

DNA topology in chromosomes: a quantitative survey and its physiological implications

Maria Barbi · Julien Mozziconacci · Hua Wong ·
Jean-Marc Victor

Received: 23 December 2011 / Revised: 22 October 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract Using a simple geometric model, we propose a general method for computing the linking number of the DNA embedded in chromatin fibers. The relevance of the method is reviewed through the single molecule experiments that have been performed *in vitro* with magnetic tweezers. We compute the linking number of the DNA in the manifold conformational states of the nucleosome which have been evidenced in these experiments and discuss the functional dynamics of chromosomes in the light of these manifold states.

Keywords DNA · Topology · Chromatine

Mathematics Subject Classification (2000) 92B99

1 Introduction

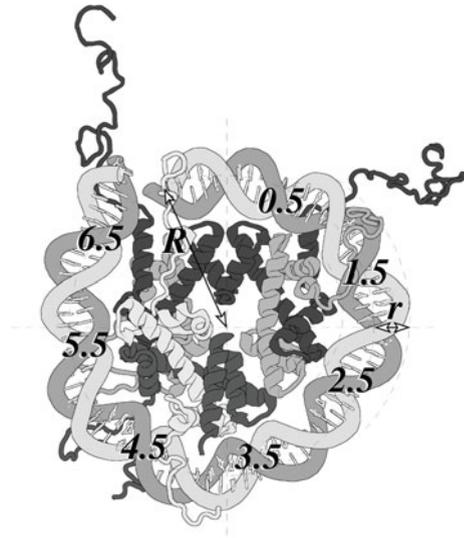
According to the “central dogma” of molecular biology, the DNA double helix codes for the sequence of all proteins in the cell. It also codes for non-messenger (siRNA,

M. Barbi (✉) · J. Mozziconacci · J.-M. Victor
Laboratoire de Physique Théorique de la Matière Condensée, CNRS UMR 7600,
and CNRS GDR 3536, Université Pierre et Marie Curie, Case courrier 121,
4 place Jussieu, 75252 Paris, France
e-mail: barbi@lptmc.jussieu.fr

H. Wong
Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France
e-mail: hua.wong@pasteur.fr

H. Wong
CNRS GDR 3536, Université Pierre et Marie Curie, Case Currier 121,
4 place Jussieu, 75252 Paris, France

Fig. 1 Structure of half a nucleosome. The Super Helix Locations (SHL) have been indicated as well as the nucleosome and DNA radii. The figure has been obtained starting from the crystallographic structure (Davey et al. 2002)



miRNA,...) and ribosomal (rRNA) RNAs. However, in eukaryotic cells, this may represent as few as 2 % of the DNA sequence. It is then tempting to hypothesize that the non coding DNA is used by the cell as a template on which a hierarchical architecture is assembled in order to regulate gene expression. This architecture has to deal with the mechanical properties of DNA. The latter can be simply described as a screw, 2 nm in diameter, up to a few meters in length, with a plexiglass-like Young modulus. Dealing with meters of this screw is a daunting problem since the nucleus is only 10 μm wide.

In the nucleus of eukaryotic cells the basic units of this complex architecture is a spool of proteins (histones) on which 147 DNA base pairs (bp) are wrapped in 1.7 turns, thus forming a nucleosome (Fig. 1). Nucleosomes are more or less regularly spaced on the genome every 200 bp, forming a bead on a string array. Attractive interactions between nucleosomes fold this array into a fiber about 30 nm in diameter (Wolffe 1998). This fiber is then cross-linked to form loops (Byrd and Corces 2003; Labrador and Corces 2002). Chromosomal loops can enclose one single gene or a set of genes. When a gene is silenced, the fiber in the loop is condensed and buried inside the chromosome. When a gene is expressed, the corresponding loop is translocated to a transcription factory (Cook 1999), consisting of hundreds of different proteins required for DNA transcription and RNA maturation. Among them, RNA polymerases are devoted to the transcription of DNA into RNA. These enzymes are loaded on the gene promoter in the initiation step. Then they start to transcribe DNA into RNA following the double helix groove. During this elongation process the DNA has to be screwed inside the RNA polymerase embedded in the transcription factory and the chromatin fiber has to absorb the topological and mechanical constraints upstream and downstream from the transcription starting site (Lavelle 2007). Indeed, when the RNA polymerase progresses along the DNA, the double helix in front of it becomes overwound (positively supercoiled) whereas the DNA behind

it becomes underwound (negatively supercoiled). This effect is usually referred to as twin-supercoiled-domain (TSD) model, first introduced by Liu & Wang (Lavelle 2007; Liu and Wang 1987). These authors originally considered the case of prokaryotes; however, the TSD model has been shown to be potentially relevant for eukaryotes as well (Giaever and Wang 1988; Ljungman and Hanawalt 1992). Transcription-coupled negative supercoiling in chromatin has been recently observed even in the presence of active topoisomerases (Matsumoto and Hirose 2004), supporting the idea that supercoiling accumulation *in vivo* may have a functional role and should not be removed. In a previous work, we suggested e.g. that the relaxation of the supercoiling could be used to rewind genes inside the RNA-polymerase, up to the promoter, in order to initiate a new transcription round (Wong et al. 2009). Another hypothesis is that supercoiling storage, and the concomitant chromatin conformational change, could be used by the cell as a label to remind that the corresponding gene has already been transcribed.

The other chromosomal transactions, in particular replication, repair and recombination, all have to deal with such topological constraints. Understanding the dynamics of the chromosome therefore requires to quantitatively evaluate the linking number of the DNA in the chromatin fiber.

Topology of DNA is a long standing problem, since Vinograd et al. introduced the idea that the conformation of DNA both in eukaryotes and prokaryotes is related to topological quantities (Vinograd et al. 1965) about 50 years ago. The mathematical definition of the three relevant topological quantities, i.e. the twist Tw , the writhe Wr and the linking number Lk has been stated soon after by (Fuller 1971, 1978).

After these original publications, a rather large amount of work have been made to study the relationship between twist and writhe for closed and open curves, especially in relation to torsionally stressed DNA including in presence of binding proteins (Maggs 2000, 2001; Rossetto and Maggs 2003; Starostin 2005; White and Bauer 1988a,b; White et al. 1986). However, the question of defining the DNA topology in the context of the chromatin fiber has never been addressed until we have shown that it is possible to define a twist and a writhe for the 30 nm fiber (Barbi et al. 2005). These two quantities share the same properties as the twist and the writhe of the DNA double helix. In particular, their sum $Wr_{fiber} + Tw_{fiber}$ is equal to the linking number of the fiber Lk_{fiber} which is itself equal to the linking number of the DNA up to a constant. One important question that we addressed in this first work is how to define a linking number *per nucleosome* in a regular fiber. In subsequent publications we used this results to interpret experimental data (Bancaud et al. 2006, 2007; Recouvreux et al. 2011) and to discuss the role of topological constraints in some biological processes (Bécavin et al. 2010; Lavelle et al. 2011; Mozziconacci et al. 2006; Wong et al. 2007). The full derivation of the fiber linking number was however not addressed in enough detail in any of the previous publications. The purpose of the present paper is to give the complete framework allowing for the calculation of the fiber topological quantities. In particular, we explicitly derive one convenient expression for the linking number *per nucleosome* in a regular fiber. Such definition can be obtained by choosing the fiber axis as a preferential direction and by defining the DNA writhe with respect to this special direction. This requires however to overcome the limitations related to the non-additivity of the writhe, a question that has been extensively discuss by Fuller (1978), and that we revisit from a slightly different point of view. This first part

of our paper is organized as follows: we start by introducing the two-angle model used to describe the fiber structure (Sect. 2). We then calculate the DNA linking number in the framework of this model (Sect. 3) and we finally propose a fast, alternative method for this calculation (Sect. 4).

The second part of the paper is devoted to the discussion of some biological questions and experimental results. We apply indeed our calculations to recent single molecule experiments performed *in vitro* with magnetic tweezers (Sect. 5). In this context, we also take into account the polymorphism of the nucleosome and evaluate the topological quantities of the various nucleosome conformations. We finally switch to a more biological perspective (Sect. 6) to propose physically plausible scenarios for the different steps involved in transcription, focusing on: (1) the chromatin fiber elongation and unwinding, associated with the translocation towards a transcription factory during the initiation step, and (2) the absorption of topological constraints upstream and downstream from the transcription starting site during the elongation step.

2 The two-angle model

2.1 Definition of the model

DNA parameters We will start by recalling the main characteristics of DNA and chromatin geometry and by defining the notation used in the following. We note x the distance between two adjacent base pairs along the DNA axis, h the number of base pairs in a DNA helix period, and $p = xh$ the corresponding length along the axis, or double helix *pitch*. In a relaxed DNA, the distance x equals 0.32 nm, but it can vary, in particular for DNA under constraint. We denote r the double helix radius and $\theta = 2\pi/h$ the twist angle between adjacent bps for DNA at rest. Note that only one of the three variable p , h and θ needs to be assigned. We will indicate with h_0 , p_0 and θ_0 the number of base pairs, length and rotation angle for a straight and relaxed DNA in the classical B form, i.e. without additional torsional or stretching stress (see Table 1 for numerical estimates).

Nucleosome structure Chromatin fibers are formed by a sequence of units called *nucleosomes*. Each nucleosome contains a stretch of DNA of $N_r \sim 200$ bp (called the nucleosome *repeat*). Among them, n_{NCP} base pairs are wrapped onto a protein octamer, formed by proteins called histones. This results in a quasi-cylindrical structure called the nucleosome core particle (NCP). In the NCP, the DNA double helix follows a solenoidal left-handed helix for roughly $1 \frac{3}{4}$ turns, this leading to a superhelical geometry. We introduce the radius R of the solenoid, roughly equal to the sum of the octamer radius and the DNA radius, and its pitch P .

The wrapped DNA segment is maintained in its conformation by tight interactions with the histone proteins at 14 independent minor groove locations, usually referred to as SHLs (for *superhelix locations*) $\pm 0.5, \pm 1.5 \dots \pm 6.5$ (Luger et al. 1997) (see Fig. 1). The semi-integer SHL numbers correspond to the crossing in the minor groove, close to the octamer core (for the superhelix viewed from along its axis). Conversely, integral SHL numbers correspond to crossings in the major groove.

Table 1 Definition of the geometrical parameters used throughout the paper with their reference values

Parameter	Symbol	Value
Distance between adjacent bps	x	0.32 nm/bp ^a
DNA double helix radius	r	1 nm
Number of bps in a <i>linker DNA</i> period	h	
<i>Linker DNA</i> double helix period	p	
Twist angle in <i>linker DNA</i>	θ	
Number of bps in a <i>NCP DNA</i> period	h'	
<i>NCP DNA</i> double helix period	p'	
Twist angle in <i>NCP DNA</i>	θ'	
Number of bps in a DNA helix period at rest	h_0	10.5 bp/turn
DNA double helix period at rest	p_0	3.4 nm/turn
Twist angle between adjacent bps at rest	θ_0	0.60 rad/pb
Nucleosome (histones+dna) radius	R	4.18 nm
Nucleosome helix pitch	P	2.39 nm
Number of bps in a nucleosome (repeat length)	N_r	200 bp ^b
Number of bps in the NCP	n_{NCP}	
Number of bps in the crystallographic NCP	n_{NCP}^0	133 bp ^b
Total NCP overtwist	$\Delta t w_{\text{NCP}}$	
Dihedral angle between entering and exiting linkers	α	
Angle α in the crystallographic structure	α_0	0.94 rad
Dihedral angle between neighboring NCP axis	β	
Angle β for relaxed (untwisted) linkers at $\alpha = \alpha_0$	β_0	46.2 rad ^b

^a For free B-DNA

^b Indicative value, see the main text for a derivation

The remaining $N_r - n_{\text{NCP}}$ base pairs form a DNA segment, called *linker DNA*, which is almost straight and connects two neighboring NCP, this forming the characteristic “beads on a string” structure of dilute fibers.

Helical periodicity and NCP overtwisting In the following, we will indicate with h (resp. h'), $p = xh$ (resp. p') and $\theta = 2\pi/h$ (resp. θ') the DNA geometrical parameters in the linker (resp. in the NCP). The action of external constraints can twist the linker DNA, and thus modify its helical periodicity h .

More delicate, the DNA periodicity in the NCP is also influenced by the constraints imposed by the SHL anchoring points on the DNA. In its crystallographic structure, the nucleosomal DNA is bound at all the SHL so that its minor groove is perfectly phased and oriented toward the NCP axis at the entering and exiting points. Nevertheless, the precise number of wrapped base pairs n_{NCP} can vary depending on the precise DNA sequence. As a consequence, the helical periodicity h' in the NCP (sometimes referred to as the *local* h_{local} parameter) can vary from nucleosome to nucleosome.

