

# Affinity, stoichiometry and cooperativity of heterochromatin protein 1 (HP1) binding to nucleosomal arrays

Vladimir B Teif<sup>1</sup>, Nick Kepper<sup>1</sup>, Klaus Yserentant<sup>1,2</sup>, Gero Wedemann<sup>3</sup> and Karsten Rippe<sup>1</sup>

<sup>1</sup> Deutsches Krebsforschungszentrum & BioQuant, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany

<sup>2</sup> Cell Networks Cluster & Institut für Physikalische Chemie, Universität Heidelberg, Im Neuenheimer Feld 267, 69120 Heidelberg (current address)

<sup>3</sup> CC Bioinformatics, University of Applied Sciences Stralsund, Zur Schwedenschanze 15, 18435 Stralsund, Germany

E-mail: [v.teif@dkfz.de](mailto:v.teif@dkfz.de) and [karsten.rippe@dkfz.de](mailto:karsten.rippe@dkfz.de)

Received 8 June 2014, revised 12 August 2014

Accepted for publication 26 August 2014

Published 7 January 2015



CrossMark

## Abstract

Heterochromatin protein 1 (HP1) participates in establishing and maintaining heterochromatin via its histone-modification-dependent chromatin interactions. In recent papers HP1 binding to nucleosomal arrays was measured *in vitro* and interpreted in terms of nearest-neighbour cooperative binding. This mode of chromatin interaction could lead to the spreading of HP1 along the nucleosome chain. Here, we reanalysed previous data by representing the nucleosome chain as a 1D binding lattice and showed how the experimental HP1 binding isotherms can be explained by a simpler model without cooperative interactions between neighboring HP1 dimers. Based on these calculations and spatial models of dinucleosomes and nucleosome chains, we propose that binding stoichiometry depends on the nucleosome repeat length (NRL) rather than protein interactions between HP1 dimers. According to our calculations, more open nucleosome arrays with long DNA linkers are characterized by a larger number of binding sites in comparison to chains with a short NRL. Furthermore, we demonstrate by Monte Carlo simulations that the NRL dependent folding of the nucleosome chain can induce allosteric changes of HP1 binding sites. Thus, HP1 chromatin interactions can be modulated by the change of binding stoichiometry and the type of binding to condensed (methylated) and non-condensed (unmethylated) nucleosome arrays in the absence of direct interactions between HP1 dimers.

Keywords: chromatin, nucleosome, HP1, DNA-protein binding, lattice models, binding stoichiometry, cooperative binding

(Some figures may appear in colour only in the online journal)

## 1. Introduction

The eukaryotic genome is organized inside the cell nucleus by histones and other chromosomal proteins into a nucleo-protein complex called chromatin. At the first level of compaction, 145–147 DNA base pairs (bp) are wrapped around the histone octamer to form the nucleosome – the elementary chromatin unit [32, 47, 74]. At the next level, nucleosomes connected by the linker DNA are organized via their interaction into more

compacted chains [4, 22, 50, 52, 74, 76]. Nucleosome fibres can have different structures and properties depending on their spatial organization and epigenetic modification, which in turn determines access to the DNA and gene expression levels. In a simplified view, two major types of chromatin states can be distinguished: heterochromatin as a more compact and less transcribed conformation, in contrast to the more open euchromatin, which is enriched in actively transcribed genes. In a more detailed classification, multiple chromatin

states can be identified for example by principal component analysis according to the chromosomal protein content [17]. One example of a chromatin protein that demarcates silenced chromatin is heterochromatin protein 1 (HP1) with the yeast homologue Swi6 [24, 33, 36]. The N-terminal chromodomain (CD) and the C-terminal chromoshadow-domain (CSD) of HP1/Swi6 are connected by a flexible linker region. The CD interacts specifically with H3 histone tails that are trimethylated at lysine 9 (H3K9me3) [18, 27], while the CSD promotes the formation of HP1 dimers [6, 45, 51, 77]. Since the interaction of HP1 with the histone methylase SUV39H (Clr4 in yeast) via the CSD has been demonstrated, HP1 binding-driven propagation of H3K9me3 and HP1 has been proposed [16, 21, 53]. Recent studies investigated the H3K9me3 dependent binding of the yeast HP1 homologue Swi6 to different nucleosome substrates *in vitro* [2, 7–9]. From these studies it was concluded that cooperative binding and the oligomerization and spreading of Swi6 on chromatin can contribute to propagating the heterochromatin state along the nucleosome chain. While a number of theoretical models have been developed to account for the switching and propagation of nucleosome states via this type of feedback loop mechanism [3, 14, 15, 25, 38, 55, 57, 59, 78], the mechanistic details of this process are far from being understood. There are also specific theoretical issues that need to be considered. For example, long-range but non contact interactions between DNA-bound proteins can lead to the phase transition of the first kind, due to protein-DNA binding [34, 66]. A series of later works applied similar ideas of long-range interactions to the spreading of the heterochromatin state [15, 38, 57]. Yet, long-range interactions up to now have no experimental support for the case of HP1 chromatin binding, while contact HP1-HP1 interactions would not be sufficient to explain a heterochromatin-euchromatin boundary in the absence of additional factors like insulators or DNA sequence-determined modification domains.

Most of the theoretical approaches consider the DNA as a 1D lattice of units, which can adopt different states depending on reversible binding or covalent modifications [1, 10, 12, 19, 20, 29, 35, 39, 40, 41, 49, 54, 56, 65, 68, 69, 70, 73, 75]. The size of the elementary lattice unit can be set either as one DNA base pair or one nucleosome, or one can use an even larger coarse graining. Thus, these models can be scaled to study genomic events both at the molecular and the systems level [64]. The major challenge, however, is the parameterization of the model based on the experimental data as these principally determine the different properties of the model.

A series of recent studies provided valuable insights into the nucleosome binding properties of HP1/Swi6 [7–9]. Swi6 was found to bind nucleosomes as a dimer with a stoichiometry of two Swi6 dimers per mononucleosome, and assembled on the nucleosomal lattice in a manner strongly affected by the H3K9 methylation state. From an analysis of the binding curves it was concluded that attractive interactions between neighbouring Swi6 dimers exist which lead to the spreading of Swi6 along the nucleosome chain. Furthermore, the length of the DNA linker between nucleosomes had a large effect on HP1 binding. In order to further dissect the mode of Swi6 binding to chromatin, we have performed calculations using

an analytical 1D lattice binding algorithm and 3D Monte Carlo simulations of nucleosome chain folding. The results prompt us to propose an alternative model for Swi6 binding that considers changes in nucleosome chain folding depending on the nucleosome repeat length.

## 2. Model

In the previous experiments, Swi6 binding to mono- and dinucleosomes and 12mer nucleosomal arrays was quantified by measuring the fraction of nucleosomal substrate without any bound Swi6 by gel electrophoresis [7]. Thus, configurations with one or more bound Swi6 proteins per nucleosomal substrate were not distinguished. This is different from the conventional representation of the degree of binding as the fraction of occupied lattice binding sites. In particular, the concentration of half-saturation of such titration curves is not directly related to the standard dissociation constant for protein binding to a nucleosomal lattice. This warrants additional consideration for dissecting binding affinity and cooperativity parameters. In a statistical-mechanical lattice binding formalism, the unbound fraction of nucleosome arrays can be simply expressed as  $1/Z$ , where  $Z$  is the partition function of the system [65, 69, 70]. Correspondingly, the experimental curves reporting ‘1 – fraction unbound’ can be expressed as  $(1 - 1/Z)$ . Next, one has to calculate the partition function, which depends on the selected model of biological interactions. Here we will use the general transfer matrix formalism, previously introduced for more complex systems including DNA loops, multiprotein multilayer assembly and long-range interactions between DNA-bound or membrane-bound proteins [61, 62].

