Physica D 68 (1993) 104–115
North-Holland

SDI: 0167-2789(93)E0174-A

PHYSICA D

Biomolecular dynamics of DNA: statistical mechanics and dynamical models

M. Peyrard[a], T. Dauxois[a], H. Hoyet[b] and C.R. Willis[c]

[a]*Laboratoire de Physique, CNRS URA 1325, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France*
[b]*Physique non linéaire: Ondes et Structures Cohérentes, Faculté des Sciences, 6 blvd Gabriel, 21000 Dijon, France*
[c]*Department of Physics, Boston University, 590 Commonwealth Ave, Boston, MA 02215, USA*

There is a growing feeling that biomolecular structure is not sufficient to determine biological activity which is also governed by large amplitude dynamics of the molecules. The transcription of DNA or its thermal denaturation are typical examples. Traditional approaches use Ising models to describe the denaturation transition of DNA. They have to introduce phenomenological "cooperativity factors" to explain the rather sharp "melting" of this quasi one-dimensional system. We present models which describe the full dynamics of the melting. Using molecular dynamics simulations and statistical analysis, we discuss the mechanism of the denaturation, including precursor effects that can be related to large amplitude localized nonlinear excitations of the molecule in which discreteness effects play a large role. We also show the microscopic origin of the cooperativity factors.

1. Introduction

Nucleic acids are the repository of genetic information and each of the units that compose DNA or RNA molecules plays an essential role in the biological functions. The famous discovery of the double helix has emphasized a strong relationship between *structure* and function in molecular biology. However this structure, which is so well designed to include the genetic code in two complementary strands and protects it against external perturbations, would also prevent the expression of the code if the molecule were static because the coding bases are not directly accessible to chemical reaction. There are however many indications that DNA is a very *dynamical* entity, undergoing very large deformations and should not be viewed merely as a solid with a particular structure.

A typical example in which the dynamics of the molecule is essential for its function is DNA transcription during which a segment of the genetic code is copied into RNA. In order to expose the coding bases to chemical reaction, the double helix unwinds locally and forms a "bubble" which is about 20 base-pair long and moves along the molecule as the transcription proceeds. This complex process, which is activated by an enzyme, is still beyond a physical analysis but it has strong similarities with the early stage of the termal denaturation, or "melting" of the double helix. The melting, which is the separation of the two complementary strands, starts locally by the formation of small denaturated regions very similar to the transcription bubble. Another important motion of the DNA molecule is its "breathing" or fluctuational opening. In these very large fluctuations, base-pairs are temporarily broken and the two bases are exposed for chemical reaction for a very short time (10^{-7} s). These fluctuational openings can be considered as intrinsic precursors for the denaturation and they could play a role in carcinogenesis by external molecules [1].

The molecular deformation involved in melting or in fluctuational openings are so large that

0167-2789/93/$06.00 © 1993 – Elsevier Science Publishers B.V. All rights reserved

they cannot be described by linear approximations. Therefore biomolecular dynamics is a fascinating topic for nonlinear science because it is related to basic phenomena of life and we *know* that it has to be fundamentally nonlinear. We discuss here some aspects of DNA dynamics and we show that simple nonlinear models can provide a good description of the large amplitude distortions of the molecule which are observed experimentally. Our basic approach can be viewed as an extension of the Ising models, which have been widely used to study the statistical mechanics of the melting, in which we treat completely the dynamics of the bases. In section 2 we present our basic model and its statistical mechanics. This section makes the connection with the usual Ising models for DNA. Section 3 studies the dynamics of the model. Molecular dynamics is used to detect the main types of large amplitude motions and connect them with the experimental observations. Then we propose two analytical investigations adapted to the description of the fluctuational openings and denaturation bubbles. Section 4 discusses more precisely the thermodynamics of the melting of this one dimensional object, the DNA molecule. We show how the introduction of nonlinear coupling terms to describe the base stacking interactions is essential to explain the sharp melting transition observed experimentally.

2. A simple model for DNA melting and its statistical mechanics

The simplest description of DNA melting represents a base-pair by an Ising-like variable which takes only two values, 0 and 1, i.e., closed and open. The denaturation transition is then analyzed by treating the statistical mechanics of this one-dimensional Ising-spin chain [2]. The structure of DNA appears in the calculation of the statistical weight of each state of the molecule which is expressed as the product of a stability parameter for each base-pair, a cooperativity parameter taking into account the fact that breaking a base-pair destroys two stacking interactions unless the pair is the terminal pair of an open region, and an entropy parameter for each loop measuring the "stiffness" of the DNA strands. Many varieties of these Ising-models have been presented. For instance the stability parameter can be assigned to base-pair doublets rather than to a single base-pair because, as we discuss in the last section, there is a strong relationship between the stacking interaction of adjacent bases and the stability of the pairs. The model parameters are determined phenomenologically in order to get the best possible agreement between the theoretical predictions and the experimental melting curves. Once a particular model has been calibrated it may be used to predict the melting curve of another DNA segment. The success of this approach to reproduce experimental melting curves is impressive, but it involves a large number of adjustable parameters. For instance, 10 parameters are used to represent the 10 possible types of base-pair doublets along the molecule. Moreover, using an Ising-variable prevents any attempt to describe the dynamics of the fluctuational openings since states intermediate between closed and open cannot be represented.

Our approach goes further but still keeps the model as simple as possible in an attempt to determine the fundamental mechanism of the melting. Therefore we consider a simplified geometry for the DNA chain in which we have neglected the asymmetry of the molecule and we represent each strand by a set of point masses that correspond to the nucleotides. The characteristics of the model are the following:

(i) The longitudinal displacements are not considered because their typical amplitudes are significantly smaller than the amplitudes of the smaller transverse ones [3]. The stretching of a base-pair in the transverse direction is represented by a real variable y_n which can therefore describe all the states of the pair from closed ($y_n = 0$) to completely broken.

