Published online 22 July 2009

Nucleic Acids Research, 2009, Vol. 37, No. 17 ***5641–5655***
doi:10.1093/nar/gkp610

Predicting nucleosome positions on the DNA: combining intrinsic sequence preferences and remodeler activities

Vladimir B. Teif[1,2,*] and Karsten Rippe[1,*]

[1]Research Group Genome Organization & Function, Deutsches Krebsforschungszentrum and BioQuant, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany and [2]Institute of Bioorganic Chemistry, Belarus National Academy of Sciences, Kuprevich 5/2, 220141, Minsk, Belarus

Received March 9, 2009; Revised July 3, 2009; Accepted July 6, 2009

ABSTRACT

Nucleosome positions on the DNA are determined by the intrinsic affinities of histone proteins to a given DNA sequence and by the ATP-dependent activities of chromatin remodeling complexes that can translocate nucleosomes with respect to the DNA. Here, we report a theoretical approach that takes into account both contributions. In the theoretical analysis two types of experiments have been considered: *in vitro* experiments with a single reconstituted nucleosome and *in vivo* genome-scale mapping of nucleosome positions. The effect of chromatin remodelers was described by iteratively redistributing the nucleosomes according to certain rules until a new steady state was reached. Three major classes of remodeler activities were identified: (i) the establishment of a regular nucleosome spacing in the vicinity of a strong positioning signal acting as a boundary, (ii) the enrichment/depletion of nucleosomes through amplification of intrinsic DNA-sequence-encoded signals and (iii) the removal of nucleosomes from high-affinity binding sites. From an analysis of data for nucleosome positions in resting and activated human CD4$^+$ T cells [Schones *et al.*, Cell 132, p. 887] it was concluded that the redistribution of a nucleosome map to a new state is greatly facilitated if the remodeler complex translocates the nucleosome with a preferred directionality.

INTRODUCTION

In eukaryotes, the DNA-binding of many protein factors involved in transcription, replication, repair and recombination is dependent on the chromatin organization. In particular, the wrapping of DNA around the histone octamer complex has long been recognized as a mechanism to make DNA sequences inaccessible for the binding of other proteins. Accordingly, changes of nucleosome positions at promoter and enhancer regions have been shown to directly affect gene expression (1–4). In addition, nucleosomes may also play an architectural role by mediating interactions between regulatory proteins bound at distant sites during transcription initiation (5). Nucleosome positions are controlled by three major contributions (1,6). First, the intrinsic binding affinity of the histone octamer depends on the DNA sequence, and a number of natural and artificial high-affinity binding sequences have been identified (7,8). A growing set of data has correlated the *in vivo* nucleosome positions with histone affinities to different DNA sequences (2,8–11). Second, the nucleosome can be displaced or recruited by the competitive or cooperative binding of other protein factors (12–14). Third, the nucleosome may be actively translocated by ATP-dependent remodeling complexes. This reaction can be modulated by a competitive displacement/binding event (9,15). Furthermore, recent experiments have shown that the result of the remodeling reactions is directed also by the DNA sequence, and that different nucleosome remodeling complexes display characteristic translocation activities in this respect (16). Several mechanisms for ATP-dependent nucleosome translocation along the DNA have been proposed (17–19). The available data argue in favor of a loop recapture model. According to this model, the partial unwrapping of a small segment of the intranucleosomal DNA (e.g. 10–50 bp) leads to formation of a loop that is subsequently propagated around the histone octamer protein core (19–21).

Several recent studies have devised strategies to predict whole-genome nucleosome positions based on the intrinsic nucleosome-DNA affinities (8,10,11,13,22–24). A strong correlation of nucleosome positions with the sites predicted from the DNA sequence has been revealed in

*To whom correspondence should be addressed. Tel: +49 6221 54 51378; Email: Vladimir.Teif@bioquant.uni-heidelberg.de
Correspondence may also be addressed to Karsten Rippe. Tel: +49 6221 54 51376; Email: Karsten.Rippe@bioquant.uni-heidelberg.de

yeast (8,25,26) and *Drosophila* (10,27,28), but not in nematodes (29). In addition, in many instances nucleosome positions can change while the DNA sequence remains the same. For example, numerous cases of removing 'repressive' nucleosomes from sites where they block promoter access by remodeling complexes have been reported (30–32). In a recent study in yeast the genome-wide effects of the RSC remodeler have been investigated (15). Significant deviations of the nucleosome maps from the DNA sequence-determined ones were identified, and it has been concluded that the RSC activity is an important determinant of nucleosome positions. Global changes in the nucleosome pattern have also been observed during activation of human T cells (33). Finally, it is known that striking differences in the nucleosome repeat length exist between human tissues with values ranging from 173 ± 6 bp (cortical neurons) to 207 ± 8 bp (cortical glial cells) (34). Thus, nucleosome-positioning patterns can differ significantly for the same DNA sequence, and the prediction of the experimentally observed nucleosome occupancies will remain incomplete as long as the contribution of chromatin remodeling complexes is not accounted for. From the above findings, the following picture emerges: In the absence of remodelers, the equilibrium nucleosome positions on the DNA are governed by their affinities to different DNA sequences (8,35). The remodelers may be viewed as molecular machines that transform this equilibrium into a different steady state upon ATP hydrolysis (16,36). Thus, the coupling of specific chromatin remodeling activities with intrinsic histone binding preferences for certain DNA sequences determines the nucleosome locations in living cells.

With respect to the available experimental data on nucleosomes positions, two types of systems can be distinguished. One is a single nucleosome that is positioned *in vitro* on a DNA fragment of known sequence in the absence or presence of a certain chromatin-remodeling complex. Typically, nucleosomes are reconstituted on a linear (16,37) or circular (36) DNA fragment that comprises several hundreds of base pairs. Nucleosome assembly is conducted through salt-dialysis in a multistage process determined by the initial recruitment of histones H3·H4 to the DNA followed by the addition of H2A·H2B as reviewed previously (38). Upon hydrolysis of ATP the remodeler repositions the nucleosome from the positions obtained by salt-dialysis. From this well-defined system, insights into the mechanism of nucleosome translocation and the activity differences of the various remodeling complexes are obtained. In the second case the *in vivo* nucleosome occupancies are determined and analyzed in a genome-wide manner (15,27,28,30,33,39,40). These studies show that in many instances the repositioning of just one or two nucleosomes provides a critical step in activating gene expression from a certain promoter (2,30). Nevertheless, for the theoretical description a multiple-nucleosome model has to be applied since there are long-range interactions between the nucleosome positions through combinatorial rearrangements. Moving a nucleosome changes the space allowed for other nucleosomes, even if they are not the nearest neighbors, through changes in the boundary conditions.

Here, we have developed a statistical thermodynamics framework to calculate nucleosome positions by taking into account both the equilibrium nucleosome arrangement according to the DNA sequence as well as an ATP-dependent repositioning. The experimental results for the cases of a single nucleosome and a multiple-nucleosome distribution in a DNA domain of several kilobases are rationalized in terms of different types of chromatin remodeling activities. These are represented in general terms by a function for the probability that a nucleosome at a given position is translocated. Furthermore, it is shown how regular phasing around a single strong positioning signal is established. For more complex patterns of global nucleosome rearrangements in a genome-wide context, our calculations suggest that translocations of nucleosomes occur with a preferred directionality that reflects a gradient in the nucleosome energy landscape.

METHODS

The model

A remodeler can remove the nucleosome from a given position on the DNA with a certain probability. This can be realized either by sliding it to the right/left, or by completely evicting the nucleosome from the DNA (Figure 1A). The DNA sequence is represented by a one-dimensional lattice with the base pair as elementary unit. Every base pair can be considered as the start of a nucleosome binding site that covers $m = 147$ bp (Figure 1B). Once the sequence-dependent affinities of the histone octamer $K(n)$ are known, the equilibrium nucleosome binding map is determined by the partition function that is a sum of statistical weights of all possible arrangements of nucleosomes on the DNA. This distribution can be derived with the help of the general transfer matrix formalism developed previously for the calculation of DNA–protein interactions (41) and protein-membrane binding (42). Multiple nucleosome types and additional proteins can be considered as well using this formalism.

To include the activity of chromatin remodeling complexes in the theoretical description, a nucleosome at a position n is considered. If associated with a remodeler, it may be moved along the DNA by a remodeler-specific step of s base pairs (Figure 1C). For example, the remodeling complex NURF or ISW2 repositions nucleosomes in apparent increments of ~10 bp, while for SWI/SNF a step length of around 50 bp has been reported (43,44). The probability function $P_{\mathrm{m}}(n)$ that the remodeler moves the nucleosome at position n allows it to account for specific types of remodeling activities (or their combination) in a general manner. For example $P_{\mathrm{m}}(n)$ may be defined to depend on the affinity of the histone octamer and the remodeler to a given DNA sequence. At each position n, a nucleosome can slide in two directions, which is accounted for by a probability $P_{-\mathrm{s}}$ to move the nucleosome to the site $(n-s)$ and the probability $P_{+\mathrm{s}} = 1 - P_{-\mathrm{s}}$ to move it to the site $(n + s)$. In the case of a single nucleosome on a DNA segment one can assume that $P_{+\mathrm{s}} = P_{-\mathrm{s}} = 0.5$ (in the absence of additional signals that lead to a

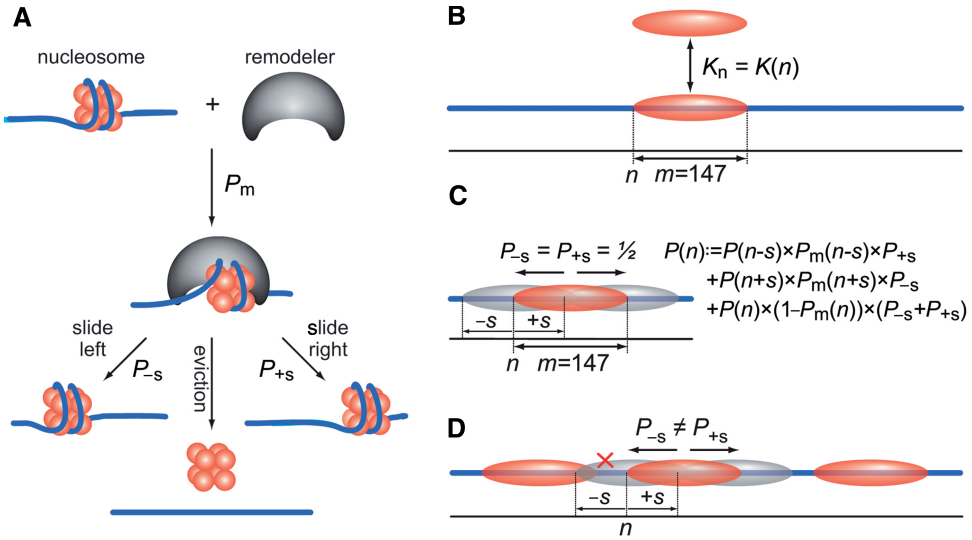

Figure 1. A schematic view of the remodeler action and the corresponding theoretical models. (**A**) Once a remodeler encounters a nucleosome, it removes it with a probability P_m, which may depend on the nucleosome and remodeler type and the DNA sequence. Nucleosome remodeling may be realized either as a sliding left/right without dissociation, or as a complete nucleosome eviction. (**B**) Nucleosome positioning may be viewed as an equilibrium thermodynamic process allowing nucleosome dissociation and rebinding to a one-dimensional DNA lattice of elementary units (base pairs) numbered by the index n. Nucleosome binding is characterized by the effective affinity $K_n = K(n)$ and the nucleosome length $m = 147$ bp. (**C**) A single-nucleosome translocation model considers one nucleosome on a short DNA segment as studied in *in vitro* experiments with reconstituted nucleosomes. Nucleosome dissociation and sliding off from the DNA ends is prohibited. The model has four parameters: the remodeling probability P_m, the elementary remodeler step s and the probabilities to move the nucleosome to the left (P_{-s}) and to the right (P_{+s}). (**D**) A multi-nucleosome translocation model describes nucleosome repositioning *in vivo*. It is based on the model (C), but the probabilities to move the nucleosome to the left and to the right now depend on the occupancy of the target site by other nucleosomes.

directionality of the translocation). In the case of multiple nucleosomes, P_{+s} and P_{-s} need to be recalculated dynamically during repositioning, since they depend on the absence or presence of another nucleosome at the target site (Figure 1D).

