View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by HAL-ENS-LYON

Dynamical DNA accessibility induced by chromatin remodeling and protein binding

F Montel, Cendrine Faivre-Moskalenko, Martin Castelnovo

To cite this version:

F Montel, Cendrine Faivre-Moskalenko, Martin Castelnovo. Dynamical DNA accessibility induced by chromatin remodeling and protein binding. Physical Review E : Statistical, Nonlinear, and Soft Matter Physics, American Physical Society, 2014, 90, pp.052717. <10.1103/PhysRevE.90.052717>. <ensl-01084485>

HAL Id: ensl-01084485

https://hal-ens-lyon.archives-ouvertes.fr/ensl-01084485

Submitted on 9 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dynamical DNA accessibility induced by chromatin remodeling and protein binding

F. Montel

Matière et Systèmes Complexes, Université Paris Diderot & CNRS (UMR 7057), 75205 Paris Cedex 13, France

C. Faivre-Moskalenko and M. Castelnovo*

Laboratoire de Physique, Ecole Normale Supérieure de Lyon, Université de Lyon and CNRS, 69364 Lyon Cedex, France

(Dated: today)

Abstract

Chromatin remodeling factors are enzymes being able to alter locally chromatin structure at the nucleosomal level and they actively participate in the regulation of gene expression. Using simple rules for individual nucleosome motion induced by a remodeling factor, we designed simulations of the remodeling of oligomeric chromatin, in order to address quantitatively collective effects in DNA accessibility upon nucleosome mobilization. Our results suggest that accessibility profiles are inhomogeneous thanks to borders effects like protein binding. Remarkably, we show that the accessibility lifetime of DNA sequence is roughly doubled in the vicinity of borders as compared to its value in bulk regions far from the borders. These results are quantitatively interpreted as resulting from the confined diffusion of a large nucleosome depleted region.

*Electronic address: martin.castelnovo@ens-lyon.fr

Introduction – Nucleosomes represent the first degree of organization of DNA within eukaryotic chromatin, in which the nucleic acid is wrapped roughly two times around a proteic core made of histones [1]. The wrapping geometry of the nucleosome constitutes on the one hand a physical barrier to DNA translation and replication, and on the other hand a way of regulating gene expression. One of the regulation strategy observed within cells is the use of remodeling factors [2]. These protein complexes are indeed able to move nucleosomes along DNA templates upon hydrolyzing ATP, and they are therefore able to generate transient windows of DNA accessibility for protein binding (transcription factors or activators for example). While chromatin remodelers have been identified almost ten years ago, their mechanisms of action are still subject of intense researches [3–5]. In particular, results about their collective effects in the modulation of nucleosome density at large scale are more scarce [6].

In this work, we show how the cooperative action of chromatin remodeling and protein binding may influence nucleosome density and DNA accessibility at large scale by mimicking the dynamics induced by remodeling factors from the swi2/snf2 family (for example RSC remodeling factor). Starting from well-positioned nucleosomal arrays, as the one observed in vivo after DNA replication, we use coarse-grained numerical simulations in order to quantify DNA accessibility induced by the presence of remodeling factors and protein binding, the number of nucleosomes being constant. The main input in the simulation is the typical mobilization rule for individual nucleosome. These rules are based on recent Atomic Force Microscopy observations on small oligo-nucleosomal templates (di- and tri-nucleosomes) remodeled by [7]. In this work, it has been observed that an individual remodeling event induces a displacement of nucleosome from its initial position towards either another nucleosome or the end of DNA template without any observable intermediates. Using this simple rule as the basic ingredient of the simulations, it is possible to investigate the structuration of nucleosomal density in the presence of remodeling factors and protein binding, and consequently the modulation of DNA accessibility. Our simulations identify the following features: (i) the remodeling by complexes of swi2/snf2 family of nucleosomal array induces the coexistence of high and low nucleosomal density regions; (ii) the lifetime of DNA accessibility is strongly inhomogeneous along DNA template, although no sequence effect is explicitly considered ; (iii) the simultaneous presence of remodelers and protein binding allows to modulate spatially the DNA accessibility.

