UNIVERSITÀ DEGLI STUDI
DI MILANO

Doctorate School in Physics, Astrophysics and Applied Physics

LIQUID CRYSTALLINE PHASES IN OLIGONUCLEOTIDE SOLUTIONS

FIS/03

Doctorate Thesis of:
Giuliano ZANCHETTA
XX cycle

Advisor:	**Prof. Tommaso BELLINI**
Co-advisor:	**Dr. Guido TIANA**
Director of the school:	**Prof. Gianpaolo BELLINI**

Academic year 2006-2007

to Michi, artist of science,

and to my women, Betta and Teresa

"Le cose tutte quante
hanno ordine tra loro, e questo è forma
che l'universo a Dio fa simigliante."

(Dante, Paradiso, I, 103)

[All things whate'er they be
Have order among themselves, and this is form,
That makes the universe resemble God]

"Il pensiero più risoluto, più ostinatamente perseguito,
é niente in confronto a ciò che avviene."

(Cesare Pavese (1908-1950))

[The most determined, the most stubbornly pursued thought
is nothing in comparison with what happens]

Abstract

In this thesis, we investigate the behavior of short complementary B-form DNA oligomers, 6 to 20 base pairs in length, exhibiting chiral nematic and columnar liquid crystal phases, even though such duplexes lack the shape anisotropy required for liquid crystal ordering. Structural characterization reveals that these phases are produced by the end-to-end stacking of the duplex oligomers into polydisperse anisotropic rod-shaped aggregates, which can order into liquid crystals. By use of polarized optical microscopy, X-ray micro-diffraction and optical interferometry, we determine the phase diagram of DNA oligomers and we estimate the stacking energy to be 4-6 K_BT. We also find that upon cooling mixed solutions of short DNA oligomers, in which only a small fraction of the present DNA is complementary, the duplex-forming oligomers phase-separate into liquid crystal droplets, leaving the unpaired single strands in isotropic solution. This spontaneous partitioning is the combined result of the free energy gain from the end-to-end stacking and LC ordering of duplexes and of depletion-type interactions favoring the segregation of the more rigid duplexes from the flexible single strands. In a chemical environment where oligomer ligation is possible, such ordering and condensation would provide an autocatalytic link whereby complementarity promotes the extended polymerization of complementary oligomers. The possible relevance of these observations for prebiotic synthesis of nucleic acids is discussed.

The work described in this thesis is partially published in:

1. M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, N. A. Clark, *End-to-End Stacking and Liquid Crystal Condensation of 6 to 20-Base Pair DNA Duplexes*, Science, **318**,1276 (2007);

2. G. Zanchetta, M. Nakata, M. Buscaglia, T. Bellini, N. A. Clark, *Phase Separation and Liquid Crystallization of Complementary Sequences in Mixtures of Nanosized DNA*, PNAS, *accepted*.

Contents

Introduction

One of the most challenging and fascinating problems in science is to encompass the wonderful variety and complexity of biological world in few simple and general laws. To this aim, it's certainly necessary to refer to notions from statistical physics and soft matter physics, whose application to biological mechanisms provided remarkable successes. Examples range from protein folding [1], to virus formation [2], to cellular organization, whose complexity can be reduced to the relative simplicity of self-assembled structures, driven by packing constraints and enforced by van der Waals, hydrophobic, and electrostatic interactions [3]. An exciting challenge is thus to search for the intrinsic self-assembly properties driving the spontaneous ordering of biological macromolecules , also in the frame of their evolution from prebiotic chemistry.

Aggregation and entropic effects are the two principal physical mechanisms regulating the self-organization in biosystems. Reversible or irreversible aggregates with elongated shapes are a distinctive property of a variety of molecules, like the stacks of polyaromatic dyes and drug molecules, stiff multistranded filaments of cytoskeleton proteins, amyloid fibers. Perhaps the most common self-assembly occurs in flexible surfactant solutions, with micelles growing in one dimension to form rod-like particles or two dimensions to form plate-like particles, depending on conditions. In such systems, 1D aggregation is often followed by orientational order.

In 1949, Onsager theoretically showed that the hard-core excluded volume interaction (crowding) alone suffices to induce orientational alignment of elongated particles in solution, other anisotropic interactions being not necessary [4]; in this way, the onset of nematic liquid crystalline (LC) phase can be explained. Since then, entropy often emerged as a source of order. In this vein, by combining colloidal hard rods and hard spheres, a phase diagram is obtained that is much richer than those occurring in separate systems of spheres and rods [5], with a variety of layered and columnar arrangements. The loss in orientational, positional and mixing entropy associated with the bulk and microphase separations are more than compensated for by the gain in entropy associated with the increased free volume in the ordered structures that arises as a consequence of these phase transitions.

Orientational order, the spontaneous alignment of elongated particles along a common axis; positional order, the spontaneous arrangement of particles in layers, columns or lattices; and demixing, the spontaneous segregation of different types of particles in different spatial domains: all of these mechanisms, well characterized in colloids, liquid crystals and polymers, appear to dominate in many biological systems too. In this respect, DNA is a particularly interesting example. Its packing *in vivo* is very tight, with high volume fractions in bacteria and in virus capsids, and it displays an astonishing hierarchical assembly into chromosomes, compacting meters of double helix down to a few micrometers. The packaging mechanisms adopted are strongly related to its natural tendency, as a long, semi-rigid polymer, to form liquid crystalline phases in concentrated solutions, as observed *in vitro* [6].

This thesis focuses on the discovery of liquid crystalline behavior in *small* DNA duplexes, 6 to 20 bp in length, that was never observed before and was unexpected because such short duplexes lack the shape anisotropy required by Onsager principle.

The origin of this research on the LC phases of DNA was the intention to use nucleic acids as a toolbox to design molecules with tunable length and rigidity. In fact, by exploiting the well known pairing mechanism of DNA and the relevant difference in flexibility between paired and unpaired sequences, molecules can be designed with various size and mechanical properties. Along this line, my supervisor, prof. Tommaso Bellini, and prof. Noel Clark (University of Colorado at Boulder) decided to first study two molecules: CGCAATTGCG, which is self-complementary, i.e. can hybridize with another strand of the same sequence and thus form a rigid helix, and CGCAATTGCGTTTTTTTTTT, only partially self-complementary, forming a helix with two flexible tails. This latter structure, a rigid core with flexible ends, bearing resemblance with usual liquid crystal molecules, was thought to *favor* the formation of liquid-crystalline phases, whereas the rigid helix without tails was the negative control. What happened was the opposite of what expected: the oligonucleotide with tails showed no phases, while the control oligo displayed wonderfully colored textures when observed in a polarized microscope, just as long DNA.
Pretty soon it became clear that understanding the very origin of LC phases in small DNA duplexes was more important than pursuing the design of DNA-based model molecules. We then characterized these phases and worked on different sequences, finding that stacking between neighboring nitrogen bases - which determines the helical shape of DNA - promotes end-to-end aggregation of helices and the formation of longer chains.
We also got interested in the effects of added impurities on LC phases, and in particular in the role of the denaturation of the helix itself, because it was observed that when enough helices were broken (and thus *ss-DNA* coils appeared) the phases vanished too. At this point, the second "mistake" came out: to obtain a system intrinsically composed of helices and coils, we mixed mutually complementary sequences, CCTCAAAACTCC and GGAGTTTTGAGG, in unbalanced ratios. At some *ss-DNA* concentration, LC phases were expected to get disturbed and disappear. Surprisingly, they didn't: a phase separation occurred, instead, between the helices and the pool of coils, with *ds-DNA* organizing in LC droplets. A delicate interplay between osmotic pressure, end-to-end

stacking and LC ordering emerged, with interesting implications on the mechanism of DNA self-assembly.
The experience of this research project is the confirmation to us that reality is surprising and stubborn and, most of times, is transparent to seeking eyes.

The chapters in the thesis are organized as follows:
CHAPTER 1 contains an introduction to the structure of DNA double helix and its properties, in particular stacking and denaturation.
CHAPTER 2 introduces various types of liquid crystals and some of the main theories describing their ordering; it also describes the so-called "living polymers", chains formed by reversible aggregation, a model system that we later use to interpret the sDNA observed behavior.
CHAPTER 3 presents materials, procedures and the experimental techniques used to characterize the samples; in particular, it is focused on the description of LC textures in polarized microscopy and on optical measurements of concentration.
After summarizing previous studies on long DNA liquid crystals, our results on sDNA ordering and its phase diagram are presented in **CHAPTER 4**.
CHAPTER 5 discusses the phase separation between rigid *ds-DNA* helices and flexible single *ss-DNA* coils, its kinetic analysis and the underlying mechanisms.
Finally, ongoing experiments on short RNA, mixtures of different helices and random sequences are presented and discussed in the frame of prebiotic scenarios in **CHAPTER 6**.

Bibliography

[1] J. R. Banavar, T. X. Hoang, A. Maritan, F. Seno, A. Trovato, *Unified perspective on proteins: A physics approach*, Phys. Rev. E **70**,041905 (2004)

[2] R. Zandi, D. Reguera, R. F. Bruinsma, W. M. Gelbart, J. Rudnick, *Origin of icosahedral symmetry in viruses*, Proc. Natl. Acad. Sci. USA **101**,15556 (2004)

[3] J. Herzfeld, *Entropically Driven Order in Crowded Solutions: From Liquid Crystals to Cell Biology*, Acc. Chem. Res. **29**,31 (1996)

[4] L. Onsager, *The effects of shape on the interaction of colloidal particles*, Ann. NY Acad. Sci. **51**,627 (1949)

[5] M. Adams, Z. Dogic, S. L. Keller, S. Fraden, *Entropically driven microphase transitions in mixtures of colloidal rods and spheres*, Nature **393**,349 (1998)

[6] F. Livolant, *Ordered phases of DNA in vivo and in vitro*, Physica A **176**,117 (1991)

CHAPTER 1

DNA

The main character of this story is Deoxyribonucleic Acid, or DNA. In living cells it is found in the form of a very narrow helical thread of diameter 2 nm, although the DNA from a single human cell has a total length of 2 m and it is compacted in the chromosome coils whose size is of a few micrometers. It is indeed in DNA's nature to stay in a tightly packed conformation. The main features of this extraordinary molecule will be briefly presented, focusing on their structural properties that are, however, strictly related to biological functions. In particular, we'll describe the fundamental role of base pairing and base stacking. In the end, evidences for the formation of mesophases by DNA *in vivo* will be reported.

1.1 DNA structure

By looking at DNA through an optical microscope, we can see chromosomes, that are 10 micrometers long structures inside the cell nucleus in which DNA is tightly packed. The hierarchical organization of chromosomes shown in figure 1.1 consists of a coiled mixture of DNA and proteins, the histones, that make "spools" (10 nm across), and the DNA wraps twice around each spool, into a series of double loops, This strategy of packing reduces the DNA length by a factor of 10^6.
When we remove the protein spools, we are left with the long, polymeric DNA molecule. It consists of two strands which coil around each other to make a double helix (from here on indicated with double stranded DNA, or *ds-DNA*). The sense of wrapping of these two strands is right-handed. The figure 1.2 reproduces the drawing in the famous 1953 paper of Watson and Crick [1].
If we uncoil the two strands, each of them may be seen to consist of a series of units called nucleotides; these are linked to one another with a certain directionality in a head-to-tail sense ($5'$-$3'$, as it is conventionally indicated from the C atoms' positions). The two strands run in opposite directions. Each nucleotide

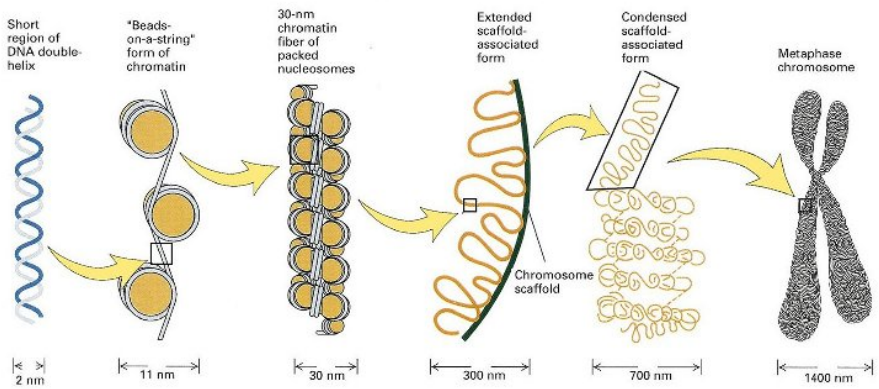

Figure 1.1 - Hierarchical organization of the DNA helix inside the chromosomes.

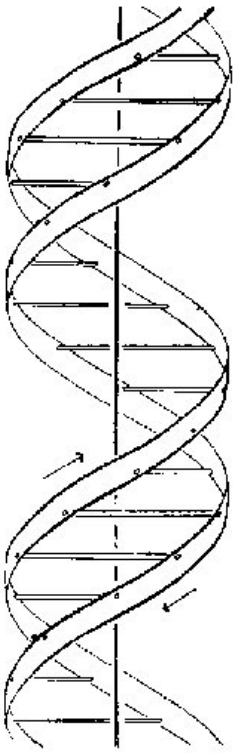

Figure 1.2 - Double helix of DNA, from [1].

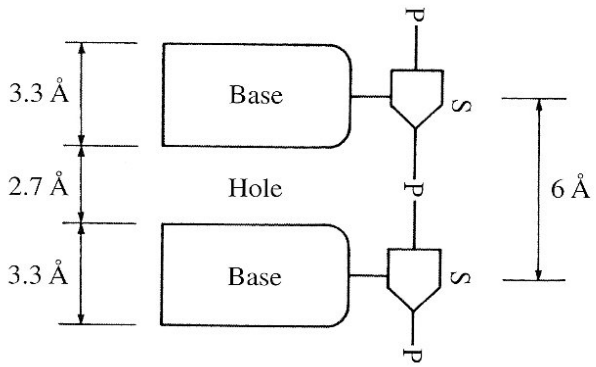

Figure 1.3 - Two nucleotides in schematic form, showing the key dimensions; from [2].

is made of about 20 atoms, grouped in the sugar (S), the phosphate (P) and the base.

1.1.1 Why an helix?

The phosphates are very soluble in water. Sugars are also soluble. What about the bases?
In figure 1.4 the four different types of DNA bases are shown: A is Adenine, T is Thymine, C is Cytosine and G is Guanine (atoms that are not labeled are carbons). A and G are called "purines" and they are bigger than T and C, which are called "pyrimidines".
All bases are strongly hydrophobic flat molecules; they become soluble in water once they are attached to a sugar and a phosphate to form a nucleotide, but their hydrophobicity places strong constraints on the overall conformation of a large DNA molecule in solution. For such a molecule to be stable in water at neutral pH, the bases will have to hide into the very center of some folded structure, so as to avoid water; while the sugars and phosphates, soluble in water, will stay on the outside. This is what happens, as we know, in the helical arrangement.
The distance between adjacent sugars or phosphates in the DNA chain is 6 Å in the usual case, and it cannot become much longer than 6.5 or shorter than 5.5 Å . The thickness of the flat part of a base is 3.3 Å , and this distance cannot change much either, because the bases are chemically rigid. This leaves us with a "hole" of 2.7 Å (figure 1.3), which some greasy object, not water, would have to fill. In brief, the bases are attached to a sugar-phosphate chain that is twice as long as the thickness of the bases themselves.
The sugar-phosphate chain is flexible, although in an indirect way. Any single phosphate group is essentially a rigid tetrahedron, having a P atom at its center and one O at each vertex. Only when we go further along the chain from the P and attach two C atoms to two of the O, can these C atoms swivel about the line of the P-O bond. All parts of the DNA sugar-phosphate chain are rigid locally, but they have this kind of indirect rotational flexibility over several bonds.

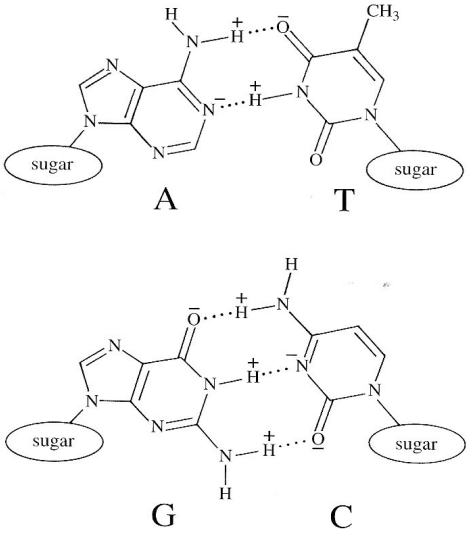

Figure 1.4 - Chemical structure of the four DNA bases, showing the Watson-Crick pairing; adapted from [2].

Given all of these constraints, how to hide the bases into the center of a DNA molecule and at the same time get rid of the "holes"?
The helix, a strongly twisted ladder, comes out to be the only shape that fulfills all the topological requirements (and at the same time avoids unacceptable contacts between neighboring atoms). In fig. 1.5 this arrangement is shown, with only one strand for sake of simplicity. We can extract that the angle θ, corresponding to the twist of one base relative to the previous one along the helix, is 32.3°, or equivalently that after $360°/32.3° \simeq 11$ phosphates a complete turn is done. Actually, although our estimate is very rough (and these slight differences are of great importance in biology), all *ds-DNA* helices have between 10 and 12 phosphates per turn, within each strand: the A-form of DNA (figure 1.6) has 11 phosphates per turn, while B has 10 and Z has 12. A and B-forms are right-handed, while Z is left-handed. To explain these differences it's necessary to build accurate space-filling models of DNA chain, and take into account the single atoms' interactions. Anyway, almost all the work done in this thesis deals with B-DNA, so we won't go in further details on it.

Pairing

The two chains of a *ds-DNA* run always in opposite, or anti-parallel, directions, because parallel double helices are not stable.
This fact results mainly from the pairing of the bases. Watson and Crick in 1953 proposed the helical structure for DNA, and they also put forward a set of rules for base-pairing: the most stable base pairs are A-T and G-C (see figure 1.4); since all four possible Watson-Crick base pairs, A-T, T-A, C-G, G-C, are of the same size, they can fit easily into the framework of a regular double helix.

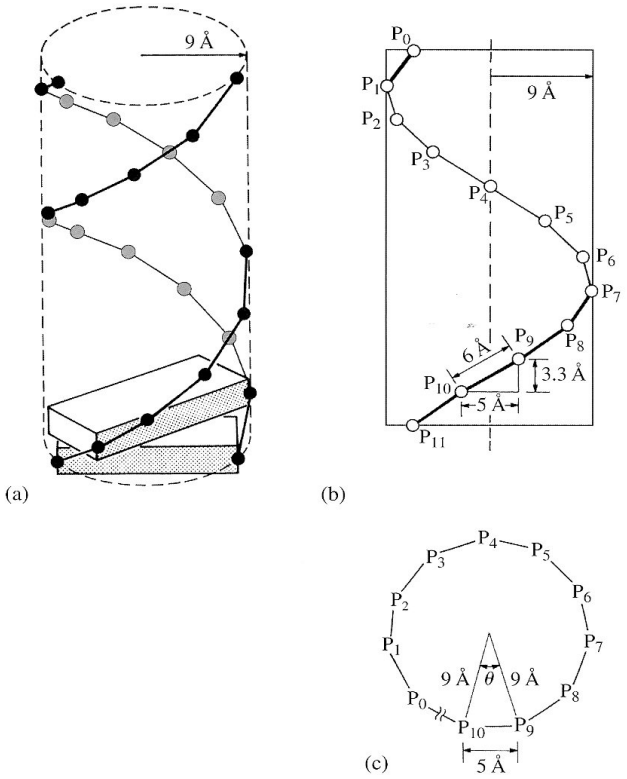

Figure 1.5 - Three views of the sugar-phosphate chain wrapped helically around a cylinder: (a) sugars are represented as circles and phosphates as thin lines; (b) Phosphates are circles and sugars thin lines; (c) top view. From [2].

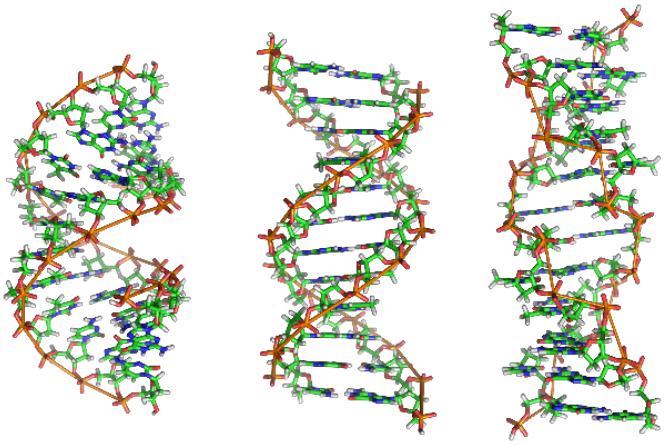

Figure 1.6 - The three main forms of DNA: A (left) and B (center) are right-handed with 10 and 11 phosphates per turn, respectively, while Z is left-handed with 12 phosphates.

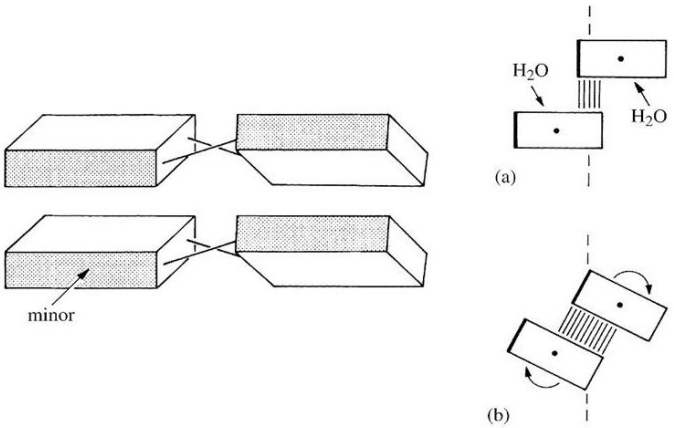

Figure 1.7 - Propeller twist (b) allows a greater overlap of bases compared to undistorted arrangement (a); adapted from [2].

In addition, it explains how the genes in DNA are duplicated (and then stably inherited) on cell division. Whenever a cell divides, and needs to duplicate its DNA, it can do so simply by splitting the DNA into two separate strands. Certain enzymes come along and use each of these strands as a "template" for the precise synthesis of a new strand, according to the Watson-Crick rules: A with T and G with C. These pairings are based on the simple fact that, within any DNA base, due to the shape of orbitals there is a small surplus of negative electric charge on N and O atoms not linked to H, while there is a small surplus of positive charge on these same atoms where they are attached to H. Thus, consider the base pair of A and T as shown in figure 1.4. Not counting the two N that are attached to sugars, there are three N on A and two O on T that have a surplus of negative charge. On the other hand, one N on A and one N on T have a surplus positive charge. So all we have to do is to put the pluses and minuses together, thereby making hydrogen bonds that are indicated as dotted lines. It's easy to see that C-G pair, because it has three H-bonds, is more stable than A-T.
Why can't there be other stable base pairs, such as G-A or C-T? Some of them are ruled out by the difficulty of making two or more H-bonds. But others, such as G-T or G-U (where U stands for Uracil, see par. 1.1.2), although less stable, are not excluded. The H-bonding produces a pair with an overall shape similar to those in figure 1.4 and stable enough. In fact, in particular G-U pair is used during recognition between messenger-RNA and transfer-RNA.
Another possible set of rules for pairing, that relies on H-bonds between different atoms in the bases, is the Hoogsteen mode [2]. It is much less stable than Watson-Crick pairing and has minor biological significance; it can be found, however, in G-quartets (see section 2.3.1).

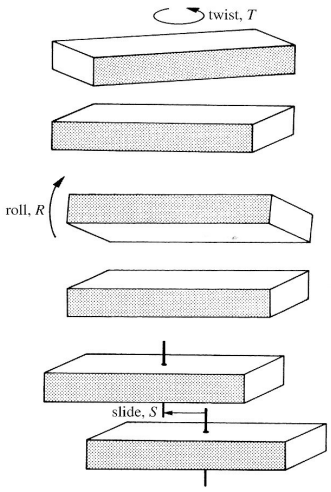

Figure 1.8 - Twist, roll and slide motions on a base step; from [2].

Higher orders

This very elementary level of description enables understanding the most basic features of the overall shape of DNA.
At a higher detail, bases, while being parallel to each other, are not necessarily normal to the helix axis, and that the twist angle ranges from 20° to 50° about a mean of 34°, instead of the value of about 32° predicted from our first-order considerations.
One of the reasons is the fact that, in order to maximize the surface hidden from water, the bases can rotate away from the plane normal to helix axis, as in figure 1.7. This motion is called "propeller twist" because the same happens to the paired bases, but it the opposite sense of rotation, and thus a propeller-like shape is assumed.
We can describe the displacement of a base-pair,now assumed as a unique flat block, relative to the neighbors, with three out of the six Euler degrees of freedom (the other three being fixed by chain constraints): twist (T), the rotation about the local twist axis; roll (R), a rotation due to the rolling open of base pairs along their long axes, which can vary from +20° to -10°; slide (S), a translation that describes the relative sliding of base pairs along their long axes, typically from +3 Å to -2 Å . These motions are showed in figure 1.8; they are not independent from each other, and their values are the result of the delicate balance among all the different favorable and unfavorable interactions and constraints. Their geometrical relationships can be found in [2].

The overall effect of these "adjustments" can be seen in figure 1.9, which represents real structures of DNA sequences, characterized with the parameters now introduced. The first on the left is close to the B-DNA structure, and the last one to the A-form: they are the most stable arrangements that DNA nucleotides can reach, depending on the sequence and on the environmental conditions (hydration, pH, ionic strength).

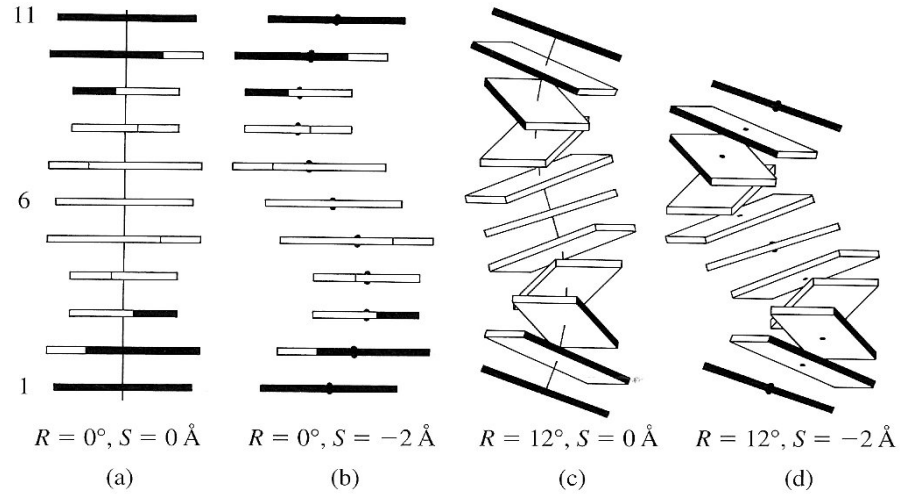

Figure 1.9 - The effects of uniform roll R or slide S on a 11 bp helix having T=36°; (a) corresponds to B-form, while (d) to A. Image was taken from [2].

One of the consequences of the different values of R and S is the change in the length of the helix. While in B-DNA the helical pitch is about 3.4 nm, corresponding to a rise (shift in the position along the helix axis) of 3.4 Å /bp, in A-DNA it is instead 2.8 nm, corresponding to a rise of $\sim$ 2.4 Å /bp.
It's worth to point out that the vertical separation between consecutive base pairs in the DNA molecule is always optimized to be the same, because any vertical compression or extension of the largely flat base pairs is associated with substantial energy penalty (see par. 1.3.2); the effective physical distance between adjacent base pairs is thus always around 3.3-3.4 Å , irrespective of the DNA form, i.e. of the inclination of the base pairs relative to the helix axis.

1.1.2 RNA

Rybonucleic Acid, RNA, is strictly related to DNA, has the same base-sugar-phosphate chain structure and is most probably older than DNA (see chapter 6).
The main difference with DNA, apart the biological function that we won't discuss here, is that the sugar of the chain is Ribose instead of Deoxyribose, having two hydroxyl groups. These hydroxyl groups make RNA less stable than DNA because it is more prone to hydrolysis.
In addition, in RNA Uracil (U) base substitutes Thymine, differing from it only in the absence of the CH_3 or methyl group.
These differences have implications on the structure: while the hydrated DNA has a B-form, RNA usually adopts a structure very close to A-DNA, with a shorter rise per base pair (2.4 Å) and higher tilt of the bases.

1.2 Duplex melting

When double-stranded native DNA is heated, the bonding energies between the strands (mainly due to H-bonds pairing) and within the strands (mainly due to stacking) can be overcome by the thermal energy, and the two strands can separate. This process, called melting or denaturation, involves breaking of three hydrogen bonds in every G-C base pair and breaking of two hydrogen bonds in A-T base pair.

1.2.1 Measuring denaturation

DNA duplex melting and formation can be easily monitored through the absorbance in the UV region, that increases as the double helix "melts" to two single strands. This effect ranges in magnitude from 20% and 40% of the low-temperature value depending on duplex length and is caused by the structural changes in the strands that result from formation of Watson-Crick base pairs and associated interactions between neighboring bases as stacking interactions. When duplex DNA unwinds, in the simplest situation (all-or-none transition) it is assumed that the changes in the extinction coefficient are proportional to the extent of denaturation. In other words, the change in absorbance at around 260 nm is directly proportional to the fraction of broken base pairs, θ_B [3]. Therefore, when θ_B increases from 0 - intact duplex - to 1 - completely melted, separated single-strands - it is presumed that the total absorbance increases proportionally (we couldn't use this method because the absorbance signal saturates at the high DNA concentration we're interested in).
Analogously, some flat dye molecules such as Ethidium Bromide [4] are known to well intercalate between two different stacked base pairs, and their fluorescence efficiency undergoes a strong enhancement when they are in the hydrophobic environment of the bases.
Thus, a decrease in the fluorescence signal of these dyes is interpreted to proportionally depend on the increasing θ_B; we will use this second probing method in chapter 4. However, the following description is valid for every measured signal proportional to θ_B.
The graph of absorbance (or fluorescence) vs. temperature, the melting curve, is analyzed to obtain melting temperature T_m, standard Gibbs free energy ΔG, standard enthalpy ΔH and standard entropy ΔS of the helix-coil transition. The melting temperature is usually defined as the temperature when half of molecules are "melted", that is the temperature at the midpoint of the double helix to single strands transition.
DNA stability is found to depend on the chain length and the percentage of G-C pairs, due to the higher number of H-bonds compared to A-T pairs.
Actually, thermodynamic stability depends not only on G-C content but also on the sequence order or context. To consider these effects, the nearest-neighbor (NN) model has been developed (see section 1.2.3).

Other techniques are also employed. The DNA denaturation process involves absorption of heat, which is necessary for breaking of hydrogen-bonding and stacking interactions, and therefore can be measured directly with calorimetry, in particular differential scanning calorimetry (DSC). This technique has the

advantage of providing direct, model-independent measurements of thermodynamic functions. DSC is not contingent on assumptions about the nature of melting transition. Therefore, calorimetric measurements can be used to test the accuracy of the two-state assumption of DNA melting. Primary data is collected as the change of excess heat capacity (ΔC_P) vs. temperature. From these curves, the enthalpy of the melting transition is found from the integrated area under the curve. Similarly, the transition entropy is equal to the area under the plot of ΔC_P vs. T. The other methods are indirect and require some assumptions about the number of different states or intermediates during the DNA melting process in order to deduce the enthalpy, entropy, free energy of the transition. The disadvantage of the calorimetric technique is that it usually requires substantially larger amounts of DNA samples than optical spectroscopic measurements.

1.2.2 Modeling denaturation

Generally, two distinct models are used to study DNA melting curves from spectroscopic measurements.
To obtain the free energy of DNA denaturation, the resulting melting curve is analyzed either in terms of a concerted two-state model or in terms of a multi-state zipper model. The two-state, also called all-or-none model assumes that DNA can reside solely in two states. Only the intact duplex and completely melted single strands are assumed to be significantly present during the melting transition. The multi-state model is based on the statistical mechanical description for duplex states with varying degrees of hydrogen-bonded base pairs. It takes into account also partially melted duplexes that can be significantly populated during melting transition. For both models, relevant thermodynamic equations required to analyze the melting profiles depend on the molecularity of melting reaction, i.e. on the number of colliding molecular entities that are involved in the reaction.
Considered the small length of the strands described in this thesis (6-20 bp), we can safely assume a two-state behavior and we will focus on it.

Both for absorbance and fluorescence measurements, in the linear regime, from collected raw curves vs. temperature T, the total fraction of melted (broken) base pairs, θ_B, is determined using the formula

$$\theta_B = \frac{S(T) - S_L}{S_U - S_L} \tag{1.1}$$

where $S(T)$ is the signal of the experimental curve at a given temperature T and S_L and S_U respectively the lower and the upper baseline. S_U is obtained by fitting of measured signal after the melting transition. Similarly, S_L is obtained from the melting curve at low temperatures, before onset of the melting transition.

We performed melting experiments on self-complementary sequences only; below we thus derive the equations describing the dependence of melting temperature on DNA concentration, ionic strength and on thermodynamic parameters

only for this situation. For a more complete description of this topic, see [5].
As already said, the two-state model of linear duplexes assumes that intermediate states between the intact and fully melted duplexes are not significantly populated throughout the melting transition. Consider the reversible equilibrium annealing reaction of two single strands, $S_1 = S_2 = S$, to form a duplex, D

$$S + S \rightleftarrows D \tag{1.2}$$

with an equilibrium constant K_D, that we can write as $K_D = \beta K_{duplex}$ with K_{duplex} reflecting the contributions from the internal degrees of freedom, as hydrogen bonding and stacking, and with β for the external ones, related to initiation penalties (or equivalently to a concentration effect).
Assuming that the transition proceeds in two-state fashion, the fraction of melted base pairs is equal to fraction of melted duplexes and can be written as

$$\theta_B = \frac{[S]}{C_T} = 1 - \alpha \tag{1.3}$$

where C_T is the total strand concentration, $C_T = [S] + 2[D]$, and α is the fraction of intact duplexes, $\alpha = 2[D]/C_T$.
The equilibrium constant is defined as

$$K_D = \frac{[D]}{[S]^2} \tag{1.4}$$

and, if we substitute equation 1.3 and the expression for the total concentration, it can be written

$$K_D = \frac{\alpha}{2(1-\alpha)^2 C_t} \tag{1.5}$$

or making α explicit:

$$\alpha = \frac{1 + 4C_T K_D - \sqrt{1 + 8C_T K_D}}{4C_T K_D} \tag{1.6}$$

.

Given the usual relation between thermodynamic parameters and the equilibrium constant [6]

$$\Delta G = \Delta H - T\Delta S = -RT \ln K_D \tag{1.7}$$

we can divide both sides by T and differentiate with respect to 1/T, obtaining van't'Hoff equation

$$\Delta H = RT^2 \frac{d \ln K_D}{dT} \tag{1.8}$$

Because duplexes are formed from two strands, the molecularity of duplex formation is greater than one. Consequently, the melting equilibrium is concentration dependent. Thermodynamic parameters can be evaluated from the dependence of T_m, defined as the temperature at which $\alpha = 0.5$, on concentration. At $T = T_m$,

$$K_D(T_m) = \frac{1}{C_t} \tag{1.9}$$

and after some rearrangements the following expression can be obtained [5]:

$$T_m = \frac{\Delta H}{\Delta S + R \ln(C_T)} \tag{1.10}$$

where $\Delta H = \Delta H_{duplex} + \Delta H_{nuc}$ and $\Delta S = \Delta S_{duplex} + \Delta S_{nuc}$ are the thermodynamic parameters from the internal and external contributions.
The reciprocal of equation 1.10

$$\frac{1}{T_m} = \frac{R}{\Delta H} \ln(C_T) + \frac{\Delta S}{\Delta H} \tag{1.11}$$

can be used to extract the parameters from the plot of $1/T_m$ vs. $\ln(C_T)$, assumed that:

- the transition is really two-state (yielding a linear plot);
- ΔH and ΔS are temperature independent;
- there is zero change in heat capacity at constant pressure for the melting transition, $\Delta C_P = 0$.

Actually, we're interested in extracting T_m from fluorescence measurements (comparing the measured temperature with that predicted from online calculators), but we usually lack the information on concentration.

We note that the description of the binding of mutually complementary strands would be the same as above, with C_T replaced by $C_T/4$ [5].

1.2.3 Predicting denaturation

As already said, the stability of an oligonucleotide is related not only to the nucleotide composition, but to the sequence order too. The nearest-neighbors (NN) model predicts the thermodynamic quantities from the correlations holding between neighboring bases, expressed as one parameter for each doublet; it also assumes that there are no significant interactions beyond NN couples. In DNA duplex, there are 10 doublets (5′-AG-3′/5′-CT-3′, AA/TT, AT/AT, etc.). Consequently, the thermodynamics of double helix formation from single strands for any sequence can be calculated from 10 nearest-neighbor thermodynamic parameters (additional parameters for the initiation of duplex formation are necessary). The NN parameters are extracted by averaging over a number of short sequences whose melting temperature and thermodynamic parameters are experimentally measured; contribution of salt concentration is also empirically taken into account [7] (the same parameters are reported with some mistakes in [8]).
These provide the basis for the following equation:

$$\Delta G = \Delta G_{initiation} + \Delta G_{symmetry} + \sum \Delta G_{NN} - 0.114 \cdot (ph/2) \ln[Na^+] \tag{1.12}$$

where $\Delta G_{initiation}$ and $\Delta G_{symmetry}$ are the entropic penalties related to end effects and possible self-complementarity of the sequence; ph is the number of phosphates in the duplex. The equation is accurate around T = 37 $°C$, over a range of salt concentration between 0.05 and 1 M.

To give an example, the estimate for pairing (and stacking) of the self-complementary ATATATAT sequence is ΔG = -0.4 kcal/mol and for CGCGCGCG is ΔG = -10.6 kcal/mol (estimated at 37 °C and 50 mM of Na^+). These are respectively the least and most stable 8-mer sequences.
Analogous NN parameters are given for ΔH and ΔS, and for temperatures above or below 37 °C ΔG is obtained from equation 1.7. These parameters are used to calculate the melting temperature, through equation 1.10, or equivalently to extract information on thermodynamic equilibria.

1.3 Nothing but a charged rod...

A complete description of DNA (or RNA) involves the aspects that we sketched in the previous paragraphs, and many more as its functioning, transcription and replication. Just focusing on a structural point of view, we should discuss twisting, bending and super-coiling of the *ds-DNA* , important for example for the assembly into chromosomes.
However, because this thesis deals with the liquid-crystalline behavior of *short* (6-20 bp) DNA helices, for our purposes we can try to "zoom out" and give a coarse-grained description of a tract of *ds-DNA* below the melting temperature as a stiff cylinder with sticky ends and a helical charge pattern on its side.
This picture will be very useful in describing the ordering properties of DNA (and the main features apply to RNA too). In the following we're going to briefly discuss the terms "stiff", "sticky ends" and "helical charge pattern".

1.3.1 Flexibility

An important issue for the description of small fragments of DNA, and of mixtures of *ds-DNA* helices and *ss-DNA* coils, is to determine their flexibility.

ds-DNA

It is well known that the conformation of *ds-DNA* in solution ranges from that of a rigid rod to that of a worm-like chain (WLC), depending on molecular weight. The "stiffness" of the worm-like coil is usually described by its persistence length, L_p; L_p is related to the distance over which the direction of orientation of two successive segments of the molecule are correlated. Many different methods have been used to determine the persistence length of defined-sequence DNA molecules. These methods include light scattering, cryo-electron microscopy, scanning force microscopy, force-measuring laser tweezers, transient electric birefringence, electrophoresis and various biochemical methods. The consensus value of the persistence length obtained for relatively short, well-characterized DNA molecules is $\sim$ 50 nm, corresponding to about 150 bp, or above, depending on temperature, pH and ionic strength [9].
Curiously, the experiments conducted so far on the LC phases of short DNA fragments had just 146 bp as lower length limit, corresponding to the length of DNA around a histone. This fact was due to the relative ease to obtain them from enzymatic methods. Thus, such experiments were done on semi-flexible

polymers.
We can safely conclude that helices in 6-20 bp range are well described as stiff rods.

ss-DNA

As regards *ss-DNA* , its persistence persistence length is much lower than that of *ds-DNA* . As a polyelectrolyte, its conformations depend critically on pH, ionic strength etc., reflecting the contributions from the intrinsic rigidity of the chain and from the electrostatic repulsion. Indeed, experimental measures [10, 11] indicate that L_p can vary between 1 and 5 nm.
A measure of the degree of packing of a semi-flexible polymer is its gyration radius, R_g; it is expressed by the Kratky-Porod equation [12, 13]:

$$R_g^2 = \frac{L_p L_d}{3}\left[1 - 3\left(\frac{L_p}{L_d}\right) + 6\left(\frac{L_p}{L_d}\right)^2 - 6\left(\frac{L_p}{L_d}\right)^3 [1 - e^{-L_d/L_p}]\right] \tag{1.13}$$

where L_p is the persistence length and $L_d = Mb$ the contour length of the polymer, with M the number of bases and b the size of a monomer (the single nucleotide). For extreme values of L_p, the equation recovers the asymptotic limits of a random coil

$$R_g^2 \sim \frac{1}{3} L_p L_d \qquad L_d \gg L_p \tag{1.14}$$

and of a rod-like polymer

$$R_g^2 \sim \frac{1}{12} L_d^2 \qquad L_d \ll L_p \tag{1.15}$$

This behavior is shown in figure 1.10, with the "consensus" values of $b = 0.43$ nm and $L_p = 3$ nm; these values can be considered as an ambiguous choice, since the contour length could be set to the chemical length of 0.6 nm per base and the persistence length could be shorter, being strongly dependent on the density of ions. However, the crossover between the two behaviors is around M=20-30. Note that eq. 1.13 doesn't take into account the excluded volume effect. We used this expression to estimate the gyration radius of *ss-DNA* in chapter 5.

1.3.2 Base stacking

As we have already pointed out in the previous paragraphs, both the structure of DNA and its stability are mainly based on the hydrophobicity of the bases that shows up in their stacking; we're now going to treat the other contributions to this phenomenon, with strong consequences on the helix structure (the higher-order terms of section 1.1.1).
We're interested in all these effects because the stacking will be showed in chapter 4 to play an important role in the appearance of unexpected liquid-crystalline phases.
There is an open debate on what is the relative weight of the contributions to the aromatic stacking: aside hydrophobicity (whose ultimate origin is itself discussed), electrostatic forces and Van der Waals dispersive forces. It's usually

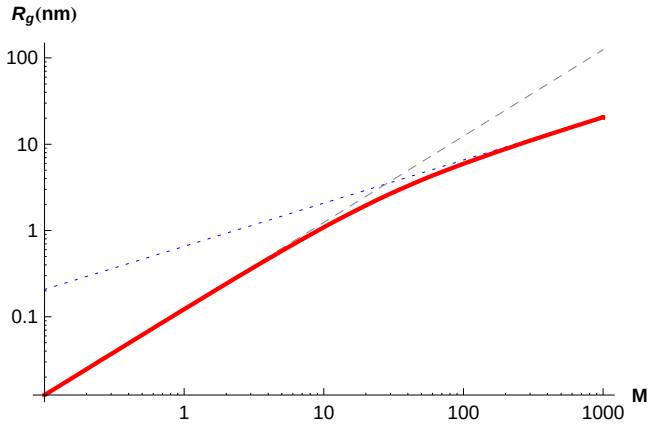

Figure 1.10 - Log-Log plot of Kratky-Porod equation (1.13) and its asymptotic behavior for small and large persistence lengths.

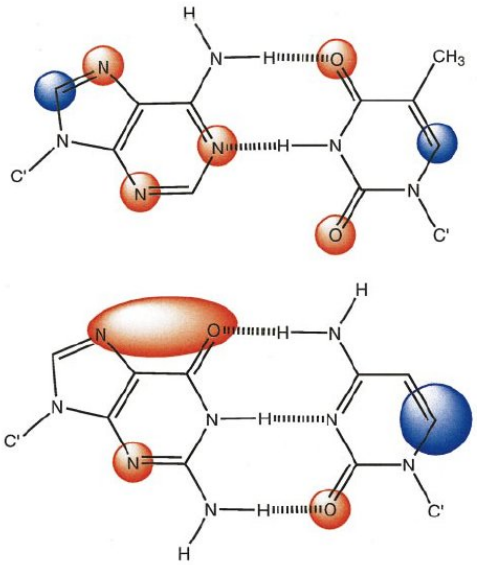

Figure 1.11 - Simplified representation of the electrostatic potential on the surfaces of the DNA base pairs; from [14].

very hard to distinguish among the various components in experimental measurements, and different models and approximations lead, of course, to a wide range of estimates.
The electrostatic contribution comes first from the fact that the upper and lower surfaces of bases have a slight negative charge, due to the electrons delocalized in the aromatic system, and thus a sandwich-like structure is obtained as regards charge. This contributes to the mutual slide of the bases to minimize the interaction surface, opposite to hydrophobic effect that is instead maximized by superposition. The relative importance of the two effects depends critically on the environment conditions, i.e. on the amount of water and the ionic strength.
In addition to the "layered" charge, a second electrostatic effect holds: the charge distribution on the plane of the rigid bases can influence their mutual orientation; thus, some otherwise favored positions are forbidden because of electrostatic repulsion. While C-G base pair has a pretty strong dipole, A-T has only small patches of isolated "partial" positive or negative charges (fig. 1.11); these localized charge patterns have a striking influence in the positioning of the flat molecules relative to one another [15], maximizing attraction and minimizing repulsion between partial charges on individual atoms of the rings. Of course, the overall electrostatic effect, although not homogeneous in base plane, is repulsive, with energies up to +4 kcal/mol for base-base interactions [14]. Van der Waals force, instead, related to dipole-induced dipole and induced dipole-induced dipole attractions, is certainly attractive [16].
The estimated range of the overall stacking energy between two base-pairs inside an helix is -6 to -17 kcal/mol ($\simeq$ 10-28 K_BT) depending on the different nucleotides [17, 14, 18]. Similar values are given for the enthalpic gain experimentally evaluated for the combination of pairing and stacking [8].
Finally, we note that base stacking is already known to be the driving force for aggregation of chromonic-type DNA complexes, as will be discussed in section 2.3.1, and appears to be relevant also in crystalline oligonucleotides [19] and in DNA/protein complexes [20, 21].
Estimated attractive interactions between two G-quartets are about -49 kcal/mol, while the electrostatic repulsive component is +28 kcal/mol, which gives an overall stacking energy of $\Delta H = -21 kcal/mol$ [22].

1.3.3 Chirality and charge

Although much of the behavior of the short DNA duplexes will be based on a hard rod description, we shouldn't forget that DNA is chiral (of definite handedness).
Due to its helical nature, it has grooves, and thus a "steric" chirality like "fusilli"-type of pasta; but the behavior of high concentrated DNA, such as the ordering in chiral nematic phase, which we'll describe in chapter 4, cannot be described by the simple steric packing [23].
However, nucleic acid duplex is a highly negatively charged polyelectrolyte due to ionized phosphate groups located in the backbone. Therefore, long-range electrostatic interactions between nucleic acids and counter-ions and co-ions in solution significantly affects the interactions among nucleic acids and their thermodynamics (see section 1.2.3).
Each phosphate group on the DNA chain has a $-e$ charge at neutral pH, which corresponds to a high surface charge density of about $-1\ e/nm^2$, helically pat-

terned. Studies incorporating the contribution of the chiral charge distribution to DNA-DNA interactions [23, 24] obtained pretty good agreement in the prediction of the cholesteric pitch.

1.4 DNA and liquid crystals: an endless love

As we are dealing with DNA and liquid crystals in this thesis, it's interesting to point out that they have an intimate connection, on one hand related to the discovery of DNA structure by Watson and Crick, and on the other hand as regards the arrangement of DNA inside the cell nucleus and its functions. We'll describe in detail the DNA liquid crystalline phases in chapter 4.

1.4.1 Liquid crystals and the structure of DNA

The history of the discovery of DNA structure is fascinating and is deeply connected to its tendency to form LC phases.
Between the late '40s and the early '50s, many new evidences provided hints to understand the structure of DNA, and yet, incredibly for us, grown up with the familiar image of the double helix, nobody had that simple idea until Watson and Crick built their model [1].
The prospects of elucidating the genetic function in terms of molecular structure were greatly improved when it was clarified that the genetic substance was DNA, which had a well-defined chemical structure, rather than a more complex nucleoprotein. There were many indications of simplicity and regularity in DNA structure. The chemists had shown that DNA was a polymer in which the phosphate and deoxyribose parts of the molecule were regularly repeated in a polynucleotide chain with $3'$-$5'$ linkages. An important regularity was discovered: although the sequence of bases along the poly-nucleotide chains was complex and the base composition of different DNAs varied considerably, the numbers of adenine and thymine groups were always equal, and so were the numbers of guanine and cytosine. In the electron microscope, DNA was seen as a uniform unbranched thread of diameter about 20 Å . It was also shown that the bases in DNA lay with their planes roughly perpendicular to the length of the thread-like molecule.
Among the first, Astbury and Bell, Wilkins and Franklin did X-rays diffraction experiments on DNA gels and pulled fibers, and found evidence of considerable regularity in DNA; Astbury correctly interpreted the strong 3.4 Å reflection as being due to planar bases stacked on each other and it was also known that the bases were hydrogen-bonded together; it was even suggested that the polynucleotide chains might be linked by these hydrogen bonds to form multi-chain micelles.
An important issue came from the dependence of DNA structure on the state of hydration, as seen in par. 1.1.1, and this in turn reflects on X-ray diffraction [25] : while in solution DNA prefers B-form, the less hydrated A-form gives a much more crystalline pattern. Not only are the reflections sharper, but more significantly, they lay on clearly distinguishable layer lines and row lines. It's easier to be indexed, and for this reason Rosalind Franklin concentrated on it.
The B-form pattern (the famous photo 51 taken from Franklin, figure 1.12), observed by Watson and described to him, was sufficient to make Francis Crick

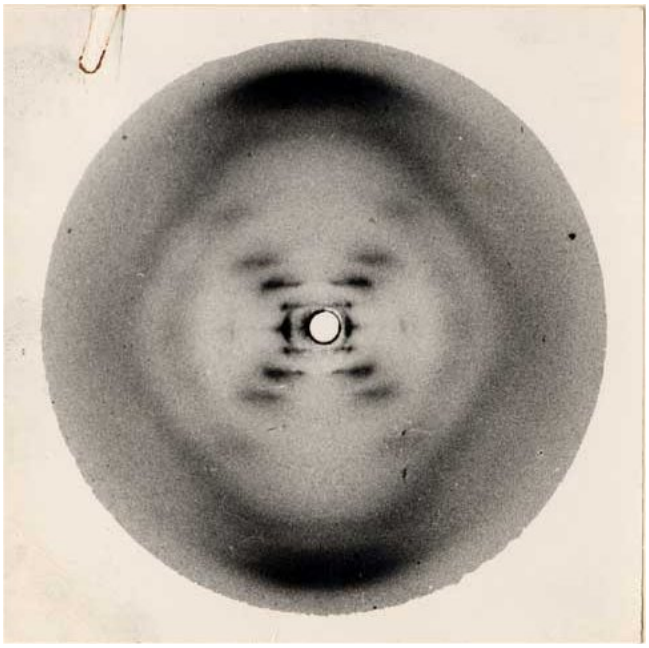

Figure 1.12 - The famous photo 51 by Rosalind Franklin, showing the diffraction pattern of B-DNA helices.

see things differently. Working on diffraction patterns of helices, he knew what the molecular transform of a helix looked like. The details of the A pattern gave plenty of information, but it concerned the lattice parameters - the packing of the strands rather than the internal structure within a strand. For the B-form it was the other way round. The lateral disorder in the specimen had blurred the reciprocal lattice, leaving the transform of the molecule that indicated, at his eyes, a 10-fold helix. Further crystallographic considerations (based upon the A-form too, and the continuous transformation between the two forms) led him to understand that two anti-parallel strands were giving that pattern. Watson's idea of A-T and C-G pairing completed the picture.

