debate. So far, the proposed models considered only the linker geometry and not the forces that mediate chromatin folding. Computer simulations of chromatin with various NRLs (118) that include internucleosome interactions (103) should further clarify the dependence of internal fiber organization upon linker length.

Yet another type of structure is observed for nucleosome arrays with short linker length typical of yeast and neuronal cells. Several previous studies have shown that such chromatin can be packed into partially folded 30 nm fibers without linker histone. When reconstituted arrays with 167 and 207 NRLs were compared under similar experimental conditions, it turned out that linker histone was indeed required to compact the long-linker arrays but not the short-linker ones (93, 116). We speculate that these differences in chromatin fiber folding arise due to stronger electrostatically screening of linker DNA by the core histone tails in short-linker arrays.

So what are the functional implications of linker length? Interestingly, yeast genome is very transcriptionally active and has short linker lengths consistent with two-start helical structure observed for similar-sized nucleosome arrays *in vitro* (96, 97). Wu *et al.*, suggest that such a structure would keep the linkers well exposed and accessible for protein binding compared to the crossed linker model that hides its linker inside the fiber (114). The length of the DNA linkers also dictates whether a linker histone is needed for maximal compaction or not. For example, the yeast genome contains a dramatically lower level of the linker histone than metazoan genomes (119, 120) as the short linkers are sufficient to allow close packing of nucleosomes. In contrast, chromatin fibers possessing longer linker DNA require linker histones for full compaction (93, 110, 116). The long-linker chromatin is generally less transcriptionally active than the more open short-linker chromatin probably due to the repressive function of linker histone that has been recently demonstrated *in vivo* (121). In addition to chromatin folding, linker histones can also facilitate chromatin fiber compaction through self-association (100), which has been suggested to present an additional level of chromatin higher-order compaction (122). For large mammalian genomes, where only a fraction of the total chromatin needs to be actively transcribed, the longer linkers may help to spatially segregate open euchromatin with lower levels of the linker histone from the repressed heterochromatin showing a higher level (123) and a longer retention time (124) of linker histone on the nucleosomes. Finally, the tight compaction of chromatin fibers with the longest NRLs found in echinoid spermatozoa reflects its folding at the higher ionic strength of the sea water (125) and may not be suitable for physiological conditions in mammalian cells. Since these larger variations in the linker may lead to significant changes in higher order folding, cells sustain roughly uniform NRLs.

Cumulatively, the existing data show that the chromatin conformation is not sensitive to rotational changes induced by a distribution of nucleosome positions about the mean NRL ( $\pm 2$  bp) but can be altered by larger linker length variations in two steps: one above 30 bp and the other above 60 bp. Perhaps, the need for certain NRLs to achieve high packing densities can explain why the ISWI family of chromatin remodelers (such as ACF (126)), which promote uniform NRLs, are necessary for the formation of highly compact heterochromatin (127). Hence, genes that are no longer required by the cell could be globally modified by organizing its nucleosomes into uniform repeats through the action of these enzymes without much dependence on the underlying DNA sequence. Alternatively, activation of genes could be aided through local nucleosome disruption, as achieved by another class of remodeling enzymes, SWI/SNF, whose function are to disorganize chromatin structure in a gene-specific manner (29, 128, 129).

In this review, we have discussed recent progress in studies of nucleosome positioning from a structural perspective. The technological breakthroughs that uncovered vast information previously hidden in the complexity of large genomes are truly fascinating. Despite all these advancements, many important questions remain open.

What are the functions of H2A.Z variant containing nucleosomes that typically assemble near the TSS sites? On the one hand, H2A.Z promotes chromatin intra-fiber compaction, and on the other hand it unfolds the nucleosome structure. Does the promoter have a unique architecture specified by these variant nucleosomes that can be transiently unfolded by the transcriptional machinery and at the same time maintain potential for compact higher-order folding?

More systematic studies are needed to quantify the contribution of the different factors dictating nucleosome positioning. It is clear that the affinity of the underlying DNA sequence for histone octamer (so-called nucleosome positioning code) cannot explain the entire complexity of nucleosome positioning (67). Will it be possible to predict nucleosome positions more precisely if sequence-specific DNA binding factors are taken in consideration? To this end, studying more complex reconstituted systems that include additional protein factors *in silico* and *in vitro* would also be highly useful.

