EUROPHYSICS LETTERS 1 April 2002

Europhys. Lett., **58** (1), pp. 140–146 (2002)

How short-ranged electrostatics controls the chromatin structure on much larger scales

H. Schiessel

Max-Planck-Institut für Polymerforschung, Theory Group
P.O. Box 3148, D-55021 Mainz, Germany

(received 3 October 2001; accepted in final form 14 January 2002)

PACS. 87.15.-v – Biomolecules: structure and physical properties.
PACS. 36.20.Ey – Conformation (statistics and dynamics).

Abstract. – We propose that the degree of swelling of the 30 nm chromatin fiber (a "measure" of its transcriptional activity) is mainly determined by the short-ranged electrostatical interaction between different sections of the "folded" DNA chain. These sections constitute only a small fraction of the chain and they are located close to the entry-exit points of the DNA chain at the nucleosome core particles. We present a model that allows to estimate the degree of swelling of chromatin fibers as a function of salt concentration, charge density of the strands, etc. Different mechanisms by which the state of chromatin can be controlled *in vitro* and *in vivo* are discussed.

DNA in eucaryotic cells is organized within a protein-DNA complex known as chromatin. In this way plant and animal genomes are packed into volumes whose linear dimensions are many orders of magnitude smaller than their contour lengths. For instance, the human genome is made up of billions of base pairs (bp) corresponding to about one meter of DNA chains. These highly charged and hard-to-bend polymers are condensed into complexes that have a characteristic size of a micron and therefore fit into the cell nucleus. *At the same time* it is of vital importance that a fraction of the genetic code stored within these tight complexes is accessible as there are gene regulatory proteins that bind to specific sequences, RNA polymerases (rather bulky protein complexes) that need to gain access to whole genes during their transcription etc. How DNA is "folded" within chromatin and how it can be "unfolded" for, say, transcription purposes is still poorly understood.

The primary structure of chromatin is known in great detail from X-ray studies [1]. The basic unit is the nucleosome consisting of the core particle and the linker DNA (of typical length 60 bp) that connects to the neighboring core particle. The resulting structure is a beads-on-chain necklace (10 nm fiber). The core particle consists of 147 bp DNA wrapped in 1 and 3/4 left-helical turns around a globular octamer of cationic proteins (two molecules each of the so-called core histones H2A, H2B, H3 and H4) —forming a squat cylinder with a radius of about 5 nm and height of about 6 nm. The higher-order secondary and tertiary structures on scales from 10 nm up to a micron are still a matter of controversy [2]. It is well known that the bead-on-a-chain folds into a thicker fiber with a diameter of roughly 30 nm but it is

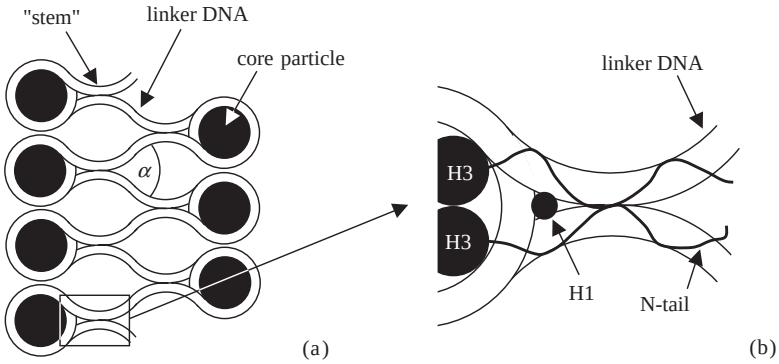

Fig. 1 – (a) Schematic view of a section of the 30 nm chromatin fiber. For simplicity, the fiber is shown as a two-dimensional zig-zag fiber. The histone octamers are displayed as black circular disks. The DNA (white) consists of wrapped parts, "stem" sections and sections that link neighboring nucleosomes. (b) Enlarged view of the stem region. As shown schematically, the stem section of the DNA interacts with cationic N-tails of the core histones and with the linker histone H1 (see text for details).

still not clear how the neighboring beads are arranged in this fiber with respect to each other and where the linker is located. In the solenoid models [3] it is assumed that the chain of nucleosomes forms a helical structure with the linker DNA being bent whereas the zig-zag- or crossed-linker models [4] postulate straight linkers that connect core particles that are located at opposite sites of the fiber. An example of that kind of structure, the two-dimensional zig-zag fiber, is shown in fig. 1(a). In general, the fiber is three-dimensional.

