(24) Farid, S.; Martic, P. A.; Thompson, D. R.; Specht, D. P.; Hartman, S. E.; Williams, J. L. R. *Pure Appl. Chem.* **1979**, *51*, 241.
(25) Rettig, W.; Wermuth, G.; Lippert, E. *Ber. Bunsenges. Phys. Chem.* **1979**, *83*, 692.
(26) Cazeau-Dubroca, C.; Peirigua, A.; Ait Lyazide, S.; Nouchi, G. *Chem. Phys. Lett.* **1983**, *98*, 511.
(27) Cazeau-Dubroca, C.; Pririgua, A.; Ait Lyazidi, S.; Nouchi, G.; Cazeau, Ph.; Lapouyade, R. *Chem. Phys. Lett.* **1986**, *124*, 110.
(28) Rotkiewicz, K.; Rubaszewska, W. *Chem. Phys. Lett.* **1980**, *70*, 444.
(29) For recent discussion, see: Lippert, E.; Rettig, W.; Bonačić-Koutecký, V.; Heisel, F.; Miehé, J. A. *Adv. Chem. Phys.* **1987**, *68*, 1.

Global Equilibrium Configurations of Supercoiled DNA†

Ming-Hong Hao and Wilma K. Olson*

Department of Chemistry, Rutgers, the State University of New Jersey, New Brunswick, New Jersey 08903. *Received December 2, 1988*

ABSTRACT: The equilibrium configurations of circular DNA with imposed linking number differences are studied by computer simulation. The configurational energy of the double helix is described by an isotropic elastic deformation model together with a pairwise potential for calculating the self-interaction between distant points along the chain. The axis of the closed double-helical trajectory of the polymer is represented by piecewise cyclic B-spline curves that are approximated by a finite number of controlling points. The global minimum of the total potential energy is obtained with an algorithm that combines Metropolis–Monte Carlo sampling with a simulated annealing procedure. Simulation conditions are examined as are the effects of adjustable energy parameters. The dependence of the algorithm on temperature is related to the actual chain length of the DNA through the well-known relationship between persistence length and temperature. The predicted three-dimensional arrangements are consistent with the configurations of supercoiled DNA observed in electron micrographs. The most stable structures of the closed chain are found to be interwound configurations that are also critically dependent on the specified linking number difference. The toroidal configuration is found to be unstable and, upon energy minimization, is transformed to an interwound form. The optimized trajectories of knotted structures, however, are independent of the linking difference over the limited range of values examined.

Introduction

It is commonly observed in physical studies of naturally occurring closed circular DNA that a sample of homogeneous molecular weight resolves into three well-distinguished components: an intact ring, a nicked ring with one broken strand, and a linear duplex created by two single-strand breaks.[1-4] The linear DNA is well-known to behave as a Porod–Kratky wormlike chain.[5,6] The same model can also be extended, with proper accounting of the ring-closure constraint, to the treatment of nicked DNA.[6,7] An intrinsic supercoiling, however, overrides the configurational fluctuations of the double helix in intact naturally occurring circular DNA, with ordinary techniques of polymer configurational statistics becoming less effective, if not invalid. This supercoiling is attributed to a linking number difference with respect to the natural linking number of the relaxed double helix. The linking number is a topological parameter equal to the number of times one strand of the DNA links through the closed circle formed by the other strand.[3,8-10] In relaxed circular DNA (e.g., the planar closed circle) the linking number is simply the winding number of the chain as measured by the number of residues divided by the intrinsic helical repeat. In supercoiled DNA, however, the linking number is a sum of the twisting of the chain about the helical axis plus the writhing of the folded helical trajectory.[8,9] The linking number difference in a supercoiled duplex can only be relieved by chemical breakage of one of its two circular strands or by perturbation of the intrinsic helical repeat of the DNA through change of environment[11-14] or interaction with other molecules.[4,15]

A useful approach in understanding DNA supercoiling is to characterize the configuration of the chain near its energy minimum. A convenient macroscopic model is provided by a long elastic rod, which when overtwisted and closed into a circle, folds into a unique three-dimensional shape. Various treatments[16-19] have accordingly been offered to treat the supercoiling of DNA in a manner analogous to the treatment of the mechanical equilibrium of a twisted and closed elastic rod.[20,21] Solutions to the three-dimensional arrangement of a linear rod have been explicitly obtained,[17,22] and the solution to the circular rod problem has been discussed.[22,23] The elastic equilibrium configurations of the closed rod are reported to be toroidal-like forms,[23] n-fold knots,[22] and n-leafed roses.[22] These elastic treatments, however, are still incomplete in the context of the DNA supercoiling problem in that no forces or torques other than those acting on the ends of the rod have been considered. Because closed circular DNA is usually interwound when examined under the electron microscope,[24] there is the possibility that the solution configuration of a supercoiled molecule may also adopt an interwound form.

A different approach to the supercoiling question is to consider the configuration of minimum deformation energy of a closed circular DNA with given linking number difference. Fuller has developed such a theory, predicting the interwound configuration to be a possible solution.[9] Two different attempts, based on this formalism, have been recently offered to obtain solutions to the problem.[25,26] The minimization of the deformation energy, however, leads to highly nonlinear equations that have been treatd only by simplified variational methods (in which the helical parameters of the interwound helices are purposely specified as variational parameters). Even though such treatments are useful, they are clearly not the complete

† From the thesis submitted by Ming-Hong Hao in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Rutgers, the State University of New Jersey, 1988.

0024-9297/89/2222-3292$01.50/0

solution. In addition, the self-intersection of the DNA molecule due to intrachain contacts has not been considered in earlier treatments. There are noticeable self-contact points that must be included as basic energy contributions in an interwound DNA supercoil. It is therefore of interest to find the explicit solution of the supercoiling problem under more complete conditions.

In this paper, we present a new method to study the supercoiled configuration of minimum deformation energy. We use the usual elastic model to describe the twisting and bending energies of the double helix, and we explicitly include chain self-interaction through a pairwise function of the contour points. We treat the linking number difference as an independent variable in the determination of the energy minimum. To minimize the potential energy, we represent the exact trajectory of the supercoiled DNA by piecewise B-spline curves. In the limit of a sufficiently large number of controlling points, these curves provide an exact representation of the axis of the supercoiled trajectory. We approximate the trajectory, however, with a finite number of controlling points and therefore smooth the actual supercoil to some extent. We find that, upon optimization of the total energy, the B-spline curves reproduce the correct global features of the interwound configuration with only a small number of controlling points. We combine a Monte Carlo method with a simulated annealing procedure to minimize the total energy of the above model. We demonstrate the efficiency of the method with a number of examples, and we detail explicit solutions to the energy minimization problem. Finally, we rationalize the minimization method in terms of the flexibility of supercoiled DNA with respect to its chain length, demonstrating the close relationship of the method to the configurational statistics of linear and nicked chains.

Methodology

Potential Energy. From general considerations, the factors that are most important in determining the configuration of a supercoiled DNA may be listed as follows: (1) the deformation energy of the DNA molecule along a supercoiled trajectory; (2) the fixed linking number difference; (3) the self-interaction of distant parts of the chain; (4) the exact ring closure constraint; (5) the given chain length. In what follows, we are interested in those supercoiled configurations that minimize the deformation energy while simultaneously subject to the specified constraints.

A locally rodlike model is adopted here, and the deformation energy of the DNA is calculated in terms of the elastic energy of a slightly deformed rod.[21] Such a model has been used in a number of studies of the static[25-27] and dynamic[28] properties of DNA. The DNA is usually assumed to have a circular cross section so that the total deformation energy E can be expressed as[9,20,21,25]

$$E = (A/2)\int_0^L \kappa(s)^2\,ds + (C/2)\int_0^L (\omega - \omega_0)^2\,ds \quad (1)$$

where κ is the curvature of the centroid of the rod, ω the rate of local twist, ω_0 the intrinsic rate of twist, A the coefficient of flexural (e.g., bending) rigidity, C the twisting stiffness constant, s the arc length, and L the total contour length of the supercoil.

The linking number difference, denoted by ΔLk, must be a constant when the two ends of the DNA rod are rigidly abutted to form a circular chain. Using the well-known topological relationship[8,9] Lk = Wr + Tw, between the linking number Lk, the writhing number Wr, and the total twist number Tw of circular DNA, ΔLk can be expressed as

$$\Delta\mathrm{Lk} = \mathrm{Wr} + (1/2\pi)\int_0^L (\omega - \omega_0)\,ds \quad (2)$$

where $\mathrm{Tw} = (1/2\pi)\int_0^L \omega\,ds$ is the number (not necessarily an integer) of revolutions of one of the DNA strands about the supercoiled trajectory, and Wr is a parameter describing the three-dimensional folding of the helical pathway. The latter quantity can be obtained from the double-contour Gauss integral[29-31]

$$\mathrm{Wr} = \frac{1}{4\pi}\int_0^L \int_0^L \frac{(\mathbf{t}(s_1) \times \mathbf{t}(s_2))\cdot(\mathbf{r}(s_1) - \mathbf{r}(s_2))}{|\mathbf{r}(s_1) - \mathbf{r}(s_2)|^3}\,ds_1\,ds_2 \quad (3)$$

where $\mathbf{r}(s)$ is the expression for the trajectory and $\mathbf{t}(s) = d\mathbf{r}/ds$ is the unit tangent of the trajectory. As noted below, eq 2 can be used to eliminate the twisting component from eq 1.

