

Semiflexible random $A-B$ block copolymers under tension

Pallavi Debnath and Binny J. Cherayil*

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

In this paper we explore the conformational properties of random $A-B$ block copolymers with semiflexible segments in a constant force field as a model of the behavior of biopolymers under tension. The degree of semiflexibility of individual blocks, which is characterized by a persistence length l , may range from values that correspond to complete flexibility of the block to values that correspond to nearly complete rigidity. The distribution of blocks is described by a two-state Markov process that generates the statistics governing the process of living polymerization in the steady state. Force-extension curves for this model of the polymer are calculated as an average over the chains conformations (which are described by a finitely extensible Gaussian model), as well as over both quenched and annealed distributions of the sequence of A and B blocks along the chain backbone. The results are sensitive to the nature of the sequence distribution: in the annealed case, the force-extension curves are sigmoidal for essentially all values of the parameters that characterize the conformational properties of the A and B blocks and their statistical correlations, while in the quenched case, the curves exhibit plateau regions that are reminiscent of various kinds of transitions in proteins and DNA.

I. INTRODUCTION

Experimental techniques to apply carefully calibrated forces to isolated molecules have now made it possible to explore *in vitro* many of the mechanical events that in living systems govern the pathways of different biochemical processes.^{1,2} The applied forces can vary in size from a few piconewtons to several hundreds of piconewtons, and can induce several different kinds of response, ranging from the simple disentanglement associated with the loss of entropic freedom to more complex transitions associated with the appearance of morphologically distinct conformational phases.

Many of these responses can be understood in terms of simple models of the macromolecule,³ especially if the mutual interactions between a small number of distinct structural elements are the principal controlling factors behind the observed behavior.⁴ It is often adequate under such circumstances to represent the molecule as a random connected sequence of two or more pointlike residues with different strengths of interaction. These so-called random $A-B$ heteropolymer models have been very successful in identifying many of the broad features of the folding pathways in biopolymers like proteins and polypeptides.^{5,6} But, they also ignore many of the structural motifs that are known to influence certain kinds of conformational change in these molecules. α -helices and β -sheets are just two such motifs, and they can seldom be modeled successfully within the existing formulations of $A-B$ heteropolymers. However, somewhat more faithful representations of biopolymers with secondary structural elements can be developed along the same lines starting from random *block copolymer* models. Block copolymers, by virtue of containing subunits that are themselves polymeric, can be made to mimic some of these sec-

ondary structural elements by suitable choice of the parameters that control the conformational properties of the individual blocks.

This idea forms the basis for the model developed recently by Buhot and Halperin⁷ to study the response of biopolymers to stretching. Here, the polymer is represented as a linked chain of rods and coils arranged at random along the backbone. Two distinct sequence distributions are considered: one where the randomness is quenched, and the other where it is annealed. The force-extension profiles for these two scenarios are found to be quite different; the quenched case shows a pronounced change in slope, while the annealed case is sigmoidal. The predictions are in qualitative agreement with certain experimental results, but they are largely derived from phenomenological considerations, and so are not easily justified microscopically.

We have recently formulated a similar model of biopolymers.⁸ In our approach, the macromolecule is represented as a random linked sequence of two different kinds of “prepolymer,” A or B , which may span a range of different degrees of flexibility, from completely flexible to almost completely stiff. Each prepolymer, itself a chain of identical residues, is modeled as an inextensible curve using the path integral formalism of Saitô, Takahashi, and Yunoki.⁹ Within this approach, and in the absence of excluded volume interactions and stretching forces, the mean size of the chain, averaged over the sequence of A and B prepolymers using a two-state Markov process,¹⁰ can be determined exactly as a function of the overall length of the heteropolymer, the relative amounts of A to B components in the chain, and the degree of stiffness of these components.

This model is based on a well-defined microscopic Hamiltonian that is easily generalized, *formally*, to the case where a nonzero force \mathbf{f} acts at either end of the polymer. Unfortunately, the generalization no longer admits of an ex-

*Electronic mail: cherayil@ipc.iisc.ernet.in

act analytical solution under the original constraint of finite extensibility that must be imposed on the Hamiltonian. As we show in this paper, an exact solution of the problem can still be derived if this constraint is relaxed.

The general features of the model are introduced in the following section. The model is then applied to the calculation of the mean square end-to-end distance of a random block copolymer that is acted on by a force. The calculation involves an average over the conformational degrees of freedom of the chain as well as over the sequence distribution of A and B prepolymer segments. The sequence average is carried out for both quenched and annealed randomness using the same two-state Markov process mentioned above, which was introduced by Fredrickson, Milner, and Leibler in their study of microphase ordering in block copolymers melts.¹⁰ The calculation of the average size of the chain is completed by enforcing the requirement of inextensibility. This is done approximately (but self-consistently) using a conformation-dependent spring constant to connect near-neighbor monomers along the chain. The results of the calculation are presented in Sec. III, along with a discussion of their implications.

II. THEORY

A. Definitions

In these calculations, the heteropolymer is modeled as a random block copolymer of n “prepolymers” (to use the terminology of Fredrickson, Milner, and Leibler¹⁰) of contour length N . Each prepolymer is a semiflexible homopolymer made up of either A or B residues that is characterized by a persistence length l_A or l_B , and that is represented as the continuum limit of a discrete set of coupled harmonic oscillators with bond strengths b_A or b_B . In this continuum representation,¹¹ the prepolymer is the locus of a curve described by the set of distances $\mathbf{r}(\tau)$. These distances are the vectorial locations of the monomers that are imagined to lie at the points τ defining the backbone of the chain. The total contour length of the heteropolymer, nN , is denoted M , so, in general, the parameter τ varies from 0 to M . A constant force \mathbf{f} is assumed to act at either end of the chain. In units where the thermal energy $k_B T$ is 1, the Hamiltonian of this heteropolymer is given by

$$H = b_A \sum_{i=1}^n \delta_{\theta_i, 1} \int_{(i-1)N}^{iN} d\tau \dot{\mathbf{r}}(\tau)^2 + b_B \sum_{i=1}^n \delta_{\theta_i, -1} \int_{(i-1)N}^{iN} d\tau \dot{\mathbf{r}}(\tau)^2 - \mathbf{f} \cdot \sum_{i=1}^n \mathbf{R}_i. \quad (1)$$

Here, θ_i is a dichotomous random variable that is defined to be $+1$ if the i th prepolymer is made up of A residues, and that is defined to be -1 if the i th prepolymer is made up of B residues. The symbol $\dot{\mathbf{r}}(\tau)$ stands for the derivative $d\mathbf{r}(\tau)/d\tau$, while \mathbf{R}_i defines the end-to-end distance of the i th prepolymer, and is given by

$$\mathbf{R}_i = \int_{(i-1)N}^{iN} d\tau \dot{\mathbf{r}}(\tau) \quad (2)$$

$$= \mathbf{r}(iN) - \mathbf{r}((i-1)N) \quad (3)$$

$$\equiv \mathbf{r}_i - \mathbf{r}_{i-1}. \quad (4)$$

As mentioned in the Introduction, this model of the heteropolymer is different from our earlier model,⁸ which was based on the approach of Saitô, Takahashi, and Yunoki⁹ (STY), and which used the set of *tangent* vectors $\mathbf{u}(\tau) = d\mathbf{r}(\tau)/d\tau$ at the points τ , and not the position vectors $\mathbf{r}(\tau)$ themselves, to define the trajectory of the continuum chain. Although the STY approach can be implemented exactly to calculate the mean square end-to-end distance of the random heteropolymer when $\mathbf{f}=0$, the corresponding calculations when $\mathbf{f} \neq 0$, are highly nontrivial. The present model, defined by Eq. (1), is an attempt to retain the basic structure of the earlier approach without rendering it analytically intractable. In particular, we shall treat the constraint of inextensibility, expressed earlier by the requirement that $|\mathbf{u}(\tau)|=1$ everywhere, in an alternative self-consistent approximation to be discussed later.

