Chapter 4

Studies of thermal denaturation

4.1 Introduction

In this chapter we present some preliminary results obtained when studying the *twist-opening* model in contact with a thermal bath. In the previous chapters we always referred essentially to a microcanonical scheme, with constant energy conditions. We start here a new approach in which constant temperature conditions are chosen. Canonical ensemble simulations are performed in order to study the thermal behavior of a model chain and to compare the results with the real DNA statistical features in similar conditions. We investigate here whether the model is to reproduce the DNA denaturation transition and we attempt to adjust the model parameters on the basis of available experimental data.

We will consider both the original and the improved form of the *twist-opening* model. As we will see, in the first case there is not good agreement with experimental results on DNA, while for the improved model the agreement is much better.

We have performed simulations in the canonical ensemble using the completely deterministic Nosé Hoover thermostat technique. We have also looked at the specific dynamical mechanism that leads to the denaturation transition: this allows to investigate the kind of dynamical distortion that are thermally induced in the chain and to compare them with the analytical solutions introduced in this work.

In this context, we will come back to the results obtained in Chapter 3 for what concerns the existence of small amplitude "breathing" solutions. In that chapter we found that, with our choice of parameters, only slowly oscillating solutions, arising from small q acoustic modes, can be obtained for the original model. This phenomenology is confirmed by constant temperature studies of the dynamical behavior of the chain.

Anyway, we discussed a possible improvement of the model, which allows to recover oscillating approximate solutions related to optical modes. We will see that *the improved model features in constant temperature conditions correspond much better to real DNA properties.*

The addition of the direct stacking term has in fact a major role even in the context of denaturation. As already stressed by Dauxois and co-workers in [33], base pairs situated in the vicinity of an open base pair have a larger probability to open, because of a change in the electronic overlap which leads to to the release of the stacking interaction. In the original *twist-opening* model, such an effect was absent. Moreover, in the small displacement limit, the interaction between neighboring radial stretches arises from terms of the type $(y_n + y_{n-1})$: such terms have the opposite effect of decreasing the stretch distortion of pairs in vicinity of an open pair. The introduction of W_{stack} allows to recover the correct stacking interaction, and represents an important modification which acts in denaturation as to favor bubble formation.

In Section 4.2 we remind the fundamental properties of the denaturation process. Section 4.3 is devoted to a summary of the Nosé Hoover thermostating technique. We stress the main difficulties of the original method and the improved technique more recently introduced by Martyna and co-workers using Nosé Hoover thermostat chains. In Section 4.4 we briefly explain the results, problems and possible solutions introduced by the authors in performing canonical studies for the PB model. We then use the Nosé Hoover chain technique in Section 4.5 in order to obtain constant temperature conditions for the model. In this section we first refer to the original model (which corresponds to put $S = 0$ in the improved Lagrangian (3.68)) then to the improved one ($S \neq 0$). We then discuss the results in order to have indications for the choice of the parameter G_0, which we cannot fix on the basis of other experimental results.

Furthermore, in Section 4.5 we show space-time graphs illustrating the dynamical configurations at different temperatures, as functions of site number and time. This allows to understand the dynamical mechanism that leads to denaturation. We also discuss the correspondence with thermal behavior of the model and the known analytical solutions, for the original and improved versions of the *twist-opening* model.

4.2 The denaturation of DNA

Denaturation is the complete strand separation arising when DNA is heated up to a certain temperature that we denote by T_{den}. This melting temperature T_{den} depends on the base sequence, because of the different strength of double and triple hydrogen bonds connecting bases in the pairs AT and GC respectively. Anyway, the range of variability of T_{den} is not too large: an approximate estimate for natural DNA fixes the melting temperature T_{den} in the range $326 \div 370\,K$ [18].

One can obtain experimentally a curve that represents the percentage of denaturated base pairs of DNA in solution as a function of the temperature by heating slowly the solution and measuring its absorbtion in the UV range, near $260\,nm$. This depends on the fact that free nucleotides absorb in this range, while their absorbtion is strongly reduced when they are stacked in the double helix [18].

Denaturation arises in a very narrow temperature range of *a few degrees*. Denaturation curves have the shape schematically indicated in Figure 1.3.

For a DNA model which attempts to reproduce the opening features of the molecule it is thus interesting to study its ability to reproduce denaturation curves. For this, one has to simulate a constant temperature condition, and to calculate, for each temperature, the percentage of denaturated base pairs in the chain. The first request is obviously that the model displays denaturation, *i.e.* that all the base pairs open above a certain temperature threshold. Then, one would like to reproduce approximately the correct T_{den}, and finally have denaturation in a sufficiently narrow temperature range. Model parameters can be adjusted in order to approach as well as possible the main features of real denaturation curves.

4.3 The simulation of the canonical ensemble with the Nosé Hoover method.

The Nosé thermostat

Molecular dynamics studies of our model can help in understanding its specific features by describing its behavior when in contact with a thermal bath. The statistical properties which define the denaturation curves can be numerically calculated, and the results can be used to correct the Lagrangian and to fix its parameters. Furthermore, this can help in understanding the mechanism by which the model is excited by thermal energy, and thus what kind of fluctuations spontaneously arise in the chain. The observed dynamics can be compared with the known properties of real DNA at different temperatures, so that the efficiency of the model itself in reproducing the DNA dynamical behavior can be usefully tested.

The numerical simulation of a system in the canonical ensemble, *i.e.* with fixed particle number, volume and temperature, can be obtained essentially through two different approaches: either using stochastic forces as the Langevin one, or modifying the Hamilton equation in such a way to satisfy the condition of constant temperature, preserving at the same time the deterministic character of the model. We will use the latter method. More precisely we will refer to the Hoover reformulation [39] of the Nosé method [38], with the improvement introduced by Martyna *et al* in [58].

The Nosé method [38, 40] is based on the idea of extending the physical system by introducing one more degree of freedom which represents the thermal bath[1]. Let us denote the system variables by q_i and their masses by m_i. The thermal interaction between the particles of the system and the bath is expressed as a rescaling of the time by a factor s. The kinetic energy can be controlled by controlling the particles velocities. The introduction of s can be viewed as the introduction of a flexible time: the change of the time length implies in fact a change in the velocities. In the extended system the real velocities v_i are thus substituted by virtual velocities $V_i = sv_i$. This leads to define virtual momenta P_i defined as functions of the real ones p_i as $P_i = sp_i$. The variable s regulating the energy flux is the only added degree of freedom, with a parameter Q which represents the thermal inertia and fixes the temperature fluctuations.

The Lagrangian of the extended system is postulated to be of the form

$$L = \sum_{i=1}^{N} \big(\frac{m_i}{2} s^2 \, \dot{q_i}^2 - V(\{q_i\})\big) + \frac{Q}{2}\dot{s}^2 - (N+1)k_B T \ln s, \tag{4.1}$$

where T is the temperature and k_B the Boltzmann constant.