(ii) Two neighboring nucleotides of the same strands are connected by an harmonic potential to keep the model as simple as possible. On the other hand, the bonds connecting the two bases belonging to different strands are extremely stretched when the double helix opens locally so that their nonlinearity must not be ignored. We use a Morse potential to represent the transverse interaction of the bases in a pair. It describes not only the hydrogen bonds but the repulsive interactions of the phosphate groups, partly screened by the surrounding solvent as well. The Hamiltonian of the model is then the following:

$$H = \sum_n [\tfrac{1}{2} m \dot{y}_n^2 + \tfrac{1}{2} K (y_n - y_{n-1})^2 + D(\mathrm{e}^{-a y_n} - 1)^2] . \tag{1}$$

Since we are interested in the thermal denaturation transition of the molecule, the natural approach is to investigate the statistical mechanics of the model. Due to the one-dimensional character of the system, and because the interactions are restricted to nearest neighbor interactions, it can be treated exactly, including fully the nonlinearities, with the transfer operator method [4].

For a chain containing N units with nearest neighbor coupling, the classical partition function, given in terms of the Hamiltonian (1), can be expressed as

$$\begin{aligned} \mathscr{Z} &= \int_{-\infty}^{+\infty} \prod_{n=1}^{N} \mathrm{d}y_n \, \mathrm{d}p_n \, \mathrm{e}^{-\beta H} \\ &= \mathscr{Z}_p \times \int_{-\infty}^{+\infty} \mathrm{d}y_n \, \mathrm{e}^{-\beta f(y_n, y_{n-1})} \\ &= \mathscr{Z}_p \mathscr{Z}_y , \end{aligned} \tag{2}$$

where $f(y_n, y_{n-1})$ is the potential part of the Hamiltonian. The momentum part are readily integrated to give the usual kinetic factor for N particles $\mathscr{Z}_p = (2\pi m k_\mathrm{B} T)^{N/2}$. The potential part can be evaluated exactly [5–7] in the thermodynamic limit of a large system ($N \to \infty$) using the eigenfunctions and eigenvalues of the transfer integral operator

$$\int \mathrm{d}y_{n-1} \, \mathrm{e}^{-\beta f(y_n, y_{n-1})} \, \phi_i(y_{n-1}) = \mathrm{e}^{-\beta \epsilon_i} \, \phi_i(y_n) . \tag{3}$$

The calculation is similar to the one performed by Krumhansl and Schrieffer [6] for the statistical mechanics of the ϕ^4 field. It yields $\mathscr{Z}_y = \exp(-N\beta\epsilon_0)$, where ϵ_0 is the lowest eigenvalue of the operator. We can then compute the free energy of the model as $\mathscr{F} = -k_\mathrm{B} T \ln \mathscr{Z} = -(\tfrac{1}{2} N k_\mathrm{B} T) \ln(2\pi m k_\mathrm{B} T) + N\epsilon_0$ and the specific heat $C_\mathrm{v} = -T(\partial^2 \mathscr{F} / \partial T^2)$. The quantity which gives a measure of the extent of the denaturation of the molecule is the mean stretching $\langle y_\mathrm{m} \rangle$ of the hydrogen bonds, which can also be calculated with the transfer integral method [5] and yields

$$\begin{aligned} \langle y \rangle = \langle y_\mathrm{m} \rangle &= \frac{\sum_{i=1}^{N} \langle \phi_i(y) | y | \phi_i(y) \rangle \, \mathrm{e}^{-N\beta\epsilon_i}}{\sum_{i=1}^{N} \langle \phi_i(y) | \phi_i(y) \rangle \, \mathrm{e}^{-N\beta\epsilon_i}} \\ &= \langle \phi_0(y) | y | \phi_0(y) \rangle \\ &= \int \phi_0^2(y) y \, \mathrm{d}y , \end{aligned} \tag{4}$$

since in the limit of large N the result is again dominated by the lowest eigenvalue ϵ_0 associated with the normalized eigenfunction $\phi_0(y)$.

In the continuum limit approximation, the TI eigenvalue problem can be solved exactly, but experiments on proton exchange in DNA [8] show some evidence of exchange limited to a single base pair which suggests that discreteness effects can be extremely large in DNA. Therefore we have solved numerically the eigenvalue equation of the transfer operator [9] without approximations. The TI operator is symmetrized and the integral is replaced by sums of discrete increments, using summation formulas at different orders. The problem is then equivalent to finding the eigenvalues and eigenvectors of a symmetric matrix.

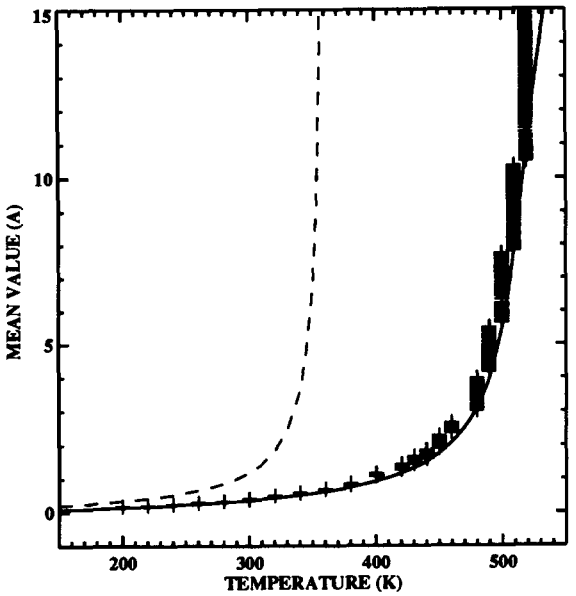

Fig. 1. Variation of $\langle y \rangle$ versus temperature: the dash line corresponds to the TI results in the continuum limit, the solid line gives the exact TI results obtained by numerical solution of the TI operator, and the plus signs correspond to molecular dynamics simulations.

