A Brownian Dynamics Program for the Simulation of Linear and Circular DNA and Other Wormlike Chain Polyelectrolytes

Konstantin Klenin, Holger Merlitz, and Jörg Langowski
Division Biophysics of Macromolecules, German Cancer Research Center, D-69120 Heidelberg, Germany

ABSTRACT For the interpretation of solution structural and dynamic data of linear and circular DNA molecules in the kb range, and for the prediction of the effect of local structural changes on the global conformation of such DNAs, we have developed an efficient and easy way to set up a program based on a second-order explicit Brownian dynamics algorithm. The DNA is modeled by a chain of rigid segments interacting through harmonic spring potentials for bending, torsion, and stretching. The electrostatics are handled using precalculated energy tables for the interactions between DNA segments as a function of relative orientation and distance. Hydrodynamic interactions are treated using the Rotne-Prager tensor. While maintaining acceptable precision, the simulation can be accelerated by recalculating this tensor only once in a certain number of steps.

INTRODUCTION

Understanding the role of DNA three-dimensional (3D) structure in its biological function requires a quantitative description of the structure and dynamics of long DNA chains in their various states. Double-helical DNA is a wormlike polyelectrolyte chain, whose conformation—neglecting atomic detail—can be described by a simple set of parameters: bending and twisting elasticity, hydrodynamic diameter, and electrostatic interactions between different parts of the same molecule.

The bending elasticity of B-DNA is given by its persistence length of 50 nm and does not strongly depend on monovalent ion concentration between 0.01 and 1 M. In the computations referred to in this paper, we work with a constant persistence length of 50 nm (Hagerman, 1988). For the torsional elasticity we take $C = 2.0 \cdot 10^{-19}$ erg cm (Schurr et al., 1992). The hydrodynamic properties of DNA can be well described by a hydrodynamic diameter of 2.4 nm (Hagerman and Zimm, 1981). Long-range intramolecular interactions are otherwise due to electrostatic repulsion between the negatively charged phosphate groups on the helix backbone. One of the main points of this paper is to develop an efficient procedure to compute electrostatic interactions in DNA.

A numerical simulation of the time-dependent structure of a long chain molecule like DNA in solution is equivalent to solving the equations of motion for this molecule plus the surrounding solvent. Given the size of the molecule—some 1000 basepairs of DNA, corresponding to $\sim 10^5$ atoms—a numerical description at the atomic level on present-day machines is unthinkable over any reasonable time scale. But since one is mainly interested in global features of the structure of the DNA molecule, such as cyclization probabilities, as computed in the accompanying paper, it is possible to view the DNA as a linear chain of segments, each including several tens of basepairs, and to model the solvent as a continuous viscous fluid. The thermal motion of the solvent molecules is included in the equations of motion through a random force.

This approach of describing the dynamics of a system is called *Brownian dynamics* (BD), a method that has been applied in several cases to modeling DNA dynamics (Allison, 1986; Allison et al., 1989, 1990; Chirico and Langowski, 1994). Up to now, in most cases new BD models were developed for each particular case. However, the power of the BD method and the range of problems in the 3D structure of DNA and other wormlike polymers where it can be successfully applied justifies the implementation of a general-purpose BD package for computing polymer dynamics, with special regard for DNA. In this paper we describe the implementation of such a program. An accompanying paper (Merlitz et al., 1998) shows its application to the problem of DNA loop closure probabilities in the presence of bound proteins and/or bends.

METHODS

The DNA molecule is modeled as an elastic chain with electrostatic interactions. The principal approach taken here is similar for linear and circular DNA; specific differences between the two cases will be mentioned where they apply. The conformation of a chain of N straight segments is specified by the space positions of its vertices, $\boldsymbol{r}_i$, $i = 0, \ldots, N$. The segments are represented by the vectors $\boldsymbol{s}_i = \boldsymbol{r}_{i+1} - \boldsymbol{r}_i$, $i = 0, \ldots, N-1$. To each segment a system of three orthogonal vectors of unit length is attached ($\boldsymbol{f}_i$, $\boldsymbol{g}_i$, $\boldsymbol{e}_i$), so that the $\boldsymbol{e}_i$ direction coincides with the direction of the ith segment: $\mathbf{e}_i = \mathbf{s}_i/s_i$, $s_i = |\mathbf{s}_i|$ (Fig. 1). For a circular chain we have the additional restriction that $\boldsymbol{r}_N = \boldsymbol{r}_0$.

Received for publication 19 June 1997 and in final form 3 November 1997.
Address reprint requests to Dr. Jörg Langowski, Deutsches Krebsforschungszentrum, Abt. Biophysik der Makromolekule (0830), Postfach 101949, D-69009 Heidelberg, Germany. Tel.: 49-6221-423390; Fax: 49-6221-423391; E-mail: joerg.langowski@dkfz-heidelberg.de.

0006-3495/98/02/780/09 $2.00

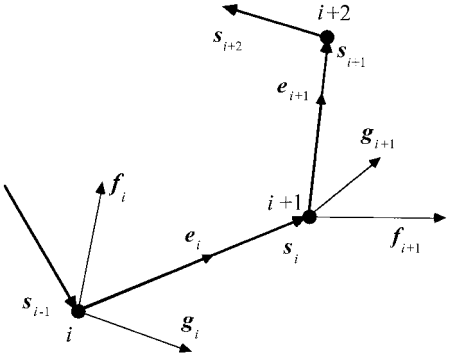

FIGURE 1 Chain geometry with the segment vectors $\mathbf{s}_i$ and the segment coordinate systems ($\mathbf{f}_i$, $\mathbf{g}_i$, $\mathbf{e}_i$) that define the relative orientation between segments.

The total energy of a given chain conformation is given as the sum of the stretching, bending, twist, and electrostatic energies.

The stretching energy is defined for each segment i:

$$\frac{E_i^{(s)}}{k_B T} = \frac{1}{2(l_0\delta)^2}(l_0 - s_i)^2 \tag{1}$$

Here k_B is the Boltzmann constant, T the temperature, l_0 the segment equilibrium length, and δ the stiffness parameter, so that $(l_0\delta)^2$ is approximately equal to the variance of the segment length distribution.

The bending energy is defined for each chain joint. We call a joint *unbent* if it connects segments that form a straight line at equilibrium, and *bent* if the angle θ_i^* between the segments at equilibrium is nonzero. To each bent joint i we attach an auxiliary unit vector $\boldsymbol{b}_i$ that is fixed in the coordinate system ($\boldsymbol{f}_i$, $\boldsymbol{g}_i$, $\boldsymbol{e}_i$), its polar coordinates being (θ_i^*, φ_i^*). Note that this formalism is different from our first implementation (Chirico and Langowski, 1996) where a heuristic "kink potential" was used that was given in terms of the Euler angles for rotating one segment into the next. Here the bending energy of the ith joint is:

$$\frac{E_i^{(b)}}{k_B T} = \alpha_b \beta_i^2 \tag{2}$$

where β_i is the angle between $\boldsymbol{e}_{i-1}$ and $\boldsymbol{e}_i$ for an unbent joint, or between $\boldsymbol{e}_{i-1}$ and $\boldsymbol{b}_i$ for a bent joint; α_b is the bending rigidity parameter chosen in such a way that the Kuhn length is equal to:

$$B = l_0 \frac{1 + \langle \cos\beta \rangle}{1 - \langle \cos\beta \rangle}, \tag{3}$$

where

$$\langle \cos\beta \rangle = \frac{\int_0^\pi \cos\beta \sin\beta \exp(-\alpha_b\beta^2) d\beta}{\int_0^\pi \sin\beta \exp(-\alpha_b\beta^2) d\beta}.$$

The twist energy is defined for each adjacent segment pair:

$$\frac{E_i^{(t)}}{k_B T} = \frac{1}{2k_B T}\frac{C}{l_0}\tau_i^2, \tag{4}$$

where C is the torsional rigidity constant and τ_i is the twist angle between the $(i-1)$th and ith segments.