The nucleosome array is considered as a 1D lattice of binding sites with appropriate mathematical rules to make the 1D model equivalent to the corresponding 3D structure under consideration. Previously we have worked with models where the elementary lattice unit is the DNA base pair, protein amino acid or the nucleosome [65, 67–69]. Here for simplicity we will consider one half-wrap of the nucleosome as the elementary binding unit of one Swi6 dimer to one H3 tail of the nucleosome. The transfer matrices are constructed so that each matrix element  $Q_n(i, j)$  contains the probabilities to find the lattice unit  $n$  in a state  $i$  provided the unit  $n + 1$  is in state  $j$ . Prohibited combinations of states are characterized by weights of zero. The partition function is given by sequential multiplication of all transfer matrices enclosed between two unit vectors:

$$Z = (1 \ 1 \ \dots \ 1) \times \prod_{n=1}^N Q_n \times \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \quad (1)$$

Furthermore, we define  $N$  as the number of Swi6-dimer binding sites on a nucleosome array,  $c_0$  as the molar concentration of free Swi6 dimers,  $K_d(n)$  as the Swi6-nucleosome dissociation constant per Swi6 dimer (the inverse value is the Swi6 binding constant  $K(n)$ ) and  $w$  as the contact cooperativity constant for interactions between neighbouring Swi6 dimers [34, 37, 66]. For heterogeneous binding sites of a single protein type to

DNA without long-range interactions, the transfer matrix can be constructed as follows:

$$Q_{ij}(n) = \begin{pmatrix} K(n)c_0w & K(n)c_0 \\ 1 & 1 \end{pmatrix} \quad (2)$$

Here, the rows of the matrix list the states of the current lattice unit  $i$  and the columns list the states of the next lattice unit  $j$ , where only two states are allowed, bound (first row) or free (second row). The situation becomes more complicated if we recollect that the geometric distances between two Swi6 dimers located on the same nucleosome and those located on the neighbouring nucleosomes are not equivalent (up to now we did not make any assumptions about the way in which Swi6 proteins interact with each other). The transfer matrix formalism can account for these features. For Swi6 dimers bound to the same nucleosomes, we can consider two protein binding modes numbered as ‘1’ and ‘2’ characterized by the same protein concentration and binding constant. In this case each lattice unit can be in three states (bound by protein type 1, bound by protein type 2, or free), and the corresponding transfer matrix has 9 elements:

$$Q_{ij}(n) = \begin{pmatrix} K(n)c_0 & K(n)c_0w & K(n)c_0 \\ K(n)c_0w & K(n)c_0 & K(n)c_0w \\ 1 & 1 & 1 \end{pmatrix} \quad (3)$$

By introducing binding constants  $K_1$  and  $K_2$  for different HP1 binding modes, different binding stoichiometries  $m_1$  and  $m_2$  and different cooperativity parameters  $w_{11}$ ,  $w_{12}$ ,  $w_{21}$ ,  $w_{22}$  the model can be made more complex as previously described [61]. However, here the simple case of standard McGhee-von Hippel contact cooperativity (equation (2)) and pair-wise cooperativity (equation (3)) are sufficient. The difference between the latter two types of cooperativity is that according to equation (2) each protein can have interactions with two neighbours at saturation, while in equation (3) each protein can interact only with one of its neighbours which is bound to the same nucleosome.

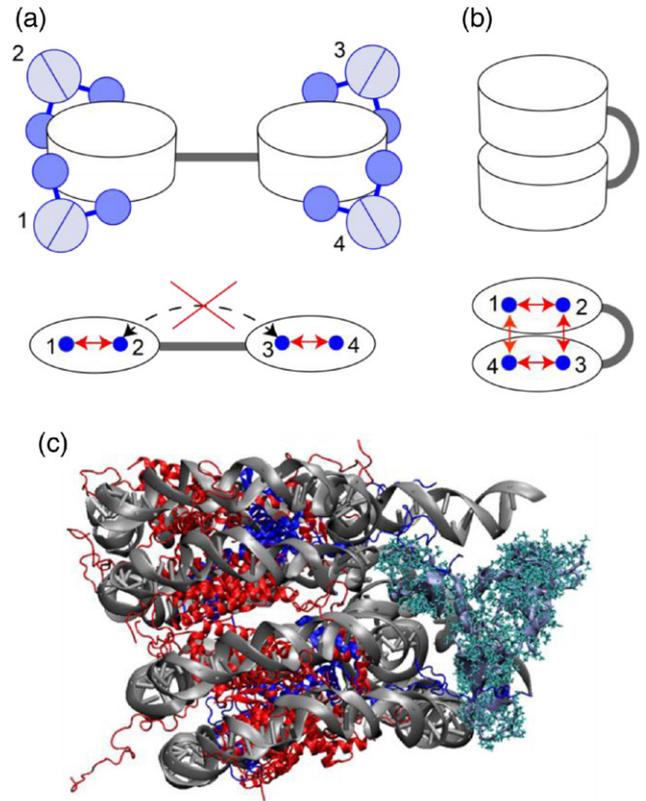
Figure 1 illustrates this concept for Swi6 binding to mono- and dinucleosomes. Strictly speaking, the model in equation (3) does not take into account that proteins #1 and #4 interact, which would require a more complicated model (see below). However, this affects only boundary conditions and thus the approximation by the matrix in equation (2) is justified. The only change is that the value of  $w$  is squared (energy doubled) upon transition from the interaction of a protein with a single-neighbour to two-neighbours.

In addition, we have considered a model in which binding is not site-specific and binding sites are independent. In terms of the molecular features it represents the case that *all* nucleosomes are equally (un)methylated and that proteins bound to neighbouring sites do not interact. Under these conditions, equations (1)–(2) can be solved analytically and the partition function is given as

$$Z = (1 + Kc_0)^N, \quad (4)$$

where  $N$  is the number of binding sites. In a specific case when the system undergoes a conformational transition depicted in figure 1, the partition function is given by equation (5) [48]:

$$(1 + Kc_0)^{N^2} Z = (1 + Kc_0)^{N^1} + K_T(1 + Kc_0)^{N^2}, \quad (5)$$



**Figure 1.** A scheme of the lattice models for different types of contact interactions between nucleosome-bound Swi6 dimers. (a) Pair-wise cooperativity. Swi6 dimers interact only within one nucleosome. At saturation each Swi6 dimer has contact with only one neighbor. (b) When two nucleosomes are stacked on top of each other, neighboring Swi6 dimers could potentially interact in two conformations. Either they are bound to two different nucleosomes as in (a) or they are both bound to two different nucleosomes, which requires their close proximity. (c) Molecular model of a HP1 dimer binding to H3 tails from two different nucleosomes in a stacked dinucleosome structure. The DNA is depicted in grey, histone H3 in dark blue, H2A, H2B and H4 in red and the HP1 dimer in light blue.

where  $N_1$  and  $N_2$  are the numbers of sites in the extended and folded dinucleosome conformation correspondingly and  $K_T$  is the equilibrium constant for the process of dinucleosome folding (figure 1(b)). If dinucleosome folding does not affect the stoichiometry of HP1 binding and all binding sites remain equivalent, equation (5) can be simplified to

$$Z = (1 + wKc_0)^N, \quad (6)$$

where the cost of dinucleosome folding is included in  $w$  as the allosteric interaction between bound HP1s. The latter type of protein binding cooperativity has the same mathematical description as the McGhee-von Hippel cooperativity, but does not require direct protein-protein contact and is induced through the conformational transition of the dinucleosome.