1.4.2 Liquid crystalline organization of DNA inside the cell nucleus

The familiar image (figure 1.1) of chromosomes can be actually quite confusing, because they become compact and assume their famous X shape only when cells are about to divide.
At other stages in the life of the cell, far from cell division, the chromosomes are generally more extended and less condensed. Nevertheless, DNA's concentration inside the nucleus is incredibly high, up to 400 mg/ml [26] (even higher concentrations are claimed [27]).
Liquid crystalline organization of chromatin, similar to that described in chapter 4, can be found *in vivo*: hexagonal packing in bacteriophages and certain sperm heads [27], cholesteric organization in dinoflagellate chromosomes, bac-

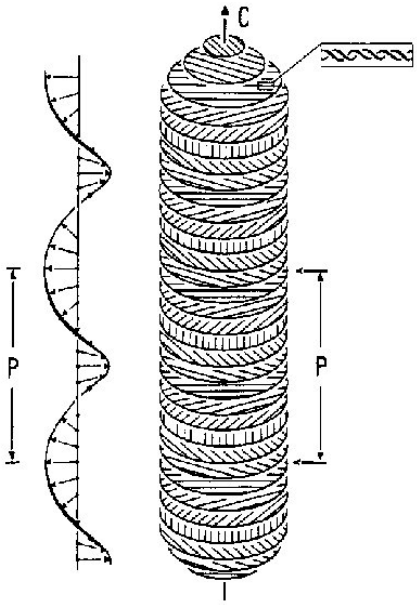

Figure 1.13 - Schematic drawing of the proposed cholesteric structure of a Dinoflagellate chromosome, from [28].

terial nucleoids and mitochondrial DNA, helical-shaped chromosomes in many species [28]. The different forms of condensed chromatin seem to be related to different local concentrations of DNA, a observed *in vitro*; actually, a parallel can be made between the geometries of liquid crystalline DNA *in vitro* and the organization of chromatin *in vivo*.
A particularly well studied example is given by dinoflagellates, which are primitive unicellular algae. In the nucleus, the main part of chromatin remains condensed throughout the cell cycle in the form of elongated chromosomes. Different patterns can be recognized: parallel alignment of chromatin filaments and periodic patterns with and without series of arcs. They correspond to different orientations of the chromosome axis (and cholesteric axis) relative to the section plane, as sketched in figure 1.13. Unfortunately, most of the microphotographs from samples *in vivo* found in literature are very poor.

The dense packing of DNA is interpreted to be the solution to store the genetic material into a small volume and, at the same time, to also preserve its integrity at particular moments of the cell cycle or of the life of the organism. According to this view, the liquid crystalline arrangement could be the best compromise between packing and ability to decondense and to regain activity.
An interesting example is given in [29]: bacteria lack nucleosomal organization and, during starvation, the enzymes promoting endurance to damaging factors get degraded and inactivated. Nonetheless, bacteria like *E.coli* show a remarkable resistance also in nutrient-depleted habitats, a result of finely tuned and fully reversible intracellular phase transitions. These nonenzymatic transitions, detected and studied in bacteria as well as in defined *in vitro* systems, result

in DNA sequestration and generic protection within tightly packed and highly ordered assemblies. This behavior is proven in *E.coli* with or without Dps, an important DNA-binding protein. In presence of Dps, DNA packs and segregates within DNA-protein micro-crystalline assemblies, while DNA protection in bacteria that lack Dps is shown to derive from a collapse into a cholesteric liquid-crystalline phase, appearing as a series of nested arcs [30] in figure 1.14. These processes result in ordered DNA structures whose condensed organization limits accessibility to damaging factors.
DNA reorganization into a cholesteric phase is also accompanied by a segregation of the chromatin and the ribosomes. This phenomenon can be directly attributed to entropic effects. The entropic cost of packing spherical particles (i.e. ribosomes) with rod-like molecules is large enough to exceed the entropic cost of total demixing (see chapter 5). Consequently, a concentrated cholesteric phase of the DNA rods is formed, which completely excludes ribosomes.
A similar strategy, though not involving LC but almost crystalline packing, is pursued by the bacterium *Deinococcus radiodurans*, that survives ionizing and UV radiations at higher doses than any other organism. It was proposed [31] that the bacterium is able to accurately reconstruct its genome from hundreds of radiation-generated fragments, although in the absence of an intact template, thanks to its DNA conformation. Indeed, its genome assumes an unusual toroidal morphology that may contribute to its radioresistance: because of restricted diffusion within the tightly packed and laterally ordered DNA toroids, forming a rigid matrix, radiation-generated free DNA ends are held together, which may facilitate template-independent yet error-free joining of DNA breaks.

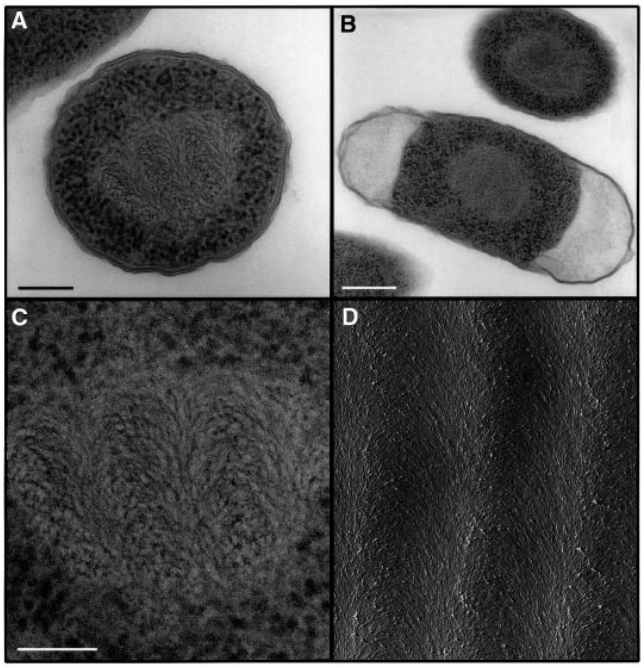

Figure 1.14 - (A-C) Electron microscopy images of bacteria cells and (D) of *in vitro* DNA cholesteric phase: in all the images nested arcs are clearly visible. Scale bars are 150, 300 and 100 nm in A,B,C; from [29].

Bibliography

[1] J. D. Watson, F. H. C. Crick, *Molecular structure of nuclei acids - A structure for Deoxyribose Nucleic Acid*, Nature **171**,737 (1953)

[2] C. R. Calladine, H. R. Drew , B. F. Luisi, A. A. Travers, *Understanding DNA - The molecule and how it works*, 3rd edition, Elsevier Academic Press (2004)

[3] C. R. Cantor, P. R. Schimmel, *Biophysical chemistry*, Freeman (1980)

[4] J. B. LePecq, C. Paoletti, *A flourescent complex between ethidium bromide and nucleic acids: physical-chemical characterization*, J. Mol. Biol. **27**,87 (1967)

[5] R. Owczarzy, *PhD Thesis*, University of Illinois at Chicago (1999)

[6] F. Reif, *Statistical and thermal physics*, McGraw-Hill international editions (1985)

[7] J. Santalucia, *A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics*, Proc. Natl. Acad. Sci. USA **95**,1460 (1998)

[8] J. Santalucia, D. Hicks, *The Thermodynamics of DNA Structural Motifs*, Annu. Rev. Biophys. Biomol. Struct **33**,415 (2004)

[9] Y. Lu, B. Weers, N. C. Stellwagen, *DNA Persistence Length Revisited*, Biopolymers **61**,261 (2002)

[10] B. Tinland, A. Pluen, G. Sturm, G. Weill, *Persistence Length of Single-Stranded DNA*, Macromolecules **30**,5763 (1997)

[11] M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, T. Ha, *Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy*, Biophys. J. **86**,2530 (2004)

[12] O. Kratky, G. Porod, *Röntgenuntersushung gelöster Fagenmoleküle*, Rec. Trav. Chim. Pays-Bas **68**,1106 (1949)

[13] K. Kopecka, G. Drouin, G. W. Slater, *Capillary electrophoresis sequencing of small ssDNA molecules versus the Ogston regime: Fitting data and interpreting parameters*, Electrophoresis **25**,2177 (2004)

[14] M. J. Packer, M. P. Dauncey, C. A. Hunter, *Sequence-dependent DNA structure: dinucleotide conformational maps*, J. Mol. Biol. **295**,71 (2000)

[15] C. A. Hunter, *Sequence-dependent DNA structure: the role of base stacking interactions*, J. Mol. Biol. **230**,1025 (1993)

[16] E. T. Kool, *Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replication*, Annu. Rev. Biophys. Biomol. Struct **30**,1 (2001)

[17] J. Šponer, J. Leszczyński, P. Hobza, *Nature of Nucleic Acid-Base Stacking: Nonempirical ab Initio and Empirical Potential Characterization of 10 Stacked Base Dimers. Comparison of Stacked and H-Bonded Base Pairs*, J. Chem. Phys. **100**,5590 (1996)

[18] J. Šponer, P. Jurečka, I. Marchan, F. J. Luque, M. Orozco, P. Hobza, *Nature of Base Stacking: Reference Quantum-Chemical Stacking Energies in Ten Unique B-DNA Base-Pair Steps*, Chem. Eur. J. **12**,2854 (2006)

[19] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Crystal structure analysis of a complete turn of B-DNA*, Nature **287**,755 (1980)

[20] M. R. Redinbo, L. Stewart, P. Kuhn, J. J. Champoux, W. G. J. Hol, *Crystal Structures of Human Topoisomerase I in Covalent and Noncovalent Complexes with DNA*, Science **279**,1504 (1998)

[21] C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, T. J. Richmond, *Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution*, J. Mol. Biol. **319**,1097 (2002)

[22] J. T. Davis, *G-Quartets 40 Years Later: From* 5′*-GMP to Molecular Biology and Supramolecular Chemistry*, Angew. Chem Int. Ed. **43**,668 (2004)

[23] F. Tombolato, A. Ferrarini, *From the double-stranded helix to the chiral nematic phase of B-DNA: A molecular model*, J. Chem. Phys. **122**,054908 (2005)

[24] A. A. Kornyshev, S. Leikin, S. V. Malinin, *Chiral electrostatic interaction and cholesteric liquid crystals of DNA*, Eur. Phys. J. E **7**,83 (2002)

[25] J. E. Lydon, *The DNA double helix - the untold story*, Liq. Cryst. Today **12**,1 (2003)

[26] J. Pelta, D. Durand, J. Doucet, F. Livolant, *DNA Mesophases Induced by Spermidine: Structural Properties and Biological Implications*, Biophys. J. **71**,48 (1996)

[27] F. Livolant, *Ordered phases of DNA in vivo and in vitro*, Physica A **176**,117 (1991)

[28] F. Livolant, M. F. Maestre, *Circular Dichroism Microscopy of Compact Forms of DNA and Chromatin in Vivo and in Vitro: Cholesteric Liquid-Crystalline Phases of DNA and Single Dinoflagellate Nuclei*, Biochemistry **27**,3056 (1998)

[29] A. Frenkiel-Krispin, S. Levin-Zaidman, E. Shimoni, S. G. Wolf, E. J. Wachtel, T. Arad, S. E. Finkel, R. Kolter, A. Minsky, *Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection*, EMBO J.**20**,1184 (2001)

[30] F. Livolant, A. Leforestier, *Condensed phases of DNA: structures and phase transitions*, Prog. Polym. Sci. **21**,1115 (1996)

[31] S. Levin-Zaidman, J. Englander, E. Shimoni, A. K. Sharma, K. W. Minton, A. Minsky, *Ringlike Structure of the Deinococcus radiodurans Genome: A Key to Radioresistance?*, Science **299**,254 (2003)

CHAPTER 2

Liquid crystalline ordering and living polymerization

Liquid crystals (LC) are a beautiful example of how the shape of the molecules, and thus entropic contributions, in addition to and sometimes above any other interaction, can lead to a wide variety of mesophases. The two liquid crystalline classes, thermotropics and lyotropics, and the main phases will be described.
We'll describe the Onsager theory of LC ordering, with some of its refinements, because it will be used in chapter 4 to analyze our data.
In addition, the so-called "living polymers" systems, which combine chain formation and various forms of ordering, will also be introduced, since they are a good model for the experimental observations described in this thesis. A particular emphasis will be given to the equations relating average length, concentration and sticking energy.

2.1 Liquid crystals

All matter exists in one of several states such as, for example, solid, liquid or gas having different degree and type of positional and orientational order. In the isotropic liquid or gas phases, molecules without order occupy the whole space and have no long-range positional or orientational order. The solid state can be divided to several categories, such as glass or crystal. In a typical crystalline phase, molecules or atoms have three-dimensional long-range positional order and in molecular crystals there is also long-range orientational order. The mesomorphic phases (or mesophases) between the isotropic liquid and the crystalline phases, called liquid crystals (LC), are fluids with some degrees of positional and/or orientational ordering of the molecules [1, 2, 3].
Liquid crystals are usually made of strongly anisotropic organic molecules, either elongated (calamitic, rod-like molecules) or disk-like (discotic molecules). As a rule, the inner part of mesogenic molecules is rigid (e.g. phenyl groups) and the

outer part flexible (aliphatic chains). This double character explains altogether the existence of steric interactions (between rod-like or disk-like cores) yielding orientational order and the fluidity of the mesomorphic phases. Typical examples are cyanobiphenyls and MBBA. These produce thermotropic mesophases, i.e. phases with a single component, whose phase transitions can be induced by a change in temperature.
The other broad LC class is constituted by the lyotropic mesophases: they occur when anisotropic amphiphilic molecules (soaps, phospholipids, various types of surfactant molecules and biomolecules) are added to a solvent. Because amphiphiles have two distinct parts, a polar head and a non-polar tail, the building units of lyotropic phases are usually aggregates of many molecules (micelles) rather than single molecules. This microphase separation dominating the lyotropic behavior is partly present also in thermotropic LC, as for example in the smectic phases, where polar and non polar portions of the molecules form distinct alternatinig planes in the system.
A typical example of lyotropics is a water solution of SDS, sodium dodecyl sulphate. For concentrations above the critical micellar concentration, *cmc*, these molecules form aggregates of different shapes, spherical or cylindrical micelles, bilayers, inverse cylinders, and inverse micelles. These complex elements are in turn building blocks for various phases with long-range order, that we'll consider later.
Synthetic macromolecules, made of mesogenic monomers, attached either chain-like or comb-like to a backbone, may also be building blocks for LC phases; biological polymers (DNA, PBLG, Xanthane), due to the rigidity of their backbones, and some viruses with highly anisotropic shape, such as TMV, form liquid crystal phases in solution in vitro. DNA LC phases will be the subject of chapter 4.
Another interesting example of lyotropic mesomorphism is presented by the so-called chromonic LCs, comprising a range of dyes and *nucleic acids*. Due to the importance of this class, we'll treat it separately in par. 2.17.
The arrangement of LC structures is the result of molecular anisotropy and of anisotropic interactions. Within a coarse-grained description, though, a LC is an elastic continuum, whose local configuration comes from the competition of elastic energy and the boundary conditions on the surfaces [2].

2.1.1 Classification of LC phases

LCs show many possible structures, which can belong to the same compound (polymorphism). There are four basic types of liquid crystalline phases, classified accordingly to the dimensionality of the translational correlations of building units: nematic (no translational correlations), smectic (1D correlation), columnar (2D) and various 3D-correlated structures, such as cubic phases.
The emphasis given in this introduction to some of the phases does not reflect their absolute importance, rather their importance in the following. Their textures under polarizing microscopy will be introduced in section 3.2.2.

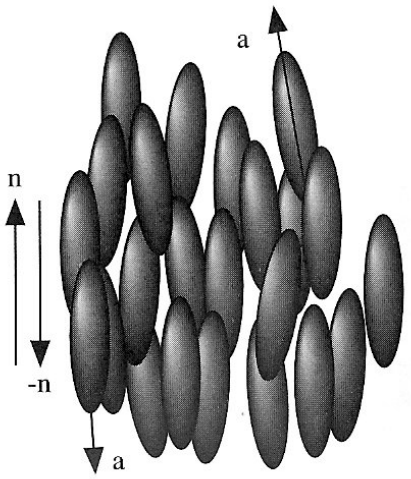

Figure 2.1 - Uniaxial nematic phase, with molecular axis **a** and director **n**; from [2].

Nematic and chiral nematic phase

Uniaxial nematics, indicated with N, are optically uniaxial phases (see fig. 2.1 and chapter 3). The unit vector **n** along the optical axis is called the director. Even when the building units are polar (as cyanobiphenyls), molecular flip-flops and head-to-head overlapping establish centrosymmetric average arrangement in the nematic bulk. Thus, **n** and $-\mathbf{n}$ are equivalent notations. The director is an axis of continuous rotational symmetry: the symmetry point group of the N phase is the same as that of a homogeneous circular cylinder, $D_{\infty h}$ [4]. The molecules, which are anisotropic in shape, align on average parallel to **n**.
Anyway it must be recalled that N phase is fluid: the centers of mass of the molecules are not correlated, like in a liquid.
A measure of the alignment of the molecules is given by the order parameter S, defined as:

$$S = \frac{1}{2}\langle 3\cos^2\theta - 1\rangle \tag{2.1}$$

where θ is the angle between the long molecular axis and **n** and the brackets denote a statistical (spatial and temporal) average. It changes continuously from $S = 0$, for random orientations, to $S = 1$, for perfect alignment.

When the building block (molecule or aggregate) is chiral, the nematic phase may turn to cholesteric, N^*, which also possesses solely orientational order of the long molecular axis. This phase, locally identical to the N phase, differs from it because of a a spontaneous macroscopic helical superstructure with a twist axis perpendicular to the local director. Thus the phase consists of local nematic "layers", which are continuously twisted with respect to each other. Assuming the helix axis to be along the z direction, the N^* phase exhibits a director field $\mathbf{n}(r) = \mathbf{n}(\mathbf{z})$, which is described by:

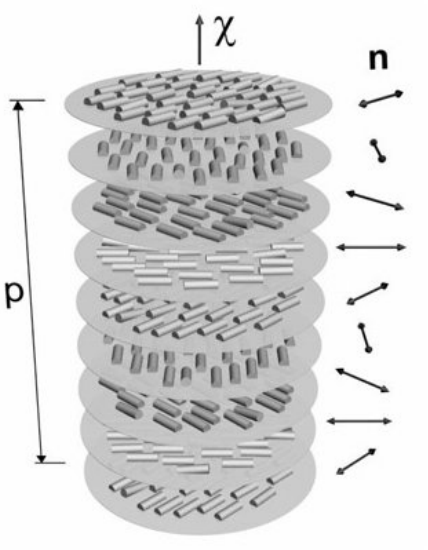

Figure 2.2 - Director configuration in the cholesteric phase; from http://dept.kent.edu/spie/liquidcrystals/index.html.

$$\mathbf{n}(\mathbf{z}) = \begin{pmatrix} \cos(2\pi z/p + \phi_0) \\ \sin(2\pi z/p + \phi_0) \\ 0 \end{pmatrix} \tag{2.2}$$

where p is the pitch of the helical superstructure (positive for a right-handed helix and negative for a left-handed one) and ϕ_0 is a constant that depends on the boundary conditions. The magnitude of the pitch of a cholesteric phase can vary from values as small as $p \simeq 100nm$ to large values of p equal to many micrometers. A schematic view of the corresponding director configuration is shown in figure 2.2. Note that, owing to the head-tail symmetry of the molecules ($\mathbf{n} = -\mathbf{n}$), the periodicity of the phase along the helix axis is given by half the pitch, $p/2$.

Being a chirality dependent quantity, the cholesteric twist can be either right-handed or left-handed, depending (also) on the configuration of the chiral elements within the molecule and on elastic constants. However, the relationship between molecular chirality and macroscopic phase chirality still remains an open question.

Smectic phases

Smectics are layered phases with quasi-long-range 1D translational order of centers of molecules in a direction normal to the layers, with a sinusoidal distribution of centers of mass along this direction. Within the layers, instead, the molecules show mainly fluid-like arrangement, some of them having also a bond-orientational order. Accordingly, Smectics can be regarded as stacks of 2D liquids or else as 1D crystals. They are of three main types, described below.

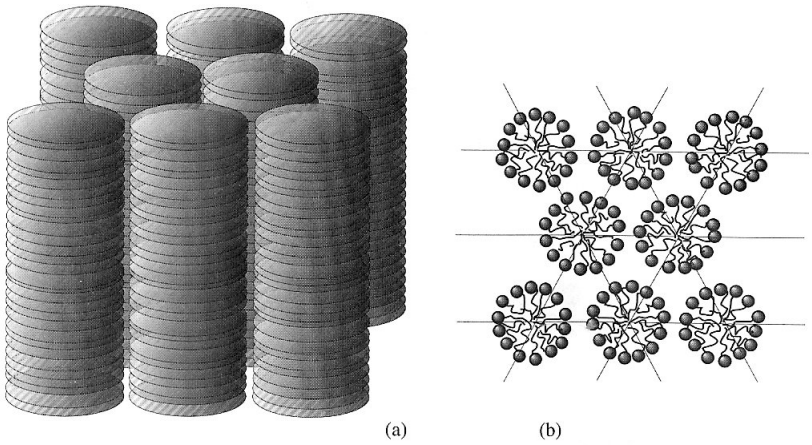

Figure 2.3 - (a) Thermotropic hexagonal columnar phase formed by disk-like molecules; (b) Lyotropic hexagonal columnar phase formed by cylindrical micelles, 2D cut normal to cylinders. From [2].

- Smectic A (SmA) is a uniaxial medium with the optic axis perpendicular to the layers; the director **n** is along the normal to the layers. There is no long-range positional order within the layers: each layer is a 2D fluid.
- Smectic C (SmC) is also composed of a 1D stack of fluid layers; however, it is a biaxial phase because the long axes **a** of the molecules are tilted with respect to the layers' normal **t**. The axes **a** average to the "nematic" director **n**, if no attention is paid to the layers.
- Smectic C^* (SmC^*), composed of chiral molecules, is a chiral version of SmC. The molecular tilt precesses around the normal to the layers.

Some smectic phases with liquid layers display hexagonal bond-orientational order in the layers. They are called hexatic smectics. For example, in hexatic smectic B, the molecules are normal to the layers and have no positional order within the layers, as in SmA. However, they show long-range hexagonal ordering of the directions that link the molecules (bond ordering). Tilted versions of hexatic B phase are hexatic smectics F and I. Note that the layers in A,C,C^* and hexatic smectics are not correlated: these phases have been indeed used as experimental models to verify theories of 2D media.

Columnar phases

Columnar phase can appear either in lyotropic systems, or with discotic molecules, or in chromonics (examples are given in figure 2.3). In its simplest form, it consists of molecules (discs or rods) stacked one on top of the other aperiodically to form liquid-like columns, the different columns constituting a 2D lattice. The structure is often the highly packed hexagonal, but a number of variants has been identified: hortorombic, rectangular, tilted, etc.
In hexagonal phase, the one of interest in this thesis, molecules are unidirectionally aligned with a lateral hexagonal order. However, the structure is not

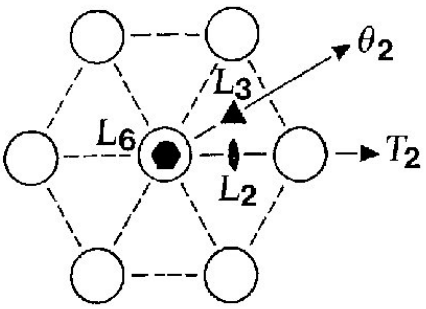

Figure 2.4 - The symmetries of columnar hexagonal phase; from [5].

that of a true crystal: molecules present some disorder around their position in the hexagonal array and the columns of molecules generally show a parallel continuous bend and are able to slide with respect to each other. Moreover, each molecule is free to rotate around its longitudinal axis. It is possible to define three series of molecular layers, separated by an angle of 60°, according to the three main directions of the hexagonal network. Molecules lie in the planes of these layers, whose thickness d is equal to $a \sin 60°$, a being the spacing between adjacent molecules.
The symmetries of the columnar hexagonal liquids are illustrated in fig. 2.4. L_6, L_3 and L_2 are, respectively, 6-fold, 3-fold and 2-fold axes parallel to the molecular direction; the other 2- fold axes (T_2 and θ_2) are normal to the molecular direction. Each point of L_6 or L_2 is a center of symmetry and each plane normal to the molecules is a plane of symmetry as well as (L_6, T_2) and (L_6, θ_2).

Tridimensional phases

- Lyotropic cubic phases are formed of bilayers that extend along three directions of space, as a result of the minimization of the bending energy of bilayers. The periodic arrangement characterizes the layers, while the molecules are free to move [2].

- Blue phases (figure 2.5) are made of chiral molecules that organize in an inhomogeneous way on the following basis. Let $\mathbf{n_0}$ be some director. In the local state of the smallest energy, the chiral molecules in the vicinity of $\mathbf{n_0}$ have the tendency to rotate helically along all directions perpendicular to $\mathbf{n_0}$, not only along one direction $\mathbf{c}$, as in the N^* phase. This geometry, which is called a double-twist, is energetically preferable to the 1D twist, at least for some chiral materials. However, as the distance from the director $\mathbf{n_0}$ increases, the cholesteric cylindrical shells become flatter and the double twist smoothly disappears. The director far-field configuration comes closer to the 1D twist of the N^* phase; the energy gain is reduced. Thus, the double twist cannot extend over the whole 3D space. A typical radius of the energy-gaining cylindrical region about the $\mathbf{n_0}$ axis is the half-pitch $p/2$. These cylinders of finite radius cannot tile space continuously. The situation is reminiscent of the phenomenon of frustration,

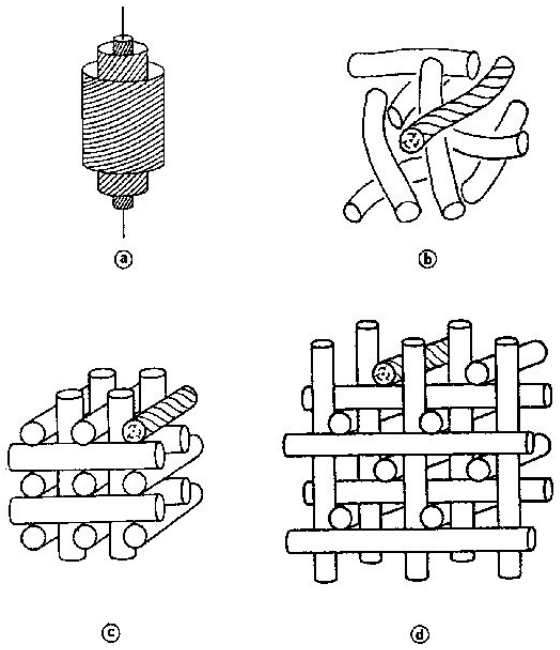

Figure 2.5 - Blue phases double twist cylinder (a) in which the molecular orientations are drawn onto concentric cylindrical surfaces. The cylinders are either disordered in an isotropic liquid (b) or packed into a cubic structure (c-d); from [5].

which is relieved by disclination lines, either regularly distributed or in disorder.
BP of type I and II are modeled as regular networks of disclination lines with periodicity of the order of p. Indeed, the 3D periodic structure of these phases is revealed in their nonzero shear moduli, ability to grow well-faceted monocrystals, and ability of Bragg reflection in the visible part of the spectrum (thus, when viewed between crossed polarizers, the phases often appear blue). BP are indicated as the pre-cholesteric ordering in DNA liquid crystals.

· The isotropic L_3 sponge phase is described as a bilayer of surfactant that extends through all space in a random fashion and divides the solvent into two connected continuous domains [2].

2.2 Descriptions of LC ordering

The more or less complex arrangement of LC structures is the result of molecular anisotropy and of anisotropic interactions. The simplest phase transition, and thus the first challenge for theories aiming to describe the behavior of LCs, is the onset of long-range orientational ordering, i.e., the isotropic to nematic transition, and its explanation was first exposed by Onsager in 1949 [6]. We'll deal with the calculations in paragraph 2.2.2, let's now introduce a simple physical description of his arguments [7]. The phenomenon of alignment transition turns out to be intimately related to that underlying the freezing transition for hard spheres. Consider a suspension of hard rods, i.e., prolate objects which exert no forces whatsoever on one another until they touch, at which point they experience infinite repulsions. Accordingly, the system is athermal: energetic considerations never enter into the problem, and the free energy consists entirely of entropic contributions. The first entropic term is that arising from the distribution of orientations of the particles, and it favors the isotropic solution in which all rod orientations are equally likely. The second is associated with the number of ways one can pack the hard rods, and it favors the aligned fluid state in which neighboring rods are more likely to be parallel. This is because the volume excluded to one hard rod by another is a minimum when their axes lie along the same direction and a maximum when they are perpendicular. Indeed, the fact that rodlike objects take up less space - can pack more efficiently (in more ways) - when they are parallel is obvious to anybody who has tried to fit spilled wooden matches back into their box without first aligning them. Since this packing entropy contribution to the free energy involves pairs of particles in lowest order, its dependence on density is quadratic and higher, whereas the orientational term, being a single-particle contribution, varies linearly with density. Hence, at high enough concentrations the packing entropy term will dominate, and the overall free energy of the system will be minimized for the aligned fluid state. Onsager showed in particular that this transition should be first order and that for long rods the limiting concentration - particle volume fraction - for the stability of the isotropic phase should be of order D/L, where D and L describe the diameter and length of the rods.
We'll now start from the formal expression of the entropy in a system of spheres,

and then we'll extend it to hard rods.

2.2.1 Entropy in disordered systems

The number η_i of ways to introduce one particle in a volume V of solvent much larger than the volume displaced by the particle is proportional to V, $\eta_i = AV$, where A is some constant of proportionality. Let us introduce n identical solute particles and make the hypothesis that A does not depend on the number of particles already present, a reasonable assumption until the volume occupied by n particles is small compared with V. Then, the number of ways to introduce these n particles is [2]

$$\eta_n = \frac{(AV)^n}{n!} \tag{2.3}$$

where the number of equivalent permutations $n!$ takes into account that the particles are identical. The associated entropy $S = nK_B \ln(\eta_n)$, using Stirling formula for large n, is given by

$$S \simeq K_B n \ln(AVe/n) \tag{2.4}$$

The free energy of the solution of noninteracting particles is then

$$F = U - TS = nF_0(T) - nK_BT \ln(V/n) \tag{2.5}$$

where the A and the e terms are now into the internal energy $nF_0(T)$. The reduced pressure can be expressed as

$$p \equiv -\left(\frac{\partial F}{\partial V}\right)_{T,n} = \frac{nK_BT}{V} = cK_BT \tag{2.6}$$

with $c = n/V$ concentration of particles. The physical meaning of p, the partial derivative of the free energy at constant n, can be understood as follows: imagine that the solution of particles and a pure solvent are in contact through a membrane that is permeable for the solvent but not for the particles. The quantity $-pdV$ is the free energy variation of the system in a reversible process in which the volume of the solution is varied by an infinitesimal amount dV. Therefore, p is the normal force per unit area exerted by the solvent on the membrane, measured along the outer normal; p is called *osmotic pressure*. Note that with the free energy 2.5, eq. 2.6 can be written as $pV = RT$, i.e. as the equation of state of an ideal gas, if V is considered as the volume of solution that contains 1 mole of the solute.
Equation 2.5 is modified when one takes into account, at the first order, interactions between couples of particles; hence, with the use of the virial expansion we can rewrite the free energy (per particle) as

$$F = F_0 + K_BT[\ln(c) + cv/2 + O(c^2)] \tag{2.7}$$

where the second virial coefficient v has the dimensions of volume and comes from pair interactions between particles. Terms of higher order in Eq. 2.7 would correspond to interactions involving three or more particles. If no attractive interactions exist and the particles are approximated by hard spheres of radius

r_0, then it can be shown that v is an excluded volume (the volume that is not allowed for the center of sphere "1" when sphere "2" is fixed at the origin):

$$v = \frac{4}{3}\pi(2r_0)^3 = 8v_p \tag{2.8}$$

where v_p is the volume of one particle. Generally, v accounts for both repulsions (not only steric) and attractions and is calculated from the pair interaction potential $w_{12}(r)$ function of the distance, assuming central forces

$$v = \int_{r=0}^{\infty} [1 - \exp(-w_{12}/K_BT)]4\pi r^2 dr \tag{2.9}$$

Note that in the presence of solvent, the potential $w_{12}(r)$ depends not only on the direct particle-particle interactions, but also on particle-solvent and solvent-solvent interactions. Thus, the last expression is justified only when the solvent can be treated as a continuum medium whose presence can be accounted for in the potential.

2.2.2 Onsager's theory

Onsager [6] discussed the statistics of a solution of hard rods of length L and diameter D, with the following assumptions [1]:

1. The only forces of importance correspond to steric repulsion; the rods cannot interpenetrate each other (the effect of electrostatic repulsion between charged rods can also be included by increasing the effective diameter of the rods [8], see par. 2.2.4);
2. The volume fraction $\Phi = c\frac{1}{4}\pi LD^2$, where c is the concentration of rods, is much smaller than unity;
3. The rods are very long ($L \gg D$). It will turn out in practice that the values of Φ of interest near the isotropic-nematic transition are such that $\Phi \sim 4D/L$. Thus requirements 2 and 3 are, in fact, linked.

Differently from spheres (eq. 2.7), for rods we must specify not only the overall concentration c, but also their angular distribution; let us call $cf_{\mathbf{a}}d\Omega$ the number of rods per unit volume pointing in a small solid angle $d\Omega$ around a direction labelled with the unit vector $\mathbf{a}$. Note incidentally that the sum of these solid angles must give the total concentration c, i.e.

$$\int f_{\mathbf{a}} d\Omega = 1 \tag{2.10}$$

The free energy is now given by a natural extension of equation 2.7:

$$F = F_0 + K_BT\left(\int f_{\mathbf{a}} \ln(4\pi f_{\mathbf{a}} c)d\Omega + \frac{1}{2}c\int\int f_{\mathbf{a}} f_{\mathbf{a}'} v(\mathbf{a}\mathbf{a}')d\Omega d\Omega'\right) + O(c^2) \tag{2.11}$$

The second term in equation 2.11 describes the drop in entropy associated with molecular alignment (a non-constant f). The third term describes the excluded

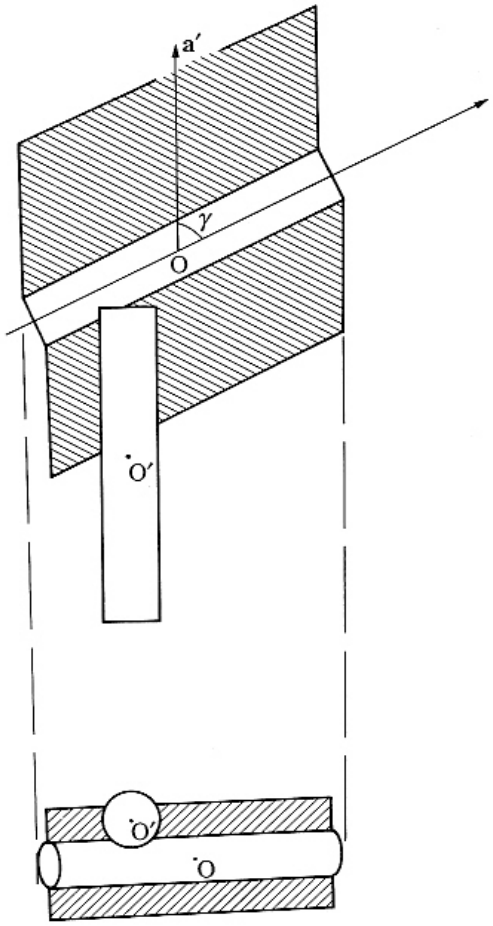

Figure 2.6 - Excluded volume for two long hard rods (end effects are not considered); from [1].

volume effects: $v(\mathbf{aa}')$ is the volume excluded by one rod in direction $\mathbf{a}$ as seen by another rod in direction $\mathbf{a}'$. The calculation of v is simple for the long rods where end effects are ignored and is explained in fig 2.6. The result is

$$v = 2L^2 D|\sin\gamma| \quad (L \gg D) \tag{2.12}$$

where γ is the angle between $\mathbf{a}$ and $\mathbf{a}'$.
It must be emphasized that equation 2.11, limited to order c, is a mean-field approximation: correlations between different rods are not taken into account. We can obtain a self-consistent equation for the distribution function $f_{\mathbf{a}}$ by specifying that the free energy 2.11 is a minimum for all variations of $f_{\mathbf{a}}$ that satisfy the constraint 2.10. This corresponds to writing

$$\delta F = K_B T \lambda \int \delta f_{\mathbf{a}} d\Omega \tag{2.13}$$

where λ is an unknown Lagrange multiplier and gives the self-consistent equation

$$\ln(4\pi f_{\mathbf{a}}) = \lambda - 1 - c \int v(\mathbf{aa}') f_{\mathbf{a}'} d\Omega' \tag{2.14}$$

λ is then determined by the normalization condition 2.10.
Equations 2.12 and 2.14 show that the concentration c enters the problem only through the combination $cL^2D = const \cdot \Phi L/D$.
Equation 2.14 always has an "isotropic" solution ($f_{\mathbf{a}} = 1/4\pi$, independent of $\mathbf{a}$) but, *if* $\Phi L/D$ *is large enough*, it may also have anisotropic solutions, describing a nematic phase. It is difficult to solve the non-linear integral equation 2.14 exactly. Onsager used a variational approach, based on a trial function of the form

$$f_{\mathbf{a}} = const \cdot \cosh(\alpha \cos\theta) \tag{2.15}$$

where α is a variational parameter, and θ is the angle between $\mathbf{a}$ and the nematic axis (the constant factor is chosen to satisfy 2.10). In the region of interest, α turns out to be large (~ 20) and the function f is strongly peaked around $\theta = 0$ and $\theta = \pi$. The order parameter is

$$S = \frac{1}{2} \int f_{\mathbf{a}} (3\cos^2\theta - 1) \sin\theta d\theta \simeq 1 - 3/\alpha \quad (\alpha \gg 1) \tag{2.16}$$

Minimizing the energy F (2.14) with respect to α, one obtains a function $F(c)$ showing a first-order phase transition from isotropic ($\alpha = 0$) to nematic ($\alpha \sim 20$). The volume fraction Φ occupied by the rods, in the nematic phase, just at the transition point, is $\Phi_N^c = 4.5D/L$, while for the isotropic phase in coexistence with the nematic $\Phi_{Iso}^c = 3.3D/L$.
Note that Φ_N^c and Φ_{Iso}^c are independent of T in this model because hard rods are an "athermal" system. Of particular interest is the value of the order parameter S_c in the nematic phase just at the transition. This turns out to be quite high ($S_c \simeq 0.84$). Thus the Onsager solution leads to a rather abrupt transition between a strongly ordered nematic and a completely disordered isotropic phase.

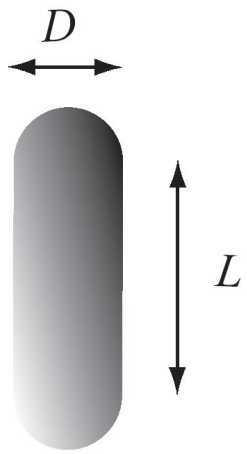

Figure 2.7 - Schematic representation of a spherocylinder of length L and diameter D.

The theory has been subjected to many experimental tests, involving aqueous solutions of long, rigid, rod-like, macromolecules such as TMV [9], or sterically stabilized, inorganic rodlike particles [10]: its reliability, of course, decreases with decreasing aspect ratios and when interactions others than excluded volume cannot be neglected.

2.2.3 Simulations

Besides theory and experiments, computer simulations are a useful method to investigate phenomena, because they have the possibility to capture, on the basis of simple assumptions, a wide range of complex behaviors. Indeed, from simulations a wonderful confirmation of Onsager concept (although, as we'll see, much has to be refined) was obtained.
Spherocylinders are a natural anisotropic extension of spheres. They have been shown to reproduce many liquid-crystalline phases, including smectics (unlike simple ellipsoids). This is probably the principal reason for their wide use in molecular simulations, in spite of their high level of simplication with respect to real LCs.
A spherocylinder (figure 2.7) is defined as a cylinder of diameter D and length L, capped with two hemispheres of diameter D. The total spherocylinder length is $L+D$, while quantities better used to parametrize the system are the anisotropy parameter L/D or the length-to-width ratio $L/D+1$ (these are respectively 0 and 1 for spheres).
The simplest nontrivial molecular interaction we can conceive is the hard repulsive potential:

$$u(r_m) = \begin{cases} \infty, & r_m \leq D \\ 0, & r_m > D \end{cases} \quad . \tag{2.17}$$

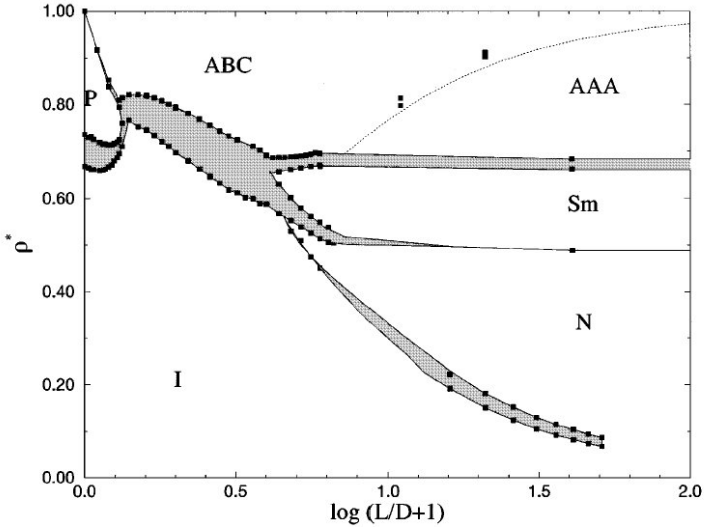

Figure 2.8 - Phase diagram of hard spherocylinders with L/D between 0 and 100, from [11]. The volume fraction ρ^* is plotted as a function of $\log(L/D+1)$ in order to give equal emphasis to all parts of the phase diagram. Coexistence regions are shaded. The following phases can be distinguished: the low-density isotropic liquid, the high density orientationally ordered solid, the low-L/D plastic solid, and for $L/D > 3.7$ the nematic and the smectic-A phases.

which prevents overlap of molecules whose separation is less than r_m.
The hard repulsive potential is a short range potential since molecules do not interact until they overlap. The physical origin of this kind of interaction may be found in the Pauli exclusion volume. Of course this approach is actually neglecting the long range interactions, which may play an important role in the formation of mesophases characterized by a long range order. Nevertheless the consequences of this approximation are quantitative rather than qualitative, so that the hard repulsive model remain one of the most appealing in the study of mesophases.
We should also remind one more time that the definition of the pair potential u in equation 2.17 implies system independence of temperature, since temperature will appear just as a scaling quantity in the equation of state. Consequently, the hard repulsive potential may be seen as an appropriate model potential primarily for lyotropic liquid crystals, whose phase transitions are driven by density changes and not by temperature changes.
Despite its simplicity, the hard spherocylinder model is able to reproduce many of the principal phases of matter; figures 2.8 and 2.9 show the complete phase diagram for hard spherocylinders, calculated by Bolhuis and Frenkel [11].
In particular, note that for $L/D < 3.7$ there is no nematic phase, in pretty good accordance with Onsager model.

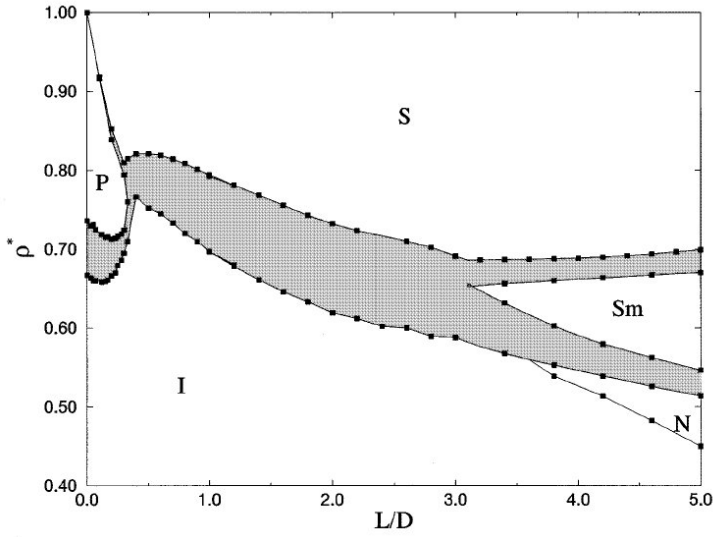

Figure 2.9 - Phase diagram for hard spherocylinders of aspect ratio $L/D \leq 5$, from [11].

2.2.4 Beyond Onsager

Despite its huge importance for establishing the concept of pure entropic ordering, Onsager theory's validity is limited to monodisperse, long hard rods, without any other interaction than excluded volume, and can account for the nematic phase formation only.
In this section, we will briefly describe some of the extensions to Onsager's description, in particular for the role of flexibility or polydispersity of rods on the phase diagram. Without going into detailed calculations and comparison with experiments (for which we refer to original papers), we'll show the main effect of these added features.

Charged rods

Besides the hard core repulsion, charged rods have a long range repulsive interaction of the following form

$$\frac{U_{el}(x)}{K_B T} = \frac{A' e^{-\kappa(x-D)}}{\sin\gamma} \tag{2.18}$$

where x is the closest distance between two charged rods, A' is the proportionality constant obtained by solving the Poisson-Boltzman equation (it depends on surface charge density), κ^{-1} is the Debye screening length which depends on ionic strength and γ is the angle between two rods.
In the case of charged rods there are contributions to the second virial coefficient v from both the hard core excluded volume interaction (eq. 2.12) and the long range electrostatic repulsion interaction. These two contribution lead to the expression [8]:

$$v = \frac{1}{4}\pi L^2 D_{eff} = \frac{1}{4}\pi L^2 D + \frac{1}{4}\pi L^2 \kappa^{-1}(\ln A' + C) \tag{2.19}$$

with C a numerical constant. Therefore the thermodynamics of charged rods will be equivalent to the thermodynamics of thicker hard rods with the effective diameter D_{eff}. This holds exactly for the isotropic phase, while, for the nematic, it doesn't because the electrostatic energy is lower for perpendicular rods then for parallel rods. Therefore the charge effectively destabilizes the nematic phase by shifting the I-N transition to higher concentrations and reducing the order parameter of the nematic phase coexisting with the isotropic phase.

Finite length

For the theory to work at higher densities (or, equivalently, for shorter aspect ratios) the free energy of the unperturbed liquid of rods needs to take into account third and higher virial coefficients. An alternative theory that accomplished this uses scaled particle free energy of hard rods [12] (used among others, in [13], [14] and [15].
"Scaling" allows the general difficulty of evaluating many-body interactions to be avoided by interpolating between two extremes that are easy to evaluate. Since, on the one hand, an infinitesimally small particle can be excluded by only one other particle at a time, the probability of being able to insert it into a crowded solution is easy to calculate. At the other extreme, the difficulty of inserting a macroscopically large particle into a crowded solution simply corresponds to the pressure-volume work of excavating the necessary hole. The wonder of scaled particle theory is that interpolation between these two extremes, with care to include enough scaling dimensions and preserve thermodynamic consistency, yields remarkably accurate estimates of the configurational entropy as assessed by computer simulations and experimental studies.
Without going into details (see [14]), we show, in figure 2.10, how scaled particle theory can better deal with smaller aspect ratios than Onsager's model, in good accordance with simulated results obtained by [11].

Flexible rods

The effects of flexibility on the Iso-N transition of semiflexible "wormlike" chain molecules are fairly well understood, shifting the densities - or concentrations - at the transition to higher values and decreasing the phase gap with increasing degree of flexibility. These findings agree qualitatively with experimental results for suspensions of virus particles and also have been applied to thermotropic liquid-crystalline polymers.
The effects of flexibility on transitions to smectic phases (not directly addressed by Onsager, but interesting for our system) have also been studied. Computer simulations indicate that increasing flexibility also shifts the nematic-smectic or the nematic-columnar transition (depending on the systems) to higher densities. Similar findings have been obtained by extensions of the Khokhlov-Semenov [16] theory for wormlike chains. In [17], hard flexible rod-like polymers are considered, including the treatment of phases with long-range order as hexagonal columnar. The conclusion is that, as the rods become more flexible the Iso-N recedes to higher concentrations and finally, for small enough persistence length, the N phase disappears, and there is a direct transition from the isotropic to

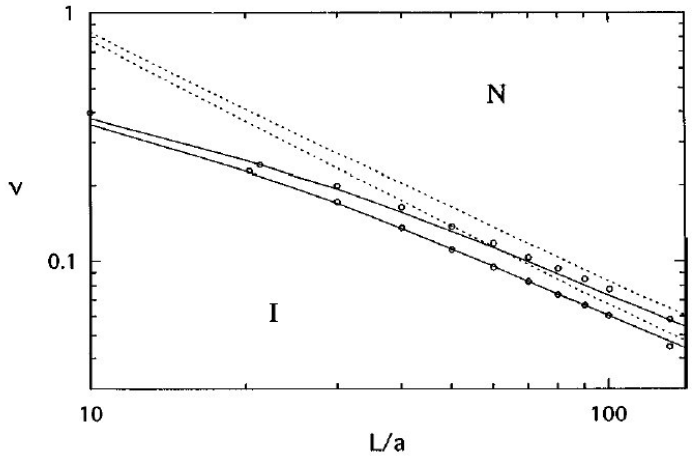

Figure 2.10 - Solute volume fractions at isotropic-nematic coexistence as a function of spherocylinder aspect ratio. Circles are simulation results of Bolhuis and Frenkel [11]. Solid lines are calculated using scaled particle theory. Dashed lines are calculated using Onsager's second virial approximation; from [14].

the columnar phase.
Suppression of N phase and direct Iso-C transition are also found for high flexibility in [18], as shown in the phase diagram of figure 2.11.
Anyway, such different theoretical approaches and approximations can lead to very different results, and the comparison with experiments is often difficult; for example, [19] considers end-effects within the second-virial approximation and finds that the N phase would eventually disappear, leaving a direct Iso-Sm transition.
Soft-repulsive worm-like rods are simulated in [20]; an hexatic phase instead of columnar appears at high concentrations. With the increase of internal flexibility, the smectic A phase results to be shrunk by the nematic phase, at lower densities, and the hexatic, at higher densities; it eventually disappears in systems of sufficiently flexible rods.
In the end of this incomplete excursion in the field, the overall well-accepted conclusion can be that a higher flexibility of rods leads to a higher Iso-N concentration, possibly leading to N disappearance.