The difficulties in determining the structure of the 30 nm fibers are due to the lack of reliable experimental methods. Electron cryomicroscopy allows to visualize fibers *in vitro* but the fiber geometry can only be detected at ionic strengths well beyond physiological conditions where the structures are much more open [5]. Fibers at such low ionic strengths show straight linkers, a fact that supports the second class of the above-mentioned chromatin models —at least for fibers at these unphysiological conditions. At higher ionic strengths, however, the fiber is so dense that their internal structure remains obscure.

A different approach to obtain structural information on the chromatin fiber is to stretch an individual fiber using micromanipulation devices as recently achieved by Cui and Bustamante [6]. They found chromatin fibers to be extremely soft with respect to their longitudinal extension (compared to naked DNA). Katritch *et al.* [7] performed computer simulations of crossed-linker fibers and showed that this model is capable of reproducing the experimental data quite well for several sets of values of the "free" parameters. Schiessel *et al.* [8] gave an analytical treatment of the model and —based on an optimization criteria (see below)— they were able to arrive at reasonable predictions of the mechanical properties of the chromatin fiber *without the use of any adjustable parameters* (an extension of these results has been recently given by Ben-Haïm *et al.* [9]). These studies do not invalidate the solenoid model but demonstrate that the crossed-linker model successfully describes several features of the 30 nm fiber.

The main idea of the crossed-linker model as introduced by Woodcock and coworkers [4] is that the three-dimensional structure of the 30 nm chromatin fiber is determined by two quantities only: the entry-exit angle α of the DNA at the core particle (cf. fig. 1(a)) and the dihedral angle ϕ that describes the rotational setting of neighboring nucleosomes. The

entry-exit angle α follows from the local properties of the nucleosome and should have a constant value throughout the fiber as long as there are homogeneous conditions. The second angle, ϕ, surprisingly turns out to be also fairly constant throughout the fiber, a fact that is corroborated by the phenomenon of the so-called quantization of the linker length (cf. ref. [10] for a discussion of this effect). Therefore it seems to be a reasonable approximation to consider both α and ϕ to be constant throughout the fiber (as well as the linker length b that is also fairly constant for a given chromatin complex [10]). The geometrical properties of the fiber, *e.g.*, its radius R and the linear density Λ of nucleosomes along the fiber, are then a function of the two angles α and ϕ only. In ref. [8] we termed this geometrical model, therefore, the *two-angle model*. A special case is $\phi = 180°$ leading to zig-zag fibers as the one depicted in fig. 1(a).

In our study of the geometrical and mechanical properties of the two-angle model [8] we suggested that the values of α and ϕ of the "native" 30 nm chromatin fiber in its "silenced" form (no transcriptional activity) are chosen in such a way that the three-dimensional *density* of the fiber (defined by $\Lambda/\pi R^2$) and its "*accessibility*" are optimized. Accessibility means here how rapid the "silenced" fiber opens up when its geometrical properties (characterized by the two angles ϕ and α) are changed for the purpose of, say, transcription. As long as the positions of the nucleosomes at the DNA are fixed (as assumed here), ϕ and b are always constant and hence it is the angle α with which the fiber geometry, *i.e.* its degree of swelling, is controlled. High accessibility means that the fiber opens up strongly with an increasing value of α, leading to a strong reduction of the nucleosome line density. Hence our measure for accessibility, as introduced in ref. [8], is the quantity $-\partial\Lambda/\partial\alpha$.

The purpose of this paper is to study how the entry-exit angle α of the DNA at the nucleoseomes can be controlled via electrostatics. It can been seen in the cryo EM studies [5] that the fibers open up and become therefore more accessible when the ionic strength is reduced and that this opening is directly linked to an increase in α. It was suggested that via other mechanisms (for instance, the acetylation of so-called histone tails [11], as explained in more detail below) the angle α and therefore the degree of swelling can be changed for a given section of the fiber and that this constitutes a biochemical means to control the transcriptional activity of genes.

Whereas the X-ray studies of the core particle [1] allow a detailed knowledge of the wrapped part of DNA, it does not give insight into the conformational properties of the entering and exiting strands since the core particles were constituted from 146 bp DNA only. One has therefore to refer to the electron cryomicrographs. In these micrographs it can be clearly seen that 10 nm stretches of the entering and exiting DNA strands are glued together forming a unique "stem motif" [5] (cf. also fig. 1(a)). The gluing of the two equally charged chains is accomplished —amongst other things— via a special cationic protein, the so-called linker histone H1 as shown schematically in fig. 1(b).

