

# STATISTICAL PHYSICS MODELS OF DNA DENATURATION

M. TIERZ

ABSTRACT. In this work we present some physical approaches to the study of DNA denaturation. After a short motivation for studying DNA denaturation, we present the main features of recent models that implement volume exclusion interactions in several different ways. The inclusion of this type of interactions are seen to be rather crucial in the determination of the type of phase transition. Some conclusions are presented and we discuss several rather subtle issues concerning the type of excluding volume considerations and the type of phase transitions.

## 1. INTRODUCTION

Natural DNA exists as a double-helix strand. Upon heating, the two strands may separate. This unbinding transition is called DNA denaturation. Denaturation is currently induced in laboratory to obtain single stranded DNA (ssDNA). Experimentally, DNA denaturation has been investigated since the sixties: the amount of UV light (260 nm) absorbed by the DNA is known to be proportional to the fraction of unbound base pairs in the chain. By increasing the temperature, the absorbed UV light increases through a series of sharp jumps, which are reminiscent of a first order transition. Quite a lot of efforts have been devoted to understand the origin of such sharpness, in particular recent attention was devoted to the effects of the double helix stiffness and of excluded volume interactions. In addition to the interest of the process *per se*, its role in biotechnical applications is noteworthy. A clear example is the key role played by thermal denaturation in the famous Polymerase Chain Reaction amplification procedure (see Appendix for a little bit more of Information on this topic).

Interestingly enough, in a statistical physics context, a simple model of the DNA denaturation transition was introduced already in 1966 by Poland and Scheraga [1]. Thus, simple models of DNA thermal denaturation have attracted attention for quite a long time.

The model consists of an alternating sequence (chain) of straight paths and loops, which idealize the denaturation of DNA, consisting of a sequence of double stranded and single stranded molecules. An attractive energy is associated with paths. Within this scheme the full partition function  $Z_N$  for a chain whose constituent strands have lengths  $N$ , is factorized in terms of elementary partition functions for the loops and segments.

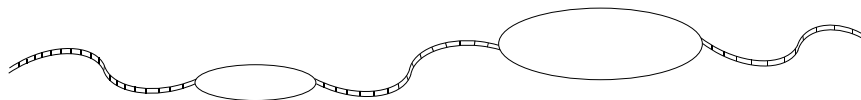


FIGURE 1. Schematic representation of the model

Thermal denaturation of DNA is considered to be representative of a typical competition between energy and entropy. When we are at low temperatures, the binding energy of the base pairs dominates the thermodynamics while as soon as the temperature is raised, sections of the DNA separate, because this is a configuration with greater entropy.

In principle, the original model ignored interactions between different parts of a chain. This includes of course all details regarding real DNA such as chemical composition, stiffness or torsion. While we shall describe the model and the recent modifications that make the model somewhat more realistic in the next Section, let us advance here some of the results obtained in [1]. The main result we would like to emphasize is that the phase transition was found to be determined by a parameter  $c$ , which characterizes the statistical weight of a loop. The number of configurations of a loop of length  $\ell$  behaves as  $s^\ell/\ell^c$  for large  $\ell$ , with  $s$  a non-universal constant. In this early models, it was shown [1] that the phase transition is first order if  $c > 2$  and second order if  $1 < c \leq 2$  while for  $c < 1$  no transition takes place and the strands are always bound. The loop is typically modelled using random-walk configurations that lead to  $c = d/2$  in  $d$  dimensions. Therefore, note that the transition turns out to be continuous in  $d = 2$  and  $d = 3$  dimensions.

Therefore, this result is in contrast with the experimental observations above mentioned. In this sense, the model was later generalized by Fisher to take into account the self-avoiding interactions within each loop [3, 2]. In this situation, while the loop entropy remains the same, the exponent  $c$  takes the value  $d\nu$ , where  $\nu$  is the correlation length exponent of a self-avoiding random walk, which is typically well-known. These values of  $\nu$ , lead to a  $c$  larger than the case of a random-walk model, but still they were not big enough, since these values are still smaller than two, both in  $d = 2$  and  $d = 3$  dimensions, and therefore the phase transition predicted is still continuous (somewhat sharper, but still continuous). Fisher himself suggested [3] that self-avoiding interactions *between* the various parts of the chain and not just within loops, would further sharpen the transition possibly making it first order.