The twist angle τ_i is calculated by defining a vector $\boldsymbol{p}_i = \mathbf{s}_{i-1} \times \mathbf{s}_i$, which is normal both to $\boldsymbol{s}_{i-1}$ and $\boldsymbol{s}_i$. Now, we can easily calculate the angles α_i between $\boldsymbol{f}_{i-1}$ and $\boldsymbol{p}_i$, and γ_i between $\boldsymbol{p}_i$ and $\boldsymbol{f}_i$. Then, $\tau_i = \alpha_i + \gamma_i$ (Fig. 2). During the BD simulation we assume that at time t, $\tau_i(t - \Delta t) - \pi \le \tau_i(t) \le \tau_i(t - \Delta t) + \pi$, where $\tau_i(t - \Delta t)$ is the twist angle one simulation step ago (Δt is chosen such that the probability that the twist angle changes by more than $\pm\pi$ becomes negligible).

The starting point for the electrostatic energy is the expression for the energy of interaction between two uniformly charged nonadjacent segments (i, j) in a 1:1 salt solution in the Debye-Hückel approximation:

$$\frac{E_{ij}^{(e)}}{k_B T} = \frac{\nu^2}{k_B T D}\int d\lambda_i \int d\lambda_j \frac{\exp(-\kappa r_{ij})}{r_{ij}}. \tag{5}$$

The integration is done along the two segments; λ_i and λ_j are the distances from the segment beginnings, r_{ij} is the distance between the current positions at the segments to which the integration parameters λ_i and λ_j correspond; κ is the inverse of the Debye length, so that $\kappa^2 = 8\pi e^2 I/k_B T D$, I is the ionic strength, e the proton charge, D the dielectric constant of water, ν the linear charge density which for DNA is equal to $\nu_{DNA} = -2e/\Delta$, where $\Delta = 0.34$ nm is the distance between basepairs.

There are two problems to be solved for the Eq. 5: 1) the linear density ν should be renormalized from that of DNA to a smaller value in order to ensure the correct excluded volume effects; 2) the integration should be approximated by a more simple procedure to save computation time.

The renormalization of the linear density was done as in Stigter (1977). As pointed out in Schellman and Stigter (1977), the Gouy layer of immobile counterions reduces the effective charge density by a factor of $q = 0.73$ for NaCl concentrations between 1 and 500 mM. Next, the Debye-Hückel approximation is a linearization of the Poisson-Boltzmann equation and valid only for a very small electric potential: $\varphi \ll k_B T/e$. We choose the renormalized charge

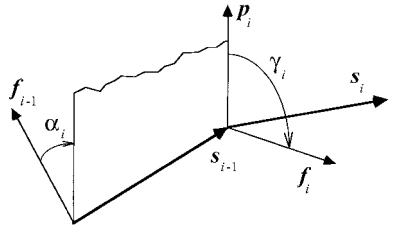

FIGURE 2 Definition of the twist angle $\tau_i = \alpha_i + \gamma_i$. $\mathbf{p}_i$ is perpendicular to $\mathbf{s}_{i-1}$ and $\mathbf{s}_i$; α_i is measured between $\mathbf{f}_{i-1}$ and $\mathbf{p}_i$, γ_i between $\mathbf{p}_i$ and $\mathbf{f}_i$.

density v^* in such a way that the known solution of the Debye-Hückel equation for a straight thin line with charge density v^* coincides with the solution of the Poisson-Boltzmann equation for a cylinder of the DNA radius r_{ES} and the charge density qv_{DNA} in the regions where $\varphi \ll k_BT/e$ (Fig. 3).

In order to save computation time, a tabulation of the double integral (Eq. 5) was used. For simplicity we assume that each segment has the same length l_0. For each segment pair (i, j) we can define a vector $l_0\boldsymbol{R}_{ij}$ connecting the middle points of the two segments. Then the mutual position of the segments can be defined by the following four dimensionless parameters:

$$\rho_{ij} = |\boldsymbol{R}_{ij}|,$$

center-to-center distance in L_0 units, (6a)

$$\rho_{ij} \geq 0$$

$$\gamma_{ij} = (1/\rho_{ij})\boldsymbol{e}_i \cdot \boldsymbol{R}_{ij},$$

tilt angle cosine for the ith segment, (6b)

$$-1 \leq \gamma_{ij} \leq 1$$

$$\gamma_{ji} = -(1/\rho_{ij})\boldsymbol{e}_j \cdot \boldsymbol{R}_{ij},$$

tilt angle cosine for the jth segment, (6c)

$$-1 \leq \gamma_{ji} \leq 1$$

$$\sigma_{ij} = \frac{(\boldsymbol{e}_i \times \boldsymbol{R}_{ij}) \cdot (\boldsymbol{e}_j \times \boldsymbol{R}_{ij})}{|\boldsymbol{e}_i \times \boldsymbol{R}_{ij}||\boldsymbol{e}_j \times \boldsymbol{R}_{ij}|},$$

twist angle cosine, $-1 \leq \sigma_{ij} \leq 1$ (6d)

Equation 5 can be rewritten in the form:

$$\frac{E^{(e)}_{ij}}{k_BT} = \alpha_e f(\rho_{ij}, \gamma_{ij}, \gamma_{ji}, \sigma_{ij}) \tag{7}$$

where $\alpha_e = v^{*2} l_0/k_BTD$ and

$$f(\rho, \gamma_1, \gamma_2, \sigma) = \int_{-1/2}^{1/2} dx_1 \int_{-1/2}^{1/2} dx_2 \frac{\exp(-\kappa|\boldsymbol{R}|)}{|\boldsymbol{R}|}, \tag{8}$$

$$\mathbf{R} = \rho\mathbf{v}_0 - x_1\mathbf{v}_1 + x_2\mathbf{v}_2 \tag{9}$$

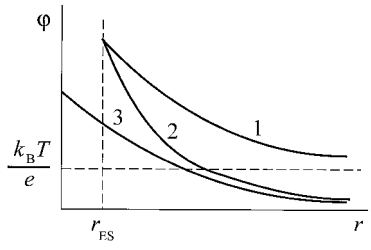

FIGURE 3 Electrostatic potential for the interaction between DNA segments. A Debye-Hückel potential (1) was renormalized (3) such as to coincide at large distances with the nonlinear Poisson-Boltzmann equation (2).

where $\mathbf{v}_0$, $\mathbf{v}_1$, and $\mathbf{v}_2$ are unit vectors oriented in such a way that $\mathbf{v}_0\mathbf{v}_1 = \gamma_1$, $-\mathbf{v}_0\mathbf{v}_2 = \gamma_2$, and $(\mathbf{v}_1 \times \mathbf{v}_0) \cdot (\mathbf{v}_2 \times \mathbf{v}_0)/|\mathbf{v}_1 \times \mathbf{v}_0||\mathbf{v}_2 \times \mathbf{v}_0| = \sigma$.