If the random block copolymer defined above contained just a single *flexible* block, of, say, type A , the bond strength b_A would be $3/2l_A$, and the Hamiltonian of Eq. (1) would reduce to the well-known Hamiltonian of the continuum Gaussian chain.¹¹

It proves convenient to eliminate the Kronecker deltas in Eq. (1), and to rewrite H identically as

$$H = \sum_{i=1}^n \Delta_i \int_{(i-1)N}^{iN} d\tau \dot{\mathbf{r}}(\tau)^2 - \mathbf{f} \cdot \sum_{i=1}^n \mathbf{R}_i, \quad (5)$$

where

$$\Delta_i \equiv \frac{1}{2}(b_A + b_B) + \frac{1}{2}(b_A - b_B)\theta_i, \quad (6)$$

$$\equiv D_1 + D_2\theta_i. \quad (7)$$

This is the form of the Hamiltonian that will be used in the subsequent calculations.

B. End-to-end distance

A useful measure of the radial dimensions of the polymer subject to the force \mathbf{f} is its mean square end-to-end distance, defined as

$$\begin{aligned} \langle\langle \mathbf{R}^2 \rangle\rangle_s &= \sum_{i=1}^n \sum_{j=1}^n \langle\langle \mathbf{R}_i \cdot \mathbf{R}_j \rangle\rangle_s \\ &= \sum_{i=1}^n \langle\langle \mathbf{R}_i^2 \rangle\rangle_s + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle\langle \mathbf{R}_i \cdot \mathbf{R}_j \rangle\rangle_s, \end{aligned} \quad (8)$$

where the outer brackets, with subscript s , denote the average over the sequence of A s and B s along the backbone, which may either be quenched (subscript q) or annealed (subscript a), while the inner brackets, without a subscript, denote the average over the conformational degrees of freedom alone. These distinct averages are defined, respectively, by

$$\begin{aligned} \langle\langle X \rangle\rangle_q &\equiv \sum_{\{\theta_i\}} p(\{\theta_i\}) \frac{1}{Q(\{\theta_i\})} \int \mathcal{D}[\mathbf{r}(\tau)] X \\ &\quad \times \exp(-H[\{\theta_i\}, \mathbf{r}(\tau)]), \end{aligned} \quad (9)$$

$$\langle\langle X \rangle\rangle_a \equiv \frac{\sum_{\{\theta_i\}} p(\{\theta_i\}) \int \mathcal{D}[\mathbf{r}(\tau)] X \exp(-H[\{\theta_i\}, \mathbf{r}(\tau)])}{\sum_{\{\theta_i\}} p(\{\theta_i\}) Q(\{\theta_i\})}. \quad (10)$$

In these two equations, X is any sequence and conformation-dependent property of the chain, $p(\{\theta_i\})$ is the probability that the first prepolymer is of type θ_1 , the second is of type θ_2 , and so on, and $Q(\{\theta_i\})$ is the chain partition function, defined as

$$Q(\{\theta_i\}) = \int \mathcal{D}[\mathbf{r}(\tau)] \exp(-H[\{\theta_i\}, \mathbf{r}(\tau)]). \quad (11)$$

The symbol $\int \mathcal{D}[\mathbf{r}(\tau)]$ stands for the functional integral over the continuous conformations of the chain, while the sum $\sum_{\{\theta_i\}}$ stands for the sum over the values of $\theta_1, \theta_2, \dots, \theta_n$.

The functional integrals in Eqs. (9) and (10) that define the average over chain conformations can be rewritten in terms of ordinary integrals over a product of chain ‘‘propagators’’ according to the general prescriptions outlined in, for instance, Ref. 11. In this way, one can show that

$$\begin{aligned} & \int \mathcal{D}[\mathbf{r}(\tau)] \mathbf{R}_i \cdot \mathbf{R}_j e^{-H} \\ &= \int d\mathbf{R}_1 \int d\mathbf{R}_2 \cdots \int d\mathbf{R}_i \cdots \int d\mathbf{R}_j \cdots \int d\mathbf{R}_{n-1} \\ & \quad \times \int d\mathbf{R}_n \mathbf{R}_i \cdot \mathbf{R}_j G_f(\mathbf{R}_1, N) \cdots G_f(\mathbf{R}_i, N) \cdots \\ & \quad \times G_f(\mathbf{R}_j, N) \cdots G_f(\mathbf{R}_{n-1}, N) G_f(\mathbf{R}_n, N), \end{aligned} \quad (12)$$

where

$$\begin{aligned} G_f(\mathbf{R}_k, N) &\equiv \int_{\mathbf{r}((k-1)N)=\mathbf{r}_{k-1}}^{\mathbf{r}(kN)=\mathbf{r}_k} \mathcal{D}[\mathbf{r}(\tau)] \\ & \quad \times \exp\left(-\Delta_k \int_{(k-1)N}^{kN} d\tau \dot{\mathbf{r}}(\tau)^2 + \mathbf{f} \cdot \mathbf{R}_k\right) \\ &= \left(\frac{\Delta_k}{\pi N}\right)^{3/2} \exp(-\Delta_k \mathbf{R}_k^2 / N + \mathbf{f} \cdot \mathbf{R}_k). \end{aligned} \quad (13)$$

In the same way, the partition function Q can be reduced to the following product of ordinary integrals:

$$Q(\{\theta_i\}) = \prod_k \int d\mathbf{R}_k G_f(\mathbf{R}_k, N). \quad (15)$$

From the definition of the propagator in Eq. (14), this expression for the partition function is easily evaluated as

$$Q = \exp\left[f^2 N^2 \sum_{k=1}^n 1/4\Delta_k\right]. \quad (16)$$

The simple Gaussian integrals in Eq. (12) are also easily carried out to produce

$$\int \mathcal{D}[\mathbf{r}(\tau)] \mathbf{R}_i^2 e^{-H} = Q \Theta_1, \quad (17)$$

$$\int \mathcal{D}[\mathbf{r}(\tau)] \mathbf{R}_i \cdot \mathbf{R}_j e^{-H} = Q \Theta_2, \quad i \neq j \quad (18)$$

where

$$\Theta_1 = \left(\frac{3N}{2\Delta_i} + \frac{f^2 N^2}{4\Delta_i^2}\right), \quad (19)$$

$$\Theta_2 = \frac{f^2 N^2}{4\Delta_i \Delta_j} \quad i \neq j. \quad (20)$$

From these equations, one sees that

$$\langle\langle \mathbf{R}^2 \rangle\rangle_q = \sum_{i=1}^n \langle \Theta_1 \rangle_s + 2 \sum_{i=1}^n \sum_{j \neq i}^n \langle \Theta_2 \rangle_s, \quad (21)$$

$$\langle\langle \mathbf{R}^2 \rangle\rangle_a = \frac{1}{\langle Q \rangle_s} \left[\sum_{i=1}^n \langle Q \Theta_1 \rangle_s + 2 \sum_{i=1}^n \sum_{j \neq i}^n \langle Q \Theta_2 \rangle_s \right]. \quad (22)$$

The sequence averages in Eqs. (21) and (22) are calculated in the next section.