Simulation rules – We consider an oligomeric array composed of N nucleosomes initially

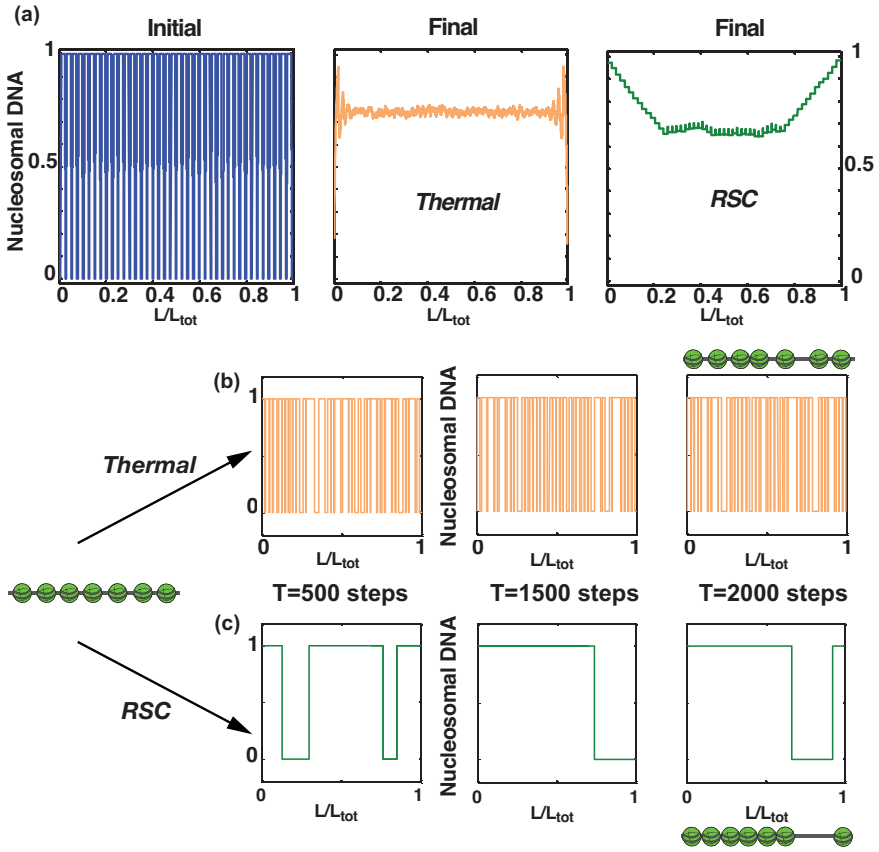

FIG. 1: Comparison of nucleosomal profiles after thermal and RSC mobilization. *(a)* Average nucleosomal DNA, in the common starting configuration (left), after 2000 steps of thermal (center) and RSC (right) mobilization. Snapshots of nucleosomal DNA at different times for thermal *(b)* and for RSC *(c)* mobilization. Parameters for this figure are $N = 20, L_k = 50bp, 2000$ runs. Cartoons have illustrative purposes for typical stationary configurations for thermal and RSC mobilization, but not at proper scale.

ordered with a linker length L_k , therebye imposing a fixed nucleosome density. The total length of DNA is therefore $L_{tot} = NL_c + (N+1)L_k$, where L_c is the DNA length covered by a single nucleosome. For the sake of simplicity, we choose the canonical value $L_c = 147bp$. Choosing linker lengths between $20bp$ and $50bp$, the nucleosome occupancy is modulated between 75% and 88% [8]. This choice of fixed density allows to focus on events occuring between DNA replication phases, unlike Parmar *et al.* who investigated the evolution of overall nucleosome density upon nucleosome deposition starting from naked DNA [6]. The general algorithm used in the simulation is the following: at each simulation step k nucleosomes (typically $k = N$) are chosen randomly and a local motion rule (to be described below) is applied successively on these k nucleosomes. Only moves avoiding sterical clash

