Introduction

Biology is today a source of a large number of experimental results, which refer to features of biological processes in a more or less detailed way. Anyway, observations are often related to indirect properties. Problems of interpretation arise in relation with specific features that cannot be directly investigated with the available experimental techniques. Therefore, one must reconstruct the underlying mechanisms by making hypotheses that have then to be verified, maybe through experiments involving other secondary effects.

Modeling represents an important *tool* in order to investigate biological processes that are not yet understood because of experimental difficulties. A model can be tested by careful attempts to reproduce known experimental results via numerical simulations. Once having assured that the model is capable of reproducing the known processes in the context of interest, it is possible to look at the detailed mechanisms which lead to this reproduction. This often gives a better understanding of the processes themselves. On the other hand, one can at this point formulate hypotheses on other aspects of the process on the basis of the model features, suggesting new directions of experimental research. In this sense, we can see modeling in biology as an intermediate step between available results and open problems.

DNA offers by itself a good example of the complexity mentioned above. This macromolecule, which can have a length of the order of a kilometer, exists in one copy in (almost) every living cells, where it has fundamental functions. The relation between *structure* and *function* is today a central paradigm in DNA studies. During all processes involving DNA the spatial configuration and the local shape of the molecule play very important roles, being implied in interactions with auxiliary proteins and external machineries. The extreme molecule length, the limited available space inside the cell, and the structural constraints related to the helical geometry of the molecule are all elements which contribute to the complexity of various phases of DNA functioning. The many supercoiling theoretical and experimental studies pointing out the relations between the double helix geometrical configuration and the molecular functioning are examples of these investigations [1] - [6].

The fundamental function of DNA is to contain the coded instructions for the synthesis of proteins, which are used to build all the cell structures and which mediate most of the cell living processes. These instructions are stored in the sequence of the four bases (adenosine, guanine, cytosine and thymine) which compose the DNA core, protected by the external double helix backbone. Therefore, DNA has to be opened, by separating the two strands which are coiled into the double helical structure, in order to expose the base sequence for reading.

DNA opening is a *dynamical process*. This process is strongly constrained by the molecular structure and by specific local interactions with enzymes. It is regulated in a complex manner. Moreover, it is necessary in order to achieve protein synthesis. Nevertheless, its actual mechanism is far from being understood. In order to have a clearer picture of the processes which lead to protein synthesis, a deeper study of the dynamical features which characterizes DNA opening would represent an important contribution.

There is, in fact, a growing interest in DNA *dynamical modeling*. Many works published in this field [7] - [14] have the merit of extending DNA theoretical studies to the physical analysis of the molecular motion, which had been neglected before. Many models have been proposed to explain the energetics and specific mechanisms of important processes like denaturation, transcription and duplication of DNA. These studies often refer to *non linear models* which possess soliton solutions: solitons could in fact explain the strong energy localization and the existence of moving localized distortions which characterize some of the biological features under study. The work we present in this Thesis belongs to this line of research.

The aim of this kind of theoretical works is not that of a detailed reproduction of the whole DNA structure (which is instead the aim of molecular dynamics simulation techniques), but that of a mesoscopic description which can take into account some central structural elements, allowing thus a deeper understanding of the true dynamical mechanisms involved. Anyway, some of the structural constraints, which have a strong effect on these mechanisms, have to be included to obtain a more realistic description.

We know from biology that a local opening of the two strands which form the double helix is not possible without a local uncoiling of the helix itself, *i.e.* a local decrease of the angle that describes the torsion of the structure around the main molecular axis. *In this Thesis we propose a helical dynamical model of DNA.*

We will start from the planar model proposed by Peyrard and Bishop (*PB*) [14]. The PB model just focus on the distance between the two bases in a pair which is stretched when the two strands open. Even if it does not represent the main geometrical constraints related to the helical structure, it has led to several results in explaining some of the dynamical features of DNA. Nevertheless, it cannot be considered as a realistic model, because it neglects constraints that, in real DNA, are known to have strong and even dramatic effects on base pair opening.

Our aim is to build a more realistic model, which takes into account the heli-

cal structure and then the torsional deformations induced by base pairs opening. Starting from the PB model, we add to the base pair opening one more *angular* degree of freedom per site, which allows the description of the helical geometry of the molecule. We then complete our model by explicitly referring to the helical constraints in order to obtain an appropriate Lagrangian. *The resulting* twist-opening model [15, 16] *represents more exactly the molecular shape, and is an attempt to include in a unique scheme dynamics, structure and functioning of DNA.*

In order to give a general outline of this work, let us summarize the results we present.

Chapter 1 is devoted to a general description of *DNA structure, functions and known dynamical properties.* We also describe in detail the main features of the transcription process, pointing out the mechanisms involved in its activation and stressing the features which still remain unknown. This chapter will lead to a definition of the biological processes and of the interpretative problems we are interested in. Furthermore, it introduces the main structural parameters and degrees of freedom of the DNA structure which are involved in the dynamics of the biological processes of interest.

