

## Statistical Mechanics of a Nonlinear Model for DNA Denaturation

M. Peyrard

*Laboratoire d'Optique du Réseau Cristallin, Faculté des Sciences, 6, Boulevard Gabriel, 21100 Dijon, France  
and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545*

A. R. Bishop

*Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory,  
Los Alamos, New Mexico 87545*

(Received 19 December 1988)

We investigate the statistical mechanics of a simple lattice model for the denaturation of the DNA double helix. The model consists of two chains connected by Morse potentials representing the H bonds. We determine the temperature dependence of the interstrand separation and we show that a mechanism involving an energy localization analogous to self-focusing may initiate the denaturation.

PACS numbers: 87.10.+e, 03.40.Kf, 05.90.+m

The dynamics of DNA transcription is one of the most fascinating problems of modern biophysics because it is at the basis of life. However, it is also a very difficult problem due to the complex roles played by RNA polymerases in the process. It is now well established<sup>1</sup> that a local denaturation of DNA is involved so that it is interesting to investigate the thermal denaturation of the double helix as a preliminary step for understanding the transcription.

In the last few years the idea that nonlinear excitations could play a role in the dynamics of DNA has become increasingly popular. Englander *et al.*<sup>2</sup> first suggested a theory of soliton excitations as an explanation of the open states of DNA. Later Yomosa<sup>3</sup> proposed another soliton theory using a plane base-rotator model that was further refined by Takeno and Homma,<sup>4</sup> who introduced a model allowing some discreteness effects to be taken into account, and by Zhang,<sup>5</sup> who improved the model for base coupling. However, although all these models exhibit solitary-wave solutions with a kink shape that could represent open states, the solutions were always obtained in a continuum-limit approximation, which is probably a severe approximation for DNA,<sup>4</sup> and they were not related to the thermal denaturation of DNA since no statistical mechanics of the models was considered. At the same time, an alternative approach, more closely connected to experimental data was developed: Vibrational normal-mode analysis of infrared and Raman experiments suggested that local melting could be achieved through breathing modes<sup>6</sup> and a self-consistent phonon theory was proposed for the hydrogen-bond melting.<sup>7-9</sup> The essential role of strong nonlinearities was pointed out recently by Prohofsky<sup>10</sup> and the aim of this work is to treat these intrinsically in a model similar to those used to analyze the infrared and Raman experiments. We employ a transfer integral technique to analyze the statistical mechanics of the model and determine the interstrand separation in the double helix as a function of temperature. This method is well suited for

discrete models so that no continuum approximation is necessary.

For each base pair, our model includes two degrees of freedom  $u_n$  and  $v_n$  which correspond to the displacements of the bases from their equilibrium positions along the direction of the hydrogen bonds that connect the two bases in a pair. Following previous investigations on DNA,<sup>7,8</sup> the potential for the hydrogen bonds is approximated by a Morse potential. (Our discussion below is readily adapted to any pair potential, but the Morse potential is especially convenient analytically.) A harmonic coupling due to the stacking is assumed between neighboring bases so that the Hamiltonian for the model is

$$H = \sum_n \frac{1}{2} m (\dot{u}_n^2 + \dot{v}_n^2) + \frac{1}{2} k [(u_n - u_{n-1})^2 + (v_n - v_{n-1})^2] + V(u_n - v_n), \quad (1)$$

with

$$V(u_n - v_n) = D \{ \exp[-a(u_n - v_n)] - 1 \}^2.$$

For simplicity we have neglected the inhomogeneities due to the base sequence and the asymmetry of the two strands: We use a common mass  $m$  for the bases and the same coupling constant  $k$  along each strand. The Morse potential  $V$  is an average potential representing the two or three bonds which connect the two bases in a pair. It can be estimated from the parameters obtained for the individual bonds by the analysis of the vibrational modes of DNA.<sup>8</sup>

With these assumptions, the motions of the two strands can be described in terms of the variables

$$x_n = (u_n + v_n)/\sqrt{2}, \quad y_n = (u_n - v_n)/\sqrt{2},$$

which represent the in-phase and out-of-phase motions, respectively. Only the out-of-phase displacements  $y_n$  stretch the hydrogen bonds. The Hamiltonian (1) be-

comes

$$H = H(x) + H(y) = \sum_n \left\{ \frac{p_n^2}{2m} + \frac{1}{2} k (x_n - x_{n-1})^2 \right\} + \sum_n \left\{ \frac{q_n^2}{2m} + \frac{1}{2} k (y_n - y_{n-1})^2 + D [\exp(-a\sqrt{2}y_n) - 1]^2 \right\}, \quad (2)$$

where  $p_n = m\dot{x}_n$  and  $q_n = m\dot{y}_n$ .

