How the chromatin fiber deals with topological constraints

Maria Barbi, Julien Mozziconacci, and Jean-Marc Victor*
Laboratoire de Physique Theorique des Liquides, CNRS UMR 7600, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex 05, France
(Received 13 April 2004; published 23 March 2005)

In the nuclei of eukaryotic cells, DNA is packaged through several levels of compaction in an orderly retrievable way that enables the correct regulation of gene expression. The functional dynamics of this assembly involves the unwinding of the so-called 30-nm chromatin fiber and accordingly imposes strong topological constraints. We present a general method for computing both the twist and the writhe of any winding pattern. An explicit derivation is implemented for the chromatin fiber which provides the linking number of DNA in eukaryotic chromosomes. We show that there exists one and only one unwinding path which satisfies both topological and mechanical constraints that DNA has to deal with during condensation/decondensation processes.

DOI: 10.1103/PhysRevE.71.031910 PACS number(s): 87.16.Sr

In the cells of higher eukaryotes, e.g., animals or plants, meters of DNA are packaged by means of proteins into a nucleus of a few micrometer diameter, providing an extreme level of compaction. Coding sequences (*genes*) are therefore dispersed in a mess of folded DNA and proteins (*chromatin*) and should be retrieved at will in order to enable a correct genetic expression and therefore the cell survival. This leads to the need for an orderly and dynamically retrievable structure, which is actually achieved by means of a chromatin partition into functional domains, each containing one or a group of genes. In each domain, DNA is folded in a hierarchical structure, including several winding levels. It is first wrapped around spools of proteins thus forming a "beads on a string" assembly, which is in turn folded into a 30-nm-diam fiber. This fiber is further organized into a three-dimensional cross-linked network [1]. In this network, two neighboring nodes are connected by a chromatin fiber loop whose typical length is about 50 000 base pairs (bp). In order to provide the transcription machinery with an access to specific genomic regions, the corresponding loop has to be selectively decondensed, via a reversible unwinding process that elongates the fiber [2]. The dynamics of this process involves strong mechanical and topological constraints, the former due to DNA elasticity [3], the latter due to the conservation of the linking number Lk of DNA [4] in a loop during the unwinding. Of course, topological constraints could be released by the intervention of topoisomerases, but it has been shown *in vivo* that chromatin decondensation could take place even when those enzymes were inhibited [5]. Moreover, the classic experiment on *simian virus 40* (SV40) minichromosomes clearly demonstrates that the linking number is unaffected by the decondensation process *in vitro* [6].

In this paper, we address the issue of how evolution has dealt with the extremely difficult problem of finding an efficient winding pattern fulfilling both mechanical and topological constraints at a time. To answer this question, we start by giving an analytical formula for the linking number of DNA in a generic bent fiber. We show that, despite the fact that Lk is known to be a nonextensive quantity, it is yet possible to express it in terms of the mean linking number of the constitutive elements of the beads on a string assembly. This allows us to set about an exhaustive numerical exploration of Lk for all the possible fiber conformations. By analyzing the results of this exploration, we are able to infer the existence and the *uniqueness* of a relevant winding/unwinding path satisfying all the constraints. We show furthermore that this engineering problem is solved at the *local* level by the design of the constitutive elements of the fiber.

In a typical chromatin fiber, the beads, called *nucleosome core particles* (NCP), are spaced at intervals of ~200 bp. Each NCP contains ~150 DNA bp, forming $1\frac{3}{4}$ turns of a left-handed superhelix [7], and two neighboring beads are connected by ~50 bp stretches of DNA, called *linkers*. The unit of a NCP and a linker is called a *nucleosome*. The number of DNA bp in a nucleosome is known as the *repeat length*. In order to describe the structural polymorphism of a chromatin fiber loop, we use the original two-angle model (Fig. 1) introduced by Woodcock *et al.* [8]. In this model, where the linkers are assumed to be straight, the geometry of a fiber made up of N nucleosomes is fully characterized by two sets of angles α_i and β_i $(i=1,\ldots,N)$, specific to the

