

Dynamics of chain closure: Approximate treatment of nonlocal interactions

Pallavi Debnath and Binny J. Cherayil^{a)}

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

The Wilemski–Fixman model of diffusion controlled-reactions [J. Chem. Phys. **58**, 4009 (1973)] is combined with a generalized random walk description of chain conformations to predict the dependence of the closure time τ on the chain length N of polymers with reactive end groups and nonlocal interactions. The nonlocal interactions are modeled by a modification to the connectivity term in the Edwards continuum representation of the polymer. The modification involves a parameter h lying between 0 and 1 that is a measure of the extent of correlation between adjacent monomers on the chain backbone. Different choices of h correspond to chain conformations of different average radial dimensions. In particular, the values $1/3$, $1/2$ and $3/5$ provide approximations to the statistics of polymers in poor, theta and good solvents, respectively. The closure time τ of such chains is calculated analytically for different N . In all cases, τ is found to vary as a power law in N , N^b , with b a function of h . For the special case $h = 1/3$, which models collapsed polymers and globular proteins, b is about 1.6–1.7.

I. INTRODUCTION

Loop formation in polymers can often serve as a paradigm for biological processes in which distant parts of a macromolecule come together and initiate a series of physical or chemical events that lead to permanent conformational changes. Such processes frequently occur through the association of specific pairs of residues or domains within the macromolecule, which are brought into proximity with each other through the random thermal fluctuations of the medium. If the first encounter of the reactive pair is rate-determining, knowledge of the mean first passage time τ for chain closure can provide insights into the nature of the conformational transition.¹

Although biopolymers are structurally complex, even simple statistical models can sometimes describe their properties successfully when short-range interactions are the dominant controlling factor in their behavior.² In theoretical descriptions of molecules like DNA, RNA, and homopolypeptides, many of the structural details at the level of the repeat unit can often be ignored, and the molecule represented as a succession of elementary units with interactions that extend at most to the nearest or next nearest neighbors. If these units are further reduced to points, the molecule can in fact be represented as a continuous one-dimensional string. Continuum models of this kind have recently been used with some success to calculate closure times in synthetic sticky-ended single-stranded DNA.^{3,4}

The extension of such calculations to more general situations, in particular to polymers with non-negligible long range interactions, is nontrivial. In these cases, even the problem of simple cyclization involving the ends of the chain, much less the problem of contact pairing between in-

terior elements and other more realistic folding pathways, has not been systematically investigated. A number of questions about the problem have, therefore, not been fully addressed. It is not yet known, for instance, how τ , the mean closure time, varies with the molecular weight M in chains with long-range interactions. Such knowledge may provide clues to the factors behind important conformational changes, including those that direct the course of protein folding, which can be accomplished within microseconds or less, after somehow circumventing the problem of multiple local minima.⁵

It is, therefore, of considerable interest to calculate τ for chains with long-range interactions, particularly attractive interactions, since such interactions govern the formation of compact globular configurations. However, a full-scale attack on the problem, starting from even a minimal continuum representation of the chain—such as the Edwards model, with two and possibly three-body interactions included⁶—is generally intractable under any but the most severe approximations. One way to overcome the difficulties associated with nonlocal effects is to incorporate these effects into a description of the chain that is effectively local, and that is, therefore, simpler to treat, in principle. A realization of this idea may be found in polymer models that are formulated around the statistics of “generalized” random walks,⁷ rather than the simple random walks that underlie more conventional approaches. In the path integral representation of these models, the Hamiltonian is a generalization of the connectivity term in the Edwards model, all the effects of excluded volume and other long-range interactions being contained in a single parameter h that appears there. By suitable choice of h , long wavelength properties of the chain, such as the end-to-end distance, or the radius of gyration, can be calculated in fairly close agreement with results obtained

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

by more sophisticated models, without the explicit inclusion of excluded volume effects.

A similar calculation of τ may therefore provide a useful *preliminary* theory of its behavior under conditions where other models would be difficult or impossible to treat. The subject of this paper is the calculation of τ within such an approach.

As reviewed in Sec. II below, the dynamics of chains with reactive end-groups is described by a Smoluchowski equation with an added sink term, the sink term representing the loss of probability from chain closure. When this term is chosen to be a function solely of the magnitude of the separation between the ends of the chain, its time correlation function $C(t)$ can be expressed in terms of the time correlation function of the end-to-end distance. Further, the integral of $C(t)$ is related to the mean first passage time τ . $C(t)$ depends critically on the details of the chosen model, and is therefore an important intermediate step in the calculation of τ . The calculation of $C(t)$, and from it τ , is discussed in Sec. III within the formalism of generalized random walks. Section IV presents the principal results of this calculation.

II. DYNAMICS OF CHAIN CLOSURE

In a neutral solvent, the dynamics of an isolated polymer having reactive end-groups can be described by the following Smoluchowski equation:^{8,9}

$$\frac{\partial \psi(\{\mathbf{r}\}, t)}{\partial t} = \mathcal{D}\psi(\{\mathbf{r}\}, t) - kS(\{\mathbf{r}\})\psi(\{\mathbf{r}\}, t). \quad (1)$$

Here, $\psi(\{\mathbf{r}\}, t)$ 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; k is a rate constant; S is a sink function that prescribes the conditions under which reaction between the ends of the chain takes place; and \mathcal{D} is a differential operator, defined as

$$\mathcal{D} \equiv D_0 \sum_{i=1}^n \frac{\partial}{\partial \mathbf{r}_i} \cdot \left[\frac{\partial}{\partial \mathbf{r}_i} + \frac{\partial U}{\partial \mathbf{r}_i} \right] \psi(\{\mathbf{r}\}, t), \quad (2)$$

with D_0 the diffusion constant, defined as the inverse of the friction coefficient per unit length of the polymer, and U the potential energy of the chain.