In the following we need the partial derivatives of Eq. 8. At first, a four-dimensional table for the values of the integral (Eq. 8) and each of its partial derivatives was constructed *numerically*. Then, during the simulation, a linear interpolation was used to obtain the values of $(\partial f/\partial\rho)$, $(\partial f/\partial\gamma_1)$, $(\partial f/\partial\gamma_2)$, $(\partial f/\partial s)$ at particular points of the $(\rho, \gamma_1, \gamma_2, \sigma)$ space. The table steps were chosen in such a way that the values of ρ, arccos γ_1, arccos γ_2, and arccos σ had constant increments. The tabulation range for γ_1, γ_2, and σ is $[-1,1]$. The range of ρ, $[\rho_{min}, \rho_{max}]$, and the table size for each argument are parameters of the approximation, which we chose by the following criteria. For the minimum distance, ρ_{min}, all possible values of the electrostatic energy should be large enough (e.g., $>10\ k_BT$), so that this distance is practically unreachable during the simulations. For distances $>\rho_{max}$ all possible energy values should be negligible (say, $<0.01\ k_BT$). The mutual displacement of segments corresponding to one ρ-step should not exceed the Debye length, and the same restriction is applied to the displacement of segment ends corresponding to one step in the γ_1, γ_2, and σ dimensions. These criteria are rather "soft," in order to keep the total table size within reasonable limits (the minimum table size for $l_0 = 10$ nm and $I = 1$ M is 14 MB of memory, including all partial derivatives of Eq. 8). We should note, however, that even crude hard-core potentials for electrostatic repulsion in DNA can be applied in many cases to predict statistical properties of DNA to good precision (Vologodskii and Cozzarelli, 1995); therefore, the representation of the electrostatic potential according to Eqs. 7 and 8 is probably a good approximation.

So far, we neglected the fact that the segment length is only approximately equal to l_0. That means that the "charged" segment does not coincide exactly with the "geometrical" segment. This seems to be a good approximation for the excluded volume effects, since the chain is supposed to be stiff with respect to stretching ($\delta \ll 1$). One has to be careful, however, to avoid that during the simulation parts of the chain cross each other through discontinuities between adjacent charged segments, if the length of their phantom geometrical counterparts is greater than l_0. In order to exclude this possibility we define the $\boldsymbol{R}_{ij}$ vector for two segments (i, j) of arbitrary length in the following way:

$$\boldsymbol{R}_{ij} = \frac{1}{l_0}(\boldsymbol{r}^{(m)}_{j(i)} - \boldsymbol{r}^{(m)}_{i(j)}) \tag{10}$$

where $\boldsymbol{r}^{(m)}_{i(j)}$ is a "shifted" middle point of the segment:

$$\boldsymbol{r}^{(m)}_{i(j)} = \begin{cases} \boldsymbol{r}_i + (l_0/2)\boldsymbol{e}_i, & \text{if } |\boldsymbol{r}^{(m)}_j - \boldsymbol{r}_i| \leq |\boldsymbol{r}^{(m)}_j - \boldsymbol{r}_{i+1}|; \quad (a) \\ \boldsymbol{r}_{i+1} - (l_0/2)\boldsymbol{e}_i, & \text{else;} \quad (b) \end{cases} \tag{11}$$

and $\boldsymbol{r}^{(m)}_j = (\boldsymbol{r}_j + \boldsymbol{r}_{j+1})/2$ is the actual middle point of the segment. This means that the geometrical segment that is used to calculate the excluded volume interaction is shifted

with its end toward the segment joint that is closest to the other segment in the interaction. Thereby, any gaps that might appear at the joint due to stretching are automatically closed and the chains cannot cross.

Since forces and torques are the partial derivatives of the energy over the system coordinates, the latter should be specified formally. For each segment i we chose the following four coordinates: the three space coordinates $\boldsymbol{r}_i$ of the segment beginning and the angle φ_i of rotation of the local vector basis ($\boldsymbol{f}_i$, $\boldsymbol{g}_i$, $\boldsymbol{e}_i$) around the $\boldsymbol{e}_i$ axis. For the angle coordinate the zero position is not defined. Therefore, in addition, we need to specify how to keep the φ_i coordinates unchanged while a displacement of the $\boldsymbol{r}_i$ coordinates takes place. The following rule was used to derive forces and torques and to perform moves in the simulation procedure. If, as a result of $\boldsymbol{r}_i$ displacement, the $\boldsymbol{e}_i$ vector has a new value $\boldsymbol{e}_i'$ and $\boldsymbol{A}_i$ is a rotation matrix, so that $\boldsymbol{e}_i' = \boldsymbol{A}_i\boldsymbol{e}_i$, then the new value of the $\boldsymbol{f}_i$ vector is $\boldsymbol{f}'_i = \boldsymbol{A}_i\boldsymbol{f}_i$.

In a linear chain three additional degrees of freedom $\boldsymbol{r}_N$ are required for the last segment. Here are the expressions for the forces and torques for the ith segment obtained by differentiating Eqs. 1, 2, 4, and 7:

Stretching force

The stretching force acting on the ith vertex from the $(i+1)$th one is parallel to the ith segment:

$$\frac{\mathbf{F}^{(s)}_{i,next}}{k_BT} = \frac{s_i - l_0}{(l_0\delta)^2}\,\mathbf{e}_i \tag{12}$$

The total stretching force for the ith vertex is therefore

$$\mathbf{F}^{(s)}_i = -\mathbf{F}^{(s)}_{i-1,next} + \mathbf{F}^{(s)}_{i,next} \tag{13}$$

Bending force

The contribution of the energy stored in the bending angle β_i to the bending force acting on the ith vertex from the $(i+1)$th one is perpendicular to the ith segment and lies in the bend plane:

$$\frac{\mathbf{F}^{(b)}_{i,next}}{k_BT} = \frac{2\alpha_b\beta_i}{s_i}\,\tilde{\mathbf{p}}_i \times \mathbf{e}_i \tag{14a}$$

for an unbent joint, and

$$\frac{\mathbf{F}^{(b)}_{i,next}}{k_BT} = \frac{2\alpha_b\beta_i}{s_i\sin\beta_i}\,(\mathbf{e}_{i-1} \times \mathbf{b}_i) \times \mathbf{e}_i \tag{14b}$$

for a bent joint. Here $\tilde{\mathbf{p}}_i = \mathbf{p}_i/p_i$; $\mathbf{p}_i = \mathbf{s}_{i-1} \times \mathbf{s}_i$, so that $\tilde{\mathbf{p}}_i = \mathbf{e}_{i-1} \times \mathbf{e}_i/\sin\beta_i^*$, where β_i^* is the angle between $\mathbf{e}_{i-1}$ and $\mathbf{e}_i$ (for an unbent joint, $\beta_i^* = \beta_i$).

The analogous contribution to the bending force acting on the ith vertex from the $(i-1)$th one is perpendicular to the $(i-1)$th segment and also lies in the bend plane:

$$\frac{\mathbf{F}^{(b)}_{i,prev}}{k_BT} = \frac{2\alpha_b\beta_i}{s_{i-1}}\,\tilde{\mathbf{p}}_i \times \mathbf{e}_{i-1} \tag{15a}$$

for an unbent joint, and

$$\frac{\mathbf{F}^{(b)}_{i,prev}}{k_BT} = \frac{2\alpha_b\beta_i}{s_{i-1}\sin\beta_i}\,[\mathbf{b}_i - (\mathbf{e}_{i-1}\cdot\mathbf{b}_i)\mathbf{e}_{i-1}] \tag{15b}$$

for a bent joint.

