Interaction, recognition, and condensation of DNA duplexes

I n a u g u r a l - D i s s e r t a t i o n

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Andrey G. Cherstvy

aus Byhov, Belarus

Jülich

2002

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf

Referent: Prof.Dr. A.A. Kornyshev
Korreferent: Prof.Dr. H.Löwen
Tag der mündlichen Prüfung: 20.11, 21.11, und 25.11.2002.

Abstract

The helical symmetry of DNA plays the key role in many cell processes. This symmetry is also of the paramount importance for description of many DNA properties observed *in vitro*. In the first part we give an explanation of one such phenomenon: temperature-induced DNA condensation in the presence of Mn^{2+}, detected by osmotic stress measurements in dense columnar hexagonal DNA assembly. Our analysis is based on the theory of electrostatic interaction of *ideal* DNA duplexes developed by A.A.Kornyshev and S.Leikin several years ago. We have modified this theory to take into account the *non-fixed*, self-consistently determined, distribution of adsorbed cations on DNA surfaces. DNAs thus adopt the patterns of adsorbed cations to minimize the total free energy. With temperature increase and with compression of the aggregate cations occupy the major DNA groove that stimulates attraction between DNAs due to better charge separation along the helices. When DNAs are pushed together, the entropy of the aggregate increases due to higher binding entropy of cations in the major groove. For plausible model parameters both the calculated force curve and entropy change agree quantitatively with experimental data. It confirms the hypothesis that in dense assemblies DNAs interact predominantly electrostatically.

In the second part we explore the basic consequences of electrostatic interaction of *non-ideally helical* DNA duplexes of a finite torsional rigidity. We generalize the theory of interaction of ideal duplexes for the case of *non-rigid* DNA molecules with sequence-dependent twist patterns. It is shown that the interaction energies of *rigid* homologous and non-homologous DNA sequences differ dramatically: the latter cannot attract. Non-rigid DNAs with random base-pair sequences adjust their incommensurate twist patterns, soft enough sequences can profit from restored strand-groove alignment and attract each other. The recognition energy of *homologous sequences* increases linearly with their length, similarly to the frequency of recombination between homologous DNA fragments in many organisms. It suggests that *direct* electrostatic interaction of juxtaposed *intact duplexes* of 100-200bp-long can be responsible for their recognition from a distance. We have determined the set of parameters which can affect the recognition energy. It gives a hint to control homologous recombination. Experimental verification of the predicted tendencies could help understanding how homologous DNA sequences recognize each other in gene shuffling reaction and DNA repair, the (most) important phenomena for existence and evolution of life.

Contents

Introduction **4**

1 Temperature-induced aggregation of Mn−DNA **7**
1.1 DNA properties in solutions 7
1.1.1 DNA as a polyelectrolyte molecule: counterion condensation 7
1.1.2 DNA condensation 9
1.2 Specificity of Mn^{2+}-binding to DNA 13
1.2.1 Puzzle of Mn-DNA condensation 14
1.2.2 Condensation of Mn-DNA upon heating is entropy-driven 15
1.3 Distribution of cations on DNA 15
1.3.1 The model of Mn-DNA aggregation 15
1.3.2 Two-state adsorption model for Mn 16
1.4 Kornyshev-Leikin theory of electrostatic interaction of helical macromolecules 19
1.4.1 Debye-Hückel-Bjerrum approximation 19
1.4.2 The field equations and their solution 20
1.4.3 Interaction free energy 21
1.4.4 DNA-DNA interaction 23
1.4.5 Predictions of the theory and their experimental confirmation 23
1.5 The free energy of DNA assembly 25
1.5.1 Homogeneously-charged rods in the cell model 25
1.5.2 Linearized and nonlinear Poisson-Boltzmann equations 28
1.6 Total interaction pressure 30
1.6.1 The minimum of the free energy 30
1.6.2 Intermolecular forces 32
1.7 Entropy change and the force 33
1.7.1 Partition function of adsorbed cation 33
1.7.2 Parameters of the model 35
1.7.3 Calculated pressure curve and cations re-distribution . 36