An approximate expression for the mean first passage time τ can be derived from Eq. (1). The derivation has been discussed at length in earlier references,³ so only the final result will be given here:

$$\tau = \int_0^\infty dt \left(\frac{C(t)}{C(\infty)} - 1 \right), \quad (3)$$

where

$$C(t) = \int d\mathbf{R} \int d\mathbf{R}_0 S(R) G(\mathbf{R}, t | \mathbf{R}_0, 0) S(R_0) \psi_{\text{eq}}(\mathbf{R}_0). \quad (4)$$

In the above expression, $G(\mathbf{R}, t | \mathbf{R}_0, 0)$ 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 the sink function, which is assumed here to depend only

on the magnitude of the separation between the chain ends. An expression for $S(R)$ will be provided later. Equation (3) is not exact; it makes use of a self-consistent approximation introduced by Wilemski and Fixman, and also imposes the limit $k \rightarrow \infty$.⁸ The first approximation corresponds to the assumption of local equilibrium, while the second corresponds to the assumption that the chain ends react immediately and irreversibly whenever they satisfy the distance constraints defined by the sink function.

The calculation of $C(t)$ requires knowledge of the conditional probability $G(\mathbf{R}, t | \mathbf{R}_0, 0)$. For a completely flexible polymer, this probability can be taken as Gaussian, because \mathbf{R} is the sum of a large number of independent, random bond vectors, and the central limit theorem may therefore be invoked.⁹ For a self-avoiding or self-attracting chain, or a chain of limited flexibility, the probability is no longer necessarily Gaussian, but it will be assumed to be so in order to render the calculations tractable. A future publication will discuss how the assumption may be lifted.¹⁰

The requisite Gaussian form of the conditional probability in Eq. (4) is⁹

$$G(\mathbf{R}, t | \mathbf{R}_0, 0) = \left(\frac{3}{2\pi \langle R^2 \rangle_{\text{eq}}} \right)^{3/2} \frac{1}{(1 - \phi^2(t))^{3/2}} \times \exp \left[- \frac{3(\mathbf{R} - \phi(t)\mathbf{R}_0)^2}{2\langle R^2 \rangle_{\text{eq}}(1 - \phi^2(t))} \right], \quad (5)$$

where the function $\phi(t)$ is defined as

$$\phi(t) = \frac{\langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle_{\text{eq}}}{\langle R^2 \rangle_{\text{eq}}}, \quad (6)$$

the angle brackets $\langle (\dots) \rangle_{\text{eq}}$ denoting an ensemble average over ψ_{eq} .

After substituting Eq. (5) into Eq. (4), and carrying out the angular part of the integrations over the vectors \mathbf{R} and \mathbf{R}_0 , $C(t)$ becomes

$$C(t) = 16\pi^2 \left(\frac{3}{2\pi \langle R^2 \rangle_{\text{eq}}} \right)^3 \frac{1}{(1 - \phi^2(t))^{3/2}} \times \int_0^\infty dR R^2 S(R) \int_0^\infty dR_0 R_0^2 S(R_0) \times \exp \left[- \frac{3}{2\langle R^2 \rangle_{\text{eq}}} \frac{(R^2 + R_0^2)}{(1 - \phi^2(t))} \right] \times \frac{\sinh[3\phi R R_0 / \langle R^2 \rangle_{\text{eq}} (1 - \phi^2(t))]}{3\phi R R_0 / \langle R^2 \rangle_{\text{eq}} (1 - \phi^2(t))}. \quad (7)$$

To complete the calculation of $C(t)$, the sink function $S(R)$ must be specified. Here, $S(R)$ is chosen to be the Heaviside sink, defined as

$$S(R) = 1; \quad R \leq a \\ = 0; \quad R > a \quad (8)$$

where a is the distance of closest approach of the chain ends prior to irreversible cyclization. When $S(R)$ is chosen in this way, the integrals in Eq. (7) cannot be carried out exactly. However, the expansion of the integrand in the small parameter

$$x_0 = \frac{3a^2}{2\langle R^2 \rangle_{\text{eq}}}, \quad (9)$$

followed by term by term integration leads to an approximate result that is expected to be accurate.⁹ In this way, it is found that

$$C(t) = \frac{16x_0^3}{9\pi(1-\phi^2)^{3/2}} \left[1 - \frac{6x_0}{5(1-\phi^2)} + \dots \right]. \quad (10)$$

To the same order of approximation, this equation may be resummed as

$$C(t) \cong \frac{16x_0^3}{9\pi} \left(1 - \phi^2 + \frac{4}{5}x_0 \right)^{-3/2}, \quad (11)$$

which is of exactly the same functional form as the expression derived by Doi¹¹ using the following Gaussian sink function:

$$S(R) = \mathcal{N} \exp\left(-\frac{3R^2}{2a^2}\right), \quad (12)$$

where \mathcal{N} is a normalization constant. Equation (11) differs from the corresponding result in Ref. 11 in the value of the coefficient of x_0 ; it is 4/5 above, and 4/3 in Doi's calculation.

It only remains to determine $\phi(t)$ to proceed further with the calculation of τ . The calculation of $\phi(t)$ is discussed in the following section.