With this choice, the conjugate momenta are

$$\begin{aligned} P_i &= \frac{\partial L}{\partial \dot{q_i}} = m_i s^2 \dot{q_i} \ , \\ p_s &= \frac{\partial L}{\partial \dot{s}} = Q\dot{s} \end{aligned}$$

so that $\frac{m}{2} s^2 \dot{q_i}^2 = \frac{1}{2m} P_i^2 s^{-2} = \frac{1}{2m} p_i^2$ and the system Hamiltonian is

$$H(q_i, P_i, s, p_s) = H_0 + \frac{p_s^2}{2Q} + (N+1)k_B T \ln s \tag{4.2}$$

where H_0 is the original system Hamiltonian. In this Hamiltonian, P_i are the virtual momenta, but they always appear in the form P_i/s, representing the real ones p_i.

[1]In standard statistical mechanics the canonical ensemble is obtained by making the system exchange energy with a macroscopic reservoir with many particles. In this scheme the system still exchange energy with an external object, with the difference that it is represented by a degree of freedom. But this degree of freedom interacts with all the particles and its own dynamics is defined in such a way that this interaction leads to a canonical evolution for the physical system of interest.

The equations of motion derived for q_i, p_i, s and p_s are then

$$\begin{aligned}
\dot{q}_i &= \frac{P_i}{m_i s^2}\,, \\
\dot{p}_i &= -\frac{\partial V}{\partial q_i}\,, \\
\dot{s} &= \frac{p_s}{Q}\,, \\
\dot{p}_s &= Q\Big(\sum_{i=1}^{N} \frac{P_i^2}{m_i s^2} - (N+1)k_B T\Big)\,.
\end{aligned} \tag{4.3}$$

These equations guarantee that in the extended phase space the microcanonical conditions for the whole system are equivalent to canonical ensemble conditions for the original system of particles. The microcanonical partition function is in fact defined as:

$$\mathcal{Z} = \int dp_s ds \prod_i dP_i dq_i \; \delta(H - E)\,. \tag{4.4}$$

Rescaling $\mathcal{Z}$ in order to express it in terms of real momenta $p_i = P_i/s$, one obtains

$$\mathcal{Z} = \int dp_s ds\, s^N \prod_i dp_i dq_i \; \delta(H - E)\ . \tag{4.5}$$

The integral over s can be easily calculated to obtain

$$\begin{aligned}
\int ds\, s^N\, \delta\Big(H_0 + \frac{p_s^2}{2Q} + (N+1)k_B T \ln s - E\Big) &= \\
\frac{1}{(N+1)k_B T} \exp\Big(\frac{H_0 + (p_s^2/2Q) - E}{k_B T}\Big)\ ;
\end{aligned} \tag{4.6}$$

the partition function is then of the form

$$\mathcal{Z} = C \cdot \int \prod_i dp_i dq_i \exp\big(\frac{H_0}{k_B T}\big)\ , \tag{4.7}$$

and this, as desired, is the expression of the canonical partition function for the initial system.

Making the ergodic hypothesis, one has that the time average is equivalent to the average on the extended microcanonical ensemble, and thus to the canonical average for the system of interest:

$$\lim_{t\to\infty} \frac{1}{t} \int^t A\, dt = \langle A(p/s, q)\rangle_{ext.sys.} = \langle A(p,q)\rangle_{can.} \tag{4.8}$$

The Nosé-Hoover thermostat

In the Hoover reformulation [39] of the Nosé method a negative feedback mechanism is more apparent. Hoover defines a new variable

$$\zeta = s\frac{p_s}{Q} \tag{4.9}$$

which represents a friction coefficient; rescaling the time unit as $dt \to dt/s$ and using the original variables Equations (4.3) can be simplified into the following form:

$$\begin{aligned}
\dot{q}_i &= \frac{p_i}{m_i}\,, \\
\dot{p}_i &= -\frac{\partial V}{\partial q_i} - \zeta p_i\,, \\
\frac{\partial}{\partial t}\ln s &= \zeta\,, \\
\frac{\partial}{\partial t}\zeta &= \left(\sum_{i=1}^{N}\frac{p_i^2}{m_i} - Nk_BT\right)/Q
\end{aligned} \tag{4.10}$$

(the constant $(N+1)$ is substituted by N in this formulation).

This formulation is called Nosé-Hoover thermostat. Friction coefficient ζ, which fluctuates between positive and negative values, stabilizes the kinetic energy of the real system, $\sum_{i=1}^{N}(p_i^2/m_i)$, to the canonical expectation value Nk_BT. In this case the s variable is decoupled from the others, and can be neglected, so that the time evolution of the system is given simply by the first three equation in the phase space formed by the variables (q_i, p_i, ζ). Equations of motion (4.10) cannot be derived from an Hamiltonian, even if a *pseudo-Hamiltonian* exists, *i.e.* there is a conserved quantity, obtained by solving the last equation in (4.10).

The inertia parameter Q and the ergodic assumption

The main difficulty of this method, in both formulations, is in the choice of the inertia parameter Q. If its value is too large, the external "bath", represented in the Nosé idea by an additional particle, is too heavy and thus the energy exchange between the bath and the system is too small. On the contrary, if Q is too small, the bath tends to "follow" the rest of the system and oscillates too fast, making the equilibrium state for the whole system hard to approach. The problem of choosing a good value for Q is often quite delicate, and the consequence of a bad choice is that kinetic energy oscillations (namely the mean squared deviation of the kinetic energy) are quite different from what expected for the canonical ensemble. The best choice for Q is usually obtained when the thermostat is in good "resonance" with the system, *i.e.* one must impose for the thermostat approximatively the same linear frequency of the system.

Furthermore, as pointed out by Martyna and *et al* [58], for small or stiff system the ergodic assumption fails. This represents a more fundamental problem. This depends essentially on the fact that, while positions and momenta of the initial system are driven by the thermostat, there is nothing to drive the fluctuations of ζ. For this reason, in some cases ζ falls onto periodic or quasi-periodic orbits, instead of exploring the whole phase, thus breaking the ergodic assumption.

The Nosé Hoover chains

Martyna and co-workers [58] propose a new method to assure the extended system ergodicity, which gives good results. In order to drive the thermostat fluctuations, they introduce a second thermostat, which is coupled just with the first one; they add then a third thermostat to control the second one, and so on. These Nosé-Hoover thermostat chains usually assure ergodicity already with three or four thermostats.