Note that Eqs. 14a and 15a for an unbent joint are symmetric with respect to renumbering the vertices in opposite order (one force can be obtained from the other by substituting $\mathbf{s}_i \leftrightarrow -\mathbf{s}_{i-1}$, $\mathbf{e}_i \leftrightarrow -\mathbf{e}_{i-1}$, $\tilde{\mathbf{p}}_i \leftrightarrow -\tilde{\mathbf{p}}_{i-1}$). Obviously, there is no such symmetry for a bent joint because the $\mathbf{b}_i$ vectors are not defined in a symmetrical way. We note also that the expressions for a bent joint (14b, 15b) become those for the unbent joint upon substituting $\mathbf{b}_i \to \mathbf{e}_i$.

The total bending force for the ith vertex is

$$\mathbf{F}^{(b)} = -\mathbf{F}^{(b)}_{(i-1),next} + \mathbf{F}^{(b)}_{i,prev} + \mathbf{F}^{(b)}_{i,next} - \mathbf{F}^{(b)}_{(i+1),prev} \tag{16}$$

Bending torque

The torque on segment i induced by bending is for a bent joint:

$$\frac{T^{(b)}_i}{k_BT} = \frac{2\alpha_b\beta_i}{\sin\beta_i}\,\mathbf{e}_{i-1}\cdot(\mathbf{e}_i \times \mathbf{b}_i) \tag{17}$$

For an unbent joint, this torque is equal to zero.

Twisting force

The force on the ith vertex from the $(i+1)$th one induced by mutual twisting of the $(i-1)$th and ith segments by an angle τ_i is perpendicular to the plane of the ith bend:

$$\frac{\mathbf{F}^{(t)}_{i,next}}{k_BT} = \frac{C}{k_BT\,l_0}\,\frac{\tau_i s_{i-1}}{p_i}\,(1 - \mathbf{e}_{i-1}\cdot\mathbf{e}_i)\tilde{\mathbf{p}}_i \tag{18}$$

The symmetric expression for the twisting force acting on the ith vertex from the $(i-1)$th one is

$$\frac{\mathbf{F}^{(t)}_{i,prev}}{k_BT} = -\frac{C}{k_BT\,l_0}\,\frac{\tau_i s_i}{p_i}\,(1 - \mathbf{e}_{i-1}\cdot\mathbf{e}_i)\mathbf{p}_i \tag{19}$$

The total twisting force for the ith vertex is then

$$\mathbf{F}^{(t)}_i = -\mathbf{F}^{(t)}_{(i-1),next} + \mathbf{F}^{(t)}_{i,prev} + \mathbf{F}^{(t)}_{i,next} - \mathbf{F}^{(t)}_{(i+1),prev} \tag{20}$$

Twisting torque

The torque on the ith segment induced by twisting the $(i+1)$th segment by an angle τ_{i+1} with respect to the ith one is

$$\frac{T^{(t)}_{i,next}}{k_BT} = \frac{C}{k_BT\,l_0}\,\tau_{i+1} \tag{21}$$

The total twisting torque for the ith segment is then

$$T_i^{(t)} = -T_{(i-1),next}^{(t)} + T_{i,next}^{(t)} \quad (22)$$

Electrostatic force

The contribution of the electrostatic interaction between segments i and j to the force acting on the ith vertex is

$$\frac{\mathbf{F}_{ij,1}^{(e)}}{k_B T} = -\frac{1}{k_B T}\frac{\partial \mathbf{E}_{ij}^{(e)}}{\partial \mathbf{r}}_i = \alpha_e\left(\frac{\partial f_{ij}}{\partial \rho_{ij}}\frac{\partial \rho_{ij}}{\partial \mathbf{r}_i} + \frac{\partial f_{ij}}{\partial \gamma_{ij}}\frac{\partial \gamma_{ij}}{\partial \mathbf{r}_i} + \frac{\partial f_{ij}}{\partial \gamma_{ji}}\frac{\partial \gamma_{ji}}{\partial \mathbf{r}_i} + \frac{\partial f_{ij}}{\partial \sigma_{ij}}\frac{\partial \sigma_{ij}}{\partial \mathbf{r}_i}\right)$$

where $f_{ij} = f(\rho_{ij}, \gamma_{ij}, \gamma_{ji}, \sigma_{ij})$. An analogous relationship is valid for the force $\mathbf{F}_{ij,2}^{(e)}$ acting on the $(i + 1)$th vertex from segment j. We assume that $\mathbf{F}_{ij,1}^{(e)} = \mathbf{F}_{ij,2}^{(e)} = \mathbf{0}$ when $i = j$ or $i = j \pm 1$. The partial derivatives of f (Eq. 8) are tabulated as described above. The expressions for the derivatives of ρ_{ij}, γ_{ij}, γ_{ji}, and σ_{ij} with respect to $\mathbf{r}_i$ and $\mathbf{r}_{i+1}$ are given below, where the following auxiliary notations are used: $\tilde{\mathbf{R}}_{ij} = \mathbf{R}_{ij}/\rho_{ij}$ is a unit vector in the $\mathbf{R}_{ij}$ direction; $\alpha_{i(j)}$ and $\beta_{i(j)}$ are the coefficients in the expression for the "shifted" middle-point of a segment:

$$\boldsymbol{r}_{i(j)}^{(m)} = \alpha_{i(j)}\boldsymbol{r}_i + \beta_{i(j)}\boldsymbol{r}_{i+1},$$

so that (see Eq. 11):

$$\alpha_{i(j)} = (1/2)(1 \pm (1 - l_0/s_i));$$

$$\beta_{i(j)} = (1/2)(1 \mp (1 - l_0/s_i))$$

The upper variant corresponds to the condition (a) in Eq. 11, the lower variant corresponds to the condition (b).

The partial derivatives are thus:

$$\frac{\partial \rho_{ij}}{\partial \boldsymbol{r}_i} = -\frac{1}{l_0}\alpha_{i(j)}\tilde{\boldsymbol{R}}_{ij} \mp \frac{1}{2}\frac{\gamma_{ij}}{s_i}\boldsymbol{e}_i \quad (23)$$

$$\frac{\partial \rho_{ij}}{\partial \boldsymbol{r}_{i+1}} = -\frac{1}{l_0}\beta_{i(j)}\tilde{\boldsymbol{R}}_{ij} \pm \frac{1}{2}\frac{\gamma_{ij}}{s_i}\boldsymbol{e}_i \quad (24)$$

$$\frac{\partial \gamma_{ij}}{\partial \boldsymbol{r}_i} = \frac{1}{s_i}(\gamma_{ij}\boldsymbol{e}_i - \tilde{\boldsymbol{R}}_{ij}) - \frac{1}{\rho_{ij}}\left[\frac{1 \pm 1}{2l_0}\boldsymbol{e}_i + \gamma_{ij}\frac{\partial \rho_{ij}}{\partial \boldsymbol{r}_i}\right] \quad (25)$$

$$\frac{\partial \gamma_{ij}}{\partial \boldsymbol{r}_{i+1}} = -\frac{1}{s_i}(\gamma_{ij}\boldsymbol{e}_i - \tilde{\boldsymbol{R}}_{ij}) - \frac{1}{\rho_{ij}}\left[\frac{1 \mp 1}{2l_0}\boldsymbol{e}_i + \gamma_{ij}\frac{\partial \rho_{ij}}{\partial \boldsymbol{r}_{i+1}}\right] \quad (26)$$