C. Sequence average

The average over the distribution of A and B prepolymers can be carried out once the function $p(\{\theta_i\})$ is specified. If one makes the Markovian assumption that the identity of the i th prepolymer in the chain is determined only by the identity of the preceding prepolymer, then, as Fredrickson, Milner, and Leibler (FML) have shown,¹⁰ the probability of realizing a given sequence of A s and B s is determined, in general, by a set of four conditional probabilities p_{KL} , $K, L = A, B$, where p_{KL} is the conditional probability of observing K given L . If one further assumes that this sequence results from the process of living polymerization under steady-state conditions, then the p_{KL} can be related to the mole fractions φ and $1 - \varphi$ that define the composition of A and B in the chain so generated. When combined with the Markov condition, these two assumptions imply that only *two* parameters are needed to fix the overall sequence distribution in the chain: one is φ itself, and the other may be chosen to be the nontrivial eigenvalue λ of the matrix of conditional probabilities. The conditional probabilities p_{KL} can now be expressed in terms of these parameters as

$$p_{AA} = \varphi(1 - \lambda) + \lambda, \quad (23)$$

$$p_{BB} = \varphi(\lambda - 1) + 1, \quad (24)$$

$$p_{AB} = 1 - p_{BB}, \quad (25)$$

$$p_{BA} = 1 - p_{AA}. \quad (26)$$

The parameter λ , which lies between -1 and $+1$, controls the degree of correlation between A and B prepolymers in the chain; when $\lambda \rightarrow -1$, the A and B prepolymers tend to succeed each other in alternation, when $\lambda \rightarrow +1$, the A s tend to succeed A s, and B s tend to succeed B s, and, finally, when $\lambda = 0$, A s and B s follow each other entirely randomly.

From these considerations, one can show that

$$\sum_{\{\theta_i\}} p(\{\theta_i\}) \theta_i = 2\varphi - 1, \quad (27)$$

$$\sum_{\{\theta_i\}} p(\{\theta_i\}) \theta_l \theta_m = (2\varphi - 1)^2 + 4\varphi(1 - \varphi)\lambda^{|l-m|}. \quad (28)$$

Results for higher moments of the variable θ_i are also easily determined.

1. Quenched disorder

For this case, referring to Eq. (21), and using Eq. (27), one may verify that

$$\sum_{i=1}^n \langle \Theta_1 \rangle_s = n[\varphi \langle \mathbf{R}_A^2 \rangle + (1 - \varphi) \langle \mathbf{R}_B^2 \rangle], \quad (29)$$

where $\langle \mathbf{R}_A^2 \rangle$ and $\langle \mathbf{R}_B^2 \rangle$ are the mean-square end-to-end distances of the i th prepolymer averaged over the conformations of the remaining $n - 1$ prepolymers when the i th prepolymer is of type A and when it is of type B , respectively. From the definitions of quenched and annealed averages in Eqs. (9) and (10), one can show that when a given prepolymer is of a *definite* type (A or B), the sequence average over the remainder of the chain yields the same factor of unity for both the quenched and annealed operations. Thus, $\langle \mathbf{R}_A^2 \rangle$ and $\langle \mathbf{R}_B^2 \rangle$ actually correspond to the dimensions of a homopolymer of length N as calculated from the Hamiltonian of Eq. (5) in which $n = 1$ and the variable θ_i in Eq. (6) is assigned the value $+1$ (for a chain of A residues) or the value -1 (for a chain of B residues.)

Returning to Eq. (21), we have, similarly

$$\sum_{i=1}^n \sum_{j \neq i}^n \langle \Theta_2 \rangle_s = \frac{f^2 N^2}{4(D_1^2 - D_2^2)^2} [\{D_1 - (2\varphi - 1)D_2\}^2 S_1 + 4\varphi(1 - \varphi)D_2^2 S_2], \quad (30)$$

with S_1 and S_2 given by

$$S_1 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n = \frac{1}{2} n(n-1), \quad (31)$$

$$S_2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \lambda^{|i-j|} = \frac{\lambda}{1-\lambda} \left[n - \frac{1-\lambda^n}{1-\lambda} \right]. \quad (32)$$

2. Annealed disorder

This case is actually harder to treat than the case of quenched disorder, in contrast to most other related calculations in condensed matter. Fortunately, it still admits an exact result. Referring now to Eq. (22), the first sum on the right-hand side can be re-expressed as

$$\frac{1}{\langle Q \rangle_s} \sum_{i=1}^n \langle Q \Theta_1 \rangle_s = \frac{3}{2} n \alpha N + \frac{1}{4} n f^2 N^2 (\alpha^2 + \beta^2) - \frac{3}{2\Xi} N \beta \left(1 + \frac{1}{3} \alpha f^2 N \right) T_1, \quad (33)$$

where

$$\alpha = \frac{D_1}{D_1^2 - D_2^2}, \quad \beta = \frac{D_2}{D_1^2 - D_2^2}, \quad X = \frac{1}{4} \beta f N^2, \quad (34)$$

and

$$\Xi = \left\langle \exp \left(-X \sum_{k=1}^n \theta_k \right) \right\rangle_s, \quad (35)$$

$$T_1 = \sum_{i=1}^n \left\langle \theta_i \exp \left(-X \sum_{k=1}^n \theta_k \right) \right\rangle_s. \quad (36)$$

Similarly, the second sum on the right-hand side of Eq. (22) can be written in the form

$$\frac{1}{\langle Q \rangle_s} \sum_{i=1}^n \sum_{j \neq i}^n \langle Q \Theta_2 \rangle_s = \frac{1}{8} n(n-1) \alpha^2 f^2 N^2 - \frac{1}{4\Xi} f^2 N^2 \times [\alpha \beta (T_2 + T_3) + \beta^2 T_4], \quad (37)$$

where

$$T_2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\langle \theta_i \exp \left(-X \sum_{k=1}^n \theta_k \right) \right\rangle_s, \quad (38)$$

$$T_3 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\langle \theta_j \exp \left(-X \sum_{k=1}^n \theta_k \right) \right\rangle_s, \quad (39)$$

$$T_4 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\langle \theta_i \theta_j \exp \left(-X \sum_{k=1}^n \theta_k \right) \right\rangle_s. \quad (40)$$

The sequence averages in the expressions Ξ , T_1 , T_2 , T_3 , and T_4 as well as the associated summations are easily calculated following the matrix methods described in detail in our earlier paper, Ref. 8, which we do not repeat here. A summary of the principal results of these calculations (which are extremely lengthy) is provided in the Appendix. Graphs of the fractional extension of the heteropolymer as a function of different chain parameters are presented later, after a discussion, in the next section, of the treatment of inextensibility.