### 3. Results

As the initial reference state, we analyzed Swi6 binding to methylated and non-methylated mononucleosomes using

**Table 1.** Concentration dependent sedimentation coefficients and molecular weights of mouse HP1 $\beta$  determined by analytical ultracentrifugation. Sedimentation coefficients and molecular weights of recombinant mouse His-tagged full-length HP1 $\beta$  were measured and analyzed at the indicated concentrations of monomer as described previously (Müller-Ott *et al.*, 2014). Values refer to H<sub>2</sub>O and a temperature of 20 °C as standard state.

Concentration ( $\mu$ M)	$S_1$ (S) <sup>a</sup>	$S_2$ (S) <sup>a</sup>	$M_{ave}$ (kDa) <sup>b</sup>
1.1	$2.0 \pm 0.1$	–	n.d.
3.5	n.d.	n.d.	$21 \pm 1^c$
4.3	$2.0 \pm 0.1$	$2.8 \pm 0.2$	n.d.
7.2	–	$2.8 \pm 0.1$	n.d.
10–23 $\mu$ M <sup>d</sup>	–	$2.8 \pm 0.2$	$47 \pm 1$

<sup>a</sup> Determined by sedimentation velocity ultracentrifugation and data analysis as described previously.

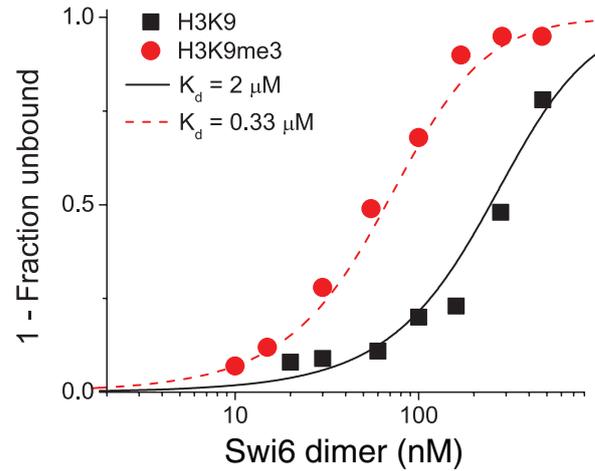
<sup>b</sup> Determined from sedimentation equilibrium ultracentrifugation at 10000 and 15000 rpm and 230 nm and 280 nm absorbance. Data were fitted to a one-component model.

<sup>c</sup> Sedimentation equilibrium ultracentrifugation showed systematic deviations of residues when fit to a one component model. A good fit was obtained with a two-component model with molecular weights of  $M_1 = 18.8$  kDa and  $M_2 = 42.6$  kDa indicative of a significant amount of HP1 monomer being present at this protein concentration.

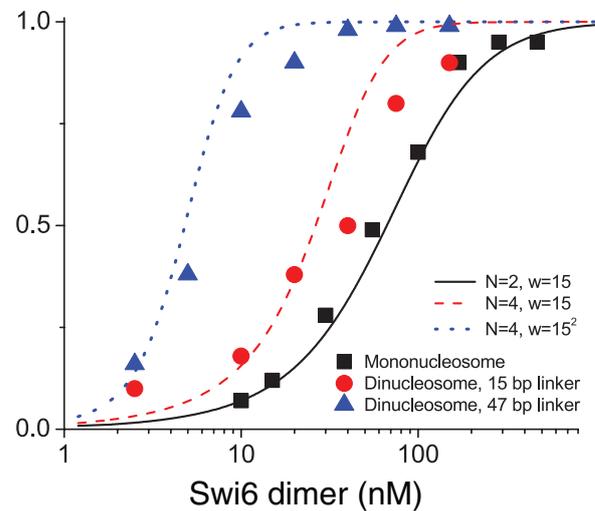
<sup>d</sup> Average values from measurements at 10.4  $\mu$ M and 23.4  $\mu$ M which showed only a single dimeric HP1 species.

equations (1)–(3). For the chromoshadow-domain of Swi6 a value of  $K_d < 17$  nM for its self-association was determined [7]. Thus, it is justified to consider Swi6 as a single dimeric species in the following theoretical analysis of its binding to different nucleosome substrates. However, it is noted that the analysis of mouse HP1 $\beta$  by analytical ultracentrifugation shows that the free protein has a dimerization dissociation constant of about 1  $\mu$ M (table 1) [42]. Accordingly, a HP1 monomer-dimer equilibrium would have to be taken into account in *in vitro* experiments conducted at protein concentrations in the range of  $10^{-6}$  M and below for the mouse HP1 protein. We compared two models: a simple model allowing one Swi6 dimer interacting with one H3 tail and a more complex model, in which one Swi6 dimer interacts either with one or two H3 tails. The fit did not improve when using the more complex model. Accordingly, we used the simpler model with one mode of binding (equation (2)). With this model, we found that Swi6 binding to mononucleosomes can be described well with a contact cooperativity parameter  $w = 15$  and dissociation constant  $K_d = 0.33 \mu$ M for trimethylated and  $K_d = 2 \mu$ M for unmodified nucleosomes. This yields a 1.8 kT ( $\sim 1$  kcal mol<sup>-1</sup>) energy difference per one Swi6 dimer binding to H3K9me3 versus the unmodified H3K9 tail, which is comparable to  $\sim 2.6$  kcal mol<sup>-1</sup> energy of HP1 chromodomain binding to H3K9me3 peptides [26] (figure 2).

Next, we analyzed the observed differences in binding to mononucleosomes and dinucleosomes with different linker lengths (figure 3). We started with the mononucleosome set of parameters determined above. These reproduced the experimental binding curve for the 15 bp linker dinucleosome when doubling the number of mononucleosome binding sites  $N$  and by additionally raising the cooperativity parameter  $w$  to the second power. The molecular mechanism behind this is that with the 15 bp linker each Swi6 dimer had one contact in a



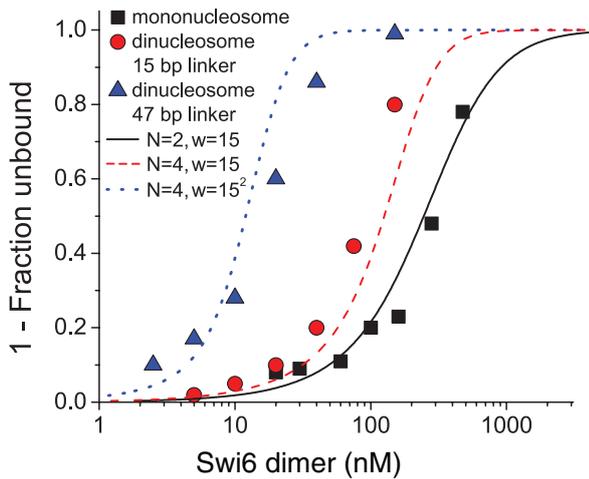
**Figure 2.** Swi6 binding to methylated (red) and non-methylated (black) mononucleosomes. Binding curves depict the fit to the data for  $N = 2$ ,  $w = 15$  according to equation (2) and the dissociation constant  $K_d$  values indicated in the figure. Data points are from previous *in vitro* binding experiments [7].