Due to the superhelical wrapping around a cylindrical support, the resulting helical periodicity h' differs in any case from the intrinsic DNA periodicity as measured in the laboratory frame (Sivolob and Prunell 2004). In the case of a left-handed

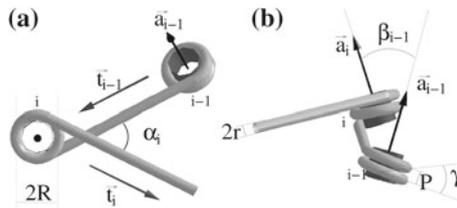


Fig. 2 The two angle model. **a** View down the NCP axis \mathbf{a}_i . **b** View down the linker direction \mathbf{t}_{i-1} . The angle α_i is the angle $(\mathbf{t}_{i-1}, \mathbf{a}_i, \mathbf{t}_i)$ and β_{i-1} the angle $(\mathbf{a}_{i-1}, \mathbf{t}_{i-1}, \mathbf{a}_i)$ standing for the twist (*modulo* 2π) of the DNA linker. We also indicate the DNA radius r , the NCP radius R and pitch P , and the superhelix azimuthal angle γ

superhelix and in the *absence* of additional twist in the wrapped DNA, it can be calculated (Le Bret 1988) and is smaller than the intrinsic helical periodicity; if the intrinsic helical periodicity equals $h_0 = 10.5$ bp/turn in free DNA, one obtains $h' = 10.35$ bp/turn.

In several cases h' can be experimentally estimated. It corresponds indeed to the spacing between the most internal regions that are in contact with the surface, or, equivalently, between the most external regions of each strand. These latter regions are maximally exposed to external attack, and h' can e.g. be determined by in situ digestion with DNase I. The lengths of the single-stranded fragments obtained after histone extraction, measured by gel electrophoresis, are multiples of h' . Experimental estimates of h' spread over a wide range (9.7–11 bp/turn), even if most of them are between 10.2 and 10.3, the mean of all values being $\langle h' \rangle = 10.24$ bp/turn (Prunell 1998). Interestingly, these values indicate that some amount of *overtwisting* does exist in NCPs ($h' < 10.35$ bp/turn). This extra-twist in the NCP derives from the precise number of base pairs included between the most external SHL: if the anchoring point is move to the next DNA base pair on both the ± 6.5 SHL, then n_{NCP} increases by 2 and consequently h' increases so to keep fixed the entering and exiting DNA phasing, by screwing an additional DNA length in the NCP. Note that, if the DNA axis path is conserved, this also implies a decrease of the local inter-basepair distance x .

The 30 nm fiber In the following, we will describe the fiber geometry by means of the original two-angle model introduced by Woodcock et al (Woodcock and Dimitrov 2001), which assumes that all linkers are straight. In order to define the fiber geometry, therefore, one only need to assign two angles per nucleosome, α_i and β_i (Fig. 2). In the following, we will often deal with homogeneous fibers. In that case, we will note $\alpha_i = \alpha$, $\beta_i = \beta$. Let us use this simplified notation for the following general consideration.

The precise definition of α is the angle between the projections of the entering and exiting linkers onto the plane perpendicular to the axis of the NCP; β is the dihedral angle between neighboring NCP axis with respect to their connecting linker. In other words, α is related to the wrapping angle of DNA around NCP (which amounts to $3\pi + \alpha$), while β accounts for the rotation of the next NCP with respect to the previous one, depending on the linker helical phasing and length.

The model allows α and β to assume any value in a continuous interval. For real nucleosomes, the α angle is indeed quantified and assume discrete values depending on how many SHL anchoring points are formed. A continuous α can somehow account for the relative flexibility of the structure, induced by a slight linker bending, for instance as a consequence of the binding of linker histones (proteins known to make a chromatin fiber more compact by binding, and bending, its linkers). The classical two-angle model can be adapted to these deformed fiber geometries by slightly varying the α angle with respect to its canonical discrete values.

The reference α_0 and β_0 An important reference value for α can be obtained from the X-ray crystallographic data. The best fit of the double helix axis as obtained by crystallography leads to a superhelix path with a 4.18 nm radius and a 2.39 nm pitch (Luger et al. 1997). The crystallographic superhelix corresponds thus to 1.65 turns, this covering 126 base pairs (the remaining 10 bp segments at both ends are essentially straight). As a consequence, the angle α for this reference crystallographic NCP structure, to which we will refer as α_0 , amounts to $\alpha_0 = 1.65 \cdot 2\pi - 3\pi = 0.3 \pi$ or 0.94 rad (54°).

In this crystallographic configuration all the 14 SHL anchoring points are bound, this leading to 13 double helix periods in the wrapped DNA. Therefore, the total number of wrapped base pairs in the crystallographic NCP structure can be derived from the h' estimate as $n_{\text{NCP}}^0 = 13 h'$ (where the apex indicate that we are considering the reference $\alpha = \alpha_0$ state. For the mean value $\langle h' \rangle = 10.24$ bp/turn one obtains $n_{\text{NCP}}^0 \simeq 133$ bp.

The angle β depends on the degree of overtwist of the linker (accounted for equivalently by θ or p), but also on the linker length itself. Evidently, the addition *at fixed* α of one more base pair in the linker will lead to rotation of the next NCP of θ , this increasing β of the same amount. Nevertheless, β depends on α in a non trivial way: the position of the DNA grooves at the exiting point of the previous NCP introduces indeed a phase to be taken into account. In the following we will deal in fact with the problem of deriving the correct dependence of β on p . It will be useful to introduce a reference value β_0 for β , corresponding to the the crystallographic NCP structure with *relaxed* (untwisted) linkers. As we have seen, in this case α equals α_0 , and exactly $n_{\text{NCP}}^0 = 13 h'$ bp are wrapped into the NCP. All the anchoring points are bound, so that the DNA minor groove exactly face the inner NCP core at both the entry and exit points of each NCP. Thus, no phase shift affect the calculation of the linker twist and we can write

$$\beta_0 = (N_r - n_{\text{NCP}}^0) \theta_0 = 2\pi \frac{(N_r - n_{\text{NCP}}^0)}{h_0}. \quad (1)$$

By taking e.g. a repeat of $N_r = 200$ bp as a a reference value, we obtain $\beta_0 = 77 \theta_0 \simeq 14.67 \pi$ (or 46.1 rad, corresponding to 7 turns plus about 120°) Table 2.

2.2 Parameterization of the NCP left-handed helix

In the two-angle model, the repeat is composed of two parts: a superhelical DNA with solenoidal shape in the NCP, followed by a straight linker DNA, eventually twisted.

Table 2 NCP parameters for some different NCP configurations^(*)

Parameter	Symbol	5S	Mean values ^(*)	Untwisted NCP
Number of bps in a DNA period in the NCP	h' (bp/turn)	10.13	10.24	10.38
DNA double helix period in the NCP	p' (nm/turn)	3.24	3.28	3.32
Twist angle in NCP DNA	θ' (rad/bp)	0.62	0.61	0.60
Total NCP overtwist	$\Delta t w_{\text{NCP}}$ (turns)	0.32	0.17	0.

Based on the averaged $\langle h' \rangle$ given in [Prunell \(1998\)](#)

In order to calculate the contribution of these two parts, we first need to evaluate their relative length. Given the repeat length N_r , the corresponding total nucleosomal DNA length is $L = N_r \cdot x$. Let us denote by $\ell(\alpha)$ the length of DNA wrapped on the octamer whose wrapping angle is $3\pi + \alpha$. The left-handed solenoid described by the DNA axis is represented here by the parametric curve

$$\mathbf{r}(\sigma) = \left(-R \cos\left(\frac{2\pi\sigma}{P}\right), R \sin\left(\frac{2\pi\sigma}{P}\right), \sigma \right) \tag{2}$$

with σ the arc-length along the NCP axis, $0 \leq \sigma \leq \frac{3\pi+\alpha}{2\pi} P$. Let s be the arc-length along the solenoidal curve, with $ds = \left| \frac{d\mathbf{r}}{d\sigma} \right| d\sigma$. The length of the wrapped DNA can therefore be calculated as

$$\begin{aligned} \ell(\alpha) &= \int ds = \int_0^{\frac{3\pi+\alpha}{2\pi} P} \sqrt{\left(\frac{2\pi R}{P}\right)^2 + 1} d\sigma = \int_0^{\frac{3\pi+\alpha}{2\pi} P} \frac{\sqrt{P^2 + 4\pi^4 R^2}}{P} d\sigma \\ &= \sqrt{4\pi^2 R^2 + P^2} \frac{3\pi + \alpha}{2\pi}, \end{aligned} \tag{3}$$

while the linker length amounts obviously to $L - \ell(\alpha)$.

The other important quantity that will be useful for the following calculations is the the tortuosity τ of the DNA axis wrapped in the NCP. In order to define it, we need to introduce first the tangent, normal and binormal vector relative to the DNA path, defined respectively by

$$\mathbf{t} = \frac{d\mathbf{r}}{d\sigma} \Big/ \left| \frac{d\mathbf{r}}{d\sigma} \right| = \frac{P}{\sqrt{4\pi^2 R^2 + P^2}} \frac{2\pi R}{P} \left(\sin\left(\frac{2\pi\sigma}{P}\right), \cos\left(\frac{2\pi\sigma}{P}\right), 1 \right), \tag{4}$$

$$\mathbf{n} = \frac{d\mathbf{t}}{d\sigma} \Big/ \left| \frac{d\mathbf{t}}{d\sigma} \right| = \left(\cos\left(\frac{2\pi\sigma}{P}\right), -\sin\left(\frac{2\pi\sigma}{P}\right), 0 \right), \tag{5}$$

$$\mathbf{b} = \mathbf{t} \wedge \mathbf{n} = \frac{P}{\sqrt{4\pi^2 R^2 + P^2}} \left(\sin\left(\frac{2\pi\sigma}{P}\right), \cos\left(\frac{2\pi\sigma}{P}\right), -\frac{2\pi R}{P} \right). \tag{6}$$

The local torsion τ_L has been defined by White and Bauer (White and Bauer 1988a,b; White et al. 1986) by the relation

$$\frac{d\mathbf{b}}{ds} = \frac{d\sigma}{ds} \frac{d\mathbf{b}}{d\sigma} = -\tau_L \mathbf{n} \tag{7}$$

that gives, after substitution of Eqs. 4–6

$$\tau_L = -\frac{2\pi P}{4\pi^2 R^2 + P^2}. \tag{8}$$

The tortuosity of the DNA axis wrapped in the NCP corresponds then to the absolute value of the total torsion T obtained by integrating the local torsion τ_L over *one* helical turn of the DNA axis path. Following White and Bauer in (White and Bauer 1988a,b; White et al. 1986):

$$\begin{aligned} \tau = |T| &= \left| \frac{1}{2\pi} \int_0^\ell \tau_L ds \right| = \left| \frac{1}{2\pi} \int_0^P -\frac{2\pi P}{4\pi^2 R^2 + P^2} \frac{\sqrt{P^2 + 4\pi^4 R^2}}{P} d\sigma \right| \\ &= \frac{P}{\sqrt{4\pi^2 R^2 + P^2}} \end{aligned} \tag{9}$$

3 Calculating the DNA linking number in the fiber

3.1 DNA twist

Although the twist is an extensive quantity, its calculation for DNA in a fiber is rather subtle. Nevertheless, it can be done at the level of a single nucleosome as a function of angles α_i and β_i , without need of considering the whole structure. For this reason, we will neglect again the index i in most part of this section, for the aim of simplicity.

The calculation of the nucleosome twist presented here follows approximatively White and Bauer’s derivation and use the results presented in their work (White and Bauer 1988a,b; White et al. 1986). Let us denote by H the curve in space describes by one DNA strand, and by C the DNA axis trajectory. In order to apply the White–Fuller theorem to DNA in the fiber, one should therefore calculate the twist of H about C. The sum of this quantity and of the writhe of the DNA axis C gives the topological invariant linking number.

Nucleosomal twist and NCP overtwist Following again White and Bauer, and using the notation introduced before, it is then easy to express the two twist contributions as follows:

$$tw_{\text{linker}} = \frac{L - \ell(\alpha)}{p} = \frac{N_f - n_{\text{NCP}}}{h} \tag{10}$$

$$tw_{\text{NCP}} = \frac{\ell(\alpha)}{p'} - \tau = \frac{3\pi + \alpha}{2\pi} = \frac{n_{\text{NCP}}}{h'} - \tau = \frac{3\pi + \alpha}{2\pi} \tag{11}$$

The right side expressions in term of base pair numbers are more convenient in our case, because of the fact that x , the proportionality parameter between h and p , can vary in the overtwisted NCP.