The dependence of P_m and s on the DNA sequence describes different 'remodeling rules' that represent features of different classes of remodelers. It is noted that, mathematically, the situation of a single remodeler moving a nucleosome continuously step by step is equivalent to the situation when several remodelers attach/detach a nucleosome after each elementary translocation step s. While s is an experimentally observable parameter, it is more difficult to define P_m, which may have a complex dependence on the DNA sequence as given by n. This may include a direct DNA sequence-dependence of the remodeling reaction or an indirect dependence through the intrinsic histone-DNA affinities $K(n)$. The equilibrium constant $K(n)$ for the histone octamer binding at position n is a function of the DNA sequence and can be derived from the analysis of recent experiments (2,8,9,25,45). In addition, one can try to define $P_m(n)$ on the basis of explicit kinetic or thermodynamic considerations as done previously (16). In the latter study, the remodeler activity was described from the point of view of enzyme kinetics. Nucleosome positioning to a certain site was explained either by lowering the remodeler binding affinity to the nucleosomes at the target DNA sequence (the 'release' model) or by a reduced translocation rate away from this site (the 'arrest' model). Here, the remodeler binding/unbinding reactions are not treated explicitly, and all mechanistic differences are cast into $P_m(n)$, $P_{+s}(n)$ and $P_{-s}(n)$.

Calculating equilibrium nucleosome distributions

The equilibrium nucleosome distributions are calculated with the transfer matrix formalism, which has been previously introduced as a systematic way to calculate DNA–protein–drug binding in gene regulation (41) and protein–membrane binding in signal transduction (42). In this method, the DNA is represented as a one-dimensional lattice of units (base pairs) numbered from left to right as $1 \ldots N$. All microstates allowed for each individual unit are enumerated, and transfer matrices are constructed that contain the conditional probabilities $Q_n(i, j)$ of having unit n in state i and unit $n + 1$ in state j. This approach can be extended as needed to include not only the binding of the histone octamer but also linker histones and other chromosomal proteins or transcription factors. In addition, different thermodynamic features could be assigned to a certain type of nucleosome that for example contains the H2A.Z histone variant instead of the canonical H2A histone. In a system with f types of interacting DNA binders, each DNA unit may be in R states shown in the Supplementary Table S1, where R is given by Equation (1) (41):

$$R = \sum_{g=1}^{f} (m_g + V_g) + \max(V_g) + 3, \qquad \mathbf{1}$$

Here, m_g is the length of a protein of type g, and $\max(V_g)$ is the maximum protein–protein interaction distance

along the DNA. A histone octamer may be considered as a single protein covering $m = 147$ DNA units (Figure 1B). In the current study, all nucleosomes are of the same type (one DNA binder type, $f = 1$), and do not interact with each other ($V = 0$). This model is based on $m + 3$ states, according to which the histone octamer complex may be bound or not to a given DNA sequence (Supplementary Table S2). Each elementary DNA unit is either free or inside a nucleosome. Three different free DNA states are distinguished, correspondingly before, after and between the nucleosomes. A DNA unit inside a nucleosome may be in one of m states depending on its location relative to the nucleosome start. This model is equivalent to the nucleosome model based on the dynamic programming approach of bioinformatics (8), also known in biophysics as the recurrent relations approach (46). Only one type of DNA ligand is present, but many states of the complex exist, i.e. a histone octamer can be present at any of the $N-m+1$ binding sites of the DNA lattice.

The nucleosome translocation model

In the *in vitro* experiments mentioned above, typically only one histone octamer is present that can slide along the DNA in the absence of any competing proteins. The internucleosomal DNA state thus cannot be realized, and we need to consider only $m + 2$ states. Furthermore, only certain combinations of states of consecutive units listed in Supplementary Table S2 can be present, which is reflected by a very sparse transfer matrix (Supplementary Table S3). Prohibited combinations of states are assigned a zero statistical weight in the transfer matrix. For this model, the following nonzero matrix elements exist: A free DNA unit may be followed by another free unit with statistical weight 1. In particular, a 'left free end' unit may be followed by another 'left free end' unit, and a 'right free end' unit may be followed by another 'right free end' unit. However, the 'left free end' cannot be followed by the 'right free end', which assures that a nucleosome is bound somewhere between. A bound DNA unit may be followed by another bound unit or by the 'right free end'. The matrix element $Q_n(1, 2)$, which corresponds to the first bound unit, is characterized by the statistical weight equal to the sequence-specific binding constant, $K(n)$:

$$Q_{\mathrm{n}}(1,2) = K(n) = \exp(-\Delta G_{\mathrm{n}}/RT), \qquad \mathbf{2}$$

where ΔG_{n} is the energy of the histone octamer-DNA interaction for a binding site starting at nucleotide n, R the gas constant and T the temperature. ΔG_{n} could be taken as a difference between the energies at nonspecific and specific nucleosome-binding sites.

Dissociation of nucleosomes

In extension of the above model histone octamer particles may dissociate from the DNA and rebind at thermodynamic equilibrium. The number of nucleosomes on the DNA is not fixed as in the nucleosome translocation model. Here, the total number of proteins in the system (bound to the DNA and in solution) is fixed. Considering the limit of an excess of free proteins over free binding sites, the bulk concentrations may be considered as constant. Thus, the nucleosome-dissociation model requires an additional parameter c_0 for the effective histone octamer concentration, and Equation (2) is rewritten as follows:

$$Q_{\mathrm{n}}(1,2) = K(n) \times c_0, \qquad \mathbf{2a}$$

The transfer matrix for the nucleosome-dissociation model is shown in Table S4. This matrix is based on $m + 3$ states listed in Table S2. Five matrix elements distinguish this model from the single-nucleosome translocation model: $Q_{\mathrm{n}}(1, 2)$ is given by Equation (2a) as discussed above. $Q_{\mathrm{n}}(m, 1) = 1$ allows a bound nucleosome to be followed by another bound nucleosome; $Q_{\mathrm{n}}(m + 1, m + 2) = 1$ allows a situation when there are no nucleosomes on the DNA; $Q_{\mathrm{n}}(m + 3, 1) = 1$, $Q_{\mathrm{n}}(m + 3, m + 3) = 1$ and $Q_{\mathrm{n}}(m, m + 3) = 1$ allow free DNA gaps between the nucleosomes. The transfer matrices described above correspond to DNA units far from the boundaries. The boundary conditions imply additional constraints on the transfer matrices within m units close to the DNA ends. These do not allow incomplete histone octamer binding at the DNA ends.

Calculation of binding maps

For both the nucleosome-translocation and nucleosome-dissociation models, the partition function Z is given by the multiplication of the corresponding transfer matrices Q_{n} according to the DNA sequence (47):

$$Z = \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix} \times \prod_{n=1}^{N} Q_n \times \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \qquad \mathbf{3}$$

It is convenient to calculate the partition function Z and its derivatives recursively according to Equations (4) and (5) (48).

$$Z = A_N \times \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}, \quad A_i = A_{i-1} \times Q_n, \quad A_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix} \qquad \mathbf{4}$$

$$\frac{\partial Z}{\partial X} = \frac{\partial A_N}{\partial X} \times \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}, \quad \frac{\partial A_n}{\partial X} = \frac{\partial A_n - 1}{\partial X} \times Q_n + A_{n-1} \times \frac{\partial Q_n}{\partial X} \qquad \mathbf{5}$$

The derivatives $\partial Z/\partial K_{\mathrm{n}}$ yield the binding map, i.e. the probabilities $P(0, n)$ to start a nucleosome at base pair n in the absence of remodelers (41). The probability that a given DNA unit is inside a nucleosome in the absence of remodelers is given by $C(0, n)$:

$$C(0,n) = \sum_{n-m+1}^{n} P(0,k) \qquad \mathbf{6}$$

Repositioning of a single nucleosome

The equilibrium distribution of nucleosomes $P(0,n)$ is calculated in the absence of remodelers by the transfer matrix formalism as described above. The nucleosome map in the presence of a remodeler is calculated iteratively and is denoted as $P(j,n)$, where j is the iteration number. For each iteration, the nucleosome start site probabilities are changed, proceeding from the first to the last DNA nucleotide according to the following procedure:

$$\begin{aligned}
&(i)\ n \le s: \\
&\quad P(j,n) = P(j-1,n+s)\times P_{\mathrm{m}}(n+s)\times P_{-\mathrm{s}}(n+s) \\
&\qquad + P(j-1,n)\times(1-P_{\mathrm{m}}(n))\times P_{+\mathrm{s}}(n) \\
&(ii)\ s < n \le N-m-s+1: \\
&\quad P(j,n) = P(j-1,n-s)\times P_{\mathrm{m}}(n-s)\times P_{+\mathrm{s}}(n-s) \\
&\qquad + P(j-1,n+s)\times P_{\mathrm{m}}(n+s) \\
&\qquad \times P_{-s}(n+s) + P(j-1,n)\times(1-P_{\mathrm{m}}(n)) \\
&\qquad \times (P_{-\mathrm{s}}(n) + P+s(n)) \\
&(iii)\ n > N-m-s+1: \\
&\quad P(j,n) = P(j-1,n-s)\times P_{\mathrm{m}}(n-s)\times P_{+\mathrm{s}}(n-s) \\
&\qquad + P(j-1,n)\times(1-P_{\mathrm{m}}(n))\times P_{-\mathrm{s}}(n)
\end{aligned} \quad \mathbf{7}$$

$P_{+\mathrm{s}}(n) = P_{-\mathrm{s}}(n) = 0.5$ for a single-nucleosome sliding model. Due to the fixed boundary conditions in Equation (7) a new steady state is obtained after a sufficient number of iterations that represents the final map of nucleosome positions in the presence of a remodeler.