In *Chapter 2* we first remind briefly the results of the *Peyrard Bishop model.* Like other DNA models, the PB model possesses soliton like solutions; namely, *discrete breather solutions*, characterized by an internal fast oscillation modulated by a localized envelope soliton, which is solution of a Non Linear Schrödinger (NLS) equation and moves along the chain with constant velocity. These are approximated solutions, obtained by application of the *Multiple Scale Expansion* technique. Nevertheless they are characterized by a good stability. Furthermore, typical oscillation frequency and extent of the obtained distortions agrees with spectroscopy results on DNA dynamical properties at room temperature.

We then explain why the PB model does not describe in a realistic way DNA dynamics.

We present the *twist-opening* model through a "step by step" building of its Lagrangian. The model has the correct helical shape at equilibrium and allows the correct coupling between torsional deformation and base pair opening: we will confirm this fact by the numerical integrations of the equations of motion with simple initial conditions. We will then discuss the choice of the model constants on the basis of experimental data on real DNA. Finally, we will write down a rescaled approximated version of the equations of motion for the *twist-opening* model, that will be used in the following chapter in order to find small amplitude solutions.

The existence of moving breathers in the PB model suggests to look for analogous solutions in the generalized helicoidal model. In *Chapter 3*, we deal with the problem of *finding small amplitude soliton-like solutions for the model.* This is one of the central topics of our work.

In order to solve this problem we need the extension of the same mathematical tool used for the PB model, namely the Multiple Scale Expansion (MSE), to the case of *vectorial lattices* with more than one degree of freedom per site. The extension, which is not trivial, can be performed in a general way. The aim is the derivation of small amplitude soliton-like solutions whose envelopes obey a NLS equation: the central point of the derivation is to look for a wave packet solution where a weak dispersion effect is balanced by a weak nonlinearity. The expansion in multiple scales consists essentially in introducing sets of independent variables to describe on one hand the slow envelope time-space behavior and on the other the central mode fast oscillatory motion. With respect to the case of systems with only one degree of freedom per site, where a scalar equation has to be solved, in the case of vectorial fields the linear part corresponds to an eigenvalue-eigenvector problem, so that one has to perform a *perturbative expansion* instead of a Taylor series expansion. To combine the two effects of dispersion and nonlinearity one has then to perform two parallel expansions. The two expansions can be combined in a unique operatorial technique, easily applicable to complex systems, which allows to obtain the NLS equation for the envelope function in the general case of Klein-Gordon vectorial lattices [17].

We will apply this technique to the *twist-opening* model. We will find in this way special soliton envelope solutions characterized by a *local oscillating opening coupled with an untwist of the helix.* This result confirms that the model behaves as DNA, for what concerns the coupling between strand separation and untwist of the helical structure.

We have performed numerical integrations of the equations of motion of the model, over long times, in order to test the stability of the approximate solution, which turns out to be good as long as their amplitude does not increase too much.

With the chosen set of parameters, if we want to obtain solutions that reproduce the known DNA features, we have to introduce a very simple *improvement of the model.* The Lagrangian of the model has to be correct by introducing *one more coupling term*, which is justified on the basis of neighboring base pairs *stacking* interaction in real DNA. We then show that, with this correction, our model reproduces quite well the known DNA dynamical features. The analytical solutions arising from the application of the expansion technique and the numerical results of direct simulations agree now with the expected behavior.

The importance of the introduction of the new term is confirmed in *Chapter 4*. Here we test in fact our model from a quite different point of view: we perform *canonical ensemble simulations.* This allows to investigate the properties of the model at constant temperature, *i.e.* in more natural condition. We visualize the radial and angular chain configurations at various temperatures and compare the results with the expected behavior. When heated, DNA tends to form local openings or bubbles, whose extension increases with the increase of the temperature up to a complete strand separation, called *denaturation*. The constant temperature simulations also allow to compare the model behavior at increasing temperatures with the available data on DNA denaturation. The temperature at which denaturation arises and the temperature interval which characterizes the transition are well confirmed experimental data. The comparison of their value with those obtained for the model using simulations represents a good test for the model itself.

The resulting thermal behavior for the model without correction term is again quite different from real DNA dynamical and statistical properties. Instead, the improved model shows better agreement: denaturation temperature is much closer to the real one. The specific dynamical mechanisms allowing the transition show features that are related to the analytical results of Chapter 3. This means that the analytical solutions we have found actually describe distortions which can be thermally activated, and which therefore can represent spontaneous excitations in DNA chains.

The improved *twist-opening* model represents thus quite well DNA dynamics. The results presented in this Thesis allow to conclude that this model can be a useful tool in future DNA studies.

Note: The most important quantities defined in this Thesis, related to the model or to the calculation technique, will be indicated in the page margins. We do not indicate systematically all constants and functions, but focus on those that are recalled elsewhere, to facilitate the reader.