At physiological concentrations (roughly 100 mM salt corresponding to a screening length of about 10 Å) the electrostatics is essentially short-ranged (note that the diameter of DNA is 20 Å). It seems therefore reasonable to assume that α is set within the small region where the two linker DNA are in close contact, *i.e.* in the stem region [12]. This value of α in turn controls the large-scale secondary structure of chromatin, the 30 nm fiber. To mimic this situation, we assume in the following two parallel DNA strands that are held together tightly at $y = 0$ for $x \leq 0$ and that are free for $x > 0$, cf. fig. 2. Because of their mutual electrostatic repulsion the two strands bend away from each other. When the two strands are far enough from each other their interaction is screened so that they asymptotically approach straight lines (neglecting thermal fluctuations as is appropriate for the length scales under consideration) which define the opening angle α as indicated in fig. 2.

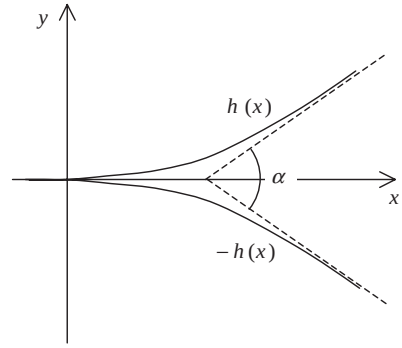

Fig. 2 – Idealized model for the entry-exit region of the DNA at the nucleosome. $h(x)$ and $-h(x)$ describe the conformations of the entering and exiting DNA sections, respectively. The asymptotic slope at large x determines the entry-exit angle α and, in turn, the overall conformation of the chromatin fiber (as depicted in fig. 1(a)).

We describe the conformation of the upper DNA chain by the height function $h(x)$. By symmetry the position of the lower strand is then given by $-h(x)$. From the above given considerations two boundary conditions at $x = 0$ follow:

$$h(0) = h'(0) = 0. \tag{1}$$

The entry-exit angle α is related to the slope of $h(x)$ at infinity as follows:

$$\tan(\alpha/2) = h'(\infty). \tag{2}$$

We model the two DNA chains as semiflexible polymers with persistence length l_{P} and line-charge density λ immersed in a salt solution characterized by the Bjerrum length $l_{\mathrm{B}} \equiv e^2/\epsilon k_{\mathrm{B}}T$ and the Debye screening length $\kappa^{-1} = (8\pi c_{\mathrm{s}} l_{\mathrm{B}})^{-1/2}$ (ϵ dielectric constant of the solvent, T temperature, k_{B} Boltzmann constant and c_{s} salt concentration). We assume that the charges on the DNA strands interact via a screened electrostatic potential. The free energy of the system is then given by

$$\frac{F}{k_{\mathrm{B}}T} \simeq \int_0^\infty \mathrm{d}x \left[l_{\mathrm{P}} \left(\frac{\mathrm{d}^2 h}{\mathrm{d}x^2} \right)^2 + 2 l_{\mathrm{B}} \lambda^2 K_0(2\kappa h(x)) \right]. \tag{3}$$

The first term in the integral accounts for the bending of the *two* DNA strands and the second one describes the interaction between the two chains ($K_0(x)$ being the 0th order modified Bessel function). Here the interaction of a given charge on one chain with all the charges on the other chain is approximated by the interaction of this charge with a *straight* chain at the distance $2h$. As shown below, the integral, eq. (3), is a good approximation as long as the value of α that follows from its minimization is sufficiently small. The conformation of the upper chain, $h(x)$, is then the solution of the corresponding Euler-Lagrange equation (using $K_0'(x) = -K_1(x)$)

$$l_{\mathrm{P}} \frac{\mathrm{d}^4 h}{\mathrm{d}x^4} - 2 l_{\mathrm{B}} \lambda^2 \kappa K_1(2\kappa h) = 0, \tag{4}$$

together with four boundary conditions: Two are given at the origin, eq. (1), and two follow from the condition of straight "linkers" at infinity:

$$h''(\infty) = h'''(\infty) = 0. \tag{5}$$

Defining $\tilde{h} = 2\kappa h$ and introducing the dimensionless quantity $\tilde{x} = (4l_{\rm B}\lambda^2\kappa^2 l_{\rm P}^{-1})^{1/4}x$, eq. (4) can be rewritten as

$$\frac{{\rm d}^4\tilde{h}}{{\rm d}\tilde{x}^4} - K_1(\tilde{h}) = 0. \tag{6}$$

We denote the dimensionless asymptotic slope by $c_0 = {\rm d}\tilde{h}/{\rm d}\tilde{x}|_{\tilde{x}=\infty}$. It follows immediately that $\tan(\alpha/2)$ is given by

$$\tan(\alpha/2) = h'(\infty) = \frac{c_0}{\sqrt{2}}\left(\frac{l_{\rm B}}{l_{\rm P}}\right)^{1/4}\left(\frac{\lambda}{\kappa}\right)^{1/2}. \tag{7}$$

