



## A novel strategy of transcription regulation by intragenic nucleosome ordering

Cédric Vaillant, Leonor Palmeira, Guillaume Chevereau, et al.

*Genome Res.* 2010 20: 59-67 originally published online October 26, 2009

Access the most recent version at doi:[10.1101/gr.096644.109](https://doi.org/10.1101/gr.096644.109)

---

**References** This article cites 47 articles, 9 of which can be accessed free at:  
<http://genome.cshlp.org/content/20/1/59.full.html#ref-list-1>

### License

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

An advertisement banner with a teal background. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." in white. In the center, there is a white-bordered box containing the text "LEARN MORE". On the right, there is a photograph of a woman wearing a red and white superhero cape and mask, and the Cellecta logo, which consists of a cluster of green dots and the word "CELLECTA" in white.

---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

---

Copyright © 2010 by Cold Spring Harbor Laboratory Press

# A novel strategy of transcription regulation by intragenic nucleosome ordering

Cédric Vaillant,<sup>1,2,5</sup> Leonor Palmeira,<sup>1,2,3,5</sup> Guillaume Chevereau,<sup>1,2</sup> Benjamin Audit,<sup>1,2</sup> Yves d'Aubenton-Carafa,<sup>4</sup> Claude Thermes,<sup>4</sup> and Alain Arneodo<sup>1,2,6</sup>

<sup>1</sup>Université de Lyon, F-69000 Lyon, France; <sup>2</sup>Laboratoire Joliot-Curie and Laboratoire de Physique, CNRS, ENS-Lyon, 69364 Lyon Cedex 07, France; <sup>3</sup>Université Lyon 1, CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622 Villeurbanne, France; <sup>4</sup>Centre de Génétique Moléculaire, CNRS, 91198 Gif-sur-Yvette, France

Numerous studies of chromatin structure showed that nucleosome free regions (NFRs) located at 5' gene ends contribute to transcription initiation regulation. Here, we determine the role of intragenic chromatin structure on gene expression regulation. We show that, along *Saccharomyces cerevisiae* genes, nucleosomes are highly organized following two types of architecture that depend only on the distance between the NFRs located at the 5' and 3' gene ends. In the first type, this distance constrains in vivo the positioning of  $n$  nucleosomes regularly organized in a "crystal-like" array. In the second type, this distance is such that the corresponding genes can accommodate either  $n$  or  $(n + 1)$  nucleosomes, thereby displaying two possible crystal-like arrays of  $n$  weakly compacted or  $n + 1$  highly compacted nucleosomes. This adaptability confers "bi-stable" properties to chromatin and is a key to its dynamics. Compared to crystal-like genes, bi-stable genes present higher transcriptional plasticity, higher sensitivity to chromatin regulators, higher H3 turnover rate, and lower H2A.Z enrichment. The results strongly suggest that transcription elongation is facilitated by higher chromatin compaction. The data allow us to propose a new paradigm of transcriptional control mediated by the stability and the level of compaction of the intragenic chromatin architecture and open new ways for investigating eukaryotic gene expression regulation.

[Supplemental material is available online at <http://www.genome.org>.]

The small-scale chromatin structure, as defined by the local nucleosome occupancy, conditions the regulation of transcription by modulating the accessibility of transcription factors to their cognate regulatory sites (Kornberg and Lorch 1999; Li et al. 2007; Morse 2007; Rando and Ahmad 2007; Segal and Widom 2009). In yeast, recent high-resolution experiments revealed a specific gene promoter nucleosome pattern characterized by a nucleosome depleted region (NFR) of typical size 100–200 base pairs (bp) upstream of the transcription start site (TSS) and, to a lesser extent, at the gene 3' end (Yuan et al. 2005; Albert et al. 2007; Lee et al. 2007; Mavrich et al. 2008; Shivaswamy et al. 2008). Interestingly, gene expression level is negatively correlated with nucleosome occupancy at promoter, whereas relative nucleosome occupancy would rather control the ability for genes to adapt their expression under environmental changes (Lam et al. 2008; Tirosh and Barkai 2008). For example, genes with higher relative occupancy upstream of the TSS present higher transcriptional plasticity (Tirosh and Barkai 2008; Radman-Livaja and Rando 2009). In contrast, genes with a pronounced nucleosome depleted region at the TSS are generally constitutively expressed. Some genes, thus, have a rather static chromatin pattern facilitating the permanent access/recruitment of activators/repressors, while others present a more homogeneous and dynamic chromatin reflecting a fine regulation by nucleosome positioning at regulatory sites (Morse 2007; Boeger et al. 2008; Lam et al. 2008; Tirosh and Barkai 2008).

These observations raise the issue of determining if intragenic nucleosome organization might also be a determinant of gene

expression regulation and, if this is the case, to what extent this organization would be directly encoded (hard-wired) in the underlying gene sequence? At promoters, NFRs are enriched in poly(dA:dT) motifs that are known to alter nucleosomal DNA structure (Bao et al. 2006) and to favor nucleosome disassembly (Iyer and Struhl 1995; Suter et al. 2000). Periodically well-positioned nucleosomes reveal a 10 bp periodic distribution of AA/TT dinucleotides (Ioshikhes et al. 2006; Mavrich et al. 2008; Shivaswamy et al. 2008), but this periodicity mainly concentrates at the two proximal nucleosomes on both sides of the TSS (Mavrich et al. 2008; Shivaswamy et al. 2008). It is unlikely that this sequence specificity accounts for the periodic stretches of nucleosomes observed genome-wide and, in particular, along genes, since most positioned nucleosomes are not associated with strong nucleosome positioning sequences (Peckham et al. 2007; Yuan and Liu 2008). Along this line, in vitro experiments showed that more than 95% of genomic sequences have the same affinity for histones than random DNA sequences (Lowary and Widom 1997). Periodic ordering of nucleosomes along gene sequences could then result from nonlocal effects induced by the NFRs located at both gene extremities (Kornberg and Stryer 1988). Here, we analyze yeast genome-wide nucleosome mapping and reveal a strikingly organized intragenic chromatin architecture. We show that the distance between the 5' and 3' NFRs controls this architecture and that this boundaries-directed nucleosome organization constitutes a newly evidenced strategy of gene expression regulation.

## Results

### Yeast genes display a highly organized nucleosomal pattern

In order to determine how nucleosomes are organized along gene length, we ordered the genes by the distance  $L$  that separates

<sup>5</sup>These authors contributed equally to this work.

<sup>6</sup>Corresponding author.

E-mail [alain.arneodo@ens-lyon.fr](mailto:alain.arneodo@ens-lyon.fr); fax 33-4-72-72-80-80.