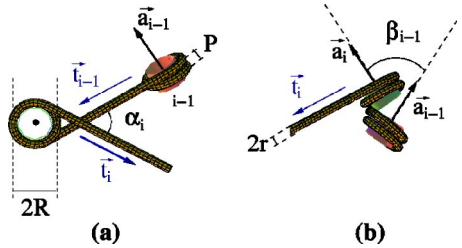

FIG. 1. (Color online) Schematic of the DNA winding pattern along two neighboring nucleosomes in the two-angle model. (a) View down the NCP axis $\vec{a}_i$. (b) View down the linker direction $\vec{t}_{i-1}$. The angle α_i is the dihedral angle $(-\vec{t}_{i-1},\vec{a}_i,\vec{t}_i)$ and β_{i-1} is the dihedral angle $(\vec{a}_{i-1},\vec{t}_{i-1},\vec{a}_i)$ standing for the twist (*modulo* 2π) of the DNA linker. We also indicate the DNA radius $r\simeq 1.0$ nm, NCP radius $R\simeq 5.3$ nm, and NCP pitch $P\simeq 2.4$ nm.

*Author to whom all correspondence should be addressed. Email address: victor@lptl.jussieu.fr

NCPs and the linkers, respectively [9,10]. Varying α_i corresponds to changing the number of bp wrapped in the ith NCP, whereas varying β_i corresponds to twisting the ith linker.

In vivo, the fiber is most probably homogeneous within a given loop [11], i.e., that the distributions of α_i and β_i are peaked around mean values $\overline{\alpha}$ and $\overline{\beta}$. During the cell cycle, the biological activity is tuned by biophysical [12] and biochemical [13] parameters, through a variation of these mean values. In turn, a variation of $\overline{\alpha}$ or $\overline{\beta}$ results in a change of the fiber internal structure and hence affects its degree of compaction. This conformational change is necessarily accompanied by a change of the writhe, Wr, and the twist, Tw, of the DNA double helix. Consequently, the DNA linking number is *a priori* bound to change according to the White-Fuller theorem Lk=Tw+Wr [14,15]. This poses the rather important question of whether and how the angles $\overline{\alpha}$ and $\overline{\beta}$ can be changed, while maintaining a fixed DNA linking number. To answer this question, we set about computing the DNA linking number in a fiber as a function of the sets α_i and β_i. This involves the computation of both the twist and the writhe of the DNA double helix in the fiber.

Concerning the computation of the writhe, Fuller's first theorem [16] states that, for a closed curve $\vec{r}(s)$ with $s \in [0,L]$, the writhe is related to the *signed* area A enclosed by the tangent indicatrix $T(s)$, namely, $\mathrm{Wr}=A/2\pi-1$ (*modulo 2*). $T(s)$ is the curve on the unit sphere traced out by the tangent vector $\vec{t}(s)$ to the curve $\vec{r}(s)$. Fuller's second theorem [16] permits us to get rid of the congruence *modulo 2* by computing Wr from the area S swept out by the unique shortest geodesic arc from an arbitrarily fixed point C on the unit sphere to the running point $T(s)$, namely, $\mathrm{Wr}=S/2\pi$. However, the calculation of the writhe of an open curve whose initial and final tangent vectors $\vec{t}(0)$ and $\vec{t}(L)$ do not coincide is rather subtle. We follow the procedure recently outlined by Maggs [17] to obtain a consistent measure for the writhe by closing the open tangent indicatrix with a geodesic arc. This being rather complicated for a general fiber conformation, it would be helpful if one could calculate the writhe from the writhes of the individual nucleosome units. However, the writhe is known to be nonextensive and, as shown by Starostin [18], the total writhe of a curve is given by the sum of the writhes of its parts *plus* the surface of the spherical polygon composed by the set of geodesics closing each part. In the case of a chromatin fiber, this nonextensivity may be overcome in the following way. As shown schematically in Fig. 2, the DNA indicatrix $T(s)$ can be naturally divided into N parts, namely, the N nucleosomes, in correspondence to $\vec{t}_i=\vec{t}(s_i)$, the direction of the ith linker. We denote by T_i the point on the unit sphere that corresponds to $\vec{t}_i$. The total writhe of the curve, Wr, is then given by the following addition rule:

$$\mathrm{Wr}=\sum_{i=1}^{N}\mathrm{Wr}_i+\frac{S_{T_0T_1\cdots T_N}}{2\pi}. \tag{1}$$