Equation (6) can be solved asymptotically for small values of $\tilde{h}$, where $K_1(\tilde{h}) \simeq 1/\tilde{h}$. In leading order the asymptotic solution is given by $\tilde{h}(\tilde{x}) \simeq \tilde{h}_0(\tilde{x}) = \tilde{x}^2\sqrt{(-\ln\tilde{x})}$. This solution fulfills the boundary conditions at the origin, eq. (1). The boundary conditions at infinity, eq. (5), are only marginally affected by the solution for small values of $\tilde{x}$ as can be seen by adding terms like $a\tilde{x}^2$ and $b\tilde{x}^3$. We can use this approximate solution to give a rough estimate for c_0. $\tilde{h}_0(\tilde{x})$ reaches its maximal slope 0.56 at $\tilde{x} \simeq 0.41$. For larger values of $\tilde{x}$ the approximate solution $\tilde{h}_0(\tilde{x})$ becomes more and more unreliable. This asymptotic analysis indicates that c_0 is of the order one [13].

A consequence of the rescaling $h \to \tilde{h}$ and $x \to \tilde{x}$ is that both the bending and the electrostatic contribution to the free energy, eq. (3), scale as $l_{\rm B}^{3/4}\lambda^{3/2}l_{\rm P}^{1/4}\kappa^{-1/2}$. Note that this scaling behavior (and likewise all the other above given results) only holds as long as the approximations leading to eq. (3) are reasonable. The key approximation is to replace the original curved section by a straight chain that lies parallel to the x-axis (see above). In order to find the range of validity of this approximation, we consider now the first-order correction in $h'(x)$: The interaction between a charge on the lower chain at $(x, -h(x))$ and the upper chain is now estimated by replacing that chain by its tangent at the point $(x, h(x))$. This means nothing but replacing that chain by a straight chain at distance $d(x) = 2h(x)\cos(\alpha(x)/2)$ with $\tan(\alpha(x)/2) = h'(x)$. We now use the fact that most of the bending and electrostatic energies are localized close to the origin. Thus we use the asymptotic solution for small x, *i.e.* we replace $h(x)$ in eq. (3) by $h_0(x) \approx (\gamma x)^2/2\kappa$, with $\gamma = (4l_{\rm B}\lambda^2\kappa^2/l_{\rm P})^{1/4}$. Furthermore, we replace $K_0(2\kappa h(x))$ by $K_0(2\kappa d(x))$ and calculate the free energy by integrating from $x = 0$ to $x = \gamma^{-1}$. In the zeroth order in $h'(x)$ we recover the above-given scaling behavior of the free energy, especially the electrostatic energy is given by $E_{\rm el}^{(0)}/kT \approx 4^{3/4}l_{\rm B}^{3/4}\lambda^{3/2}l_{\rm P}^{1/4}\kappa^{-1/2}$. The first-order correction in $h'(x)$ leads to the following additional electrostatic repulsion: $E_{\rm el}^{(1)}/kT \approx 2^{1/2}l_{\rm B}^{5/4}\lambda^{5/2}/\kappa^{3/2}l_{\rm P}^{1/4}/3$. The approximation made in eq. (3) is good as long as $E_{\rm el}^{(1)}/E_{\rm el}^{(0)} = l_{\rm B}^{1/2}\lambda/2\kappa l_{\rm P}^{1/2} \approx \tan^2(\alpha/2)/3$ is much smaller than one. For example, for $\alpha = 45\,^\circ$ we find $E_{\rm el}^{(1)}/E_{\rm el}^{(0)} \approx 0.06$, *i.e.* the approximation is still reasonable whereas for $\alpha = 90\,^\circ$, where $E_{\rm el}^{(1)}/E_{\rm el}^{(0)} \approx 0.3$, we expect appreciable deviations of the theory from experiments since the chain-chain repulsion is underestimated by ≈ 20 percent.