1.7.4 Other interpretations of experimental data 38
1.8 Auxiliary results of the model and beyond the model 39
1.8.1 Adsorption isotherm and its effect on interaction pressure 39
1.8.2 Abrupt transitions of cations between the grooves . . . 41
1.8.3 Behavior of DNAs of two types on the lattice 43
1.8.4 Ions of the finite size and Poisson-Boltzmann equation 45
1.8.5 Are DNA melting and aggregation coupled? 47
1.8.6 Aggregation of non-ideal duplexes, the role of bp fluctuations 51
1.9 Azimuthally frustrated DNA lattices 52
1.9.1 Experimental observations and puzzles 52
1.9.2 Optimal azimuthal angles on hexagonal lattice 52
1.9.3 Observed lattice structures: Hexagonal *vs.* orthorhombic 56

2 Interaction and recognition of non-ideal DNA duplexes 58
2.1 DNA-DNA recognition: the role and description 58
2.1.1 Homologous recombination and the length of homology 58
2.1.2 Electrostatic mechanism of homology recognition . . . 59
2.1.3 Recognition energy and DNA torsional softness 60
2.2 The free energy functional 61
2.3 Rigid DNA duplexes 64
2.4 Soft DNA duplexes 65
2.4.1 Infinite sequences 65
2.4.2 Finite sequences 66
2.5 Approximate solution in parabolic approximation 67
2.6 How to affect the recognition energy? 68
2.7 Beyond $\cos(\delta\phi)$ interaction potential: ready to kink 72
2.7.1 Rigid chains 72
2.7.2 Soft chains: Two-well-approximation 72
2.7.3 The role of kink-solitons 73
2.8 Azimuthal kinks on ideal non-rigid DNA duplexes 74
2.8.1 Free energy functional and basic equations 74
2.8.2 Big kink 75
2.8.3 Small kink 76
2.8.4 Outlook 78

Conclusions 80

Acknowledgement 81

A Non-rigid duplexes: The free energy functional 82

B Non-rigid duplexes: perturbation theory in small κ **85**

Bibliography **86**

Introduction

This work is devoted to understanding of how DNA helices interact and recognize each other, how and under which conditions it can lead to their self-assembly. In general, assemblies of long helical macromolecules are common building blocks of living organisms. Bundles of α-helices form domains in many proteins. Bundles of collagen triple helices form tendons, cornea, matrix of skin and bone, and other connective tissues. DNA duplexes are tightly assembled in viral capsids and wrapped around histones to form the nucleosome infrastructure of eukaryotic chromatin. Some biopolymers adopt even a multimolecular structure, like tobacco mosaic virus. In other words, many biological interactions involve helices, and many complex biological structures consist of helices. In *in vitro* experiments spontaneous and induced assembly of many biohelices has also been observed (DNA, collagen, hydroxypropylcellulose, guanosine four-stranded helices, myosin, etc.).

DNA is one of the most important biohelices present in nature. It plays the paramount role for functioning of all living organisms. After its structure became known, DNA attracted the attention of many scientists, but still a lot of puzzles related to the interaction of DNA helices remain unsolved.

One of them is DNA condensation in the presence of specific agents into bundles with 6-10Å of water between the DNA surfaces. Why do many biologically important 3^+ and 4^+ cations in (multi)milli-molar concentrations condense double-stranded DNA(s)? Why do only two divalent cations (Mn^{2+} and Cd^{2+}) and none of monovalent cations condense DNA in aqueous solutions? Why does attraction occur between like-charged objects, such as DNAs? The second puzzle is the transition of DNA from B to A form in dense assemblies with decrease of humidity or in the presence of special cations. Why does DNA move its phosphates closer and how does it depend on the type of cation? The third puzzle is the variety of DNA lattice structures in dense assemblies (hexagonal, orthorhombic, etc.), detected by X-ray diffraction, with different axial shifts between the nearest DNAs. The appearance of cholesteric DNA liquid-crystal phase at 15Å separation between the surfaces of 150-base-pairs-long DNA fragments is also not completely understood. The fourth puzzle is why do B-DNAs have the integer number of base pair per helical turn in fibers and in crystals, as measured by X-ray crystallography, and non-integer number in solutions? Fifth is the short decay length of the force between many biohelices in the last 10Å between their surfaces. This is a far from complete list.