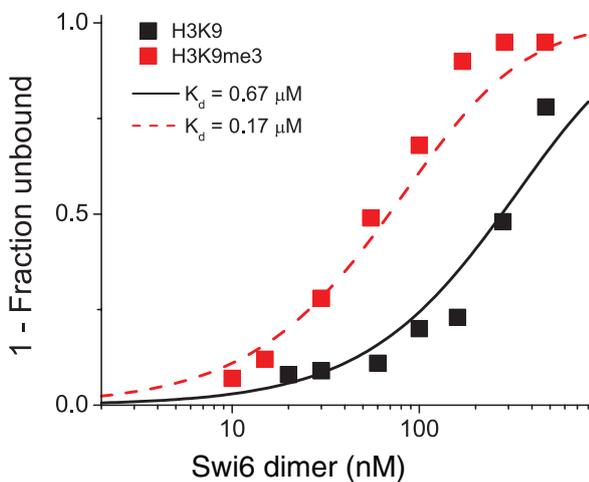
**Figure 3.** Swi6 binding to methylated mono- and dinucleosomes. Binding curves were calculated according to equation (1) for a value of  $K_d = 0.33 \mu$ M. Data points are from previous *in vitro* binding experiments [7].

tetrameric Swi6 structure. In contrast with the 47 bp linker, two Swi6 tetramers can come close to each other in 3D so that each Swi6 dimer would interact with two, not with one neighbouring dimer (see figure 1).

The same binding model fits for non-methylated mono- and dinucleosomes, albeit with an about 6-fold lower binding affinity (figure 4). In this case the experimental data for the mononucleosome were described using the algorithm in equations (1) and (2) with the following parameters:  $N = 2$ ,  $K_d = 2 \mu$ M,  $w = 15$ . By changing the stoichiometry to allow binding of twice as many proteins ( $N = 4$ ), the dinucleosome binding curve was described well. For the dinucleosome with the short linker (15 bp), the binding cooperativity was the same as in the case of the mononucleosome. For the dinucleosome with the long linker (47 bp) additional protein–protein contacts form as manifested by the two-fold increase in the protein–protein



**Figure 4.** Swi6 binding to non-methylated mono- and dinucleosomes. Binding curves were calculated according to equation (1).  $K_d = 2 \mu\text{M}$ . Other parameters are indicated in the figure. Data points are from previous *in vitro* binding experiments [7].

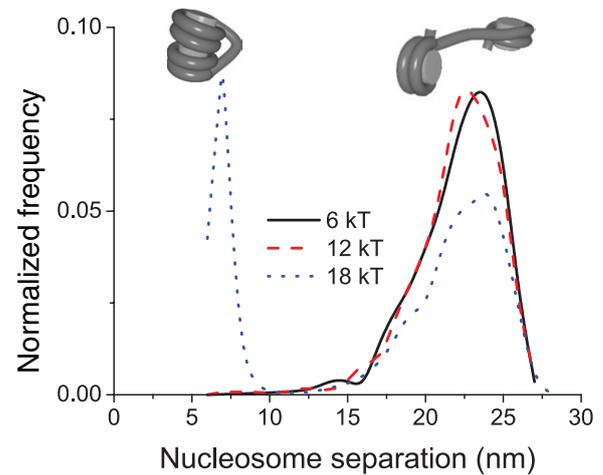


**Figure 5.** Swi6 binding to methylated (red) and non-methylated (black) mononucleosomes. Binding curves represent a theoretical fit by a non-cooperative model (equation (4)) with  $N = 2$  and dissociation constants indicated in the figure. Data points are from previous *in vitro* binding experiments [7].

interaction energy. The corresponding cooperativity constants multiply ( $w^2$  instead of  $w$ ). The latter can be explained if a long linker allows the dinucleosome to fold back so that two core particles make contact with each other and allow all four Swi6 dimers to interact (see figure 1).

Since data fitting in figures (2)–(4) required only weak contact cooperativity, we asked whether it is possible to fit the data without contact cooperativity between HP1 dimers. Figure 5 shows the recalculation using equation (4), assuming that HP1 dimers do not form energetically stable complexes with each other. This fit results in higher dissociation constants ( $K_d = 0.67 \mu\text{M}$  for unmethylated and  $0.17 \mu\text{M}$  for methylated mononucleosomes).

For the dinucleosome our Monte Carlo simulations indicated a folding transition for 47 bp DNA linkers to a stacked structure (figure 1(c)) but not for 15 bp linkers. To quantify the probability of this process, we performed a series of MC

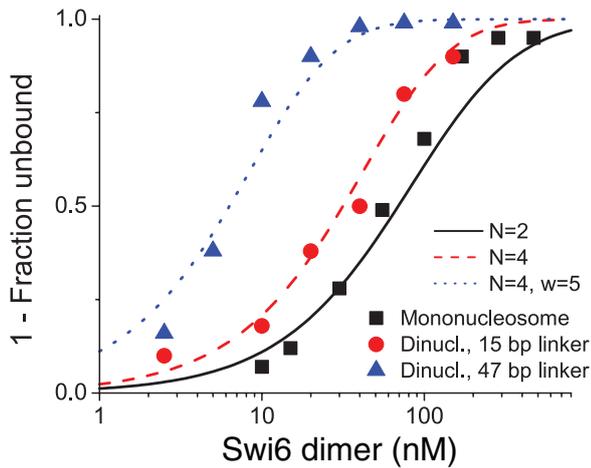


**Figure 6.** Monte Carlo simulation of the dinucleosome folding. The frequency of conformations as a function of the distance between two nucleosomes is plotted for different energies of nucleosome-nucleosome stacking interactions from 6 kT to 18 kT (values indicated on the figure). The first peak at a separation distance of around 6–7 nm corresponds to the folded dinucleosome conformation, where two nucleosomes are stacked on top of each other (see inset). The peak at around 23 nm separations corresponds to the stretched DNA linker between nucleosomes (linker length 47 bp).

simulations of chromatin based on our previous 3D model [31]. In the simulations shown in figure 6, the energy for the nucleosome stacking interactions at an optimal distance and alignment was varied between 6 and 18 kT. As discussed previously this range of values corresponds to those derived from experiments [30].