Our calculation does not account for *intra*molecular electrostatic contributions. These can be included by adding the electrostatic persistence length $l_{\rm OSF} = l_{\rm B}\lambda^2/4\kappa^2$ (the Odijk-Skolnick-Fixman length, cf. ref. [14]) to the bare persistence length, *i.e.* by replacing $l_{\rm P}$ by $l_{\rm P}+l_{\rm OSF}$. This leads to $\tan(\alpha/2) \propto (l_{\rm OSF}/(l_{\rm P}+l_{\rm OSF}))^{1/4}$. Note that the above-stated condition for the validity of eq. (3) can be reformulated in terms of $l_{\rm OSF}$: $E_{\rm el}^{(1)}/E_{\rm el}^{(0)} = \sqrt{l_{\rm OSF}/l_{\rm P}} \ll 1$, *i.e.* we require $l_{\rm OSF} \ll l_{\rm P}$.

We apply now our results to the problem of how the geometry of the chromatin fiber is controlled *in vitro* and, on a more tentative level, *in vivo*. The *in vitro* experiments show that

chromatin fibers "open up" with a decreasing salt concentration. From electron cryomicrographs it was estimated that $\alpha_{\rm exp} \approx 85^\circ$ for $c_{\rm s} = 5\,{\rm mM}$ and $\alpha_{\rm exp} \approx 45^\circ$ for $c_{\rm s} = 15\,{\rm mM}$ and from electron cryotomography that $\alpha_{\rm exp} \approx 35^\circ$ for $c_{\rm s} = 80\,{\rm mM}$ [5]. We expect from eq. (7) that $\alpha \simeq 2\arctan\left(Cc_{\rm s}^{-1/4}\right)$ with C being a constant. Let us take the angle at the highest salt concentration, $c_{\rm s} = 80\,{\rm mM}$, as the reference value. From this $C = 0.94$ follows. With this value of C we predict $\alpha \approx 51^\circ$ for $c_{\rm s} = 15\,{\rm mM}$ and $\alpha \approx 64^\circ$ for $c_{\rm s} = 5\,{\rm mM}$. Whereas the predicted value $\alpha \approx 51^\circ$ at an intermediate ionic strength is close to $\alpha_{\rm exp} \approx 45^\circ$, the value $\alpha \approx 64^\circ$ for low salt concentrations is noticeably too low ($\alpha_{\rm exp} \approx 85^\circ$). However, as shown above, for such a large value of α the chain-chain repulsion is underestimated by ≈ 20 percent.

How can the degree of swelling of the chromatin fiber be controlled *in vivo*? Under the assumption that the above-mentioned geometry is valid, the only parameter that might be under biochemical control is the linear charge density λ. It is known that the formation of a dense chromatin fiber is dependent on the presence of several components, especially on the presence of the cationic linker histones and of some of the lysine-rich (*i.e.* cationic) N-tails of the core histones that appear to be long, flexible polyelectrolyte chains [1]. In fig. 1(b) we give a tentative picture of the conformation of two N-tails that protrude from the histone core (we assume here that these are the tails of the two H3 core histones that are located close to the entry-exit point). It is known that if either of these components is missing, the fiber does not fold properly (cf. ref. [11] and references therein). As indicated in fig. 1(b) the tails might form a complex with the entering and exiting linker DNA in such a way that they effectively reduce its linear charge density λ. It is known that transcriptionally active regions in chromatin show an acetylation of the core histone tails (*i.e.* the cationic groups of the lysines are neutralized). In our tentative picture this acetylation mechanism would increase λ and according to eq. (7) this would lead to an opening of the entry-exit angle α. The acetylation might therefore be the first step in the decondensation of a stretch of the chromatin fiber that needs to be accessed for transcription [15].

Further steps might then involve the loss of the linker histones leading to $\alpha \approx \pi$, *i.e.* a bead-on-string filament. After the loss of the linker histones the nucleosomes might even become mobile as was reported in ref. [19]. We recently suggested that this "nucleosome sliding" results from DNA "reptating" around the histone core with the help of intranucleosomal loops [20]. Loop formation might also be crucial for the actual transcription itself: It was suggested that RNA polymerase might elongate through the nucleosome by passing it in a loop [21]. However, up to now a detailed knowledge of how transcription through chromatin is possible is still unclear and a matter of current research.

* * *

I wish to thank R. Bruinsma, A. Johner, A. Lesne and G. Migliorini for useful discussions.

