

The dynamics of chain closure in semiflexible polymers

Arti Dua and Binny J. Cherayil^{a)}

Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore-560012, India

The mean first passage time of cyclization τ of a semiflexible polymer with reactive ends is calculated using the diffusion-reaction formalism of Wilemski and Fixman [J. Chem. Phys. **60**, 866 (1974)]. The approach is based on a Smoluchowski-type equation for the time evolution, in the presence of a sink, of a many-body probability distribution function. In the present calculations, which are an extension of work carried out by Pastor *et al.* [J. Chem. Phys. **105**, 3878 (1996)] on completely flexible Gaussian chains, the polymer is modeled as a continuous curve with a nonzero energy of bending. Inextensibility is enforced on average through chain-end contributions that suppress the excess fluctuations that lead to departures from the Kratky–Porod result for the mean-square end-to-end distance. The sink term in the generalized diffusion equation that describes the dynamics of the chain is modeled as a modified step function along the lines suggested by Pastor *et al.* Detailed calculations of τ as a function of the chain length N , the reaction distance a , and the stiffness parameter z are presented. Among other results, τ is found to be a power law in N , with a z -dependent scaling exponent that ranges between about 2.2–2.4.

I. INTRODUCTION

A number of important genetic processes, including transcription, replication, and recombination, are believed in some way to be under the structural control of the hair-pin loops in DNA and RNA that are produced by base-pairing within single-stranded polynucleotide segments.^{1–4} Loop formation itself is controlled by such factors as solvent quality, chain rigidity and degree of polymerization.⁵ These factors, therefore, indirectly influence the course of gene expression and regulation. Their effects on single chain dynamics have recently been studied in the experiments of Libchaber and co-workers⁶ on the thermodynamics and kinetics of chain closure in synthetic single stranded DNA (ssDNA). In these experiments, single-stranded poly(deoxyadenosine) or poly(deoxythymidine) of between 8 and 30 repeat units are designed with terminal 5-base complementary sequences having a fluorophore and a quencher at either end. In aqueous solution, the DNA strands cyclize whenever the ends are close enough to allow base-pairing. The average time of chain closure—a measure of the ease of loop formation—can be determined from the extent of fluorescence quenching, which varies as a function of the distance between the fluorophore and the quencher.

The results are sensitive to details of backbone rigidity, and cannot be rationalized solely by treating the DNA molecule as a Gaussian coil. Such a model is expected to apply only to chains whose overall length is many times larger than the average persistence length, a condition unlikely to hold for relatively short polynucleotide strands, which are intrinsically semiflexible. A theory of chain closure in these systems must therefore account explicitly for the effects of chain stiffness to be generally valid. We are, therefore, mo-

tivated to develop a model of chain dynamics that can be used to determine cyclization times of semiflexible polymers. We do this following essentially the same methodology as Pastor *et al.*,⁷ who adapted the many-body diffusion–reaction formalism of Wilemski and Fixman^{8–10} to treat flexible polymers; our approach is different only in that the polymer is now treated as an essentially inextensible continuous curve of variable persistence length l . The small l limit of this model describes the flexible Gaussian chain.

As discussed in Sec. II below, the dynamics of chains with reactive, loop-forming end-groups can be described by a Smoluchowski equation to which a sink term, representing the loss of probability from chain closure, has been added. Within this formalism, Wilemski and Fixman have shown⁹ that the closure time τ can be expressed as the integral of an equilibrium time correlation function of the sink term, which in turn can be related to the time correlation function of the end-to-end distance. Section III discusses the evaluation of this latter correlation function, following the introduction of a continuum model for a semiflexible polymer. The calculation of τ is then shown to reduce to the numerical evaluation of a single integral involving a sum over certain normal modes. Section IV presents the principal results of this calculation. Derivations of some of the key expressions used in the text are included in the Appendices.

II. TIME EVOLUTION OF CHAIN CONFORMATIONS

The rate of cyclization of a chain with reactive ends in a neutral solvent (assumed to be at the theta temperature so that excluded volume interactions can be neglected) is governed by the time evolution of the chain conformations as a whole. Since chain closure takes place on relatively long time scales (on the order of microseconds to milliseconds for the DNA samples of Ref. 6, for instance), the dynamics of

^{a)}Electronic mail: cherayil@ipc.iisc.ernet.in

the polymer—in the absence of reactions between its ends—is effectively overdamped, and can be adequately described by the Smoluchowski equation. If $\psi(\{\mathbf{r}\}, t)$ denotes the probability density that the chain has the conformation $\{\mathbf{r}\} \equiv \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$ at time t , where \mathbf{r}_i is the position of the i th monomer in a chain of n monomers, then neglecting hydrodynamic interactions, this equation takes the form

$$\begin{aligned} \frac{\partial \psi(\{\mathbf{r}\}, t)}{\partial t} &= D_0 \sum_{i=1}^n \frac{\partial}{\partial \mathbf{r}_i} \cdot \left[\frac{\partial}{\partial \mathbf{r}_i} + \frac{\partial U}{\partial \mathbf{r}_i} \right] \psi(\{\mathbf{r}\}, t) \\ &\equiv \mathcal{D} \psi(\{\mathbf{r}\}, t), \end{aligned} \quad (1)$$

where D_0 is the diffusion constant, defined as the inverse of the friction coefficient per unit length of the polymer, and U is the potential energy of the chain. The second line of this equation provides the definition of the generalized diffusion operator \mathcal{D} .

This equation must be modified when the ends of the chain are reactive. If the reaction (in this case linkage of the ends of the chain) is assumed to occur instantaneously and irreversibly whenever the ends come within a distance a of each other, the overall rate of reaction will be governed by the rate of approach of the end groups. This process is still diffusion controlled, so it can continue to be described by a Smoluchowski-type equation, as in Eq. (1) above. However, Eq. (1) is conservative, implying that the total time derivative of $\psi(\{\mathbf{r}\}, t)$ is zero. This is no longer the case when reactions at the ends of the chain permanently eliminate closed loops from the distribution of chain conformations. To account for this loss of probability over time, $\psi(\{\mathbf{r}\}, t)$ can either be made to satisfy appropriate boundary conditions,¹¹ or an additional “sink” term can be introduced into the diffusion equation. The latter approach, which is somewhat simpler to implement, is adopted here, and it leads to the following equation:

$$\frac{\partial \psi(\{\mathbf{r}\}, t)}{\partial t} = \mathcal{D} \psi(\{\mathbf{r}\}, t) - k S(\{\mathbf{r}\}) \psi(\{\mathbf{r}\}, t), \quad (2)$$

where k is a rate constant, and S is the sink function, the choice of which prescribes the minimum contact distance of two ends of the chain when reaction takes place.

The chain dynamics described by Eq. (2) is not necessarily equivalent to the experimental situation that obtains in Ref. 6, where cyclization is generally reversible, and measurements are made under steady-state conditions. The treatment of steady-state chain closure will be discussed elsewhere.

To derive an expression for the closure time τ , we introduce the probability $\chi(t)$ that the chain is unreacted at time t ; this probability is given by

$$\chi(t) = \int d\{\mathbf{r}\} \psi(\{\mathbf{r}\}, t). \quad (3)$$

By differentiating Eq. (3) with respect to time, and making use of Eq. (2) along with the result $\int d\{\mathbf{r}\} \mathcal{D} \psi = 0$ (which holds because of the vanishing of ψ at the boundaries), one finds that

$$\frac{d\chi(t)}{dt} = -k\mu(t), \quad (4)$$

where

$$\mu(t) = \int d\{\mathbf{r}\} S(\{\mathbf{r}\}) \psi(\{\mathbf{r}\}, t). \quad (5)$$

If $\chi(t)$ is now assumed to decay as a single exponential, such that

$$\chi(t) \approx e^{-t/\tau}, \quad (6)$$

then the closure time is given by

$$\tau = \int_0^\infty dt \chi(t). \quad (7)$$

Equation (7) can be re-expressed as an integral over an equilibrium time correlation function involving the sink function, as shown in Ref. 7, using results from Ref. 9. For completeness, a sketch of the derivation is provided in Appendix A. In this section, we only state the result

$$\tau = \int_0^\infty dt \left(\frac{C(t)}{C(\infty)} - 1 \right), \quad (8)$$

where

$$C(t) = \int d\mathbf{R} \int d\mathbf{R}_0 S(R) G(\mathbf{R}, t | \mathbf{R}_0, 0) S(R_0) \psi_{\text{eq}}(\mathbf{R}_0). \quad (9)$$

Here $G(\mathbf{R}, t | \mathbf{R}_0, 0)$ is the conditional probability that a chain with the end-to-distance \mathbf{R}_0 at time $t=0$ has the end-to-end distance \mathbf{R} at time t ; $\psi_{\text{eq}}(\mathbf{R}_0)$ is the equilibrium distribution of the end-to-end distance; and $S(R)$ is a sink function, which is assumed to depend only the magnitude of the separation between the ends of the chain. An explicit expression for this sink function will be provided later. As shown in Appendix A, the derivation of this equation assumes the limit $k \rightarrow \infty$, which in physical terms means that chain ends are assumed to react immediately and irreversibly when they satisfy the distance constraints imposed by the sink function. The closure time is therefore independent of k , and is actually a first passage time.