The structure determines the interactions of biohelices, the interactions in turn determine the functionality of the helices. Electrostatics is a major contribution to the interaction between many biohelices, since virtually many of them have high density of surface charges. For instance, B-DNA has one elementary charge per each 1.7Å axial separation. It is evident that the distribution of phosphates and associated cations on DNA molecule is extremely important for electrostatic DNA-DNA interaction. The standard DNA polyelectrolyte "rod-model" does not account for this specificity and thus is not able to describe many effects, especially at short DNA surface-to-surface separations.

The first theory that takes into account the helical symmetry of DNA and discreteness of its charges has been developed several years ago (Section 1.4). On the basis of the exact Hamiltonian and the linear Poisson-Boltzmann equation A.A.Kornyshev and S.Leikin have calculated the electrostatic interaction energy of two long ideal DNA duplexes, when phosphates and adsorbed cations are treated *explicitly*.

This theory predicts the driving forces for all above-described puzzles, and also gives a hint for the solution of many others. First, DNA condensation occurs due to predominant adsorption of (multivalent) cations in major DNA groove. It produces a sharp charge separation along the helices and consequent zipper-like attraction. Second, B-to-A transition occurs due to lower electrostatic interaction energy of A-DNAs, since their charges are better separated. The preferential binding of cations in the minor groove thus stabilizes the B-form of DNA. Third, at close separations the optimal axial shift between two DNAs is nonzero. This gives a hint for the description of variety of observed frustrated DNA lattices. The calculated interaction energy of skewed helices gives a close estimate for the cholesteric pitch of DNA liquid-crystal phases. Fourth, it was shown that the gain in electrostatic interaction energy of B-DNAs in dense aggregates due to the integer number of base pairs per turn is larger than the torsional energy DNA overwinding from 10.5 to 10.0 base pairs per helical turn. Fifth, the theoretically predicted decay length depends on Debye screening length and the number and positioning of helical strands. For the interaction of DNA, collagen, and guanosine four-stranded helices the predicted values are in excellent agreement with experimentally measured.

The Kornyshev-Leikin theory of DNA-DNA electrostatic interaction may, however, need modification to elucidate other DNA puzzles. Two such modifications are made in the present thesis.

The *first assumption* of the theory was that the pattern of adsorbed cations is fixed and does not change either with temperature or with DNA-DNA separation. But it is plausible that with increasing temperature two DNAs can adjust the distributions of adsorbed cations between the grooves to make the interaction more profitable.

Phenomenon: The measurements of forces columnar hexagonal assemblies (osmotic stress technique coupled with X-ray diffraction) show that DNA molecules condense spontaneously upon heating only in the presence of Mn^{2+} and Cd^{2+}. It was suggested that the temperature-induced DNA condensation is entropy-driven, and that the redistribution of adsorbed cations between different DNA adsorption sites can contribute to that.

In Chapter 1 we construct a model for distribution of cations on the DNA surface on the basis of a two-state adsorption model. It allows to change the partitioning of cations between the grooves with temperature and upon DNA-DNA interaction. The portions of adsorbed cations in the grooves are determined self-consistently by free energy minimization. As a result, the observed temperature dependence of the DNA-force-curve is reproduced, and a new source of entropy increase – the entropy of adsorbed cations – is suggested.

The *second assumption* of the Kornyshev-Leikin theory was the ideality of the DNA helix and its infinite torsional rigidity. The modified theory shows that DNA non-ideality reduces the DNA-DNA attraction, or can even turn it into repulsion (the strand-groove register along the helices is disrupted by sequence-dependent twist variation). The theory also predicts the dramatic difference in interaction of homologous and non-homologous DNA sequences which suggests the electrostatic mechanism of DNA-DNA recognition. However, the role of torsional flexibility has not been explored.

Phenomenon: (i) Recognition of homologous DNA sequences is a precursor stage of homologous recombination. It was measured that for bacteriophages, bacteria, and mammalian cells

a 50-200bp-long DNA homology is the minimal length for efficient homologous recombination. For longer sequences the dependence of recombination frequency on the length of homology is linear. (ii) Spontaneous precipitation of DNA sequences with unrelated texts was detected, that confirms that real non-rigid DNA can attract.

In Chapter 2 we generalize the Kornyshev-Leikin theory of electrostatic interaction of imperfect duplexes to account for finite DNA torsional rigidity. We find that real DNA sequences may adjust their twist patterns and attract each other. The calculated recognition energy increases linearly with the length of homology, both for rigid and soft sequences. For 200bp-long DNA fragments the predicted recognition energy is several $k_{\mathrm{B}}T$. We discuss also the possibility of torsional static kinks induced by DNA-DNA interaction and their effect on the interaction energy.

