

- (14) Similar irregularities have been noted in the persistence vector analysis on the polymethylene chain by Yoon and Flory (Yoon, D. Y.; Flory, P. J. *J. Chem. Phys.* 1974, 61, 5566.). In their study, extension of the persistence vector is expressed in the XY plane of the coordinate system defined by the first two

bonds of the chain (cf. Figure 1 of this reference). Variation of the X component of the vector with the number of bonds closely relates to the change in $\langle \cos \theta \rangle$.

- (15) Bunn, C. W.; Howells, E. R. *Nature (London)* 1954, 174, 549.
(16) Bates, T. W. *Trans. Faraday Soc.* 1967, 63, 1825.

Ring-Closure Probabilities for Twisted Wormlike Chains. Application to DNA[†]

Jiro Shimada and Hiromi Yamakawa*

Department of Polymer Chemistry, Kyoto University, Kyoto 606, Japan.

Received June 30, 1983

ABSTRACT: The ring-closure probability with the end orientations specified, or the corresponding J factor, as defined as the ratio of equilibrium constants for cyclization and bimolecular association, is evaluated for a twisted wormlike chain, i.e., a special case of the helical wormlike chain. For large length L of the chain contour, evaluation is carried out by the use of the Daniels approximation and the weighting function method previously developed. For small L , the continuous chain is replaced by an equivalent discrete chain in order to treat directly the configuration integral, taking proper account of fluctuations in the configuration of the closed ring around its most probable one, the continuous limit being taken at the final stage of calculation. This is a refinement of the previous approach of Yamakawa and Stockmayer. The evaluation for small L is carried out through (topological) linking-number-dependent J factors, the sum of which is equal to the desired J . It is predicted that the J factor as a function of L stays at zero for very small L , then increases oscillating, and finally decreases monotonically. The derived equations are applied to an analysis of experimental data obtained by Shore et al. for DNA to determine its bending and torsional elastic constants, or its persistence length and Poisson ratio, and also the helix repeat.

I. Introduction

For DNA fragments having cohesive ends, Shore et al.¹ have recently measured the ring-closure probability, or the Jacobson-Stockmayer (J) factor,² as defined as the ratio of equilibrium constants for cyclization and bimolecular association. The results are in semiquantitative agreement with the theoretical prediction for the angle-independent ring-closure probability derived by Yamakawa and Stockmayer³ on the basis of the Kratky-Porod (KP) wormlike chain.⁴ Strictly, however, the J factor for DNA is not a smooth function of chain length but seems to exhibit oscillations for fragments below 500 base pairs. This may be regarded as arising from the fact that if the number of base pairs in the DNA fragment is not an integral multiple of the number of base pairs per helix turn, i.e., the helix repeat, then the need to twist the DNA helix in order to make strand ends meet decreases the J factor significantly for sizes less than 500 base pairs.¹ Thus we must consider a more general ring-closure probability with the end orientations specified. The object of the present paper is to evaluate it by modeling DNA as a special case of our helical wormlike (HW) model,⁵⁻⁷ i.e., a twisted wormlike chain.

The HW chain is defined as an elastic wire model with both bending and torsional energies such that at the minimum of its total configurational (elastic) energy, its contour (as a space curve) becomes a regular helix specified by the constant curvature κ_0 and torsion τ_0 . We can then affix a localized Cartesian coordinate system [$\mathbf{e}_\xi(s)$, $\mathbf{e}_\eta(s)$, $\mathbf{e}_\zeta(s)$] to the chain at the contour point s , with \mathbf{e}_ζ being identical with the unit tangent vector $\mathbf{u}(s)$ and with \mathbf{e}_ξ and \mathbf{e}_η being in the directions of the principal axes of inertia of its cross section at s . For the present purpose, we may represent DNA by a special case of the HW chain with κ_0

= 0 and $\tau_0 \neq 0$, i.e., the KP1 chain,⁸ affixing properly a localized coordinate system to each base pair.^{7,9} Then, at the minimum of energy, the regular helix above reduces to a straight line; as the contour distance s is changed, the localized coordinate system at s moves, rotating about its ζ axis (contour), the locus of the end of the ξ or η axis being a regular helix of pitch $2\pi/\tau_0$ (see Figure 2 of SMHWC-IX⁸) and this helix corresponding to one of the sugar-phosphate backbones. (It is right-handed for $\tau_0 > 0$ and left-handed for $\tau_0 < 0$.) In order to completely define the model, we need two other parameters, i.e., the bending and torsional force constants, but it is convenient to use, instead of them, the stiffness parameter λ^{-1} and Poisson's ratio σ (see the next section). In the present case ($\kappa_0 = 0$), λ^{-1} is equal to the Kuhn segment length and twice the persistence length.⁵ We note that similar elastic models have been adopted by Fuller,¹⁰ Benham,¹¹ and Le Bret¹² in the study of the supercoiling of DNA and by Barkley and Zimm¹³ in the study of its dynamics.

Now let $\mathbf{r}(s)$ be the radius vector of the contour point s ($0 \leq s \leq L$) of the chain of total contour length L and let $\Omega(s) = [\theta(s), \phi(s), \psi(s)]$ ($0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$, $0 \leq \psi \leq 2\pi$) be the Euler angles defining the orientation of the localized coordinate system at s with respect to an external coordinate system. We can define the conditional distribution function, i.e., the Green's function $G(\mathbf{R}, \Omega | \Omega_0; L)$ of $\mathbf{r}(L) = \mathbf{R}$ and $\Omega(L) = \Omega$ when $\mathbf{r}(0) = 0$ and $\Omega(0) = \Omega_0$, \mathbf{R} being the end-to-end distance. Following the Jacobson-Stockmayer theory² and its extension,^{14,15} the desired J factor (in molecules per unit volume) may be related to the ring-closure probability $G(0, \Omega_0 | \Omega_0; L)$ as

$$J = 8\pi^2 G(0, \Omega_0 | \Omega_0; L) \quad (1)$$

It is, however, important to recall that the reaction product is not a homogeneous species but rather a mixture of closed circular DNA with different topological linking numbers.^{10,16,17} The linking number, which is an integer

[†] This paper is dedicated to Professor Walter H. Stockmayer on his 70th birthday.

and which we designate by N , is defined as follows. Suppose that a given (unknotted) closed circular DNA is deformed smoothly so that its contour (helix axis) may form a circle in a plane. Then N is the number of rotations the localized coordinate system at s completes about the contour as s is changed from 0 to L . We note that it is a topological parameter and is independent of chain configuration (or deformation). We can then consider the N -dependent ring-closure probability $G(0, \Omega_0 | \Omega_0; N, L)$, so that

$$G(0, \Omega_0 | \Omega_0; L) = \sum_{N=-\infty}^{\infty} G(0, \Omega_0 | \Omega_0; N, L) \quad (2)$$

It is clear that this G depends on N as $|\Delta N|$, where

$$\Delta N = N - \tau_0 L / 2\pi \quad (3)$$

Note that $\tau_0 L / 2\pi$ is the number of helix turns in the (linear) DNA fragment of length L in its undeformed state (at the minimum of energy), so that ΔN is not necessarily an integer.

The evaluation of the N -dependent ring-closure probability is, in general, very difficult, but it is rather easy for short DNA, since the fraction of species with large $|\Delta N|$ will be very small. In fact, this can then be achieved by devising a new approach (a refinement of the method of Yamakawa and Stockmayer³), i.e., by replacing the continuous chain by an equivalent discrete chain to treat directly the configuration integral without recourse to a solution of the differential equation for the Green's function. Thus, for small L , the J factor may be evaluated from eq 1 and 2. For large L , the Daniels approximation^{3,18,19} and the weighting function method,²⁰ which are invalid for small L , may be used for the evaluation of the N -independent ring-closure probability $G(0, \Omega_0 | \Omega_0; L)$, i.e., the J factor, as done previously for the angle-independent ring-closure probability $G(0; L)$.^{3,20} (These methods are not applicable at all to the N -dependent ring-closure probability.)

The plan of the present paper is as follows. In section II, we evaluate $G(0, \Omega_0 | \Omega_0; N, L)$ for small L . The new method is also applied to $G(0; L)$ and $G(0, \mathbf{u}_0 | \mathbf{u}_0; L)$. The latter is defined from the Green's function $G(\mathbf{R}, \mathbf{u} | \mathbf{u}_0; L)$ of $\mathbf{r}(L) = \mathbf{R}$ and $\mathbf{u}(L) = \mathbf{u}$ when $\mathbf{r}(0) = 0$ and $\mathbf{u}(0) = \mathbf{u}_0$ irrespective of the Euler angle ψ , where $\mathbf{u} = (1, \theta, \phi)$ in spherical polar coordinates. In section III, we evaluate $G(0, \Omega_0 | \Omega_0; L)$ for large L in the Daniels approximation and by the weighting function method. In section IV, we present numerical results for the J factor. In section V, we make an analysis of the data obtained by Shore et al. In the appendices, details of some mathematical aspects are given.