At 12 kT the dinucleosome with a 47 bp linker length showed the first fold back events (distances between the centres of the two nucleosomes were below 11 nm). The folded back conformation was essentially absent for 15 bp linker dinucleosomes. The energy of 12 to 15 kT is comparable to the energy of the two Swi6s binding to the methylated nucleosome (using  $K_d = 0.33 \mu\text{M}$  as determined in figures (2)–(4)). It is noted that the binding of one HP1 can stabilize the stacked nucleosome conformation and would facilitate the binding of the second dimer via an allosteric change of the nucleosome substrate. Swi6 binding curves with a dinucleosome substrate were well described in equation (4) (HP1 dimers do not interact) (figure 7) and an allosteric cooperativity,  $w = 5$  (equation (6)).

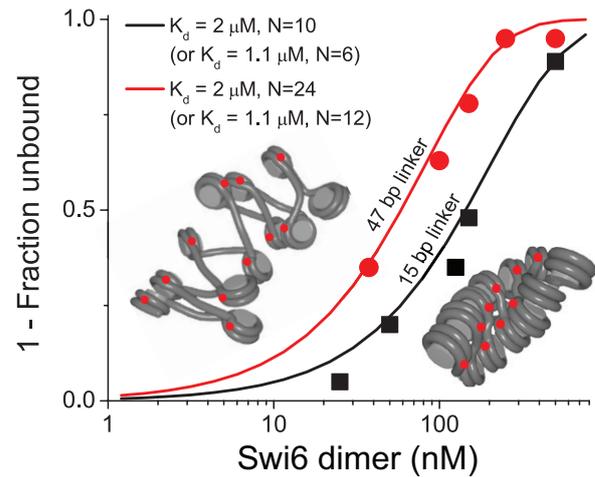
Finally, we analyzed Swi6 binding to the methylated 12 nucleosome arrays (figure 8). A simple model was able to quantitatively describe the experimental data in the absence of Swi6 binding cooperativity. Binding to the 12 nucleosome array with 47 bp linkers was fit with a dissociation constant  $K_d = 2 \mu\text{M}$  and  $N = 24$  binding sites for Swi6 dimers (or alternatively with a dissociation constant  $K_d = 1.1 \mu\text{M}$  and  $N = 12$  binding sites). Keeping the  $K_d$  value and half the number of binding sites, we were able to fit the binding curve corresponding to the 12 nucleosomal array with a 15 bp linker. In particular, if the nucleosomal array with 47 bp linkers can bind 24 Swi6 dimers (two dimers per nucleosome, as established above), the nucleosomal array with a 15 bp linker can bind only 10 dimers, which



**Figure 7.** Swi6 binding to methylated mono- and dinucleosomes. Binding curves for mononucleosomes ( $N = 2$ ) and dinucleosomes with 15 bp linkers ( $N = 4$ ) were calculated according to a non-cooperative model (equation (4)). The curve for dinucleosomes with 47 bp linkers was calculated using equation (6) assuming that the energy of the dinucleosome folding is cast into allosteric binding cooperativity ( $w = 5$ ).  $K_d = 0.17 \mu\text{M}$  for all curves. Data points are from previous *in vitro* binding experiments [7].

connect each second nucleosome (10 Swi6 bridges between 12 nucleosomes).

How can this model be connected to the available structural information? Our Monte Carlo simulations based on the previously described methodology [31] show that increased linker length is related to a more decondensed nucleosomal array conformation. As shown in figure 8 (inset), regular nucleosomal arrays formed with 15 bp DNA linkers have less space for Swi6 binding in comparison with more open chromatin fibers with 47 bp linkers. Furthermore, two nucleosomes which appear to be close to each other in 3D are not necessarily two sequential neighbours in the 1D lattice, as discussed previously in the context of lattice models for chromatin [69]. The latter effect is demonstrated in more detail in figure 9, where the stoichiometry of HP1-chromatin binding is justified (24 HP1 dimers for an amorphous 12 nucleosome array, versus 10 HP1 dimers for a 12 nucleosome array). Note that our MC model for long DNA linkers is fully compatible with the recent high-resolution nucleosomal array structure determined by cryo-EM [58]. This might explain the counterintuitive experimental result that dinucleosomes bind Swi6 similar to mononucleosomes, while longer nucleosomal arrays show significant differences. The considerations above are based on a geometry of the chromatin fibre with mostly straight and cross-linked linker DNA as observed previously in a tetranucleosome crystal structure (NRL = 169 bp) without linker histones [6] and cryo-EM based structures of 12 mer arrays with a linker histone H1 and NRL = 177 or 187 bp [58]. As shown here and in our previous work [31], using a local nucleosome geometry that is derived from the tetranucleosome structure will promote a more open and irregular fibre conformation for longer linker length as found in the 47 bp linker (NRL = 194) nucleosome array versus the 15 bp linker nucleosome chain (NRL = 162 bp). This conformational difference results from an increased electrostatic repulsion between negative charges



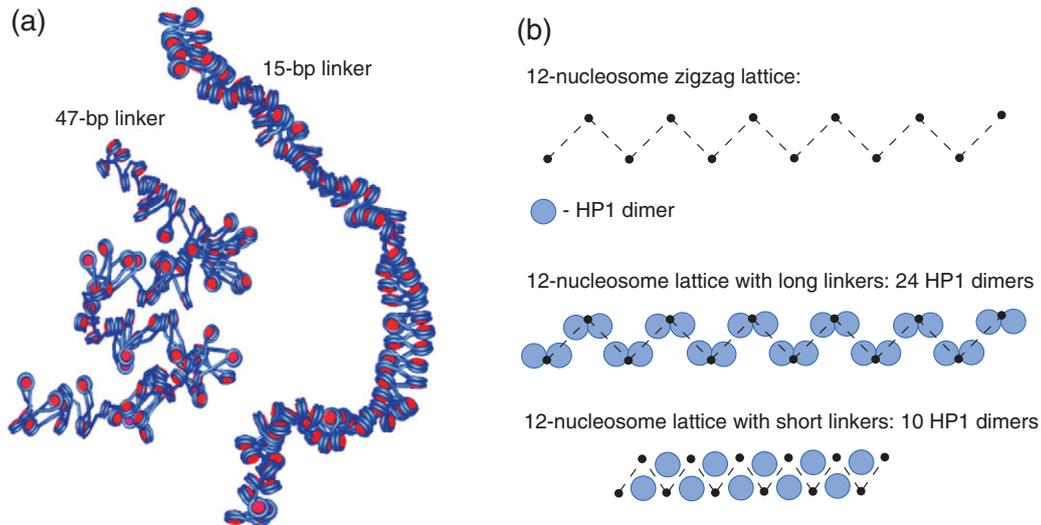
**Figure 8.** Swi6 binding to methylated arrays of 12 nucleosomes. Binding curves were calculated according to equation (1). Other parameters are given in the figure. Insets show representative snapshots from the Monte Carlo simulations of the 12 nucleosome arrays with the linker length 47 bp (left) and 15 bp (right). HP1 proteins were not included in the MC simulations of the chain and are shown here to illustrate how the longer linker could double the number of potential binding sites with accessible H3 N-terminal tails for nucleosomes in spatial proximity. Data points are from previous *in vitro* binding experiments [7].

in the DNA linker. The addition of linker histones would favor fibre compaction as concluded from the appearance of folded nucleosome arrays with a conformation similar to that found in the tetranucleosome crystal structure but with longer linker lengths of 30 and 40 bp [58].