Wr_i is the writhe of the ith nucleosome and is given by the surface S enclosed by the restriction of the tangent indicatrix between T_{i-1} and T_i and the geodesic $T_{i-1}T_i$. The nonextensive correction $S_{T_0T_1\cdots T_N}$ is the signed area of the spherical polygon connecting all points T_i that correspond to the linker directions $\vec{t}_i$. It can be computed as the sum of the areas of the spherical triangles $T_0T_iT_{i+1}$. It can also be expressed with respect to any point C on the unit sphere as $\Sigma_{i=1}^{N}S_{CT_{i-1}T_i}$ plus a boundary term $S_{CT_NT_0}$. Up to this term, the total writhe can be written as $\mathrm{Wr}=\Sigma_{i=1}^{N}\mathrm{Wr}_i|_C$, where

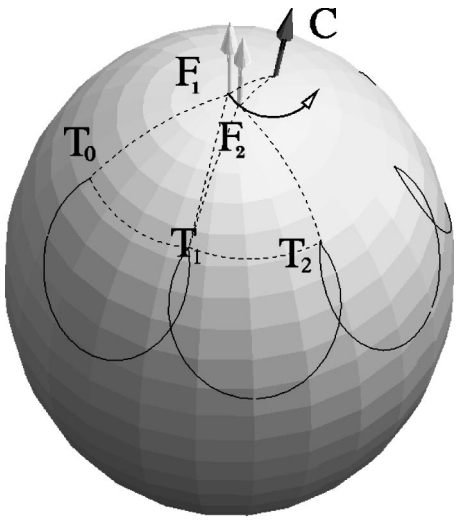

FIG. 2. Schematic of the tangent indicatrix of DNA in a chromatin fiber with the indication of the particular points used in the calculations: each solid-line arc corresponds to a nucleosome, with linker directions given by the connecting points T_i. Points F_i indicate the local fiber axis directions. C is an arbitrarily fixed point. All dashed lines indicate geodesic arcs. For the sake of clarity, scales and lengths have been arbitrarily chosen.

$$\mathrm{Wr}_i|_C=\mathrm{Wr}_i+\frac{S_{CT_{i-1}T_i}}{2\pi} \tag{2}$$

is the writhe of nucleosome i with respect to point C, corresponding to the writhe of the nucleosome i alone plus the surface of the spherical triangle $CT_{i-1}T_i$. We see that an appropriate redefinition of the writhe of a nucleosome permits us to express the total writhe as an extensive quantity.

For the simplest case of a straight regular fiber, all the nucleosomes are evenly dispersed on a regular helix, with the angles $\alpha_i=\alpha$ and $\beta_i=\beta\ \forall i$. In this case, though Wr_i is the same for all nucleosomes, $\mathrm{Wr}_i|_C$ varies in general from nucleosome to nucleosome. Nonetheless, there exists a special point $C=F$, defined by the director of the fiber axis, for which $\mathrm{Wr}_i|_F$ is independent of i, so that we can define an effective writhe *per nucleosome* $\mathrm{Wr}(\alpha,\beta)=\mathrm{Wr}_i|_F$, such that $\mathrm{Wr}=N\,\mathrm{Wr}(\alpha,\beta)$ [24]. We reiterate that the effective writhe *per nucleosome* $\mathrm{Wr}(\alpha,\beta)$ is not the bare writhe *of one individual nucleosome*, the difference being the surface of the spherical triangle FT_0T_1.