The key unknown in Eq. (9) is the conditional probability $G(\mathbf{R}, t | \mathbf{R}_0, 0)$. For a completely flexible polymer, this probability is Gaussian, because \mathbf{R} is the sum of a large number of independent, random bond vectors, and the central limit theorem applies.⁷ For a semiflexible polymer, on the other hand, successive bond vectors are no longer uncorrelated, but \mathbf{R} can still be written as the sum of a large number of independent normal modes,¹² so in this case too, for moderate to relatively high degrees of stiffness, it is reasonable to describe $G(\mathbf{R}, t | \mathbf{R}_0, 0)$ by a Gaussian.¹³ (Without this assumption the resulting mathematics would become quite intractable.)

A Gaussian distribution is completely specified by its mean μ_m and variance σ^2 ; for a polymer, these take the form

$$\mu_m = \langle \mathbf{R}(t) \rangle_{\text{eq}} = \phi(t) \mathbf{R}_0, \quad (10)$$

$$\sigma^2 = \langle \mathbf{R}^2(t) \rangle_{\text{eq}} - \langle \mathbf{R}(t) \rangle_{\text{eq}} \cdot \langle \mathbf{R}(t) \rangle_{\text{eq}}, \quad (11)$$

where the angle brackets $\langle(\dots)\rangle_{\text{eq}}$ denote an ensemble average over ψ_{eq} . The function $\phi(t)$ is defined as

$$\phi(t) = \frac{\langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle_{\text{eq}}}{\langle R^2 \rangle_{\text{eq}}}. \quad (12)$$

From these definitions one can show that

$$G(\mathbf{R}, t | \mathbf{R}_0, 0) = \left(\frac{3}{2\pi \langle R^2 \rangle_{\text{eq}}} \right)^{3/2} \frac{1}{(1 - \phi^2(t))^{3/2}} \times \exp \left[- \frac{3(\mathbf{R} - \phi(t)\mathbf{R}_0)^2}{2\langle R^2 \rangle_{\text{eq}}(1 - \phi^2(t))} \right]. \quad (13)$$

For a semiflexible polymer $\langle R^2 \rangle_{\text{eq}}$ is given by¹⁴

$$\langle R^2 \rangle_{\text{eq}} = \frac{N}{p} \left[1 - \frac{1}{2pN} (1 - e^{-2pN}) \right], \quad (14)$$

where $N \equiv nl$ is the contour length of the chain, and $p \equiv 1/l$ is the inverse persistence length.

Substituting Eq. (13) into the expression for the correlation function $C(t)$ [Eq. (9)], and carrying out the angular part of the integrations over the vectors \mathbf{R} and \mathbf{R}_0 , one can show that

$$C(t) = 16\pi^2 \left(\frac{3}{2\pi \langle R^2 \rangle_{\text{eq}}} \right)^3 \frac{1}{(1 - \phi^2(t))^{3/2}} \times \int_0^\infty dR R^2 S(R) \int_0^\infty dR_0 R_0^2 S(R_0) \times \exp \left[- \frac{3}{2\langle R^2 \rangle_{\text{eq}}} \frac{(R^2 + R_0^2)}{(1 - \phi^2(t))} \right] \times \frac{\sinh[3\phi R R_0 / \langle R^2 \rangle_{\text{eq}} (1 - \phi^2(t))]}{3\phi R R_0 / \langle R^2 \rangle_{\text{eq}} (1 - \phi^2(t))}. \quad (15)$$

The further evaluation of $C(t)$ requires a specific choice for the sink function. Here we restrict our attention to the Heaviside sink, defined as

$$S(R) = 1; \quad R \leq a \\ = 0; \quad R > a, \quad (16)$$

where a is the closest the two ends of the chain may approach before irreversible ring formation takes place. Although this choice does not permit the analytic evaluation of the integrals in Eq. (15), it is possible to obtain what is expected to be an accurate approximate expression for $C(t)$ by expanding the integrand to first order in the small parameter

$$x_0 = \frac{3a^2}{2\langle R^2 \rangle_{\text{eq}}}, \quad (17)$$

and then integrating term by term. This produces

$$C(t) = \frac{16x_0^3}{9\pi(1 - \phi^2)^{3/2}} \left[1 - \frac{6x_0}{5(1 - \phi^2)} + \dots \right]. \quad (18)$$

To the same order of approximation, this expression may be rewritten in the more convenient resummed form as

$$C(t) \cong \frac{16x_0^3}{9\pi} \left(1 - \phi^2 + \frac{4}{5}x_0 \right)^{-3/2}, \quad (19)$$

which is of exactly the same functional form as the expression derived by Doi¹⁵ using the following Gaussian sink function:

$$S(R) = \mathcal{N} \exp \left(- \frac{3R^2}{2a^2} \right), \quad (20)$$

where \mathcal{N} is a normalization constant. The two expressions are not identical, however; the coefficient of x_0 in Ref. 15 is 4/3, not 4/5, as it is above. In our calculations, $C(t)$ is determined exclusively from the resummed form of the Heaviside sink, Eq. (19).

To calculate τ it now only remains to specify the form of the correlation function $\phi(t)$. This is discussed in the following section.

III. TIME CORRELATION FUNCTION OF THE END-TO-END DISTANCE

The semiflexible polymer is modelled as a continuous curve with a nonzero energy of bending. Points on this curve, at time t , are described by the Cartesian vector $\mathbf{r}(s, t)$, where s is a real variable that defines where a given point (monomer) is on the curve with respect to an origin located at its mid-point; for a chain of contour length N , s lies between $-N/2$ and $N/2$. In terms of the monomer coordinates $\mathbf{r}(s, t)$, and in units where the thermal energy $k_B T = 1$, the configurational part of the Hamiltonian of this model of a semiflexible chain can be written as¹⁶⁻¹⁹

$$H = \nu \int_{-N/2}^{N/2} ds \left| \frac{\partial \mathbf{r}(s, t)}{\partial s} \right|^2 + \eta \int_{-N/2}^{N/2} ds \left| \frac{\partial^2 \mathbf{r}(s, t)}{\partial s^2} \right|^2 + \nu_0 (|\mathbf{u}_{-N/2}|^2 + |\mathbf{u}_N|^2), \quad (21)$$

where

$$\nu = \frac{3p}{2}, \quad \eta = \frac{3}{8p}, \quad \nu_0 = \frac{3}{4} \quad (22)$$

and

$$\mathbf{u}(s, t) \equiv \frac{\partial \mathbf{r}(s, t)}{\partial s}. \quad (23)$$

The first term in Eq. (21) describes the connectivity of the chain, (which is purely entropic in origin), while the second describes its bending energy. By representing the chain in this form, in terms of first and second derivatives of the monomer position, the tangent vector $\mathbf{u}(s)$ [Eq. (23)] should, strictly speaking, be a unit vector, meaning $|\mathbf{u}(s)| = 1$. This constraint is difficult to enforce rigorously, so as a matter of convenience it is replaced by the mathematically less demanding constraint $\langle |\mathbf{u}(s)|^2 \rangle = 1$, where the angular brackets refer to an equilibrium average over the conformations of the chain. This constraint is now handled by the last two terms in Eq. (21), which act effectively to suppress the excess fluctuations in the chain ends that lead to departures from the Kratky-Porod result for the end-to-end distance.¹⁹

In the foregoing continuum representation of the chain, the end-to-end distance $\mathbf{R}(t)$ at time t can be written as

$$\mathbf{R}(t) = \mathbf{r}(N/2, t) - \mathbf{r}(-N/2, t). \quad (24)$$

The equilibrium correlation function of this distance at two different times, therefore, becomes

$$\begin{aligned} \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle &= \langle \mathbf{r}(N/2, t) \cdot \mathbf{r}(N/2, 0) \rangle + \langle \mathbf{r}(-N/2, t) \\ &\quad \cdot \mathbf{r}(-N/2, 0) \rangle - \langle \mathbf{r}(N/2, t) \cdot \mathbf{r}(-N/2, 0) \rangle \\ &\quad - \langle \mathbf{r}(-N/2, t) \cdot \mathbf{r}(N/2, 0) \rangle. \end{aligned} \quad (25)$$