Figure 1 compares the thermal evolution of $\langle y_m \rangle$ obtained with the continuum approximation and the exact numerical calculation, for the model parameters discussed in the next section. Both methods show a divergence of the hydrogen bond stretching over a given temperature, but the melting temperature given by the numerical treatment is significantly higher, pointing out the large role of discreteness in DNA dynamics if one uses realistic parameters for the model. The TI calculation shows that the specific heat has a broad maximum around the denaturation temperature.

3. Dynamics of the DNA molecule

The thermodynamics of our DNA model shows that it exhibits a thermal evolution that is qualitatively similar to the denaturation of the molecule observed experimentally. But this statistical approach does not give information on the *mechanism* of the denaturation, and in particular, does it start locally by the formation of denaturation bubbles in agreement with the experiments. In order to study this aspect, we have investigated the dynamics of the model in contact with a thermal bath by molecular dynamics simulation with the Nose method [10–11].

Beginning with the Hamiltonian H and the $2N$-dimensional phase space of a chain of N base-pairs with periodic boundary conditions, the fixed temperature canonical ensemble can be simulated by the addition of a single variable s, which regulates the energy flows, and an additional parameter M, which fixes the scale of the temperature fluctuations. Nose demonstrated that in this phase space of the extended Hamiltonian H', the microcanonical ensemble of H' is precisely the canonical ensemble of H at temperature T. This property is only exact for equilibrium properties, but investigations currently in progress [12] show that, provided that the characteristic time of the Nose thermostat, controlled by M, is properly chosen, it can also give reliable results for the dynamical properties. Most of the simulations have been performed with a chain of 256 base-pairs with periodic boundary conditions, but in order to achieve better statistics, some simulations have been performed on a Connection Machine-200 with 16384 base-pairs. Equations are integrated with a 4th order Runge–Kutta scheme with a time-step chosen to conserve H' to an accuracy better than 0.001% during a run.

We have chosen a system of units adapted to the energy and time scales of the problem. Energies are expressed in eV, masses in atomic mass unit (a.m.u.) and length in Å. The resulting time unit is $1\,\text{t.u.} = 1.0214 \times 10^{-14}\,\text{s}$. The choice of appropriate model parameters is a very controversial topic, as attested by the debate in the literature [13]. There are well established force fields for molecular dynamics of biological molecules, but they have been designed to provide a good description of the *small* amplitude motions of the molecule and are not reliable for the very *large* amplitude motions involved in the denaturation. In our model, the Morse

potential is an effective potential which links the two strands. It results from a combination of an attractive part due to the hydrogen bonds between two bases in a pair and the repulsive interaction between the charged phosphate groups on the two strands. The potential for the hydrogen bonds can be rather well estimated [14] but the repulsive part is harder to determine because the repulsion is partly screened by ions of the solvent. Consequently we had to rely on estimations. The parameters that we use have been chosen to give realistic properties for the model in terms of vibrational frequencies, size of the open regions, etc., but future work will be needed to confirm our choice. We do not expect however that a better choice would change *qualitatively* the results presented here. The parameters that we have chosen are: a dissociation energy $D = 0.04$ eV, a spatial scale factor of the Morse potential $a = 4.45\ \text{Å}^{-1}$, a coupling constant $K = 0.06$ eV/Å, a mass $m = 300$ a.m.u. The constant of the Nose thermostat has been set to $M = 1000$.

A first scan of the dynamics of the model is obtained by imposing a slow temperature ramp (200–540 K) that generates sets of states which are approximately equilibrated and are used then as initial states for simulations at constant temperature. Figure 2 shows a time evolution of the dynamics of the model at three temperatures. The stretching of the base-pairs is indicated by a grey scale, darker dots corresponding to larger stretching. Looking at this figure, one notices immediately two major features. First, as one moves along an horizontal direction, i.e., along the molecule for a given time, the amplitude of the stretching varies very much from site to site. This is especially true at high temperature, but it is still noticeable at 150 K, well below the melting temperature. This shows that there is no equipartition of energy in this nonlinear system, but on the contrary a tendency for the energy to localize at some points which is more and more pronounced as temperature increases. At high temperature, the figure shows large black regions which correspond to denaturated regions of the molecule. These black areas are the denaturation bubbles observed experimentally. At the highest temperature shown here (fig. 2c) they extend over 20 to 50 base pairs and their boundaries are sharp.

If the temperature is raised slightly above 540 K, the bubbles grow even more and finally extend over the whole chain: the molecule is completely denaturated. The second remarkable feature on fig. 2 can be observed by moving along a vertical line on the figure, i.e., following the time evolution of a given base pair. If one choses one region of the molecule in which the energy is concentrated, one can see alternating black and light-grey dots. This is due to an *internal breathing* of the localized excitations that oscillate between a large amplitude (black dots in the figure) and a small amplitude state (light dots) in a regular manner. These motions are the fluctuational openings of DNA. They exist even well below the denaturation temperature and coexist with denaturated bubbles in the high temperature range. Figure 2b shows that they play the role of precursor motions for the formation of the bubbles.