II. Ring-Closure Probabilities for Small Contour Length

We divide the continuous chain into $n + 1$ segments to replace it by an equivalent discrete chain and evaluate the ring-closure probability for the latter, extrapolation to $n = \infty$ being made at the final stage. Mathematically, the procedure is equivalent to the direct evaluation of the functional integral representation of a distribution or wave function.²¹ For small L , the energy may be expanded around its minimum in terms of fluctuations in Ω , and therefore in twist and writhe,^{10,16,17} which must be very small. The configuration integral over the fluctuations may then be evaluated by an application of the well-known technique in random flight statistics.²² We note that the linking number N may then be specified only for the most probable configuration at the minimum of energy.

A. Potential of the Discrete Chain. We begin by deriving an expression for the total potential energy U of

the discrete chain from that for the continuous chain. For this purpose, we first give a brief description of the latter. The localized coordinate system at $s + \Delta s$ is obtained by an infinitesimal rotation $\Delta \bar{\Omega}$ of the system at s , and the deformed state of the chain may be determined by the vector $\omega(s) = (\omega_\xi, \omega_\eta, \omega_\zeta) = \Delta \bar{\Omega} / \Delta s$ ($\Delta s \rightarrow 0$). For the KP1 chain under consideration, U is given by²³

$$U = \frac{1}{2} \int_0^L [\alpha(\omega_\xi^2 + \omega_\eta^2) + \beta(\omega_\zeta - \tau_0)^2] ds \quad (4)$$

where α and β are the bending and torsional force constants, respectively, and are related to each other by the equation $\alpha/\beta = 1 + \sigma$. The parameter λ^{-1} is defined by

$$\alpha = \frac{1}{2} k_B T \lambda^{-1} \quad (5)$$

where k_B is the Boltzmann constant and T is the absolute temperature. Note that $\omega_\xi^2 + \omega_\eta^2 = \dot{\mathbf{u}}^2$, the dot indicating the derivative with respect to s , so that the term proportional to β drops for the original KP chain.²⁴ Recall also that ω may be expressed in terms of the Euler angles as²³

$$\begin{bmatrix} \omega_\xi \\ \omega_\eta \\ \omega_\zeta \end{bmatrix} = \begin{bmatrix} \sin \psi & -\sin \theta \cos \psi & 0 \\ \cos \psi & \sin \theta \sin \psi & 0 \\ 0 & \cos \theta & 1 \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\psi} \\ \dot{\theta} \end{bmatrix} \quad (6)$$

In what follows, all lengths are measured in units of λ^{-1} , and $k_B T$ is chosen to be unity unless noted otherwise.

Now we consider the discrete chain composed of $n + 1$ segments numbered 0, 1, ..., n , each having length L/n except the two end ones of length $L/2n$. We can affix a localized coordinate system ($\mathbf{e}_{\xi p}, \mathbf{e}_{\eta p}, \mathbf{e}_{\zeta p}$) to the p th segment ($p = 0, 1, \dots, n$) corresponding to the system [$\mathbf{e}_\xi(s), \mathbf{e}_\eta(s), \mathbf{e}_\zeta(s)$] at $s = pL/n$ of the continuous chain and designate the associated Euler angles by $\Omega_p = (\theta_p, \phi_p, \psi_p)$. The total potential energy $U(\Omega_1, \dots, \Omega_n)$ (in units of $k_B T$) of the discrete chain with Ω_0 fixed may be written, from eq 4 and 6, as

$$U(\{\Omega_n\}) = \sum_{p=1}^n u(\Omega_{p-1}, \Omega_p) \quad (7)$$

where $\{\Omega_n\} \equiv \Omega_1, \dots, \Omega_n$ and

$$u(\Omega_{p-1}, \Omega_p) = u^{(0)}(\Omega_{p-1}, \Omega_p) - \frac{1}{2} (\sin \theta_{p-1})^{-2} (L/n) \quad (8)$$

with

$$u^{(0)}(\Omega_{p-1}, \Omega_p) = (n/4L) [(\Delta \theta_p)^2 + (\Delta \phi_p)^2 \sin^2 [\frac{1}{2}(\theta_p + \theta_{p-1})] + (1 + \sigma)^{-1} \{\Delta \phi_p \cos [\frac{1}{2}(\theta_p + \theta_{p-1})] + \Delta \psi_p - \tau_0 L/n\}^2] \quad (9)$$

It is important to note here that Ω_p should rather be determined successively from Ω_0 with given $\Delta \Omega_p \equiv \Omega_p - \Omega_{p-1} \equiv (\Delta \theta_p, \Delta \phi_p, \Delta \psi_p) \equiv (\theta_p - \theta_{p-1}, \phi_p - \phi_{p-1}, \psi_p - \psi_{p-1})$ ($p = 1, \dots, n$), so that $-\infty < \theta_p, \phi_p, \psi_p < +\infty$. We have written down the bending and torsional pair potential $u^{(0)}$ following the maneuver of Feynman and Hibbs²¹ and moreover added to it the infinitesimally small potential given by the second term on the right-hand side of eq 8 in order to make it possible to evaluate the configuration integral over $\{\Omega_n\}$. The potential U given by eq 7 with eq 8 and 9 is justified by the fact that it gives a known correct result, as shown later.

The partition function Z is then given by

$$Z = \int \exp[-U(\{\Omega_n\})] d\{\Omega_n\} \quad (10)$$

with $d\{\Omega_n\} = d\Omega_1 \dots d\Omega_n$ and $d\Omega_p = |\sin \theta_p| d\theta_p d\phi_p d\psi_p$. If we change variables from Ω_p to $\Delta \Omega'_p = (\Delta \theta'_p, \Delta \phi'_p, \Delta \psi'_p) = [\Delta \theta_p, \Delta \phi_p \sin (\theta_{p-1} + \frac{1}{2} \Delta \theta_p), \Delta \phi_p \cos (\theta_{p-1} + \frac{1}{2} \Delta \theta_p) +$

$$U = U^* + U_0(\{\delta\Omega'_n\}) + \tilde{U}(\{\delta\Omega'_n\}) \quad (23)$$

$$R_\alpha = R_{\alpha,0}(\{\delta\Theta'_n\}) + \tilde{R}_\alpha(\{\delta\Theta'_n\}) \quad (24)$$

$$h = h^* + \tilde{h}(\{\delta\Theta'_n\}) \quad (25)$$

where U_0 and $R_{\alpha,0}$ are the main parts of the fluctuations and are given by

$$U_0 = \frac{n}{4L} \sum_{p=1}^n \{(\delta\theta_p - \delta\theta_{p-1})^2 + (\delta\phi'_p - \delta\phi'_{p-1})^2 - \frac{1}{4}(\Delta\phi^*)^2(\delta\theta_p + \delta\theta_{p-1})^2 + (1 + \sigma)^{-1}[(\delta\psi_p - \delta\psi_{p-1})^2 + \frac{1}{4}(\Delta\phi^*)^2(\delta\theta_p + \delta\theta_{p-1})^2 - \Delta\phi^*(\delta\theta_p + \delta\theta_{p-1})(\delta\psi_p - \delta\psi_{p-1}) - \Delta\psi^*(\delta\theta_p + \delta\theta_{p-1})(\delta\phi'_p - \delta\phi'_{p-1})]\} \quad (26)$$

with $\Delta\phi^* \equiv \phi^*_p - \phi^*_{p-1} = 2\pi/n$ and $\Delta\psi^* \equiv \psi^*_p - \psi^*_{p-1} - \tau_0 L/n = 2\pi\Delta N/n$, and by

$$\begin{aligned} R_{x,0} &= -\frac{L}{n} \sum_{p=1}^n (1 - \frac{1}{2}\delta_{pn})\delta\phi'_p \sin \phi^*_p \\ R_{y,0} &= \frac{L}{n} \sum_{p=1}^n (1 - \frac{1}{2}\delta_{pn})\delta\phi'_p \cos \phi^*_p \\ R_{z,0} &= -\frac{L}{n} \sum_{p=1}^n (1 - \frac{1}{2}\delta_{pn})\delta\theta_p \\ R_{M,0} &= -\frac{L}{n} \sum_{p=1}^n (1 - \frac{1}{2}\delta_{p,n/2})\delta\theta_p \end{aligned} \quad (27)$$

with δ_{pn} the Kronecker delta. In eq 23–25, \tilde{U} , \tilde{R}_α , and \tilde{h} are higher order terms, for which we omit explicit expressions.