Article published online before print. Article and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.096644.109>.

the first 5' nucleosome downstream of the 5' NFR and the last 3' nucleosome upstream of the 3' NFR (Mavrich et al. 2008; Venters and Pugh 2009). Note that we used  $L$  as a substitute for the distance  $\mathcal{L}$  between the 5' and 3' NFRs, which was more difficult to measure accurately because of the NFR shape variability. We plotted vertically, from top to bottom, from the shortest to the longest gene, the organized nucleosome occupancy profiles observed in vivo (Lee et al. 2007) (Methods) and obtained a two-dimensional (2D) map that reveals a strikingly organized nucleosome ordering (Fig. 1A). Small genes ( $L \leq 1.5$  kbp) present a clear periodic packing between the two bordering NFRs with a well-defined number  $n$  of regularly spaced nucleosomes (Fig. 1C). As  $L$  increases, these "crystallized" genes cluster in domains of genes displaying the same number  $n$  of nucleosomes, from  $n = 2$  to about 10 nucleosomes ( $L \leq 1.5$  kbp). For larger gene size, the periodic nucleosome positioning remains visible, but appears fuzzy in the region where the confinement induced by both boundaries is probably too weak to strongly constrain the positioning of central nucleosomes. Close inspection of this 2D map shows that the range of influence of the 5' NFR extends to about seven nucleosomes as compared to about five nucleosomes for the 3' NFR. This likely results from NFR nucleosome occupancy profiles that are more pronounced, on average, at the 5' than at the 3' gene extremities (Supplemental Fig.

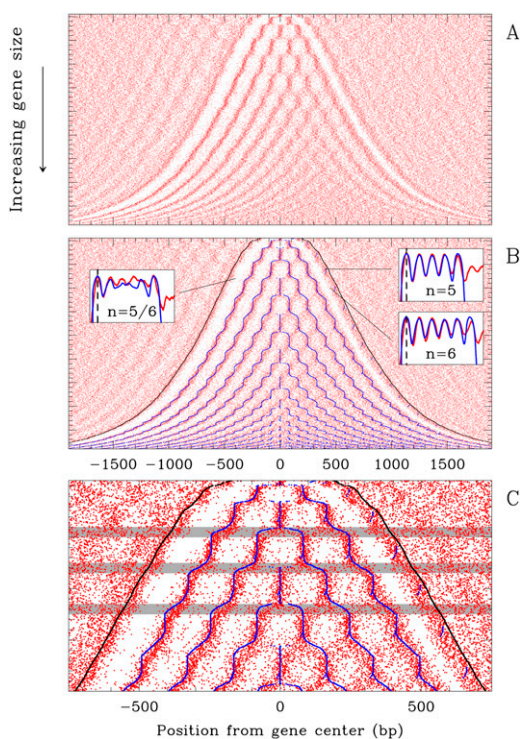
S1). We thus observe a periodic ordering that extends inside the genes, consistently with a statistical ordering mechanism induced by exclusion from the boundaries. In this ordering mechanism, the strength, period, and range depend on the degree of nucleosome confinement that increases with the nucleosome exclusion strength of the boundaries and with the average nucleosome density (Kornberg and Stryer 1988). The more nucleosomes are confined, the more they adopt a long-range and compact periodic organization, with the interboundary distance as an additional control parameter.

This pattern was compared to a 2D map of intergenic regions that displayed totally different nucleosome distributions (Supplemental Fig. S2). Along intergenic regions, which consist in regions located between the NFRs of neighbor genes, only the bordering nucleosomes are well-positioned. We observe no propagation of nucleosome phasing from bordering to internal nucleosomes that present a fuzzy distribution. This lack of intergenic crystal-like nucleosomal organization is likely related to the asymmetric shape of the NFRs that are significantly sharper on the gene side than on the intergenic side, where their range of influence appears to be limited to the closest nucleosome (Supplemental Fig. S1).

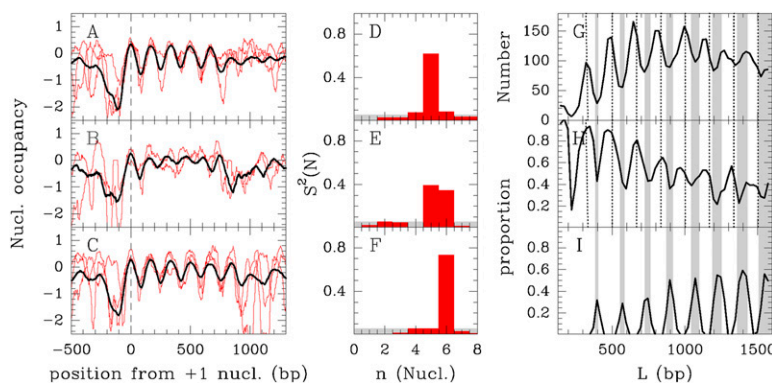
### Two classes of genic chromatin architecture: crystal-like versus bi-stable structures

To determine the contributions to the observed nucleosomal distribution of successive periodicities corresponding to  $n$ ,  $(n + 1)$ , ... nucleosomal arrays, we performed the spectral analysis of gene nucleosome occupancy profiles (see Methods) and assigned them to one of three categories, crystal-like (1940/4554), bi-stable (946/4554) and other (1668/4554), depending on the crystallization state of their chromatin (the lists of these genes are given in the Supplemental material). Genes presenting a profile with a single and dominant periodic contribution were considered as "crystal genes" (Fig. 2A,C,D,F). Genes presenting two periodic contributions were considered as "bi-stable" genes. Among large genes, we also detected genes presenting more than two periodic contributions, thus displaying properties of multi-stable genes.

We first observed a quantized distribution of  $L$  values for all genes, with maxima corresponding to crystal-like genes with  $L = n \times 167$  ( $n = 2, 3$ , etc. . .) (Fig. 2G). It results that genes with  $L$  values corresponding to bi-stability are underrepresented. Indeed, our criteria to select bi-stable genes from their power spectrum was chosen very stringent on purpose (Methods) in order to ensure no contamination by "crystal" or "other" genes; this certainly contributed to some underestimation of the bi-stable category to the benefit of higher confidence levels. In agreement with analyses of the 2D map (Fig. 1A), we found that the mean proportion of crystal genes clearly presents a 167 bp periodicity as a function of  $L$  up to around 1.5 kbp (Fig. 2H). In addition, this proportion is globally larger at small  $L$  values and decreases when increasing  $L$ . This actually defines a crystallization regime associated with  $L$  values, in which large domains of crystal-like genes alternate with smaller domains, where crystal-gene density locally drops. At transitions between  $n$  and  $n + 1$  domains (Fig. 2I), intermediate  $L$  windows can be identified where the nucleosome occupancy profile becomes seemingly fuzzy without clear nucleosome positioning (Fig. 2B). As revealed by the power spectrum analysis, this fuzziness does not mean that the crystallization mechanisms are no longer at work inside those genes. The presence of two dominating peaks in the power spectrum reveals the statistical coexistence of two crystal-like chromatin states with  $n$  and  $n + 1$  nucleosomes (Fig. 2B,E).



**Figure 1.** 2D map of nucleosomes along yeast genes. (A) The 4554 genes are ordered vertically by the distance  $L$  between the NFRs first (5') and last (3') nucleosomes. The nucleosome occupancy profile of each gene is figured along a horizontal line: red dots correspond to the minima of nucleosome occupancy; nucleosomes occupy the white zones; in vivo data are retrieved from (Lee et al. 2007). (B) Predictions of our physical modeling (blue) are drawn on top of experimental data (red). (Insets) Mean experimental (red) and theoretical (blue) nucleosome occupancy profiles for crystal genes harboring 5 nucleosomes (right, top), 6 nucleosomes (right, bottom), and for bi-stable genes with 5/6 nucleosomes (left). (C) Zoom on the first 2000 genes in B; gray-shaded areas correspond to some bi-stable  $L$ -domains. In B and C, the black curves indicate the 5' - and 3' -end positions of the theoretical excluding nucleosome energy barriers.