$$\frac{\partial \gamma_{ji}}{\partial \boldsymbol{r}_i} = \frac{1}{\rho_{ij}}\left[\frac{1}{l_0}\alpha_{i(j)}\boldsymbol{e}_j \pm \frac{(\boldsymbol{e}_i \cdot \boldsymbol{e}_j)}{2s_i}\boldsymbol{e}_i - \gamma_{ji}\frac{\partial \rho_{ij}}{\partial \boldsymbol{r}_i}\right] \quad (27)$$

$$\frac{\partial \gamma_{ji}}{\partial \boldsymbol{r}_{i+1}} = \frac{1}{\rho_{ij}}\left[\frac{1}{l_0}\beta_{i(j)}\boldsymbol{e}_j \mp \frac{(\boldsymbol{e}_i \cdot \boldsymbol{e}_j)}{2s_i}\boldsymbol{e}_i - \gamma_{ji}\frac{\partial \rho_{ij}}{\partial \boldsymbol{r}_{i+1}}\right] \quad (28)$$

$$\frac{\partial \sigma_{ij}}{\partial \boldsymbol{r}_i} = \sqrt{1 - \sigma_{ij}^2}\left\{\frac{s_j\gamma_{ji}}{|\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_j|^2\rho_{ij}}\left[\frac{1}{l_0}\alpha_{i(j)}(\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_j) \pm \frac{1}{2}\frac{(\tilde{\boldsymbol{R}}_{ij} \cdot (\boldsymbol{s}_j \times \boldsymbol{s}_i))}{s_i^2}\boldsymbol{e}_i\right] - \frac{\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_i}{|\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_i|^2}\left(1 - \frac{\alpha_{i(j)}s_i\gamma_{ij}}{l_0\rho_{ij}}\right)\right\}; \quad (29)$$

$$\frac{\partial \sigma_{ij}}{\partial \boldsymbol{r}_{i+1}} = \sqrt{1 - \sigma_{ij}^2}\left\{\frac{s_j\gamma_{ji}}{|\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_j|^2\rho_{ij}}\left[\frac{1}{l_0}\beta_{i(j)}(\tilde{\boldsymbol{R}} \times \boldsymbol{s}_j) \mp \frac{1}{2}\frac{(\tilde{\boldsymbol{R}}_{ij} \cdot (\boldsymbol{s}_j \times s_i))}{s_i^2}\boldsymbol{e}_i\right] - \frac{\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_i}{|\tilde{\boldsymbol{R}}_{ij} \times \boldsymbol{s}_i|^2}\left(1 + \frac{\beta_{i(j)}s_i\gamma_{ij}}{l_0\rho_{ij}}\right)\right\}; \quad (30)$$

In Eqs. 29 and 30 the positive value of the square root $\sqrt{1 - \sigma_{ij}^2}$ is taken when $\mathbf{R}_{ij} \cdot (\mathbf{s}_i \times \mathbf{s}_j) > 0$, and its negative value otherwise.

The total electrostatic force acting on the ith vertex is

$$\mathbf{F}_i^{(e)} = \sum_{j=0}^{N-1}(\mathbf{F}_{ij,1}^{(e)} + \mathbf{F}_{(i-1)j,2}^{(e)}) \quad (31)$$

The boundary conditions for the force expressions, Eqs. 12–31, are different for linear and circular chains. For a linear chain all parameters with indexes out of range do not exist and can be formally set to zero. The allowed ranges are $0 \le i \le N - 1$ for $\mathbf{F}_{i,next}^{(s)}$, $\boldsymbol{T}_{i,next}^{(t)}$; $1 \le i \le N - 1$ for $\mathbf{F}_{i,next}^{(b)}$, $\mathbf{F}_{i,prev}^{(b)}$, $\boldsymbol{T}_i^{(b)}$, $\mathbf{F}_{i,next}^{(t)}$, $\mathbf{F}_{i,prev}^{(t)}$; $0 \le i \le N - 1$, $0 \le j \le N - 1$ for $\mathbf{F}_{ij,1}^{(e)}$, $\mathbf{F}_{ij,2}^{(e)}$. For a closed circular chain circular boundary conditions are in effect, such that all indices have to be taken modulo N.

Hydrodynamic interactions

In order to model hydrodynamic interactions defined between spherical objects, a bead with radius a was attached to each chain vertex. With $N' = N + 1$ beads for a linear and $N' = N$ beads for a circular chain, we describe the hydrodynamic interaction between beads i and j by a 3×3 Rotne-Prager tensor:

$$\boldsymbol{D}_{ij} = D_0\frac{3a}{4r_{ij}}\left[\boldsymbol{I} + \frac{\boldsymbol{r}_{ij} \otimes \boldsymbol{r}_{ij}}{r_{ij}^2} + \frac{2a^2}{3r_{ij}^2}\left(\boldsymbol{I} - 3\frac{\boldsymbol{r}_{ij} \otimes \boldsymbol{r}_{ij}}{r_{ij}^2}\right)\right], \quad \text{if} \quad r_{ij} \ge 2a, \quad i \ne j; \quad (32a)$$

$$\boldsymbol{D}_{ij} = D_0\left[\left(1 - \frac{9}{32}\frac{r_{ij}}{a}\right)\boldsymbol{I} + \frac{3}{32}\frac{\boldsymbol{r}_{ij} \otimes r_{ij}}{ar_{ij}}\right], \quad \text{if} \quad r_{ij} \le 2a, \quad i \ne j; \quad (32b)$$

$$\boldsymbol{D}_{ii} = D_0\boldsymbol{I} \quad (32c)$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$, $r_{ij} = |\boldsymbol{r}_{ij}|$, $D_0 = k_B T/6\pi\eta a$, η is the water viscosity, $\boldsymbol{I}$ is a unit 3×3 matrix, and $\boldsymbol{r} \otimes \boldsymbol{r}$ denotes a matrix with the components $(r_\alpha r_\beta)$; $\alpha, \beta = x, y, z$.

The bead radius a is connected with the DNA hydrodynamic radius r_{HD} in the following way. Let us consider two

geometrical objects, both modeling a piece of DNA of Kuhn length B. The first object is a straight line with n beads attached to it so that the distance between neighboring beads is l_0. The number of the beads is equal to the closest integer to B/L_0, and the bead radius is a. The second object is a cylinder of the radius r_{HD} and the length nl_0. We choose the bead radius a in such a way that the diffusion coefficients of the both objects have the same value. The diffusion coefficient for the string of beads is [see, for example, Hagerman and Zimm (1981)]:

$$D^{row} = \frac{1}{3}\,\mathrm{tr}\left(\sum_{i,j=1}^{n} \boldsymbol{S}_{ij}\right)^{-1}$$

where the tensors $\boldsymbol{S}_{ij}$ are related to the diffusion tensors $\boldsymbol{D}_{ij}$ for the bead row in the following way: $\Sigma_{j=1}^{n}\,\boldsymbol{S}_{ij}\boldsymbol{D}_{jk} = \delta_{ik}\boldsymbol{I}$, δ_{ik} is the Kronecker delta symbol.