III. CHAIN CONFORMATIONS AS GENERALIZED RANDOM WALKS

A. The model

The defining ingredient in most models of the the statistical mechanics of polymers is the chain Hamiltonian H . In models where the polymer is depicted as the locus of a string of contour length N , and in units where $k_B T$ is unity, H is typically written as¹²

$$H = H_0 + H_v, \quad (13)$$

$$\begin{aligned} &= \frac{3}{2l} \int_0^N d\tau \left(\frac{\partial \mathbf{r}(\tau)}{\partial \tau} \right)^2 + \frac{1}{2} \int_0^N d\tau \\ &\times \int_0^N d\tau' V[\mathbf{r}(\tau) - \mathbf{r}(\tau')], \end{aligned} \quad (14)$$

where l corresponds, effectively, to a bond length, $\mathbf{r}(\tau)$ defines the trajectory of the string in three-dimensional space in terms of a continuous variable τ , which specifies the location of individual monomers (imagined to be points) along the backbone of the chain, and $V[\mathbf{r}(\tau) - \mathbf{r}(\tau')]$ is a potential energy term that accounts for the effects of pair interactions—both attractive and repulsive—between non-bonded segments. Higher body contributions to the potential, if non-negligible, may also be included in Eq. (14) as additional additive terms, along with any other terms that may be necessary to describe the interactions of the polymer with its environment.

Calculations with this and related Hamiltonians are generally nontrivial, even for relatively simple approximations

to V , such as the delta function pseudopotential $v_0 \delta(\mathbf{r}(\tau) - \mathbf{r}(\tau'))$ introduced by Edwards, but they simplify considerably when the nonbonded interaction terms are neglected altogether. The neglect of these interactions yields a model in which the chain is completely flexible, with no penalty for unphysical self-intersections. The statistics of such a chain are described by a Gaussian distribution of chain segments, so that the mean square of the distance \mathbf{R} from one end of the chain to other is given by $\langle R^2 \rangle = Nl$.

Although the term H_0 in Eq. (14) is not strictly an energetic contribution, it is commonly interpreted as the energy of a chain of n harmonic oscillators with near neighbor couplings in the limit $n \rightarrow \infty$, $l \rightarrow 0$, $nl \rightarrow N$. But it may also be interpreted as the weight function for the trajectory of a point particle that evolves in “time” τ under the action of a stochastic force with white noise statistics.⁷ The equation of motion of this particle is

$$\frac{\partial r_\alpha(\tau)}{\partial \tau} = \theta_\alpha(\tau), \quad \alpha = x, y, z, \quad (15)$$

where the random variable $\theta(\tau)$ is defined by

$$\langle \theta_\alpha(\tau) \rangle = 0, \quad (16)$$

$$\langle \theta_\alpha(\tau) \theta_\beta(\tau') \rangle = \frac{l}{3} \delta_{\alpha\beta} \delta(\tau - \tau'). \quad (17)$$

These “dynamics” may be reformulated in terms of path integrals; one then obtains an expression for the partition function of the particle trajectories in which the Boltzmann factor is found to be $\exp(-H_0)$, with H_0 identical to the expression given in Eq. (14).

Since H_0 may be interpreted in terms of particle trajectories (which are equivalent, at the coarse-grained level, to individual polymer configurations), modifications to the dynamics defined by Eq. (15) can be expected to yield alternative representations of the polymer backbone. This idea was explored in Ref. 7 by assuming that the noise term in Eq. (15) is not white, but colored, and is defined completely by the following values of its first two moments:

$$\langle \theta_\alpha(\tau) \rangle = 0, \quad (18)$$

$$\begin{aligned} \langle \theta_\alpha(\tau) \theta_\beta(\tau') \rangle &= \frac{l^{2-2h}}{6\Gamma^2(h+1/2)} \delta_{\alpha\beta} \delta(\tau - \tau') \\ &\times \{ \tau^{2h-1} + (\tau')^{2h-1} \}, \end{aligned} \quad (19)$$

where $\Gamma(x)$ is the gamma function, and h is a parameter lying between 0 and 1 that characterizes the degree of correlation between particle positions at successive instants of “time” τ . The same path integral analysis as used earlier now leads to the Boltzmann factor $\exp(-H_G)$ in the expression for the partition function of the particle trajectories, where H_G is given by:⁷

$$H_G = \frac{3\Gamma^2(h+1/2)}{2l^{2-2h}} \int_0^N d\tau \left(\frac{\partial \mathbf{r}(\tau)}{\partial \tau} \right)^2 \tau^{1-2h}. \quad (20)$$

If H_G is regarded as the Hamiltonian of the equivalent polymer configuration, the average square of the distance between the ends of the chain in these configurations is easily calculated formally. The result is

$$\langle R^2 \rangle = \frac{1}{2h\Gamma^2(h+1/2)} N^{2h} l^{2-2h}. \quad (21)$$

Equation (21) is a generalization of the corresponding result for random walks, and reduces to it when $h=1/2$. Other choices of h can model the scaling behavior of polymers with fully developed repulsive or attractive long-ranged interactions. For instance, self-avoiding walks, i.e., chains with strong excluded volume forces between nonbonded segments, can be modeled by $h=3/5$,^{7,13} and collapsed polymers, such as compact globular proteins and the like, can be modeled by $h=1/3$.^{14,15}

The utility of the above generalized random walk model of chain statistics has been investigated in a series of papers from this laboratory.^{7,16-18} Although the model is not realistic at the microscopic level, it does provide a satisfactory description of *macroscopic* chain properties for different values of h . Apart from $\langle R^2 \rangle$ itself, these include the scaling exponents of the thermodynamic anomalies characterizing phase separating polymer solutions,¹⁶ scaling exponents for relaxation times in the Rouse and Rouse-Zimm models,¹⁷ and various quantitative descriptors of chain shape.¹⁸ In the present calculation, our interest is essentially confined to the chain length dependence of the mean time of chain closure, which is governed by the gross conformational dynamics of the chain, and is, therefore, expected to be adequately described by such a model. At any rate, in the absence of other models of chain dynamics that could treat the effects of long-range interactions analytically, the generalized random walk model is a convenient zeroth-order approximation.