D. Inextensibility

In polymer models where the bonds between adjacent residues act effectively like Hookean springs, the chain can be stretched indefinitely when a constant force is applied to either end. This produces, among other results, an unphysical molecular weight dependence of the mean-square end-to-end distance at large values of the force,^{12,13} and seriously limits the model's ability to treat problems in which the polymer is not unperturbed. However, without invoking the constraint⁹ $|\mathbf{u}(\tau)| = 1$, finite extensibility at all values of the force can be ensured in Gaussian chains by making the spring depend nonlinearly on the applied force in such a way that an effectively infinite force is required to stretch the chain beyond its contour length. This is the principle used in the so-called finitely extensible nonlinear elastic (FENE) models of a polymer,¹⁴ and it is used here, in modified form to achieve the same ends. The idea is to replace the spring constants b_A or b_B in the present model by the following spring constant:¹²

$$b_i = \frac{3}{2l_i} \frac{1 - \langle R_i^2 \rangle_0 / \langle R_i^2 \rangle_m}{1 - \langle R_i^2 \rangle / \langle R_i^2 \rangle_m}, \quad i = A, B. \quad (41)$$

Here, $\langle R_i^2 \rangle_0$, $\langle R_i^2 \rangle_m$, and $\langle R_i^2 \rangle$ denote the mean-square end-to-end distances of the i th prepolymer when it is, respectively, unperturbed, fully stretched, and under the action of a constant force \mathbf{f} . The rationale behind this definition of b_i is the following: when $f \rightarrow 0$, $\langle R_i^2 \rangle \rightarrow \langle R_i^2 \rangle_0$, so $b_i \rightarrow 3/2l_i$, which is the ‘‘spring constant’’ for the Gaussian chain, whereas when $f \rightarrow \infty$, $\langle R_i^2 \rangle \rightarrow \langle R_i^2 \rangle_m$, so $b_i \rightarrow \infty$, implying that an infinite force is required to extend the chain to its full contour length. In this way, the chain segment cannot be unphysically extended beyond N .

Equation (41) is now solved self-consistently for $\langle R_i^2 \rangle$ by calculating $\langle R_i^2 \rangle$ from the following Hamiltonian:

$$H_i = b_i \int_{(i-1)N}^{iN} d\tau \dot{\mathbf{r}}(\tau)^2 + \mathbf{f} \cdot \mathbf{R}_i. \quad (42)$$

The result is

$$\langle R_i^2 \rangle = \frac{3N}{2b_i} [1 + f^2 N / 6b_i]. \quad (43)$$

Substituting the expression for b_i from Eq. (41) into Eq. (43), and introducing the definitions $\alpha_i \equiv \langle R_i^2 \rangle_0 / \langle R_i^2 \rangle_m = Nl_i / N^2 = l_i / N$ and $z_i = \langle R_i^2 \rangle / \langle R_i^2 \rangle_m = \langle R_i^2 \rangle / N^2$, one finds that the mean-square end-to-end distance of the i th prepolymer under the force \mathbf{f} is obtained from the solution to the following quadratic equation in z_i :

$$z_i = \frac{\alpha_i}{1 - \alpha_i} (1 - z_i) \left[1 + \frac{\alpha_i}{9(1 - \alpha_i)} f^2 N^2 (1 - z_i) \right], \quad i = A, B. \quad (44)$$

From this equation, it is easily shown that

$$\sqrt{z_i} = \frac{\langle R_i^2 \rangle^{1/2}}{N} = \left[1 + \frac{9(1 - \alpha_i)}{2\alpha_i^2 f^2 N^2} \times \left\{ 1 - \sqrt{1 + \frac{4}{9} \alpha_i^2 f^2 N^2} \right\} \right]^{1/2}. \quad (45)$$

The above expression for z_i is now used in Eq. (41) to determine the spring constant b_i

$$b_i = \frac{l_i f^2}{3[\sqrt{1 + 4l_i^2 f^2 / 9} - 1]}, \quad i = A, B. \quad (46)$$

As is easily confirmed, the above result produces $b_i \rightarrow 3/2l_i$ as $f \rightarrow 0$, while $b_i \rightarrow \infty$ as $f \rightarrow \infty$.

III. RESULTS AND DISCUSSION

A. Homopolymer case

Before proceeding to a discussion of the behavior of the random heteropolymer under the action of the force \mathbf{f} , we first consider the results for the case of a single homopolymer to test the utility of the FENE approach against literature data. As it happens, these results need not be derived separately; they are already contained in Eq. (45) for the fractional extension of the i th prepolymer, which is nothing but a

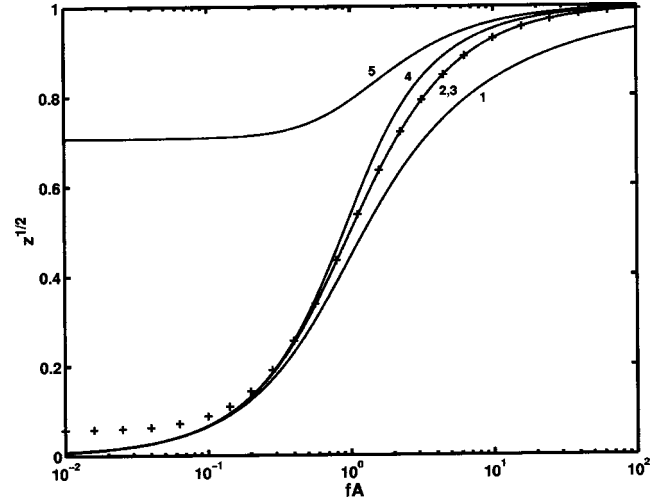


FIG. 1. Fractional extension \sqrt{z} of a homopolymer as a function of the dimensionless force $fA \equiv fl/2$ for different chain models. Curves 1 and 4 correspond to the Marko–Siggia and freely jointed chain models, respectively. Curves 2 and 5 and the curve defined by the symbol + (curve 3) are calculated from Eq. (32) with, respectively, $l/N = 3.24 \times 10^{-5}$, 0.5, and 3.24×10^{-3} .

homopolymer of contour length N . N is essentially arbitrary, and it is useful to let it be large, while l_i is kept fixed. In this limit, Eq. (45) simplifies to

$$\frac{\langle R^2 \rangle^{1/2}}{N} = \left[1 + \frac{9}{2f^2 l^2} \left\{ 1 - \sqrt{1 + \frac{4}{9} f^2 l^2} \right\} \right]^{1/2}, \quad (47)$$

where we have omitted the subscript i , since it is no longer necessary to distinguish between the A and B prepolymers. The above expression has two further interesting analytic limits. When $f \rightarrow 0$, the fractional extension becomes

$$\frac{\langle R^2 \rangle^{1/2}}{N} \sim \frac{1}{3} fl. \quad (48)$$

This result is identical to the small- f prediction of a highly successful model of the elastic response of DNA to an applied force by Marko and Siggia¹⁵ using an analytic interpolation formula derived from a Kratky–Porod representation of a semiflexible chain.

