NOLTA, IEICE

Invited Paper

Intrinsic localized modes in nonlinear models inspired by DNA

Michel Peyrard [1 a)] *and Guillaume James* [2 b)]

[1] *Ecole Normale Supérieure de Lyon, Laboratoire de Physique, 46 allée d'Italie, 69364 Lyon Cedex 07, France*

[2] *Laboratoire Jean Kuntzmann, Université de Grenoble and CNRS, BP 53, 38041 Grenoble Cedex 9, France*

[a)] *Michel.Peyrard@ens-lyon.fr*
[b)] *Guillaume.James@imag.fr*

Received April 18, 2011; Revised July 27, 2011; Published January 1, 2012

Abstract: We discuss nonlinear dynamic models for the fluctuational opening of the base pairs in DNA and show that a standard model which is satisfactory for time-independent properties has to be improved to properly describe the time scales of the fluctuations. The existence of an energy barrier for the closing of the base pairs has to be taken into account. This introduces a model which sustains a new class of Intrinsically Localized Modes (ILMs). We investigate their properties numerically, and then consider two simplified versions of the improved DNA model allowing an analytical study of some properties of those ILMs. The models are different because the effective barrier necessary for the existence of this new class of ILMs is obtained either through the on-site potential or through the nonlinear stacking interaction, but they nevertheless sustain similar nonlinear localized excitations. An extension of the usual anti–continuum limit has to be introduced for the analysis, and relies on a continuation of localized equilibria from infinity.

Key Words: DNA, energy localization, intrinsic localized modes

1. Introduction

The famous paper of A.J. Sievers and S. Takeno [1] pointed out the possibility that nonlinear coupling between the sites of an atomic lattice could lead to the localization of vibrations even if the absence of any defect or disorder. As noticed in this paper such modes may be thermally generated. However atomic crystals are usually characterized by strong interatomic bonds so that, in order to excite the nonlinearity one has to reach high excitation levels or high temperatures [2]. On the contrary biological molecules are "soft" objects, which often have to undergo large conformational changes at physiological temperature in order to perform their function, so that they are good candidates to observe nonlinear phenomena. Among them DNA is particularly interesting as an example of a nonlinear lattice because it has a regular structure. In his book *What is life?*, published in 1944 before the discovery of the DNA structure [3], Erwin Schrödinger introduced the concept of "aperiodic crystal" to describe a gene. The famous double helix of DNA, proposed by J.D. Watson and F.H.C.

Nonlinear Theory and Its Applications, IEICE, vol. 3, no. 1, pp. 27–51 DOI: 10.1587/nolta.3.27

Crick [4] confirmed this remarkable intuition. The molecule, with its two strands rolling around a stack of base pairs, appears as a one-dimensional lattice in which the perfect periodicity of the structure is only broken by the existence of different possibilities for the base pairs. The four bases, A, T, G, C, are associated in two types of pairs only, AT and GC. These "Watson–Crick pairs" have remarkably similar sizes in spite of the difference in their chemical formula, allowing the regular, crystal-like, structure of DNA. Within a pair the bases are linked by hydrogen bonds (two for AT, three for GC) which hold the two helices together.

However this remarkable structure is only the *average* structure. DNA is a highly dynamic entity. The hydrogen bonds linking the bases in a pair can break, allowing the bases to move out of the stack, into the surrounding solvent, before the base pair closes again. The lifetime of a closed base pair is only of the order of a few milli-seconds [5], and the lifetime of the open state of the pair is of the order of 100 nano-seconds. Experiments show that a base pair can open while its neighbors stay closed. Owing to its structure and dynamics DNA can therefore be viewed as a lattice of coupled oscillators, the base pairs, undergoing large amplitude fluctuations which are strongly localized.

Section 2 presents a simple model for the statistical physics and fluctuations of DNA. In this framework the local openings of the base pairs appear as *discrete breathers.* It is interesting that the biologists who first noticed these large fluctuations called them the "breathing of DNA" [6]. Independently from each other, the scientists in biology and theoretical physics had converged to the same vocable.

However, although the model quantitatively describes the statistical properties of DNA, when its dynamics is tested against experiments, the model presented in Sec. 2.1 appears to lead to time scales of the base-pair fluctuations which do not match experiments. Section 2.2 shows how it can be improved by modifying the potential that describes the interactions between the two bases in a pair. Instead of a simple well representing the bonding between the bases, one must introduce a barrier for the pair closing to take into account the interactions of the open bases with the solvent. This new model turns out to be interesting for nonlinear science because it sustains a new class of Intrinsic Localized Modes (ILM).

The last section examines the properties of these new ILMs more thoroughly by considering two special cases: i) the intra-pair potential has a barrier but the stacking interaction is assumed to be harmonic (Sec. 3.1), ii) the intra-pair potential is a Morse potential without barrier, but a strong nonlinearity in the stacking interaction introduces nevertheless a barrier for the closing of a base pair (Sec. 3.2).

2. A model for nonlinear dynamics and statistical mechanics of DNA

2.1 The PBD model and its statistical physics

When DNA is heated in the range 70° to 90° Celsius, depending on the base-pair sequence, the occurrence of base pair breaking increases so much that segments of consecutive pairs open together, giving rise to the so-called "DNA bubbles" and then the open regions grow until they extend to the full molecule, leading to a complete separation of the two strands. This thermal denaturation, also called "DNA melting" started to attract the attention of theoreticians soon after the discovery of the structure of the molecule [7–10] because it poses the question of a phase transition in a quasi one-dimensional system, where it is unexpected. The first models described a base pair as a two-state system, which can be either closed or open. Such models are convenient for statistical physics, but they cannot be used to study the dynamics of the opening and closing events because they do not describe the intermediate states. The PB model [11, 12] is the simplest model which tries to go beyond Ising-like descriptions. It does not intend to describe the helicoidal structure of the molecule and defines the status of base pair n in terms of a single real number y_n which measures the stretching of the bond linking the two bases. This allows a description of intermediate states between fully closed and fully open and therefore the investigation of the dynamics of the fluctuations of DNA at all amplitudes. The Hamiltonian of the PB model is

$$H = \sum_n \frac{p_n^2}{2m} + W_l(y_n, y_{n-1}) + V(y_n). \tag{1}$$

The first term corresponds to the kinetic energy of bases with momentum p_n and mass m. The next term $W_l(y_n, y_{n-1})$ describes the stacking interactions between the pairs. If base n moves out of the stack, i.e. when y_n increases, it tends to pull the adjacent bases too. In the PB model this coupling is described by its simplest form, the harmonic approximation

$$W_l(y_n, y_{n-1}) = \frac{1}{2}K(y_n - y_{n-1})^2. \tag{2}$$

The last term of the Hamiltonian is the intra-pair potential that links the two bases in a pair. It comes from the hydrogen bonds between the bases but is also affected by the repulsion between the charged phosphates of the backbone. This binding corresponds to a potential well around the equilibrium position $y_n = 0$. At large distance, the force between the bases vanish, i.e. the potential tends to a constant value, while at very short distance the two bases strongly repel each other, which corresponds to a sharp rise of the potential. A potential function which satisfies these conditions is the Morse potential (Fig. 1)

$$V_M(y_n) = D\left(e^{-\alpha y_n} - 1\right)^2. \tag{3}$$

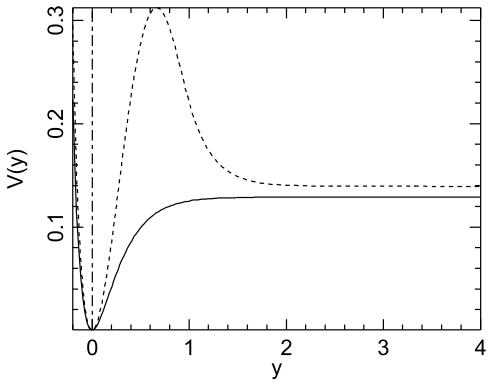

Fig. 1. Intra-pair potentials used in the DNA models. The full line shows the Morse potential of the PB and PBD models. The dotted line shows the potential with a barrier introduced in Sec. 2.2.

For this one-dimensional model, with nearest neighbor coupling, the statistical physics properties can be determined by an explicit calculation of the partition function

$$Z = \int \prod_n dp_n dy_n \; e^{-\beta H(\{p_n, y_n\})}. \tag{4}$$

In the strong coupling limit (large K) this can be done analytically by the transfer integral method [11, 13]. For lower values of K, which are appropriate for DNA, an efficient numerical integration can be performed [14, 15] to get exact results. A thermal denaturation, with a divergence of the mean value of the base-pair stretching $\langle y \rangle$ at a critical temperature T_c is found for this model. However, for realistic parameters, the results show that the model leads to a broad, second order, transition (Fig. 2) while experiments indicate that the transition is very sharp for a DNA homopolymer, where all base pairs are the same.

Therefore the simplest nonlinear model is not sufficient to quantitatively describe the physics of DNA. It can be improved by modifying the stacking potential [12, 16] W_l into a nonlinear expression

$$W(y_n, y_{n-1}) = \frac{1}{2}K\left[1 + \rho\, e^{-\delta(y_n + y_{n-1})}\right](y_n - y_{n-1})^2. \tag{5}$$

In this form the model is known as the PBD model. This expression for W means that the effective stacking interaction decreases from $K(1+\rho)$, when both interacting pairs are closed, to K when

Table I. Parameters for the calculations using the PBD DNA model. With energies in eV, lengths in Å and masses in atomic mass units, the time unit is 10^{-14} s.

$D_{AT} = 0.12905$ eV	$D_{GC} = 0.16805$ eV	
$\alpha_{AT} = 4.2$ Å^{-1}	$\alpha_{GC} = 6.9$ Å^{-1}	
$K = 0.00045$ eV/Å^2	$\rho = 50$	$\delta = 0.2$ Å^{-1}
$m = 300$ atomic mass units		

either one is open. This is reasonable because the opening of a base pair leads to a decrease of the overlap of the π–electrons of adjacent bases, which is mostly responsible of the stacking interaction. Equation (5) means that the DNA chain becomes much more flexible when base pairs are open. This is consistent with the observed sharp decrease of the persistence length of DNA when the two strands separate. This is the characteristic length of the decay of the angular correlations along the molecule, which measures the rigidity of the molecule. It is of the order of 50 nm for double stranded DNA, but it drops by a factor of about 50 when single strands of DNA are considered. This extra flexibility allows large fluctuations of the open segments, which rise the entropy of the system, hence decrease its free energy $F = E - TS$, making the melting transition easier. As a consequence, when open regions start to appear by heating, this leads to an entropy driven transition and a sharp "melting" of the double helix (Fig. 2).

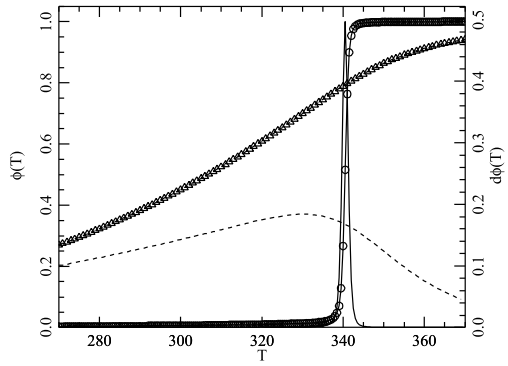

Fig. 2. Melting curves of homopolymer poly-A DNA calculated with the PBD and PB models. Full line with circles: fraction $\phi(T)$ of open base pairs versus temperature calculated with the PBD model using the parameters of Table I. A base pair is defined as open when its average stretching is beyond 1.5 Å. The full line without symbols is the derivative versus T of $\phi(T)$.
Dash line with triangles: fraction of open base pairs calculated with the PB model, i.e. by setting $\rho = 0$ in the parameters. As this also slightly changes the transition temperature, the parameter $D_{AT} = 0.105$ eV has been selected for this calculation instead of $D_{AT} = 0.12905$ eV for the PBD model to get similar transition temperatures for both cases. The dash line without symbols shows the derivative of the melting curve for the PB model, magnified by a factor 20 to make it visible on the same vertical scale as $d\phi(T)/dT$ in the PBD case.

The change in the stacking potential to introduce the anharmonic potential W is sufficient to give a model that quantitatively describes the thermal denaturation of DNA. The analysis of the melting profiles of many long DNA sequences shows that they can be calculated to a good accuracy with a single set of 7 parameters. This gives a predictive power to the PBD model to determine melting profiles of various sequences, provided they are long enough to avoid subtle effects which are observed when the thermal denaturation is studied with a high resolution on DNA segments of a few tens of base pairs [17]. The parameter set used in our study is listed in Table I.