Repositioning in a multi-nucleosome context

Similarly to the single-nucleosome case, we start from the equilibrium-binding map and iteratively recalculate the distribution according to Equation (7). Due to the presence of multiple nucleosomes the parameters $P_{-\mathrm{s}}(n)$ and $P_{+\mathrm{s}}(n)$ have to be adjusted at each step because the boundary conditions for each nucleosome change dynamically with repositioning of the neighboring nucleosomes. The probabilities that a potential repositioning site is already occupied, $C(j,n)$, are given by Equation (6). Then the probability to move a nucleosome by one unit to the right is $P_{+1}(n) = 0.5\times(1-C(j,n+m))$. Correspondingly, the probabilities to move a nucleosome by s units, $P_{+\mathrm{s}}(n)$ and $P_{-\mathrm{s}}(n)$ are given by the following expressions:

$$P_{+s}(n) = 0.5\times\prod_{k=n+m}^{n+m+s}(1-C(j,k)) \quad \mathbf{8}$$

$$P_{-s}(n) = 0.5\times\prod_{k=n-1}^{n+s-1}(1-C(j,k)) \quad \mathbf{9}$$

For the case when a remodeler complex senses the intrinsic histone-DNA energy landscape, the following expressions were used:

$$P_{+s}(n) = 0.5\times\frac{K(n+s)}{K(n)}\prod_{k=n+m}^{n+m+s}(1-C(j,k)) \quad \mathbf{10}$$

$$P_{-s}(n) = 0.5\times\frac{K(n-s)}{K(n)}\prod_{k=n-1}^{n+s-1}(1-C(j,k)) \quad \mathbf{11}$$

The iterative repositioning was performed similarly to the single-nucleosome model, with the parameters $P_{+\mathrm{s}}(n)$ and $P_{-\mathrm{s}}(n)$ being recalculated at each step before applying Equation (7).

Boundary conditions

Three possibilities exist to account for the boundary conditions in the calculations. Open boundary conditions would allow nucleosomes to slide off the ends of the DNA. Unless this is compensated by new incoming nucleosomes, this would lead to the depletion of nucleosomes after a sufficiently large number of remodeling iterations, i.e. a steady state is not reached. Circular boundary conditions circumvent this problem by equating the flux of nucleosomes sliding off one DNA end and entering the other end. They would be appropriate to describe *in vitro* experiments with circular DNA substrates (49), but cannot be applied to linear DNA segments or to the analysis of a region within a chromosome. Accordingly, fixed boundary conditions (nucleosomes cannot leave the ends of the linear DNA) are used here. This can potentially lead to artificial ordering of several nucleosomes close to the boundary as concluded previously for equilibrium ligand binding models (50–52). However, by including sufficiently long DNA regions that flank the genomic locus of interest any potential disturbance due to fixed boundary conditions can be avoided.

Transformations between start site probabilities and occupancy maps

The probability that the DNA unit n is covered by a nucleosome is referred to as $C(n)$ and the probability that a nucleosome starts at a DNA unit n as $P(n)$. The values of $C(n)$ and $P(n)$ are connected by the following relations:

$$n \le m,\ C(n) = \sum_{1}^{n} P(k) \quad \mathbf{12}$$

$$n < m,\ P(n) = C(n) - \sum_{1}^{n-1} P(k), P(1) = C(1) \quad \mathbf{13}$$

$$m < n < L,\ C(n) = \sum_{n-m+1}^{n} P(k) \quad \mathbf{14}$$

$$m < n \le L-m+1, P(n) = C(n) - \sum_{n-m+1}^{n-1} P(k) \quad \mathbf{15}$$

The binding maps were recalculated recursively according to Equations (12–15) as explained in Supplementary Data.

RESULTS AND DISCUSSION

Comparing predicted and experimentally determined nucleosome occupancies

There are two types of experiments available for nucleosome positioning. In the nucleosome reconstitution experiments *in vitro*, one can map exact positions of the nucleosomes, while in the whole-genome experiments the output is usually an averaged nucleosome occupancy score. These two types of data need to be treated differently to extract nucleosome affinities and remodeler rules. Figure 2 shows single-nucleosome reconstitution experiments at a DNA fragment of 359 bp in length corresponding to the *Drosophila hsp70* promoter (16,37). Previous footprinting allowed establishing exact positions of five preferred nucleosome start sites on this fragment (Figure 2A) (16,37). A quantitative electrophoretic analysis yields relative occupancies of these positions, which, after mirroring and scaling for the length of the DNA fragment, give the nucleosome start site probabilities (black line in Figure 2B). The probability that a given DNA base pair is occupied by the nucleosome (red line in Figure 2B) is given by the sum of the probabilities to start the nucleosome at the neighboring sites as detailed in the 'Methods' section and Supplementary Data. The blue line shows the probabilities to start a nucleosome predicted by the DNA sequence according to a previously published algorithm (25,45). Due to the boundary conditions, the method fails at the distal DNA sites, but in the middle of the DNA it is still valid. The positions of the N1 and N3 peaks exactly coincide with small peaks of the predicted start site positions. However, the relative occupancies of the peaks expected from the DNA sequence do not correlate with the experimentally observed ones. This picture may change significantly *in vivo* due to different boundary conditions (stronger positioning sites nearby) and the remodeler action. In fact, the region of the *hsp70* promoter shown in Figure 2 is depleted of nucleosomes *in vivo* (53).

Figure 3A depicts the *in vivo* nucleosome occupancies in a 2-kb enhancer region of chromosome 5 involved in the activation of the human $CD4^+$ T cells as derived from MNase digestion and subsequent Solexa sequencing (33). It is representative for the reported changes in the nucleosome position pattern that occur all through the genome upon T-cell activation (33). The locus contains a conserved noncoding sequence termed CNS1 that is required for the maturation of naive $CD4^+$ T cells upon induction (54). Its precise function is currently not known but it is apparent that it contains several nucleosomes that are subject to chromatin remodeling activity upon activation. The distributions of nucleosome occupancy probabilities

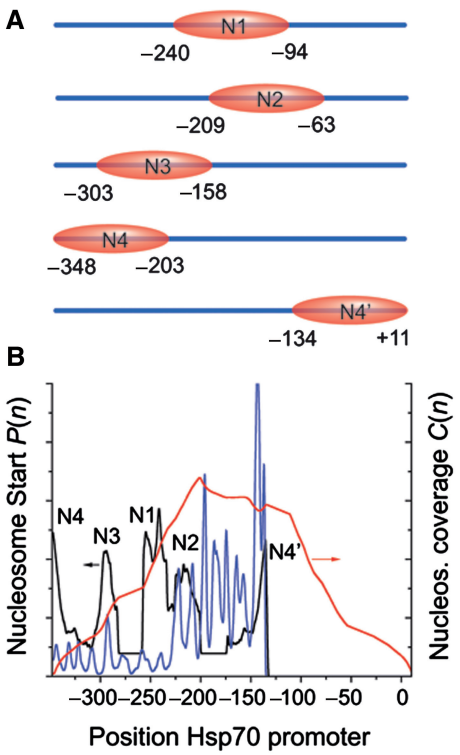

Figure 2. Nucleosome positioning *in vitro* at the *Drosophila hsp70* promoter region (359 bp). (**A**) Nucleosome positions determined by footprinting (37). (**B**) Relative intensities of nucleosome positioning peaks determined from our scan of the electrophoresis experiments (16). The black line shows nucleosome start site probabilities, and the red line is a calculated probability that a given DNA base pair is occupied by a nucleosome. The blue line shows a predicted start site probability assuming that the positioning is determined solely by the DNA sequence according to the algorithm by Segal, Widom and co-workers (25,45).

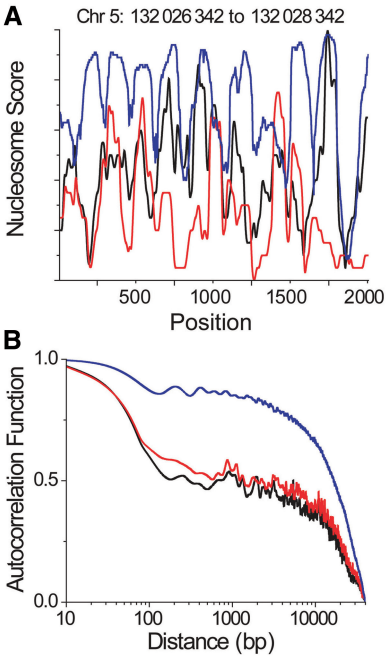

Figure 3. Experimental nucleosome occupancy scores in the resting (black lines) and activated (red lines) human $CD4P^{+P}$ T cells (33) compared with the predictions for nucleosome positioning based on the DNA sequence (blue lines) according to the algorithm by Segal, Widom and coworkers (25,45). (**A**) Nucleosome occupancies for a 2-kb enhancer region, which is involved in the IL-4 and IL-13 gene regulation during T-cell activation (Chr5: 132 026 342 to 132 1028 342). (**B**) Autocorrelation functions calculated for a 40-kb interval including the enhancer region above.

of the resting (black) and activated (red) $CD4^+$ T cells reported in ref. 33 are superimposed with the nucleosome occupancies predicted from the DNA sequence (blue). The two experimental distributions show clear differences with some nucleosome-binding sites displaying changes in occupancies, while others appear to be shifted to another position. Apart from some exceptions the experimentally found positions mostly coincide with those predicted from the DNA sequence, but the relative occupancies of these sites do not correlate. The deviations of experimental nucleosome occupancies from sequence-predicted maps (Figure 3A), are not restricted to the region shown in Figure 3A, but represent a general feature of the nucleosome positions from this data set. This is inferred from an autocorrelation analysis (Figure 3B), similar to that used previously to evaluate nucleosome positions (8,13). The comparison of the three nucleosome distributions for a 40-kb genomic region including the region shown in Figure 3A reveals that the theoretical nucleosome map has much more regular nucleosome spacing than the experimental distributions. The autocorrelation curves for the resting and activated cell state were very similar on a short scale (<80 bp). Other genomic regions of a size around 40 kb showed a similar behavior. Significantly larger regions could not be computed due to the presence of gaps in the experimental data set.

Repositioning of a single nucleosome

Calculations were performed for a DNA fragment of 359 bp in length corresponding to the *Drosophila hsp70* DNA fragment used in previous *in vitro* experiments (16,37). Based on the experimental data shown in Figure 2, five preferred nucleosome positions were identified and the relative histone octamer–DNA-binding affinities corresponding to these positions were assigned. For the calculations the nonspecific equilibrium binding constant was normalized to $K = 1\,M^{-1}$. The energy difference for site-specific sites was around $2\,kT$ ($K \approx 7\,M^{-1}$). This is consistent with experimental estimates of $\Delta\Delta G \approx 3$ kcal/mol for nucleosome positioning by 5 S rDNA sequence as compared to the random sequence (55). As long as only the translocation of nucleosomes but not their dissociation is considered the absolute values of K are not relevant. Accordingly, the distribution depends only on the ratio between equilibrium binding constants for different DNA sites.

In the calculations for the positioning of a single nucleosome, two cases were distinguished. In one series of calculations, it was assumed that the remodeler probability $P_m(n)$ to move the nucleosome from site n by a step of s base pairs is not affected by the nucleosome position, and that the outcome of the translocation reaction is determined only by the nucleosome–DNA interactions. Alternatively, it was investigated how the nucleosome position distribution changes if remodelers display a site-dependent activity in addition to intrinsic histone–DNA affinity differences. Results of these calculations are presented in Figure 4, where the red lines depict equilibrium distributions of nucleosomes in the absence of remodelers, and the blue lines reflect the activity of different remodeler types.