For a chain containing  $N$  base pairs, the classical partition function, given in terms of the Hamiltonian (2), may be factored as

$$Z = \int_{-\infty}^{+\infty} \prod_{n=1}^N dx_n dy_n dp_n dq_n e^{-\beta H(p_n, x_n, q_n, y_n)} \\ = Z_p Z_x Z_q Z_y. \quad (3)$$

The two momentum parts are readily integrated to give the familiar kinetic factor for  $N$  particles:

$$Z_p = Z_q = (2\pi mk_B T)^{N/2}.$$

Since the motions in  $x$  and  $y$  are decoupled,  $Z_x$  is simply the contribution of the potential energy in a harmonic chain, so that it is given by

$$Z_x = (2\pi/\beta k)^{N/2},$$

and, since the coupling involves only nearest-neighbors interactions,  $Z_y$  can be expressed in the form

$$Z_y = \int_{-\infty}^{+\infty} \prod_{n=1}^N dy_n e^{-\beta f(y_n, y_{n-1})}, \quad (4)$$

where  $f$  denotes the potential-energy component  $H(y)$ . This integral (4) can be evaluated exactly in the thermodynamic limit of a large system ( $N \rightarrow \infty$ ) using the eigenfunctions and eigenvalues of a transfer integral

$$\varphi_0(y) = (\sqrt{2}a)^{1/2} \frac{(2d)^{d-1/2}}{[\Gamma(2d-1)]^{1/2}} \exp(-de^{-\sqrt{2}ay}) \exp[-(d-\frac{1}{2})\sqrt{2}ay]. \quad (8)$$

The results (7) and (8) can be used to compute the free energy  $F = -k_B T \ln(Z)$  of our model. It is obtained as the sum of the different contributions in  $Z$  [Eq. (3)].

More interesting for the study of DNA denaturation is the mean stretching  $\langle y_m \rangle$  of the hydrogen bonds. It is given by

$$\langle y_m \rangle = \frac{1}{Z} \int \prod_{n=1}^N y_m e^{-\beta H} dx_n dy_n dp_n dq_n,$$

but, since  $H$  [Eq. (2)] separates into a sum over the  $x$ ,  $y$ ,  $p$ , and  $q$  variables,  $\langle y_m \rangle$  reduces to

$$\langle y_m \rangle = \frac{1}{Z_y} \int \prod_{i=1}^N y_m e^{-\beta f(y_n, y_{n-1})} dy_n.$$

As the model is assumed to be homogeneous, the result does not depend on the particular site  $m$  considered. The integral can again be calculated with the transfer integral method<sup>11</sup> and yields

$$\langle y \rangle = \frac{\sum_{i=1}^N \langle \varphi_i(y) | y | \varphi_i(y) \rangle e^{-N\beta\epsilon_i}}{\sum_{i=1}^N \langle \varphi_i(y) | \varphi_i(y) \rangle e^{-N\beta\epsilon_i}}.$$

operator<sup>11,12</sup>

$$\int dy_{n-1} e^{-\beta f(y_n, y_{n-1})} \varphi_i(y_{n-1}) = e^{-\beta\epsilon_i} \varphi_i(y_n). \quad (5)$$

The calculation is similar to the one performed by Krumhansl and Schrieffer<sup>12</sup> for the statistical mechanics of the  $\phi^4$  field. It yields  $Z_y = \exp(-N\beta\epsilon_0)$ , where  $\epsilon_0$  is the lowest eigenvalue of a Schrödinger-type equation which determines the eigenfunctions of the transfer integral operator (5),

$$-\frac{1}{2\beta^2 k} \frac{\partial^2 \varphi_i}{\partial y^2} + D(e^{-4ay} - 2e^{-2ay}) \varphi_i(y) \\ = (\epsilon_i - s_0 - D) \varphi_i(y), \quad (6)$$

with  $s_0 = (1/2\beta) \ln(\beta k/2\pi)$ .

