Computation of Writhe in Modeling of Supercoiled DNA

Abstract: We describe four previously unpublished methods allowing the computation of the writhe for a supercoiled DNA molecule modeled by a polymer chain consisting of straight segments. These methods are compared with each other in terms of computational efficiency and the scope of their applicability is discussed. © 2000 John Wiley & Sons, Inc. Biopoly 54: 307–317, 2000

Keywords: Brownian dynamics; linking number; molecular dynamics; Monte Carlo; twist

INTRODUCTION

The concept of writhe is at the crossroads of differential geometry and molecular biology. For a mathematician, the writhe \( W_r \) is the Gauss double integral over a closed curve \( C \) in the three-dimensional space:

\[
W_r = \frac{1}{4\pi} \int_C \int_C \frac{(dr_2 \times dr_1)r_{12}}{r_{12}} \tag{1}
\]

where \( r_1 \) and \( r_2 \) are the points passing along the curve \( C, r_{12} = r_2 - r_1, r_{12} = |r_{12}| \). A biologist will use the writhe to describe the supercoiling of a closed circular DNA. One can get a visual idea of the writhe by looking, for example, at the image of a supercoiled plasmid obtained by means of a scanning force microscope (Figure 1). Here the writhe is the number of the crossings that the molecule axis forms with itself. For the plasmid shown in Figure 1, the writhe is equal to \(-4\). Unfortunately, the relation between this quantity and the Gauss integral [Eq. (1)] is not immediately obvious. We will consider this relation again in the next section.

In computer simulations of supercoiled DNA by Monte Carlo, Brownian dynamics, and molecular dynamics methods, it is often necessary to calculate the writhe for a given molecular conformation.\(^2\)\(^-\)\(^12\) In this case the numerical integration of Eq. (1) is very inefficient in terms of the computational time. It is possible, however, to take advantage of the fact that a DNA molecule is usually approximated by a chain consisting of straight segments. The problem of computation of the writhe for such a chain was first realized, and in principle, solved by Vologodskii et al.\(^2\) and le Bret.\(^13\) However, the elucidation of this problem in the literature is far from being satisfactory. Vologodskii et al.\(^2\) just mentioned the possibility to calculate the Gauss integral [Eq. (1)] analytically, without adding any further details. Le Bret,\(^13\) who used a more elaborate approach, gave a final expression but omitted some essential points in the derivation. Later le Bret reformulated his result in clearer terms,\(^3\) and this was the last publication concerning the subject.

Correspondence to: Jörg Langowski
© 2000 John Wiley & Sons, Inc.
During our experience with DNA modeling we realized that other fruitful approaches to the problem under consideration are possible. We present here a full description of four different methods allowing computation of the writhe for a circular DNA molecule modeled by a polymer chain consisting of straight segments. Two of them are new. Among the two others, one has been used by Vologodskii et al.\(^2\) (we found it appropriate to give it here, because the details are not available anywhere else in the literature) and the other is a simplified modification of the le Bret algorithm.\(^13\) The presented methods are compared with each other in terms of computational efficiency and the scope of their applicability is discussed.

**THEORETICAL BACKGROUND**

The strands of a closed circular DNA form two linked closed curves that cannot pass through each other by any possible deformations, as long as the covalent bonds are kept intact. Such a molecule can be characterized by a linking number. The linking number for two given closed curves \(C_1\) and \(C_2\) is defined in the following way.

First of all, we select an arbitrary direction of travel along the both curves. Then we choose an arbitrary spatial direction in which we will look at them. At a sufficiently large distance we can see the projection of the curves into a plane perpendicular to the view direction. In this two-dimensional image, the curves may intersect one another and each curve may intersect itself. For each crossing we can discern which curve (or which part of the curve) lies closer to the viewpoint and which one is more distant. The crossings can be divided in two classes: right-handed and left-handed ones. The right-handed crossing can be completed to a right-handed helix turn by connecting the crossing parts according to the selected direction of travel (Figure 2). Similarly, the left-handed crossing can be completed to a left-handed helix turn. We assign the weight \(w = +1\) to each right-handed crossing and the weight \(w = -1\) to each left-handed one. The total number of crossings between different curves is always even. Therefore the sum of their weights is even, too. The integer equal to half this sum is called the linking number: \(L_k = \frac{1}{2} \sum_i w_i\) (summation over all crossings between different curves). The linking number defined in such a way does not depend on the direction of the view. Indeed, by smooth variations of this direction, some new crossings may appear and some old ones may vanish, but the crossings always appear and vanish in pairs and the total number of weights for each pair is always equal to zero. For the same reason the linking number is invariant with respect to any deformations of the closed curves, provided that they do not pass through each other. The linking number changes its sign, however, if the direction of travel along one of the curves is altered.