Polydisperse rods

It is widely accepted that length polydispersity of the rods suppresses the smectic phase. We give a couple of examples. For example in [22] long hard spherocylinders are studied. If the polydispersity is small (standard deviation $\sigma < 0.08$), the phase behavior is essentially unchanged from that observed in monodisperse systems; thus nematic, smectic, and crystal phases are exhibited. For an intermediate range of polydispersities ($0.08 < \sigma < 0.18$), the smectic phase is found to become increasingly destabilized with respect to the nematic phase at low densities and a columnar phase at high densities. This eventually

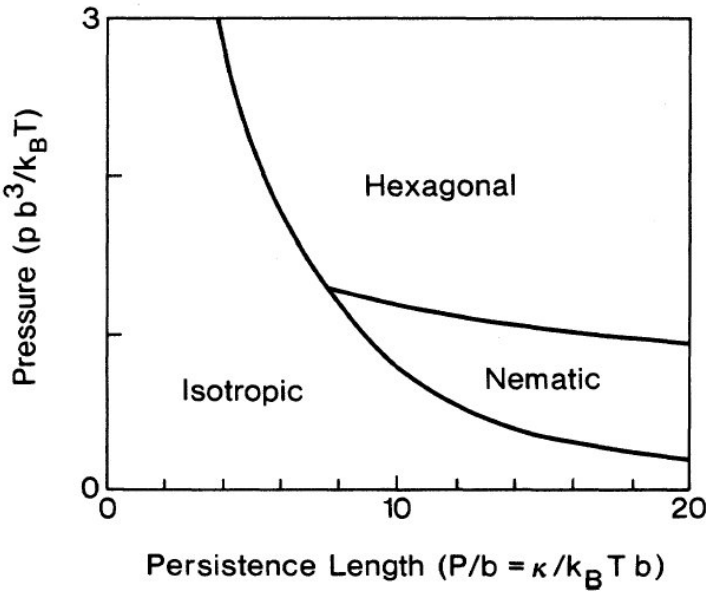

Figure 2.11 - Calculated phase diagram for infinite flexible rods, showing suppression of nematic phase; from [18].

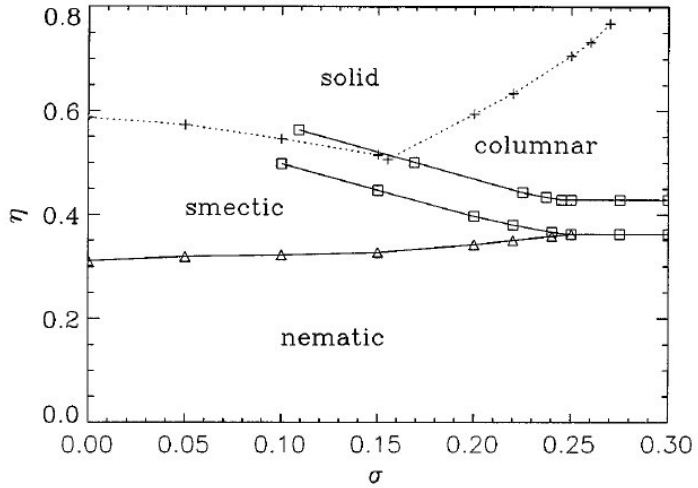

Figure 2.12 - Phase diagram for a polydisperse system of parallel hard rods. The distribution is modeled as a gaussian of standard deviation σ. Squares denote coexistence of phases, triangles a second-order phase transition, and crosses represent instabilities; from [21].

leads to a terminal polydispersity ($\sigma \approx 0.18$) above which the smectic phase is no longer stable.
In [21], density functional theory is applied to study the influence of polydispersity on the stability of columnar, smectic, and solid ordering in solutions of rods. For $\sigma > 0.25$ a direct first-order nematic-columnar transition is found (see fig. 2.12), while for smaller σ there is a continuous nematic-smectic and first-order smectic-columnar transition.

To summarize, we give an example of the combination of (some of) the features described above. In figures 2.13 to 2.15 (from [23]), the phase diagram of hard rods, calculated with scaled particle theory and the approach of [24], similar to that obtained in [11], changes by the addition of charge and polydispersity: I-N transition shifts to higher concentrations, Sm is suppressed and coexistence regions are narrower.

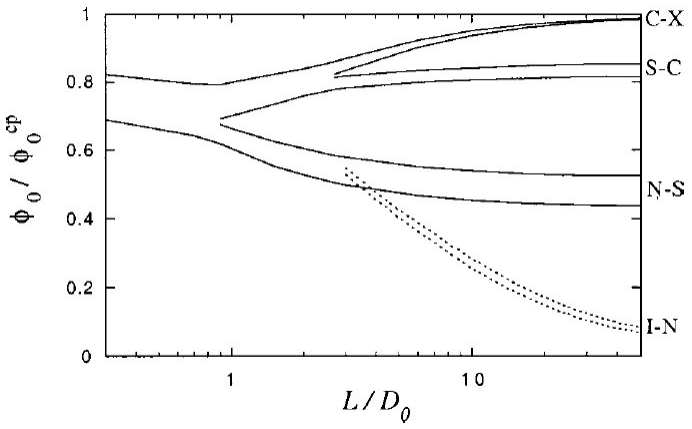

Figure 2.13 - Coexistence volume fractions vs. aspect ratio for uncharged spherocylinders. Solid lines are positional ordering predictions from [24], dashed lines indicate I-N transition from scaled particle theory. From [23].

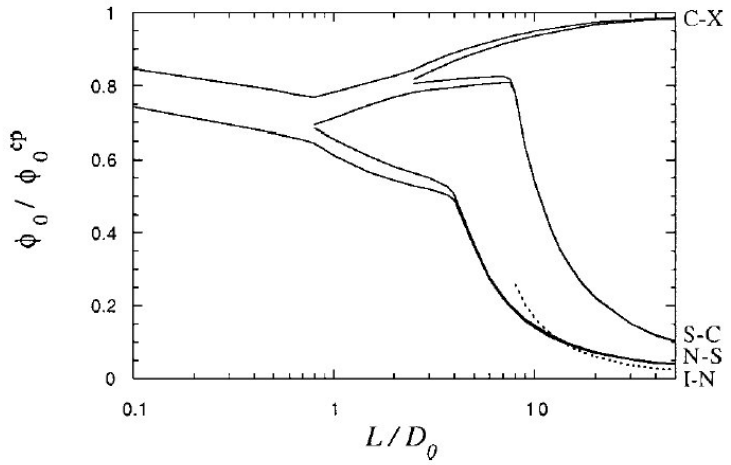

Figure 2.14 - Same as fig. 2.13, for charged spherocylinders; from [23].

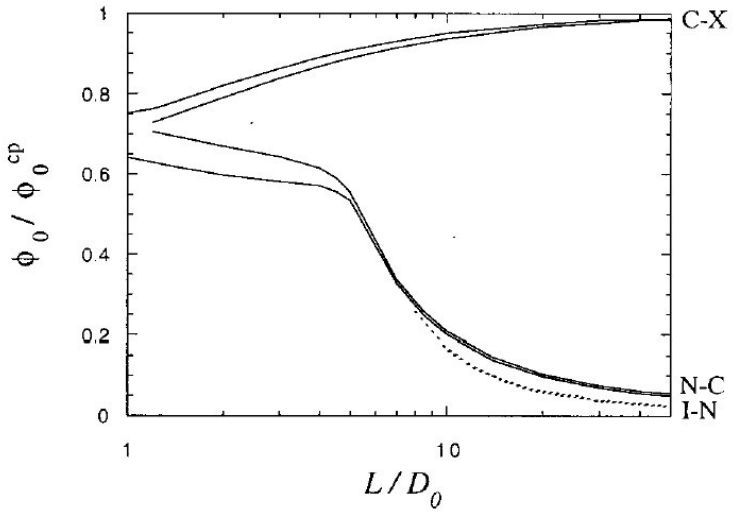

Figure 2.15 - Same as fig. 2.14, with the added effect of polydispersity; from [23].

2.3 Living polymers

Reversible self-assembly to form elongated aggregates is a distinctive property of a variety of molecules [25]. For example, poly-aromatic dye and drug molecules have the capability of reversibly aggregating into cylindrical stacks of indeterminate length (see par. 2.17). Analogously, some proteins have the functionally important capacity to reversibly aggregate into stiff multi-stranded filaments. Normal examples include the proteins that constitute the filamentous skeleton that is responsible for giving most cells their shape and strength, while sickle

cell hemoglobin constitutes a pathological example in which filaments distort and stiffen the red blood cells causing occlusion of blood vessels.
Perhaps the most common self-assembly occurs in surfactant solutions. For rigid surfactants (such as the perfluorinated fatty acids), micelles tend to grow with a disk-like morphology. However, greater polymorphism is possible for flexible surfactants, with micelles growing in one dimension to form rod-like particles or two dimensions to form plate-like particles, depending on conditions. An intrinsic property of self-assembling systems is the broad size distribution of the particles that are formed and its variability with concentration and temperature and other conditions; it arises from a competition between the free energy of association of monomers and the entropy loss due to the formation of polymers. As the temperature is lowered, the energetic gain overcomes the entropic loss, thereby providing the physical origin of polymerization upon cooling. For all these reasons, such systems are called "living polymers".
The other interesting feature of such systems is that, once they are concentrated enough, they can show ordered phases [26,27] thanks to the resulting high axial ratio; we have already described this behavior, of course, dealing with lyotropic mesophases formed by surfactants.
These two properties, variable length and ordering, are linked together: it was already noted in the very first works combining reversible polymerization and LC transition, to our knowledge [28] and [29] (extending respectively Onsager's and Flory's theories), that, since increasing rod length favors alignment, it is to be expected that alignment will also favor polymerization and a reciprocal coupling of orientation and polymerization will result. This mutual advantage is of fundamental interest for the discussion of the observed behavior of sDNA liquid-crystallization.
Figure 2.16, from [30], shows the phase diagram of rod-like aggregates depending on free energy Φ of monomer addition to a chain (or inversely on temperature). There is a low-concentration isotropic phase whose region of stability grows with increasing temperature (weaker aggregation). When Φ is large, aggregation is strong enough that long enough aggregates are formed at low concentrations to result in spontaneous orientational ordering (a nematic phase). At higher concentrations, positional ordering also sets in. For rod-like aggregates, the smectic phase is suppressed because, in that configuration, the variable lengths of the particles would waste a lot of space. On the other hand, the polydispersity of rod-like aggregates does not present a problem in the columnar phase. Thus for rod-like aggregates the nematic gives way to a columnar phase with increasing concentration. For most of the phase diagram the columnar phase is stable up to close packing ($v_p \sim 0.9$). However, when aggregation is very weak (small Φ or high T) the C phase becomes unstable with respect to a crystalline phase of short aggregates.
Analogous behavior is exhibited by a disk-like aggregating system, but for high concentrations the situation is the reverse, with polydispersity in diameter problematic in the columnar phase but accommodated well in the smectic phase. Thus, for disk-like aggregates the nematic gives way to a smectic phase at increasing concentrations. For both rods and disks, a jump occurs in the average aggregate size at each ordering transition, which reflects the coupling between spatial ordering and aggregate growth [30].
The topology of these phase diagrams has been observed in laboratory studies of polyaromatic molecules that self-assemble to form rod-like aggregates, or rigid

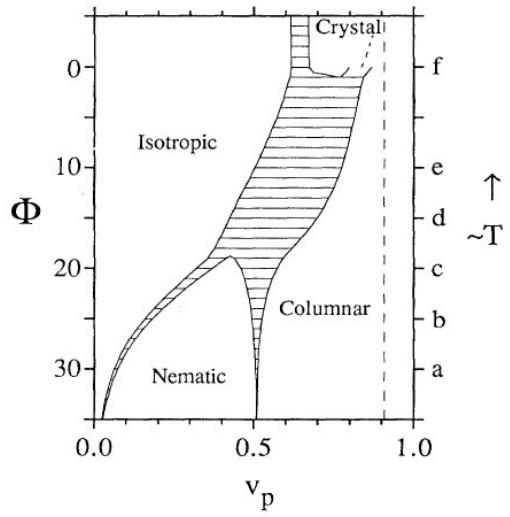

Figure 2.16 - Phase diagram of a self-assembling rodlike system, from [30]. Shaded areas indicate coexistence, while the vertical dashed line is the hard-rod close-packing limit. Points a-f represent different effective temperatures of the system.

surfactants that self-assemble to form disk-like aggregates.

2.3.1 Chromonics

An interesting subclass of living polymers are chromonics: they emerged as a family of lyotropic mesogens with properties distinct from those of conventional amphiphiles [31]. This family consists of various drugs (as disodium chromoglycate), dyes (as Sunset Yellow), nucleic acids and antibiotics. In almost every respect the properties are different from those of ordinary lyotropic mesogens of the phosphilipid type. The molecules have aromatic rather than aliphatic structures. They are rigid rather than flexible and planar disc-like or plank-like, rather than rod-like. The hydrophilic solubilising groups are disposed around the periphery of the molecules rather than at one end. The molecules aggregate in solution, not into micelles, but into columns, and they have distinctive optical textures [32].
The name chromonic was derived from the bischromone structure of the anti-asthmatic drug disodium chromoglycate, chosen because of the fortuitous combination of connotations of the word, with both colour (with reference to dyes) and with chromosomes (with reference to nucleic acids).
Chromonic mesogens can be regarded as being insoluble in one dimension. The classic chromonic phases are the nematic, N phase and the hexagonal, M phase (see figure 2.17). In both of these, the molecules are stacked in columns. In the N phase, these lie in a nematic array (i.e. the columns are more or less parallel, but there is no positional order and there is no orientational order of the columns about their long axes). In the M phase, the columns lie on a lattice with statistical hexagonal symmetry and have long-range order.

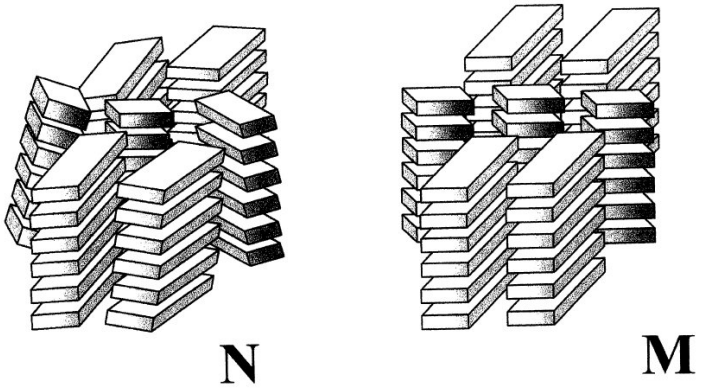

Figure 2.17 - The structure of chromonic N and M Phases, from [31].

Chromonic mesophases are the lyotropic counterparts of the discotic mesophases; in many aspects, chromonic systems are closer to thermotropic systems than to conventional amphiphiles. In both cases, the driving force causing liquid crystalline phase formation is the face-to-face aggregation of molecules forming columns and the geometrical aspects of the packing of these columns are more or less the same in the two cases. The difference being that in one case the columns lie in a sea of alkyl chains and in the other, they lie in a sea of water. The driving force for the aggregation is enthalpic rather than entropic, and in general, there is a strong tendency for chromonic molecules to aggregate into columns, even in very dilute solution - just as conventional lyotropic mesogens form micelles before a mesophase is formed [33].
Although there is a threshold concentration before significant aggregation begins to occur, there is no specific optimum column length and, therefore no analogue of the critical micelle concentration of conventional amphiphiles. The term "isodesmic" (first used in the study of the aggregation of nucleic acids in solution) has been applied to the steady build up of chromonic aggregates where the addition or removal of one molecule to a stack is always associated with the same increment of free energy.
This is in direct contrast with the situation for conventional amphiphilic association, where the micelle represents a free energy minimum - and there is a cost to the system in having either larger or smaller units.

Stacking of nucleic acids

Guanosine and its derivatives are known to be insoluble in water, often forming a gel: their edges have self-complementary hydrogen-bond donors and acceptors, and its polarizable aromatic surface, with a strong molecular dipole is ideal for stacking (see section 1.3.2). This basis for self-association was established in 1962 [34], when 3'-GMP (guanosine mono-phosphate) and 5'-GMP were found to form layers of hydrogen-bonded tetramers, planar macro-cycles held together by eight H-bonds of the Hoogsteen type. Their structure is shown in figure 2.18.

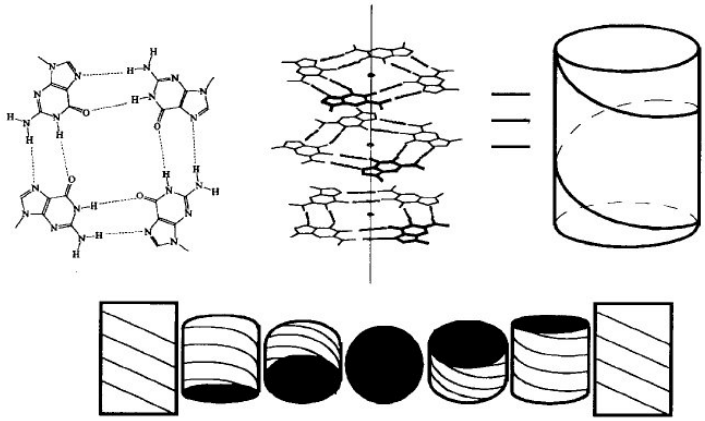

Figure 2.18 - Assembly of GMP into quartets, stacking and formation of cholesteric phase, from [36].

The reason why we're dealing with guanosine self-assemblies in chromonics section is that tetramers do not always form a gel, but different G-quartets can stack on each other (see [35] for a comprehensive review), forming helices of finite length. While stacking energy is too small for single Guanosines, the bigger surface of quartets promotes the aggregation. They do not stack in register, and are instead rotated one with respect to the other. The columns are similar to the four-stranded helix formed by poly(G).
These chiral columnar aggregates, already present in isotropic solutions, at some concentration can be in turn the building blocks for lyotropic mesophases, as can be seen in figure 2.18. They self-correlate to give cholesteric phases and, at higher concentrations, more ordered phases as columnar are obtained [36]. This feature is common to GMPs and guanosine oligomers up to 5'-GGGGGG-3', the basic repeat always being the G-quartet.

Interestingly, the tendency of nucleic acids to supramolecular self-assembly is replicated on a bigger length-scale in Nucleosome Core Particles, the elementary chromatin units (each of them is a protein complex with 146bp of *ds-DNA* around it) that we can see as beads on a string.
Their behavior has been studied under different ionic strength and osmotic pressure conditions [37], mimicking physiological conditions. Above a critical concentration, NCPs stack on top of each other to form columns that further organize into multiple columnar phases (see figure 2.19). An isotropic (and in some cases a nematic) phase of columns is observed in the moderate concentration range. Under higher concentration conditions, a lamello-columnar phase and an inverse hexagonal phase form under low salt conditions, whereas a 2D hexagonal phase or a 3D orthorhombic phase is found at higher salt concentration.
This spontaneous ordering behavior may be relevant on the highly dynamic hierarchical organization of chromosomes.

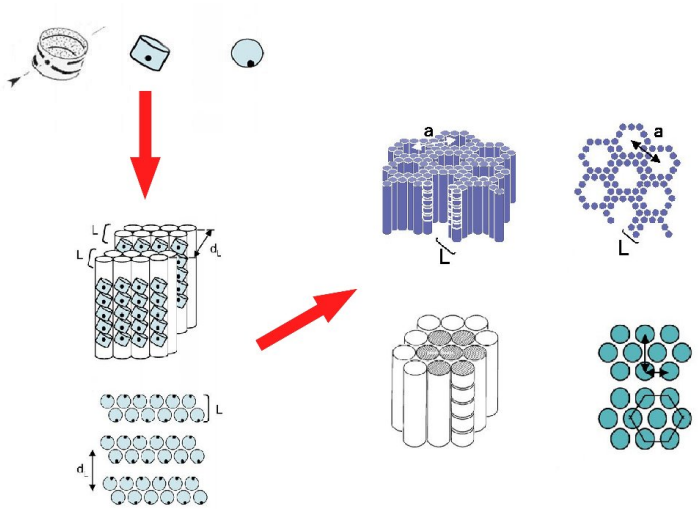

Figure 2.19 - Stacking of Nucleosome Core Particles and their supramolecular ordering; redrawn from [37].

2.4 Energy & average length

How is the equilibrium average length of "living" chains related to the energy of bonding? A great number of very different approaches can be found in literature, based on mean-field theories or scaling arguments, with independent chains or within models developed up to the second or higher order virial expansion. These predict pretty different behaviors and some of them suffer obscurity of approximations, dimensions and notations.
Early works on surfactant micelles (see [38]) employed either a Flory-Huggins mean-field approach which gives:

$$\bar{L} \simeq \varphi^{1/2} \exp(\varepsilon/2) \tag{2.20}$$

or made use of scaling arguments, good for strong excluded volume interactions:

$$\bar{L} \simeq \varphi^{y} \exp(\varepsilon/2) \tag{2.21}$$

with $y \simeq 0.6$. In the previous equations, $\bar{L}$ is the average length, φ the volume fraction and ε the energy cost (in units of $K_B T$) of breaking a long micelle into two pieces somewhere along the "chain"; in the case of sDNA, instead, the system is composed of discrete monomers (the short helices) aggregating in oligomers. That these estimates are not accurate for short equilibrium chains is confirmed by the fact that, for vanishing φ, $\bar{L}$ tends to 0, incorrect for a chain composed of monomers.
We now come to some of the possible approaches applicable in the regime of linear aggregates formed by few monomers.

2.4.1 Ideal gas approximation

We begin by considering an ideal gas of N_{tot} monomers in a volume V and at temperature T, capable of reversible self-assembly into linear aggregates (with energy gain $\varepsilon K_B T$ for each bond). The free energy of the family of aggregates with aggregation number n is

$$\frac{G_n}{VK_BT} = \rho_n[\log(\rho_n\lambda^3) - 1] - \rho_n(n-1)\varepsilon \tag{2.22}$$

where ρ_n is the number density of molecules belonging to that family and $\lambda = \sqrt{h^2/2\pi nmK_BT}$ is the thermal wavelength, with m the mass of a monomer. We can thus obtain the chemical potential as

$$\frac{\mu_n}{K_BT} = \frac{\partial}{\partial\rho_n}\left(\frac{G_n}{VK_BT}\right) = \log(\rho_n\lambda^3) - (n-1)\varepsilon \tag{2.23}$$

By integrating over the kinetic degrees of freedom, the thermal wavelength can be factorized and simplified, and by equating to the chemical potential of monomers, the concentration of molecules in n-mers is obtained in function of total concentration.
In Ref. [33] the description of the behavior of a chromonic system is based on this approach or, equivalently, on the law of mass action as invoked in [39]. Assuming an isodesmic aggregation, the following equation for the volume fractions is extracted:

$$\varphi_n = n\left[\varphi_1 \exp(\varepsilon)\right]^n \exp(-\varepsilon) \tag{2.24}$$

Summing over n, the total volume fraction is:

$$\sum_{n=1}^{\infty}\varphi_n = \frac{\varphi_1}{[1-\varphi_1\exp(\varepsilon)]^2} = \varphi \tag{2.25}$$

Eq. 2.25 gives a relation between distribution of monomers, energy per bond and total volume fraction. From the distribution, it is also possible to extract the mean chain length as

$$M = \frac{\sum_{n=1}^{\infty} n\frac{\varphi_n}{n}}{\sum_{n=1}^{\infty}\frac{\varphi_n}{n}} = \frac{2\varphi\exp(\varepsilon)}{\sqrt{1+4\varphi\exp(\varepsilon)}-1} \tag{2.26}$$

With the same type of approach for the free energy of semi-flexible chains, plus a term accounting for interactions among chains, Ref. [40] describes the behavior of dipolar fluids, forming linear aggregates, and obtains for M

$$M = \frac{1}{2}\left[1+\sqrt{1+4\rho\exp(S_0)}\right] \tag{2.27}$$

where S_0K_BT is interpreted as the mean free energy for each bond.
If monomer indistinguishability within a chain is taken into consideration (but there is no reason not to consider it for every monomer!), M can be obtained by the combination of the two equations below

$$\rho = \rho_1 \exp[\rho_1 \exp(S_0)]$$

$$M = \frac{\rho \exp(S_0)}{\exp[\rho_1 \exp(S_0)] - 1} \tag{2.28}$$

with ρ_1 the number density of free monomers.
We observe that equations 2.27 and 2.28 are not dimensionally consistent, if ρ, according to authors' definition, is the number density of monomers.

2.4.2 Mixing entropy

The sDNA solution can also be described as a multi-component mixture of reactive components, the sDNA n-mers of arbitrary size n. This is quite an old idea when applied to the self-assembly of amphiphilic molecules, as may be seen in [41]. The problem reduces, somehow, to the choice of the "right" form of the entropy of mixing.

Ideal mixture

An expression formally equivalent to equation 2.26 can be obtained from the following calculations, for which we are indebted to F. Nallet (Centre de recherche P.Pascal, Bordeaux).
The entropy of mixing is here computed on a mole fraction basis, as is usually done in the theory of ideal solutions [39], instead of on a volume fraction basis, as is more correct for large differences in the volumes of solute and solvent (like sDNA in water). However, according to Flory himself [42] this is, asymptotically at least, correct in the limit of vanishing concentrations even in the case of macromolecular solutions.
In general terms, taking into account the Gibbs-Duhem relation for a multi-component mixture, the free enthalpy for the small DNA solution - viewed as made of mixture of n-mers - may be written as

$$G = N_s \mu_s + \sum_{n=1}^{\infty} N_n \mu_n \tag{2.29}$$

where N_s is the total number of solvent molecules with chemical potential μ_s and N_n is the number of aggregates of size n, with chemical potential μ_n, considered as independent species.
In the ideal mixing approximation, and neglecting orientational entropy, etc. the relevant chemical potentials are expressed - for a solution both incompressible and dilute, and on a mole fraction basis - in the following way

$$\mu_n = n v_1 p + \bar{\mu}_n(T) + K_B T \ln \frac{N_n}{N_s} \tag{2.30}$$

where p is the pressure, T the temperature and v_1 the volume of monomers. Note that the "true" mole fraction of n-mers should be expressed as $N_n/(N_s + \sum_n N_n)$ but, at low enough solute content, this does not differ too much from N_n/N_s. Besides, the function of the temperature $\bar{\mu}_n(T)$ is the "reference" chemical potential for n-mers, which corresponds to the free energy to insert an aggregate

in a dilute solution, depending on hydration and internal energy.
Mass conservation requires

$$\sum_{n=1}^{\infty} nN_n = N_{tot} \tag{2.31}$$

where N_{tot} is the total number of molecules introduced into the solution, as a constraint on the size distribution $\{N_n\}$ of n-mers.
With A_k representing a k-mer, the self-assembly of molecules may be represented by chemical equilibria

$$A_n + A_q \rightleftharpoons A_{n+q} \tag{2.32}$$

with corresponding law of mass action relations

$$\mu_n + \mu_q = \mu_{n+q} \tag{2.33}$$

between the relevant chemical potentials.
Now comes the main assumption. The self-assembly of molecules, eq. 2.32, into one-dimensional aggregates is accounted for by expressing "reference" chemical potentials in the following form:

$$\bar{\mu}_n(T) = n\bar{\mu}_1(T) - (n-1)\varepsilon K_B T \tag{2.34}$$

where ε is a positive coefficient, of order 1 to 10. Eq. 2.34 reflects the linearity of hydration energy and of the whole oligomer free energy with the number of monomers, plus the contribution of interactions among monomers.
With this assumption, introducing eq. 2.30 and eq. 2.34 into eq. 2.33 leads to

$$\frac{x_n x_q}{x_{n+q}} = c^* \tag{2.35}$$

where x_k denotes the (approximate) mole fraction N_k/N_s of k-mers and c^* is defined by the relation

$$c^* = \exp(-\varepsilon) \tag{2.36}$$

This number is the analogue of the critical micellar concentration in classical micellization problems.
The central equation for self-assembly in the present context, eq. 2.35, is readily solved - taking into account mass conservation eq. 2.31 - in the following form

$$x_n = c^* X^n \tag{2.37}$$

with X, as a function of the total solute amount $c \equiv N_{tot}/N_s$, given by

$$X = 1 + \frac{c^*}{2c}\left(1 - \sqrt{1 + \frac{4c}{c^*}}\right) \tag{2.38}$$

In the limit where $c \ll c^*$, the ratio X of the geometric series reduces to the simple expression

$$X \approx \frac{c}{c^*} - 2\left(\frac{c}{c^*}\right)^2 \tag{2.39}$$

whereas at higher concentrations $c \gg c^*$ (to ensure self-consistency, c must still be much smaller than 1)

$$X \approx 1 - \sqrt{\frac{c^*}{c}} \tag{2.40}$$

The average size M of the aggregates is readily evaluated from the size distribution, eq. 2.37, as

$$M = \frac{\sum nN_n}{\sum N_n} = \frac{c}{c^*}\frac{1-X}{X} = \frac{c\exp(\varepsilon)\left[-1+\sqrt{1+4c\exp(\varepsilon)}\right]}{1+2c\exp(\varepsilon)-\sqrt{1+4c\exp(\varepsilon)}} \tag{2.41}$$

As may be deduced from the asymptotic limits, equations 2.39 and 2.40, as appropriate, the average size of the aggregates increases at first linearly with slope $1/c^*$ at low concentrations and crosses over to a square root behavior at higher concentrations.
We can rearrange equation 2.41 to obtain, as already observed, the very same expression of equation 2.26, except for the fact that volume fraction φ is substituted by molar fraction c. Since $c/\varphi \simeq V_{water}/V_{DNA} \sim 10^{-3}$, it is evident that the two equations will yield very different predictions.

Flory-Huggins

If, more appropriately, we consider solute molecules much larger than the solvent molecules, according to Flory [42] the mixing entropy is expressed in terms of volume fractions of i solutes and solvent:

$$\Delta S = N_s \ln \varphi_s + N_1 \ln \varphi_1 + ... + N_i \ln \varphi_i \tag{2.42}$$

where subscript s is for solvent. Thus the chemical potential regulating the equilibrium is

$$\mu_i = i\mu_0 - (i-1)\varepsilon K_B T + K_B T \ln \varphi_i \tag{2.43}$$

where μ_0 is a reference chemical potential and $\varepsilon K_B T$ is the energy of association of a pair of molecules. i in our case indicates the number of molecules associated in a chain.
Since for the aggregation process, if $i + j = m$,

$$\mu_i + \mu_j = \mu_m$$

holds, we obtain that

$$\frac{\varphi_i \varphi_j}{\varphi_m} = \exp(-\varepsilon)$$

Since $\sum_i \varphi_i = \varphi$, by assuming for the volume fractions the functional form $\varphi_i = \exp(-\varepsilon)X^i$, we obtain

$$X = \frac{\varphi}{\exp(-\varepsilon) + \varphi}$$

We therefore find the equilibrium mean length of the aggregates to be

$$M = \frac{\sum iN_i}{\sum N_i} = \frac{\varphi \exp(\varepsilon)}{\ln[1 + \varphi \exp(\varepsilon)]} \tag{2.44}$$

2.4.3 Second virial approximation

Compared to the treatments of the previous paragraphs, different approximations are used in the approach described in this section: on one hand, the system is not actually a mixture of aggregating monomers, with its length distribution, but instead, once obtained the length that minimizes the free energy, monodisperse chains *all* of that length. On the other hand, we go beyond the "ideal gas" treatment and add to the free energy both terms that account for mixing entropy and polydispersity, and a contribution from steric repulsion and end effects within second virial approximation [43].
In [44] (see also [27]), again describing surfactant micelles, a second virial term is added to the free energy of semi-flexible spherocylinders. The following relation is obtained:

$$\bar{L} = \sqrt{\rho\sigma^3 \exp(\varepsilon + \kappa_I \varphi)} \tag{2.45}$$

ρ is the number density of surfactant molecules, $\rho = \varphi 6t/\pi\sigma^3$, with σ the diameter of the micelles and t the number of molecules in capping hemispherical ends; κ_I is a constant related to the virial expansion, estimated 8/3 from geometrical considerations, but "the value [...] should not be taken too seriously" [44].
As observed for equations 2.20 and 2.21, $\bar{L} \to 0$ for vanishing concentrations (but not for vanishing binding energy).
In [43], this treatment is adapted to chains of spheres, comparing it to Monte-Carlo simulations. As is usual in simulations, all the variables are rescaled to σ, the diameter of particles, in particular the number density of particles ρ is expressed in units of σ^{-3}. They thus have:

$$M = \frac{1}{2}\left[1 + \sqrt{1 + 4\rho \exp(\varepsilon + \kappa_I \rho)}\right] \tag{2.46}$$

with $\kappa_I = 1.45\sigma^3$ obtained fitting the MC results. Equation 2.46 is the direct "discrete" version of eq. 2.45, if $m = 1$; indeed, it is 1 for $\rho = 0$ but goes parallel to 2.45 for $\rho > 0$. It is also close to eq. 2.27, with the additional concentration dependence of the exponent.
However, we note again that the expression is not dimensionally consistent, although numerically it is (we can think σ to be 1).

2.4.4 Perturbative theory

In [45], hard spheres with sticky poles are described with Wertheim perturbation theory and M is expressed as

$$M = \frac{1}{2}\left(1 + \sqrt{1 + 8\rho\Delta}\right) \tag{2.47}$$

where ρ is number density and Δ, which has the dimensions of a volume, reflects the interactions (excluded volume plus sticky poles) once averaged with correlation function and geometry of the system. At low concentrations, $\Delta \sim V_b[\exp(\varepsilon/K_BT) - 1]$, with V_b the spherically averaged bonding volume. Δ can

also be read as a chemical equilibrium binding constant, $2\Delta = K_b$, which is in turn related to the free energy; therefore, equation 2.47 can be rewritten as

$$M = \frac{1}{2}\left[1 + \sqrt{1 + 4\exp[(-\Delta U + T\Delta S)/K_B T]}\right]$$

where $\Delta U/K_B T = -\varepsilon$. Following authors' definition of the equilibrium constant, in this last expression - otherwise similar to 2.27 and 2.46 - the concentration dependence has disappeared, leading to the unphysical conclusion that the average length of the chain does not depend on monomers concentration.

To conclude, we found that also from the same type of approach different predictions can be made, and viceversa formally equivalent expressions are obtained within different conceptual frames. Anyway, comparison is often made hard by the lack of clarity in units and dimensions. It should be noted, however, that all the models cannot be expected to keep their validity when interactions between chains are no longer negligible. This may happen at quite low concentrations, because the overlap volume fraction of rod-like particles of length L and radius R (with $R \ll L$) scales like R^2/L^2.
In chapter 4, we'll make use of those consistent expressions that allow a comparison with experimental data to extract the stacking energy in sDNA system from the inferred average length. These relations are plotted in figures 2.20 and 2.21, either vs. bonding energy at fixed volume fraction or viceversa, with typical parameter values. We can see that, despite their significant functional and conceptual differences, the above equations give roughly similar results for the energy-length range we're interested in.

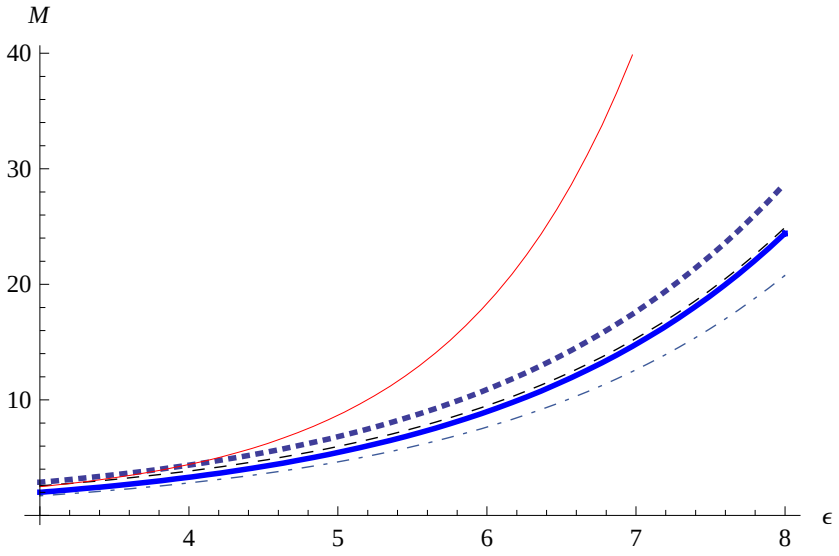

Figure 2.20 - Dependence of average aggregation number M on bonding energy ε according to various models (dotted line: eq. 2.46; dashed: eq. 2.26; thin: eq. 2.44; thick: eq. 2.20; dot-dashed: eq. 2.21); $\varphi = 0.2$.

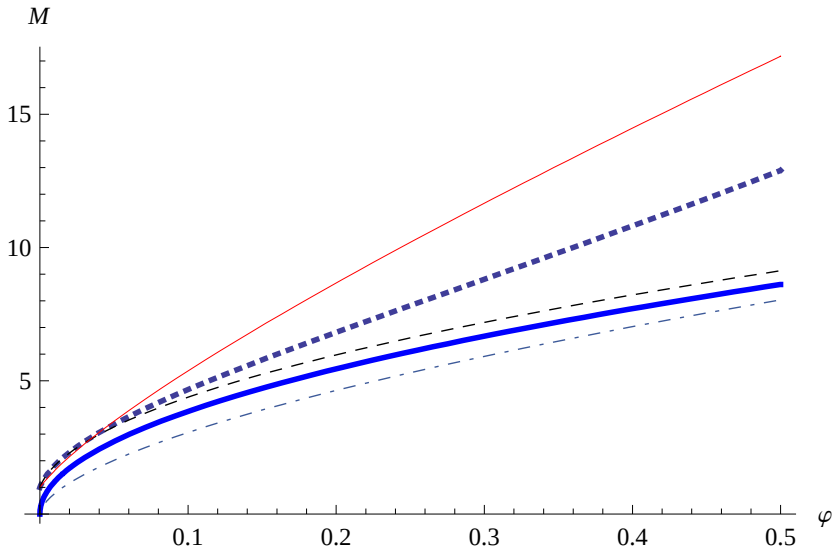

Figure 2.21 - Dependence of average aggregation number M on volume fraction φ according to various models, with $\varepsilon = 5$; legend as in figure 2.20.

Bibliography

[1] P. G. De Gennes, J. Prost, *The Physics of Liquid Crystals - 2nd edition*, Oxford University Press (1993)

[2] M. Kleman, O. D. Lavrentovich, *Soft Matter Physics*, Springer (2003)

[3] S. Chandrasekhar, *Liquid Crystals - 2nd edition*, Cambridge University Press (1992)

[4] P. W. Atkins, *Physical chemistry*, Oxford University Press (1999)

[5] F. Livolant, A. Leforestier, *Condensed phases of DNA: structures and phase transitions*, Prog. Polym. Sci. **21**,1115 (1996)

[6] L. Onsager, *The effects of shape on the interaction of colloidal particles*, Ann. NY Acad. Sci. **51**,627 (1949)

[7] W. M. Gelbart, A. Ben-Shaul, *The "New" Science of "Complex Fluids"*, J. Phys. Chem. **100**,13169 (1996)

[8] A. Stroobants, H. N. W. Lekkerkerker, T. Odijk *Effect of electrostatic interaction on the liquid crystal phase transition in solutions of rodlike polyelectrolytes*, Macromolecules **19**,2232 (1986)

[9] Z. Dogic, K. R. Purdy, E. Grelet, M. Adams, S. Fraden, *Isotropic-nematic phase transition in suspensions of filamentous virus and the neutral polymer Dextran*, Phys. Rev. E **69**,051702 (2004)

[10] P. A. Buining, H. N. W. Lekkerkerker, *Isotropic-Nematic Phase Separation of a Dispersion of Organophilic Boehmite Rods*, J. Phys. Chem. **97**,11510 (1993)

[11] P. Bolhuis, D. Frenkel, *Tracing the phase boundaries of hard spherocylinders*, J. Chem. Phys. **106**,666 (1997)

[12] M. A. Cotter, D. C. Wacker *Van der Waals theory of nematogenic solutions. I. Derivation of the general equations*, Phys. Rev. A **18**,2669 (1978)

[13] H. N. W. Lekkerkerker, A. Stroobants, *Phase behaviour of rod-like colloid+flexible polymer mixtures*, Nuovo Cimento D **16**,949 (1994)

[14] E. M. Kramer, J. Herzfeld, *Distribution functions for reversibly self-assembling spherocylinders*, Phys. Rev. E **58**,5934 (1998)

[15] G. Lasher *Nematic Ordering of Hard Rods Derived from a Scaled Particle Treatment*, J. Chem. Phys. **53**,4141 (1970)

[16] A. R. Khokhlov, A. N. Semenov *Liquid-crystalline ordering in the solution of partially flexible macromolecules*, Physica A **112**,605 (1982)

[17] R. Hentschke, J. Herzfeld *Isotropic, nematic, and columnar ordering in systems of persistent flexible hard rods*, Phys. Rev. A **44**,1148 (1991)

[18] J. V. Selinger, R. F. Bruinsma *Hexagonal and nematic phases of chains. II. Phase transitions*, Phys. Rev. A **43**,2922 (1991)

[19] R. C. Hidalgo, D. E. Sullivan, J. Z. Y. Chen *Smectic ordering of homogeneous semiflexible polymers*, Phys. Rev. E **71**,041804 (2005)

[20] G. Cinacchi, L. De Gaetani, Y. Martinez-Ratón, L. Mederos, E. Velasco, *Theory and Computer Simulations of Model Colloidal Liquid Crystals*, oral presentation at European Conference on Liquid Crystals 2007

[21] A. M. Bohle, R. Holyst, T. Vilgis *Polydispersity and Ordered Phases in Solutions of Rodlike Macromolecules*, Phys. Rev. Lett. **76**,1396 (1996)

[22] M. A. Bates, D. Frenkel, *Influence of polydispersity on the phase behavior of colloidal liquid crystals: A Monte Carlo simulation study*, J. Chem. Phys. **109**,6193 (1998)

[23] E. M. Kramer, J. Herzfeld, *Avoidance model for soft particles. II. Positional ordering of charged rods*, Phys. Rev. E **61**,6872 (2000)

[24] M. Taylor, R. Hentschke, J. Herzfeld *Theory of Ordered Phases in a System of Parallel Hard Spherocylinders*, Phys. Rev. Lett. **62**,800 (1989)

[25] J. Herzfeld, *Entropically Driven Order in Crowded Solutions: From Liquid Crystals to Cell Biology*, Acc. Chem. Res. **29**,31 (1996)

[26] M. P. Taylor, J. Herzfeld, *A Model for Nematic and Columnar Ordering in a Self -Assembling System*, Langmuir **6**,911 (1990)

[27] P. van der Schoot, M. E. Cates, *Growth, Static Light Scattering, and Spontaneous Ordering of Rodlike Micelles*, Langmuir **10**,670 (1994)

[28] R. Briehl, J. Herzfeld *Tactoidal state and phase transitions in systems of linear polymers of variable length*, Proc. Natl. Acad. Sci. USA **76**,2740 (1979)

[29] W. E. McMullen, W. M. Gelbart, A. Ben-Shaul, *Isotropic-nematic transition in micellized solutions*, J. Phys. Chem. **82**,5616 (1985)

[30] M. P. Taylor, J. Herzfeld, *Shape anisotropy and ordered phases in reversibly assembling lyotropic systems*, Phys. Rev. A **43**,1892 (1991)

[31] J. E. Lydon, *Chromonic mesophases*, Curr. Opin. Colloid Interface Sci. **8**,480 (2004)

[32] Yu. A. Nastishin, H. Liu, T. Schneider, V. Nazarenko, R. Vasyuta, S. V.Shiyanovskii, O. D. Lavrentovich *Optical characterization of the nematic lyotropic chromonic liquid crystals: Light absorption, birefringence, and scalar order parameter*, Phys. Rev. E **72**,041711 (2005)

[33] V. R. Horowitz, L. A. Janowitz, A. L. Modic, P. A. Heiney, P. J. Collings *Aggregation behavior and chromonic liquid crystal properties of an anionic monoazo dye*, Phys. Rev. E **72**,041710 (2005)

[34] M. Gellert, M. N. Lipsett, D. R. Davies, *Helix Formation by Guanylic Acid*, Proc. Natl. Acad. Sci. USA **48**,2013 (1962)

[35] J. T. Davis, *G-Quartets 40 Years Later: From* 5′*-GMP to Molecular Biology and Supramolecular Chemistry*, Angew. Chem Int. Ed. **43**,668 (2004)

[36] G. Gottarelli, G. Proni, G. P. Spada *The self-assembly and lyotropic mesomorphism of riboguanylic acids (GMP)*, Liq. Cryst. **22**,563 (1997)

[37] S. Mangenot, A. Leforestier, D. Durand, F. Livolant *Phase Diagram of Nucleosome Core Particles* , J. Mol. Biol. **333**,907 (2003)

[38] M. E. Cates, S. J. Candau, *Statics and dynamics of worm-like surfactant micelles*, J. Phys.: Condens. Matter **2**,6869 (1990)

[39] J. Israelachvili, *Intermolecular & Surface Forces - 2nd edition*, Academic Press (1992)

[40] P. I. C. Teixeira, J. M. Tavares, M. M. Telo da Gama, *The effect of dipolar forces on the structure and thermodynamics of classical fluids*, J. Phys.: Condens. Matter **12**,R411 (2000)

[41] D. Blankschtein, G. M. Thurston, G. B. Benedek, *Phenomenological theory of equlibrium thermodynamic properties and phase separation of micellar solutions*, J. Chem. Phys. **85**,7268 (1968)

[42] P. J. Flory, *Principles of Polymer Chemistry*, Cornell University Press (1953)

[43] X. Lü, J. T. Kindt, *Monte Carlo simulation of the self-assembly and phase behavior of semiflexible equilibrium polymers*, J. Chem. Phys. **120**,10328 (2004)

[44] P. van der Schoot, M. E. Cates, *The Isotropic-to-Nematic Transition in Semi-Flexible Micellar Solutions*, Europhys. Lett. **25**,515 (1994)

[45] F. Sciortino, E. Bianchi, J. F. Douglas, P. Tartaglia, *Self-assembly of patchy particles into polymer chains: A parameter-free comparison between Wertheim theory and Monte Carlo simulation*, J. Chem. Phys. **126**,194903 (2007)

CHAPTER 3

Materials and experimental techniques

An important, though not unique, part of the work of characterization of the sDNA LC phases was done through the observation of their textures with Polarized Transmission Optical Microscopy (PTOM), Polarized Reflection Optical Microscopy (PROM) and Fluorescence Optical Microscopy (FOM) over a wide range of experimental conditions. The correspondence between optical textures and molecular ordering will be introduced for the phases displayed by DNA samples. We get the occasion to show some of the textures observed in sDNA, not discussed in detail in chapter 4.
This chapter also describes principles and procedures of the other experimental techniques employed to study the samples: optical reflection interferometry (ORI) for the determination of the concentration, fluorescence microscopy for melting studies and X-ray diffraction to probe the phase structure.

3.1 Materials

3.1.1 Oligomers

sDNA oligomers (HPLC purified and ion free) of 6 to 24 base pairs in length were purchased from Bionexus Inc., (Oakland, CA, USA) and Primm s.r.l. (Milano, Italy) and obtained as lyophilized powder, usually in 0.2-0.3 mg aliquots. To obtain liquid crystal phases, oligomers were mixed with distilled deionized water (pH $\lesssim 5$); some experiments were also conducted with MES pH 7 buffer. The sequences studied in this thesis are listed below, from 5' terminal to 3', grouped by common features.

Self-complementary sequences

CGATCG (6 bp, SC), CGCATGCG (8 bp, SC), CGCAATTGCG (10 bp, SC), CGCGAATTCGCG ("Dickerson dodecamer" D12, SC), ACGCGAATTCGCGT (14 bp, SC), CGCGAAAATTTTCGCG (16 bp, SC), ACGCAGAATTCTGCGT (16 bp, SC), AACGCAAAGATCTTTGCGTT (20 bp, SC);
Different versions of SC 12 bp: AACGCATGCGTT, AACGAATTCGTT.

Partially self-complementary sequences/modified terminals

CGCGAATTCGCGT (12T), CGCGAATTCGCGTT (12TT), CGCAATTGCGTTTTTTTTTT (10-10T);
CGCGCGAATTCGCG (CG12); pCGCGAATTCGCG (p12, "p" stands for phosphorylated terminal), pCGCGAATTCGCGp (p12p);

Mutually complementary sequences

GGAGTTTTGAGG (A), CCTCAAAACTCC (B), CCTTCCAAAACCTTCC (C), GGAAGGTTTTGGAAGG (D);
Fluorescein-labeled versions of A: GGAGT(Fluorescein-dT)TTGAGG, Fluorescein-GGAGTTTTGAGG;
Modified versions of B: CCTCAGGACTCC, AATCAAAACTCC, ACTCAAAACTCA.

Random sequences

nnnnnnnn ("n" stands for random nucleotide, with equal 0.25 probability for the four bases), CGnnnnCG, CGnnnnnnCG, CnnnnnnG, CnnnnnnnnG;
selection of random *single stranded* 20-mers to 24-mers (SS-MIX, PCR-grade purification).

RNA sequences

CGCAUGCG (RNA 8 bp), CGCAAUUGCG (RNA 10 bp), CGCGAAUUCGCG (RNA 12 bp).

3.1.2 Other materials

"lDNA" was generated from DNA sodium salt from salmon TASTES (Sigma-Aldrich D1626) dissolved into pure water and sonicated for 3 hours by a Branson sonicator (SONIFIER 250) to reduce the molecular weight to $500 < N < 900$ bp (sample "S3h").
Poly-(ethylene glycol) (PEG) chains with average molecular weights of 200, 400, 600, 1000, 8000, 20000 and the fluorophore Ethidium Bromide were purchased from Sigma-Aldrich.