The self-interaction of the chain molecule from self-contacts is treated by an empirical function with the familiar 12-6 potential. The interaction of chain contour points i and j at a distance r_{ij} is expressed as

$$V_{ij} = \epsilon[(d/r_{ij})^{12} - (d/r_{ij})^6] \qquad r_{ij} \le d$$
$$V_{ij} = 0 \qquad r_{ij} > d \quad (4)$$

where ϵ is an empirical parameter describing the depth of the pairwise energy well and d is the effective interaction (e.g., repulsive) radius of the DNA. The total self-interaction energy is summed over all pairs of contour points that are within a distance d. This function is justified by simple considerations of the excluded-volume effect between distant parts of a long polymer chain. A hard-sphere potential would be sufficient to take account of the excluded-volume effect. A better account of the chain self-interaction, however, is provided by the general Lennard-Jones potential $(a/r)^n - (b/r)^m$. The 12-6 potential is one of the Lennard-Jones functions. It is expected that when only short-range repulsive interactions are considered, the final configuration of supercoiled DNA should not be sensitive to the chosen potential form. Equation 4 is thereby used to approximate the continuous repulsive interactions between parts of the DNA backbone that come within a predetermined contact range. The exact nature of this repulsive interaction is not considered.

Since the curvature integral in eq 1 and the potential in eq 4 are independent of the local twist of the DNA rod, the minimization of the total energy of the supercoiled duplex can be carried out in two steps in a manner similar to Fuller's approach.[9] First, the variable ω is eliminated from a given configuration by minimizing the twist energy subject to the constraint of eq 2. Second, the resulting energy function is minimized to give the spatial configuration of the supercoiled DNA. When the first step is completed, eq 4 is unaffected, but eq 1 is transformed to

$$E = (A/2)\int_0^L \kappa(s)^2\,ds + (2\pi^2 C/L)(\Delta\mathrm{Lk} - \mathrm{Wr})^2 \quad (5)$$

where ω is replaced by the constant value $(2\pi^2/L)\mathrm{Tw}$. Minimization of the combined potential of eq 4 and 5 under the constraints of ring closure and chain length is expected to produce the structures that dominate the configurations of supercoiled DNA. It is noted that eq 5 is a function only of the spatial trajectory of the double-helical axis. This is due to the constant twist rate along the whole trajectory and the consequent simplification of the twisting energy contribution to the form of the second term in eq 5.

B-Spline Curve Smoothing. A straightforward treatment of the problem of DNA supercoiling would be

to consider the detailed trajectory itself. The analytical solution to the minimization problem of eq 5 alone is sufficiently difficult to warrant a numerical treatment, and the introduction of the self-interaction term further complicates the situation. A method that can facilitate the convergence of the minimization of the energy function will therefore be helpful in solving the problem.

It is known that the problem of finding the global minimum of a badly behaved function with multiple local minima can be simplified by the use of a piecewise smoothing function to represent the problematic function.[32] If the smoothing function is properly defined and is a proper representation of the global features of the original function, the solution to the global minimum of the modified function is easier to obtain and identical with the solution to the original one. It has been recently pointed out[33] that B-spline curves provide a convenient formalism for smoothing the trajectories of supercoiled DNA. The cyclic form of the curves[34,35] can be employed to satisfy the ring-closure constraints of the chain molecule. If $\mathbf{p}_1$, $\mathbf{p}_2$, ..., $\mathbf{p}_n$ are a series of ordered points taken from the actual trajectory of a curve, the order 4 cyclic B-spline curve $\mathbf{R}$ approximates the primitive trajectory with the piecewise expression

$$\mathbf{R}_i = F_1(u)\mathbf{p}_{i-1} + F_2(u)\mathbf{p}_i + F_3(u)\mathbf{p}_{i+1} + F_4(u)\mathbf{p}_{i+2} \quad (6)$$

where $0 \leq u \leq 1$, $i = 2, 3, ..., n+1$ (modulo n), and the $F_i(u)$ are basis functions defined by

$$F_1(u) = (-u^3 + 3u^2 - 3u + 1)/6$$
$$F_2(u) = (3u^3 - 6u^2 + 4)/6$$
$$F_3(u) = (-3u^3 + 3u^2 + 3u + 1)/6$$
$$F_4(u) = u^3/6 \quad (7)$$

The points $\mathbf{p}_i$ are called the controlling points of the B-spline curves.

These B-spline curves have the following known properties: (1) they lie in the convex hull of the polygon formed by the points $\mathbf{p}_i$ and form a smooth curve according to the shape of the polygon;[35] (2) they have the property of variance-diminishing (i.e., they represent a straight line exactly, and the number of intersections of the B-spline approximation to a curve with any straight line does not exceed the number of crossings of that straight line by the original curve);[35-37] (3) they represent the exact trajectory of the original curve when the maximum distance $|\mathbf{p}_{i+1} - \mathbf{p}_i|$ approaches zero.[37] The B-spline curves also have convenient analytical properties. Order k B-spline curves, for example, have continuous derivatives up to order $k-2$. Because the current model of DNA supercoiling requires knowledge only of first and second derivatives (see below), order 4 B-spline curves are employed in this work.

Property 3 indicates that only in the limiting case when the maximum distance $|\mathbf{p}_{i+1} - \mathbf{p}_i|$ approaches zero and the number of chosen controlling points becomes infinite do the B-spline curves exactly represent the trajectory of the original curve. These curves, however, are known to give good approximations to the shapes of regular curves with only a few controlling points. For example, the relative deviations in the curvature of the B-spline approximation to a planar circle are only ±2% with an octagon as the determining polygon.[33] In principle, it should be possible to choose the number of controlling points sufficiently large so that the resulting B-spline curve is sufficiently close to the exact trajectory of a supercoiled DNA. The most computationally effective B-spline curves, however, are those with a modest number of controlling points. In what follows, B-spline curves with a small to modest number of controlling points are carefully investigated. In such an approach, the local fluctuations of the supercoiled DNA are largely omitted, and only the global features of DNA supercoiling are revealed.

With the B-spline representation, the curvature integral and the writhing number appearing in eq 5 can be expressed, respectively, as parametric integrals over u. Specifically

$$\int_0^L \kappa(s)^2\, ds = \sum_{i=2}^{n+1} \int_0^1 \frac{|\mathbf{R}'_i(u) \times \mathbf{R}''_i(u)|^2}{|\mathbf{R}'_i(u)|^5}\, du \quad (8a)$$

$$\mathrm{Wr} = \frac{1}{4\pi} \sum_{i=2}^{n+1} \sum_{j=2}^{n+1} \int_0^1 \int_0^1 \frac{[\mathbf{R}'_i(u_1) \times \mathbf{R}'_j(u_2)]\cdot[\mathbf{R}_i(u_1) - \mathbf{R}_j(u_2)]}{|\mathbf{R}_i(u_1) - \mathbf{R}_j(u_2)|^3} du_1\, du_2 \quad (8b)$$

where $\mathbf{R}' = d\mathbf{R}/du$ and $\mathbf{R}'' = d^2\mathbf{R}/du^2$. The elastic energy is then a function of the controlling points that define the configuration of the supercoiled DNA.

Equation 8 is a double summation, which is time-consuming to evaluate by computer. The formula, however, can be reexpressed as a single summation based on a theorem first proposed by Fuller.[38] Only a special case, which is related to an algorithm described below, is considered. If C_2 is the curve formed from C_1 by moving a single controlling point $\mathbf{p}_i$ by a small amount, and Wr_1 and Wr_2 are the writhing numbers of C_1 and C_2, respectively, the difference in writhing number between the two curves can be written as

$$\mathrm{Wr}_2 - \mathrm{Wr}_1 = \frac{1}{2\pi} \sum_{k=i-2}^{i+1} \int_0^1 \frac{\mathbf{R}'_{1k} \times \mathbf{R}'_{2k}}{|\mathbf{R}'_{1k}||\mathbf{R}'_{2k}| + \mathbf{R}'_{1k}\cdot\mathbf{R}'_{2k}} \left[\frac{\mathbf{R}''_{1k}}{|\mathbf{R}'_{1k}|} - \frac{\mathbf{R}'_{1k}(\mathbf{R}''_{1k}\cdot\mathbf{R}'_{1k})}{|\mathbf{R}'_{1k}|^3} + \frac{\mathbf{R}''_{2k}}{|\mathbf{R}'_{2k}|} - \frac{\mathbf{R}'_{2k}(\mathbf{R}''_{2k}\cdot\mathbf{R}'_{2k})}{|\mathbf{R}'_{2k}|^3} \right] du \quad (9)$$

where subscripts 1 and 2 refer to curves C_1 and C_2, respectively. The result is the area enclosed by the tangent vectors of C_1 and C_2 defined on the surface of a unit sphere divided by 2π (e.g., the difference in area between the indicatrixes of the two curves).

Numerical Approaches. The B-spline representation parametrizes a continuous supercoiled trajectory, which is easily treated numerically. The optimization of the configuration is a nonlinear minimization problem. While there are numerous computational techniques that can be used to find the local minima of a multidimensional function,[39] we have used the following relatively new approach for reasons that become apparent later.