For a generic bent fiber, a local fiber axis F_i can be defined for each i as the axis of the straight fiber characterized by constant angles $\alpha=\alpha_i$ and $\beta=\beta_i$ and containing the ith nucleosome. As shown in Fig. 2, this allows us to subdivide

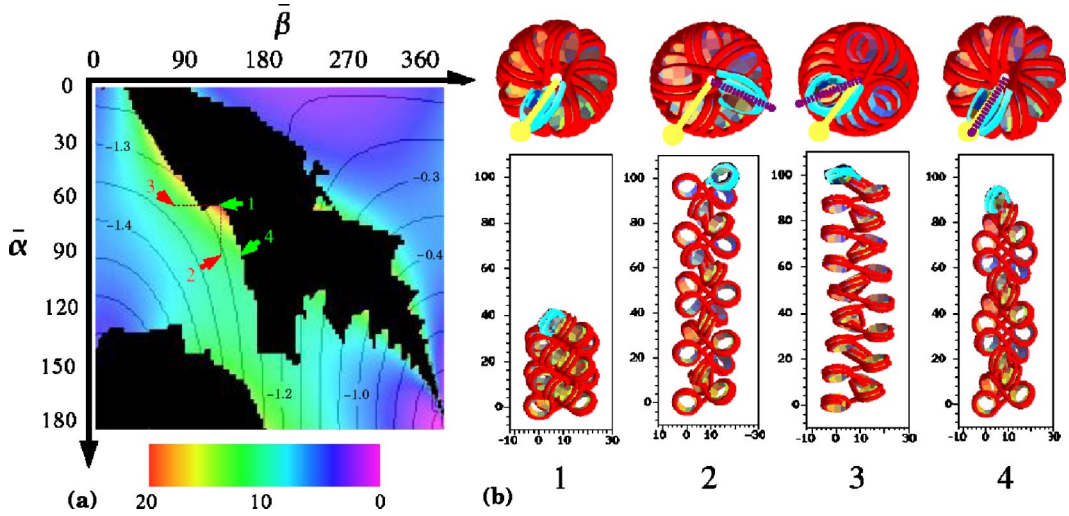

FIG. 3. (Color online) (a) Solid contour lines (spaced at intervals of 0.05) represent the linking number $\mathrm{Lk}(\bar{\alpha},\bar{\beta})$ computed for $\bar{\alpha}$ ranging from 0 to 180°, $\bar{\beta}$ from 0 to 360°, and $\bar{n}=192$ bp. Color scale gives the fiber compaction defined as the number of nucleosomes per unit volume $C=1/\pi dR^2$, in units of NCP/(30 nm)^3. Black regions correspond to sterically forbidden regions. Points 1 to 4 are particular reference states: structure 1 is the most compact fiber structure allowed; paths 1–2, 1–3, and 1–4 correspond, respectively, to decondensation at constant $\bar{\beta}$, constant $\bar{\alpha}$, and constant $\mathrm{Lk}(\bar{\alpha},\bar{\beta})$. (b) Fiber structures corresponding to points 1 to 4 in (a), top view and frontal view drawn for 24 nucleosomes. For clarity, the first and the last NCP orientations are outlined by a solid (yellow) and a dashed (violet) line. The twist change between any two fiber structures is the angle swept out by the dashed line (namely, −0.70 between 1 and 2, −1.10 between 1 and 3, and 0 between 1 and 4).

the surface of the spherical polygon $T_0T_1\cdots T_N$ into spherical triangles defined with respect to the points F_i, so that the total writhe finally reads

$$\mathrm{Wr}=\sum_{i=1}^{N}\mathrm{Wr}(\alpha_i,\beta_i)+\sum_{i=1}^{N-1}\frac{S_{F_iT_iF_{i+1}}}{2\pi}+\sum_{i=1}^{N-1}\frac{S_{CF_iF_{i+1}}}{2\pi} \tag{3}$$

with

$$\mathrm{Wr}(\alpha_i,\beta_i)=\mathrm{Wr}_i+\frac{S_{F_iT_{i-1}T_i}}{2\pi} \tag{4}$$

up to the boundary terms $S_{CF_NF_1}+S_{F_NT_NT_0F_1}$.