The Nosé-Hoover chain method can be expressed, in the case of M thermostats, by the following equations:

$$\begin{aligned}
\dot{q}_i &= \frac{p_i}{m_i}, \\
\dot{\eta}_i &= \frac{p_{\eta_i}}{Q_i}, \\
\dot{p}_i &= -\frac{\partial V}{\partial q_i} - \frac{p_{\eta_1} p_i}{Q_1}, \\
\dot{p_{\eta_1}} &= \Big(\sum_{i=1}^{N} \frac{p_i^2}{m_i} - N k_B T\Big) - \frac{p_{\eta_1} p_{\eta_2}}{Q_2}, \\
&\dots \\
\\
\dot{p_{\eta_j}} &= \Big(\frac{p_{\eta_{j-1}}^2}{Q_{j-1}} - k_B T\Big) - \frac{p_{\eta_j} p_{\eta_{j+1}}}{Q_{j+1}}, \\
&\dots \\
\dot{p_{\eta_M}} &= \Big(\frac{p_{\eta_{M-1}}^2}{Q_{M-1}} - k_B T\Big).
\end{aligned} \tag{4.11}$$

The equations for η_i are added for completeness.

4.4 Anharmonic stacking in the PB model

Before going on, it is useful to recall which results have been obtained for the denaturation properties of the PB model. Peyrard and co-workers focus on DNA denaturation from the very beginning [14, 33]. They perform numerical simulation at constrained temperature with a (single) Nosé Hoover thermostat, and statistical-mechanics analysis with the transfer-integral method to obtain analytical and numerical results on the denaturation statistical properties. Furthermore, molecular dynamics simulations allow to investigate the *mechanism* underlying denaturation. By visualizing the local state of the base pairs as a function of the time and of the site number, at fixed temperature, these authors are in fact able to detect in the model the behavior of the characteristic phases of real DNA denaturation:

1. in the low temperature range, the model presents oscillations localized on a few sites, consistently with the fluctuational openings ("breathing") observed experimentally and with the existence of small amplitude breather solution for this model;

2. at higher temperatures, the small amplitude oscillations grow to give larger oscillating bubbles localized on a number of sites that increases with temperature, that well represent the denaturation bubbles observed in real DNA;

3. at the highest investigated temperature bubbles grow to cover all the chain until it reaches complete melting.

The estimation of the denaturation temperature T_{den} depends on the chosen denaturation threshold for the base-pairs, *i.e.* the hydrogen bond length at which one base pair is considered as denaturated. This threshold can be chosen by referring to the Morse potential shape or to the hydrogen bond equilibrium length. Anyway, any reasonable choice is acceptable in order to give an approximate estimation of T_{den}, and in fact, a change in the threshold is found not affect very much T_{den}.

For the PB model the authors find, numerically and analytically, that T_{den} is of the order of $500K$. This deviation from typical values may depend on the choice of parameters.

There is, anyway, a more fundamental deviation from the real situation: *denaturation occurs in fact over a too large temperature range* (of the order of $300K$).

In [33][2], Dauxois, Peyrard and Bishop propose an improvement of the PB model to correct this behavior. Following the suggestions made within an Ising model description of DNA in [59], they argue that such an extremely sharp transition in a one-dimensional model can occur as a consequence of cooperativity effects. These can be introduced in the PB model at the microscopic level by an appropriate *anharmonic stacking* interaction potential that reflects the change in

[2]See also [50].

the electronic distribution on the bases when the hydrogen bonds are broken:

$$W(y_n, y_{n-1}) = \frac{K}{2}\Big(1 + \rho e^{-\eta(y_n+y_{n-1})}\Big)(y_n - y_{n-1}). \tag{4.12}$$

The new anharmonic inter-site coupling can be justified on the basis of the observation that the stacking energy is actually a property of the whole base pair, and depends on the state of the neighboring pairs. When the hydrogen bond breaks, the electronic distribution on the base pair is modified, causing a decrease of the stacking interaction with the adjacent bases. The prefactor to the usual term $(y_n - y_{n-1})$ in (4.12) depends on the *sum* of adjacent stretches $(y_n + y_{n-1})$, so that it decrease from $(1+\rho)K/2$ to $K/2$ when either one or both base pairs n and $n-1$ open. This means that a base pair which is placed in the immediate vicinity of a denaturated bubble has an higher probability to open.

The addition of the anharmonicity in the model does not qualitatively affect the dynamics of the system near transition, but the dependence on temperature is strongly modified. New denaturation studies performed with an appropriate parameter choice show transition in a temperature window of less then $30K$ with denaturation temperature estimated to be $361.5K$ [33].

4.5 Denaturation for the two versions of the ***twist-opening*** model with different values of G_0

4.5.1 Model without the direct stacking term W_{stack}

Denaturation percentages

We have performed numerical simulations at constant temperature for the system of equations describing the original and the improved versions of the *twist-opening* model for a chain of $N = 256\, bps$. Let us start by presenting results of the former case. We integrate the equations of motion using a fourth order Runge-Kutta algorithm with integration step of $0.05\, t.u.$ We simulate the canonical conditions by introducing for each set of degrees of freedom $\{r_n\}$ and $\{\varphi_n\}$ a chain of three Nosé Hoover thermostats, which turns out to be sufficient to obtain the correct canonical energy averages and fluctuations:

$$\langle E^{r,\varphi}_{kin}\rangle = \frac{1}{2}\, N\, k_B T \tag{4.13}$$

$$\langle \Big(E^{r,\varphi}_{kin} - \frac{1}{2}\, N\, k_B T\Big)^2 \rangle = \frac{2}{N}\langle E^{r,\varphi}_{kin}\rangle^2 = \frac{1}{2} N k_B^2 T^2 \tag{4.14}$$

where $E^{r,\varphi}_{kin}$ is the total kinetic energy of the N degrees of freedom $\{r_n\}$ and $\{\varphi_n\}$ respectively, k_B the Boltzmann constant and T the fixed temperature.

The system takes a long time to reach equilibrium, defined when kinetic energies satisfy Equations (4.13), (4.14). After an initial transient of $125000\,t.u.$, we calculate the value of kinetic energy and its fluctuations, and the percentage of denaturated base pairs, averaged on an equal time. The averaged values are in fact, choosing these time intervals, well stabilized and coherent with canonical expectation values. Equations (4.13) and (4.14) are satisfied with an error of less then 0.5%.

The averaged number of open pairs is calculated by considering as denaturated a pair with $r_n > 8\ln 2/\alpha \approx 1.25$ Å, where $y = \ln 2/\alpha$ corresponds to the inflection point of the Morse potential. All simulations were performed with *free boundary conditions*. We report in Figure 4.1 the resulting denaturation curves for the *twist-opening* model with the parameter choice already discussed in the previous chapters, and with different choices of the unknown G_0 constant.