Nevertheless, it was not known how to implement analytically this more realistic scenario into the model. This task has been partly accomplished in recent years [4, 5], employing results -that we shall also briefly describe- on polymer networks of fixed topology [6].

This work is organized as follows. In the next Section we attempt to explain recent work on the Poland Scheraga model that in addition incorporates excluded volume interactions, and explore the corresponding effect on the type of phase transition. We recall the basic formalism for the computation and the necessary results -to account for interactions- in the context polymer networks of fixed topology are briefly introduced as well. In the next Section, we present a rather different model that includes only intrachain excluded volume interactions and reduces the problem to a quantum-mechanical problem. In the last chapter we briefly summarize the work and discuss about the role of various type of excluded volume effects in the order of the phase transition.

## 2. POLAND-SCHERAGA MODEL WITH EXCLUDED VOLUME INTERACTIONS.

As explained in the introduction, the old models have been recently revisited with the inclusion of volume excluded effects and the corresponding impact on the order of the phase transition has been addressed. This phase transition is expected to be,

on experimental grounds, of first order, and not continuous as the original models imply. First of all, we present the features of the model that are independent of the inclusion of these more realistic terms. Namely, the -grand canonical- partition function and the possible nature of the phase transition, that turns out to depend in the type of entropic contribution provided by the denaturated loops. Then, we briefly sketch results on polymer networks of fixed topology, that were employed in [6] to consider the interaction of the denaturated loops with the rest of the chain. The main conclusion is that, for loops whose size is small in comparison with the length of the chain, their entropic contribution possesses the same functional form, but with a different form of the parameter  $c$  that characterizes the strength of this entropic contribution. This parameter turns out to be bigger than in the non-interacting case (signaling, as expected, a smaller entropic contribution from the denaturated loops) and this value is enough to render the phase transition of first order for  $d \geq 2$ .

**2.1. Partition function and phase transitions.** We recall here the basic formalism to study the model. A more detailed and mathematical approach has been carried out in [9], but we content here only with the essential results for the presentation of the model and the posterior extension to include interactions. We have already shown in the figure in the Introduction how the model considers a typical DNA configuration: sequences of bound monomers separated by denaturated loops. The statistical weight of a bound sequence of length  $k$  is  $\omega^k = \exp(-kE_0/T)$ , with  $T$  the temperature. Then, the statistical weight of a denaturated sequence of length  $k$  is given by the change in entropy due to the added configurations arising from a loop of length  $2k$ . For large  $k$  this has the form  $s^k/k^c$ , where  $s$  is a non-universal constant and the exponent  $c$  is determined by the properties of the loop configurations. The computation is done in the grand canonical ensemble, where the total chain length  $L$  is allowed to fluctuate. The grand canonical partition function,  $\mathcal{Z}$ , is given by [4, 5]:

$$(2.1) \quad \mathcal{Z} = \sum_{M=0}^{\infty} G(M)z^M = \frac{V_0(z)U_L(z)}{1 - U(z)V(z)},$$

where  $G(M)$  is the canonical partition function of a chain of length  $M$ ,  $z$  is the fugacity and

$$(2.2) \quad U(z) = \sum_{k=1}^{\infty} \frac{s^k}{k^c} z^k, \quad V(z) = \sum_{k=1}^{\infty} \omega^k z^k,$$

$V_0(z) = 1 + V(z)$  and  $U_L(z) = 1 + U(z)$ . The factors  $V_0(z)$  and  $U_L(z)$  properly account for the boundaries, but we shall not discuss their contribution here (see [5], for example). The average chain length,  $\langle L \rangle$ , is set by choosing a fugacity such that  $\langle L \rangle = \partial \ln \mathcal{Z} / \partial \ln z$ .