Suppose that $\mathbf{X}$ is the three-dimensional configuration of a supercoil. In the B-spline representation, $\mathbf{X} = \mathbf{X}(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n)$, where the $\mathbf{p}_i$ ($i = 1-n$) are the controlling points. If $\mathbf{X}^*$ is the unique configuration of the global energy minimum, Pincus has proven that $\mathbf{X}^*$ can be analytically expressed as[40]

$$\mathbf{X}^* = \lim_{\beta\to\infty} \frac{\int \mathbf{X} \exp(-\beta E(\mathbf{X}))\, d\mathbf{X}}{\int \exp(-\beta E(\mathbf{X}))\, d\mathbf{X}} \quad (10)$$

where β is a parameter and the integral is evaluated over all possible configurations. The limiting integral in eq 10, however, can be explicitly evaluated for only a few simple analytical functions. For a complicated system like a supercoiled trajectory, the integral must be determined nu-

merically. Pincus has also noted the connection between the minimization problem formulated by eq 10 and the Metropolis–Monte Carlo method.[41] In the Metropolis approach,[42] the configurational integral is evaluated by a Monte Carlo procedure with the parameter β equal to $1/k_bT$, where k_b is the Boltzmann constant and T the absolute temperature. In this connection, an increase of β is equivalent to a decrease of T. A practical procedure for finding the global minimum of a multidimensional function based on eq 10 is to evaluate the integral by Monte Carlo methods through a series of gradually decreasing temperatures until the temperature is brought to a sufficiently low range. This procedure is also known in the literature as simulated annealing.[43-45]

The algorithm used here is as follows. The Monte Carlo integration of eq 10 is started at a sufficiently high temperature (or small β). The starting configuration $\mathbf{X}^0 = \mathbf{X}(\mathbf{p}_1^0, \mathbf{p}_2^0, ..., \mathbf{p}_n^0)$ is arbitrarily but reasonably chosen. The positions of the controlling points are adjusted in small steps, randomly but guided by the laws of probability. A new configuration is generated from an old one by displacing a randomly chosen controlling point. The energy of the new configuration is then evaluated and compared with that of the old. The transition rule in the Monte Carlo procedure is determined according to the Metropolis method. If the energy of the new configuration is lower than that of its predecessor, it is automatically accepted (i.e., $\Delta E < 0$). If the energy is greater, the probability $w = \exp(-\Delta E/k_bT)$ is calculated and compared with a random number between zero and unity. The new configuration is accepted only if w is greater than or equal to the random number. Otherwise, the old configuration is retained for the next move. It can be shown that after a large number of such moves, the system is able to evolve into a Boltzmann distribution at a given temperature.[42]

As the system approaches the equilibrium distribution at a given temperature, the temperature is then lowered by a small amount. The system is again moved at random according to the Metropolis procedure until a new equilibrium distribution (at the lower temperature) is reached. The cooling process is repeated several times until the temperature is below the freezing point, where the configuration is no longer changed. The simulation is analogous to the growth of a single crystal from a melt by carefully cooling the solution, hence the term "simulated annealing". The outcome of the Monte Carlo simulation is the limiting integral of eq 10. According to Pincus theory, the result is the global minimum energy configuration of the supercoil under the given potential. A deeper exposition of the algorithm is provided below when numerical results are presented.

The constraint of given chain length is treated here in such a way that the numerical algorithm can be executed. If the total contour length of the DNA is kept strictly constant, few moves can be made, even at high temperatures. This constraint is therefore relaxed so that the total contour length of circular DNA is able to fluctuate to a small degree about its given or starting value L_0. The pseudopotential

$$V_L = K(L - L_0)^2 \tag{11}$$

where K is a constant and L the actual contour length, is used to ensure that the chain length is flexible yet close to its given value during the simulation. Equation 11 is not a unique way to maintain the given chain length. An alternative method might be to accept all configurations generated with chain length $L_0 \pm \Delta L$, where ΔL is a small number. The distribution of acceptable chain lengths is described in this case by a step function rather than the exponential distribution associated with eq 11.

The starting temperature of the simulation can be judged by the percentage of trial moves that are accepted. The temperature is chosen so that about 30–50% of the moves among all trials are accepted. The probability of moving is correlated with the magnitude of the moving step. Step size must be small enough to locate the minimum with sufficient accuracy but large enough to reach the equilibrium distribution in a reasonable time. The step size is chosen here to be roughly 0.5–1.0% of the total contour length of the DNA supercoil. Once the step size is determined, a few hundred test runs are performed to determine the appropriate starting temperature. All integrals in the simulation are evaluated numerically. The parameter u is varied between 0 and 1 at increments of 0.05. This increment is found to give sufficiently accurate numerical values of the integrals. It is noted that only one controlling point is changed at each move. If $\mathbf{p}_i$ is moved, the altered portion of the supercoiled pathway is limited to that between $\mathbf{R}_{i-2}$ and $\mathbf{R}_{i+1}$ (referring to the notation of eq 6). To increase computational efficiency, numerical integrals are evaluated in separate regions of i and the components are saved for subsequent moves. When a controlling point is moved, only the four components involved need to be updated.

The reason for using the above seemingly cumbersome Monte Carlo procedure is its suitability in identifying the unique supercoiled configuration at the global energy minimum. For many minimization problems, if the starting configuration is well chosen, algorithms such as the steepest descent and conjugate gradient methods[39] are well-known to converge much faster to the energy minimum than the Monte Carlo method. Gradient methods, however, are very likely to be trapped by local minima, and the results obtained from such methods are highly dependent on the starting configuration. The B-spline curves, on the other hand, are themselves the local minimum (i.e., the elastic line) of the configuration determined by the set of controlling points. A given spline trajectory is improved, in our experience, only slightly by the steepest descent method. Furthermore, there is generally no predetermined information as to whether or not a starting configuration is close to the global minimum. The appropriate procedure is one that can overcome the local minima and find the global minimum independent of the starting configuration. This is the precise advantage of the simulated annealing method. At a given temperature there is always a finite probability that the configuration will move out of a local minimum.

Results and Discussion

Minimization by Simulated Annealing. The temperature decrease is the key to the minimization of the energy function in the present algorithm. The following examples are illustrative of the cooling procedure of the algorithm. Two possible procedures for decreasing the temperature during the simulation are considered. One is to use a constant cooling rate whereby the current temperature is multiplied by a constant factor f of less than unity at each temperature change. Another is to employ an accelerative cooling procedure in which the temperature is decreased, at the mth stage, by a factor of $(f_0)^m$ with $f_0 < 1$ being the initial cooling rate. A comparison of the variation of energy with temperature for the two procedures is shown in Figure 1. These quantities are reported in reduced form as $\tilde{E}$, a unitless quantity related to the total energy $E + V_{ij}$ by the factor $2\pi^2C/L$, and $\tilde{T}$, a quantity proportional to the temperature. The starting temperature T_0 is defined by $k_bT_0 = 0.1E_0$, where E_0 is the

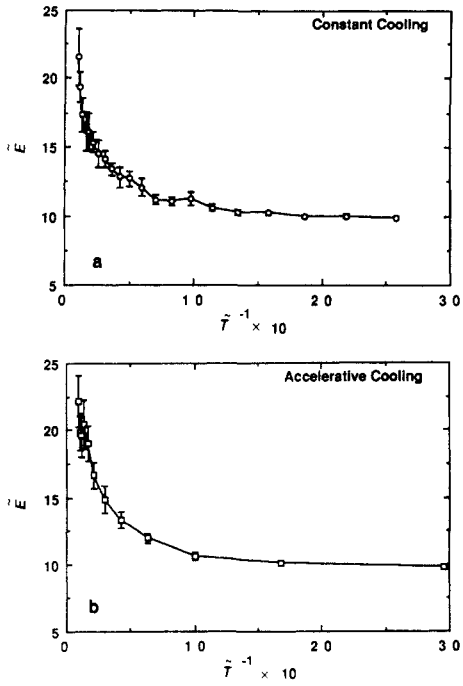

Figure 1. Variation of reduced energy $\tilde{E}$ with respect to reduced temperature $\tilde{T}$ in the simulated annealing process. Total elastic and self-contact energies are reduced by the factor $2\pi^2C/L$ and $\tilde{T}$ is the fraction of the initial temperature. (a) Constant cooling and (b) accelerative cooling of an elastic chain with $\Delta Lk = 4$, $A/C = 1.5$, $\tilde{\epsilon} = 3$, $d = L_0/90$, and $K = 0.3$. Refer to the text for further details.

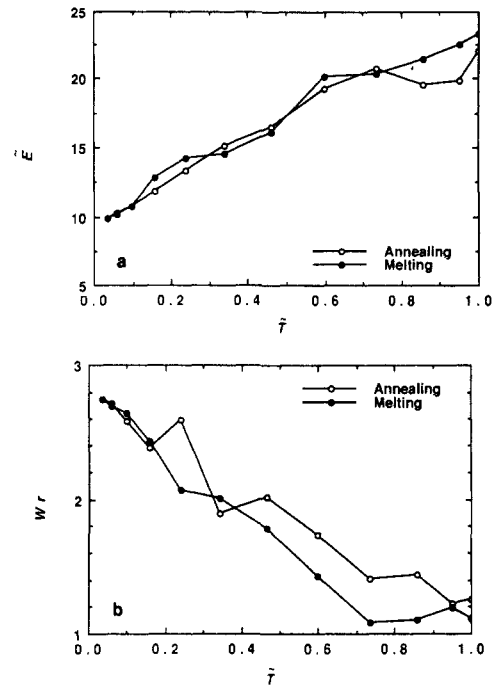

Figure 2. Comparison of the cooling and heating processes in the simulated annealing procedure illustrated by the variation of (a) reduced energy and (b) writhing number with respect to reduced temperature. The simulated supercoil is described by the parameters listed in the legend to Figure 1.

initial energy, and $\tilde{T}$ is defined by the ratio T/T_0. An arbitrary configuration defined by 14 controlling points (see the legend to Figure 3 for coordinates and the example in Figure 5a for the pattern of the supercoil) is used as the starting configuration for the two simulations. The reduced energy is independent of the chain length, so that L_0 can be arbitrarily chosen as the starting chain length in the simulation. The linking number difference ΔLk is fixed at 4, the ratio of bending to twisting force constants A/C at 1.5, the coefficient of the reduced self-contact energy $\tilde{\epsilon}$ at 3, the self-contact distatnce d at $L_0/90$, and the stretching force constant K at 0.3. At each temperature stage, the configuration is relaxed by 12000 Monte Carlo moves. (For mere minimization purposes, however, fewer Monte Carlo runs are sufficient.) A cooling factor of 0.85 is used for the constant cooling procedure in Figure 1a, and an initial value of $f_0 = 0.95$ for the accelerative cooling procedure in Figure 1b. The energies reported in the figure are average values, and the error bars are the standard deviations of the energies at the specified temperatures. The global minimum is given by the energy value at the lowest temperature.

