nature structural & molecular biology

Determinants of nucleosome positioning

Kevin Struhl[1] & Eran Segal[2,3]

Nucleosome positioning is critical for gene expression and most DNA-related processes. Here we review the dominant patterns of nucleosome positioning that have been observed and summarize the current understanding of their underlying determinants. The genome-wide pattern of nucleosome positioning is determined by the combination of DNA sequence, ATP-dependent nucleosome remodeling enzymes and transcription factors that include activators, components of the preinitiation complex and elongating RNA polymerase II. These determinants influence each other such that the resulting nucleosome positioning patterns are likely to differ among genes and among cells in a population, with consequent effects on gene expression.

Eukaryotic genomes are packaged into chromatin, whose basic repeating unit is a nucleosome that consists of a histone octamer wrapped around 147 base pairs (bp) of DNA[1]. Nucleosomes are arranged into regularly spaced arrays, with the length of the linker region between nucleosomes varying among species and cell types. Although initially nucleosomes were believed to provide a universal, nonspecific coating of genomic DNA, it has long been known that nucleosomes occupy favored positions throughout the genome. High-resolution, genome-wide analyses have revealed a common pattern: nucleosomes are depleted at many (but not all) enhancer, promoter and terminator regions, and they typically occupy preferred positions in genes and non-gene regions[2–9]. In yeast, the −1 and +1 nucleosomes flanking the promoter are located at highly preferred positions, and the extent of preferred nucleosome positioning gradually decreases from the 5′ to 3′ end of the coding region[4,7].

In this Review we will consider the mechanisms by which the genomic pattern of nucleosome positioning is achieved. Intense research efforts collectively have revealed that nucleosome positioning is not determined by any single factor but rather by the combined effects of several factors including DNA sequence, DNA-binding proteins, nucleosome remodelers and the RNA polymerase II (Pol II) transcription machinery. One aim of this Review is to resolve the apparent controversy (in large part generated by us and the late Jonathan Widom, to whom we dedicate this Review) about the role of DNA sequence in establishing the genomic pattern of nucleosome positioning. This controversy stems from differences in interpretation and imprecision of terms, as the experimental observations have been remarkably consistent and are uncontested. In this Review we will focus on nucleosome positioning in yeast, particularly the biochemical and genetic analyses that have been rarely performed in multicellular organisms. We suspect that many aspects of nucleosome positioning are conserved among eukaryotes, but one important difference is that the linker histone H1 in yeast is structurally atypical, and it is present at very low concentrations in comparison to core histones[10].

Definition of nucleosome positioning and occupancy

We define the term 'nucleosome positioning' broadly to indicate where nucleosomes are located with respect to the genomic DNA sequence. Nucleosome positioning is a dynamic process, but sequencing-based mapping approaches identify the positions of individual nucleosomes in a single cell at a specific time. Nevertheless, nucleosome positions are typically discussed on a cell- and time-averaged basis. Nucleosome positioning can vary from perfect positioning, in which a nucleosome is located at a given 147-bp stretch in all DNA molecules in a cell population, to no positioning, in which nucleosomes are located at all possible genomic positions with equal frequency across a cell population (**Fig. 1a**).

Nucleosome positioning is related to, but is distinct from, 'nucleosome occupancy', which reflects the fraction of cells from the population in which a given region of DNA is occupied by a histone octamer (**Fig. 1b**). Although most genomic DNA is occupied by nucleosomes, many functional regions (promoters, enhancers and terminators) are depleted of nucleosomes (have low occupancy) and some regions are largely nucleosome-free.

Nucleosome occupancy and positioning are critical to biological outcomes[4,7,11,12], primarily because nucleosomes inhibit the access of other DNA-binding proteins to DNA. Highly accessible regions in genomes are identified by preferential restriction endonuclease cleavage[13] and DNase I hypersensitivity[14] analyses. Quantitative analysis of HinfI cleavage[15] or Leu3 binding[16] in yeast cells indicates that access of factors to target sites in nucleosome-depleted promoter regions is ~10–20-fold higher than to identical sequences that are associated with nucleosomes. More generally, as a consequence of their modes of binding, transcription factors differ a great deal with respect to how much their binding is inhibited by nucleosomes. A special class of 'pioneer' transcription factors (for example, FoxA and GATA) can bind their target sites in the context of nucleosomal DNA[17]. Such pioneer factors, via recruitment of nucleosome remodelers, can open up the local chromatin, thereby facilitating the binding of other transcription factors that otherwise would be blocked by nucleosomes. The TATA-binding protein, and hence the entire basic Pol II transcription machinery, virtually cannot bind

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA. [2]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. [3]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel. Correspondence should be addressed to E.S. (eran.segal@weizmann.ac.il) or K.S. (kevin@hms.harvard.edu).

Received 3 December 2012; accepted 3 January 2013; published online 5 March 2013; doi:10.1038/nsmb.2506