The quantity of interest is the fraction of bounded pairs, denoted by  $\theta$ , which is the order parameter. The average number of bounded pairs in a chain is given by  $\langle m \rangle = \partial \ln \mathcal{Z} / \partial \ln \omega$ , therefore we have:

$$(2.3) \quad \theta = \lim_{L \rightarrow \infty} \frac{\langle m \rangle}{\langle L \rangle} = \frac{\partial \ln z^*}{\partial \ln \omega}.$$

Here  $z^*$  is the value of the fugacity in the limit  $\langle L \rangle \rightarrow \infty$ . This is the lowest value of the fugacity for which the partition function (2.1) diverges. That is,  $U(z^*)V(z^*) =$

1. Using  $V(z) = \omega z / (1 - \omega z)$  one has

$$(2.4) \quad U(z^*) = 1/(\omega z^*) - 1.$$

Then, the nature of the denaturation transition is characterized by the dependence of  $z^*$  on  $\omega$ . The function  $U(z)$  is independent of  $\omega$ . It is finite for  $z < 1/s$  and diverges when  $z > 1/s$ . On the other hand the function  $1/V(z)$  increases continuously as  $\omega$  decreases -corresponding to an increase in  $T$ -. Thus, as the temperature increases from  $T = 0$ , the fraction of attached monomers decreases and  $z^*$  increases. However if  $z^*$  reaches the value  $z_c = 1/s$  (so that  $1/V(1/s) \geq U(1/s)$ ) any further increase of the temperature leaves  $z^*$  unchanged so that  $\theta = 0$ . Consequently, the transition occurs at  $z^* = z_c = 1/s$ . Its nature is determined by the behavior of  $U(z)$  around  $z_c$ , and this behavior is controlled by the value of the exponent  $c$ . The three possibilities are:

1. Divergence of  $U(z_c)$ , therefore  $z^*$  is an analytic function of  $\omega$  and no phase transition takes place. This happens when  $c \leq 1$ .
2. Convergence of  $U(z_c)$  while its derivative  $U'(z)$  diverges at  $z^* = z_c$ . Thus  $\theta \propto \partial z^* / \partial \omega$  vanishes continuously ( $\theta \sim |T - T_c|^{(2-c)/(c-1)}$ ) at the transition and the transition is *continuous*. This happens for  $1 < c \leq 2$ ,
3. Convergence of both  $U(z)$  and its derivative at  $z^* = z_c$ . Then the transition is *first order*. In contrast to the continuous case, the average size of a denaturated loop is finite at the transition. This happens for  $c > 2$ .

**2.2. Key result to implement volume-excluded interactions: polymer networks of fixed topology.** In [6], Duplantier considered a general self-avoiding network made of identical long polymer chains linked together. This type of networks can be easily characterized by the set of numbers  $\{n_L\}$ , which are the number of vertices connecting  $L$  chains, with the case  $L = 1$  representing external legs. In [6], the exact value of the critical exponents controlling the asymptotic number of configurations  $\omega_G$  of a generic network above described was given. An example of such a network can be seen in the following figure:

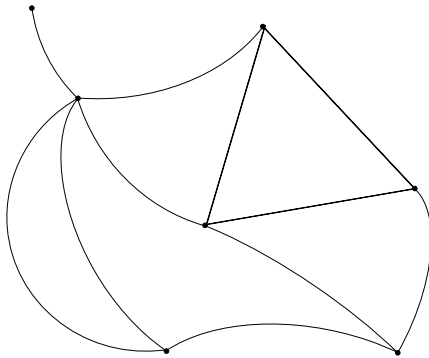


FIGURE 2. A network with 11 chains, 5 loops, and several types of vertices

The results of Duplantier were given in  $d = 4 - \varepsilon$  and in  $d = 2$  dimensions. As can be readily guessed, the first case was computed with renormalization group techniques, while the two-dimensional case exploits conformal invariance. In the

asymptotic limit  $N \rightarrow \infty$ ,  $s_i \rightarrow \infty$ , the total number of configurations for the network scales as [6]:

$$(2.5) \quad \Gamma \sim \mu^N N^{\gamma_G - 1} f\left(\frac{s_1}{N}, \frac{s_2}{N}, \dots\right)$$