It is observed from Figure 1 that the energy fluctuations are largest near the starting temperature of both energy trajectories, consistent with increased flexibility of the superhelix at high temperature. A slow cooling rate is therefore helpful to the system in reaching an equilibrium distribution in the high-temperature range. Below a certain temperature, however, the energy of the system decreases very slowly with decrease in temperature. If a slow cooling rate is used in the low-temperature region, the result is not significantly improved over the more computationally efficient accelerative cooling procedure. With the latter algorithm, the cooling rate is slow at the start of the simulation (high-temperature region) to guarantee that the equilibrium distribution is maintained at each decrease of temperature, but fast in the lower temperature range to locate the required energy minimum quickly. Because of these advantages, most of the simulations discussed below are performed with the accelerative cooling procedure.

The simulated annealing procedure is basically a reversible process. If the cooling has been carried out so that an equilibrium is attained at each temperature, a basically similar energy trajectory is obtained when the process is reversed (i.e., the system is heated at a similar rate). The result of such an annealing and melting cycle is illustrated in Figure 2, where the average reduced energy and the writhing number are plotted as functions of the reduced temperature for the heating and cooling trajectories. The cooling schedule adopted is the accelerative cooling procedure of Figure 1. The heating process is started from the final frozen configuration of the cooling sequence with the heating factor given by the reciprocal of the cooling factor at each temperature. As observed in Figure 2a, the energy of the melting trajectory is very similar to that of the cooling (e.g., annealing) trajectory. The variation of Wr with $\tilde{T}$ in Figure 2b is also comparable along the heating and cooling curves. During the annealing process in Figure 2, an irregular coil is frozen into a regular interwound helix with roughly three helical turns (see Figure 11 below for representative configurations as a function of temperature), and the process is reversed during the heating process. The observed hysteresis in Figure 2 is attributed to the random features of the algorithm and the finite number of configurations sampled.

A Simple Supercoil: The Figure-8 Configuration. The closed figure-8 configuration is found in the classic elastica problem, in which an elastic thin rod is allowed to bend in a plane subject to isolated forces or torques acting at its ends.[20] The problem can be solved analytically with the closed figure-8 configuration occurring at a special bending angle. This problem, however, is not exactly the same as the DNA supercoiling problem. Ring closure is an intrinsic constraint in the DNA problem, and supercoiling is due to a linking number difference in the mol-

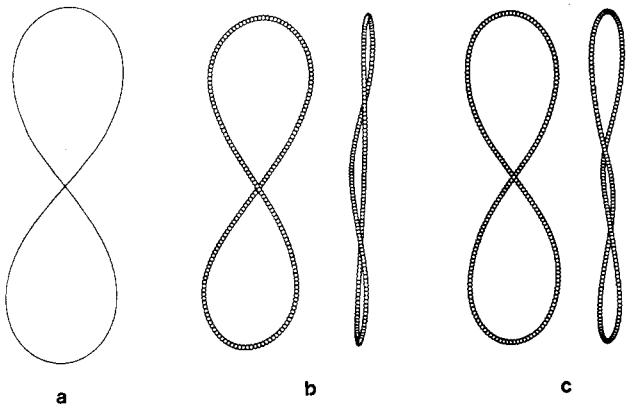

Figure 3. Comparison of (a) the exact elastic equilibrium configuration[20] of a thin rod in the figure-8 form with the minimum energy configurations obtained from B-spline approximations with (b) 8 and (c) 14 controlling points. Controlling points of the starting configuration of the simulated 8-point structure are (0.0, 4.0, 0.0), (−2.0, 2.0, 0.5), (0.0, 0.0, 1.0), (2.0, −2.0, 0.5), (0.0, −4.0, 0.0), (−2.0, −2.0, −0.5), (0.0, 0.0, −1.0), (2.0, 2.0, −0.5), while those of the initial 14-point model are (0.6, 0.6, −1.0), (−1.4, −0.6, −1.0), (−3.0, −3.0, −0.60), (−2.8, −4.8, −0.2), (0.0, −6.0, 0.0), (2.0, −5.2, 0.2), (3.0, −3.0, 0.6), (1.4, −1.4, 1.0), (−1.4, 0.6, 1.0), (−3.6, 3.0, 0.6), (−2.8, 4.8, 0.2), (1.0, 6.0, 0.0), (2.0, 4.8, −0.2), (3.0, 3.0, −0.6).

ecule. Nevertheless, it has been shown by Tsuru and Wadati[26] that the same figure-8 form is a local minimum energy configuration that becomes more stable than the planar circle when the linking number difference is greater than a critical value determined by the ratio A/C of bending to twisting force constants.

As a simple example, we have compared the minimum energy figure-8 configuration obtained with the current simulated annealing procedure with that given by the exact solution to the elastica problem.[20] The configuration of the exact solution is illustrated in Figure 3a and the trajectories obtained from the simulated annealing of B-spline trajectories with 8 and 14 controlling points in Figure 3b,c, respectively. The starting configurations in the numerical procedure are arbitrarily chosen figure-8 shapes generated with the controlling points given the figure legend. The chain with fewer controlling points is annealed by a constant cooling procedure with $f = 0.80$ and the chain with more points by an accelerative procedure with $f_0 = 0.93$. The systems are relaxed with 4000 moves at each cooling temperature. In both cases ΔLk is fixed at 2, the ratio A/C at 1.5, $\tilde{\epsilon}$ at 3, d at $L_0/90$, and K at 0.3. Because of the self-contact energy contribution, the energy-minimized figure-8's are three-dimensional structures, in contrast to the planar structure of the elastica. The computationally generated structures are therefore shown in two different planar projections in Figure 3b,c. The two trajectories are essentially identical in the main plane projection but are different from the planar configuration of the exact solution when viewed in a perpendicular projection. The writhing numbers of the optimized 8- and 14-point closed curves are 1.01 and 1.04, respectively. Despite the small number of controlling points in the B-spline approximations, the configurations are very similar to the elastica structure.

The simulated annealing method has also been used to estimate the critical value of ΔLk at which the equilibrium figure-8 configuration becomes more stable than the planar circle. The latter is a local minimum configuration over a certain range of ΔLk.[18] The results of the minimization of a series of closed curves of variable linking number difference are reported in Figure 4. The reduced energy and the writhing number are plotted as functions of ΔLk over the range 0–2. The starting trajectory in all cases is

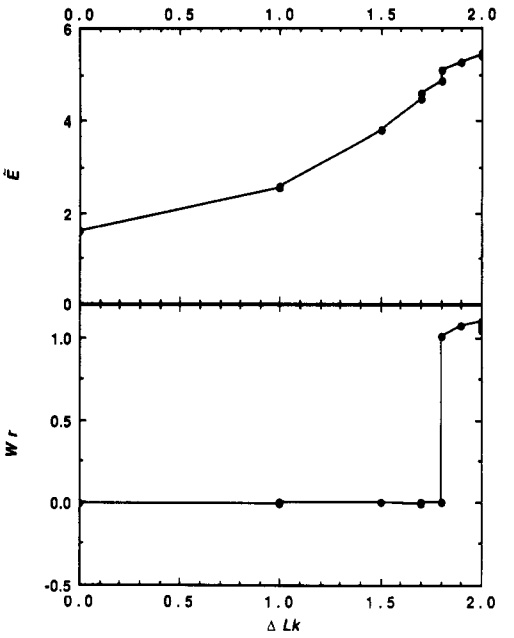

Figure 4. Variation of reduced energy and writhing number of the global minimum configuration of an elastic chain as a function of linking number difference. Energy parameters are given in the legend to Figure 1.

the 14-point B-spline curve detailed in Figure 3, and the energy parameters are the values specified above. Each simulation was performed with a 12-step accelerative cooling procedure ($f_0 = 0.95$) entailing 12 000 Monte Carlo moves per step. According to the figure, the energy-minimized structures with ΔLk < 1.7 are planar circles with writhing numbers of approximately zero, while the structures with ΔLk > 1.9 are figure-8 forms with writhing numbers close to unity. Repeated simulations with ΔLk = 1.8 are found to yield both circular and figure-8 structures. Interestingly, the reduced energy of the annealed circle (4.9) is somewhat lower than that (5.1) of the figure-8 with ΔLk = 1.8, suggesting that the critical linking number difference at which the circle becomes less stable is between 1.8 and 1.9. This value is in agreement with the analytical results of Tsuru and Wadati,[26] who find with the same ratio of A/C that the circular configuration is less stable than the figure-8 configuration when the critical linking number difference is greater than 1.89. The correspondence between the numerical simulation and the analytical solution is further improved if the self-contact energy contribution is reduced in value (data not shown).