are effectively performed. This elementary step is then iterated. Elementary runs are then reproduced in order to reach statistical significance of the results. This versatile design of the simulation allows addressing the comparison of different local motion rules. The results presented in this work were obtained by two different elementary motion rules: *(i) RSC*-like mobilization (swi2/snf2 family), for which the nucleosome is moved processively with equal probability to its right or to its left until it reaches either another nucleosome or the end of DNA template, and *(ii)thermal*-like mobilization, for which the nucleosome travels a fixed distance to its left or to its right with equal probability. For this last mobilization scenario, the system is then equivalent to a Tonk's gas. In our study, we chose a thermal jump size of 10bp, which is consistent with rotational positioning of nucleosomes [9]. The elementary simulation step during which k motion of nucleosomes are tried is considered as our time unit. Therefore our results will be valid for time larger than this elementary time scale. Following Narlikar *et al.* [10], we consider that the rate limiting step associated to nucleosome remodeling is the sliding of the nucleosome. Note that the quantitative comparison of time scales for both mobilization scenarios is not addressed within the present work since it is strongly correlated with the rate of individual mobilization events. This question goes beyond the scope of our investigations. For the sake of simplicity, we chose the same rate for RSC and thermal mobilization since it allows to highlight differences in nucleosomal patterns. Additionally, within our simulation design, detailed balance for the different rates is expected to breakdown, because each individual motion of RSC-mobilized nucleosomes is performed thanks to the consumption of external energy (ATP). This external forcing of nucleosome motion breaks the detailed balance, as it has been already suggested and discussed in the literature in the context of molecular motors or self-assembly modeling [11–13].

In order to analyze the evolution of DNA accessibility during the simulation, we used probes of variable size (typically $10bp$) located at the dyad of every nucleosomes in the starting configuration. Each probe will produce a unit value whenever it is not covered by a nucleosome, and the probe status (free or occupied) is updated at each simulation step. This permits to monitor precisely the evolution of accessibility. For the investigation of protein binding on oligo-chromatin template, we chose the simplest rule for binding in order to analyze the asymptotic stationary state of the simulations: after an initial period of RSC-mobilization in order to reach stationary state before binding, the protein will bind at once and irreversibly at the first simulation step when the site is free of nucleosome. The simulation is then continued for sufficient time in order to reach a new stationary state in

the presence of the protein.

Results – Using the simulations described in the previous section, we first compare nucleosome occupancy within an oligomeric array when nucleosomes are mobilized according to different dynamical scenarii. Following most works dealing with nucleosome positioning, we define the nucleosome occupancy, or equivalently nucleosomal DNA, as a local variable whose value is 1 if the base pair is covered by a nucleosome and 0 otherwise. In particular, we compare remodeling patterns generated by RSC with patterns observed upon thermal mobilization. The result of this comparison is shown in figure 1. The average nucleosome occupancy for the two mobilization scenarios exhibit very different traits. Thermal mobilization of nucleosomes exhibit well-known "statistical positioning effects", i.e. there are damped oscillations of occupancy close to the DNA boundaries [5, 14, 15].

On the contrary, RSC patterns shows accumulation of nucleosomes close to boundaries when averaged over many simulation runs and almost constant nucleosome occupancy in the bulk region far from boundaries. Individual snapshots exhibit a 1D dynamical close-compact ordering. Indeed, the remodeling leads asymptotically to the formation of two large clusters separated by a single mobile nucleosome depleted region of size $(N+1)L_k$. As a consequence, the number of available nucleosome configurations in the large time limit is quantized: there are exactly $N+1$ distinct configurations of nuclesomes. Another consequence at the level of average nucleosomal occupancy is the fact that this last quantity exhibits a train of regions of constant occupancy of length L_c, and of vertical jumps of size $1/(N+1)$. Similarly, simple arguments allows to show that the range of the nucleosome accumulation seen on the average nucleosome occupancy is also estimated by $L_{border} = (N+1)L_k$.

Next we investigate the influence of number of nucleosomes and linker length on the global shape of average nucleosome occupancy for both thermal and RSC mobilization. The results are shown in figure 2. It is observed that average nucleosome occupancy is lower in the plateau region for RSC than for thermal mobilization, reflecting the presence of the large nucleosome depleted region identified in the previous section for RSC mobilization. In the case of thermal mobilization, the average nucleosome occupancy far from the boundaries is simply estimated by the asymptotic value for an infinite system, namely $\rho^{(\infty)}_{nuc,T} = L_c/(L_c+L_k)$, which is independent in number of nucleosomes, up to leading order in $1/N$. Similarly in the case of RSC mobilization, the average nucleosome occupancy is estimated by considering the different configurations of the nucleosome depleted region far from the boundaries. The value is therefore estimated as $\rho^{(\infty)}_{nuc,RSC} = 1 - L_k/L_c$. Comparison of these two formula