where the factor  $\mu$  is the effective connectivity constant for the walks and  $f$  is a scaling function. the important exponent  $\gamma_G$  only depends on the topology of the network and is:

$$(2.6) \quad \gamma_G = 1 - \mathcal{L}d\nu + \sum_k \eta_k n_k$$

where  $\eta_k$  are exponents associated to a vertex with  $k$  outgoing legs.  $\mathcal{L}$  denotes the number of independent loops. Interestingly enough, the scaling dimensions  $\eta_k$  are known in several dimensions, as we mentioned above. In  $2d$  the result is exact thanks to conformal invariance and in  $d = 4 - \varepsilon$  is known up to order  $\varepsilon^2$ .

Why are this -fairly general- results, useful in the context of DNA denaturation ? As mentioned, the answer was provided in [4, 5], where they profitably employed these results, valid for very general network topologies, and applied it to a network geometry as the one depicted in the following figure:



FIGURE 3. Polymer network that represents the loop embedded in a chain

That is, a loop of total length  $2l$  connected to two polymers of lengths  $N/2$  each one. At this stage one can already guess which is the novelty provided in [4, 5]. By considering (2.5), the total number of configurations of the previous simple network can be computed. Thus, they are taking into account the interaction of a denaturated loop with the rest of the chain. However, it is manifest, and clearly stated in [4, 5], that this implements the interaction in a simplified way: the loop should interact with bound segments and other loops as well. Nevertheless, this full picture can not be implemented analytically but, in any case, as explained in [4, 5], such additional consideration of the interactions only makes the  $c$  bigger and already with the simple interaction described above, the value of  $c$  is big enough in order to make the transition of first order.

Following the above notation one has  $\mathcal{L} = 1$ ,  $n_1 = 2$  and  $n_3 = 2$  and thus the exponent is  $\gamma_{loop} = 1 - d\nu + 2\nu(\eta_1 + \eta_3)$ . Now, from (2.5) the total number of configurations scale as:

$$(2.7) \quad \Gamma \sim \mu^{N+l} (N+l)^{\gamma_{loop} - 1} f\left(\frac{l}{N}\right)$$

Now, in the limit  $l \ll N$  one should recover the partition function for a self-avoiding walk of total length  $N$  which has an entropic exponent  $\gamma = 1 + 2\nu\eta_1$  (the network then is just a line with the two endpoints as vertices). Since then the number of configurations -to leading order- is  $s^N L^{\gamma-1}$ . Then, this requires that:

$$(2.8) \quad f(x) \sim x^{\gamma_{loop} - \gamma} = x^{2\nu\eta_3 - d\nu}$$

in the limit  $x \rightarrow 0$ .

The excluded volume effect between these two parts is encoded in the value of the effective exponent  $c$  associated with the loop entropy. Thus, as we had already advanced, within the approximations presented, we can account for the interaction, keeping the usual formalism of the Poland-Scheraga models, just considering an effective value for the parameter  $c$  in the entropic contribution of the loops. From the previous equation the effective exponent  $c$  is:

$$(2.9a) \quad c = \gamma - \gamma_{\text{loop}} = d\nu - 2\sigma_3 .$$

We have just to consider the known values of  $\sigma_3$  and  $\nu$  in different dimension. From [6], the resulting parameter  $c$  can be computed. We summarize the results in the following table:

	$\sigma_3$	$\nu$	$c$
$d = 2$	$-29/64$	$3/4$	$2 + 13/32$
$d = 4 - \varepsilon$	$-3\varepsilon/16 + 9\varepsilon^2/512$	$1/2(1 + \varepsilon/8 + 15/4(\varepsilon/8)^2)$	$2 + \varepsilon/8 + 5\varepsilon^2/256$
$d = 3$	$-0.175$	$0.588$	$2.115$

Thus, since the exponent  $c$  is found to be larger than 2 in  $d = 2, 3$  and  $4 - \varepsilon$  then, the transition is first order in these dimensions. To conclude, note that to be able to express the number of configurations of the loop with the same functional but effective parameter as in the non-interacting case, it has been considered that the size of the loop is much smaller than the size of the chain. Interestingly, the resulting phase transition in the model turns out to be of first order and then the average size of denaturated loop is finite at the transition. This provides a necessary self-consistency to the method.