Remodeler-spacer. A remodeler without any DNA sequence specificity repositions a nucleosome in steps of s base pairs from any initial position with equal probability ($P_m = 1$). The boundary conditions prevent a nucleosome from leaving the DNA at the ends of the fragment. The calculations in Figure 4A demonstrate that the action of such a remodeler is to break the initial sequence-specific nucleosome distribution, and to redistribute the nucleosomes to a state with equal spacing, which is a multiple of the remodeler step s. Yet, the strongest initial nucleosome binding sites remain enriched. This effect results from the fixed translocation step size and a finite DNA segment length. The spacing pattern would become less well defined if the step size was a distribution centered on some average value, e.g. $s = 50 \pm x$, or if the DNA length $N >> s$. This remodeler type allows coping with the periodicity constraints required for the higher-order chromatin packing, as well as with the fact that most of the accessible nucleosome positions appear to be predefined by the DNA sequence. The latter is inferred from our previous observation that *in vivo* remodelers change the relative occupancy of the intrinsically favored nucleosome binding sites rather than translocating a nucleosome to a 'new' site (Figure 3). To verify this mechanism *in vitro*, one could examine DNAs in which different known nucleosome-positioning sequences are combined. In the absence of a remodeler, the nucleosome will be positioned according to its intrinsic binding affinity. Upon addition of a remodeler of the 'spacer' type defined above, one would obtain nucleosome-binding maps similar to those shown in Figure 4A, with regular nucleosome spacing and some remaining preference for high-affinity binding sites. Fitting of the experimental distributions with this model would allow extracting the elementary remodeler step length s, a parameter otherwise not easily accessible.

Remodeler-amplifier. To translocate nucleosomes, remodelers need to break histone–DNA contacts, e.g. form a DNA loop and propagate it along the nucleosome to reposition the nucleosome in steps that are determined by the loop size (19). Since this involves the unwrapping of DNA from the histone octamer core, the efficiency of the remodeler reaction could be dependent on the DNA affinity of the histone octamer, i.e. P_m is a function of $K(n)$. At first, a simple relation $P_m = 1/K(n)$ was investigated. A remodeler that operates in this manner would accelerate nucleosome movements without affecting the sequence preferences. Indeed, the corresponding binding map coincides with the initial thermal equilibrium distribution of nucleosomes on the DNA (red line in Figure 4B). If the remodeler binding to nucleosome-positioning DNA sequences is facilitated as well (for example because nucleosome positioning sequences are more bendable), the repositioning will display a pronounced DNA sequence dependence, e.g. $P_m = 1/K^2$. Figure 4B shows that for this case the effect of differences in the intrinsic nucleosome–DNA affinities was amplified. The probability to find a nucleosome at a high-affinity DNA site was

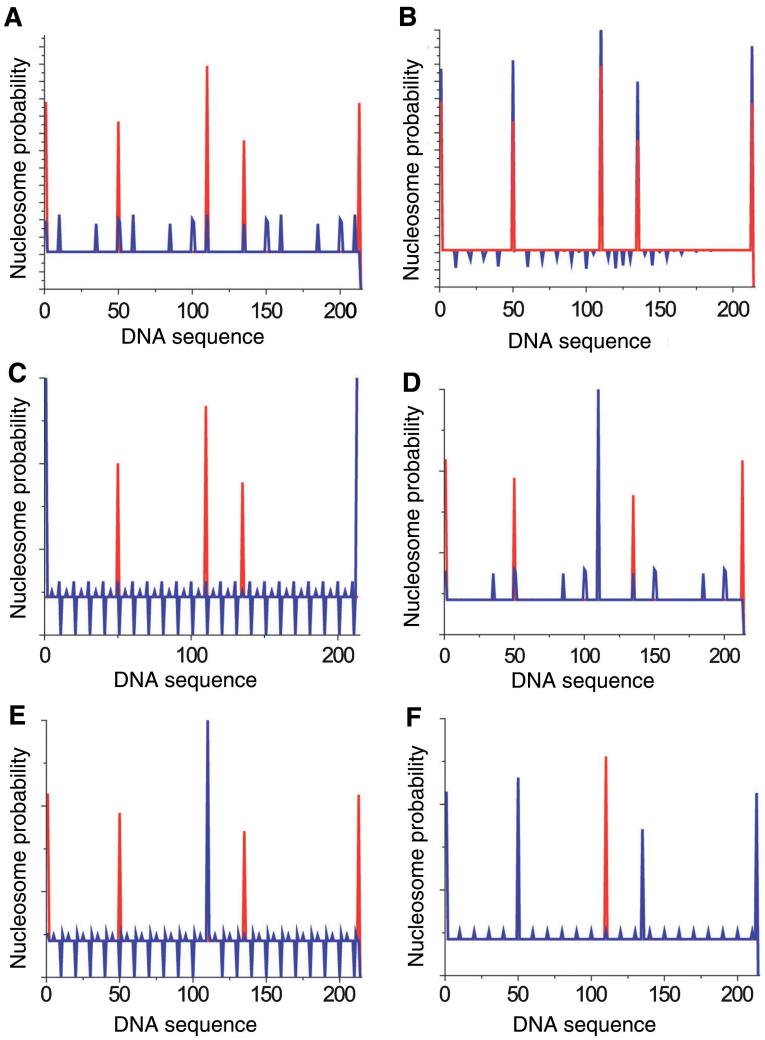

Figure 4. Probabilities that a nucleosome starts at the DNA site n, calculated in the frame of the single-nucleosome translocation model in the absence of remodelers (red lines) and in the presence of remodelers (blue lines), 10 000 iterations. (**A**) Repositioning probabilities are equal for all sites: $P_m = 1$, $s = 50$. (**B**) Repositioning probabilities correlate with thermal nucleosome positioning probabilities: $P_m = 1/K^2$, $s = 10$. (**C**) The initial map is calculated assuming that there is no specific binding at the distal DNA sites. The remodeler can move nucleosomes from any site but the DNA ends [$P_m(1 < n < N) = 1$, $P_m(n = 1, N) = 1$, $s = 10$]. (**D** and **E**) Remodeler-positioner does not translocate a nucleosome from a position $n = 110$ [$P_m(n = 110) = 0$, $P_m(n \neq 110) = 1$, $s = 50$] (D) and $s = 10$ (E). (**F**) Remodeler removes a nucleosome from site $n = 110$ [$P_m(n = 110) = 1$, $P_m(n \neq 110) = 1/K$, $s = 10$].

increased, while nonspecific sites were depleted. ISWI and NURF complexes are examples of remodelers that redistribute nucleosomes away from their initial thermal equilibrium, but only within the other positions that are closely related to those selected during thermal equilibration (16,18). Accordingly, for this remodeler class, one could attempt to measure the $P_m(K)$ dependence.

Directionality and end-effects. The nucleosome is a symmetric object, and as such, it cannot account for any directional movement unless directionality is DNA sequence-directed and/or appropriately localized modification of the histone octamer core (like post-translational histone modification or incorporation of histone variants) are present. In *in vitro* experiments a directionality of the translocations was observed in the case of SWI/SNF, ISWI and Rad54 remodelers tracking in the 3′-5′ direction along one strand of the DNA duplex (17). In our calculations, such directionality can be accounted for by different probabilities of moving a nucleosome from the site n toward the left and right DNA ends. This leads to a saturation of sites close to one of the DNA ends, while the sites closer to the other end are depleted. In addition, the probability to find a nucleosome at nonspecific sites closer to the favored DNA end is increased (data not shown). *In vivo*, some remodelers appear to act directionally, e.g. by moving the nucleosomes from/to regulatory sequences. For example, Isw2 functions adjacent to promoter regions where it directionally repositions nucleosomes at the interface between genic and intergenic sequences. This results in increased nucleosome occupancy of the intergenic region (2). However, this type of directionality may also be described by remodeler sequence-specificity (see below) rather than by asymmetric tracking along

the DNA. On the other hand, as demonstrated in Figure 2, some of the nucleosome sites at a linear DNA fragment *in vitro* do not correspond to the inherent nucleosome–DNA affinities, but are specific to the experimental setup. For example, all nucleosome-reconstitution experiments report high nucleosome populations at the DNA ends independently of the DNA sequence (37). This effect could be explained in part by entropic effects due to the restrictions at the polymer ends as suggested by Brownian dynamics simulation of nucleosome sliding in the absence of a remodeler (56). By setting $P_m = 1$ at both DNA ends, we can ensure that the remodeler removes the distal nucleosomes. As a result, other binding sites become more populated. Such a remodeling activity was reported for ISW1a, Chd1 and ACF, which showed a strong preference for mobilizing nucleosomes from the DNA ends to more central sites (16,57,58). Another extreme would be to set $P_m = 0$ at both DNA ends, which leads to trapping the nucleosomes as they reach this position. Figure 4C shows the results of such calculations, where the initial nucleosome map did not contain the nucleosomes at the ends, and after remodeling, only the distal nucleosome positions remain. A distal nucleosome lacking linker DNA on one site appears to be indeed a bad substrate for some remodeling complexes as observed for Isw1b, which moves nucleosomes to the end (58).

It is noted that for the calculations described in the context of Figure 4, a model was used where nucleosomes cannot slide off the DNA end. However, *in vivo* nucleosomes could be translocated over larger distances or nucleosomes could be evicted. If sliding off the DNA ends is allowed, the total nucleosome binding probability is not conserved and the map of binding depends on the number of iterations. After a large number of iterations only the strongest binding sites remain, and other binding sites are completely depleted (data not shown). This mechanism would allow for a discrimination of the strongest binding site out of a number of competing weaker sites.

Remodeler-remover and remodeler-positioner. Sequence-specific remodelers can rewrite the nucleosome position pattern and enrich or deplete nucleosomes at the binding sites that are defined by the intrinsic DNA sequence preference of the histone octamer. Although chromatin remodeling complexes can translocate nucleosomes also to sites with an intrinsically low affinity, this seems to be the case only for certain locations as discussed above (16). The DNA sequence-dependent translocation has been analyzed in Figure 4D–F. First, calculations were conducted for a large-step remodeler ($s = 50$), which repositions a nucleosome from all sites except for $n = 110$ [$P_m(n \neq 110) = 1$, $P_m(n = 110) = 0$]. After 10 000 iterations, the corresponding binding map in Figure 4D becomes similar to that in Figure 4A. By removing nucleosomes from all preferred positions except for $n = 110$ an enrichment at this site is observed. Thus, the remodeler acts also as an 'amplifier' at these positions (albeit according to a different mechanism than in Figure 2B), while the basic features of the 'spacer' activity are retained. Lowering the step size to $s = 10$ makes it a less effective spacer and a more effective amplifier (Figure 4E). The model with the parameters $P_m(n \neq 110) = 1/K$ and $P_m(n = 110) = 0$ yielded similar distributions (data not shown). In Figure 4F the calculated binding map for a remodeler is depicted, which has a higher repositioning probability for the site $n = 110$ [$P_m(n \neq 110) = 1/K$, $P_m(n = 110) = 1$]. As expected, such a remodeler removes a nucleosome from this single high-affinity site to any of the weaker binding sites. This remodeler type is referred to here as a 'remover'. It would act locally by redistributing one or two nucleosomes to other sites. This behavior has been observed in nucleosome repositioning experiments where the human SWI/SNF removed two nucleosomes from their high-affinity sites at the C-myc promoter to adjacent low-affinity sites (36) as well as for the hsp70 and rDNA promoter fragments (16).

Equilibrium nucleosome distributions *in vivo*

Recently, whole-genome nucleosome occupancies have been reported for several organisms (15,27,28,30, 33,39,40). Figure 5A shows data from Schones *et al.* (33) for nucleosome occupancies in human $CD4^+$ T cells. Exemplary distributions of nucleosome occupancy probabilities in a segment of 2 kb from the chromosome 5 of the resting and activated $CD4^+$ T cells are depicted. As discussed above, this type of data reflect DNA occupancies $C(n)$, not the nucleosome start site probabilities $P(n)$. The nucleosome start site probabilities can be calculated from the DNA occupancy probabilities with a minimal loss of information using the recursive procedure described under 'Methods' section. Figure 5B shows the start site probabilities calculated for the chromosomal region shown above. One can see, that the activated state distribution (red lines) becomes more regular due to the removal of several small peaks from the resting state distribution (black lines). In analogy with a standard biochemical equilibrium, the start site probabilities $P(n)$ can be normalized to [0, 1], and the nucleosome affinities $K(n)$ can be defined as $K(n) = P(n)/(1 - P(n))$. Here it is assumed that the effective nucleosome concentration is equal to 1.