B. Dynamics

Given the above expression for H_G , the actual dynamics of the chain (the dynamics in real time t as opposed to dynamics in the space of monomer positions τ) can be described by the following Langevin equation:

$$\frac{\partial r_\alpha(\tau, t)}{\partial \tau} = -\frac{1}{\zeta} \frac{\delta H_G}{\delta r_\alpha(\tau, t)} + \eta_\alpha(\tau, t), \quad \alpha = x, y, z, \quad (22)$$

where ζ is the monomer friction coefficient, and η is a Gaussian random variable with statistics defined by

$$\langle \eta_\alpha(\tau, t) \rangle = 0, \quad (23)$$

$$\langle \eta_\alpha(\tau, t) \eta_\beta(\tau', t') \rangle = \frac{2}{\zeta} \delta_{\alpha, \beta} \delta(\tau - \tau') \delta(t - t'). \quad (24)$$

Equation (22), along with Eq. (20) for H_G , may be solved by a normal mode decomposition if a complete orthonormal basis set can be found that separates the dynamics of the coupled chain segments into independent motions of the normal modes. Based on results derived by Maccone,¹⁹ the appropriate basis set, denoted here as $\phi_n(\tau)$, was identified in Ref. 7 as being proportional to Bessel functions $J_\nu(z)$. Specifically,

$$\phi_n(\tau) = A_n \tau^h J_\nu(\gamma_n(\tau/N)^{h+1/2}), \quad (25)$$

where ν , A_n , and γ_n are constants related to the parameters of the polymer model; in Ref. 7, the first two parameters were given expressly by

$$\nu = \frac{2h}{2h+1}, \quad (26)$$

$$A_n = \frac{\sqrt{2h+1}}{N^{h+1/2} J_\nu(\gamma_n)}, \quad (27)$$

while the third was found to be determined by the solution of the equation

$$J_{\nu-1}(\gamma_n) = 0. \quad (28)$$

The parameter γ_n is, therefore, the n th zero of the Bessel function of order $\nu-1$.

From the properties of Bessel functions,²⁰ one sees that in the limit $h=1/2$, corresponding to the case of the simple random walk (or a polymer at the theta point), the normal modes of the generalized random walk reduce to sines. Although sine functions are an acceptable basis set with which to decouple the dynamics of the Rouse model [Eq. (22) with $h=1/2$], the boundary conditions they satisfy do not correctly describe the behavior of the ends of the chain; their estimates of quantities related to relaxation times, therefore, tend to differ from conventional treatments based on cosines. Either basis set yields exactly the same exponents in the molecular weight dependence of the relaxation times, but not the same amplitudes.¹⁷ The differences can be minimized by normalizing the amplitudes to some standard state, as shown in Ref. 17, but it is preferable that the formalism exactly recover the Rouse limit under the appropriate conditions.

This can be achieved by a simple change of sign in Eq. (26), i.e., by defining ν not as $+2h/(2h+1)$ but as

$$\nu = -\frac{2h}{2h+1}. \quad (29)$$

The change is permitted, as either definition provides a legitimate solution to the Bessel equation that formed the basis for Maccone's original treatment of extensions to Brownian motion.¹⁹ The basis functions $\phi_n(\tau)$ continue to be given by Eq. (25), with A_n given by Eq. (27), but the parameter γ_n , as may be shown, is now determined, not by the solution of Eq. (28), but by the solution of

$$J_{\nu+1}(\gamma_n) = 0. \quad (30)$$

In other words, the γ_n are now defined as the zeros of the Bessel function of order $\nu+1$, with ν given by Eq. (29). This choice of basis set yields cosines in the limit $h=1/2$, and *exactly* reproduces the relaxation times of the Rouse chain, as will be shown later.

C. Time correlation function of the end-to-end distance

At any given time t , the distance \mathbf{R} between one end of the chain and the other is given by

$$\mathbf{R}(t) = \mathbf{r}(N, t) - \mathbf{r}(0, t). \quad (31)$$

Introducing the normal mode expansion

$$\mathbf{r}(\tau, t) = \sum_{n=1}^{\infty} \mathbf{X}_n(t) \phi_n(\tau), \quad (32)$$

into this definition, where the $\{\mathbf{X}\}$ are the normal modes, and the $\phi_n(t)$ are given by Eq. (25), one can show that

$$\mathbf{R}(t) = \sum_{n=1}^{\infty} \mathbf{X}_n(t) A_n N^h \left[J_\nu(\gamma_n) - \frac{(\gamma_n/2)^\nu}{\Gamma(\nu+1)} \right]. \quad (33)$$