According to Tirado and Garcia de la Torre (1979, 1980), the diffusion coefficient of a cylinder is:

$$D^{cyl} = (1/3)(D_{\parallel}^{cyl} + 2D_{\perp}^{cyl}) \tag{33a}$$

with

$$D_{\parallel}^{cyl} = \frac{k_B T}{2\pi\eta(nl_0)}\,(\ln p - 0.207 + 0.980\,p^{-1} - 0.133\,p^{-2}) \tag{33b}$$

$$D_{\perp}^{cyl} = \frac{k_B T}{4\pi\eta(nl_0)}(\ln p + 0.839 + 0.185\,p^{-1} + 0.233\,p^{-2}) \tag{33c}$$

$$p = \frac{(nl_0)}{2r_{HD}} \tag{33d}$$

Brownian dynamics algorithm

The simulations were performed using a second-order Brownian dynamics algorithm. A chain displacement from a given conformation $\{\boldsymbol{r}_i(t), \varphi_i(t)\}$ at time t was done in two half-steps. A tentative first-order displacement is:

$$r_i'(t + \Delta t) - \mathbf{r}_i(t) = \sum_{j=0}^{N'} \mathbf{D}_{ij}(t)\,\frac{\mathbf{F}_j(t)}{k_B T}\,\Delta t + \mathbf{R}_i,$$

$$i = 0, \ldots, N' \tag{34a}$$

$$\varphi_i'(t + \Delta t) - \varphi_i(t) = D_{rot}\,\frac{T_{i(t)}}{k_B T}\,\Delta t + \Phi_i,$$

$$i = 0, \ldots, N' - 1 \tag{34b}$$

where $\boldsymbol{D}_{ij}(t)$, $\boldsymbol{F}_j(t)$, $T_i(t)$ are calculated for the given conformation as outlined above; $D^{rot} = k_B T/4\pi\eta r_{RD}^2 l_0$ is the rotational diffusion coefficient assumed to be equal to that of a cylinder of the DNA radius r_{RD} and the length l_0 (without end-effect corrections), Δt is the time step, and the random displacements $\boldsymbol{R}_i$, Φ_i have the following covariances:

$$\langle \boldsymbol{R}_i \rangle = 0, \quad \langle \boldsymbol{R}_i \otimes \boldsymbol{R}_j \rangle = 2\boldsymbol{D}_{ij}\Delta t \tag{35a}$$

$$\langle \boldsymbol{\Phi}_i \rangle = 0, \quad \langle \boldsymbol{\Phi}_i \boldsymbol{\Phi}_j \rangle = 2D^{rot}\,\Delta t\,\delta_{ij} \tag{35b}$$

The $\boldsymbol{R}_i$ values were obtained as $\boldsymbol{R}_i = \Sigma_{i=0}^{N}\,\boldsymbol{A}_{ik}\boldsymbol{X}_k$, where $\boldsymbol{X}_k$ are uncorrelated Gaussian-distributed random vectors with zero mean and unit variance for each component. The $\boldsymbol{A}_{ik}$ tensors are related to the $\boldsymbol{D}_{ij}$ tensors as: $\Sigma_{k=0}^{N}\,\boldsymbol{A}_{ik}\boldsymbol{A}_{jk}^{T} = \boldsymbol{D}_{ij}$. $\boldsymbol{A}_{ik}$ was obtained through a Cholesky factorization.

The final half-step is:

$$\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i'(t + \Delta t) = \sum_{j=0}^{N'} \mathbf{D}_{ij}(t)\,\frac{-\mathbf{F}_j(t) + \mathbf{F}_j'(t)}{k_B T}\,\Delta t,$$

$$i = 0, \ldots, N' \tag{36a}$$

$$\varphi_i(t + \Delta t) - \varphi_i'(t + \Delta t) = \mathbf{D}_{rot}(t)\,\frac{-T_i(t) + T_i'(t + \Delta t)}{2k_B T}\Delta t,$$

$$i = 0, \ldots, N' - 1 \tag{36b}$$

Here $\mathbf{F}_j'(t + \Delta t)$ and $T_i'(t + \Delta t)$ are the forces and the torques calculated for the conformation $\{\mathbf{r}_i'(t + \Delta t), \varphi_i'(t + \Delta t)\}$. Since the values of the diffusion tensor $\mathbf{D}_{ij}$ depend on the chain conformation much weaker than forces and torques, they are kept the same as in the first half-step.

The initial chain conformation is of no importance for a linear chain. However, in modeling a closed circular DNA, the initial conformation determines the linking number difference (White, 1989)

$$\Delta Lk = \Delta Tw + Wr, \tag{37}$$

that, being a topological invariant, does not change during the simulation procedure. In Eq. 37 $\Delta Tw = (1/2\pi)\,\Sigma_{i=0}^{N-1}\,\tau_i$ is the deviation of the DNA twist from its equilibrium value in a linear unstressed chain and Wr is the writhe of the chain.

As an initial conformation we use a flat regular N-sided polygon with $Wr = 0$ and $\tau_i = 2\pi\Delta Lk/N$ for all segments. ΔLk does not necessarily have to be an integer; in that case, $\tau_0 = (\alpha_0 + \gamma_0 + 2\pi\,\Delta Lk) \bmod 2\pi - \pi$. The invariance of ΔLk is explicitly checked during the simulation (by using Eq. 37), as well as the fact that the chain remains unknotted. The algorithm for calculation of writhe can be found in le Bret (1979). The algorithm for the knot checking is described in Frank-Kamenetskil and Vologodskil (1981).

Simulation parameters

Implementation and availability of the program

The BD program *corchy* was implemented in C on a Silicon Graphics Indy R4400. The simulation is initialized from an input parameter file, and the trajectory is written in successive "snapshots" to an output file at predefined simulation time intervals. The program is available as a UNIX tar file on E-mail request from the authors.

The required parameters of DNA for a BD simulation are (the default values, used in the tests below, are given in parentheses): the contour length L, the Kuhn length B (100 nm), the torsional rigidity C ($2.0 \cdot 10^{-19}$ erg cm), the linking number difference ΔLk (for closed circular DNA), the screening factor q for the phosphate charge (0.73), the effective radii: for hydrodynamic interactions r_{HD} (1.2 nm), for rotational diffusion r_{RD} (1.2 nm), and for electrostatic interactions r_{ES} (1.2 nm). To model a DNA molecule with permanent bends a pair of spherical coordinate angles (θ_i^*, φ_i^*) is required for each bent joint i. Environment parameters include the temperature T (293 K), the ionic strength (0.1 M), the dielectric constant of solvent (water) D (80.2), and the solvent viscosity η (1.0 cp).

Another group of parameters is intrinsic to the model but without much influence on its underlying physics. These are the segment equilibrium length l_0 (10 nm), the stretching stiffness parameter δ (0.08), and the time step Δt. To l_0 and δ, in fact, defined values can be assigned, so that l_0 corresponds to the length per basepair and δ to the DNA stretching elasticity as recently measured (Smith et al., 1996). However, calculating to basepair precision is not necessary for the computation of global structural properties of DNA, and we found that the choice of the stretching elasticity did not influence the dynamics significantly. Since these three parameters strongly influence computation time, they should be carefully adapted to each particular task.

The common rule to choose a parameter $P = \{l_0, \delta, \Delta t\}$ is the following. If we are interested in a DNA property X, a formal dependence of X on P can be obtained, and the chosen P value should be large enough to keep the computation time within acceptable limits and small enough so that the dependence of X on P remains very weak.