To describe the two types of base pairs in actual DNA sequences, the intra-base potential $V(y)$ uses different parameters for the AT and GC pairs. In this work, as we want to emphasize nonlinear phenomena to localize the fluctuations in DNA, we will only consider homopolymers of DNA made of AT base pairs. Therefore *all results of the present paper correspond to homogeneous systems, without*

spatial disorder. They are obtained with $D = D_{AT}$, $\alpha = \alpha_{AT}$.

The parameter set of Table I has been applied to analyze a neutron scattering investigation of the thermal denaturation of films made of oriented DNA molecules [18]. A quantitative analysis of the neutron data, and further experiments to get some insight on the length of the fragments that stay closed immediately before the full melting transition of DNA [19], show that the PBD model gives a good quantitative analysis of the experiments. This indicates that this model can provide an accurate description of the statistical physics of DNA, i.e. of its *equilibrium* statistical properties.

2.2 Improving the model to describe the dynamics of the opening

Let us now examine the time-dependence of the base-pair fluctuations given by the PBD model. To compare with experimental data we need to study the dynamics of the model in contact with a thermal bath. This can be done with molecular dynamics (MD) simulations by coupling the model to suitably chosen fluctuating degrees of freedom. We use the Nose-Hoover thermostat for this study [20] with a chain of 5 thermostats to achieve a good ergodicity.

Figure 3 shows the amplitude of the fluctuations of the base-pair stretching at 300 K along a DNA homopolymer of 256 AT pairs, over a time interval of 0.4 ns. The blue background of this figure, which corresponds to closed base pairs according to the color scheme used to display the stretching of the base pairs, indicates that most of the base pairs are closed, as expected at 300 K. However the

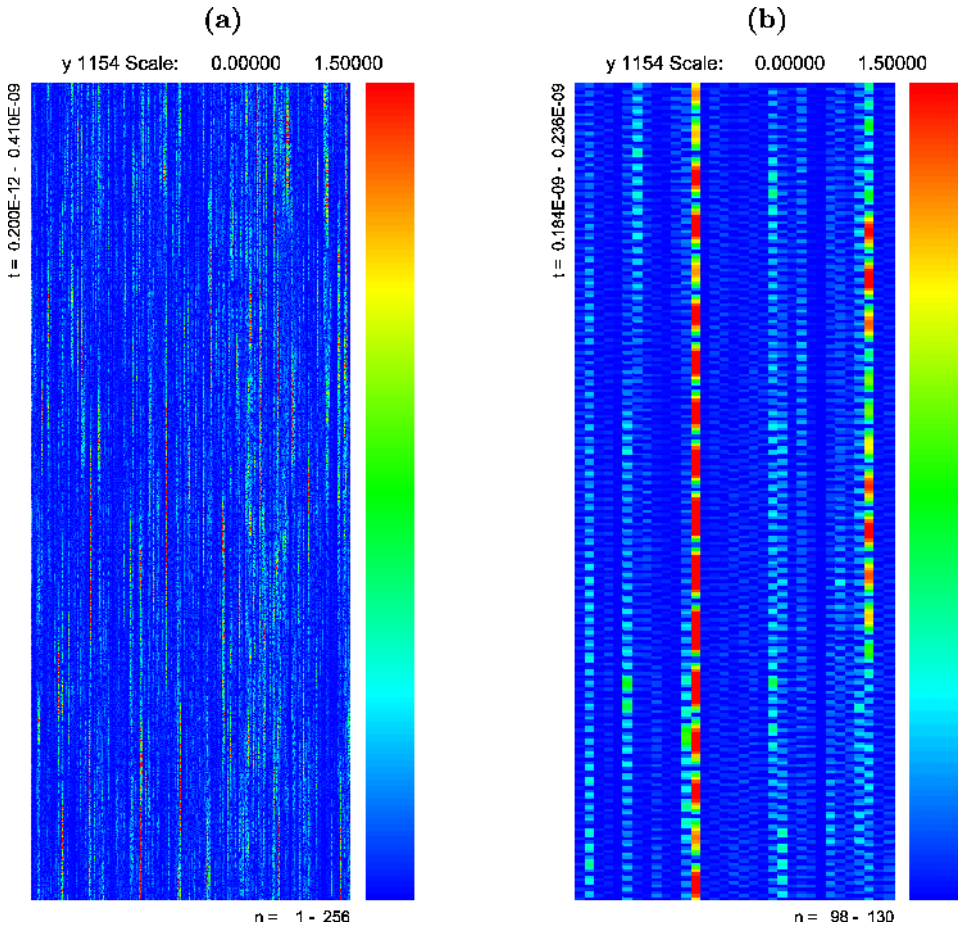

Fig. 3. Fluctuations of the PBD DNA model for a DNA segment of 256 AT base pairs at 300K. The horizontal axis corresponds to the position along the molecule and the vertical axis corresponds to time. It extends over a time range of 0.4 ns. The amplitude of the base-pair stretching is shown by a color scale, indicated on the right of each figure. Blue corresponds to closed base pairs ($y \leq 0$) while red shows the base pairs which have a stretching at least equal to 1.5 Å. The left Fig (a) shows the fluctuations of the whole DNA segment, while the right Fig (b) is a magnification by a factor 8 in space and time, only showing the central part of the left figure.

figure shows vertical lines in colors corresponding to large stretchings. They correspond to sites which open while their neighbors stay closed, in agreement with the experimental observations on DNA [5]. The magnification of one part of this diagram (Fig. 3(b)) shows how this opening occurs. The colored lines are actually dotted lines including blue segments i.e. $y \leq 0$, which alternate with segments where $y > 0$. This indicates that, at those particular sites, the base pair stretching oscillates, or, in other words that the *thermal fluctuations have generated discrete breathers in the DNA lattice.* As shown in Fig. 3(b) the amplitude of those breathers can exceed 1.5 Å (red color on the figure). Figure 1 indicates that this value corresponds to the beginning of the plateau of the Morse intra-pair potential, which means that the corresponding base pair can be considered as *temporarily open.*

Therefore the molecular dynamics simulations show that the PBD model of DNA has a small number of base pairs which temporarily open at temperatures well below the melting temperature of the double helix. Those openings are similar to the breathing of DNA observed by biologists. They occur locally, generally concerning only one base pair at a time at 300 K while the adjacent pairs only have small amplitude displacements. Qualitatively the results seem to agree with experimental observations, but, to determine the validity of the model one must check the quantitative properties of this DNA "breathing". This can be done by performing long MD simulations and making statistics of the times during which a base pair stays closed (the lifetime of a base pair) and how long does an average opening event lasts. Figure 4 shows the results of such a calculation. In this study a base pair is defined as "opened" when its stretching exceeds 1.5 Å, which is the beginning of the plateau on the Morse intra-pair potential, and closed again when y falls below 0.3 Å. The quantitative results slightly depend on the definition of the threshold but tests show that, as long as the threshold stays in the same order, the variation is only marginal.

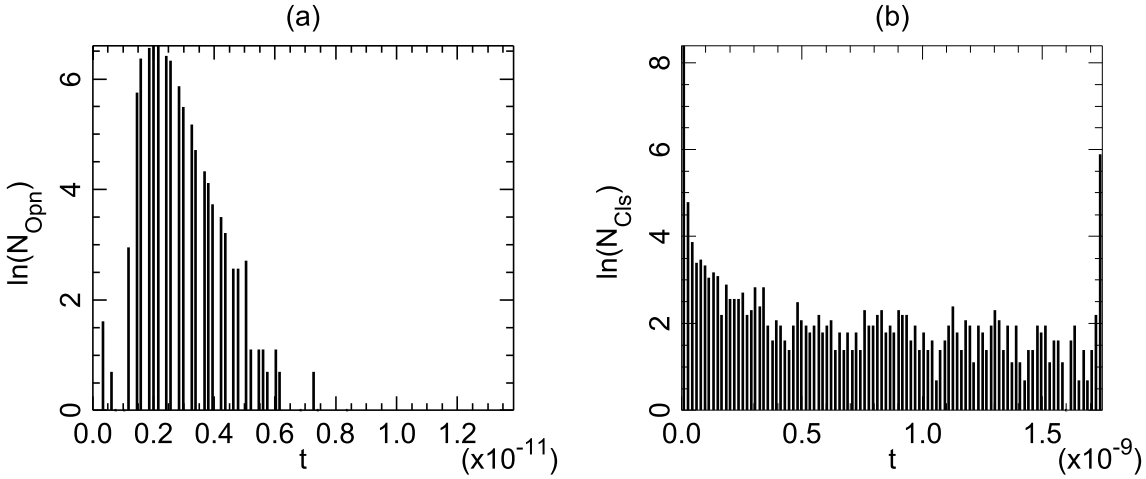

Fig. 4. Histogram of the duration of the open state (a) and lifetime of the closed base pairs (b) in the PBD model at 300 K, computed by simulation of DNA segment of 256 pairs for a total time of 6.5 ns.

Figure 4 shows that the durations of the opening events lie in a rather narrow range of 1 to 7×10^{-12} s. The average duration of the open state is 2.3×10^{-12} s. This time, in the ps range, is very short. It is of the order of the typical periods of the lowest frequency vibrational modes in DNA. This is consistent with the numerical observation that the opening of the base pairs occurs via discrete breathers. The durations of the opening events given by the PBD model are much smaller than the lifetime of the open base pairs of about 100 ns [5] inferred from experiments. The lifetimes of closed pairs observed in the MD simulations have a much broader distribution, but they are found to lie in the ns range (Fig. 4) with an average value of 2.9 ns, which is again far below the few ms estimated lifetime of an AT pair [5].

This disagreement with experiments, by orders of magnitude, shows that, although it is satisfactory when static properties of DNA are concerned, the PBD model is missing an essential ingredient to describe the dynamics of the molecule. Some ideas to find what is missing can come from all-atom

molecular dynamics simulations [21] biased to observe the free energy pathway associated to base-pair openings, in spite of the fact that it is a very rare event. Adding a geometrical constraint with a bias potential, it is possible to force the opening and observe the probability distribution of the fluctuations to reconstruct a free energy corrected from the bias potential. The results are highly sensitive to the detail of the solvent and counterion dynamics but they show nevertheless that the free energy of the open state may have a shallow minimum, which explains why a base pair can stay open for a long time rather than vibrating as a breather around the closed state. In defining the potentials of a mesoscopic model like the PBD model one should think that the model only describes a subset of all degrees of freedom. Therefore one micro-state of the model actually corresponds to a set of microscopic states of the real systems, where all the degrees of freedom which are not present in the mesoscopic model are fluctuating. It means that the "potentials" of a mesoscopic model are actually free energies averaged over the degrees of freedom which are not represented at the mesoscopic scale. This suggests that the Morse intra-pair potential, which had been motivated by only considering the stretching of the base pairs, should be modified to take into account other hidden degrees of freedom which generate a barrier for reclosing. This barrier could have different origins. The open bases can make hydrogen bonds with the solvent, which must be broken before the base pair can close again. Moreover, when they are out of the stack that they form in the double helix, the bases gain new degrees of freedom; Their plane can rotate, which makes reclosing difficult.

To model these effects we have therefore introduced a new intra-pair potential

$$V_h(y) = \begin{cases} A\left[e^{-\alpha y} - 1\right]^2 & \text{if} \quad y < 0, \\ ay^2 + by^3 + cy^4 & \text{if} \quad 0 \leq \sigma y \leq d, \\ D + Ee^{-\sigma y}\left(y + \frac{1}{\sigma}\right) & \text{if} \quad \sigma y > d. \end{cases} \tag{6}$$

This potential, which replaces the Morse potential of the PBD model is plotted in Fig. 1 (dotted line). It is determined by the parameters D, E, α, σ, and the choice of d. The other parameters a, b, c, A, are derived by imposing the continuity of the potential and its two first derivatives in $y = 0$ and $y = d/\sigma$. The analytical form has been chosen to preserve the expression of the Morse potential for $y < 0$, and for analytical convenience in the range $\sigma y > d$, while giving the required shape of a potential that tends to a constant value for large y, and with a barrier between the open and closed states. To study DNA with this new model we have used $D = 0.13945$ eV, $E = 10$ eV Å^{-1}, $\alpha = 4.2$ Å^{-1}, $\sigma = 5.0$ Å^{-1} and $d = \sigma$.

2.3 Properties of the improved DNA model

Let us now examine the validity of this new model to describe the statistical and dynamic properties of the base pair fluctuations in DNA.