### 3. A MODEL WITH INTER-CHAIN EXCLUDED VOLUME INTERACTIONS

Here we present a rather different model, perhaps simpler, that implements volume exclude effects in a rather particular way but still also lead to first order phase transitions. More precisely, in [10] it has been argued that taking into account excluded-volume interactions *between* the two strands while neglecting these interactions within each strand effectively leads to a long range potential between the two strands [10]. This model, which takes into account the excluded volume interactions only partially also leads to a first order transition. Thus, the model is rather different from the previous one and this seems an interesting reason to present it but rather, it also will allow to make some comments regarding the various possible effects of excluded volume terms, depending on how they are implemented.

The partition function of the two chains of length  $N$  reads :

$$(3.1) \quad Z = \int \mathcal{D}\vec{r}_1(s) \mathcal{D}\vec{r}_2(s) e^{-\frac{d}{2a^2} \int_0^N ds \left( \left( \frac{d\vec{r}_1}{ds} \right)^2 + \left( \frac{d\vec{r}_2}{ds} \right)^2 \right) - \beta \int_0^N ds v(\vec{r}_1(s) - \vec{r}_2(s))} \\ \times e^{-g \int_0^N ds \int_0^N ds' \delta(\vec{r}_1(s) - \vec{r}_2(s'))} .$$

where  $d$  is the space dimension,  $a$  is the Kuhn length<sup>1</sup> of the monomers,  $\beta$  is the inverse temperature,  $g$  is the excluded volume parameter and  $v(\vec{r}_1 - \vec{r}_2)$  is the short range binding potential for monomers  $s$  of chain 1 and 2.

<sup>1</sup>The Kuhn length is a microscopic parameter, with the physical meaning of a monomer size, as mentioned. In field theoretic approaches to polymers, it provides a natural cutoff ultraviolet scale.

To be able to treat (3.1), several approximations were carried out in [10]. The main point was to approximate the excluded volume term by a long range interaction between the strands. As mentioned before, the transition is still of first order, as it happens with models with a more complete treatment of volume effects. Let us briefly sketch the approximations in [10], that lead to a simplification of (3.1), studying it as a (relatively) simple quantum mechanical model.

It is clear from equation (3.1), that one should focus on the relative coordinate ( $\vec{r}_1(s) - \vec{r}_2(s)$ ) of the two chain system to study a possible adsorption transition. Then, first step is to introduce the right coordinates:

$$(3.2) \quad \vec{R}(s) = \frac{\vec{r}_1(s) + \vec{r}_2(s)}{2}$$

$$(3.3) \quad \vec{r}(s) = \vec{r}_1(s) - \vec{r}_2(s)$$

with this change of variables in (3.1), we have

$$(3.4) \quad Z = \int \mathcal{D}\vec{r}(s) e^{-\frac{d}{4a^2} \int_0^N ds \left(\frac{d\vec{r}}{ds}\right)^2 - \beta \int_0^N ds v(\vec{r}(s)) - W(\{\vec{r}(s)\})}$$

where the interaction term  $W(\{\vec{r}(s)\})$  can be treated perturbatively -under the assumption that the two chains are far apart and consequently their interaction is very weak, assumption that is valid in the denaturated phase- and then reads:

$$(3.5) \quad W(\{\vec{r}(s)\}) \simeq g \int_0^N ds \int_0^N ds' \left( \frac{d}{\pi |s - s'| a^2} \right)^{d/2} e^{-\frac{d(\vec{r}(s) + \vec{r}(s'))^2}{4a^2 |s - s'|}}$$

valid for dimensions  $d > 2$ . This expression is asymptotically exact in the denaturated phase, where very few monomers of the two chains come close together.