To evaluate the averages that appear on the right-hand side of the above equation, we need an expression for $\mathbf{r}(s, t)$, the position of the s th monomer at time t . This function satisfies a differential equation that can be derived from the application of Hamilton's principle to the chain Lagrangian, $\mathcal{L} = (\frac{1}{2}m) \int_{-N/2}^{N/2} ds \dot{\mathbf{p}}^2 - H$, where $\mathbf{p}(s)$ is the momentum of the monomer at s . This differential equation, as shown for instance in Ref. 17, is given by

$$\left(\frac{\partial}{\partial t} + 2D_0 \eta \frac{\partial^4}{\partial s^4} - 2D_0 \nu \frac{\partial^2}{\partial s^2} \right) \mathbf{r}(s, t) = \boldsymbol{\theta}(s, t), \quad (26)$$

and satisfies the following boundary conditions:

$$\nu \frac{\partial \mathbf{r}(s, t)}{\partial s} - \eta \frac{\partial^3 \mathbf{r}(s, t)}{\partial s^3} \Big|_{s=\pm N/2} = 0, \quad (27)$$

$$\nu_0 \frac{\partial \mathbf{r}(s, t)}{\partial s} + \eta \frac{\partial^2 \mathbf{r}(s, t)}{\partial s^2} \Big|_{s=N/2} = 0, \quad (28)$$

$$\nu_0 \frac{\partial \mathbf{r}(s, t)}{\partial s} - \eta \frac{\partial^2 \mathbf{r}(s, t)}{\partial s^2} \Big|_{s=-N/2} = 0. \quad (29)$$

The variable $\boldsymbol{\theta}(s, t)$ is a random force with white noise statistics whose correlation satisfy

$$\langle \boldsymbol{\theta}(s, t) \boldsymbol{\theta}(s', t') \rangle = \frac{2}{\zeta} \delta(s-s') \delta(t-t') \mathbf{I}, \quad (30)$$

\mathbf{I} being the unit tensor.

Equation (26) is solved by

$$\begin{aligned} \mathbf{r}(s, t) &= \int_{-N/2}^{N/2} ds_1 G_0(s, s_1 | t) \mathbf{r}(s_1) \\ &\quad + \int_{-N/2}^{N/2} ds_1 \int_0^t dt_1 G_0(s, s_1 | t-t_1) \boldsymbol{\theta}(s_1, t_1), \end{aligned} \quad (31)$$

where the Green's function G_0 is the solution to

$$\left(\frac{\partial}{\partial t} + 2D_0 \eta \frac{\partial^4}{\partial s^4} - 2D_0 \nu \frac{\partial^2}{\partial s^2} \right) G_0(s, s_1 | t) = \delta(s-s_1) \delta(t). \quad (32)$$

By direct substitution, the solution to Eq. (32) is found to be

$$G_0(s, s_1 | t) = \theta(t) \sum_{n=0}^{\infty} Q_n(s) Q_n(s_1) e^{-D_0 \lambda_n t}, \quad (33)$$

where $\theta(t)$ is the step function, and $Q_n(s)$, $n=0,1,2,\dots$ are a complete orthonormal set of eigenfunctions that are the solutions to the following eigenvalue equation:

$$L Q_n(s) = \lambda_n Q_n(s), \quad n=0,1,2,\dots, \quad (34)$$

where λ_n are the eigenvalues, and L is the Hermitian operator

$$L = 2\eta \frac{\partial^4}{\partial s^4} - 2\nu \frac{\partial^2}{\partial s^2}. \quad (35)$$

The eigenfunctions $Q_n(s)$ satisfy the same boundary conditions as defined in Eqs. (27)–(29).

Equation (34) is a simple fourth-order differential equation that may be solved by rewriting it as a system of four first-order differential equations in a new set of dependent variables. In matrix notation, this system is of the form $\partial \mathbf{X}(s)/\partial s = \mathbf{M} \cdot \mathbf{X}(s)$. From the solution to this equation, one can show that¹²

$$\begin{aligned} Q_n(s) &= A \cos(\beta_n s) + B \cosh(\alpha_n s) + C \sin(\beta_n s) \\ &\quad + D \sinh(\alpha_n s), \end{aligned} \quad (36)$$

where A , B , C , and D are constants to be determined from the boundary conditions, and α_n and β_n are parameters that are obtained from the characteristic equation of the matrix \mathbf{M} . They are defined as

$$\alpha_n = \left[\frac{\nu}{2\eta} \pm \left[\left(\frac{\nu}{2\eta} \right)^2 + \frac{\lambda_n}{2\eta} \right]^{1/2} \right]^{1/2}, \quad (37)$$

$$i\beta_n = - \left[\frac{\nu}{2\eta} \mp \left[\left(\frac{\nu}{2\eta} \right)^2 + \frac{\lambda_n}{2\eta} \right]^{1/2} \right]^{1/2}. \quad (38)$$

The parameters α_n and β_n are related to each other and to the eigenvalues λ_n by the equations

$$\alpha_n^2 - \beta_n^2 = \frac{\nu}{\eta}, \quad (39)$$

and

$$\lambda_n = 2\eta \alpha_n^2 \beta_n^2. \quad (40)$$

Because $Q_n(-s)$ is also a solution Eq. (34), the linear combinations $Q_n(s) + Q_n(-s)$ and $Q_n(s) - Q_n(-s)$ are solutions as well, so the eigenfunctions may be chosen with definite parity. In general, then, they are given by $Q_n(s) = A \cos(\beta_n s) + B \cosh(\alpha_n s)$ (even parity solutions), and by $Q_n(s) = C \sin(\beta_n s) + D \sinh(\alpha_n s)$ (odd parity solutions). The application of the boundary conditions in Eq. (27), gives the following expression for $Q_n(s)$:

$$\begin{aligned} Q_n(s) &= \left(\frac{C_n}{N} \right)^{1/2} \left[- \frac{\cos(\beta_n s)}{\sin(\beta_n N/2)} \right. \\ &\quad \left. + \left(\frac{\alpha_n}{\beta_n} \right)^2 \frac{\cosh(\alpha_n s)}{\sinh(\alpha_n N/2)} \right], \quad (\text{even parity}), \end{aligned} \quad (41)$$

$$\begin{aligned} Q_n(s) &= \left(\frac{D_n}{N} \right)^{1/2} \left[\frac{\sin(\beta_n s)}{\cos(\beta_n N/2)} \right. \\ &\quad \left. + \left(\frac{\alpha_n}{\beta_n} \right)^2 \frac{\sinh(\alpha_n s)}{\cosh(\alpha_n N/2)} \right], \quad (\text{odd parity}), \end{aligned} \quad (42)$$

where C_n and D_n are unknown, but can be chosen to ensure normalization of the eigenfunctions. The further application of the boundary conditions in Eqs. (28) and (29) leads to a pair of eigenvalue equations that correspond to the even and odd parity solutions; these equations are

$$\begin{aligned} & \beta_n^3 \sinh(\alpha_n N/2) \cos(\beta_n N/2) \\ & + \alpha_n^3 \cosh(\alpha_n N/2) \sin(\beta_n N/2) + 2p(\alpha_n^2 + \beta_n^2) \\ & \times \sin(\beta_n N/2) \sinh(\alpha_n N/2) = 0, \quad (\text{even parity}), \end{aligned} \quad (43)$$

$$\begin{aligned} & \beta_n^3 \cosh(\alpha_n N/2) \sin(\beta_n N/2) \\ & - \alpha_n^3 \sinh(\alpha_n N/2) \cos(\beta_n N/2) - 2p(\alpha_n^2 + \beta_n^2) \\ & \times \cos(\beta_n N/2) \cosh(\alpha_n N/2) = 0, \quad (\text{odd parity}). \end{aligned} \quad (44)$$

From Eqs. (31) and (33), an expression for the time correlation function of the monomer position can now be derived as

$$\begin{aligned} & \langle \mathbf{r}(s, t) \cdot \mathbf{r}(s, 0) \rangle \\ & = \sum_{n=0}^{\infty} \int_{-N/2}^{N/2} ds_1 \mathcal{Q}_n(s) \mathcal{Q}_n(s_1) \langle \mathbf{r}(s_1, 0) \cdot \mathbf{r}(s, 0) \rangle \\ & \times e^{-D_0 \lambda_n t} + \sum_{n=0}^{\infty} \int_{-N/2}^{N/2} ds_1 \int_0^t dt_1 \mathcal{Q}_n(s) \mathcal{Q}_n(s_1) \\ & \times \langle \boldsymbol{\theta}(s_1, t_1) \cdot \mathbf{r}(s, 0) \rangle e^{-D_0 \lambda_n (t-t_1)}. \end{aligned} \quad (45)$$