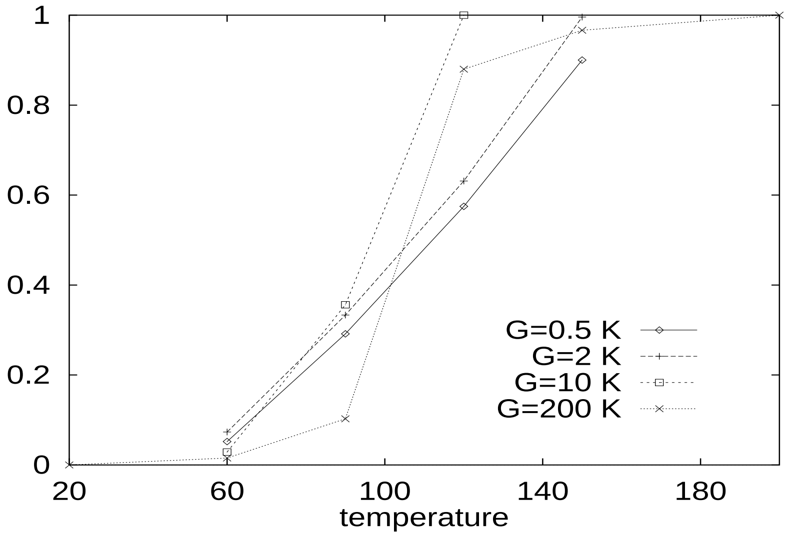


Figure 4.1: Denaturation curves obtained by numerical integration of the *twist-opening* model with constant temperature conditions. We report the fraction of base pairs which satisfy the relation $r_n > 8\ln 2/\alpha \approx 1.25$ Å. Different curves are obtained with $G/K = G_0/(R_0^2 K)$ values as indicated in the legend. We calculate six points for each curve, corresponding to temperatures $T = 20, 60, 90, 120, 150$ and $190K$ (with one more point for the case $G = 0.5K$).

The results show that, with the chosen set of parameters, the transition temperature T_{den} is approximatively in the range $90 \div 110\,K$. This values are actually well below the real ones. This could suggest, once more, relevant errors in the choice of the parameters, or, otherwise, that we need to improve our model in some way. We will see in Section 4.5.3 that the introduction of W_{stack} affects strongly the denaturation features of the model, with a relevant improvement of the value of the transition temperature.

The dependence on G_0

Moreover, we observe that the curves presented in Figure 4.1 show a dependence of some features of transition on the value of G_0. As a general result, we can say that the temperature interval in which denaturation arises is narrower for greater values of G_0. According to our predictions, denaturation is thus affected by the rigidity of the two backbone strands with respect to bending. The effect is such that the more rigid are the two helices that forms the duplex, the more rapid is the denaturation transition as a function of the temperature. This can be easily understood. We know that the formation of a denaturation bubble in a given region implies an untwist of the same region. The term of the Lagrangian depending on G_0 prevents strongly localized changes of the twist angle along the chain, with increasing efficiency as G_0 increase. More precisely, larger G_0's do not affect the possibility to large twist distortions, but favor smooth distortions, *i.e.* they request that various twists change happen quite simultaneously along the chain. Small bubbles will thus form and grow more easily in the case of a "soft" helix, while a rigid one will tend to stay close and to open as a whole when the total kinetic energy reaches a global denaturation threshold.

In the context of this first results it seems that a large value of G_0 will be more efficient in order to reproduce the very narrow temperature range of the real DNA denaturation transition. This indication is anyway too vague to allow for a specific choice. We can expect that denaturation becomes hard to achieve for too rigid helices and try then to fix an intermediate value[3], but this poses problems for the simulations, that become longer and longer as the G_0 parameter increases, due to the fact that a sort coordination in the untwisting motion is needed. Interestingly, this "temporal" behavior could be also taken into account in order to mimic real denaturation experiments. Anyway, the study of some more direct structural features will turn out to be more useful in order to obtain an evaluation of G_0.

Dynamics at constant temperature

Radial configurations

In order to understand better the dynamical mechanisms underlying the denaturation transition we register the configurations of the chain at thermal equilibrium for different temperatures, for a time interval of about $1950\, t.u.$ Radial configurations can be conveniently displayed in a grey-scale density plot, where the horizontal axis corresponds to the spatial index n along the chain, the vertical

[3]Furthermore, we remind that the shape of the dispersion relation curves depends on G_0. For large values of G_0 (approximatively above $R_0^2 K$) the gap between the two branches disappear. This implies that, in principle, values of G_0 larger than 1 has to be excluded in order to have gap breathers. These remarks must be taken into account in order to make a definitive choice of G_0.

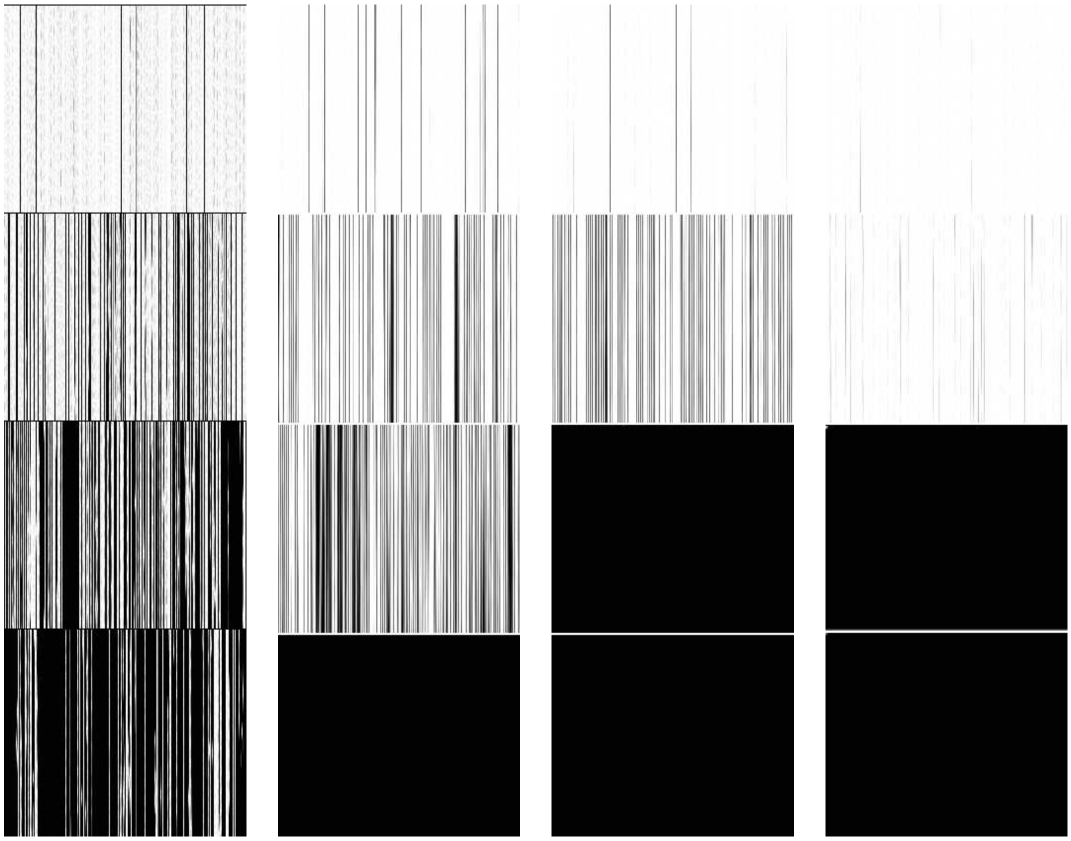
axis to time, and the grey level indicates the hydrogen bond length: from white, for close base pairs, to black, which corresponds to open base pairs with bond stretch greater than 1.2 Å.