Chapter 1

Temperature-induced aggregation of Mn−DNA

The mechanisms of temperature-induced entropy-driven aggregation of many biopolymers are not completely understood. It is commonly believed that hydration forces and water release govern the spontaneous precipitation of DNA at ambient temperature. However, no direct experimental evidences have been found, and the specificity of cations which cause aggregation is still puzzling. In this chapter this specificity and the helicity of DNA charge distribution are first considered together. We explore the role of DNA-DNA electrostatic interaction and complementarity of the patterns of adsorbed cations for temperature-induced DNA condensation and entropy change upon it.

1.1 DNA properties in solutions

1.1.1 DNA as a polyelectrolyte molecule: counterion condensation

Nucleic acids, and DNA in particular, play a crucial role for functioning of all living organisms. DNA is the key molecule responsible for storage, duplication, and realization of genetic material. In cells DNAs prefer to adopt a double-helical right-spiral *B*-form discovered by Watson and Crick [1] in conjunction with X-ray fibre diffraction data of Wilkins and Franklin [1]. Many other DNA forms have also been observed (*A*, *C*, *D*, *Z*, *Psi*, etc.). The transitions between them can be triggered by humidity, salt concentration, special agents, DNA base-pair (bp) sequence, etc. [2].

B-DNA is a highly negatively charged polyelectrolyte molecule. It bears ionizable groups, which in polar solvent can dissociate into charged macroion and small counterions. "Bare" *B*-DNA has one elementary charge per 1.7Å [2] and thus attracts positive charges from salt solution. The competition between energy minimization and entropy maximization results in a particular equilibrium of the chains with cations around them.

Manning-Oosawa counterion condensation: Simplest version. The behavior of counterions around a cylindrical linear homogeneously-charged macromolecule have been investigated first by Manning [3,4] and Oosawa [5] in the framework of Counterion Condensation theory. Manning was interesting to derive the limiting law for polyelectrolyte like Debye-Hückel law for simple electrolytes. As was noted by Onsager, the energy of counterions in electrostatic potential of

[1] In 1962 Crick, Watson, and Wilkins have been awarded by Nobel Price for Physiology or Medicine for discovery of DNA.

the cylinder and the entropy of counterions in solution, have the logarithmic dependence on the volume. It can lead to the association of a part of cations with cylindrical polyelectrolyte. The same energy-entropy arguments result in no counterion condensation around a charged sphere, and accumulation of all cations near a highly charged plane.

The Manning theory describes the condensation of counterions as a transition between two states, the bound ("condensed") state and a free state. In bound state the cations are however free to move along the surface of polyelectrolyte. The cloud of condensed counterions interact with the molecule electrostatically with the energies higher than $k_{\rm B}T$, screens the electric field of the cylinder and allows the rest counterions to leave the molecule. The thickness of the cloud is of the order of several Å for uniformly charged cylinder with B-DNA parameters. This theory predicts that the linear charge density of macroions cannot thus exceed some critical value.

According to the Manning theory [6], in *salt-free* solution the fraction θ of polyelectrolyte charge, compensated by condensed counterions, is determined *only* by charge density of polyelectrolyte and ions' valence, z,

$$\theta = 1 - \frac{1}{z\xi}, \text{ if } \xi > 1, \tag{1.1}$$

and $\theta = 0$ otherwise (no counterion condensation exists for $\xi < 1$). Here $\xi = l_B/b$ is the dimensionless charge parameter, which is the ratio of the Bjerrum length,

$$l_B = \frac{e_0^2}{\varepsilon k_{\rm B} T} \tag{1.2}$$

(= 7.1Å for water at room temperature) to the average distance between the nearest unit charges along the axis of polyelectrolyte molecule, b. Here e_0 is the charge of the proton, $k_{\rm B}$, the Boltzmann constant, T, the absolute temperature, ε, the solvent dielectric constant. If $\xi > 1$ the polyelectrolyte is considered to be highly-charged and its charge density is reduced by condensed cations down to $\xi_{th} = 1$. Such charge reduction eliminates the divergency of polyion-counterion partition function that was used by Manning as the main argument for existence of counterion condensation. The modified version of the theory shows that in a salt solution fraction θ increases with salt concentration [6]. The modifications of the theory for simultaneous association of mono- and multi-valent cations have been developed [4].