The calculation of the dynamical structure factor from the molecular dynamics results exhibits two types of excitations. In the high frequency range, one recognizes the phonon modes corresponding to linear motions of the chain. At low temperature their dispersion curve is well described by the linear dispersion curve resulting from the equations of motions of the model. Close to melting, on the contrary, most of the chain is on the plateau of the Morse potential and therefore experiences almost no restoring force that brings it back to $y = 0$. The dispersion curve is then the dispersion relation of a chain of harmonically coupled particles, without a substrate potential, i.e., a dispersion relation without gap. The variation versus temperature of the frequency of main phonon peak at wavevector $q = \frac{1}{2}\pi$ plotted on fig. 3 shows clearly the transition between a frequency belonging to the original dispersion curve at low T toward that of a gapless dispersion curve. This phonon softening

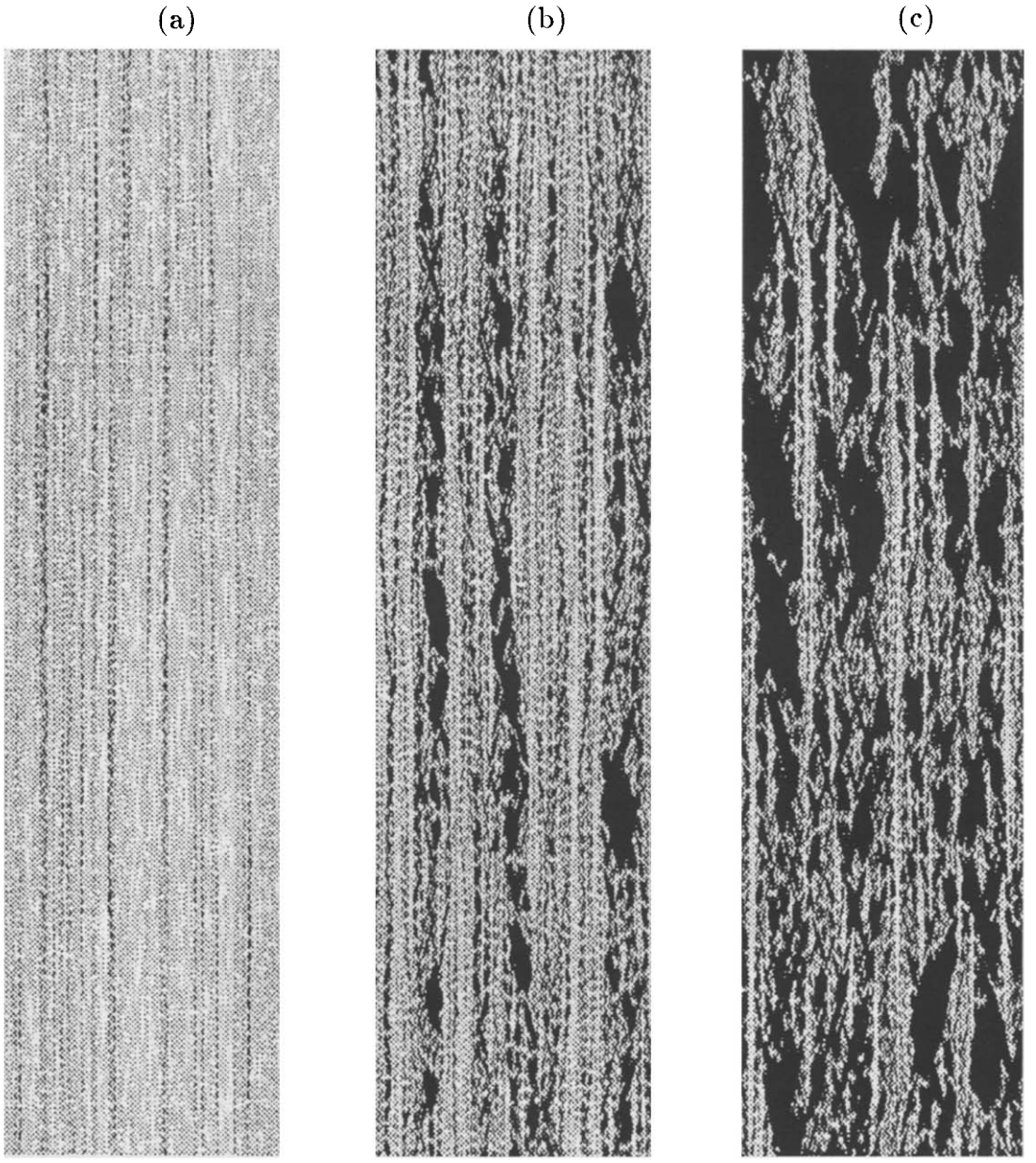

Fig. 2. Results of molecular dynamics simulations at three different temperatures (a) $T = 150$ K, (b) $T = 340$ K, (c) $T = 450$ K. The horizontal axis indicates the position along the 256 cells of the molecule and the vertical axis indicates time. The stretching y_n of the base-pairs along the molecule is indicated by a grey scale, the lighter grey corresponding to $y \leq -0.1$ Å and black indicating $y \geq 1$ Å. Therefore black regions show broken base-pairs.

should be observable experimentally in the vicinity of DNA melting transition. The second characteristic feature of the dynamical structure factor is a low frequency peak, associated to the fluctuational opening, which shifts to zero frequency as the denaturation bubbles form near the melting point.