Thus, the change of the binding behaviour in 15 bp versus 47 bp linkers can be explained by a change of binding stoichiometry due to differential compaction of the fibre. In addition, the conformation of nucleosome arrays imposes additional constraints on the HP1 interactions that are not present for mono- and dinucleosomes. It should be noted that while the relative stoichiometry difference between 15 bp versus 47 bp arrays is a feature that can be robustly obtained from the analysis of the data, the absolute stoichiometry changes between mononucleosomes and arrays might be obscured by concomitant changes in binding affinities.

#### 4. Discussion

Here, we conducted an analysis of previously published binding isotherms [7] to dissect the binding of Swi6/HP1 dimers to nucleosomal substrates. In our analysis, we confirm a several-fold increased affinity upon H3K9 trimethylation, which is compatible with other *in vitro* measurement estimates [28, 44]. Furthermore, we found that accounting for the different stoichiometry in the absence of cooperativity is sufficient to rationalize Swi6/HP1 binding to chromatin. Notably, the observed allosteric cooperativity of binding for dinucleosome substrates with long but not with short DNA linkers was quantitatively consistent with the stacking of nucleosomes as predicted from MC simulations. This conformation of two nucleosomes allows for simultaneous interactions of



**Figure 9.** (a) MC simulation of nucleosome arrays with 15 bp and 47 bp linkers. (b) Schematic drawing explaining the potential change of the HP1 binding stoichiometry due to the geometrical constraints imposed by the shorter DNA linker length.

one HP1 dimer with two H3 tails as depicted in the all-atom model in figure 1(c). Interestingly, targeting of HP1 $\beta$  to heterochromatin appears to require the simultaneous recognition of two H3K9-methylated histone H3 molecules [71], which suggests that conformations of the nucleosome chain with two stacked nucleosomes are the preferred HP1 binding sites. The allosteric cooperativity arising due to the dinucleosome conformation change increased the Swi6 binding constant five-fold. This would be significantly larger than nonspecific interactions of Swi6 with DNA, to which an up to two-fold increase in the case of 47 bp linkers was attributed previously [7]. A similar allosteric cooperativity without direct interactions between DNA-bound proteins was described previously for the process of DNA condensation *in vitro* [60]. Thus, we conclude that protein–protein contact cooperativity between Swi6/HP1 dimers is not required to rationalize the available experimental binding curve. Furthermore, our recent analytical ultracentrifugation measurements of mouse HP1 $\beta$  proteins argue against the existence of HP1 tetramers or higher order complexes in solutions since the protein is present in a dimeric state at protein monomer concentrations up to the range of 30  $\mu$ M (table 1) [42]. Accordingly, cooperative interactions between chromatin-bound HP1 dimers are likely to be very weak. Furthermore, Swi6/HP1 binding to nucleosome arrays was compatible with a simple model where the stoichiometry of HP1 binding (but not the cooperativity of the HP1-HP1 interaction) changes depending on the chromatin compaction state (figures 8 and 9). Thus, a contiguous coverage of a genomic region by HP1 spreading would not be expected. This is consistent with recent experimental results which showed that about 90% of mouse HP1 proteins interact with chromatin very transiently with residence time on the second scale, while only  $\sim$ 10% were bound for a period of minutes [11, 42, 43]. At the same time only a moderate enrichment of HP1 and the H3K9me3 modification for pericentric heterochromatin of 2–3 fold above other chromatin loci was observed that followed the higher DNA density in this nuclear

**Table 2.** Parameters of the chain of nucleosomes in the Monte Carlo simulation.

Fibre Type	linker (bp)	$\alpha(^{\circ})$	$\beta(^{\circ})$	$\gamma(^{\circ})$	$\delta(^{\circ})$	$\epsilon(^{\circ})$	$\phi(^{\circ})$	c(nm)	d(nm)
cl	15	70	140	-70	20	0	0	3.3	8
cl	47	25	140	-25	20	0	0	3.3	8

subcompartment [42]. These observations are difficult to reconcile with a model in which HP1 recognition of H3K9me3 is coupled to the spreading of the protein along chromatin due to protein-protein contact cooperativity. Interestingly, the Clr4/SUV39H methylases contain a chromodomain that can recognize H3K9 methylation, which could be important for the propagation of the H3K9me3 state and could compete with the binding of HP1 [2]. Thus, it will be important to further dissect, how Swi6/HP1 and Clr4/SUV39H (and possibly other factors) conspire via protein–protein interactions, recognition of the H3K9me2 and H3K9me3 modifications and long-range interactions between bound proteins along the nucleosome chain due to chromatin looping, to establish, maintain and propagate heterochromatin.

## 5. Materials and Methods

Calculations of the binding curves were performed using the *TFnuc* software suite [5, 61, 63, 64]) as described in the Model section.

MC simulations were performed as described previously [31]. The 12mers were simulated with a maximal nucleosome interaction energy of 6kT for the optimal stacking of two nucleosomes. The dinucleosomes were simulated with maximal interaction energies of 6, 12, 18kT. It is noted that the effective average interaction strength between two nucleosomes is significantly lower as the constraints imposed by the DNA linker partly counteract optimal stacking. The main simulation parameters are summarized in table 2.

The linker length, six angles and two distances are used to define the local nucleosome geometry in the Monte Carlo simulations. The angles  $\alpha$ ,  $\beta$  and  $\gamma$  define the conformation of lowest energy in which the linker DNA is positioned relative to the nucleosome. The parameter  $d$  describes the distance of the linker DNA entering and leaving the nucleosome. The angles  $\delta$ ,  $\epsilon$  and  $\phi$  define the orientation of the nucleosome relative to the DNA entry and exit points. The distance  $c$  describes the distance between the DNA entry and exit points and the centre of the histone core. For details of the parameterization see [31].

The all-atom model of HP1 $\beta$  was built using the pdb structure coordinates 1DZ1 of the chromoshadow domain [6] and 1GUW of the chromo domain [46]. The connecting hinge region was homology modelled. This HP1 structure was then attached to two stacked nucleosomes from the tetranucleosome structure [6] with their H3 tail binding to the HP1 $\beta$  chromo-domain according to the previously reported interactions [46]. The initial structures were refined by energy minimization and molecular dynamics with the Amber 10.0 software package [71]. Steric clashes of the model structures were cleaned up by energy minimization and short molecular dynamics (MD) simulations (100ps) with a generalized Born solvent accessible surface model (ionic strength 150mM,  $\epsilon_{\text{internal}} = 1.0$ ,  $\epsilon_{\text{external}} = 78.5$ ) [23, 72]. The particle mesh Ewald method [13] for non-periodic-calculations was used for the treatment of the electrostatic interactions. As the convergence criterion for the energy gradient, a root-mean-square of the Cartesian elements of a gradient of less than  $0.05 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$  was chosen. Further molecular dynamics simulations were carried out in explicit water, at a physiological salt condition (150 mM) and periodic boundary conditions in an NPT (constant pressure and constant temperature) ensemble for 5 ns.