Figure 4.2 reproduces the radial configurations for the cases $G_0/(R_0^2 K) =$ 0.5, 2, 10, 100 (columns (a), (b), (c), (d) respectively) for the four temperatures $T = 60$, 90, 120, $150K$ (from top to bottom). Dark vertical strips correspond to stable opening. These openings appear at some time and remain well localized in space for a long time. Their amplitude is much larger and they are much more localized in comparison with analytical stable "acoustic" solutions we are able to build in the context of the multiple scale expansion technique. Anyway, they show properties which correspond in some way to those of the analytical solutions: they do not show internal oscillations, in agreement with the results on the existence of optical breather with low wave number in the model.

We notice that the absence of internal oscillation is true mainly for larger distortions, and, therefore, for higher temperatures. For low temperatures there are some smaller oscillating distortions, which are maybe in the phonon regime, slightly visible in the upper part of graphs reported in Figure 4.2. We can confirm the presence of these smaller oscillations by using two different colors for opening and compression of base pairs. The resulting space-time plot is reported in Figure 4.3. It corresponds to a piece of the graph of Figure 4.2 with $G_0 = 0.5\, R_0^2 K$, taken between temperatures $60K$ and $90K$: we show in green the slightly stretched regions and in red the compressed ones, while the larger distortions are again in black and equilibrium corresponds to white.

As for the PB model, we can say that in our model energy equipartition is not present at constant temperature, at least for small and medium time scales. Energy in fact strongly localizes on some lattice sites instead of distributing homogeneously on the chain. Equipartition will be probably restored on longer time scales, in a statistical sense, because local openings can appear and disappear randomly on each site along the chain. Anyway, persistence reveals a slow decay of correlation.

Localized openings increase in number and extensions as the temperature increases. This is clearer in the first plot (a), where the transition is smoother. Open regions extend progressively over the whole chain, up to a complete denaturation. For the last three G_0 values, in fact, the chain is completely denaturated at $T = 150K$; in the last two cases it denaturates already at $120K$. The plots in Figure 4.2, in fact, confirm the fact that increasing G_0 one gets a sharper transitions. Even if they denaturate already at $T = 120K$, the chains with larger G_0 values displays for lower temperature much less open base pairs with respect to the lower G_0 cases, in agreement with the presence of a more rapid transition.



(a) (b) (c) (d)

Figure 4.2: **Radial** configurations for various choices of the parameter G_0: **(a)**: $G_0 = 0.5KR_0^2$, **(b)**: $G_0 = 2KR_0^2$, **(c)**: $G_0 = 10KR_0^2$, **(d)**: $G_0 = 100KR_0^2$. From top to bottom: $T = 60$, $T = 90$, $T = 120$, $T = 150K$. Each image shows the system evolution on a time interval of $1950\,t.u.$ (time increase from top to bottom). The horizontal coordinate corresponds to the position along the chain n; the bright regions to closed base pairs, the darker ones to base pairs stretched over 1.2 Å.

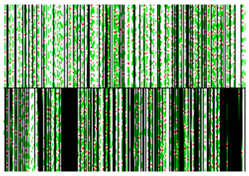


Figure 4.3: Radial configurations: smaller oscillating distortions. We report a fragment of the previous image with $G_0 = 0.5KR_0^2$ and, from top to bottom, $T = 90$, $T = 120$. We show in green the slightly stretched regions and in red the compressed ones. the darker regions correspond to base pairs stretched over 1.2 Å.

Twist configurations

Using the same graphical technique we can also visualize *twist* distortion along the chain. Furthermore, by using two colors, it is possible to distinguish untwisted regions from overtwisted ones. The general behavior of twist in the thermalized chains whose radial configurations are reported in Figure 4.4 shows a clear tendency to untwist the helix. The global untwist increases with temperature, in evident correlation with the global percentage of open base pairs. Large untwisted regions alternates with small regions where twist is increased, due to the geometrical constraints on the winding of the two strands. When the chain opens completely, the twist decreases by more than $0.8\,\Theta_0$ (which is the value chosen for the threshold to full red in the color plots) everywhere along the chain.

It is interesting to discuss the role of the strand rigidity on the thermally induced twist distortion. For the two chains with$G_0/(R_0^2K) = 10$ and 100 and for the lower temperatures, in fact, twist remains much closer to its equilibrium value and no bubble formation is observed. Instead, extremely localized, small, positive twist distortions appear. This overtwist points seem more movable and their density is approximately equivalent to that of the negative twist distortions, which are much less important with respect to the small G_0 cases. This confirms that an increasing strand rigidity prevents twist deformation, and, thus, base pairs opening.

When temperature increases enough, *i.e.* for $T = 120$, $150K$, and we wait long enough, all base pairs open and we observe the global untwist of the whole chain as for the other cases.