Since the molecular dynamics simulations have found the two types of motions observed experimentally, fluctuational openings and bub-

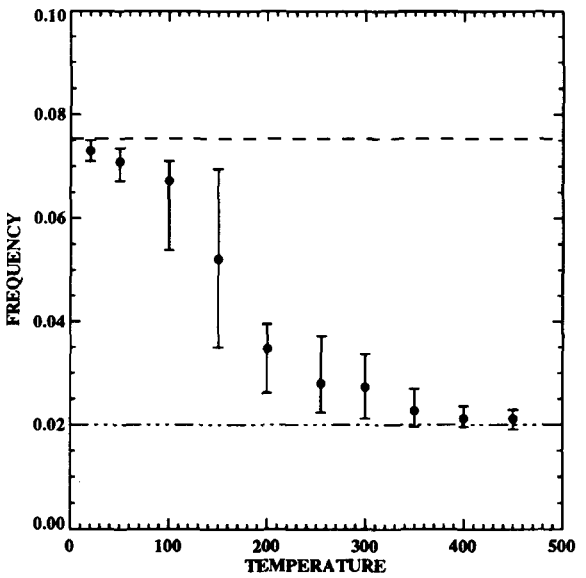

Fig. 3. Frequency of the phonon modes versus temperature for $q = \frac{1}{2}\pi$. The error bars indicate an interval where the frequency lies, at those temperatures for which a single mode cannot be identified. The horizontal lines correspond to the phonon frequency at the bottom of the well (dashed) and on the Morse plateau (dash-dot-dot-dot).

bles, it is interesting to see whether an analysis in terms of nonlinear excitations can explain these motions.

The fluctuational openings correspond to large amplitude breathing modes which are localized by nonlinear effects. As they extend only over a few base pairs, they are intrinsically discrete. Their existence and long term stability poses the general question of the existence of breathers in discrete Klein–Gordon models which has already attracted a great deal of attention without receiving a definite answer [15]. An analytical investigation of the DNA model presented above is difficult due to the Morse potential, but, since the fluctuational openings are intermediate amplitude oscillations, their study can be performed with a simpler potential that has qualitatively the shape of the Morse potential for small and intermediate amplitude but is more suitable analytically. The potential

$$V(y_n) = D(\tfrac{1}{2}y_n^2 - \alpha \cdot \tfrac{1}{3}y_n^3)\,, \tag{5}$$

for $y_n \leq 1/\alpha$ is convenient. Introducing the variable $w_n = \alpha y_n$ and the scaled time $\tau = \sqrt{K/m}\,t$, and defining the parameter $\omega_d^2 = D/K$, which measures the discreteness, the Hamiltonian of this simplified models becomes

$$H = \frac{K}{\alpha^2}\sum_n [\tfrac{1}{2}\dot{w}_n^2 + \tfrac{1}{2}(w_n - w_{n-1})^2 + \omega_d^2(\tfrac{1}{2}w_n^2 - \tfrac{1}{3}w_n^3)]\,. \tag{6}$$

Large amplitude localized breathers can be obtained for this model [16] by the lattice-Green functions method [17] developed recently to study local modes in nonlinear lattices. We seek long-lived oscillatory solutions under the form

$$w_n = \sum_{i=0}^{\infty} \phi_n^i \cos(\mathrm{i}\omega_b t) \tag{7}$$

where ω_b is the eigenfrequency of the fundamental mode ($i = 1$) and ϕ_n^i is the time independent amplitude of the ith mode (the ansatz has a dc part simply because of the asymmetry of the potential). Inserting this ansatz in the equations of motion and setting the coefficients of $\cos(\mathrm{i}\omega_b t)$ equal to one another and retaining only the first three terms, which gives already a good approximation, we obtain:

$$\omega_d^2\phi_n^0 - (\phi_{n+1}^0 + \phi_{n-1}^0 - 2\phi_n^0) = \omega_d^2[\phi_n^{0^2} + \tfrac{1}{2}(\phi_n^{1^2} + \phi_n^{2^2}]\,, \tag{8}$$

$$(\omega_d^2 - \omega_b^2)\phi_n^1 - (\phi_{n+1}^1 + \phi_{n-1}^1 - 2\phi_n^1) = \omega_d^2(2\phi_n^0 + \phi_n^2)\phi_n^1\,, \tag{9}$$

$$(\omega_d^2 - 4\omega_b^2)\phi_n^2 - (\phi_{n+1}^2 + \phi_{n-1}^2 - 2\phi_n^2) = \omega_d^2(2\phi_n^0\phi_n^2 + \tfrac{1}{2}\phi_n^{1^2})\,. \tag{10}$$

Using the lattice Green's function of the linear part of the above equations,

$$G(n-m, \omega_b) = \omega_d^2 \frac{1}{N}\sum_q \frac{\mathrm{e}^{\mathrm{i}q(n-m)}}{\omega^2(q) - \omega_b^2}\,, \tag{11}$$

where q is the wave vector inside the first Brillouin zone, ϕ_n^0, ϕ_n^1, ϕ_n^2 can be expressed in terms of the r.h.s. of eqs. (8)–(10), which gives a set of simultaneous nonlinear eigenvalue equations determining the eigenfrequency ω_b and the eigenfunctions ϕ_n^i. They can be solved numerically by iterations, starting with appropriate initial conditions. We fix the dc term, which amounts to choosing a particular amplitude for the solution. Only 15 iterations are necessary to determine the values with an accuracy of 10^{-4}. A numerical simulation of the dynamics of the chain with the solution determined as above as an initial condition shows that a large amplitude breather localized on very few lattice sites is extremely long lived, in agreement with the results of the molecular dynamics simulations of our DNA model. Consequently the very narrow fluctuational openings observed experimentally in DNA [8] and in our molecular dynamics investigations could well be discrete breathers stabilized by nonlinearity.