When we look at a closed circular DNA in some direction we can distinguish two possible types of crossings between its strands, as illustrated in Figure 3. The *twist* crossings result from helical winding of

![FIGURE 1 Scanning force microscope image of a supercoiled DNA obtained as described in Ref. 1. (Protocol I, scanning in air). The DNA length is 3736 base pairs (bp). The scale is 100 nm.](image)

![FIGURE 2 Right-handed (a) and left-handed (b) crossings of oriented curves. The more distant curve is shown with a break. Following the directions of the curves, one can complete a right-handed crossing to a right-handed helix turn (a’) and a left-handed crossing to a left-handed helix turn (b’).](image)
one strand around the other. For right-handed helical forms of DNA their individual weight is equal to 1, provided that the directions of travel along the two strands are chosen to be parallel (that is also assumed in all our further considerations). The writhe crossings are caused by self-intersections of the DNA axis. The total number of crossings of each type is even.

(Strictly speaking, these two crossing types are distinguishable only in the projections in which the curvature radius of the image of the molecule axis exceeds the radius of DNA everywhere. If it is not the case, our considerations remain, however, valid: one only has to scale down the DNA radius to get a sufficiently “thin” image with the same axis as before. In the limit of infinitely thin molecule, when the two strands coincide, the twist crossings vanish and each pair of matching writhe crossings contracts to a single point.)

We define the directional twist \( T_w \) as the half total weight of the twist crossings and the directional writhe \( W_r \) as half the total weight of the writhe crossings. In other words, the directional twist is the total number of the visible turns of one of the strands about the DNA axis. The directional writhe is the total weight of the self-crossings of the DNA axis. The attribute directional is used here to emphasize that both quantities taken separately depend on the view direction, just in contrast to their sum

\[
Lk = T_w + W_r
\]  

(2)

Averaging Eq. (2) over all possible view directions, we get

\[
Lk = \langle T_w \rangle + \langle W_r \rangle
\]

(3)

where \( T_w = \langle T_w \rangle \) is the twist and \( W_r = \langle W_r \rangle \) is the writhe of a DNA molecule.

We can prove the equivalence of this definition of writhe to Eq. (1) in the following way. Let us consider two arbitrary points in the space, \( r_1 \) and \( r_2 \), as the origins of two infinitesimal vectors, \( dr_1 \) and \( dr_2 \), respectively. Let the solid angle \( d\Omega^*(r_1, r_2) = d\Omega^*_{12} \) be formed by all those view directions in which the vectors \( dr_1 \) and \( dr_2 \) apparently cross, with the vector \( dr_1 \) being the one closest to the viewpoint. Then (see Figure 5)

\[
d\Omega^*_1 = |dr_1^\perp| \cdot |dr_2^\perp| \cdot \sin \alpha / r_{12}^2
\]

(4)

where \( dr_1^\perp \) and \( dr_2^\perp \) are the components of \( dr_1 \) and \( dr_2 \) perpendicular to the vector \( r_{12} = r_2 - r_1, r_{12} = \|r_{12}\| \). Note that \( \Omega^*_1 = \Omega^*_{21} \). We can easily find the relationship between the absolute value of the integrated quantity from Eq. (1) and the solid angle \( \Omega^*_{12} \):

\[
\frac{|(dr_2 \times dr_1)_{r_{12}}|}{r_{12}^2} = \frac{|(dr_2^\perp \times dr_1^\perp)_{r_{12}}|}{r_{12}^2}
\]

\[
= \frac{|dr_1^\perp| \cdot |dr_2^\perp| \cdot \sin \alpha}{r_{12}^2} = d\Omega^*_{12}
\]

We define now the “signed solid angle” \( d\Omega(r_1, r_2) \) in such a way that \( d\Omega(r_1, r_2) = +d\Omega^*(r_1, r_2) \), if the apparent crossing of the vectors \( dr_1 \) and \( dr_2 \) is right-handed (the weight \( w_{12} \) is equal to +1), and

\[
\text{FIGURE 4} \quad \text{Two views of the same pair of closed curves:}
\]

(a) \( T_{w_D} = +3, W_{r_D} = 0 \); (b) \( T_{w_D} = 0, W_{r_D} = +3 \). The view directions differ by an angle of 30°.
FIGURE 5  The view on the two infinitesimal vectors $dr_1$ and $dr_2$ in the direction close to that of the vector $r_{12}$. The vectors $dr_1$ and $dr_2$ appear as their projections $dr_1^\perp$ and $dr_2^\perp$. By slight variations of the view direction, one can still see the crossing, if the relative position of the point $r_1$ with respect to the point $r_2$ lies within the parallelogram shown by the dashed line. The parallelogram area is $|dr_1^\perp| \cdot |dr_2^\perp| \cdot \sin \alpha$, where $\alpha$ is the angle between the vectors $dr_1^\perp$ and $dr_2^\perp$.