3.1.3 Cells

Very small volumes were necessary to reach high concentration from the tiny amounts of sDNA studied. Rectangular glass slides, typically 15-20 × 7-10 mm,

1 mm thick, were used to fabricate homemade "sandwich" cells, with stripes of Poly-ethylene terephthalate (PET) or Poly-imide (Kapton) film, 4 to 20 μm in thickness, as spacers. The resulting channel hosting the sample solution, about 10 mm long and 2 mm wide, was sealed with epoxy glue.
2 mm thick, high refractive index F2 glass (Schott AG, Mainz, Germany, n=1.62) was used in concentration measurements, while 50 μm float glass cells were fabricated for X-ray microbeam experiments.
When, instead of a concentration gradient, the experimental conditions required to keep the sample sealed for long periods or at high temperatures, usual sandwich cells, with thin film or silica rods (10 or 20 μm of diameter) as spacers, were sealed with a Fluorinated oil (Fomblin Y25, Ausimont, Italy), immiscible with water, highly viscous and with boiling point much above $100°C$.

3.2 Optical microscopy

Optical microscopy accomplishes three tasks: produces a magnified image of the specimen (magnification), separates the details in the image (resolution), and renders the details visible to the eye, camera, or other imaging device (contrast) [1, 2].
sDNA molecules are too small to be observed in a microscope: however, detailed information about their ordering is obtained through the use of polarizers, that make anisotropies visible. This is the topic of the present section.

3.2.1 LC optical anisotropy

A simple liquid is isotropic and its refractive index is independent on the direction of the incident light beam. An incident ray is refracted according to Snell's law of refraction. In contrast, most of the crystalline materials are anisotropic, i.e. their index depends on the direction of light propagation. This property is shared by liquid crystals and reflects the different polarizabilities along the axes of the molecules and the supra-molecular organization.
A uniaxial material or a phase is said to have positive birefringence (or to be optically positive) if the refractive index parallel to the optical axis, $n_{\parallel}$, is larger than that perpendicular to it, $n_{\perp}$, or $\Delta n = n_{\parallel} - n_{\perp} > 0$. The reverse is true for negative birefringent media. This is summarized in the index ellipsoid shown in fig. 3.1, with long and short axes equal to $n_{\parallel}$ and $n_{\perp}$ for a positive medium and viceversa for a negative medium.

Uniaxial phases

What said above is valid for a uniaxial phase, a medium with two principal refractive indices and only one optical axis, given by the normal to a plane whose intersection with the ellipsoids is a circle.
An incoming light beam is split into two components, the ordinary ray and the extraordinary. The two beams propagate in different directions, and their polarizations are perpendicular to each other; they also have different velocities, because they experience different refractive indices, n_o and n_e. This leads to a phase difference δ, given by

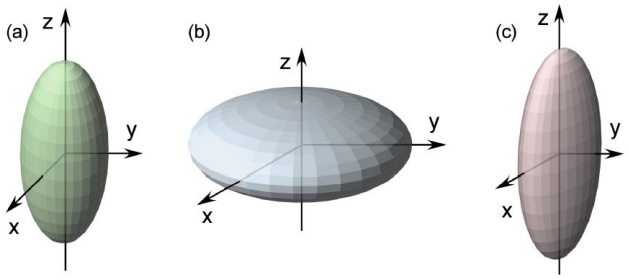

Figure 3.1 - Index ellipsoids (a) uniaxial (positive), (b) uniaxial (negative), and (c) biaxial case; from [3].

$$\delta = \frac{2\pi}{\lambda}(n_e - n_o)d \tag{3.1}$$

where λ is the vacuum wavelength and d is the distance into the medium. The indices n_o and n_e are related to the principal axes by

$$\begin{aligned} n_o &= n_\perp \\ n_e &= \frac{n_\parallel n_\perp}{\sqrt{n_\parallel^2 \cos^2\phi + n_\perp^2 \sin^2\phi}} \end{aligned} \tag{3.2}$$

where ϕ is the angle between the optical axis and the direction of light propagation.
In the general case, the incident linearly polarized light is converted to elliptically polarized light, with a component that can pass through the crossed analyzer and thus the sample appears bright. The transmitted intensity is

$$I = I_0 \sin^2(2\varphi)\sin^2(\delta/2) \tag{3.3}$$

with I_0 the light intensity after the polarizer and φ the azimuthal angle, the angle between the analyzer and the projection of the optical axis onto the sample plane. The first sin factor in 3.3 describes the changes in transmitted light intensity when rotating a birefringent sample between crossed polarizers. The second is responsible for the beautiful colors of the liquid crystal textures, because δ, and thus the transmitted intensity, is different depending on the wavelength.
Let's consider a well aligned uniaxial nematic sample between crossed polarizers. For homeotropic orientation (nematic director perpendicular to the glass plates), the direction of light propagation coincides with the optical axis. In this case $\phi = 0$ and thus we have $n_e = n_o$ and $\delta = 0$. From equation 3.3, we see that the transmitted intensity will always be equal to $I = 0$ for all positions φ of the sample between crossed polarizers. For this reason the uniform homeotropic nematic is called pseudo-isotropic. For uniform planar alignment $\phi = 90°$, $n_o = n_\perp$ and $n_e = n_\parallel$, i.e. $\Delta n = n_\parallel - n_\perp$. The transmitted intensity changes with $\sin^2(2\varphi)$. It has a maximum for $\varphi = 45°$ and a minimum for $\varphi = 0°$ and $\varphi = 90°$. The sample appears dark whenever the optical axis is parallel to one of the polarizer directions. The texture appearance has a periodicity of 90° when rotating the sample through crossed polarizers.

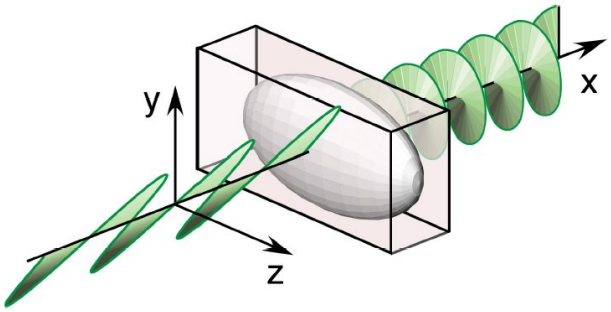

Figure 3.2 - Linearly polarized light becomes elliptically polarized passing through a birefringent medium; from [3].

When a phase, as SmC, doesn't have rotational symmetry around the optical axis, C_∞, the optics becomes more complicated, as we have a biaxial medium. It is characterized by three refractive indices and the index ellipsoid is shown in fig. 3.1. However, in liquid crystal materials the biaxiality is usually very small, in the order of 10^{-3}.

Optical rotation and activity

Chiral liquid crystals exhibit optical activity and optical rotation, i.e. they turn the polarization plane of propagating light. This phenomenon can be also observed in the isotropic phase, due to the single molecules effect. The molecular optical activity of chiral liquid crystals is generally very small and can be usually disregarded, while the optical rotation of the phase, for N^* or SmC^*, can be huge, of thousands of degrees.
We'll go back to light propagation inside an helical medium in the next section, introducing the N^* textures.

3.2.2 Optical textures of LC phases

The optical properties of liquid crystal phases often directly reflect the symmetry of their structures. Birefringence, anisotropy of the refractive index, allows the visualization of the macroscopic molecular orientation. In thin liquid crystal sample cells placed between two crossed polarizers under an optical microscope, various textures and birefringence colors will be observed. These textures and colors not only look beautiful but also provide a lot of information about the structure of the LC phases. Although there are many experimental techniques to investigate the structure and physical properties of LC phases, microscope observations often give enough information to determine the structure even if a well-aligned domain would not be obtained.
Textures of the LC phases of interest in the next chapters are summarized below. Other textures and detailed explanations about phase characterization from polarized microscopy can be found in [4].

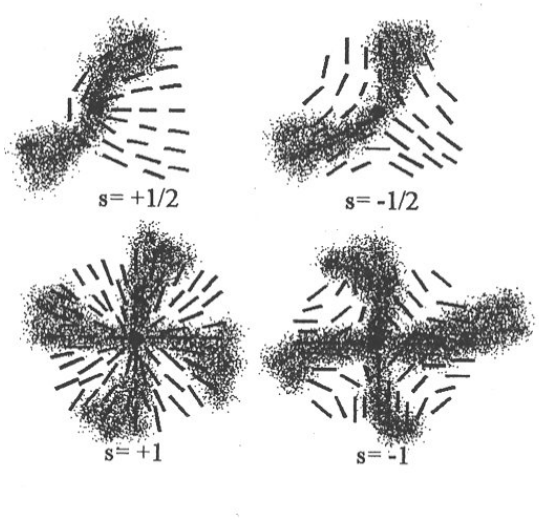

Figure 3.3 - Local orientation of the director in *schlieren* texture, with different strength defects; from [4].

Nematics

Nematics between untreated glass plates often orient with their director parallel to the substrates. If this orientation is not homogeneous, and varies slowly in the plane of the substrate, the so-called *Schlieren* textures appear between crossed polarizers, exhibiting characteristic sets of curved dark brushes, as seen in fig. A.1 (color plate in Appendix A).
These brushes correspond to the extinction position of the nematic director field, with $\mathbf{n}(r)$ coinciding with the direction either of the polarizer or the analyzer. The brushes come together, twofold or fourfold, in a singular point, a topological defect.
Figure 3.3 reports the corresponding director configurations, with different values of the strength of the defects, defined as $s = \pm$ number of brushes/4, with + sign if, rotating the polarizer, the brushes rotate in the same direction and - if the opposite happens.

By subjecting a nematic to uniform planar boundary conditions, we obtain a sample that optically behaves like a uniaxial birefringent plate of a crystal with its optical axis in the plane of substrate. According to equation 3.3, the transmitted light intensity between crossed polarizers varies proportionally to $\sin^2(2\varphi)$: whenever the director is oriented along one of the polarizer directions, the sample appears black, while maximum of intensity is observed when $\mathbf{n}$ is oriented at 45° between P and A.
When, instead, nematic has homeotropic boundary conditions, the director orients perpendicular to the substrates and the direction of light propagation coincides with the optical axis of the phase. Therefore, no birefringence is present and the sample is black between crossed polarizers, for every azimuthal angle. This is called pseudo-isotropic texture. The nematic features become apparent if we perturb the homogeneous ordering of molecules by pushing some part of the cell, or if we simply tilt the cell.

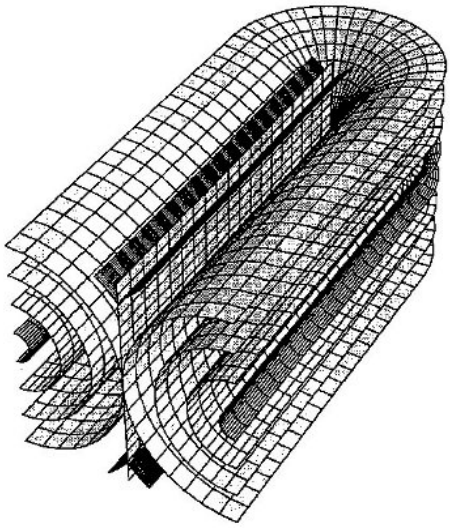

Figure 3.4 - Structure of *oily streaks* texture showing the folding of layers in a defect area; from [5].

Cholesterics

The most commonly observed texture of cholesteric phase (between untreated substrates) is the *oily streaks*, shown in figures A.2 and A.3 (color plates in Appendix A). The long molecular axis is basically parallel to the substrates and the N^* super-helix axis is thus perpendicular to the glass plates. The texture can be seen as a network of defect lines dispersed in uniformly helical regions; this is reconstructed in fig. 3.4, where the surfaces are parallel to the molecules. The defect areas can also show focal conic domains, as in fig. A.2.
Thanks to its helical arrangement of the molecules, the cholesteric phase can exhibit some unique optical properties; the most striking is the selective reflection: circularly polarized light of a specific handedness and wavelength is reflected. When the pitch is in the visible range, as it is the case in fig. A.2, the effect can be seen at naked eye as beautiful iridescent colors (while in fig. A.3 the pitch is above visible range).
If we follow light propagation along the helix axis of a continuously twisted birefringent medium, we can obtain an expression for the optical rotation φ (radians per unit length) as a function of wavelength. Neglecting the molecular optical activity and the interface effects, and assuming that the material is non-absorbing, we have the following simplified expression valid in the regime $\lambda \approx p$:

$$\varphi = \frac{\pi \Delta n^2 p}{4\lambda^2} \frac{1}{1-(\lambda/\lambda_0)^2} \tag{3.4}$$

where Δn is the birefringence of a "layer", p is the cholesteric pitch and $\lambda_0 = \tilde{n}p$ is the wavelength of selective reflection ($\tilde{n}$ is the refractive index of a quasi-nematic layer). Note that for $\lambda = \lambda_0$ eq. 3.4 diverges, and the sign of φ for $\lambda < \lambda_0$ gives the handedness of the phase. We can consider a linearly polarized incident beam as consisting of a left and a right circularly polarized component. At λ_0 the component that doesn't match the spatial handedness of the

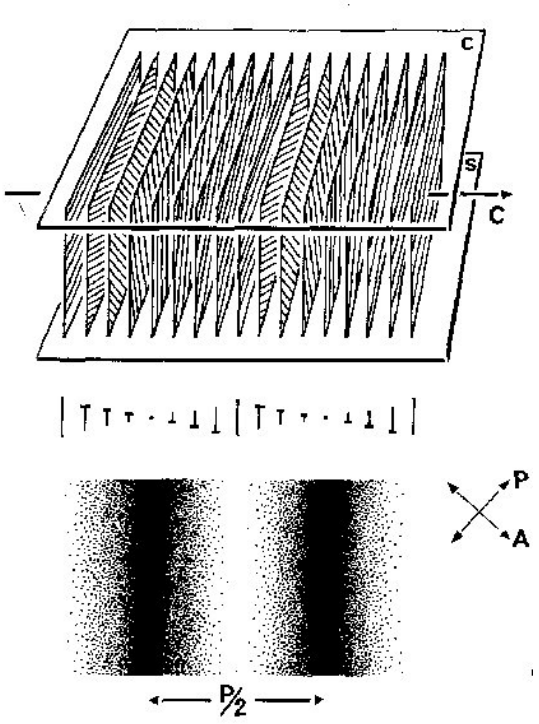

Figure 3.5 - Schematic view of fingerprints texture; from [6].

cholesteric helix is completely reflected without phase change.
When $p \gg \lambda$, the so-called adiabatic limit, the already mentioned optical rotation is instead the dominating effect.

When a cholesteric phase is subject to homeotropic boundary conditions, the helix axis is oriented parallel to the substrate. In general, every in-plane orientation of the helix axis is possible, and thus it will vary slowly over macroscopic distances, similar to the director of the nematic phase under planar boundary conditions. If the pitch is under 1 μm, we cannot microscopically resolve the modulations due to the director twist, and thus the system behaves under crossed polarizers as a uniform planar nematic.
When longer pitch cholesterics are observed with the same boundary conditions, it is possible to see a pattern of dark lines (fig. 3.5), where the director points along the direction of the light, and bright ones; the periodicity of the equidistant stripes is given by $p/2$. For an evident reason, these are called "fingerprints" (figs. A.3 and A.4 in Appendix A).

Columnar phases

Focal conic textures, or developable domains, with numerous disclinations are frequently observed in columnar mesophases [7,8], as shown in figs. A.5 and A.6 (color plates in Appendix A). The texture is planar and molecular orientations rotate by about 180° around $+\pi$ disclination lines which are normal to the

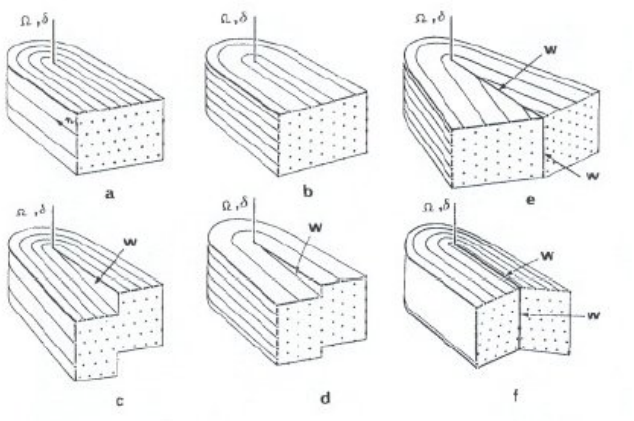

Figure 3.6 - Several example of $+\pi$ disclinations when the defect line is parallel to rotation axis; from [9].

preparation plane. The rotation may occur without formation of a wall but more often walls are associated with these $+\pi$ lines because it is unlikely that the two hexagonal networks coincide exactly after a rotation of 180° (see fig. 3.6). Other walls are also observed separating domains of different molecular orientations. They frequently join to form $-\pi$ disclination lines. A schematic representation of the molecular organizations with defect lines and walls is given in fig. 3.7. The structure is fluid and walls can be distorted and moved by pressing the coverslip.
Developable domains can also be showed by Sm and N^*, and indeed they can be seen in the "borders" of fig. A.2. All of the three phases can be viewed as lamellar structures. In a smectic phase, molecules are normal (or oblique) to the layer plane and each layer is close to the length of the molecule; while we can assign the cholesteric phase a virtual layering corresponding to $p/2$, with molecules parallel to layers. Finally, in a columnar hexagonal phase, the three equivalent directions of the hexagonal network define three series of parallel layers; molecules are lying in the plane of these layers whose thickness is close to the inter-helix spacing. Then, the molecular arrangement of a focal conic, shown in fig. 3.8 for the layers of a smectic phase, also apply to "bundles" of columns in hexagonal columnar phase.

Undulating textures are also possible (see fig. A.7 in Appendix A), due to the undulation of the molecular columns of long helical molecules such as DNA. This process produces multiple textures depending on the orientation of the plane of undulation of molecules with respect to the preparation plane. Some possible arrangements, when disclination lines are parallel or normal to the substrates, are shown in fig. 3.9. In order to simplify the drawing, layers of molecules have been shown instead of molecules themselves. Note that one layer corresponds to one series of molecules along one of the three main directions of the hexagonal network.
Undulations come from rotations of molecular orientations of less than 180° around disclination lines (δ); if these are normal to the preparation plane, molecular orientations can follow sinusoidal paths in projection onto the observation

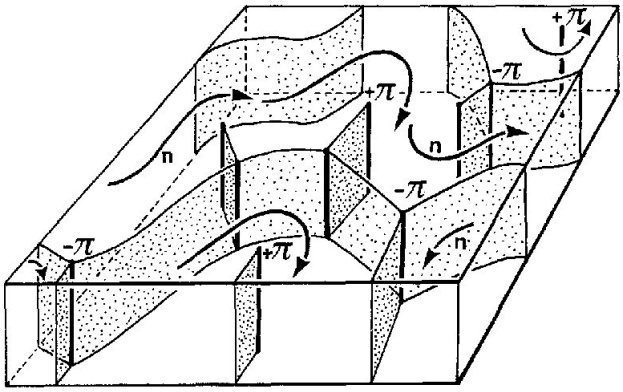

Figure 3.7 - Hexagonal columnar phase structure, with inversion walls; from [6].

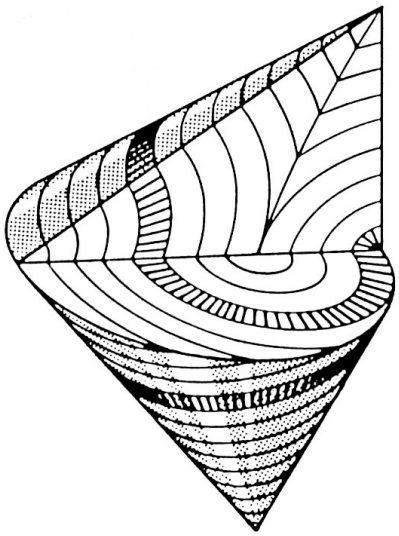

Figure 3.8 - Arrangement of the molecular planes in (chiral nematic, columnar or smectic) focal conics; from [10].

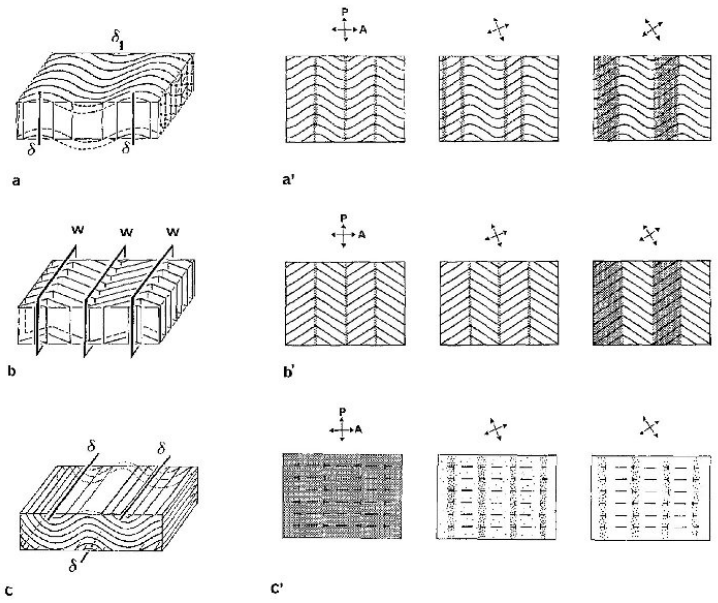

Figure 3.9 - Arrangement of molecular planes corresponding to undulating textures; from [9].

plane, leading between crossed polarizers to striated patterns which turn to banded patterns when the stage is rotated. Or else the regular curvature of the layers can disappear and discontinuity walls are created. Striated and banding patterns are still observed but the position of the extinguished bands does not move continuously anymore when rotating the stage. Disclination lines may also be lying in the preparation plane. The preparations are then either fully extinguished when the molecular directions are parallel to one of the polar directions, or banded for other orientations of the sample. In this third situation, the banding patterns are due to variations in the inclination of molecular orientations with respect to the microscope axis and not to variations in their orientations with respect to polarizer directions.
Another texture exhibited by columnar phases is the mosaic texture (fig. A.8 in Appendix A). It is commonly observed in soft crystals, evolution at higher concentrations of Sm hexatic phases, but with long range positional order. Macroscopic domains have uniform orientation of the optical axis, and thus homogeneous colors, while discontinuous changes are encountered from grain to grain. Columnar phases forming from isotropic melt often exhibit also dendritic-like, homeotropic growing domains (fig. A.9 in Appendix A); these directly reflect the sixfold symmetry of the underlying hexagonal packed molecule piles [7].

3.2.3 Microscopy procedures

All the optical observations were performed through depolarized transmitted and reflected light microscopy (PTOM, PROM) on a polarized optical microscope (Nikon Eclipse E400 POL and Nikon TE200 inverted) as a function of cell temperature, controlled by a heat stage (Instec, STC200D). Images were

acquired on a Nikon DS-5M or Olympus C5050 camera. Fluorescence optical microscopy (FOM) was performed on the TE200 microscope using filter sets (Chroma, VT, USA) appropriate to the fluorophores.
After loading the sample, typically $\sim 0.1-0.3$ mg of DNA in a few μl of distilled water, into the channel, cell was closed using epoxy glue to slow down the water evaporation rate. When necessary, only one end of the channel was sealed, so that water evaporating slower from one end of the channel induced gradients of DNA concentration. Samples were thermally cycled several times to enhance the concentration gradient, which could be made quite smooth.

To make the "contact" preparation of sDNA (10bp) and lDNA S3h (chapter 4), two small droplets of each sample on a glass slide were covered by a cover slip, gently brought in touch and left diffusing one into the other for about one hour. Water was subsequently slowly dried from the edge of the cover slip until liquid crystal phases appeared.

Experiments on nucleation and growth of LC domains (chapters 5 and 6) were performed on samples that had been kept at $T = 70°C > T_m$ for tens of minutes, so to ensure spatial uniformity of the concentration, and then quenched to the desired temperature, close to or well below T_m. As the samples cool, LC domains appear and gradually grow in area to fill a fraction of the cell that ultimately saturate. Depending of the quenching depth and on the time left to the sample to equilibrate, LC domains grow fewer and bigger, or more and smaller. Image analysis was performed with ImageJ and Matlab softwares on black and white 2560x1920 pixel images, thresholded to exclude camera noise.

Determination of diffusion coefficient

An independent estimate of the diffusion coefficient of a short duplex in a *ss-DNA* pool was obtained by following the diffusion of A sequence (GGAGTTTTGAGG) in a 1:10 mixture with B sequences (CCTCAAAACTCC) after a contact with a sample of only B sequences at the same concentration. About 2.5% of A was tagged at 5' with Fluorescein. The B solution was introduced by capillarity to fill completely a 2 μm-thick homemade cell, and the A solution was put in contact with it on the edge of the cell. The fluorescence signal and its time evolution inside the cell were monitored with fluorescence microscopy imaging. Given the unbalanced concentration of the two strands we can assume that almost all the A sequences are in the form of helices, thus the measured fluorescence intensity is proportional to the concentration $c(x,t)$ of duplexes. The tail of the fluorescence profile (far enough from the edge to exclude evaporation effects) at different times is fitted by a solution of the diffusion equation for two semi-infinite bodies with a "step" initial condition:

$$c(x,t) = \frac{c_0}{2}\left[1 - erf\left(\frac{x}{\sqrt{4Dt}}\right)\right] \tag{3.5}$$

where c_0 is the starting concentration of A. From the fitting the diffusion coefficient D is obtained. Two fluorescence profiles are reported in figure 3.10 at different times few minutes after the contact, with the corresponding fits.

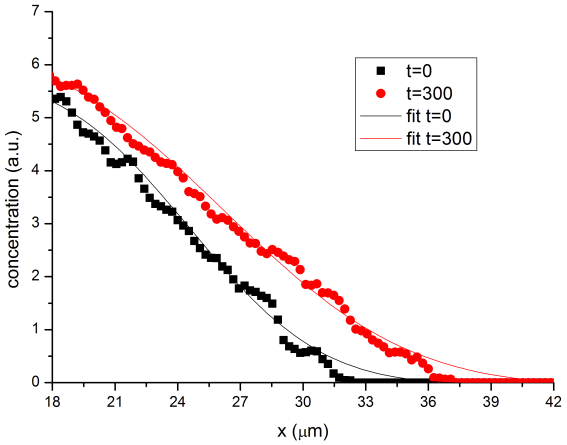

Figure 3.10 - Concentration profiles, obtained from fluorescence microscopy images, of A strand labeled with Fluorescein diffusing in a B strands pool. Two profiles are shown at different times, with the fit lines (eq. 3.5).

3.3 Helix denaturation measurements

The relationship of the LC phase transitions to the duplex denaturation transition was probed by simultaneously monitoring polarized transmitted light and measuring the fluorescence emission of Ethidium Bromide [11] mixed into the sDNA solution at a concentration of about one molecule per duplex. Upon increasing T, as the sDNA duplexes unbind, the Ethiudium Bromide fluorescence is progressively reduced, enabling the determination of the fraction of paired strands. As discussed in section 1.2, in our conditions we can assume a linear correspondence between the fluorescence signal and the fraction of bound helices. Therefore, after subtracting the background noise (mainly due to the residual fluorescence of unbound Ethidium and camera noise), we can obtain the helices fraction α and the melting temperature, corresponding to $\alpha = 1/2$. In figure 3.11, we report an example of melting curve. We can see that, while reaching a plateau at the lowest temperatures, corresponding to complete binding of the helices, at the highest temperatures investigated the curve doesn't reach a stable plateau. The high T region is difficult to explore because of thermal expansion and convective flows inside the cell, and formation of steam. The resulting contrast is lower (never more than 5), and the curves are slightly broader, than in usual experimental curves reported in literature [12]. Nonetheless, it was possible to analyze the curves according to the model of section 1.2 to extract the melting temperature, with an estimated error of 5 degrees.

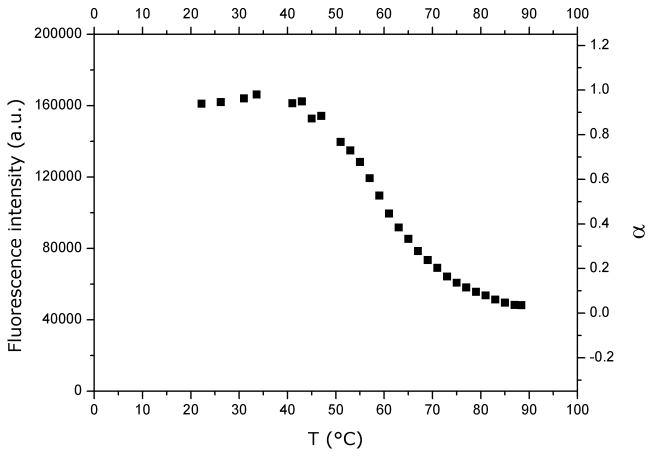

Figure 3.11 - Curve obtained in a typical melting temperature experiment, from a sDNA 16 bp sample in C_U phase. The fraction of formed duplexes α - right axis - is extracted from Ethidium Bromide fluorescence intensity - left axis - as explained in the text.

3.4 Concentration measurements

Concentration of sDNA solutions was measured through the determination of their refractive index n, which in turn was obtained through optical reflection interferometry (ORI) measurements, thanks to the so-called Fabry-Perot interference.

3.4.1 Multiple interference

Let's consider a layer of thickness d of a medium (our DNA solution) with refractive index n_2, sandwiched between media of index n_1 (the glass windows); we can forget about the air-glass interfaces because the thickness of glass windows is much bigger than d (1 mm versus a few micrometers) and we collect light from the thin slab only.
The transmittance and the reflectance of light through this multi-layer can be calculated as [13]:

$$T = \frac{1}{1 + F\sin^2(\delta/2)} \tag{3.6}$$

$$R = 1 - T = \frac{F\sin^2(\delta/2)}{1 + F\sin^2(\delta/2)} \tag{3.7}$$

where $F = 4R_1/(1-R_1)^2$ is called contrast, or *finesse*, of the fringe pattern; R_1 is the reflectance at the first 1-2 interface. The phase difference δ is given by

$$\delta = \frac{4\pi n_2 d}{\lambda}\cos\theta \tag{3.8}$$

with θ the angle of incidence of light and λ the light wavelength. δ doesn't depend on n_1 and it depends on the optical path, the product of the refractive index and the thickness of the slab. Thus, if we know one of the two, we can obtain the other one. In fact, we observe that T and R have their minima when δ is respectively $\pi \pm 2\pi j$ and $0 \pm 2\pi j$ with j integer number, thus we can fit the minima positions (in wavelength) in the intensity curves and extract the $n_2 d$ product:

$$\lambda_m = \frac{4n_2 d \cos\theta}{m} \tag{3.9}$$

where m assumes progressive integer values [14].

In the above discussion, we have considered the refractive index to be constant in the range 450-750 nm, but actually there is a weak wavelength dependence: n can be expressed by the empirical equation [15]

$$n(\lambda) = \sqrt{A + \frac{B}{1 - \frac{C}{\lambda^2}}} \tag{3.10}$$

with sample-dependent parameters (in [15] the reported parameters show a 1% variation among various samples). Anyway, in the relatively narrow range of our measurements, using a constant value for n without taking into account this dispersion introduces an error of less than 1%, smaller than other sources in our experimental conditions.

3.4.2 Procedure

Spectra were recorded in the interval 450-750 nm, where the DNA absorption spectrum is almost flat, either at 45° or at normal incidence. In the latter case the spectrometer (Oceanoptics S200 UV-VIS) was coupled to the microscope through a homemade optical system. Measures were usually taken at room temperature; the light spot (around 0.5 mm^2, big enough to average over many differently oriented domains) was focused on the cell at various locations and the obtained spectra were analyzed to extract the optical path within the cell through the spacing of Fabry-Perot fringes. To enhance the contrast of the fringes, sandwich cells were fabricated, as described in par. 3.1.3, with high n glass; before introducing by capillarity the DNA solution, the thickness of every cell was carefully measured along the whole channel, using both air and water as inner media of known refractive indexes. Some recorded spectra with different media (air, water and DNA solution) filling the channel are showed in figure 3.12.
To verify the independence of the measured refractive index from the arrangement of molecules, we also performed some measures at different temperatures, up to the LC-Iso transition. By taking into account the effect of temperature on cell thickness and water index, substantial agreement was observed. The measured refractive indexes are also similar to values reported in literature [16, 15, 17].

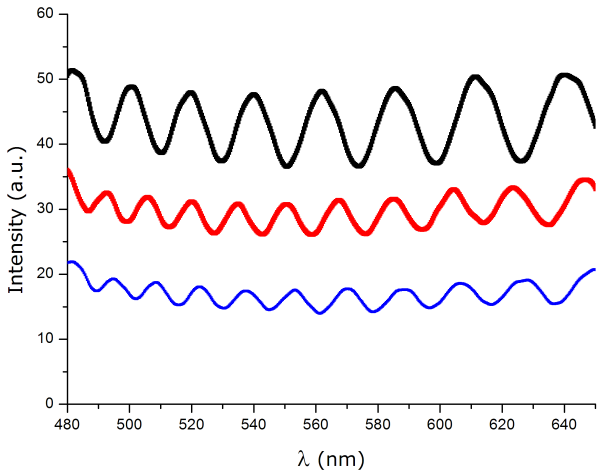

Figure 3.12 - Example of interference fringes observed in light reflected by cell channel filled with air (top, black), water (middle, red) and sDNA 16 bp solution (bottom, blue); the intensities of top and bottom curves have been shifted for clarity.

3.4.3 Concentration-refractive index relation

A critical aspect of this kind of measurement is the connection between the measured refractive index and the quantity of interest, the concentration of DNA solution; this is usually expressed with the quantity dn/dc:

$$n = n_{H_2O} + \frac{dn}{dc}c \tag{3.11}$$

Due to the small amounts of samples, we couldn't measure dn/dc directly, and we had to rely on reported experiments, usually performed in more dilute solutions and for DNA chains longer than ours. To what extent dn/dc is constant (i.e. equation 3.11 is linear in concentration) is an open question in literature. Because of these discrepancies, we chose to use the average value, independent on concentration, of 0.175 cm^3/g [18, 19, 20], and the commonly accepted value of 1.8 g/cm^3 for DNA density [21, 22, 23, 24, 25].

We estimate the overall experimental error for this optical method to be within 10% (a variation of 0.01 in the refractive index can lead to an error of 50 mg/ml in mass concentration). In chap. 4, we'll compare the results obtained with this method with those from X-ray measures, finding a pretty good agreement, within estimated error.

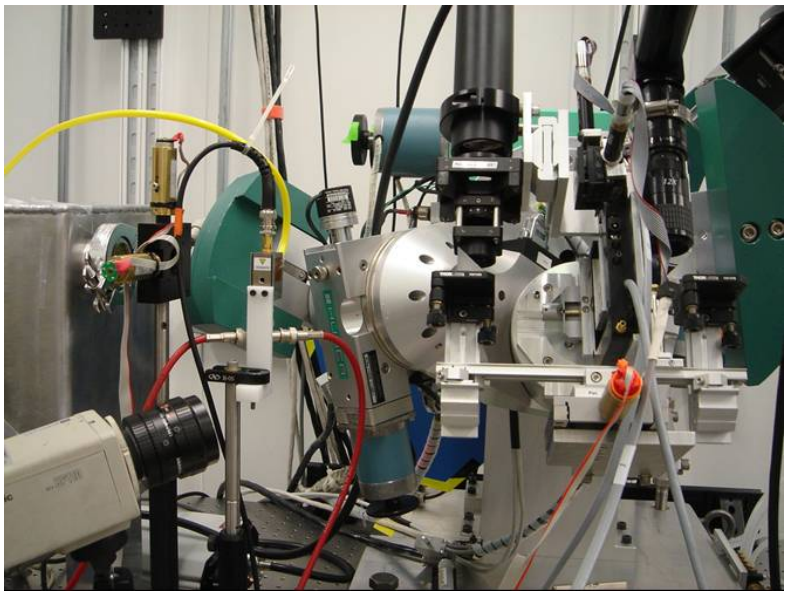

Figure 3.13 - Picture of the μ-XRD apparatus at APS.

Characteristic length	Origin
$\sim 3.4\mathring{A}$	1 bp thickness
$\sim N \cdot 3.4\mathring{A}$	Oligomer length
$\sim 13\mathring{A}$	Minor groove
$\sim 19\mathring{A}$	Major groove
$\sim 20 - 24\mathring{A}$	B-DNA diameter
$\sim 34\mathring{A}$	Helical pitch
$\sim 24 - 30\mathring{A}$	Inter-helix spacing

Table 3.1 - Typical length scales for DNA double helices.

3.5 X-ray diffraction

To investigate the structure of the liquid crystal phases, X-ray diffraction (XRD) measurements were performed with a synchrotron X-ray micro-beam (20 Bend Magnet Beamline at the Advanced Photon Source, Argonne National Laboratory,USA), in conjunction with an *in-situ* polarizing optical microscope to collect diffraction patterns from selected areas. The apparatus was developed by Brandon Chapman and prof. Ronald Pindak from Brookhaven National Lab and Julie Cross from Argonne Photon Source [26].
The beam was micro-focused by Kirkpatrick-Baez mirrors, its cross section at the sample was approximately 14 $\mu m \times 14\ \mu m$, smaller than the average uniaxial columnar domain size. An X-ray energy of 16 keV ($\sim 0.77\mathring{A}$; flux: $\approx 3 \cdot 10^7$ photons/sec) was used in order to minimize absorption by the 50 μm thick glass slides confining the sample. The total glass thickness (comprised the hot-stage apertures) was around 300 μm, with an overall transmission of 75%.
X-ray scattering from the sample was detected by a 2-dimensional area detector Bruker AXS CCD and a polarizing microscope (Navitar lens, Pulnix CCD camera), with $3.5\mu m$ resolution, was coupled to the system, as shown in figures 3.13 and 3.14.
In table 3.1 we show the expected periodicities in sDNA, corresponding, in our experimental setup, to a range of scattering vectors between 0.18 and $1.8\mathring{A}^{-1}$. The resolution of the apparatus was estimated to be about $5 \cdot 10^{-3}\mathring{A}^{-1}$.
The system was calibrated with a 8CB sample: assuming the Sm layer thickness to be $d \sim 31.4\mathring{A}$, $q = 0.2\mathring{A}^{-1}$, the sample-CCD distance was calculated from the measured diffraction spots.
To compare XRD results with optical measurements, we extracted an estimate of the concentration. Since in C_U phase we deal with locally parallel, hexagonally packed columns, and we can assume the distance between two adjacent, stacked duplexes inside a column to be similar to that inside a continuous helix, the concentration of helices whose spacing (center-to-center distance) is h can be obtained as:

$$c = \rho_{DNA} * \frac{\pi}{\sqrt{12}} \left(\frac{D}{2h}\right)^2 \tag{3.12}$$

where D is the diameter of DNA helix.

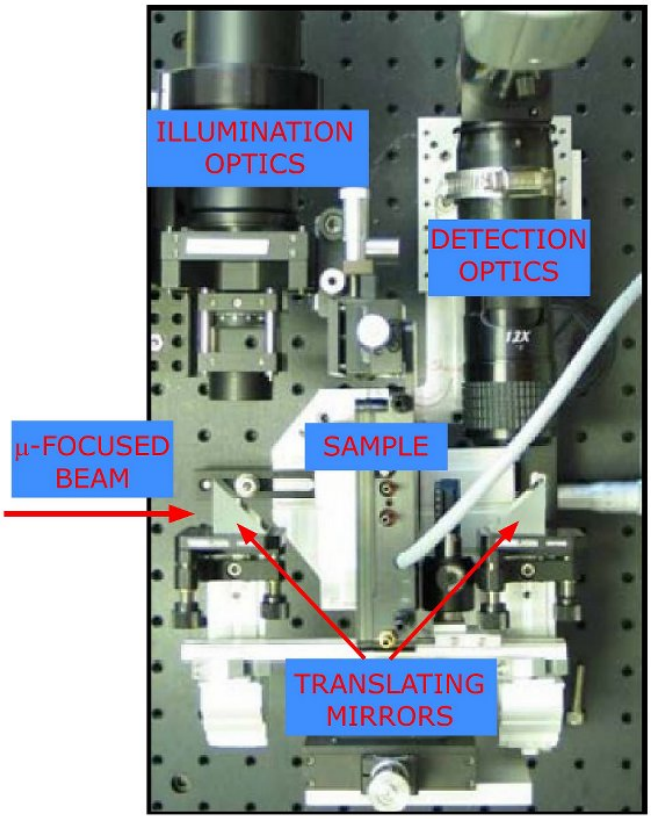

Figure 3.14 - Side view of the optical microscope used in conjunction with X-ray diffraction apparatus. The main parts are indicated; adapted from [26].

Bibliography

[1] M. Abramowitz, *Microscope: Basics and Beyond*, Olympus America, Inc. (Revised Edition 2003)

[2] *http* : *//microscopy.fsu.edu*

[3] M. Nakata, *PhD Thesis*, Tokyo Institute of Technology (2004)

[4] I. Dierking, *Textures of Liquid Crystals*, Wiley-VCH Verlag (2003)

[5] O. D. Lavrentovich, M. Kleman, in *Chirality in Liquid Crystals*, H. S. Kitzerow, C. Bahr, Eds. (Springer-Verlag,2001), chap.5

[6] F. Livolant, A. Leforestier, *Condensed phases of DNA: structures and phase transitions*, Prog. Polym. Sci. **21**,1115 (1996)

[7] Y. Bouligand, *Defects and textures of hexagonal discotics*, J. Physique **41**,1307 (1980)

[8] P. Oswald, M. Kléman, *Défauts dans une mésophase hexagonale discotique: disinclinaisons et parois*, J. Physique **42**,1461 (1981)

[9] F. Livolant, Y. Bouligand, *Liquid crystalline phases given by helical biological polymers (DNA, PBLG and xanthan). Columnar textures*, J. Physique France **47**,1813 (1986)

[10] G. W. Gray, J. W. G. Goodby, *Smectic Liquid Crystals, Textures and Structures*, Leonard Hill (1984)

[11] J. B. LePecq, C. Paoletti, *A flourescent comple ethidium bromide and nucleic acids: physical-chemical characterization*, J. Mol. Biol. **27**,87 (1967)

[12] R. Owczarzy, *PhD Thesis*, University of Illinois at Chicago (1999)

[13] M. V. Klein, T. E. Furtak, *Optics - 2nd edition*, Wiley (1986)

[14] R. Swanepoel, *Determining refractive index and thickness of thin films from wavelength measurements only*, J. Opt. Soc. Am. A **2**,1339 (1985)

[15] A. Samoc, A. Miniewicz, M. Samoc, J. G. Grote, *Refractive-Index Anisotropy and Optical Dispersion in Films of Deoxyribonucleic Acid*, J. Appl. Polym. Sci. **105**,236 (2007)

[16] A. Kagemoto, M. Nakazaki, S. Kimura, Y. Momohara, K. Ueno, Y. Baba, *Calorimetric investigation on liquid crystals of deoxyribonucleic acid in concentrated solutions*, Termochim. Acta **284**,309 (1996)

[17] R. Oldenbourg, T. Ruiz, *Birefringence of macromolecules. Wiener's theory revisited, with applications to DNA and tobacco mosaic virus*, Biophys. J. **56**,195 (1989)

[18] D. Jolly, H. Eisenberg, *Photon correlation spectroscopy, total intensity light scattering with laser radiation and hydrodynamic studies of a well fractioned DNA sample*, Biopolymers **15**,61 (1976)

[19] T. Nicolai, M. Mandel, *Ionic Strength Dependence of the Second Virial Coefficient of Low Molar Mass DNA Fragments in Aqueous Solutions*, Macromolecules **22**,438 (1989)

[20] P. Wissenburg, T. Odijk, P. Cirkel, M. Mandel *Multimolecular Aggregation in Concentrated Isotropic Solutions of Mononucleosomal DNA in 1 M Sodium Chloride*, Macromolecules **27**,306 (1994)

[21] T. Nicolai, M. Mandel, *Dynamic Light Scattering by Aqueous Solutions of Low Molar Mass DNA Fragments in the Presence of NaCl*, Macromolecules **22**,2348 (1989)

[22] H. T. Goinga, R. Pecora, *Dynamics of Low Molecular Weight DNA Fragments in Dilute and Semidilute Solutions*, Macromolecules **24**,6128 (1991)

[23] M. E. Ferrari, V. A. Bloomfield, *Dynamics of Low Molecular Weight DNA Fragments in Dilute and Semidilute Solutions*, Macromolecules **25**,5266 (1992)

[24] D. M. Hatters, L. Wilson, B. W. Atcliffe, T. D. Mulhern, N. Guzzo-Pernell, G. J. Howlett, *Sedimentation Analysis of Novel DNA Structures Formed by Homo-Oligonucleotides*, Biophys. J. **81**,371 (2001)

[25] H. Durchschlag, in *Thermodynamic Data for Biochemistry and Biotechnology*, H. J. Hinz, Ed. (Springer-Verlag,1986)

[26] B. D. Chapman, R. Pindak, J. O. Cross, *Using In-Situ Optical Microscopy and Micro-Focus X-ray Diffraction to Probe Textured Liquid-Crystal Samples*, PNC/XOR-CAT & NSLS poster.

CHAPTER 4

Liquid crystalline phases of short DNA

4.1 Introduction

The ability of duplex DNA to form liquid crystal (LC) phases when hydrated has been known since the late 1940s and played a crucial role in the decipherment of its structure, enabling alignment of the DNA chains and measurement of the X-ray structure factor of a single chain uncomplicated by inter-chain correlations [1, 2].
Since that time the LC phases of solutions of duplex B-DNA have been extensively characterized by optical [3, 4, 5, 6, 7, 8], X-ray [9, 10], and magnetic resonance [11, 12] methods for chain lengths N ranging from 10^7 base pairs (bp) semiflexible polymers down to N $\sim$ 100 bp rigid rod-like segments, comparable in size to the B-DNA bend persistence length, $L_p \approx 50$ nm (see section 1.3.1). These studies of long DNA (lDNA) have revealed isotropic (I); chiral nematic LC (N); uniaxial columnar LC (C_U) (or line hexatic [10]); higher-ordered columnar LC (C_2); and crystal (X) phases vs. increasing DNA concentration.
The appearance of such LC phases has been accounted for theoretically by modeling B-DNA as a repulsive rigid or semi-flexible rod-shaped solute. The basic model is Onsager's treatment of monodisperse repulsive hard rods (length L, diameter D) [13], which, if they are sufficiently anisotropic in shape, nematic order for volume fraction $\phi > \phi_{IN} = 4D/L \approx 24/N$ (D $\sim$ 2 nm, L $\sim$ N/3 nm for B-DNA). The complete computer simulated phase diagram for hard rods by Bolhuis and Frenkel [14] quantitatively confirms this prediction for L/D $>$ 4.7 (N $>$ 28 bp), and also shows that for L/D $<$ 4.7 there should be no LC phases at any ϕ.
We were therefore surprised to find nematic LC ordering in short B-DNA duplexes (sDNA) of length $6bp < N < 20bp$, with the nematic phase appearing for N values an order of magnitude smaller than those predicted from ϕ_{IN}, precluding ordering by the Onsager-Bolhius-Frenkel (OBF) criterion. A similar conclusion was reached by Alam and Drobny in attempting to account for a

magnetically reorientable, orientationally ordered phase in NMR studies of a B-DNA dodecamer [12].
Additionally, we have observed columnar LC ordering for these oligomers, which is also remarkable since in the hard rod models [14], the only translationally ordered LC phase appearing is the lamellar SmA.
In this chapter, we show that the observation of the nematic and columnar LC phases provides clear evidence for end-to-end stacking of the sDNA into rod-shaped aggregates. Such interaction, sketched in figure A.10 (color plate in Appendix A), can have a substantial effect on sDNA organization and phase behavior. While the individual duplexes are not anisotropic enough in steric shape to produce LC phases, the hydrophobic ends cause the formation of much more anisotropic assemblies that can orientationally and positionally order.

4.2 Textures and phase identification

We studied the series of self-complementary sDNA duplex forming "palindromic" oligomers whose sequences are shown in figures A.11 and A.12 (color plates in Appendix A), along with a variety of non-complementary and partially complementary oligomers, discussed later in section 4.4.2. sDNA solutions in 4-8 μm gaps between glass plates were observed by: (i) polarized transmission optical microscopy (PTOM) to probe optical textures; (ii) optical reflection interferometry (ORI) to measure refractive indices and thus DNA concentration c (mg solute/ml solution); and (iii) synchrotron microbeam X-ray diffraction (XRD) to probe local molecular organization, as described in chapter 3. In spite of the challenges presented by the extremely small sDNA sample quantities available, these techniques provided unambiguous evidence for chiral nematic (N) and uniaxial columnar (C_U) liquid crystal phases in the sDNA solutions. At higher concentration, more ordered columnar (C_2) and crystal-like (X) phases were also found.
Figures A.11 and A.12 show PTOM images of the typical textures for the whole range of oligomers length. The phase behavior observed for the sDNAs will be discussed in detail and compared with long DNA in the following sections. We start by the identification of the various phases, with a description of the observed textures (some other, bigger pictures of sDNA LC phases can be found in chapter 3).

4.2.1 N phase

Since the DNAs are chiral molecules the N phase exhibits macroscopic chirality in the form of a helical precession in space of the local optical axis (the local DNA molecular helix axis). This macroscopic helix is of pitch p, the distance required for a 2π optical axis reorientation. The N phase appears with its chiral helix axis z either parallel (N_{PAR}) or normal (N_{NOR}) to the plates, forming either the focal conic (if $p < 1\mu m$) or fingerprint (if $p > 1\mu m$) textures in the PAR case; and forming the Grandjean texture [15] in the NOR case.
The cholesteric pitch was observed to increase with sDNA concentration, as beautifully displayed in figure A.13 (by M.Nakata, see color plate in Appendix A), and with T, approaching the I-N transition. This dependence corresponds

to the behavior of lDNA [16,8,17]. Also, as is clear from the selective reflection colors in the visible exhibited by the Grandjean textures, e.g., for N = 8 bp and 10 bp in figure A.12 and in figure A.13, the pitch was found for several of the sDNA nematics to be $p \sim 0.3\ \mu m$, considerably smaller than the $p \sim 2.5\ \mu m$ typically observed in $N \geq 147$ bp DNA and explained by the current models [18,19].
Birefringence measurements on the parallel textures using a Berek compensator enabled determination of Δn, the birefringence of the N phase, revealing that the optical principal axis parallel to the helix axis has the higher refractive index, as is also the case for the N phase in $N \geq 146$ bp DNA [5], i.e., that the nematic helix axis is parallel to the planes of the base pairs of the sDNA duplexes and normal to their double helix axes.