The second term on the r.h.s. of this equation uses the series representation of the Bessel function²⁰ to define the expansion of $\mathbf{r}(0, t)$. It now follows that:

$$\begin{aligned} \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \langle \mathbf{X}_n(t) \cdot \mathbf{X}_m(0) \rangle A_n A_m N^{2h} \\ &\quad \times \left[J_\nu(\gamma_n) - \frac{(\gamma_n/2)^\nu}{\Gamma(\nu+1)} \right] \\ &\quad \times \left[J_\nu(\gamma_m) - \frac{(\gamma_m/2)^\nu}{\Gamma(\nu+1)} \right]. \end{aligned} \quad (34)$$

The correlation function of the normal modes in the above equation is obtained from the solution of Eq. (22) re-expressed in terms of the variables $\{\mathbf{X}\}$. After several steps of algebra involving various Bessel identities,²⁰ one can show that this equation is

$$\begin{aligned} \frac{\partial X_{n,\alpha}(t)}{\partial t} &= - \frac{3\Gamma^2(h+1/2)\gamma_n^2}{4(1+\nu)^2 \zeta l^{2-2h} N^{2h+1}} X_{n,\alpha}(t) \\ &\quad + \theta_\alpha(t), \quad \alpha = x, y, z, \end{aligned} \quad (35)$$

where

$$\theta_\alpha(t) = \int_0^N d\tau \eta_\alpha(\tau, t) A_n \tau^h J_\nu(\gamma_n(\tau/N)^{h+1/2}). \quad (36)$$

From Eq. (35), the time correlation function of \mathbf{X}_n can be calculated as

$$\langle X_{n,\alpha}(t) X_{m,\beta}(0) \rangle = \langle X_{n,\alpha}(0) X_{m,\beta}(0) \rangle e^{-\mu_n t}, \quad (37)$$

where

$$\begin{aligned} \mu_n &= \frac{3\Gamma^2(h+1/2)\gamma_n^2}{4(1+\nu)^2 \zeta l^{2-2h} N^{2h+1}} \\ &= \frac{3\Gamma^2(h+1/2)\gamma_n^2(h+1/2)^2}{\zeta l^{2-2h} N^{2h+1}}. \end{aligned} \quad (38)$$

The use of the normal mode expansion [Eq. (32)] in the expression for the Hamiltonian H_G [Eq. (20)] reduces the latter to a quadratic form:

$$H_G = \frac{3(h+1/2)^2 \Gamma^2(h+1/2)}{2l^{2-2h} N^{2h+1}} \sum_{\alpha=x,y,z} \sum_{n=1}^{\infty} X_{n,\alpha}^2 \gamma_n^2, \quad (39)$$

so that

$$\langle X_{n,\alpha}(0) X_{m,\beta}(0) \rangle = \delta_{n,m} \delta_{\alpha,\beta} \frac{N^{2h+1} l^{2-2h}}{3(h+1/2)^2 \gamma_n^2 \Gamma^2(h+1/2)}, \quad (40)$$

and, hence, eventually

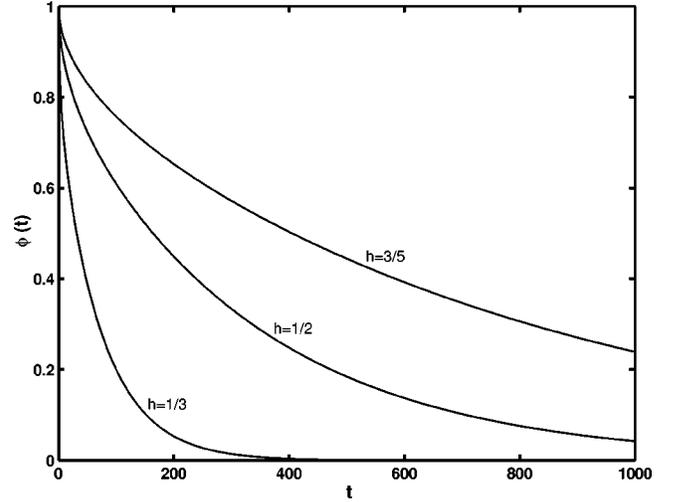


FIG. 1. The normalized time correlation function of the end-to-end distance $\phi(t)$ as calculated from Eqs. (41) and (21) vs time t for three different values of the parameter h (3/5, 1/2, and 1/3) at a fixed chain length of $N = 100$.

$$\begin{aligned} \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle &= \frac{2l^{2-2h} N^{2h}}{(h+1/2)\Gamma^2(h+1/2)} \\ &\quad \times \sum_{n=1}^{\infty} \frac{1}{\gamma_n^2} \left[1 - \frac{(\gamma_n/2)^\nu}{J_\nu(\gamma_n)\Gamma(\nu+1)} \right]^2 \\ &\quad \times \exp\left\{ - \frac{3\gamma_n^2(h+1/2)^2 \Gamma^2(h+1/2)}{l^{2-2h} N^{2h+1} \zeta} t \right\}. \end{aligned} \quad (41)$$

As may be verified, in the limit $h=1/2$, corresponding to simple Brownian motion, $\nu = -1/2$, $\gamma_n = n\pi$, and

$$\langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle = 8Nl \sum_{n:\text{odd}} \frac{1}{n^2 \pi^2} \exp(-3n^2 \pi^2 t / N^2 l \zeta), \quad (42)$$

which is the well-known expression for the decay of the end-to-end correlations for the Rouse chain.²¹

IV. RESULTS AND DISCUSSION

If the time correlation function of the end-to-end distance [Eq. (41)] is normalized by its equilibrium value [Eq. (21)], one obtains the function $\phi(t)$ [Eq. (6)] required for the calculation of τ . The variation of $\phi(t)$ with time t (expressed in units in which the monomer friction coefficient ζ and the Kuhn length l are unity) at fixed $N=100$ is shown in Fig. 1 for the three values of h that model polymers in poor, theta and good solvents, viz., 1/3, 1/2, and 3/5, respectively. For these values of h , the characteristic relaxation times [the time it takes, roughly, for $\phi(t)$ to fall to $1/e$ of its initial value] are significantly different. Given that h is expected to provide a measure of how strongly chain segments attract or repel one another, the differences can be taken to reflect the importance of nonbonded interactions in the rate of conformational decay.