$$d\Omega(r_1, r_2) = -d\Omega^s(r_1, r_2), \text{ if the crossing is left-handed (} w_{12} = -1):$$

$$d\Omega(r_1, r_2) = w_{12}d\Omega^s(r_1, r_2) = \frac{(dr_2 \times dr_1)r_{12}}{r_{12}}$$

If we divide Eq. (5) by the value of the full solid angle, $4\pi$, then the obtained quantity, $d\Omega(r_1, r_2)/4\pi$, can be treated as half the total weight of the apparent crossings between the vectors $dr_1$ and $dr_2$ averaged over all possible directions. Hence, the linking number of two closed curves $C_1$ and $C_2$ is

$$Lk = \frac{1}{4\pi} \int_{C_1} \int_{C_2} d\Omega(r_1, r_2), \quad (6)$$

where $d\Omega(r_1, r_2)$ is defined by Eq. (5), $r_1$ and $r_2$ pass along the curves $C_1$ and $C_2$, respectively. The writhe of a curve $C$ is

$$Wr = \frac{1}{4\pi} \int_C \int_C d\Omega(r_1, r_2) \quad (7)$$

in full agreement with Eq. (1). Note that in Eq. (7) the factor before the integration sign is still $1/4\pi$, as in Eq. (6), although the writhe is defined as the average total weight of the self-crossings not divided by two.

The additional factor $1/2$ in Eq. (7) is due to the fact that in double integration over the same curve $C$ every pair of vectors $dr_1$ and $dr_2$ is taken into account twice.

For our further purposes, we also need an integral representation of the twist. We assume that the space curve of the DNA axis is parameterized by the distance $t$ traveled along its contour: $r = r(t)$, $0 \leq t \leq L$, $r(0) = r(L)$, $L$ being the total DNA length. Then one of the strands can be represented in the form $s = s(t) = r(t) + sa(t)$, where $a(t)$ is a unit vector perpendicular to the axis $r(t)$, $a(0) = a(L)$, and $s$ has the meaning of the DNA radius. As the twist does not depend on $s$, we can formally pass to the limit $s \to 0$, so that the curves $r(t)$ and $s(t)$ almost coincide. In the expression for the twist,

$$Tw = Lk - Wr = \frac{1}{4\pi} \int_{t=0}^{L} \int_{q=0}^{L} [d\Omega(s(t), r(q))$$

$$- d\Omega(r(t), r(q))] \quad (8)$$

the difference between the terms $d\Omega(s(t), r(q))$ and $d\Omega(r(t), r(q))$ is essential only when $q \approx t$. Hence, we can define a small interval $(t - \delta, t + \delta)$ in such a way that only the $q$ values belonging to this interval contribute to the integral in Eq. (8). Note that $\delta$ depends on $s$ and $\delta \gg s$. It is possible, however, to choose such a small value of $s$ that $\delta$ is also small enough, so that we can treat the axis at the distances comparable to $\delta$ as a straight line. The straight line does not form any self-crossings, therefore $d\Omega(r(t), r(q)) = 0$, if $t - \delta < q < p + \delta$, and

$$Tw = \int_{t=0}^{L} \frac{1}{4\pi} \int_{q=t-\delta}^{t+\delta} d\Omega(s(t), r(q)) \quad (9)$$

Here the quantity

$$dT_w = \frac{1}{4\pi} \int_{q=t-\delta}^{t+\delta} d\Omega(s(t), r(q)) \quad (10)$$

is practically independent of $\delta$, if $\delta \gg s$, and can be easily calculated analytically. It can also be interpreted as half the average total weight of apparent crossings between the vector $ds(t)$ and the tangent line to the axis at the point $r(t)$. As shown in Figure 6, $dT_w(t) = d\varphi(t)/2\pi$, where $d\varphi(t)$ is the angle of rotation of the vector $a(t)$ about the axis in the interval $(t, t + dt)$. As $a(t)$ is a unit vector, the absolute value
The twists of its constituent parts. i.e., the total twist of a molecule is equal to the sum of
The form of Eq. (12) implies that the twist is additive,

\[ d \varphi = \left( \frac{dr}{dt} \times a \right) da \]  

(11)

The angle \( d \varphi \) is positive, if the rotation of the vector \( a \) around the axis forms a right-handed helix, and negative otherwise. The resulting expression for the twist is

\[ T_w = \frac{1}{2\pi} \int_{t=0}^{t} \left( \frac{dr}{dt} \times a \right) \frac{da}{dt} \, dt \]  

(12)

The form of Eq. (12) implies that the twist is additive, i.e., the total twist of a molecule is equal to the sum of the twists of its constituent parts.

Note that in a closed chain the fractional part of \( T_w \) does not depend on \( a(t) \), as follows from Eq. (3).