Contact experiment

The PTOM observations clearly identify a N phase in the sDNA solutions but leave open the question of its relation to the N phase observed in lDNA. An approach widely used to identify unknown but possibly related LC phases is to fabricate a "contact" cell in which a quasi-linear concentration gradient in composition between the two compounds with the phases in question is established. If in PTOM observation a texture spans the concentration range without interruption by another phase, then the structural identity of the phases is established [15]. To compare the sDNA LC phases to those in lDNA, we prepared a contact cell with two distinct concentration gradients along orthogonal axes: a gradient between N = 10 bp sDNA and $N \sim 900\ bp$ DNA (S3h) along one axis, and a gradient in overall DNA concentration in the normal direction, as shown in figure A.14 (Appendix A). The N phase (bordering the I phase) exhibits on the N = 10 bp end, "focal-conic" N_{PAR} and selective reflecting N_{NOR} textures, characteristic of a nematic with $p \sim 600$ nm. These textures evolve in a continuous way to the fingerprint texture of the N = 900 bp end, indicating a single phase with p gradually increasing up to $p \sim 10\ \mu m$ for N = 900 bp. From this experiment, we conclude that the symmetry and structure of the sDNA N phase is the same as the N phase of lDNA: a helical winding along a helix axis ($\mathbf{z}$) of a nematic director $\mathbf{n}(\mathbf{z})$ (in this case the optical polarization direction of low refractive index) in which $\mathbf{n}(\mathbf{z})$ is everywhere normal to $\mathbf{z}$, and where $\mathbf{n}(\mathbf{z})$ gives the local molecular orientation of the DNA double helix axis.

4.2.2 C_U phase

At higher sDNA concentration the uniaxial columnar phase (C_U) grows from the isotropic upon cooling as a texture of two kinds of areas: developable domains [20], as seen in figures A.11 and A.12, having birefringence substantially larger than that of the N phase ($n_{C_U} \simeq 2n_N$); and domains that are apparently isotropic, i.e., that have no in-plane birefringence (see also fig. A.15).
Developable domains are clear indications of either fluid lamellar smectic or fluid columnar order with respectively the layer normal or column axis parallel to the glass plane. Study of the developable domain birefringence with Berek compensator showed that the high refractive index direction is radial where the domains are circular, indicating that the DNA base pair planes are also radial

and thus that the DNA double helix axes lie parallel to circles about the domain center (white rings in fig. A.15). Following the arguments of Livolant formulated for lDNA [5], the structure is therefore columnar, with the double helix axes parallel to the columns, and in the birefringent domains the column axis is parallel to the glass plane ($C_{U\ PAR}$). The apparently isotropic domains have instead the column axes normal to the glass plane: they become birefringent upon tilting the sample, indicating that they are optically uniaxial with the column axis normal to the glass ($C_{U\ NOR}$). The PTOM observations indicate unambiguously that the C_U phase is a uniaxial fluid columnar liquid crystal. This is also confirmed by measuring the specific birefringence for the C_U phase at the N boundary ($n \simeq 0.025$ at 350 mg/ml), which yields values in close agreement with those measured for the lDNA LC phases, once properly rescaled for concentration, thus supporting the notion that the packing of the sDNA phase is strictly analogous to that of the lDNA LC phases. This conclusion is also supported by the contact preparation of fig. A.14, wherein the C_U phase is also continuous along the 10-900 bp concentration gradient.

X-ray diffraction

Further structural characterization of the C_U phase was carried out using simultaneous PTOM and synchrotron-based microbeam X-ray diffraction. In these experiments the $14\mu m \times 14\mu m$ X-ray beam size was smaller than the texture domain size, so that area detection enabled efficient collection of the diffraction patterns of single $14\mu m \times 14\mu m \times 6\mu m$ thick sDNA LC domains in a PTOM texture. Figure A.15 (Appendix A) summarizes the key results, showing diffraction from the principal C_U phase orientations: birefringent (C_{U-PAR}) and isotropic (C_{U-NOR}), confirming the nature of this phase as a hexagonal packing of uniaxial columns which lacks intracolumn positional correlation along the columns. These diffraction spots are resolution limited, as expected for the scattering from a C_U single domain [21], and for this reason we tend to exclude hexatic phase, whose peaks have potentially resolvable width, as was observed in lDNA [10]. The 2.8 nm [10] lattice spacing for 16bp C_U (figure A.15) corresponds to an inter-helix spacing of 3.23 nm. Thus, through equation 3.12, we obtain a volume fraction of 0.5 and a mass concentration of 900 mg/ml, in good agreement with optical measurements (figure A.17).

4.2.3 Higher concentration phases

The next higher concentration phase is another fluid liquid crystal phase (C_2) which forms dendrite or tree-branch shaped domains of rectangular morphology and higher birefringence than the C_U phase, a result of the higher sDNA concentration. Since DNA is a linear polymer it must be some kind of columnar phase but the texture exhibits two distinct birefringence colors indicative of lower symmetry than the C_U. This is possibly the orthorhombic phase identified in N = 146 bp DNA [22], although we lack evidences from XRD.
At yet higher concentration a solid-like phase (X) of lower birefringence appears, which might be either a crystal or an amorphous glassy phase. This phase sequence was observed for all of the 6 bp $<$ N $<$ 20 bp palindromic sDNAs, with the exception that the C_2 phase does not show up for the N = 6 bp oligomer.

Our guess is that at the high concentration required for the 6 bp duplex, the interactions of the helices in adjacent chains may induce crystallization instead.

At the concentration necessary for ordering into LC phases (300 mg/ml < c < 1400 mg/ml), the estimated Na^+ counterion molarity m_c is in the range $2M < m_c < 10M$. Thus, although prepared in deionized water, the LC phases are obtained at effectively very large salt concentration. Addition of 0.2 M NaCl or 0.1 M $MgCl_2$ produced no apparent variation of the phase behavior or stability. Also, the usual experiments were done with unbuffered solutions of milliQ water (somewhat acid), but experiments with MES pH 7 buffer showed the same phases with similar stability.

4.3 Comparison of phase diagrams of lDNA and sDNA

Observation of the local optical texture and ORI measurement of the local concentration of the palindromic oligomers (6 < N < 20) enabled construction of the N-c phase diagram of figure A.16, which combines phase boundaries measured for sDNA with those obtained from the literature for long DNA (figure A.17), along with the predictions from the Onsager and other models of interacting semiflexible rod-shaped particles and aggregate solutes (figure A.18). For sake of clarity, we display both the separate and unified phase diagrams in Appendix A.

As noted above, LC phases are found in sDNA under conditions where strictly repulsive hard rods of similar steric shape would be expected to be isotropic. Here we discuss the observed phase behavior of both lDNA and sDNA, plotted in figure A.16, in the broader context of current models of LC formation in solutions of rod-shaped objects, including effects of flexibility, aggregation, and orientation (see section 2.2.4). In the figure, experimental variables c (DNA concentration c in mg solute/ml solution) and N are related to theoretical variables L and ϕ by considering the B-DNA double strands to be cylinders of mass density DNA $= 1.8 g/cm^3$ [23], of an effective diameter $D \approx 4.0$ nm, chosen to fit the $I - N$ and $N - C_U$ phase boundaries measured for lDNA to models [7], as will be discussed below.

The lDNA I-N phase boundary measured for 147 bp < N < 8000 bp [7], and for N = 100 [11] is shown in fig. A.16-A by the red triangles, the two lines bounding the coexistence range. The phase boundary shifts up in c with decreasing L for $L \lesssim L_p \approx 50\ nm$, in reasonable agreement with the simple Onsager rigid-rod limit (OBF line in figure A.16), if the effective double helix diameter is taken to be $D_{eff} = 4.0\ nm$ to account for the electrostatic repulsion between chains [7]. Thus, while the lDNA I-N phase boundary can be interpreted in terms of an electrostatically swollen DNA diameter, it is impossible to do so for the sDNA I-N data on the basis of shape factors alone since any reasonable diameter yields axial ratios where there are no LC phases in any of the models.

The experimental $N - C_U$ phase boundary for lDNA (dotted red line and solid circles) can be obtained from the data of Rill, (N = 147 bp [6]) and Podgornik (N = 150 Kbp [24]). We note that for N > 147 bp the concentration c_{NCU} for

the $N - C_U$ transition depends only weakly on N, as predicted by several of the models [14, 25, 26, 27]. The choice of effective DNA diameter $D_{eff} = 4.0\ nm$ also yields an effective $N - C_U$ transition volume fraction of $\phi_{NC} = 0.55$, close to the model predictions [14, 26, 27] (TBB line in fig. A.16-A). Note that the simple hard rod model of rigid monodisperse particles [14] organizes into the SmA phase at high c, but changes to the C_U phase upon the introduction of either polydispersity in hard rod length [26,27] or by molecular flexibility [29,34], as these effects reduce the free volume gain if the system were to organize into layers. In the limit of large L/D ratio, a length polydispersity above 20% is sufficient to eliminate the SmA phase, replaced by a hexagonal C_U phase for $\phi > \phi_{NC} \sim 0.6$ [26].

Among the sDNAs the LC phases appear in the N = 20 oligomer duplexes at concentrations in the range of those of the lDNAs, in spite of lacking a L/D sufficiently large to enable LC ordering in the hard rod model. Of particular interest to note in this regard is that the N = 20 $N - C_U$ transition concentration ($\sim$ 450 mg/ml) is nearly the same as that of the lDNA ($c_{NCU} \sim 370 mg/ml$). Given the lack of dependence of c_{NCU} on L for lDNA, this fact suggests that the phase behavior of the N = 20 oligomers might be understandable on the basis of the hard rod model, if their effective length L_{eff} were appropriately adjusted. In the simplest picture this adjustment assumes end-to-end aggregation into units of total length $L_{eff} = \langle a \rangle \cdot$(base height), from the aggregation of $\langle a \rangle / N$ oligomers; this length is sufficient to increase the L_{eff} of 20 bp duplex to contact the Onsager line, as shown in the construction in fig. A.16-B. For the N = 20 case this implies aggregates of length $\langle N \rangle \sim 200$, i.e., consisting of $\langle a \rangle \sim 10$ oligomer duplexes.

Justification for obtaining LC formation from such end-to-end aggregation can be found in model "living polymer"-type systems, where monomers with no steric anisotropy can reversibly aggregate into linear chains [25,28]. These chains have an exponential or broader distribution of lengths and thus are intrinsically polydisperse. Computer simulations of the LC behavior of spheres, reversibly aggregating into linear chains [25, 28], and those of flexible rods [29] show that for large bending rigidity the aggregate I-N transition occurs according to the Onsager prediction if the average aggregate length $\langle L \rangle$ is used in the Onsager model. In fig. A.16-A we report the I-N phase boundary of aggregating spheres according to the results of ref. [28] and show that the axial ratio of the aggregates, of average size $\langle N \rangle$, matches well that expected for hard rods (green dots and construction). This accordance indicates that the model aggregates behave effectively as hard rods, justifying the similar construction in Fig. A.16-B.

Furthermore, the unavoidable length polydispersity accompanying such aggregation replaces smectic phases by columnar. Thus, if we can assume that a similar criterion would apply to similarly aggregated rod-shaped particles, then the estimate described in the previous paragraph for LC formation in the N = 20 case is obtained. At higher concentration the system of spheres aggregating in rigid chains, studied in [25], exhibits a transition to the C_U phase. The $N - C_U$ phase boundary is predicted to be at about $\phi_{NC} \sim 0.5$ when the adhesion strength is large enough to yield aggregates of axial ratio $L/D > 5$ (dashed black line in fig. A.16 [25]).

Such end-to-end stacking of DNA duplexes is a structural theme in both DNA crystals [30, 31] and in DNA/protein hybrid crystals [32, 33] where DNA in adjacent unit cells stacks end-to-end.

As the oligomer length is decreased the C_U phase and the nematic phase persist for duplexes as short as N = 6, although the concentrations required to obtain these phases increase sharply. For N = 6 the $N - C_U$ transition is found at c = 1200 mg/ml, about 2/3 of that of neat duplex DNA $c_{DNA} = \rho_{DNA} = 1800 mg/ml$. Thus the LC phases of the oligomers of smallest N may be better viewed as being like thermotropic LC phases, rather than like those of colloidal particles.

The most notable feature of the sDNA phase behavior is the presence of the N phase even for the smallest sDNA studied. The model systems, ranging from flexible aggregate chains [26] to infinitely long repulsive chains [34] clearly show a requirement for adequate rigidity ($L_p \lesssim 10D$) for the nematic to appear. For the shortest sDNAs, the systems are of sufficiently high concentration to behave like single component thermotropic LC-forming systems, where axial ratios $L/D \gtrsim 5$ are required to exhibit nematic order. The only possible scenario to account for the sDNA behaving as anisotropic particles with L/D sufficiently large is end-to-end aggregation of the duplexes into oligomer chains of rigidity sufficient to enable nematic ordering. Both the infinite length and aggregate flexible rod models show that in addition to suppressing the N phase: (i) increasing flexibility elevates the concentrations needed to get both the N and C_U phases, possibly accounting for the increasing concentration scale for decreasing N; and (ii) since flexibility suppresses LC ordering, estimates of aggregate length using the Onsager line or other rigid chain phase boundaries are lower limits since flexible rods would have to be longer to give the LC phases [26,34]. At this high concentration where LC phases are found for the shortest oligomers, the inter-axial distance approaches the chemical diameter [24,35] where steric repulsion dominates the inter-chain interactions. Hence, with decreasing N and increasing concentration, the effective chain diameter evolves from $D \sim 4$ nm to $D \sim 2.4$ nm. This change shifts the model phase boundaries in fig. A.16-A to higher concentration (from black ϕ scale to orange ϕ scale), accounting for the increased concentration necessary for LC phase formation. This effective diameter variation is scaled out in fig. A.16-B, where the concentration axis is normalized by c_{NCU}, assuming the DNA concentration at the $N - C_U$ transition (c_{NCU}) to correspond, for each oligomer, to an effective volume fraction $\phi_{NC} \approx 0.55$ (the estimate of ϕ_{NCU} obtained by averaging over the results from the various models discussed above [14,26,27]). That is, the D appropriate for a given oligomer is that which makes its $N - C_U$ transition occur at $\phi_{NCU} = 0.55$, which yields D = 4.0 nm from the lDNA $N - C_U$ data, comparable to that required to also fit the lDNA I-N data. The resulting scaled phase diagram in fig. A.16-B illustrates that the nematic range in the sDNAs is shrinking as N decreases, indicating that the $L_p \sim 5D$ limit is being approached for small N. Once the effective volume fraction is held by the strict requirement placed by the $N - C_U$ transition, the amount of linear aggregation can be evaluated by horizontally projecting the I-N data points onto the expected I-N L/D for linear aggregation. For the N = 6, 8, 10, 12, 14, 16, 20 bp oligomers the mean numbers of oligomer duplexes in the aggregates, $\langle a \rangle$, and total aggregate lengths in basepairs, $\langle N \rangle$, are respectively $\langle a \rangle$ = 12, 9, 9, 8, 11, 6, 9 and $\langle N \rangle$ = 75, 80, 87, 97, 100, 160, 180 bp. The decreasing aggregate length needed for nematic order for shorter oligomers is a result of their higher concentration. The aggregation number does not depend strongly on N, implying that the adhesion energies are similar at the I-N transitions in the various oligomers.

Fig. A.16-B shows clearly that the nematic range is decreasing with decreasing N, which the computer simulations (fig. A.18) show is a sign of increasing flexibility of the aggregates [29, 34].

4.3.1 Stacking energy

At present, the organization of anisotropic solutes with sticky ends and high concentration has not been dealt with theoretically, so only semi-quantitative evaluation of ΔE_S can be made. We present here three possible ways to give such an estimate.

Orientational entropy

Since X-ray experiments at different inter-helix distances [9] show that significant osmotic pressure is necessary to achieve the concentrations required for the lDNA (and sDNA) LC phases, indicating a net repulsive interaction (electrostatic and steric) between the DNA duplexes, ΔE_S represents the energy difference for a pair of aligned and stacked duplexes relative to their mean repulsive interaction. In the systems of fixed concentration studied here the net repulsion is balanced by the confining walls and ΔE_S represents the energy available to internally structure the phase. In order to add single duplex oligomers to a semi-rigid duplex chain ΔE_S must be large enough to overcome the loss of orientational entropy ΔS of each new member, requiring

$$\Delta E_S > T\Delta S \sim K_B T \ln\left(\frac{\Delta\Omega}{4\pi}\right) \tag{4.1}$$

where $\Delta\Omega$ is the orientational phase space of the fluctuations in relative orientation of the duplexes in the chain. If the aggregated duplexes are not free to reorient about the helix axis, then an estimate for $\Delta\Omega$ is $\Delta\Omega \sim \pi\Delta\theta^2$, where $\Delta\theta$ is the mean square fluctuation in the relative orientation of helix axes in an aggregate. If the aggregates are of rigidity comparable to that of lDNA, then $\Delta\theta \approx 0.1$ radian and $T\Delta S \approx 6\ K_B T$, to be considered a minimum possible free energy requirement for aggregation, i.e., of imposing common orientation within an aggregate.

Living polymers models

From the estimate of the length of the sDNA duplex chains obtained above, it is, in principle, possible to estimate the end-to-end stacking energy ΔE_S between the duplexes. This can be obtained from various expressions proposed to describe the living polymer size distribution and mean aggregate length either in dilute [36] or in semi-dilute [28] solutions (discussed in section 2.4). Of particular relevance are the experiments on the chromonic mesogen Sunset Yellow, in Ref. [36]. Sunset yellow is an aromatic molecular sheet of area comparable to that of a DNA base pair and in aqueous solution exhibits a stacking energy of 7.2 $K_B T$.
The equations 2.46, 2.26 and 2.44, although they give different relations between

Equation	Reference	ΔE_S $[K_BT]$	σ $[K_BT]$
2.46	[28]	5.0	0.5
2.26	[36]	4.3	0.5
2.44	[39]	3.9	0.4

Table 4.1 - ΔE_S values obtained from living polymerization models.

the stacking energy and the average chain length, lead to similar estimates, ranging from 4 to 5 K_BT (see table 4.1). Noteworthy, for every given model, the energy value extracted from the data does not show any dependence on N, except for the experimental uncertainties.
More recent approaches consider reversible aggregation in high concentration system [37, 38] and obtain corrections to the scaling laws expected for dilute system, without however enabling explicit energy estimates.

Hydrophobicity

Another approach to estimating ΔE_S is simply from the terminal structure of the DNA duplex itself, noting that the H-bonded base pairs form hydrophobic planar structures which can lower their interfacial energy with adjacent water by moving close enough to another to expel the intervening water (the basic origin of the DNA helicity). This leads to an estimate of $\Delta E_S \approx 2500 cal/nm^2 mol$, frequently used e.g. to roughly estimate the contact energy of biological macromolecules interacting because of hydrophobic pockets [40, 41]. Since the area of a H-bonded pair of nitrogen bases is about 0.7 nm^2, we estimate a stacking energy $\Delta E_S \approx 6 K_BT$, in line with living polymerization estimates.

Our estimated values are about half of the base stacking energy expected for duplets of free H-bonded base pairs [42] and for the stacking of G-quartets [43], and about half of the enthalpic gain experimentally evaluated for the combination of pairing and stacking [44]. Our determination of ΔE_S is however larger than the base stacking free energy within DNA molecules [45], as expected because of the stacking entropic penalties contributing to the free energy and because of the higher constraints imposed by the backbone to the positional and orientational freedom of the stacked bases.
Anyway, this value is also surprisingly close to twice the reported estimates for coaxial stacking of two single bases [44, 46, 47], i.e. the free energy gain associated with the stacking of two contiguous strands hybridized on a third strand.
Also worth noting is that the stacking energy, the relative preference for end-to-end vs. side-by-side ordering, is likely to be largest under the low monovalent salt conditions of our experiments where the polymer backbones are highly repulsive. With multivalent cations the DNA tends toward electrostatic aggregation and the contrast between duplex side and end is reduced.

4.4 Factors affecting the phase behavior

4.4.1 Effect of temperature

The sDNA solutions exhibited thermotropic mesomorphism, melting at sufficiently high temperature, T, to the optically isotropic liquid (I) phase. This is shown in figure A.19 (Appendix A), where we plot T_{LC}, the largest T at which the N and C_U phases are found. T_{LC} grows with N and, for each oligomer, is larger for the C_U phase.
The relationship of the LC phase transitions to the duplex denaturation transition was probed. As discussed in chapters 1 and 3, the latter can be monitored by doping the sDNA solutions with Ethidium Bromide.
We find that the sDNA denaturation is spread over a range of about $20°C$ centered on the melting temperature T_m, which grows with the density of sDNA, as is observed for lDNA [48] (in addition to the concentration effect, a stabilizing effect of the LC template has been reported [49]).
Systematic measurements were done on 10bp and 16bp self-complementary oligos. We found $T_m \approx T_{I-C}$, that is when c is large enough to give the C_U phase, then the DNA denaturation occurs along with the LC melting transition to the I phase. At lower c we find that $T_m \approx T_{I-N} + 10°$, implying that the N phase is more readily disrupted by unpaired strands than the C_U phase.
The link between the stability of the helices and that of LC phase was also demonstrated by comparing the behavior of the couple of complementary sequences A + B (behaving as D12, see next section) with A + B', where B' contains internal mismatched nucleotides. The measured T_{I-N} and T_{I-C} are lower for the second couple, according to the disturbing effect of the mismatches on pairing.
We'll discuss more in detail the effects of helix renaturation on the LC phase behavior in the next chapter.

4.4.2 Modification of sequence and terminals

The requirement of developing sufficient shape anisotropy and rigidity in the aggregate units puts significant constraints on the sDNA structure in order for LC phases to appear. Thus, sequence substitutions reducing the ability of the sDNA duplexes to form linear rigid aggregates will reduce the stability of the LC phases. A variety of additional experiments were performed to probe the dependence of the sDNA LC behavior on oligomer base pair sequence and chain termination.
These included study of 12mers that were either palindromic, but with a sequence different from that of the Dickerson oligomer (AACGCATGCGTT), or mixtures of 12mers with different sequences that were complementary, e.g., CCTCAAAACTCC + GGAGTTTTGAGG (see chapter 5). Each of these exhibits the N and CU phases with concentration ranges comparable to those of the Dickerson 12mer in fig. A.17.

Effect of terminals

In an effort to influence the duplex end-to-end adhesion we carried out further experiments on the Dickerson 12-mers (D12), but with added unpaired tails, either 1T, 2T or 10T groups added at the 3' terminals. We found that this modification suppresses the LC phases with the exception of the C_2 phase, which is still observed in the case of D12 with 1T and 2T groups. We interpret this result as indicating that dangling ends reduce end-to-end adhesion, in line with the base-stacking concept.
By contrast, D12 duplexes phosphorylated on the 5' end exhibited nearly the same phase behavior as that of the $-OH$ terminated D12 duplexes described so far. Such duplexes, bearing one phosphor group per end, have the same composition as obtained by fracturing a lDNA duplex. This is obviously not true for solutions of duplexes with no phosphors at the ends and for solutions of duplexes phosphorylated at both 3' and 5' terminals. We tested also this last case with D12 duplexes phosphorylated on both ends, and found that the N and C_U phases were suppressed and only the C_2 observed, as with the 1T and 2T terminations.
Interestingly, the already mentioned coaxial stacking has also been studied by the modification of the terminal bases [46]; it was shown that two phosphate groups poorly influenced the stacking, while a complete inhibition comes from the steric hindrance of a fluorescein molecule attached to one terminal (and the same happens for LC formation, as we'll see in chapter 5). Stacking is also found to weakly depend on pH (explored in the range 6-8) and salt concentration (0.1-1 M). All these results seem to describe pretty well our experimental observations too, further proving that stacking is the leading mechanism.

A different effect of the terminals modification is obtained through the sequence CG12, the same as D12 but bearing an additional CG at 5'. This sequence, when forming a duplex, has dangling ends capable to form links with neighboring duplexes. For this reason, such sequence is not disfavored, but even forced to form chains somewhat analogous to sDNA living polymers. Indeed, this oligomer shows the very same phase behavior as sDNA (see figure A.20, color plate in Appendix A), with a similar thermal stability.

Bibliography

[1] M. H. F. Wilkins, A. R. Stokes, H. R. Wilson, *Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids*, Nature **171**,738 (1953)

[2] J. E. Lydon, *The DNA double helix - the untold story*, Liq. Cryst. Today **12**,1 (2003)

[3] V. Luzzati, V. Nicolaieff, A. Nicolaieff, *Small-angle x-ray diffraction study of gels of deoxyribonucleic acid and nucleoproteins*, J. Mol. Biol. **1**,127 (1959)

[4] C. Robinson, *Liquid-crystalline structures in polypeptide solutions*, Tetrahedron **13**,219 (1961)

[5] F. Livolant, A. M. Levelut, J. Doucet, J. P. Benoit *The highly concentrated liquid-crystalline phase of DNA is columnar hexagonal*, Nature **339**,724 (1989)

[6] R. L. Rill, T. E. Strzelecka, M. W. Davidson, D. H. Van Winkle, *Ordered Phases in Concentrated DNA Solutions*, Physica A **176**,87 (1991)

[7] K. Merchant, R. L. Rill, *DNA Length and Concentration Dependencies of Anisotropic Phase Transitions of DNA Solutions*, Biophys. J. **73**,3154 (1997)

[8] F. Livolant, A. Leforestier, *Condensed phases of DNA: structures and phase transitions*, Prog. Polym. Sci. **21**,1115 (1996)

[9] R. Podgornik, H. H. Strey, K. Gawrisch, D. C. Rau, A. Rupprecht, V. A. Parsegian, *Bond orientational order, molecular motion, and free energy of high-density DNA mesophases*, Proc. Natl. Acad. Sci. USA **93**,4261 (1996)

[10] H. H. Strey, J. Wang, R. Podgornik, A. Rupprecht, L. Yu, V. A. Parsegian, E. B. Sirota *Refusing to Twist: Demonstration of a Line Hexatic Phase in DNA Liquid Crystals*, Phys. Rev. Lett. **84**,3105 (2000)

[11] R. Brandes, M. Kearns, *Magnetic Ordering of DNA Liquid Crystals*, Biochemistry **25**,5890 (1986)

[12] T. M. Alam, G. Drobny, *Magnetic ordering in synthetic oligonucleotides. A deuterium nuclear magnetic resonance investigation*, J. Chem. Phys. **92**,6840 (1990)

[13] L. Onsager, *The effects of shape on the interaction of colloidal particles*, Ann. NY Acad. Sci. **51**,627 (1949)

[14] P. Bolhuis, D. Frenkel, *Tracing the phase boundaries of hard spherocylinders*, J. Chem. Phys. **106**,666 (1997)

[15] I. Dierking, *Textures of Liquid Crystals*, Wiley-VCH Verlag (2003)

[16] D. H. Van Winkle, M. W. Davidson, W. Chen, R. L. Rill, *Cholesteric Helical Pitch of Near Persistence Length DNA*, Macromolecules **23**,4140 (1990)

[17] C. B. Stanley, H. Hong, H. H. Strey, *DNA Cholesteric Pitch as a Function of Density and Ionic Strength*, Biophys. J. **89**,2552 (2005)

[18] F. Tombolato, A. Ferrarini, *From the double-stranded helix to the chiral nematic phase of B-DNA: A molecular model*, J. Chem. Phys. **122**,054908 (2005)

[19] A. A. Kornyshev, S. Leikin, S. V. Malinin, *Chiral electrostatic interaction and cholesteric liquid crystals of DNA*, Eur. Phys. J. E **7**,83 (2002)

[20] M. Kleman, *Defects in liquid crystals*, Rep. Prog. Phys. **52**,555 (1989)

[21] C. R. Safinya, K. S. Liang, W. A. Varady, N. A. Clark, G. Andersson *Synchrotron X-Ray Study of the Orientational Ordering D2-D1 Structural Phase Transition of Freely Suspended Discotic Strands in Triphenylene Hexa-n-dodecanoate*, Phys. Rev. Lett. **53**,1172 (1984)

[22] D. Durand, J. Doucet, F. Livolant, *A study of the structure of highly concentrated phases of DNA by X-ray diffraction*, J. Phys. II France **2**,1769 (1992)

[23] H. Durchschlag, in *Thermodynamic Data for Biochemistry and Biotechnology*, H. J. Hinz, Ed. (Springer-Verlag,1986)

[24] R. Podgornik, H. H. Strey, V. A. Parsegian, *Colloidal DNA*, Curr. Opin. Colloid Interface Sci. **3**,534 (1998)

[25] M. P. Taylor, J. Herzfeld, *A Model for Nematic and Columnar Ordering in a Self -Assembling System*, Langmuir **6**,911 (1990)

[26] M. A. Bates, D. Frenkel, *Influence of polydispersity on the phase behavior of colloidal liquid crystals: A Monte Carlo simulation study*, J. Chem. Phys. **109**,6193 (1998)

[27] A. M. Bohle, R. Holyst, T. Vilgis *Polydispersity and Ordered Phases in Solutions of Rodlike Macromolecules*, Phys. Rev. Lett. **76**,1396 (1996)

[28] X. Lü, J. T. Kindt, *Monte Carlo simulation of the self-assembly and phase behavior of semiflexible equilibrium polymers*, J. Chem. Phys. **120**,10328 (2004)

[29] R. Hentschke, J. Herzfeld *Isotropic, nematic, and columnar ordering in systems of persistent flexible hard rods*, Phys. Rev. A **44**,1148 (1991)

[30] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Crystal structure analysis of a complete turn of B-DNA*, Nature **287**,755 (1980)

[31] H. R. Drew, R. E. Dickerson, K. Itakura, *A Salt-induced Conformational Change in Crystals of the Synthetic DNA Tetramer d(CpGpCpG)*, J. Mol. Biol. **125**,535 (1978)

[32] M. R. Redinbo, L. Stewart, P. Kuhn, J. J. Champoux, W. G. J. Hol, *Crystal Structures of Human Topoisomerase I in Covalent and Noncovalent Complexes with DNA*, Science **279**,1504 (1998)

[33] C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder, T. J. Richmond, *Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution*, J. Mol. Biol. **319**,1097 (2002)

[34] J. V. Selinger, R. F. Bruinsma *Hexagonal and nematic phases of chains. II. Phase transitions*, Phys. Rev. A **43**,2922 (1991)

[35] M. Mandelkern, J. G. Elias, D. Eden, D. M. Crothers, *The dimensions of DNA in solution*, J. Mol. Biol. **152**,153 (1981)

[36] V. R. Horowitz, L. A. Janowitz, A. L. Modic, P. A. Heiney, P. J. Collings *Aggregation behavior and chromonic liquid crystal properties of an anionic monoazo dye*, Phys. Rev. E **72**,041710 (2005)

[37] Y. Rouault, *Concentration-induced growth of polymerlike micelles*, Phys. Rev. E **58**,6155 (1998)

[38] S. C. Greer, *Reversible polymerizations and aggregations*, Annu. Rev. Phys. Chem. **53**,173 (2002)

[39] P. J. Flory, *Principles of Polymer Chemistry*, Cornell University Press (1953)

[40] J. A. Reynolds, D. B. Gilbert, C. Tanford, *Empirical Correlation Between Hydrophobic Free Energy and Aqueous Cavity Surface Area*, Proc. Natl. Acad. Sci. USA **71**,2925 (1974)

[41] A. R. Fersht, *Structure and Mechanism in Protein Science*, Freeman & Co. (1999)

[42] J. Šponer, P. Jurečka, I. Marchan, F. J. Luque, M. Orozco, P. Hobza, *Nature of Base Stacking: Reference Quantum-Chemical Stacking Energies in Ten Unique B-DNA Base-Pair Steps*, Chem. Eur. J. **12**,2854 (2006)

[43] J. T. Davis, *G-Quartets 40 Years Later: From 5′-GMP to Molecular Biology and Supramolecular Chemistry*, Angew. Chem Int. Ed. **43**,668 (2004)

[44] J. Santalucia, D. Hicks, *The Thermodynamics of DNA Structural Motifs*, Annu. Rev. Biophys. Biomol. Struct **33**,415 (2004)

[45] E. T. Kool, *Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replication*, Annu. Rev. Biophys. Biomol. Struct **30**,1 (2001)

[46] V. A. Vasiliskov, D. V. Prokopenko, A. D. Mirzabechov, *Parallel multiplex thermodynamic analysis of coaxial base stacking in DNA duplexes by oligodeoxyribonuclotide microchips*, Nucl. Acids Res. **29**,2303 (2001)

[47] P. Yakovchuk, E. Protozanova, M. D. Frank-Kamenetskii, *Base-stacking and base-pairing contributions into thermal stability of the DNA double helix*, Nucl. Acids Res. **34**,564 (2006)

[48] J. Santalucia, *A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics*, Proc. Natl. Acad. Sci. USA **95**,1460 (1998)

[49] D. Grasso, S. Fasone, C. La Rosa, V. Salyanov, *A calorimetric study of the different thermal behaviour of DNA in the isotropic and liquid-crystalline states*, Liq. Cryst. **9**,299 (1991)

CHAPTER 5

Phase separation and liquid crystallization of complementary sequences in sDNA mixtures

5.1 Introduction

Molecular crowding of cellular interior is known to play a role in the spontaneous organization of biological macromolecules [1]. Several biochemical and physiological processes are found to be influenced by strong packing constraints [2, 3] and complex ordered arrangements, such as liquid-crystalline mesophases, have been shown to arise from highly packed bio-molecules [4]. Of particular interest is the ordering of highly concentrated DNA in cell nuclei which can lead to a variety of mesophases in vivo [5]. However, whether such packing constraints had represented an evolutionary advantage, and how biological macromolecules may have evolved to function in highly crowded environments, remain outstanding questions.

As we saw in the previous chapter, concentrated solutions of fully hybridized short DNA (sDNA) exhibit various liquid crystalline forms of supramolecular ordering, promoted by end-to-end adhesion of the paired bases at the terminals of the duplexes. This behavior has been described for a wide set of self-complementary 6-20 bp sDNA sequences and for mutually complementary sequences in the same length range. Here we explore the phase behavior of mixtures of sDNA where only some of the sequences are complementary, thus able to pair in duplexes, and part are not and thus always remain in the solution as single strands (SSs). We have found that in concentrated mixtures of duplex and SS sDNA, the system phase separates into duplex rich liquid crystal (LC) domains coexisting with a duplex-poor isotropic phase, leading to the physical segregation of 4-6 nm long complementary chains from non-complementary ones (figure A.21, color plate in Appendix A). This phase separation is a collective

effect of the duplex and SS sDNA, since the duplexes alone would not form LCs in such solutions, their concentration being well below that required for LC formation. We show that this interesting behavior is effectively described in terms of nucleation and growth of liquid crystals, wherein, upon cooling an isotropic solution of single stranded sDNA, a destabilizing transformation is produced by the hybridization of the complementary oligos as the temperature is lowered below the oligo melting temperature. The phase separation is due to a combination of duplex formation and end-to-end stacking, leading to depletion-type forces that arise from the entropy gain of depletants (single strands) upon demixing of solute particles (duplex aggregates). This manifestation of the depletion interaction is unique in that above the duplex melting temperature the depletants and solutes are chemically homogeneous and completely miscible, the difference in molecular size and flexibility enabling a depletion interaction only at temperatures where duplexes can form. The phase separation here described leads to a spontaneous form of purification of well-paired strands from unpaired or badly paired strands. This phenomenon could have provided the basis for a prebiotic molecular selection mechanism and synthesis of extended nucleic polymers by a novel form of LC-promoted autocatalysis, as will be further explored in next chapter.

5.2 Phase separation in sDNA mixtures

As for the case of self-complementary oligonucleotides, solutions of concentrated mutually complementary sDNA self-assemble and organize into LC phases. In figure A.22 (color plate in Appendix A) we show Polarized Transmission Optical Microscopy (PTOM) images for the chiral nematic and columnar LC phases found in sDNA-A + sDNA-B (A-B) 1:1 mixtures and we sketch the main features of the phase diagram. We find that the sDNA concentration c at the phase boundaries is $c(isotropic - nematic) \simeq 600 mg/ml$, $c(nematic - columnar) \simeq 780 mg/ml$, $c(columnar - crystal) \simeq 1180 mg/ml$. The concentration of neat DNA is $c \simeq 1800 mg/ml$. As the temperature T is raised, phases melt. At the same time, when T grows above the sDNA melting temperature T_m, duplexes unbind. Measures reported in previous chapter indicate that T_m is equal or slightly larger than the temperature at which LC domains melt, both of them showing an increase with sDNA concentration.

In figure A.23-A and A.23-B (Appendix A) we show PTOM images of cells containing respectively an A-B sample with molar ratio [B]/[A] = 3 and a mixture of 16-mer self complementary (SC) sDNA with mixed single strands (MIX) sample with molar ratio [MIX]/[SC] = 4, after equilibration at room temperature. We observe in both cases an isotropic majority phase rich in SS sDNA (black in transmission), coexisting with LC domains rich in sDNA duplexes. Analogously to what described previously, the optical textures of the LC domains observed in the experiments here reported indicate either chiral nematic or columnar ordering of the duplexes depending on the sDNA concentration.
The emergence of these phases is explained by the presence of attractive interactions between the duplex paired terminal bases, yielding end-to-end stacking. In the majority of our observations in A-B and SC-MIX samples we could de-

tect domains of columnar phases, as those reported in figure A.23, while the chiral nematic was observed only for values of [B]/[A] and [MIX]/[SC] smaller than 3. The partitioning of the duplex sDNA into the LC domains is confirmed by doping the sample with a small percentage of sDNA-A labeled with a fluorescein group externally to the duplex. This is shown in figure A.24 (in color plates in Appendix A), where it is apparent that the fluorescence is larger within the LC domains. We cannot extract quantitative information from this measurement because the fluorescent moiety disturbs the LC ordering and thus FITC-conjugate 12-mers are less concentrated in the domains than the "clean" 12-mers. Interestingly, no LC phase was ever detected in mixtures of sDNA-B and FITC-conjugate sDNA-A, when the FITC group is on the 5' terminus.

When equilibrium is reached, the LC domains fill a given fraction of the cell, ϕ_{LC}, which we find to be approximately equal to the DS volume fraction $\phi_{DS} = v_{DS}/(v_{SS} + v_{DS})$ (figure 5.1), where v_{DS} and v_{SS} are the total volumes occupied by duplexes and single strands, respectively. The values of ϕ_{DS} reported in the figure are calculated from the weight fraction of the various sDNA sequences by modeling the duplexes as cylinders of diameter 2.4 nm and height 4 nm (12bp) and 5.5 nm (16bp) and the single strands as semi-flexible coils, whose radius of gyration can be computed according to the Krakty-Porod equation [6,7]. On this basis, v_{SS} has been computed by using the "consensus" values of 4.3 Å per base for the SS contour length and 3 nm for the SS persistence length. However, as already seen in chapter 1.3.1, these values are not an unambiguous choice, since the contour length could be set to the chemical length of 6 Å per base and the persistence length could be shorter because of the large density of chains and ions [8]. We exploit this difference to quantify the uncertainty in the determination in ϕ_{DS} (error bars in figure 5.1). Hence, data suffer of rather large uncertainty on which should be added the difficulty in controlling preparations of sub-milligram cells. Furthermore, being a fraction of the domains smaller than the cell thickness, the determination of ϕ_{LC} is also approximate, the values in the figure being an upper boundary since they are extracted by assuming a cylindrical shape (across the cell surfaces) of the domains. However, the linear dependence of ϕ_{LC} on ϕ_{DS} and the close match in the values of the two volume fractions, indicate that the segregation of the duplexes inside the LC domains is almost complete. Actually, the fact that the slope is smaller than 1 (0.8 according to the linear fit - continuous line in the figure), could reflect the better packing - and higher concentration - of sDNA in the LC than in the Isotropic phase.
Overall, the notion of a nearly complete segregation of duplexes from sDNA mixtures is expected. First, it is consistent with the expected equilibrium between free duplexes (of molar concentration $[DS_{free}]$) and duplexes bound to the columnar structure ($[DS_{LC}]$), attained at the domain boundaries. $[DS_{free}] = k_{as}[DS_{LC}]$, where k_{as}, the equilibrium constant, expresses the free energy difference ΔG_{DS} between associated and free DS. Previous analysis of the sDNA phase diagram has led us to determine the double strand end-to end adhesion energy to be about $E \sim 6 - 8K_BT$, a figure that constitutes a lower estimate for ΔG_{DS}, given the presence of additional entropic forces, as described below. Hence $\Delta G_{DS} > 6 - 8K_BT$ implies $[DS_{free}]/[DS_{LC}] = k_{as} < \exp(-\Delta G_{DS}/K_BT) \simeq 10^{-3}$, an evaluation supporting the notion of an almost complete segregation of DS into the LC domains.

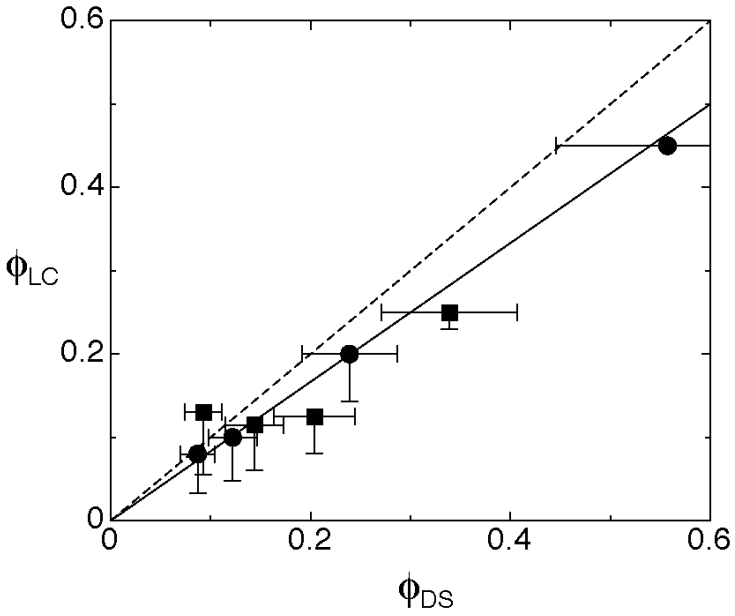

Figure 5.1 - Volume fraction ϕ_{LC} taken in the sample by LC domains as a function of the unbalancing of the [DS]/[SS] ratio in A-B mixtures (dots) and SC-MIX mixtures (squares). ϕ_{DS} is the ratio between the total volumes taken by duplexes and the total volume of sDNA, as computed by the sample preparation. Lines represent $\phi_{LC} = \phi_{DS}$ (dashed line) and $\phi_{LC} = 0.8\phi_{DS}$ (continuous line). Error bars are evaluated as described in the text.

Moreover, to obtain a nearly complete segregation of the sDNA duplexes it is necessary that nearly all complementary oligomers are associated in DSs. Since the sDNA concentration for a columnar phase for 12-mers is around 800-1000 mg/ml (see figure A.22), from the estimate above we expect $[DS_{free}] \simeq 0.2 - 0.3mM$. Given the fact that the binding constant for the oligomers under study at T = $35°C$ is of the order of $10^7 mol^{-1}$ (see equation 1.12 and section 6.6), we deduce that, even in the MIX SS sDNA rich phase, the large majority of the SC 16-mers is bound in duplexes. The same is even more true for the A-B mixture, where the large concentration of B strands in the SS sDNA rich phase strongly favors the binding of the residual A strands. Hence, in both the A-B and in SC-MIX mixtures we can assume the duplex to be fully associated.

5.2.1 Nucleation and growth of LC domains

The association of complementary strands into duplexes drastically alters their statistical physical properties, transforming the highly flexible single strands,

whose persistence length is of the order of a few bases, to rigid segments of double helix, their persistence length being much longer than their length (see section 1.3.1). As this association proceeds, the concentration of duplexes increases and the system is brought into a metastable state, a condition that leads to the formation and growth of nuclei of a new stable equilibrium phase. As T is raised above the duplex melting temperature, the nuclei dissolve and, if T is lowered again, form again in different positions. The temperature quenching necessary to trigger the onset of the nucleation process is rather mild, enough to ensure the duplex formation, which effectively yields a deep quenching from an hypothetical high T state with associated duplexes but melted LC domains, a condition which cannot be achieved upon heating because as T is raised the duplexes melt. The almost complete phase separation of duplexes and single strands is a consequence of this effectively deep quenching.

Inspection of figures A.23 and A.24 reveals two main geometries for the developing domains: either spherical focal conic domains with four perpendicular brushes or domains with an uniform optical axis and variable birefringence. These last domains appear to have an elongated shape in PTOM, but are actually approximately spherical as it can be observed by fluorescence microscopy (fig. A.24-B), or by uncrossing the polarizers in PTOM. In both domain geometries the stacked sDNA columns run parallel to the I-C interface. In the focal conic domains, they circle the central defect. In the uniaxial domains, there are regions (black in PTOM) where the columns run mainly perpendicularly to the cell surfaces, connecting up in the central part of the domain in vertical loops that produce the observed birefringence. This notion is supported by the fact that the refractive index is larger along the long axis of the elongated birefringent part of the uniaxial domains, indicating that the sDNA columns run across the short axis. Domains often have a smoothed polygonal shape, with "corners" normally located in the dark homeotropic regions of the uniaxial domains, probably related to the hexagonal symmetry of the C phase.
Indeed, our textures closely resemble those reported in [9, 10] for hexagonal discotic mesophases: Bouligand [9] describes both the growth of homeotropic domains, evolving from spherical to polygonal to dendritic shape - which reflects hexagonal packing; and the growth of focal conic domains in drops with a free upper surface. In the latter case, he observes initially circular domains to grow faster in the direction *normal* to columns, then developing lateral fans.
It is most likely that some of our LC domains nucleated on the cell surface. We basically observe the same type of behavior in sDNA, but usually limited to the first stages described by Bouligand: in fact, in all the domains described above, there was no evidence of anisotropic growth along a preferred axis, neither parallel nor normal to helices. Indeed, a few times we observed dendritic growth in weakly unbalanced mixtures, but, probably because of the high surface tension, most of the domains assumed compact shapes.

To gain insight into the nucleation process, we performed a detailed analysis of the early stages of the formation and growth of nuclei, extracting the number of nuclei N_N and their average size (radius r_N) as a function of time t. The results are shown in figure 5.2, where N_N and r_N vs. t are shown for an A-B mixture for [B]/[A] = 10, with t = 0 the time when the hot stage reaches T = 35 $^\circ C$.
Data indicate that both N_N and r_N reach an approximately steady state value

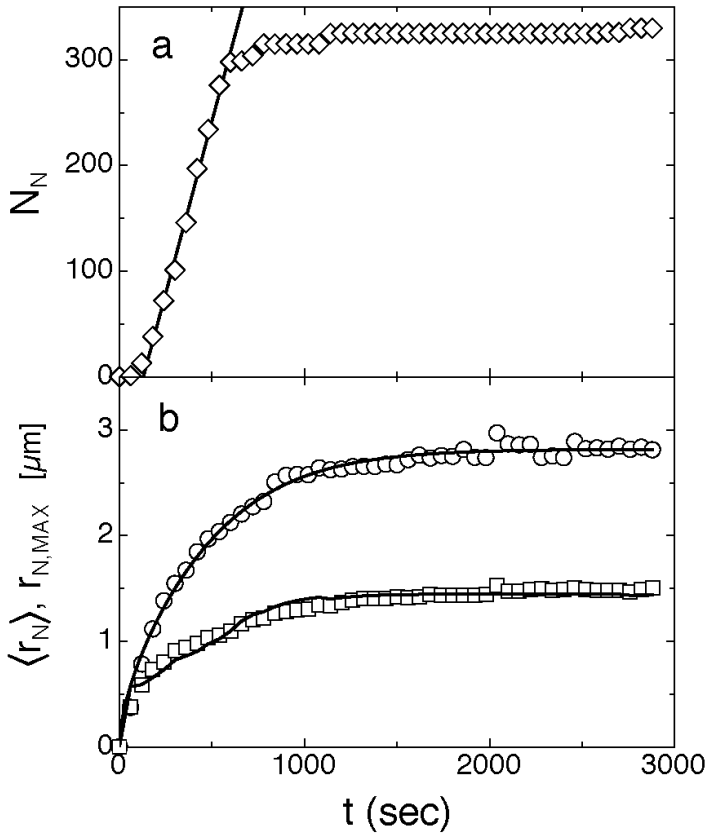

Figure 5.2 - Time dependence of the number of LC nuclei N_N (diamonds), of the radius of the largest (early nucleated) LC domains $r_{N,MAX}$ (dots) and of the mean nuclei radius $\langle r_N \rangle$ (squares). Lines are best fits of the Smoluchowski-type model to the data. The noise in the fitting curve for $\langle r_N \rangle$ is due to the convolution of the model with the derivative of N_N, performed to take into account the time distribution in the first appearance of the nuclei.

after about 20 minutes. On a larger time scale (days), domains undergo an additional growth that we interpret as due to the aggregation of LC nuclei having a size too small to be optically detected.
Nucleation involves the ordering of sDNA duplexes in a large enough quantity that the free energy gain associated to the formation of the stable phase overcomes the penalty associated to the surface tension of the interface. As the nuclei grow, the mixture is depleted of monomers and the nucleation rate decreases and eventually vanishes. We have extracted from the data in figure 5.2-A the nucleation rate ρ. In doing so, we have neglected the first data points, since the real onset of the nucleation is rather undetermined because of the $\sim 0.5\mu m$ microscopy resolution limits, a length much larger than the nucleation critical radius estimated below. However, assuming negligible merging of nuclei in the early stages of nucleation, the value for ρ, even if evaluated at a later stage, is correct. We find a nucleation rate of 0.8 nuclei/sec in the observed volume, corresponding to a rate $\rho = 4 \cdot 10^{-6}$ nuclei/$\mu m^3 sec$. To discuss the implications of this value, we need to first discuss the growth mechanism.
Once nuclei of supercritical diameter are formed, their growth is controlled by the diffusion either of single duplexes or of stacked duplexes that contact the nuclei and merge into them. Given the concentration of duplexes in the initially homogeneous isotropic sample and the end-to-end adhesion energy $\varepsilon \sim 6 - 8\ K_B T$, we expect the mean aggregation number N of stacked duplexes to be $N = \frac{1}{2} + \frac{1}{2}\sqrt{1 + 4\phi \exp(\varepsilon \kappa_I \phi)} \approx 3$ (equation 2.46), where ϕ is the volume fraction of DS sDNA and $\kappa_I \sim 1.45$ is a virial coefficient taking into account the steric repulsion between monomers [11]. In the simplest scenario the merging of DS sDNA into LC nuclei takes place with no activation barrier, i.e. in a diffusion limited regime. This process is easy to model using the approach originally proposed by Smoluchowski to describe the kinetics of colloidal aggregation [12]. Accordingly, we have computed the growth rate of a nucleus on the basis of the following assumptions:

1. the duplexes, insulated or grouped in short stacked chains, have an average diffusion coefficient D, and when they contact the nucleus, they become part of it;

2. at any time, diffusion of duplexes is a steady state process determined by their concentration gradient;

3. the local concentration of duplexes in the proximity of the nuclei is determined by the bound/free equilibrium k_{as}, and hence is effectively negligible;

4. the total concentration of duplexes is conserved within a spherical basin of radius R_B, whose size is experimentally determined from the mean distance of the nuclei.