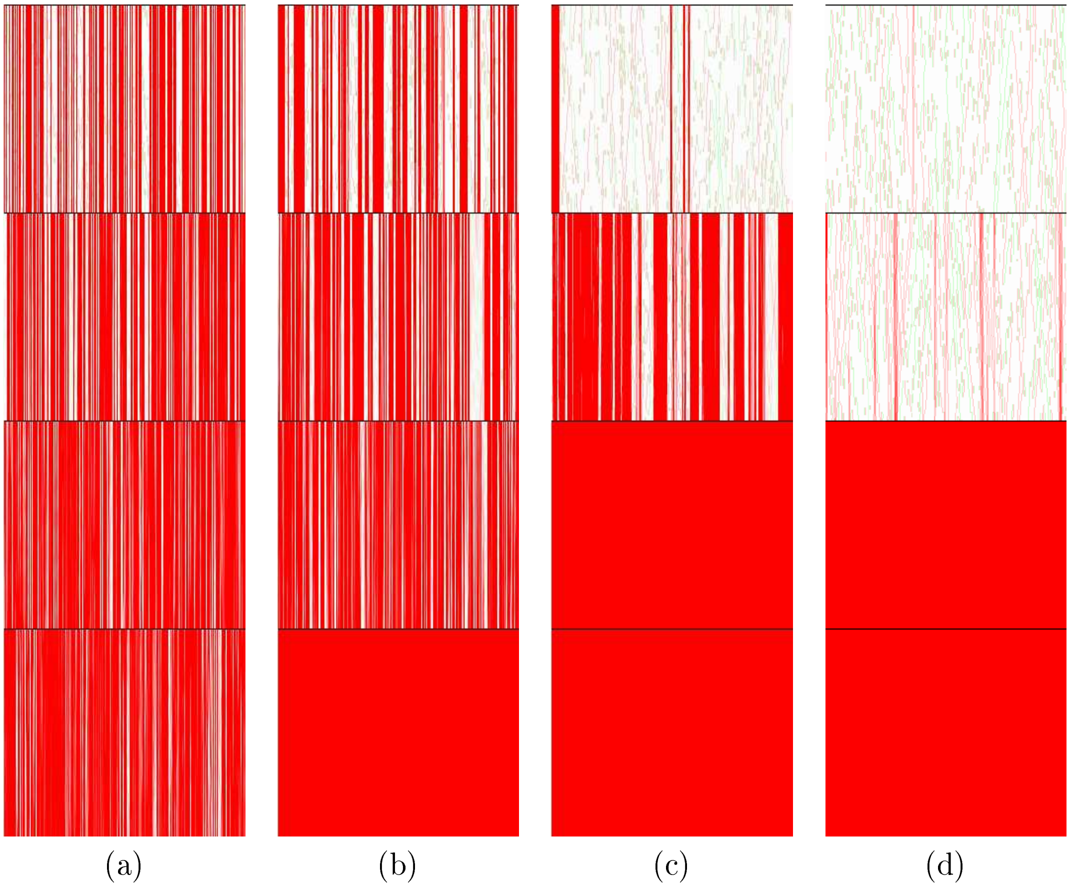


Figure 4.4: **Twist** configurations for various choices of the parameter G_0: **(a)**: $G_0 = 0.5KR_0^2$, **(b)**: $G_0 = 2KR_0^2$, **(c)**: $G_0 = 10KR_0^2$, **(d)**: $G_0 = 100KR_0^2$. From top to bottom: $T = 60$, $T = 90$, $T = 120$, $T = 150K$. Each image shows the system evolution on a time interval of $1950\,t.u.$ (time increase from top to bottom). The horizontal coordinate corresponds to the position along the chain n; the red regions to untwist, the green ones to overtwist distortions.

Incorrect results in reproducing DNA dynamical features without W_{stack}

From discussions in Section 3.6 we know that our first version of the *twist-opening* model, even if it reproduces the helical geometry of DNA in a better way with respect to previous models, has some important drawbacks. Namely, it does not possess the approximate breather solutions which can be derived from linear modes on the acoustic branch. These solutions are especially interesting because they reproduce the DNA "breathing" experimentally observed at room temperatures: it is in fact the possibility of a thermal activation that confers to these kind of solution their importance in the context of DNA dynamical studies. In this section we have seen that thermal energy actually does not activate oscillating solution, but distortions that remain opened or oscillate very very slowly.

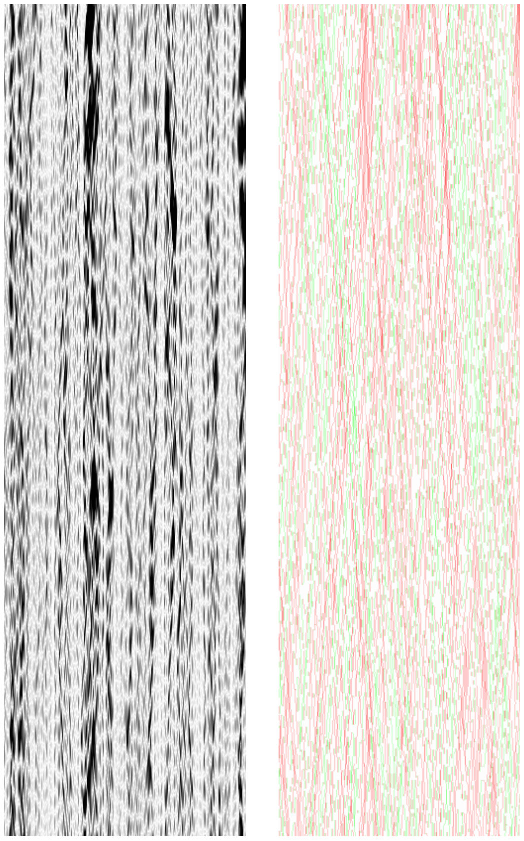
4.5.2 "Breathing" in the improved *twist-opening* model

We know from Chapter 3 that the addition to the model of the direct stacking term W_{stack} allows to recover analytical oscillating solutions. Furthermore, the oscillations have frequency of the same order of magnitude of that experimentally observed in DNA "breathing". We look now at the properties of the improved model when put in equilibrium with a thermal bath at room temperatures, in order to understand if its behavior actually reproduces these known DNA dynamical properties.

We consider here just the case $G_0 = 2\ R_0^2 K$. Being the model constants not yet fixed, we fix the other parameters to the same values we used for the original model and we arbitrarily take $S = 2K$. We fix $T = 300K$ and we plot the radial and twist chain configurations, with the same graphics of previous figures. The plots are reported in Figure 4.5.

From Figure 4.5 we immediately see that the improved model shows a different thermal induced behavior with respect to the original one. Oscillating distortions are now visible; their frequency can be approximately calculated from the picture and turns out to be that of real DNA breathing modes. We will discuss more in detail this result in the concluding section 4.5.3.

Furthermore, we observe that the improved model, at a temperature of $300K$, is still below the denaturation transition. Its thermal behavior is thus in better agreement with real DNA properties: DNA denaturates in fact in the range of $326 \div 370\,K$.



(a) (b)

Figure 4.5: Model configurations for the improved model at $T = 300K$, with $G_0 = 2\ R_0^2 K$, $S = 2K$: **(a)**: radial configuration; **(b)**: twist configuration. Each image shows the system evolution on a time interval of $7800\,t.u.$ (time increase from top to bottom). The horizontal coordinate corresponds to the position along the chain n. For radii, the bright regions correspond to closed base pairs, the darker ones to base pairs stretched over $1.2\,\text{Å}$. For angles, the red regions correspond to untwist, the green ones to over-twist distortions.