Dependence of Supercoiled Configuration on Linking Number Difference. Because supercoiled configurations are expected to depend critically on the linking number difference, the numerical procedure has been extended to chains of ΔLk greater than those that can be studied by exact analytical methods. The preferred geometries of long interwound DNA's were treated only approximately by Tsuru and Wadati.[26] The simulated annealing procedure was accordingly repeated on the 14-point starting structure detailed above for values of ΔLk ranging from 0 to 5. The trajectories were relaxed with an accelerative cooling scheme involving 2000–4000 moves per temperature step. Other parameters were held at the values employed previously. The initial and final curves of the various simulations are illustrated in Figure 5. The starting structure in Figure 5a, a deformed figure-8, is completely arbitrary. As evident from Figures 5b–g, the simulated supercoils are found to relax to nearly perfect circles when ΔLk = 0–1 but to form various interwound helices for ΔLk ≥ 2. At ΔLk = 2, the minimized config-

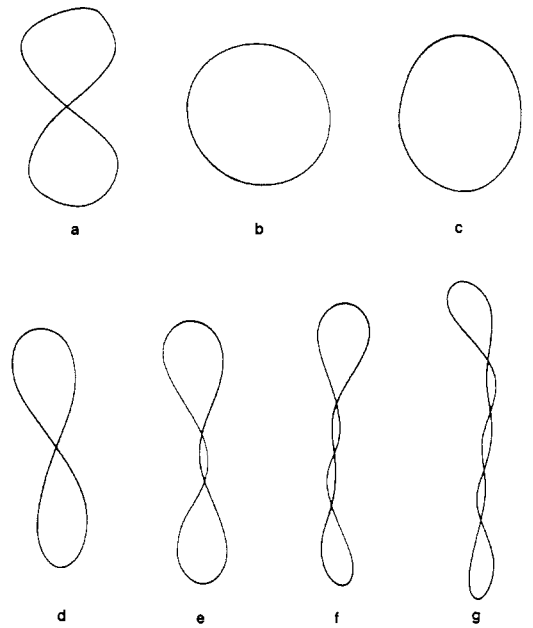

Figure 5. Dependence of optimized supercoiled configuration on linking number difference. The starting trajectory (a) is defined by the 14 controlling points specified in the legend to Figure 3. Optimized configurations obtained with (b) $\Delta Lk = 0$, (c) $\Delta Lk = 1$, (d) $\Delta Lk = 2$, (e) $\Delta Lk = 3$, (f) $\Delta Lk = 4$, (g) $\Delta Lk = 5$ and energy parameters listed in the legend to Figure 1.

uration is a figure-8, at $\Delta Lk = 3$ an interwound form with one helical turn, and at $\Delta Lk = 4$ and 5 interwound helices with two and three helical turns, respectively. It is very clear that as ΔLk increases, the energetically favored configurations are superhelices with additional helical turns. Extension of the 14-point B-spline model, however, to curves of greater ΔLk is not possible. More complicated interwound trajectories cannot be approximated with such a small number of controlling points.

Two points can be made from the energy-minimized curves in Figure 5. First, the final configurations are all dramatically different from the single starting structure. It is therefore unlikely that local minima have trapped the algorithm and prevented finding the global minimum at each ΔLk. Second, the critical influence of ΔLk on the supercoiled configurations is manifested through the various optimized supercoiled structures. When ΔLk is less than or equal to unity, the most stable configurations are planar circles with near zero writhe. This conclusion has been reached by others.[25,26] When $\Delta Lk \geq 2$, however, the most stable configurations are exclusively interwound supercoils. This conclusion has not been directly reached in past studies. The experimentally observed supercoiling of closed circular DNA with linking deviations[4] is correctly predicted by the present approach.

Table I details the computational conditions of these simulations and some properties of the final configurations. The first row specifies ΔLk, while the second row includes the percentage p of successful moves out of the total trials at the starting temperature. The latter values are related to the temperatures at which the simulations start. The third row lists the numbers of total trial moves required to reach an equilibrium at each temperature, and the fourth row, the initial cooling factors f_0 at the starting temperature. These parameters have been adjusted for different configurations. Empirically, a larger number of moves and a slower cooling rate are required for more complicated configurations to reach an equilibrium distribution at each temperature and to maintain an equilibrium along the cooling trajectory. The simple final

Table I
Computational Parameters for Optimized Supercoiled Configurations[a]

Figure-8 Starting State (Figure 5a)						
ΔLk	0	1	2	3	4	5
p	36	32	29	38	37	42
moves	2000	2000	2000	3000	4000	4000
f_0	0.93	0.93	0.93	0.95	0.95	0.95
steps	6	6	7	8	9	10
Wr	−0.01	0.00	1.04	1.84	2.91	3.33
$\tilde{E}$	1.59	2.59	5.38	7.54	9.55	12.24

Doubly Interwound Starting State (Figure 6a)						
ΔLk			3	4	5	6	6'[b]
p			35	55	55	55	28
moves			3000	5000	5000	5000	5000
f_0			0.95	0.95	0.95	0.95	0.95
steps			11	12	13	13	7
Wr			1.66	2.84	3.61	4.02	4.57
$\tilde{E}$			7.98	8.66	11.2	14.5	13.5

[a] ΔLk is the fixed linking deviation of each configuration, p the percentage of successful moves out of the total trial moves, moves the number of total trial moves at each temperature, f_0 the cooling rate at each decrease of temperature, steps the number of temperature lowerings until the freezing temperature is reached, Wr the writhing number, and $\tilde{E}$ the reduced energy of the corresponding configurations. [b] Configuration 6' was generated starting from an intermediate configuration during the optimization of configuration 6. The intermediate configuration was first heated above the existing temperature and then simulated with the normal annealing procedure. 6' is therefore the consequence of a variable cooling schedule.

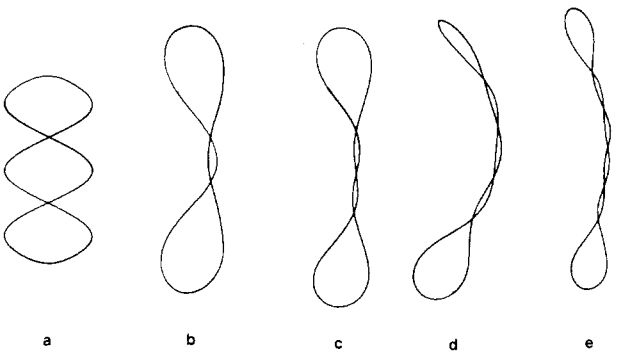

Figure 6. Series of supercoiled configurations obtained upon the simulated annealing of the 24-point starting trajectory defined by the following controlling points: (2.1, −2.1, 1.5), (0.0, −3.0, 3.0), (−2.1, −2.1, 4.5), (−3.0, 0.0, 6.0), (−2.1, 2.1, 7.5), (0.0, 3.0, 9.0), (2.1, 2.1, 10.5), (3.0, 0.0, 12.0), (2.1, −2.1, 13.5), (0.0, −3.0, 15.0), (−2.1, −2.1, 16.5), (0.0, 0.0, 18.0), (2.1, 2.1, 16.5), (0.0, 3.0, 15.0), (−2.1, 2.1, 13.5), (−3.0, 0.0, 12.0), (−2.1, −2.1, 10.5), (0.0, −3.0, 9.0), (2.1, −2.1, 7.5), (3.0, 0.0, 6.0), (2.1, 2.1, 4.5), (0.0, 3.0, 3.0), (−2.1, 2.1, 1.5), (0.0, 0.0, 0.0). Optimized configurations obtained with (b) $\Delta Lk = 3$, (c) $\Delta Lk = 4$, (d) $\Delta Lk = 5$, (e) $\Delta Lk = 6$ and energy parameters listed in the legend to Figure 1.

configurations for $\Delta Lk \leq 2$ are not appreciatively changed when a greater number of moves are made and a slower cooling rate is used in the simulations. The more complicated configurations for $\Delta Lk \geq 2$, however, will be quite different if these chains are cooled more quickly. The fifth row of Table I specifies the number of cooling steps needed to reach the freezing temperature of the different configurations. On average, the freezing temperature is about 0.6 times the starting temperature. The last two rows of the table list the writhing numbers and the reduced energies of the different minimized configurations. The writhing number of each configuration is correlated with the imposed linking difference of the supercoil (i.e., a larger Wr is found in the simulated configurations of larger ΔLk).

Figure 6 shows a second series of supercoiled configurations defined by 24 controlling points. The example

illustrates and clarifies the effects of starting structure and the number of controlling points on the simulated annealing procedure. The starting structure is an interwound helix (generated from the coordinates listed in the figure legend). The standard set of energy parameters is employed, and ΔLk is varied at integral increments from 3 to 6. The annealing is achieved with an accelerative procedure involving 3000–5000 moves per cooling step. A comparison of corresponding curves in Figures 5 and 6 reveals that basically identical configurations are obtained for the same ΔLk. These results support the conclusions drawn above regarding the effect of linking number variations on chain trajectory. The similarities further demonstrate that the predictions do not depend on the starting configurations or the number of controlling points for the B-spline curves. Numerical data describing the optimized configurations in Figure 6 are also included in Table I.

Because the present approach is a Monte Carlo numerical procedure, the starting configuration and the number of controlling points (or variables) affect the efficiency of the algorithm. Simulations starting from configurations close to the final state locate the energy minimum more easily than those initiated from more distant states. The latter calculations require much more time in intermediate searching stages. As pointed out above, B-spline curves more closely match actual spatial trajectories when additional controlling points are used. B-spline curves defined by more controlling points should therefore provide a better representation of real DNA supercoils. A particular supercoiled trajectory, however, also requires a minimum number of controlling points. For example, the highly interwound structures with ΔLk = 5 in Figures 5g and 6d entail three full helical turns plus two end loops. It is more difficult to approximate the trajectory in Figure 5g with only 14 controlling points. The trajectory in Figure 6d defined by 24 points appears smoother and, indeed, has a lower reduced energy ($\tilde{E}$ = 11.2 versus 12.2) than that in Figure 5g. On the other hand, the number of controlling points affects the efficiency of the simulated annealing algorithm and therefore the computed energy minima. It is more difficult for the Monte Carlo simulation to reach an equilibrium distribution at a given temperature with a greater number of controlling points. The configuration must be annealed more slowly to maintain an equilibrium. It is also observed that with 24 controlling points some of the interwound helices in Figure 5 are slightly bent. This phenomenon is thought to be due to the incomplete optimization of these configurations. There is an upper practical limit on the number of controlling points that can be treated within reasonable computational time.