We call r_N the radius of the growing domain, whose volume is proportional, through the number density n_{in}, to the number of contained duplexes. In the frame of this Smoluchowski-type model, a first order differential equation is derived for the dimensionless quantity $Y(t) = r_N(t)/R_B$:

$$\frac{dY}{dt} = \frac{2D_0(Y_0 - Y^3)}{Y(Y+2)(Y-1)^2} \tag{5.1}$$

where $D_0 = D/R_B^2$, $Y_0 = n_0/n_{in}$ and n_0 is the number density of sDNA duplexes in the basin of radius R_B before the nucleation. The value of the parameter R_B is determined from the average distance between the nuclei at equilibrium. The parameter D_0 is thus obtained from the fitting of Y(t) to the experimental data by means of two kind of analysis: (i) Y(t) is directly fitted to the measured size of the five largest domains (corresponding to those first nucleated in the cell) $r_{N,MAX}(t)/R_B$; (ii) the calculated Y(t) is convoluted with the measured number of nuclei appeared in each time interval and the resulting curve is fitted to the measured value of $\langle r_N(t)\rangle/R_B$. As a result, D is the only free parameter, the other ones being set on the basis of the concentration of nuclei in the cell and of the concentrations of duplexes in the cell and in the nuclei.We have determined D = 0.16 $\mu m^2/sec$ (fit to $r_{N,MAX}$) and D = 0.11 $\mu m^2/sec$ (fit to $r_N(t)$). The growth process is shown in figure 5.2-B.

Fluorescence diffusion experiment

Given the difficulty in determining the (large) viscosity of such a tiny quantity of material, to confirm our picture of the growth process we determined D for diluted A-B sDNA duplexes in a concentrated solution of B single strands, by making a contact cell with a pure B solution and a B solution doped with Fluorescein-A sDNA. As described in chapter 3, we acquired fluorescence profiles as a function of time across the contact line, fitted them with the Error Function spatial dependence expected for the diffusional evolution of a step function in concentration and extracted $D = 0.06 \pm 0.01 \mu m^2/s$.
Although the comparison suffers of various uncertainties, the most severe being the difficulty in determining the basin radius R_B, the rough agreement between the two different determinations indicates that the growth of the nuclei is indeed due to a diffusion limited process of single or weakly aggregated duplexes. This analysis also enables to roughly estimate the viscosity, $\eta = 1.2 \pm 0.5\ Pa \cdot sec$, of the concentrated sDNA solutions here investigated. The value obtained is large but not uncommon in aqueous solution of polymers [13].

Nucleation

Phase separation through nucleation can be described as an activated process, for which

$$\rho = \rho_0 \exp\left(-\frac{E^*}{K_B T}\right) \tag{5.2}$$

where ρ_0 is a space-time density of homogeneous nucleation "trials"; and E^* is the maximum in $E(r)$, the work necessary to establish a crystallite of radius r via a reversible thermodynamic path [14]. In the case of homogeneous nucleation, the attempt rate ρ_0 can be estimated as n/τ, where $n \sim 1/l^3$ is the duplex density in the sample just after the quenching, l is the mean spacing between duplexes in the same conditions, and $\tau \sim l^2/D$ is the time for a duplex to diffuse the distance l. Since in our experiments $n \sim 10^7$ molecules/μm^3 and $\tau \sim 0.4$ msec, we find $\rho_0 \sim 3{\cdot}10^{10}/\mu m^3 sec$, which with the measured $\rho \sim 4{\cdot}10^{-6}/\mu m^3 sec$ enables us to extract $E^* \simeq 36\ K_B T$ from equation 5.2.

Classical nucleation theory [15] provides an estimate for the activation barrier E^* in terms of γ, the surface tension at the liquid-LC interface. Namely

$$\frac{E^*}{K_BT} = \frac{16\pi}{3}\left(\frac{\gamma l_{LC}^2}{K_BT}\right)^3 \frac{1}{s^2} \tag{5.3}$$

where $l_{LC} \sim 3nm$ is the mean (isotropically averaged) duplex spacing in the LC phase. $s = \Delta\mu/K_BT = \ln([DS]/[DS_{free}])$ expresses the difference $\Delta\mu$ in bulk Gibbs free energy per particle between the droplet interior and the quenched isotropic phase. $[DS]$ and $[DS_{free}]$ are the molar concentration of DS in the quenched isotropic phase and in the isotropic phase coexisting with the LC domains. Since $[DS_{free}] \sim [DS_{LC}]/1000$ (see above) and $[DS] \sim [DSLC]/10$ (from the [A]/[B] ratio), $s \sim 4.6$. From the value of E^* extracted from the data, we obtain $\gamma l_{LC}^2 = 3.5K_BT$ and hence $\gamma \sim 1.5mJ/m^2$ for the LC droplet surface tension, which is comparable to that of protein crystals in solution [16]. These figures also imply an estimate for the critical radius for nucleation R^*. According to classic nucleation theory, $R^* = 2\gamma l_{LC}^3/\Delta\mu = 2\gamma l_{LC}^3/sK_BT \sim 5nm$, a value well below the resolution of PTOM observations.

5.3 Phase separation in sDNA-PEG mixtures

The behavior described so far for sDNA mixtures has interesting analogies with the very well studied mixtures of proteins and flexible polymers such as poly-(ethylene glycol) PEG, where protein crystals nucleate in a PEG-rich isotropic background [17]. In both systems, the mixture made of compact molecular structures (proteins, DNA duplexes) and flexible chains (PEG, SS sDNA) is unstable, and nucleates domains where the more rigid of the two species develops supramolecular ordering. To further explore the generality of the phenomenon, and to gain clues to interpret the observations, we have studied mixtures of sDNA duplexes and PEG of molecular weight ranging from 200 to 20000 with a corresponding gyration radius between 0.5 and 6 nm [18]. The mixtures were handled and observed in the same way as the sDNA mixtures.

At the lowest PEG molecular weights we could not observe any phase separation. Separation and LC domain formation instead occur with PEG having molecular weight equal to 1000 or larger, leading to a state of the system apparently quite similar to the one observed in sDNA mixtures, as shown in figure A.25-A (Appendix A).

The separated phase is indeed rich in sDNA, as we could check by doping the starting solution with Ethidium Bromide, which has a strong interaction with the DNA bases and hence marks the sDNA rich domains with a larger fluorescence. This is shown in figure A.25-B.

Despite the apparent similarity, a quite relevant difference can be detected by heating the sample above the DS sDNA melting temperature. We find that the phase separation does not depend on the formation of duplexes but on the chemical or physical mismatch between PEG and SS sDNA. This is shown in figures A.25-C and D, respectively PTOM and fluorescence microscopy images of the same sample held at $T = 75°C$ for 1 hour, a time sufficient to thoroughly disperse sDNA LC domains into the SS sDNA background. Given the fact that even highly concentrated PEG solutions are less viscous than the sDNA mixtures [13], we expected the sDNA domains to fade away. Instead, they don't

dissolve, so that by cooling again, domains are formed in the same locations as they were previously, an occurrence never observed in sDNA mixtures. To further prove this point we mixed PEG with non-complementary sDNA, obtaining even in this case phase separation (data not shown).
Hence we understand the formation of LC domains in sDNA-PEG mixtures as a consequence of a liquid-liquid phase separation, in which one of the phases is rich in complementary SS sDNA, which then hybridize, end-to-end couple and order.
Recent systematic observation of lysozyme and PEG indicate that, quite similarly, the formation of protein crystal is anticipated by a liquid-liquid phase separation where droplets of amorphous concentrated protein solutions are formed, in which crystals later develop [17]. On the contrary, and despite careful checking, in none among the many sDNA mixtures we have studied we could recognize phase separation of domains not associated with LC ordering.

5.4 Entropy and energy combine to drive the sDNA phase separation

In the A-B and SC-MIX mixtures the fraction of the sDNA making duplexes can be as small as 10% of the total sDNA content, in turn typically in the range 500-1000 mg/ml. If we imagine for a moment that the single strands were to be absent, the duplex concentration would be too small to enable significant end-to-end duplex adhesion, even with the 6 K_BT attractive adhesion energy inferred from the experiments on self complementary oligomers. In this case, given the phase diagram of figure A.22, LC phases should not be expected to form. The phase separation of such dilute duplex mixtures leading to the condensation of duplexes into high concentration domains indicates the presence of an additional effective attractive interaction between the duplexes that can only be due to the single strands. Since DNA is phase separating from DNA this interaction is not likely to be enthalpic in nature, but rather due to some fundamental incompatibility of the single stranded and duplex DNA with respect to their mixing behavior, that therefore must be entropic in nature. Given the fact that duplex DNA is much more rigid than single stranded DNA, one's attention is immediately drawn to the known entropic immiscibility of rigid and flexible polymers in solution [19], or to the entropic depletion interaction forces arising in binary mixtures of particles of different sizes or penetrability [20, 21]. Hence depletion-type forces provided by SS sDNA are necessary for the DS condensation to take place. At the same time, the enthalpic contribution from the duplex stacking energy is necessary as well. This is demonstrated by the suppression of the phase separation when stacking is impaired as in the case of DS sDNA terminated by a Fluorescein moiety.

Estimate of the osmotic pressure

A rough estimate of the total force holding the sDNA duplexes in the domains can be obtained by combining the typical inter-columnar distance in the sDNA C_U phase (chapter 4) with the corresponding osmotic pressure Π measured, for

the same packing, in long DNA columnar phases [22,23], i.e. $\Pi \sim 10^6 - 10^7 Pa$. The energy for detaching a sDNA duplex from the LC domain is of the order $\Pi v_{DS} \sim 3 - 30 K_B T$, where v_{DS} is the volume of a single DS. This energy goes into unstacking the sDNA paired terminal bases (6-8 $K_B T$) and into winning the entropic force, whose simplest estimate is $\Pi_{SS} \sim [SS]RT \sim 6 \cdot 10^5 Pa$, i.e. the osmotic pressure of the SS. Accordingly, the energy required to win the entropic forces is $\Pi_{SS} v_{DS} \sim 2 K_B T$.

5.4.1 Entropic forces

The depletion attraction between particles was first found theoretically by Asakura and Oosawa in their description of the interaction of hard particles in the presence of non-adsorbing polymers, modeled as penetrable spheres [20]. They found an increase of the entropy of the polymer spheres when the hard particles were moved together, leading to an effective attraction between the particles. Depletion forces have been observed and analyzed theoretically in a variety of systems, including mixtures of large and small colloidal particles, both with spheres and rods, mixtures of colloidal particles and polymers, and mixtures of proteins and polymers, inducing either or both particle aggregation and phase separation [21,24]. Depletion forces are also credited responsible for the formation of protein crystals in mixtures of proteins and polymers such as PEG [25].

The depletion effect is most easily understood in systems of large and small hard (repulsive) particles wherein the volume available to and thus the entropy of the small particles is increased by aggregation of the large particles. The size difference becomes less relevant when rigid particles are mixed with flexible polymers. In this case, phase separation of a particle-rich phase and a polymer-rich phase is expected at every ratio between particle radius R and coil gyration radius R_g. In the sDNA mixtures here described, R of duplexes and R_g of SS sDNA are similar. In this regime, according to depletion models [26], liquid-liquid phase separation should be expected. However, trying to account for our experimental data on the basis of hard sphere and penetrable polymer coil model is obviously an oversimplification. DNA strands are highly charged, and hence the system is in a regime of high repulsion and high osmotic pressure. Moreover, it has been shown that in the so-called "protein limit" of the depletion forces (i.e. for $R_g \gg R$), specific interactions play a major role, yielding for example to polymer length dependences not predicted by the simplified theories [27].

Phase separation and orientational ordering in one of coexisting phases in mixtures of rods and polymers were first predicted by Flory [19] on the basis of a lattice model. Accordingly, the entropic gain in aligning long rodlike particles may drive a phase separation where the random coils are almost completely expelled by the nematic phase. Later models [28,29,30] focus on the depletion forces present in polymer and rod-like colloids mixtures and also predict phase coexistence of isotropic and anisotropic phases. These predictions partially account for the observed behavior of mixtures of polymers and bohemite rods [31], cellulose rods [32] and filamentous viruses [24]. Even more generally, entropic forces are expected to play a major role in concentrated solutions of mixed solutes, in which the solutes differ not only in size or penetrability, as in the case of depletion forces, but also in flexibility and tendency to self-assemble, giving rise to rich phase diagrams, which include nematic LC ordering [33]. These forces

are considered responsible for various mechanisms of self-assembly and spatial organization in living cells [1,34]. Indeed, mismatch in flexibility and tendency to aggregate into elongated chains are exactly the key properties of the sDNA system.
Unfortunately, the range of the ratio persistence/contour lengths is, in the case of SSs, too small with respect to the range explored in these models, while its width, comparable to duplexes, its limited flexibility, and the electrostatic repulsion make the penetrable sphere description of Asakura-Oosawa's depletion interaction also hard to apply. However, although the theoretical models cited here do not fit the range of statistical physical properties of duplex and SS sDNA, their combination strongly support the notion that entropic forces act to segregate duplexes from SS sDNA.

5.5 Conclusions

The phase separation of sDNA mixtures here described can be understood as a combination of entropic forces, due to flexibility mismatch, and energy gained upon stacking the paired terminals of DS sDNA. The subtle combination of elements appears finely tuned in sDNA since neither LC ordering nor liquid-liquid phase separation occurs if end-to-end adhesion of duplexes is suppressed, by introduction of steric hindrance at the duplex ends, for example. This is in contrast to other depletants such as short chain PEG polymers, that we find can induce liquid-liquid phase separation without LC ordering. Indeed, interesting analogies emerge between PEG-promoted protein crystallization and PEG-promoted liquid crystallization of sDNA duplexes, since in both cases a liquid-liquid phase separation appears to precede molecular ordering, an effect probably due to strong entropic forces, in turn due to the strong mismatch in molecular properties (size and flexibility) between the chemically heterogeneous solutes. By contrast, the chemical homogeneity of sDNA mixtures makes the entropic forces tenuous enough not to be sufficient to provide phase separation by themselves. Such a delicate balance may be related to molecular selection processes in the first emergence of poly-nucleic acids, as we'll further investigate in the next chapter. Indeed, the phase separation here described leading to a spontaneous form of purification of well-paired strands from unpaired or badly paired strands, could have provided the basis for a pre-biotic molecular selection mechanism. Should the partitioning of complementary sDNA duplexes shown here combine with conditions favoring ligation, a novel form of liquid crystal-promoted autocatalysis would be established.
Indeed, into semi-rigid linear aggregates the terminal groups on neighboring oligomers are in close proximity, and thus that their effective concentration, c_{tLC}, is high. The most conservative estimate for c_t assumes that the duplexes in a chain are uncorrelated with respect to their orientation about the chain axis, and thus that the terminal groups occupy a toroidal-shaped volume v_{tLC} about the chain axis. We estimate $v_{tLC} \sim 0.8nm^3$, yielding $c_{tLC} = 1/v_{tLC} \sim 2$ Molar. The end group concentration in the isotropic phase in the case N=12 (c = 400mg/ml), is $c_{tISO} = 0.006$ Molar, yielding a significant end group concentration enhancement, $c_{tLC}/c_{tISO} \sim 320$, upon LC condensation.
The overall process would favor the preferential growth of the oligomers well

aligned inside the liquid crystallites, and thus promote the polymerization of complementary oligos.

At the same time, duplex segregation and liquid crystallization driven by SS sDNA appear as a new test bench for the exploration and understanding of entropic forces in the low nano-scale size range. This system appears particularly attractive because of its chemical homogeneity which simplifies interpretation, reducing the complexity associated to the variability of intermolecular forces among the various chemical species, typically relevant in the case of concentrated heterogeneous solutions of macromolecules [27].

Bibliography

[1] J. Herzfeld, *Entropically Driven Order in Crowded Solutions: From Liquid Crystals to Cell Biology*, Acc. Chem. Res. **29**,31 (1996)

[2] A. P. Minton, *The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media*, J. Biol. Chem. **276**,10577 (2001)

[3] R. J. Ellis, *Macromolecular crowding: obvious but underappreciated*, Trends Biochem. Sci. **26**,597 (2001)

[4] Y. Bouligand, *Twisted fibrous arrangements in biological materials and cholesteric mesophases*, Tissue Cell **4**,189 (1972)

[5] F. Livolant, *Ordered phases of DNA in vivo and in vitro*, Physica A **176**,117 (1991)

[6] O. Kratky, G. Porod, *Röntgenuntersushung gelöster Fagenmoleküle*, Rec. Trav. Chim. Pays-Bas **68**,1106 (1949)

[7] K. Kopecka, G. Drouin, G. W. Slater, *Capillary electrophoresis sequencing of small ssDNA molecules versus the Ogston regime: Fitting data and interpreting parameters*, Electrophoresis **25**,2177 (2004)

[8] B. Tinland, A. Pluen, G. Sturm, G. Weill, *Persistence Length of Single-Stranded DNA*, Macromolecules **30**,5763 (1997)

[9] Y. Bouligand, *Defects and textures of hexagonal discotics*, J. Physique **41**,1307 (1980)

[10] P. Oswald, M. Kléman, *Défauts dans une mésophase hexagonale discotique: disinclinaisons et parois*, J. Physique **42**,1461 (1981)

[11] X. Lü, J. T. Kindt, *Monte Carlo simulation of the self-assembly and phase behavior of semiflexible equilibrium polymers*, J. Chem. Phys. **120**,10328 (2004)

[12] M. Von Smoluchowski, *Versuch einer mathematischen Theorie der Koagulationskinetic kolloider Lösungen*, Z. Phys. Chem. **92**,129 (1917)

[13] P. Gonzàlez-Tello, F. Camacho, G. Blàzquez, *Density and Viscosity of Concentrated Aqueous Solutions of Polyethylene Glycol*, J. Chem. Eng. Data **39**,611 (1994)

[14] L. D. Landau, E. M. Lifshitz, in *Statistical Physics*, Addison Wesley (1958), chap. 12

[15] P. G. Debenedetti, *Metastable Liquids*, Princeton University Press (1996)

[16] R. P. Sear, *On the Interpretation of Quantitative Experimental Data on Nucleation Rates Using Classical Nucleation Theory*, J. Phys. Chem. B **110**,21944 (2006)

[17] O. Galkin, P. G. Vekilov, *Control of protein crystal nucleation around the metastable liquid-liquid phase boundary*, Proc. Natl. Acad. Sci. USA **97**,6277 (2000)

[18] R. Bhat, S. N. Timasheff, *Steric exclusion is the principal source of the preferential hydration of proteins in the presence of polyethylene glycols*, Prot. Sci. **1**,1133 (1992)

[19] P. J. Flory, *Statistical Thermodynamics of Mixtures of Rodlike Particles. 5. Mixtures with Random Coils*, Macromolecules **11**,1138 (1978)

[20] S. Asakura, F. Oosawa, *Interaction between particles suspended in solutions of macromolecules*, J. Polym. Sci. **33**,183 (1958)

[21] R. Tuinier, J. Rieger, C. G. de Kruif, *Depletion-induced phase separation in colloid-polymer mixtures*, Adv. Colloid Interface Sci. **103**,1 (2003)

[22] R. Podgornik, H. H. Strey, K. Gawrisch, D. C. Rau, A. Rupprecht, V. A. Parsegian, *Bond orientational order, molecular motion, and free energy of high-density DNA mesophases*, Proc. Natl. Acad. Sci. USA **93**,4261 (1996)

[23] H. H. Strey, V. A. Parsegian, R. Podgornik, *Equation of state for polymer liquid crystals: Theory and experiment*, Phys. Rev. E **59**,999 (1999)

[24] Z. Dogic, K. R. Purdy, E. Grelet, M. Adams, S. Fraden, *Isotropic-nematic phase transition in suspensions of filamentous virus and the neutral polymer Dextran*, Phys. Rev. E **69**,051702 (2004)

[25] A. McPherson, *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press (1999)

[26] H. N. W. Lekkerkerker, W. C. K. Poon, P. N. Pusey, A. Stroobants, P. B. Warren, *Phase behaviour of colloid + polymer mixtures*, Europhys. Lett. **20**,559 (1992)

[27] J. Bloustine, T. Virmani, G. M. Thurston, S. Fraden, *Light Scattering and Phase Behavior of Lysozyme-Poly(Ethylene Glycol) Mixtures*, Phys. Rev. Lett. **96**,087803 (2006)

[28] H. N. W. Lekkerkerker, A. Stroobants, *Phase behaviour of rod-like colloid+flexible polymer mixtures*, Nuovo Cimento D **16**,949 (1994)

[29] P. G. Bolhuis, A. Stroobants, D. Frenkel, H. N. W. Lekkerkerker, *Numerical study of the phase behavior of rodlike colloids with attractive interactions*, J. Chem. Phys. **107**,1551 (1997)

[30] A. Matsuyama, T. Kato, *Orientation-dependent depletion interaction in rodlike colloid-polymer mixtures*, Eur. Phys. J. E **6**,15 (2001)

[31] M. P. B. van Bruggen, H. N. W. Lekkerkerker, *Morphology and kinetics of the isotropic-nematic phase transition in dispersions of hard rods*, Macromolecules **33**,5532 (2000)

[32] C. D. Edgar, D. G. Gray, *Influence of Dextran on the Phase Behavior of Suspensions of Cellulose Nanocrystals*, Macromolecules **35**,7400 (2002)

[33] D. T. Kulp, J. Herzfeld, *Crowding-induced organization of cytoskeletal elements.III. Spontaneous bundling and sorting of self-assembled filaments with different flexibilities*, Biophys. Chem. **57**,93 (1995)

[34] J. Herzfeld, *Crowding-induced organization in cells: spontaneous alignment and sorting of filaments with physiological control points*, J. Mol. Recognit. **17**,376 (2004)

CHAPTER 6

Oligonucleotide LC ordering and prebiotic scenarios

Man has always been striving for knowing the mystery behind the living systems and his own existence. As men and scientists, we can pursue everybody's questions with an additional, particular point of view: well aware that we can't give an exhaustive answer, we can try to elucidate how complex phenomena, such as the development of biotic macromolecules and their self-organization, emerged, and we can look for their roots in simple and general physical principles (if any). We argue that the discovered behavior of small DNA described in the previous chapters, i.e. phase separation, end-to-end stacking and liquid-crystalline ordering of helices with respect to *ss-DNA* , can shed light on the steps that lead to the formation of the present DNA molecule, the fundamental information carrier in cell's life. Did this tendency to ordering lead to - or at least favor - ligation and elongation of DNA? How general is this mechanism of self-purification? In the following, we want to address these and other questions enlightened from our work on sDNA, to test the universality of this phenomenon.
This chapter divides into two main parts: the first part reviews the current studies on early-life scenarios, in particular on the origin and development of RNA and DNA; it shows how the combination of phase separation, self-assembly and LC ordering could solve some open issues; it presents some other examples of liquid crystalline models of biological systems.
To gain insight in the main features and in the potentiality of the proposed mechanism, in the second part we present some ongoing experiments on the LC ordering of short RNA oligomers, on LC-mediated purification of long DNA helices and on LC phases in random sequences.

PART I

6.1 Prebiotic scenarios

In the last decades, many attempts were undertaken (and some significant advances were obtained) to clarify some critical steps in life's origin and evolution, such as the synthesis of first building blocks, the origin of RNA and DNA or the first cellular organization [1]. However, also given the difficulty to verify some of environmental conditions on the early Earth, most of the issues are still debated [2], the first being just about the basic requirements to define a living system [3]: the ability of autonomous replication and the possibility to keep and propagate information (and thus to take advantage on natural selection).

In figure 6.1, from [4], the timeline of the main events regarding the origin of life is reported, showing its possible emergence from the clutter of prebiotic compounds through nucleic acids and proteins, over a period of about one billion of years, as suggested by evolution evidences.
Of course, to trace this path, many problems have to be considered: these include the non-enzymatic synthesis of biological monomers in the pre-biotic environment, the search for the first energy source of the primitive metabolism, the exploration of mechanisms for sequestering biomolecules on a surface or within vesicles (compartmentation), and above all the search for simple organic replicating systems.
As already mentioned, replication is a key property at the basis of the possibility of evolution: if a polymer is able to give rise to additional polymer molecules of the same sequence and if the rate of production of new copies exceeds the rate of degradation of existing copies, then a particular polymer sequence will persist over time. Natural environments are subject to fluctuating conditions, both periodic (night and day, seasons) and unpredictable (cataclysmic events). When the environment is altered, the special properties associated with a particular polymer may no longer apply and the capacity for self-replication may be lost. Persistence in a changing environment requires a more general mechanism for self-replication that allows the polymer sequence to change somewhat over time, but retain its heritage in most of the sequence that is unchanged. The polymer must be replicated in essentially the same manner regardless of

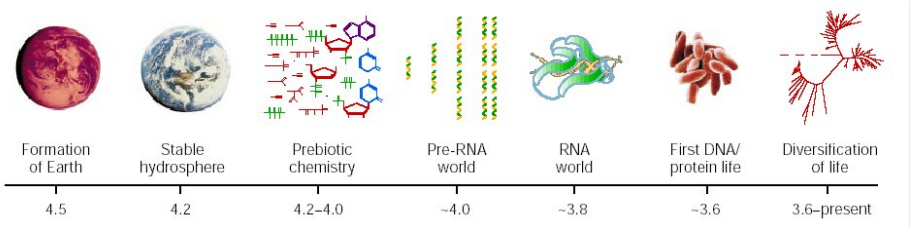

Figure 6.1 - Timeline of the early history of life on Earth, with (approximate) dates in billions of years before the present; from [4].

its sequence. Variation will arise owing to inevitable copying errors, and those variants too must be amenable to replication. Once a general mechanism existed for self-replication, allowing the introduction of variation and the ability to replicate those variants, darwinian evolution could start to operate: the special properties of a particular polymer sequence then were defined by its net rate of accumulation (rate of production minus rate of degradation), and sequences that were associated with the most favorable survival rates would have come to dominate their local environment. From that point onward, the natural history of life on Earth played out as a succession of dominant polymer sequences and their associated functional properties [4].
This role, of course, is nowadays held by DNA transcription mechanism, but indeed the origin of a self-replicating genetic code probably represents the main inescapable issue to be resolved before a really plausible scenario for the emergence of life on the early Earth can be defined.

6.1.1 The RNA-world

The RNA molecule has a pervasive role in contemporary biology, especially with regard to the most fundamental and highly conserved cellular processes. It is involved as a primer in DNA replication, as a messenger that carries genetic information to the translation machinery, as a catalyst that lies at the heart of the ribosome and in many other roles. Some catalytic RNAs, called ribozymes (ribonucleic enzymes), were also found to assist in RNA processing events and the replication of viral genomes.
In principle, replication of a ribozyme requires only a single macromolecular activity: an RNA-dependent RNA polymerase (i.e., an enzyme promoting polymerization) that synthesizes first a complement, and then a copy of the ribozyme. If this RNA polymerase were itself a ribozyme, then a simple ensemble of molecules might be capable of self-replication and eventually, in the course of evolution, give rise to the protein-nucleic acid world of contemporary biology.
It thus seems likely that RNA was the first molecule having the capability to support life based on RNA genomes that are copied and maintained through the catalytic function of RNA itself, giving rise to the so-called "RNA-world" [5,4], later replaced by the present machinery of DNA and proteins.
There are several examples of ribozymes, evolved *in vitro* from a random pool, that catalyze the template-directed joining of an oligonucleotide 3'-hydroxyl and oligonucleotide 5'-triphosphate [6], although there is no known ribozyme in biology with similar properties.
However, the RNA-world view has been disputed because of issues intrinsic to RNA chemistry. Although ribose, phosphate, purines and pyrimidines all may have been available in prebiotic environment, these had to combine to form nucleotides in very low yield, complicated by the presence of a much larger amount of various nucleotide analogues.
The nucleotides (and their analogues) may even have joined to form polymers, with a combinatorial mixture of 2',5', 3',5' and 5',5'-phosphodiester linkages, a variable number of phosphates between the sugars, D and L-stereoisomers of the sugars, and assorted modifications of the sugars, phosphates and bases. The self-replication mechanism had somehow to accommodate these compositional differences and select the "right" nucleic acids. In addition, only conveniently

activated nucleotides can be ligated to a chain. Actually, the phosphorylation of mononucleotides and the synthesis of short oligomers was demonstrated in suitable extreme environmental conditions [7,8], but today the usual laboratory route is to use phosphorimidazolides of nucleosides or other activating groups [9], favoring polymerization, whose presence in prebiotic environment has not been proved.

Another class of objections raised against the RNA world hypothesis pertains to the activities of RNA catalysts, i.e. to the mechanisms that must have led to the emergence of specific, rather long (despite the relative fragility of long RNA polymers in aqueous solutions), active sequences over all possible sequences. Indeed, although it was demonstrated that oligo(C)s as short as four monomer units in length can serve as efficient templates for the synthesis of oligo(G)s from activated monomers [10], a RNA fragment length of 30 is assumed to be required for a good catalytic activity. In this case, approximately 10^{18} different sequences (0.02 g RNA) are possible and could be synthesized; if a 50-mer length is needed, 10^{30} RNAs ($3.5 \cdot 10^{7}$ kg RNA) theoretically exist, but cannot likely be made at once. Although several sequences could have had a similar activity, concentration issues associated with the total catalytic activity must not be overlooked.
In summary, *if* the building blocks of RNA were available in the prebiotic environment, *if* these combined to form polynucleotides, and *if* some of the polynucleotides began to self-replicate, then the RNA world may have emerged as the first form of life on Earth.
Assuming its validity, the RNA-world somehow solves the "chicken or the egg" problem between nucleic acids and proteins, but still leaves the following question unanswered: how did the *first* poly-nucleotides arise from monomers, without *any* enzyme, of whatever nature?

In contrast to the "information first" scenario sketched so far, the other main theory, named "metabolism first", claims that life arose from autocatalytic self-organizing chemical cycles [11]. For complex mixtures of reactants and products to move in the direction of life, a process of self-organization would be necessary. This process would enhance the concentration of certain components of the mixture, either at the expense of others, or by new synthesis from raw materials, with these changes driven by an external source of energy. Despite the absence of a genetic polymer, a transformed mixture of this type could be considered to hold hereditary information, which would be represented by the identity and concentration of its constituents ("compositional genome"). Evolution would be represented by changes in the composition of the system and in the reactions used to sustain it, in response to changes in the surrounding environment. Growth of the system would take place through the acquisition or synthesis of additional quantities of the key components, and reproduction would occur when physical forces split the enlarged system into two or more fragments.
Unfortunately, no plausible self-sustaining chemical cycles have been found so far, and thus even the proof of principle is still missing. Therefore, in the current absence of any other reasonable precursor, the RNA model represents a system that allows us to explore essential aspects of the emergence of a polymeric, genetic system without the requirement of a complex metabolism.

6.1.2 Required conditions for the development of the first poly-nucleotides

Whatever theory is thought to be the most adequate, the critical step for the dawn of RNA (or DNA) as the information carrier lies in its elongation from single nucleotides, or oligomers, to the long polymer we know nowadays.
RNA self-replication cannot be achieved by one single molecule - at least two are needed - and generally for any workable chemical system one needs local concentrations of at least femto-moles of sequences and larger for the mono-nucleotides. Clearly, a relatively high concentration is a key requirement for the polymerization to take place and not to be overcome by chain degradation.
High concentration alone is not sufficient, though, to sustain polymerization: like most of the present enzymes work by geometrical, physical constraint - by keeping close together active groups and thus enhancing reaction rates -, a template mechanism is invoked in prebiotic times to favor ligation of different nucleotides. In the words of C. De Duve [1]: "The need seems inescapable for some autocatalytic process such that each lengthening step favors subsequent lengthening.[...] Only in this way could the enormous kinetic obstacle to chain elongation be surmounted.[...] any invoked catalytic mechanism must accommodate the participation of a template, for there can have been no emergence of true RNA molecules without replication."
We're going to briefly review some of the current studies on how these two requirements could have been accomplished on early Earth. Each of them is supported by some experimental evidences (or more often by laboratory proofs of principle), but none of them can overcome the others with clear evidence.
We highlight that there is a subtle analogy between, on one hand, the basic requirements for life's definition - replication, requiring on its turn a mechanism for copy and polymerization, and competition/selection, requiring individual, compartmented entities - and, on the other hand, the conditions that could have facilitated the evolution of poly-nucleotides - a template and local high concentration. The involvement of nucleic acids in the deep roots of life's origin is again evidenced.

High concentration

Both to keep concentration high enough to enable encounters and reactions, and to make selection and evolution possible, some kind of compartmentation necessarily had to start. Of what nature? Life is now organized into cells, very complex "worlds" basically separating genetic material (and various degrees of machineries) from the outside through selectively permeable lipid membranes.
The most natural and conservative approach is thus to imagine some simple proto-cells, combinations of RNA and surfactants, achieving the same results and able to replicate [3]. This possibility is made more interesting by the fact that some surfactant micelles and vesicles were found to spontaneously split (and thus "self-replicate") under appropriate conditions [12, 13].
Although some similar phenomenon had certainly to occur for the birth of the first cell to take place, that this was the real driving force of the first RNA segregation and elongation is questionable, since the replication of vesicles appears as a continuous statistical process, hardly bearing resemblance with the real, discontinuous cell replication.

Although there is evidence that prebiotic oceans were as dilute as contemporary ones, hydrotermal marine environments, characterized by heat currents flowing through porous minerals, likely played an important role in the development of life, providing a heat source and fluctuating conditions.
Indeed, the possibility to obtain a spontaneous concentration of RNA strands near hydrothermal mounds is suggested by two simple physical phenomena related to the motion of fluids inside a channel and of particles within fluids, namely convection and thermo-diffusion [14]. To accumulate molecules from a highly diluted prebiotic ocean, a considerable entropic gap has to be bridged. In a rough estimate, a 10^6-fold accumulation is required for small protobiomolecules to interact. It was shown that these two effects, convection and thermo-diffusion, together can provide a robust method of accumulation inside hydrothermal marine pores [14].

Finally, some structured environments, constituted by either self-assembling molecules or mineral surfaces, could perform at the same time the two major functions which can enhance biopolymer synthesis: they can physically confine or compartmentalize molecular reactions; and they can also promote reactions due to chemical and physical properties that differ from those of bulk aqueous media, permitting the assembly of reagents in normally unstable configurations. We'll mention the proposed systems in the following section.

Template

The appeal of RNA as the first self-replicating molecule relies on the fact that, by definition, it would be capable of acting autocatalytically for its own synthesis and, at the same time, such autocatalytic molecule would act as a template to bind the precursors by non-covalent forces and organize them in such a way that the reactive groups come in close proximity (see figure 6.2). Studies with activated trimers and hexamers showed that template autocatalysis can only occur if the sequences of both trimers match the sequence of the hexamer according to the Watson-Crick base-pairing rules. They also showed that the condensation reactions are predominantly controlled by the stacking of nucleic acid bases flanking the newly formed internucleotide link [15].
As already said, however, we are left with a new "chicken or the egg" problem: how did the first templating RNA oligomer arise, without a template?

Bulk condensation polymerization reactions are usually thermodynamically driven towards hydrolysis in dilute aqueous solutions. Therefore, in the absence of high concentration, and possibly besides it, a surface-promoted mechanism is required to enhance the polymerization rates.
Some mineral surfaces have been proposed as good templates for nucleotide polymerization. The most credited candidate is montmorillonite (figure 6.3), a clay mineral with layer structure. An excess negative charge present in the layers is counter-balanced by cations held between them. Reversible hydration or solvation of the cations cause the layers to expand, favoring the entrance of certain molecules [16].
A big number of experiments verified the binding of mononucleotides onto the montmorillonite surface or inside its layers and its ability to promote the formation of the phosphodiester bond (in suitably activated monomers) and thus

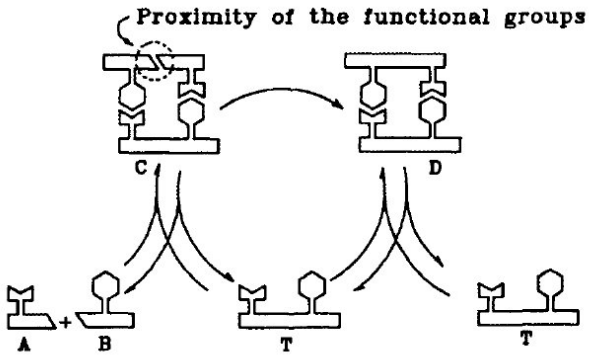

Figure 6.2 - Mechanism of template-directed self-replication; from [15].

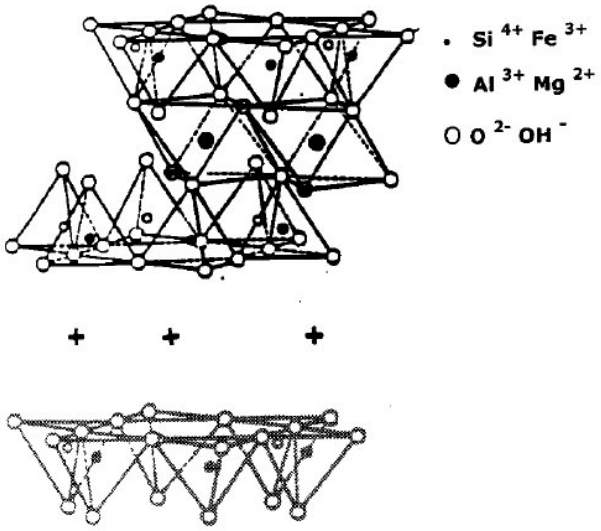

Figure 6.3 - The structure of montmorillonite, with inter-layer ions; from [16].

the elongation of nucleotide polymers [17, 9, 18]. Interestingly, montmorillonite was also found to favor the homo-chiral selection of nucleotides [19], a critical step in the development of longer molecules [1].

However, in clays the catalytic surface is a liquid-solid interface, and the lack of fluidity in this interface could be poorly compatible with an efficient surface diffusion [20]. Such surfaces also lack the flexibility which is found in present-day enzymes, and which is known to be crucial to catalysis. Coacervation is a liquid-liquid phase separation involving polymeric chains. Oparin [21] suggested that prebiotic polymerization reactions took place in a heterogeneous, coacervated system, rather than in the bulk of a homogeneous phase.
Coacervation was considered as an essential concentrating process by which mixtures of randomly formed prebiotic polymers initially in dilute solutions were condensed into concentrated assemblies. The phase separation of the polymers into separate coacervate droplets - Oparin claimed - provided the appropriate medium required for the evolution of these prebiotic systems. Although naive (he thought that coacervated droplets directly lead to cells), his theory introduced for the first time the idea that a physical phase separation process could lead to locally enhanced concentration of nucleotides and thus more favorable elongation.
The proposal of the catalytic role of a liquid-liquid interface, namely between an oil slick and salty water, was put forward by Lars Onsager [22], and the renaturation process of DNA oligomers was found to be enhanced at the interface between phenol and water [20].
Other proposed systems acting as oligomerization template are gel matrices [23] or eutectic ice-water mixtures [24,25]: when an aqueous solution is slowly frozen at a temperature above the eutectic-point temperature, solutes are excluded from the ice crystals and remain within the liquid fraction (eutectic phase), resulting in the concentration of reaction mixtures. According to this view, polymerization could have occurred, rather than by adsorption to ice crystal surfaces, in these cavities between ice crystals, filled with highly concentrated solutes in liquid phase, thanks to a dehydration effect.

We conclude this section by quoting two of the most authoritative scientists in the prebiotic field, C. De Duve and G.F. Joyce, both struggling to fill the gap between the *clutter* (both use this term) of organic molecules and the ordered polynucleotides: a robust explanation is still missing. We think that our research on short DNA ordering, of which we are going to describe some recent progresses, can make a new step toward the right direction.

"How RNA could possibly have emerged from the clutter without a "guiding hand" would baffle any chemist.[...] It seems possible only by selection, [...] a process that presupposes replication." C. De Duve [1];

"The chief obstacle to understanding the origin of RNA-based life is identifying a plausible mechanism for overcoming the clutter wrought by prebiotic chemistry." G.F. Joyce [4].

6.2 sDNA behavior: a good model for DNA formation

The phenomena that we have described - pairing, segregation, stacking and LC ordering of short duplexes - can provide in a single shot a number of distinctive features:

- a mechanism of selection based on flexibility: only paired strands segregate to form LCs;
- a locally enhanced concentration of complementary strands, up to 10^3-fold, and of terminal concentration - see estimate in chapter 4);
- a liquid crystalline matrix which provides a flexible template, possibly favoring chemical ligation;
- a process of selection based on length: longer and longer helices are favored, because they can more easily fit in the aligned environment.

The association of some of these features had already emerged in studies devoted to surfactant living polymerization or fibrillar assembly. As already discussed, in a solution of rods the greater the particle axial ratio, the lower the volume fraction at which spontaneous, entropically-driven spatial ordering sets in [26]. For a given volume fraction, ordering is more advantageous when there are fewer, longer particles than when there are more, shorter particles. The important point here, learnt from living (reversible) polymers [27], is that the reciprocal will also be true if the particles are able to reconfigure themselves: if they stay out of each others' way by orientational or positional alignment, then it will be easier for them to compose themselves into fewer, longer rods than it would be in the absence of order. Thus, in systems capable of aggregation, ordering promotes self-assembly and both degrees of freedom, spatial and self-assembling, will jointly equilibrate to minimize the free energy. This mutual coupling of orientation and polymerization can result, even in the absence of inter-chain interactions, in a very wide two-phase region in which a very dilute isotropic phase consisting of relatively short molecules coexists with a very concentrated and very highly ordered anisotropic phase consisting of very long molecules [28].
We could also make one more step backward: we know (see section 2.3.1) that GMP, while showing only a weak propensity to stack as a mono-nucleotide, can form quartets that, destabilized because of the enlarged hydrophobic surface, are instead strongly prone to the formation of stacked aggregates and LC organization. The "switch" of the stacking advantage, leading to more and more elongated assemblies, is the onset of H-bonds between the bases, that creates a surface too large to allow solubility.
In this picture, the molecules that best self-structured by a cascade of pairing and stacking in ordered structures prone to chemical ligation, enabled self replication and hence information storage.

6.3 Other LC-related models of biological organization

It's worth to mention that the one just described is not the only case where flexibility-based phase separation and/or spontaneous LC ordering have been recognized as appropriate frames for biological complex behaviors, involving either DNA or other cell constituents.
We have already mentioned, in section 1.4.2, the high concentration of DNA inside cell nucleus, its liquid crystalline arrangement and the possible role of DNA protection and repair held by the cholesteric phase of the DNA itself in bacterial cells exposed to damaging factors [29].
The development of the eukaryotic kingdom was accompanied by two major events related to macromolecular crowding: a substantial increase in the amount of DNA and its confinement within a space much more defined than in prokaryotic cells. These two events necessarily resulted in an extreme DNA crowding which - as discussed above - could strongly promote DNA self-assembly processes and the formation of mesophases. Whereas crowding and the formation of a cholesteric phase are considered not necessarily impeding DNA biological functions, the denser columnar phase obtained at higher DNA concentrations could completely suppress them.
It was thus suggested [30] that, just because of the strong tendency of DNA to self-order, the nucleosome assembly was rendered essential to counter-balance this tendency by introducing a steric barrier to such a collapse: the hierarchical assembly of DNA would thus be a consequence of the delicate balance between LC ordering and functional needs.

Another important example comes from the self-assembly mechanisms of collagen and chitin, respectively the protein and the polysaccharide forming fibers of animal connective tissues, and of cellulose, a polysaccharide present in plant cell walls. Collagen and chitin predominate in the extra-cellular matrix and support the long-range 3D organization of their cells and organisms.
Such macromolecules are produced by specialized cells, excreted into the extracellular space in a dense state, and then they escape cellular control to self-organize into highly ordered biological tissues. Their intrinsic ability to self-assemble and spontaneously form complex ordered networks, without the intervention of cells, is demonstrated by the fact that these molecules are able to reassemble *in vitro*. For example, collagen solutions form cholesteric phases; the chiral fluid phase can be stabilized by an increase in pH and the resulting dense fibrillar gels still exhibit the helical symmetry observed *in vivo* in bones. Therefore, a structural similarity is suggested between their supramolecular arrangement in connective tissues and the LC ordering [31]. Although connective tissues exhibit no or very little fluidity in their functional state, an ordered fluid state (cholesteric) could have played a role at some point in their morphogenesis, thus imprinting chirality on the solid tissue formed outside the cell, composed of triple-helices.

The cited work by Judith Herzfeld and coworkers about living polymerization, phase separation and ordering (see e.g. [32]) was originated by the description of the independent organization of microtubules and F-actin inside the

crowded cell interior. Indeed, the cytoskeleton is a structured system of protein filaments that plays an important role in cell morphology, rheology, motility, division and intracellular transport. The cytoskeleton is highly dynamic, with components constantly disassembling and reassembling in new arrangements throughout the cell. Among its constituents, tubulin forms hollow microtubules with 13 monomers around the circumference (a relatively thick and stiff structure), while actin forms microfilaments comprising a twisted double strand of monomers, a relatively thin and flexible structure. Filaments of such types were shown to phase separate and independently align, forming N phases, just because of entropic reasons [32].

Microtubules, among other functions, play an important role during the cell division. At first sight, cell mitosis appears to be an over-complicated process (see figure A.26 in Appendix A): compacted chromosomes move into position in a plane across the center of the cell; at some signal, all the chromosomes break into two daughter chromatids, which then separate and move, as if they were being pulled by invisible strings fastened to their centers, towards "poles" at opposite ends of a "spindle". Finally, when the two sets of chromatids are separated, the cell breaks into two daughter cells, the chromosome progressively unwinds until the strands become so fine that they become indistinguishable again. One of the most impressive features of this process is that in some cases it is completed in times as short as four minutes.
John Lydon [34] suggested that in the whole process there is the evidence of a LC behavior: the way in which the cytoplasm is structured yet fluid, the spontaneous appearance of the poles, the way in which the chromosomes "know" where to assemble. The key players in this process appear to be self-assembling mesogenic units, the microtubules, too small to be seen directly in the optical microscope, but whose arrangement can be highlighted with fluorescent dyes. In Lydon's model, all the cytoplasm (microtubules plus other smaller molecules) becomes a single domain of a lyotropic nematic phase, with two +1 disclinations that spontaneously move apart to equilibrium positions forming the bipolar spindle structure. In this director field pattern, alien bodies will drift towards the equatorial plane and then move toward the poles in exactly the same way that polymer strands tend to accumulate at disclinations.

PART II

6.4 RNA liquid-crystalline phases

Because of its relevance to prebiotic studies, and to compare its behavior with DNA, we turned our attention to RNA or, better, to sRNA too. As described in section 1.1.2, RNA shares most of its features with DNA so that we can describe it as a stiff, charged rod with hydrophobic ends. As regards helical organization, two main differences hold: RNA is characterized by an enhanced stability of duplexes [35] and it usually adopts an A-conformation.
To investigate the effect of these differences on the ability to form LC phases,

we selected three short self-complementary sequences corresponding to those studied in DNA:

- CGCAUGCG
- CGCAAUUGCG
- CGCGAAUUCGCG

In polarized microscopy, aqueous solutions of such sequences with a concentration gradient show textures very similar to DNA oligomers. N phase is, however, limited to a very narrow region between Isotropic and C_U: although we haven't measured yet the concentration in the various phases, we can infer that this corresponds to a narrow range of concentrations of RNA in N. No textures corresponding to C_2 were observed in the 8 bp strand. We show some textures in figures A.27, A.28 and A.29 (color plates in Appendix A).

The temperature of transition from both sRNA LC phases to Iso is moderately lower than in the case of DNA (chapter 4), in the order of 10 degrees. Since RNA duplexes have higher duplex stability, the opposite trend could have been expected. However, other RNA features could account for the observed behavior: first, the rise per bp in A-form is about 2/3 of that in B-form, so that we should probably compare the behavior of RNA duplexes with 6 and 8 bp DNA oligos instead of 8, 10 and 12 bp. As we can infer from the sDNA phase diagram, shorter oligos form less stable LC phases.
In addition, in A-form bases are much more tilted than in B and this asymmetry could disturb the end-to-end stacking, by limiting the allowed relative orientations of neighboring helices and therefore reducing the overall stability of the complexes. Actually, it is known from literature data that the dodecamer studied in our experiments assumes an A form [36], while we lack evidence of the conformation of the shorter sequences (and reciprocally we can't exclude that also in DNA oligomers the A-form prevails over B, when concentration is high and thus hydration is low, except for the extensively studied Dickerson dodecamer).
The observed behavior could be the result of the composition of such different trends. Careful structure characterization in both nucleic acids and concentration measurements in sRNA will be necessary to address these points.

To explore the possible phase separation mechanisms in RNA, we also mixed the RNA self-complementary dodecamer with a mixture composed of *ss-DNA* and we repeatedly observed the nucleation of crystals, like those shown in figure A.30 (Appendix A). Although we cannot prove that *only* RNA is in the crystals, we can certainly conclude that it is the major component of the observed aggregates. In fact, the *ss-DNA* alone doesn't show neither crystal nor liquid crystal phases, and RNA alone at similar concentrations doesn't display any birefringence. Thus, some kind of phase separation is acting, favored by the presence of *ss-DNA* . Again, further experiments are needed, with RNA single strands, to verify in RNA the observed behavior of *ds*/*ss*-DNA systems.

6.5 Mixtures of helices of different length

Inspired by our results of chapter 5 and by the phase behavior predicted by Herzfeld [32] (chains of different size and flexibility independently segregate in ordered phases), we decided to investigate the potentiality of the combined effect of phase separation, stacking and LC ordering in the "purification" of helices. To this aim, we worked on more complex mixtures, with (self-)complementary sequences of different length and non-matched single strands, to get a system of two different helices and single stranded coils. In this system, however, we have one more degree of freedom, the temperature dependent denaturation/renaturation process of the helices, that transforms two flexible coils into a single rigid rod, and thus allows to control the relative amount of the populations.

6.5.1 A+B+C+D

The natural extension of A+B system (section 5.2) is obtained by the addition of two other mutually complementary oligos, CCTTCCAAAACCTTCC (C) and GGAAGGTTTTGGAAGG (D). Due to their sequences, A can only form a stable helix with B, not with C or D, and the same holds for the others. By mixing the four strands, either in unbalanced ratios or adding to them a mixture of *ss-DNA* , we actually obtain two kind of helices, of length 12 and 16, in a bath of coils. In both cases, below T_{LC}, LC droplets appear, in fractions similar to the helices volume fractions (figure A.31, color plate in Appendix A). However, since there is a small gap between the temperatures of renaturation of the A-B and C-D couples, we couldn't distinguish the onset of one species from the other. Also, doping the A species with Fluorescein-labeled strands, we couldn't measure any difference in the fluorescence of the various domains, suggesting that in the system, in our experimental conditions, the LC phases are composed of equally distributed 12 and 16 long helices.