Can nucleosome positioning prediction tools be useful in narrowing down the search range for predicting sites of TF occupancy, with the idea that the nucleosome bound sites are not available for TF binding? We are aware of one such tool that has been developed for obtaining a nucleosome-guided map of transcription factor binding sites in yeast (130).

Does nucleosome positioning reflect the higher order structure and/or distinguish heterochromatin and euchromatin domains? While some studies of NRL have suggested that this could be true, genome-wide nucleosome-mapping studies have not yet uncovered distinct nucleosome positioning signatures over condensed heterochromatic regions.

In the near future, we expect further development of cost-efficient high throughput sequencing technologies and sequence analysis software that, together with molecular genetic techniques for altering nucleosome-positioning sequences *in vivo*, will elucidate the molecular mechanisms governing nucleosome positions. Will the new data lead to a unified picture or at least a finite number of rules allowing one to predict nucleosome positions from the DNA sequence? This is not certain. In view of the complexity of epigenetic mechanisms, multitude of histone modifications, and hundreds of millions of nucleosomes that make up the human genome, we still may have to deal with “One million Hows, two million Wheres, and seven million Whys”. This research was reported by the authors in part at Albany 2009: The 16th Conversation (131).

### **Acknowledgments**

This work was supported by Hellman Fellowship to G. Arya and NSF grant MCB-0615536 to S. Grigoryev.

### **References and Footnotes**

1. G. Felsenfeld and M. Groudine. *Nature* 421, 448-453 (2003).
2. C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, and T. J. Richmond. *J Mol Biol* 319, 1097-1113 (2002).
3. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond. *Nature* 389, 251-260 (1997).
4. E. Segal and J. Widom. *Trends Genet* 25, 335-343 (2009).