Many small monovalent (and divalent) cations indeed have nearly identical binding constants to native DNA, that confirms the validity of Eq. (1.1) for them (see description in Ref. [6]). The binding of other cations, which may interact strongly with DNA (like Ag^+ and Cu^+), does not obey Eq. (1.1) [6]. However, due to its essential simplicity and the ability to describe DNA colligative, transport and osmotic properties [6], as well as binding equilibria of DNA in solutions [3, 4, 6–8], the Manning-Oosawa theory became a point of reference in the theory of highly-charged polyelectrolyte solutions.

Binding of cations: Since DNA can forms bonds with many cations, the latter are not distributed randomly in a close-shell around DNA. Some cations may interact with special sites in the grooves, another – with sites on the phosphates. Cations form hydrogen bonds or/and taking part in other specific interactions. Cation-DNA interaction can be strong enough to initiate changes in the DNA conformational structure [9]. Such interactions and their consequences became recently a subject of intensive experimental investigations. Indeed, high-resolution X-ray-diffraction experiments have shown that many mono- [10–12], some di- [13, 14] and almost all tri- [15, 16] valent biologically important cations reside preferentially in DNA grooves. Some

divalent cations bind on the phosphate strands (Ca^{2+}, Mg^{2+}). Molecular dynamics simulations have also shown that distribution of cations around DNA is not uniform, but follow the DNA helical symmetry [17, 18].

It means that the Manning's assumption of mobile cloud of condensed counterions is questionable for these cations, and counterion condensation theory needs modification to take into account the adsorption (or even chemisorption) of cations on DNA. Such effects was suggested for example for site-specific binding of divalent cations [19]. Indeed, many biological DNA-condensing cations usually have three or more amine groups that form site-specific hydrogen bonds with DNA, their estimated binding energy is $\sim$10 $k_{\rm B}T$/bp. The adsorption isotherms of binding of many cations to DNA have been measured and many model adsorption isotherm have been suggested [14]. In Section 1.3.2, as a possible modification of Manning counterion condensation theory, we explore plausible adsorption isotherm of cations into *two* DNA grooves.

Poisson-Boltzmann equation: The more straightforward tool for investigation of cations distribution around homogeneously-charged rods is the Poisson-Boltzmann (PB) equation. Usually the problem of potential distribution is formulated for hexagonal array of charged cylinders [2]. Then the potential distribution in the lattice is *approximated* by potential distribution around a *single* cylinder in its cell, the so-called "cell model" [22, 23], Section 1.5.1.

Bare DNA is however highly-charged and the linearized PB equation fails near the surface of the molecule – the nonlinear PB equation should be considered. The classical (and till now unique) solution of the nonlinear PB equation in cylindrical cell in "counterion-only" case was derived half a century ago independently by Fuoss *et al.* [24] and Alfrey *et al.* [25].

However, both linear and nonlinear PB equations lead to smooth decrease of concentration of cations with departure from the macroion axis and *do not reproduce the Manning's "cloud hypothesis"*. Moreover, the radius of the cylinder around DNA, where the fraction θ of cations is situated, diverges $\propto \sqrt{\text{cell radius}}$ [26]. Thus, the Manning theory gives incorrect description of counterion distribution near the polyion surface and also of the electrostatic free energy of polyelectrolyte solution [27].

Thus, the description of polyelectrolyte solutions on the basis of the (linearized) PB equation appears to be more self-consistent and theoretically justified. It is used therefore throughout this work. Known limitation of the PB equation for description of point-like charges near highly-charged surfaces ([28, 29] and Section 1.8.4), appears to be insignificant for DNA, which is highly-neutralized by condensed/adsorbed counterions, Section 1.5.1.

1.1.2 DNA condensation

Experimental observations: DNA condensation plays an extremely important role for storage of genetic information. Genomic DNA is usually a very long molecule and therefore it must be tightly packaged to fit into small space in a cell or virus particles. For example, the DNA of T7 bacteriophage is 10^{-4} times smaller in the phage head than in unpacked form [30], the chromosome DNA of *Escherichia coli* has 1.5×10^3 compression ratio. Genetic DNAs of eukatyotes form more complex chromatid structure, using histones and other proteins for packaging [2]. Knowledge of condensation process is essential also for understanding of such biological processes as transcription and replication, it can also be helpful for the purposes of gene therapy.