Now we are in a position to carry out the integration. We first consider the asymptotic form G_0 of G , which is obtained with neglect of \tilde{U} , \tilde{R}_α , and \tilde{h} ; i.e.

$$G_0(0, \Omega_0 | \Omega_0; N, n) = (2\pi)^{-3} h^* L Z^{-1} \exp(-U^*) \times \int \exp(-U_0 + i \sum_{\alpha} k_{\alpha} R_{\alpha,0}) d\mathbf{k} d\{\delta\Omega'_{n-1}\} \quad (28)$$

where $d\{\delta\Omega'_{n-1}\} = d(\delta\Omega'_1) \dots d(\delta\Omega'_{n-1})$, with $d(\delta\Omega'_p) = d(\delta\theta_p) d(\delta\phi'_p) d(\delta\psi_p)$ and $d(\delta\phi'_p) = \cos \delta\theta_p d(\delta\phi_p) = \sin \theta_p d(\delta\phi_p)$. As seen from eq 26, U_0 is a quadratic form in $\{\delta\Omega'_{n-1}\}$, its associated matrix being given in Appendix A. If we transform it to a diagonal form by orthogonal transformations Q and $\{\delta\Omega'_{n-1}\}$ to $\{\delta\xi_{3n-3}\} = \delta\xi_1, \dots, \delta\xi_{3n-3}$ by

$$(\delta\theta_p, \delta\phi'_p, \delta\psi_p) = \sum_{j=1}^{3n-3} \delta\xi_j (Q_{3p-2,j}, Q_{3p-1,j}, Q_{3p,j}) \quad (p = 1, \dots, n-1) \quad (29)$$

then G_0 may be written as

$$G_0(0, \Omega_0 | \Omega_0; N, n) = (2\pi)^{-3} h^* L Z^{-1} \exp(-U^*) \times \int \exp\left\{\sum_{j=1}^{3n-3} \left[-\frac{n}{4L} \lambda_j (\delta\xi_j)^2 + iL \sum_{\alpha} k_{\alpha} \mu_{\alpha j} \delta\xi_j\right]\right\} d\mathbf{k} d\{\delta\xi_{3n-3}\} \quad (30)$$

where λ_j are the eigenvalues of the $(3n-3) \times (3n-3)$ matrix \mathbf{A} given by eq A2, arranged in increasing order, and $\mu_{\alpha j}$ are given by

$$\begin{aligned} \mu_{xj} &= -n^{-1} \sum_{p=1}^{n-1} Q_{3p-1,j} \sin \phi^*_p \\ \mu_{yj} &= n^{-1} \sum_{p=1}^{n-1} Q_{3p-1,j} \cos \phi^*_p \\ \mu_{zj} &= -n^{-1} \sum_{p=1}^{n-1} Q_{3p-2,j} \\ \mu_{Mj} &= -n^{-1} \sum_{p=1}^{n/2} (1 - \frac{1}{2}\delta_{p,n/2}) Q_{3p-2,j} \end{aligned} \quad (31)$$

The integral in eq 30 is of the form often encountered in random flight statistics²² and can readily be evaluated.

However, this requires some comments. Care must be expended upon the order of integrations, since all λ_j are not always positive. Suppose that the first ν eigenvalues $\lambda_1, \dots, \lambda_{\nu}$ are negative and the remaining ones are positive. We have found that if integrations are carried out first over $\delta\xi_{\nu+1}, \dots, \delta\xi_{3n-3}$, then over \mathbf{k} , and finally over $\delta\xi_1, \dots, \delta\xi_{\nu}$, the integral is convergent for $|\Delta N|/(1 + \sigma) < 3^{1/2}$ but not for $|\Delta N|/(1 + \sigma) > 3^{1/2}$, consistent with the prediction of Le Bret.¹² Fortunately, however, it can be shown that if integrations are carried out formally first over $\delta\xi_1, \dots, \delta\xi_{3n-3}$ as if all λ_j were positive, the same result is obtained, provided that the integral from the former route is convergent. (We can remove zero eigenvalues by properly choosing n .) In any case, we have

$$G_0(0, \Omega_0 | \Omega_0; N, n) = C_0(\Delta N, \sigma) L^{-13/2} \exp(-U^*) [1 + \frac{1}{4}L + \mathcal{O}(L^2)] \quad (32)$$

with

$$C_0 = 2(4\pi)^{-7/2} [(1 + \sigma)^n n^{-3} K \prod_{j=1}^{3n-3} \lambda_j]^{-1/2} \quad (33)$$

where K is the determinant of the 4×4 matrix whose $\alpha\beta$ elements $K_{\alpha\beta}$ are given by

$$K_{\alpha\beta} = \sum_{j=1}^{3n-3} \mu_{\alpha j} \mu_{\beta j} / \lambda_j n \quad (34)$$

Next we consider the correction to G_0 . It may be evaluated from eq 28 with the integrand multiplied by the additional factor $(1 + \tilde{h}/h^*) \exp(-\tilde{U} + i \sum_{\alpha} k_{\alpha} \tilde{R}_{\alpha}) - 1$. We may expand this factor in terms of $\{\delta\Omega'_{n-1}\}$ and k_{α} , retaining terms of $\mathcal{O}(L)$, where we note that $(\delta\Omega'_p)^2 = \mathcal{O}(L)$ and $k_{\alpha} = \mathcal{O}(L^{-3/2})$ as deduced from eq 28 with eq 26 and 27. It is then necessary to evaluate integrals of the form

$$\int \prod_{q=1}^m T_q \exp\left\{\sum_{j=1}^{3n-3} \left[-\frac{n}{4L} \lambda_j (\delta\xi_j)^2 + iL \sum_{\alpha} k_{\alpha} \mu_{\alpha j} \delta\xi_j\right]\right\} d\{\delta\xi_{3n-3}\} \quad (35)$$

with $1 \leq m \leq 6$, where $T_q(\{\delta\Omega'_{n-1}\})$ is any of $\delta\theta_p$, $\delta\theta_p \pm \delta\theta_{p-1}$, $\delta\phi'_p$, $\delta\phi'_p \pm \delta\phi'_{p-1}$, $\delta\psi_p$, and $\delta\psi_p \pm \delta\psi_{p-1}$. In other words, we must evaluate the moments $\langle \prod_q T_q \rangle$ of a distribution function which is proportional to the exponential in eq 35. Since T_q may be expressed as linear combinations of $\{\delta\xi_{3n-3}\}$, this distribution function is reduced to a multivariate Gaussian distribution of $T_q - \langle T_q \rangle$ by a slight modification of the Wang-Uhlenbeck theorem,²² where the first and second moments ($\langle T_q \rangle$ and $\langle T_q T_{q'} \rangle$) are readily found. With these moments, the desired higher ones may be evaluated by the method of cumulants.²⁵ The integration over \mathbf{k} that remains can easily be performed.

Thus, G may be written in the form

$$G = G_0 [1 + C_1(\Delta N, \sigma) L + \mathcal{O}(L^2)] = C_0 L^{-13/2} \exp(-U^*) [1 + (C_1 + \frac{1}{4})L + \mathcal{O}(L^2)] \quad (36)$$

where we omit an explicit expression for C_1 as a function of ΔN and σ because of its length. Note that the parameter τ_0 appears only in ΔN . Finally, eq 36 may be better approximated by

$$G(0, \Omega_0 | \Omega_0; N, L) = C_0 L^{-13/2} \exp\left\{-\frac{\pi^2}{L} \left[1 + \frac{(\Delta N)^2}{1 + \sigma}\right] + (C_1 + \frac{1}{4})L\right\} \quad (37)$$

where we assume that the limit $n \rightarrow \infty$ has been taken.

The problem that remains is a determination of λ_j , Q , C_0 , and C_1 , followed by extrapolation to $n = \infty$. We find $C_0 = \pi^{3/2}/2(1 + \sigma)^{1/2}$ for $\Delta N = 0$, but, in general, the

Table I
Values of the Numerical Constants in Eq 38 and 39

j	a_{0j}	$a_{1j}^{(0)}$	$a_{1j}^{(1)}$
0	2.784	0.2639	-0.0383
1	2.113	0.1399	-0.0827
2	0.6558	-0.1131	0.0125
3	1.719	0.6500	-0.2170
4	-2.478	-1.1223	0.3961
5	2.588	1.0320	-0.3991
6	-1.210	-0.4601	0.1899
7	0.2437	0.0829	-0.0367

analytical treatment is almost impossible. Therefore, we have numerically solved the eigenvalue problem and evaluated C_0 for $n = 8, 16, 32$, and 64 and C_1 for $n = 4, 8, 16$, and 32 . The final values of C_0 and C_1 have been established by extrapolation to $n^{-1} = 0$. We note that the error of extrapolation becomes very large as $|\Delta N|/(1 + \sigma)$ is increased to its critical value $3^{1/2}$. Good interpolation formulas for C_0 and C_1 have been found to be

$$C_0 = (1 + \sigma)^{-1/2} \sum_{j=0}^7 a_{0j} [\Delta N / (1 + \sigma)]^{2j} \quad (38)$$

$$C_1 = \sum_{j=0}^7 [a_{1j}^{(0)} + a_{1j}^{(1)} / (1 + \sigma)] [\Delta N / (1 + \sigma)]^{2j} \quad (39)$$

where a_{0j} , $a_{1j}^{(0)}$, and $a_{1j}^{(1)}$ are numerical constants. These have been determined by the method of least squares for $0 \leq |\Delta N|/(1 + \sigma) \leq 1.45$, and the results are given in Table I. The values of C_0 and C_1 calculated from eq 38 and 39 agree with the original extrapolated values to within 0.2 and 0.5%, respectively. For $|\Delta N|/(1 + \sigma) > 1.45$, $G(0, \Omega_0 | \Omega_0; N, L)$ almost vanishes for $L \lesssim 3$, as shown numerically in section IV, so that we may then put $C_0 = 0$. All numerical work has been done by the use of a FACOM M-382 digital computer in this University.

C. Other Ring-Closure Probabilities. In this subsection, we apply our method to the ring-closure probabilities $G(0, \mathbf{u}_0 | \mathbf{u}_0; L)$ and $G(0; L)$ and also to the quantity $G(-\mathbf{u}_0 | \mathbf{u}_0; L)$ defined from the Green's function $G(\mathbf{u} | \mathbf{u}_0; L)$. Although $G(-\mathbf{u}_0 | \mathbf{u}_0; L)$ is not a ring-closure probability, its evaluation serves as a justification of our treatment. Since all of them are concerned with only the behavior of the chain contour, the final results may be obtained correctly even if we do not consider the torsional energy from the start as in the case of the original KP chain. Therefore, we drop the term proportional to $(1 + \sigma)^{-1}$ from the U of the discrete chain and designate the rest by $U_B(|\Theta_n|)$. We give only the results with a brief description of the derivation.