5. T. N. Mavrich, I. P. Ioshikhes, B. J. Venters, C. Jiang, L. P. Tomsho, J. Qi, S. C. Schuster, I. Albert, and B. F. Pugh. *Genome Res* 18, 1073-1083 (2008).
6. T. N. Mavrich, C. Z. Jiang, I. P. Ioshikhes, X. Y. Li, B. J. Venters, S. J. Zanton, L. P. Tomsho, J. Qi, R. L. Glaser, S. C. Schuster, D. S. Gilmour, I. Albert, and B. F. Pugh. *Nature* 453, 358-362 (2008).
7. D. E. Schones, K. R. Cui, S. Cuddapah, T. Y. Roh, A. Barski, Z. B. Wang, G. Wei, and K. J. Zhao. *Cell* 132, 887-898 (2008).
8. H. Weintraub. *Nucleic Acids Res* 5, 1179-1188 (1978).
9. B. D. Athey, M. F. Smith, D. A. Rankert, S. P. Williams, and J. P. Langmore. *J Cell Biol* 111, 795-806 (1990).
10. T. A. Blank and P. B. Becker. *J Mol Biol* 252, 305-313 (1995).
11. E. Segal, Y. Fondufe-Mittendorf, L. Y. Chen, A. Thastrom, Y. Field, I. K. Moore, J. P. Z. Wang, and J. Widom. *Nature* 442, 772-778 (2006).
12. S. Mattei, B. Sampaolese, P. De Santis, and M. Savino. *Biophys Chem* 97, 173-187 (2002).
13. W. Lee, D. Tillo, N. Bray, R. H. Morse, R. W. Davis, T. R. Hughes, and C. Nislow. *Nat Genet* 39, 1235-1244 (2007).
14. R. Belotserkovskaya, S. Oh, V. A. Bondarenko, G. Orphanides, V. M. Studitsky, and D. Reinberg. *Science* 301, 1090-1093 (2003).
15. D. S. Johnson, A. Mortazavi, R. M. Myers, and B. Wold. *Science* 316, 1497-1502 (2007).
16. G. Robertson, M. Hirst, M. Bainbridge, M. Bilenky, Y. J. Zhao, T. Zeng, G. Euskirchen, B. Bernier, R. Varhol, A. Delaney, N. Thiessen, O. L. Griffith, A. He, M. Marra, M. Snyder, and S. Jones. *Nat Methods* 4, 651-657 (2007).
17. J. D. Lieb, X. Liu, D. Botstein, and P. O. Brown. *Nat Genet* 29, 327-333 (2001).
18. B. Ren, F. Robert, J. J. Wyrick, O. Aparicio, E. G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T. L. Volkert, C. J. Wilson, S. P. Bell, and R. A. Young. *Science* 290, 2306-2309 (2000).
19. B. E. Bernstein, C. L. Liu, E. L. Humphrey, E. O. Perlestein, and S. L. Schreiber. *Genome Biol* 5, R62 (2004).
20. G. C. Yuan, Y. J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, and O. J. Rando. *Science* 309, 626-630 (2005).
21. N. D. Heintzman, G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, Z. Ye, L. K. Lee, R. K. Stuart, C. W. Ching, K. A. Ching, J. E. Antosiewicz-Bourget, H. Liu, X. M. Zhang, R. D. Green, V. V. Lobanenkov, R. Stewart, J. A. Thomson, G. E. Crawford, M. Kellis, and B. Ren. *Nature* 459, 108-112 (2009).
22. B. R. Cairns. *Nature* 461, 193-198 (2009).