The parameters of the potential V_h have been chosen to give the same thermal denaturation temperature as the PBD model for a DNA homopolymer made of AT base pairs. Figure 5 shows that this

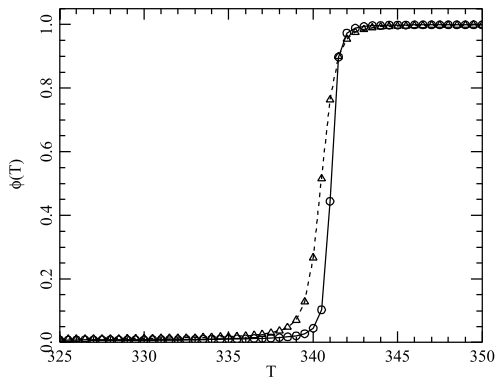

Fig. 5. Comparison of the melting profiles for poly-A DNA sequences computed with the modified PBD model with a closing barrier (full line with circles) and the PBD model with the Morse potential (dotted line with triangles).

choice leads to melting profiles for DNA homopolymers which are very similar for the two models, although the potential with a barrier gives a slightly sharper transition. This feature has also been observed in another statistical physics study of DNA where a barrier attributed to the solvent has been introduced in order to sharpen the transition [22]. Studies of inhomogeneous sequences show that this extra sharpness of the transition for a homopolymer also leads to melting profiles which are more sensitive to the details of the local sequence than with the PBD model, however when thermal averages are concerned the two models give results which are comparable.

2.4 Dynamics of the fluctuations

Conversely strong differences between the two models appear when one considers dynamical effects, i.e. time dependent properties.

Figure 6 shows the amplitude of the fluctuations of the base-pair stretching at 300 K in the model with potential V_h. It should be compared with Fig. 3 for the PBD model. Except for the intra-pair potential all conditions are the same for the two figures, for the MD simulations and for the plots which use the same color scale. Instead of many short-lived open states (corresponding to red spots on the figures) for the PBD model, Fig. 6 shows a single, long-lived, open state that extends over the whole time interval shown in the figure, i.e. 0.41 ns. The magnified image provides further information

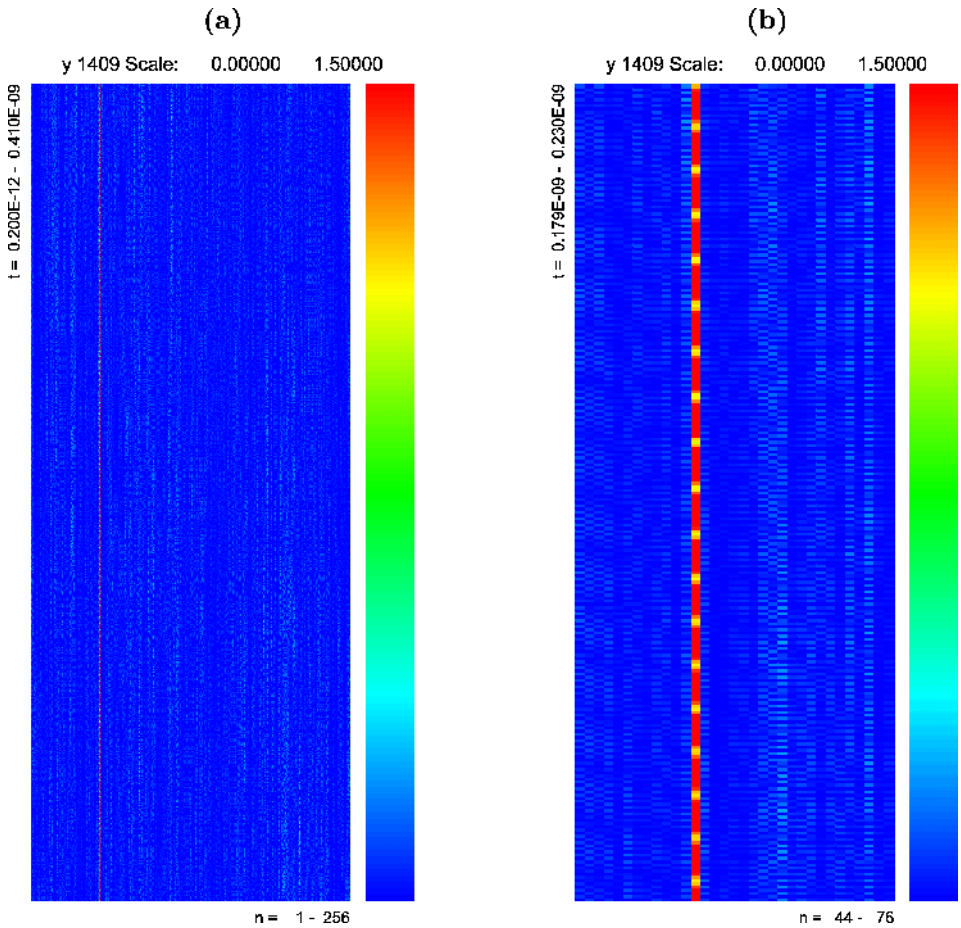

Fig. 6. Fluctuations of the DNA model with a barrier in the intra-pair potential, for a DNA segment of 256 AT base pairs at 300 K. The horizontal axis corresponds to the position along the molecule and the vertical axis corresponds to time. It extends over a time range of 0.4 ns. The amplitude of the base-pair stretching is shown by a color scale, indicated on the right of each figure. Blue corresponds to closed base pairs ($y \leq 0$) while red shows the base pairs which have a stretching at least equal to 1.5 Å. The left Fig (a) shows the fluctuations of the whole DNA segment, while the right figure is a magnification by a factor 8 in space and time, showing only one part of the left figure.

about the nature of this open state. The maximum stretching at the open site is not constant. It varies between a value larger than 1.5 Å (red color) and a lower value of about 1.1 Å (yellow color). Therefore this figure suggests that, in the model with a barrier in the intra-pair potential, the thermal fluctuations have generated a new kind of nonlinear localized excitation, different from the breathers oscillating around the minimum of the intra-pair potential ($y = 0$) observed in the PBD model.

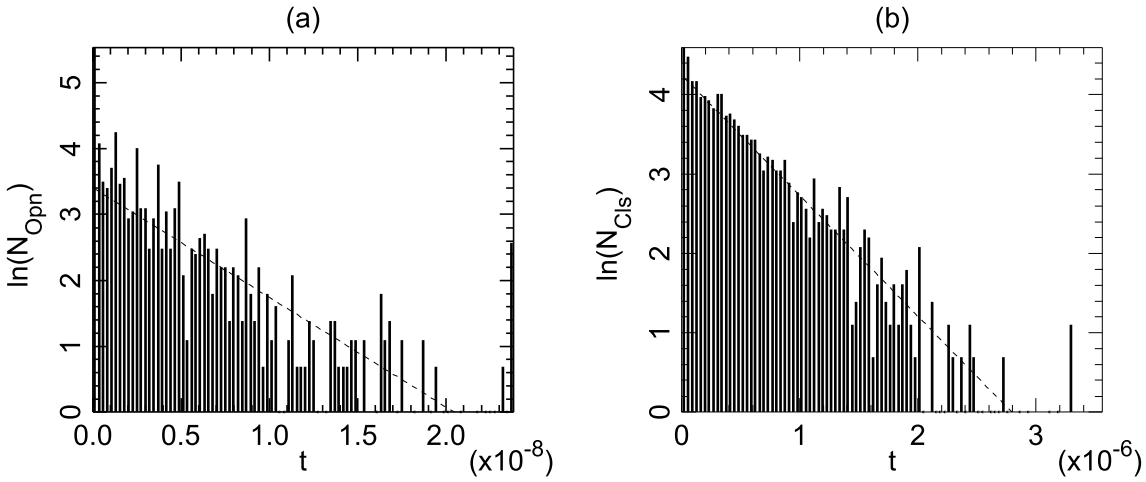

Fig. 7. Histogram of the duration of the open state (a) and lifetime of the closed base pairs (b) in the PBD model at 300K, computed by simulation of DNA segment of 256 pairs for a total time of 3.28 μs.

Before studying these excitations in the next subsection, let us examine the consequences of the barrier in the potential on the time scales of the open and closed states of DNA. The histograms of the duration of the open states and lifetimes of the closed base pairs in the new model (Fig. 7) show that the distributions of these characteristic times is qualitatively changed by the barrier in the intra-pair potential. Both the duration of the open states and the lifetimes of the closed pairs have an exponentially decaying distribution. The average value of the duration of the open state is 3.3 ns i.e. more than three orders of magnitude longer than for the PBD model with a Morse potential. Similarly the lifetime of a closed base pair is 0.59 μs in the presence of the potential barrier, i.e. 2000 times longer than for the PBD models. The duration of the open state is still about 30 times smaller than the measured value, while the lifetime of a closed pair is still about three orders of magnitude smaller than the values measured experimentally. The model is still not quantitatively correct to describe the dynamics of the base-pair fluctuations in DNA, but the improvement brought by the new intra-pair potential is nevertheless really significant. The discrepancy with experiments shows the limits of an oversimplified model where a single degree of freedom is used to describe the complex phenomena that take place in the actual molecular structure of DNA in contact with the solvent. Using an effective potential such as V_h can correct some of the weaknesses of the simplified description but not all of them.

Figure 8 shows that the temperature dependence of the lifetime of a closed base pair t_c and the duration of the open state t_o follow Arrhenius laws

$$t_o = t_1 e^{E_1/k_B T} \qquad t_c = t_0 e^{E_0/k_B T}, \tag{7}$$

which are characteristic of a thermal activation. As discussed in the next subsection the activation energies $E_0 = 0.250$ eV and $E_1 = 0.086$ eV can be estimated from the properties of the ILM of the model.

2.4.1 A new class of ILM

As shown by Fig. 6(b) the MD simulations of the improved DNA model find that thermal fluctuations can lead to highly localized nonlinear excitations in which the base-pair stretching at one site is large

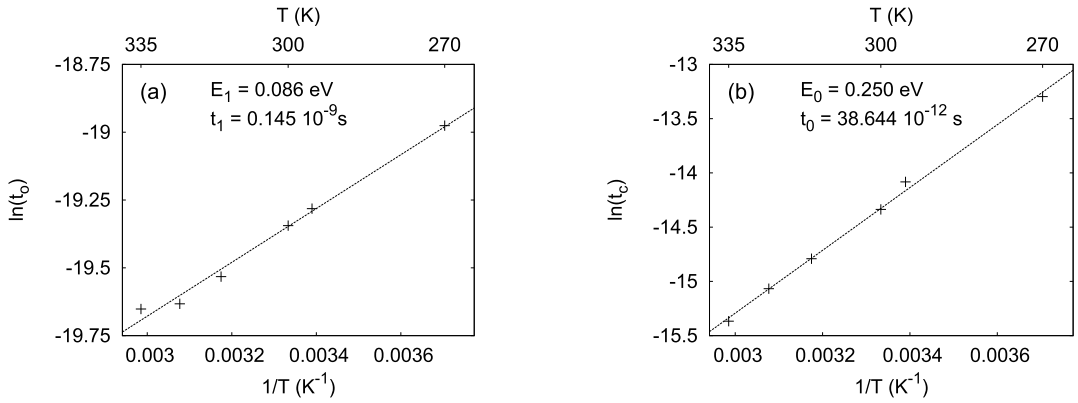

Fig. 8. Arrhenius plot of the duration of the open state t_o (a) and of the lifetime of a closed base pair t_c (b) in the model with a barrier in the intra-pair potential. The figure shows $\ln(t_0)$ and $\ln(t_c)$ versus $1/T$.

and oscillates around a non-zero average, while the neighboring sites only show a very small stretching. To understand the nature of those ILM it is useful to consider a *reduced system* made of only 3 sites, the central site i_0 of the ILM, and its two adjacent sites $i_{0\pm1}$. In a first approximation the two adjacent sites are assumed to be fixed at their equilibrium value $y_{i_{0\pm1}} = 0$. The potential energy of this reduced system is

$$V_{\rm eff}(y_{i_0}) = W(y_{i_0}, 0) + W(0, y_{i_0}) + V_h(y_{i_0}). \tag{8}$$

It can be viewed as an *effective potential* for the stretching at site i_0.

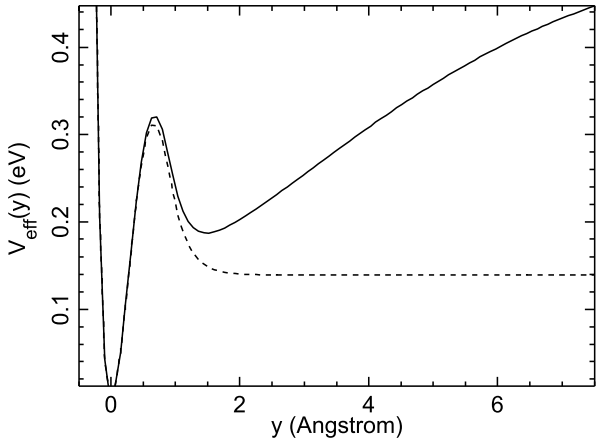

Fig. 9. The effective potential $V_{\rm eff}(y)$ defined by Eq. (8) (full line) and the intra-pair potential $V_h(y)$ (dotted line). The parameters for W and V_h are those listed above for the improved DNA model with potential V_h.