The problems of choosing the parameters l_0 and δ, as well as the problem of estimating the minimal number of steps, are similar to those occurring in the Metropolis Monte Carlo simulations of DNA (Klenin et al., 1995, 1991). Here, as an illustration, we apply the common rule to estimate the time step Δt—the problem specific to the BD approach.

Time step

The proper choice of the time step should provide the correct statistical properties of a chain, in particular the correct distributions of segment length, of twist and bend angles between neighboring segments. As characteristics of the distributions we took the variance of the segment length $\langle (s - \langle s \rangle)^2 \rangle$, the variance of the twist angle $\langle \tau^2 \rangle$, and the average cosine of the bend angle $\langle \cos \beta \rangle$. (The averaging $\langle \ldots \rangle$ is performed over the simulation time as well as over all the segments.) The exact values for a linear chain in the absence of electrostatic interaction are known:

$$\frac{\langle (s - \langle s \rangle)^2 \rangle}{l_0^2} = \frac{\delta^2 + 3\delta^6}{(1 + \delta^2)^2} \approx \delta^2;$$

$$\langle \tau^2 \rangle = \frac{k_B T l_0}{C}; \quad \langle \cos \beta \rangle = \frac{B - l_0}{B + l_0}.$$

A crude estimation of the upper limit of the time step Δt can be obtained from the condition that the variance of the random terms in Eq. 34 should be smaller than the variance of the thermal fluctuations:

$$2D_0 \Delta t < \langle (s - \langle s \rangle)^2 \rangle; \quad 2D^{rot} \Delta t < \langle \tau^2 \rangle;$$

$$\frac{4D_0 \Delta t}{l_0^2} < 1 - \langle \cos \beta \rangle.$$

The corresponding values for the default parameters are: $\Delta t < 3.5$ ns; $\Delta t < 4.5$ ns; $\Delta t < 90$ ns.

Table 1 shows the dependency of $\langle (s - \langle s \rangle)^2 \rangle$, $\langle \tau^2 \rangle$, and $\langle \cos\beta \rangle$ on the time step Δt for simulations of a 100-nm linear chain without electrostatic interactions ($q = 0$). In the same table dependencies are given for the first-order version of the BD algorithm, in which the second half-step (Eq. 36) is omitted, so that $\mathbf{r}_i(t + \Delta t) \equiv \mathbf{r}_i'(t + \Delta t)$ and $\varphi_i(t + \Delta t) \equiv \varphi_i'(t + \Delta t)$.

The parameter most sensitive to the time step is the variance of the segment length, $\langle (s - \langle s \rangle)^2 \rangle$. For the second-order version, a time step $\Delta t = 1.0$ ns gives a deviation of

TABLE 1 Mean squared fluctuations of the bond length and of the twisting angle, and average bending cosine for linear DNA chains as a function of length of the simulation time step

Second-Order Algorithm								
Δt [ns]	0 (exact)	0.2	0.5	1.0	2.0	3.0	3.5	4.0
$\langle s\text{-}\langle s \rangle \rangle^2 / l_0^2$	0.00632	0.00637	0.00627	0.00613	0.00539	0.00431	0.00377	0.0618
$\langle \tau^2 \rangle$	0.2024	0.206	0.206	0.202	0.188	0.165	0.151	0.139
$\langle \cos \beta \rangle$	0.8182	0.82	0.82	0.82	0.82	0.82	0.81	0.71
First-Order Algorithm								
Δt [ns]	0 (exact)	0.1	0.2	0.5	1.0	2.0		
$\langle s\text{-}\langle s \rangle \rangle^2 / l_0^2$	0.00632	0.00645	0.00658	0.00683	0.00751	0.00947		
$\langle \tau^2 \rangle$	0.2024	0.210	0.211	0.219	0.233	0.274		
$\langle \cos \beta \rangle$	0.8182	0.81	0.82	0.81	0.82	0.81		

3% from the theoretical value for this parameter. To obtain comparable accuracy, the time step for the first-order version has to be set to a value between 0.1 and 0.2 ns. Similar limits apply for the value of the twisting variance, $\langle\tau^2\rangle$. Thus, the gain factor in terms of Δt for the second-order algorithm is ~5–10. At the same time, the loss in terms of computational effort is small, since the most time-consuming calculations of the random displacements are performed only once per step in both algorithms. That means that the second-order version is significantly more efficient than the first-order one. If the approximation is made that the HI matrix is recalculated only once in a certain number of steps (see below), the difference in efficiency between the algorithms is less pronounced.

It remains to check that the electrostatic energy term does not require smaller Δt values than those estimated above. As an example of a property that is very sensitive to the electrostatic interactions we took the superhelix diameter for a closed circular chain. Formally, the mean superhelix diameter is defined as the minimal distance from a chain vertex to all segments with the exception of the m neighboring ones, averaged over all vertices and over the time. The simulations were performed with $m = 2$ for a closed circular chain 100 nm long with linking number difference $\Delta Lk = -4$ and Kuhn length B equal to the equilibrium segment length $l_0 = 10$ nm: the bending energy was switched off to provide maximum elasticity. The chain structure in this case is very tight, but we did not observe any significant dependence of the superhelix diameter on the time step for $\Delta t \leq 2$ ns. That verifies our previous estimations of the time step.

The discretization of the chain into rigid segments can change the statistical behavior as soon as the number of segments is smaller than a certain limit. As shown in the accompanying paper, 10-nm segments are inadequate for the simulation of short fragments with permanent bends. In these cases a touching bead model with a segment length of 3.4 nm was used instead. Here a smaller time step is needed, since the stretching potential per segment is now more rigid, and at the same time the fluctuations increase due to the larger diffusion constant for smaller subunits. In order to prevent a blow-up of the chain, the time step had to be reduced to $\Delta t = 2 \cdot 10^{-10}$ s for touching beads.

Hydrodynamic interactions (HI)

The evaluation of the HI is the most time-consuming part in a Brownian dynamics simulation, since it includes the Cholesky factorization of the HI matrix. This factorization is necessary in order to generate fluctuating forces with the correct covariances (Eq. 35a). The simulation can be accelerated significantly by using the same HI matrix over several time steps instead of recalculating it step by step. It has to be checked, of course, whether this procedure significantly changes the chain dynamics. As a measure for the dynamics of our model DNA we have used the time evolution of the writhe of the superhelix.

In our test simulations a closed chain with 40 segments of 10-nm length and a linking number of $Lk = -4$ was used. Starting with a flat circular conformation, where $Wr = 0$, the writhing number decays exponentially down to a negative value of $Wr = -2.6$ during the relaxation of the conformation into equilibrium. The relaxation time τ is defined as the time the writhe needs to reach $1/e$ of the equilibrium value. Both the relaxation time as well as the average fluctuations of the writhing number at equilibrium were measured for different update intervals of the HI matrix. The results were compared with test simulations without HI in order to check whether the observables were sensitive to HI at all. They are presented in Table 2.

Updating the HI tensor on every step or every 50 steps gave no difference for the parameters listed above exceeding the uncertainties due to statistical errors. To be on the safe side, we used a 10-step HI update interval for the calculation of dynamical properties of superhelical and linear DNAs in the accompanying paper (Merlitz et al., 1998); however, we found that even a 100-step update interval was very satisfactory in most cases. For the evaluation of the end-to-end distribution function, which is an average static quantity and insensitive to dynamical properties of the chain, no HI are needed at all; in fact, the end-to-end distribution function could also be obtained by Monte Carlo simulations.