23. S. Cuddapah, R. Jothi, D. E. Schones, T. Y. Roh, K. R. Cui, and K. J. Zhao. *Genome Res* 19, 24-32 (2009).
24. N. Spies, C. B. Nielsen, R. A. Padgett, and C. B. Burge. *Mol Cell* 36, 245-254 (2009).
25. S. Schwartz, E. Meshorer, and G. Ast. *Nat Struc Mol Biol* 16, 990-U117 (2009).
26. H. Tilgner, C. Nikolaou, S. Althammer, M. Sammeth, M. Beato, J. Valcarcel, and R. Guigo. *Nat Struc Mol Biol* 16, 996-1001 (2009).
27. I. P. Ioshikhes, I. Albert, S. J. Zanton, and B. F. Pugh. *Nat Genet* 38, 1210-1215 (2006).
28. I. Albert, T. N. Mavrich, L. P. Tomsho, J. Qi, S. J. Zanton, S. C. Schuster, and B. F. Pugh. *Nature* 446, 572-576 (2007).
29. B. R. Cairns. *Nat Struc Mol Biol* 14, 989-996 (2007).
30. K. Struhl. *Proc Natl Acad Sci USA* 82, 8419-8423 (1985).
31. A. Almer, H. Rudolph, A. Hinnen, and W. Horz. *EMBO J* 5, 2689-2696 (1986).
32. C. Z. Jiang and B. F. Pugh. *Nature Reviews Genetics* 10, 161-172 (2009).
33. J. E. Phillips and V. G. Corces. *Cell* 137, 1194-1211 (2009).
34. Y. T. Fu, M. Sinha, C. L. Peterson, and Z. P. Weng. *PLoS Genet* 4, e1000138 (2008).
35. P. T. Lowary and J. Widom. *J Mol Biol* 276, 19-42 (1998).
36. Y. Zhang, Z. Moqtaderi, B. P. Rattner, G. Euskirchen, M. Snyder, J. T. Kadonaga, X. S. Liu, and K. Struhl. *Nat Struc Mol Biol* 16, 847-853 (2009).
37. F. Salih, B. Salih, and E. N. Trifonov. *J Biomol Struct Dyn* 26, 273-281 (2008).
38. F. Salih, B. Salih, S. Kogan, and E. N. Trifonov. *J Biomol Struct Dyn* 26, 9-15 (2008).
39. I. Ioshikhes, A. Bolshoy, K. Derenshteyn, M. Borodovsky, and E. N. Trifonov. *J Mol Biol* 262, 129-139 (1996).
40. I. Gabdank, D. Barash, and E. N. Trifonov. *J Biomol Struct Dyn* 26, 403-412 (2009).
41. V. A. Bondarenko, L. M. Steele, A. Ujvari, D. A. Gaykalova, O. I. Kulaeva, Y. S. Polikanov, D. S. Luse, and V. M. Studitsky. *Mol Cell* 24, 469-479 (2006).
42. Abstracts of Albany 2009: 16<sup>th</sup> Conversation. June 16-20, Albany, New York, USA; P. De Santis, A. Scipioni. A Statistical Thermodynamic Approach for Predicting The Sequence-Dependent Nucleosome Positioning along Genomes, Abstract #195. *J Biomol Struct Dyn* 26, 787-927 (2009).
43. G. C. Yuan and J. S. Liu. *PLoS Comput Biol* 4, e13 (2008).
44. H. E. Peckham, R. E. Thurman, Y. T. Fu, J. A. Stamatoyannopoulos, W. S. Noble, K. Struhl, and Z. P. Weng. *Genome Res* 17, 1170-1177 (2007).
45. N. Kaplan, I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, D. Tillo, Y. Field, E. M. LeProust, T. R. Hughes, J. D. Lieb, J. Widom, and E. Segal. *Nature* 458, 362-366 (2009).