The closure time τ is calculated from Eq. (3) using the above expression for $\phi(t)$ and the expression for $C(t)$ given

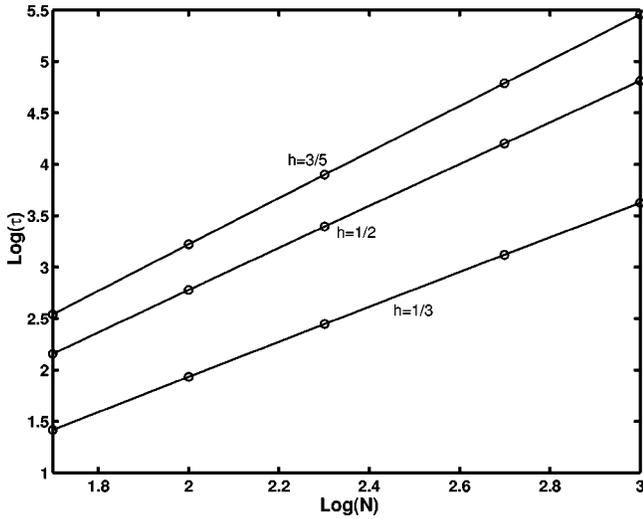


FIG. 2. Log–log plot of the mean first passage time of chain closure τ as calculated from Eq. (3) [using Eq. (11) for $C(t)$ and Eqs. (41) and (21) for $\phi(t)$] vs chain length N for three different values of the parameter h ($3/5$, $1/2$, and $1/3$) at a fixed value of 0.5 for the reaction radius a .

by Eq. (11). The calculations at this stage are done numerically. The results are a function of the following parameters: The chain length N , the reaction radius a , and the parameter h (l and ζ being set to 1.) Figure 2 shows the variation of τ with N on a log–log scale for $h = 3/5$, $1/2$, and $1/3$ at an arbitrary fixed value of 0.5 for the reaction radius a . The curves of τ vs N for these values of h at two other values of a considered in the work of Pastor *et al.*,⁹ viz., 0.1 and 1.0 , are virtually the same, and are not included in the figure for clarity. The three curves in Fig. 2 are all linear for essentially the entire range of N , from about 50 to 1000 , but have slightly different slopes. The time τ is, therefore, a power law in N of the form N^b , with b a function of h that is not strongly dependent on a in the large N regime of interest in our calculations. The slopes b of these various curves have been determined, and are shown in Table I for the given values of h and the different values of a . [Because of sensitivity to numerical errors, results for still smaller values of a ($a < \mathcal{O}(10^{-2})$) do not seem to be especially reliable.] The implications of these results for the cases $h = 1/2$, $3/5$, and $1/3$ are now considered in turn.

TABLE I. Calculated scaling exponents b in the relation $\tau \sim N^b$ for different reaction distances a and different correlation parameters h .

a	h	Exponent
0.1	$1/3$	1.67
	$1/2$	2.02
	$3/5$	2.24
0.5	$1/3$	1.70
	$1/2$	2.04
	$3/5$	2.25
1.0	$1/3$	1.76
	$1/2$	2.09
	$3/5$	2.28

A. Polymers at the theta point ($h = 1/2$)

The case $h = 1/2$ can be taken to correspond to polymers in theta solvents, where the counterbalancing effects of the attractive and repulsive components of the intermolecular potential lead to chain configurations that on average are unperturbed and follow Gaussian statistics. This case has been discussed at length in Refs. 3 and 4, and will not be discussed again in detail here. However, a number of additional observations are worth making.

The calculated values of b in Table I at the indicated values of a are not in perfect agreement with the previous results derived from our model of semiflexible chains in the limit of large flexibility,^{3,4} (although the respective curves of $\phi(t)$ vs t agree nearly quantitatively.) For instance, the present estimate of b for the case $a = 0.1$ is 2.02 , whereas the previous estimate³ is 2.18 . The difference, we believe, can be ascribed to round-off and other errors in the numerical routines that are used in the two calculations. The earlier calculation, which required solutions to highly nonlinear coupled equations, was considerably more involved, and therefore, likely to have been subject to greater uncertainty.