**Figure 2.** Average experimental nucleosome occupancy profiles of crystal-like and bi-stable genes (black), together with some examples of individual gene profiles (red): (A) crystal genes with  $n = 5$  nucleosomes; (B) bi-stable genes at the transition between the  $n = 5$  and  $n = 6$  crystal gene domains; (C) crystal genes with  $n = 6$  nucleosomes. Power spectrum (PS) (see Methods) of nucleosome occupancy profiles: (D) PS of one crystal gene profile with a dominant peak at  $n = 5$  nucleosomes; (E) PS of one bi-stable gene with two main peaks of comparable magnitude at  $n = 5$  and  $n = 6$  nucleosomes; (F) PS of one crystal gene profile with a dominant peak at  $n = 6$  nucleosomes. Total number of genes (G), proportion (density) of crystal genes (H), and proportion (density) of bi-stable genes (I) as a function of the gene size  $L$  (see Methods); the plotted data correspond to average values computed in a 50 bp sliding window. In G–I, the vertical gray bands define bi-stability domains and the vertical dotted lines indicate the successive  $L = n \times 167$ , with  $n = 2, 3, \dots$

Indeed when increasing  $L$ , more peaks can appear in the power spectrum suggesting some evolution from a bi-stable to a multi-stable structure.

This quantized distribution of  $L$  values observed for all genes classes (Fig. 2G,H,I) actually reflects specific crystallization mechanisms at work inside genes. A naive way to order nucleosomes in a crystal-like fashion is to impose a strict period of 167 bp, which will lead to a trivially quantized distribution of  $L$  distances:  $L = n \times 167$  bp. A more realistic interpretation of the observed quantized distributions is to relax this very strict constraint allowing the nucleosomal repeat length (NRL) to fluctuate around a given value with some lower and upper bounds:  $l_{min} < NRL = L/(n-1) \simeq \mathcal{L}/n < l_{max}$ . Then  $n$ -crystal states will exist in the range  $(n-1)l_{min} < L < (n-1)l_{max}$ . When  $L$  slightly increases [ $nl_{min} < L < (n-1)l_{max}$ ], there is possible coexistence of  $n$ - and  $(n+1)$ -crystal states. Similarly, when  $L$  slightly decreases [ $(n-1)l_{min} < L < (n-2)l_{max}$ ], there is possible coexistence of  $(n-1)$ - and  $n$ -crystal states. This very tentative scenario raises the issue of the crystallization mechanisms that would produce such NRL distribution and, in turn, the observed succession of  $n$ -crystal genes separated by bi-stable (multi-stable) genes (Fig. 1A).

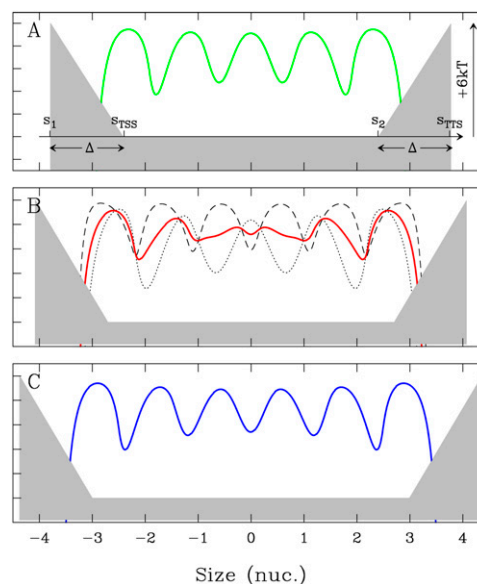
We investigated whether crystal-like and bi-stable gene classes display some enrichment in specific functions (Supplemental Table S1). We only observed a slightly significant enrichment in cytoskeleton organization and biogenesis genes among bi-stable genes (56%,  $P = 10^{-2}$ ) and in genes involved in translation among crystal-like genes (74%,  $P = 10^{-2}$ ).

### Intragenic chromatin architecture conforms to equilibrium statistical ordering principles

Does the intragenic nucleosome organization only depend on the properties (distance and nucleosome exclusion strength) of the 5' and 3' NFRs? To test this hypothesis, we predicted the nucleosome occupancy profile inside the 4554 yeast genes using a thermodynamical model of nucleosome assembly that depends on the nucleosome size  $l$ , the chemical potential  $\mu$  (that fixes the aver-

age nucleosome density) and the nucleosome energy landscape  $E(s)$  (see Methods, Equation 1). To test whether a statistical ordering mechanism induced by excluding boundaries at gene extremities may account for the observed crystallization features (Fig. 1), we imposed a fixed force at gene extremities in the form of linear energy barriers (see Methods, Equation 2) to mimic the effect of experimental NFRs (Fig. 3A). For simplicity, we considered identical 5' and 3' energy barrier shapes. In addition, no sequence effect was added in the energy profile  $E(s)$  that remained constant between the bordering walls. This, coupled with a hard-core interaction between nucleosomes, allowed us to compute the thermodynamical equilibrium density of nucleosomes along genes (see Methods, Equation 1) and to build the corresponding intragenic nucleosome occupancy profile (see Methods, Equation 3). By fixing the nucleosome size to  $l = 146$  bp and adjusting the chemical potential so that nucleosomes cover 75% of the yeast genome, we obtained the averaged nucleosome occupancy profiles, as exemplified for  $n = 5$  and 6 nucleosomes in Figure 3. The resulting 2D map (Fig. 1B) is in remarkable agreement with the in vivo data. This simple model predicts the existence of  $L$  domains of crystallization characterized by a defined number of nucleosomes (Fig. 3A,C) alternating with bi-stable domains presenting a seemingly fuzzy occupancy profile (Fig. 3B). Globally, crystallization is observed up to gene sizes of at

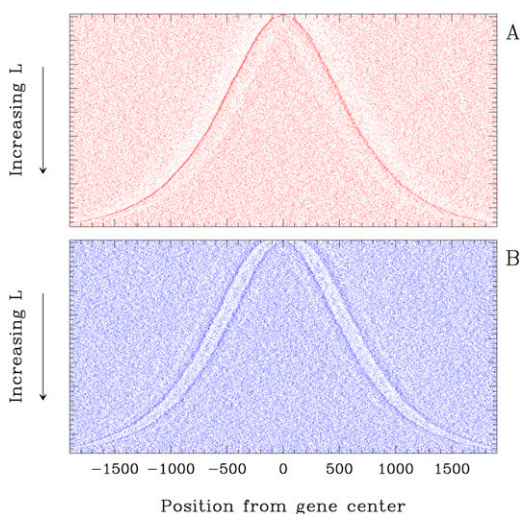
nucleosomes cover 75% of the yeast genome, we obtained the averaged nucleosome occupancy profiles, as exemplified for  $n = 5$  and 6 nucleosomes in Figure 3. The resulting 2D map (Fig. 1B) is in remarkable agreement with the in vivo data. This simple model predicts the existence of  $L$  domains of crystallization characterized by a defined number of nucleosomes (Fig. 3A,C) alternating with bi-stable domains presenting a seemingly fuzzy occupancy profile (Fig. 3B). Globally, crystallization is observed up to gene sizes of at



**Figure 3.** Theoretical probability of nucleosome occupancy at each point of a box bordered by two linear energy barriers at gene extremities (Methods, Equation 2): (A) Box large enough to shelter five nucleosomes (green); (B) larger box where the two dotted and dashed configurations are possible; the weighted average of the five and six nucleosome crystal-like profiles yields a fuzzy-looking average profile (red); and (C) larger box where six nucleosomes can be accommodated.

least 10 nucleosomes, as shown in the *in vivo* 2D map. In our modeling, the 2D-map geometry is symmetrical with respect to gene center by construction although *in vivo* the 3' NFR range of influence is slightly smaller than on the 5' side. This elementary thermodynamical modeling thus provides a simple interpretation of the crucial role of the interbarrier distance as a fine control parameter of chromatin structure, and demonstrates that the complex pattern of intragenic nucleosome distribution mainly depends on the presence of the bordering NFRs and not so much on the gene sequence.