The strict mathematical theory involving the linking number, the twist, and the writhe was developed by Călugăreanu\textsuperscript{15} and White.\textsuperscript{16} Fuller\textsuperscript{17} was the first to apply the theory to a supercoiled DNA. A simplified introduction to the theory adapted for physicists can be found in Ref. 14 and that adapted for biologists can be found in Ref. 18.

**METHODS OF COMPUTATION OF THE WRITHE**

In computer simulations a DNA molecule is usually modeled by a polymer chain consisting of straight segments. In this section, we present four methods allowing the calculation of the writhe for such chains.

**The Writhe as the Gauss Integral**

The Gauss double integral [Eq. (7)] along a polygon of \( N \) segments can be expressed as the double sum,

\[ W_r = 2 \sum_{i=2}^{N} \sum_{j<i} \frac{\Omega_{ij}}{4\pi} \]  

(13)

where \( \Omega_{ij}/4\pi \) is the Gauss integral along the segments \( i, j \) \( (\Omega_{ii} = \Omega_{jj}, \Omega_{i,i+1} = \Omega_{ii} = 0) \). Our goal now is to obtain an analytical expression for the quantity \( \Omega_{ij} \) for a pair of segments arbitrary oriented in the space.

**Method 1a.** One possibility is to apply a pure geometrical approach. Let the points 1 and 2 be the beginning and the end of the first segment, \( r_{12} \), and the points 3 and 4 be the beginning and the end of the second segment, \( r_{34} \) (Figure 7). In this case, the absolute value of the Gauss integral multiplied by \( 4\pi \) is the solid angle \( \Omega^* \) formed by all those view directions in which the vectors \( r_{12} \) and \( r_{34} \) apparently cross, with the vector \( r_{12} \) being the nearest to the viewpoint [see Eqs. (4), (5), and (7)]. As shown in Figure 7, the solid angle \( \Omega^* \) is bounded by the four planes: (134), (124), (234), and (123). In other words, the \( \Omega^* \) value is the area of the quadrangle on a unit sphere with the apexes formed by the intersections of the sphere with the rays originating at the center of the sphere and parallel to the lines (31), (41), (42), and (32). The area of a quadrangle on a unit sphere is

\[ \Omega^* = \alpha + \beta + \gamma + \delta - 2\pi \]  

(14)

where \( \alpha, \beta, \gamma, \) and \( \delta \) are the angles of the quadrangle, which can be found in the following way. Let

\[ n_1 = \frac{r_{13} \times r_{14}}{|r_{13} \times r_{14}|}, \quad n_2 = \frac{r_{14} \times r_{24}}{|r_{14} \times r_{24}|}, \quad n_3 = \frac{r_{24} \times r_{23}}{|r_{24} \times r_{23}|}, \quad n_4 = \frac{r_{23} \times r_{13}}{|r_{23} \times r_{13}|} \]  

(15)
be the unit vectors normal to the planes bounding the solid angle $\Omega^*$. Then, taking into account their mutual orientation, $\alpha = \arccos(-\mathbf{n}_1 \cdot \mathbf{n}_2) = \pi/2 + \arcsin(\mathbf{n}_1 \cdot \mathbf{n}_3)$, $\beta = \pi/2 + \arcsin(\mathbf{n}_2 \cdot \mathbf{n}_3)$ and so on. Finally, we have

$$
\Omega^* = \arcsin(\mathbf{n}_1 \cdot \mathbf{n}_2) + \arcsin(\mathbf{n}_1 \cdot \mathbf{n}_3) + \arcsin(\mathbf{n}_2 \cdot \mathbf{n}_3) + \arcsin(\mathbf{n}_3 \cdot \mathbf{n}_1) \quad (16a)
$$

The Gauss integral $\Omega/4\pi$ is equal to $+\Omega^*/4\pi$, if the two segments form a right-handed crossing, i.e., $(\mathbf{r}_3 \times \mathbf{r}_4) \cdot \mathbf{r}_{13} > 0$, and $-\Omega^*/4\pi$ otherwise:

$$
\frac{\Omega}{4\pi} = \frac{\Omega^*}{4\pi} \text{sign}((\mathbf{r}_3 \times \mathbf{r}_4) \cdot \mathbf{r}_{13}) \quad (16b)
$$

**Method 1b.** Another possibility to find the $\Omega/4\pi$ value is to calculate the integral from Eq. (1) analytically. This approach was first used by Vologodskii et al., although the details were not published. Here it is more convenient to denote the segments as $s_1$ and $s_2$ and their starting points as $r_1$ and $r_2$, respectively. The segment lengths are $s_1 = |\mathbf{s}_1|$ and $s_2 = |\mathbf{s}_2|$. The correspondent unit vectors are $\mathbf{e}_1 = \mathbf{s}_1/|\mathbf{s}_1|$ and $\mathbf{e}_2 = \mathbf{s}_2/|\mathbf{s}_2|$. Let us resolve the vector $r_{12} = r_2 - r_1$ in the (not orthonormal) coordinate frame $(-\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 \times \mathbf{e}_2)$:

$$
\mathbf{r}_{12} = -a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_0(\mathbf{e}_1 \times \mathbf{e}_2) \quad (17)
$$