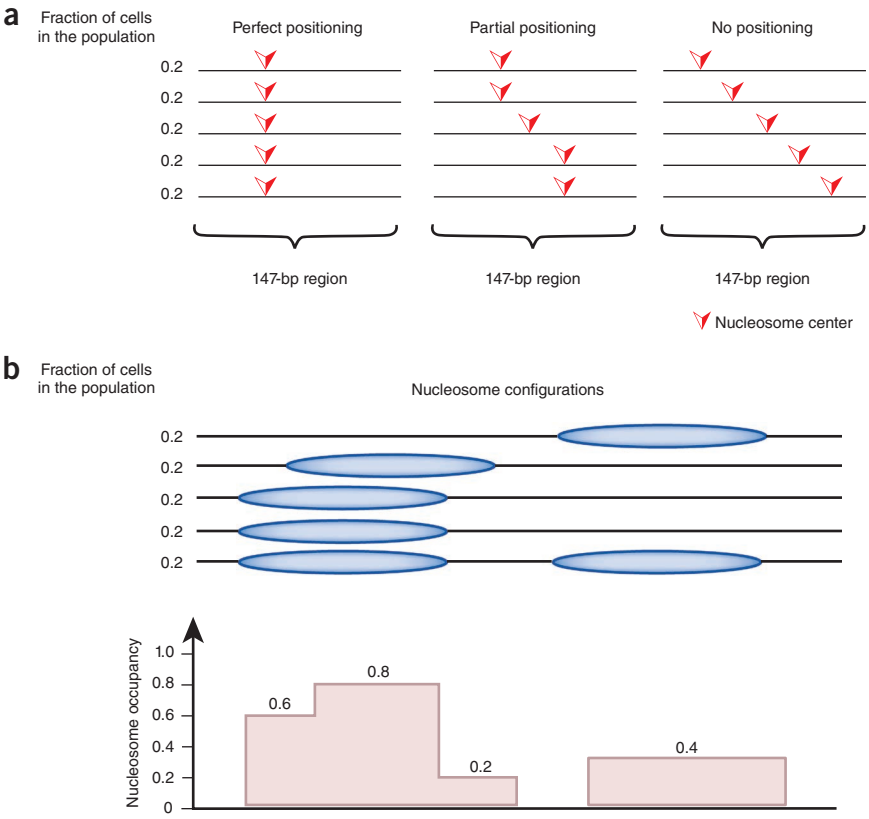

Figure 1 Illustration of the concepts of nucleosome positioning and nucleosome occupancy. (**a**) Nucleosome positioning along every base pair in the genome is defined as the fraction of cells from the population in which that base pair is at the center of a 147-bp nucleosome. For a given 147-bp region in a cell population, illustrated are perfect positioning, in which the center of the nucleosome is at the same base pair across all cells; partial positioning, in which there is a preference for some locations; and no positioning, in which all locations have equal probability. (**b**) Nucleosome occupancy along every base pair in the genome is defined as the fraction of cells from the population in which the base pair is occupied by any histone octamer. Illustrated are nucleosome locations across a cell population (top) and the resulting nucleosome occupancy per base pair (bottom).

nucleosomal DNA and hence requires a nucleosome-free region to bind core promoters and initiate transcription[18].

Nucleosome positioning is strongly affected by DNA sequence

The debate about the role of DNA sequence in nucleosome positioning *in vivo* revolves around the intrinsic sequence preferences of the histone octamer in the absence of any other component. The affinity of histone octamers for a given 147-bp sequence varies over more than three orders of magnitude[19]. As such, histone octamers exhibit considerable DNA sequence specificity, albeit lower than that of a classical specific DNA-binding protein. However, unlike DNA-binding proteins that achieve specificity by virtue of direct and strong interactions between a few base pairs and amino acids, the specificity of nucleosome formation largely reflects the overall ability of a given 147-bp sequence to bend around the histone octamer[20]. For optimal nucleosome formation, more bendable sequences are in contact with the histones, and less bendable sequences are solvent-exposed.

Two major sequence determinants affect bending and hence nucleosome formation. First, dinucleotides vary considerably with respect to their bending properties. Optimal nucleosome formation occurs when bendable dinucleotides (AT and TA) occur on the face of the helical repeat (every 10 bp) that directly interacts with histones[21,22]. Recent mapping of nucleosomes in the yeast *Saccharomyces cerevisiae*—using a H4-S47C–mediated cleavage technique that allows precise mapping of nucleosomes with respect to the DNA sequence—revealed 10-bp periodicity of bendable dinucleotides throughout nearly the entire 147-bp region[23]. The exact position of the histone octamer with respect to the ~10-bp helical repeat is termed 'rotational positioning', and thus, DNA sequence is a critical determinant of how nucleosomes are rotationally positioned (**Fig. 2**).

Second, the homopolymeric sequences poly(dA:dT) and poly (dG:dC) are intrinsically stiff (for different structural reasons) and are strongly inhibitory to nucleosome formation[24–27]. Notably, poly (dA:dT) tracts are abundant in eukaryotic genomes[28] and are particularly prevalent in promoters of certain organisms such as *S. cerevisiae*[29,30]. As we will discuss below, intrinsic properties of poly(dA:dT) are important for nucleosome depletion, promoter accessibility and transcriptional activity[12,15,26].

By analogy to DNA-binding proteins, the above DNA sequence preferences for nucleosomes can be expressed as a position-weight matrix that can generate a nucleosome score for any 147-bp region of DNA[22]. These nucleosome scores correspond to relative affinities of 147-bp sequences for histone octamers; the differences in affinities must affect nucleosome positioning *in vivo*. In particular, as dinucleotide preferences vary along the 147-bp region of the nucleosome, most DNA sequences will have considerably different nucleosome