4.5.3 Denaturation of the improved *twist-opening* model

We have then considered higher temperatures and we have looked at the denaturation transition in the improved model. We have examined four cases which correspond to the choices $G_0 = 0.5,\ 2,\ 10,\ 100\ R_0^2 K$, fixing the other parameters to the values we used for the original model and S to $2K$. We have calculated denaturation percentages at various temperatures between $T = 370K$ and $T = 580K$. Time intervals used in simulations are the same as for the previous case.

Because of the necessity of a better choice for the parameters of the improved model, we can only refer about preliminary results. Anyway, these results allow to conclude that the *twist-opening* model, in its final form, reproduces quite well the DNA denaturation features and lead to important conclusions on the relevance of the helical geometry in this process. Let us discuss the model behavior in with respect to the various features of the denaturation process.

Denaturation transition.

From our first calculations we deduce that the improvement of the model leads to important results for what concerns the denaturation curves. Denaturation temperature can be now situated in the range $430 \div 520K$, which is much closer to the true DNA value that the value found for the original model, even if it is, this time, too large. Furthermore, T_{den} seems now to depend more on G_0, and thus it may be adjust through a correct parameter choice.

The interval of temperature in which denaturation arises also depends on the choice of G_0, and is narrower for larger values of this parameter, as for the not improved model. The dependence of the transition rapidity on the parameter G_0 can be again related to the greater difficulty to locally untwist the helix if the strands forming its structure are more rigid with respect to bending. This effect results in a minor density of untwisted regions in rigid helices, and therefore in a decrease of the number of open base pairs.

The denaturation interval can be roughly estimated of the order of $100K$, or less for the sharper cases: it is thus again too large in comparison to the few degrees of real denaturation curves. Anyway, an improvement of this behavior appears possible at present. We probably need a more detailed reconstruction of the denaturation curves in order to evaluate in a precise manner the temperature interval. We stress that the two properties which characterize denaturation curves, namely T_{den} and the curve width, allow the determination, by curve fitting, of two parameters. Accurate curve calculations is quite cumbersome, but can lead in principle to a determination of our unknown parameters S and G_0.

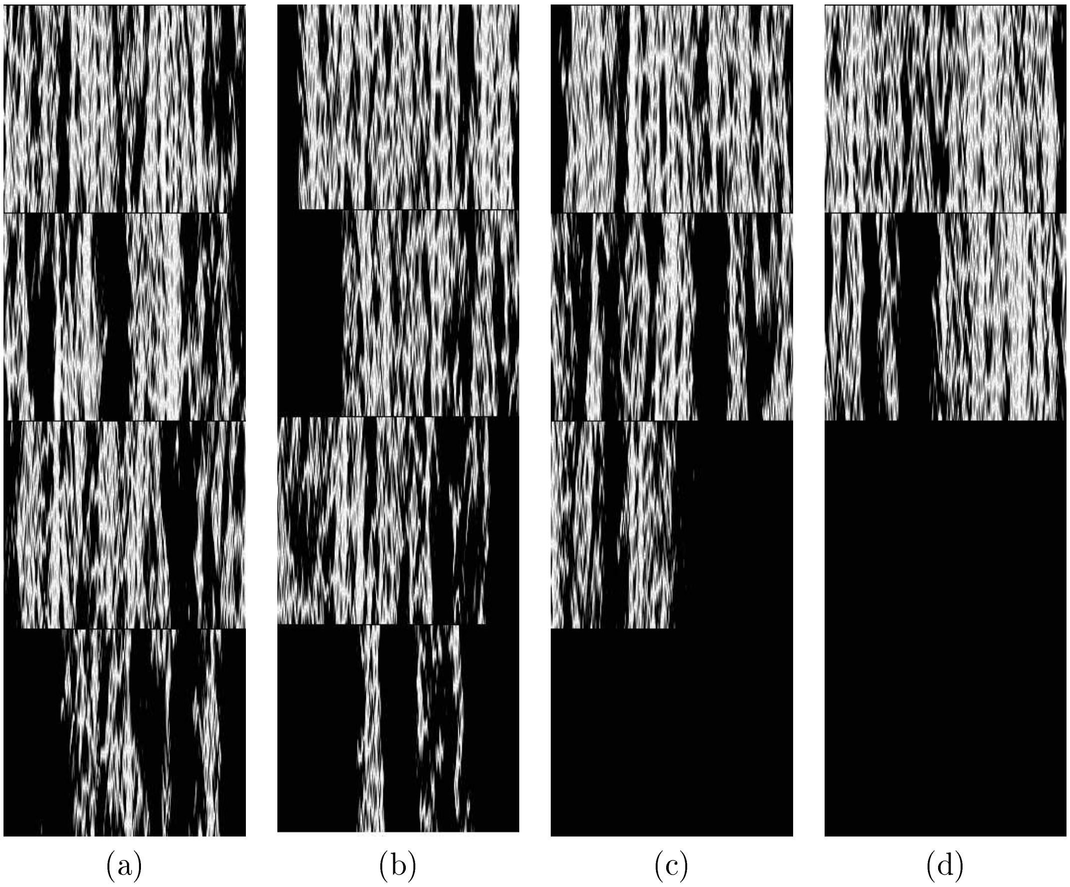
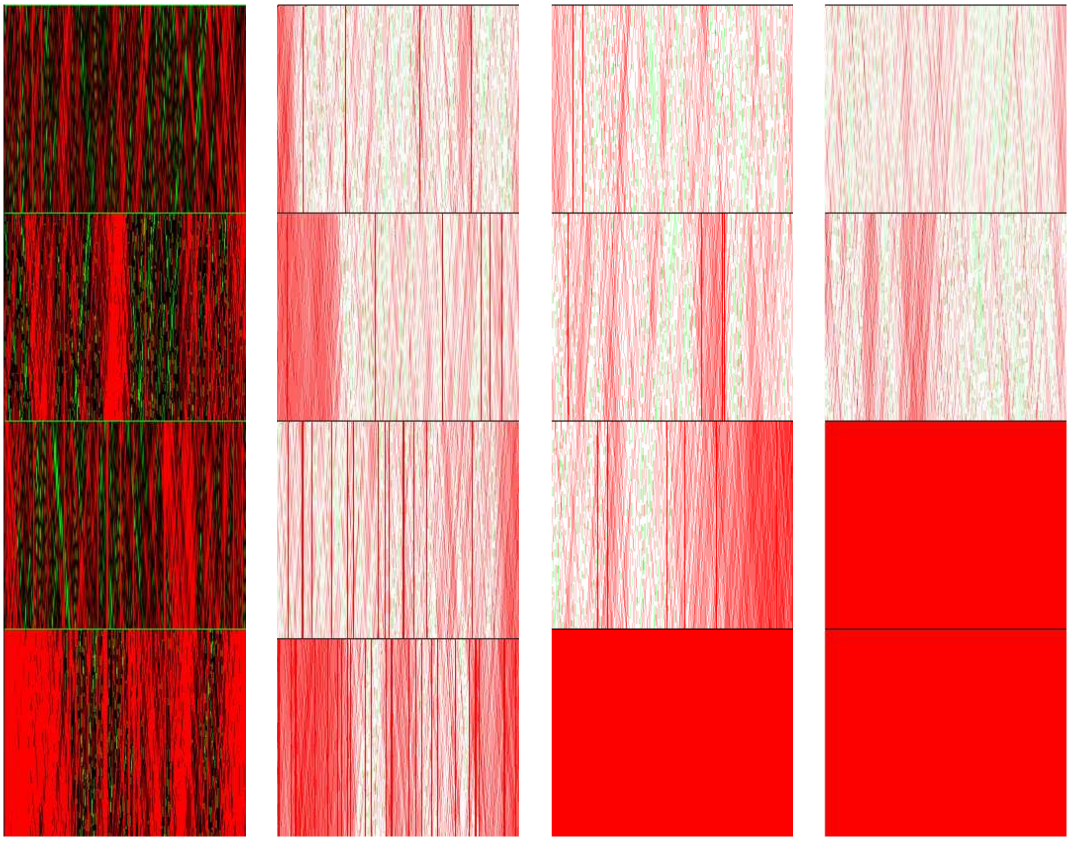