This result is not in contradiction with the predictions of models recently proposed to account for sequence effect on nucleosome positioning, such as the Field et al. (2008) model, based on statistical learning (model I), and the Vaillant et al. (2007) model, based on the computation of the free-energy cost of bending a DNA fragment of a given sequence from its natural curvature to the final superhelical structure around the histone core (model II). These models predict nucleosome occupancy profiles that account remarkably well for the recent nucleosome positioning *in vitro* data (Kaplan et al. 2009) obtained at lower nucleosome density (30% as compared to 75% *in vivo*): the Pearson's correlation coefficient is  $r = 0.77$  for model I and  $r = 0.70$  for model II (correlation between the two predicted profiles,  $r = 0.81$ ). But when comparing the 2D maps of nucleosome occupancy along the yeast genes constructed with the *in vitro* data (Fig. 4A) and predicted by model II (Fig. 4B), we observe, as expected, some positioning of the 5' and 3' nucleosomes at gene extremities. However, in the intragenic regions located between these bordering nucleosomes, none of these 2D maps display the regular crystal-like and bi-stable nucleosomal patterns displayed *in vivo* (Fig. 1A). Consistently with the previous studies, these results demonstrate that the intragenic sequences are unable to generate these strong regular nucleosome positioning patterns *in vitro* (at low nucleosome density and in the absence of external factors such as transcription factors, remodelers, etc.). They confirm that the crystal-like and bi-stable nucleosomal patterns result from the thermodynamical ordering at high nucleosome density imposed by excluding energy barriers at gene extremities.



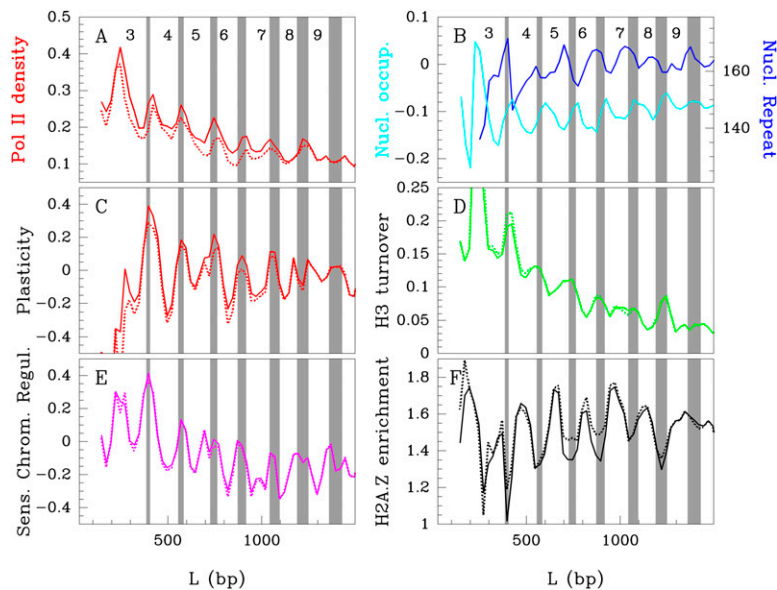
**Figure 4.** 2D map of nucleosome occupancy along yeast genes (same coding as in Fig. 1A). (A) *In vitro* data retrieved from Kaplan et al. (2009). (B) Theoretical nucleosome energy gain  $[-E(s)]$  landscape predicted by the model II (Vaillant et al. 2007).

This understanding of intragenic nucleosome organization, in terms of statistical ordering, is further supported by the nucleosome positioning observed *in vivo* in the central region (1 kbp) of the  $N = 483$  longest yeast genes with  $L > 3000$  bp. In these regions far away from the influence of the bordering inhibitory energy barriers, 1270 nucleosomes were found to be well-positioned (see Methods) corresponding to  $\sim 39\%$  coverage of the sequence (data from Lee et al. 2007), i.e., a number smaller than the coverage  $\sim 80\%$  observed close to the gene extremities due to boundary confining, but significantly larger than zero as an indication that the sequence plays some role in the statistical positioning of central nucleosomes (Supplemental Fig. S3). Importantly, among this set of well-positioned central nucleosomes observed *in vivo*, only 191 correspond to (at a 35 bp precision) well-positioned nucleosomes also observed in the *in vitro* data (Kaplan et al. 2009). Since intrinsic histone–DNA interactions are likely to result into well-positioned nucleosomes at the lower *in vitro* nucleosome density, this means that less than 15% (191/1270) of *in vivo* well-positioned nucleosomes in central intragenic regions can be attributed to strongly positioning DNA sequences. Again the major determinant of nucleosome positioning is the statistical ordering no longer imposed by the bordering excluding barriers, but by the fluctuations in the central energy landscape induced by the DNA sequence. As expected, our toy model with constant energy profile  $E(s)$  between the bordering barriers did not generate any well-positioned nucleosomes in the central region of large genes. When artificially introducing in between the energy barriers, the sequence-dependent energy profile predicted by the model II (Vaillant et al. 2007), we numerically generated nucleosome occupancy profiles (Supplemental Fig. S3) that exhibit some central nucleosome positioning. Indeed 1309 nucleosomes are predicted to be well-positioned (corresponding to  $\sim 40\%$  coverage of the sequence) among which we recover 28% (360/1270) of the *in vivo* central nucleosomes including 53% (100/191) of the ones also observed *in vitro*. Moreover, when shuffling the gene sequence, we recovered a similar number  $\sim 1140$  of well-positioned nucleosomes as the signature of statistical ordering in a noisy energy landscape, but we lost a majority (74/100) of the nucleosomes intrinsically positioned by sequences with high affinity to histones.

Altogether our results agree with the central conclusions of recent experimental *in vitro* and *in vivo* studies of nucleosome positioning in *S. cerevisiae* (Zhang et al. 2009), which provide additional evidence that intrinsic histone–DNA interactions make only a modest contribution to the *in vivo* intragenic nucleosome positioning pattern, the major determinant being the statistical ordering mainly induced by the excluding NFRs located at gene extremities.