The study of the denaturation bubbles turns out to be much more difficult because they cannot be studied independently of the thermal effects. Due to the shape of the Morse potential, at $T=0$, any large amplitude opening bringing a set of neighboring base-pairs on the plateau of the Morse potential is unstable. This initial state would oscillate at very low frequency, but cannot stay in an open state since the bases in the open region are called back to their closing state by the bases which are still closed in the molecule and by the small downward slope of the Morse plateau. Therefore the existence of long lived open bubbles in our model, as they are seen in the molecular dynamics simulations, is fundamentally a thermal effect which cannot be studied by Hamiltonian mechanics, as we did for the fluctuational openings. One way around this difficulty is to *include temperature effects in the potential itself*. This is exactly what the self-consistent phonon (SCP) method does using a trial Hamiltonian for the calculation of the free energy of the system at temperature T [18]. Introducing $u_n = y_n - \langle y \rangle = y_n - \eta$ and two parameters Ω^2 and ϕ, we apply the SCP method [18] by considering the trial harmonic Hamiltonian

$$H_o = \sum_n \left[\tfrac{1}{2} m \dot{u}_n^2 + \tfrac{1}{2}\phi(u_n - u_{n+1})^2 + \tfrac{1}{2}\Omega^2 u_n^2 \right] . \tag{12}$$

The canonical partition function $\mathscr{Z}$ can be expressed as the product of the unperturbed partition function $\mathscr{Z}_0$ and a perturbation factor $\mathscr{Z}_1$:

$$\begin{aligned} \mathscr{Z} &= \int \prod_i \mathrm{d}u_i \, \mathrm{e}^{-\beta V} \\ &= \left(\int \prod_i \mathrm{d}u_i \, \mathrm{e}^{-\beta H_o} \right) \langle \mathrm{e}^{-\beta(H-H_o)} \rangle_0 \\ &= \mathscr{Z}_0 \mathscr{Z}_1 . \end{aligned}$$

The perturbation series for $\mathscr{Z}_1$ is calculated by expanding the exponential, and the logarithmic function; the coefficient of $(-\beta)^n/n!$ in the expansion of $\ln \mathscr{Z}_1$ is called the nth cumulant and is written [19] $\langle (H-H_o)^n \rangle_{0,c}$. Thus

$$\begin{aligned} \mathscr{F} &= -k_B T \ln \mathscr{Z} \\ &= -k_B T \ln \mathscr{Z}_0 - k_B T \ln \langle \mathrm{e}^{-\beta(H-H_o)} \rangle_0 \\ &= \mathscr{F}_0 - k_B T \sum_{n=1}^{\infty} \frac{(-\beta)^n}{n!} \langle (H-H_o)^n \rangle_{0,c} \\ &= \mathscr{F}_0 + \mathscr{F}_1 + \mathscr{F}_2 + \cdots . \end{aligned}$$

The first contribution $\mathscr{F}_0$ is the contribution of N harmonic oscillators,

$$\mathscr{F}_0 = -\tfrac{1}{2} k_B T \sum_{q=0}^{N-1} \ln \frac{2\pi k_B T}{\omega^2(q)} . \tag{13}$$

It can be shown [20] that the variational free energy $\mathscr{F} \leq \mathscr{F}_0 + \mathscr{F}_1 = \mathscr{F}_v$ gives an upper bound for the actual free energy. The only difficulty in evaluating $\mathscr{F}_v$ is the self-consistent substrate potential. It can be evaluated and gives an "effective" potential that keeps the shape of the Morse potential, but with a minimum which is

temperature dependent. The expression for the first order correction for the free energy is then:

$$\mathcal{F}_1 = (K - \phi)(\langle u^2 \rangle - \langle v^2 \rangle) - \tfrac{1}{2}\Omega^2 \langle u^2 \rangle + D(1 + \mathrm{e}^{-2a\eta + 2a^2\langle u^2 \rangle} - 2\,\mathrm{e}^{-a\eta + (a^2/2)\langle u^2 \rangle}), \tag{14}$$

with $\langle u^2 \rangle = \langle u_n^2 \rangle$ and $\langle v^2 \rangle = \langle u_n u_{n+1} \rangle$. Considering η, $\langle u^2 \rangle$ and $\langle v^2 \rangle$ as *variational* parameters, the conditions for $\mathcal{F}_v$ to be stationary give $\eta = \frac{3}{2}a\langle u^2 \rangle$, $\phi = K$ and $\Omega^2 = 2a^2 D \exp(-\frac{2}{3}a\eta)$. As η is an increasing function of temperature, Ω^2 decreases with T, which corresponds to the mode softening observed in the molecular dynamics simulations. Since $\langle u^2 \rangle$ and η are related, the minimization of $\mathcal{F}_1$ amounts simply to solving the equation

$$\eta = \frac{3ak_B T}{2N} \sum_p [2a^2 D\, \mathrm{e}^{-2a\eta/3} + 4K \sin^2(\pi p/N)]^{-1}. \tag{15}$$

In practice, eq. (15) is solved by a simple bisection method and fig. 4 shows the free energy $\mathcal{F}_v$ versus η at different temperatures. We see clearly that the self-consistent solution, which corresponds to the metastable minimum of the dashed and dash-dotted curves, disappears at $T_c = 411$ K to give a strictly decreasing function of η. Over T_c the only minimum is obtained for an infinite value of η, which does not correspond to a self-consistent solution of the problem. Consequently T_c can be identified as the denaturation temperature given by the SCP. Moreover, assuming that the free energy can play the role of an effective potential at temperature T, one can see that, as T approaches the melting temperature, the free energy exhibits a maximum followed by a decreasing part. In a "mechanical" equation of motion, such a potential would give stable open bubbles.