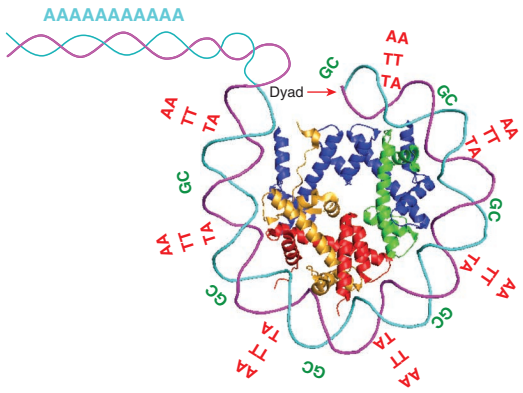

Figure 2 Nucleosome sequence preferences. Within the 147 bp that are wrapped around the histone octamer, there is a preference for distinctive dinucleotides that recur periodically at the DNA helical repeat and are known to facilitate the sharp bending of DNA around the nucleosome. These include ~10-bp periodic AA, TT or TA dinucleotides that oscillate in phase with each other and out of phase with ~10-bp periodic GC dinucleotides. The linker regions exhibit a strong preference for sequences that resist DNA bending and thus disfavor nucleosome formation. Among these, poly(dA:dT) tracts and their variants are most dominant and highly enriched in eukaryotic promoters.

scores at neighboring positions in an individual helical repeat. As a consequence, there will usually be a preferred position(s) in a helical repeat, and such rotational positioning is clearly observed in eukaryotic genomes. However, if one considers a favored rotational position, genomic locations that are 10 bp apart (1 helical turn) will have similar (although not identical) nucleosome formation scores, unless the nearby locations contain strongly inhibitory sequences[31]. Thus, in many genomic regions, DNA sequence alone is unlikely to strongly favor one nucleosome position over nearby positions that are rotationally equivalent.

Poly(dA:dT) tracts are important for nucleosome depletion

The role of DNA sequence in establishing nucleosome positions *in vivo* has been addressed in experiments in which nucleosomes are assembled *in vitro* by salt dialysis using purified histones and genomic DNA[32,33]. Under these conditions, many yeast promoter and terminator regions are depleted of nucleosomes, indicating that the DNA sequences intrinsically disfavor formation of nucleosomes. Consistent with this view, depletion of nucleosomes at promoters is observed under a wide variety of conditions *in vivo* and is unaffected by transcriptional activity[32,34]. In addition, when artificial chromosomes containing large genomic regions from heterologous yeast species are analyzed in *S. cerevisiae*, depletion of nucleosomes at promoters is maintained in a manner that depends on the number, length and fully homopolymeric nature of poly(dA:dT) sequences[35]. Thus, depletion of nucleosomes at most promoter sequences is strongly influenced by intrinsic properties of poly(dA:dT) sequences.

A variety of detailed functional analyses indicate that the intrinsic properties of poly(dA:dT) are important for nucleosome depletion, promoter accessibility and transcriptional activity[12,15,26,29,34–37]. In principle, poly(dA:dT)-mediated depletion could be due to a poly(dA:dT)-binding activator protein, but the one known factor with such DNA-binding specificity (Datin) actually inhibits transcription[15]. Consistent with an earlier study[15], expression measurements of a large-scale promoter library designed with systematic manipulations to the properties and spatial arrangement of poly (dA:dT) tracts showed that longer and more perfect tracts induce transcription regardless of the identity of the binding transcription factor[12]. This effect of the tracts on transcription is mediated by reduced nucleosome occupancy and thus increased accessibility that these tracts confer on nearby promoter elements, such as transcription factor–binding sites.

Alteration of poly(dA:dT) tracts offers a general evolutionary mechanism, applicable to promoters regulated by different transcription factors, for tuning expression in a predictable manner, with resolution that can be even finer than that attained by altering transcription factor–binding sites[12]. Indeed, genomes have likely used these tracts to regulate expression, for example, to partly compensate for differences in gene copy number that exist among ribosomal protein genes in *S. cerevisiae*[38] and to alter expression of cellular respiration genes across yeast species[39]. Moreover, partitioning yeast genes into two categories on the basis of the extent to which their promoters favor nucleosome formation in a manner largely dependent on the occurrences of poly(dA:dT) tracts results in functionally distinct gene classes. Specifically, the class of genes whose promoters have a higher intrinsic preference for nucleosomes is enriched for stress-response genes, exhibits higher rates of histone turnover and transcriptional noise, and contains more targets of chromatin remodelers, consistent with an ongoing dynamic competition between nucleosome assembly and factor binding[30,40]. Thus, poly(dA:dT) tracts are important determinants of nucleosome depletion *in vivo* and have likely been used by organisms to obtain biological outcomes.

Yeast terminator regions are also depleted of nucleosomes both *in vivo* and *in vitro*. In contrast to the situation at promoters, depletion of nucleosomes at terminator regions is strongly correlated with the orientation of and distance to neighboring genes, and it is strongly affected by growth conditions and transcriptional elongation by Pol II (ref. 41). Thus, the contribution of DNA sequence to depletion at terminators requires additional examination.

A minority of *S. cerevisiae* promoters are depleted for nucleosomes *in vivo*, yet lack poly(dA:dT) sequences and are not depleted in nucleosome assembly experiments *in vitro*. More generally, depletion of nucleosomes at promoters is observed in many species in which poly(dA:dT) sequences are less prevalent than in *S. cerevisiae* or are rare[42–44]. At these promoters, depletion of nucleosomes is not determined by DNA sequence but is likely due to activator-mediated recruitment of nucleosome remodelers that evict histones in promoter regions.

Aspects of positioning not determined by DNA sequence

Although nucleosome depletion can be reconstituted at most promoter sequences in nucleosome assembly experiments with purified histones and DNA, other aspects of the *in vivo* pattern of nucleosome positioning cannot be reconstituted[32,33]. In particular, strong positioning of the +1 nucleosome is not observed, and there is no evidence for any favored position in this region under conditions in which histone concentrations are either limiting or saturating. These observations suggest that the DNA sequence is not the main determinant of the position of the +1 nucleosome. Furthermore, positions of the +2 nucleosome and further downstream nucleosomes are determined primarily by the length of the linker region and hence by the position of the +1 nucleosome, and they are also not reconstituted *in vitro*.