Equation (6) is formally identical to the Schrödinger equation for a particle in a Morse potential, so that it can be solved exactly.<sup>13,14</sup> This equation has a discrete spectrum when  $d \equiv (\beta/a)(kD)^{1/2} > \frac{1}{2}$  and the eigenvalue and the normalized eigenfunction for the ground state are then

$$\epsilon_0 = \frac{1}{2\beta} \ln \left[ \frac{\beta k}{2\pi} \right] + \frac{a}{\beta} \left[ \frac{D}{k} \right]^{1/2} - \frac{a^2}{4\beta^2 k} \quad (7)$$

and

In the limit of large  $N$ , the result is again dominated by the lowest eigenvalue  $\epsilon_0$  so that  $\langle y \rangle$  is given by

$$\langle y \rangle = \langle \varphi_0(y) | y | \varphi_0(y) \rangle = \int \varphi_0^2(y) y dy \quad (9)$$

for the normalized eigenfunction  $\varphi_0(y)$ . This integral has been evaluated numerically with the expression (8) of  $\varphi_0(y)$ . The results are shown in Fig. 1 for three values of the coupling constant  $k$ . In the calculation, we have used  $D = 0.33$  eV and  $a = 1.8 \text{ \AA}^{-1}$  which corresponds to mean values for the N-H...N and N-H...O bonds in the A-T and G-C base pairs.<sup>8</sup> This figure shows a rapid increase of  $\langle y \rangle$  around a particular temperature which is a characteristic of DNA denaturation as observed for instance by measuring its absorbance of ultraviolet light at 260 nm.<sup>1</sup> The denaturation temperature is indeed sensitive to the parameters of the hydrogen bonds which bind the two strands, but Fig. 1 shows that it is also very sensitive to the *intrastrand* interaction constant, a parameter which is not so well determined exper-

imentally. As  $k$  increases the denaturation temperature increases. This is consistent with the increase observed experimentally in the presence of reagents that increase the hydrophobic interactions.<sup>1</sup> Our results indicate that  $k$  must be of the order of  $3.0 \times 10^{-3} \text{ eV/\AA}^2$  to obtain a reasonable denaturation temperature. This value indicates a weak coupling between the transverse base motions in DNA indicating that discreteness effects have

to be taken into account in the models describing DNA denaturation in terms of solitary waves.<sup>2-5</sup>

Although it indicates *at which temperature* the denaturation occurs, the calculation of  $\langle y \rangle$  does not indicate *how* it occurs. To explore this aspect, it is interesting to relate our model to those involving nonlinear excitations. Since the nonlinearities appear in terms of the variable  $y$ , we consider the equation of motion which derives from  $H(y)$  [Eq. (2)]:

$$m \partial^2 y_n / \partial t^2 - k(y_{n+1} + y_{n-1} - 2y_n) - 2\sqrt{2}Da e^{-\sqrt{2}ay_n} (e^{-\sqrt{2}ay_n} - 1) = 0. \quad (10)$$

The phenomena which are precursors to denaturation can be investigated by expanding Eq. (10) for small  $y$  as

$$m \partial^2 y_n / \partial t^2 - k(y_{n+1} + y_{n-1} - 2y_n) + 4Da^2 y_n - 6\sqrt{2}Da^3 y_n^2 + \frac{28}{3} Da^4 y_n^3 = 0. \quad (11)$$

A solution of this equation can be obtained via a multiple-scale expansion<sup>15</sup> as

$$y_n = F_1(X_n, \tau) e^{i\theta_n} + \text{c.c.} + \epsilon [F_0(X_n, \tau) + F_2(X_n, \tau) e^{2i\theta_n} + \text{c.c.}], \quad (12)$$

with  $X_n = \epsilon x = \epsilon n l$  and  $\tau = \epsilon t$ ,  $l$  being the lattice spacing, i.e., the mean distance between adjacent bases. In this expression  $y$  is written as a modulated wave in which the carrier wave  $\exp(i\theta_n) = \exp[i(qnl - \omega t)]$ , with  $\omega$  and  $q$  related by the dispersion relation of the lattice, includes the discreteness effect while the modulation factor is treated in a continuum limit (the so-called *semidiscrete* approximation<sup>16</sup>). However, for a qualitative understanding of the denaturation, we can restrict ourselves to a continuum approximation for the carrier wave as well, which significantly simplifies the calculations. The dc and first-harmonic terms in (12),  $F_0$  and  $F_2$ , are necessary because Eq. (11) contains an even power of  $y$ . The multiple-scale expansion yields a nonlinear Schrödinger