Figure 9 shows the shape of this effective potential. It is an asymmetric double-well (DW) potential. The deeper well is the ground state at $y = 0$, but the combination of the stacking interaction W and intra-pair potential V_h generates a second, shallower, minimum, with a higher energy, at $y = y_0 \neq 0$.

The shape of the potential immediately suggests the possibility of two types of localized modes. The first one is a "conventional" breather, oscillating with a large amplitude around the minimum $y = 0$ of the on-site potential V_h. The second one is a new kind of breather, which oscillates around y_0 which is an unstable point for the on-site potential V_h, stabilized in the effective potential by the coupling with the two adjacent sites. We call Double-Well breather (DW-breather) this type of ILM as a reminder that its existence comes from the double well shape of the effective potential. As discussed below it can oscillate in the second well only or even span both wells when its amplitude becomes very large.

However this conjecture is based on the reduced system of three atoms, and it must be checked for the full nonlinear lattice. We have used MD simulations to confirm its validity. These calculations

are performed *without any thermal bath* as we want to study an intrinsic property of the model. We simulate the equations of motions that derive from the Hamiltonian of the DNA model

$$m\frac{d^2y_n}{dt^2} + \frac{\partial W(y_n, y_{n-1})}{\partial y_n} + \frac{\partial W(y_n, y_{n+1})}{\partial y_n} + V_h'(y_n) = 0. \tag{9}$$

The lattice comprises $N = 512$ sites. The initial condition imposes a given amplitude to the central site $y_{N/2}(t = 0) = y_c$, all other sites starting from $y_n = 0$. All initial velocities are equal to 0.

This initial condition does not correspond to an exact periodic solution of the equations of motions and we observe the emission of small amplitude waves from the central site, together with a large oscillatory motion at this site. In the finite lattice these waves would be reflected at the boundaries and perturb the long term evolution of the dynamics near the center. This can be avoided by introducing a damping term $\gamma(n)dy_n/dt$ in Eq. (9) in two regions near the end. In our calculation these regions extend over 64 sites at both ends and the damping coefficient decays linearly from $\gamma(n) = 0.2$ at the boundaries of the lattice to 0 at the boarder of the damped regions.

Figure 10 shows that, after a transient during which the central site emits very small amplitude waves which are damped out at the boundaries, the dynamics evolves towards a stable oscillatory

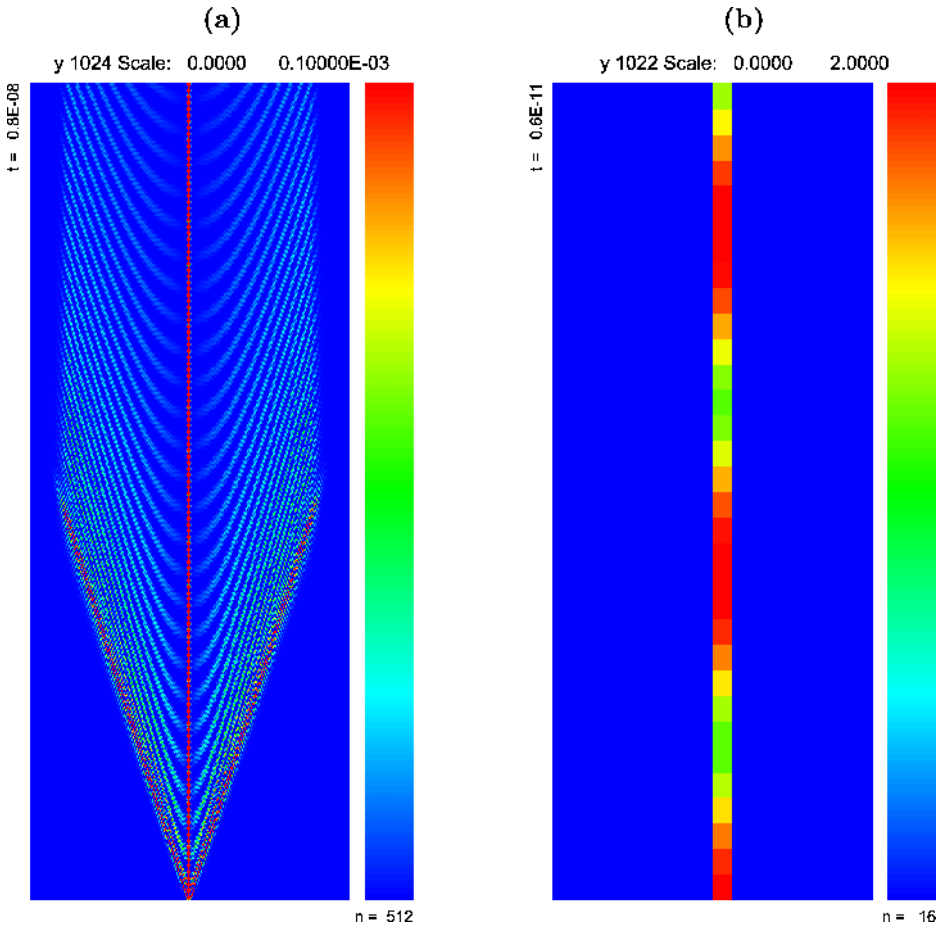

Fig. 10. Dynamics of the improved DNA model when only the central particle is excited, in the presence of damping at both ends of the lattice. The horizontal axis corresponds to the position along the lattice. The vertical axis corresponds to time. (a) Initial stage of the dynamics, showing the complete lattice (512 particles) between $t = 0$ and $t = 0.8\ 10^{-8}$ s. The color scale has been selected to show very small amplitude waves (in the range from 0 to $0.1\ 10^{-3}$ Å. (b) steady state solution obtained after $t = 1.2\ 10^{-8}$ s. The picture is magnified to show only the 16 central particles of the lattice. The color scale extends from 0 to 2 Å and the time scale is also magnified to display a few periods of the breather only (the time interval shown extends over 6 ps only). This figure shows that the DW-breather oscillates between a minimum of about 1 Å (green color) and a maximum reaching 2 Å (red color).

state, with a DW-breather at the excited site. We can then record the properties of this breather, its energy, amplitude, frequency, as a function of the amplitude of the initial excitation. The results are shown in Figs. 11 and 12.

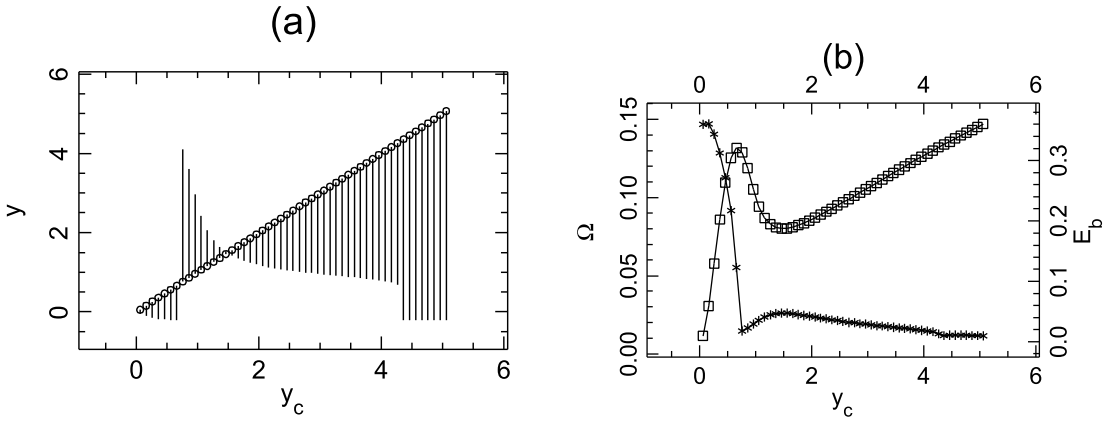

Fig. 11. Properties of the breathers created by exciting the central particle of the lattice. (a) amplitude of its oscillations versus the amplitude y_c of the excitation of the central particle. For each value of y_c the vertical line extends between the extrema of the oscillations of the breather. The circles show the initial value $y_{N/2} = y_c$. (b) Frequency (stars) and energy (squares) of the breather versus y_c.

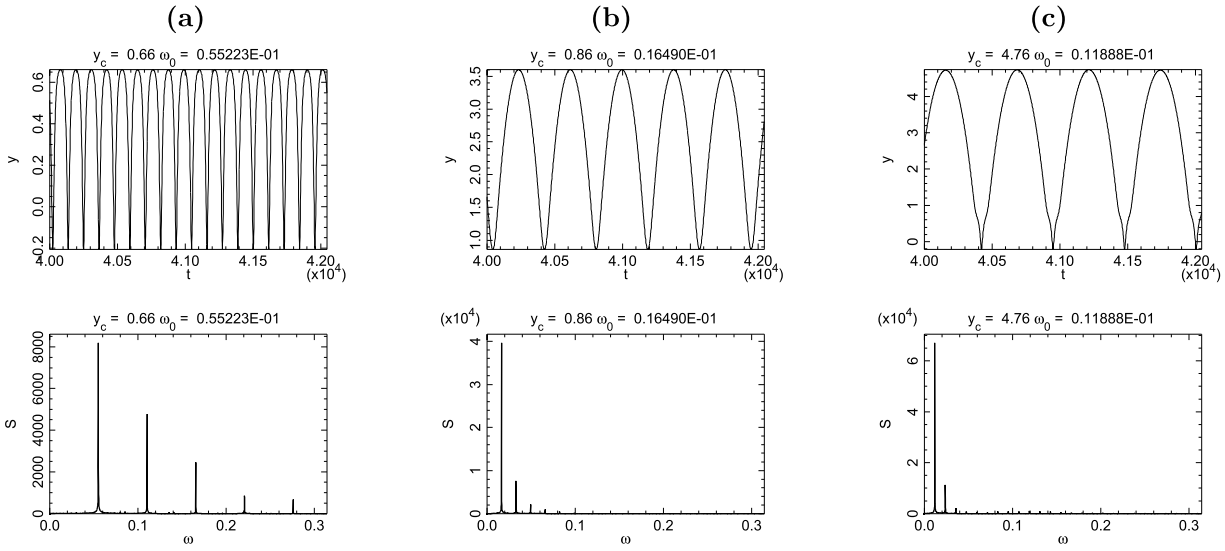

Fig. 12. Some examples of the the breathers created by exciting the central particle of the lattice. In each case the top image shows the time dependence of the oscillations $y_{N/2}(t)$ of the central particle, and the bottom image shows the Fourier transform of $y_{N/2}(t)$, i.e. the spectrum of the breather. (a) for an excitation of amplitude $y_c = 0.66$ Å the breather oscillates in the main potential well. (b) an excitation of amplitude $y_c = 0.86$ Å creates a DW-breather, which oscillates in the second well. Its frequency is much lower than the frequency of the breather in the main well. (c) a very large excitation $y_c = 4.76$ Å creates a DW-breather that spans the two wells.

As shown by these figures the model has a rich variety of ILMs. When the excitation is sufficiently small, it generates an ordinary breather that oscillates around the bottom of the main well ($y = 0$). Its spectrum, which is quasi-harmonic at very low amplitude, becomes richer when the amplitude increases, as shown in Fig. 12(a). When y_c exceeds a threshold defined by the maximum of the effective potential of Fig. 9, the nature of the breather changes. It turns into a DW-breather which oscillates in the second well. Its frequency drops sharply, in agreement with the shallow shape of the

second well. As shown by Fig. 11(b) the frequency of the DW-breather always stays small, and shows a maximum when the amplitude drops to 0 around the minimum of the second well. In the vicinity of this point the calculation of the spectrum of the breather shows that it is again single-frequency, i.e. the DW-breather is quasi-harmonic. When y_c passes a second threshold $y_c \approx 4.7$ Å, the nature of the oscillation of the DW-breather changes because it spans both wells. This is accompanied by a new decrease of the frequency. Figure 12 shows that the oscillation $y_{N/2}(t)$ of the central particle is highly anharmonic in this case, and this shows up in the complex spectrum of the breather.

Figure 11(b), showing the energies of the DW-breathers, provides some insight on the activation energies measured in MD simulations for the lifetime of a base pair $E_0 = 0.250$ eV and for the duration of the open state $E_1 = 0.086$ eV. It appears that E_0 is slightly above the minimum of the DW-breathers energy (≈ 0.2 eV) while E_1 corresponds to the energy difference between E_0 and the maximum energy of the DW-breather (≈ 0.32 eV). This is reasonable because it corresponds to the energy that should be given to the DW-breather associated with the open state to allow it to overcome the energy barrier which separates the DW-breathers in the second minimum from the breathers around the closed state (Fig. 11(b)).