As concerns the twist of the DNA double helix in the fiber, its extensivity enables a more straightforward calculation. First of all, note that, if the repeat length n_i is allowed to change from nucleosome to nucleosome, the ith linker will be relaxed (untwisted) only for a particular equilibrium value $\beta^{eq}(\alpha_i,n_i)$ of β_i. Then, let $\mathrm{Tw}(\alpha_i,\beta_i)$ be the twist of the ith nucleosome, with n_i as a repeat length. At fixed α_i, the variation $\Delta\,\mathrm{Tw}(\alpha_i,\beta_i)$ with respect to β_i is equal to $\Delta\beta_i/2\pi$, whereas, at fixed β_i, the variation with α_i is equal to the variation of the tortuosity $\tau R\Delta\alpha_i/2\pi$, where $\tau=2\pi P/(4\pi^2R^2+P^2)$ is the torsion of the double helix axis in the NCP [25]. In order to have an expression for the total twist, we just need to know, therefore, its value on a particular reference state. We take as a reference a nucleosome with 146 bp in the NCP [7] and a relaxed linker. In this state, $\alpha_i=\alpha^0\simeq 60°$ and $\beta_i=\beta_i^0=\beta^{eq}(\alpha^0,n_i)=2\pi(n_i-146)/h_0$. The twist is then $\mathrm{Tw}(\alpha^0,\beta_i^0)=146/h+(n_i-146)/h_0$, where h_0 and h are, respectively, the helical repeat of the double helix free in solution and overtwisted in the NCP [23]. The total twist of the DNA double helix in the fiber is therefore

$$\mathrm{Tw}=\sum_{i=1}^{N}\tau R\frac{\alpha_i-\alpha^0}{2\pi}+\frac{\beta_i-\beta_i^0}{2\pi}+146\left(\frac{1}{h}-\frac{1}{h_0}\right)+\frac{n_i}{h_0}. \tag{5}$$

For the DNA double helix in a straight and relaxed state, the writhe is zero and the total twist simply amounts to the last term in Eq. (5), hence $\mathrm{Lk}^0=\mathrm{Tw}^0=\Sigma_{i=1}^{N}n_i/h_0=N\bar{n}/h_0$, where $\bar{n}$ is the mean repeat length of the fiber. Therefore, defining the local twist with respect to the relaxed DNA as

$$\mathrm{Tw}(\alpha_i,\beta_i)=\tau R\frac{\alpha_i-\alpha^0}{2\pi}+\frac{\beta_i-\beta_i^0}{2\pi}+146\left(\frac{1}{h}-\frac{1}{h_0}\right), \tag{6}$$

we obtain from Eq. (3) and Eq. (5) the following expression for the DNA *excess* linking number of a generic fiber:

$$\mathrm{Lk}-\mathrm{Lk}^0=\sum_{i=1}^{N}\mathrm{Lk}(\alpha_i,\beta_i)+\sum_{i=1}^{N-1}\frac{S_{F_iT_iF_{i+1}}}{2\pi}+\sum_{i=1}^{N-1}\frac{S_{CF_iF_{i+1}}}{2\pi}, \tag{7}$$

where

$$\mathrm{Lk}(\alpha_i,\beta_i)=\mathrm{Wr}(\alpha_i,\beta_i)+\mathrm{Tw}(\alpha_i,\beta_i) \tag{8}$$

is the local DNA linking number *per nucleosome*.