Setting  $S = (s + s')/2$  and  $\sigma = s - s'$  in equation (3.5), we get for large  $N$

$$(3.6) \quad W(\{\vec{r}(s)\}) \simeq g \int_0^N dS \int_{-\infty}^{+\infty} d\sigma \left( \frac{d}{\pi |\sigma| a^2} \right)^{d/2} e^{-\frac{d(\vec{r}(S + \sigma/2) + \vec{r}(S - \sigma/2))^2}{4a^2 |\sigma|}}.$$

The range of  $|\sigma|$  which contributes to the integral is given by  $|\sigma|_{\text{typ}} \sim d(\vec{r}(S + \sigma/2) + \vec{r}(S - \sigma/2))^2/4a^2$ . Thus, for small separations where the interaction is sizeable, only small  $\sigma$  contribute significantly to the integral and we may expand  $\vec{r}(S + \sigma/2) + \vec{r}(S - \sigma/2)$  in powers of  $\sigma$ , and obtain

$$(3.7) \quad \begin{aligned} W(\{\vec{r}(s)\}) &\simeq g \int_0^N dS \int_{-\infty}^{+\infty} d\sigma \left( \frac{d}{\pi |\sigma| a^2} \right)^{d/2} e^{-\frac{d\vec{r}^2(S)}{a^2 |\sigma|}} \\ &\simeq \int_0^N ds \frac{\alpha_d}{|\vec{r}(s)|^{d-2}} \end{aligned}$$

where

$$(3.8) \quad \alpha_d = 2g \left( \frac{1}{\pi} \right)^{d/2} \Gamma\left(\frac{d}{2} - 1\right) \frac{d}{a^2}$$

Thus, the effect of the inter-chain excluded volume interaction is to generate a repulsive term, characterized by a singular potential, in the relative coordinate.

Now, for long chains, it is possible to employ the usual quantum mechanical methods [11] in polymer physics. One is then interested in the study of the ground state of the Hamiltonian:

$$(3.9) \quad H = -\frac{a^2}{d}\nabla^2 + \beta v(\vec{r}) + \frac{\alpha_d}{|\vec{r}|^{d-2}}$$

This ground state (wavefunction  $\phi_0$ , energy  $E_0$ ) is given by the Schrödinger equation  $H\phi_0 = E_0\phi_0$  and the free energy per monomer is  $f = E_0 T$ . Therefore, by exploring the ground-state wavefunction of the above quantum problems in the respective dimensions, the order of the unbinding transition (that occurs at the critical temperature and  $E_0 = 0$ ) can be found. This is easily done in [10], where from the (normalizable) form of the ground-state wavefunction is known that the phase transition is of first order. This follows from the fact that  $|\phi_0|^2$  represents the probability density of finding the strands a distance  $r$  from each other. Equivalently, the order parameter  $\theta$  is inversely proportional to  $\int_0^\infty |\phi_0(r)|^2 dr$ . Since the respective  $\phi_0$  turn out to be normalizable for  $d \geq 2$ , then the phase transition turns out to be of first order in these dimensions.

To conclude we note in passing that the study of singular potentials such the ones in [10] goes back to early days in quantum mechanics. In this sense, a very special case is . This model is especially subtle since it also posses a conformal symmetry. Note that, due to the concomittant symmetry the coupling constant in this case turns out to be dimensionless. This apparently innocent fact leads to many remarkable properties of the model. Of ourse, it may very well be that the subtle issues associated to this potential have no relevance in the denaturation transition in  $d = 4$  dimensions -in the context of this particular model, of course- but even in this negative case, it may be worth to explore a little bit further than in [10].

#### 4. CONCLUSIONS AND COMMENTS

As we have explained, DNA denaturation is an important process from the biological point of view, both for theoretical reasons and most notably in practical applications. Interestingly enough, attempts of a detailed understanding of this particular process have been carried out since many years ago. In particular, a statistical physics approach to the problems is not only feasible but has been seriously considered since the sixties. As we have seen, this interest has been considerably renewed in recent years. Of course, the experimental results guided the model-building process in a very clear way: a first order phase transition was expected, while the original statistical physics models, not taking into account volume-excluded effects (among many other simplifications of course), were predicting a continuous phase transition. In recent years, a successful account of volume excluded effects, employing results in the theory of polymer networks of fixed topology, have shown that indeed a phase transition of first order takes place. Nevertheless, the subject is tricky and this point is not without many subtle issues. For example, as explained in [7], it is possible to implement another type of restricted self-avoidance, different (and to a certain extent, complementary) to the one in the model of the last Section. Namely, to neglect the mutual avoidance between the strands but keeping the self-avoidance within the strand. This has been done in [8], and the result has been a continuous transition. Thus, it is noteworthy that, as pointed out in [7], to relax the self-avoidance in the two ways mentioned leads either to a sharpening of