3. ILMs in nonlinear models inspired by DNA

The results on the DNA improved model have shown that it can sustain a rich variety of intrinsic localized modes. However this model is complex because it has both an on-site (intra-pair) potential V_h with a barrier and a coupling W which is nonlinear. This is consistent with the properties of DNA, and necessary to quantitatively describe the experimental results, but it makes analytical studies difficult. In this section we show that this complexity is not necessary if we are interested in the new class of excitations, the DW-breathers, that the DNA model has exhibited. We examine two models, inspired by the DNA model, but simpler, and show that they can also sustain DW-breathers.

3.1 Model 1: on-site potential with a barrier, and harmonic coupling

In this first simplified version of the DNA model, the intra-pair potential with a barrier, V_h (Eq. (6)) is preserved but the stacking interaction W is replaced by the linear approximation W_l (Eq. (2)) so that the equations describing the dynamics of the model are

$$m\,\frac{d^2y_n}{dt^2} + V_h'(y_n) = K\,(y_{n+1} - 2y_n + y_{n-1}), \quad n \in \mathbb{Z}. \tag{10}$$

A preliminary step to find time-dependent solutions is to determine the stationary points, i.e. the equilibria of the system which are determined by the system

$$K\,(y_{n+1} - 2y_n + y_{n-1}) = V_h'(y_n), \quad n \in \mathbb{Z}. \tag{11}$$

If they are spatially localized they must moreover satisfy

$$\lim_{n\to\pm\infty} y_n = 0. \tag{12}$$

A standard method to analyze spatially localized equilibria in nonlinear lattices is the anti–continuum limit technique (see [23–25] and section 9 of [31]). The usual anti–continuum limit starts from an exact solution y_n^0 of (11)–(12) obtained for $K = 0$. For a finite number of sites n, y_n^0 is set to the local maximum of V_h at the top of the potential barrier, and to the global minimum $y = 0$ at all the remaining sites (in infinite number). In a second step, one continues this solution for $K \approx 0$ using the implicit function theorem, which yields an exponentially localized solution of Eq. (11). In the present case, all these solutions are unstable at small coupling because the only nontrivial critical point of V_h is a local maximum. Consequently, the usual anti–continuum limit technique cannot be readily applied to the model that we consider here in order to find stable localized equilibria and it must be extended.

Localized equilibria can nevertheless be obtained from an extension of the anti–continuum limit, introduced in [33]. This method relies on the observation that V_h admits a "critical point at infinity",

which should allow to construct highly localized equilibria of (11) at small coupling. This requires to extend the above perturbative analysis to the singular situation when the excited site lies at infinity on the plateau of the potential V_h for $K = 0$. In [33], localized equilibria are obtained by "continuation from infinity" for arbitrarily small values of K, but without reaching $K = 0$. The perturbative analysis is performed near an approximate solution $\tilde{y}_n$ of Eqs. (11) and (12) for $K \approx 0$, instead of an exact solution for $K = 0$ as in the classical anti–continuum limit. This initial guess satisfies $\tilde{y}_n = 0$ for $n \neq 0$, and $\tilde{y}_0$ is a critical point of the modified potential

$$V_K(y) = V_h(y) + Ky^2,$$

which incorporates a restoring force due to nearest-neighbors. This is the effective potential introduced in Eq. (8) written in the context of the particular model that we study here. As shown by Fig. 13, it has a double-well shape, qualitatively similar to the shape of the effective potential found for the improved DNA model (Fig. 9) so that we can expect to find the same kind of ILMs in the two models.

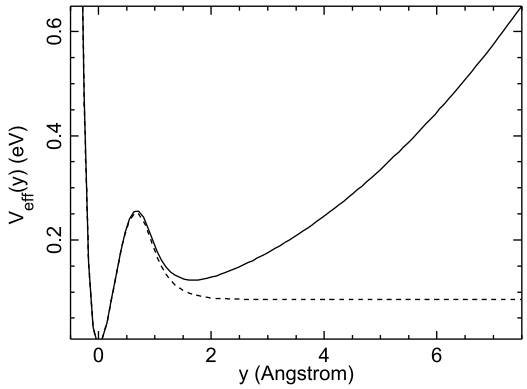

Fig. 13. Effective potential for model 1, with the linear stacking interaction W_l and the intra-pair potential V_h. The parameters are $K = 0.01$ eV Å^{-2}, $D = 0.0857$ eV, $E = 4$ eV Å^{-1}, $\alpha = 4$ Å^{-1}, $\sigma = 4$ Å^{-1}.

The special potential V_h defined by Eq. (6) allows simple computations because (for K small enough) the solution of $V_K'(\tilde{y}_0) = 0$ in the interval $[d\,\sigma^{-1}, \infty)$ reads explicitly

$$\tilde{y}_0 = -\frac{1}{\sigma}\,\ln\,(\frac{2K}{\sigma E}). \tag{13}$$

An exact solution $y_n^{\rm eq}$ of (11)–(12) is obtained in [33] for $K \approx 0$, using the contraction mapping theorem in some neighborhood of $\tilde{y}_n$.

Theorem 1 There exists a constant $K_0 > 0$ such that for all $K \in (0, K_0)$, problem (11)–(12) admits a solution $y_n^{\rm eq}(K)$ satisfying

$$\sup_{n\in\mathbb{Z}}\ |y_n^{\rm eq}(K) - \tilde{y}_n(K)\,| = O(\,|\,K\,\ln(K)\,|\,), \quad K \to 0. \tag{14}$$

Moreover, $y_n^{\rm eq}$ decays to 0 exponentially as $n \to \pm\infty$ and has the symmetry $y_{-n}^{\rm eq} = y_n^{\rm eq}$.

Note that estimate (14) implies

$$\lim_{K\to 0} y_0^{\rm eq}(K) = +\infty, \quad \lim_{K\to 0} y_n^{\rm eq}(K) = 0 \ \text{ for } \ n \neq 0, \tag{15}$$

i.e. the equilibrium is highly localized when K is small. The equilibrium is spatially symmetric, but more general (not necessarily symmetric) solutions could be obtained by the same method. Figure 14 shows a numerical test of this theorem. The equilibrium solution has been found by minimizing the energy of the system, with the condition that the central site should have y_0 larger than the value that corresponds to the barrier of the effective potential. It shows that the localized equilibrium exists and that it decays exponentially away from the center as indicated by theorem 1.

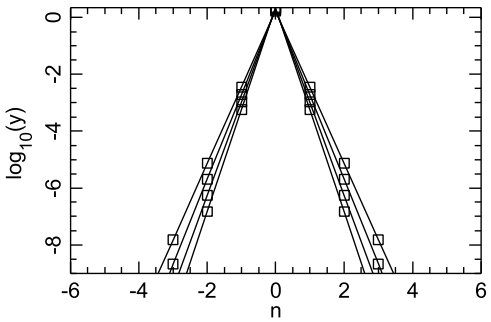

Fig. 14. Shape of the equilibrium solution in model 1 with the parameters for V_h indicated in the caption of Fig. 13, and 4 values of K: $K = 0.01,\ 0.005,\ 0.0025,\ 0.00125$ eV Å^{-2}. The different curves correspond to the different values of K, the curves decaying more rapidly corresponding to smaller K. This solution, obtained from a numerical solution of the equilibrium equations Eq. (11) confirms the results resulting from Theorem 1: $y_n^{\rm eq}$ decays exponentially to 0 away from the center and moreover smaller Ks lead to a more localized state.

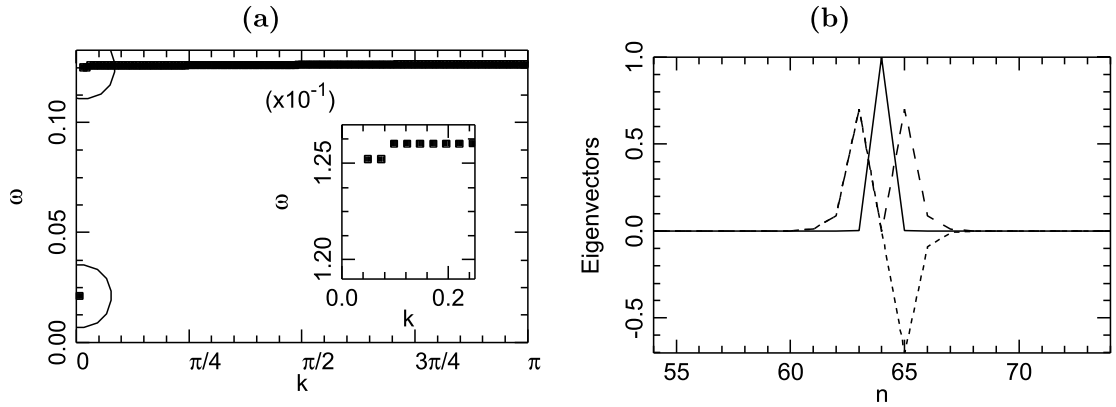

Fig. 15. (a) Linear spectrum of the equilibrium solution shown in Fig. 14 for the case $K = 0.01$ eV Å^{-2}. The circles attract the attention on three particular frequencies which are outside of the phonon band. One of them is far below the band. It is the frequency ω_b of the discrete breather. The two others also correspond to localized modes. The inset shows a magnification of the lower k part of the phonon band. The two frequencies right below the band are clearly visible. (b) Plot of the eigenvectors associated with the three modes which are outside of the phonon spectrum. The full line corresponds to the lowest mode, and, as expected it has the shape on the discrete breather. The dotted line corresponds to mode 2 and the dash line to mode 3. These numerical results show that modes 2 and 3 are also highly localized in spite of the fact that their frequencies are close to the phonon band.

The linear spectrum of this equilibrium solution shows the existence of a low frequency with an associated localized eigenvector, and suggests the existence of Lyapunov periodic orbits emanating from the localized equilibrium. These solutions correspond to the discrete breathers numerically computed in [29, 33]. Figure 15 shows the numerical calculation of the spectrum of one of the equilibrium solutions shown in Fig. 14 (case $K = 0.01$ eV Å^{-2}). As predicted theoretically it exhibits a frequency well below the bottom of the phonon band, which is associated with a localized symmetric eigenvector which corresponds to the shape of the corresponding discrete breather in the small amplitude limit. Moreover the spectrum shows two other frequencies slightly below the phonon band, which are associated with localized modes. All other frequencies belong to the phonon band and correspond to extended modes.

An existence theorem for discrete breathers oscillating in some finite-size neighborhood of $y_n^{\rm eq}$ has been proved in [29]. As above, the proof requires to extend the classical anti–continuum limit for periodic solutions [31, 34] to the case when the excited site lies at infinity for $K = 0$, with the additional difficulty that breather frequencies depend on K and vanish as $K \to 0$. The existence

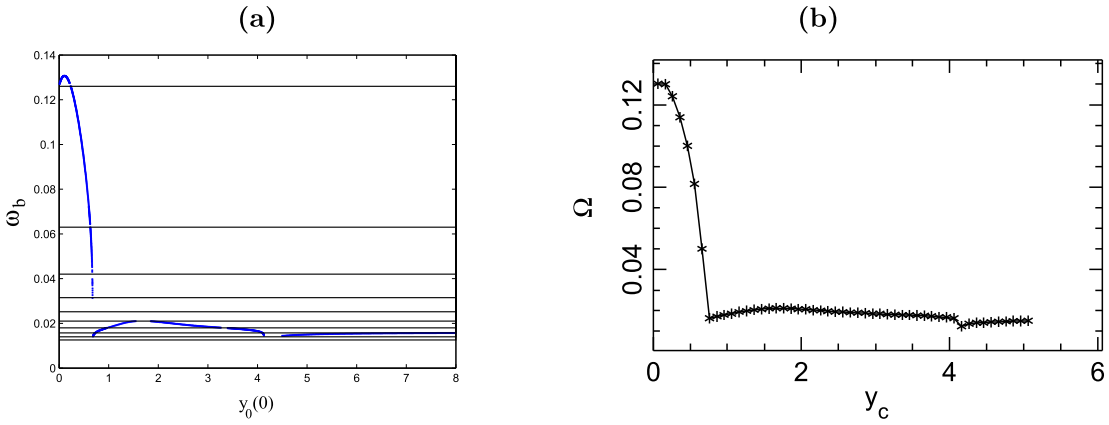

Fig. 16. (a) Frequency ω_b of breather solutions of Eq. (10) numerically computed by the Newton's method [29], as a function of breather amplitude $y_0(0)$ (expressed in Å). Resonances with phonons, represented by black bands, prevent the existence of breathers (only the first ten resonance bands have been represented for clarity). Parameters are $\alpha = \sigma = 4\,\text{Å}^{-1}$, $D = 0.0857\,\text{eV}$, $E = 4\,\text{eV Å}^{-1}$. Particle masses are set to $m = 300$ atomic mass units and $K = 0.01$ eV Å^{-2}. (b) Frequency of the breathers obtained numerically by exciting only the central particle of the lattice, as explained in Sec. 2.4. The numerical results given by the two different methods are in excellent agreement.

result of [29] is valid for small values of K, and under a non-resonance condition implying that all multiples of the breather frequency ω_b lie outside the phonon band. More precisely, Eq. (10) linearized at $y_n = 0$ admits solutions in the form of linear waves (phonons) $y_n(t) = C\cos(kn - \omega_q t)$, whose frequency satisfies the dispersion relation

$$m\,\omega_q^2 = 2\,A\alpha^2 + 2K(1 - \cos k). \tag{16}$$

Phonon frequencies belong to the band $[\omega_{min}, \omega_{max}]$, where $\omega_{min} = \alpha(2A/m)^{1/2}$ and $\omega_{max} = (\omega_{min}^2 + 4K/m)^{1/2}$. For a fixed value of K, Fig. 16 shows the forbidden frequency bands for discrete breathers, consisting of the phonon band and its submultiples.