It has been suggested that many bacteriophages use *multivalent* cations to package their DNAs. Indeed, when polyamines known to exist in host bacteria are added to DNA in dilute so-

[2] Assemblies of 500Å long DNAs with density $\sim$350mg/ml (DNA-DNA interaxial separation R $\sim$33Å), indeed adopt hexagonal lattice [20, 21].

lution, DNA chains spontaneously form highly-concentrated toroids with similar shape and size as *in vivo* packaged DNA [31,32]. The addition of alcohol (methanol, ethanol) tends to produce more rod-like structures when condensation is provoked by $Co(NH_3)_6^{3+}$ [33] or spermidine^{3+} [34].

Generality: Many other highly-charged polyelectrolytes, such as F-actin [35], tobacco mosaic and fd virus [36], also condense into dense bundles. In the condensed state the helical segments are locally aligned, the volume fraction of solvent and DNA are compatible, and DNA helices may be separated by just one or two layers of water [37]. The aggregation does not depend crucially on the structure of polyions and condensing cations, and is similar for many biopolymers.

DNA condensation has been observed by a variety of techniques that detect changes in polymer size or chirality, including various forms of electron microscopy, total intensity and dynamic laser light scattering, sedimentation, viscometry, linear optical dichroism and circular dichroism, and fluorescence microscopy. In the latter case DNA condensation has been provoked by high concentration of PEG. High molecular wight PEG is also used in osmotic stress experiments. The condensation of single DNA in very dilute solutions is observed for long molecules [37], with plasmid-sized or smaller DNAs, several molecules are incorporated into condensed structure. Condensation is therefore difficult to distinguish form aggregation and precipitation [37, 38].

Multivalent cations: Many multivalent cations cause precipitation of DNA in solutions (polyamines spermidine^{3+} [30, 39] and spermine^{4+} [40, 41], inorganic cations $Co(NH_3)_6^{3+}$ [31, 42–44]). The basic proteins found in association with DNA (protamine, polylysine, histones, etc) are also known to facilitate DNA condensation. The measured DNA persistence length decreases down to ∼200Å at 25mM of $Co(NH_3)_6^{3+}$ (compared to 500Å at physiological concentration of NaCl) [45], that may help DNAs to condense into dense toroids. Experiments show that in solutions of multivalent cations DNAs condense when nearly 90% of their charge is neutralized by condensed cations [40], that is in agreement with Manning' $\theta \approx 0.92$.

Linkers: The dependence on linker length and charge was detected for DNA condensation with diaminoalkane $NH_3^+(CH_2)_nNH_3^+$. The condensation has been observed only for $n = 3, 5$ but not for $n = 2, 4, 6$ [46]. The formation of phosphate-cation-phosphate-links was also suggested as a mechanism of DNA collapse by tri-and di-amines [47], where the correlation between the equilibrium spacing of DNA lattice and the chains length of polyamines has also been observed.

Divalent cations: Recent experiments show that di-valent cations cause DNA condensation. For instance, Mn^{2+} can produce toroidal condensates of supercoiled plasmid DNA [3] [48]. The capability of Mn^{2+} to condense long linear DNA array has also been confirmed by the measurements of DNA-DNA force in osmotic stress experiments [49]. No DNA precipitation was observed in DNA solution with Mg^{2+} and Ca^{2+} salts [37].

Some di-valent cations (Mg^{2+} and putrescine^{2+}), which unable to cause DNA precipitation in water, do it however in alcohol-water mixtures [40]. Alcohols are known to decrease the dielectric constant of solution, that weakens the repulsion between DNA phosphates and provokes DNA condensation. Condensing agents generally act either by decreasing DNA charge (adsorption of cations) or by making DNA-solvent interaction less favorable (alcohols are excluded from DNA phase and effectively exert an osmotic pressure on it, like PEG). In alcohol-water solution also a lower concentration of $Co(NH_3)_6^{3+}$ is required to cause DNA condensation [33], whereas higher concentration are necessary in the presence of organic osmolyte glycine, which increase

[3] But not of linearized plasmid. It was detected that millimolar concentrations of Mn and supercoiling can cooperate to drive DNA condensation at room temperature in aqueous solutions [37].

the dielectric constant [50]. It supports the *electrostatic* basis of DNA condensation.