Multiplying Eq. (17) separately by $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_1 \times \mathbf{e}_2$, we get a system of three independent equations, from which the coefficients $a_1$, $a_2$, and $a_0$ can easily be found:

$$
\begin{align*}
   a_1 &= r_{12}(\mathbf{e}_2 \cos \beta - \mathbf{e}_1)/\sin^2 \beta \quad (18a) \\
   a_2 &= r_{12}(\mathbf{e}_1 - \mathbf{e}_2 \cos \beta)/\sin^2 \beta \quad (18b) \\
   a_0 &= r_{12}(\mathbf{e}_1 \times \mathbf{e}_2)/\sin^2 \beta \quad (18c)
\end{align*}
$$

where $\beta$ is the angle between $\mathbf{e}_1$ and $\mathbf{e}_2$:

$$
\cos \beta = \mathbf{e}_1 \cdot \mathbf{e}_2, \quad \sin^2 \beta = 1 - (\mathbf{e}_1 \cdot \mathbf{e}_2)^2 \quad (19)
$$

The segments can be represented parametrically as

$$
\begin{align*}
   \mathbf{R}_1 &= \mathbf{r}_1 + x_1\mathbf{e}_1, \quad 0 \leq x_1 \leq s_1 \quad (20a) \\
   \mathbf{R}_2 &= \mathbf{r}_2 + x_2\mathbf{e}_2, \quad 0 \leq x_2 \leq s_2 \quad (20b)
\end{align*}
$$

Then

$$
\mathbf{R}_{12} = \mathbf{R}_2 - \mathbf{R}_1 = -((a_1 + x_1)\mathbf{e}_1 + (a_2 + x_2)\mathbf{e}_2) + a_0(\mathbf{e}_1 \times \mathbf{e}_2) \quad (21)
$$

and

$$
R_{12}^2 = |\mathbf{R}_{12}|^2 = (a_1 + x_1)^2 + (a_2 + x_2)^2 - 2(a_1 + x_1)(a_2 + x_2)\cos \beta + a_0^2\sin^2 \beta \quad (22)
$$

In these notations the Gauss integral over the two segments is

$$
\frac{\Omega}{4\pi} = -\frac{1}{4\pi} \int_{a_1}^{a_2} dt_1 \int_{a_2}^{a_1} dt_2 \frac{a_0\sin^2 \beta}{(t_1^2 + t_2^2 - 2t_1t_2\cos \beta + a_0^2\sin^2 \beta)^{3/2}} \quad (23)
$$

Here the integration parameters are $t_1 = a_1 + x_1$ and $t_2 = a_2 + x_2$. This integral can be calculated analytically by the standard methods and the result is

$$
\frac{\Omega}{4\pi} = F(a_1 + s_1, a_2 + s_2) - F(a_1 + s_1, a_2) - F(a_1, a_2 + s_2) + F(a_1, a_2) \quad (24)
$$
The idea is to attach to the curve \( C \) parameter \( s \) of the curves \( C \). The closed curve \( C \) can be chosen arbitrary, provided that it runs close enough to the curve \( C \) along its entire length.

Practically, by modeling the DNA with the molecule axis approximated by a chain of \( N \) straight segments, it is convenient to represent the auxiliary curve in the following way. Let us attach to each segment \( s_i \) a unit vector \( a_i \) perpendicular to it. Then we can imagine that the auxiliary curve passes through the nodes \( r_i + ks_i + \epsilon a_i \), where \( r_i \) is the initial point of the \( i \)th segment, \( k \) belongs to the interval (0, 1), and \( \epsilon \) is an arbitrary small parameter (Figure 8). The twist of such a double chain depends neither on the exact values of the parameters \( k \) and \( \epsilon \) nor on the particular shape of the auxiliary curve between the nodes. It is completely defined by the vectors \( s_i \) and \( a_i \), \( i = 1, \ldots, N \), provided that the twist between successive segments (successive nodes) is assumed to lie in the interval \((-1/2, 1/2)\). The fractional part of the total twist is independent of \( a_i \).