In both twist amounts, the main contribution simply accounts for the number of turns one DNA strand makes about the double helix axis. Note that, for $\alpha = \alpha_0$, the contribution to the NCP twist always amounts exactly to $n_{\text{NCP}}^0/h' = 13$, because there are 13 double helix periods in the wrapped DNA. In order to obtain an expression for any α we will assume that the DNA twist is uniform inside the NCP and write

$$tw_{\text{NCP}} = 13 \frac{n_{\text{NCP}}}{n_{\text{NCP}}^0} - \tau \frac{3\pi + \alpha}{2\pi} \quad (12)$$

The additional term for the NCP twist arises from the fact that the DNA axis is not straight. This contribution corresponds to the twist of the NCP straight axis about the DNA helix axis (White and Bauer 1988a,b; White et al. 1986), or, in other words, to the twist induced by the shape of the DNA axis path on a untwisted (Frenet) ribbon following the same path. It is accounted for by the tortuosity parameter, and its negative sign derives from the left-handedness of the NCP helix.

It can be useful to determine an expression for the overtwist of a particular nucleosome to compare it with experimental data. In general, the experimental result can be expressed in terms of the overall NCP overtwisting Δtw_{NCP} , i.e. the difference between the twist of the wrapped DNA (taken at $\alpha = \alpha_0$) and the twist of a straight and relaxed DNA stretch of same length. From Eq. 11 one gets

$$\Delta tw_{\text{NCP}} = 13 - \frac{n_{\text{NCP}}^0}{h_0} - \tau \frac{3\pi + \alpha_0}{2\pi} \quad (13)$$

(see Table 1 for reference values). We stress again that, given the crystallographic form of the considered NCP, Δtw_{NCP} is a constant, which only depends on the exact number n_{NCP}^0 of base pairs between the two external SHL at $\alpha = \alpha_0$.

Relation between β and the twist

The total twist of a nucleosome is therefore given by the sum of the two contributions Eqs. 10 and 11. In order to introduce an explicit twist dependence on β , we will calculate tw and β as a function of the wrapped base pairs n_{NCP} , then recombine the two expressions.

Let us start with the simpler case when $\alpha = \alpha_0$. From Eqs. 10 and 11 we get immediately, for a given linker period h ,

$$tw(\alpha_0, h) = (N_r - n_{\text{NCP}}^0) \frac{1}{h} + 13 - \tau \frac{3\pi + \alpha_0}{2\pi} \quad (14)$$

where we had explicitly indicated the dependence on α and p . As already noticed for β_0 , the angle β is simply expressed at $\alpha = \alpha_0$ in terms of the linker length as

$$\beta(\alpha_0, h) = (N_r - n_{\text{NCP}}^0) \frac{2\pi}{h} \pmod{2\pi}. \quad (15)$$

Therefore, at $\alpha = \alpha_0$ we can write

$$tw(\alpha_0, h) = \frac{\beta(\alpha_0, p)}{2\pi} + 13 - \tau \frac{3\pi + \alpha_0}{2\pi} \pmod{2\pi}. \tag{16}$$

It is possible to get rid of the (modulo 2π) limitation in the previous expression if $\beta(\alpha_0, p)$ is close enough to β_0 . In this case, we have at a time

$$(N_r - n_{\text{NCP}}^0) \frac{2\pi}{h} = \beta(\alpha_0, h) + 2k\pi \quad \text{and} \tag{17}$$

$$(N_r - n_{\text{NCP}}^0) \frac{2\pi}{h_0} = \beta_0 + 2k\pi \tag{18}$$

with the same k .

After a little algebra, this leads to

$$tw(\alpha_0, h) = \frac{\beta(\alpha_0, h) - \beta_0}{2\pi} + \Delta tw_{\text{NCP}} + \frac{N_r}{h_0} \tag{19}$$

where Δtw_{NCP} is given by Eq. 13 and N_r/h_0 represents the twist of a free DNA segment of equal length.

We need now to address the general case of $\alpha \neq \alpha_0$. The question is how the twist tw and β angle will vary if DNA is further wrapped, or unwrapped.

For aim of clarity, we start by considering the case of a double helix wrapped on a cylinder *without tortuosity*, i.e. wrapped on a perfect circular path, with $\tau = 0$. In this case the variation $d\beta/2\pi$ equals the variation of twist $d tw$, as we will now demonstrate. If no additional twist exists in the NCP, the angle β is always equal to the rotation angle around the linker axis leading the contact reference frame in $\ell(\alpha)$ to coincide to the reference frame in L . The wrapping of a DNA length $d\ell$ will simply rotate the contact frame from the position ℓ to $\ell + d\ell$ around the NCP axis. This transformation induces a rotation of the $(\ell + d\ell, L)$ linker segment around the NCP axis too, and therefore it does not modify the angle β . Nevertheless, in presence of an additional NCP twist, i.e. when $h \neq h'$ and $p \neq p'$, the segment $(\ell, \ell + d\ell)$ change its twist rate while wrapped, this leading to a rotation of the next par of the linker and therefore to a variation of β of

$$\frac{d\beta}{2\pi} = \left(\frac{1}{p'} - \frac{1}{p} \right) d\ell = \left(\frac{1}{h'} - \frac{1}{h} \right) dn_{\text{NCP}}. \tag{20}$$

If $\tau = 0$, this is exactly equal to the twist variation: $d\beta/2\pi = d tw$. Let now consider the current case where the double helix wraps following a left-handed helical path and the tortuosity τ is therefore nonzero. The twist variation should thus account for the tortuosity effect and we simply get

$$d tw = \left(\frac{1}{h'} - \frac{1}{h} \right) dn_{\text{NCP}} - \frac{\tau}{2\pi} d\alpha. \tag{21}$$

On the other hand, the variation $d\beta$ is *not affected* by the left-handed tortuosity, so that we obtain

$$d\,tw = \frac{d\beta}{2\pi} - \frac{\tau}{2\pi}d\alpha. \quad (22)$$

To understand why β is unchanged, it could be useful to refer to the case of a ribbon wrapping onto a cylinder following a left-handed helical path. As long as the helical path is regular, the extremities of the ribbon stay oriented exactly in the same way with respect to the cylinder axis when wrapped. In the same way, wrapping DNA onto a NCP does not affect *by itself* the orientation β of the next NCP with respect to the axis of the previous one. At given p , only the variation of twist rate between the NCP and the linker ($h \neq h'$) influences the value of β while α is changed.

By integrating Eq. 22 we get

$$tw(\alpha, \beta) = \frac{\beta}{2\pi} - \frac{\tau}{2\pi}(3\pi + \alpha) + C, \quad (23)$$

To get the integration constant C we calculate Eq. 23 at $\alpha = \alpha_0$ and compare the result to Eq. 19. This leads to $C = \Delta tw_{\text{NCP}} - \beta_0/2\pi + N_r/h_0 + \tau(3\pi + \alpha_0)/2\pi$, and we finally get the expected expression of the nucleosome twist as a function of α and β :

$$tw(\alpha, \beta) = \frac{\beta - \beta_0}{2\pi} - \tau \frac{\alpha - \alpha_0}{2\pi} + \Delta tw_{\text{NCP}} + \frac{N_r}{h_0}. \quad (24)$$

Note that the previous result is immediately generalized to the case of local $\alpha_i \neq \alpha$ and $\beta_i \neq \beta$, simply by replacing α and β in Eq. 24.

Relaxed linker constraint If DNA linkers are too short with respect to the twist persistence length, then they cannot absorb a relevant amount of twist because of the too high energy cost. Therefore, one can be interested in being able to change the fiber parameters in such a way that the linker twist is always zero. This constraint is of course simply expressed by the condition $h = h_0$. In terms of β , we find from Eqs. 1 and 20 that one should choose the relaxed β angle following the relation

$$\beta_{\text{relaxed}} = \beta_0 + 2\pi \left(\frac{1}{h'} - \frac{1}{h_0} \right) [n_{\text{NCP}} - n_{\text{NCP}}^0]. \quad (25)$$

Relative twist In general biological applications it is usual to refer to a *relative* DNA twist, i.e. to calculate the difference between the total twist tw and a reference twist tw_0 , which corresponds to the twist of a straight and relaxed B-DNA of the same length. In our case, the reference nucleosomal value can be written as

$$tw_0 = \frac{N_r}{h_0} \quad (26)$$

i.e. as the last term in Eq. 24. The relative twist is directly obtained as

$$\Delta t w(\alpha, \beta) = t w(\alpha, \beta) - t w_0. \quad (27)$$

We can now come back to the case of an extended fiber containing N nucleosomes, not necessarily identical. As previously mentioned, the twist is an extensive quantity: therefore, the total relative fiber twist is simply the sum of all the nucleosome twists. If α_i and β_i are the two-angle model parameters characterizing the i -th nucleosome, then, using Eq. 24, the total relative twist of DNA in the fiber is:

$$\Delta T w = \sum_{i=1}^N \left(\frac{\beta_i - \beta_0}{2\pi} - \tau \frac{\alpha_i - \alpha_0}{2\pi} + \Delta t w_{\text{NCP}} \right). \quad (28)$$

3.2 DNA writhe

Generalities about writhe calculation Let us denote $\mathbf{r}(s)$ any closed curve with curvilinear coordinate $s \in [0, L]$. As a measure of the curling of the curve in space, writhe essentially depends on its unit tangent vector, $\mathbf{t}(s) \propto \frac{d\mathbf{r}}{ds}(s)$. The computation of the writhe for a closed curve $\mathbf{r}(s)$ of length L and tangent vector $\mathbf{t}(s)$ is already a rather complicated task. It is generally evaluated through a double integral according to a method based on Gauss theorem, whose reads

$$W_r = \frac{1}{4\pi} \int_0^L \int_0^L \frac{(\mathbf{r}(s_1) - \mathbf{r}(s_2)) \cdot (\mathbf{t}(s_1) - \mathbf{t}(s_2))}{|\mathbf{r}(s_1) - \mathbf{r}(s_2)|^3} ds_1 ds_2. \quad (29)$$

The Gauss integral can be efficiently integrated numerically. Nevertheless, we will use here the alternative method introduced by Fuller (Fuller 1978), which has the advantage of a direct geometrical interpretation and of a more straightforward generalization to open curves (Starostin 2005). We will show how the Fuller approach can be developed, for the case of composed, quasi-periodic curves as those of interest here, so to obtain an effective scale separation that simplifies the calculation considerably.

Consider again the unit vector tangent to the curve, $\mathbf{t}(s)$. The tangent vector vertex, $T(s)$, lies on the unit sphere and, as s varies from 0 to L , it describes on this sphere a closed curve, called *tangent indicatrix* (Maggs 2000, 2001). Note in particular that a straight segment in space corresponds to a fixed point on the sphere, and that a local bent is represented by a geodesic arc connecting the entering and the exiting tangent vectors at the bending point (Maggs 2000, 2001). Fuller's first theorem states that the writhe of the curve $\mathbf{r}(s)$ can be calculated from the signed area \mathcal{A} enclosed by the tangent indicatrix $T(s)$, namely $W_r = \mathcal{A}/2\pi - 1 \pmod{2}$. The congruence modulo 2 represents a first difficulty for a complete writhe calculation. Anyway, Fuller's second theorem (Fuller 1978) permits to get rid of this congruence, if one calculates the writhe difference between two closed curves, $\mathbf{r}_1(s)$ and $\mathbf{r}_2(s)$, under a set of hypotheses that

we will discuss in a while. This difference can be expressed as a single integral, which has the form

$$\Delta Wr = \frac{1}{2\pi} \int_0^L \frac{\mathbf{t}_1(s) \wedge \mathbf{t}_2(s)}{1 + \mathbf{t}_1(s) \cdot \mathbf{t}_2(s)} \frac{d}{ds} (\mathbf{t}_1(s) + \mathbf{t}_2(s)) ds \quad (30)$$

Some attention has to be paid in applying this formula. The hypotheses of the theorem impose indeed that one of the curves can be obtained from the other by a continuous deformation in space such that (1) at any point along the curve and at any moment during the transformation, the tangent vector should not assume opposite direction with respect to any other intermediate state, (2) the curve is non-self-intersecting all along the transformation, and (3) the tangent to the curve varies continuously during the transformation. We will come back to these hypotheses and discuss the applicability of the theorem to our case in Sect. 3.4.

Provided these conditions, it is possible to show that the integral to be calculated corresponds geometrically to the area \mathcal{S} swept out by the unique shortest geodesic arc from the running point $T_1(s)$ to the running point $T_2(s)$ on the unit sphere (Fuller 1978). The writhe difference of Eq. 30 simply amounts therefore to $\mathcal{S}/2\pi$.

It is useful to note that, if one of the curves can be chosen as a straight line, then its tangent vector defines a unique point C on the sphere. In this case, the difference of writhes is given by the area swept out by the unique shortest geodesic arc from the fixed point C to the running point $T(s)$. Of course, the direction of the straight line (the point C) can be chosen arbitrarily to some extent, provided that the theorem hypotheses are not violated.