the transition (the previous model) or a softer transition (situation above described and studied in [8]).

#### REFERENCES

- [1] D. Poland and H.A. Scheraga, *J. Chem. Phys.*, **45**, 1456; **45** 1464 (1966).
- [2] M. E. Fisher, Walks, walls, wetting, and melting, *J. Stat. Phys.* **34** (1984), 667–729.
- [3] M. E. Fisher, *J. Chem. Phys.* **45** (1966), 1469–1473..
- [4] Y. Kafri, D. Mukamel and L. Peliti, *Phys. Rev. Lett.*, **85**, 4988 (2000)
- [5] Y. Kafri, D. Mukamel and L. Peliti, *Eur. Phys. Jour. B* **27**, 135 (2002).
- [6] B. Duplantier, *Phys. Rev. Lett.* **57**, 941 (1986); *J. Stat. Phys.* **54**, 581 (1989).
- [7] M. Baiesi, E. Carlon, E. Orlandini and A.L. Stella, *Eur. Phys. J. B* **29**, 129 (2002).
- [8] E. Carlon, E. Orlandini and A.L. Stella, *Phys. Rev. Lett.*, **88**, 198101 (2002)
- [9] C. Richard and A.J. Guttmann, arXiv:cond-mat/0302514.
- [10] T. Garel, C. Monthus and H. Orland, *Europhys. Lett.*, **55**, 132 (2001).
- [11] P.G.de Gennes, *Scaling concepts in polymer physics* , Cornell University Press, Ithaca, (1979).

## 5. APPENDIX: SOME ADDITIONAL BIOLOGICAL INFORMATION

Here we provide a bit more of (elementary) information regarding DNA denaturation and some related processes and its applications. As explained, if we heat up a tube of DNA dissolved in water, the energy of the heat can pull the two strands of DNA apart. This the *denaturation* process that we have attempted to describe employing tools of statistical physics. The two strands still have the same nucleotide sequences and therefore they are still complementary.. Now, if we cool the tube again, then in the course of the normal, random molecular motion they will eventually bump into each other ... and stick tightly, reforming double-stranded DNA. This process is named *annealing* or *hybridization*, and it is very specific; only complementary strands will come together if it is done in the right way. This process is used in many crime labs to identify specific strands of DNA in a mixture. Then, as soon as the DNA is denaturated in two strands, it is possible to replicate the DNA. Why ? because any single-stranded piece of DNA can only hybridize with another if their sequences are complementary. If we have just one strand, we can actually build another strand to match it.

But how is this achieved from a practical point of view ? It can be done either in a test tube or in a live cell: The DNA strands are separated (for example, by heating them in a test tube). For each strand, we provide a primer, which is a short piece of DNA that sticks to one end of the strand. Then we add an enzyme. This is a specific type of protein called a "DNA polymerase" that can "read" the bases on one strand and can attach the complementary base to the growing strand. The polymerase "walks" down the template strand and creates its exact complement as it goes. The same thing happens to the other original strand. Thus, when we started, we had one double-stranded piece of DNA. After polymerase is done, we end up with two identical pieces - exact copies of each other.

INSTITUT D'ESTUDIS ESPACIALS DE CATALUNYA (IEEC/CSIC)., INSTITUTO DE CIENCIAS DEL ESPACIO (ICE/CSIC). EDIFICI NEXUS, GRAN CAPITÀ, 2-4, 08034 BARCELONA, SPAIN.

*E-mail address:* tierz@ieec.fcr.es