The twist \( T_w \) between the segments \( i - 1 \) and \( i \) can be easily found. If

\[
p_i = \frac{s_{i-1} \times s_i}{|s_{i-1} \times s_i|} \tag{27}
\]

is the unit vector perpendicular to both segments, then, as shown in Figure 9,

\[
T_w = \frac{1}{2\pi} \left[ \alpha_i + \gamma_i \right]_{2\pi} \tag{28}
\]

where \( \alpha_i \) is the angle of rotation between the vectors \( a_{i-1} \) and \( p_i \) about the segment \( s_{i-1} \), \( -\pi < \alpha_i \leq \pi \), and \( \gamma_i \) is the angle of rotation between the vectors \( p_i \) and \( a_i \) about the segment \( s_i \), \( -\pi < \gamma_i \leq \pi \). The positive direction of rotation is defined by the right-hand rule. The notation \([\cdot \cdot \cdot]_{2\pi}\) means that the value of the angle in the square brackets should be taken from the interval \((-\pi, \pi)\). Note that \( \alpha_i \) and \( \gamma_i \) are two of the three Eulerian angles for the successive local frames \((a_{i-1}, a_{i-1} \times e_{i-1}, e_{i-1})\) and \((a_i, a_i \times e_i, e_i)\);

\[
e_i = s_i/|s_i| \tag{27}
\]

**Method 2a.** Let us choose the \( a_i \) vectors in such a way that \( a_i = e_i \times p_{i+1} \), where \( e_i = s_i/|s_i| \). The vector \( a_i \) can be obtained from the vector \( p_{i+1} \) by rotation about the segment \( s_i \) by the angle \( +\pi/2 \). The twist angle between the vectors \( a_{i-1} \) and \( a_i \) is the same as that between the vectors \( p_i \) and \( p_{i+1} \):

\[
W_r = Lk - T_w \tag{26}
\]
\[ \alpha_i + \gamma_i = \arccos(p_i \cdot p_{i+1}) \text{sign}((p_i \times p_{i+1}) \cdot s_i) \]

\[ = \arccos(p_i \cdot p_{i+1}) \text{sign}(p_i \cdot s_{i+1}). \quad (29) \]

Here, to simplify the sign factor, we use the fact that \( p_{i+1} \) is parallel to \( s_i \times s_{i+1} \) and apply the formula for a double vector product.

The total twist is

\[ Tw = \frac{1}{2\pi} \sum_{i=1}^{N} \arccos(p_i \cdot p_{i+1}) \text{sign}(p_i \cdot s_{i+1}) \quad (30) \]

(Here and further we imply the cyclic substitution for the indexes exceeding the range 1, \ldots, \( N \).)

The linking number can be found as the sum of the directional twist and the directional writhe for the projection on the \( xy \) plane.

To find the directional twist \( Tw_i \), we have to count the crossings between segments and the adjoining parts of the auxiliary curve. Let the auxiliary curve pass through the nodes \( r_i + e a_{i-1} \) and \( r_i + e b_i \), where \( b_i = e_i \times p_i \) (Figure 10). This pair of nodes lie in the same plane as the segments \( s_{i-1} \) and \( s_i \). If, in the chosen projection, the \( b_i \) vector is pointing to one side of the segment \( s_i \) and the \( b_{i+1} \) vector is pointing to the other side of the segment \( s_{i+1} \), then a crossing takes place somewhere within the segment \( s_i \) (Figure 10). Note that if \( b_i \) is pointing to the right of the segment \( s_i \), then the z-component of their vector product is positive: \( \text{sign}(b_i \times s_i)_z = \text{sign}(b_i \times e_i)_z = \text{sign}(p_i)_z = 1 \). If \( b_i \) is pointing to the left then \( (p_i)_z < 0 \). The crossing weight is \( w_i = \text{sign}(p_i \times p_{i+1} \cdot s_i) = \text{sign}(p_i \times p_{i+1} \cdot s_i) = \text{sign}(p_i \cdot s_{i+1}) \). Hence, the directional twist is

\[ Tw_i = \frac{1}{2} \sum_{i=1}^{N} \tau_i \quad (31a) \]

\[ \tau_i = \begin{cases} 
\text{sign}(p_i \cdot s_{i+1}), & \text{if } (p_i)_z \cdot (p_{i+1})_z < 0 \\
0, & \text{else (also for adjacent segments)}
\end{cases} \quad (31b) \]

The directional writhe is

\[ Wr_z = \sum_{i=2}^{N} \sum_{j<i} w_{ij} \quad (32a) \]

\[ w_{ij} = \begin{cases} 
\text{sign}((s_i \times s_j) \cdot (r_j - r_i)), & \text{if the segments } i, j \text{ apparently cross} \\
0, & \text{else (also for adjacent segments)}
\end{cases} \quad (32b) \]

The criterion for apparent crossing of the segment is following. Let \( r_i + k s_j, -\infty < k < \infty \), be the straight line containing the \( i \)th segment. Its projection
on the $xy$ plane is $r_i^{(p)} + k_i s_i^{(p)}$, where $r_i^{(p)}$ and $s_i^{(p)}$ are the two-dimensional projections of the vectors $r_i$ and $s_i$, respectively. At the apparent crossing point of the lines $i$ and $j$, the parameters $k_i$ and $k_j$ satisfy the condition

$$r_j^{(p)} + k_j s_j^{(p)} = r_i^{(p)} + k_i s_i^{(p)}$$  \hspace{1cm} (33)

The segments $i$ and $j$ apparently cross only in the case when $0 < k_i \leq 1$ and $0 < k_j \leq 1$, where $k_i$ and $k_j$ are the solution of Eq. (33).