Further difficulties arise for the calculation of the writhe of an open curve. Except for the case when the initial and final tangent vectors $\mathbf{t}(0)$ and $\mathbf{t}(L)$ coincide, the calculation is rather subtle. We follow the procedure outlined by Maggs (Maggs 2000, 2001) to obtain a consistent measure for the writhe by closing the open tangent indicatrix with a geodesic arc.

Writhe of a straight regular fiber The chromatin fiber is a quite regular structure composed by almost identical repetitive nucleosomes. It would therefore be helpful if one could calculate its writhe from the writhes of the individual units. However, the writhe is known to be non-extensive and, as shown by Starostin (Starostin 2002)¹, 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 (Fig. 3) closing each part. However, as shown by Fuller (Fuller 1978), under some hypotheses (that reflect those of the second Fuller theorem discussed above) the writhe of a given curve can be expressed as the writhe of a reference curve plus a sum of locally determined writhe differences,

¹ The 2002 work of Starostin has been published in 2005 (Starostin 2005), but the preprint version contains more material than its published version, and in particular a whole section devoted to the writhe additivity, to which we refer in our paper.

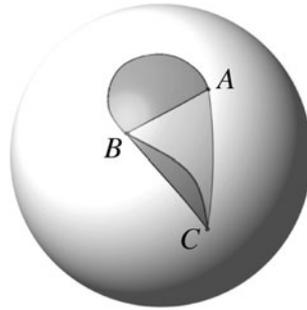


Fig. 3 A schematic view of the writhe calculation for an open curve. The *curved line* on the sphere represents the tangent indicatrix of the corresponding open curve in space, the initial and final tangent vector vertices going from point *A* to point *C*. The curve writhe is calculated by closing it with the geodesic *AC*. If the curve is separated into two parts, then the writhe (times 2π) can be calculated as the sum of the surfaces of the two parts, each closed by an indicatrix (dark grey areas) plus the surface of the spherical triangle formed by the three geodesics *AB*, *BC*, *CA*. Note that all the surfaces are signed: positive if the border curve is followed anti-clockwise, negative if it is followed clockwise. The overall surface for the case in figure is therefore the area enclosed by the tangent indicatrix *ABC* and the geodesic *CA*, as expected

each of which expresses the effect of altering a single section of the reference curve. In the following, we will derive an expression for the DNA writhe *per nucleosomes* by following an original procedure, based on the Starostin approach. The final result is in agreement with the mentioned Fuller theorem and allows to the calculation of the DNA linking number in a fiber as the sum of the linking number of the fiber “carrying” structure plus the nucleosomes contributions.

In the case of a chromatin fiber, the DNA indicatrix $T(s)$ can be naturally divided into N parts, namely the N nucleosomes, in correspondence to $\mathbf{t}_i = \mathbf{t}(s_i)$, the direction of the (straight) linker i . Denote by T_i the point on the unit sphere that corresponds to \mathbf{t}_i . The total writhe of the DNA in the fiber is then given, accordingly to Starostin (2005), by the following addition rule:

$$Wr = \sum_{i=1}^N wr_{NCP}(\alpha_i) + \frac{S_{T_0 T_1 \dots T_N}}{2\pi}. \tag{31}$$

Here, $wr_{NCP}(\alpha_i)$ is the writhe of the i -th nucleosome *whose indicatrix had been closed by a geodesic*, while $S_{T_0 T_1 \dots T_N}/2\pi$ is the area enclosed in the spherical polygon of vertices $T_0, T_1 \dots T_N$, and represents the writhe of the “carrying” structure formed by the sequence of linkers. In the following, we will refer to this structure as to the *linker skeleton*. Striking speaking, the previous formula is only valid modulo 1. Nevertheless, if $wr_{NCP}(\alpha_i)$ can be calculated exactly and $S_{T_0 T_1 \dots T_N} < 2\pi$, the writhe results to be a continuous function of i . In absence of self intersections, it is therefore possible to use Eq. 31 to get the exact writhe of the curve (Starostin 2005). We will now discuss separately the two terms of Eq. 31.

The single nucleosome writhe $wr_{NCP}(\alpha_i)$ should be calculated as the surface enclosed by the restriction of the tangent indicatrix between T_{i-1} and T_i and the closing geodesic $T_{i-1}T_i$. The calculation is easily carried out as it reduces to the

writhe of a left-handed helix segment. For a nucleosome of a given entry-exit angle α_i , given the radius R and the pitch P , one gets²

$$wr_{\text{NCP}}(\alpha_i) = \frac{1}{\pi} \left[\arctan \left(\tau \tan \left(\frac{3\pi + \alpha_i}{2} \right) \right) - \tau \frac{3\pi + \alpha_i}{2} \right] + \text{round} \left(\frac{3\pi + \alpha_i}{2\pi} \right), \quad (32)$$

where $\text{round}(x)$ is the nearest integer to x . The term $\text{round}(3\pi + \alpha_i/2\pi)$ in Eq. 32 accounts for the integer part of the twist, i.e. for the full number of turns made by the DNA axis around the NCP. Note that $wr_{\text{NCP}}(\alpha_i)$ only depends on α_i .

Let us proceed to a simple verification of the validity of Eq. 32. When the wrapping angle $3\pi + \alpha_i$ is set to 2π , one obtains $wr_{\text{NCP}}(\alpha_i) = 1 - \tau$ with $\tau = P/\sqrt{P^2 + 4\pi^2 R^2}$. This correctly amounts, as expected, to $2\pi^{-1}$ times the surface of the spherical segment delimited by the circle described by $T(s)$, i.e. the “parallel” identified by the superhelix azimuthal angle γ , with $\tan(\gamma) = P/(2\pi R)$, and $\tau = \sin(\gamma)$.

An alternative expression for the nucleosome writhe arises from a different calculation of the corresponding surface. This surface can be evaluated indeed by taking a fraction $(3\pi + \alpha_i)/2\pi$ of the spherical segment area, and by completing it by the surface of the spherical triangle identified by the entry and exit linker tangent vectors and by the NCP axis. Following this procedure, one can write

$$wr_{\text{NCP}}(\alpha_i) = (1 - \tau) \frac{3\pi + \alpha_i}{2\pi} + \frac{\mathcal{S}_{T_{i-1}T_iA_i}}{2\pi}. \quad (33)$$

The area of the spherical triangle $T_{i-1}T_iA_i$ could then be calculated by the usual procedure: the signed area of a spherical triangle ABC (unit vectors \mathbf{a} , \mathbf{b} , \mathbf{c}) has absolute value equal to the sum of the angles ($\angle ABC + \angle BCA + \angle CAB - \pi$) and sign given by $\text{sgn}((\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c})$:

$$\mathcal{S}_{ABC} = \text{sgn}((\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c}) (\angle ABC + \angle BCA + \angle CAB - \pi). \quad (34)$$

The second term $\mathcal{S}_{T_0T_1\dots T_N}$ in Eq. 32 is the non-extensive contribution, given by the signed area of the spherical polygon connecting all points T_i that correspond to the linker directions. Nevertheless, it should be very useful, at this point, to “remove” the non-extensivity by an appropriate partition of this area. In principle, $\mathcal{S}_{T_0T_1\dots T_N}$ can be computed as the sum of the areas of the spherical triangles $T_0T_{i-1}T_i$ ($1 < i \leq N$). We observed moreover that, once chosen an arbitrary point C on the unit sphere (corresponding to some unit vector \mathbf{c}), the area $\mathcal{S}_{T_0T_1\dots T_N}$ can also be written as

$$\mathcal{S}_{T_0T_1\dots T_N} = \sum_{i=1}^N \mathcal{S}_{CT_{i-1}T_i} + \mathcal{S}_{CT_NT_0}, \quad (35)$$

² Compare with Eq. 16 in Starostin (2005). Incidentally, we point out an error in the round term of that formula, that should be rewritten as $+\text{round}(aL/2\pi)$.

$CT_{i-1}T_i$ being the spherical triangles formed by the three geodesics joining points C , T_{i-1} and T_i . This partition allows a natural decomposition of the writhe into single nucleosome contributions: up to a boundary term $\mathcal{S}_{CT_N T_0}/2\pi$, the writhe can be expressed as

$$Wr = \sum_{i=1}^N wr_i|_C, \tag{36}$$

where we define the writhe $wr_i|_C$ of nucleosome i with respect to point C as

$$wr_i|_C = wr_{NCP}(\alpha_i) + \frac{\mathcal{S}_{CT_{i-1}T_i}}{2\pi}. \tag{37}$$

To summarize, the introduction of a reference point C and an appropriate redefinition of the writhe of each nucleosome with respect to this point allow us to express the total writhe as an *extensive quantity*.

Moreover, the introduction of this reference point allows a further and important simplification. Up to now, the spherical triangle contributions $\mathcal{S}_{CT_{i-1}T_i}$ differ from each other. Let consider, as the simplest case, a straight, perfectly regular fiber, with all the angles α_i (respectively, β_i) equal to a common value α (respectively, β). Even in this case, $wr_i|_C$ varies from nucleosome to nucleosome, in spite of the fact that $wr_{NCP}(\alpha_i)$ are the same for all nucleosomes. Nonetheless, there exists a special point, $C = F$, defined by the director of the *fiber axis*, for which

$$\mathcal{S}_{FT_{i-1}T_i} = \mathcal{S}(\alpha, \beta) \quad \forall i. \tag{38}$$

For this particular choice, then, we can define an effective *writhe per nucleosome*, $wr(\alpha, \beta)$ such that

$$wr_i|_F = wr(\alpha, \beta) = wr_{NCP}(\alpha) + \frac{\mathcal{S}(\alpha, \beta)}{2\pi} \tag{39}$$

and
$$Wr = \sum_{i=1}^N wr_i|_F = N wr(\alpha, \beta). \tag{40}$$

We reiterate that the effective writhe *per nucleosome* $wr(\alpha, \beta)$ is not the bare writhe of one individual nucleosome, the difference being the surface of the spherical triangle FT_0T_1 .

The particular choice of the reference point F ³ leads thus to an expression for the writhe which is not only additive, but also provides identical contributions for each nucleosome. This supplies, at the same time, a consistent definition of the writhe *per nucleosome*, a quantity which is independent of the fiber length N and is usually referred to in the biological literature. The result, summarized by Eqs. 39 and 40, is

³ We note that the same viewpoint F was introduced by Crick (1976).

also very useful from a practical point of view: it allows for the calculation the DNA writhe in a fiber without need of integration or sum.

3.3 Writhe of a bent fiber

We will show now how the previous definitions can be extended to this more general case of a smoothly bent fiber, with angles α_i and β_i slightly varying around their mean values $\bar{\alpha}$ and $\bar{\beta}$. In Sect. 3.5, we will derive a corresponding linking number. We will see how a separation of scale naturally arises from the generalization to bent fibers, leading in turn to introduce a consistent definition for the writhe and twist of the *fiber* itself.

We start by noting that, for a generic fiber, a local fiber axis F_i can be defined for each i as the axis of a straight fiber characterized by constant angles $\alpha = \alpha_i$ and $\beta = \beta_i$, and containing i th nucleosome. It is therefore still possible to define an *effective* nucleosome writhe as

$$wr(\alpha_i, \beta_i) = wr_i|_{F_i} = wr_{\text{NCP}}(\alpha_i) + \frac{\mathcal{S}_{F_i T_{i-1} T_i}}{2\pi}. \quad (41)$$

In contrast with the case of a regular fiber (Eq. 40), the total writhe is no more the sum of the contributions $wr(\alpha_i, \beta_i)$. Indeed, the surface of the spherical polygon $T_0 T_1 \dots T_N$ indeed does not simply amounts to the sum of the spherical triangles $F_i T_i T_{i+1}$ defined with respect to points F_i . We can anyway introduce again an external point C , and write the spherical polygon surface as a sum of spherical triangles areas. The total writhe finally reads

$$Wr = \sum_{i=1}^N wr(\alpha_i, \beta_i) + \sum_{i=1}^{N-1} \frac{\mathcal{S}_{F_i T_i F_{i+1}}}{2\pi} + \sum_{i=1}^{N-1} \frac{\mathcal{S}_{C F_i F_{i+1}}}{2\pi} \quad (42)$$

up to the boundary term $\mathcal{S}_{C F_N F_0} + \mathcal{S}_{F_N T_N T_0 F_0}$. Equation 42 generalizes the calculation of the relative writhe to the case of an irregular fiber. In Sect. 3.6 we will be able to give a geometrical meaning to the two additional terms in this equation. Note that these terms vanish as $F_i \rightarrow F \forall i$, i.e. for a straight fiber. In this case, Eqs. 41 and 42 coincide with 39 and 40, respectively.