(NLS) equation for  $F_1(X, \tau)$  (Ref. 17),

$$iF_{1s} + PF_{1ZZ} + Q|F_1|^2 F_1 = 0,$$

with  $Z = X - V_g T$ ,  $s = \epsilon \tau$ , and

$$V_g = \frac{d\omega}{dq} = \frac{q}{\omega} \frac{kl^2}{m}, \quad P = \frac{kl^2/m - V_g^2}{2\omega}, \quad Q = \frac{16Da^4}{\omega m}.$$

The solitary waves in  $y$  resulting from this equation are the breathing modes suggested by the infrared and Raman experiments.<sup>6</sup>

The statistical mechanics of the NLS equation has been investigated recently by Lebowitz, Rose, and Speer.<sup>18</sup> They show that, since the associated Hamiltonian is unbounded below in energy, the systems can develop singularities in a finite time. Such singularities, which correspond to self-focusing phenomena in plasmas, may be responsible for nucleation of DNA denaturation because they occur when the  $L^2$  norm of the field

$$N(F_1) = \int_0^L |F_1|^2 dx$$

(where  $L$  denotes the size of the system) is such that  $N\beta$  exceeds some threshold value. In our calculation,  $N(F_1)$  is given at temperature  $T$  by  $\langle y^2 \rangle$  which can be calculated similarly to  $\langle y \rangle$  and is given by an expression analogous to Eq. (9):

$$\langle y^2 \rangle = \langle \varphi_0(y) | y^2 | \varphi_0(y) \rangle = \int \varphi_0^2(y) y^2 dy. \quad (13)$$

Figure 2 shows that  $\langle y^2 \rangle$  rises by several orders of magnitude in a small temperature range around DNA denaturation so that the threshold for energy localization could be reached. This suggests that the denaturation "bubble" observed experimentally at the beginning of the denaturation process could be created by *energy localization due to nonlinear effects*. It is, however, difficult to provide a more quantitative analysis of this phenomenon within the framework of the present model

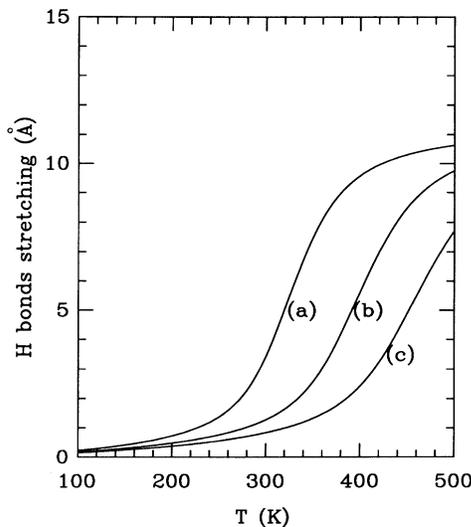


FIG. 1. Variation of  $\langle y \rangle$  as a function of temperature for three values of the coupling constant  $k$ : (a)  $k = 2.0 \times 10^{-3} \text{ eV/\AA}^2$ , (b)  $k = 3.0 \times 10^{-3} \text{ eV/\AA}^2$ , and (c)  $k = 4.0 \times 10^{-3} \text{ eV/\AA}^2$ .

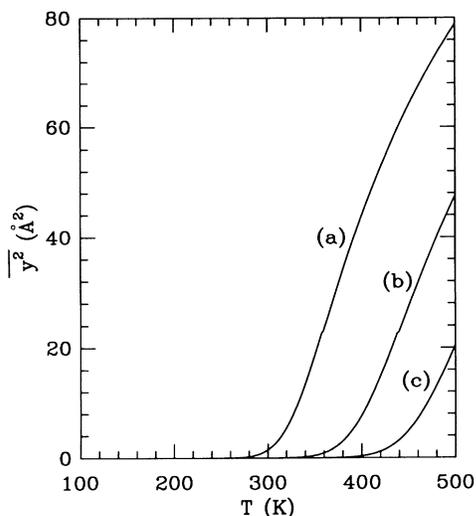


FIG. 2. Variation of  $\langle y^2 \rangle$  as a function of temperature for three values of the coupling constant  $k$ : (a)  $k = 2.0 \times 10^{-3}$  eV/ $\text{\AA}^2$ , (b)  $k = 3.0 \times 10^{-3}$  eV/ $\text{\AA}^2$ , and (c)  $k = 4.0 \times 10^{-3}$  eV/ $\text{\AA}^2$ .

because the coefficients of the NLS derived in the limit of small  $y$  depend on the frequency of the carrier wave introduced in the expansion (12). Although this frequency must lie within the frequencies of the lattice phonon modes, it is not well defined for the thermal fluctuations that we consider here.