#### Intragenic nucleosome density correlates positively with transcription rate

We next assessed the functional implications of the intragenic chromatin structure by analyzing how different features relate to the two types of nucleosomal organizations identified in the *in vivo* data. We first examined the way transcription rate, estimated by the RNA polymerase II (Pol II) density (Steinmetz et al. 2006), relates to  $L$  distance (Fig. 5A). In individual crystal  $n$ -domains, the transcription rate decreases when increasing  $L$ . Some significant anti-correlation is actually observed between internucleosome distance (Fig. 5B) and transcription rate in  $n$ -crystal domains: for  $n = 3$ ,  $r = -0.2$ ,  $P = 3.6 \times 10^{-2}$ ;  $n = 4$ ,  $r = -0.24$ ,  $P = 7 \times 10^{-4}$ ;  $n = 5$ ,  $r = -0.21$ ,  $P = 3.5 \times 10^{-3}$ ;  $n = 6$ ,  $r = -0.3$ ,  $P = 1.3 \times 10^{-3}$ , and  $n = 7$ ,



**Figure 5.** Bi-stable nucleosome organization controls gene expression. Sliding window (50 bp) analysis of average transcription rate estimated by Pol II density (Steinmetz et al. 2006) (A), nucleosome repeat length (dark-blue) and intragenic mean nucleosome occupancy (light-blue) (B), transcriptional plasticity (Tirosh and Barkai 2008) (C), H3 turnover rate (Dion et al. 2007) (D), sensitivity to chromatin regulators disruption (Steinfeld et al. 2007) (E), and H2A.Z occupancy (Zhang et al. 2005; Tirosh and Barkai 2008) (F), as a function of the distance  $L$ . The vertical gray bands define bi-stability domains. The dotted curves correspond to the results obtained when excluding the 175 (out of 4554) ribosomal protein genes from the analysis; no significant changes are observed.

$r = -0.23$ ,  $P = 5 \times 10^{-3}$ . Hence, the more compact the nucleosomal array, the shorter the linker size, the higher the transcription rate. Consistently, in the intermediate bi-stable domains between the  $n$  - and  $(n + 1)$  -crystal domains, we observed a local and sharp increase of the rate coinciding with an increasing proportion of the compact  $(n + 1)$  -crystal pattern that coexists with the more diluted  $n$  -crystal chromatin state.

#### Genes showing a bi-stable dynamic chromatin structure are expression regulated genes

The distribution of transcriptional plasticity, which quantifies for each gene the variation of expression level as the result of environmental condition changes (Tirosh and Barkai 2008), presents large variations with the  $L$  distance, bi-stable (crystal) genes presenting high (low) plasticity values (Fig. 5C). Bi-stable genes present, at least, two crystal chromatin states, a diluted ( $NRL = L/(n-1)$ ) and a compact ( $NRL = L/n$ ) state. According to our model, their transcription rate profile results from the contributions of high and low transcription rates. These genes may thus adopt one of the two crystal states and switch from one to the other state under some perturbation or condition changes, e.g., a very small change in the  $L$  distance, that would in turn lead to strong changes of transcription rate and to a high level of transcriptional plasticity. In full agreement with this suggestion, we observed that bi-stable gene density is significantly correlated to transcriptional plasticity ( $r = 0.42$ ,  $P = 1.1 \times 10^{-3}$ ). These properties of the plasticity distribution strongly sustain our model of the nucleosomal array compaction dependence of transcription rate.

In a recent study, it has been shown that gene expression can substantially vary when disrupting chromatin regulators, including chromatin modifiers, such as HATs, HDACs, HMTs, and

ATP-dependent chromatin remodeling factors (Steinfeld et al. 2007). The distribution of sensitivity to chromatin regulators reveals that bi-stable genes are, on average, significantly more sensitive than crystal genes due to such a lack of structural regulation (Fig. 5E). This is consistent with the fact that chromatin regulators may control the interbarrier distance  $\mathcal{L}$  and in turn the distance  $L$  ( $\sim \mathcal{L} - 188$  bp) via the statistical ordering imposed by nucleosome excluding barriers. Note that in the precise case of Isw2 remodeling (Whitehouse et al. 2007), we did not obtain conclusive results concerning the crystal/bi-stable nature of the target genes (data not shown). When compared to crystal genes, bi-stable genes are thus highly plastic, with a wide dynamic range of expression level as the signature of a more dynamic and regulated chromatin structure.

#### Intragenic chromatin structure and epigenetic marks

We next examined the histone variant H2A.Z occupancy (Zhang et al. 2005; Tirosh and Barkai 2008) and the histone H3 turnover rate (Dion et al. 2007) within

crystal and bi-stable gene domains. Crystal genes present a high enrichment in H2A.Z (Fig. 5F), whereas bi-stable genes present a high H3 turnover rate (Fig. 5D). These results are in agreement with the hypothesis that the H2A.Z histone variant, which is mainly deposited at the 5' nucleosome position (Albert et al. 2007), contributes to stabilize this nucleosome, thereby reinforcing the position and possibly the nucleosome exclusion strength of the 5' end NFR. This would, in turn, enhance the crystallization properties (periodicity and phasing) of the nucleosome array along the gene. Consequently H2A.Z incorporation may contribute to lock the chromatin into a very stable crystal state. By contrast, low H2A.Z enrichment in the 5' nucleosome of bi-stable genes may favor the ability for these genes to change from one to the other of the  $n$  - and  $(n + 1)$  -nucleosome states. On the other hand, the high H3 turnover rate observed for bi-stable genes likely reflects the dynamical nature of their chromatin resulting from a facilitated transition between the  $n$  - and  $(n + 1)$  -nucleosome states. More frequent eviction and reassembly of one nucleosome may thus improve a de novo histone H3 deposition.

#### Discussion

The role of chromatin structure (nucleosome density and positioning) on gene expression regulation has been mostly investigated at the level of transcription initiation and different regulation strategies associated to different kinds of promoter structural design have been revealed (Lee et al. 2007; Morse 2007; Boeger et al. 2008; Lam et al. 2008; Tirosh and Barkai 2008; Segal and Widom 2009). Here, we show that the intragenic chromatin architecture is significantly associated with the regulation of gene expression, very likely at the level of transcription elongation, and that this architecture is, to a main part, a consequence of a statistical ordering induced by