We first consider $G(0, \mathbf{u}_0 | \mathbf{u}_0; L)$. The corresponding $G(0, \mathbf{u}_0 | \mathbf{u}_0; n)$ for the discrete chain may be written in a form similar to eq 17

$$G(0, \mathbf{u}_0 | \mathbf{u}_0; n) = (2\pi)^{-3} L Z_B^{-1} \int h(|\Theta_n|) \exp[-U_B(|\Theta_n|) + i \sum_{\alpha} k_{\alpha} R_{\alpha}(|\Theta_n|)] d\mathbf{k} d|\Theta_{n-1}| \quad (40)$$

where Z_B is the partition function given by

$$Z_B = \int \exp[-U_B(|\Theta_n|)] d|\Theta_n| = \left(\frac{4\pi L}{n} \right)^n [1 - 1/4 L + \mathcal{O}(L^2)] \quad (41)$$

with $d|\Theta_n| = d\Theta_1 \dots d\Theta_n$ and $d\Theta_p = |\sin \theta_p| d\theta_p d\phi_p$. The most probable (closed) configuration is a circle of radius $L/2\pi$; i.e.

$$\Theta_p^* = \left(\frac{\pi}{2}, \frac{2\pi p}{n} \right) \quad (42)$$

with

$$\gamma_{\alpha} = 0 \quad \text{for all } \alpha \quad (43)$$

and then we have

$$U_B^* = \pi^2/L, \quad h^* = \pi^{-1} \quad (44)$$

The matrix associated with a quadratic form $U_{B,0}$ from U_B is given in Appendix A. In this case, the factor C_0 in front of the exponential can be evaluated analytically, and we obtain the final result

$$G(0, \mathbf{u}_0 | \mathbf{u}_0; L) = \pi^2 L^{-6} \exp \left(-\frac{\pi^2}{L} + 0.514L \right) \quad (45)$$

Next we consider $G(0; L)$. If $G(0, \mathbf{u}_0; L)$ is the integral of $G(0, \mathbf{u} | \mathbf{u}_0; L)$ over \mathbf{u} , it is clear that $G(0; L) = G(0, \mathbf{u}_0; L)$. The corresponding $G(0, \mathbf{u}_0; n)$ for the discrete chain is given by eq 40 with $d|\Theta_n|$ in place of $d|\Theta_{n-1}|$. For the continuous chain, the associated variational problem has already been solved by Yamakawa and Stockmayer.³ If we choose \mathbf{u}_0 and \mathbf{u}^* in the xy plane so that the x axis bisects the angle between \mathbf{u}_0 and \mathbf{u}^* , as depicted in Figure 1 of ref 3, then Θ_p^* may be expressed in terms of the Jacobian elliptic function cn (whose parameter is k_0^2)²⁶ with $s = pL/n$

$$\Theta_p^* = \pi/2$$

$$\cos \phi_p^* = 1 - 2 \text{cn}^2(v | k_0^2) \quad (46)$$

with

$$v = 2k_0^{-1} K(k_0^{-2}) \left(\frac{p}{n} - \frac{1}{2} \right) \quad (47)$$

$$\gamma_x = -2[K(k_0^{-2})]^2, \quad \gamma_y = \gamma_z = \gamma_M = 0 \quad (48)$$

where $K(k_0^{-2})$ is the complete elliptic integral of the first kind.²⁶ Since $k_0 = 1.100$,³ we have $K(k_0^{-2}) = 2.321$, $\gamma_x = -10.77$, and

$$U_B^* = 7.027/L, \quad h^* = 0.2554 \quad (49)$$

We note that $-4\gamma_x$ is equal to the reduced Lagrange multiplier γ in ref 3 and that Lh is the distance of M from the initial tangent \mathbf{u}_0 (not the x axis). According to Yamakawa and Stockmayer, the most probable configuration is characterized by the vanishing curvature at the chain ends. This condition is equivalent to a certain boundary condition imposed in the treatment of the buckling of a rod.²⁷

In this case, γ_x does not vanish, and therefore $U_{B,0}$ and also terms involving γ_x form a quadratic form, which we designate by $\tilde{U}_{B,0}$. The matrix associated with it is given in Appendix A. Thus we obtain the final result

$$G(0; L) = 28.01 L^{-5} \exp \left(-\frac{7.027}{L} + 0.492L \right) \quad (50)$$

In eq 62 with eq 60 of ref 3, the factor L^{-1} appears in place of L^{-5} . This is due to the fact that the previous treatment of the fluctuation is only approximate.

Finally, we consider $G(-\mathbf{u}_0 | \mathbf{u}_0; L)$. Since \mathbf{R} is not fixed, the corresponding $G(-\mathbf{u}_0 | \mathbf{u}_0; n)$ for the discrete chain may be written simply as

$$G(-\mathbf{u}_0 | \mathbf{u}_0; n) = L Z_B^{-1} \int h(|\Theta_n|) \exp[-U_B(|\Theta_n|) + i k_M R_M(|\Theta_n|)] d\mathbf{k}_M d|\Theta_{n-1}| \quad (51)$$

The most probable configuration is a semicircle of radius L/π ; i.e.

$$\Theta_p^* = \left(\frac{\pi}{2}, \frac{\pi p}{n} \right) \quad (52)$$

with

$$\gamma_M = 0 \quad (53)$$

and then we have

$$U_B^* = \pi^2/4L, \quad h^* = \pi^{-1} \quad (54)$$

In this case, the quadratic form is obtained by a slight modification of $U_{B,0}$, as noted in Appendix A, and the factor C_0 in front of the exponential can be evaluated analytically. Thus we obtain the final result

$$G(-\mathbf{u}_0|\mathbf{u}_0;L) = \frac{1}{4}\pi^{1/2}L^{-3/2} \exp\left(-\frac{\pi^2}{4L} + 0.250L\right) \quad (55)$$

Alternatively the exact expansion of $G(-\mathbf{u}_0|\mathbf{u}_0;L)$ around $L = 0$ is found to be

$$G(-\mathbf{u}_0|\mathbf{u}_0;L) = \frac{1}{4}\pi^{1/2}L^{-3/2} \sum_{l=0}^{\infty} (-1)^l (2l+1) \exp[-\pi^2(2l+1)^2/4L + \frac{1}{4}L] \quad (56)$$

as shown in Appendix B. It is seen that eq 55 agrees with eq 56 with only the $l = 0$ term retained. This is a justification of the present approach, especially the approximation of eq 36 by eq 37 and similar ones.

III. Ring-Closure Probabilities for Large Contour Length

Our starting equation is the expansion of the Green's function $G(\mathbf{R}, \Omega|\Omega_0;L)$ for the HW chain, as given by eq 27 of SMHWC-V¹⁹

$$G(\mathbf{R}, \Omega|\Omega_0;L) = \sum_{l_3 \geq 0} \sum_{|m_l| \leq l_l} \sum_{|j_l| \leq l_l} \mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3}(R;L) \mathcal{D}_{l_1}^{m_1 j_1}(\Omega) \mathcal{D}_{l_2}^{m_2 j_2*}(\Omega_0) Y_{l_3}^{m_3 - m_1}(\Theta, \Phi) \quad (57)$$

where $\mathbf{R} = (R, \Theta, \Phi)$ in spherical polar coordinates, Y_l^m are the normalized spherical harmonics, $\mathcal{D}_l^{mj}(\Omega)$ ($0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$, $0 \leq \psi \leq 2\pi$) are the normalized Wigner functions, with the asterisk indicating the complex conjugate, and l_i satisfy the triangular inequalities. If we put $\mathbf{R} = 0$ and $\Omega = \Omega_0 = (0, 0, 0)$ in eq 57 and use the relations $\mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3}(0;L) = 0$ for $l_3 \neq 0$, $\mathcal{D}_l^{mj}(0,0,0) = [(2l+1)/8\pi^2]^{1/2} \delta_{mj}$, $Y_0^0 = (4\pi)^{-1/2}$ with eq 28 of SMHWC-V,¹⁹ and eq 29 and 32 of SMHWC-IX,⁸ we obtain

$$G(0, \Omega_0|\Omega_0;L) = (8\pi^2)^{-1} (4\pi)^{-1/2} \sum_{j=0}^{\infty} \sum_{l=j}^{\infty} (2 - \delta_{j0}) (2l+1) \mathcal{G}_{l00}^{00jj}(0;L) \quad (58)$$

where \mathcal{G}_{l00}^{00jj} is the real part of \mathcal{G}_{l00}^{00jj} .