Structure of condensed DNA: DNA structure can change during condensation. Raman spectroscopy of protamine-calf thymus DNA complexes shows a modified B-form with bp unstacking [51]. Divalent transition metals cause deviations from normal B-form vibrational spectra [52, 53]; changes are also seen with cobalt hexammine and cobalt pentammine. Some sequences may be converted from B- to Z-conformation by $Co(NH_3)_6^{3+}$ [54]. Insertion of blocks capable to proceed B-to-Z transition stimulates DNA condensation with $Co(NH_3)_6^{3+}$, spermidine and spermine. It was shown that $MnCl_2$ can also induce right(B)-to-left(Z) helical transformation of GC-rich DNA sequences both at low and high ($\sim$ M) concentrations [55]. Runs of adenines (A-tracts) are considered as sources of sequence-induced DNA bending that produces unusually small toroids.

In the osmotic stress experiments on dense DNA arrays with transition cations [49] DNA however remains much more stable. Usual B-DNA helical structure follows from analysis of x-ray diffraction pictures, with no indication on DNA melting or backbone destabilization, Section 1.7.4.

Mono-valent cations: DNA condensation with multivalent cations depends on the presence of monovalent salt in solution. Indeed, the concentration of spermine^{4+} required for DNA condensation is linearly proportional to DNA concentration in the region of $c_{\rm DNA}$ =0.1$-$10 mM and increases also with concentration of NaCl in solution [44]. The concentration of $Co(NH_3)_6^{3+}$ required to cause DNA condensation also increases with addition of univalent salt in solution [56]. In *dilute* DNA solution NaCl can prevent DNA condensation: at $\sim$ 0.2M of NaCl DNA does not precipitate in the presence of spermine^{4+} [44]. The competitive adsorption of mono- and multi-valent cations as well as screening effects may cause such behavior. Note that in very dilute DNA solutions, $c_{\rm DNA} \sim 10^{-2}$mM, DNA precipitation is almost independent of spermine^{4+} concentration, but dependent on DNA *molecular weight* [56]. In Ref. [31] was also shown that DNAs shorter than 400bp do not condense into ordered structures. All these facts suggest that DNA condensation is far from being governed by a simple mechanism.

2D DNA condensation: Divalent cations can provoke DNA condensation also in 2 dimensions. Its understanding is very important since cationic-lipid-DNA complexes are one of the most promising synthetic vehicles for gene delivery applications. 2D array of parallel DNAs adsorbed on the surface of cationic membranes reveals an abrupt shrinking with increase of concentration of divalent Co^{2+}, Mg^{2+}, Ca^{2+}, and Mn^{2+} [57]. The equilibrium DNA-DNA spacing varies from $\approx$27Å for Mn^{2+} to $\approx$29Å for other ionic species, that is very close to $\sim$ 28Å measured in columnar DNA assemblies with Mn^{2+} [49]. The fraction of neutralized DNA charge is low, $\approx$ 63%, compared to 90% for 3D-DNA condensation [42]. Spermine and spermidine also stimulate 2D DNA condensation, no condensation with univalent salts was detected.

Intermolecular Forces. Osmotic Stress: Electrostatic [58] and hydration [59, 60] forces [4] govern the interaction of DNAs in dense assemblies. Forces between DNA double helices in *dense* hexagonal array have been measured directly as a function of distance, pushing the

[4]Indeed, many biomolecular surfaces tightly bind and organize at least one shell of water. Water molecules in this layer have the orientation of hydrogen bonds different than that of the tetrahedral hydrogen-bond network of bulk water. The perturbation of water in second and further water layers may occur. The overlap of these layers make the approach of molecules energetically very unfavorable, the cost per water molecule may reach a substantial fraction of the energy of hydrogen bond [60]. This produce a strong, usually, *repulsive* force.
"Hydration force magnitudes depend on the strength of surface water ordering, while the decay length and sign, attraction or repulsion, depend on the mutual structuring of water on the two surfaces. Attraction results from a complementary ordering, while repulsion is due to symmetrical structuring." [66]