The second limit of interest corresponds to $f \rightarrow \infty$; here, Eq. (45) reduces to

$$\frac{\langle R^2 \rangle^{1/2}}{N} \sim \left[1 - \frac{3}{fl} \right]^{1/2}. \quad (49)$$

This result may be compared with the large- f predictions of the above Marko–Siggia (MS) model, and another prediction based on the freely jointed chain (FJC) model,¹⁶ which are given, respectively, by

$$\frac{\langle R^2 \rangle^{1/2}}{N} \sim 1 - \frac{1}{\sqrt{2} fl}, \quad \text{MS} \quad (50)$$

$$\frac{\langle R^2 \rangle^{1/2}}{N} \sim 1 - \frac{1}{fl}, \quad \text{FJC}. \quad (51)$$

Figure 1 shows the variation of $z^{1/2} \equiv \langle R^2 \rangle^{1/2} / N$ with the logarithm of a dimensionless force $fA \equiv fl/2$ for different

models of the chain. Curve 1 corresponds to the results of the Marko–Siggia calculation, in which the contour length N and the persistence length l have been estimated from experiment to be $32.8 \mu\text{m}$ and 106 nm , respectively. This curve provides an excellent fit to experimental measurements of DNA stretching.¹⁵ Curve 4 corresponds to the results of the FJC model. Curves 2 and 5 and the curve defined by the + signs (curve 3) are derived from Eq. (45), in which the parameters f , N and l have been grouped into the dimensionless combinations fA and l/N . In curve 3, l/N is assigned the value 3.24×10^{-3} based on the estimates for l and N determined by Marko and Siggia.¹⁵ As is clear, with this assignment for l/N , our model reproduces the small- f limit of the MS model (and hence of the experimental data) somewhat poorly. In curve 2, l/N is arbitrarily assigned the value 3.24×10^{-5} , which significantly improves the agreement with the MS model. In all subsequent calculations involving the heteropolymer, this value for l/N will be assigned to one of the components of the copolymer. Curve 5 shows the results of yet another choice for l/N , in this case the choice 0.5, which makes the polymer relatively stiff to begin with, even in the absence of the force.

The results depicted in Fig. 1 do indeed demonstrate that our FENE approximation provides a useful approach to the behavior of semiflexible polymers under tension. We now turn to a discussion of the heteropolymer case.

B. Heteropolymer case

1. Quenched disorder

Equations (29), (30), and (45), when used in Eq. (21), determine the force-extension behavior of the random A – B block copolymer as a function of the following parameters: (i) the number of prepolymers n in the chain (which is kept fixed); (ii) the contour length N in a given prepolymer (assumed to be the same for both A and B and also kept fixed); (iii) the fraction φ of A in the polymer; (iv) the persistence lengths l_A and l_B of the two prepolymers; and (v) the degree of statistical correlation λ between A and B along the chain backbone. The nature of this behavior is best illustrated graphically.

Figure 2 is a plot of the fractional extension $z^{1/2} = \langle \langle R^2 \rangle \rangle_q^{1/2} / M$ vs the force fN for the following fixed values of the parameter λ : 0.0 (curve 2), 0.8 (curve 3), and 0.99 (curve 4). In each of these curves, the parameters n , φ , l_A/N , and l_B/N are kept fixed at these respective values: 50, 0.5, 0.99, and 3.24×10^{-5} . Curves 2–4 therefore describe the behavior of a moderately long copolymer with roughly equal amounts of A and B in which the A prepolymer is quite stiff ($N/l_A \approx 1$), the B prepolymer is very flexible ($N/l_B \gg 1$) and the arrangement of A and B becomes increasingly random in the direction of decreasing λ . (λ can also assume negative values, with $\lambda = -1$ corresponding to an arrangement in which A and B alternate with each other in perfect regularity. The force-extension curves for a chain with $\lambda < 0$ at the above values of n , φ , l_A/N , and l_B/N are almost indistinguishable from curve 2, and are therefore not shown.) Curve 1 is the force-extension curve for a homopolymer of contour length M , obtained from the heteropolymer results by setting

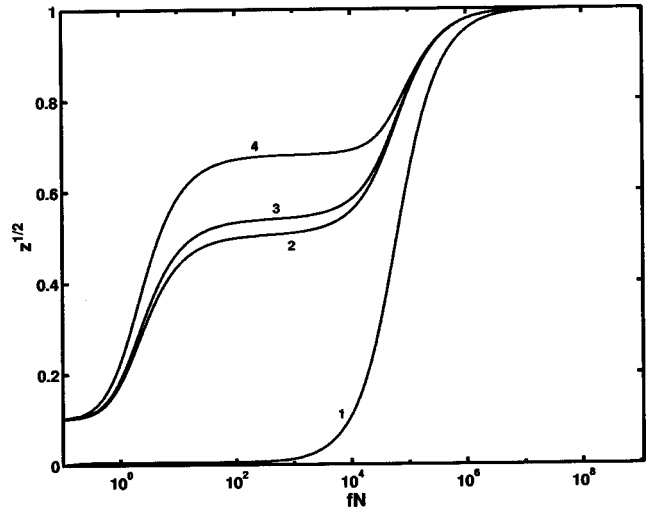


FIG. 2. Fractional extension vs force for the quenched random heteropolymer at different values of the correlation parameter λ . The curves are calculated from Eqs. (21), (29), (30), and (45) at the λ values of 0.0 (curve 2), 0.8 (curve 3), and 0.99 (curve 4) at fixed values of n (50), φ (0.5), l_A/N (0.99), and l_B/N (3.24×10^{-5}). Curve 1 is the corresponding result for a homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$.

$l_A/N = l_B/N = 3.24 \times 10^{-5}$. The notable feature of curves 2–4 is, of course, the appearance of a plateau region for certain values of the force. Numerous experiments and simulations on polypeptides and polynucleotides have revealed similar plateaus,¹⁷ and they have been generally ascribed to the existence of transitions between distinct conformational substates of the polymer.