Now, for the KP1 chain, if we follow the development of SMHWC-V, recalling that the angular correlation function $g^{ij'}(s)$ is given by eq 55 of SMHWC-VIII,²⁸ then we obtain the interesting relation

$$\mathcal{G}_{l00}^{00jj}(R;L;\sigma, \tau_0) = \exp[-(\sigma j^2 + i j \tau_0)L] \times \mathcal{G}_{l00}^{00jj}(R;L;\sigma = \tau_0 = 0) \quad (59)$$

where we note that \mathcal{G}_{l00}^{00jj} is real for $\sigma = \tau_0 = 0$, as deduced from eq 8, 23, and 28 of SMHWC-V and eq 50 of SMHWC-IX with $l_3 = 0$. Substitution of eq 59 into eq 58 leads to

$$G(0, \Omega_0|\Omega_0;L) = (8\pi^2)^{-1} \sum_{j=0}^{\infty} F_j(L) \cos(j\tau_0 L) \quad (60)$$

with

$$F_j(L) = (2 - \delta_{j0}) \times (4\pi)^{-1/2} \exp(-\sigma j^2 L) \sum_{l=j}^{\infty} (2l+1) \mathcal{G}_{l00}^{00jj}(0;L;\sigma = \tau_0 = 0) \quad (61)$$

Next, integrating both sides of eq 57 over ψ and ψ_0 , dividing by 2π , and putting $\mathbf{R} = 0$ and $\mathbf{u} = \mathbf{u}_0 = \mathbf{e}_z$, we obtain

$$G(0, \mathbf{u}_0|\mathbf{u}_0;L) = (4\pi)^{-1} F_0(L) \quad (62)$$

Finally, integrating both sides of eq 57 over Ω and Ω_0 , dividing by $8\pi^2$, and putting $\mathbf{R} = 0$, we obtain

$$G(0;L) = (4\pi)^{-1/2} \mathcal{G}_{000}^{0000}(0;L) \quad (63)$$

A. The Daniels Approximation. In the Daniels approximation, \mathcal{G}_{l00}^{00jj} is expanded in inverse powers of L , suppressing all exponential terms of order $\exp(-\text{const} \cdot L)$, so that for the KP1 chain $\mathcal{G}_{l00}^{00jj}(R;L)$ may be set equal to zero for $j \neq 0$, as seen from eq 59. We therefore have

$$8\pi^2 G(0, \Omega_0|\Omega_0;L) = 4\pi G(0, \mathbf{u}_0|\mathbf{u}_0;L) = \left(\frac{3}{2\pi L}\right)^{3/2} \left[1 - \frac{11}{8L} + \frac{103}{1920L^2} + \mathcal{O}(L^{-3})\right] \quad (64)$$

where the second equality has been obtained from eq 16 of ref 3 with $\mathbf{u} = \mathbf{u}_0$. Thus, in this approximation, $G(0, \Omega_0|\Omega_0;L)$ is independent of σ and τ_0 . In the same approximation, $G(0;L)$ is given by³

$$G(0;L) = \left(\frac{3}{2\pi L}\right)^{3/2} \left[1 - \frac{5}{8L} - \frac{79}{640L^2} + \mathcal{O}(L^{-3})\right] \quad (65)$$

B. Weighting Function Method. This method provides a least-squares approximation of \mathcal{G}_{l00}^{00jj} by a polynomial with a suitable weighting function as in eq 16 of SMHWC-VI.²⁰ Then the m th approximation to $\mathcal{G}_{l00}^{00jj}(R;L)$ may be written in the form

$$\mathcal{G}_{l00}^{00jj}(R;L) = \left(\frac{3}{2\langle R^2 \rangle}\right)^{3/2} w(\rho) \sum_{k=0}^m M_{l,k}^j(L) \rho^{2k} \quad (66)$$

where we choose, as the weighting function $w(\rho)$, $w_{II}(\rho)$ of SMHWC-VI, i.e.

$$w(\rho) = \exp[-a_1 \rho^2 - a_2 \rho^4 - (b\rho^2)^5] \quad (67)$$

with

$$\rho = \left(\frac{3}{2\langle R^2 \rangle}\right)^{1/2} R \quad (68)$$

Note that $w(\rho)$ with $a_2 = 0$ is the weighting function of Fixman and Skolnick,²⁹ who originally proposed this method. The coefficients a_1 , a_2 , b , and $M_{l,k}^j$ are functions of L . The first three of them are first determined in such a way that the normalized weighting function alone gives the exact moments $\langle R^2 \rangle$, $\langle R^4 \rangle$, and $\langle R^6 \rangle$, and then $M_{l,k}^j(L)$ ($k = 0-m$) are determined in such a way that the $G(\mathbf{R}, \Omega|\Omega_0;L)$ given by eq 57 with eq 66 and this w gives the exact moments $\langle R^{2m'} \mathcal{D}_{l00}^{0j'j*}(\Omega_0) \rangle$ ($m' = 0-m$). All these moments may be evaluated following the method previously developed.^{19,30,31}

With the values of $\mathcal{G}_{l00}^{00jj}(0;L;\sigma = \tau_0 = 0)$ thus evaluated with $m = 6$ and for $0 \leq j \leq l \leq 5$, $G(0, \Omega_0|\Omega_0;L)$ has been found to be convergent for $L > 2.8-3.0$. On the basis of these values, we have then constructed interpolation formulas for $F_j(L)$, since all the coefficients above can only be determined numerically. The results are

$$F_0(L) = \sum_{k=0}^3 f_{0k} L^{-k-3/2}$$

$$F_1(L) = \exp[-(2 + \sigma)L] \sum_{k=0}^4 f_{1k} L^{-k}$$

$$F_j(L) = 0 \quad (j \geq 2) \quad (69)$$

Table II
Values of the Numerical Constants in Eq 69

k	f_{0k}	f_{1k}
0	0.3346	-0.1856
1	-0.4810	2.353
2	-0.04212	2.344
3	0.1495	-18.47
4		16.37

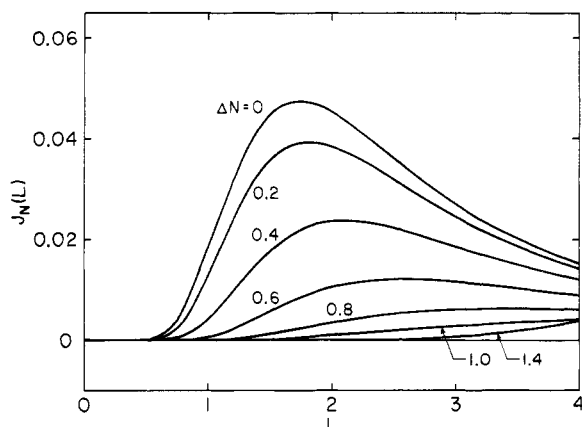


Figure 2. Linking-number-dependent J factor J_N plotted against reduced contour length L for various values of ΔN . The values are those from the configuration integral.

where f_{0k} and f_{1k} are numerical constants, and their values determined by the method of least squares are given in Table II. The values of G calculated from eq 60 or 62 with eq 69 agree with the original values by the weighting function method to within 1% for $2 \leq L \leq 4$ (although the latter are not reliable for $L \leq 2.8$). For $L \geq 4$, we may use the Daniels approximation, i.e., eq 64, since the values by the weighting function method agree well with those from eq 64. Recall that the weighting function method is valid over a wider range of L .²⁰

The values of $G(0;L)$ by the weighting function method agree with those from eq 50 for $2 \leq L \leq 4$ and those from eq 65 for $L \geq 4$ to within 1%.

IV. Numerical Results

In this section, we examine numerically the behavior of the ring-closure probabilities as functions of total contour length L , all lengths being still reduced.

We first consider the linking-number-dependent J factor $J_N(L)$ defined by

$$J_N(L) = 8\pi^2 G(0, \Omega_0 | \Omega_0; N, L) \quad (70)$$

The values of $J_N(L)$ calculated from eq 70 with eq 37–39 for $\sigma = 0$ and for the indicated values of ΔN are plotted against L in Figure 2. It is seen that $J_N(L)$ exhibits a maximum for $0 \leq |\Delta N| \leq 1$ and that at constant $L \leq 3$, it decreases with increasing $|\Delta N|$ and becomes negligibly small for $|\Delta N| \geq 1.4$. In this connection, recall that the circular configuration is never stable for $|\Delta N|/(1 + \sigma) > 3^{1/2}$. We note that as σ is increased from 0 to 0.5, J_N with $\Delta N = 0$ and also its dependence on ΔN become small. In any case, it may be concluded that J_N only with $N = N^*$ and $N^* \pm 1$ make significant contributions, where N^* is an integer closest to $\tau_0 L / 2\pi$.