3.4 Fuller theorem applicability and some precisions

Let us now come back to a more technical point. In the derivation of Eqs. 39 and 40, we used areas on the unit sphere to calculate writhe. We are indeed in the frame of the second Fuller's theorem. In order to validate *a posteriori* our result for the straight fiber, we have then to check if the Fuller's theorem hypotheses are valid in this case.

The second Fuller's theorem is supposed to give the writhe *difference* between two states: we start therefore by defining, as a reference state, a straight DNA parallel to the fiber axis, and we try to continuously modify it to obtain the final DNA path in the

fiber. The first step of the transformation will be the bending of our DNA line to form a sequence of straight segments oriented as the fiber linkers, i.e. forming the “carrying” linker skeleton structure. Now, it is possible to show that all the possible *regular* linker skeleton structures that can be obtained in the framework of the two-angle model are of solenoidal shape (Ben-Haim et al. 2001), where the fiber axis direction \mathbf{f} (of vertex F) is the direction of the growing helix. The linker directions T_i and the fiber direction F then always belong to the same hemisphere. The transformation from the initial state to this intermediate configuration is therefore possible without infringing the theorem hypotheses: that tangent vector $\mathbf{t}(s)$ never happens to be oppositely directed to its direction at any other transformation step, $\forall s$, without self-intersections all along the transformation, and with the tangent to the curve varying continuously⁴.

The second intermediate step is the addition of small segments parallel to the fiber axis at the location of each nucleosome. Again, this can be done in the framework of the Fuller’s hypotheses due to the helical orientation of the linkers. The second reference structure has therefore all linkers parallel to their final direction and all the NCP DNAs replaced by straight DNA segments parallel to the fiber axis. Note that, on the unit sphere, this leads to a polygon of zero surface, because all odd vertices coincide with F . The writhe of this intermediate step is therefore zero, as that of the initial straight DNA.

To obtain the complete fiber, we have now to replace each of the NCP segments by the corresponding DNA path in the nucleosome. This can be done in two further steps: first, by a distortion of the straight segments into NCP helices whose axis is parallel to the fiber axis ; then by rotating them to obtain their final orientation.

Now, NCP DNAs are also described by helices: the first distortion is therefore allowed, since the NCP tangent vector describes a growing curve. The rotation step may, instead, induce some violation of the theorem hypotheses. Note however that, in the spirit of the approach outlined by Fuller (1978), the theorem only has to be respected “locally”, since what is evaluated here is a writhe difference between a referring and a final state and nucleosomes may be “treated” sequentially. The point is therefore to ensure that during the rotation of one nucleosome to its final orientation, the tangent vector to any of his points never happens to be oppositely directed to the fiber axis. This is in fact ensured, with one single exception, by a geometrical properties of the fibers. The NCP exhibits indeed pseudo twofold symmetry. The axis of symmetry lies in the plane of the nucleosome disk (perpendicular to the superhelical axis), and is called the dyad axis. In a fiber, the dyad axis is always perpendicular to the fiber axis (Ben-Haim et al. 2001). Therefore, the tangent vector always have a non-zero component along the dyad axis, that is orthogonal to the fiber axis, with the exception of the two points where the DNA path intersect the dyad axis. These two points are the closest and farthest points on DNA with respect to the fiber axis, and are the unique points where the tangent vector can be opposite with respect to the fiber axis. But this only happens for a particular fiber structure, with almost horizontal nucleosomes axis. With the exception of this pathological case, the required condition is verified.

⁴ The fact that in the skeleton structure the tangent vector presents discontinuous points may appear as a limitation to the application of the theorem, but this problem may be easily solved by defining a correct limiting procedure (Starostin 2005)

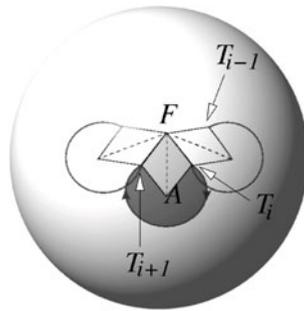


Fig. 4 A schematic view of the writhe calculation for a nucleosome in a regular fiber. Scales and length have been arbitrarily chosen for the sake of clarity. The *curve* represents the tangent indicatrix of DNA in the fiber. Point T_i represent the vertex of the i th linker tangent vector. The *circular part* correspond to the indicatrix of the DNA wrapping around the nucleosome. The point F is the direction of the fiber axis. The surface $\mathcal{S}(\alpha, \beta)$ of Eq. 38 corresponds to the spherical triangle FT_iT_{i+1} . If the nucleosome axis A is taken into account, the writhe per nucleosome $wr(\alpha, \beta)$ of Eq. 39 is thus given by 2π times the sum of the *dark grey* and *light grey areas* in the figure

Having started from the straight fiber axis, we then finally obtain the actual DNA path in the fiber. Since the intermediate state has zero writhe, the final writhe is calculated as the sum of the writhes of all the nucleosomes with respect to their reference (intermediate) state, which coincides with the fiber axis, as in the definition of the writhe *per nucleosome* of Eq. 39 and as shown in Fig. 4.

We can conclude that Fuller's second theorem applies, and the quantity Wr in Eq. 40 correctly gives the difference in writhe between the DNA in the given fiber and that of a straight DNA, parallel to the fiber axis. This straight DNA will represent our *referring state*. It is interesting to note that a further rotation in space of the reference straight DNA again give no variation in writhe. This is also the reason why, in our derivation of the total fiber writhe, any arbitrary point C (never opposite to linker directions) can be used, in principle. It is easy to show by geometrical considerations that the choice of the reference point only effects the boundary terms in Eq. 36.

3.5 The DNA linking number and the fiber twist

An expression for the DNA linking number According to our definitions, the linking number of a straight unconstrained DNA of total length L_{tot} is $Lk_0 = N_{tot}/h_0$, i.e. it just amounts to its natural twist. It is therefore straightforward to consider the relative linking number of DNA in the fiber, $Lk - Lk_0$. By using Eqs. 28 and 42 we get the following expression for the relative DNA linking number

$$\Delta Lk = Lk - Lk_0 = \sum_{i=1}^N \left[wr(\alpha_i, \beta_i) + \frac{\beta_i - \beta_{0i}(\alpha_i)}{2\pi} - \tau \frac{\alpha_i - \alpha_0}{2\pi} + \Delta t w_{NCP} \right] + \sum_{i=1}^{N-1} \left[\frac{\mathcal{S}_{F_i T_i F_{i+1}}}{2\pi} + \frac{\mathcal{S}_{C F_i F_{i+1}}}{2\pi} \right]. \quad (43)$$

The expression of ΔLk can be putted in a simpler form by noting that the results of the preceding paragraphs allow the definition of a relative DNA linking number of the i^{th} nucleosome as

$$\Delta lk(\alpha_i, \beta_i) = \Delta tw(\alpha_i, \beta_i) + wr(\alpha_i, \beta_i), \tag{44}$$

where $\Delta tw(\alpha_i, \beta_i)$ comes from by Eqs. 24 and 27 and $\Delta wr(\alpha_i, \beta_i) = wr(\alpha_i, \beta_i)$ is given Eq. 41.

We can therefore write the relative linking number of DNA in the fiber as

$$\Delta Lk = \sum_{i=1}^N \Delta lk(\alpha_i, \beta_i) + \sum_{i=1}^{N-1} \frac{\mathcal{S}_{F_i T_i F_{i+1}}}{2\pi} + \sum_{i=1}^{N-1} \frac{\mathcal{S}_{C F_i F_{i+1}}}{2\pi}, \tag{45}$$

where $\Delta lk(\alpha_i, \beta_i)$ is the relative linking number of nucleosome i calculated with respect of the local fiber axis, and the latter sum two latter sums account for the fiber axis variations.

3.6 From DNA to fiber topology

It is interesting to look for the topological meaning of the additional terms in Eq. 45. A comparison with Eq. 35 shows that the very last term has exactly the same structure of the skeleton contribution to the writhe, with linker directions T_i replaced by local fiber axis directions F_i . Therefore, this term is nothing but the writhe of the *fiber axis* with respect to the fixed point C , i.e. to a straight fiber. We will call it the *fiber writhe* and denote it as $Wr^{\mathcal{F}}$:

$$Wr^{\mathcal{F}} = \sum_{i=1}^{N-1} \frac{\mathcal{S}_{C F_i F_{i+1}}}{2\pi}. \tag{46}$$

We can therefore identify ΔLk with the linking number of the fiber itself, $Lk^{\mathcal{F}}$. Hence, the remaining terms in Eq. 45 must correspond to the twist of the fiber around its axis,

$$Tw^{\mathcal{F}} = \sum_{i=1}^N \Delta lk(\alpha_i, \beta_i) + \sum_{i=1}^{N-1} \frac{\mathcal{S}_{F_i T_i F_{i+1}}}{2\pi}, \tag{47}$$

in order to satisfy the White–Fuller theorem at the level of the fiber, $Lk^{\mathcal{F}} = Wr^{\mathcal{F}} + Tw^{\mathcal{F}}$.

The fiber twist thus contains two terms. The first one amount to the sum of the DNA linking number of the i th nucleosome $lk(\alpha_i, \beta_i)$, i.e. the nucleosome lk contributions “as if” it was inserted in the *straight* fiber of given α_i and β_i . Interestingly, this first contribution can be interpreted therefore as a *local* fiber twist, similar to what discussed for the DNA twist in the NCP. This suggests that the second term contribution to

$T w^{\mathcal{F}}$ may be associated to the fiber tortuosity. Indeed, the tortuosity τ is obtained by integrating the local torsion $\tau_L = d\mathbf{b}/ds$ over the helical curve (Eq. 9).

We stress that the previous definition of the fiber topological quantities has an important meaning either *in vitro* and *in vivo*. In magnetic tweezers micromanipulation experiments, a fiber is fixed at both ends, to the glass support and the magnetic bead respectively. The fiber linking number is then modified by rotating the beads of a given number of turns. The fiber rearrangement should thus be compatible with the imposed torsion. In living cells, on the contrary, transitions between compacted and decompact chromatin loops are observed. These loops are likely to be clamped at their ends: in this case, fiber rearrangements should thus modify the fiber length at constant linking number. It is therefore tempting to consider the fiber as a continuous rod, so to define its twist and writhe in this standard frame. Nevertheless, the chromatin fiber is a deformable object with internal degrees of freedom. The formal expressions 46 and 47 show how complicated a correct definition of the fiber writhe and twist is.

However, at least for the case of a straight fiber, simple topological considerations suggest that the previous expressions should coincide with an intuitive definition of the fiber twist and writhe. For instance, an intuitive definition of the fiber twist can be given, at least for the case on a straight fiber. In this case, it should correspond to the rotation angle of the fiber “top” with respect to the fiber “bottom”: one can consider for instance the rotation of the last linker with respect to the first one, once projected on the plane perpendicular to the fiber axis. For a curved fiber, it is instead the writhe of the fiber axis which is easier defined, at least formally; on the contrary, it is more delicate to obtain an intuitive but correct definition of the fiber twist. Our expressions 46 and 47 allows for a compromise between intuition and accuracy. They lead when possible to results that can easily compared with the overall behavior of the fiber.

4 A fast method for the calculation of fiber linking number variations

4.1 Decimal part of a fiber linking number

The linking number of a given fiber is a number roughly of the order of the number of nucleosomes and should be calculated as described in the previous section. Nevertheless, if one is only interested in the *decimal part* of this quantity, an alternative, faster calculation can be performed. Moreover, this calculation applies to any open DNA chain, following a given geometry in space, and therefore can be useful in studying very inhomogeneous fibers as e.g. fibers with an arbitrary variable spacing between nucleosomes.

Consider such an open DNA chain arranged in space following a given geometry and possibly twisted. The curve described by the DNA axis is $\mathbf{r}(s)$ with s the arc-length along the DNA. The tangent vector $\mathbf{t}(s)$ is defined as in Eq. 4, and follows the tangent indicatrix on the unit sphere. If $\mathbf{t}(s)$ is discontinuous in some points, the vertices corresponding to the two limit vectors will be closed by a geodesic arc. Note that the DNA path $\mathbf{r}(s)$ is only needed in order to define the normal and binormal vectors, defined in Eqs. 5 and 6. The normal vector $\mathbf{n}(s)$ in particular allows to define an *intrinsic* ribbon following the DNA path $\mathbf{r}(s)$, and called the Frenet ribbon, a special

ribbon whose normal coincide with the principal normal vector to the curve. The DNA ribbon, intended as the ribbon whose normal is e.g. directed from the DNA axis toward the minor groove (or any other equivalent choice), does not coincide with the Frenet one, and is generally much more twisted due to the DNA helical shape. Let us indicate the DNA normal vector as $\mathbf{x}(s)$.