The second term in this equation vanishes by virtue of the Gaussian character of the random force $\boldsymbol{\theta}(s, t)$, which has a zero mean, so the sole contribution to the monomer time correlation function comes from the first term. The evaluation of this term requires an expression for the equilibrium average $\langle \mathbf{r}(s_1, 0) \cdot \mathbf{r}(s, 0) \rangle$, which is shown in Appendix C to be given by

$$\begin{aligned} & \langle \mathbf{r}(s_1, 0) \cdot \mathbf{r}(s, 0) \rangle \\ & = \frac{\min\{s, s_1\}}{p} - \frac{1}{4p^2} [1 - (e^{-2ps} + e^{-2ps_1} - e^{-2p|s-s_1|})]. \end{aligned} \quad (46)$$

Substituting Eqs. (43)–(46) into Eq. (25) and carrying out the integration s_1 , we obtain the following expression for the time correlation function of the end-to-end distance:

$$\begin{aligned} \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle & = \frac{D_n}{N} \sum_{\substack{n=1 \\ \text{odd}}}^{\infty} \left(\frac{2T_1 e^{-D_0 \lambda_n t}}{p} \right) \\ & \times \left[\frac{2T_2}{\beta_n^2} - \frac{(1 + e^{-2pN})T_3}{\alpha_n^2} + \frac{\beta_n(1 - e^{-2pN})}{2p\alpha_n^2} \right] \\ & \times \left(1 - \frac{\alpha_n^4}{\beta_n^4} \right), \end{aligned} \quad (47)$$

where

$$\begin{aligned} T_1 & = \tan(\beta_n N/2) + \frac{\alpha_n}{\beta_n} \tanh(\alpha_n N/2), \\ T_2 & = \tan(\beta_n N/2) - \frac{\beta_n}{\alpha_n} \tanh(\alpha_n N/2), \\ T_3 & = \tan(\beta_n N/2) - \frac{\alpha_n^3}{\beta_n^3} \tanh(\alpha_n N/2), \end{aligned} \quad (48)$$

with

$$\begin{aligned} \left(\frac{D_n}{N} \right) & = \left[\frac{N}{2} \left(\sec^2(\beta_n N/2) - \frac{\alpha_n^2}{\beta_n^2} \operatorname{sech}^2(\alpha_n N/2) \right) \right. \\ & \left. + \frac{(3\alpha_n^2 - \beta_n^2)T_2}{(\alpha_n^2 + \beta_n^2)\beta_n} \right]^{-1}. \end{aligned} \quad (49)$$

There are no contributions to $\langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle$ from n even.

The explicit expression for $\phi(t)$, as given by Eq. (12), can be found from Eqs. (47)–(49). This is then substituted into Eq. (19) to determine $C(t)$, which along with $C(\infty)$, is used in Eq. (8) to obtain the closure time τ .

The above procedure can only be implemented numerically. It begins by assigning a definite value to the stiffness parameter $z \equiv pN$ by assigning definite values to the contour length N and the inverse persistence length p . Variations in z are effected by varying p at constant N . Having fixed z , Eq. (44) is solved for the parameter α_n , using Eq. (39) to express β_n in terms of α_n , at different fixed values of the mode number n . The solutions are obtained numerically using Mathematica. The n th eigenvalue λ_n is then calculated from Eq. (40). The time correlation function of the end-to-end distance is next calculated from Eq. (47) by first assigning a definite value to the time t (its units are determined by D_0 , which is arbitrarily set to 1, as in Ref. 7), and then summing over modes using the previously determined values of α_n , β_n , and λ_n in conjunction with Eqs. (48) and (49). The sum typically converges within the first 500 modes. The function $\phi(t)$, for the given value of t , is now obtained from Eqs. (47) and (12); $\phi(t)$ in turn is used to determine $C(t)$ from Eq. (19) after assigning a definite value to the reaction distance a (which, like N , is expressed in units of l , the persistence length.) The process is repeated for different fixed values of t . $C(\infty)$ is likewise calculated from Eq. (19) by taking its $t \rightarrow \infty$ limit analytically. Knowing $C(t)$ and $C(\infty)$, the closure time τ is now calculated from Eq. (8) by numerical integration using routines available in Mathematica. The entire calculation can be repeated for different values of z , N , and a .

IV. RESULTS AND DISCUSSION

A. The flexible limit

As a check of the general methodology developed in the foregoing sections, we first study our model in the flexible limit, where prior analytical and simulation results are available for comparison.^{7,11,15,20–24} This limit corresponds, in general, to values of z much larger than unity. In practice, we set z to N , thus fixing the persistence length l at 1, the fundamental bond length. That this assignment does correspond to the flexible chain limit is verified by noting that the eigenvalues λ_n calculated from Eqs. (39), (40), and (44) are nearly identical to the known eigenvalues of the continuum Rouse model. With z set to N , we now estimate τ for different values of N and a . The results are shown in Table I in the column marked DC, along with results from the simulations of Pastor *et al.*⁷ (the column marked Sim.), the results calculated by Pastor *et al.* using a certain approximation in the Wilemski–Fixman model^{7–10} (the column marked WF2) and the results of two other calculations using the Wilemski–Fixman model to be discussed shortly (the columns marked

TABLE I. Closure time τ for different N and a as estimated by different theoretical approaches.

N	a	τ				
		Sim.	WF2 ^a	WF3 ^b	WF4 ^c	DC ^d
50.0	0.1	570±30	677	393	162	233
	0.5	174±10	233	169	143	168
	1.0	110±5	163	130	123	136
75.0	0.1	1070±70	1313	797	365	500
	0.5	410±20	498	381	331	378
	1.0	250±10	367	306	293	317
100.0	0.1	1800±80	2104	1306	644	863
	0.5	680±30	848	673	591	674
	1.0	450±20	646	558	530	578

^aPastor–Zwanzig–Szabo model (PZS), Ref. 7, in the Wilemski–Fixman approximation, with $\phi(t)$ and λ_n calculated from Eqs. (50) and (51), respectively.

^bPZS model with $\phi(t)$ calculated from Eq. (52). The eigenvalues used in the calculation are the discrete Rouse eigenvalues.

^cPZS model with $\phi(t)$ calculated from Eq. (53). The eigenvalues used in the calculation are the continuum Rouse eigenvalues.

^dPresent calculations.

WF3 and WF4, respectively.) The values of N and a used in our calculations were chosen to coincide with the values reported in Ref. 7.

It is clear from the table that there can be significant differences between all five sets of predictions, although they are all consistent in suggesting (a) that τ decreases with increase in a at fixed N , and (b) that τ increases with increase in N at fixed a . These trends are physically reasonable: On the one hand, increasing a at fixed N brings the reaction radius closer to the equilibrium end-to-end separation, making it *easier* on average for the chain to cyclize; and on the other, increasing N at fixed a takes the reaction radius farther away from the end-to-end separation, making it *harder* on average for the chain to cyclize.

The quantitative differences between these predictions are less easily rationalized, however. One reason for the differences may be the function $\phi(t)$, the time correlation function of the end-to-end distance. In Ref. 7, $\phi(t)$ is written as

$$\phi(t) = \frac{8}{N(N+1)} \sum_{\text{odd } n} \left(\frac{1}{\lambda_n} - \frac{1}{4} \right) \exp(-D_0 \lambda_n t), \quad (50)$$

where the eigenvalues λ_n are given by

$$\lambda_n = \frac{3}{b^2} (2 \sin \theta_n/2)^2. \quad (51)$$

Here b is the bond length, θ_n is $n\pi/(N+1)$, and the summation is carried out over a finite number of modes. The factor of $(2 \sin \theta_n/2)^2$ in Eq. (51) coincides with the eigenvalues of the discrete Rouse chain, but the coefficient $3/b^2$ is extraneous, and has been put in by hand. (On dimensional grounds alone the presence of the bond length b in this equation would be suspect, but presumably it is really a dimensionless length of some kind. In any event, it is set to 1 in the calculations, so it effectively drops out of the calculations.) Although the factor of 3 in Pastor *et al.*'s definition of the discrete Rouse eigenvalues should be omitted when these eigenvalues are used in the *denominator* of Eq. (50), this

factor *should* be retained when the eigenvalues are used in the argument of the exponential (provided D_0 is set to 1).