In Fig. 3, the stages by which the plateau develops are more clearly seen. Here, $z^{1/2}$ is plotted against the logarithm of fN at the following fixed values of l_A/N : 3.24×10^{-5} (curve 1), 0.001 (curve 2), 0.01 (curve 3), and 0.5 (curve 4) at fixed n (50), φ (0.5), l_B/N (3.24×10^{-5}) and λ (0.0). So,

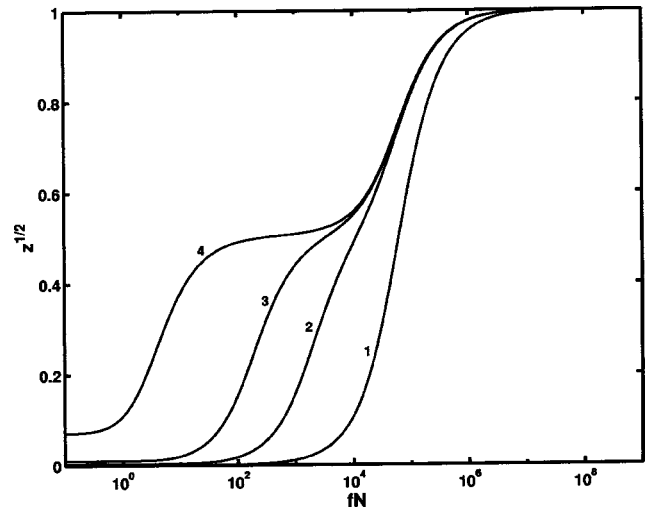


FIG. 3. Fractional extension vs force for the quenched random heteropolymer at different values of the degree of stiffness l_A/N of one of the copolymer components. The curves are calculated from Eqs. (21), (29), (30), and (45) at the l_A/N values of 3.24×10^{-5} (curve 1), 0.001 (curve 2), 0.01 (curve 3), and 0.5 (curve 4) at fixed values of n (50), φ (0.5), l_B/N (3.24×10^{-5}), and λ (0.0).

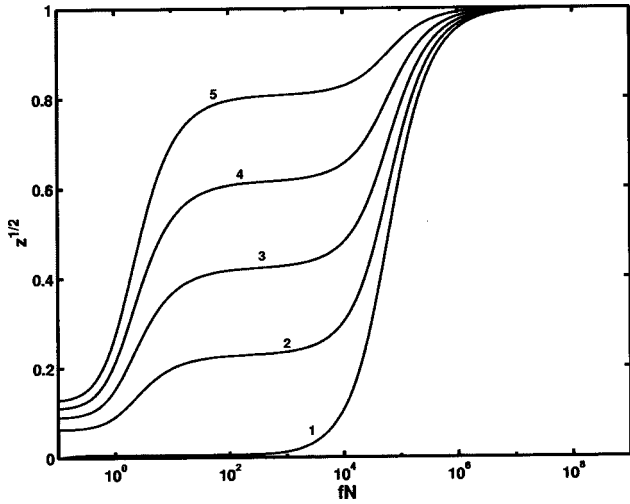


FIG. 4. Fractional extension vs force for the quenched random heteropolymer at different values of the fraction φ of one of the copolymer components. The curves are calculated from Eqs. (21), (29), (30), and (45) at the φ values of 0.2 (curve 2), 0.4 (curve 3), 0.6 (curve 4), and 0.8 (curve 5) at fixed values of n (50), l_A/N (0.99), l_B/N (3.24×10^{-5}), and λ (0.6). Curve 1 refers to the homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$.

the curves describe a completely random copolymer with equal amounts of A and B in which one component is flexible and the other becomes increasingly stiff in the right to left direction. The increase of stiffness in one component therefore tends to produce greater chain extension for a given force, but thereafter, as the force continues to increase, the further increase in chain extension is less for those chains with the more rigid component.

Figure 4 is a plot of $z^{1/2}$ against fN for the following fixed values of the parameter φ : 0.2 (curve 2), 0.4 (curve 3), 0.6 (curve 4), and 0.8 (curve 5) at fixed n (50), l_A/N (0.99), l_B/N (3.24×10^{-5}) and λ (0.6). So, the curves describe a moderately “blocky” copolymer (A s tend to follow A s, on average, and B s tend to follow B s) in which one component is flexible and the other relatively stiff, and the proportion of stiff to flexible increases from curves 2 to 4. Curve 1 corresponds to a homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$. (Had the persistence length of the homopolymer corresponded to a stiff chain, the resulting curve would have been essentially a straight through $z^{1/2} = 1$ for all fN .)

2. Annealed disorder

When the distribution of A and B along the heteropolymer corresponds to annealed disorder, the force-extension curves are calculated from Eq. (22) using Eqs. (33) and (37) and the results given in the Appendix. Figure 5 is a plot of the fractional extension $z^{1/2} = \langle \langle R^2 \rangle \rangle_a^{1/2} / M$ vs the force fN for 3 values of λ : 0.0 (curve 2), 0.8 (curve 3), and 0.99 (curve 4) at fixed values of n (50), φ (0.5), l_A/N (0.99) and l_B/N (3.24×10^{-5}). For these parameter values, the curves in Fig. 5 represent the annealed analogs of the curves in Fig. 2. Curve 1 corresponds to a homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$. The sigmoidal nature of the curves in

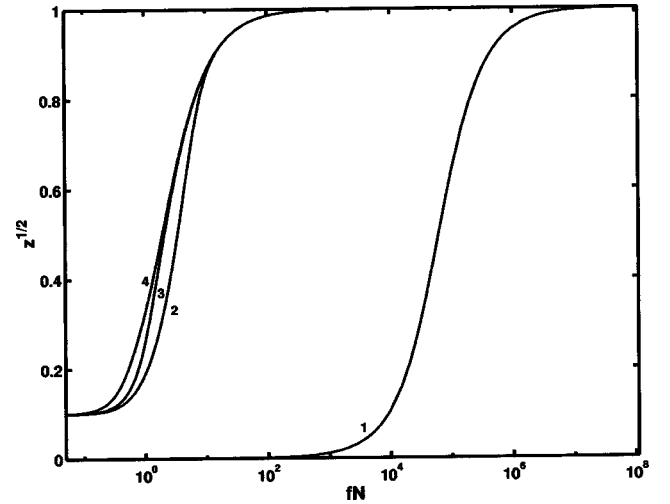


FIG. 5. Fractional extension vs force for the annealed random heteropolymer at different values of the correlation parameter λ . The curves are calculated from Eqs. (22), (33), (37) and results from the Appendix at the λ values of 0.0 (curve 2), 0.8 (curve 3), and 0.99 (curve 4) at fixed values of n (50), φ (0.5), l_A/N (0.99), and l_B/N (3.24×10^{-5}). Curve 1 is the corresponding result for a homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$.

Fig. 5, which contrasts with the steplike behavior of the curves in Fig. 2, is in qualitative agreement with the trends observed in the more phenomenological model of Buhot and Halperin, and may be the result of equilibrium between conformational intermediates that could otherwise have formed in the absence of annealing. Figure 5 also suggests that the negligible dependence of $z^{1/2}$ on λ may have a similar origin.

Figure 6 shows the variation of $z^{1/2}$ with fN at 4 fixed values of l_A/N : 3.24×10^{-5} (curve 1), 0.001 (curve 2), 0.01 (curve 3), and 0.5 (curve 4) at constant values of n (50), φ (0.5), l_B/N (3.24×10^{-5}) and λ (0.0). The curves here represent the annealed analogs of the curves in Fig. 3.