It is interesting to note that, in the very last term of Eq. (7), $\Sigma_{i=1}^{N-1}S_{CF_iF_{i+1}}$ is the area swept out by the geodesic arc from C to the point F_i as i increases. Therefore, this term is nothing but the writhe of the fiber *axis*, $\mathrm{Wr}^{\mathcal{F}}$. The excess linking number $\mathrm{Lk}-\mathrm{Lk}^0$ is the contribution of the winding pattern of the double helix to the total linking number Lk. It should therefore be interpreted as the linking number of the fiber, $\mathrm{Lk}^{\mathcal{F}}$. The difference $\mathrm{Lk}^{\mathcal{F}}-\mathrm{Wr}^{\mathcal{F}}$ must in turn be identi-

fied with the twist of the fiber around its axis, $\mathrm{Tw}^{\mathcal{F}}$, in order to satisfy the White-Fuller theorem *at the level* of the fiber $\mathrm{Lk}^{\mathcal{F}}=\mathrm{Wr}^{\mathcal{F}}+\mathrm{Tw}^{\mathcal{F}}$. This formal definition of $\mathrm{Tw}^{\mathcal{F}}$ actually matches the intuitive definition of the fiber twist as the rotation angle of the fiber "top" with respect to the fiber "bottom," as shown in Fig. 3(b), where we give the twist variations between different fibers.

We now evaluate the relative contributions of the three terms in Eq. (7). The first one can be evaluated in a mean-field approximation as $\Sigma_{i=1}^{N}\mathrm{Lk}(\alpha_i,\beta_i)\simeq N\,\mathrm{Lk}(\overline{\alpha},\overline{\beta})$. We have therefore computed $\mathrm{Lk}(\overline{\alpha},\overline{\beta})$ for $\overline{\alpha}$ ranging from 0° (corresponding to "open" NCP wrapped with 1.5 turns of superhelix [19]) to 180° (corresponding to "closed" NCP, with two complete turns) and $\overline{\beta}$ from 0° to 360° (one period). A contour plot of our results obtained for $\overline{n}=192$ bp is shown in Fig. 3(a). According to our definition, $\mathrm{Lk}(\overline{\alpha},\overline{\beta})$ also depends on the mean repeat length $\overline{n}$ through the term $\overline{\beta}^0/2\pi=(\overline{n}-146)/h_0$ (*modulo* 2π). Anyway, changing $\overline{n}$ simply shifts the levels of the contour lines without affecting their shapes. The last two terms of Eq. (7) arise from the bending of the fiber. Taking into account that the mean radius of curvature of the fiber axis is equal to the fiber persistence length, which is $\simeq$30 nm [10], it is possible to show that their net contribution *per nucleosome* never exceeds 0.01. We can therefore conclude that an Lk-conserving path in the $(\overline{\alpha},\overline{\beta})$ plane practically never deviates from a given contour line.

We finally come to the issue of whether a biologically relevant Lk-conserving decondensation path in the $(\overline{\alpha},\overline{\beta})$ plane can actually exist. We know that such a path must remain close to a contour line. Moreover, it should be capable of transforming a highly compact structure into a decondensed one. A fiber compaction map is displayed in Fig. 3(a). An exhaustive scan of this map shows that there is only one small region, around point 1, which provides a degree of compaction high enough to match the maximal densities observed *in vitro*, i.e., $\sim$6 NCP/10 nm. Hence, there can be only one Lk-conserving decondensation path, namely the contour line going through point 1 and leading to point 4. It is clear from Fig. 3(a) that varying only $\overline{\alpha}$ (leading to point 2) or only $\overline{\beta}$ (leading to point 3) is forbidden at constant Lk.

Moreover, energy considerations have to be taken into account. As shown by Langowski in a recent simulation [20], the internucleosomal interactions play a minor role in the control of the fiber compaction while the elastic energy (bending and twisting) of the linkers plays the major part. It comes out that the linkers should remain relaxed all over the decondensation path. First, in order to have relaxed linkers at point 1, $\overline{\beta}$ must be equal to $\beta^{eq}(\alpha^0,\overline{n})$: this leads to selecting a special value for the mean repeat length $\overline{n}$, equal to 192 bp (modulo 10.6 bp). Second, during the decondensation (path 1–4), $\overline{\alpha}$ varies from $\sim$60 to $\sim$90°. This variation corresponds to the wrapping of $\sim$10 bp of the linker onto the protein spool. As the DNA helical repeat is lower in the NCP ($h=10.2$ bp/turn) than in the linkers ($h_0=10.6$ bp/turn), this wrapping is accompanied by a change of the equilibrium value $\beta^{eq}(\overline{\alpha},\overline{n})$ from 120 to 140°. Remarkably, along the decondensation path 1–4, $\overline{\beta}$ varies as well from 120 to 140°. Hence $\overline{\beta}=\beta^{eq}(\overline{\alpha},\overline{n})$ all over the path 1–4, provided that $\overline{n}=192$ bp, which is the repeat length measured in HeLa cells.