Then it is convenient to introduce a parameter (auxiliary variable) r ($0 \leq r \leq 1/2$) defined by

$$r = |\tau_0 L / 2\pi - N^*| \quad (71)$$

Its meaning is the following: If the linear chain, which is initially in the undeformed state, is deformed so that its contour is always confined in a plane, we must twist or

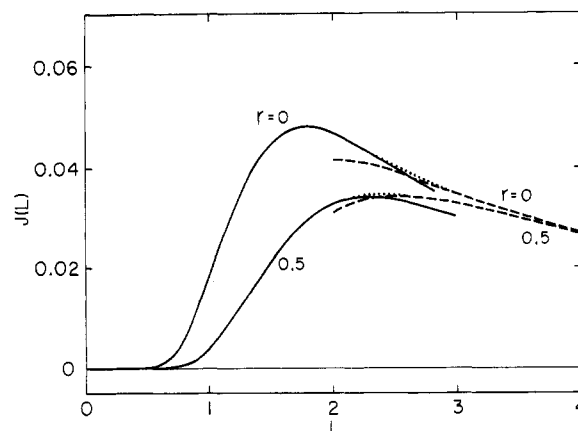


Figure 3. J factor plotted against reduced contour length L for $\sigma = 0$ and $r = 0$ and 0.5. The full curves represent the values from the configuration integral for small L , and the broken curves those by the weighting function method for large L . The dotted curves indicate empirical interpolations.

untwist one end with respect to another by at least r turns in order to join them so that $\Omega = \Omega_0$. Since N^* is a step function of L , r is a periodic function of L , i.e.

$$r(L) = \frac{|\tau_0|L}{2\pi} - k \quad \text{for } k \leq \frac{|\tau_0|L}{2\pi} \leq k + \frac{1}{2}$$

$$r(L) = 1 - \frac{|\tau_0|L}{2\pi} + k \quad \text{for } k + \frac{1}{2} < \frac{|\tau_0|L}{2\pi} < k + 1 \quad (72)$$

with k being nonnegative integers.

Now, the three values of N mentioned above correspond to $\Delta N = r$ and $r \pm 1$, and therefore, the J factor defined by eq 1 may be expressed as

$$J(L) = \sum_{\Delta N=r-1}^1 J_N(L) \quad (73)$$

provided that L is small. The values of $J(L)$ calculated from eq 70 and 73 with eq 37–39 for $\sigma = 0$ and at $r = 0$ and 0.5 are represented by the full curves in Figure 3. The broken curves represent the corresponding values (by the weighting function method) calculated from eq 1 with eq 60 and 69. Note that the dependence on r of the latter values arises from the factor $\cos(j\tau_0 L)$ in eq 60, which is seen, from eq 71, to equal $\cos(2\pi jr)$. The values for small L (full curves) and for large L (broken curves) may be interpolated smoothly, as indicated by the dotted curves. If necessary, an empirical interpolation formula, as given in Appendix C, may be used to do this.

The values of $J(L)$ thus calculated as an explicit function of L with eq 72 for $\tau_0 = 5\pi$ are represented by the full curves in Figures 4 and 5 for $\sigma = 0$ and 0.5, respectively. The dotted curves with $r = 0$ and 0.5 indicate the upper and lower bounds, respectively. The broken curves represent the values of the J factor $J_{(1)}(L)$ defined by

$$J_{(1)}(L) = 4\pi G(0, \mathbf{u}_0 | \mathbf{u}_0; L) \quad (74)$$

which have been calculated from eq 45 for $L < 1.9$ and eq 62 with eq 69 for $2.8 < L < 4$ (and also eq C4). It is seen that $J(L)$ oscillates around $J_{(1)}(L)$, the amplitudes of oscillation being larger for $\sigma = 0$ than for $\sigma = 0.5$. This seems reasonable, since the effects of the torsional energy disappear in the formal limit $\sigma \rightarrow \infty$. We note that for DNA, $\tau_0 \approx 200$ if we assume $\lambda^{-1} = 1100 \text{ \AA}$, the helix repeat 10.4, and the distance 3.4 \AA between base pairs. As is seen from eq 72, the period of oscillation of $J(L)$ becomes short as $|\tau_0|$ is increased.

Finally, we consider the J factor $J_{(0)}(L)$ defined by

$$J_{(0)}(L) = G(0; L) \quad (75)$$

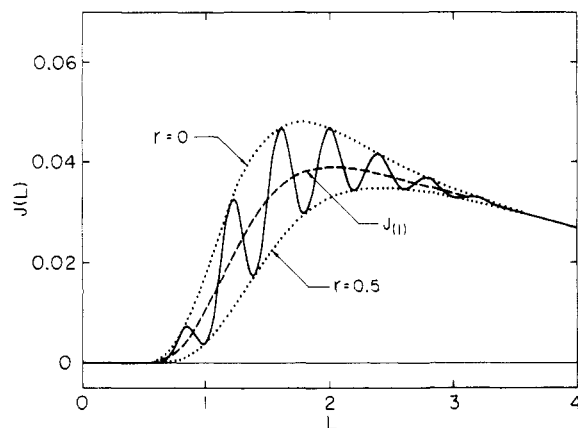


Figure 4. J factor plotted against reduced contour length L for $\sigma = 0$ and $\tau_0 = 5\pi$. The full curve represents the values as an explicit function of L , and the broken curve those of the J factor $J_{(1)}$ with the end tangents specified. The dotted curves indicate the upper ($r = 0$) and lower ($r = 0.5$) bounds of J .

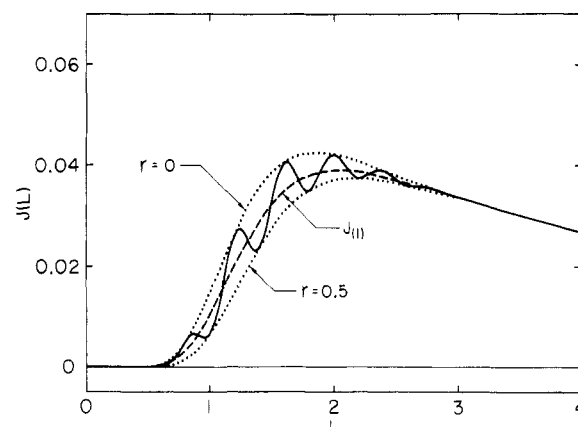


Figure 5. J factor plotted against reduced contour length L for $\sigma = 0.5$ and $\tau_0 = 5\pi$; see the caption of Figure 4.

Its values calculated from eq 50 for $L < 4$ are represented by the full curve in Figure 6. Recall that they agree well with the values by the weighting function method for $2 \lesssim L \lesssim 4$. As is expected, $J_{(0)}(L)$ is larger than $J_{(1)}(L)$ and $J(L)$ in the range displayed. For comparison, the corresponding Yamakawa–Stockmayer (YS) values calculated from eq 62 with eq 60 and 61 of ref 3 are represented by the chain curve, and the recent Monte Carlo values of Hagerman³² by the open circles. The YS values are somewhat overestimated near the peak, while the Hagerman values are rather in good agreement with ours calculated from eq 50.

V. Comparison with Experiment

In this section, we make an analysis of the experimental data for the J factor for DNA obtained by Shore et al.¹ and very recently by Shore and Baldwin.³³ In principle, λ^{-1} , σ , and τ_0 may be determined conveniently from a best fit of a double-logarithmic plot of the theoretical $J(L)$ against L to that of the observed J (in mol/L) against the number n_{bp} of base pairs in the DNA fragment. Note that τ_0 is related to the helix repeat n_0 by the equation $\tau_0 = 2\pi/n_0 l_{bp}$, where l_{bp} is the distance between base pairs, for which we assume 3.4 Å.

With the data of Shore et al.,¹ however, it is difficult to make such a precise analysis. Therefore, we attempt to determine only λ^{-1} and σ , considering only the upper ($r = 0$) and lower ($r = 0.5$) bounds of the theoretical $J(L)$. Figure 7 shows the results of the analysis. The points represent the observed values.¹ (These values are twice the original ones,¹ which are in error.³³) The full curves

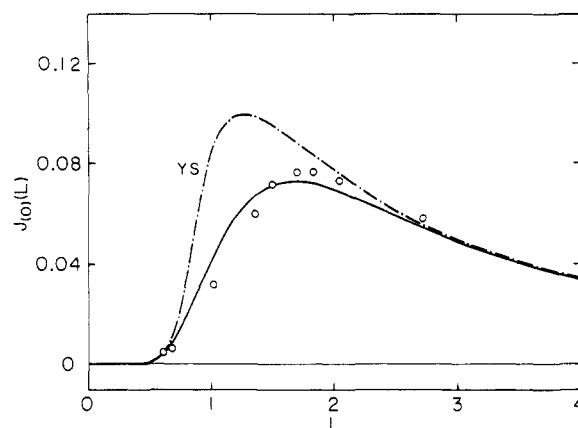


Figure 6. Angle-independent J factor $J_{(0)}$ plotted against reduced contour length L . The full curve represents the present values from the configuration integral, the chain curve Yamakawa and Stockmayer's values,³ and the open circles Hagerman's Monte Carlo values.³²

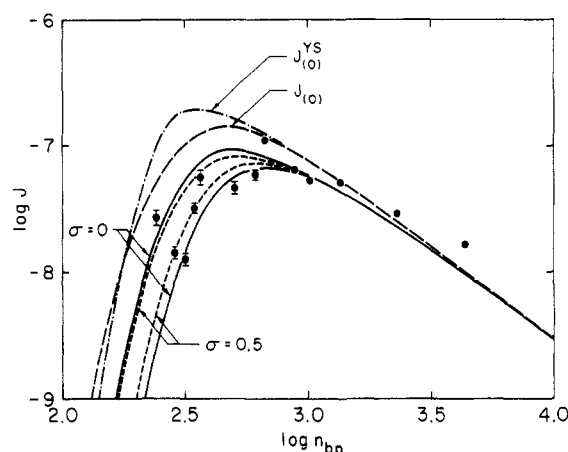


Figure 7. Double-logarithmic plots of the J factor (in mol/L) against the number n_{bp} of base pairs for DNA. The filled circles represent the experimental data of Shore et al.^{1,33} The full curves represent the best fit theoretical values of the upper and lower bounds with $\lambda^{-1} = 950$ Å for $\sigma = 0$, and the broken curves those for $\sigma = 0.5$. The present and Yamakawa–Stockmayer theoretical values of the angle-independent J factor $J_{(0)}$ with the same λ^{-1} are also included.

represent the best fit theoretical values of the upper and lower bounds with $\lambda^{-1} = 950$ Å for $\sigma = 0$, and the broken curves those for $\sigma = 0.5$. For comparison, the present and YS theoretical values of $J_{(0)}$ with the same λ^{-1} are also included. First, it is seen that the agreement between theory and experiment is very much improved if we use the J factor $J(L)$ with the orientations specified instead of the angle-independent one $J_{(0)}(L)$, as is expected. Further, the assignment of $\sigma = 0$ is then seen to be much better than $\sigma = 0.5$.