Theoretically, the combination of a boundary constraint and high nucleosome concentrations along the DNA could generate an ordered array that begins with the +1 nucleosome and decays with increasing distance downstream. This 'statistical positioning' phenomenon could permit nucleosome-depleted regions mediated by poly(dA:dT) tracts to act as barriers and indirectly contribute to the long-range positioning of nucleosomes that flank promoters. However, statistically positioned arrays flanking poly(dA:dT) tracts were not observed even under *in vitro* conditions in which arrays were clearly generated[33], providing evidence against this theoretical possibility.

Overall, positions of nucleosomes that are assembled *in vitro* resemble those observed *in vivo* beyond random expectation, indicating that DNA sequence does contribute significantly to nucleosome positioning. However, in many cases, other factors can override the sequence preferences for nucleosome formation. Such overriding is likely to occur in the many genomic regions in which the dynamic range of affinities of the histone octamer to the underlying sequence is narrower than the more than three orders of magnitude observed for specific DNA sequences.

Role of nucleosome remodelers in nucleosome positioning

Several aspects of the *in vivo* nucleosome positioning pattern can be reconstituted *in vitro* if a yeast crude extract and ATP are added to purified histones and DNA[45]. Specifically, the combination of ATP-dependent nucleosome remodelers in a cell-free extract enhances depletion of nucleosomes at promoters to the extent observed *in vivo*, and it also generates positioned nucleosomes flanking the nucleosome-depleted regions (that is, the +1 and −1 nucleosomes). In principle, remodeling enzymes may simply allow nucleosomes to sample

positions rapidly, resulting in a thermodynamic equilibrium that is not achieved when nucleosome assembly is performed by salt dialysis. However, nucleosome assembly reactions containing *Drosophila melanogaster* ACF[33] or the yeast RSC[46] do not generate the *in vivo* pattern, even though they efficiently mobilize nucleosomes. These observations suggest that nucleosome-remodeling enzymes do not simply facilitate the movement of nucleosomes to preferred intrinsic positions, but rather are critical in establishing the specificity of where nucleosomes are located.

The mechanism by which nucleosome remodelers override intrinsic DNA sequence preferences of histone octamers is unclear. However, RSC[47,48] and perhaps other nucleosome remodelers can bind specific DNA sequences, and they are likely to have other sequence preferences for positioning nucleosomes[49,50]. In addition, nucleosome remodelers may be influenced by the boundary of nucleosome-disfavoring sequences, such that in the course of moving nucleosomes kinetic effects override thermodynamic equilibrium[50]. In this regard, nucleosome remodelers may use nucleosome-depleted regions as a measuring device to position the +1 and −1 nucleosomes. Notably, multiple nucleosome remodelers in yeast cell-free extracts are necessary to reconstitute the genome-wide pattern observed *in vivo*. Reactions involving individual nucleosome remodelers reconstitute only part of the pattern, either being restricted to subsets of genes or generating less precise positioning[46].

Although yeast nucleosome remodeling activities in themselves can reconstitute some aspects of the *in vivo* pattern, there are two important aspects of the pattern that they do not faithfully generate. First, the precise locations of the +1 nucleosomes generated *in vitro* poorly match those *in vivo*. Second, the extent of positioning of more downstream nucleosomes (for example, +3 and beyond) is substantially less in the *in vitro* reactions than what is observed *in vivo*. As a consequence, the nucleosome arrays over coding regions are less pronounced and appear shorter. These observations indicate that nucleosome remodelers are necessary but not sufficient to establish the *in vivo* pattern of nucleosome positioning. As we will discuss below, aspects of Pol II transcription also have critical roles in establishing the *in vivo* pattern.

In addition to data from *in vitro* reconstitution experiments, there is considerable genetic evidence for the critical role of nucleosome remodelers in establishing nucleosome positioning *in vivo*. In *S. cerevisiae*, a strain lacking the Isw2 remodeler shows altered nucleosome positioning adjacent to the promoter at the interface of genic and intergenic sequences[51]. Specifically, Isw2 helps to position the +1 nucleosome onto unfavorable DNA near the promoter in a directional manner, and it suppresses antisense transcription[52]. In addition, the RSC remodeling complex also mobilizes nucleosomes onto unfavorable sequences in the vicinity of the promoter[53].

Role of nucleosome remodelers in nucleosome spacing

As nucleosomes are typically arranged in regularly spaced arrays with a relatively constant linker length between nucleosomes, nucleosome spacing is a key aspect of nucleosome positioning. Drastic alteration of nucleosome positioning patterns is observed in an *S. cerevisiae* strain lacking both the Isw1 and Chd1 remodelers[54] or in an *S. pombe* strain lacking two related CHD remodelers (Hrp1 and Hrp3)[55–57]. In each of these evolutionarily diverged yeast species, positioning of the +2 nucleosome is much lower than that observed in the wild-type strain, and positioning of the +3 nucleosome and more downstream nucleosomes is essentially lost. This dramatic alteration of nucleosome positioning in coding regions is not due to histone depletion but rather to a defect in nucleosome spacing. Thus, a combination of the Isw1 and Chd1 (or Hrp1 and Hrp3) nucleosome remodeling enzymes is required for the correct nucleosome spacing that is the basis of positioned nucleosome arrays. Notably, a small subset of nucleosomes in downstream portions of coding regions are correctly positioned, presumably because of intrinsic DNA sequence preferences that facilitate or override the effect of nucleosome remodelers[54]. Positions of the +1 and −1 nucleosomes are essentially unaffected by the combined loss of Isw1 and Chd1 function, indicating that their positioning occurs by a distinct mechanism that may involve Isw2.