The normalization of $\phi(t)$ in Eq. (50) appears to have been taken care of by the additive factor involving the coefficient of $1/4$. [That is to say, $\phi(0)$, as obtained from Eq. (50), is unity.] But it would have been more natural to normalize $\phi(t)$ by writing it as

$$\phi(t) = \frac{\sum_{\text{odd } n} \lambda_n^{-1} \exp(-3D_0 \lambda_n t)}{\sum_{\text{odd } n} \lambda_n^{-1}}, \quad (52)$$

with $D_0 = 1$ and λ_n given by $(2 \sin \theta_n/2)^2$. The closure times τ predicted by the use of *this* expression for $\phi(t)$ in the Wilemski–Fixman model along with the resummation approximation of Ref. 7 lead to the values shown under column WF3 in Table I. These values are all considerably lower than those in WF2, particularly at $a = 0.1$, indicating a fairly sensitive dependence of τ on $\phi(t)$.

In our calculations, $\phi(t)$ is obtained from the large z limit of Eq. (47) after normalization. It cannot be expressed in closed form, but for a given time t and with $z = N$, it is found to yield the same numerical value as the following expression:

$$\phi(t) = 8 \sum_{\text{odd } n} \frac{1}{n^2 \pi^2} e^{-3n^2 \pi^2 D_0 t / N^2}. \quad (53)$$

Since the calculations use a continuum model of the chain, it may be more appropriate to compare our estimates of τ with the estimates obtained using Eq. (53) in the Wilemski–Fixman model, rather than either Eqs. (50) or (52). Values of τ calculated with Eq. (53) are shown in column WF4 in Table I, and they are seen to be quite comparable to the figures in the column marked DC. (Because of the way $\phi(t)$ enters into the expression for τ [Eqs. (8) and (19)], small numerical differences in its value can be magnified. So although the eigenvalues that generate WF4 and DC are nearly the same, the final estimates for τ need not be.)

Further inspection of Table I indicates that the greatest deviations between the various estimates of τ occur for the smallest value of a . These differences tend to become smaller as a is increased beyond 1. Table II shows this trend quantitatively in four of the above models (WF2, WF3, WF4, and DC.)

A simulation study by Srinivas *et al.*²⁵ on fluorescence energy transfer between chromophores attached terminally to Rouse chains has recently been carried out using the Wilemski–Fixman approach. The simulation methodology was checked by applying it to the polymer cyclization problem and comparing the results with those obtained by Pastor *et al.*⁷ for the Heaviside sink of infinite strength. The two sets of results were found to be in agreement. The results in Ref. 25 are obtained with a $\phi(t)$ based on the following formula derived by Wilemski and Fixman:

$$\phi(t) = \frac{8}{\pi^2} \sum_{\text{odd } n} \frac{4}{n^2} \exp(-\lambda_n t). \quad (54)$$

The eigenvalues here are defined as $\lambda_n = 3D_0 n^2 \pi^2 / N^2 b^2$, with D_0 and b set to 1.

TABLE II. Dependence of closure time τ on reaction radius a for $l=1$ at $N=100$.

a	τ			
	WF2 ^a	WF3 ^b	WF4 ^c	DC ^d
0.1	2104	1306	644	863
0.5	848	673	591	674
1.0	646	558	530	578
1.5	547	484	476	510
2.0	477	432	428	454
2.5	422	387	385	406
3.0	376	348	347	365
3.5	337	314	313	328
4.0	303	284	283	297
4.5	273	257	257	269
5.0	247	233	233	244

^aPZS model as defined in footnote^a Table I.

^bPZS model as defined in footnote^b Table I.

^cPZS model as defined in footnote^c Table I.

^dPresent calculations.

B. Semiflexible chains

The effects of stiffness on chain closure times are studied by progressively decreasing the value of z from its flexible limit of $z=100$. The limit of complete rigidity is reached when $z \rightarrow 0$, but we do not study this limit since rigid rods do not cyclize. Nor do we consider the case $z \ll 1$ (the limit of very restricted flexibility) since it is not clear that our assumption of a Gaussian distribution for the end-to-end distance remains valid under these conditions. Moreover, our equations tend to become numerically unstable in this region of z values. Accordingly, we restrict the further calculation of τ to the regime $1 \leq z \leq 100$. In this regime, we investigate the effect of semiflexibility on τ by plotting τ versus N (on a log-log scale) for different fixed values of z and a . Figure 1 shows one such plot for $z=1$ and $a=0.1, 0.5$, and 1.0 . The graphs are seen to be linear, so one can calculate a scaling

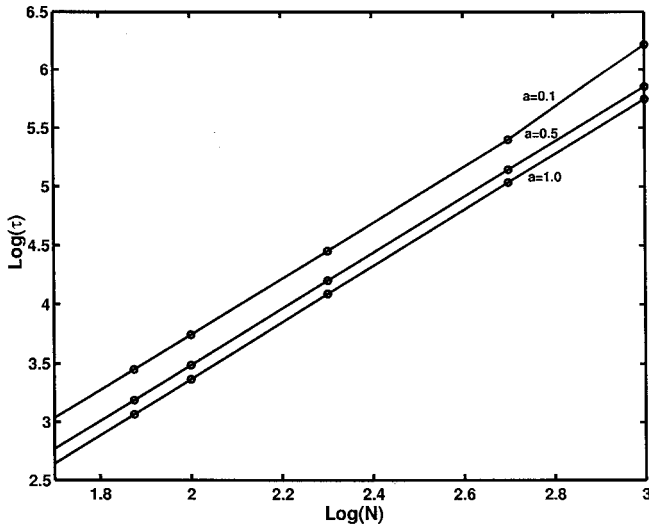


FIG. 1. Variation of closure time τ with chain length N at fixed $z=1$ (corresponding to a chain of limited flexibility) for different values of the reaction distance a (0.1, 0.5, and 1.0). Open circles are the calculated values of τ using Eq. (8) along with Eqs. (12)–(19). Full lines are lines of best fit through the calculated points.

TABLE III. Scaling exponents in the variation of τ with N at different z and a .

z	a	Scaling exponent
		DC ^a
1.0	0.1	2.43
	0.5	2.37
	1.0	2.39
10.0	0.1	2.30
	0.5	2.28
	1.0	2.30
50.0	0.1	2.22
	0.5	2.20
	1.0	2.22
100.0	0.1	2.18
	0.5	2.16
	1.0	2.20

^aPresent calculations.

exponent for the variation of τ with N . For the chosen values of z and a , this exponent is roughly 2.4. At different z (but the same set of a values), τ continues to vary as a power law in N , but with a different scaling exponent (the graphs are not shown, but they are qualitatively the same as Fig. 1). The results are summarized in Table III, and they indicate, in general, that the rate of change of τ with N increases as stiffness increases.

In the opposite limit, i.e., as the chain becomes more flexible, the exponent is seen to approach 2. Appendix B calculates τ for the so-called harmonic spring model of the chain, which is the same as the Rouse model, except that only a single mode is retained in the summation of Eq. (54). Within the Wilemski–Fixman model using the resummed form of the Heaviside sink, τ for this model can be calculated exactly. The calculation shows that $\tau \sim N^2$, independent of a .

Stiff chains of a given persistence length tend to become more flexible as the length of the chain increases. One therefore expects the scaling exponents that describe τ for such chains to cross over to values characteristic of flexible chains as N increases at fixed l . To see if this trend emerges from our calculations, we calculate τ as a function of N in the range $50 \leq N \leq 1000$ for fixed $l=50$ and $a=1.0$. At the smallest value of N , therefore, the chain can be regarded as stiff, and thereafter, it becomes progressively more flexible. Figure 2 shows that the variation of τ with N is indeed characterized by at least two distinct scaling regimes, one lying between about $50 \leq N \leq 150$, where we estimate the exponent to be about 2.33 by linear least squares fitting, and the other lying between about $150 \leq N \leq 1000$, where we estimate the exponent to be about 1.80, again by linear least-squares fitting. Some element of subjectivity obviously enters into the selection of the data points included in the fitting routine. If the scaling exponent in the flexible limit is determined independently by studying the variation of τ with N at fixed l , rather than at fixed z , one finds that for $l=1$ and $a=1.0$, this exponent is about 2.05. The discrepancy between this result and the value of 2.20 obtained from a calculation at fixed z is due to the difference between the two constraints. Varying N at fixed z can generate l values less than 1 (when N is 50 and z is 100, for instance), which we believe are not strictly compatible with the model as defined. Nevertheless, both ap-