Very recently (just after the draft of this paper was prepared), the new, more accurate data, although in the narrow range of n_{bp} , obtained by Shore and Baldwin³³ became available to us. Figure 8 shows a precise analysis of these data. The full curve represents the best fit theoretical values with $\lambda^{-1} = 900$ Å, $\sigma = -0.4$, and $n_0 = 10.46$, while the broken curve represents those with $\lambda^{-1} = 950$ Å, $\sigma = -0.2$, and $n_0 = 10.46$, ignoring the largest three observed values.

The values of λ^{-1} above are to be compared with the estimate of $\lambda^{-1} = 1000$ Å by Hagerman,³⁴ and the value 10.46 of n_0 above with the estimates of $n_0 = 10.4$ – 10.6 by Wang³⁵ and Rhodes and Klug.³⁶ The two sets of estimates of λ^{-1} and σ in Figure 8 lead to the torsional constants C

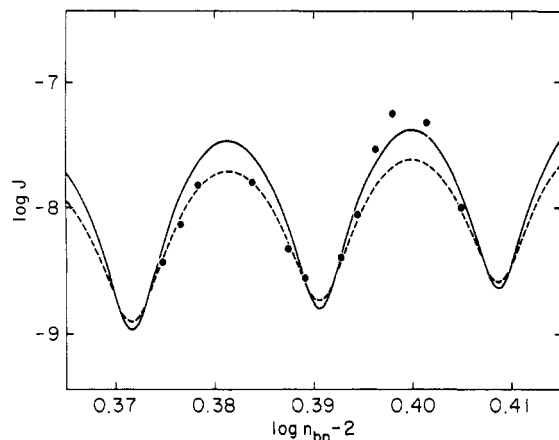


Figure 8. Double-logarithmic plots of the J factor (in mol/L) against the number n_{bp} of base pairs for DNA. The filled circles represent the experimental data of Shore and Baldwin.³³ The full curve represents the best fit theoretical values with $\lambda^{-1} = 900$ Å, $\sigma = -0.4$, and $n_0 = 10.46$, while the broken curve represents those with $\lambda^{-1} = 950$ Å, $\sigma = -0.2$, and $n_0 = 10.46$, ignoring the largest three observed values.

($\approx \beta$) equal to 3.0×10^{-19} and 2.4×10^{-19} erg cm, respectively. These are larger than the value 1.4×10^{-19} erg cm (corresponding to $\sigma = 0.5$) obtained by Thomas et al.³⁷ and Millar et al.³⁸ from fluorescence depolarization data. (Recall that $\sigma \approx 0.5$ for most bulk polymeric materials.) In this connection, we note that in the application of the HW chain to flexible chains, σ is usually assumed to be zero, since the equilibrium conformational behavior is rather insensitive to variation of σ .⁵⁻⁷ Since we are considering local elasticity on the atomic or molecular level, the assignment of $\sigma = 0$ (or <0) is not necessarily surprising.

VI. Concluding Remarks

We have evaluated the ring-closure probability with the end orientations specified for a special case of the HW chain, i.e., the KP1 chain, to primarily explain the experimental data obtained by Shore et al. for DNA. In particular, for small L , evaluation has been possible through the linking-number-dependent ring-closure probability. The specification of the linking number is closely related to the imposition of nonperiodic boundary conditions on the distribution functions. Evaluation of them with such boundary conditions is possible near the rigid rod limit¹³ and also near the rigid ring limit as in the present theory. However, it is difficult when large fluctuations are allowed, as seen from the present development. Indeed, for the HW chain, we have previously^{6,23} derived the differential equations satisfied by the Green's functions with periodic boundary conditions, so that they may be expanded in terms of the Wigner functions \mathcal{D}_l^{mj} , with l being nonnegative integers, as in eq 57. Of course, nonperiodic boundary conditions can be imposed in the mechanical (not statistical) problems such as the determination of the most stable configuration under constraints.^{11,12}

The present analysis indicates that for DNA, the J factor is also a useful quantity from which both λ^{-1} and σ , or the bending and torsional elastic constants, and also n_0 may be determined. Our equation for J_N may be applied to an analysis of topoisomer distribution data, provided that L and $|\Delta N|$ are small.

Acknowledgment. This work was motivated by the preprint Dr. R. L. Baldwin sent to H.Y. in 1981, and the analysis has been completed with the new data in the

recent preprint. These and his valuable suggestions must be acknowledged. H.Y. also thanks Dr. P. J. Hagerman for sending his unpublished Monte Carlo data.

Appendix A. Quadratic Forms

In matrix notation, if $\mathbf{x}^T = (\delta\theta_1, \dots, \delta\theta_{n-1}, \delta\phi'_1, \dots, \delta\phi'_{n-1}, \delta\psi_1, \dots, \delta\psi_{n-1})$ is the $(3n-3)$ -dimensional row vector, with the superscript T indicating the transpose, the potential U_0 given by eq 26 may be written in the quadratic form

$$U_0 = \frac{n}{4L} \mathbf{x}^T \mathbf{A} \mathbf{x} \quad (\text{A1})$$

with

$$\mathbf{A} = \begin{bmatrix} \mathbf{M} + \mathbf{N} & \Delta\phi^* \mathbf{S} & \Delta\psi^* \mathbf{S} \\ \Delta\phi^* \mathbf{S}^T & \mathbf{M} & \mathbf{0} \\ \Delta\psi^* \mathbf{S}^T & \mathbf{0} & (1 + \sigma)^{-1} \mathbf{M} \end{bmatrix} \quad (\text{A2})$$

where $\Delta\phi^* = 2\pi/n$, $\Delta\psi^* = 2\pi\Delta N/n$, and \mathbf{M} , \mathbf{N} , and \mathbf{S} are the $(n-1) \times (n-1)$ matrices whose pq elements ($1 \leq p, q \leq n-1$) are given by

$$M_{pq} = 2\delta_{pq} - \delta_{p,q+1} - \delta_{p,q-1} \quad (\text{A3})$$

$$N_{pq} = -\frac{\sigma}{4(1+\sigma)} (\Delta\phi^*)^2 (2\delta_{pq} + \delta_{p,q+1} + \delta_{p,q-1}) \quad (\text{A4})$$

$$S_{pq} = \frac{1}{2(1+\sigma)} (\delta_{p,q+1} - \delta_{p,q-1}) \quad (\text{A5})$$

In the case of $G(0, \mathbf{u}_0 | \mathbf{u}_0; L)$, if $\mathbf{x}_B^T = (\delta\theta_1, \dots, \delta\theta_{n-1}, \delta\phi'_1, \dots, \delta\phi'_{n-1})$ is the $(2n-2)$ -dimensional row vector, the quadratic form $U_{B,0}$ from the potential U_B is given by

$$U_{B,0} = \frac{n}{4L} \mathbf{x}_B^T \mathbf{A}_B \mathbf{x}_B \quad (\text{A6})$$

with

$$\mathbf{A}_B = \begin{bmatrix} \mathbf{M} + \mathbf{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \quad (\text{A7})$$

where \mathbf{M} and \mathbf{N} are the $(n-1) \times (n-1)$ matrices whose pq elements ($1 \leq p, q \leq n-1$) are given by eq A3 and A4 with $\sigma = \infty$, respectively. In the case of $G(-\mathbf{u}_0 | \mathbf{u}_0; L)$, the quadratic form is given by eq A6 with eq A7, where $\Delta\phi^* = \pi/n$ (instead of $\Delta\phi^* = 2\pi/n$).

In the case of $G(0; L)$, if $\hat{\mathbf{x}}_B^T = (\delta\theta_1, \dots, \delta\theta_n, \delta\phi'_1, \dots, \delta\phi'_n)$ is the $2n$ -dimensional row vector, the quadratic form $\hat{U}_{B,0}$ is given by

$$\hat{U}_{B,0} = \frac{n}{4L} \hat{\mathbf{x}}_B^T \hat{\mathbf{A}}_B \hat{\mathbf{x}}_B \quad (\text{A8})$$

with

$$\hat{\mathbf{A}}_B = \begin{bmatrix} \hat{\mathbf{M}} + \hat{\mathbf{N}} & \mathbf{Q} \\ \mathbf{0} & \hat{\mathbf{M}} \end{bmatrix} \quad (\text{A9})$$

where $\hat{\mathbf{M}}$ and $\hat{\mathbf{N}}$ are the $n \times n$ matrices whose pq elements ($1 \leq p, q \leq n$) are given by

$$\hat{M}_{pq} = \delta_{pq} (2 - \delta_{pn}) (1 - n^{-2} \gamma_x \cos \phi_p^*) - \delta_{p,q+1} - \delta_{p,q-1} \quad (\text{A10})$$

$$\hat{N}_{pq} = -\frac{1}{4} \{ \delta_{pq} [(\Delta\phi_p^*)^2 + (\Delta\phi_{p+1}^*)^2 (1 - \delta_{pn})] + \delta_{p,q+1} (\Delta\phi_p^*)^2 + \delta_{p,q-1} (\Delta\phi_{p+1}^*)^2 \} \quad (\text{A11})$$

with $\Delta\phi_p^* = \phi_p^* - \phi_{p-1}^*$, where ϕ_p^* is given by eq 46.