Additional support for a role of nucleosome remodelers for nucleosome spacing and positioning comes from a functional evolutionary experiment[35] that is based on the observation that nucleosome spacing varies among yeast species[43,44]. When large genomic regions from a foreign yeast species are introduced into *S. cerevisiae*, the distance between nucleosomes is characteristic of *S. cerevisiae*, not of the donor yeast species[35]. As a consequence, the vast majority of nucleosomes on the foreign DNA are not located at the positions that occur when the same DNA is present in the endogenous organism. The change in spacing could be due to differences in histone concentrations between the species, as the number of nucleosomes on genomic DNA will necessarily affect the average spacing. However, depletion of histone H3 does not generally alter nucleosome spacing, although some positioned nucleosomes are preferentially lost or maintained[55,58,59]. Or this observation may be due to species-specific differences in the spacing properties of the remodelers themselves because nucleosome assembly *in vivo* requires remodelers and remodelers have specific nucleosome-spacing properties that are independent of histone concentration[60].

In *S. cerevisiae*, histone H1 does not noticeably affect nucleosome spacing because histone H1 protein amounts are much lower than those of the core histones[10]. However, histone H1 and its various subtypes have an important role in nucleosome spacing in multicellular organisms. Overexpression or depletion experiments *in vivo* indicate that histone H1 increases the spacing between adjacent nucleosomes[61–64], and differences in linker histone subtypes might underlie cell type–specific differences in nucleosome spacing. In addition to histone H1, the HMG14, 17 proteins may also have a role in nucleosome spacing[61].

Role of transcription factors and Pol II elongation

In addition to DNA sequence and ATP-dependent nucleosome remodelers, Pol II transcription also contributes to establishing the genomic pattern of nucleosome positioning. In this respect, nucleosome positioning, transcription and perhaps other DNA-based processes such as DNA replication should be viewed as processes that reciprocally affect each other. The effect of Pol II on nucleosome positioning is achieved via transcriptional activator proteins, general transcription factors that compose the preinitiation complex and the elongating Pol II machinery.

Nucleosome depletion can generate positioning *in vivo*. Transcriptional activators, via targeted recruitment of nucleosome remodelers, can generate nucleosome-depleted regions. Some activator proteins can bind nucleosomal DNA fairly well, whereas others rely on intrinsic histone-destabilizing sequences or cooperativity with other activators to access their sites. Under standard growth conditions, the Rap1, Abf1 and Reb1 activators are important for generating nucleosome-depleted regions at subsets of *S. cerevisiae* genes[32,47,53]. When foreign DNA is introduced into *S. cerevisiae*, nucleosome-depleted regions often occur in coding regions, unlike the case in the native organisms[35]. Such *de novo* nucleosome-depleted regions

presumably arise from the fortuitous binding of *S. cerevisiae* activators to evolutionarily irrelevant target sites. Notably, many of these fortuitous nucleosome-depleted regions are associated with a positioned nucleosome array that strongly resembles the standard nucleosome positioning pattern of endogenous *S. cerevisiae* genes. Thus, generation of a nucleosome-depleted region, even in the absence of intrinsic nucleosome-destabilizing sequences, is sufficient to generate the *in vivo* pattern of nucleosome positioning.

Basal transcription factors help position the +1 nucleosome. The precise location of the +1 nucleosome is a critical determinant of the nucleosome positioning pattern because nucleosome spacing constraints are major determinants of the downstream nucleosome positions. The inability of the collection of nucleosome remodelers in crude extracts to accurately reconstitute the *in vivo* positions of +1 and −1 nucleosomes suggests that some other factor(s) is involved. The relationship between the position of the +1 nucleosome and the transcription start site (TSS) suggests that the general transcription factors have a role[33]. Furthermore, locations of the +1 nucleosome and TSS shift in concert when foreign yeast DNA is introduced into *S. cerevisiae*, such that TSS on the foreign DNA shifts to an *S. cerevisiae*–like location[35]. Given the strong *in vivo* positioning of both the preinitiation complex and the +1 nucleosome, a spacing relationship between these two entities requires that at least one of these be anchored to a specific location, thereby permitting a defined location for the second entity. Although the limited sequence specificity of nucleosome remodelers makes it unlikely that they can provide such an anchor, preinitiation complexes bound at core promoters may be sufficient, with the location of the TATA-binding protein bound to the TATA element or TATA-related sequence being the major determinant of the anchor point[65].

These considerations strongly suggest that the preinitiation complex has a role in fine-tuning the position of the +1 nucleosome. One speculative possibility is that a component(s) of the preinitiation complex is important for transiently recruiting the Isw2 and/or RSC complex that overrides inherent DNA sequence preferences to precisely position the +1 nucleosome. In addition, for organisms in which Pol II is often paused just downstream of the promoter[66], there is a strong distance relationship between the presence of paused Pol II and the NELF pausing factor and the position of the +1 nucleosome[67].

Pol II elongation and generating nucleosome arrays. Nucleosome arrays emanating from promoter regions occur unidirectionally in the transcribed direction, even though the +1 and −1 nucleosomes are well positioned. In addition, the decay of nucleosome positioning toward the center of genes displays a 5′-3′ asymmetry[68]. These initial observations suggested that the elongating Pol II machinery is important in establishing the pattern of nucleosome positioning. In accord with this idea, the Chd1 and Isw1 nucleosome remodelers that are critical for nucleosome positioning in coding regions have genome-wide association patterns that strongly resemble that of elongating Pol II. Conversely, nucleosome assembly in transcriptionally incompetent, cell-free extracts poorly recapitulates positioned nucleosome arrays in the downstream portions of coding regions, even though the +1 and −1 nucleosomes are strongly positioned. Lastly, the length of nucleosome arrays emanating from fortuitous nucleosome-depleted regions in foreign yeast DNA that act as functional promoters is strongly correlated both in direction and length with the mRNA[35].