Figure 5C depicts the binding maps for the resting (black) and activated (red) states of the abovementioned region of chromosome 5 of human $CD4^+$ T cells using the binding constants extracted from Figure 5B. This data set can be related to the three basic remodeling activities of an 'amplifier', 'remover' or 'spacer' identified from the single-nucleosome analysis described above (Figure 4). A remodeler of the 'remover' type evicts the nucleosomes at positions 5 and 10 in Figure 5C. Nucleosomes 6, 8 and 9 are subject to the action of a remodeler-'amplifier'. In addition, the nucleosomes are nonspecifically spaced between the strong positioning sites to occupy all the space available after eviction of their neighbors. For a region shown in Figure 5 and a number of other genomic regions (data not shown), we were able to interpret all nucleosome changes with the help of combinations of these three remodeler activities. For example, a remodeler-'positioner' was not required, because all potential nucleosome positions (but not occupancies) seemed to be predefined by the DNA sequence. This suggests that experimentally obtained genomic data indeed can be

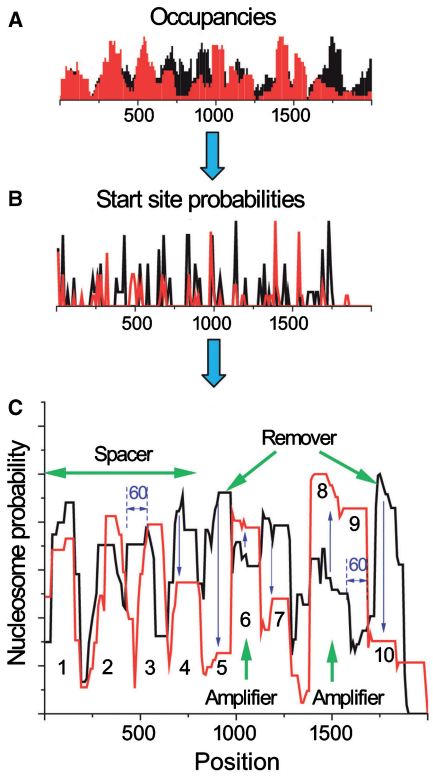

Figure 5. A quantitative analysis of *in vivo* data using the equilibrium-binding model. The experimental nucleosome occupancy scores for the resting (black) and activated (red) $CD4^+$ T cell reported by Schones *et al.* (33) (**A**) were transformed into the nucleosome start site probabilities (**B**). Then, relative binding affinities were assigned as $K(n) = P(n)/(1-P(n))$ and used to calculate equilibrium binding maps shown in the (**C**). Three main remodeler activities predicted from single-nucleosome modeling (Figure 4) are observed at this genomic region as indicated by the arrows. The nucleosomes 4 and 10 are removed, the occupancies of nucleosomes 6, 8 and 9 and the nucleosomes between these sites are amplified and shifted due to the change of their local boundary conditions as imposed by the flanking nucleosomes.

reproduced on the basis of a small set of simple remodeler rules within the theoretical framework developed here. *Vice versa*, the rules for the acting remodelers may be deduced from the comparison of the maps obtained for different cell states. The latter would require additional experimental information concerning the type(s) of remodelers participating in this particular event or the knock down of specific remodeler classes as done in yeast for ISW2 (9) and RSC (15). Equilibrium binding maps of the type shown in Figure 5 can be used to follow the fate of individual nucleosome occupancies in different cell states. By varying the effective nucleosome concentration, they could provide experimentally testable mechanisms for the implementation of 'nucleosome switches', in which large changes of the expression pattern would be mediated by relatively small changes in the chemical potential (i.e. the effective concentration) of histones (59).

Calculating the redistribution of nucleosome position patterns

The above type of equilibrium model does not allow it to dissect the dynamics of nucleosome redistributions. In order to identify how nucleosome maps change in response to different remodeler activities, we have performed calculations using the multiple-nucleosome translocation model depicted in Figure 1D.

Nonspecific remodelers establish regular nucleosome spacing. Figure 6 shows the results of the calculation for a 2-kb enhancer region of the human chromosome 5 in the context of the multiple-nucleosome translocation model. The initial nucleosome-binding map was computed from the intrinsic nucleosome-DNA occupancies according to refs. (25,45). In the calculations in Figure 6A, the nucleosomes were redistributed by a nonspecific remodeler ($P_m = 1$, $s = 10$). The probabilities to move the nucleosome to the left and to the right were calculated according to Equations (8) and (9). After 100 000 iterations, the initial binding map is transformed into a new steady state, in which the nucleosomes tend to form a tightly packed nucleosome array with equal spacing. The average nucleosome–nucleosome distance in Figure 6A is determined by the length of the histone octamer-binding site (147 bp) and the length of nucleosome–nucleosome interaction potential ($V = 0$ in Figure 6). The final nucleosome distribution does not depend on the initial nucleosome preferences and the remodeler step size s. This is true for small remodeler steps, $s << m$ (for both fixed and variable s) as it appears to be the case for nucleosome positioning in $CD4^+$ T cells (Figures 3–5). Increasing the remodeler step size results in a more homogenous distribution. If the remodeler step size is comparable to the length of the histone octamer-binding site the distribution becomes phased in intervals of the remodeler step s (data not shown).

Figure 6B shows calculations for a remodeler-amplifier ($P_m = 1/\exp(K(n))$, $s = 10$), which favors nucleosome removals from low-affinity sites and disfavors removals from high-affinity sites. These calculations demonstrate that the intrinsic DNA affinities of the histone octamer can lead to a regular nucleosome spacing in the presence of an inversely related remodeler activity. Although the map in Figure 6B is not as regular as that in Figure 6A, a well-defined pattern of groups of high-probability nucleosome start sites separated by roughly equal intervals is apparent. This ordering results from the remodeler activity, which keeps the intrinsically preferred nucleosome sites and introduces the boundary-determined spacing.

Regular nucleosome spacing has been reported for many systems *in vivo*. In particular, nucleosome phasing is required to maintain the higher-level packing of heterochromatin (60), centromeres (40) and specific regulatory regions (27,61). Regular spacing does not form spontaneously by *in vitro* reconstitution but requires the activity of chromatin remodeling factors uch as ACF (62). Accordingly, ACF and other remodelers have been proposed to act as nucleosome 'spacing engines' (63,64). *In vivo* a small number of precisely positioned nucleosomes may play a role as boundary constraints similar to the 'DNA ends' in our calculations. Weakly positioned nucleosomes would be evenly spaced between strongly positioned nucleosomes according to the mechanism revealed above. A recent study has found that the genomic

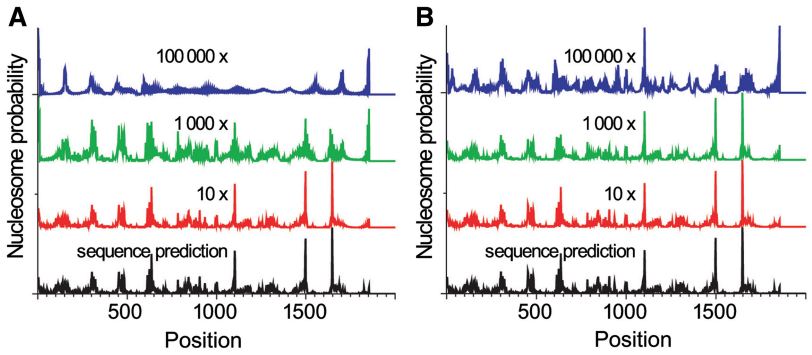

Figure 6. Modeling multi-nucleosome redistribution by a remodeler-spacer. (Bottom) Shows nucleosome start site probabilities predicted from the intrinsic affinities according to the algorithm by Segal, Widom and coworkers (25,45). (Top) Shows the nucleosome maps recalculated iteratively according to Equations (7–9) for the remodeler-spacer starting from the thermal equilibrium distribution. The number of iterations is indicated in the Figure. (**A**) The remodeling probability is equal for all sites ($P_m = 1$, $s = 10$). (**B**) The remodeling probability is negatively correlated with intrinsic histone-DNA affinities [$P_m(n) = 1/\exp(K(n))$, $s = 10$].

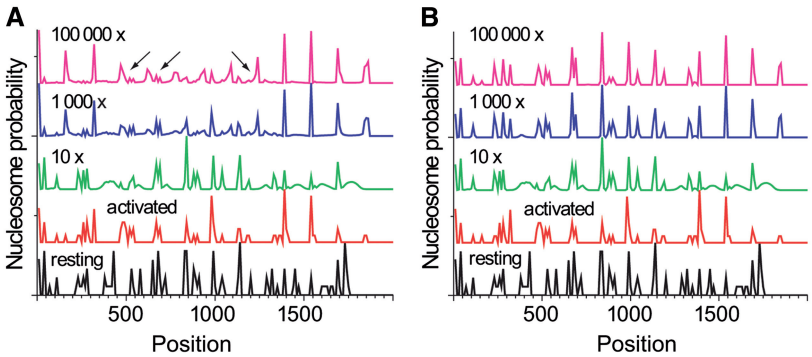

Figure 7. Calculating multi-nucleosome redistribution patterns by a sequence-specific remodeler. The two plots at the bottom show experimental nucleosome start site probabilities in the resting (black) and activated (red) $CD4^+$ T cells from the data of Schones *et al.* (33). Above them the theoretical nucleosome maps calculated iteratively starting from the resting state as an initial distribution are depicted. The remodeling probability depends on the activated state pattern [$P_m(n) = 1/\exp(P_a(n))$], where $P_a(n)$ is the nucleosome start site probability in the activated state and the step size is $s = 10$. The number of iterations is indicated in the Figure. (**A**) The probabilities to move the nucleosome to the left and to the right were determined only by the occupancies of the corresponding target sites according to Equations (8) and (9). Arrows point to the nucleosome locations where the discrepancy between the observed and expected distributions is especially large. (**B**) The probabilities to move the nucleosomes to the left and to the right were determined not only by the occupancies of the target sites, but also included a preferred directionality calculated using the nucleosome preferences given by $P_a(n)$ according to Equations (10) and (11).

sequence usually determines only the location of the –1 and +1 nucleosomes in yeast. The +1 nucleosome forms a barrier against which nucleosomes are packed, resulting in uniform positioning, which decays at increasing distances from the barrier (27,61). An even stronger barrier might be imposed by the insulator binding protein CTCF, which can induce phasing for about 10 adjacent nucleosomes on each side (65,66). This is consistent with our calculations in Figure 6 suggesting that a single strong positioning signal may position several nucleosomes. Thus, a regular nucleosome spacing can be obtained also in the absence of any linker length dependence in the remodeling reaction.

Sequence-specific repositioning. In this scenario, the remodeler repositions nucleosomes from all sites except those coinciding with the nucleosome sites in the final experimental distribution. This would correspond to the release model for nucleosome positioning proposed previously (16). The additional regulatory signals that direct such a repositioning activity could involve changes of the remodeler activity for example through the binding/release of regulatory subunits and/or modifications of epigenetic signals recognized by the remodeler. This is not considered here explicitly; instead, all nucleosome-position-specific remodeling activities are included into the probability for translocation $P_m(n)$.