Factors That Affect the Supercoiled Configuration.

1. Repulsive Radius. Besides the linking number difference, the optimized configuration of a supercoil is affected by a number of other factors. One of these is the self-interaction of the chain as characterized by the repulsive radius d. This parameter should be determined by the physical interactions of the sugar–phosphate backbone of the DNA double helix. The effect of self-interaction is actually determined by the ratio of d to the chain contour length L. These two quantities are therefore considered together.

Optimized supercoiled configurations derived from the same starting structure with different repulsive radii are compared in Figure 7. In all cases ΔLk is fixed at 4 and the starting structure is the 14-point structure defined in Figures 3 and 5a. The parameter d is set equal to $L_0/90$, $L_0/63$, $L_0/45$, and $L_0/22.5$ with all other energy parameters held at standard values. It is observed that as d increases

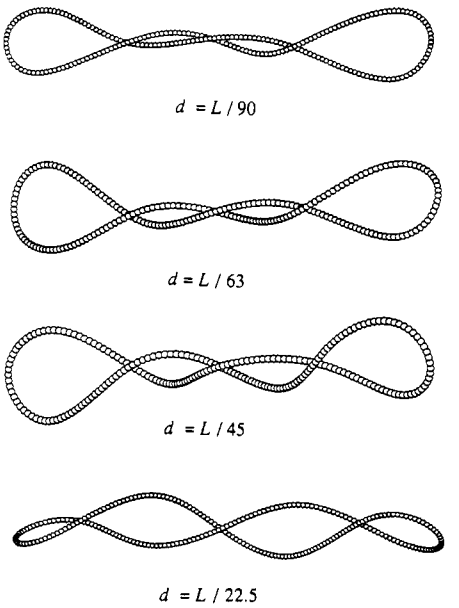

Figure 7. Structural examples illustrating the dependence of the optimized supercoiled configuration with ΔLk = 4 on the ratio of the repulsive radius d to the initial contour length L_0. Remaining energy parameters are held at the values listed in the legend to Figure 1.

and repulsive interactions force the two strands of the interwound structures greater distances apart, the pitch angles of the interwound helices become smaller and their radii larger. The writhing numbers of the optimized configurations also become smaller with increase in d, while the reduced energies become greater. The computed values of Wr for the increasing series of d are respectively 2.98, 2.91, 2.73, and 2.00, while the corresponding values of $\tilde{E}$ are 13.3, 13.7, 15.2, and 18.6. It is noted that the configuration with the most severe repulsive contacts unwinds by approximately one superhelical turn relative to the configurations with lesser self-contacts. The terminal loops of the highest energy structure are also noticeably smaller than those in the lower energy configurations. The limiting case where d = 0, however, is not evaluated because of computational difficulties in determining the writhing number with the Gauss integral when the chain trajectory self intersects.

2. Bending to Twisting Ratio. The ratio of flexural rigidity to twisting stiffness constants A/C is another factor that affects the optimized supercoiled configuration. Optimized configurations resulting from the simulated annealing of the above detailed 14-point starting structure under constraints of varying ΔLk and A/C are presented in Figure 8. Structures of the same ΔLk are found in the same columns, and those of identical A/C in the same rows of the figure. The writhing numbers are also listed below each configuration. When ΔLk = 2 (column 1), the optimized configurations are in the region between figure-8 and circular forms. The choice of A/C is crucial to the configurations. When A/C = 2, the optimized structure is a planar circle, but when A/C = 1.5 or 1.0, the optimized structure is close to a figure-8. When ΔLk = 3 (column 2), however, the supercoiling energy is relieved during the minimization process by increasing the writhing number to values of approximately 2. The resulting structures are characterized by roughly two superhelical turns. In this case the optimized configurations are not greatly affected by the different choices of A/C. When A/C is decreased, the chain trajectory is bent more easily but twisted with more difficulty. The configurational consequences are more noticeable with simple trajectories like the figure-8

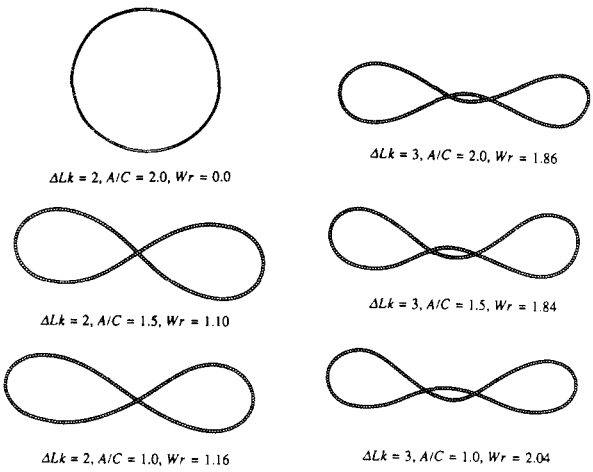

Figure 8. Structural examples illustrating the combined effects of the ratio A/C of the flexural rigidity coefficient to the twisting stiffness constant and ΔLk. The starting configuration is a circle of 14-controlling points. Other energy parameters are held at the values listed in the legend to Figure 1.

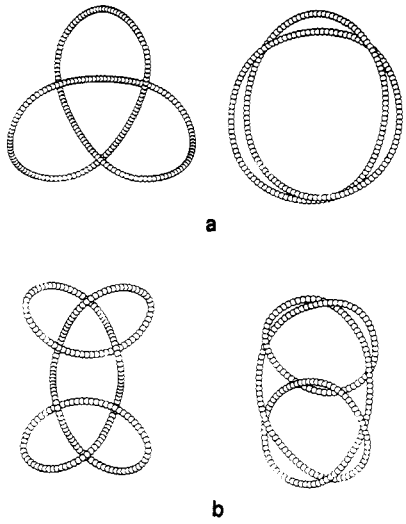

Figure 9. Initial and optimized knotted structures obtained from the simulated annealing. (a) Trefoil knot with ΔLk = 4. (b) Double trefoil knot with ΔLk = 6. Controlling points of the two starting structures are those detailed elsewhere.[33] Other energy parameters are held at the values listed in the legend to Figure 1.

than with more highly wound supercoils. At large ΔLk, the effect of A/C is small, and the minimized supercoiled configurations are all similar interwound forms.

3. Knotting Effects. When the trajectories are knotted, the optimized configurations are found to be completely different from those without knots. The pathways of the supercoil are not able to cross the knotted points during the simulated annealing process. The supercoiling pressure therefore cannot be relieved by the optimal writhing of the trajectory. The effect is illustrated in Figure 9 where two knotted structures are compared before and after the simulated annealing process. The trefoil knot in Figure 9a is described by 12 controlling points and the double trefoil in Figure 9b by 24 points. Configurational energies are determined with the standard parameters listed above. Bending deformations are clearly minimized in the final knotted trajectories, with the low-energy structures much less curved than the starting configurations. The optimized structures are as closely related as possible to planar forms given the self-contact constraints of the knot points. The final configurations of the two knots are also unaffected by small variation of ΔLk. The writhing number of the optimized trefoil knot

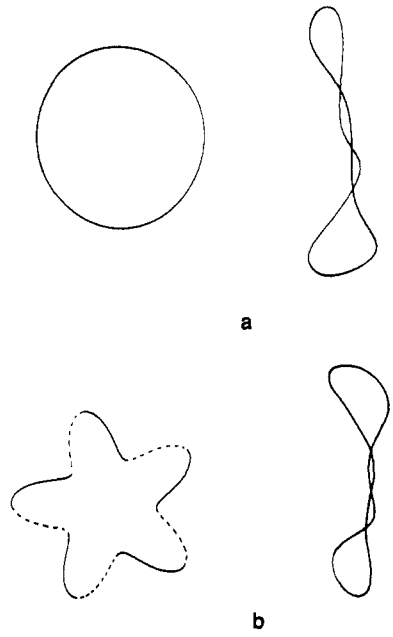

Figure 10. Examples illustrating the relative stability of (a) circular and (b) toroidal configurations with a linking number difference of 4 with respect to the interwound form. The starting circular trajectory is defined, as detailed previously,[33] by 14 evenly spaced controlling points and the starting toroidal trajectory by 30 evenly spaced points.

is found to range between 3.02 and 3.01 for ΔLk between 0 and 4, while that of the minimized double trefoil knot is found to equal 6.05 when ΔLk = 4 and 6.11 when ΔLk = 6. As expected from eq 5, the reduced energy of the trefoil knot is at its lowest (6.2) when ΔLk = 3 and that of the double trefoil knot is minimized (15.7) when ΔLk = 6.

Interwound Configurations versus Circular and Toroidal Configurations. The circular configuration is a local minimum at small ΔLk. When the linking number difference is greater than a critical value, however, the circular form is no longer the global minimum. Indeed, when the simulated annealing process is initiated from a circular form with high ΔLk, the configuration is converted into a more stable interwound form. The configuration resulting from the simulated annealing of a 14-controlling point circle with ΔLk = 4 is illustrated in Figure 10a. After 10 temperature steps (of 4000 Monte Carlo moves/step), Wr is increased to 2.54 and $\tilde{E}$ is lowered from 17.5 to 11.6. (By comparison, the minimum energy configuration found above from the simulated annealing of a 14-point interwound starting state with ΔLk = 4 and the same energy function is characterized by a writhing number of 2.91 and a reduced energy of 9.6.) The simulation starting from a circle is found to require considerable computational effort, involving a higher initial temperature and numerous trial moves, to bring the configuration to the energy minimum. Planar circles are therefore not used as starting configurations in other simulations in this work.