These observations suggest that Pol II elongation strongly affects nucleosome positioning. In particular, the unidirectionality of nucleosome arrays can be easily explained by the unidirectionality of transcription, whereas it is unclear how such unidirectionality could be imposed only by nucleosome-depleted regions and nucleosome remodelers. Furthermore, the inefficiency of yeast cell extracts to reconstitute downstream nucleosome positions in the coding region suggests that recruitment of the Chd1 and Isw1 remodeling enzymes is not the sole mechanism by which elongating Pol II affects nucleosome positioning. Nevertheless, the mechanistic connection between Pol II elongation and nucleosome arrays in coding regions remains to be established.

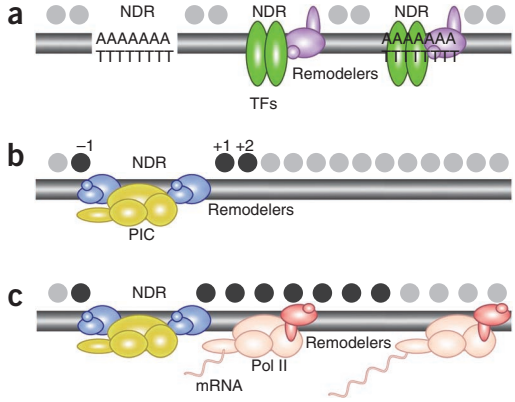

Figure 3 Determinants of nucleosome positioning. (**a**) Nucleosome-depleted regions (NDRs) are generated either by poly(dA:dT) tracts and/or by transcription factors and their recruited nucleosome remodeling complexes. Gray circles indicate nucleosomes. (**b**) Nucleosomes located at highly preferred positions (black circles) flanking the NDR are generated by nucleosome-remodeling complexes (for example, Isw2 and RSC, likely in a transcription-independent manner), and fine-tuned by the Pol II preinitiation complex (PIC) and associated factors. (**c**) Positioning of the more downstream nucleosomes depends on transcriptional elongation, and the recruitment of nucleosome-remodeling activities (for example, Chd1 and Isw1) and histone chaperones by the elongating Pol II machinery. This figure, which was modified from ref. 36, does not include DNA sequence determinants of rotational positioning (**Fig. 2**) or contributions to nucleosome spacing by histone H1.

Summary

The genetic, biochemical and informatics analyses performed in many laboratories together demonstrate that the genome-wide pattern of nucleosome positioning is determined by the combination of DNA sequence, nucleosome remodelers and transcription factors including activators, components of the preinitiation complex and elongating Pol II (**Fig. 3**). Although each of these components has discernible effects in isolation, they also reciprocally affect each other and hence affect the nucleosome positioning pattern in potentially complex ways. The DNA sequence is critical for rotational positioning along the DNA helix, and it also is an important determinant for nucleosome occupancy. In particular, poly(dA:dT) and poly(dG:dC) tracts are intrinsically inhibitory to nucleosome formation, whereas non-homopolymeric (G+C)-rich regions favor nucleosome formation. DNA sequence also contributes to the nucleosome positioning pattern, but several aspects of the *in vivo* pattern cannot be accounted for by intrinsic histone-DNA interactions.

As demonstrated by cross-species experiments, nucleosome-depleted regions are largely sufficient to generate the standard pattern of ordered nucleosome arrays[35]. At native promoters, nucleosome-depleted regions can be generated intrinsically via poly(dA:dT) tracts and/or activator-mediated recruitment of nucleosome remodelers that evict histones in the vicinity of the activator-binding site. The use of poly(dA:dT) sequences varies considerably among organisms;

they are very common in *S. cerevisiae* but not in the fission yeast *S. pombe* and in multicellular organisms[42–44].

In addition to their role in creating nucleosome-depleted regions, ATP-dependent remodelers are important in nucleosome positioning. In a manner independent of transcription, they can assemble nucleosomes flanking the depleted region to generate the +1 and −1 nucleosomes. However, these remodeling enzymes cannot accurately position the +1 nucleosome to the *in vivo* location nor can they generate properly positioned nucleosomes at more downstream locations. Precise positioning of the +1 nucleosome is strongly influenced by the location of the preinitiation complex, although the mechanistic basis for this spacing relationship is poorly understood. Lastly, generation of positioned nucleosome arrays throughout the coding region is coupled to Pol II elongation, and it involves Pol II–dependent recruitment of nucleosome remodelers and perhaps some other aspect of Pol II elongation. The putative roles of Isw2 and RSC for positioning the +1 nucleosome and of Isw1 and Chd1 for transcription-coupled positioning and spacing of nucleosomes in the coding region are strongly supported by genome-wide mapping of nucleosome remodelers on positioned mononucleosomes[69].

In some situations where a gene is activated, it is likely that establishment of the nucleosome pattern occurs stepwise: activator-mediated generation of a nucleosome-depleted region, positioning of the +1 and −1 nucleosomes, and elongation-coupled assembly of positioned and properly spaced nucleosome arrays over the coding region. However, such a stepwise process is unnecessary, and probably irrelevant, for steady-state maintenance of the nucleosome positioning pattern. Instead, the combined effects of DNA sequence, nucleosome remodelers and transcription factors make independent contributions to the pattern.