Let us indicate the Frenet ribbon, defined by $\mathbf{r}(s)$, $\mathbf{t}(s)$ and $\mathbf{n}(s)$, by $\mathcal{R}_{0,L}$. We will start by *closing* it by a Frenet ribbon $\mathcal{R}_{L,0}^c$ build in such a way that its tangent vector follows the geodesic between $\mathbf{t}(L)$ and $\mathbf{t}(0)$. The tangent vector to $\mathcal{R}_{L,0}^c$ is therefore always on a plane, and the writhe of the closing ribbon is therefore zero. Moreover, its twist also vanish because the twist of a planar Frenet ribbon is always zero. Nevertheless, a contribution to the twist of the closed ribbon $\mathcal{R}_{0,L} \cup \mathcal{R}_{L,0}^c$ arise from the two connections at $s = 0$ and $s = L$. At these positions, the normal to the closed Frenet ribbon is discontinuous: the derivative of the tangent vector is different if calculated along $\mathcal{R}_{0,L}$ or along $\mathcal{R}_{L,0}^c$. Let us indicate by $\mathbf{n}^+(0)$, $\mathbf{n}^-(L)$ the normal vectors on the $\mathcal{R}_{0,L}$ side and $\mathbf{n}^+(L)$, $\mathbf{n}^-(0)$ those on the $\mathcal{R}_{L,0}^c$ side. The connection between the two ribbons involves therefore a local twist contribution that amount to

$$Tw^c = \frac{\angle(\mathbf{n}^-(0), \mathbf{n}^+(0))}{2\pi} + \frac{\angle(\mathbf{n}^-(L), \mathbf{n}^+(L))}{2\pi} \tag{48}$$

where $\angle(\mathbf{v}, \mathbf{w})$ represents the signed angle between vectors \mathbf{v} and \mathbf{w} . This is the unique contribution of the closing ribbon to the overall linking number of the closed ribbon $\mathcal{R}_{0,L} \cup \mathcal{R}_{L,0}^c$: $Lk^c = Tw^c$. On the other hand, the linking number is an integer, as for any closed curve. As a consequence, we get for the linking number Lk_{Frenet} of the DNA Frenet ribbon $\mathcal{R}_{0,L}$

$$Lk_{Frenet} = \text{integer} - Tw^c \tag{49}$$

The writhe contribution to the Frenet linking number Lk_{Frenet} exactly equals the DNA writhe Wr , as it only depends on the geometrical path in space. On the contrary, Lk_{Frenet} evidently does not account for the actual DNA twist Tw . Indeed, the Frenet normal vector does not coincide with the real DNA normal, as we have seen. In order to correctly calculate the open DNA linking number, we therefore need to calculate its twist, i.e. how the DNA normal $\mathbf{x}(s)$ rotate with respect to the Frenet local frame defined by $\mathbf{n}(s)$. This quantity is obviously difficult to calculate, but its decimal part is simply given by the relative position of the $\mathbf{x}(s)$ and $\mathbf{n}(s)$ vectors at the ending points $s = 0$ and L . The exceeding twist ΔTw of the open DNA with respect to the Frenet ribbon thus reads

$$\begin{aligned} \Delta Tw &= \text{integer} + \frac{\angle(\mathbf{x}(0), \mathbf{n}^+(0))}{2\pi} + \frac{\angle(\mathbf{n}^-(L), \mathbf{x}(L))}{2\pi} \\ &= Lk - Lk_{Frenet} \end{aligned} \tag{50}$$

All together, Eqs. 48, 49 and 50 allow the following expression for the linking number Lk of the open DNA:

$$Lk = \text{integer} + \frac{\angle(\mathbf{x}(0), \mathbf{n}^-(0))}{2\pi} + \frac{\angle(\mathbf{n}^+(L), \mathbf{x}(L))}{2\pi} \quad (51)$$

where we have performed the difference between angles in Equation 50 and 48.

The previous formula enables us to calculate very simply the linking number of a fiber or any other DNA arrangement modulo an integer. Of course, the method does not allow for the calculation of the twist and writhe contributions separately. Nevertheless, it can be useful in several cases. One first application is e.g. an independent calculation of the NCP overtwisting of a crystallographic DNA. For this aim it is sufficient to consider the DNA path inside the NCP, built the corresponding Frenet ribbon and close it by a second Frenet ribbon as we just explained. At the entering and exiting points, the DNA minor groove is always oriented toward the NCP axis, and we know that the DNA twist around the Frenet ribbon amounts to 13 turns, which corresponds to the integer in Eq. 50. On the other hand, the writhe wr_{NCP} can in this case explicitly calculated (see Eq. 32). For $\alpha = \alpha_0$ we obtain $wr_{\text{NCP}} = -1.794$. From the DNA linking number expression 51 one can therefore deduce the NCP twist as

$$tw_{\text{NCP}(\alpha_0)} = \text{integer} + \frac{\angle(\mathbf{x}(0), \mathbf{n}^-(0))}{2\pi} + \frac{\angle(\mathbf{n}^+(L), \mathbf{x}(L))}{2\pi} - wr_{\text{NCP}}. \quad (52)$$

This expression allows to check the previous formula for the twist of the NCP Eq. 11 and represents an interesting alternative for its calculation.

4.2 Progressive calculation of a linking number variation

A second, important advantage of the calculation of the linking number decimal part just introduced is the possibility of considering very distorted fibers through a progressive calculation. Starting from a regular fiber whose linking number can be exactly calculated, one can indeed modify it slowly toward a different configuration, and calculate, step by step, a linking number variation δLk by mean of Eq. 51. If the modification step is small enough to make the linking number variation always less than 1, thus δLk can be integrated and the linking number of the final configuration obtained. Interestingly, the detailed geometry of the deformed fibers could be ignored, and only the position of the reference DNA normal and Frenet normal vectors at the entering and exiting points are needed. This is particularly adapted to the study of fiber micromanipulation experiments (see the next section).

4.3 Calculation of the linking number of a fiber without information on the linker path

A final remark concerning the linking number calculation concerns the case where the precise arrangement in space of NCPs in a given fiber is known, but the linker paths are not. This case is practically relevant since the NCP arrangement can be obtained experimentally, for instance by electron microscopy, by imaging highly compact fibers reconstructed *in vitro* (Robinson et al. 2006). However, the obtained images do not

give any information on the positions of linkers DNA and the modelling of the whole fiber architecture still represents a debated problem (Ball 2008; Wong et al. 2007). Interestingly, for the calculation of the fiber linking number, the exact arrangement of linkers inside the fiber does not matter! Imagine indeed to start with the “real” fiber geometry and to deform (bent, kink or twist) linkers, without cutting them and by keeping the NCP in their initial position. The fiber extremities will not move, and the transformation is therefore performed at constant linking number. The linking number of the fiber thus only depend on how the NCP are positioned in space (and on how they are ordered along the fiber) but not on the precise path of the linkers. A proper choice of α and β , reproducing the NCP arrangement, is therefore sufficient in order to correctly evaluate the fiber topology.

5 NCP plasticity: the four nucleosome states

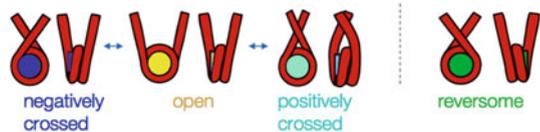
In this section, we illustrate one direct application of the previous theoretical derivation to the interpretation and fitting of experimental data. The results have been obtained in close collaboration with the experimental group of J.L. Viovy. Our contribution in modeling the fiber structure and quantifying its relevant features allowed us to interpret the results and propose new mechanisms

5.1 The nucleosome plasticity as revealed by magnetic tweezers

In 2006, Bancaud et al. obtained the first magnetic tweezers manipulation of single chromatin fibers, and showed that chromatin can accommodate surprisingly large amount of torsional stress, either negative or positive, without much change in its extension (Bancaud et al. 2006, 2007; Recouvreur et al. 2011). In these two studies, nucleosome arrays were reconstituted on 2×18 tandem repeats of the 208 bp 5S nucleosome positioning sequence. They were subsequently ligated at each end to a naked DNA spacer labeled appropriately for attachment of one end of the construct to the coated bottom of the flow cell and of the other end to a paramagnetic bead. The rotation of the paramagnetic bead exerts some torsion on the attached fiber. The fiber torsional behavior is described, at a given force, by its extension-to-rotation curve. The torsional response of naked DNA is characterized by a bell-shaped curve where the two quasi-linear compactions on either side correspond to the formation of positive (resp. negative) plectonemes upon introduction of positive (resp. negative) rotation.

Compared to its naked DNA template, the chromatin fiber is shorter and the center of rotation of its extension-to-rotation curve is shifted towards negative values. This shift is the expected consequence of the absorption of approximately one negative superhelical turn per nucleosome. The observed shortening of about 50 nm, (i.e., 150 bp) per nucleosome results from DNA wrapping around the histone core. Moreover, compared to naked DNA of the same length, the fiber in the elastic regime appears to be much more torsionally flexible: it can absorb large amounts of torsion without much shortening (Bancaud et al. 2006). This large torsional resilience can be explained by taking into account three different nucleosome states, previously identified in minicircles studies (De Lucia et al. 1999; Sivolob et al. 2003), and characterized

Fig. 5 Schematic view of the four states of the nucleosome. See main text for details



by a different pathway for the entering and exiting linker DNAs (Fig. 4). While the *negatively crossed* state corresponds to the standard crystallographic structure, with straight linker DNAs following the direction of the entering and exiting tangent to the nucleosome superhelix, in the *positively crossed* state the two DNAs slightly bind to cross each other in a positive way. Finally, in the *open* state, both $+6.5$ and -6.5 SHLs unbind and the linker DNAs do not cross any more. As will be detailed in the next section, the precise contribution to the fiber linking number of these different states is different, this being the key to interpret the observed fiber resilience. The three conformations have indeed different free energies, but their stability also depends on the applied torque: the application of a positive (resp. negative) torque to the fiber will indeed favor the states with the most positive (resp. negative) linking number *per* nucleosome, and the resulting nucleosome dynamic equilibrium between the different conformational states acts as a topological buffer, absorbing the additional torsion thanks to the rearrangement of the internal nucleosome structure (Bancaud et al. 2006).

More unexpected, chromatin fibers display a hysteretic behavior when submitted to extensive positive supercoiling (Bancaud et al. 2007). This hysteresis was interpreted as a consequence of the trapping of positive turns in individual nucleosomes through their transition to an altered form, called reversome (Bancaud et al. 2007) (for reverse nucleosome, see Fig. 5) or R-octasome (Zlatanova et al. 2009). The altered form should be metastable in order to explain the presence of hysteresis, and has to be stabilized by the application of a large positive torque. Moreover, the reversome must contribute to the fiber linking number with a large positive value, in order to explain the trapping of positive turns. The structure of this altered nucleosome is a priori not known, but we proposed a plausible structure based on suitable simulations, as we will describe in the next section.

5.2 The linking number of the four nucleosome states

The negatively crossed and open states NCPs therefore exist in four different geometries. Among them, the negatively crossed and open states simply correspond to two particular choices of the two-model α parameters. Indeed, $\alpha = \alpha_0 = 0.94$ rad (54°) reproduces the crystallographic configuration where all the 14 SHL anchoring points are bound (see Sect. 2.1). The negatively crossed NCP coincides with this structure if straight linkers are added by prolonging the natural direction of entering and exiting DNA. The open state corresponds instead to the case where the two external SHL are broken, again with straight linkers. In this case, only 12 SHL anchoring points are bound, this leading to 11 double helix periods in the wrapped DNA. This leads to the value $\alpha_{open} \simeq -0.65$ rad (-37°), corresponding to a wrapping of about 1.4 turns. The linking number of open and negatively crossed nucleosomes in homogeneous fibers

can be therefore simply determined by applying the previous results (Expression 44) to these particular values of the α parameter and a β angle corresponding to a relaxed linker DNA. The calculation gives $Lk_{neg} = -1.4$, $Lk_{open} = -0.4$, in quite good agreement with the values determined by minicircle studies (De Lucia et al. 1999; Sivolob et al. 2003), of -1.4 and -0.7 , respectively.

The positively crossed state The calculation of the topological quantities for the positively crossed state of the nucleosome is more complicated, since it involves curved linkers. In this torsion-induced structure each linker is constrained and bent in the opposite direction with respect to the nucleosomal DNA super-helix sense. The two linkers therefore pass one under the other, and are kept in this distorted conformation by the steric interaction that prevents them to cross each other.

The possibility of having bent linkers is not accounted for by the two angle model. We thus modified the model in order to evaluate the linking number of a positively crossed nucleosome. A simple way of doing this is by introducing symmetric DNA *kinks* at the entry and exit points, then adding straight linkers. The kinks should be adapted as to make the two linker DNAs cross positively. The kink angle γ is therefore set in such a way that the entering and exiting linkers “touch” each other, in order to take into account the steric interaction that keep the positively crossed NCP in its bend conformation. In other words, for a given α , the vectors tangent to DNA at the entry and exiting points are rotated with respect to a rotation axis which is orthogonal to their initial orientation and to the NCP axis. The rotation angle γ is calculated to make that the minimal distance between the two linker axes equal to twice the DNA diameter. Such a procedure allows us to obtain positively crossed linkers with the introduction of a single additional degree of freedom, which can furthermore be fixed by the condition of contact between linkers.