REFERENCES

[1] Luger K., Mäder A. W., Richmond R. K., Sargent D. F. and Richmond T. J., *Nature*, **389** (1997) 251.
[2] Woodcock C. L. and Dimitrov S., *Curr. Opin. Genet. Dev.*, **11** (2001) 130.
[3] Finch J. T. and Klug A., *Proc. Natl. Acad. Sci. U.S.A.*, **73** (1976) 1897.
[4] Woodcock C. L., Grigoryev S. A., Horowitz R. A. and Whitaker N., *Proc. Natl. Acad. Sci. U.S.A.*, **90** (1993) 9021.

[5] Bednar J., Horowitz R. A., Grigoryev S. A., Carruthers L. M., Hansen J. C., Koster A. J. and Woodcock C. L., *Proc. Natl. Acad. Sci. U.S.A.*, **95** (1998) 14173.
[6] Cui Y. and Bustamante C., *Proc. Natl. Acad. Sci. U.S.A.*, **97** (2000) 127.
[7] Katritch V., Bustamante C. and Olson W. K., *J. Mol. Biol.*, **295** (2000) 29.
[8] Schiessel H., Gelbart W. M. and Bruinsma R., *Biophys. J.*, **80** (2001) 1940.
[9] Ben-Haïm E., Lesne A. and Victor J.-M., *Phys. Rev. E*, **64** (2001) 051921.
[10] Widom J., *Proc. Natl. Acad. Sci. U.S.A.*, **89** (1992) 1095.
[11] van Holde K. and Zlatanova J., *Proc. Natl. Acad. Sci. U.S.A.*, **93** (1996) 10548.
[12] We assume here that internucleosomal interaction can be neglected. There is evidence from the stretching experiments [6] that the nucleosomes are "in contact" at a physiological ionic strength and that there is a short-range attraction between them. With the swelling of the fiber at lower ionic strength the internucleosomal interaction becomes less important and might be negligible (cf. our discussion on the internucleosomal interaction in ref. [8]).
[13] Note that our analysis is based on the assumption of a vanishing distance between the two strands for $x \leq 0$. This leads to a logarithmic singularity of the curvature at $x = 0$, *i.e.* to an infinite torque at the origin. The structure could only be stabilized for an infinite adsorption energy per length between the two strands for $x \leq 0$. A more realistic case is $\tilde{h}(x) \equiv \varepsilon \ll 0$ for $x \leq 0$. In this case the following approximate solution can be constructed: $\tilde{x}^4/24\varepsilon - \tilde{x}^3/(3\tilde{x}_0\sqrt{-\ln \tilde{x}_0}) + \tilde{x}^2\sqrt{(-\ln \tilde{x}_0)} + \varepsilon$, for $\tilde{x} < \tilde{x}_0$ and $\tilde{h}_0(x) + \varepsilon$, for $\tilde{x}_0 < \tilde{x} \ll 1$; here $\tilde{x}_0$ is the solution of $\tilde{x}_0^2\sqrt{-\ln \tilde{x}_0} = \varepsilon$. These two functions cross over smoothly at $\tilde{x} \approx \tilde{x}_0$ and obey eq. (6) for $\tilde{x} \ll \tilde{x}_0$ and for $\tilde{x} \gg \tilde{x}_0$, respectively. This indicates that the solution is only marginally affected by the value of ε as long as $\varepsilon \ll 0$ and that our original assumption, eq. (1), is reasonable.
[14] Barrat J.-L. and Joanny J.-F., *Adv. Chem. Phys.*, **94** (1996) 1.
[15] The processes that are involved in the acetylation and deacetylation might be quite specific and involved as, for instance, discussed in ref. [16]. The histone tail modifications might serve specific functions via the modification of their secondary structure that in turn modifies their interaction with certain proteins [17]. Recently there have been even attempts to decipher a specific "language" of covalent histone modifications [18]. It might be that such specific processes act in concert with the more basic charge neutralization principle discussed here.
[16] Grunstein M., *Nature*, **389** (1997) 349.
[17] Hansen J. C., Tse C. and Wolffe A. P., *Biochemistry*, **37** (1998) 17637.
[18] Strahl B. D. and Allis D., *Nature*, **403** (2000) 41.
[19] Pennings S., Meersseman G. and Bradbury E. M., *Proc. Natl. Acad. Sci. U.S.A.*, **91** (1994) 10275.
[20] Schiessel H., Widom J., Bruinsma R. F. and Gelbart W. M., *Phys. Rev. Lett.*, **86** (2001) 4414.
[21] Studitsky V. M., Clark D. J. and Felsenfeld G., *Cell*, **71** (1992) 371.