The finding that the closure time τ in these calculations and in the calculations of Refs. 3 and 4 scale roughly as N^2 agree with certain analytic results based on the Wilemski–Fixman (WF) formalism in the limit of large N at fixed reaction radius a under free-draining conditions^{9,11} and with the results of at least some simulations of long flexible polymers under similar conditions.^{8,9,22} Within the WF formalism itself, and provided excluded volume effects are absent and N/a remains large, the N dependence of τ is expected to be altered to $N^{3/2}$ when the polymer becomes nonfree draining through the incorporation of hydrodynamic effects.^{8,23}

But the above $N^{3/2}$ behavior is seen in several other alternative models of cyclization that neglect hydrodynamic interactions altogether. It is seen, for instance, in the model of Brereton and Rusli, in which the terminal monomers of the chain are treated as completely independent Brownian particles whose equilibrium interparticle separation follows the same Gaussian distribution they would have obeyed had they been joined by N other segments.²⁴ It is also seen in the harmonic spring approximation to the WF model,^{8,11} and in the Gaussian chain model of Flory *et al.*²⁵ And more recently, it is the predicted behavior in the model of Szabo, Schulten, and Schulten²⁶ (SSS), in which chain dynamics is assumed to be governed by the Smoluchowski equation for diffusion on a one-dimensional potential of mean force.

The SSS model has been widely invoked to rationalize experimental data on loop formation in small polypeptides of controlled length. The experiments typically monitor the rate at which an excited state species at one end of the chain is quenched by contact with a suitable molecular species at the other. In one such experiment, by Bieri *et al.*,²⁷ the closure rate κ for chains of between 1 and 4 repeat units of the glycine-serine dimer was found to scale as $N^{-1.36 \pm 0.26}$, in apparent agreement—as claimed—with the SSS prediction. However, in the study by Hudgins *et al.*²⁸ of a series of polypeptides of the same dimer with up to 10 repeat units, κ was found to scale as $N^{-1.05 \pm 0.06}$, a result less consistent

with the SSS model. Agreement with the SSS model was also claimed in experiments by Lapidus *et al.*²⁹ on 10–20 residue polypeptides of the alanine-glycine-glutamine trimer, where κ was found to vary as $N^{-3/2}$ in the asymptotic limit of large N , although it was found to be much less N -dependent for shorter chains, possibly reflecting the inapplicability of Gaussian statistics to stiff polymers. A later experimental study by the same group,³⁰ backed by molecular dynamics simulations using a simplified potential based on Ramachandran maps, accounted explicitly for previously neglected distance dependence effects of the quenching rate, and confirmed the $N^{-3/2}$ dependence of κ , but only if the diffusion coefficient used in the SSS formalism was made roughly tenfold smaller than expected of free diffusion between the terminal monomers.

Portman³¹ has recently argued that the SSS mean first passage time is actually a lower bound on the true closure time, and would therefore require a smaller-than-expected diffusion coefficient to fit experimental or simulation data, essentially to correct for discrepancies that result from using Gaussian statistics to describe an intrinsically non-Gaussian process. Jun *et al.*³² have used Kramers rate theory to suggest that the N -dependence of κ actually reflects the competition between energetic and entropic components of chain closure, the former being dominant for short chains and the latter for long chains. Much the same kind of argument had been advanced in Ref. 9, where different limits (small a at fixed N , on the one hand, or large N at fixed a on the other) were said to account for different N dependences of κ . However, the calculations of Jun *et al.*³² do not appear to produce $N^{-3/2}$ scaling under conditions expected to be relevant to real systems, but they do predict a minimum in the closure time at a distance of about 3 and 4 persistence lengths, as in the work of Refs. 28 and 30. At the same time, no such extremum is evident in another recent simulation, by Wang and Makarov;³³ and although $\kappa \sim N^{-1.65}$ at large N in this simulation, in fair agreement with the SSS prediction, the diffusion coefficient that fits the data is actually close to the free diffusion value.

As these results indicate, the microscopic basis for the distance dependence of closure times cannot always be identified unambiguously, and other more stringent tests may be required to discriminate between rival mechanistic scenarios.

B. Polymers in good solvents ($h=3/5$)

The case $h=3/5$ can be taken to correspond to polymers in good solvents, where repulsive interactions between monomer segments dominate, and the average radial dimensions of the chain are those of a self-avoiding walk. The prediction that $\tau \sim N^{2.2}$ (Table I) is consistent with numerous prior studies, including the independent particle model of Brereton and Rusli,²⁴ the renormalization group calculations of Friedman and O’Shaughnessy,³⁴ Brownian dynamics simulations³⁵ and at least some experimental data.³⁶ It is probably reasonable to conclude therefore that the enhanced N dependence of τ (or κ) (as compared to the Gaussian model) originates wholly in the effects of long-range excluded volume effects between different parts of the chain.

C. Collapsed polymers ($h=1/3$)

The case $h=1/3$ can be taken to correspond to collapsed polymers, where attractive interactions between monomer segments dominate, and the average end-to-end distance of the chain scales as $N^{1/3}$. At present there appear to be no experimental or simulation data on cyclization in such systems, so our prediction (Table I) that their average closure time scales approximately as $N^{-1.7}$ must remain untested.

The average radial dimensions of a number of compact globular proteins in their native state can also be said¹⁵ to scale as $N^{1/3}$, but a direct comparison of the $h=1/3$ results with these systems is not necessarily warranted. Native proteins when unfolded adopt configurations that depend strongly on how they have been denatured; high temperatures, high pressures and chemical denaturants all tend to produce open configurations that are either random coils or possibly self-avoiding walks.³⁷ It is only under “cold denaturing” conditions that these configurations are thought to be essentially collapsed, with the size scaling as $N^{1/3}$. The polypeptide data of Sec. IV A above, are therefore, probably best compared to results for the case $h=1/2$, as has been done, but the results for the case $h=1/3$ indicate that the presence of nonbonded attractive interactions can also contribute to decreasing the N -dependence of loop formation rates relative to flexible polymers in the free-draining limit.