46. R. Kiyama and E. N. Trifonov. *Febs Let* 523, 7-11 (2002).
47. V. B. Teif and K. Rippe. *Nucleic Acids Res* 17, 5651-5655 (2009).
48. Abstracts of Albany 2009: 16<sup>th</sup> Conversation. June 16-20, Albany, New York, USA; F. Cui, D. F. Wang and V. B. Zhurkin. AT-rich Fragments at the Nucleosome Ends May be Related to Linker Histone Binding: Implications for Nucleosome Positioning, Abstract #197. *J Biomol Struct Dyn* 26, 787-927 (2009).
49. F. Cui and V. B. Zhurkin. *Nucleic Acids Res* 37, 2818-2829 (2009).
50. M. J. Fedor, N. F. Lue, and R. D. Kornberg. *J Mol Biol* 204, 109-127 (1988).
51. M. J. Pazin, P. Bhargava, E. P. Geiduschek, and J. T. Kadonaga. *Science* 276, 809-812 (1997).
52. G. Chevereau, L. Palmeira, C. Thermes, A. Arneodo, and C. Vaillant. *Phys Rev Lett* 103, 188103 (2009).
53. R. T. Koerber, H. S. Rhee, C. Jiang, and B. F. Pugh. *Mol Cell* 35, 889-902 (2009).
54. G. Arya and T. Schlick. *Proc Natl Acad Sci USA* 103, 16236-16241 (2006).
55. J. C. Hansen, C. Tse, and A. P. Wolffe. *Biochemistry* 37, 17637-17641 (1998).
56. J. Yao, P. T. Lowary, and J. Widom. *Proc Natl Acad Sci USA* 90, 9364-9368 (1993).
57. J. Svaren, E. Klebanow, L. Sealy, and R. Chalkley. *J Biol Chem* 269, 9335-9344 (1994).
58. G. Almouzni, M. Mechali, and A. P. Wolffe. *EMBO J* 9, 573-582 (1990).
59. L. Mirny. *arXiv.org*, arXiv:0901.2905v1 (2009).
60. M. S. Kobor, S. Venkatasubrahmanyam, M. D. Meneghini, J. W. Gin, J. L. Jennings, A. J. Link, H. D. Madhani, and J. Rine. *PLoS Biol* 2, 587-599 (2004).
61. A. Lusser and J. T. Kadonaga. *Bioessays* 25, 1192-1200 (2003).
62. V. K. Gangaraju and B. Bartholomew. *Mut Res-Fund Mol M* 618, 3-17 (2007).
63. G. J. Narlikar, H. Y. Fan, and R. E. Kingston. *Cell* 108, 475-487 (2002).
64. M. N. Kagalwala, B. J. Glaus, W. W. Dang, M. Zofall, and B. Bartholomew. *EMBO J* 23, 2092-2104 (2004).
65. D. F. V. Corona, C. R. Clapier, P. B. Becker, and J. W. Tamkun. *EMBO Rep* 3, 242-247 (2002).
66. I. Whitehouse and T. Tsukiyama. *Nat Struc Mol Biol* 13, 633-640 (2006).
67. P. D. Partensky and G. J. Narlikar. *J Mol Biol* 391, 12-25 (2009).
68. M. Manohar, A. M. Mooney, J. A. North, R. J. Nakkula, J. W. Picking, A. Edon, R. Fishel, M. G. Poirier, and J. J. Ottesen. *J Biol Chem* 284, 23312-23321 (2009).
69. G. Arya and T. Schlick. *J Phys Chem A* 113, 4045-4059 (2009).
70. S. A. Jacobs and S. Khorasanizadeh. *Science* 295, 2080-2083 (2002).
71. B. E. Black, D. R. Foltz, S. Chakravarthy, K. Luger, V. L. Woods, and D. W. Cleveland. *Nature* 430, 578-582 (2004).
72. D. W. Abbott, V. S. Ivanova, X. Y. Wang, W. M. Bonner, and J. Ausio. *J Biol Chem* 276, 41945-41949 (2001).
73. J. Y. Fan, F. Gordon, K. Luger, J. C. Hansen, and D. J. Tremethick. *Nat Struc Mol Biol* 9, 172-176 (2002).
74. J. Y. Fan, D. Rangasamy, K. Luger, and D. J. Tremethick. *Mol Cell* 16, 655-661 (2004).
75. D. Nathan and D. M. Crothers. *J Mol Biol* 316, 7-17 (2002).
76. S. Buratowski. *Science* 322, 1804-1805 (2008).
77. S. Shivaswamy and V. R. Iyer. *Mol Cell Biol* 28, 2221-2234 (2008).
78. S. Shivaswamy, A. Bhinge, Y. J. Zhao, S. Jones, M. Hirst, and V. R. Iyer. *PLoS Biol* 6, 618-630 (2008).
79. A. D. Basehoar, S. J. Zanton, and B. F. Pugh. *Cell* 117, 847-847 (2004).
80. K. Struhl. *Ann Rev Biochem* 58, 1051-1077 (1989).
81. C. Martinez-Campa, P. Politis, J. L. Moreau, N. Kent, J. Goodall, J. Mellor, and C. R. Goding. *Mol Cell* 15, 69-81 (2004).
82. M. G. Poirier, M. Bussiek, J. Langowski, and J. Widom. *J Mol Biol* 379, 772-786 (2008).
83. M. S. Santisteban, T. Kalashnikova, and M. M. Smith. *Cell* 103, 411-422 (2000).
84. C. Y. Jin and G. Felsenfeld. *Genes Dev* 21, 1519-1529 (2007).
85. S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, and D. J. Patel. *Nat Struc Mol Biol* 14, 1025-1040 (2007).
86. S. Lomvardas and D. Thanos. *Cell* 106, 685-696 (2001).
87. G. Koutroubas, M. Merika, and D. Thanos. *Mol Cell Biol* 28, 926-938 (2008).
88. Y. Hu, I. Kireev, M. Plutz, N. Ashourian, and A. S. Belmont. *J Cell Biol* 185, 87-100 (2009).
89. A. Bassett, S. Cooper, C. Y. Wu, and A. Travers. *Curr Opin Genet Dev* 19, 159-165 (2009).
90. R. Ghirlando and G. Felsenfeld. *J Mol Biol* 376, 1417-1425 (2008).
91. F. Thoma, T. Koller, and A. Klug. *J Cell Biol* 83, 403-427 (1979).
92. S. E. Gerchman and V. Ramakrishnan. *Proc Natl Acad Sci USA* 84, 7802-7806 (1987).
93. S. A. Grigoryev, G. Arya, S. Correll, C. L. Woodcock, and T. Schlick. *Proc Natl Acad Sci USA* 106, 13317-13322 (2009).
94. C. L. Woodcock, S. A. Grigoryev, R. A. Horowitz, and N. Whitaker. *Proc Natl Acad Sci USA* 90, 9021-9025 (1993).
95. S. H. Leuba, G. L. Yang, C. Robert, B. Samori, K. Vanholde, J. Zlatanova, and C. Bustamante. *Proc Natl Acad Sci USA* 91, 11621-11625 (1994).