The condition of non-resonance with linear waves is classical in the context of discrete breathers [31, 32], but yields unusual features in the continuation from infinity of [29], because the maximal frequency of discrete breathers beyond the potential barrier goes to 0 with K. As we have seen, close to their maximal frequency, the breathers beyond the barrier correspond to (weakly-nonlinear) localized modes of a spatially localized equilibrium y_n^{eq}. When K is small, y_0^{eq} goes to infinity on the flat part of the potential, which induces low-frequency oscillations. As a consequence, breathers exist for parameters (ω_b, K) lying outside certain resonance tongues accumulating near $(0, 0)$ [29]. When a multiple of the breather frequency lies in the phonon band, numerical results of [29] indicate that breathers are replaced by "almost localized" solutions having a small oscillatory tail at infinity. Such solutions are known as phonobreathers or nanopterons (see [26, 27, 32] and references therein).

Figure 16(a) taken from [29] provides the frequency-amplitude relation for a family of discrete breathers at fixed value of K obtained using a Newton's method using continuation from infinity. The local maximum of the breather frequency at $\omega_b \approx 0.02$ corresponds to oscillations near the localized equilibrium. As discussed in Sec. 2.4.1 these breathers can also be generated in numerical simulations by exciting a single particle and letting the dynamic solution relax to a steady oscillatory solution after the radiation of small amplitude waves which are absorbed by damping near the boundaries. Figure 16(b) shows that the numerical result is in very good quantitative agreement with the frequencies of exact discrete breathers, shown in Fig. 16(a). This indicates that the breathers are stable solutions which behave as attractors for an approximate initial condition.

In addition to the discrete breathers oscillating beyond the barrier and the classical ones oscillating below [28, 31], numerical computations of [29] yield another family of breather solutions oscillating on both sides of the barrier. These solutions also appear from infinity as $K \to 0$ and the proof of their existence is still an open theoretical problem. However, similar solutions have been proved to exist

in particular classes of Klein-Gordon lattices where the local anharmonic potential is constant above some amplitude threshold [30]. Moreover numerical simulations for the model indicate that, as shown in Sec. 2.4 for the DNA model, those breathers oscillating in both wells can be easily generated by a large amplitude localized initial excitation, and they are stable.

3.2 Model 2: Morse on-site potential. Anharmonic coupling

In the improved DNA model with the intra-pair potential V_h, as well as for its simplified version discussed in the previous section, the double-well shape of the effective potential $V_{\rm eff}$ (Eq. (8)) comes from the barrier in the intra-pair potential. It is interesting to notice that, if the nonlinearity in stacking is large enough, a similar double-well shape can also appear even if the intra-pair potential does not have a barrier, as shown in Fig. 17 for the case of the Morse intra-pair potential V_M. In this case the barrier for reclosing is entirely due to the interaction between sites, i.e. the stacking potential $W(y_n, y_{n-1})$ defined in Eq. (5). It persists as long as ρ is large enough ($\rho > 40.6$ for the parameters used to draw Fig. 17) but becomes shallower when ρ decreases. It is therefore interesting to study the properties of the ILMs in this model, which is the PBD model in the particular case of a very strong nonlinearity in the stacking interaction.

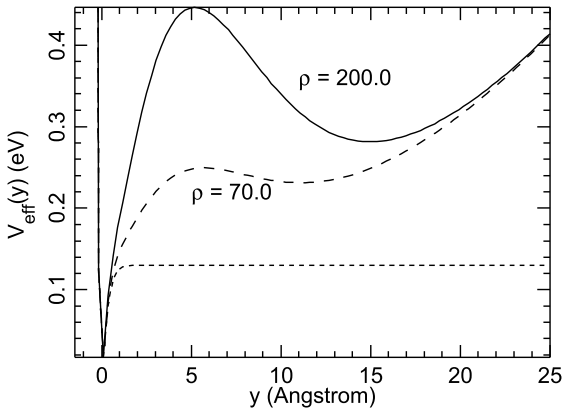

Fig. 17. Effective potential of the PBD model with a Morse intra-pair potential V_M and a nonlinear stacking interaction W for two values of ρ. The dotted line shows the Morse potential V_M for $D = 0.13$ eV, $\alpha = 4.2$ Å^{-1}. The full line shows the effective potential $V_{\rm eff}(y)$ (Eq. (8)) for $K = 0.00045$ eV Å^{-2}, $\delta = 0.4$ Å^{-1} and $\rho = 200$. The parameters of the model have been chosen to give a denaturation transition temperature $T_c \approx 340$ K, which is a realistic value for DNA. The dash line shows $V_{\rm eff}(y)$ for a smaller value $\rho = 70$, all other parameters being unchanged.

3.2.1 Main results

As a first step let us analyze spatially symmetric localized equilibria of the Morse chain with nonlinear stacking interactions, when K is small and ρ is large, with $\rho = O(K^{-1})$. Such equilibria are found as critical points in $\ell_2(\mathbb{Z})$ of the potential

$$U = \sum_{n\in\mathbb{Z}} V_M(y_n) + W(y_n, y_{n-1}). \tag{17}$$

Setting $\rho = \kappa/K$, the stacking potential (5) becomes

$$W(y_n, y_{n-1}) = \frac{1}{2}\left(K + \kappa\, e^{-\delta\,(y_n + y_{n-1})}\right)(y_n - y_{n-1})^2.$$

The equilibria are then determined by

$$V_M'(y_n) + \frac{\partial W}{\partial y_n}(y_{n+1}, y_n) + \frac{\partial W}{\partial y_n}(y_n, y_{n-1}) = 0, \quad n \in \mathbb{Z}. \tag{18}$$

The above scaling of ρ allows us to obtain localized solutions by continuation from infinity, using the approach introduced in [33] and reviewed in section 3.1. The main difference with the case of [33] is that the barrier of the effective potential originates from the nonlinear stacking and the above scaling of ρ.

We shall prove the following existence theorem giving localized solutions of Eq. (18) at small coupling K (in what follows we denote by $\delta_{m,n}$ the usual Kronecker symbol).

Theorem 2 Assume $\alpha > \delta > 0$ and set $\rho = \kappa/K$ in the nonlinear stacking potential (5), where $\kappa > 0$ is a fixed constant. For all $K > 0$ small enough, the potential (17) admits a critical point $\{y_n\}_{n\in\mathbb{Z}} \in \ell_2(\mathbb{Z})$ satisfying

$$y_n = \delta_{0,n}\, y_0^* + \psi_n, \tag{19}$$

$$y_0^* = \frac{1}{\delta}[\,-\ln(\frac{2K}{\kappa}) + \ln(-\ln(\frac{2K}{\kappa}))\,], \tag{20}$$

$$\|\{\psi_n\}\|_{\ell_2(\mathbb{Z})} = O\Big(\,-\frac{\ln(-\ln(\frac{2K}{\kappa}))}{\ln(\frac{2K}{\kappa})}\,\Big) \text{ as } K \to 0^+. \tag{21}$$

Moreover, y_n decays to 0 exponentially as $n \to \pm\infty$ and has the symmetry $y_{-n} = y_n$.

As shown by Eqs. (19)–(21), when $K \to 0$ the amplitude at site $n = 0$ diverges logarithmically, whereas the amplitudes at the other sites decay to 0 (at least extremely slowly).

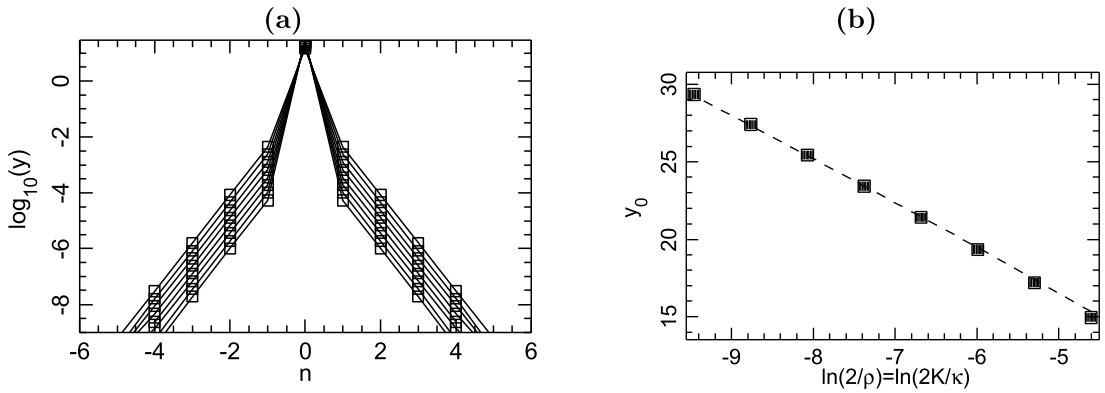

Fig. 18. (a) Numerical solution for the localized static equilibrium that exists in the PBD model with a large nonlinearity in the stacking interaction. The Morse potential parameters are those used for Fig. 17: $D = 0.13$ eV, $\alpha = 4.2$ Å^{-1}. The stacking potential W uses $\delta = 0.4$ Å^{-1}. The coupling K takes 8 values $K = 0.00045/2^p$ ($p = 0, 1, \ldots 7$) and ρ is determined by the condition $\rho = \kappa/K$ with $\kappa = 0.09$, i.e. $\rho = 200$ for the largest value of K, $K = 0.00045$ eV Å^{-2}, as in Fig. 17. (b) comparison between the numerically determined amplitude y_0 at the center of the equilibrium solution (filled squares) and the theoretical expression Eq. (20) (dash line). The results show an excellent agreement between the numerical data and the theoretical prediction as soon as K is small enough to have $\ln(2K/\kappa) \leq -6$.

Before turning to the proof of this theorem, let us check its validity by numerical analysis. Figure 18 shows that the predicted localized solution, decaying exponentially away from the center, is found by a numerical minimization of the potential energy of the system and that its amplitude evolves with K as indicated by theorem 2. The linear spectrum of the excitations around this static solution is shown in Fig. 19. As for the model studied in Sec. 3.1, it includes a low frequency mode corresponding to the low amplitude discrete breather around the equilibrium solution, and two other localized modes below the phonon band. Figure 20 shows the variation of the frequency of the lowest mode when K varies in a large range. This suggests that the DW-breathers can exist in a broad parameter range, and this can be checked by MD simulations by exciting a single site. For $\rho = 200$, starting for instance from an initial amplitude $y_c = 15.5$ Å we get a small amplitude breather in the vicinity of the shallow minimum of the effective potential (Fig 17). Its frequency, determined by the Fourier transform of

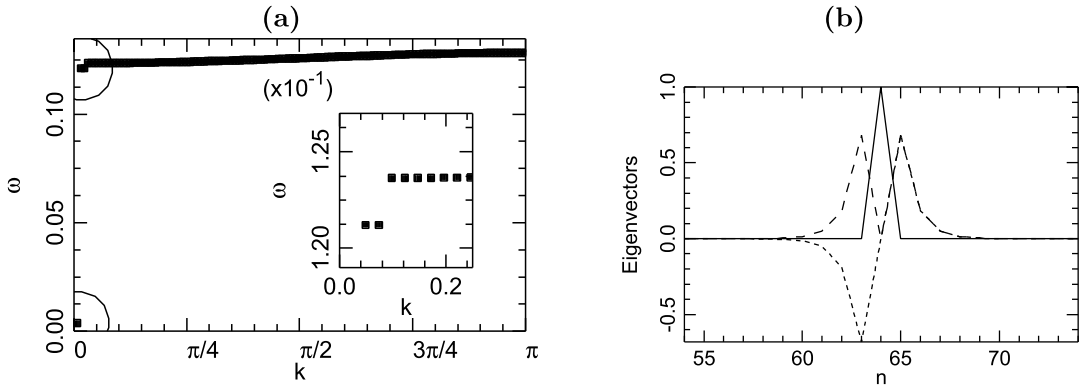

Fig. 19. (a) Linear spectrum of the equilibrium solution shown in Fig. 18 for the case $K = 0.00045$ eV Å^{-2}. The circles attract the attention on three particular frequencies which are outside of the phonon band. One of them is far below the band. It is the frequency ω_b of the discrete breather. The two others also correspond to localized modes. The inset shows a magnification of the lower k part of the phonon band. The two frequencies right below the band are clearly visible. (b) Plot of the eigenvectors associated with the three modes which are outside of the phonon spectrum. The full line corresponds to the lowest mode, and, as expected it has the shape on the discrete breather. The dotted line corresponds to mode 2 and the dash line to mode 3.