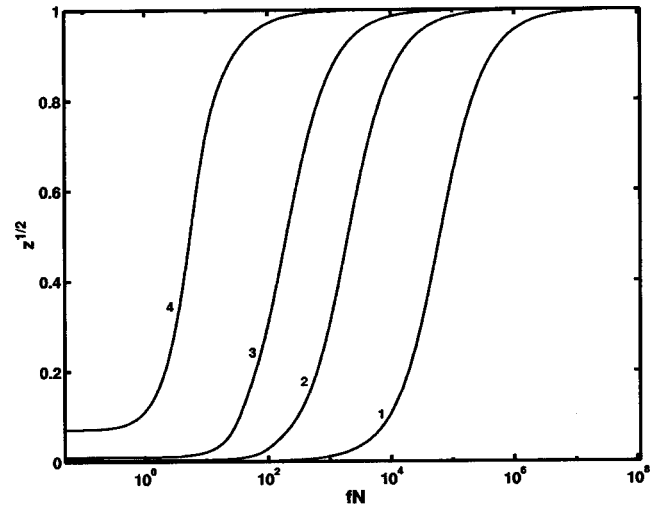


FIG. 6. Fractional extension vs force for the annealed random heteropolymer at different values of the degree of stiffness l_A/N of one of the copolymer components. The curves are calculated from Eqs. (22), (33), (37), and results from the Appendix at the l_A/N values of 3.24×10^{-5} (curve 1), 0.001 (curve 2), 0.01 (curve 3), and 0.5 (curve 4) at fixed values of n (50), φ (0.5), l_B/N (3.24×10^{-5}), and λ (0.0).

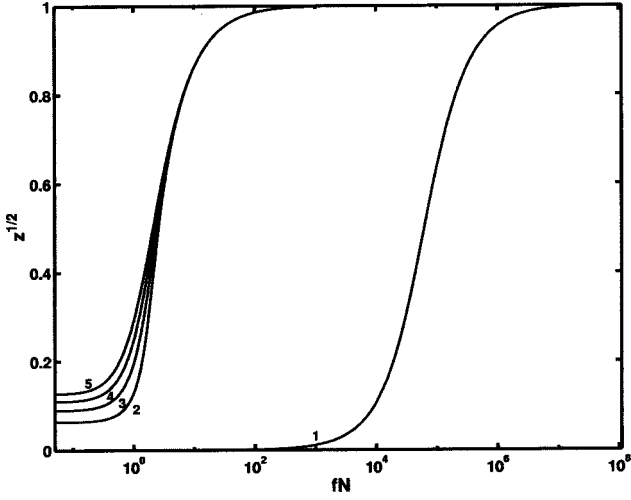


FIG. 7. Fractional extension vs force for the annealed random heteropolymer at different values of the fraction φ of one of the copolymer components. The curves are calculated from Eqs. (22), (33), (37), and results from the Appendix at the φ values of 0.2 (curve 2), 0.4 (curve 3), 0.6 (curve 4), and 0.8 (curve 5) at fixed values of n (50), l_A/N (0.99), l_B/N (3.24×10^{-5}), and λ (0.6). Curve 1 refers to the homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$.

Figure 7 is a plot of $z^{1/2}$ vs fN at 4 fixed values of φ : 0.2 (curve 2), 0.4 (curve 3), 0.6 (curve 4), and 0.8 (curve 5) at constant n (50), l_A/N (0.99), l_B/N (3.24×10^{-5}) and λ (0.6). The curves here represent the annealed analogs of the curves in Fig. 4. Curve 1 corresponds to a homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N = l_B/N = 3.24 \times 10^{-5}$.

Finally, Fig. 8 provides a comparison, in the same graph, of $z^{1/2}$ vs fN for the annealed, quenched, and homopolymer results for one representative set of parameter values. Curves 2 and 3 correspond, respectively, to the quenched and an-

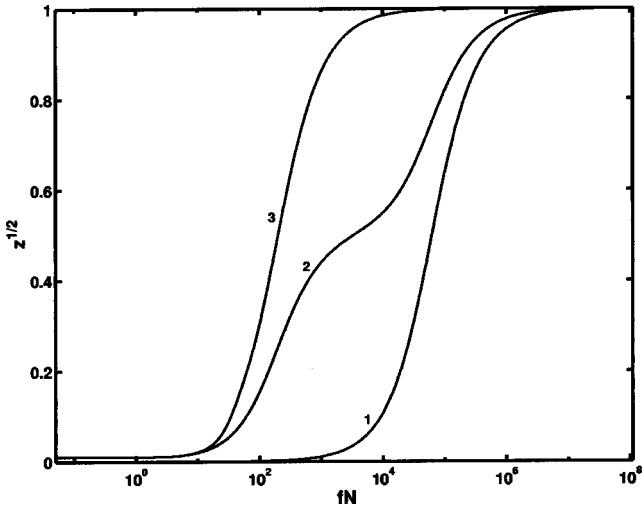


FIG. 8. Comparison of fractional extension vs force for a homopolymer (curve 1) and quenched and annealed heteropolymers (curves 2 and 3). For the heteropolymers, $n=50$, $\varphi=0.5$, $\lambda=0.0$, $l_A/N=0.01$, and $l_B/N=3.24 \times 10^{-5}$. The homopolymer of contour length M , is obtained from the heteropolymer with $l_A/N=l_B/N=3.24 \times 10^{-5}$, and with n , φ , and λ set to the above values.

nealed cases, with $n=50$, $\varphi=0.5$, $\lambda=0.0$, $l_A/N=0.01$, and $l_B/N=3.24 \times 10^{-5}$. Curve 1 corresponds to a homopolymer of contour length M , obtained from the heteropolymer results by setting $l_A/N=l_B/N=3.24 \times 10^{-5}$.

In summary, then, the stretching of random block heteropolymers under quenched and annealed averaging leads to distinct force-extension curves. The results for the quenched case suggest that as a chain unfolds under the action of increasingly large forces (eventually reaching full extension), it may pass through apparently distinct conformational intermediates. The effect depends in detail on the statistical features of the heteropolymer, but is qualitatively the same for all such polymers. The results for the annealed case provide no evidence for the existence of similar conformational transitions; all annealed heteropolymers respond to the applied force in essentially the same way as a homopolymer of comparable size.

ACKNOWLEDGMENT

One of the authors (P.D.) gratefully acknowledges financial support from the Council of Scientific and Industrial Research, India.

APPENDIX: PRINCIPAL FORMULAS FOR THE CALCULATION OF ANNEALED AVERAGES

Referring to the equations in Sec. II C. (2), and introducing the following matrices:

$$\mathbf{P} = \begin{pmatrix} p_{AA} & p_{AB} \\ p_{BA} & p_{BB} \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} e^{-X} & 0 \\ 0 & e^X \end{pmatrix},$$

$$\mathbf{F} = \begin{pmatrix} e^{-X} & 0 \\ 0 & -e^X \end{pmatrix}, \quad (\text{A1})$$

it can be shown, using the methods described in Ref. 8, that

$$\Xi = (1 \ 1) \mathbf{E} (\mathbf{PE})^{n-1} \begin{pmatrix} p_A \\ p_B \end{pmatrix}, \quad (\text{A2})$$

and that

$$Y_i = \left\langle \theta_i \exp \left(-X \sum_{k=1}^n \theta_k \right) \right\rangle_s$$

$$= (1 \ 1) (\mathbf{EP})^{n-i} \mathbf{F} (\mathbf{PE})^{i-1} \begin{pmatrix} p_A \\ p_B \end{pmatrix}, \quad (\text{A3})$$