96. B. Dorigo, T. Schalch, A. Kulangara, S. Duda, R. R. Schroeder, and T. J. Richmond. *Science* 306, 1571-1573 (2004).
97. T. Schalch, S. Duda, D. F. Sargent, and T. J. Richmond. *Nature* 436, 138-141 (2005).
98. B. Rydberg, W. R. Holley, I. S. Mian, and A. Chatterjee. *J Mol Biol* 284, 71-84 (1998).
99. J. Bednar, R. A. Horowitz, S. A. Grigoryev, L. M. Carruthers, J. C. Hansen, A. J. Koster, and C. L. Woodcock. *Proc Natl Acad Sci USA* 95, 14173-14178 (1998).
100. L. M. Carruthers, J. Bednar, C. L. Woodcock, and J. C. Hansen. *Biochemistry* 37, 14776-14787 (1998).
101. G. Wedemann and J. Langowski. *Biophys J* 82, 2847-2859 (2002).
102. R. Stehr, N. Kepper, K. Rippe, and G. Wedemann. *Biophys J* 95, 3677-3691 (2008).
103. G. Arya, Q. Zhang, and T. Schlick. *Biophys J* 91, 133-150 (2006).
104. F. Strauss and A. Prunell. *EMBO J* 2, 51-56 (1983).
105. J. Widom. *Proc Natl Acad Sci USA* 89, 1095-1099 (1992).
106. R. A. Horowitz, D. A. Agard, J. W. Sedat, and C. L. Woodcock. *J Cell Biol* 125, 1-10 (1994).
107. P. J. J. Robinson, W. An, A. Routh, F. Martino, L. Chapman, R. G. Roeder, and D. Rhodes. *J Mol Biol* 381, 816-825 (2008).
108. A. Worcel, S. Strogatz, and D. Riley. *Proc Natl Acad Sci USA* 78, 1461-1465 (1981).
109. S. P. Williams, B. D. Athey, L. J. Muglia, R. S. Schappe, A. H. Gough, and J. P. Langmore. *Biophys J* 49, 233-248 (1986).
110. P. J. J. Robinson and D. Rhodes. *Curr Opin Struc Biol* 16, 336-343 (2006).
111. B. Dorigo, T. Schalch, K. Bystricky, and T. J. Richmond. *J Mol Biol* 327, 85-96 (2003).
112. M. Shogren-Knaak, H. Ishii, J. M. Sun, M. J. Pazin, J. R. Davie, and C. L. Peterson. *Science* 311, 844-847 (2006).
113. M. Kruithof, F. T. Chien, A. Routh, C. Logie, D. Rhodes, and J. van Noort. *Nat Struc Mol Biol* 16, 534-540 (2009).
114. C. Y. Wu, A. Bassett, and A. Travers. *EMBO Rep* 8, 1129-1134 (2007).
115. H. Wong, J. M. Victor, J. Mozziconacci. *PLoS One* 2, (2007).
116. A. Routh, S. Sandin, and D. Rhodes. *Proc Natl Acad Sci USA* 105, 8872-8877 (2008).
117. D. J. Tremethick. *Cell* 128, 651-654 (2007).
118. N. Kepper, D. Foethke, R. Stehr, G. Wedemann, and K. Rippe. *Biophys J* 95, 3692-3705 (2008).
119. I. Freidkin and D. J. Katcoff. *Nucleic Acids Res* 29, 4043-4051 (2001).
120. J. A. Downs, E. Kosmidou, A. Morgan, and S. P. Jackson. *Mol Cell* 11, 1685-1692 (2003).
121. X. W. Lu, S. N. Wontakal, A. V. Emelyanov, P. Morcillo, A. Y. Konev, D. V. Fyodorov, and A. I. Skoultchi. *Genes Dev* 23, 452-465 (2009).
122. S. A. Grigoryev. *Febs Let* 564, 4-8 (2004).
123. R. T. Kamakaka and J. O. Thomas. *Embo J* 9, 3997-4006 (1990).
124. T. Misteli, A. Gunjan, R. Hock, M. Bustin, and D. T. Brown. *Nature* 408, 877-881 (2000).
125. P. J. Giannasca, R. A. Horowitz, and C. L. Woodcock. *J Cell Sci* 105, 551-561 (1993).
126. T. Ito, M. Bulger, M. J. Pazin, R. Kobayashi, and J. T. Kadonaga. *Cell* 90, 145-155 (1997).
127. D. V. Fyodorov, M. D. Blower, G. H. Karpen, and J. T. Kadonaga. *Genes Dev* 18, 170-183 (2004).
128. C. L. Peterson, J. Cote, J. Quinn, J. Workman, E. Richmond, M. Wechsler, L. Burns, L. Boyer, and K. Pollard. *J Cell Biochem*, 154-154 (1995).
129. H. Kwon, A. N. Imbalzano, P. A. Khavari, R. E. Kingston, and M. R. Green. *Nature* 370, 477-481 (1994).
130. L. Narlikar, R. Gordan, U. Ohler, and A. J. Hartemink. *Bioinformatics* 22, E384-E392 (2006).
131. Abstracts of Albany 2009: 16th Conversation. June 16-20, Albany, New York, USA; G. Arya, S. Grigoryev and T. Schlick, Mesoscale Modeling Predicts New Polymorphic Structure of Chromatin, Abstract #187. *J Biomol Struct Dyn*, 26, 908-909 (2009).

Date Received: November 14, 2009

Communicated by the Editor Ramaswamy H. Sarma