In practice, the addition of two symmetric kinks in the linkers can easily be introduced in the calculation of the DNA writhe. A kink in a straight DNA corresponds to a discontinuity in its tangent vector, and is therefore associated with a geodesic on the unit sphere. The tangent path on the sphere is such that the entry and exit tangents are no more coincident and that an intermediate point, corresponding to the linker direction, appears between them. This path is sketched in Fig. 6.

The calculation of the writhe is then extended to the positively crossed nucleosome by including two more spherical triangle areas to the standard calculation. Moreover, an additional twist term should be taken into account when DNA is locally kinked. Using the α and γ parameters to fit the experimental data from magnetic tweezers allow to fix their values with a reasonable precision. In the case of the results of Bancaud et al. (2006), with a repeat length of 208 bp, we have chosen $\alpha = 30^\circ$ and $\gamma = -10^\circ$, this leading (using Expression 44) to $Lk_{pos} = -0.2$ (to be compared to -0.4 as obtained in minicircle studies (De Lucia et al. 1999; Sivolob et al. 2003)).

The reversome The observation of an hysteresis cycle of the extension-to-rotation curve when the fiber is strongly positively supercoiled led to the assumption that the nucleosome undergoes an important structural change. The new structure has to trap a part of the torsional constraint, and to release it once the applied torque is reduced. An analysis of the difference in the linking number *per nucleosome* between the onward

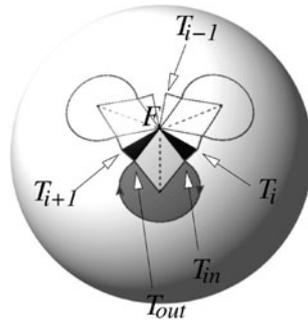


Fig. 6 A schematic view of the writhe calculation for a positively crossed nucleosome in a regular fiber. With respect to the negatively crossed state, the entering and exiting tangent vectors vertices T_{in} and T_{out} are distinct from the linker tangent vector vertices T_i and T_{i+1} . The indicatrix has therefore a more complex path, composed by the geodesic $T_i T_{in}$, the circular path corresponding to the wrapping of DNA around the nucleosome, then the geodesic $T_{out} T_{i+1}$. When calculated with respect to the fiber axis F , the writhe is therefore proportional to the sum of signed area of the *black*, *dark grey* and *light grey* parts

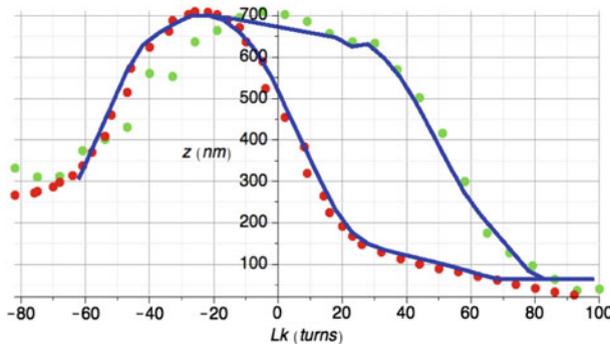


Fig. 7 Fit (blue solid line) of the experimental extension-to-rotation data obtained by Bancaud and co-workers (same of Fig. 5B in Bancaud et al. (2007)): the onward curve (dark red points) is obtained by applying an increasing torsion, the backward curve (light green points) by relaxing back the fiber by rotating its extremity in the opposite way. As for the case of a free DNA (Strick et al. 1998) the almost linear part of the onward curve corresponds to the region where the fiber wraps on itself forming plectonemes; on the corresponding linear part in the backward curve, plectonemes are unwound (Bancaud et al. 2006). In these regions the nucleosome distribution among the different states are supposed to remain constant, with only standard nucleosomes for the onward curve and only reversomes for the backward one. Therefore, the difference in the number of turns between the two linear parts allows to the estimation of the difference in linking number between the nucleosome and the reversome (Bancaud et al. 2006) (color figure online)

and backward curves (see Fig. 7) reveals that the difference in linking number between the “standard” and new conformations is of about 2 turns.

In the initial conformation, the “standard” nucleosome is a mixed population of the three states (negatively crossed, open and positively crossed), whose relative weights depend on the applied torque and can be determined at each point on the hysteresis curve by fitting the fiber extension (see Recouvreux 2011 for further details). By this procedure we obtained an averaged value of the “standard” nucleosome linking number of about -1 turn. Hence, the linking number achieved by the altered

nucleosome is close to +11 turn (more precisely, we get about 0.9 turns). This value is close to the linking number that has been calculated for the right-handed *tetrasome*, i.e. the particle formed by the most internal two (H3–H4) dimers once removed the H2A–H2B ones. Such tetrasomes are known indeed to fluctuate between “pseudo-mirror-symmetrical” left- and right-handed chiral conformations of nearly equal and opposite Lk [–0.7 and +0.6 for 5S DNA sequences (Sivolob and Prunell 2004)]. Based on this conformational change of the tetrasome, we proposed for the altered form of the reversome an almost “mirror-inversed” structure, obtained through a chiral inversion of the nucleosome and implying the breaking of internal bonds (called the *docking domains*) between the dimers H2a–H2B and the tetramer (H3–H4)₂ (Bancaud et al. 2007). The proposed structure remains largely speculative and we do not have any direct information on its geometry, but it is compatible with the experimental results. In this case, we are indeed in an inverse situation: magnetic tweezers manipulations give us an estimation of the reversome linking number, but no direct information on its molecular structure.

As a first attempt to have quantitative estimates of the quantities involved, we then built a simple model of the reversome as chiral symmetric of the standard nucleosome (Fig. 5). The best fit of the experimental data for the case of 208 bp repeat leads us to choose an α angle for the reversome configuration of 30°. In this case, the linking number contribution *per nucleosome* is 0.85. This naive model of the reversome structure matches quite remarkably the all-atom structure that we obtained alternatively by inverse kinematics (Bancaud et al. 2007; Zlatanova et al. 2009). Moreover, geometrical and topological parameters of this tentative structure can be calculated and used to fit the experimental extension-to-rotation curve. The goodness of the fit may be seen as a confirmation of the chosen structural model.

5.3 The fit of the experimental data

Homogeneous and non homogeneous fibers The interpretation and fitting procedure that we applied to the experimental results uses a statistical mechanics approach accounting for the evolution of the four state populations as a function of the applied torque, hence of the rotation applied to the fiber extremity. The fiber conformation at a given step of the experiment is a disordered mixture of nucleosome states in thermodynamic equilibrium. The linking number calculation that we developed so far for homogeneous fibers has to be extended to heterogeneous fibers. For this purpose, we use a mean field approximation, by means of which we can compute the linking number of heterogeneous fibers as the sum of the linking numbers *per nucleosome* over the set of nucleosome states in the fiber. This mean field procedure amounts to use the same value for the linking number *per nucleosome* as in a homogeneous fiber. See Refs. (Bancaud et al. 2006, 2007; Recouvreux et al. 2011) for further details.

The linking number dependence on the applied torque Finally, it is interesting to note that, since the experimental setup implies the application of an increasing torque to the fiber, the DNA torsion varies all along the extension-to-rotation curve. As a consequence, the geometrical parameter β varies accordingly: it is therefore important

to take into account the variation of the linking number of all the different states as β is modified. This can be easily done by the calculation method previously described, leading to a linking number *per nucleosome* which is a function of the applied torque. In practice, for the configurations intervening in the described experiments, this effect is only relevant for the reversome, whose linking number dependence on the applied couple can be reasonably approximated by a linear function. For all the other states, instead, the linking number is approximately constant in the range of the applied torques.

Obtained fit Figure 7 presents the resulting fit of the experimental data obtained by Bancaud and co-workers on a fiber containing 30 nucleosomes. We have obtained by the same procedure similar fits for the data obtained by Recouvreur et al. (see Fig. 4 in [Recouvreur et al. 2011](#)), where more details on the fitting procedure are also given.)

The good performances obtained in fitting the data show that this mean field approach (i.e. thermodynamic averaging of the nucleosome states *before* summing their contribution to the fiber linking number) is in good agreement with the exact calculation (i.e. thermodynamic averaging *after* summing the contribution of the nucleosome states to the fiber linking number). We wish to stress that the fit of the experimental curve is particularly sensitive to the “extremal” fiber configurations, where the nucleosomes are almost all either in the negatively crossed, or positively crossed, or reversome states. The exact conformation of the intermediate states is less crucial in determining the curve behavior. Our approximation method becomes more appropriate for these more homogeneous configurations, this explaining at least in part the success of our fitting procedure.

To summarize, in this section we have shown how our calculations of the topological properties of nucleosomes in fibers can be applied (or extended) to the three different nucleosome conformations observed in minicircle studies ([De Lucia et al. 1999](#); [Sivolob et al. 2003](#)). Moreover, they can be useful in testing the plausibility of new speculative nucleosome structures, as for the case of the reversome. A careful characterization of all these structures is indeed necessary in order to interpret experimental results that, as in the case of magnetic tweezers, only allow for the measurement of indirect and averaged quantities.

6 Physiological implications

We wish to conclude by illustrating the above calculations with two examples of applications of biological interest. In both cases, our conclusions are quite speculative, but suggest physically plausible mechanisms that could be relevant in different biological situations.

6.1 Fiber condensation at constant linking number

A first problem that can be addressed is how the DNA can be highly condensed in chromatin. From the point of view of the fiber, this packing can be achieved in two ways: a high packing ratio of DNA in the chromatin fiber and/or a winding pattern



Fig. 8 Compaction (from *left to right*) of a fiber at constant linking number in the framework of the standard two angle model. The final state is a “two-start helix” and is approximately the most compact structure that can be obtained by the standard two-angle model. Note that the number of nucleosomes of the compact fiber is larger than for the other two structures

of the fiber itself. In the context of a chromatin loop, however, the fiber is anchored at both ends by cross-linking protein complexes such as insulators, and the DNA linking number has therefore to be kept constant during those two conformational changes. Modeling the geometrical and topological properties of the chromatin fiber gives insights into how the decondensation processes can occur with the imposed constraint. Our calculation allows to propose two plausible scenarios of elongation and unwinding at constant linking number.

Starting from the homogeneous two angle model, a first level of compaction can be obtained by shortening the fiber in a homogeneous way by continuously varying the angles α and β (Fig. 8). In order to study the effect of such variations, we have obtained a chart displaying both the compaction and the linking number *per nucleosome* (Eq. 45) of regular fibers as functions of α and β for a given repeat length (Barbi et al. 2005). We have then considered the possible paths leading from decondensed to condensed fibers with a minor or negligible variation of the linking number. We have found that, starting from a decondensed beads-on-a-string-like state (see Fig. 8, left), the angle α of all the nucleosomes can be uniformly decreased while the angle β is slightly modified (compatibly with the DNA twist persistence length), in such a way that the DNA linking number is kept constant and the fiber compaction increases up to about 6 nucleosomes per 10 nm (Barbi et al. 2005). This leads close to the most compact fiber that can be obtained in the two-angle model, and which is formed by the wrapping of two columns of stacked nucleosomes one around each other (Fig. 8, right). Linker DNAs are found in the center of the fiber and are cross-linking those two columns. It was therefore named as the “two-start helix” or the “cross-linker structure”. This structure has been already reported in several experimental studies such as crystallography, neutron scattering and electron-microscopy (Routh and Sandin 2008; Schalch et al. 2005; Williams et al. 1986). The obtained compaction level corresponds to the one found for isolated fibers *in vitro* (Williams et al. 1986) and may be assumed to be relevant *in vivo* during interphase. In this scenario, the compaction/decompaction of a compact fiber is achieved at constant linking number and almost without DNA twisting, so to avoid any coiling or entangling of the fiber.