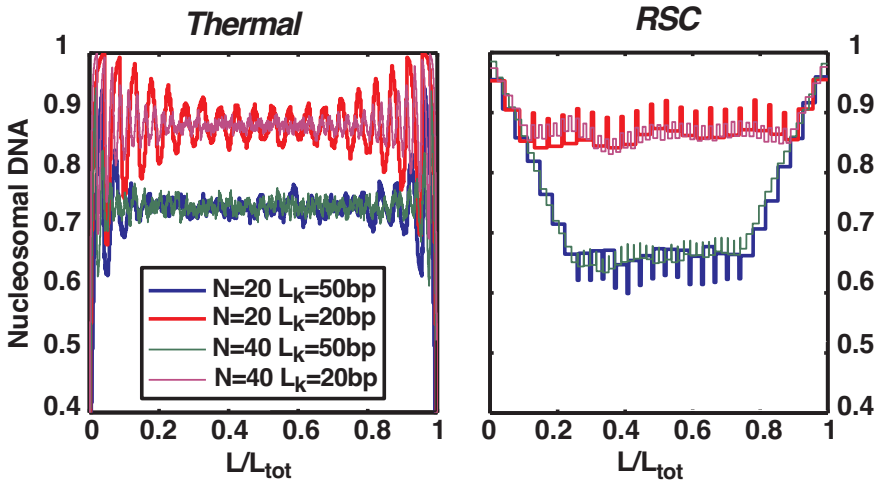

FIG. 2: (Color online) Dependence of average nucleosomal profile in N and L_k for thermal *(left panel)* and RSC *(right panel)* mobilization. The four curves on each panel are associated with parameters: $N = 20$ and $L_k = 50bp$ *(dark thick line)*, $N = 20$ and $L_k = 20bp$ *(light thick line)*, $N = 40$ and $L_k = 50bp$ *(bottom thin line)*, $N = 50$ and $L_k = 20bp$ *(upper thin line)* .Other common parameters: 2000 steps, 2000 runs.

implies that the bulk average nucleosome occupancy is always larger in thermal mobilization as compared to RSC mobilization. The two previous estimations are in good agreement with the simulation results. Note that in the case of RSC mobilization, the expression for $\rho^{(\infty)}_{nuc,RSC}$ is valid whenever $L_k < L_c$, which implies the presence of a finite size plateau. The two previous estimations for the bulk average nucleosome occupancy are valid up to $1/N$ order.

The range of nucleosome accumulation for RSC mobilization at the boundaries is estimated by the size of the nucleosome depleted region. In the rescaled units of figure 2, this range is therefore $L_{border}/L_{tot} = L_k/(L_c + L_k)$. This estimation is again in good agreement with the simulations. The range of boundaries effect are opposite for thermal mobilization : it decreases with the linker length. This effect has already been discussed in the literature [5].

The lifetime of DNA accessibility as measured by probes of given sizes along the nucleosomal array provides informations about the dynamics of remodeling in the stationary state. The accessibility lifetimes is shown for RSC mobilization in figure 3. It is observed globally that the average lifetime of DNA accessibility upon RSC mobilization increases significantly with the number of nucleosomes. As a qualitative comparison, the accessibility is roughly constant and homogeneous for thermal mobilization. As it was already noted previously in