Appendix B. Derivation of Eq 56

If we choose $\mathbf{u}_0 = \mathbf{e}_z$ and $\mathbf{u} = -\mathbf{e}_z$ in eq 13 of ref 3, we readily obtain

$$G(-u_0|u_0;L) = (4\pi)^{-1} \sum_{l=0}^{\infty} (-1)^l (2l+1) \exp[-l(l+1)L] \quad (B1)$$

This is the expansion around $L = \infty$. The expansion around $L = 0$ may be obtained by the use of Jacobi's imaginary transformation³⁹

$$\vartheta_1(z|t) = i \left(\frac{i}{t} \right)^{1/2} \exp \left(-\frac{iz^2}{\pi t} \right) \vartheta_1 \left(\frac{z}{t} \middle| -\frac{1}{t} \right) \quad (B2)$$

for the elliptic ϑ_1 function defined by

$$\vartheta_1(z|t) = 2 \sum_{l=0}^{\infty} (-1)^l \exp[i\pi t(l + 1/2)^2] \sin[(2l+1)z] \quad (B3)$$

Differentiation of both sides of eq B2 with respect to z leads to

$$\sum_{l=0}^{\infty} (-1)^l (2l+1) \exp[-(l + 1/2)^2 L] = \pi^{3/2} L^{-3/2} \sum_{l=0}^{\infty} (-1)^l (2l+1) \exp[-\pi^2 (l + 1/2)^2 / L] \quad (B4)$$

where we have put $z = 0$ and $t = iL/\pi$. From eq B1 and B4, we obtain eq 56.

Appendix C. Interpolation Formulas

We first construct an empirical interpolation formula for $J(L)$ for intermediate L . Let L_0 be the value of L at which the value of J given by eq 73 agrees with that of J given by eq 1 with eq 60 and 69, both for a given value of r . It may be determined graphically. Let J_1 and J'_1 be the values of the former J and its first derivative with respect to L at $L = L_0 - 0.4 \equiv L_1$, respectively, and let J_2 and J'_2 be those of the latter J and its first derivative at $L = L_0 + 0.4 \equiv L_2$, respectively. J'_1 and J'_2 may be calculated from

$$J'_1 = 8\pi^2 \sum_{\Delta N=r-1}^1 \left\{ -\frac{13}{2L_1} + \frac{\pi^2}{L_1^2} \left[1 + \frac{(\Delta N)^2}{1+\sigma} \right] + C_1(\Delta N, \sigma) + \frac{1}{4} \right\} G(0, \Omega_0 | \Omega_0; N, L_1) \quad (C1)$$

$$J'_2 = -\sum_{k=0}^3 \left(\frac{3}{2} + k \right) f_{0k} L_2^{-5/2-k} - \cos(2\pi r) \{ (2 + \sigma) F_1(L_2) + \exp[-(2 + \sigma)L_2] \sum_{k=1}^4 k f_{1k} L^{-1-k} \} \quad (C2)$$

where in eq C1, $G(0, \Omega_0 | \Omega_0; N, L)$ and C_1 are given by eq 37 and 39, respectively, and in eq C2, F_1 is given by eq 69 and f_{jk} are given in Table II. Then a good interpolation formula for $J(L)$ at constant r , which we designate by $J(L, r)$, is

$$J(L, r) = J_1 + J'_1(L - L_1) - 1.5625[3(J_1 - J_2) + 0.8(2J'_1 + J'_2)](L - L_1)^2 + 1.9531[2(J_1 - J_2) + 0.8(J'_1 + J'_2)](L - L_1)^3 \quad (L_1 < L < L_2) \quad (C3)$$

$J(L)$ as an explicit function of L in this range may be calculated from eq C3 and 72.

A similar interpolation formula for $J_{(1)}(L)$ defined by eq 74 is

$$J_{(1)}(L) = 0.03882 + 0.003494(L - 1.9) - 0.01618(L - 1.9)^2 + 0.008601(L - 1.9)^3 \quad (1.9 < L < 2.7) \quad (C4)$$

References and Notes

- (1) Shore, D.; Langowski, J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4833.
- (2) Jacobson, H.; Stockmayer, W. H. *J. Chem. Phys.* **1950**, *18*, 1600.
- (3) Yamakawa, H.; Stockmayer, W. H. *J. Chem. Phys.* **1972**, *57*, 2843.
- (4) Kratky, O.; Porod, G. *Recl. Trav. Chim. Pays-Bas* **1949**, *68*, 1106.
- (5) Yamakawa, H. *Macromolecules* **1977**, *10*, 692.
- (6) Yamakawa, H.; Fujii, M. *J. Chem. Phys.* **1976**, *64*, 5222 and succeeding papers. These papers are referred to as SMHWC-I, -II, and so on.
- (7) Fujii, M.; Nagasaka, K.; Shimada, J.; Yamakawa, H. *Macromolecules* **1983**, *16*, 1613.
- (8) Yamakawa, H.; Fujii, M.; Shimada, J. *J. Chem. Phys.* **1979**, *71*, 1611.
- (9) Yamakawa, H.; Yoshizaki, T. *J. Chem. Phys.* **1983**, *78*, 572.
- (10) Fuller, F. B. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 815.
- (11) Benham, C. J. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2397; *Biopolymers* **1979**, *18*, 609.
- (12) Le Bret, M. *Biopolymers* **1979**, *18*, 1709.
- (13) Barkley, M. D.; Zimm, B. H. *J. Chem. Phys.* **1979**, *70*, 2991.
- (14) Flory, P. J. "Statistical Mechanics of Chain Molecules"; Interscience: New York, 1969.
- (15) Flory, P. J.; Suter, U. W.; Mutter, M. *J. Am. Chem. Soc.* **1976**, *98*, 5733.
- (16) Crick, F. H. C. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2639.
- (17) Cantor, C. R.; Schimmel, P. R. "Biophysical Chemistry"; W. H. Freeman: San Francisco, 1980; Part III, Chapter 24.
- (18) Gobush, W.; Yamakawa, H.; Stockmayer, W. H.; Magee, W. S. *J. Chem. Phys.* **1972**, *57*, 2839.
- (19) Shimada, J.; Yamakawa, H. *J. Chem. Phys.* **1977**, *67*, 344.
- (20) Yamakawa, H.; Shimada, J.; Fujii, M. *J. Chem. Phys.* **1978**, *68*, 2140.
- (21) Feynman, R. P.; Hibbs, A. R. "Quantum Mechanics and Path Integrals"; McGraw-Hill: New York, 1965.
- (22) Yamakawa, H. "Modern Theory of Polymer Solutions"; Harper and Row: New York, 1971.
- (23) Yamakawa, H.; Shimada, J. *J. Chem. Phys.* **1978**, *68*, 4722.
- (24) Saito, N.; Takahashi, K.; Yunoki, Y. *J. Phys. Soc. Jpn.* **1967**, *22*, 219.
- (25) Rice, S. A.; Gray, P. "The Statistical Mechanics of Simple Liquids"; Interscience: New York, 1965.
- (26) Abramowitz, M.; Stegun, I. A. "Handbook of Mathematical Functions"; Dover: New York, 1965.
- (27) Love, A. E. H. "A Treatise on the Mathematical Theory of Elasticity", 4th ed.; Dover: New York, 1926.
- (28) Yamakawa, H.; Shimada, J. *J. Chem. Phys.* **1979**, *70*, 609.
- (29) Fixman, M.; Skolnick, J. *J. Chem. Phys.* **1976**, *65*, 1700.
- (30) Yamakawa, H. *J. Chem. Phys.* **1973**, *59*, 3811.
- (31) Yamakawa, H.; Fujii, M.; Shimada, J. *J. Chem. Phys.* **1976**, *65*, 2371.
- (32) Hagerman, P. J., private communication.
- (33) Shore, D.; Baldwin, R. L. *J. Mol. Biol.* **1983**, *170*, 957.
- (34) Hagerman, P. J. *Biopolymers* **1981**, *20*, 1503.
- (35) Wang, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 200.
- (36) Rhodes, D.; Klug, A. *Nature (London)* **1980**, *286*, 573.
- (37) Thomas, J. C.; Allison, S. A.; Appellof, C. J.; Schurr, J. M. *Biophys. Chem.* **1980**, *12*, 177.
- (38) Millar, D. P.; Robbins, R. J.; Zewail, A. H. *J. Chem. Phys.* **1982**, *76*, 2080.
- (39) Whittaker, E. T.; Watson, G. N. "A Course of Modern Analysis", 4th ed.; Cambridge University Press: Cambridge, 1927; Chapter XXI.