where p_A and p_B are the probabilities, respectively, that a given prepolymer is either of type A or type B, and have the values φ and $1-\varphi$. The function Y_i (the complete form of which will not be shown in the interests of brevity) can be calculated from

$$Y_i = (1, 1) \left[\frac{1}{C_1} \Lambda_1^{n-i} \mathbf{A}_1 + \frac{1}{C_2} \Lambda_2^{n-i} \mathbf{A}_2 \right]$$

$$\times \mathbf{F} \left[\frac{1}{C_1} \Lambda_1^{i-1} \mathbf{B}_1 + \frac{1}{C_2} \Lambda_2^{i-1} \mathbf{B}_2 \right] \begin{pmatrix} p_A \\ p_B \end{pmatrix}, \quad (\text{A4})$$

where

$$\Lambda_{1,2} = \frac{1}{2} [p_{AA} e^{-X} + p_{BB} e^X \pm \sqrt{(p_{AA} e^{-X} + p_{BB} e^X)^2 - 4\lambda}], \quad (\text{A5})$$

the plus sign referring to Λ_1 and the minus sign to Λ_2 . The parameters C_1 and C_2 are defined by

$$C_i = 1 + \frac{(p_{AA}e^{-X} - \Lambda_i)^2}{p_{AB}p_{BB}}, \quad i=1,2 \quad (\text{A6})$$

while the matrices \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{B}_1 , and \mathbf{B}_2 are given by

$$\mathbf{A}_1 = \begin{pmatrix} 1 & -\phi_1^* \\ -\phi_1 & \phi_1\phi_1^* \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 1 & -\phi_2^* \\ -\phi_2 & \phi_2\phi_2^* \end{pmatrix} \quad (\text{A7})$$

$$\mathbf{B}_1 = \begin{pmatrix} 1 & -\psi_1^* \\ -\psi_1 & \psi_1\psi_1^* \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} 1 & -\psi_2^* \\ -\psi_2 & \psi_2\psi_2^* \end{pmatrix}. \quad (\text{A8})$$

The elements of these matrices are

$$\phi_i = \frac{(p_{AA}e^{-X} - \Lambda_i)}{p_{AB}e^{-X}}, \quad i=1,2 \quad (\text{A9})$$

$$\phi_i^* = \frac{(p_{AA}e^{-X} - \Lambda_i)}{p_{BA}e^X}, \quad i=1,2 \quad (\text{A10})$$

$$\psi_i = \frac{(p_{AA}e^{-X} - \Lambda_i)}{p_{AB}e^X}, \quad i=1,2 \quad (\text{A11})$$

$$\psi_i^* = \frac{(p_{AA}e^{-X} - \Lambda_i)}{p_{BA}e^{-X}}, \quad i=1,2. \quad (\text{A12})$$

In terms of the above results, the expressions T_1 , T_2 , T_3 , and T_4 that appear in Sec. II C 2 are calculated from

$$T_1 = \sum_{i=1}^n Y_i, \quad (\text{A13})$$

$$T_2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_i, \quad (\text{A14})$$

and

$$T_3 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_j. \quad (\text{A15})$$

The summations in these expressions can all be obtained analytically. The expression for T_4 is somewhat more complicated. It is obtained from

$$T_4 = \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1-1) \left[\frac{1}{C_1} \Lambda_1^{n-j} \mathbf{A}_1 + \frac{1}{C_2} \Lambda_2^{n-j} \mathbf{A}_2 \right] \\ \times \mathbf{F} \left[\frac{1}{C_1} \Lambda_1^{j-i-1} \mathbf{B}_1 + \frac{1}{C_2} \Lambda_2^{j-i-1} \mathbf{B}_2 \right] \\ \times \mathbf{P} \mathbf{F} \times \left[\frac{1}{C_1} \Lambda_1^{i-1} \mathbf{B}_1 + \frac{1}{C_2} \Lambda_2^{i-1} \mathbf{B}_2 \right] \begin{pmatrix} p_A \\ p_B \end{pmatrix}. \quad (\text{A16})$$

The summations in this expression can also be obtained in closed form.

- ¹T. Strick, J.-F. Allemand, V. Croquette and D. Bensimon, *Phys. Today* **54**, 46 (2001).
- ²R. Merkel, *Phys. Rep.* **346**, 343 (2001).
- ³D. Marenduzzo, A. Maritan, and F. Seno, *J. Phys. A* **35**, L233 (2002); P. Grassberger and H. P. Hsu, *Phys. Rev. E* **65**, 031807 (2002); D. Marenduzzo, A. Maritan, A. Rosa, and F. Seno, *cond-mat/0206105* (2002).
- ⁴P. L. Geissler and E. I. Shakhnovich, *Phys. Rev. E* **65**, 056110 (2002); D. Bensimon, D. Dohmi, and M. Mézard, *Europhys. Lett.* **42**, 97 (1998).
- ⁵A. Chakraborti, *Phys. Rep.* **342**, 1 (2001).
- ⁶V. S. Pande, A. Yu. Grosberg, and T. Tanaka, *Rev. Mod. Phys.* **72**, 259 (2000).
- ⁷A. Buhot and A. Halperin, *Phys. Rev. Lett.* **84**, 2160 (2000).
- ⁸P. Debnath and B. J. Cherayil, *J. Chem. Phys.* **116**, 4330 (2002).
- ⁹N. Saitô, K. Takahashi, and Y. Yunoki, *J. Phys. Soc. Jpn.* **22**, 219 (1967).
- ¹⁰G. H. Fredrickson, S. T. Milner, and L. Leibler, *Macromolecules* **25**, 6341 (1992).
- ¹¹K. F. Freed, *Adv. Chem. Phys.* **22**, 1 (1972); K. F. Freed, *Renormalization Group Theory of Macromolecules* (Wiley, New York, 1987).
- ¹²A. Dua and B. J. Cherayil, *J. Chem. Phys.* **112**, 8707 (2000).
- ¹³B. Gaveau and L. S. Schulman, *Phys. Rev. A* **42**, 3470 (1990).
- ¹⁴R. B. Bird, C. F. Curtiss, R. C. Armstrong, and O. Hassager, *Dynamics of Polymeric Liquids* (Wiley, New York, 1987), Vol. 2.
- ¹⁵J. F. Marko and E. D. Siggia, *Macromolecules* **28**, 8759 (1995).
- ¹⁶S. B. Smith, L. Finzi, and C. Bustamante, *Science* **258**, 1122 (1992).
- ¹⁷P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron, *Science* **271**, 792 (1996); S. Smith, Y. Cui, and C. Bustamante, *ibid.* **271**, 795 (1996); M. Rief, J. M. Fernandez, and H. E. Gaub, *Phys. Rev. Lett.* **81**, 4764 (1998); E. Paci and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6521 (2000); H. Li, A. F. Oberhauser, S. B. Fowler, J. Clarke, and J. M. Fernandez, *ibid.* **97**, 6527 (2000); A. Sarkar, J.-F. Léger, D. Chatenay, and J. F. Marko, *Phys. Rev. E* **63**, 051903 (2001).