inhibitory boundaries located at both gene extremities. In particular, we demonstrate that a thermodynamical model of nucleosome assembly at equilibrium accounts very well for the evolution of the chromatin pattern as a function of the gene size measured by the distance between the 5' and 3' NFRs. We mainly identified two types of nucleosomal architectures: crystal-like genes with regularly positioned nucleosomes and bi-stable genes with a fuzzy-looking nucleosomal profile resulting from the coexistence of two possible crystal-like states with different compaction levels. As compared to crystal-like genes that present a constitutive expression level, bi-stable genes show a higher transcriptional plasticity and are more sensitive to chromatin regulators. We further show that the transcription rate tends to increase when the nucleosome linker size decreases. Accordingly, by means of a single nucleosome switching, bi-stable genes may drastically alter their expression level in response to external changes. Interestingly we obtained the same functional/structural relationship for the intragenic chromatin as those observed by Tirosh and Barkai (2008) at promoters. Crystal genes, as DPN (depleted proximal nucleosome) genes, correspond to constitutively expressed genes with a well-defined and stable chromatin pattern, i.e., a periodic nucleosome arrangement for crystal genes and a well-localized and pronounced NFR for DPN (Supplemental Fig. S4). On the other hand, bi-stable genes, as OPN (occupied proximal nucleosomes) genes, correspond to transcriptionally plastic genes with a more dynamic and regulated chromatin. We observe a seemingly fuzzy pattern for bi-stable genes (Fig. 2B) that actually corresponds to the statistical superposition of two crystal states and an extended less depleted region for OPN genes (Supplemental Fig. S4). However, despite similar structural regulation strategies, we did not observe any significant correlation between the bi-stable/crystal nature of genes and their OPN/DPN promoter architecture (Supplemental Fig. S5). For both crystal-like and bi-stable genes, nucleosome organization results from the presence of NFRs at gene extremities. These gene categories are not expected to be correlated with the OPN and DPN gene classes that reflect the presence or the absence of a deep trough in the nucleosome occupancy profile upstream of the TSS. DPN genes are significantly enriched in crystal and bi-stable genes (56% are crystal and 25% are bi-stable) as compared to OPN genes (40% are crystal and 20% are bi-stable). This latter level of enrichment indicates that OPN genes still present sufficiently strongly bound nucleosomes at the 5' and 3' ends to induce nucleosome confinement and a periodic ordering for short gene sizes, as confirmed by the periodicity observed at the 5' gene end in the average nucleosome occupancy profile (Supplemental Fig. S4).

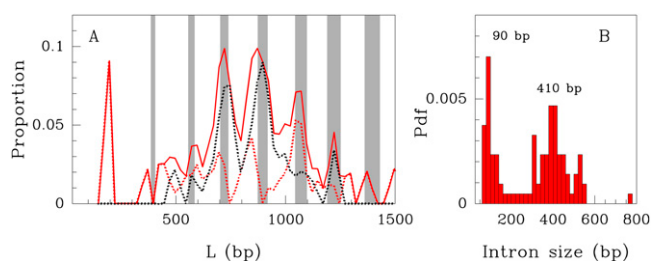
Since intragenic nucleosome arrangement may drastically affect the elongation process, our analysis raises the question of why a well-ordered and regularly compacted nucleosome array would enhance the transcription rate. As previously proposed (Morse 2007), a compact 10 nm chromatin fiber is likely to restrict aberrant binding of transcription factors or of other factors that may perturb the proper progression of Pol II and/or induce cryptic transcription. Crystallization of the nucleosome array actually means stronger linear confinement, i.e., well-defined nucleosome positioning within genes. Pol II elongation has to deal with chromatin structure and, in particular, requires structural regulation that involves nucleosome modifications and remodeling. We guess that a regular intragenic nucleosome array might, upon gene activation, enhance transcription rate by reducing intrinsic disturbances. According to geometrical modeling of the 30 nm chromatin fiber (Lesne and Victor 2006), a short linker size would rather lead to an open well-ordered chromatin secondary structure that would facilitate the sequential action of chromatin regulators

associated with Pol II progression, such as the FACT complex (Hartzog 2003), as well as the action of chromatin modifiers. Recently, reversomes (for reverse nucleosomes, built on a right-handed tetrasome) were proposed to form under the action of a DNA supercoiling wave pushed in front of the RNA polymerase (Lavelle and Prunell 2007). The nucleosome reversome would facilitate transcription elongation by giving the RNA polymerase a lever to break the docking of the H2A–H2B dimers, which otherwise exerts a stringent blocking against transcription in absence of other factors (Bancaud et al. 2007). By a dynamical domino-like effect, the elongation process of compact regularly spaced nucleosomal intragenic arrays might result in a reversome wave that would progress faster than the RNA polymerase.

To examine a possible relation between chromatin state and the presence of introns, we selected the 107 genes containing one intron from the 4554 yeast genes and observed that they preferentially localize within bi-stable domains (Fig. 6A). We suggest that intron size in yeast might work as a mechanism allowing for the transition of a gene from one to another crystal-like/bi-stable nucleosomal architecture. Observation that intronic size distribution is bimodal with peaks at 90 and 410 bp (Fig. 6B) sustains this hypothesis. Similar bimodal distributions were also obtained for *Saccharomyces kluyverii*, *Kluyveromyces thermotolerans*, and *Debaromyces hansenii* (data not shown). Indeed, 90 bp is about the distance that would allow a gene to change from bi-stable to crystal chromatin state (and vice versa). Unfortunately, when using our model to compute the nucleosome occupancy profiles on intron-containing intragenic sequences with and after removing the intron (Parenteau et al. 2008), the sampling of each category (crystal, bi-stable, other) was too small to draw significant conclusions. An alternative hypothesis relies on recent studies indicating that splicing regulation might be associated with chromatin-mediated regulation of transcriptional elongation, involving the action of remodeling factors, such as Swi/Snf (Kornblihtt 2006; Allemand et al. 2008). A proper chromatin structure might thus be favored in order to facilitate/regulate intron splicing.

Ribosomal protein genes are generally short, highly transcribed, and intron containing. Among the 107 single intron-containing genes, the subset of 54 ribosomal protein genes is significantly enriched in bi-stable genes (Fig. 6A), consistent with our previous observation concerning the high plasticity level of bi-stable genes (Fig. 5).

The nucleosomal structure of promoters and its implications in transcription initiation has been studied in various organisms from yeast to human, including the nematode and *Drosophila* (Bernstein et al. 2004; Lee et al. 2004, 2007; Yuan et al.



**Figure 6.** (A) Proportion of single intron containing yeast genes (107/4554) as a function of the gene size  $L$ ; the vertical gray bands define bi-stability domains. The dotted curves correspond to the density of ribosomal protein genes (54) (black dotted curve) and to the density of the nonribosomal protein genes (53) (red dotted curve). (B) Histogram of intron sizes in single-intron containing genes.

2005; Ozsolak et al. 2007; Mavrich et al. 2008; Miele et al. 2008; Shivaswamy et al. 2008; Tirosh and Barkai 2008; Valouev et al. 2008). In this work, we have identified new regulation mechanisms in *S. cerevisiae* that involve intragenic chromatin structure. To what extent these chromatin-mediated regulation processes generalize to distant eukaryotic species is a very challenging question for future studies.

## Methods

### Data source

#### Gene data

A selection of 4554 yeast genes for which high-confidence transcription start and termination sites could be assigned was determined as follows. We used the high resolution transcriptome analysis from David et al. (2006), in which genome-wide tiling arrays were used to define a set of transcribed segments along the *S. cerevisiae* genome. Transcribed segments overlapping known nonambiguous genes on less than 50% of the segment length were not considered. The transcription start and end sites were determined from the boundaries of the given transcribed segments. Note that the transcribed segment used to determine the 5' end could be different from the one used for the 3' end of the gene. This resulted in 4554 genes for which high-confidence transcription start and termination sites could be assigned. Genes containing introns were downloaded from the Saccharomyces Genome Database (SGD project; <http://www.yeastgenome.org/>). Only genes with one intron were retained for our analysis of intron size. Genome-wide data on yeast RNA polymerase II (Pol II) were retrieved from Steinmetz et al. (2006). A given position was considered enriched in Pol II when its score was in the highest quintile. We also considered that each polymerase has a certain occupancy in space, therefore we considered each enriched Pol II position to be included in a  $\pm 30$  bp surrounding box. Transcription plasticity was retrieved from (Tirosh and Barkai 2008): it is the average of the square  $\log_2$  expression ratio estimated from a large number of microarray experiments. Sensitivity to disruption of chromatin modifiers was obtained from Steinfeld et al. (2007) and Tirosh and Barkai (2008): it quantifies the extent by which gene expression depends on chromatin regulators activity.