Lastly, it is important to mention that the nucleosome positioning pattern described here is gene-averaged and hence represents a typical pattern. Thus, although the positioning mechanisms discussed here apply to all genes, the precise pattern at individual genes may differ from the gene-averaged pattern. Such variations will depend on the DNA sequences, the nucleosome remodelers that act at the gene and the complexity of transcription units within the gene (for example, antisense or unstable transcripts). The relative contributions of these mechanisms may differ among genes, which is likely to result in differential regulation of these genes. Furthermore, even when the typical pattern predominates at a given gene, it is highly likely that the pattern is not present in all cells. For this reason, differences in the nucleosome positioning pattern among cells in a population are likely to confer distinct transcriptional properties in these cells. Nevertheless, although several questions remain open, we believe that there is now a good understanding of how many aspects of the nucleosome positioning patterns are generated *in vivo*.

ACKNOWLEDGMENTS

This Review is dedicated to the memory of Jon Widom, our dear friend and colleague, who is sorely missed by the authors and the scientific community. We thank D. Schneider for encouraging us to write this Review, and Z. Moqtaderi, T. Raveh-Sadka, M. Levo, N. Kaplan and Y. Field for comments on the manuscript. This work was supported by grant GM 30186 from the US National Institutes of Health to K.S. and grants from the European Research Council and the US National Institutes of Health to E.S.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nsmb.2506.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Richmond, T.J. & Davey, C.A. The structure of DNA in the nucleosome core. *Nature* **423**, 145–150 (2003).
2. Yuan, G.-C. *et al.* Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**, 626–630 (2005).
3. Lee, W. *et al.* A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **39**, 1235–1244 (2007).
4. Mavrich, T.N. *et al.* A barrier nucleosome model for statistical positioning of nucleosome throughout the yeast genome. *Genome Res.* **18**, 1073–1083 (2008).
5. Mavrich, T.N. *et al.* Nucleosome organization in the *Drosophila* genome. *Nature* **453**, 358–362 (2008).
6. Schones, D.E. *et al.* Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898 (2008).
7. Shivaswamy, S. *et al.* Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* **6**, e65 (2008).
8. Valouev, A. *et al.* A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* **18**, 1051–1063 (2008).
9. Valouev, A. *et al.* Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516–520 (2011).
10. Freidkin, I. & Katcoff, D.J. Specific distribution of the *Saccharomyces cerevisiae* linker histone HHO1 in the chromatin. *Nucleic Acids Res.* **29**, 4043–4051 (2001).
11. Lam, F.H., Steger, D.J. & O'Shea, E.K. Chromatin decouples promoter threshold from dynamic range. *Nature* **453**, 246–250 (2008).
12. Raveh-Sadka, T. *et al.* Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nat. Genet.* **44**, 743–750 (2012).
13. Varshavsky, A.J., Sundin, O. & Bohn, M. A stretch of "late" SV40 viral DNA about 400 bp long which induces the origin of replication is specifically exposed in SV40 minichromosomes. *Cell* **16**, 453–466 (1979).
14. Wu, C. The 5′ ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* **286**, 854–860 (1980).
15. Iyer, V. & Struhl, K. Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic structure. *EMBO J.* **14**, 2570–2579 (1995).
16. Liu, X., Lee, C.K., Granek, J.A., Clarke, N.D. & Lieb, J.D. Whole-genome comparison of Leu3 binding *in vitro* and *in vivo* reveals the importance of nucleosome occupancy in target site selection. *Genome Res.* **16**, 1517–1528 (2006).
17. Zaret, K.S. & Carroll, J.S. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**, 2227–2241 (2011).
18. Workman, J.L. & Kingston, R.E. Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu. Rev. Biochem.* **67**, 545–579 (1998).
19. Thastrom, A. *et al.* Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* **288**, 213–229 (1999).
20. Drew, H.R. & Travers, A.A. DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* **186**, 773–790 (1985).
21. Satchwell, S.C., Drew, H.R. & Travers, A.A. Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* **191**, 659–675 (1986).
22. Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442**, 772–778 (2006).
23. Brogaard, K.R., Xi, L., Wang, J.P. & Widom, J. A map of nucleosome positions in yeast at base-pair resolution. *Nature* **486**, 496–501 (2012).
24. Nelson, H.C.M., Finch, J.T., Luisi, B.F. & Klug, A. The structure of an oligo(dA)·oligo(dT) tract and its biological implications. *Nature* **330**, 221–226 (1987).
25. Suter, B., Schnappauf, G. & Thoma, F. Poly(dA-dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters *in vivo*. *Nucleic Acids Res.* **28**, 4083–4089 (2000).
26. Segal, E. & Widom, J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **19**, 65–71 (2009).
27. McCall, M., Brown, T. & Kennard, O. The crystal structure of d(G-G-G-G-C-C-C-C): a model for poly(dG)·poly(dC). *J. Mol. Biol.* **183**, 385–396 (1985).
28. Dechering, K.J., Cuelenaere, K., Konings, R.N. & Leunissen, J.A. Distinct frequency-distributions of homopolymeric DNA tracts in different genomes. *Nucleic Acids Res.* **26**, 4056–4062 (1998).
29. Struhl, K. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* **82**, 8419–8423 (1985).
30. Field, Y. *et al.* Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* **4**, e1000216 (2008).
31. Gaffney, D.J. *et al.* Controls of nucleosome positioning in the human genome. *PLoS Genet.* **8**, e1003036 (2012).
32. Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366 (2009).
33. Zhang, Y. *et al.* Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 847–852 (2009).
34. Sekinger, E.A., Moqtaderi, Z. & Struhl, K. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* **18**, 735–748 (2005).
35. Hughes, A., Jin, Y., Rando, O.J. & Struhl, K. A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. *Mol. Cell* **48**, 5–15 (2012).