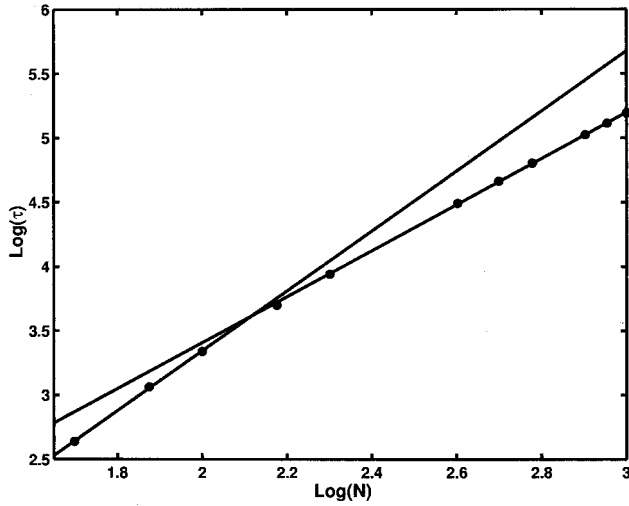


FIG. 2. Variation of closure time τ with chain length N for a reaction radius $a=1$ at fixed persistence length $l=50$ (corresponding to a chain of limited flexibility when $N=50$). Filled circles are data points calculated using Eq. (8) along with Eqs. (12)–(19). Straight lines are lines of best fit through the putative semiflexible and flexible scaling regimes.

proaches yield qualitatively the same predictions.

As a further characterization of the effects of stiffness on τ , we plot (in Fig. 3) the variation of τ with $1/z$ at $N=50$ and three different values of a (0.1, 0.5, and 1.0). As expected, τ increases as l increases at all values of a for the given N . The same qualitative behavior is observed for $N=75$ and $N=100$, but these results are not shown. In the region of relatively high flexibility, the initial increase of τ is quite steep, but thereafter it is much more gradual, and eventually τ approaches an asymptotic value as $z \rightarrow 1$. This asymptotic value is found to be an increasing function of N . Figure 3 also shows that τ is sensitive to the change in the contact distance a . The decrease in a , for a given value of N , leads to a significant increase in τ .

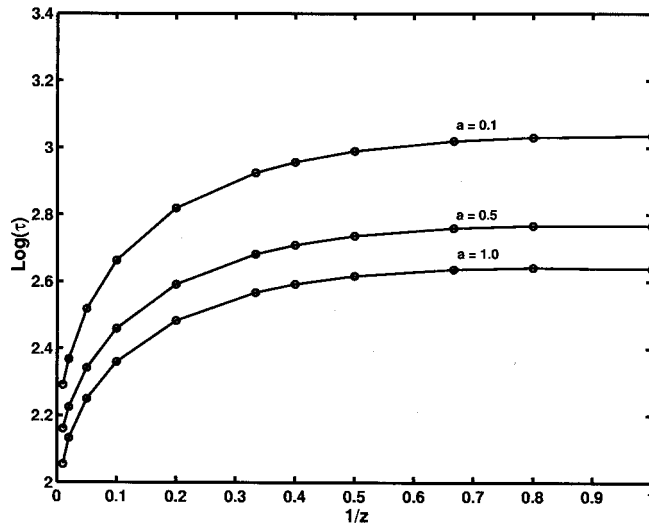


FIG. 3. Variation of τ with $1/z$ at $N=50.0$ for different values of a . Open circles represent points calculated as in Fig. 1 from Eq. (8), while the full lines are guides to the eye.

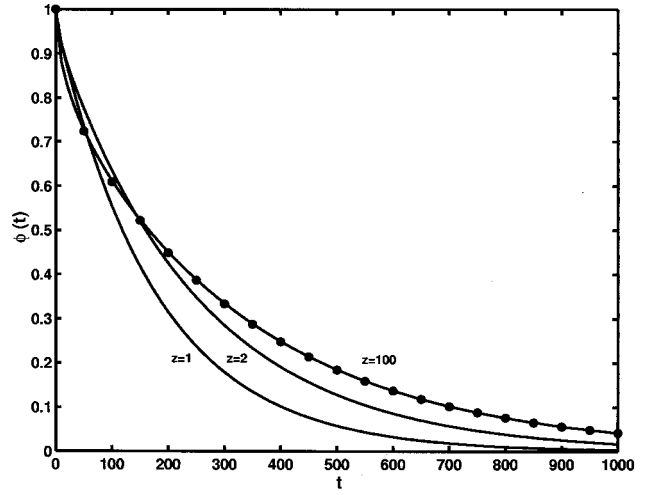


FIG. 4. Variation of $\phi(t)$ [as calculated from Eq. (47)] with t (in units where D_0 and b are unity) at three different values of z (1, 2, and 100). Full circles are data points obtained from the continuum Rouse expression for $\phi(t)$ [Eq. (53)].

One final characterization of stiffness effects is provided in Fig. 4, which shows the variation of $\phi(t)$ with t (measured in units where D_0 and b are unity) at three different values of z (1, 2, and 100). The smallest of these values corresponds to chains that are comparatively stiff, while the largest corresponds to chains of essentially complete flexibility. Also shown on the figure are values of $\phi(t)$ (the full circles) at different values of t as calculated from Eq. (53), the continuum Rouse expression for the dynamical end vector correlation function. These data points are seen to fall exactly over the curve with $z=100$. Curves for z in the range $10 \leq z \leq 100$ (not shown) also coincide almost exactly with this last curve, so chain flexibility seems to set in at about $z=10$.

Although part of the motivation for this study was to rationalize the data of Libchaber and co-workers⁶ on chain closure in DNA of different base sequences, the assumption of irreversibility in the present calculations seems to preclude a direct comparison of their results with ours. The generalization of our approach to the case of reversible chain closure is being studied.

APPENDIX A: DERIVATION OF THE CLOSURE TIME

Formally, the solution to Eq. (2) can be written as

$$\begin{aligned} \psi(\{\mathbf{r}\}, t) = & \psi_{\text{eq}}(\{\mathbf{r}\}) - k \int d\{\mathbf{r}'\} \\ & \times \int_0^t dt' G(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') S(\{\mathbf{r}'\}) \psi(\{\mathbf{r}'\}, t'), \end{aligned} \quad (\text{A1})$$

where the Green's function G satisfies

$$\mathcal{D}G(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') = \delta(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') \delta(t - t'). \quad (\text{A2})$$

If Eq. (A1) is multiplied by $S(\{\mathbf{r}\})$ and the result integrated over all monomer positions, one finds that

$$\begin{aligned} \mu(t) &= \mu_{\text{eq}} - k \int d\{\mathbf{r}\} \int d\{\mathbf{r}'\} \int_0^t dt' \\ &\quad \times S(\{\mathbf{r}\}) G(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') S(\{\mathbf{r}'\}) \psi(\{\mathbf{r}'\}, t'), \end{aligned} \quad (\text{A3})$$

where $\mu(t)$ has been defined in Eq. (5); μ_{eq} is defined in the same way except that $\psi(\{\mathbf{r}\}, t)$ is replaced by $\psi_{\text{eq}}(\{\mathbf{r}\})$.

Equation (A1) is a nonlinear integral equation for the distribution function ψ , and as such cannot be solved in closed form. Wilemski and Fixman,¹¹ therefore, introduced a closure approximation to simplify the equation; this approximation takes the form

$$\psi(\{\mathbf{r}\}, t) \approx \psi_{\text{eq}}(\{\mathbf{r}\}) \omega(t), \quad (\text{A4})$$

where

$$\omega(t) \equiv \frac{\mu(t)}{\mu_{\text{eq}}}. \quad (\text{A5})$$

It is easy to show using Eqs. (A4) and (A5) into Eq. (A3) that

$$\omega(t) = 1 - \frac{k}{\mu_{\text{eq}}} \int_0^t C(t-t') \omega(t') dt', \quad (\text{A6})$$

where $C(t-t')$, which is given by

$$\begin{aligned} C(t-t') &= \int d\{\mathbf{r}\} \int d\{\mathbf{r}'\} S(\{\mathbf{r}\}) G(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') \\ &\quad \times S(\{\mathbf{r}'\}) \psi_{\text{eq}}(\{\mathbf{r}'\}), \end{aligned} \quad (\text{A7})$$

defines a sink–sink correlation function. In principle, this function may depend on the details of the chain conformation as a whole, but in the present calculations it is a function only of the magnitude of the separation between the ends of the chain. Without loss of generality, one of the chain ends may be located at the origin of coordinates $\mathbf{0}$. In this case, $S(\{\mathbf{r}\}) = S(|\mathbf{r}_n|)$, where \mathbf{r}_n is the position of the n th monomer with respect to $\mathbf{0}$. The sink function may now be rewritten identically as

$$S(|\mathbf{r}_n|) = \int d\mathbf{R} S(|\mathbf{R}|) \delta(\mathbf{R} - \mathbf{r}_n). \quad (\text{A8})$$