#### Nucleosome data

In vivo nucleosome occupancy for 4554 yeast genes ( $\log_2$  ratio) was retrieved from microarray data. For the computation of the power spectrum (see below), we used the microarray data from (Whitehouse et al. 2007) that correspond to the detrended component of the nucleosome occupancy profile as obtained by Lee et al. (2007). In vitro nucleosome occupancy data were retrieved from Kaplan et al. (2009). OPN and DPN classes were taken from Tirosh and Barkai (2008, Supplemental Table 2). Histone variant Htz1 (also known as H2A.Z) occupancy was obtained from Zhang et al. (2005) and Tirosh and Barkai (2008). Histone H3 turnover between transcription start and end sites were determined from chromatin immunoprecipitation with microarray hybridization (ChIP-chip) genome-wide data (Dion et al. 2007).

### Nucleosome data analysis

#### 5' / 3' nucleosomes

To determine the first ("5'") and last ("3'") nucleosomes of a given gene in relation to the transcription start sites (TSS) and transcription termination sites (TTS), we smoothed the nucleosome

occupancy ( $\log_2$  ratio) profile by a ( $\sigma = 18$  bp) Gaussian window and then computed the first (respectively last) maxima of this smoothed signal in the interval  $[-50$  bp,  $+200$  bp] (respectively  $[-200$  bp,  $+50$  bp]) with respect to the TSS (respectively TTS). The extra 50 bp considered on the left (respectively right) of the TSS (respectively TTS) allowed for an inaccuracy in the annotated position of the transcript 5' (respectively 3') extremity.

#### Bi-stable and crystal-like genes

For each of the 4554 genes, we computed the power spectrum  $S^2(k)$ ,  $k = 0, \dots, L/2$  and the normalized power spectrum  $\tilde{S}^2(k) = S^2(k) / \sum_{i=1}^{L/2} S^2(i)$ , of the nucleosome occupancy profile in between the 5' and 3' bordering nucleosomes, using the fftw3 library (<http://www.fftw.org>) interfaced with LastWave signal processing software (<http://www.cmap.polytechnique.fr/~bacry/LastWave>). The main maxima of this spectrum were determined, as well as their corresponding periods. Periods were considered informative when in the range [125 bp, 210 bp]; they were otherwise discarded. Crystal-like genes were defined as having a unique maximum over a 0.06 threshold. Bi-stable genes were defined as having their first two maxima over a 0.06 threshold and the mean of the two periods in the range [160 bp, 170 bp]. This threshold was intentionally chosen stringent enough to ensure that the set of bi-stable genes was not contaminated by crystal or other genes. This led to some underestimation of the set of bi-stable genes.

#### Well-positioned nucleosomes

To determine the nucleosomes that can be considered as well-positioned in the central region of large genes (Supplemental Fig. S3), we smoothed the nucleosome occupancy data by a ( $\sigma = 18$  bp) Gaussian window and then selected the maxima above a 0.20 threshold for in vivo data from Lee et al. (2007) and a 1.5 threshold for in vitro data from Kaplan et al. (2009). For the theoretical nucleosome occupancy profile, we selected the maxima of the smoothed  $P(s)$  (see Equation 3, below) above a 0.718 threshold.

### Physical model

#### Nucleosome density profile

As proposed in Vaillant et al. (2007), when focusing on the dynamical assembly of histone octamers along the DNA chain, chromatin can be reasonably modeled by a fluid of one-dimensional rods of finite extension  $l$  (the DNA wrapping length around the octamer), binding and moving in an external potential  $E(s, l)$  (the effective nucleosome formation potential at position  $s$ ), and interacting through a hard core potential of size  $l$ . Within the grand canonical formalism, considering that the fluid is in contact with a thermal bath (at reciprocal temperature  $\beta$ ) and a histone octamer reservoir (at chemical potential  $\mu$ ), the equilibrium density  $\rho(s)$  of hard rods in an external field  $E(s, l)$  obeys the nonlinear integral equation (Percus 1976):

$$\beta\mu = \beta E(s) + \ln \rho(s) - \ln \left( 1 - \int_s^{s+l} \rho(s') ds' \right) + \int_{s-l}^s \frac{\rho(s')}{1 - \int_{s'}^{s'+l} \rho(s'') ds''} ds' \quad (1)$$

#### Nucleosome formation energy: Modeling NFRs at gene extremities

Our model of the energy landscape  $E(s)$  consists in simply taking into account the NFRs observed in vivo at yeast gene extremities, and in neglecting the effect of the gene sequence. If a significant part of the observed NFRs at yeast genes TSS and TTS are encoded in

the DNA sequence (Vaillant et al. 2007; Miele et al. 2008), some are not, probably due to the binding of external proteic factors like transcription factors or PIC/Pol II. To model the observed 5' and 3' NFRs, we locally imposed the presence of an "effective" energy barrier at both the TSS ( $s = s_{TSS}$ ) and the TTS ( $s = s_{TTS}$ ). The linear shape of these energy barriers amounts to impose a constant force on both sides of the intragenic nucleosome array. Inside the genes, we considered the energy to be constant. This led to the following energy profile (see Fig. 3A):

$$\left. \begin{array}{l} E(s) = E_M \\ E(s) = E_M(1 - (s - s_1)/\Delta) \\ E(s) = 0 \\ E(s) = E_M(s - s_2)/\Delta \\ E(s) = E_M \end{array} \right\} \text{for } \begin{array}{l} s < s_1 = s_{TSS} - \Delta \\ s_1 < s < s_{TSS} \\ s_{TSS} < s < s_2 = s_{TTS} - \Delta \\ s_2 < s < s_{TTS} \\ s > s_{TTS} \end{array}, \quad (2)$$

where the parameters  $E_M = 6kT$  and  $\Delta = 80$  bp were fixed to reproduce as faithfully as possible the 2D map of nucleosome occupancy along yeast genes (Fig. 3B,C).

To investigate the role of the DNA sequence on nucleosome positioning in the central region of the largest yeast genes, we have considered, as an alternative model, an energy profile  $E(s)$  with the same bordering energy barriers, but with the sequence-dependent nucleosome wrapping energy profile in between these barriers predicted by model II (Vaillant et al. 2007).

### Nucleosome occupancy profile

Equation 1 has an explicit solution (Vanderlick et al. 1986) that requires numerical integration. We fixed the chemical potential to  $\mu = 1$  kT and the hard-rod (nucleosome) size to  $l = 146$  bp to get the 75% nucleosome coverage observed in vivo. The nucleosome occupancy probability profile  $P(s)$  was obtained by convolving the nucleosome density  $\rho(s)$  with the rectangular function  $\Pi$  of width 146 bp:

$$P(s) = \rho \circ \Pi_{146}(s). \quad (3)$$

$P(s)$  is the probability for a base pair located at  $s$  to be occupied by a nucleosome of length 146 bp.