Thus, the analysis of the precursor phenomena in terms of a small-amplitude expansion leading to a NLS equation suggests that the process could be initiated by an energy-localization phenomenon analogous to self-focusing in plasmas. However, the results presented here are only a first step toward the understanding of DNA denaturation which is known to be very sensitive to external factors, such as salt concentration. In our model, the variation of the coupling constant  $k$  gives only a phenomenological description of these factors. Moreover, although the variables  $u_n$  and  $v_n$  could also represent base rotations provided that the parameters of the potential  $V$  are chosen appropriately, the introduction of additional degrees of freedom will be necessary because the low sensitivity of the vibrational behavior of DNA to the number of bridging hydrogen bonds suggests that the "melting" of the hydrogen-bond bridging is only part of the strand separation process. Finally, our model does

not include the inhomogeneities present in natural DNA. The transfer integral technique can be adapted to include these *extrinsic* nucleation centers which can be expected to lower the denaturation temperature.<sup>19</sup> This effect is also seen in numerical simulations<sup>20</sup> and will be investigated in future works.

One of us (M.P.) would like to thank D. K. Campbell for his hospitality at the Center for Nonlinear Studies of the Los Alamos National Laboratory, where part of this work has been performed. We would like to thank E. W. Prohofsky, H. A. Rose, A. Garcia, M. Remoissenet, and C. Reiss for helpful discussions.

<sup>1</sup>D. Freifelder, *Molecular Biology* (Jones and Bartlett, Boston, 1987).

<sup>2</sup>S. W. Englander, N. R. Kallenbach, A. J. Heeger, J. A. Krumhansl, and S. Kitwin, Proc. Natl. Acad. Sci. U.S.A. **77**, 7222 (1980).

<sup>3</sup>S. Yomosa, Phys. Rev. A **27**, 2120 (1983); **30**, 474 (1984).

<sup>4</sup>S. Takeno and S. Homma, Prog. Theor. Phys. **77**, 548 (1987).

<sup>5</sup>Chun-Ting Zhang, Phys. Rev. A **35**, 886 (1987).

<sup>6</sup>E. W. Prohofsky, K. C. Lu, L. L. Van Zandt, and B. E. Putnam, Phys. Lett. **70A**, 492 (1979).

<sup>7</sup>Y. Gao and E. W. Prohofsky, J. Chem. Phys. **80**, 2242 (1984).

<sup>8</sup>Y. Gao, K. V. Devi-Prasad, and E. W. Prohofsky, J. Chem. Phys. **80**, 6291 (1984).

<sup>9</sup>Y. Kim, K. V. Devi-Prasad, and E. W. Prohofsky, Phys. Rev. B **32**, 5185 (1985).

<sup>10</sup>E. W. Prohofsky, Phys. Rev. A **38**, 1538 (1988).

<sup>11</sup>D. J. Scalapino, M. Sears, and R. A. Ferrel, Phys. Rev. B **6**, 3409 (1972).

<sup>12</sup>J. A. Krumhansl and J. R. Schrieffer, Phys. Rev. B **11**, 3535 (1975); J. F. Currie, J. A. Krumhansl, A. R. Bishop, and S. E. Trullinger, Phys. Rev. B **22**, 477 (1980).

<sup>13</sup>P. M. Morse, Phys. Rev. **34**, 57 (1929).

<sup>14</sup>G. Dana Brabson, J. Chem. Educ. **50**, 397 (1973).

<sup>15</sup>D. J. Kaup and A. C. Newell, Phys. Rev. B **18**, 5162 (1978).

<sup>16</sup>St. Pnevmatikos, N. Flytzanis, and M. Remoissenet, Phys. Rev. B **33**, 2308 (1986).

<sup>17</sup>M. Remoissenet, Phys. Rev. B **33**, 2386 (1986).

<sup>18</sup>J. L. Lebowitz, H. A. Rose, and E. R. Speer, J. Stat. Phys. **50**, 657 (1988).

<sup>19</sup>Y. Kim and E. W. Prohofsky, Phys. Rev. B **33**, 5676 (1986).

<sup>20</sup>M. Peyrard (unpublished).