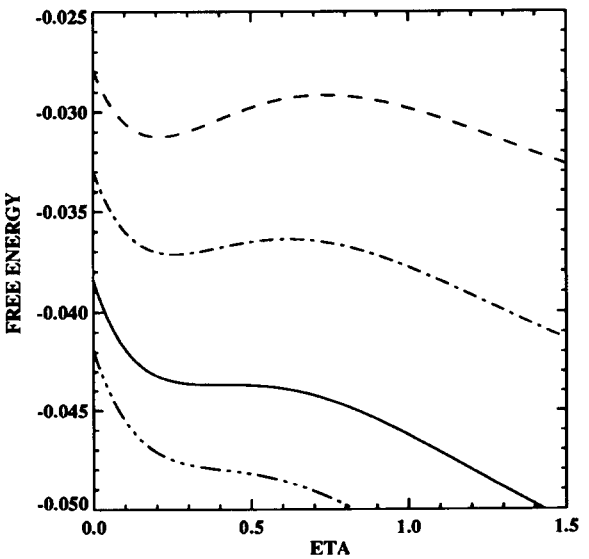

Fig. 4. Free energy $\mathcal{F}_v$ versus η at different temperatures: $T = 350$ (dashed), 380 (dash-dotted), 411 (solid), 430 (dash-dot-dot-dot) K.

4. Is DNA melting a one-dimensional phase transition?

The model discussed above has been able to describe some of the main features of DNA melting as it is observed experimentally. However there is a crucial point in which this model gives incorrect results, it is the *sharpness* of the phase transition. For an homopolymer, the experiments show that the melting occurs very abruptly over a temperature interval which is only a few K or even less. This poses a very fundamental question since DNA is basically a one-dimensional system, which is not expected to have a phase transition. We would like to conclude our paper by showing that, *within a one-dimensional model with short range interactions*, a sharp transition is possible if one takes into account properly the nonlinearity of the base stacking interaction [21]. The possibility of a phase transition in one-dimensional DNA was already examined within the Ising-model approach by Poland and Scheraga [22] and Azbel [23] who concluded that it can be attributed to cooperativity effects and to the role of the winding entropy released when the two strands separate. A simple extension of our DNA model (1)

can describe the dynamics of these effects and gives a very sharp transition in agreement with the experiments.

The stacking energy between two neighboring base pairs is described by the anharmonic potential:

$$W(y_n, y_{n-1}) = \tfrac{1}{2}K(1 + \rho\, e^{-\alpha(y_n + y_{n-1})}) \times (y_n - y_{n-1})^2 . \qquad (16)$$

This new intersite coupling, replacing the simple harmonic coupling of our previous approach, is responsible for qualitatively different properties. The choice of this potential has been motivated by the observation that the stacking energy is not a property of *individual* bases, but a character of the base *pairs* themselves [24]. When the hydrogen bonds connecting the bases break, the electronic distribution on bases are modified, causing the stacking interaction with adjacent bases to decrease. In eq. (16), this effect is enforced by the prefactor of the usual quadratic term $(y_n - y_{n-1})^2$. This prefactor depends on the *sum* of the stretchings of the two interacting base-pairs and decreases from $\frac{1}{2}K(1 + \rho)$ to $\frac{1}{2}K$ when either one (or both) base-pair is stretched. Although its form was chosen for analytical convenience, the qualitative features of potential (16) are in agreement with the properties of chemical bonds in DNA. They also provide the cooperativity effects that were introduced phenomenologically in the Ising models. A base pair that is in the vicinity of an open site has lower vibrational frequencies, which reduces its contribution to the free energy. Simultaneously a lower coupling along the strands gives the bases more freedom to move independently from each other, causing an entropy increase which drives a sharp transition. Our approach can be compared to recent views on structural phase transition in elastic media which stress that intrinsic nonlinear features characterize the physics of these transformations, and extend the standard soft mode picture [25,26]. It is important to notice that, although cooperativity in introduced through purely *nearest neighbor* coupling terms, it has a remarkable effect on the 1D transition.

Figure 5 shows the drastic change introduced by the anharmonic coupling on the specific heat of the model calculated by the transfer integral method [21]. The full curve corresponding to the anharmonic stacking interaction shows a sharp peak very similar to that one would expect from a first-order phase transition, whereas the harmonic coupling investigated before gives only a smooth maximum. This result suggests that, although DNA structure is very complicated, a simple nonlinear model is able to reproduce with a good agreement the main experimental features of its dynamics for the fluctuational openings as well as the melting curves. Indeed such a model does not pretend to give exact quantitative results that would fit exactly the experimental melting curve of an heteropolymer with all its small structure. Our feeling is that it is more important to get a basic understanding than to

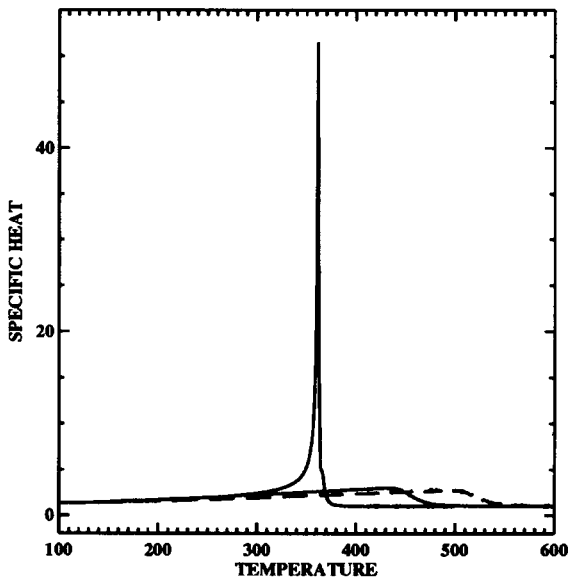

Fig. 5. Variation of the specific heat versus temperature. The very narrow peak corresponds to the anharmonic coupling case ($\alpha = 0.35$, $\rho = 0.5$), the dotted curve and the solid broad peak to harmonic coupling ($k' = 1.5k$ and $k'' = k$, respectively).

try to explain the very small experimental result. This does not mean that we think that the simple model that we discussed here is complete and should not be extended. But any extension will have to be measured for the new fundamental feature it brings against the unnecessary complications it introduces. Finally we would like to point out that, although we have discussed nonlinear dynamics of DNA, we have not introduced solitons in our picture (although the discrete breathers are probably close to being solitons). It is simply because they do not seem to be necessary to explain the dynamical features of DNA we are interested in. Solitons are marvellous objects but nonlinear science can also live without them.