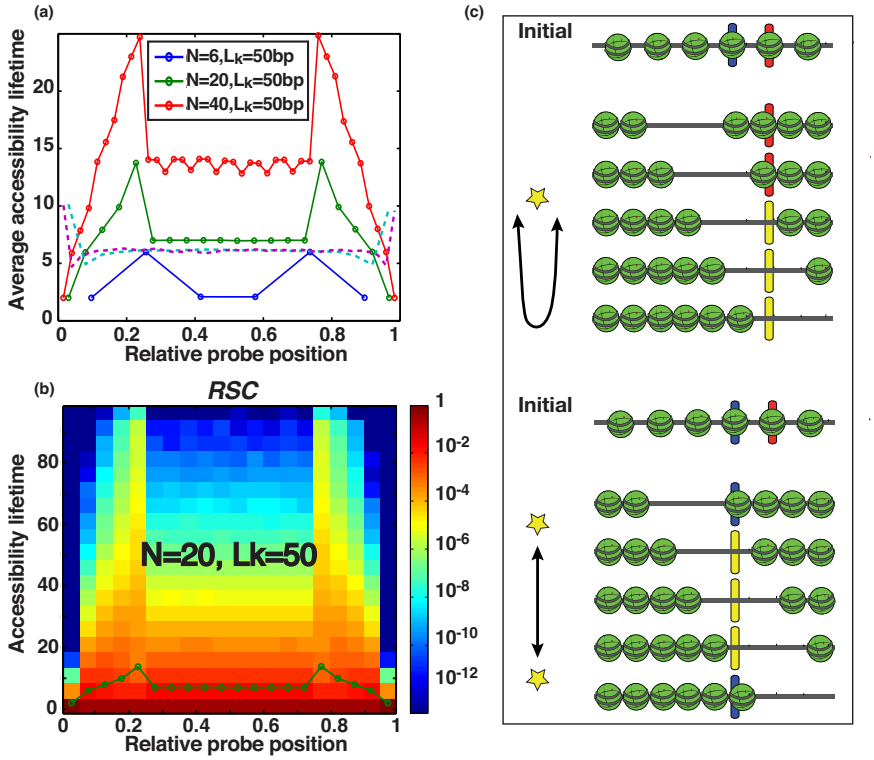

FIG. 3: DNA accessibility lifetime upon RSC mobilization at discrete sites hidden by nucleosomes in the starting configuration. *(a)* Average accessibility lifetime for $N = 6$ *(lower blue line with circles)*, $N = 20$ *(middle green line with circles)*, $N = 40$ *(upper red line with circles)*. The dotted lines correspond to the average accessibility lifetime for thermal mobilization for $N = 20$ *(light cyan)* and $N = 40$ *(dark purple)*. Other common parameters: $L_k = 50bp, 2000$ steps, 2000 runs. *(b)* Accessibility lifetime distribution in logarithmic color scale for $N = 20$. The green curve correspond to the average accessibility lifetime. *(c)* Cartoons for $N = 6$ showing only one entry/exit *(yellow star)* for accessibility event *(top, red light stick)*, and two entries/exits *(down, dark blue stick)*. The configurations allowing accessibility of sites are highlighted by the change of color of the sticks (light yellow sticks).

the introduction section of this work, the quantitative comparison of accessibility time scales is not possible within this work since it depends on the particular choice of elementary mobilization rates. Interestingly, RSC mobilization shows another striking feature, namely the relative inhomogeneity of lifetime distributions close to the boundaries of DNA template: indeed, the border regions identified in the average nucleosome occupancy profiles shows large variations of lifetime distribution. In particular, the average lifetime has a low value close to DNA boundary, reflecting the accumulation of nucleosomes at the boundaries, and

it increases rapidly towards the bulk of nucleosomal array. Remarkably, it reaches a maximum at roughly $(N+1)L_k$ then it decreases sharply towards a constant value for the bulk sites. In other words, there is a highly localized region of DNA accessibility enhancement, in which the average lifetime is roughly doubled as compared to bulk accessibility. This trend is also observed for the distribution of lifetime accessibility, as it is evidenced in figure 3b.

This result is understood by analyzing the motion of the depleted region with respect to the considered site. When a given site starts to be accessible, the depleted region overlap this site over the length a of the probe. The accessibility event stops whenever this overlapping vanishes by diffusion of the depleted region. For bulk sites, accessibility events can be nucleated and finished at similar or opposite location, as it is illustrated in figure 3c. On the contrary, for sites close to the DNA boundaries, accessibility events are necessarily nucleated and finished at the same locations. As a consequence, accessibility events might last longer for these border sites than for bulk sites. The boundary between these two regions is extremely sharp, as it is reflected by the variation of average lifetime accessibility.