When this expression is used in Eq. (A7), $C(t-t')$ becomes

$$C(t-t') = \int d\mathbf{R} \int d\mathbf{R}' S(\mathbf{R}) \bar{G}(\mathbf{R}, \mathbf{R}' | t, t') S(\mathbf{R}'), \quad (\text{A9})$$

where

$$\begin{aligned} \bar{G}(\mathbf{R}, \mathbf{R}' | t, t') &\equiv \int d\{\mathbf{r}\} \int d\{\mathbf{r}'\} \delta(\mathbf{r}_n - \mathbf{R}) \\ &\quad \times G(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') \delta(\mathbf{r}'_n - \mathbf{R}') \psi_{\text{eq}}(\{\mathbf{r}'\}). \end{aligned} \quad (\text{A10})$$

The function $\bar{G}(\mathbf{R}, \mathbf{R}' | t, t')$ satisfies the relations $\int d\mathbf{R} \bar{G}(\mathbf{R}, \mathbf{R}' | t, t') = \psi_{\text{eq}}(\mathbf{R}')$ and $\int d\mathbf{R}' \bar{G}(\mathbf{R}, \mathbf{R}' | t, t') = \psi_{\text{eq}}(\mathbf{R})$. The introduction of a Green's function $G(\mathbf{R}, \mathbf{R}' | t, t')$ defined by $\bar{G}(\mathbf{R}, \mathbf{R}' | t, t') = G(\mathbf{R}, \mathbf{R}' | t, t') \psi_{\text{eq}}(\mathbf{R}')$, satisfying the normalization condi-

tions $\int d\mathbf{R} G(\mathbf{R}, \mathbf{R}' | t, t') = \int d\mathbf{R}' G(\mathbf{R}, \mathbf{R}' | t, t') = 1$ allows the time correlation function $C(t-t')$ to be written as

$$C(t-t') = \int d\mathbf{R} \int d\mathbf{R}' S(\mathbf{R}) G(\mathbf{R}, \mathbf{R}' | t, t') S(\mathbf{R}') \psi_{\text{eq}}(\mathbf{R}'). \quad (\text{A11})$$

To proceed further, one now takes the Laplace transform of Eq. (A6), which produces

$$\tilde{\omega}(s) = \frac{1}{s} - \frac{k}{\mu_{\text{eq}}} \tilde{C}(s) \tilde{\omega}(s), \quad (\text{A12})$$

or equivalently

$$\tilde{\omega}(s) = s^{-1} \left[1 + \frac{k \tilde{C}(s)}{\mu_{\text{eq}}} \right]^{-1}. \quad (\text{A13})$$

Similarly, the Laplace transform of Eq. (4) and the definition Eq. (A5) leads to

$$\tilde{\chi}(s) = s^{-1} [1 - k \mu_{\text{eq}} \tilde{\omega}(s)]. \quad (\text{A14})$$

Now introduce the function

$$H(t) \equiv C(t) - C(\infty), \quad (\text{A15})$$

where $C(\infty)$ is the long time limit of $C(t)$. From the definition (A7) and the limit $G(\{\mathbf{r}\}, \{\mathbf{r}'\} | t, t') \rightarrow \psi_{\text{eq}}(\{\mathbf{r}\})$ when $t \rightarrow \infty$, $C(\infty)$ is seen to be given by

$$C(\infty) \equiv \lim_{t \rightarrow \infty} C(t) = \mu_{\text{eq}}^2. \quad (\text{A16})$$

Hence, from Eq. (A13) and the Laplace transform of Eq. (A15)

$$\tilde{\omega}(s) = \left[s + \frac{k}{\mu_{\text{eq}}} s \tilde{H}(s) + k \mu_{\text{eq}} \right]^{-1}, \quad (\text{A17})$$

which in turn when substituted into Eq. (A14) produces

$$\tilde{\chi}(s) = \left[1 + \frac{k}{\mu_{\text{eq}}} \tilde{H}(s) \right] \tilde{\omega}(s). \quad (\text{A18})$$

Furthermore, from Eq. (7), τ can be written as

$$\tau = \tilde{\chi}(0), \quad (\text{A19})$$

while from Eqs. (A18), (A17), and (A15), respectively, we have

$$\tilde{\chi}(0) = \left[1 + \frac{k}{\mu_{\text{eq}}} \tilde{H}(0) \right] \tilde{\omega}(0), \quad (\text{A20})$$

$$\tilde{\omega}(0) = k^{-1} \mu_{\text{eq}}^{-1}, \quad (\text{A21})$$

and

$$\tilde{H}(0) = \int_0^\infty dt [C(t) - \mu_{\text{eq}}^2]. \quad (\text{A22})$$

After substituting Eqs. (A21) and (A22) into Eq. (A20), the expression for $\tilde{\chi}(0)$ becomes

$$\tilde{\chi}(0) = k^{-1} \mu_{\text{eq}}^{-1} + \int_0^\infty dt \left[\frac{C(t)}{\mu_{\text{eq}}^2} - 1 \right]. \quad (\text{A23})$$

Since chain cyclization is assumed to take place essentially instantaneously, the rate constant k is effectively infinite, so that in this limit, using Eq. (A23) in the definition of the closure time, one finally obtains

$$\tau = \tilde{\chi}(0) = \int_0^\infty dt \left[\frac{C(t)}{C(0)} - 1 \right]. \quad (\text{A24})$$

APPENDIX B: THE HARMONIC SPRING MODEL

The closure time for the harmonic spring (HS) model within the PZS reformulation of the Wilemski–Fixman approximation is derived as follows: recall that in this approximation the time correlation function $C(t)$ is given by

$$C(t) = \frac{16x_0^3}{9\pi} \left(1 - \phi^2 + \frac{4x_0}{5} \right)^{-3/2}. \quad (\text{B1})$$

Therefore,

$$\frac{C(t)}{C(\infty)} - 1 = \sum_{k=1}^{\infty} \frac{(2k+1)!!}{k!} \left(\frac{\phi^2}{2X} \right)^k, \quad (\text{B2})$$

where $X \equiv 1 + 4x_0/5$. The harmonic spring model is defined by retaining only the first mode in Eq. (53) [which defines the function $\phi(t)$], i.e.,

$$\phi_{\text{HS}}(t) = \frac{8}{\pi^2} \exp(-3\pi^2 D_0 t / N^2) \equiv B e^{-Ct}. \quad (\text{B3})$$

From the above equation, and from the definition of the closure time as the time integral from 0 to ∞ of Eq. (B2), it follows that:

$$\tau_{\text{HS}} \equiv \frac{1}{2C} S(\alpha) = \frac{1}{2C} \sum_{k=1}^{\infty} \frac{(2k+1)!!}{kk!} \alpha^k, \quad \alpha = \frac{B^2}{2A}. \quad (\text{B4})$$

Using the identities

$$(2k+1)!! = \frac{2^{k+1}}{\sqrt{\pi}} \Gamma(k+3/2) = \frac{2^{k+1}}{\sqrt{\pi}} (k+1/2) \Gamma(k+1/2) \quad (\text{B5})$$

and

$$\Gamma(k+1/2) = \sqrt{\pi} 2^{-2k} \frac{(2k)!}{k!} \quad (\text{B6})$$

the sum $S(\alpha)$ can be written as $(2/\sqrt{\pi})S_1(\alpha) + (1/\sqrt{\pi})S_2(\alpha)$, where

$$S_1(\alpha) = \sqrt{\pi} \sum_{k=1}^{\infty} \frac{(2k)!}{(k!)^2} (\alpha/2)^k, \quad (\text{B7})$$

$$S_2(\alpha) = \sqrt{\pi} \sum_{k=1}^{\infty} \frac{(2k)!}{k(k!)^2} (\alpha/2)^k.$$

The sum S_1 is known; in closed form, it is given by

$$S_1 = \sqrt{\pi} \left[\frac{1}{\sqrt{1-2\alpha}} - 1 \right]. \quad (\text{B8})$$

Since S_1 can be expressed in terms of a derivative with respect to α of S_2 , it follows that S_2 can be calculated as:

$$S_2(x) = \int_0^x dy \frac{1}{y} \sqrt{\pi} \left(\frac{1}{\sqrt{1-y}} - 1 \right). \quad (\text{B9})$$