## Acknowledgements

V T acknowledges the support from the Heidelberg Center for Modelling and Simulation in the Biosciences (BIOMS) and a DKFZ Intramural Grant. Computational resources and data storage were provided via grants from the Bundesministerium für Bildung und Forschung (BMBF) 01IG07015G (Services@MediGRID), Centre for Modelling and Simulation in the Biosciences (BIOMS) at Heidelberg and the Deutsche Forschungsgemeinschaft (DFG) INST 295/27-1, DKFZ Intramural Grant.

## References

- [1] Afek A, Sela I, Musa-Lempel N and Lukatsky D B 2011 Nonspecific transcription-factor-DNA binding influences nucleosome occupancy in yeast *Biophys. J.* **101** 2465–75
- [2] Al-Sady B, Madhani H D and Narlikar G J 2013 Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread *Mol. Cell* **51** 80–91
- [3] Angel A, Song J, Dean C and Howard M 2011 A Polycomb-based switch underlying quantitative epigenetic memory *Nature* **476** 105–8
- [4] Belmont A S 2014 Large-scale chromatin organization: the good, the surprising, and the still perplexing *Curr. Opin. Cell Biol.* **26** 69–78
- [5] Beshnova D A, Cherstvy A G, Vainshtein Y and Teif V B 2014 Regulation of the nucleosome repeat length in vivo by the DNA sequence, protein concentrations and long-range interactions *PLoS Comput. Biol.* **10** e1003698
- [6] Brasher S V, Smith B O, Fogh R H, Nietlispach D, Thiru A, Nielsen P R, Broadhurst R W, Ball L J, Murzina N V and Laue E D 2000 The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer *EMBO J.* **19** 1587–97
- [7] Canzio D, Chang E Y, Shankar S, Kuchenbecker K M, Simon M D, Madhani H D, Narlikar G J and Al-Sady B 2011 Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly *Mol. Cell* **41** 67–81
- [8] Canzio D, Larson A and Narlikar G J 2014 Mechanisms of functional promiscuity by HP1 proteins *Trends Cell Biol.* **24** 377–86
- [9] Canzio D et al 2013 A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly *Nature* **496** 377–81
- [10] Chereji R V and Morozov A V 2014 Ubiquitous nucleosome unwrapping in the yeast genome *Proc. Natl Acad. Sci. USA* **111** 5236–41
- [11] Cheutin T, Gorski S A, May K M, Singh P B and Misteli T 2004 In vivo dynamics of Swi6 in yeast: evidence for a stochastic model of heterochromatin *Mol. Cell Biol.* **24** 3157–67
- [12] Chevereau G, Palmeira L, Thermes C, Arneodo A and Vaillant C 2009 Thermodynamics of intragenic nucleosome ordering *Phys. Rev. Lett.* **103** 188103
- [13] Darden T, York D and Pedersen L 1993 Particle mesh ewald - an  $n \log(n)$  method for ewald sums in large systems *J. Chem. Phys.* **98** 10089–92
- [14] David-Rus D, Mukhopadhyay S, Lebowitz J L and Sengupta A M 2009 Inheritance of epigenetic chromatin silencing *J. Theor. Biol.* **258** 112–20
- [15] Dodd I B, Micheelsen M A, Sneppen K and Thon G 2007 Theoretical analysis of epigenetic cell memory by nucleosome modification *Cell* **129** 813–22
- [16] Eissenberg J C and Reuter G 2009 Cellular mechanism for targeting heterochromatin formation in drosophila *Int. Rev. Cell Mol. Biol.* **273** 1–47
- [17] Filion G J et al 2010 Systematic protein location mapping reveals five principal chromatin types in drosophila cells *Cell* **143** 212–24
- [18] Fischle W, Tseng B S, Dormann H L, Ueberheide B M, Garcia B A, Shabanowitz J, Hunt D F, Funabiki H and Allis C D 2005 Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation *Nature* **438** 1116–22
- [19] Florescu A M, Schiessel H and Blossey R 2012 Kinetic control of nucleosome displacement by ISWI/ACF chromatin remodelers *Phys. Rev. Lett.* **109** 118103
- [20] Gabdank I, Barash D and Trifonov E N 2010 FineStr: a web server for single-base-resolution nucleosome positioning *Bioinformatics* **26** 845–6
- [21] Grewal S I and Jia S 2007 Heterochromatin revisited *Nat. Rev. Genet.* **8** 35–46
- [22] Hansen J C 2002 Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms and functions *Annu. Rev. Biophys. Biomol. Struct.* **31** 361–92
- [23] Hawkins G D, Cramer C J and Truhlar D G 1996 Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium *J. Phys. Chem.* **100** 19824–39