In Figure 7, the initial distribution is derived from the experimental nucleosome scores for 'resting' $CD4^+$ T cells (33). The remodeling probabilities were then estimated from the experimental nucleosome score for 'activated' $CD4^+$ T cells. Initially, these probabilities were set to be equal to the inverse values of the nucleosome start site probabilities corresponding to the activated cells. However, with this relation the nucleosome-binding map did not converge to the experimental map of the activated cells even after more than 100 000 iterations (data not shown). Therefore, a stronger (exponential) dependence of the remodeling probability was used according to $P_m(n) = 1/(\exp(P_a(n))$, where $P_a(n)$ is the start site

probability for the activated CD4$^+$ T cells (Figure 7A). The probabilities to move the nucleosome to the left and to the right, $P_{-s}(n)$ and $P_{+s}(n)$, were determined by the occupancies of the target sites according to Equations (8) and (9). As shown in Figure 7A, the calculated map tends to move from the experimental score of the resting cell to the activated one. After a large number of iterations, the theoretical nucleosome distribution captures many features of the experimentally activated nucleosome map. However, even after 100 000 iterations, there are still sites such as those indicated by arrows in Figure 7A, which do not coincide with the experimental peaks. For a even higher number of iterations (up to 1 000 000 iterations were conducted), the distribution pattern did not change significantly. When starting the calculations from the initial nucleosome distribution with some peaks deleted (mimicking the eviction of nucleosomes at these sites), the final calculated maps still did not converge to the experimental distribution in the activated cells (data not shown). Thus, the remodeling mechanism used in these calculations is not efficient with respect to mediating the change of the nucleosome position pattern between the resting to the activated cell state.

Directional nucleosome repositioning. Figure 7B shows the calculations performed as in Figure 7A, but including different probabilities to move the nucleosome to the left and to the right in the following form:

$$P_{-s}(n) = K(n-s)/K(n);\ P_{+s}(n) = K(n+s)/K(n)$$

$P_{-s}(n)$ and $P_{+s}(n)$ were normalized to [0,1] and further refined by the occupancies of the target sites according to Equations (10) and (11). Here, the binding constants were determined from the nucleosome binding map of the activated cell state with $K(n) = P_a(n)/(1-P_a(n))$. This description will result in a preferred translocation of the nucleosome in the direction, in which a more thermodynamically stable configuration is obtained. As shown in Figure 7B, after 100 000 iterations the calculated binding map converges to the experimental map for the activated CD4$^+$ T cell state. Thus, the change between nucleosome-positioning patterns as in the activation of resting CD4$^+$ T cells can be achieved much more efficiently if nucleosome translocations occur with a preferred directionality. How could this type of directional translocation be implemented into the molecular mechanism by which chromatin remodelers operate? From the above results we propose that effective remodelers are able to reposition the nucleosome by small rather than large steps and that they sample the intrinsic nucleosome–DNA interaction at different positions of the histone octamer core with respect to the DNA. According to the available structural data the remodeler usually represents a large enzymatic complex that surrounds the nucleosome as its substrate in the reactive center (67). Mechanistically, this type of interaction could provide a specific microenvironment in which DNA–histone interactions are easily being broken and re-established. This would facilitate loop propagation around the histone octamer core so that many DNA sequence positions are sampled from which the energetically most favorable ones are selected. On top of this, certain remodelers might operate through a direct readout of the DNA sequence and/or epigenetic marks, thus providing the targeting of nucleosomes to specific positions.

CONCLUSIONS

Understanding the factors that govern nucleosome positioning is an important aspect of the regulation of gene expression, since the (re)positioning of nucleosomes at regulatory regions determines, which genes are active (2,30,33). The analysis of *in vitro* and *in vivo* data on nucleosome positions, like those shown in Figures 2 and 3, confirms the previously reported correlation with energetically preferred positions that are determined by the DNA sequence. However, it is also apparent that the relative occupancies of these positions as well as the translocation of nucleosomes to certain sites cannot be accounted for by the DNA sequence alone. Therefore, a method was developed here to calculate nucleosome positions, which takes into account the intrinsic DNA affinities of the histone octamer and ATP-dependent repositioning by remodelers. In calculations based on the *in vitro* experiments using the *Drosophila hsp70* promoter several main classes of remodelers were characterized. These are referred to as 'amplifier', 'remover' and 'spacer' to facilitate the rationalization of the observed changes in the nucleosome position patterns as illustrated in Figure 5. It is noted that any type and combination of translocation activities can be included in our approach for calculating nucleosome-position patterns through an appropriately defined function $P_m(n)$ for the probability that a nucleosome at position n is translocated.

Our calculations have a number of further mechanistic implications for nucleosome remodeling. The 'spacer' type of DNA sequence-independent remodeling explains how a regular positioning of nucleosomes can be established that is defined by the presence of a few strong (intrinsic) nucleosome-positioning sequences or insulator type boundary elements as for example those formed by the CTCF protein (Figures 4A and Figure 6). This type of remodeler does not require any DNA sequence-dependent activity. On the other hand, for a remodeler-'amplifier' the probability of nucleosome translocation is dependent on the strength of histone–DNA interactions so that stronger binding sites for the histone octamer are enriched and weaker sites are depleted (Figure 4B). A number of DNA sequences have been identified as natural nucleosome-positioning and nucleosome-excluding motives (7,8,68,69). The repeats of bendable AT-rich sequences seem to favor nucleosome positioning, while more rigid GC-rich sequences tend to exclude nucleosomes. Finally, a chromatin remodeling complex can (directly or indirectly) readout DNA sequence features that do not simply reflect differences in the histone–DNA affinity (16,36,70). This observation can be used to rationalize the 'remover'-type activity identified here: nucleosomes bound at sites where they could act as particularly good substrates for the remodeling reaction are removed

according to the so-called 'release' mechanisms to other positions (16).

The question how remodelers actually move multiple nucleosomes *in vivo* to achieve a precise pattern characteristic for a new cell state was addressed in iterative calculations of the binding maps corresponding to a locus in the resting human $CD4^+$ T cell with the aim to derive the nucleosome distribution of the activated cell state. It appeared that although the different remodeler activities mentioned above can explain the changes in the nucleosome maps qualitatively, the dynamic remodeling of many nucleosomes was a relatively slow process. Accordingly, we propose that the energetic differences of the histone–DNA positioning landscape are sampled in the nucleosome-remodeling complex to select translocations that are more favorable in one direction than in the other (Figure 7).

In our model, the nucleosome core particle is considered as a single entity that occupies 147 bp of DNA. The formation of subnucleosomal particles (like a hexasome missing one H2A·H2B dimer or a H3·H4 tetramer complex bound to DNA in the absence of H2A·H2B) is not considered (71). In addition, an overlapping of nucleosome positions, i.e. a nucleosome repeat length <147 bp is prohibited. Several recent experiments suggest that the situation *in vivo* may be more complicated. Chromatin remodelers can alter the nucleosome structure (49), and nucleosomes can invade territories occupied by other nucleosomes (72). In addition, a remodeler can fully or partly evict histones from a nucleosome (32,73). These activities could provide mechanisms that speed up the large-scale transitions from one nucleosome-positioning pattern to another. It is noted that our theoretical framework is very general. By including additional states, it can be extended to account for the occurrence of subnucleosomal particles as well as a shift of the equilibrium between nucleosomes and subnucleosomal particles due to the remodeler activity.

Another mechanistic aspect to be explored in further studies is the recruitment and activation of the remodeling machinery through the sequence-specific binding of other proteins to the DNA (74). Transcription factors recognize their specific sequences on the DNA and direct remodelers to those nucleosomes that should be repositioned. The theoretical approach described here can be extended to take into account cooperative/competitive interactions of nucleosomes with transcription factors or architectural proteins such as H1, HMG-group and larger chromosome proteins involved in the epigenetic regulation (75–78). This might be particularly relevant for predicting local dynamical changes in nucleosome positioning at promoter and enhancer sequences (2,31,33). In addition, histone variants such as H2A.Z and histone modifications are likely to represent additional remodeler-targeting signals (79,80). Inasmuch as these act as different substrates for remodeling complexes, they could serve to direct the nucleosome translocation reaction to certain sites. Within the theoretical approach reported here, these types of nucleosome subtypes may be considered as additional species in the calculations (see Supplementary Data for details). Thus, our nucleosome translocation models together with the equilibrium association–dissociation description provide an extendible framework for studying the mechanistic details of multi-nucleosome rearrangements during transitions between different chromatin states.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Gernot Längst and Thomas Höfer for discussions and thank Dustin Schones for help in using the nucleosome positioning data of resting and activated T cells.

FUNDING

The Boehringer Ingelheim Foundation and a DKFZ Guest Scientist Fellowship (to V.T.), and DFG grant Ri 1283/8-1 (to K.R.). Funding for open access charge: Boehringer Ingelheim Foundation and DKFZ Guest Scientist Fellowship (to V.T.), and DFG grant Ri 1283/8-1 (to K.R.).

Conflict of interest statement. None declared.

REFERENCES

1. Schnitzler,G.R. (2008) Control of nucleosome positions by DNA sequence. *Cell. Biochem. Biophys.*, **51**, 67–80.
2. Whitehouse,I., Rando,O.J., Delrow,J. and Tsukiyama,T. (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature*, **450**, 1031–1035.
3. Schulze,S.R. and Wallrath,L.L. (2007) Gene regulation by chromatin structure: paradigms established in Drosophila melanogaster. *Annu. Rev. Entomol.*, **52**, 171–192.
4. Henikoff,S. (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.*, **9**, 15–26.
5. Zhao,X., Pendergrast,P.S. and Hernandez,N. (2001) A positioned nucleosome on the human U6 promoter allows recruitment of SNAPc by the Oct-1 POU domain. *Mol. Cell*, **7**, 539–549.
6. Richmond,T.J. (2006) Genomics: predictable packaging. *Nature*, **442**, 750–752.
7. Thåström,A., Lowary,P.T., Widlund,H.R., Cao,H., Kubista,M. and Widom,J. (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.*, **288**, 213–229.
8. Segal,E., Fondufe-Mittendorf,Y., Chen,L., Thåström,A., Field,Y., Moore,I.K., Wang,J.P. and Widom,J. (2006) A genomic code for nucleosome positioning. *Nature*, **442**, 772–778.
9. Whitehouse,I. and Tsukiyama,T. (2006) Antagonistic forces that position nucleosomes in vivo. *Nat. Struct. Mol. Biol.*, **13**, 633–640.
10. Ioshikhes,I.P., Albert,I., Zanton,S.J. and Pugh,B.F. (2006) Nucleosome positions predicted through comparative genomics. *Nat. Genet.*, **38**, 1210–1215.
11. Peckham,H.E., Thurman,R.E., Fu,Y., Stamatoyannopoulos,J.A., Noble,W.S., Struhl,K. and Weng,Z. (2007) Nucleosome positioning signals in genomic DNA. *Genome Res.*, **17**, 1170–1177.
12. Workman,J.L. and Kingston,R.E. (1992) Nucleosome core displacement in vitro via a metastable transcription factor-nucleosome complex. *Science*, **258**, 1780–1784.
13. Morozov,A.V., Fortney,K., Gaykalova,D.A., Studitsky,V.M., Widom,J. and Siggia,E.D. (2009) Using DNA mechanics to predict *in vitro* nucleosome positions and formation energies. *Nucleic Acids Res.* Advance Access, June 9, DOI 10.1093/nar/gkp1475.