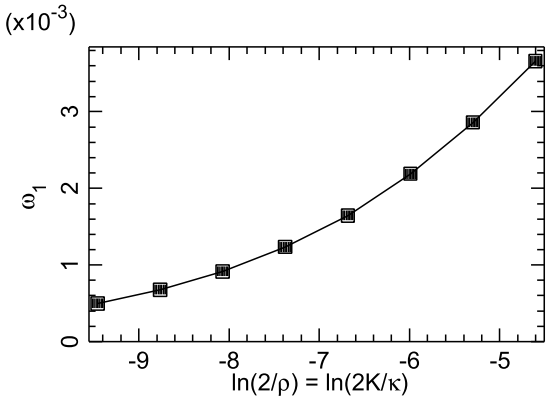

Fig. 20. Variation versus $\ln(2K/\kappa)$ of the frequency of the lowest mode around the equilibrium solution of the PBD model. The model parameters are the same as those used to draw Fig. 18.

$y_{N/2}(t)$ is found to be $\omega = 0.368\,10^{-2}$ inverse time units, which corresponds, up to numerical accuracy, to the frequency of the localized excited state of the localized equilibrium, $\omega_1 = 0.366\,10^{-2}$.

The proof of theorem 2 is performed in two steps. In section 3.3, we derive a single-site approximation of the localized equilibria, whose principal part is given by Eq. (20). In section 3.4 this initial guess is used to obtain the exact solutions of theorem 2, using the implicit function theorem.

3.3 Single-site approximate equilibrium solution

In this section we derive an approximation of the highly localized equilibrium solution of Eq. (18) described in theorem 2 for $K \approx 0$. This approximate solution $\tilde{y}_n = 0$ is localized at a single-site, i.e. satisfies $\tilde{y}_n = 0$ for $n \neq 0$.

The most natural choice is to search for $\tilde{y}_0$ as a critical point of the effective potential (8). Setting $\rho = \kappa/K$, the potential can be rewritten $V_{\text{eff}} = \tilde{V}_h + K\,y^2$, where $\tilde{V}_h(y) = V_M(y) + \kappa\,e^{-\delta\,y}\,y^2$. When $\alpha > \delta$, the potential $\tilde{V}_h$ decreases towards D when y is large, therefore it possesses the same features as the potential V_h, in particular a local maximum separating the open and closed states. Consequently, as it was the case for V_h, there exists a critical point $\tilde{y}_0$ of V_{eff} satisfying $\lim_{K\to 0} \tilde{y}_0 = +\infty$.

At this stage the problem can be further simplified, because we do not need an explicit value of $\tilde{y}_0$ as in [33], but only a sufficiently good approximation when $K \to 0$, allowing us to use the implicit

function theorem in section 3.4. Since the Morse potential component of $\tilde{V}_h$ can be neglected when $y \to +\infty$ (due to the fact that $\alpha > \delta$), one can look for a large amplitude critical point of the simplified potential

$$W_K(y) = (K + \kappa\, e^{-\delta\, y})\, y^2$$

replacing V_{eff}. Nonzero critical points of W_K satisfy

$$(-2 + \delta\, y)\, e^{-\delta\, y} = \frac{2K}{\kappa}, \tag{22}$$

which implies that the large amplitude critical point y of W_K grows at least logarithmically in K. In what follows we obtain a more precise estimate by transforming the singular perturbation problem (22) into a regular one, using several changes of variables.

The above-mentioned logarithmic singularity suggests the change of variables

$$\delta\, y = -\ln v, \tag{23}$$

which brings Eq. (22) into the form

$$-v\,(2 + \ln v) = \epsilon, \tag{24}$$

where we have set $\epsilon = \frac{2K}{\kappa}$. Now we look for a solution $v \approx 0$ of Eq. (24) for $\epsilon \approx 0$. Using a simple power-logarithmic ansatz for v suggests the change of variables

$$v = -\frac{\epsilon}{\ln \epsilon}(1 + w), \tag{25}$$

where $w \approx 0$ is a new unknown. This transforms Eq. (24) into the following problem

$$(1 + w)\,(-2\eta + 1 - \eta \ln(1 + w) + \mu) - 1 = 0, \tag{26}$$

with

$$\eta = -(\ln \epsilon)^{-1} \tag{27}$$

and

$$\mu = -\eta \ln \eta. \tag{28}$$

Using arguments from elementary calculus, η can be rewritten as a C^1 function of $\mu \in [0, e^{-1})$ satisfying $\eta(0) = \eta'(0) = 0$, which can be smoothly extended on an open neighborhood of $\mu = 0$. Considering now μ instead of K as a small parameter, Eq. (26) takes the form $f(w, \mu) = 0$, where f is C^1 in a neighborhood of $(0, 0)$ with $f(0, 0) = 0$, $\partial_w f(0, 0) = 1$. Consequently Eq. (26) can be solved locally by the implicit function theorem, which yields $w = w(\mu)$, where the function w is C^1 in $\mu \approx 0$ with $w(0) = 0$. Returning to expressions (23)–(25), this yields the following result.

Lemma 1 Fix $\delta > 0$ and $\kappa > 0$. For all $K > 0$ small enough, the potential W_K admits a critical point $\tilde{y}_0(K)$ satisfying

$$\tilde{y}_0 = \frac{1}{\delta}[-\ln(\frac{2K}{\kappa}) + \ln(-\ln(\frac{2K}{\kappa}))] + \tilde{R}(\mu), \tag{29}$$

where the function $\tilde{R}$ is C^1 in a neighborhood of 0, satisfies $\tilde{R}(\mu) = O(|\mu|)$, and

$$\mu = -\frac{\ln(-\ln(\frac{2K}{\kappa}))}{\ln(\frac{2K}{\kappa})} \to 0^+ \text{ as } K \to 0^+. \tag{30}$$

Lemma 1 provides an approximate localized solution $\tilde{y}_n = \delta_{0,n}\, \tilde{y}_0$ of Eq. (18) for fixed $\alpha > \delta > 0$, $\kappa > 0$ and $K \approx 0$. The principal parts of the exact localized equilibrium solution of theorem 2 and the approximate solution $\tilde{y}_n$ coincide (see Eq. (20)).

3.4 Proof of theorem 2

In this section we show the existence of an exact localized solution of Eq. (18), close to the approximate solution $\tilde{y}_n$ obtained in section 3.3. We can restrict ourselves to the case $\delta = 1$ and $D = 1$, up to rescaling y_n and renormalizing the constants K, κ, α, which results in the assumption $\alpha > 1$.

We assume that $y_{-n} = y_n$, which reduces Eq. (18) to an infinite system of nonlinear equations defined for $n \geq 0$. With this symmetry assumption, the case $n = 0$ of Eq. (18) becomes

$$V_M'(y_0) + W_K'(y_0 - y_1) = 0. \tag{31}$$

Now we set $y_0 = \tilde{y}_0 + u_0$ in Eq. (31), divide the resulting equation by $2K\tilde{y}_0$, and use Eq. (22) and the asymptotic expansion (29). This yields after lengthy but straightforward computations

$$u_0 + y_1 = R_0(u_0, y_1, \mu), \tag{32}$$

where $R_0 = O(\eta^{1+\alpha} K^{\alpha-1} + \eta(|u_0| + |y_1|) + u_0^2 + y_1^2)$ as $(u_0, y_1, \mu) \to 0$. We recall that $\eta \approx 0$ is a C^1 function of $\mu \approx 0$ implicitly defined through Eq. (28) and satisfying $\eta(0) = 0$. Consequently, from Eq. (27) it follows that $K = \frac{\kappa}{2} e^{-1/\eta}$ can be extended to a C^1 function of μ in a neighborhood of $\mu = 0$ with $K(0) = 0$. Using in addition the C^1 regularity of the remainder $\tilde{R}$ in lemma 1, one can check that the function R_0 is C^1 in a neighborhood of $(u_0, y_1, \mu) = 0$.

In the same way, the case $n = 1$ of Eq. (18) reads

$$V_M'(y_1) + \frac{\partial W}{\partial y_1}(y_2, y_1) + \frac{\partial W}{\partial y_1}(y_1, \tilde{y}_0 + u_0) = 0. \tag{33}$$

Using expansion (29), this yields after simple computations

$$2\alpha^2 y_1 + \kappa(y_1 - y_2) = R_1(u_0, y_1, y_2, \mu), \tag{34}$$

where $R_1 = O(|K\, \tilde{y}_0(K)| + y_1^2 + y_2^2) = O(|K\, \ln K| + y_1^2 + y_2^2)$ when $(u_0, y_1, y_2, \mu) \to 0$. After elementary manipulations we find

$$K\, \tilde{y}_0(K) = (1 + \mu) \frac{\kappa}{2\eta}\, e^{-\frac{1}{\eta}} + K\, \tilde{R}(\mu),$$

hence we have $K\, \tilde{y}_0(K) = \varphi(\mu)$ for some C^1 function φ defined on a neighborhood of $\mu = 0$. Using this property, one can check that the function R_1 is C^1 near 0.

Lastly, the case $n \geq 2$ of Eq. (18) gives

$$2\alpha^2 y_n - \kappa\, \Delta y_n = R(y_{n-1}, y_n, y_{n+1}, \mu), \quad n \geq 2, \tag{35}$$

where $\Delta y_n = y_{n+1} - 2y_n + y_{n-1}$, $R = O(y_{n-1}^2 + y_n^2 + y_{n+1}^2 + |K|(|y_{n-1}| + |y_n| + |y_{n+1}|))$ when $(y_{n-1}, y_n, y_{n+1}, \mu) \to 0$ and the function R is C^1 near 0.

The system (32)-(34)-(35) can be considered as a nonlinear equation for the infinite sequence

$$Y = (u_0, y_1, y_2, \ldots, y_n, \ldots) \in \ell_2(\mathbb{N}_0)$$

taking the form

$$\mathcal{A}\, Y = \mathcal{R}(Y, \mu), \tag{36}$$

where $\mathcal{A}$ is the bounded linear operator in $\ell_2(\mathbb{N}_0)$ defined by the right side of Eqs. (32)-(34)-(35), the nonlinear map $\mathcal{R} : \ell_2(\mathbb{N}_0) \times \mathbb{R} \to \ell_2(\mathbb{N}_0)$ is C^1 in a neighborhood of 0 and satisfies $\mathcal{R}(0, 0) = 0$, $D_Y \mathcal{R}(0, 0) = 0$ due to the above estimates on the components R_k.

Let us now prove the following:

Lemma 2 The operator $\mathcal{A} \in \mathcal{L}(\ell_2(\mathbb{N}_0))$ is invertible.

Proof. Consider $Z = (z_0, z_1, z_2, \ldots, z_n, \ldots) \in \ell_2(\mathbb{N}_0)$ and look for $Y \in \ell_2(\mathbb{N}_0)$ satisfying

$$\mathcal{A}\, Y = Z. \tag{37}$$

Equation (37) is equivalent to

$$u_0 + y_1 = z_0, \quad (2\alpha^2\mathbb{I} + \kappa\,\mathcal{L})\,\tilde{Y} = \tilde{Z}, \tag{38}$$

where we note $\tilde{Z} = (z_1, z_2, \ldots, z_n, \ldots) \in \ell_2(\mathbb{N})$. We define $\tilde{Y}$ in the same way, and introduce the Laplacian operator

$$\mathcal{L}\,\tilde{Y} = (y_1 - y_2, -\Delta y_2, -\Delta y_3, \ldots, -\Delta y_n, \ldots).$$

We begin by solving the second equation of Eq. (38) which is independent of u_0, and then obtain u_0 using the first equation. The operator $\mathcal{L}$ is self-adjoint in $\ell_2(\mathbb{N})$, and non-negative since $(\mathcal{L}\,\tilde{Y}, \tilde{Y})_{\ell_2} = \sum_{n=1}^{+\infty}(y_n - y_{n+1})^2$, therefore its spectrum $\sigma(\mathcal{L})$ is included in $\mathbb{R}^+$. Consequently, $2\alpha^2\mathbb{I} + \kappa\,\mathcal{L}$ is invertible in $\ell_2(\mathbb{N})$ and the second equation of Eq. (38) admits a unique solution $\tilde{Y}$. The first equation of Eq. (38) yields then $u_0 = z_0 - y_1$, where y_1 is the first component of the above solution $\tilde{Y}$. This completes the proof of the invertibility of $\mathcal{A}$.