It is possible to estimate the shape of the average accessibility lifetime profile by considering the mean first passage time of random walker on a line with two adsorbing boundary conditions. It is a standard result that the mean first passage time associated with a 1D random walker of diffusion coefficient D , starting at position x_0 and confined between two adsorbing walls located at $x = 0$ and $x = L$ is $\tau_{MFPT} = \frac{x_0(L-x_0)}{2D}$ [16, 17]. For the bulk sites, the average accessibility lifetime scales therefore as $t_{bulk} \sim L_c((N+1)L_k - L_c)/2D_{RSC}$, with $D_{RSC} = L_c^2/\tau_{RSC}$ with τ_{RSC} typical mobilization time by RSC. This time is independent on the position in the template. In the region close the border of the template at position X, the average accessibility lifetime is the mean first passage time of random walk between one adsorbing and one reflecting wall roughly separated by X, and starting at a distance L_c from the adsorbing wall. Using mirror symmetry in order to solve this problem, this time is equivalent to the mean first passage time of random walker between two adsorbing walls separated by $2X$. Therefore it is then given by $t_{border} \sim L_c(2X - L_c)/2D_{RSC}$. These formulas explain quantitatively the emergence of a localized region with greater average accessibility than the bulk for $(N+1)L_k/2 < X < (N+1)L_k$. In particular the maximal average accessibility lifetime is predicted to be roughly twice the lifetime for bulk sites, in agreements with the simulation results.

Finally, in order to illustrate the effect of protein binding on DNA accessibility, we allowed some sites to be bound by proteins of fixed size. This binding has two consequences on DNA

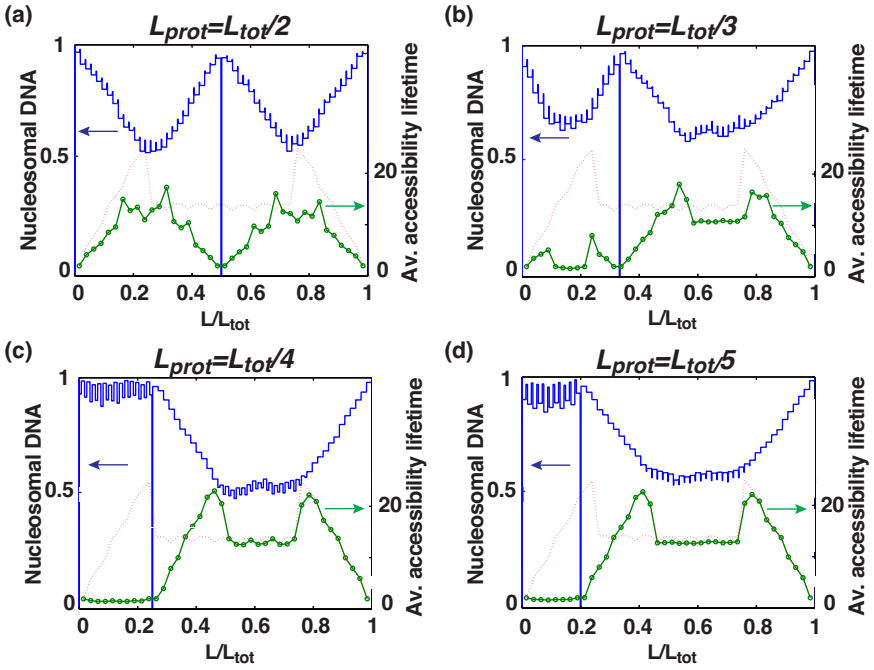

FIG. 4: Comparison of nucleosomal DNA and average accessibility lifetime after protein binding at different locations L_{prot}. *(a)* $L_{prot} = L_{tot}/2$, *(b)* $L_{prot} = L_{tot}/3$, *(c)* $L_{prot} = L_{tot}/4$, *(d)* $L_{prot} = L_{tot}/5$. For each panel, the top *(blue)* curve is the nucleosomal DNA, while the lower *(green)* curve is the average accessibility lifetime. Thin *(red)* dotted lines represent the average accessibility lifetime before binding. Other common parameters: $N = 40, L_k = 50bp, 4000$ steps, 500 runs.

accessibility. First, these sites act as effective new boundaries for DNA, and remodeled nucleosomes cannot cross them. Second, the number of nucleosomes on each side of the binding sites remain constant after the binding. The average nucleosomal profiles and the average lifetime of accessibility for all the other sites on DNA template after protein binding are shown in figure 4. The location of the binding site has a strong influence on both average nucleosomal occupancy and DNA accessibility. When this site is close to DNA boundaries, the protein binding traps closely packed nucleosomes in this extremal region, and therefore the DNA accessibility is extremely low, while the initial nucleosome depleted region is trapped on the other side of the binding site. When the protein binds at sites within the bulk nucleosomal array, DNA accessibility is changed on both sides, but in less dramatic way. These observations show that the combination of RSC remodeling and protein binding is an efficient way of separating spatially regions of high and low DNA accessibility.