14. Mollazadeh-Beidokhti,L., Deseigne,J., Lacoste,D., Mohammad-Rafiee,F. and Schiessel,H. (2009) Stochastic model for nucleosome sliding under an external force. *Phys. Rev. E*, **79**, 031922.
15. Hartley,P.D. and Madhani,H.D. (2009) Mechanisms that specify promoter nucleosome location and identity. *Cell*, **137**, 445–458.
16. Rippe,K., Schrader,A., Riede,P., Strohner,R., Lehmann,E. and Längst,G. (2007) DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc. Natl Acad. Sci. USA*, **104**, 15635–15640.
17. Cairns,B.R. (2007) Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat. Struct. Mol. Biol.*, **14**, 989–996.
18. Ferreira,H. and Owen-Hughes,T. (2006) Lighting up nucleosome spacing. *Nat. Struct. Mol. Biol.*, **13**, 1047–1049.
19. Längst,G. and Becker,P.B. (2004) Nucleosome remodeling: one mechanism, many phenomena? *Biochim. Biophys. Acta*, **1677**, 58–63.
20. Schiessel,H., Widom,J., Bruinsma,R.F. and Gelbart,W.M. (2001) Polymer reptation and nucleosome repositioning. *Phys. Rev. Lett.*, **86**, 4414–4417.
21. Chou,T. (2007) Peeling and sliding in nucleosome repositioning. *Phys. Rev. Lett.*, **99**, 058105.
22. Tolstorukov,M.Y., Choudhary,V., Olson,W.K., Zhurkin,V.B. and Park,P.J. (2008) nuScore: a web-interface for nucleosome positioning predictions. *Bioinformatics*, **24**, 1456–1458.
23. Yuan,G.-C. and Liu,J.S. (2008) Genomic sequence is highly predictive of local nucleosome depletion. *PLoS Comp. Biol.*, **4**, e13.
24. Segal,M.R. (2008) Re-cracking the nucleosome positioning code. *Stat. Applic. Gen. Mol. Biol.*, **7**, 14.
25. Kaplan,N., Moore,I.K., Fondufe-Mittendorf,Y., Gossett,A.J., Tillo,D., Field,Y., LeProust,E.M., Hughes,T.R., Lieb,J.D., Widom,J. *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, **458**, 362–366.
26. Segal,E. and Widom,J. (2009) From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat. Rev. Genet.*, **10**, 443–456.
27. Mavrich,T.N., Ioshikhes,I.P., Venters,B.J., Jiang,C., Tomsho,L.P., Qi,J., Schuster,S.C., Albert,I. and Pugh,B.F. (2008) A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.*, **18**, 1073–1083.
28. Mavrich,T.N., Jiang,C., Ioshikhes,I.P., Li,X., Venters,B.J., Zanton,S.J., Tomsho,L.P., Qi,J., Glaser,R.L., Schuster,S.C. *et al.* (2008) Nucleosome organization in the Drosophila genome. *Nature*, **453**, 358–362.
29. Valouev,A., Ichikawa,J., Tonthat,T., Stuart,J., Ranade,S., Peckham,H., Zeng,K., Malek,J.A., 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.
30. Shivaswamy,S., Bhinge,A., Zhao,Y., Jones,S., Hirst,M. and Iyer,V.R. (2008) Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.*, **6**, e65.
31. Lam,F.H., Steger,D.J. and O'Shea,E.K. (2008) Chromatin decouples promoter threshold from dynamic range. *Nature*, **453**, 246–250.
32. Boeger,H., Griesenbeck,J. and Kornberg,R.D. (2008) Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. *Cell*, **133**, 716–726.
33. Schones,D.E., Cui,K., Cuddapah,S., Roh,T.Y., Barski,A., Wang,Z., Wei,G. and Zhao,K. (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell*, **132**, 887–898.
34. van Holde,K.E. (1989) *Chromatin*, Springer, New York.
35. Solis,F.J., Bash,R., Yodh,J., Lindsay,S.M. and Lohr,D. (2004) A statistical thermodynamic model applied to experimental AFM population and location data is able to quantify DNA-histone binding strength and internucleosomal interaction differences between acetylated and unacetylated nucleosomal arrays. *Biophys. J.*, **87**, 3372–3387.
36. Sims,H.I., Lane,J.M., Ulyanova,N.P. and Schnitzler,G.R. (2007) Human SWI/SNF drives sequence-directed repositioning of nucleosomes on C-myc promoter DNA minicircles. *Biochemistry*, **46**, 11377–11388.
37. Hamiche,A., Sandaltzopoulos,R., Gdula,D.A. and Wu,C. (1999) ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell*, **97**, 833–842.
38. Rippe,K., Mazurkiewicz,J. and Kepper,N. (2008) In Dias,R.S. and Lindman,B. (eds), *DNA Interactions with Polymers and Surfactants*. Wiley, London, pp. 135–172.
39. Narlikar,L., Gordan,R. and Hartemink,A.J. (2007) A nucleosome-guided map of transcription factor binding sites in yeast. *PLoS Comput. Biol.*, **3**, e215.
40. Song,J.S., Liu,X., Liu,X.S. and He,X. (2008) A high-resolution map of nucleosome positioning on a fission yeast centromere. *Genome Res.*, **18**, 1064–1072.
41. Teif,V.B. (2007) General transfer matrix formalism to calculate DNA-protein-drug binding in gene regulation: application to O_R operator of phage lambda. *Nucleic Acids Res.*, **35**, e80.
42. Teif,V.B., Harries,D., Lando,D.Y. and Ben-Shaul,A. (2008) Matrix formalism for site-specific binding of unstructured proteins to multicomponent lipid membranes. *J. Pept. Sci.*, **14**, 368–373.
43. Schwanbeck,R., Xiao,H. and Wu,C. (2004) Spatial contacts and nucleosome step movements induced by the NURF chromatin remodeling complex. *J. Biol. Chem.*, **279**, 39933–39941.
44. Zofall,M., Persinger,J., Kassabov,S.R. and Bartholomew,B. (2006) Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat. Struct. Mol. Biol.*, **13**, 339–346.
45. Field,Y., Kaplan,N., Fondufe-Mittendorf,Y., Moore,I.K., Sharon,E., Lubling,Y., Widom,J. and Segal,E. (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.*, **4**, e1000216.
46. Nechipurenko,Y.D., Jovanovic,B., Riabokon,V.F. and Gursky,G.V. (2005) Quantitative methods of analysis of footprinting diagrams for the complexes formed by a ligand with a DNA fragment of known sequence. *Ann. NY Acad. Sci.*, **1048**, 206–214.
47. Crothers,D.M. (1968) Calculation of binding isotherms for heterogeneous polymers. *Biopolymers*, **6**, 575–584.
48. Gurskii,G.V., Zasedatelev,A.S. and Vol'kenshtein,M.V. (1972) Theory of one-dimensional absorption. II. Adsorption of small molecules on a heteropolymer. *Mol. Biol.*, **6**, 385–393.
49. Sims,H.I., Baughman,C.B. and Schnitzler,G.R. (2008) Human SWI/SNF directs sequence-specific chromatin changes on promoter polynucleosomes. *Nucleic Acids Res.*, **36**, 6118–6131.
50. Epstein,I.R. (1978) Cooperative and noncooperative binding of large ligands to a finite one-dimensional lattice. A model for ligand-oligonucleotide interactions. *Biophys. Chem.*, **8**, 327–339.
51. Kornberg,R.D. and Stryer,L. (1988) Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Res.*, **16**, 6677–6690.
52. Iovanovich,B. and Nechipurenko,Y.D. (1990) Analysis of distribution of ligands adsorbed on DNA fragments. *Mol. Biol.*, **24**, 478–486.
53. Georgel,P.T. (2005) Chromatin potentiation of the hsp70 promoter is linked to GAGA-factor recruitment. *Biochem. Cell. Biol.*, **83**, 555–565.
54. Loots,G.G., Locksley,R.M., Blankespoor,C.M., Wang,Z.E., Miller,W., Rubin,E.M. and Frazer,K.A. (2000) Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science*, **288**, 136–140.
55. Kepert,J.F., Tóth,K.F., Caudron,M., Mucke,N., Langowski,J. and Rippe,K. (2003) Conformation of reconstituted mononucleosomes and effect of linker histone H1 binding studied by scanning force microscopy. *Biophys. J.*, **85**, 4012–4022.
56. Sakaue,T., Yoshikawa,K., Yoshimura,S.H. and Takeyasu,K. (2001) Histone core slips along DNA and prefers positioning at the chain end. *Phys. Rev. Lett.*, **87**, 078105.
57. Gangaraju,V.K. and Bartholomew,B. (2007) Dependency of ISW1a chromatin remodeling on extranucleosomal DNA. *Mol. Cell. Biol.*, **27**, 3217–3225.
58. Stockdale,C., Flaus,A., Ferreira,H. and Owen-Hughes,T. (2006) Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *J. Biol. Chem.*, **281**, 16279–16288.
59. Schwab,D.J., Bruinsma,R.F., Rudnick,J. and Widom,J. (2008) Nucleosome switches. *Phys. Rev. Lett.*, **100**, 228105.
60. Dillon,N. (2004) Heterochromatin structure and function. *Biol. Cell*, **96**, 631–637.
61. Kharchenko,P.V., Woo,C.J., Tolstorukov,M.Y., Kingston,R.E. and Park,P.J. (2008) Nucleosome positioning in human *HOX* gene clusters. *Genome Res.*, **18**, 1554–1561.