The results of a simulation initiated from a toroidal configuration are presented in Figure 10b. The toroidal form is used to test the stability of this type of supercoil. Configurations of this sort are predicted by some[23] to be the low-energy form of supercoiled DNA in solution and are also said to be consistent with the low-angle X-ray scattering of closed circular duplexes.[46] Supercoiled DNA is found, however, to be interwound in electron micrographs.[24] The starting 5-fold toroidal configuration in the present simulation is defined by 30 controlling points, each helical turn of the closed curve being associated with 6

controlling points. The annealing is carried out in 12 steps, each step involving 7000 Monte Carlo moves. The linking number difference is fixed at 4, and the energy is measured in terms of the standard parameters listed above. Because of the large number of controlling points, the computational time required to reach an equilibrium distribution at each temperature is increased, and the global energy minimum is difficult to locate. During the simulation the configuration is first changed to a roughly circular shape and is later transformed to the final interwound form. The optimized configuration illustrated in Figure 10b is characterized by a writhing number of 2.5 and a reduced energy of 10.4. The starting toroidal configuration is described, in contrast, by a writhing number of 2.4 and a reduced energy of 32.3. The difference in energy between starting and final states is apparently due to the bending deformation energy. The twisting energy contributions are similar in the two forms as evidenced by the nearly identical values of Wr and hence $\Delta Tw = \Delta Lk - Wr$. Several simulations have actually been carried out starting with toroidal structures of different ΔLk. In all cases considered the final structures are found to be interwound. The optimum configuration reported here, however, is still not completely refined. The minimum energy of an interwound $\Delta Lk = 4$ trajectory is known to be lower than the current value (cf. Table I). The simulations could be carried out further to improve the exact configuration, but the results are clearly similar to those found above with interwound starting states and the same energy function.

The above results do not exclude the possibility that the toroidal form may be a mechanical equilibrium state. A mechanical equilibrium, however, does not imply a global energy minimum. Benham suggests that a toroidal configuration may be the entropically favored configuration of supercoiled DNA.[23] He draws his conclusions on the basis of the large number of local minima in toroidal-like forms predicted from the possible solutions to the mechanical equilibrium of a circular rod. On the basis of the present approach which locates the global minimum, we find the global energy configuration of supercoiled DNA to be an interwound rather than a toroidal form. Furthermore, we cannot stabilize the toroidal configuration in the Metropolis–Monte Carlo simulation. We also note that in all mechanical treatments of the supercoiled DNA problem to date, self-contact is not considered. Since self-contacts are quite possible in real DNA, we must consider the results from mechanical treatments as still incomplete.

Chain Length and Supercoiled Configuration. The chain length has two implicit effects on the supercoiled configuration. One is related to the self-interaction of the supercoiled DNA in terms of the ratio d/L as discussed earlier. Another effect, which is more fundamental to the simulated annealing method, is explored below. The flexural rigidity coefficient A has been related by Landau and Lifshitz to the persistence length a of a wormlike chain through the following formula:[47]

$$A = ak_bT \tag{12}$$

where k_b is the Boltzmann constant and T the absolute temperature. The persistence length is a measure of the tendency of a polymer to proceed in an initial direction and is equal to half its Kuhn statistical segment length. In circular DNA molecules, an additional twist component is introduced in the deformation energy. For an isotropic rodlike model, the twisting stiffness constant C is related to A by the Poisson ratio σ, $C = A/(1 + \sigma)$. The deformation energy of supercoiled DNA in eq 5 can then be rewritten as

$$E = \frac{2\pi^2C}{L}\left[\frac{L}{4\pi^2}\frac{A}{C}\int_0^L \kappa(s)^2\,ds + (\Delta Lk - Wr)^2\right] = \frac{2\pi^2C}{L}\tilde{E} \tag{13}$$

where $\tilde{E}$ is a reduced energy independent of the chain length of the molecule.

In the Metropolis–Monte Carlo procedure, the transformation probability w from one configuration to another is determined by their difference in energy (ΔE):

$$w = \exp\left[\frac{-\Delta E}{k_bT}\right] = \exp\left[\frac{-2\pi^2C}{L}\frac{\Delta\tilde{E}}{k_bT}\right] \tag{14}$$

By introducing the number of persistence lengths $x = L/a$, the transformation probability can be rewritten as

$$w = \exp\left[\frac{-2\pi^2\Delta\tilde{E}}{(1 + \sigma)x}\right] \tag{15}$$

In this formalism, there is no explicit temperature dependence. Instead, the factor x introduces a chain length measured in persistence lengths. An alternative interpretation of the simulated annealing procedure is suggested by this expression. That is, the simulation is started at a sufficiently long chain length, and the length is decreased at each stage of increase of β. It is clear that longer chains are more flexible and can be easily relaxed. By gradually reducing the chain length, the supercoiled configuration becomes more and more rigid and is eventually confined to the global energy minimum. It should be noted that in the earlier discussion of simulated annealing in terms of temperature, the reduced energy is independent of the model chain length. The chain length is arbitrarily chosen with no consideration of the physical length of the supercoiled molecule under investigation. The number of persistence lengths x defined in eq 15, however, is directly related to the statistical behavior of the supercoil that is studied.

It is interesting to compare the physical implications of simulated annealing in terms of the temperature and chain-length dependencies. When adopting the temperature interpretation, one is viewing the minimization as a search procedure on a multidimensional energy surface. To locate the global energy minimum, the potential surface must be searched and local minima overcome. For this purpose, the simulation is usually started at a high temperature to overcome the barriers of local energy minima, especially if the starting point is far away from the global minimum. Only when sufficiently close to the final point is a low temperature helpful in accelerating the speed of locating the global minimum. On the other hand, when the length interpretation is adopted, the supercoiled configuration itself is directly viewed and the problem is treated by statistical methods. Since a longer chain is more flexible and can be easily relaxed, the longer chain is used to obtain the information on the gross features of the equilibrium configuration. This gross configuration is then gradually refined by reducing the length of the chain molecule around its equilibrium distribution. As the length of the supercoil becomes shorter and shorter, the molecule becomes less and less flexible and therefore more and more confined to its global minimum energy configuration. Eventually, the length of the supercoil reaches a lower limit, and the configuration becomes virtually rigid, arriving at the unique global minimum. As such, either of the two interpretations offers a possible procedure to locate the global minimum.

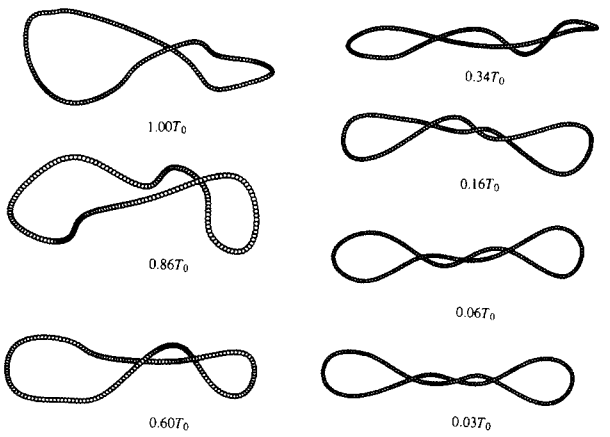

Figure 11. Variation in the equilibrium configuration of a supercoiled chain with decreasing temperature. Representative configurations with $\Delta Lk = 4$ are illustrated for a cooling process starting at T_0 and terminating at the freezing temperature $0.03T_0$. The variation of the supercoiled configuration at different chain lengths, as measured by the number of persistence lengths ($x \propto T$), is also shown by this example. Refer to text for additional details.

The relationship between chain length and supercoiled configuration is illustrated by the following example. Supercoiled configurations with $\Delta Lk = 4$ have been relaxed from the 14-controlling point starting trajectory introduced above. The accelerative cooling procedure is employed (with $f_0 = 0.95$) over several temperature steps. To obtain a reliable equilibrium sampling, the configuration is relaxed at each temperature with 12 000 Metropolis–Monte Carlo runs. Representative equilibrium configurations at different temperatures after the Monte Carlo runs are illustrated in Figure 11. Note that these are individual configurations and are subject to fluctuations during the simulations. According to the above discussion, each temperature corresponds to a specific chain length measured as the number of persistence lengths at a common temperature. At the starting temperature T_0 where $\tilde{E}_0 = 16.2$, the chain length of the supercoil is roughly 21 persistence lengths, the value determined from the relationship $x = 2\pi^2/(1 + \sigma) \times 0.1\tilde{E}_0$. The chain lengths of the remaining configurations are then calculated with the same reduction factor used for the decrease in temperature. For example, the chain length of the last configuration is $21 \times 0.03 = 0.63$ persistence lengths since the final temperature factor is 0.03. At 0.63 persistence lengths, the configuration is already rigid or frozen. Such a result is, of course, the limiting case. The significance of these results is to indicate a limiting situation. When the chain length is slower than the limiting value, the supercoiled configurations are flexible to different degrees. In the present example, when the chain length is longer than 12.6 persistence lengths (i.e., $T = 0.86T_0$), the supercoils are only partially folded. Only when the chain length is reduced to between 7.1 and 3.4 persistence lengths ($T = (0.34–0.16)T_0$) is the supercoil found to fold into a roughly interwound helix with about three helical turns. This configuration is gradually converted to the regular interwound helical form as the chain length is further decreased and then increased. Finally, it is noted that, theoretically, longer chains should have the same configuration at the global energy minimum (or the ground state) as shorter chains.