Figure 4.6: **Radial** configurations for the improved model with $S = 2K$, and for various choices of the parameter G_0: **(a)**: $G_0 = 0.5KR_0^2$, **(b)**: $G_0 = 2KR_0^2$, **(c)**: $G_0 = 10KR_0^2$, **(d)**: $G_0 = 100KR_0^2$. From top to bottom: $T = 430$, $T = 460$, $T = 490$, $T = 520K$. Each image shows the system evolution on a time interval of 1950 $t.u.$ (time increase from top to bottom). The horizontal coordinate corresponds to the position along the chain n. Grey scale ranges, from white to black, from 0 to 1.2 Å.



(a) (b) (c) (d)

Figure 4.7: **Twist** configurations for the improved model with $S = 2K$, and for various choices of the parameter G_0: **(a)**: $G_0 = 0.5KR_0^2$, **(b)**: $G_0 = 2KR_0^2$, **(c)**: $G_0 = 10KR_0^2$, **(d)**: $G_0 = 100KR_0^2$. From top to bottom: $T = 430$, $T = 460$, $T = 490$, $T = 520K$. Each image shows the system evolution on a time interval of 1950 $t.u.$ (time increase from top to bottom). The horizontal coordinate corresponds to the position along the chain n. The red regions correspond to untwist, the green ones to overtwist distortions. Black corresponds to regions where twist is at rest.

Breathing.

As expected, configuration plots at lower temperature (Figure 4.5) clearly show that oscillating openings can be thermally generated, in agreement with the presence of optical breather solutions, for what concerns the model, and with the observed "breathing" phenomena, for what concerns real DNA. In fact, on one hand, the frequency of the oscillations that we observe in the model at constant temperature has approximatively the same order of magnitude of that of the breathers analytically found in the previous section; on the other hand, the same frequency is actually observed experimentally in DNA. The agreement between the thermal behavior and the main features of analytical solutions is an indication of the fact that breathers can be thermally generated in the model. The agreement with experimental results shows in turns that the existence and biological relevance of soliton-like excitations in real DNA turns out to be confirmed as an interesting work hypothesis.

Bubble formation.

We report in Figures 4.6 and 4.7 configuration plots for $T = 430, 460, 490$ and $520K$, showing the radial and twist distortions for the improved *twist-opening* model. We use again four different values for G_0. All simulations are performed with the same integration times of the previous cases (Figures 4.2 and 4.4).

Figures 4.6 and 4.7 show that, at higher temperatures, near denaturation transition, we observe the formation of larger bubbles, that gradually cover all the chain while the temperature approaches T_{den}. Bubbles grow with temperature increase up to a complete strand separation, as expected.

Untwisting.

Figures 4.6 and 4.7 show another important characteristic, related to the twist degree of freedom and to its coupling with the base pair opening. It is evident in fact from the pictures that even a partial denaturation of the chain is possible only if a corresponding untwist is present. See, for instance, the case of $T = 490K$ and $G_0 = 10KR_0^2$, where the right part of the chain is completely opened and strongly untwisted. The chain untwisting confirms that torsion is a fundamental parameter in studying DNA opening.

Furthermore, pictures corresponding to these opening-untwisting conditions clarify that such a global untwist can be achieved only by an overall rotation of the two chain ends, which leads to eliminate extratwist from the chain. Opening is possible, *i.e.*, only with free boundary conditions. It could be therefore interesting to perform simulations with periodic boundary conditions, mimicking the case of a circular DNA. Some preliminary results show as expected that denaturation is prevented in this conditions: the two strands of a closed molecule have a fixed number of turns and cannot untwist over large regions. We can finally come back to what we have discussed in Chapters 1 and 2: the helical structure and the related geometrical constraints are unavoidable elements for a good description of dynamical opening of DNA.

All this results lead to a deeper understanding of the mechanisms underlying transition to the denaturated state in the *twist-opening* model, which confirms our expectations. Transition arise via formation and growth of bubbles characterized by an opening of base pairs coupled with a local untwist of the double helix. Small oscillating localized distortions characterize the early phase of this process, at low temperatures. The excitations extension increase with temperature, in a temperature interval that depends on the strand rigidity parameter G_0. Larger bubbles do not oscillate. Their extent increases and they finally cover the whole chain which is thus completely denaturated. As a general result, we have shown the necessity of a chain untwisting in order to open the chain, and thus the relevance of the chain boundary condition. *The proposed model mimics the known DNA dynamical properties at the different temperatures.*

A further comment on DNA modeling

To conclude this chapter, we would like to stress the importance of looking at the dynamical behavior of a DNA model in constant temperature conditions. A common criticism against DNA models which refer to soliton theory is that these studies do not consider the environmental conditions in which the real system lives. More specifically, the fact that it is not taken into account the possible disruptive effect of thermal noise is often considered as a strong limitation of their theoretical interest.

Constant temperature simulations not only allow a more realistic approach to dynamical model studies. As we have seen in this chapter, they also lead to make interesting observations of the features of the model. In the case of non linear models, in fact, thermal energy does not necessarily act in a *disruptive* way, by dispersion effects, but can have a *constructive* effect, because of energy localization. This is in agreement with the fact that in nonlinear systems, *e.g.* in our case, equipartition of energy does not hold on short and intermediate time scales, so that energy is "organized" in more or less complex structures. These structures can have a great interest in the context of studies which try to reconstruct the behavior of complex natural structures.