Statistical errors and the choice of simulation time

If an observable is measured independently N times, the statistical error decreases as $1/\sqrt{N}$. When extracting data from a trajectory of given simulation length t_{sim}, this relation, however, is only valid as long as the measurements are independent. In reality, an observable quantity X in a dynamical simulation will have a memory; its autocorrelation function

$$C(t) = \frac{\langle X(t)X(0)\rangle - \langle X\rangle^2}{\langle X^2\rangle - \langle X\rangle^2}$$

has a characteristic correlation time $t = \tau$ in which $C(t)$ decays to $1/e$.

One way to obtain independent measurements is to wait a long time, $\Delta t_{\text{m}} \gg \tau$, between two measurements to be sure that the new value is completely uncorrelated to the preceding value. This, however, leads to very poor statistics, so that we use a different approach: the measurements are taken at a much higher frequency than $1/\tau$, and correlations are tolerated in order to collect a suitable amount of data. In

TABLE 2 Influence of the updating interval of the HI matrix on the writhe relaxation time

Update interval	τ [μs]
Each step	18.5 ± 0.9
50 steps	19.6 ± 1.0
No HI	24.1 ± 1.0

TABLE 3 Some typical correlation times of the end-to-end distance for linear DNA chains of different lengths

Chain length	τ [μs]	$N_{\text{eff}} = t_{\text{sim}}/2\tau$
80 nm	1	5000
160 nm	11	455
680 nm	300	17

The effective number of measurements N_{eff} was calculated as indicated above. The simulation time t_{sim} was 10 ms and the HI matrix was updated every 10 steps.

the interpretation of the results it is taken into account that the statistical errors do not scale like $\propto 1/\sqrt{N}$, but like $\propto 1/\sqrt{N_{\text{eff}}}$, where $N_{\text{eff}} = t_{\text{sim}}/2\tau$.

In Table 3 the effective number of independent measurements N_{eff} of the end-to-end distance of a linear chain is evaluated for different chain lengths. In all cases the simulation times t_{sim} were identical (10 ms), but while the statistics are good in the case of 80 nm chains and still acceptable in the case of 160 nm chains, they are very poor for the long chain.

We can conclude that 10-ms trajectories are sufficient to describe the end-to-end distance dynamics of the chains studied in the accompanying paper, which were at most 160 nm.

CONCLUSION

We have described the implementation of a general-purpose BD program for the simulation of the dynamics of wormlike polymers that can be represented by a chain of stiff segments connected by bending, torsion, and stretching potentials. The model includes the topological constraints that are important in the descriptions of covalently closed circular DNA, and has the possibility to model permanent bends between segments.

The two most computationally intensive parts of the simulation are the computation of the inter-segment excluded volume forces (here described by a screened Coulomb potential), and the calculation of the random displacements on every Brownian dynamics step. The electrostatic part can be significantly accelerated by precalculating the electrostatic energies and their partial derivatives in a four-dimensional table whose indices are given by the values of four parameters describing the relative orientation of two segments. A further speedup is achieved by recalculating the hydrodynamic interaction matrix only after a certain number of steps; we could show that compared to a recalculation on each step this approximation does not result in appreciable errors.

Further developments to this program, notably embedding the algorithm in an object-oriented environment that facilitates implementation of other chain geometries and constraints for the simulation of larger biomolecular structures, are in progress (Ehrlich et al., 1997).

Note added in proof: The current version of the program, CORCHY++, is a recoding of the BD algorithm is C++. It also contains the earlier Monte Carlo code (Klenin et al., 1995).

This work was supported by a visiting scientist fellowship of the DKFZ to K. Klenin.

REFERENCES

Allison, S. A. 1986. Brownian dynamics simulation of wormlike chains. Fluorescence depolarization and depolarized light scattering. *Macromolecules.* 19:118–124.

Allison, S. A., R. Austin, and M. Hogan. 1989. Bending and twisting dynamics of short linear DNAs: analysis of the triplet anisotropy decay of a 209-basepair fragment by Brownian simulation. *J. Chem. Phys.* 90:3843–3854.

Allison, S. A., S. S. Sorlie, and R. Pecora. 1990. Brownian dynamics simulations of wormlike chains: dynamic light scattering from a 2311 basepair DNA fragment. *Macromolecules.* 23:1110–1118.

Chirico, G., and J. Langowski. 1994. Kinetics of DNA supercoiling studied by Brownian dynamics simulation. *Biopolymers.* 34:415–433.

Chirico, G., and J. Langowski. 1996. Brownian dynamics simulations of supercoiled DNA with bent sequences. *Biophys. J.* 71:955–971.

Ehrlich, L., C. Münkel, G. Chirico, and J. Langowski. 1997. A Brownian dynamics model for the chromatin fiber. *Computer Applications in the Biosciences.* 13:271–279.

Frank-Kamenetskil, M. D., and A. V. Vologodskil. 1981. Topological aspects of the physics of polymers: the theory and its biological applications. *Sov. Phys. Usp.* 24:679–696.

Hagerman, P. J. 1988. Flexibility of DNA. *Annu. Rev. Biophys. Biophys. Chem.* 17:265–286.

Hagerman, P. J., and B. H. Zimm. 1981. Monte Carlo approach to the analysis of the rotational diffusion of wormlike chains. *Biopolymers.* 20:1481–1502.

Klenin, K., M. D. Frank-Kamenetskii, and J. Langowski. 1995. Modulation of intramolecular interactions in superhelical DNA by curved sequences. A Monte Carlo simulation study. *Biophys. J.* 68:81–88.

Klenin, K. V., A. V. Vologodskii, V. V. Anshelevich, V. Y. Klishko, A. M. Dykhne, and M. D. Frank-Kamenetskii. 1991. Computer simulation of DNA supercoiling. *J. Mol. Biol.* 217:413–419.

le Bret, M. 1979. Catastrophic variation of twist and writhing of circular DNAs with constraint? *Biopolymers.* 18:1709–1725.

Merlitz, H., K. Rippe, K. Klenin, and J. Langowski. 1998. Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation. *Biophys. J.* 74:773–779.

Schellman, J. A., and D. Stigter. 1977. Electrical double layer, zeta potential, and electrophoretic charge of double-stranded DNA. *Biopolymers.* 16:1415–1434.

Schurr, J. M., B. S. Fujimoto, P. Wu, and L. Song. 1992. Fluorescence studies of nucleic acids: dynamics, rigidities, and structures. *In* Topics in Fluorescence Spectroscopy, Vol. 3, J. R. Lakowicz, editor. Plenum, New York. 137–229.

Smith, S., Y. Cui, and C. Bustamante. 1996. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science.* 271:795–799.

Stigter, D. 1977. Interactions of highly charged colloidal cylinders with applications to double stranded DNA. *Biopolymers.* 16:1435–1448.

Tirado, M. M., and J. Garcia de la Torre. 1979. Translational friction coefficients of rigid, symmetric top molecules. Application to circular cylinders. *J. Chem. Phys.* 71:2581–2587.

Tirado, M. M., and J. Garcia de la Torre. 1980. Rotational dynamics of rigid, symmetric top molecules. Application to circular cylinders. *J. Chem. Phys.* 73:1986–1993.

Vologodskii, A., and N. Cozzarelli. 1995. Modeling of long range electrostatic interactions in DNA. *Biopolymers.* 35:289–296.

White, J. H. 1989. An introduction to the geometry and topology of DNA structure. *In* Mathematical Methods for DNA Sequences. CRC Press, Boca Raton, FL.