Conclusions

The present study generates the supercoiled configuration of an elastic chain molecule corrected for excluded volume at its global energy minimum. The potential form adopted here is more complete than those in earlier treatments. The chain self-interaction is explicitly treated with a model parameter, which can be adjusted to fit theoretical results to experimental observations. The potential energy of the supercoiled DNA is directly minimized by the combined B-spline/simulated annealing method developed in this work, and optimized configurations are directly obtained. It is shown that under a given linking number difference, the most probable configurations are interwound superhelices. The number of superhelical turns is proportional to the imposed linking number difference. Moreover, it is explicitly demonstrated that the toroidal configuration is unstable and is transformed into an interwound form upon minimization of the energy. The results obtained here are consistent with the macroscopic example of a twisted and closed elastic rod and the electron microscopic patterns of supercoiled DNA. Such a direct result has not been obtained in previous studies.

A major approximation introduced in the treatment is the use of a finite number of controlling points to define a supercoiled trajectory. A space curve can be exactly represented by B-spline curves only with an infinite number of controlling points. It is shown that within a certain region of configuration space the number of controlling points does not greatly affect the global minimum predicted by this method. With a finite number of controlling points, the actual configurations of supercoiled DNA are smoothed, and consequently the global features of the supercoiled configurations are more easily observed. However, to reproduce the actual fluctuations of the supercoiled configurations, the number of controlling points must be sufficiently large.

Acknowledgment. We are grateful to Professor Irwin Tobias for numerous helpful discussions and critical comments. This work was supported by the U.S. Public Health Service under Grant GM34809. Calculations were performed at the Rutgers Center for Computational Chemistry.

References and Notes

(1) Voordouw, G.; Kam, Z.; Borochov, N.; Eisenberg, H. *Biophys. Chem.* **1978**, *8*, 171.
(2) Weil, R.; Vinograd, J. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *50*, 730.
(3) Vinograd, J.; Lebowitz, J.; Radloff, R.; Watson, R.; Laipis, P. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *53*, 1104.
(4) Bauer, W.; Vinograd, J. *Basic Principles in Nucleic Acid Chemistry*; Academic Press: New York, 1974; Vol. 2, Chapter 4.
(5) Kratky, O.; Porod, G. *Recl. Trav. Chim. Pays-Bas* **1949**, *68*, 1106.
(6) Gray, H. B., Jr.; Bloomfield, V. A.; Hearst, J. E. *J. Chem. Phys.* **1967**, *46*, 1493.
(7) Fujii, M.; Yamakawa, H. *Macromolecules* **1975**, *8*, 792.
(8) White, J. H. *Am. J. Math.* **1969**, *91*, 693.
(9) Fuller, F. B. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 815.
(10) Crick, F. H. C. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2639.
(11) Anderson, P.; Bauer, W. *Biochemistry* **1978**, *17*, 594.
(12) Singleton, C. K. *J. Biol. Chem.* **1983**, *258*, 7661.
(13) Brady, G. W.; Satkowski, M.; Foos, D.; Benham, C. J. In *Biomolecular Stereodynamics IV*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: Guilderland, NY, 1986; p 63.
(14) Depew, R. E.; Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 4275.
(15) Bauer, W. *Annu. Rev. Biophys. Bioeng.* **1978**, *7*, 287.
(16) Benham, C. J. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2397.
(17) Benham, C. J. *Biopolymers* **1978**, *18*, 609.
(18) Le Bret, M. *Biopolymers* **1979**, *18*, 1709.
(19) Tobias, I., unpublished work.
(20) Love, A. E. H. *Treatise on the Mathematical Theory of Elasticity*, 4th ed.; Dover: New York, 1944; Chapters 18, 19.
(21) Landau, L. D.; Lifshitz, E. M. *Theory of Elasticity*; Pergamon Press: Oxford, 1975; Chapter 2.

(22) Le Bret, M. *Biopolymers* **1984**, *23*, 1835.
(23) Benham, C. J. *Biopolymers* **1983**, *22*, 2477.
(24) Wang, J. C. *Trends Biochem. Sci. (Pers. Ed.)* **1980**, *5*, 219.
(25) Tanaka, F.; Takahashi, H. *J. Chem. Phys.* **1986**, *83*, 6017.
(26) Tsuru, H.; Wadati, M. *Biopolymers* **1986**, *25*, 2083.
(27) Shimada, J.; Yamakawa, H. *Macromolecules* **1984**, *17*, 689.
(28) Barkley, M. D.; Zimm, B. H. *J. Chem. Phys.* **1979**, *70*, 2991.
(29) Pohl, W. *J. Math. Mech.* **1968**, *27*, 975.
(30) Pohl, W. *Math. Intelligencer* **1980**, *3*, 20.
(31) Braun, W. *J. Mol. Biol.* **1983**, *163*, 613.
(32) Katkovnik, V. Ya.; Kulchitsky, Yu. *Autom. Remote Control (Engl. Transl.)* **1972**, *8*, 1321.
(33) Hao, M.-H.; Olson, W. K. *Biopolymers* **1989**, *28*, 873.
(34) Gordon, W. J.; Riesenfeld, R. F. In *Computer Aided Geometric Design*; Barnhill, R. E., Riesenfeld, R. F., Eds.; Academic Press: New York, 1974; p 90.
(35) Mortenson, M. E. *Geometric Modeling*; Wiley: New York, 1985; Chapter 2.
(36) Schoenberg, I. J. In *On Numerical Approximation*; Langer, R. E., Ed.; University of Wisconsin Press: Madison, 1959; p 249.
(37) Marsden, M. J. *J. Approx. Theory* **1970**, *3*, 7.
(38) Fuller, F. B. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 3557.
(39) Himmelblau, D. M. *Applied Nonlinear Programming*; McGraw-Hill: New York, 1972.
(40) Pincus, M. *Oper. Res.* **1968**, *16*, 690.
(41) Pincus, M. *Oper. Res.* **1970**, *18*, 690.
(42) Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
(43) Kirkpatrick, S.; Gelatt, C. D., Jr.; Vecchi, M. P. *Science* **1983**, *220*, 671.
(44) Wille, L. T. *Nature* **1986**, *324*, 46.
(45) Bounds, D. G. *Nature* **1987**, *329*, 215.
(46) Brady, G. M.; Fein, D. B.; Lambertson, H.; Grassian, V.; Foos, D.; Benham, C. J. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 741.
(47) Landau, L. D.; Lifshitz, E. M. *Statistical Physics*; Pergamon: London, 1959; Chapter 12.

Characteristic Structural Effects of Cross-Linked Polymer Supports in Solid-Phase Hypochlorite Oxidations

K. Sreekumar and V. N. Rajasekharan Pillai*

Department of Polymer Chemistry, Mahatma Gandhi University, Kottayam, Kerala, India 686 001. *Received October 17, 1988; Revised Manuscript Received January 25, 1989*

ABSTRACT: A series of spacer-modified polystyrene-supported tertiary hypochlorites were prepared starting from 2% divinylbenzene cross-linked polystyrene through multistage analogous polymer reactions. The oxidation efficiencies of these polymer-supported analogues of *tert*-butyl hypochlorites containing spacer arms of varying lengths between the cross-linked matrix and the reactive function were determined. This gives a correlation between the reactivity of the functional groups and the distance of the functional group from the cross-linked macromolecular matrix. The method involved the study of the course of the oxidation reaction of cholesterol to cholestenone by determining (i) the reduction in the capacity of the hypochlorite resin iodometrically, (ii) the uptake of the reactive function for oxidation, and (iii) the amount of the product cholestenone spectrophotometrically. These investigations revealed a sharp increase in reactivity on going from a spacer of one carbon unit to that of two carbon units. Thereafter the reactivity was found to increase gradually with a leveling off at six carbon units. A similar reactivity pattern was also observable in the various analogous polymer reactions employed for the preparation of the polymeric spacer-modified tertiary hypochlorites.

Introduction

In polymer-supported solid-phase organic reactions, it was originally assumed that the polymer matrix acts only as a passive support for binding the active functional groups. But recent studies have revealed that the macromolecular matrix has a profound influence on the nature and course of the reactions. The reactivity of functional groups attached to a polymer network is dependent on the variables of gel preparation.[1] Characteristic features of the polymer support, such as its hydrophobic or hydrophilic nature, linear or cross-linked state, degree of functionalization, and separation of the active function from the matrix, have to be taken into account in the design of an effective reagent. The high local concentration of reagents in a polymer matrix greatly influences the kinetics.[2,3] Here the thermodynamic distribution of a substrate between the support and the bulk reaction medium is a factor that favors interactions between the support phase and the substrate, which in turn enhances the effective concentration of the latter within the support, producing a rate acceleration.[4] The incompatibility of the support and the substrate leads to a diminution of the concentration of the substrate within the support volume. This is very significant with cross-linked insoluble supports.[5]

The topographical nature of the polymer matrix exerts a significant influence on the reactivity of the attached functional groups. The topography of the polymer matrix is determined by the chemical nature of the monomers, the molar percentage of cross-links, and the presence of flexible spacers between the matrix and the reactive function.[6,7]

With a view to understanding the significance of these factors on the reactivity of polymer-bound hypohalite oxidizing reagents,[8] we have prepared a series of spacer-modified polystyrene-supported tertiary hypohalite reagents. In this paper we describe the analogous polymer reactions used for their synthesis and the correlation between the length of the spacer arm and the reactivity of the hypochlorite oxidizing function in these insoluble polymeric reagents.

Results and Discussion

The spacer-modified polymeric reagents were prepared from 2% cross-linked divinylbenzene (DVB)–polystyrene by a seven-step reaction (Scheme I). A ketonic acid function, $CO(CH_2)_{n-1}COOH$, was introduced by Friedel–Crafts reaction of polystyrene (1) with aliphatic dibasic acid anhydrides. The anhydrides used were succinic (n = 3) and glutaric (n = 4). For the preparation of the ketonic acid with n = 5, the diacid chloride was used instead of the anhydride due to the nonavailability of the

* Author to whom all correspondence should be addressed.

0024-9297/89/2222-3303$01.50/0