□

For $\mu = 0$, $Y = 0$ is a solution of Eq. (36). Thanks to lemma 2, the implicit function theorem guarantees the local continuation of this solution to $\mu \approx 0$. More precisely, Eq. (36) admits a solution $Y = \mathcal{Y}(\mu) \in \ell_2(\mathbb{N}_0)$, where the function $\mathcal{Y}$ is C^1 in a neighborhood of $\mu = 0$ and satisfies $\mathcal{Y}(0) = 0$. Going back to the original variable y_n, this immediately yields the existence result of theorem 2 and estimates (20), (21).

To complete the proof of theorem 2, there remains to check the exponential decay of the localized equilibrium. Since the equilibrium is spatially symmetric, it suffices to consider the case $n \to +\infty$. Let us consider any solution y_n of Eq. (18) decaying to 0 as $n \to +\infty$ and prove its exponential decay. Setting $z_n = y_n - y_{n-1}$, Eq. (18) can be rewritten (recall that D has been normalized to unity)

$$2\,\alpha^2 y_n + (\kappa + K)(z_n - z_{n+1}) + N(y_n, z_n, z_{n+1}) = 0, \tag{39}$$

where $N(y_n, z_n, z_{n+1}) = O(y_n^2 + z_n^2 + z_{n+1}^2)$ in the neighborhood of $(0, 0, 0)$. Consequently, using the fact that y_n is spatially localized and the implicit function theorem, equality (39) can be rewritten in the following form provided n is large enough

$$z_{n+1} = f(y_n, z_n), \tag{40}$$

where f is analytic in a neighborhood of $(0, 0)$ and

$$f(y_n, z_n) = \frac{2\alpha^2}{\kappa + K} y_n + z_n + O(y_n^2 + z_n^2).$$

Consequently, we have in addition

$$y_{n+1} = y_n + f(y_n, z_n). \tag{41}$$

One can check that the fixed point $(y_n, z_n) = (0, 0)$ of the two-dimensional mapping defined by (40)–(41) is a hyperbolic saddle. Since localized solutions of Eq. (18) correspond to orbits lying on its local stable manifold when n is large enough, this guarantees their exponential decay by the stable manifold theorem. This completes the proof of theorem 2.

Remark 1 For the exact localized solution of theorem 2, expression (19) and estimate (21) imply that $\{y_n\}_{n\geq 1}$ is small in $\ell_\infty(\mathbb{N})$ when $K \approx 0$. Consequently, the sequence $\{(y_n, z_n)\}_{n\geq 2}$ belongs to the local stable manifold of the origin when K is small enough. However, (y_1, z_1) does not belong to this local stable manifold (indeed $|z_1| = |y_1 - y_0|$ is large, i.e. diverges logarithmically when $K \to 0$). This unusual property comes from the fact that the reformulation of Eq. (18) as a map (40)–(41) is only local, whereas theorem 2 describes another class of solutions which cannot be completely described by the orbit of a map.

4. Conclusions

DNA is undoubtedly a molecule that has inspired physicists interested in nonlinear science. S. Takeno and S. Homma [35] examined the anharmonic rotation of the bases in a model that lead them to a sine-Gordon equation. We have shown here that studies of the dynamics of the fluctuational opening of the base pairs of DNA naturally lead to a model which describes the temporary breaking of the pairs in terms of discrete breathers. It is this observation that stimulated the interest of S. Aubry and lead to the first mathematical proof of existence of discrete breathers by R.S. MaKay and S. Aubry [31]. However, although it can give satisfactory results for equilibrium properties, the simplest nonlinear model of DNA fluctuations does not properly describe the time scales of the fluctuations. To correct this deficiency we introduced a barrier for base-pair closing. The resulting model turns out to be particularly interesting because it sustains a new class of ILMs which oscillate around a position which is not an equilibrium position of the on-site potential. As a result the original anti–continuum limit introduced in the paper of R.S. MacKay and S. Aubry [31] has to be extended to study them. It is interesting that two types of models can lead to very similar localized excitations because the barrier allowing this new class of ILMs can come either from the on-site potential or from nonlinear stacking interactions.

For applications to DNA the models that we have introduced are however still incomplete. In spite of an improvement by several orders of magnitude with respect to the first PB model, they are not yet giving the correct time scales for the dynamics of the opening and closing of the base pairs. In these models the geometry of the molecule is not described and the missing degrees of freedom, particularly the torsion of the double helix, may be important. However another missing aspect is the strong damping exerted by the solvent on the bases when they are out of the stack of the double helix, while closed base pairs are protected form the solvent and only weakly damped as shown by their vibrational modes which can be detected by spectroscopy. Due to the strong damping of the solvent the equilibrium solutions that we derived as a first step towards the calculation of breathers may actually be relevant for themselves because they could correspond to long-lived open states of the base pairs in solution. Further studies of DNA in contact with its environment are necessary to better approach the subtleties of its dynamics.

Acknowledgments

MP would like to thank S. Takeno, colleague and friend, who invited him for a long stay in Japan which initiated a fruitful collaboration, and moreover allowed him to make many contacts which had a strong influence on the development of his research.

References

[1] A.J. Sievers and S. Takeno, "Intrinsic Localized Modes in Anharmonic Lattices," *Phy. Rev. Lett.*, vol. 61, no. 8, pp. 970–973, August 1988.

[2] M.E. Manley, A.J. Sievers, J.W. Lynn, S.A. Kiselev, N.I. Agladze, Y. Chen, A. Llobet, and A. Alatas, "Intrinsic localized modes observed in the high-temperature vibrational spectrum of NaI," *Phys. Rev. B*, vol. 79, no. 13, 134304-1-5, April 2009.

[3] E. Schrödinger, *What is Life?* Cambridge University Press (1967) (First published 1944).

[4] J.D. Watson and F.H.C. Crick, "A structure for Deoxyribose Nucleic Acid," *Nature*, vol. 171, pp. 737–738, April 1953.

[5] M. Gueron, M. Kochoyan, and J.-L. Leroy, "A single mode of DNA base-pair opening drives imino proton exchange," *Nature*, vol. 328, pp. 89–92, July 1987.

[6] B. McConnel and P.H. von Hippel, "Hydrogen exchange as a Probe of the Dynamic Structure of DNA," *J. Mol. Biol.*, vol. 50, pp. 297–316, 1970.

[7] B.H. Zimm and J.K. Bragg, "Theory of the phase transition between helix and random coil in polypeptide chains," *J. Chem. Phys.*, vol. 31, no. 2, pp.526–535, 1959.

[8] B.H. Zimm, "Theory of "Melting" in the Helical Form in Double Chains of the DNA Type," *J. Chem. Phys.*, vol. 33, no. 6, pp. 1349–1356, 1960.

[9] D. Poland and H.A. Scheraga, "Phase transitions in one-dimension and the Helix-Coil Transition in Polyamino-Acids," *J. Chem. Phys.*, vol. 45, no. 5, pp. 1456–1463, 1966.
[10] D. Poland and H.A. Scheraga, "Occurence of a phase transition in nucleic acid models," *J. Chem. Phys.*, vol. 45, no. 5, pp. 1464–1469, 1966.
[11] M. Peyrard and A.R. Bishop, "Statistical Mechanics of a Nonlinear Model for DNA Denaturation," *Phys. Rev. Lett.*, vol. 62, no. 23, pp. 2755–2758, June 1989.
[12] M. Peyrard, "Nonlinear dynamics and statistical physics of DNA," *Nonlinearity*, vol. 17, no. 2, pp. R1–R40, March 2004.
[13] T. Dauxois, N. Theodorakopoulos, and M. Peyrard, "Thermodynamic instabilities in one dimension: correlations, scaling and solitons," *J. Stat. Phys.*, vol. 107, no. 3–4, pp.869–891, May 2002.
[14] T.S. van Erp, S. Cuesta-López, and M. Peyrard, "Bubbles and the denaturation in DNA," *Eur. Phys. J. E*, vol. 20, no. 4, pp. 421–434, August 2006.
[15] N. Theodorakopoulos, "Melting of genomic DNA: Predictive modeling by nonlinear lattice dynamics," *Phys. Rev. E*, vol. 82, no. 2, pp. 021905-1-4, August 2010.
[16] T. Dauxois, M. Peyrard, and A.R. Bishop, "Entropy driven DNA denaturation," *Phys. Rev. E*, vol. 47, no. 1, pp. R44–R47, January 1993.
[17] M. Peyrard, S. Cuesta López, and D. Angelov, "Experimental and theoretical studies of sequence effects on the fluctuations and melting of short DNA molecules," *J. Phys. Condensed Matter*, vol. 21, no. 3, pp. 034103-1-13, January 2009.
[18] A. Wildes, N. Theodorakopoulos, J. Valle-Orero, S. Cuesta-Lopez, J.-L. Garden, and M. Peyrard, "The thermal denaturation of DNA studied with neutron scattering," *PRL*, vol. 106, no. 4, pp. 048101-1-4, January 2011.
[19] A. Wildes, N. Theodorakopoulos, J. Valle-Orero, S. Cuesta-Lopez, J.-L. Garden, and M. Peyrard, unpublished.
[20] G.J. Martyna, M.L. Klein, and M. Tuckerman, "Nosé-Hoover chains: The canonical ensemble via continuous dynamics," *J. Chem. Phys.*, vol. 97, no. 4, pp. 2635–2643, August 1992.
[21] E. Giudice, P. Várnai, and R. Lavery, "Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations," *Nucleic Acid Research*, vol. 31, no. 5, pp. 1434–1443, 2003.
[22] G. Weber, "Sharp DNA denaturation due to solvent interaction," *Europhys. Lett.*, vol. 73, no. 5, pp.806–811, February 2006.
[23] G.L. Alfimov, V.A. Brazhnyi, and V.V. Konotop, "On classification of intrinsic localized modes for the discrete nonlinear Schrödinger equation," *Physica D*, vol. 194, no. 1–2, pp. 127–150, July 2004.
[24] S. Aubry and G. Abramovici. "Chaotic trajectories in the standard map : the concept of anti-integrability," *Physica D*, vol. 43, no. 2–3, pp. 199–219, July 1990.
[25] S. Aubry, "Anti-integrability in dynamical and variational problems," *Physica D*, vol. 86, no. 1-2, pp. 284–296, September 1995.
[26] S. Aubry, "Breathers in nonlinear lattices : Existence, linear stability and quantization,", *Physica D*, vol. 103, no. 1–4, pp. 201–250, April 1997.
[27] M. Georgi, "Bifurcations from localised steady states to generalised breather solutions in the Klein-Gordon lattice," *International journal of dynamical systems and differential equations*, vol. 2, no. 1-2, pp. 66–97, 2009.
[28] G. James, B. Sánchez-Rey, and J. Cuevas, "Breathers in inhomogeneous lattices : an analysis via center manifold reduction," *Reviews in Mathematical Physics*, vol. 21, no. 1, pp. 1–59, February 2009.
[29] G. James, A. Levitt, and C. Ferreira, "Continuation of discrete breathers from infinity in a nonlinear model for DNA breathing," *Applicable Analysis*, vol. 89, no. 9, pp. 1447–1465, 2010.
[30] G. James and D. Pelinovsky, "Breather continuation from infinity in nonlinear oscillator chains," arXiv:1011.5927v1 [nlin.PS], p.27, 2010.
[31] R.S. MacKay and S. Aubry, "Proof of existence of breathers for time-reversible or Hamiltonian

networks of weakly coupled oscillators," *Nonlinearity*, vol. 7, no. 6, pp. 1623–1643, November 1994.

[32] A.M. Morgante, M. Johansson, S. Aubry, and G. Kopidakis, "Breather-phonon resonances in finite-size lattices: 'phantom breathers'?," *J. Phys. A: Math. Gen.*, vol. 35, no. 24, pp.4999–5021, June 2002.

[33] M. Peyrard, S. Cuesta-López, and G. James, "Nonlinear analysis of the dynamics of DNA breathing," *J. Biol. Phys.*, vol. 35, no. 1, pp. 73–89, February 2009.

[34] J.-A. Sepulchre and R.S. MacKay, "Localized oscillations in conservative or dissipative networks of weakly coupled autonomous oscillators," *Nonlinearity*, vol. 10, no. 3, pp. 679–713, May 1997.

[35] S. Takeno and S. Homma "Self-localized anharmonic rotational modes of bases in DNA," *J. Phys. Soc. Japan*, vol. 59, no. 5, pp. 1890–1901, May 1990.