- [24] Hiragami K and Festeinstein R 2005 Heterochromatin protein 1: a pervasive controlling influence *Cell. Mol. Life Sci.: CMLS* **62** 2711–26
- [25] Hong T, Xing J, Li L and Tyson J J 2011 A mathematical model for the reciprocal differentiation of T helper 17 cells and induced regulatory T cells *PLoS Comput. Biol.* **7** e1002122
- [26] Hughes R M, Wiggins K R, Khorasanizadeh S and Waters M L 2007 Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect *Proc. Natl Acad. Sci. USA* **104** 11184–8
- [27] Jacobs S A, Fischle W and Khorasanizadeh S 2004 Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides *Methods Enzymol.* **376** 131–48
- [28] Jacobs S A, Taverna S D, Zhang Y, Briggs S D, Li J, Eissenberg J C, Allis C D and Khorasanizadeh S 2001 Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3 *EMBO J.* **20** 5232–41
- [29] Kaplan N *et al* 2009 The DNA-encoded nucleosome organization of a eukaryotic genome *Nature* **458** 362–6
- [30] Kepper N, Ettig R, Stehr R, Marnach S, Wedemann G and Rippe K 2011 Force spectroscopy of chromatin fibers: extracting energetics and structural information from Monte Carlo simulations *Biopolymers* **95** 435–47
- [31] Kepper N, Foethke D, Stehr R, Wedemann G and Rippe K 2008 Nucleosome geometry and internucleosomal interactions control the chromatin fiber conformation *Biophys. J.* **95** 3692–705
- [32] Kornberg R D 1974 Chromatin structure: a repeating unit of histones and DNA *Science* **184** 868–71
- [33] Kwon S H and Workman J L 2008 The heterochromatin protein 1 (HP1) family: put away a bias toward HP1 *Mol. Cells* **26** 217–27
- [34] Lando D Y and Teif V B 2000 Long-range interactions between ligands bound to a DNA molecule give rise to adsorption with the character of phase transition of the first kind *J. Biomol. Struct. Dyn.* **17** 903–11
- [35] Lubliner S and Segal E 2009 Modeling interactions between adjacent nucleosomes improves genome-wide predictions of nucleosome occupancy *Bioinformatics* **25** i348–55
- [36] Maison C and Almouzni G 2004 HP1 and the dynamics of heterochromatin maintenance *Nat. Rev. Mol. Cell Biol.* **5** 296–304
- [37] McGhee J D and von Hippel P H 1974 Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice *J. Mol. Biol.* **86** 469–89
- [38] Micheelsen M A, Mitarai N, Sneppen K and Dodd I B 2010 Theory for the stability and regulation of epigenetic landscapes *Phys. Biol.* **7** 026010
- [39] Mirny L A 2010 Nucleosome-mediated cooperativity between transcription factors *Proc. Natl Acad. Sci. USA* **107** 22534–9
- [40] Mobius W, Osberg B, Tsankov A M, Rando O J and Gerland U 2013 Toward a unified physical model of nucleosome patterns flanking transcription start sites *Proc. Natl Acad. Sci. USA* **110** 5719–24
- [41] 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.* **37** 4707–22
- [42] Müller-Ott K *et al* 2014 Specificity, propagation and memory of pericentric heterochromatin *Mol. Syst. Biol.* **10** 746
- [43] Müller K P, Erdel F, Caudron-Herger M, Marth C, Fodor B D, Richter M, Scaranaro M, Beaudouin J, Wachsmuth M and Rippe K 2009 Multiscale analysis of dynamics and interactions of heterochromatin protein 1 by fluorescence fluctuation microscopy *Biophys. J.* **97** 2876–85
- [44] Munari F *et al* 2012 Methylation of lysine 9 in histone H3 directs alternative modes of highly dynamic interaction of heterochromatin protein hHP1beta with the nucleosome *J. Biol. Chem.* **287** 33756–65
- [45] Nielsen A L, Oulad-Abdelghani M, Ortiz J A, Remboutsika E, Chambon P and Losson R 2001 Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins *Mol. Cell* **7** 729–39
- [46] Nielsen P R, Nietlispach D, Mott H R, Callaghan J, Bannister A, Kouzarides T, Murzin A G, Murzina N V and Laue E D 2002 Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9 *Nature* **416** 103–7
- [47] Olins A L and Olins D E 1974 Spheroid chromatin units (v bodies) *Science* **183** 330–2
- [48] Poland D 2001 Ligand binding distributions in nucleic acids *Biopolymers* **58** 477–90
- [49] Reynolds S M, Bilmes J A and Noble W S 2010 Learning a weighted sequence model of the nucleosome core and linker yields more accurate predictions in *Saccharomyces cerevisiae* and *Homo sapiens* *PLoS Comput. Biol.* **6** e1000834
- [50] Robinson P J and Rhodes D 2006 Structure of the ‘30 nm’ chromatin fibre: a key role for the linker histone *Curr. Opin. Struct. Biol.* **16** 336–43
- [51] Rosnoblet C, Vandamme J, Volkel P and Angrand P O 2011 Analysis of the human HP1 interactome reveals novel binding partners *Biochem. Biophys. Res. Commun.* **413** 206–11
- [52] Schiessel H 2003 The physics of chromatin *J. Phys.: Condens. Matter* **15** R699
- [53] Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R and Reuter G 2002 Central role of drosophila SU(VAR)3–9 in histone H3–K9 methylation and heterochromatic gene silencing *EMBO J.* **21** 1121–31
- [54] Schwab D J, Bruinsma R F, Rudnick J and Widom J 2008 Nucleosome switches *Phys. Rev. Lett.* **100** 228105
- [55] Sedighi M and Sengupta A M 2007 Epigenetic chromatin silencing: bistability and front propagation *Phys. Biol.* **4** 246–55
- [56] Segal E, Fondufe-Mittendorf Y, Chen L, Thastrom A, Field Y, Moore I K, Wang J P and Widom J 2006 A genomic code for nucleosome positioning *Nature* **442** 772–8
- [57] Sneppen K, Micheelsen M A and Dodd I B 2008 Ultrasensitive gene regulation by positive feedback loops in nucleosome modification *Mol. Syst. Biol.* **4** 182
- [58] Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu R M, Zhu P and Li G 2014 Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units *Science* **344** 376–80
- [59] Stein W D 1980 The epigenetic address: a model for embryological development *J. Theor. Biol.* **82** 663–77
- [60] Teif V B 2005 Ligand-induced DNA condensation: choosing the model *Biophys. J.* **89** 2574–87
- [61] Teif V B 2007 General transfer matrix formalism to calculate DNA-protein-drug binding in gene regulation: application to OR operator of phage  $\lambda$  *Nucleic Acids Res.* **35** e80
- [62] Teif V B 2010 Predicting gene-regulation functions: lessons from temperate bacteriophages *Biophys. J.* **98** 1247–56
- [63] Teif V B, Beshnova D A, Vainshtein Y, Marth C, Mallm J P, Höfer T and Rippe K 2014 Nucleosome repositioning links DNA (de)methylation and differential CTCF binding during stem cell development *Genome Res.* **24** 1285–95
- [64] Teif V B, Erdel F, Beshnova D A, Vainshtein Y, Mallm J P and Rippe K 2013 Taking into account nucleosomes for predicting gene expression *Methods* **62** 26–38
- [65] Teif V B, Ettig R and Rippe K 2010 A lattice model for transcription factor access to nucleosomal DNA *Biophys. J.* **99** 2597–607

- [66] Teif V B, Haroutiunian S G, Vorob'ev V I and Lando D Y 2002 Short-range interactions and size of ligands bound to DNA strongly influence adsorptive phase transition caused by long-range interactions *J. Biomol. Struct. Dyn.* **19** 1093–100
- [67] 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** 368–73
- [68] Teif V B and Rippe K 2009 Predicting nucleosome positions on the DNA: combining intrinsic sequence preferences and remodeler activities *Nucleic Acids Res.* **37** 5641–55
- [69] Teif V B and Rippe K 2010 Statistical-mechanical lattice models for protein-DNA binding in chromatin *J. Phys.: Condens. Matter* **22** 414105
- [70] Teif V B and Rippe K 2011 Nucleosome mediated crosstalk between transcription factors at eukaryotic enhancers *Phys. Biol.* **8** 044001
- [71] Thiru A, Nietlispach D, Mott H R, Okuwaki M, Lyon D, Nielsen P R, Hirshberg M, Verreault A, Murzina N V and Laue E D 2004 Structural basis of HP1/PXVXL motif peptide interactions and HP1 localisation to heterochromatin *EMBO J.* **23** 489–99
- [72] Tsui V and Case D A 2000 Theory and applications of the generalized Born solvation model in macromolecular simulations *Biopolymers* **56** 275–91
- [73] Vaillant C, Palmeira L, Chevereau G, Audit B, d'Aubenton-Carafa Y, Thermes C and Arneodo A 2010 A novel strategy of transcription regulation by intragenic nucleosome ordering *Genome Res.* **20** 59–67
- [74] van Holde K E 1989 *Chromatin* (New York: Springer)
- [75] Wang J P, Fondufe-Mittendorf Y, Xi L, Tsai G F, Segal E and Widom J 2008 Preferentially quantized linker DNA lengths in *Saccharomyces cerevisiae* *PLoS Comput. Biol.* **4** e1000175
- [76] Woodcock C L 2006 *Chromatin architecture Curr. Opin. Struct. Biol.* **16** 213
- [77] Yamamoto K and Sonoda M 2003 Self-interaction of heterochromatin protein 1 is required for direct binding to histone methyltransferase, SUV39H1 *Biochem. Biophys. Res. Commun.* **301** 287–92
- [78] Zhao-Wen D, Wei L, Ping X, Shuo-Xing D and Peng-Ye W 2011 Brownian dynamics simulation of the cross-talking effect among modified histones on conformations of nucleosomes *Chin. Phys. B* **19** 048701