### Acknowledgments

This work was supported by the Conseil Régional Rhône-Alpes (Emergence 2005) and the Agence Nationale de la Recherche under projects DNAnucl (ANR-06-PCVI-0026) and HUGOREP (NT05-3\_41825).

### References

Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF. 2007. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**: 572–576.

Allemand E, Batsché E, Muchardt C. 2008. Splicing, transcription, and chromatin: A ménage à trois. *Curr Opin Genet Dev* **18**: 145–151.

Bancaud A, Wagner G, Conde E, Silva N, Lavelle C, Wong H, Mozziconacci J, Barbi M, Sivolob A, Le Cam E, et al. 2007. Nucleosome chiral transition under positive torsional stress in single chromatin fibers. *Mol Cell* **27**: 135–147.

Bao Y, White CL, Luger K. 2006. Nucleosome core particles containing a poly(dA-dT) sequence element exhibit a locally distorted DNA structure. *J Mol Biol* **361**: 617–624.

Bernstein BE, Liu CL, Humphrey EL, Perlstein EO, Schreiber SL. 2004. Global nucleosome occupancy in yeast. *Genome Biol* **5**: R62. doi: 10.1186/gb-2004-5-9-r62.

Boeger H, Griesenbeck J, Kornberg RD. 2008. Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. *Cell* **133**: 716–726.

David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, Steinmetz LM. 2006. A high-resolution map of transcription in the yeast genome. *Proc Natl Acad Sci* **103**: 5320–5325.

Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ. 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**: 1405–1408.

Field Y, Kaplan N, Fondufe-Mittendorf Y, Moore IK, Sharon E, Lubling Y, Widom J, Segal E. 2008. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput Biol* **4**: e1000216. doi: 10.1371/journal.pcbi.1000216.

Hartzog GA. 2003. Transcription elongation by RNA polymerase II. *Curr Opin Genet Dev* **13**: 119–126.

Ioshikhes IP, Albert I, Zanton SJ, Pugh BF. 2006. Nucleosome positions predicted through comparative genomics. *Nat Genet* **38**: 1210–1215.

Iyer V, Struhl K. 1995. Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J* **14**: 2570–2579.

Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, et al. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**: 362–366.

Kornberg RD, Lorch Y. 1999. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**: 285–294.

Kornberg RD, Stryer L. 1988. Statistical distributions of nucleosomes: Nonrandom locations by a stochastic mechanism. *Nucleic Acids Res* **16**: 6677–6690.

Kornblihtt AR. 2006. Chromatin, transcript elongation and alternative splicing. *Nat Struct Mol Biol* **13**: 5–7.

Lam FH, Steger DJ, O'Shea EK. 2008. Chromatin decouples promoter threshold from dynamic range. *Nature* **453**: 246–250.

Lavelle C, Prunell A. 2007. Chromatin polymorphism and the nucleosome superfamily: A genealogy. *Cell Cycle* **6**: 2113–2119.

Lee C-K, Shibata Y, Rao B, Strahl BD, Lieb JD. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* **36**: 900–905.

Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. 2007. A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39**: 1235–1244.

Lesne A, Victor J-M. 2006. Chromatin fiber functional organization: Some plausible models. *Eur Phys J E Soft Matter* **19**: 279–290.

Li B, Carey M, Workman JL. 2007. The role of chromatin during transcription. *Cell* **128**: 707–719.

Lowary PT, Widom J. 1997. Nucleosome packaging and nucleosome positioning of genomic DNA. *Proc Natl Acad Sci* **94**: 1183–1188.

Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* **18**: 1073–1083.

Miele V, Vaillant C, d'Aubenton-Carafa Y, Thermes C, Grange T. 2008. DNA physical properties determine nucleosome occupancy from yeast to fly. *Nucleic Acids Res* **36**: 3746–3756.

Morse RH. 2007. Transcription factor access to promoter elements. *J Cell Biochem* **102**: 560–570.

Ozsolak F, Song JS, Liu XS, Fisher DE. 2007. High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* **25**: 244–248.

Parenteau J, Durand M, Véronneau S, Lacombe A-A, Morin G, Guérin V, Cecez B, Gervais-Bird J, Koh C-S, Brunelle D, et al. 2008. Deletion of many yeast introns reveals a minority of genes that require splicing for function. *Mol Biol Cell* **19**: 1932–1941.

Peckham HE, Thurman RE, Fu Y, Stamatoiyannopoulos JA, Noble WS, Struhl K, Weng Z. 2007. Nucleosome positioning signals in genomic DNA. *Genome Res* **17**: 1170–1177.

Percus J. 1976. Equilibrium state of a classical fluid of hard rods in an external field. *J Stat Phys* **15**: 505–511.

Radman-Livaja M, Rando OJ. 2009. Nucleosome positioning: How is it established, and why does it matter? *Dev Biol*. doi: 10.1016/j.ydbio.2009.06.012.

Rando OJ, Ahmad K. 2007. Rules and regulation in the primary structure of chromatin. *Curr Opin Cell Biol* **19**: 250–256.

Segal E, Widom J. 2009. What controls nucleosome positions? *Trends Genet* **25**: 335–343.

Shivaswamy S, Bhinge A, Zhao Y, Jones S, Hirst M, Iyer VR. 2008. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol* **6**: e65. doi: 10.1371/journal.pbio.0060065.

Steinfeld I, Shamir R, Kupiec M. 2007. A genome-wide analysis in *Saccharomyces cerevisiae* demonstrates the influence of chromatin modifiers on transcription. *Nat Genet* **39**: 303–309.

- Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. 2006. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell* **24**: 735–746.
- Suter B, Schnappauf G, Thoma F. 2000. Poly(dA-dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters *in vivo*. *Nucleic Acids Res* **28**: 4083–4089.
- Tirosh I, Barkai N. 2008. Two strategies for gene regulation by promoter nucleosomes. *Genome Res* **18**: 1084–1091.
- Vaillant C, Audit B, Arneodo A. 2007. Experiments confirm the influence of genome long-range correlations on nucleosome positioning. *Phys Rev Lett* **99**: 218103. doi: 10.1103/PhysRevLett.99.218103.
- Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, et al. 2008. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res* **18**: 1051–1063.
- Vanderlick TK, Scriven LE, Davis HT. 1986. Solution of Percus's equation for the density of hard rods in an external field. *Phys Rev A* **34**: 5130–5131.
- Venters BJ, Pugh BF. 2009. How eukaryotic genes are transcribed. *Crit Rev Biochem Mol Biol* **44**: 117–141.
- Whitehouse I, Rando OJ, Delrow J, Tsukiyama T. 2007. Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**: 1031–1035.
- Yuan G-C, Liu JS. 2008. Genomic sequence is highly predictive of local nucleosome depletion. *PLoS Comput Biol* **4**: e13. doi: 10.1371/journal.pcbi.0040013.
- Yuan G-C, Liu Y-J, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. 2005. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**: 626–630.
- Zhang H, Roberts DN, Cairns BR. 2005. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123**: 219–231.
- Zhang Y, Moqtaderi Z, Rattner BP, Euskirchen G, Snyder M, Kadonaga JT, Liu XS, Struhl K. 2009. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nat Struct Mol Biol* **16**: 847–852.

Received May 29, 2009; accepted in revised form October 5, 2009.