Our results provide a way of understanding the quantization of the repeat lengths [21]. The average repeat length has to be selected, e.g., thanks to nucleosome positioning sequences, in order to fix the mean nucleosome orientation $\beta^{eq}(\overline{\alpha},\overline{n})$. This particular orientation not only provides the maximal fiber compaction with relaxed linkers [21], but also enables the global conservation of the linking number of the fiber during its winding/unwinding process driven by the nucleosome internal dynamics itself.

We thank A. Lesne, A. Paldi, A. Prunell, and C. Lavelle for fruitful discussions, R. Chitra for a very careful reading of the paper, and M. Quaggetto for technical support.

[1] M. G. Poirier and J. F. Marko, Proc. Natl. Acad. Sci. U.S.A. **99**, 15 393 (2002).
[2] L. A. Cirillo *et al.*, Mol. Cell **9**, 279 (2002).
[3] J. Yao, P. T. Lowary, and J. Widom, Proc. Natl. Acad. Sci. U.S.A. **90**, 9364 (1993).
[4] F. H. C. Crick, Proc. Natl. Acad. Sci. U.S.A. **73**, 2639 (1976).
[5] S. J. Wright and G. Schatten, Dev. Biol. **142**, 224 (1990).
[6] W. Keller *et al.*, Cold Spring Harbor Symp. Quant. Biol. **42**, 227 (1978).
[7] K. Luger *et al.*, Nature (London) **389**, 251 (1997).
[8] C. L. Woodcock *et al.*, Proc. Natl. Acad. Sci. U.S.A. **90**, 9021 (1993).
[9] H. Schiessel, W. M. Gelbart, and R. Bruinsma, Biophys. J. **80**, 1940 (2001).
[10] E. Ben-Haïm, A. Lesne, and J. M. Victor, Phys. Rev. E **64**, 051921 (2001).
[11] T. Weidemann *et al.*, J. Mol. Biol. **334**, 229 (2003).
[12] J. Bednar *et al.*, Proc. Natl. Acad. Sci. U.S.A. **95**, 14 173 (1998).
[13] T. Jenuwein and C. D. Allis, Science **293**, 1074 (2001).
[14] J. H. White, Am. J. Math. **91**, 693 (1969).
[15] F. B. Fuller, Proc. Natl. Acad. Sci. U.S.A. **68**, 815 (1971).
[16] F. B. Fuller, Proc. Natl. Acad. Sci. U.S.A. **75**, 3557 (1978).
[17] A. C. Maggs, J. Chem. Phys. **114**, 5888 (2001).
[18] E. L. Starostin, e-print physics/0212095.
[19] A. Prunell and A. Sivolob, in *Chromatin Structure and Dynamics: State of the Art* (Elsevier Science, Amsterdam, 2003).
[20] G. Wedemann and J. Langowski, Biophys. J. **82**, 2847 (2002).
[21] J. Widom, Proc. Natl. Acad. Sci. U.S.A. **89**, 1095 (1992).
[22] M. Le Bret, J. Mol. Biol. **200**, 285 (1988).
[23] J. J. Hayes, T. D. Tullius, and A. P. Wolffe, Proc. Natl. Acad. Sci. U.S.A. **87**, 7405 (1990).
[24] Note that the same viewpoint *F* was introduced by Crick [4].
[25] The tortuosity accounts for the twist arising from the wrapping of the DNA axis into a superhelix, in the absence of torsional constraint on the double helix itself [22].