Acknowledgements

The authors wish to acknowledge helpful and instructive discussions with S. Aubry and C. Reiss and a fruitful cooperation with A.R. Bishop. Two of us (T.D. and M.P.) are grateful for the hospitality of the center for Nonlinear Studies of Los Alamos National Laboratory. This work was performed with computing resources of the Advanced Computing Laboratory of Los Alamos National Laboratory. Part of the work has been supported by the CEC under contract SC1-CT91-0705. Work at Los Alamos is under the auspices of the DOE.

References

[1] J. Ladik and J. Cizek, Probable physical mechanisms of the activation of oncogenes through carcinogens, Int. J. Quantum Chem. 26 (1984) 67.
[2] R.M. Wartell and A.S. Benight, Thermal denaturation of DNA molecules: a comparison of theory with experiment, Phys. Rep. 126 (1985) 67.
[3] J.A. MacCammon and S.C. Harvey, in: Dynamics of Proteins and Nucleic Acids (Cambridge Univ. Press, Cambridge, 1988).
[4] M. Peyrard and A. Bishop, Statistical mechanics of a nonlinear model for DNA denaturation, Phys. Rev. Lett. 62 (1989) 2755.
[5] D.J. Scalapino, M. Sears and R.A. Ferrel, Statistical mechanics of one-dimensional Ginzburg–Landau fields, Phys. Rev. B 6 (1972) 3409.
[6] J.A. Krumhansl and J.R. Schrieffer, Dynamics and statistical mechanics of a one-dimensional model Hamiltonian for structural phase transitions, Phys. Rev. B 11 (1975) 3535.
[7] J.F. Currie, J.A. Krumhansl, A.R. Bishop and S.E. Trullinger, Statistical mechanics of one-dimensional solitary-wave-bearing scalar fields: exact results and ideal-gas phenomenology, Phys. Rev. B 22 (1980) 477.
[8] J.L. Leroy, M. Kochoyan, T. Huynh-Dinh and M. Gueron, Characterization of base-pair opening in deoxynucleotide duplex using catalyzed exchange of the imino proton, Mol. Biol. 200 (1988) 223.
[9] T. Schneider and E. Stoll, Classical statistical mechanics of the sine-Gordon and ϕ^4 chains. Static properties, Phys. Rev. B 22 (1980) 5317.
[10] S. Nose, A molecular dynamics method for simulations in the canonical ensemble, Mol. Phys. 52 (1984) 255.
[11] S. Nose, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1984) 511.
[12] N. Theodorakopulos and M. Peyrard, unpublished.
[13] L.L. Van Zandt, DNA solitons with realistic parameters, Phys. Rev. A 40 (1989) 6134;
M.A Techera, L.L. Daemen and E.W. Prohofsky, Comment on "DNA solitons with realistic parameters", Phys. Rev. A 42 (1990) 5033;
L.L. Van Zandt, Reply to "Comment on DNA solitons with realistic parameters", Phys. Rev. A 42 (1990) 5036, and references therein.
[14] Y. Gao, K.V. Devi Prasad and E.W. Prohofsky, A self-consistent microscopic theory of hydrogen bond melting with application to poly(dG)–poly(dC), J. Chem. Phys. 80 (1984) 6291.
[15] D.K. Campbell and M. Peyrard, Chaos and order in nonintegrable field theories, in: Chaos, eds. D.K. Campbell (AIP, New York, 1990), and references therein.
[16] T. Dauxois, M. Peyrard and C.R. Willis, Localized breather-like solution in a discrete Klein–Gordon model and application to DNA, Physica D 57 (1992) 267.
[17] S. Takeno, K. Kisoda and A.J. Sievers, Intrinsic localized vibrational modes in anharmonic crystals, Prog. Theor. Phys. Suppl. 94 (1988) 242.
[18] N. Bocarra and G. Sarma, Théorie microscopique des transitions s'accompagnant d'une modification de la structure cristalline, Physics (Long Island City) 1 (1965) 219.
[19] R. Kubo, The fluctuation–dissipation theorem, J. Phys. Soc. Japan 17 (1962) 1100.
[20] R.P. Feynman, Statistical Mechanics (Benjamin, 1972).
[21] T. Dauxois, M. Peyrard and A.R. Bishop, Entropy driven DNA denaturation, Phys. Rev. E 47 (1993) R44.
[22] D. Poland and H. Scheraga, Phase transitions in one-

dimension and the helix-coil transition in polyamino acids, J. Chem. Phys. 45 (1966) 1456.
[23] M. Ya. Azbel, Generalized one-dimensional Ising model for polymer thermodynamics, J. Chem. Phys. 62 (1975) 3635.
[24] C. Reiss, private communication (1991).
[25] J.A. Krumhansl and D.M. Alexander, in: Nucleic Acids and Proteins, eds. E. Clementi and R.H. Sarma (Academic Press, 1983).
[26] W.C. Kerr, A.M. Hawthorne, R.J. Gooding, A.R. Bishop and J.A. Krumhansl, First-order displacive structural phase transition studied by computer simulation, Phys. Rev. B 45 (1992) 7036.