The compact fiber of Fig. 8 can be further condensed by changing its axis trajectory in order to form a coiled structure. However, if the fiber coiling should be performed at constant linking number, a structural rearrangement of the internal fiber geometry is needed in order to compensate the fiber writhe. We have proposed (Mozziconacci et al. 2006) that an internal structural change of the nucleosome, which we named “gaping”, could solve this problem. The gaping structural change corresponds to detaching two

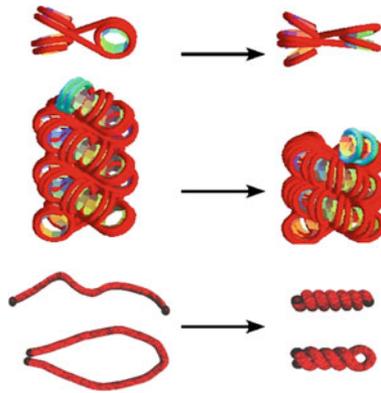


Fig. 9 Models of nucleosome and fiber before gapping (on the *left*) and after gapping (on the *right*). *Top* Three consecutive nucleosomes in a fiber. The gapping process induces the perfect stacking of two neighboring nucleosomes. *Center* side view of the chromatin fiber. The DNA of the top nucleosome has been highlighted in *cyan*. *Bottom* Models of chromatin loop before and after gapping. Depending on the geometry of the chromosome network, loops can either coil into solenoids (*upper figure*) or plectonemes (*lower figure*)

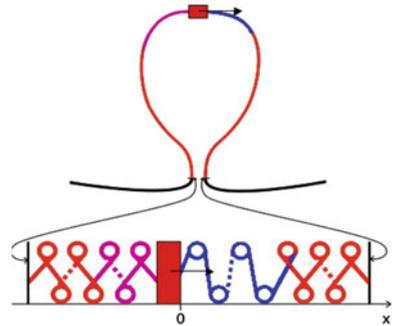
histone dimers from each other, leading to an opening of the nucleosome in the manner of a gaping oyster (Fig. 9). If this opening is accompanied by a twist of 2 bp per linker, the external faces of neighboring nucleosomes come into close contact, so to interact and to stabilize the fiber. To model this conformational change of the nucleosome we added a third angle, corresponding to the gaping angle, to the two angle model. The calculation carried in this context have been obtained using the fast method described in Sect. 4 and gives a change in the fiber Lk , induced by the gaping of all the nucleosomes, of about 7 turns per 1,000 nucleosomes, i.e. approximately a fiber loop. Hence, in order to preserve its global linking number, the compact fiber has to be writhed at an inverse rate of about -7 turns per loop. This compensatory writhe can either result in a toroidal or a plectonemic supercoiling (Fig. 9). Remarkably, the changes in the writhe evaluated for the two conformations in Fig. 9 perfectly match the twist induced by the gaping of all nucleosomes within the loop. Nucleosome gaping may thus not only compact the fiber, but may also be the driving mechanism for supercoiling the fiber loop in a condensed higher-order structure.

6.2 The fiber can accommodate transcription-induced topological constraints

The second question that we can address to illustrate the use of the method described in this paper is the problem of how transcription can be achieved in the context of chromatin. We discussed this problem in details in Bécavin et al. (2010). DNA transcription induces important topological constraints on the fiber. In fact, the DNA double helix should rotate through the RNA-polymerase, in order to allow this enzyme to open, read and copy one of the two strands.

Once the gene promoter has been located to a transcription factory, RNA-polymerases are loaded onto the gene. Transcription elongation leads to an accumulation of DNA positive supercoiling upstream and negative supercoiling down-

Fig. 10 Schematic representation of the progress of a reversome front (blue, on the right) along a compact fiber of nucleosomes, induced by the advancing RNAP (red square and rectangle). The two ends of the fiber are supposed to be clamped in a loop, as frequently observed in cells (color figure online)



stream from the elongation site, while the RNA-polymerase progresses at a speed of about 2 turns per second (Uptain et al. 1997). The negative supercoiling can be simply absorbed by denaturing DNA. The positive supercoiling on the other hand has to be absorbed by a change in the fiber twist and/or writhe. This positive supercoiling is produced at the impressive rate of one turn per 10 base pairs (one helical period), and cannot therefore be absorbed locally. An efficient manner to absorb this constraint may be to reverse all nucleosomes in the upstream fiber into reversomes. Since the linking number of the standard nucleosome is on average about -1 turn and that of the reversome $+1$ turn, about two turns can be absorbed for each nucleosome. Let us take an example to illustrate this point. In the human β -globin gene cluster, a fiber of about 100 nucleosomes (representing the distance between the transcription site and the loop end) can absorb approximately 200 turns corresponding to a transcription length of about 2 kbp, the approximative length of a β -globin gene (Wong et al. 2009).

The degree of compaction of the fiber may play here an important role. “Condensed” fibers (with more than 0.5 nucleosome per nm along the fiber) are expected to have a regular structure, since it is favoured energetically by stacking interactions between nucleosomes. It has been shown that a small amount of nucleosome positioning is enough to get a regular structure *in vitro* (Weidemann et al. 2003). Accordingly, we shall consider the regular structure of chromatin fiber established in a previous work as the generic setting (Wong et al. 2007). In this model, nucleosomes faces are stacked producing a strong steric hindrance. Steric hindrance between nucleosomes, in turn, enforces a domino effect where nucleosomes are subjected to the applied torque one by one, while in the case of a decondensed fiber transitions to reversomes will arise at random all along the fiber. A schematic picture of this reversome progressing wave is given in (Fig. 10).

Although this scenario remains for the moment only a speculative suggestion, recent experimental results (Petesch and Lis 2008) showed a wave of destabilization of nucleosomes, which grows very rapidly along the fiber, about 10 times faster than the motion of the RNA polymerase. These observations can be interpreted in the frame of our hypothesis as a reversome wave cleaning the way so that transcription elongation can proceed (Zlatanova and Victor 2009).

References

- Ball P (2008) Pulling our strings. *Chemistry*, World May:50–54
- Bancaud A, Conde e Silva N, Barbi M, Wagner G, Allemand J, Mozziconacci J, Lavelle C, Croquette V, Victor JM, Prunell A, Viovy J (2006) Structural plasticity of single chromatin fibers revealed by torsional manipulation. *Nat Struct Mol Biol* 13:444–450
- Bancaud A, Wagner G, Conde E, Lavelle C, Wong H, Mozziconacci J, Barbi M, Sivolob A, Le Cam E, Mouawad L, Viovy J, Victor J, Prunell A (2007) Nucleosome chiral transition under positive torsional stress in single chromatin fibers. *Mol Cell* 27:135–147
- Barbi M, Mozziconacci J, Victor JM (2005) How the chromatin fiber deals with topological constraints. *Phys Rev E* 71(031):910
- Bécavin C, Barbi M, Victor JM, Lesne A (2010) Transcription within condensed chromatin: Steric hindrance facilitates elongation. *Biophysical J* 98:824–833
- Ben-Haim E, Lesne A, Victor JM (2001) Chromatin: a tunable spring at work inside chromosomes. *Phys Rev E* 64(051):921
- Byrd K, Corces VG (2003) Visualisation of chromatin domains created by the gypsy insulator of *Drosophila*. *J Cell Biol* 162:565–574
- Cook PR (1999) The organization of replication and transcription. *Science* 284:1790–1795
- Crick FHC (1976) Linking numbers and nucleosomes. *Proc Natl Acad Sci USA* 73:2639–2643
- Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* 319:1097–1113
- De Lucia F, Alilat M, Sivolob A, Prunell A (1999) Nucleosome dynamics. iii. histone tail-dependent fluctuation of nucleosomes between open and closed DNA conformations. implications for chromatin dynamics and the linking number paradox. *J Mol Biol* 285:1101–1119
- Fuller FB (1971) The writhing number of a space curve. *Proc Natl Acad Sci USA* 68:815–819
- Fuller FB (1978) Decomposition of the linking number of a closed ribbon: a problem from molecular biology. *Proc Natl Acad Sci USA* 75:3557–3561
- Giaever G, Wang J (1988) Supercoiling of intracellular DNA can occur in eukaryotic cells. *Cell* 55:849–856
- Labrador M, Corces VG (2002) Setting the boundaries of chromatin domains and nuclear organization. *Cell* 111:151–154
- Lavelle C (2007) Transcription elongation through a chromatin template. *Biochimie* 89:516–527
- Lavelle C, Bancaud A, Recouvreur P, Barbi M, Victor JM, Viovy JL (2011) Chromatin topological transitions. *Prog Theor Phys* 191:30–39
- Le Bret M (1988) Computation of the helical twist of nucleosomal DNA. *J Mol Biol* 200:285–290
- Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci USA* 84:7024–7027
- Ljungman M, Hanawalt P (1992) Localized torsional tension in the DNA of human cells. *Proc Natl Acad Sci USA* 89:6055–6059
- Luger K, Maeder AW, Richmond RK, Sargent DF, Richmond TJ (1997) X-ray structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251–259
- Maggs AC (2000) Twist and writhe dynamics of stiff polymers. *Phys Rev Lett* 85:5472–5475. doi:[10.1103/PhysRevLett.85.5472](https://doi.org/10.1103/PhysRevLett.85.5472)
- Maggs AC (2001) Writhing geometry at finite temperature: Random walks and geometric phases for stiff polymers. *J Chem Phys* 114:5888–5896
- Matsumoto K, Hirose S (2004) Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of *drosophila*. *J Cell Sci* 117:3797–3805
- Mozziconacci J, Lavelle C, Barbi M, Lesne A, Victor JM (2006) A physical model for the condensation and decondensation of eukaryotic chromosomes. *FEBS Lett* 580:368–372
- Petesich SJ, Lis JT (2008) Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at hsp70 loci. *Cell* 134:74–84
- Prunell A (1998) A topological approach to nucleosome structure and dynamics: the linking number paradox and other issues. *Biophys J* 74(5):2531–2544
- Recouvreur P, Lavelle C, Barbi M, Conde e Silva N, Le Cam E, Victor JM, Viovy JL (2011) Linkers histones incorporation maintains chromatin fiber plasticity. *Biophysical J* 100:2726–2735
- Robinson PJ, Fairall L, Huynh VA, Rhodes D (2006) Em measurements define the dimensions of the 30-nm chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci USA* 103:6506–6511

- Rossetto V, Maggs AC (2003) Writhing geometry of open DNA. *J Chem Phys* 118:9864–9874
- Routh A, Sandin DS, Rhodes, (2008) Nucleosome repeat length and linker histone stoichiometry determine chromatin. *Proc Natl Acad Sci USA* 105:8872–8877
- Schalch T, Duda S, Sargent DF, Richmond T (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436:138–141
- Sivolob A, Prunell A (2004) Nucleosome conformational flexibility and implications for chromatin dynamics. *Phil Trans R Soc Lond A* 362:1519–1547
- Sivolob A, Lavelle C, Prunell A (2003) Sequence-dependent nucleosome structural and dynamic polymorphism. potential involvement of histone H2B n-terminal tail proximal domain. *J Mol Biol* 326:49–63
- Starostin EL (2002) On the writhe of non-closed curves, arXiv:physics/0212095v1
- Starostin EL (2005) On the writhe of non-closed curves. ch. 26. In: *Physical and numerical models in Knot Theory including applications to the life sciences*, vol 36. World Scientific, Singapore, pp. 525–545
- Strick TR, Allemand JF, Bensimon D, Croquette V (1998) The behavior of supercoiled DNA. *Biophys J* 74:2016–2028
- Uptain S, Kane C, Chamberlin M (1997) Basic mechanisms of transcript elongation and its regulation. *Annu Rev Biochem* 66:117–1726
- Vinograd J, Lebowitz J, Radloff R, Watson R, Laipis P (1965) The twisted circular form of polyoma viral DNA. *Proc Natl Acad Sci USA* 53:1104–1111
- Weidemann T, Wachsmuth M, Knoch TA, Müller G, Waldeck W, Langowski J (2003) Counting nucleosomes in living cells with a combination of fluorescence correlation spectroscopy and confocal imaging. *J Mol Biol* (334)
- White JH, Bauer WR (1988a) Applications of the twist difference to DNA structural analysis. *Proc Natl Acad Sci USA* 85:772–776
- White JH, Bauer WR (1988b) Dependence of the linking deficiency of supercoiled minichromosomes upon nucleosome distortion. *Nucl Acids Res* 17:8527–8535
- White JH, Gallo R, Bauer WR (1986) Calculation of the twist and writhe for representative models of DNA. *J Mol Biol* 189:329–341
- Williams S, Athey B, Muglia L, Schappe R, Gough A, Langmore J (1986) Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys J* 49
- Wolffe A (1998) *Chromatin: structure and function*. Academic Press, London
- Wong H, Victor JM, Mozziconacci J (2007) An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length. *PLoS ONE* 2:e877
- Wong H, Winn PJ, Mozziconacci J (2009) A molecular model of chromatin organisation and transcription: how a multi-RNA polymerase II machine transcribes and remodels the beta-globin locus during development. *BioEssays* 31(12):1357–1366
- Woodcock LC, Dimitrov S (2001) Higher-order structure of chromatin and chromosomes. *Curr Opin Genet Dev* 11(130–135)
- Zlatanova J, Victor JM (2009) How are nucleosomes disrupted during transcription elongation? *HFSP* 3:373–378
- Zlatanova J, Bishop TC, Victor J, Jackson V, van Holde K (2009) The nucleosome family: dynamic and growing. *Structure* 17:160–171