Discussion – The coarse-grained simulations presented in this work show the typical DNA accessibility within an oligomeric array of nucleosomes mobilized by the remodeling factor

RSC. The elementary rule for RSC action produces on the one hand a large diffusive nucleosome depleted region, and on the other hand introduces strong inhomogeneities in the average accessibility lifetimes. Interestingly, recent experiments done by Tolstorukov *et al.* demonstrate that the induction of swi/snf remodeler induces mean nucleosome occupancy profiles very similar to the one produced by our simulation [18]: the nucleosome occupancy increases as the so-called transcription start site (TSS) and its particular nucleosome architecture (NFR) is approached for wild-type conditions, while this increase is strongly reduced if remodelers are inhibited. In addition, we predict within our simulation a localized region for which the accessibility lifetime is roughly doubled as compared to the other regions. Interestingly, this last feature is quite a general result from 1D-diffusion of an accessibility windows confined between reflecting boundaries. The only requirement in order to observe a region of accessibility enhancement close to reflecting boundaries is that the size of the accessibility windows W is larger than the size of the probe a. As a consequence, the accessibility lifetime of sites located in the range $[W/2; W]$ are larger than bulk sites in the range $[W + a; L - W - a]$.

In order to progress further towards the global understanding of DNA accessibility within chromatin, both sequence effects and protein binding kinetics are to be taken into account in future works. The former effect might essentially influence the diffusion of the nucleosome depleted region on a bumpy energetic landscape, while the latter might influence the lifetime of accessibility patterns. The present results strongly suggest that simulations of the dynamics are efficient tools in order to investigate such complex molecular scenarios.

---

[1] K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond., Nature **389**, 251 (1997).

[2] L. Mohrmann and C. P. Verrijzer, Biochim. Biophys. **1681**, 59 (2005).

[3] B. Cairns, Nature **461**, 193 (2009).

[4] V. B. Teif and K. Rippe, Nuc. Ac. Research **37**, 5641 (2009).

[5] W. Mobius, B. Osberg, A. Tsankov, O. Rando, and U. Gerland, Proc. Nat. Acad. Sci. USA **110**, 5719 (2013).

[6] J. Parmar, J. Marko, and R. Padinhateeri, Nuc. Ac. Research **42**, 128 (2014).

[7] F. Montel, M. Castelnovo, H. Menoni, D. Angelov, S. Dimitrov, and C. Faivre-Moskalenko, Nuc. Ac. Research **39**, 2571 (2011).

[8] J. Wang, Y. Fondufe-Mittendorf, L. Xi, G. Tsai, E. Segal, and J. Widom, Plos Comp. Biol. **4**, e1000175 (2008).

[9] H. Schiessel, J. Phys.: Condens. Matter **15**, R699 (2003).

[10] G. Narlikar, M. Phelan, and R. Kingston, Mol. Cell **8**, 1219 (2001).

[11] F. Julicher, J. Prost, and A. Ajdari, Rev. Mod. Phys. **69**, 1269 (1997).

[12] M. Klein, Phys. Rev. **97**, 1446 (1955).

[13] M. Castelnovo, T. Verdier, and L. Foret, Europhys. Lett. **105**, 28006 (2014).

[14] L. Tonks, Phys. Rev. **50**, 955 (1936).

[15] R. Kornberg and L. Stryer, Nuc. Ac. Research **16**, 6677 (1988).

[16] C. Gardiner, *Handbook of stochastic methods: for physics, chemistry and natural sciences* (Springer Verlag, 1997).

[17] M. Gitterman, Phys. Rev. E **62**, 6065 (2000).

[18] M. Tolstorukov, C. Sansam, P. Lu, E. Koellhoffer, K. Helming, B. Alver, E. Tillman, J. Evans, B. Wilson, P. Park, , and C. Roberts, Proc. Nat. Acad. Sci. USA **110**, 10165 (2013).