36. Chen, W., Tabor, S. & Struhl, K. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. *Cell* **50**, 1047–1055 (1987).
37. Zhu, Z. & Thiele, D.J. A specialized nucleosome modulates transcription factor access to a *C. glabrata* metal responsive promoter. *Cell* **87**, 459–470 (1996).
38. Zeevi, D. *et al.* Compensation for differences in gene copy number among yeast ribosomal proteins is encoded within their promoters. *Genome Res.* **21**, 2114–2128 (2011).
39. Field, Y. *et al.* Gene expression divergence in yeast is coupled to evolution of DNA-encoded nucleosome organization. *Nat. Genet.* **41**, 438–445 (2009).
40. Tirosh, I. & Barkai, N. Two strategies for gene regulation by promoter nucleosomes. *Genome Res.* **18**, 1084–1091 (2008).
41. Fan, X. *et al.* Nucleosome depletion in yeast terminator regions is not intrinsic and can occur by a transcriptional mechanism linked to 3′ end formation. *Proc. Natl. Acad. Sci. USA* **107**, 17945–17950 (2010).
42. Lantermann, A.B. *et al. Schizosaccharomyces pombe* gene-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*. *Nat. Struct. Mol. Biol.* **17**, 251–257 (2010).
43. Tsankov, A.M., Thompson, D.A., Socha, A., Regev, A. & Rando, O.J. The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol.* **8**, e1000414 (2010).
44. Tsankov, A., Yanagisawa, Y., Rhind, N., Regev, A. & Rando, O.J. Evolutionary divergence of intrinsic and trans-regulated nucleosome positioning sequences reveals plastic rules for chromatin organization. *Genome Res.* **21**, 1851–1862 (2011).
45. Zhang, Z. *et al.* A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332**, 977–980 (2011).
46. Wippo, C.J. *et al.* The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *EMBO J.* **30**, 1277–1288 (2011).
47. Badis, G. *et al.* A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol. Cell* **32**, 878–887 (2008).
48. Floer, M. *et al.* A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell* **141**, 407–418 (2010).
49. Rippe, K. *et al.* DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc. Natl. Acad. Sci. USA* **104**, 15635–15640 (2007).
50. van Vugt, J.J. *et al.* Multiple aspects of ATP-dependent nucleosome translocation by RSC and Mi-2 are directed by the underlying DNA sequence. *PLoS ONE* **4**, e6345 (2009).
51. Whitehouse, I. & Tsukiyama, T. Antagonistic forces that position nucleosomes *in vivo*. *Nat. Struct. Mol. Biol.* **13**, 633–640 (2006).
52. Whitehouse, I., Rando, O.J., Delrow, J. & Tsukiyama, T. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031–1035 (2007).
53. Hartley, P.D. & Madhani, H.D. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**, 445–458 (2009).
54. Gkikopoulos, T. *et al.* A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. *Science* **333**, 1758–1760 (2011).
55. Hennig, B.P., Bendrin, K., Zhou, Y. & Fischer, T. Chd1 chromatin remodelers maintain nucleosome organization and repress cryptic transcription. *EMBO Rep.* **13**, 997–1003 (2012).
56. Pointner, J. *et al.* CHD1 remodelers regulate nucleosome spacing *in vitro* and align nucleosomal arrays over gene coding regions in *S. pombe*. *EMBO J.* **31**, 4388–4403 (2012).
57. Shim, Y.S. *et al.* Hrp3 controls nucleosome positioning to suppress non-coding transcription in eu- and heterochromatin. *EMBO J.* **31**, 4375–4387 (2012).
58. Celona, B. *et al.* Substantial histone reduction modulates genomewide nucleosomal occupancy and global transcriptional output. *PLoS Biol.* **9**, e1001086 (2011).
59. Gossett, A.J. & Lieb, J.D. *In vivo* effects of histone H3 depletion on nucleosome occupancy and position in *Saccharomyces cerevisiae*. *PLoS Genet.* **8**, e1002771 (2012).
60. Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R. & Kadonaga, J.T. ACF, an ISW1-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**, 145–155 (1997).
61. Drew, H.R. Reconstitution of short-spaced chromatin from the histone octamer and either HMG14,17 or histone H1. *J. Mol. Biol.* **230**, 824–836 (1993).
62. Fan, Y. *et al.* Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123**, 1199–1212 (2005).
63. Hashimoto, H. *et al.* Histone H1 null vertebrate cells exhibit altered nucleosome architecture. *Nucleic Acids Res.* **38**, 3533–3545 (2010).
64. Oberg, C., Izzo, A., Schneider, R., Wrange, O. & Belikov, S. Linker histone subtypes differ in their effect on nucleosome spacing *in vivo*. *J. Mol. Biol.* **419**, 183–197 (2012).
65. Rhee, H.S. & Pugh, B.F. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**, 295–301 (2012).
66. Adelman, K. & Lis, J.T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).
67. Chang, G.S. *et al.* Unusual combinatorial involvement of poly-A/T tracts in organizing genes and chromatin in *Dictyostelium*. *Genome Res.* **22**, 1098–1106 (2012).
68. Vaillant, C. *et al.* A novel strategy of transcription regulation by intragenic nucleosome ordering. *Genome Res.* **20**, 59–67 (2010).
69. Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T. & Pugh, B.F. Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* **149**, 1461–1473 (2012).