6.5.2 6bp+20bp

A deeper insight can be obtained by studying a second system: self-complementary 6 bp and 20 bp duplexes were mixed, alone or with the addition of *ss-DNA* . In this way, since the melting temperatures of the two strands are pretty different, we also have the possibility to "switch on" the presence of helices at various temperatures.
The knowledge of the absolute melting temperature of the two species is impaired, because we can't precisely control neither DNA nor salt concentration. However, we can estimate $T_{m-20} \simeq T_{m-6} + 25°$ [37]. This gap equals the difference in thermal stability of the LC phases of 6bp and 20bp alone, i.e. between the temperatures T_{Iso-LC} as reported in chapter 4).
If oligos of the two different lengths are mixed together without *ss-DNA* , so that each strand present in the cell can bind to helices, at low temperature they form continuous LC phases in the whole volume, similar to those showed by the single components. This observation is consistent with the end-to-end stacking mechanism, which doesn't distinguish among helices of different lengths and can arrange them together in the same chains; it also corresponds to the observed continuous phase behavior of sDNA mixed with long DNA, as shown in section

4.2.1.
When cooling down the solution from isotropic, the temperature of appearance of the first LC domains, T_{Iso-LC}, though depending on the phase and on the concentration of DNA in the sample, is intermediate between those observed in solutions of the single components, namely around 30 degrees for Nematic (it is 50° for 20bp only and 15° for 6bp) and 40° for Columnar (vs. almost 75° for 20bp and 35° for 6bp). Therefore, we can reasonably argue that at these temperatures the first LC domains can be ascribed to the 20 bp species only. By further cooling below T_{Iso-LC}, the existing domains grow bigger and some new ones appear, until at some stage, not clearly determined, the helices of the second species begin to form and add to the already formed domains. Finally, at the lowest temperatures, the whole sample is filled with LCs. There is no discontinuity between the growth of the first domains and the final condition: we thus lack the unambiguous evidence that the formation of the helices (and the consequent phase separation) is well separated for the two species. However, reheating the sample leads to the melting of the LC domains in a reverse order: the domains that first appeared are the last to fade away. This is again consistent with the fact the most stable domains are only composed of the most stable helices, i.e. of the longest.

In a sample with 6bp, 20bp and mixed *ss-DNA* , kept just below T_{Iso-LC}, nucleation of a few small domains occurs. Since all the longer helices are already there, in principle all of them can join the nuclei but, given the high viscosity of the solution, long time is needed for each of the present helices to diffuse and aggregate to LCs. Indeed, after hundreds of hours, these domains eventually reach a plateau in their size, around 100 μm in radius. Figure 6.4 shows the time evolution of such domains, together with the fraction ϕ_{LC} of LCs in cell volume. As in chapter 5, we can estimate the volume fraction ϕ_{20} occupied by the 20 bp helices in the cell (this estimate in affected by the approximations in calculating the volume of single stranded coils, see discussion in the cited chapter). The two values, the plateau of ϕ_{LC} and ϕ_{20}, agree within the experimental error. We can conclude, therefore, that the domains are almost entirely formed by 20 bp strands and the possibility of selective phase separation is demonstrated.

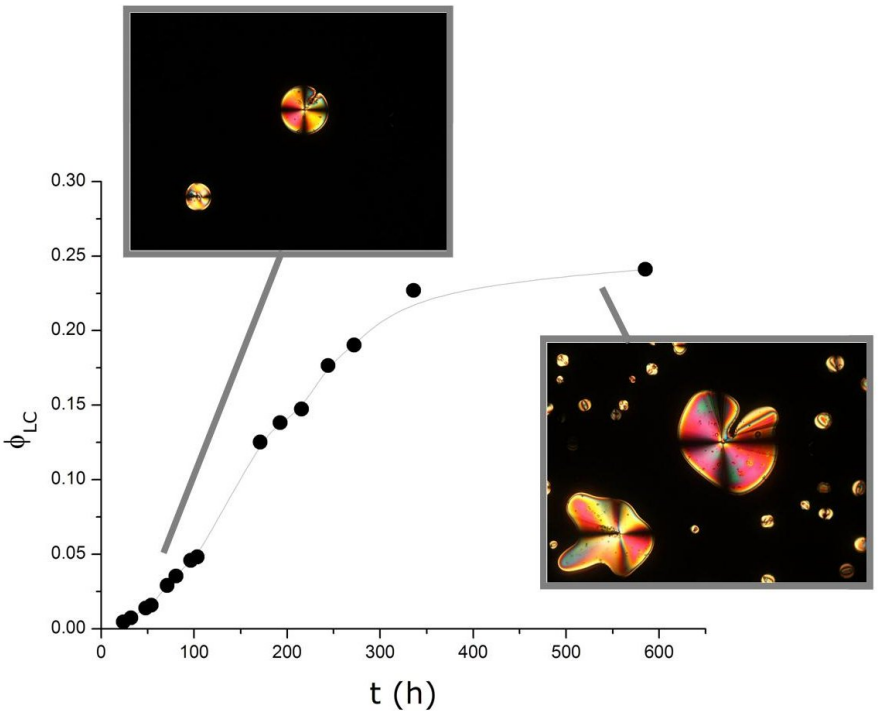

Figure 6.4 - LC fraction vs. time in a mixture of 6 bp and 20 bp duplexes with *ss-DNA* , kept at $38°C$. After hundreds of hours, ϕ_{LC} reaches a plateau, consistent with the 20 bp occupancy. Line connecting the data points is a guide to the eye.

6.6 Random pools

We don't have evidence, yet, that single nucleotides or oligonucleotides shorter than 6 base pairs show any propensity to LC ordering because of stacking. The first challenge is to observe pairing, phase separation, stacking and LC ordering - and eventually chemical ligation - of complementary oligos arising from a random pool of short sequences.
To start approaching this goal, we performed a series of experiments on sequences of different lengths and with different fractions of random nucleotides - with the same 1/4 probability for each base - ranging from 0 to 1.

6.6.1 Experimental observations

We already know, from the results reported in the previous chapters, that complementary DNA sequences, as short as 6 bp, exhibit LC behavior at high concentrations.
At the opposite extreme, we studied completely random sequences (only of length 8, so far). None of the solutions of such sequences, though highly concentrated, shows any LC phase. Even when seeded with self-complementary sequences at concentrations high enough to form LCs, random sequences do not join liquid crystalline nuclei.
We also worked on sequences having some determined and some random bases. In this way, we can progressively reduce the number of possible sequences (and thus increase the effective concentration of each of them - see below) and enhance the helix stability by adding CG pairs. The sequences here studied belong to the following classes (according to the standard nomenclature, we call n a random nucleotide), somehow related to the sequences studied in chapter 4:

- CGn...nCG: with terminal CG at both 5' and 3' and central random nucleotides;
- Cn...nG: with C at 5' and G at 3', and central random nucleotides;

In figures A.32 and A.33 (color plates in Appendix A) we report the textures observed in concentrated solutions of the first class of sequences, CGn...nCG, for the shortest studied length of the random tract. For both this sample and the longer one, a close resemblance with the previously observed phases is evident, with LC columnar domains emerging from an isotropic background. However, cholesteric phase is not present, like in A-B mixtures with $[B]/[A] > 3$ (see chapter 4). Another behavior analogous to A-B samples is that, depending on the total DNA concentration (not measured yet), there is a fairly broad range in the volume fractions occupied by LCs.
The thermal stability of the two sequences is similar, a few degrees lower than self-complementary 8bp and 10bp LC-Iso temperature. This is consistent with the lower average stability of helices.
In the second class of sequences, no phases appeared in highly concentrated solutions of CnnnnnnG, while for CnnnnnnnnG textures similar to the first class (shown in figure A.34 in Appendix A) were observed; such LC phases exhibited a LC-Iso transition temperature more than 10 degrees lower than in self-complementary 10bp.

What is the difference between the self-complementary and the random sequences of length 8? And between CnnnnnnG and CGnnnnCG? To interpret these results in the light of thermodynamic and statistical analysis, we start by considering a completely random sequence.

6.6.2 Random sequences

In a batch composed of random nucleotides, in principle each of them has its complementary sequence and is able to find and bind to it, but three main questions can arise:

1. Are we sampling all the possible sequences?
2. Are these sequences actually stable in helical form?
3. How important are the partial alignments (and thus the unpaired tails)?

1) Are we sampling all the possible sequences?

It's easy to answer "yes" to the first question. For example, there are $4^8 = 65536$ different possible sequences composed of 8 nucleotides, 4^4 of which are self-complementary. In the standard amount of DNA used in the experiments (3 O.D.= 0.1 mg), since the average molecular weight of an 8-mer sequence is $330 \cdot 8 = 2640$ g/mol, there are $N_{tot} \approx N_A 10^{-4}/2640 \simeq 2.3 \cdot 10^{16}$ 8-mer sequences. Thus, in every cell we have $N_{seq} \approx 2.3 \cdot 10^{16}/(6.5 \cdot 10^4) \approx 3.5 \cdot 10^{11}$ oligomers for every possible sequence. The number fluctuation can be estimated as $\sim \sqrt{N_{seq}} \simeq 6 \cdot 10^5$.

Before addressing the two other questions, let's consider the equilibrium between each of the duplexes and the single strands from which it is formed; their coupling can be modeled as a bimolecular reaction:

$$[rr'] = K[r][r'] = K[r]^2 \tag{6.1}$$

where we have assumed $[r] = [r']$ (true, on average, for every sequence); the binding constant K is related to the free energy difference by the usual relation:

$$\Delta G = -RT \ln K \tag{6.2}$$

An accurate prediction of the stability of a helix, as seen in section 1.2, requires both H-bonds and stacking contributions, and needs to incorporate correlations among neighboring nucleotides.
To calculate thermodynamic parameters, in the following we'll use the approximated equation 1.12 and NN parameters from [37], one of the most updated and statistically significant sets of parameters. Several other sets are reported in literature, and the reliability and accuracy of the NN model itself are questionable for very short duplexes (for the detailed comparison of the sets, the analysis of the discrepancies and the introduction of next-nearest-neighbors parameters, see [38]). Nonetheless, the general predictions of different models are similar, and we're more interested in the analysis of the trends with oligomer length and salt concentration than in the absolute numbers, that are hard to be

experimentally verified in our conditions.
A particular comment is required for the dependence on salt concentration in equation 1.12: it has a strong influence on the stability of duplexes and it is thus able to shift the equilibrium constant of several orders of magnitude. In our experimental conditions, as described in chapter 4, we have no added salt but a very high counterion concentration; we therefore set, in all the calculations of free energy, T=25 $°C$, $c_{DNA} = 700mg/ml$ (a typical concentration measured in LC phases of self-complementary sDNA in 6-20 bp range) and [salt]=1 M.

2) Are these sequences actually stable in helical form?

We can now go back to our second, important question, whether all, or at least most, of the possible helices are formed.
According to the calculations introduced above, an average ΔG of -9.8 kcal/mol is found for a random 8-mer. Thus, with an equilibrium constant $K = 1.6 \cdot 10^7 mol^{-1}$ and DNA at 700 mg/ml, corresponding to about 0.26 M, the effective molar concentration of each sequence is $0.26/4^8 \simeq 4{\cdot}10^{-6} mol^{-1}$, and thus the ratio between bound and free strands, given by $[bound]/[free] = [rr']/[r] = K[r']$, is around 64, i.e. the equilibrium balance is favorable to pairing. As already observed, the absolute values of such kind of estimates strongly depend on DNA and salt concentrations, experimental parameters that we can't easily control in the ultra-small samples. Indeed, the same calculation as above, with $c_{DNA} = 700mg/ml$ but [salt]=0.1 M, would yield [bound]/[free]=3.2, a value much closer to the critical value of 1.
It is therefore more useful and reliable to study, for fixed values of concentrations, the dependence of stability on N. Since a random sequence is made of "average" mononucleotides, the absolute value of the (negative) free energy grows linearly with N. Therefore, there's a competition between the decreasing effective concentration (whose leading term is 4^{-N}) and the exponentially increasing binding constant. With our parameters, this product results in a monotonously growing function of N in the whole range of interest ($N > 1$), i.e. stability is higher for longer random helices than for the shorter ones. ΔG and $[bound]/[free]$ are plotted in figures 6.5 and 6.6) vs. N.

3) How important are the partial alignments?

Let's now consider a given random N-mer sequence r (e.g. ACGTGCAA): there is only one complementary sequence r' (TTGCACGT) and many more partial partners, i.e. sequences that have a partial possible pairing to the first one (like, for example, TTGCTGGT, AATTGCAC or GTGCACGA), possibly competing with the main binding. We can divide them in three main classes.

a) Among these, those with internal mispairings (such as ACGTGCAA and TTGCTGGT) are 3^i, with i the number of mispairings. We can assume that they are either stable enough to behave equivalently to our sequence, or too unstable to survive. It's a hard task to establish a threshold value of i between these two situations, because stability also depends on specific sequence and on how many consecutive mispairings are present [39].
For these reasons, we won't consider them in the stability balance. Taking

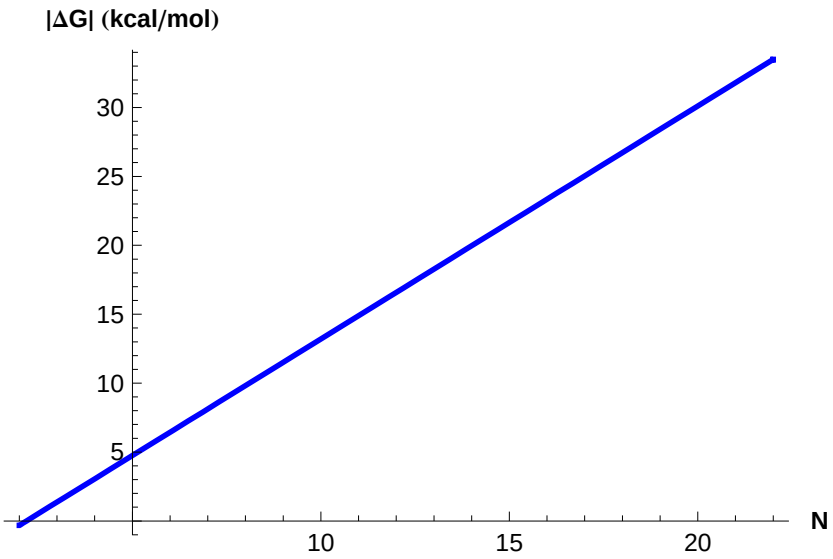

Figure 6.5 - Dependence of the free energy of duplex formation on N for random sequences.

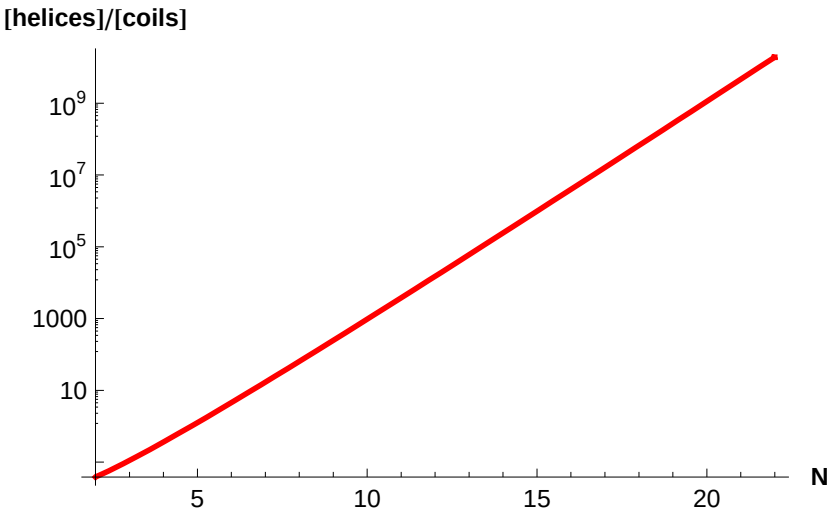

Figure 6.6 - Log plot of the bound-to-free strands ratio VS N for random sequences.

them into account will be important and could be critical to build a complete description of the overall equilibrium behavior of a pool of random oligos, but we're now going to estimate an upper value for the competitors for stacking and therefore we won't consider them, just keeping in mind that we're not counting helices virtually equivalent to the complete ones.

b) We focus, instead, on those sequences that have a correct pairing on one side of the strand and unpaired nucleotides on the other, so that the resulting complex is a central core and flexible tails on both sides - and then a disturbance to the stacking of different oligomers, as seen in section 4.4.2. An example is the couple ACGTGCAA and AATTGCAC.
Such partial duplexes have, on average, a smaller contribution from pairing than completely paired sequences. However, it has been found that dangling ends have an overall stabilizing effect on the duplex, due to the additional stacking of the dangling nucleotide on the last complete base pair. In [39] additional parameters for this effect are reported; they show that only the first dangling nucleotide actually decreases the free energy, and this contribution is usually not enough to compensate for the lost Watson-Crick pairing.
Since the end contributions, both the initiation penalty and the dangling end advantage, are equal for all the lengths, the difference between the free energies of a N-mer and a *(N-j)*-mer is proportional to the difference in the number of nucleotides, j, that is also the number of nucleotides in the flexible terminal "tails".
The number of these partial partners can be obtained from all the possible sequences different from the "right" one, either at 5' or 3' end, in j positions, $2(4^j - 1)$.
We have modeled the equilibrium of a sequence r with its unique fully paired partner r' with equation 6.1; analogously, we describe its equilibrium with the partial partners r_j as different chemical species with their binding constant K_{rj}:

$$[rr_j] = K_{rj}[r][r_j] = K_{rj}[r]^2 2(4^j - 1) \quad j = 1, 2, ...N - 1 \tag{6.3}$$

Thus, to estimate the importance of partial alignments relative to complete helices, we combine a Boltzmann factor for the energy balance with the statistical weight of the number of sequences, dividing eq. 6.3 by eq. 6.1:

$$\frac{[rr_j]}{[rr']} = \frac{K_{rj}}{K}\frac{[r_j]}{[r]} = 2(4^j - 1)\exp\left(\frac{\Delta G - \Delta G_j}{RT}\right) \tag{6.4}$$

The exponent in the right-hand side of the equation is just the average NN energy times j, plus the average dangling contribution, that is a constant. This ratio is therefore independent from N, and for $j = 1, 2, 3$ it assumes respectively the values $1.2, 0.35, 0.08$: random helices with tails of length 1 are equally important as complete helices, while longer tails have less and less importance (and we won't consider $j > 3$ because their contribution becomes negligible).

c) We should finally consider terminal mispairings at *both* ends - without shifts - because they can equally disturb the stacking (see section 4.4.2): for instance, such a couple is ACGTGCAA and GTGCACGA.
Their number can be estimated as 3^{2t}, with $t = 1, 2$ the length of the mispaired terminal tract. Although the contribution to the overall stability of terminal

mismatches is not as well characterized as for internal mismatches [39], we can estimate the free energy of sequences. Thus, the balance between mismatched and fully paired sequences, analogously to equation 6.5, is given by the ratio of equilibrium constants times the ratio of statistical weights. Calculations yield $[rr_1]/[rr'] \simeq 0.3$ and $[rr_2]/[rr'] \simeq 0.01$: the balance is thus strongly displaced toward the fully paired helices.

6.6.3 Partially random sequences

We now give a detailed analysis of the stability of the duplexes formed by partially random sequences and of their competitors. Among the terminal mispairings, only those with dangling ends are possible in such sequences.
Relative to completely random strands, they are all characterized by a reduced number of possible sequences and enhanced helix stability.

CGn...nCG

Among the sequences we studied, the closest to determined ones is CGnnnnCG, since the terminal CGs provide stability to the helices and there are relatively few possible sequences, $4^4 = 256$. Its ΔG is -12.1 kcal/mol and the [bound]/[free] ratio is around $8{\cdot}10^5$. For ΔG and [bound]/[free], the same trend with N holds as for completely random sequences. In fact, for CGnnnnnnCG there are $4^6 = 4096$ possible sequences, with an average ΔG of -15.5 kcal/mol. This corresponds to a bound-to-free ratio of $1.2 \cdot 10^7$. Indeed, the addition of fixed terminal CGs enhances the stability of the helices and their effective concentration.
In addition, contrary to the case of completely random sequences, partial mispairings are disfavored in such kind of sequences, because they have to fit to fixed ends. For this reason, only a subset of the possible sequences can stably bind to a partner to form a helix with flexible tails; this fraction is expressed as f_{seq}. Also, within this subset, the number of possible equivalent partners n_{part} depends on the length of the tail. These numbers combine to give the relative concentration of possible competitors. The ratio $[p]/[c]$ between partial and complete helices is obtained, as above, as the ratio of the binding constants times the ratio of molar concentrations:

$$\frac{[p]}{[c]} = \frac{K_{rj}}{K}\frac{[r_j]}{[r]} = f_{seq} n_{part} \exp\left(\frac{\Delta G - \Delta G_j}{RT}\right) \tag{6.5}$$

For the different lengths of the tails and for various total lengths of the strands, tables 6.1, 6.2 and 6.3 report the free energy, calculated as above according to [37], the relative concentrations and the ratio [p]/[c].
We also note that this kind of mispaired sequences - with terminal CG - would *favor*, and not disfavor, linear aggregation between different helices and thus LC formation (as in the case of CG12, see section 4.4.2).

Cn...nG

A higher degree of randomness is represented by this second class of sequences. For CnnnnnnG, $\Delta G = -10.6$ kcal/mol and [bound]/[free] is about 4100, while

N	ΔG	[p]/[c]
6	-4.2	$2.5 \cdot 10^{-4}$
8	-7.6	$2.5 \cdot 10^{-4}$
10	-11.0	$2.5 \cdot 10^{-4}$
12	-14.3	$2.1 \cdot 10^{-4}$

Table 6.1 - Free energy (in kcal/mol) and relative abundance of helices with j=1 mispairings in sequences of the class CGn...nCG; the subclass is of the type: CGn...nCCG/CGGn...nCG, $f_{seq} = 1/2$ and $n_{part} = 1$.

N	ΔG	[p]/[c]
6	-4.7	$1.5 \cdot 10^{-4}$
8	-9.5	$1.5 \cdot 10^{-3}$
10	-12.8	$1.3 \cdot 10^{-3}$
12	-16.2	$1.3 \cdot 10^{-3}$

Table 6.2 - Free energy (in kcal/mol) and relative abundance of helices with j=2 mispairings in sequences of the class CGn...nCG; the subclass is of the type: CGn...nCGCG/CGCGn...nCG, $f_{seq} = 1/8$ and $n_{part} = 1$.

N	ΔG	[p]/[c]
6	-4.4	$3.5 \cdot 10^{-4}$
8	-7.8	$3.5 \cdot 10^{-4}$
10	-11.2	$3.5 \cdot 10^{-4}$
12	-14.5	$3 \cdot 10^{-4}$

Table 6.3 - Free energy (in kcal/mol) and relative abundance of helices with j=3 mispairings in sequences of the class CGn...nCG; the subclass is of the type: CGn...nCGnCG/CGnCGn...nCG, $f_{seq} = 1/8$ and $n_{part} = 4$.

N	ΔG	[p]/[c]
6	-6.2	0.09
8	-9.6	0.09
10	-12.9	0.08
12	-16.3	0.08

Table 6.4 - Free energy (in kcal/mol) and relative abundance of helices with j=1 mispairings in sequences of the class Cn...nG; the subclass is of the type: Cn...nGG/CCn...nG, $f_{seq} = 1/2$ and $n_{part} = 1$.

N	ΔG	[p]/[c]
6	-4.5	0.02
8	-7.9	0.02
10	-11.2	0.02
12	-14.6	0.02

Table 6.5 - Free energy (in kcal/mol) and relative abundance of helices with j=2 mispairings in sequences of the class Cn...nG; the subclass is of the type: Cn...nGnG/CnCn...nG, $f_{seq} = 1/2$ and $n_{part} = 4$.

for CnnnnnnnnG $\Delta G = -14$ kcal/mol and [bound]/[free] is about 61000.
In tables 6.4, 6.5 and 6.6 the equilibria between complete and partial pairing are given, showing that complete pairing is strongly favored.

6.6.4 Summary

To summarize and compare the thermodynamic and statistical features of the various sequences, as calculated in the previous paragraphs, we can introduce a quality parameter, or score, to globally evaluate them:

$$S = \log\left(\frac{[bound]/[free]}{1 + \sum_j N_j + \sum_t N_t}\right) \tag{6.6}$$

where $\sum_j N_j$ and $\sum_t N_t$ are the total numbers of partially mispaired sequences (respectively with and without dangling ends) for every fully paired helix ($N_{j,t} = [p]/[c]$ for the different values of j, t and N).

N	ΔG	[p]/[c]
6	-2.8	$5 \cdot 10^{-3}$
8	-6.2	$5 \cdot 10^{-3}$
10	-9.6	$5 \cdot 10^{-3}$
12	-12.9	$4 \cdot 10^{-3}$

Table 6.6 - Free energy (in kcal/mol) and relative abundance of helices with j=3 mispairings in sequences of the class Cn...nG; the subclass is of the type: Cn...nGnnG/CnnCn...nG, $f_{seq} = 1/2$ and $n_{part} = 16$.

Sequence	N	ρ	S	Symbol
CGCG	4	0	3.5	A
CGATCG	6	0	4.7	B
CGCATGCG	8	0	8.5	C
CGCAATTGCG	10	0	10.2	D
CGnnCG	6	0.33	4.8	-
CGnnnnCG	8	0.5	5.7	E
CGnnnnnnCG	10	0.6	7.1	F
CGnnnnnnnnCG	12	0.67	8.3	-
CnnnnG	6	0.67	2.4	-
CnnnnnnG	8	0.75	3.6	G
CnnnnnnnnG	10	0.8	4.7	H
CnnnnnnnnnnG	12	0.83	5.9	-
nnnnnn	6	1	0.3	-
nnnnnnnn	8	1	1.3	I
nnnnnnnnnn	10	1	2.5	-
nnnnnnnnnnnn	12	1	3.7	-
nnnnnnnnnnnnnn	14	1	4.9	-

Table 6.7 - Summary of the predicted stability, free energy and score, of families of sequences ranging from N=4 to N=14 and to $\rho = 0$ to $\rho = 1$; letters are assigned to the sequences represented in figure 6.6.

S, therefore, represents (the logarithm of) the effective number of stable, complete helices free of competitors in the given conditions. Indeed, if $\sum_j N_j + \sum_t N_t \ll 1$, i.e. the total number of disturbing helices is negligible, $S \equiv \log([bound]/[free])$.

Table 6.7 and figure A.35 (in Appendix A) report the calculated values of S for the families of strands described above. Each sequence has a position on the plane, that corresponds to its length and degree of randomness. Since in the various families of sequences we have a fixed number of defined bases and variable total length, we defined the randomness parameter as $\rho = (N - d)/N$, with d the number of defined bases: it goes from 0 to 1 when d goes from N to 0. Although there isn't this ambiguity in the sequences we're presenting, this definition of randomness does not account for the *position* of defined/random nucleotides: different sequences having a different order of bases along the sequence - thus a different stability, and consequently a different score - can be in the same point of the (ρ,N) plane.
Figure 6.6 also gives the phase behavior of the experimentally studied strands. The two key observations are:

1. the lower is the randomness, the lower is the length necessary to have high values of S;

2. a correlation holds between the quality parameter S and the phase behavior: sequences with $S \gtrsim 5$ show LC behavior, while those with lower values don't.

Despite the still restricted range of experimental results, a clear trend is emerging from the presented results: whenever a sequence (or a pool of sequences) is stable enough, because of the length and/or the composition, and the possible mispaired helices are not numerous or stable enough to overcome the complete ones, we can observe a strong tendency of short oligos to align in LC phases.

As previously observed, a lower salt concentration would shift downward the whole estimate of S values, so that its physical meaning is not completely transparent. It is striking, however, that this simple parameter is able to catch the main features of the behavior of sequences spanning a wide range of composition and length. The result is consistent with the fact that 4bp self-complementary strand ($\rho = 0$, S= 3.5) doesn't show phases. It also suggests that CGnnCG, a sequence not yet characterized, will.

More significantly, it predicts that a completely random sequence would require a length above 12-14 nucleotides to LC order; ongoing experiments are checking these predictions, together with the effect of the position of random nucleotides. If these predictions will turn out to be true, the suggested scenario would be that random sequences need, to show liquid crystalline phases, to be long enough to form stable helices. In that case, the proposed mechanism of stacking and alignment would start to apply to already formed, though short, strands, and not to single nucleotides.

Bibliography

[1] C. de Duve, *Singularities - Landmarks on the pathways of life*, Cambridge University Press (2005)

[2] P. Stano, P. L. Luisi, *Basic Questions About the Origins of Life: Proceedings of the Erice International School of Complexity (Fourth Course)*, Orig. Life Evol. Biosph. **37**,303 (2007)

[3] J. W. Szostak, D. P. Bartel, P. L. Luisi, *Synthesizing life*, Nature **409**,387 (2001)

[4] G. F. Joyce, *The antiquity of RNA-based evolution*, Nature **418**,214 (2002)

[5] W. Gilbert, *The RNA world*, Nature **319**,618 (1986)

[6] D. P. Bartel, J. W. Szostak, *Isolation of New Ribozymes from a Large Pool of Random Sequences*, Science **261**,1411 (1993)

[7] K. Ozawa, A. Nemoto, E. Imai, H. Honda, K. Hatori, K. Matsuno, *Phosphorylation of nucleotide molecules in hydrothermal environments*, Orig. Life Evol. Biosph. **34**,465 (2004)

[8] H. Ogasawara, A. Yoshida, E. Imai, H. Honda, K. Hatori, K. Matsuno, *Synthesizing oligomers from monomeric nucleotides in simulated hydrothermal environments*, Orig. Life Evol. Biosph. **30**,519 (2000)

[9] S. Miyakawa, J. P. Ferris, *Sequence- and Regioselectivity in the Montmorillonite-Catalyzed Synthesis of RNA*, J. Am. Chem. Soc. **125**,8202 (2003)

[10] K. Grzeskowiak, L. Orgel, *Template-Directed Synthesis on Short Oligoribocytidylates*, J. Mol. Evol. **23**,287 (1986)

[11] R. Shapiro, *Small molecule interactions were central to the origin of life*, Quart. Rev. Biol. **81**,105 (2006)

[12] P. Walde, R. Wick, M. Fresta, A. Marangone, P. L. Luisi, *Autopoietic Self-Reproduction of Fatty Acid Vesicles*, J. Am. Chem. Soc. **116**,11649 (1994)

[13] R. Pool, P. G. Bolhuis, *Prediction of an Autocatalytic Replication Mechanism for Micelle Formation*, Phys. Rev. Lett. **97**,018302 (2006)

[14] P. Baaske, F. M. Weinert, S. Duhr, K. H. Lemke, M. J. Russell, D. Braun, *Extreme accumulation of nucleotides in simulated hydrothermal pore systems*, Proc. Natl. Acad. Sci. USA **104**,9346 (2007)

[15] B. G. Bag, G. von Kiedrowski, *Templates, autocatalysis and molecular replication*, Pure & Appl. Chem. **68**,2145 (1996)

[16] G. Ertem, *Montmorillonite, oligonucleotides, RNA and origin of life*, Orig. Life Evol. Biosph. **34**,549 (2004)

[17] J. P. Ferris, G. Ertem, *Oligomerization Reactions of Ribonucleotides on Montmorillonite: Reaction of the 5'-Phosphorimidazolide of Adenosine*, Science **257**,1387 (1992)

[18] W. Huang, J. P. Ferris, *One-Step, Regioselective Synthesis of up to 50-mers of RNA Oligomers by Montmorillonite Catalysis*, J. Am. Chem. Soc. **128**,8914 (2006)

[19] P. C. Joshi, S. Pitschb, J. P. Ferris *Homochiral selection in the montmorillonite-catalyzed and uncatalyzed Prebiotic synthesis of RNA*, Chem. Commun. ,2497 (2000)

[20] A. Goldar, J. L. Sikorav, *DNA renaturation at the water-phenol interface*, Eur. Phys. J. E **14**,211 (2004)

[21] A. I. Oparin, *The origin of life and the origin of enzymes*, Adv. Enzymol. **27**,347 (1965)

[22] L. Onsager, *Life in the early days*, in *Quantum Statistical Mechanics in the Natural Sciences*, S. L. Mintz, S. M. Widmayer, Eds. (Plenum Press, 1974), pp.1-14

[23] T. R. Samatov, H. V. Chetverina, A. B. Chetverin *Expressible molecular colonies*, Nucleic Acids Res. **33**,e145 (2005)

[24] P. A. Monnard, A. Kanavarioti, D. W. Deamer *Eutectic Phase Polymerization of Activated Ribonucleotide Mixtures Yields Quasi-Equimolar Incorporation of Purine and Pyrimidine Nucleobases*, J. Am. Chem. Soc. **125**,13734 (2003)

[25] P. A. Monnard, *Catalysis in abiotic structured media: an approach to selective synthesis of biopolymers*, J. Cell. Mol. Life Sci. **62**,520 (2005)

[26] L. Onsager, *The effects of shape on the interaction of colloidal particles*, Ann. NY Acad. Sci. **51**,627 (1949)

[27] J. Herzfeld, *Entropically Driven Order in Crowded Solutions: From Liquid Crystals to Cell Biology*, Acc. Chem. Res. **29**,31 (1996)

[28] R. Briehl, J. Herzfeld *Tactoidal state and phase transitions in systems of linear polymers of variable length*, Proc. Natl. Acad. Sci. USA **76**,2740 (1979)

[29] A. Frenkiel-Krispin, S. Levin-Zaidman, E. Shimoni, S. G. Wolf, E. J. Wachtel, T. Arad, S. E. Finkel, R. Kolter, A. Minsky, *Regulated phase transitions of bacterial chromatin: a non-enzymatic pathway for generic DNA protection*, EMBO J.**20**,1184 (2001)

[30] A. Minsky, R. Ghirlando, Z. Reich, *Nucleosomes: a Solution to a Crowded Intracellular Environment?*, J. Theor. Biol.**188**,379 (1997)

[31] E. Belamie, G. Mosser, F. Gobeaux, M. M. Giraud-Guille, *Possible transient liquid crystal phase during the laying out of connective tissues: α-chitin and collagen as models*, J. Phys.: Condens. Matter **18**,S115 (2006)

[32] D. T. Kulp, J. Herzfeld, *Crowding-induced organization of cytoskeletal elements.III. Spontaneous bundling and sorting of self-assembled filaments with different flexibilities*, Biophys. Chem. **57**,93 (1995)

[33] S. L. Rogers, G. C. Rogers, D. J. Sharp, R. D. Vale, *Drosophila EB1 is important for proper assembly, dynamics, and positioning of the mitotic spindle*, J. Cell Biol. **158**,873 (2002)

[34] J. Lydon, *Microtubules: Nature's smartest mesogens - a liquid crystal model for cell division*, Liq. Cryst. Today **15**,1 (2006)

[35] N. Sugimoto, S. Nakano, N. Yoneyama, K. Honda, *Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes*, Nucl. Acids Res. **24**,4501 (1996)

[36] S. Chou, P. Flynn, B. Reid, *Solid-Phase Synthesis and High-Resolution NMR Studies of Two Synthetic Double-Helical RNA Dodecamers: r(CGCGAAUUCGCG) and r(CGCGUAUACGCG)*, Biochemistry **28**,2422 (1989)

[37] J. Santalucia, *A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics*, Proc. Natl. Acad. Sci. USA **95**,1460 (1998)

[38] R. Owczarzy, *PhD Thesis*, University of Illinois at Chicago (1999)

[39] J. Santalucia, D. Hicks, *The Thermodynamics of DNA Structural Motifs*, Annu. Rev. Biophys. Biomol. Struct **33**,415 (2004)

APPENDIX A

Color tables

Figures of chapter 3

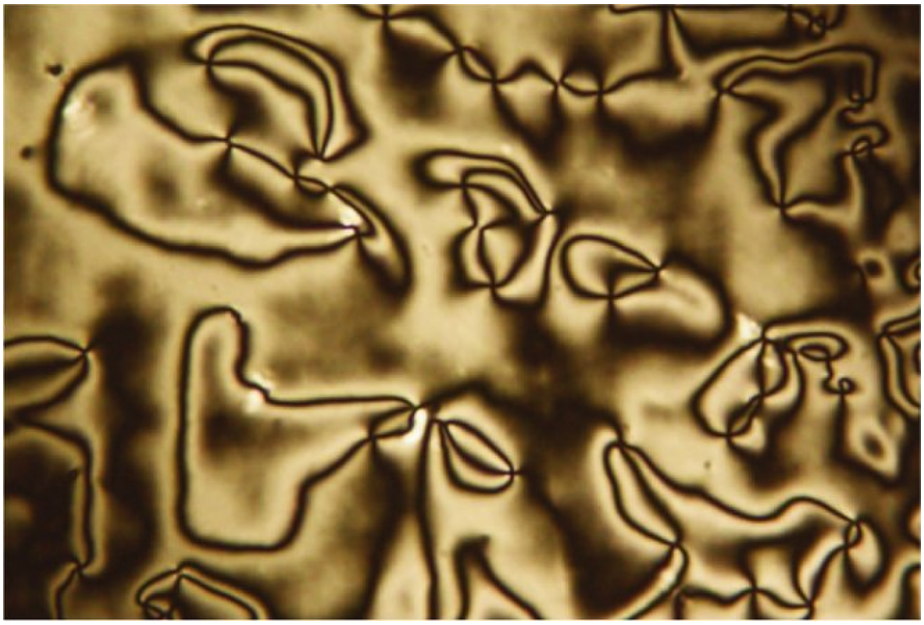

Figure A.1 - *Schlieren* texture in a sDNA 10bp sample.

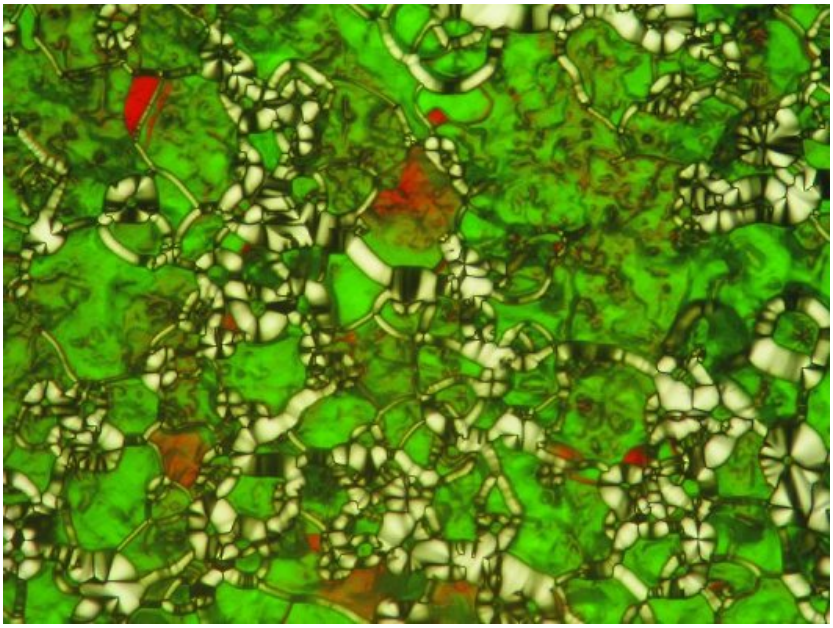

Figure A.2 - Combined polarized transmission and reflection microscopy image of *oily streaks* texture in a cholesteric sDNA 10 bp sample. Focal conics are present in defect regions ($p < 1\mu m$). Photo by Michi Nakata.

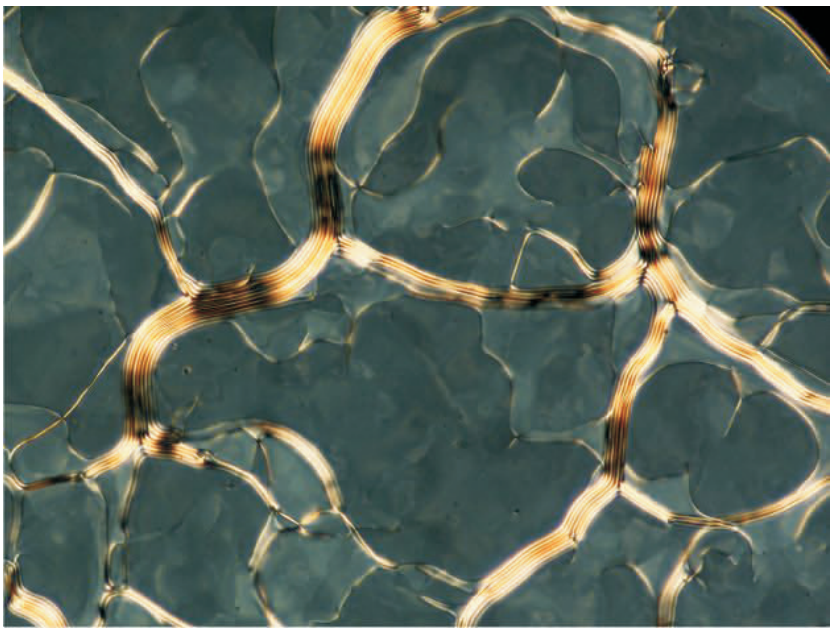

Figure A.3 - Polarized microscopy image of *oily streaks* texture in a cholesteric sDNA 10 bp sample. Fingerprints are present in defect regions ($p \simeq 3\mu m$).

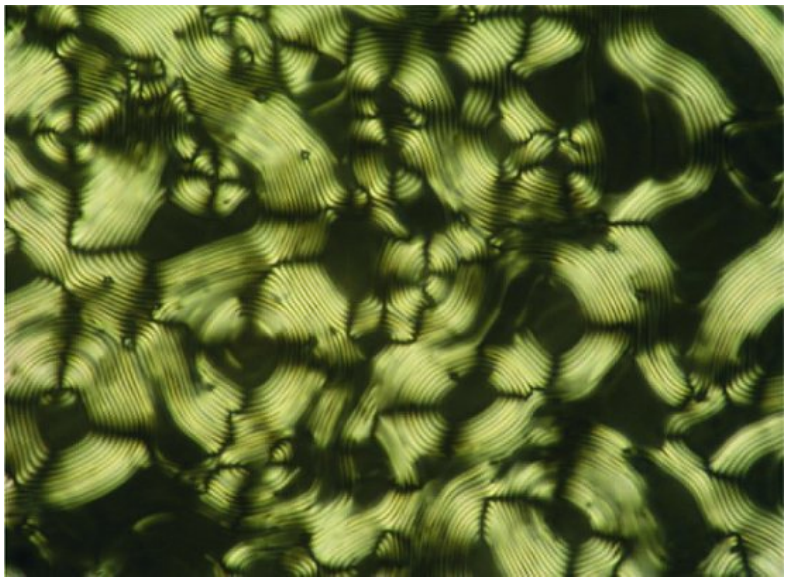

Figure A.4 - Fingerprints texture in a sDNA CG12 sample.

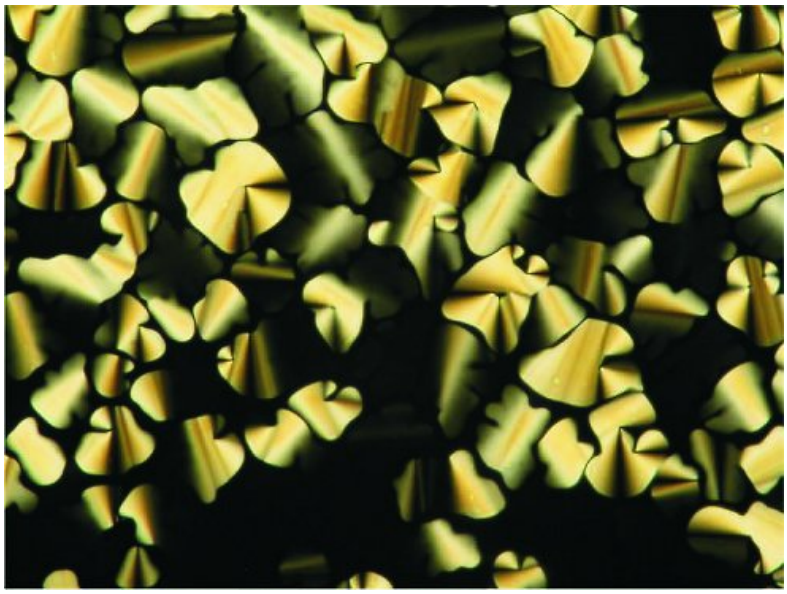

Figure A.5 - Focal conics in a sDNA 8 bp sample. The average size of domains is $50\mu m$. Photo by Michi Nakata.

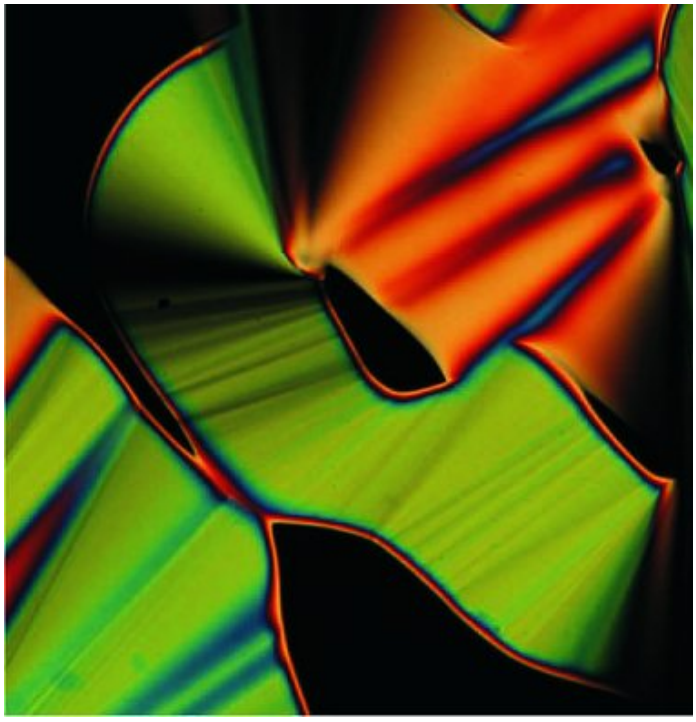

Figure A.6 - Focal conics in a sDNA 12 bp sample. The birefringence is the same as in figure A.5, but the higher thickness of the cell produces different colors.

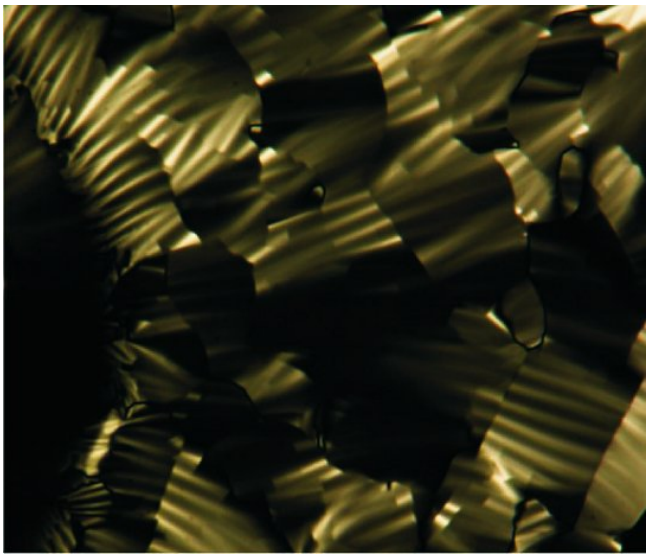

Figure A.7 - Undulating texture in a 1:1 A+B mixture.

Figure A.8 - Mosaic texture in a sDNA D12+12TT mixture.

Figure A.9 - Dendritic growth of columnar C_2 phase in a sDNA 12TT sample. Photo by Michi Nakata.

Figures of chapter 4

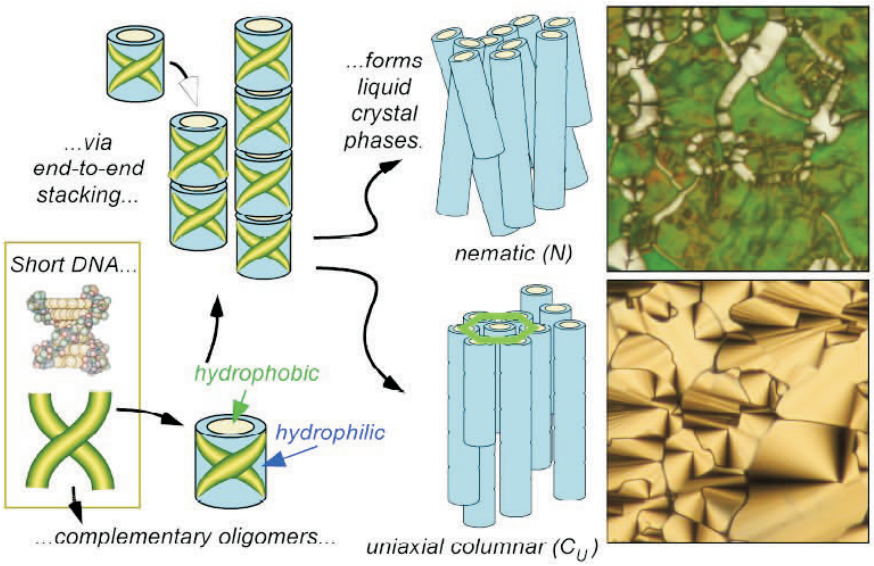

Figure A.10 - Ultra-short DNA duplexes can be idealized as hydrophilic cylinders with hydrophobic ends capable of end-to-end adhesion and stacking into units sufficiently anisotropic to orientationally and positionally order into liquid crystal phases. The nematic (N) phase is formed at lower concentration and the uniaxial columnar (C_U) at higher concentration.