ACKNOWLEDGMENT

P.D. is grateful to the Council of Scientific and Industrial Research, Government of India, for financial support.

- ¹D. Thirumalai, *J. Phys. Chem. B* **103**, 608 (1999); Z. Guo and D. Thirumalai, *Biopolymers* **36**, 83 (1995).
- ²P. Flory, *Statistical Mechanics of Chain Molecules* (Hanser, Munich, 1989); D. Poland and H. A. Scheraga, *Theory of Helix-Coil Transitions in Biopolymers* (Academic, New York, 1970).
- ³A. Dua and B. J. Cherayil, *J. Chem. Phys.* **116**, 399 (2002).
- ⁴A. Dua and B. J. Cherayil, *J. Chem. Phys.* **117**, 7765 (2002).
- ⁵W. J. Wedemeyer and H. A. Scheraga, *Encyclopedia of Life Sciences* **15**, 334 (2002).
- ⁶S. F. Edwards, *Proc. Phys. Soc. London* **85**, 613 (1965).
- ⁷B. J. Cherayil and P. Biswas, *J. Chem. Phys.* **99**, 9230 (1993).
- ⁸G. Wilemski and M. Fixman, *J. Chem. Phys.* **58**, 4009 (1973); **60**, 866 (1974); **60**, 878 (1974).
- ⁹R. W. Pastor, R. Zwanzig, and A. Szabo, *J. Chem. Phys.* **105**, 3878 (1996).
- ¹⁰P. Debnath and B. J. Cherayil (unpublished).
- ¹¹M. Doi, *Chem. Phys.* **9**, 455 (1975).
- ¹²K. F. Freed, *Adv. Chem. Phys.* **22**, 1 (1972); K. F. Freed, *Renormalization Group Theory of Macromolecules* (Wiley, N.Y., 1987).
- ¹³P. J. Flory, *J. Chem. Phys.* **17**, 303 (1949); P. J. Flory and T. G. Fox, *J. Am. Chem. Soc.* **73**, 1904 (1951); M. E. Fisher, *J. Phys. Soc. Jpn.* **26**, 44 (1968).
- ¹⁴P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979).
- ¹⁵T. G. Dewey, *J. Chem. Phys.* **98**, 2250 (1993).
- ¹⁶P. Biswas and B. J. Cherayil, *J. Chem. Phys.* **100**, 4665 (1994).
- ¹⁷P. Biswas and B. J. Cherayil, *J. Phys. Chem.* **99**, 816 (1995).
- ¹⁸P. Biswas, A. Paramekanti, and B. J. Cherayil, *J. Chem. Phys.* **104**, 3360 (1996).
- ¹⁹C. Maccone, *Nuovo Cimento* **B61**, 229 (1981); **B65**, 259 (1981).
- ²⁰I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, San Diego, 1980).
- ²¹M. Doi and S. F. Edwards, *Theory of Polymer Dynamics* (Clarendon, Oxford, 1986).
- ²²M. Sakata and M. Doi, *Polym. J.* **8**, 409 (1976).

- ²³A. Perico and C. Cuniberti, *J. Polym. Sci., Polym. Phys. Ed.* **15**, 1435 (1977).
- ²⁴M. G. Brereton and A. Rusli, *Polymer* **17**, 395 (1976).
- ²⁵P. J. Flory, U. W. Suter, and M. Mutter, *J. Am. Chem. Soc.* **98**, 5740 (1976); U. W. Suter, M. Mutter, and P. J. Flory, *ibid.* **98**, 5745 (1976).
- ²⁶A. Szabo, K. Schulten, and Z. Schulten, *J. Chem. Phys.* **72**, 4350 (1980).
- ²⁷O. Bieri, J. Wirz, B. Hollrung, M. Schutkowski, M. Drewello, and T. Keifhaber, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9597 (1999).
- ²⁸R. R. Hudgins, F. Huang, G. Gramlich, and W. M. Nau, *J. Am. Chem. Soc.* **124**, 556 (2002).
- ²⁹L. J. Lapidus, W. A. Eaton, and J. Hofrichter, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7720 (2000).
- ³⁰L. J. Lapidus, P. J. Steinbach, W. A. Eaton, A. Szabo, and J. Hofrichter, *J. Phys. Chem. B* **106**, 11628 (2002); M. Buscaglia, B. Schuler, L. Lapidus, W. A. Eaton, and J. Hofrichter, *J. Mol. Biol.* **332**, 9 (2003).
- ³¹J. J. Portman, *J. Chem. Phys.* **118**, 2381 (2003).
- ³²S. Jun, J. Bechhoefer, and B.-Y. Ha, *Europhys. Lett.* (to be published).
- ³³Z. Wang and D. E. Makarov, *J. Chem. Phys.* **117**, 4591 (2002).
- ³⁴B. Friedman and B. O'Shaughnessy, *Phys. Rev. A* **40**, 5950 (1989); B. Friedman and B. O'Shaughnessy, *Macromolecules* **26**, 4888 (1993).
- ³⁵A. Rey and J. J. Freire, *Macromolecules* **24**, 4673 (1991); A. Podtelezhnikov and A. Vologodskii, *ibid.* **30**, 6668 (1997).
- ³⁶W. A. Winnik, in *Cyclic Polymers*, edited by J. A. Semlyen (Elsevier, New York, 1986).
- ³⁷I. S. Millet, S. Doniach, and K. W. Plaxco, *Adv. Protein Chem.* **62**, 241 (2002).