The writhe of the whole chain is

$$Wr = Tw_z + Wr_z - Tw$$  \hspace{1cm} (34)

where $Tw$, $Tw_z$, and $Wr_z$ are defined by Eqs. (30), (31), and (32).

**Method 2b.** Le Bret\textsuperscript{13} proposed to take the $a_i$ vectors in the form

$$a_i = \frac{k \times s_i}{|k \times s_i|}$$  \hspace{1cm} (35)

where $k$ is the unit vector parallel to the $z$ axis. The advantage of this choice is that, in the projection on the $xy$ plane, the $a_i$ vector is always pointing to the left of the segment $s_i$: $(a_i \times s_i)_z < 0$. Consequently, if one moves down the main chain axis, the auxiliary curve runs always to the left and never forms crossings of the twist type. Hence, $Tw_z = 0$ and $Lk = Wr_z$. The twist $Tw_i$ between the segments $i - 1$ and $i$, according to Eq. (27), is

$$Tw_i = \frac{1}{2\pi} \left[ \arccos(a_{i-1,p}) \text{sign}((a_{i-1,p} \times p_i)s_{i-1}) ight.$$  
$$+ \arccos(p_i,a_i) \text{sign}(p_i \times a_i)s_i \right]_{2\pi}$$  \hspace{1cm} (36)

Here the sign factors can be simplified: $\text{sign}((a_{i-1} \times p_is_{i-1}) = \text{sign}(p_i)$, and $\text{sign}(p_i \times a_i)s_i = -\text{sign} (p_i)_z$, where $(p_i)_z$ is the $z$ component of the $p_i$ vector.

The total twist is, therefore,

$$Tw = \frac{1}{2\pi} \sum_{i=1}^{N} \left[ \arccos(a_{i-1,p}) ight.$$  
$$- \arccos(p_i,a_i) \text{sign}(p_i)_z \right]$$  \hspace{1cm} (37)

The writhe of the chain is

$$Wr = Wr_z - Tw$$  \hspace{1cm} (38)

where the directional writhe $Wr_z$ is defined, as before, by Eq. (32).

To obtain an expression for the twist, le Bret\textsuperscript{13} used another approach based on the analytical evaluation of the integral [Eq. (12)] for smoothed joints between segments. His result is equivalent to Eq. (37) but somewhat more complicated.

**COMPARISON OF THE METHODS**

We have considered four methods of calculation of the writhe for a closed chain consisting of straight segments. Which of them is more preferable for use in computer modeling of supercoiled DNA?

In terms of CPU time, methods 2a and 2b are more efficient. The number of operations required for the computation of $Tw$ and $Tw_z$ depends linearly on the number of segments $N$. In the calculation of $Wr_z$, for which the computational time is proportional to $N^2$, only the four basic arithmetic operations are used. In methods 1a and 1b the number of terms to be evaluated is also proportional to $N^2$, but every term contains square roots and inverse trigonometric functions.

The essential disadvantage of methods 2a and 2b is, however, the fact that they are not applicable to every possible chain orientation with respect to the global coordinate system. The expressions in Eqs. (31b) and (37) are not defined if one of the $p_i$ vectors is strictly perpendicular to the $z$ axis $[(p_i)_z = 0]$. Because of computational errors, the criterion for apparent crossing of the segments $i, j$ is not reliable, if the value of $k_i$ or $k_j$ from Eq. (33) is equal to 0 or 1. Fortunately, in a random chain, these unpleasant situations practically never occur, provided that double precision is used.

Methods 1a and 1b are always reliable. Singularities in the expressions for the Gauss integral $\Omega/4\pi$ over a pair of segments take place only if the both segments lie in the same plane. But in this case $\Omega = 0$. In particular, $\Omega = 0$ for any pair of adjacent segments.