The change of variables $y = \sin^2 \theta$ and the use of the trigonometric identities $\cos \theta = 1 - 2 \sin^2(\theta/2)$ and $\sin \theta = 2 \sin(\theta/2) \cos(\theta/2)$ allows this integral to be evaluated at once. The result is

$$S_2(\alpha) = -4\sqrt{\pi} \ln(\cos(\frac{1}{2} \sin^{-1}(\sqrt{2\alpha}))). \quad (\text{B10})$$

Therefore,

$$\tau_{\text{HS}} = \frac{1}{C} \left[\frac{1}{\sqrt{1-2\alpha}} - 1 - 2 \ln \left(\cos \left(\frac{1}{2} \sin^{-1}(\sqrt{2\alpha}) \right) \right) \right]. \quad (\text{B11})$$

APPENDIX C: DERIVATION OF THE CORRELATION FUNCTION OF THE MONOMER POSITION

The quantity of interest is the average $\langle \mathbf{r}(s_1, 0) \cdot \mathbf{r}(s, 0) \rangle$ [see Eq. (46)]. It can be obtained from the expression for the mean-square distance between any two monomers on the chain, which can be written in terms of the correlation of tangent vectors as

$$\langle [\mathbf{r}(s) - \mathbf{r}(s')]^2 \rangle = \int_{s'}^s ds_1 \int_{s'}^s ds_2 \langle \mathbf{u}(s_1) \cdot \mathbf{u}(s_2) \rangle. \quad (\text{C1})$$

The tangent vector correlation function in turn is given by

$$\begin{aligned} \langle \mathbf{u}(s_1) \cdot \mathbf{u}(s_2) \rangle &= \frac{1}{Z} \int d\mathbf{u}_f \int d\mathbf{u}_2 \int d\mathbf{u}_1 \int d\mathbf{u}_i \mathbf{u}_2 \cdot \mathbf{u}_1 K(\mathbf{u}_f, \mathbf{u}_2; N/2, s_2) \\ &\quad \times K(\mathbf{u}_2, \mathbf{u}_1; s_2, s_1) K(\mathbf{u}_1, \mathbf{u}_i; s_1, -N/2), \end{aligned} \quad (\text{C2})$$

where the propagator K , in general, is given by

$$K(\mathbf{u}_f, \mathbf{u}_i; N/2, -N/2) = \int_{\mathbf{u}(-N/2)=\mathbf{u}_i}^{\mathbf{u}(N/2)=\mathbf{u}_f} D[\mathbf{u}(s)] \exp[-S[\mathbf{u}]], \quad (\text{C3})$$

with

$$\begin{aligned} S[\mathbf{u}] &= \nu \int_{-N/2}^{N/2} ds |\mathbf{u}(s)|^2 + \eta \int_{-N/2}^{N/2} ds \left| \frac{\partial \mathbf{u}(s)}{\partial s} \right|^2 \\ &\quad + \nu_0 (|\mathbf{u}_{N/2}|^2 + |\mathbf{u}_{-N/2}|^2). \end{aligned} \quad (\text{C4})$$

Z is the configurational partition function of the semiflexible chain, and is identical to the numerator of Eq. (C2) except for the absence of the factor $\mathbf{u}_1 \cdot \mathbf{u}_2$.

The propagator $K(\mathbf{u}_f, \mathbf{u}_i; N/2, -N/2)$ can be evaluated by minimizing the action $S[\mathbf{u}]$, i.e., by finding the trajectory $\bar{\mathbf{u}}(s)$ such that $\delta S[\mathbf{u}] / \delta \mathbf{u} = 0$. The required trajectory satisfies the equation

$$\frac{\partial^2 \bar{\mathbf{u}}(s)}{\partial s^2} - \frac{\nu}{\eta} \bar{\mathbf{u}}(s) = 0, \quad (\text{C5})$$

which is readily solved. The solution, when substituted in Eq. (C4), gives the following expression for the minimized action:

$$S[\bar{\mathbf{u}}] = \frac{3}{4 \sinh(2p(s_2 - s_1))} [(\mathbf{u}_1^2 + \mathbf{u}_2^2) \cosh(2p(s_2 - s_1)) - 2\mathbf{u}_1 \cdot \mathbf{u}_2] \quad (\text{C6})$$

Since S is quadratic in \mathbf{u} , the substitution $\mathbf{u}(s) \rightarrow \mathbf{u}(s) + \bar{\mathbf{u}}(s)$ in Eq. (C3) leads to

$$K(\mathbf{u}_f, \mathbf{u}_i; N/2, -N/2) = \exp[-S[\bar{\mathbf{u}}]] K(0, 0; N/2, -N/2). \quad (\text{C7})$$

The above expression for the propagator can be used in Eq. (C2) to calculate $\langle \mathbf{u}(s_1) \cdot \mathbf{u}(s_2) \rangle$ by carrying out the integration over \mathbf{u}_i , \mathbf{u}_f , \mathbf{u}_1 , and \mathbf{u}_2 . The result is

$$\langle \mathbf{u}(s_1) \cdot \mathbf{u}(s_2) \rangle = e^{-2p|s_1 - s_2|}. \quad (\text{C8})$$

Integration over s_1 and s_2 , as given in Eq. (C1), produces the following expression

$$\langle [\mathbf{r}(s) - \mathbf{r}(s')]^2 \rangle = \frac{|s - s'|}{p} - \frac{1}{2p^2} [1 - e^{-2p|s - s'|}]. \quad (\text{C9})$$

The left-hand side of the above expression can be expanded and rearranged to obtain Eq. (46).

¹D. Shore and R. L. Baldwin, J. Mol. Biol. **170**, 958 (1983).

²W. H. Taylor and P. Hagerman, J. Mol. Biol. **212**, 363 (1990).

³T. M. Paner, M. Amaratunga, M. J. Doktycz, and S. Benight, Biopolymers **29**, 1715 (1990).

⁴L. Ying, M. I. Wallace, and D. Klenerman, Chem. Phys. Lett. **334**, 145 (2001).

⁵T. B. Liverpool and S. F. Edwards, J. Chem. Phys. **103**, 6716 (1995).

⁶N. L. Goddard, G. Bonnet, O. Krichevsky, and A. Libchaber, Phys. Rev. Lett. **85**, 2400 (2000).

⁷R. W. Pastor, R. Zwanzig, and A. Szabo, J. Chem. Phys. **105**, 3878 (1996).

⁸G. Wilemski and M. Fixman, J. Chem. Phys. **58**, 4009 (1973).

⁹G. Wilemski and M. Fixman, J. Chem. Phys. **60**, 866 (1974).

¹⁰G. Wilemski and M. Fixman, J. Chem. Phys. **58**, 878 (1974).

¹¹A. Szabo, K. Schulten, and Z. Schulten, J. Chem. Phys. **72**, 4350 (1980).

¹²A. Dua and B. J. Cherayil, J. Chem. Phys. **113**, 10776 (2000).

¹³D. T. Gillespie, Am. J. Phys. **64**, 225 (1996).

¹⁴N. Saitô, K. Takahashi, and Y. Yunoki, J. Phys. Soc. Jpn. **22**, 219 (1967).

¹⁵M. Doi, Chem. Phys. **9**, 455 (1975).

¹⁶R. A. Harris and J. E. Hearst, J. Chem. Phys. **44**, 2595 (1966).

¹⁷L. Harnau, R. G. Winkler, and P. Reineker, J. Chem. Phys. **102**, 7750 (1995).

¹⁸M. G. Bawendi and K. F. Freed, J. Chem. Phys. **83**, 2491 (1985).

¹⁹J. B. Lagowski and J. Noolandi, J. Chem. Phys. **95**, 1266 (1991).

²⁰A. Podtelezhnikov and A. Vologodskii, Macromolecules **30**, 6668 (1997).

²¹A. Perico and C. Cuniberti, J. Polym. Sci., Polym. Phys. Ed. **15**, 1435 (1977); M. Battezzati and A. Perico, J. Chem. Phys. **74**, 4527 (1981); C. Cuniberti and A. Perico, Prog. Polym. Sci. **10**, 271 (1984); A. Perico and M. Beggiano, Macromolecules **23**, 797 (1990).

²²J. L. Garcia Fernández, A. Rey, J. J. Freire, and I. F. de Piérola, Macromolecules **23**, 2057 (1990); A. Rey and J. J. Freire *ibid.* **24**, 4673 (1991).

²³B. Friedman and B. O'Shaughnessy, Phys. Rev. A **40**, 5950 (1989).

²⁴J. Stampe and I. M. Sokolov, J. Chem. Phys. **114**, 5043 (2001).

²⁵G. Srinivas, A. Yethiraj, and B. Bagchi, J. Chem. Phys. **114**, 9170 (2001).