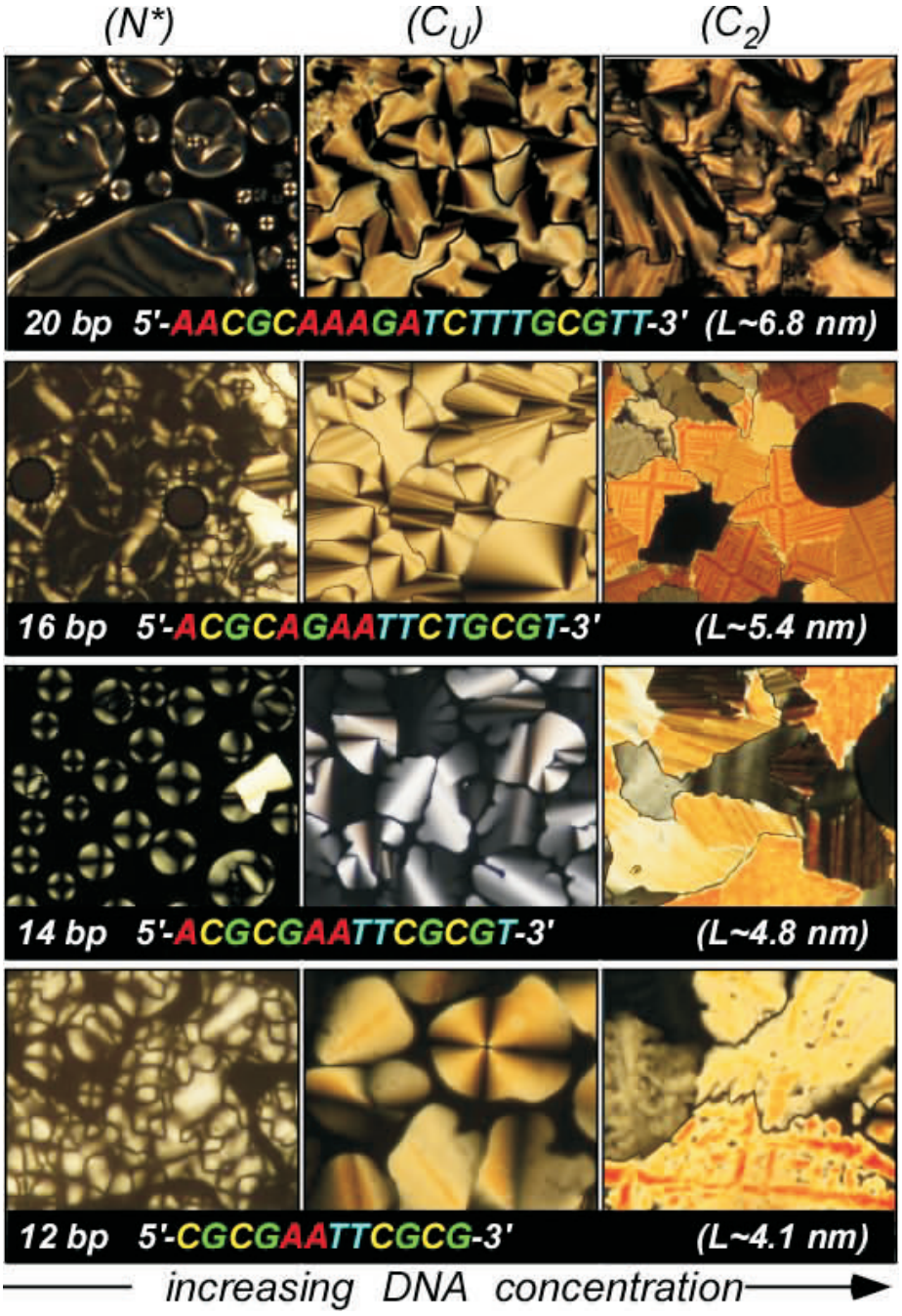

Figure A.11 - Optical textures of the LC phases of a series of solutions of sDNA of increasing length (12-20 bp) obtained by polarized transmission microscopy. Textures are described in the text. The width of each image is 120 μm. Continued in figure A.12.

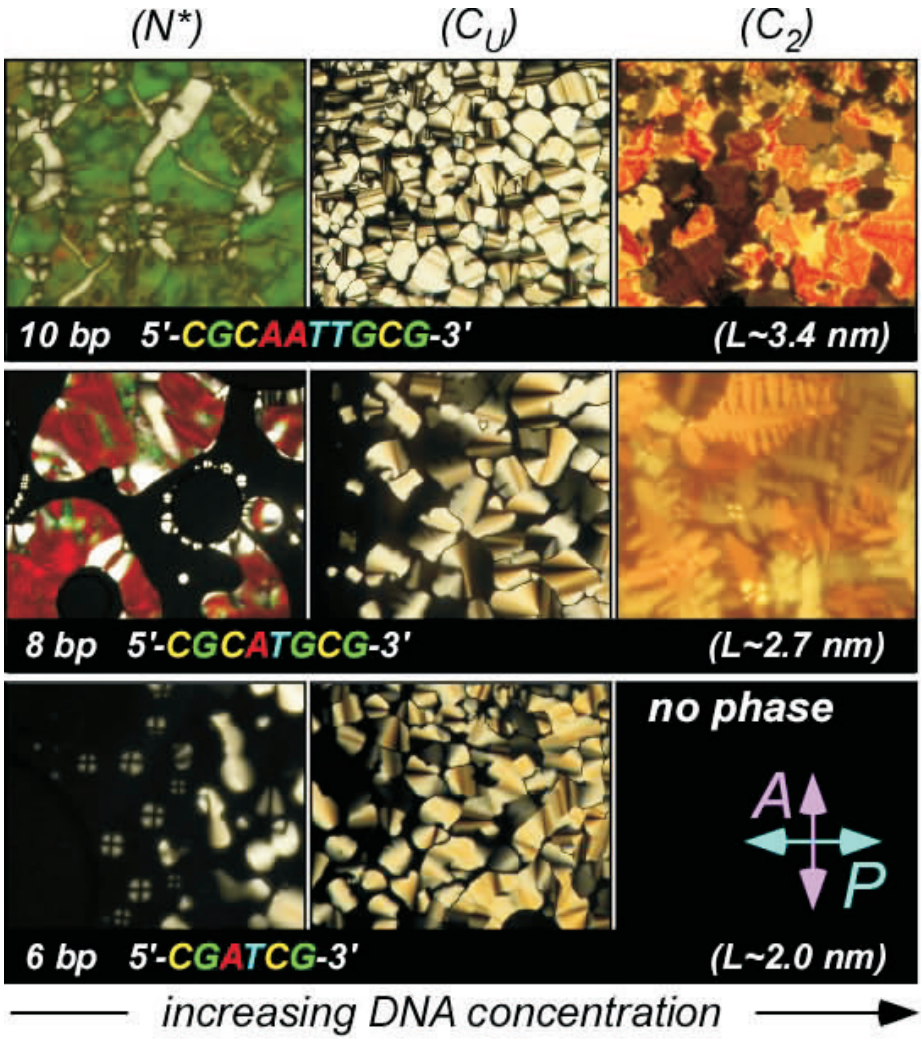

Figure A.12 - Continued from figure A.11, for N=6-10.

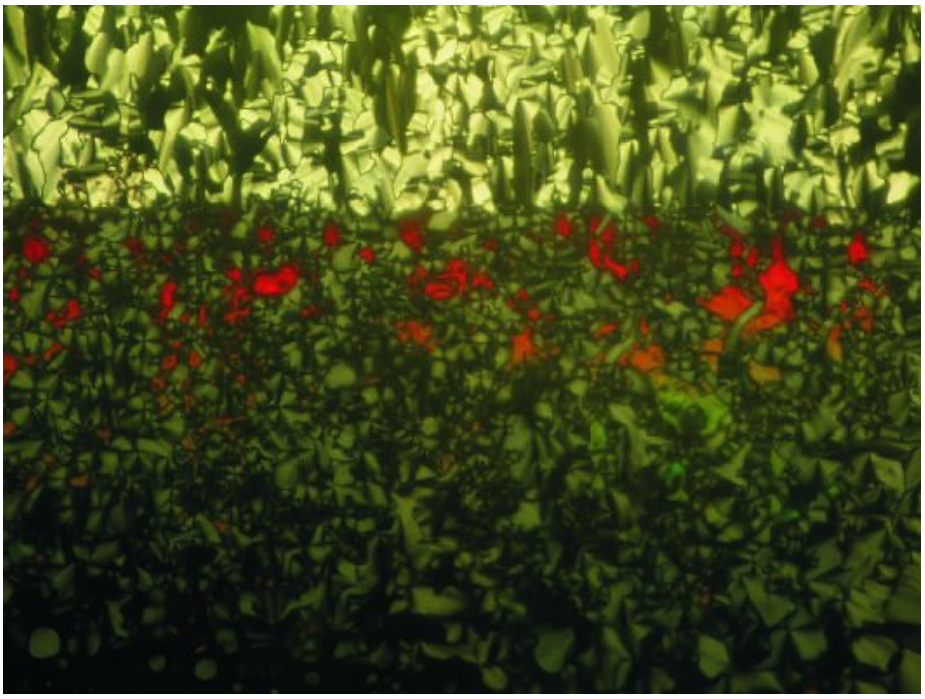

Figure A.13 - Combined polarized transmission and reflection optical texture of a solution of the N=10 self-complementary oligomer with concentration decreasing from top to bottom. The high birefringence region in the upper quarter is C_U and the rest is N^*. The reflection colors show that the N^* pitch increases with increasing concentration. The peak in the red selective back-reflection band at the $N^* - C_U$ interface is for incident light of wavelength $\lambda_{max} = 625\ nm$, yielding a maximum pitch for N = 10 of $p = \lambda_{max}/n \sim 400\ nm$. Photo by Michi Nakata.

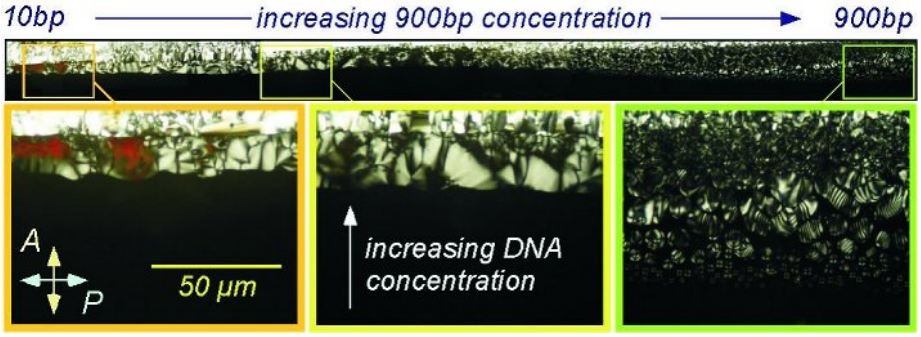

Figure A.14 - Optical texture of a 10 bp sDNA / 900 bp lDNA contact cell obtained by polarized microscopy. In this cell the relative concentration of the 10 bp and 900 bp DNA varies from left to right, and the overall DNA concentration, c, increases from bottom to top. The chiral nematic phase appears first from the I phase (black) with increasing c and continuously spans the full 10/900 concentration range, indicating that the 10 bp nematic is the same phase as the 900 bp nematic. Similar remarks apply to the C_U phase. In the 900 bp DNA the helix is visible (pitch $p \sim 3\ \mu m$), decreasing with increasing decamer concentration to exhibit selective reflection in the red (pitch $p \sim 0.4\ \mu m$) at the N = 10 end.

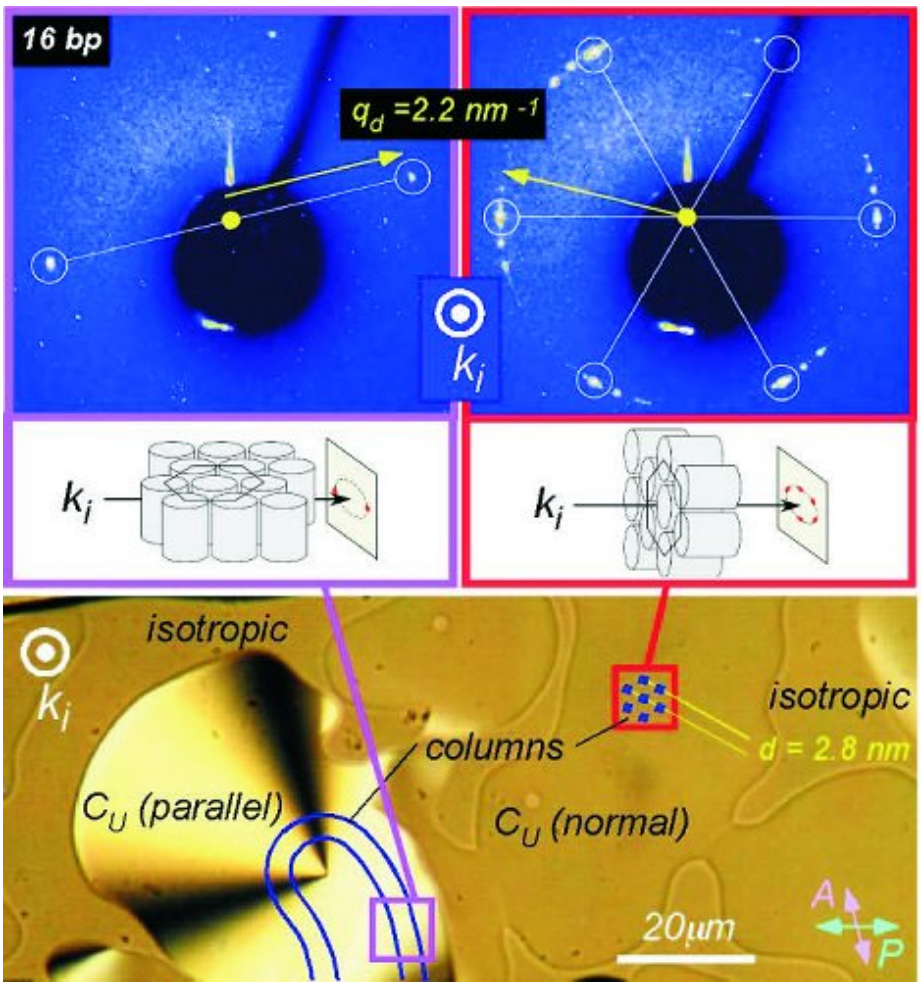

Figure A.15 - Polarized microscopy optical image of a uniaxial columnar phase texture between glass plates and synchrotron microbeam X-ray diffraction patterns of selected monodomains. Optical microscopy simultaneous with the X-ray scattering enables probing monodomains of either planar aligned developable domains (purple area) with the columns parallel to the glass, showing an array of linear columns (blue lines) in the X-ray, or domains with optical axis normal to the glass (red area), showing the hexagonal column lattice (blue dots) in the X-ray. The uniaxial hexagonal columnar structure of the sDNA C_U phase is confirmed by this experiment.

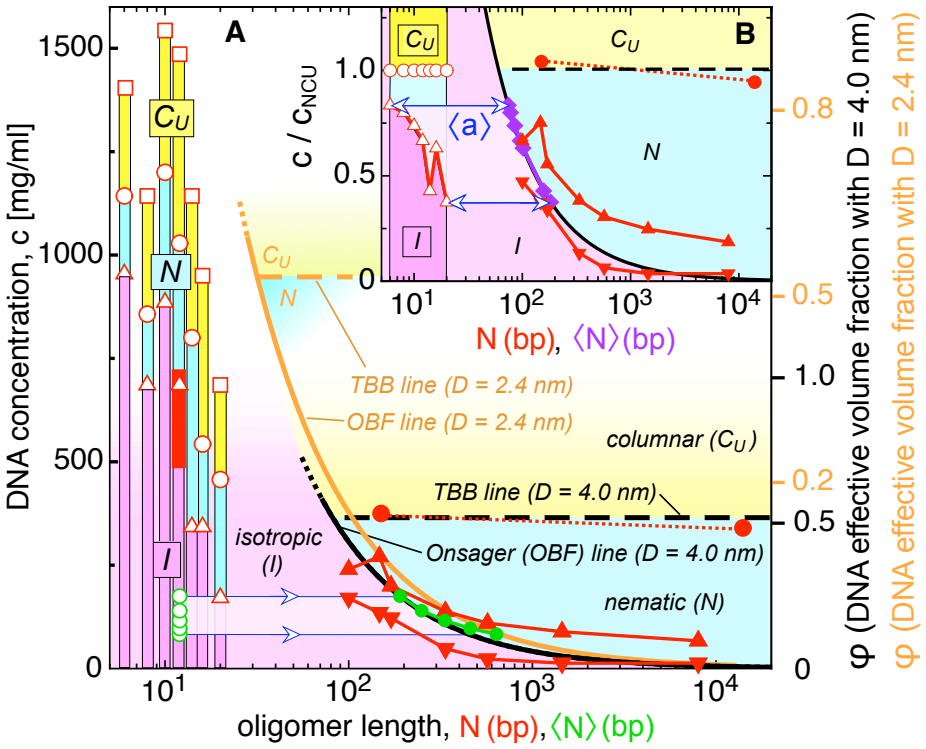

Figure A.16 - Experimental c (DNA concentration, mg solute/ml solution) - N (oligomer length) phase behavior for sDNA and lDNA, along with the theoretical behavior from several models of interacting rod-like particles. (A) The solid red triangles and solid red curve bind the measured I-N phase coexistence for lDNA (N > 100). The solid red circles and red dotted line give the measured $N - C_U$ phase boundary of lDNA. For N < 20, phase transitions from our data are marked by red open symbols (I-N: △; N-C_U: ◯; $C_U - C_2$: □) and the range of each phase indicated by colored columns: magenta (I), cyan (N), yellow (C_U), at T = 20 °C for 20 > N > 8 and T = 10 °C for N = 6. The range of the N = 12 LC phase of ref. [12] is given by the solid red rectangle. Theoretical phase boundaries for these transitions from model systems are shown for two choices of the volume fraction φ axis, with DNA effective diameter D = 4.0 nm (heavy black lines/labels) and with DNA chemical diameter D = 2.4 nm (heavy orange lines/labels). OBF I-N line is shown in black and TBB N-C_U line in dashed black. The open and closed green dots represent respectively the spherical particles of ref. [28], and their effective aggregate lengths at the I-N transition. (B) The c-N phase diagram of (A) but with c scaled with respect to c_{NCU}, enabling an estimate of the length $\langle N \rangle$ (purple diamonds), and aggregation number $\langle a \rangle$ (blue arrows) in the sDNA aggregates. See text for more details.

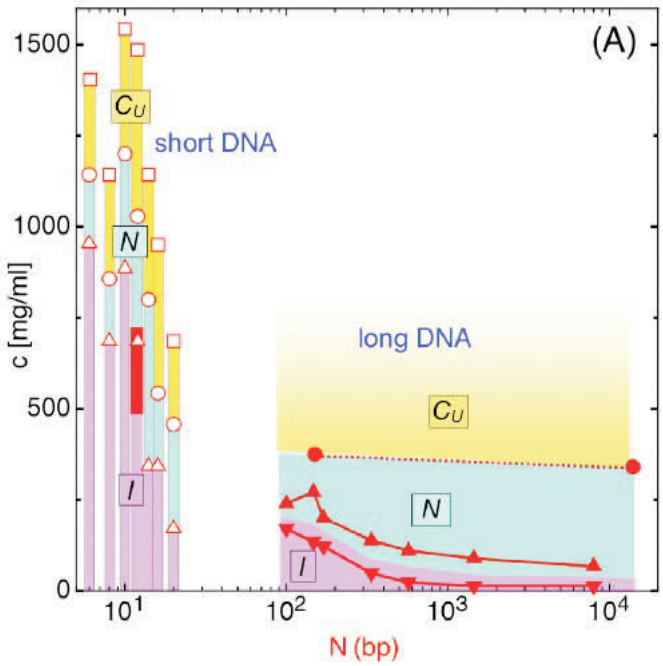

Figure A.17 - Experimental c-N phase behavior for short and long DNA. Legend: see caption to figure A.16

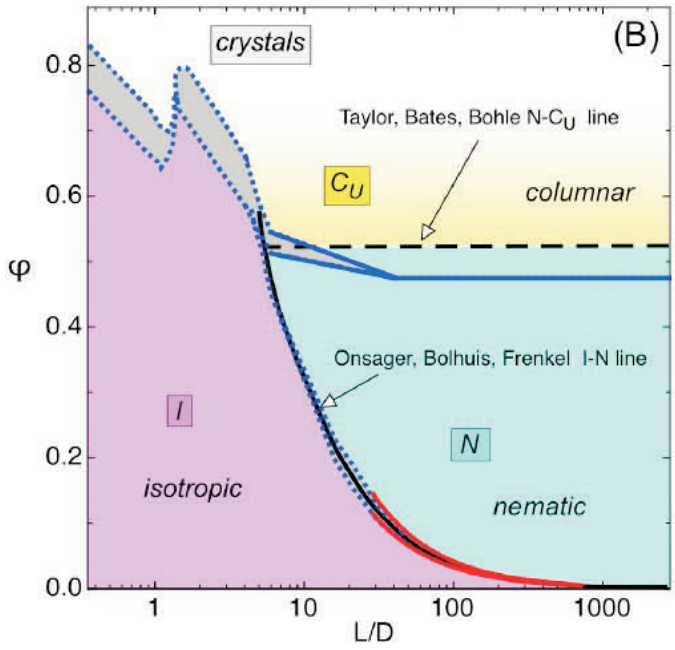

Figure A.18 - Phase diagrams from several models of interacting rod-like particles of axial ratio L/D filling a fraction φ of the total volume. Red lines are the I-N coexistence boundaries according to Onsager, yielding 4.19D/L and 3.29D/L. Dotted blue lines mark the I-N and I-X phase transitions according to the Bolhuis-Frenkel (BF) simulations of monodisperse hard sphero-cylinders (HSC). The solid blue lines are the BF N-SmA transition for monodisperse HSC, which evolves to the N-C_U transition as polydispersity or flexibility is introduced. When the SmA phase is suppressed, the system undergoes a transition from the N phase toward the C_U phase at $\varphi \sim 0.55$, according to the modeling of Taylor and Herzfeld, Bates and Frenkel, and Bohle, Holyst, and Vilgis. See text for comments and references.

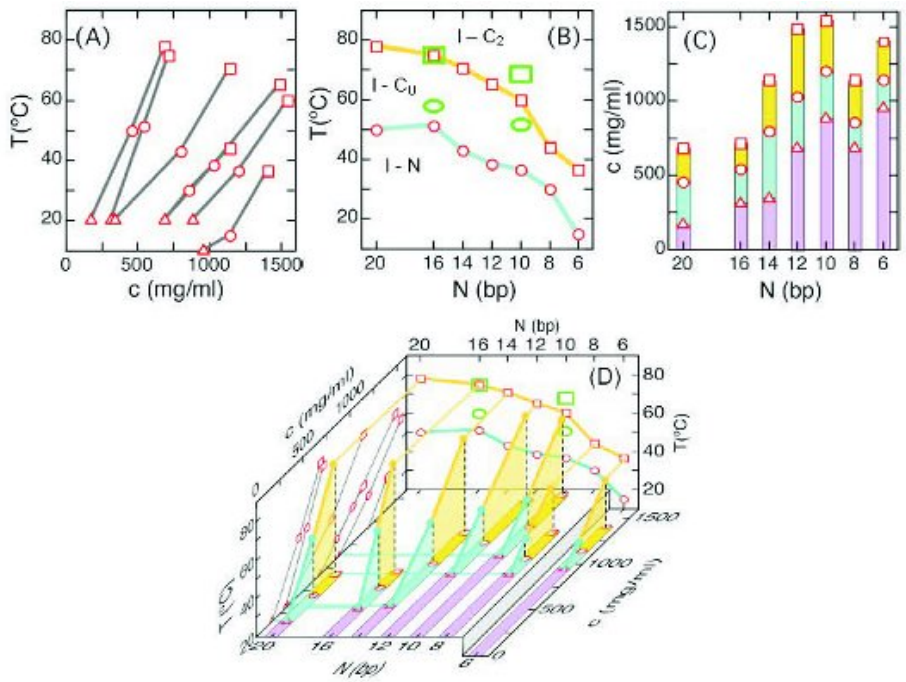

Figure A.19 - Thermal LC melting behavior vs. N and c for the series of sDNA oligomers. (A) I-LC phase transition temperatures vs. c. For each N the LC phase is N at c below the open circles, C_U for c between the open circles and open squares, and C_2 for c above the open squares. Transition temperatures measured on heating and cooling were not significantly different. (B) I-LC phase transition temperature ranges where the transition from the I is respectively to the N, C_U, or C_2 phase. Also shown are the duplex denaturation temperatures of the N=10 and N=16 oligos in the N and C_U concentration range (open green circles and squares, respectively). (C) See fig. A.17. (D) Composite 3D plot. The basal plane is the N-c diagram of C. The blue and yellow shaded polygons indicate where the N and C_U phases are observed. The polygons have nearly vertical N-C_U and C_U-C_2 boundaries indicating that these phase boundaries were not dependent on T, but only on c.

Figure A.20 - Chiral nematic fingerprints and C_U developable domains in a CG12 sample.

Figures of chapter 5

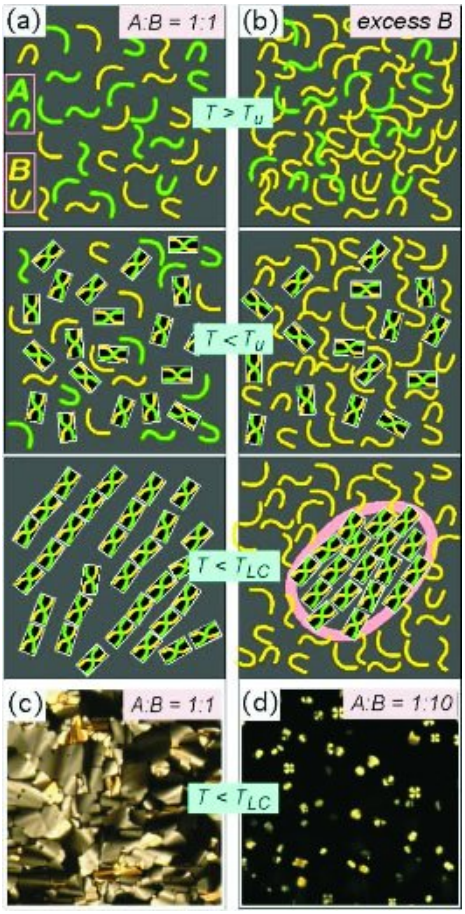

Figure A.21 - In a 1:1 mixture of complementary sDNA oligomers (a), duplexes form upon cooling below their unbinding temperature T_m and the LC phase appears below T_{LC}, filling nearly the whole area with LC domains (c). With one of the oligomer species in excess (b), the transition to the LC phase is marked by the appearance of isolated LC domains that sequester all of the other complementary oligomer into LC droplets (d).

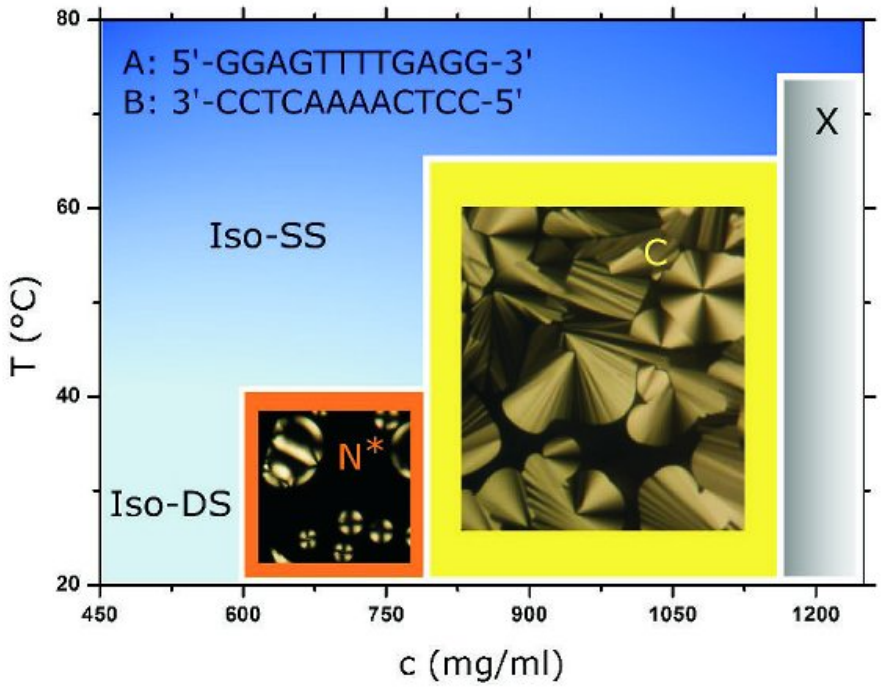

Figure A.22 - Phase diagram of aqueous solution of a 1:1 mixture of sDNA-A and sDNA-B. Upon increasing the total DNA concentration c, a chiral nematic (N^*), a columnar (C) and a crystalline phase (X) appear, whose textures from polarized microscopy are shown. Upon increasing the temperature T, phases melt into an isotropic phase of single strands (Iso-SS). When c $<$ 600 mg/ml and at T below the sDNA duplex unbinding temperature, an isotropic phase of double strands is found (Iso-DS).

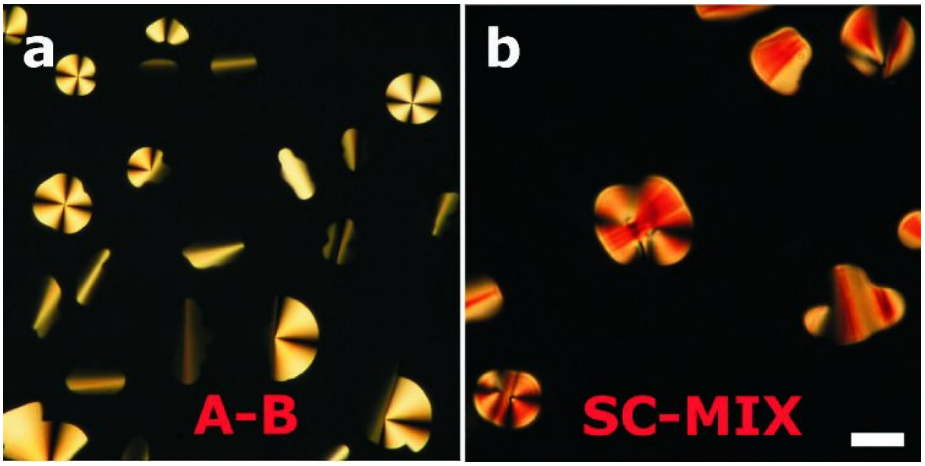

Figure A.23 - Polarized microscopy images of liquid crystal domains in unbalanced mixtures of sDNA-A and sDNA-B (a) and of sDNA-SC in a solution containing a mixture of SS sDNA sequences (b). The molar ratios are [B]/[A] = 3 and [MIX]/[SC] = 4. Scale bar is 20 μm.

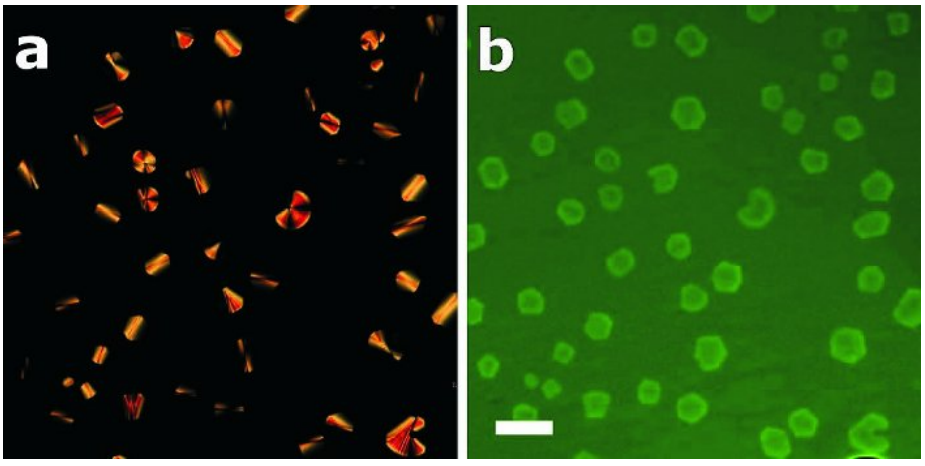

Figure A.24 - Polarized microscopy (a) and fluorescence (b) images of the same columnar domains in an unbalanced A-B mixture ([B]/[A]=10) in which a small fraction (1/40) of A sequences was labeled with fluorescein, bound externally to the double helix. Duplex-rich LC domains appear brighter on the coil-rich background. Scale bar is 50 μm.

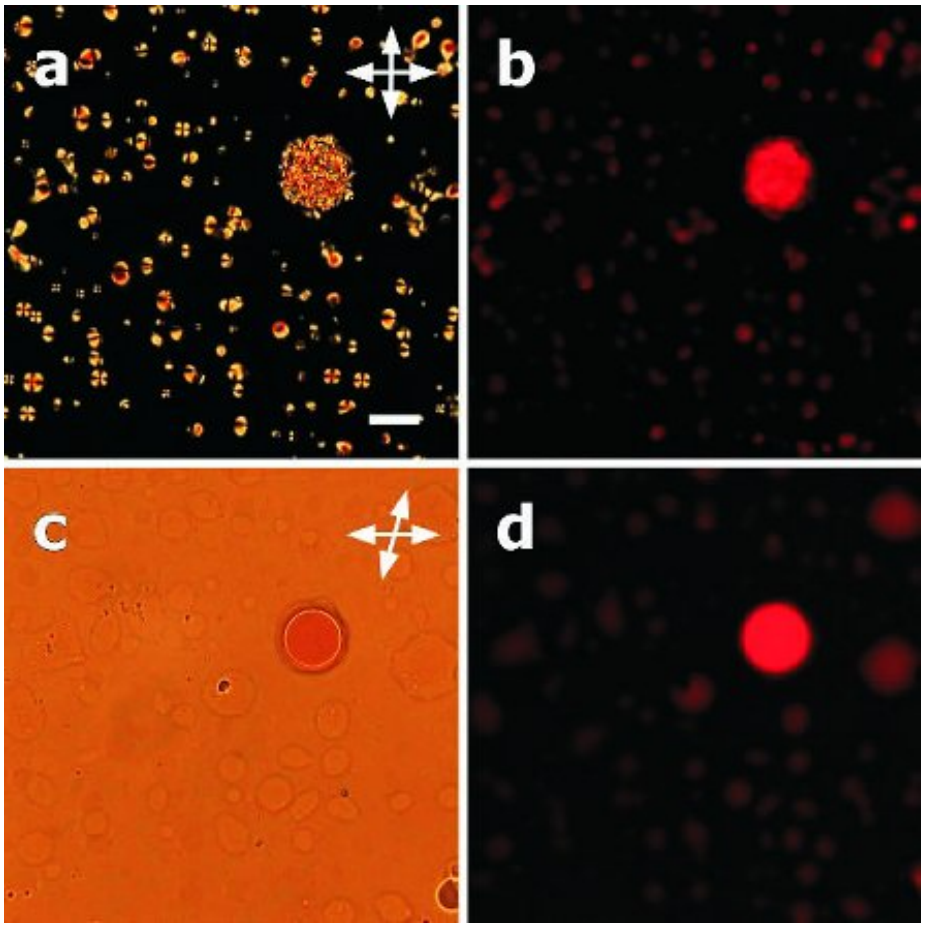

Figure A.25 - Polarized microscopy (a,c) and fluorescence (b,d) images of droplets of a self-complementary sDNA-SC mixed with PEG (MW 8000). Birefringent domains are visible at room temperature under crossed polarizers (a), and the segregation of DNA helices is manifested through the fluorescent dye Ethidium Bromide (b). Samples held for hours at high temperature ($75°C$), where LC domains are melt, still show DNA-rich domains, visible both under slightly decrossed polarizers (c) and in fluorescence image (d). Scale bar is 50 μm.

Figures of chapter 6

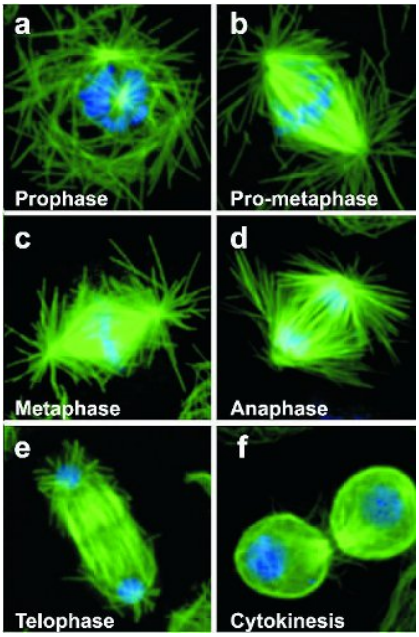

Figure A.26 - Microtubule organization during mitosis. Cells are stained for tubulin (green) and DNA (blue). Cycle goes from prophase (a) to prometaphase (b), metaphase (c), anaphase (d), telophase (e), and cytokinesis (f). From [33].

.

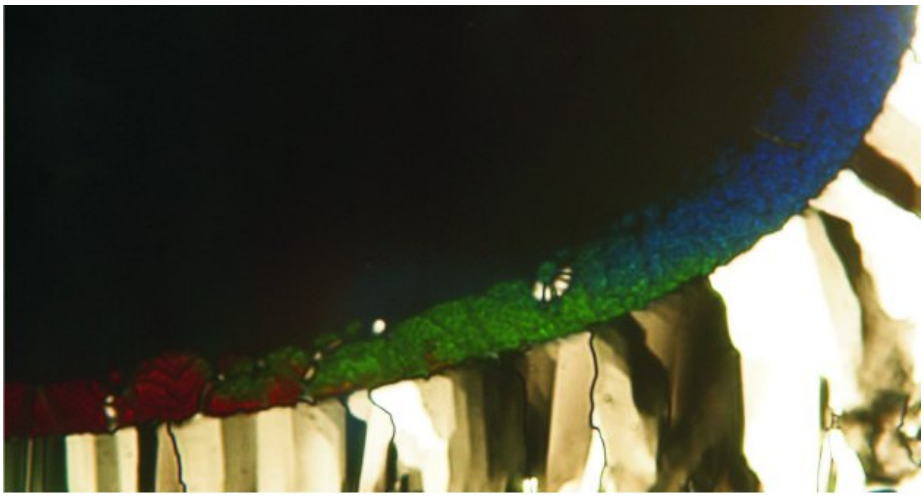

Figure A.27 - Polarized microscopy texture of cholesteric phase in CGC-GAAUUCGCG RNA duplex.

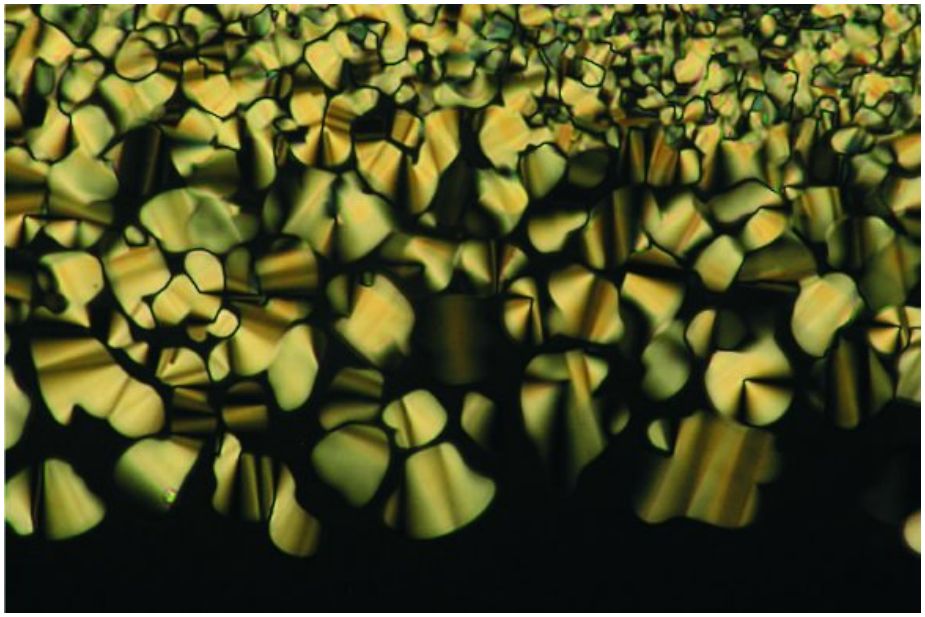

Figure A.28 - Polarized microscopy texture of C_U columnar phase in CGCAAUUGCG RNA duplex.

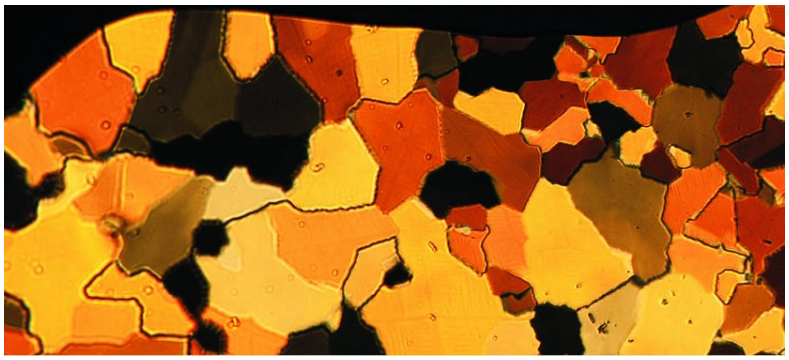

Figure A.29 - Polarized microscopy texture of C_2 phase in CGCGAAUUCGCG RNA duplex.

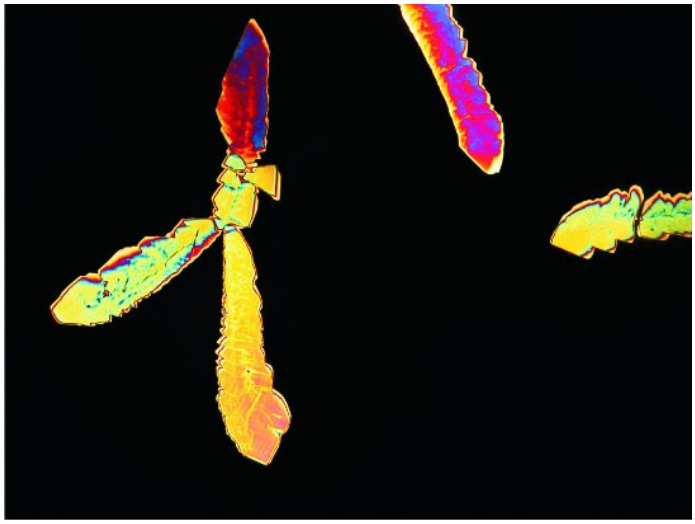

Figure A.30 - Polarized microscopy image of the crystals grown in mixtures of RNA dodecamer and *ss-DNA* .

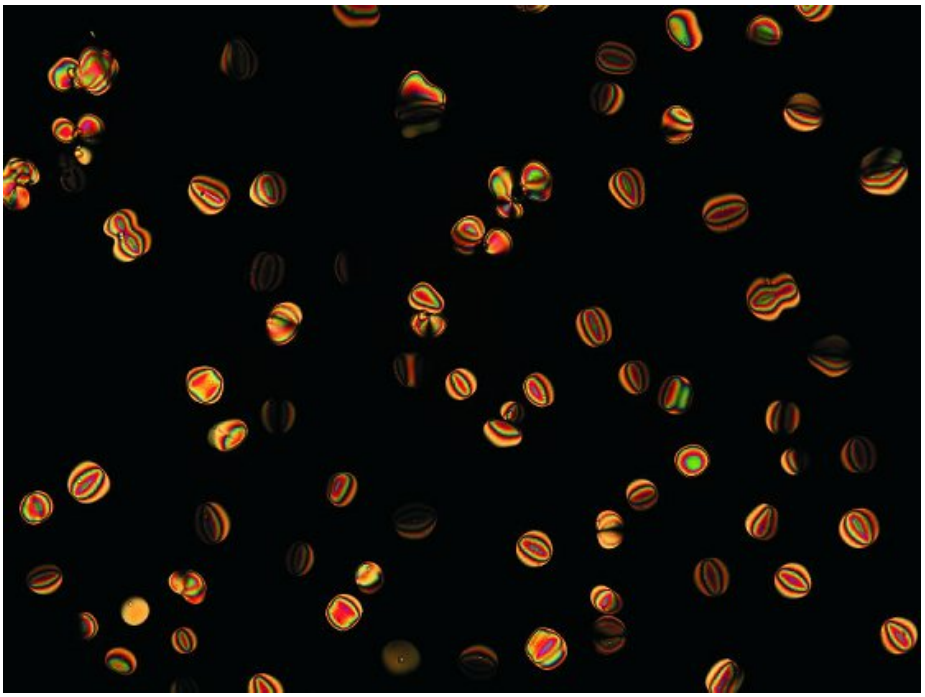

Figure A.31 - LC domains of A+B and C+D duplexes in the Isotropic background of single strands. Typical size of domains is $\simeq 30\mu m$.

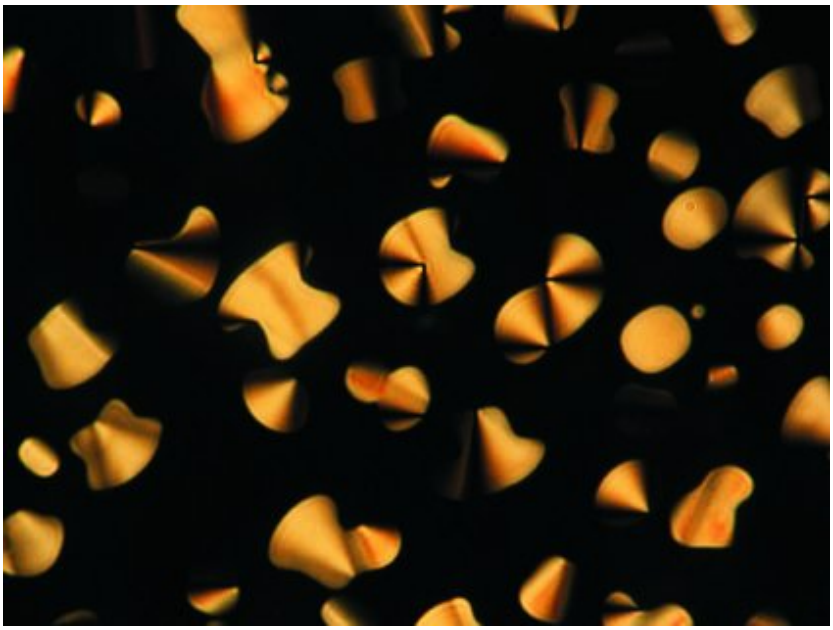

Figure A.32 - Columnar domains of CGnnnnCG; their average size is $30\mu m$.

Figure A.33 - Columnar domains of CGnnnnCG nucleated on the glass surface.

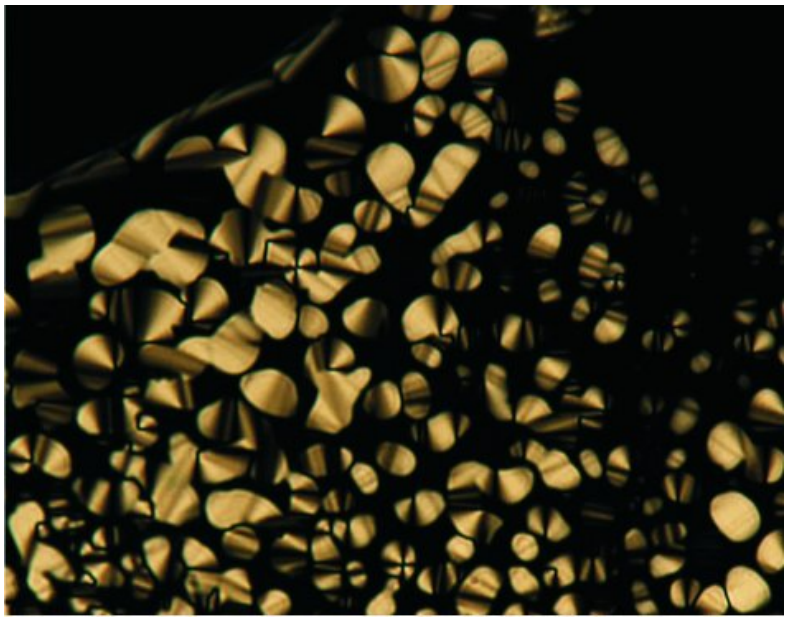

Figure A.34 - Columnar domains of CnnnnnnnnG; their average size is $20\mu m$.

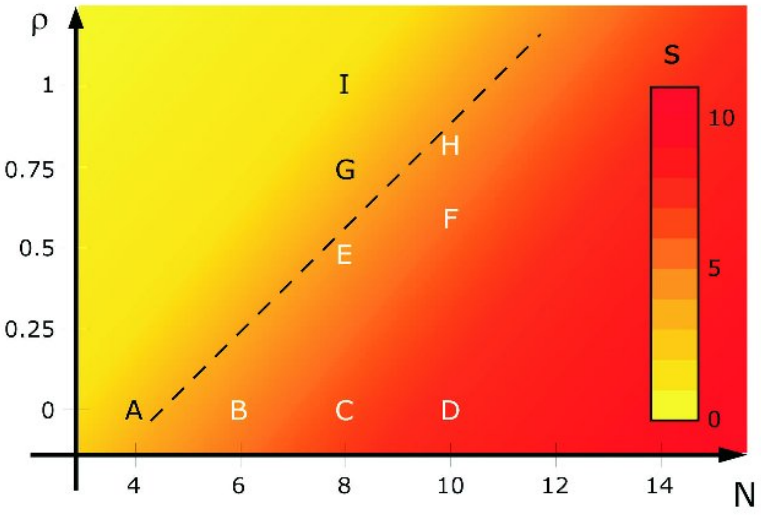

Figure A.35 - Overall thermodynamic stability of sequences of various lengths and degrees of randomness, expressed by the score S, is compared with the observed phase behavior: white sequences display LC behavior, black don't (the link between letters and sequences is provided in table 6.7).

Acknowledgements

I feel like I've just begun, but actually I have been working in "Complex Fluids and Molecular Biophysics" lab for more than four years! My first thanks goes to my supervisor, prof. Tommaso Bellini, because he involved me in the DNA project and he always patiently supported and stimulated me during these years, usually having much more trust and estimate on me than myself. He taught me to wonder and never take anything for granted - and quite a lot of physics!
Most part of this thesis wouldn't have been possible without Michi Nakata, who welcomed me to Boulder and spent a lot of time with me during the work, always kind and humble although she was one of the most brilliant minds I've ever met. She tragically left us, but I'm sure she's still taking her beautiful, colored pictures.
Thanks to prof. Noel Clark for inviting me to Boulder, for his estimate and for the really enlightening conversations on DNA LCs, that will go on next year!
Since I crossed the Ocean, let me remind all the people in the Liquid Crystal Materials Research Center: working there was an exciting experience also thanks to Chenhui, Loren, Michaelscott and all the others. And thanks to the wonderful Colorado community, the Curtos above all.
Let's come back to Italy. Thanks to my co-advisor Dr. Guido Tiana, whose insightful questions and comments helped me to focus on essential points.
Every single member of my lab would deserve my thanks for a bunch of reasons: help, advice, the right question, alpine songs (...), a smile of encouragement, in one word the company in my everyday job. It's hard to find this atmosphere around. Thanks to (random order!): Stefano "kein" Pezzati, Lucia, Vale, Ben, Cecco, Marisa, Marco, Domenico, Carlo, Silvia, Silvia, Matteo, Lollo, teens, Roberto, Cri, Tossico, Caggio and Meli.
Thanks to my PhD fellows Diego, Simona, in particular my "twin" Tommaso, and to all the people that I met in my university years, older and younger, because they witness to me that it's worth to spend life, time and energies to follow what makes me a real man.
Thanks to my brothers, my grandma and especially to my parents, for always supporting me, although they certainly still wonder what is the link between DNA and liquid crystals, and why should I spend time in studying them...

Thanks to my second "family" Paolo, Simo, Bore, Chiara, Maria, Ceci, Tommi, Sara and Francesca, for too many reasons to write them down.
Finally, thanks to my daughter Teresa, for keeping me awake in the nights to write this thesis, and to Betta, my wonderful and patient wife: the relationship with you two is a new adventure every day.