As an illustrative example for quantitative comparison of the methods, we have chosen the problem of calculation of the entropy of branch formation in a supercoiled DNA. In solving this problem we used Monte Carlo computer simulations following Vologodskii et al.\textsuperscript{8} The DNA molecule was modeled by a chain consisting of $N$ straight segments with a fixed length $l_0$. Harmonic potentials with respect to the angles between adjacent segments and to the twist deficiency $\Delta Tw$ were used. The energy of a knotted conformation was assumed to be infinity. The volume interaction between segments were given by the Debye–Hückel approximation, as described in Ref. 12.
was plasmid (2687 bp). The linking number deficiency and tures, respectively, in the total set of conformations, the parts of unbranched and singly branched struc-

\textbf{Related Quantities}a

Methods for Calculation of the Writhe or the Related Quantities

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Method(s)</th>
<th>Equation(s)</th>
<th>Dependence on N</th>
<th>t, ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>(W_r)</td>
<td>1a</td>
<td>13, 16</td>
<td>(t \propto N^2)</td>
<td>53.3</td>
</tr>
<tr>
<td>(W_r)</td>
<td>1b</td>
<td>13, 24, 25</td>
<td>(t \propto N^2)</td>
<td>52.5</td>
</tr>
<tr>
<td>(W_{r_2})</td>
<td>2a, 2b</td>
<td>32</td>
<td>(t \propto N^2)</td>
<td>8.49</td>
</tr>
<tr>
<td>(-T_w)</td>
<td>2a</td>
<td>30, 31</td>
<td>(t \propto N)</td>
<td>0.24</td>
</tr>
<tr>
<td>(-T_w)</td>
<td>2b</td>
<td>37</td>
<td>(t \propto N)</td>
<td>0.59</td>
</tr>
</tbody>
</table>

\footnotesize a The calculations were performed on a PC under Linux with Pentium Pro processor (200 MHz) for a chain of \(N = 91\) segments.

The evolution of the chain was performed by two types of trial steps: (1) pivoting of a randomly chosen subchain about its end-to-end vector by a random angle; (2) exchange of two randomly chosen subchains, of 4 and 5 segments, with proper readjustment of the end-to-end distances.\(^8\) The trial step was accepted or rejected, depending on the change in the energy according to the Metropolis rule.

We used the following parameters: The molecule length was \(L = 910\) nm, corresponding to the pUC18 plasmid (2687 bp). The linking number deficiency was \(\Delta L k = -10\) (superhelical density \(\sigma = -0.04\)), the persistence length \(L_p = 50\) nm, the torsional rigidity \(C = 2.0 \times 10^{-19}\) erg \cdot cm, the NaCl concentration \(I = 0.1 M\), the temperature \(T = 293\) K. The segment length was \(l_0 = 10\) nm, so that the number of segments \(N = 91\). The total number of the Monte Carlo steps was \(2 \cdot 10^8\). This corresponds approximately to \(10^3\) independent conformations, judging from the autocorrelation function of number of branches.\(^9\) The program was written in C++ and executed on a PC under Linux with a Pentium Pro processor (200 MHz).

The free energy of branch formation was calculated as \(\Delta F = -k_B T \ln(p_1/p_0)\), where \(p_0\) and \(p_1\) are the parts of unbranched and singly branched structures, respectively, in the total set of conformations, and \(k_B\) is the Boltzmann constant. The entropy of branch formation is then \(\Delta S = (\Delta E - \Delta F)/T\), where \(\Delta E = E_1 - E_0\) is the difference of the mean energies of chains with one branch and without branches. The results are \(\Delta F/k_B T = 0.9\); \(\Delta E/k_B T = 2.4\); \(\Delta S/k_B = 1.5\).

In the present context, two points are of particular interest. First, the chain energy contains a term proportional to \((\Delta T w)^2\), the value of \(\Delta T w\) being obtained as \(\Delta T w = \Delta L k - W_r\). Thus, the writhe \(W_r\) should be calculated at each Monte Carlo step and the corre-

\textbf{FIGURE 11} An example of the \(\omega\) function for a typical chain conformation. In the image of the chain the angles between segments are smoothed. The chain thickness corresponds to the effective diameter of DNA, which for the given NaCl concentration (0.1M) is equal to 5.6 nm.\(^{21}\)
that form superhelix end loops have relatively higher values of $v$. We assume that the $j$th segment belongs to an end loop if $v(j) \geq v_1$, where the “threshold” $v_1$ is equal to 0.5. Further, the $j$th segment does not belong to an end loop if $v(j) \leq v_2$, where the second threshold $v_2$ is equal to 0.3. A continuous sequence of segments for which $v_2 < v < v_1$ belongs to an end loop if and only if it is bounded on the both sides by segments with $v > v_1$. The number of end loops is the number of continuous sequences of segments satisfying this criterion. The number of branches is less by 2. According to the assumed value of $v_1$, a branch contains, on the average, at least one apparent crossing of the molecule axis. An example of the $v$-function for a typical chain conformation is presented in Figure 11.

For realizing this algorithm, only methods 1a and 1b are applicable, because they allow calculation of each separate term in Eq. (13). Thus, although methods 2a and 2b are more efficient, methods 1a and 1b are more universal. We do not consider here distinctions between methods 1a and 1b or between methods 2a and 2b, since these distinctions are not principal.

We thank N. Mücke for placing Figure 1 at our disposal and A. Vologodskii for valuable remarks.

REFERENCES