62. Ito,T., Levenstein,M.E., Fyodorov,D.V., Kutach,A.K., Kobayashi,R. and Kadonaga,J.T. (1999) ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. *Genes Dev.*, **13**, 1529–1539.
63. Racki,L.R. and Narlikar,G.J. (2008) ATP-dependent chromatin remodeling enzymes: two heads are not better, just different. *Curr. Opin. Genet. Dev.*, **18**, 137–144.
64. Yang,J.G., Madrid,T.S., Sevastopoulos,E. and Narlikar,G.J. (2006) The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nat. Struct. Mol. Biol.*, **13**, 1078–1083.
65. Cuddapah,S., Jothi,R., Schones,D.E., Roh,T.-Y., Cui,K. and Zhao,K. (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.*, **19**, 24–32.
66. Fu,Y., Sinha,M., Peterson,C.L. and Weng,Z. (2008) The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.*, **4**, e1000138.
67. Chaban,Y., Ezeokonkwo,C., Chung,W.-H., Zhang,F., Kornberg,R.D., Maier-Davis,B., Lorch,Y. and Asturias,F.J. (2008) Structure of a RSC-nucleosome complex and insights into chromatin remodeling. *Nat. Struct. Mol. Biol.*, **15**, 1272–1277.
68. Radwan,A., Younis,A., Luykx,P. and Khuri,S. (2008) Prediction and analysis of nucleosome exclusion regions in the human genome. *BMC Genomics*, **9**, 186.
69. Gutierrez,J., Paredes,R., Cruzat,F., Hill,D.A., van Wijnen,A.J., Lian,J.B., Stein,G.S., Stein,J.L., Imbalzano,A.N. and Montecino,M. (2007) Chromatin remodeling by SWI/SNF results in nucleosome mobilization to preferential positions in the rat osteocalcin gene promoter. *J. Biol. Chem.*, **282**, 9445–9457.
70. He,X., Fan,H.-Y., Garlick,J.D. and Kingston,R.E. (2008) Diverse regulation of SNF2h chromatin remodeling by noncatalytic subunits. *Biochemistry*, **47**, 7025–7033.
71. Mazurkiewicz,J., Kepert,J.F. and Rippe,K. (2006) On the mechanism of nucleosome assembly by histone chaperone NAP1. *J. Biol. Chem.*, **281**, 16462–16472.
72. Engeholm,M., de Jager,M., Flaus,A., Brenk,R., van Noort,J. and Owen-Hughes,T. (2009) Nucleosomes can invade DNA territories occupied by their neighbors. *Nat. Struct. Mol. Biol.*, **16**, 151–158.
73. Mizuguchi,G., Shen,X., Landry,J., Wu,W.H., Sen,S. and Wu,C. (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science*, **303**, 343–348.
74. Narlikar,G.J., Fan,H.Y. and Kingston,R.E. (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, **108**, 475–487.
75. Bonaldi,T., Längst,G., Strohner,R., Becker,P.B. and Bianchi,M.E. (2002) The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. *EMBO J.*, **21**, 6865–6873.
76. Ishii,H. (2000) A statistical-mechanical model for regulation of long-range chromatin structure and gene expression. *J. Theor. Biol.*, **203**, 215–228.
77. Miller,J.A. and Widom,J. (2003) Collaborative competition mechanism for gene activation in vivo. *Mol. Cell Biol.*, **23**, 1623–1632.
78. Adkins,N.L., McBryant,S.J., Johnson,C.N., Leidy,J.M., Woodcock,C.L., Robert,C.H., Hansen,J.C. and Georgel,P.T. (2009) Role of nucleic acid Binding in Sir3p-dependent interactions with chromatin fibers. *Biochemistry*, **48**, 276–288.
79. Li,B., Carey,M. and Workman,J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707–719.
80. Wysocka,J., Swigut,T., Xiao,H., Milne,T.A., Kwon,S.Y., Landry,J., Kauer,M., Tackett,A.J., Chait,B.T., Badenhorst,P. *et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, **442**, 86–90.

SUPPLEMENTARY MATERIALS

Predicting nucleosome positions on the DNA: combining intrinsic sequence preferences and remodeler activities

Vladimir B. Teif[1,2,*] and Karsten Rippe[1,*]

[1]Research Group Genome Organization & Function, Deutsches Krebsforschungszentrum and BioQuant, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany and [2]Institute of Bioorganic Chemistry, Belarus National Academy of Sciences, Kuprevich 5/2, 220141, Minsk, Belarus

*Corresponding authors:
Vladimir Teif, Tel.: +49-6221-54-51378; E-mail: Vladimir.Teif@bioquant.uni-heidelberg.de
Karsten Rippe, Tel.: +49-6221-5451376; E-mail: Karsten.Rippe@bioquant.uni-heidelberg.de

Table S1. Enumeration of states of a DNA unit in the multi-protein binding model. A unit may be either free from proteins (nucleosomes) or covered by a protein bound to DNA. There are f different types of protein-DNA complexes, $g = 1,..f$. A g-type complex involves m_g DNA units. The position of a DNA unit inside the complex is numbered from left to right $(1,\ldots m_g)$. The first unit in the complex is assigned the binding constant K_{ng}, where n is the number of the DNA unit. The statistical weight of the last unit in the complex depends on the type and the presence of the protein at the next DNA unit. The protein-protein contact is characterized by the cooperativity parameter $w = w(0, g_n, g_{n+1})$. The weights of the DNA units inside the binding site are equal to one if unit n in state i is followed by unit $n+1$ in state $i+1$, otherwise the weight is equal to zero. Each protein type g is assigned a maximum interaction length, V_g. A gap of l units ($l \le V_g$) between the proteins g_1 and g_2 is assigned a statistical weight $w(l, g_1, g_2)$. The gaps longer than V_g between the proteins g_1 and g_2 are assigned the weights $w(0\ldots V_{g+1}, 0, g_2) = 1$ and $w(0\ldots V_{g+1}, g_1, 0) = 1$.

Number of state of the *n*'th unit, *i*	**State description**		**Statistical weight**
	type of complex	position of unit	
1	1	1	K_{n1}
…		…	…
m_1		m_1	$w(0, 1, g_{n+1})$
…	…	…	…
$m_1+\ldots+m_{g-1}+1$	g	1	K_{ng}
…		…	…
$m_1+\ldots+m_g$		m_g	$w(0, g, g_{n+1})$
…	…	…	…
$m_1+\ldots+m_{f-1}+1$	f	1	K_{nf}
…		…	…
$m_1+\ldots+m_f$		m_f	$w(0, f, g_{n+1})$
$\sum_{g=1}^{f} m_g + 1$	left free DNA end		1
$\sum_{g=1}^{f} m_g + 2$	right free DNA end		1
…	…		…
$\sum_{g=1}^{f} m_g + 2 + \sum_{g=1}^{g_2-1} V_g + l$	g_1-l-g_2 gap (l free units before next g_2 protein), $l \le V_{g2}$		$w(l, g_1, g_2)$
…	…		…
$\sum_{g=1}^{f} (m_g + V_g) + 2 + l$	g_1-l-g_2 gap (l free units before next g_2 protein), $l > V_{g2}$		1
…	…		…
$\sum_{g=1}^{f} (m_g + V_g) + 2 + + \max(V_g) + 1$	free unit out of protein-protein interactions, not at the DNA ends		1

Table S2. States enumeration for a DNA unit in a nucleosome-DNA binding system:

State, *i*	Description	
1	Nucleosome start site	Bound
…	Nucleosome continues	
m	Nucleosome end site	
m + 1	Left DNA end	Free
m + 2	Right DNA end	
m + 3	Free between nucleosomes	

Table S3. The transfer matrix for the single-nucleosome translocation model. Zero elements are shaded:

State #	1	2	3	…	*m*	*m* +1	*m* +2
1		K_n		…			
2			1	…			
3				…			
…	…	…	…	…	…	…	…
m				…			1
m + 1	1			…		1	
m + 2				…			1

Table S4. The transfer matrix for a reversible binding model allowing nucleosome dissociation. All nucleosomes belong to one type with the same DNA affinities:

State #	1	2	3	…	*m*	*m* +1	*m* +2	*m* +3
1		c_0K_n		…				
2			1	…				
3				…				
…	…	…	…	…	…	…	…	…
m	1			…			1	1
m + 1	1			…		1	1	
m + 2				…			1	
m + 3	1			…				1

Transformations between the occupancy maps and the start site maps.

Let $C(n)$ is the probability that the DNA unit n is covered by a nucleosome, and $P(n)$ is the probability that a nucleosome starts at a DNA unit n. Then the values of $C(n)$ and $P(n)$ are connected by the following recurrent relations:

$$n \le m,\ C(n) = \sum_{1}^{n} P(k) \tag{1}$$

$$n < m,\ P(n) = C(n) - \sum_{1}^{n-1} P(k), P(1) = C(1) \tag{2}$$

$$m < n < L,\ C(n) = \sum_{n-m+1}^{n} P(k) \tag{3}$$

$$m < n \le L - m + 1,\ P(n) = C(n) - \sum_{n-m+1}^{n-1} P(k) \tag{4}$$

As follows from equations 1-4, the transformation from $P(n)$ to $C(n)$ is straightforward. On the other hand, the transformation from C(n) to $P(n)$ is possible only when the boundary conditions are well-defined. Figures S1 and S2 explain how the transformations are performed in practice:

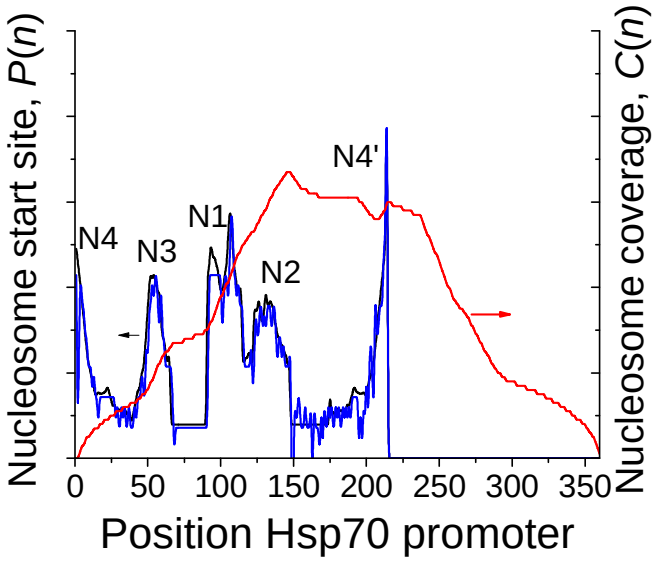

Figure S1. Going from the start site probabilities to the occupancy probabilities and back. The probabilities to start a nucleosome at a given position (black line) were determined from a densitometric scan of the experimental electrophoresis gels of the hsp70 promoter region studied in previous in vitro experiments (1). Using equations 1 and 3, the occupancies $C(n)$ (red line) were calculated. Then, using calculated $C(n)$ as an input, we recalculated the probabilities of nucleosome start sites $P(n)$ (blue line). Ideally, it should exactly coincide with the black line.

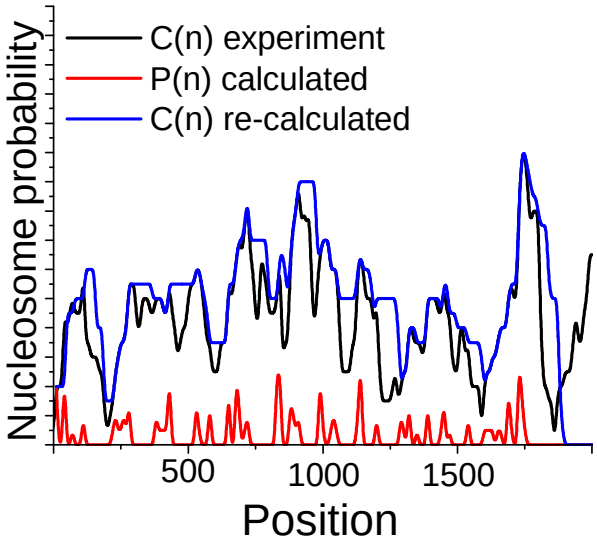

Figure S2. Going from the occupancy probabilities to the start site probabilities and back. A region of the human chromosome 5 (position 132 026 342 to 132 1028 342) is shown. We start from the experimental data of ref. (2) to obtain the probability of occupancy of a given DNA base pair, $C(n)$ (black line). These data were used to calculate the probabilities $P(n)$ to start the nucleosome at a given DNA position (red line, not to scale). Then the nucleosome coverage probabilities $C(n)$ were recalculated using the calculated $P(n)$ values as input (blue line). The black and green lines do not completely coincide due to the lost of some information during the recalculation process. Note that in Figure S1 it was possible to transform the occupancies to start site probabilities and visa versa almost without any loss of information, while that is not the case for the genomic region shown in Figure S2. This is due to the boundaries at the start and end of this genomic region, which do not allow treating the nucleosomes as only partially belonging to this region (see for example the last (right) nucleosome). Nevertheless, this type of transformation is quite accurate considering the experimental uncertainty of the nucleosome positioning data.

References

1. Rippe, K., Schrader, A., Riede, P., Strohner, R., Lehmann, E. and Längst, G. (2007) DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc Natl Acad Sci U S A*, **104**, 15635-15640.
2. Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G. and Zhao, K. (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell*, **132**, 887-898.