molecules by external osmotic stress of PEG [61]. [5] In NaCl solution exponentially decaying forces are observed, that is consistent with DLVO-like electrostatic as well as with hydration forces [63]. In solutions of $Co(NH_3)_6^{3+}$ [64] and Mn^{2+} [49,66] an attraction force between DNAs can appear. The attraction becomes stronger with cation concentration in multimillimolar range and DNA columnar array shrinks spontaneously in the region of $8-10$Å DNA surface-to-surface separations. Monte Carlo simulations of DNA columnar hexagonal aggregate with divalent salts also reveal a DNA-DNA attraction at similar interaxial separations [65]. In Mn-DNA assembly the attraction is also stimulated by temperature increase. The specificity of cation binding to DNA determines the charge patterns on the molecules and can affect DNA-DNA electrostatic interaction. The change in hydration of DNA and cations upon binding can also influence the hydration forces.

Mechanisms of DNA condensation: What are the forces which overcome the configurational entropy of DNA molecule in solution, tight bending of the helix upon toroid-like condensation, and a huge electrostatic repulsion of DNA phosphates upon DNA condensation?

Statistical mechanics of *free polymer* collapse has been developed in the works of Lifshitz, Grosberg, and co-workers [67]. It goes back to the Onsager-Isihara model of appearance of spontaneous ordering in the system of long rigid cylindrical molecules [68] and to Flory-Huggins theory of polymer solution with polymer-solvent interactions [69]. The consideration of the finite DNA bending rigidity together with polymer-polymer, solvent-solvent, and polymer-solvent interaction leads to a close estimate of the size of toroids [70].

Columnar assembly: fluctuations and correlations: In dense columnar assembly the configurational entropy of DNA molecules is small and intermolecular interaction determine the behavior of the system. The mean-field Poisson-Boltzmann equation does not however lead to attraction between two uniformly-charges cylinders regardless of cation valence in solution. Thus, correlations and fluctuations effects beyond the mean-field may be responsible for DNA-DNA attraction.

Many theoretical models have been suggested for description of attraction between like-charged cylinders. The first idea of fluctuations of associated charges was put forward by Oosawa. In this picture the correlations between the non-uniform profiles of fluctuating cations may lead to attraction of like-charged rods [5,71]. This van der Waals-like attraction increases with decreasing temperature. The idea of correlated charge fluctuations has also been used in the "ion-bridging" model, which results in short-ranged attraction due to an alternation of positive and negative charges along the polyelectrolytes [72]. The mechanism of collective charge fluctuations in dense DNA assembly has also been developed [81].

The "zero-temperature" model of ion correlations proposed in [73] and developed on the

[5] The osmotic stress technique has been described in detail [62]. DNA exposed to a polymer solution (polyethylene glycol (PEG)) forms an ordered phase. PEG with high molecular weight is almost completely excluded from DNA phase exerting an effective osmotic pressure on the latter. In equilibrium, the activities of the exchanging water and salt are equal in the DNA and PEG phase. pH and chemical potential of salt are held fixed. The osmotic stress method is based on the measurement of interaxial spacing in this phase as a function of applied osmotic stress, varied by changing the PEG concentration.
Phases: X-ray-diffraction shows that DNA assembly at densities ∼400mg/ml is packed into 2D hexagonal lattice (line hexatic). In densier lattices ($\gtrsim$ 600mg/ml : interaxial separation $R \lesssim 25$Å) B-to-A transition can take place because of the dehydration of the sample [20]. With dilution, in solution of 500Å-long-DNA a liquid crystalline phase is observed (at ∼340mg/ml: R =32–34Å) and then a "blue" isotropic phase ($\lesssim$100mg/ml). In our model we assume that DNAs are long enough to adopt hexagonal lattice for 22Å$< R <$40Å interaxial separation, where we plot our results.
Note, no force measurements can be done in the region of abrupt shrinking of DNA assembly. Thus, this unstable region can correspond to either DNA-DNA attraction or to a much weaker repulsion.

basis of ionic crystal arguments for rods [74, 75] can also lead to attraction. The density functional approach [76] and the hypernetted chain approximation [77] were also shown to be capable to produce attraction between like-charged objects. The incorporation of finite ionic radius into the charge fluctuation approach [78] has made a connection between the theories of high- [5] and low- [73] temperature fluctuations, and also led to counterion-mediated DNA attraction. This theory was also successful in description of the formation of DNA bundles and their optimal size [79, 80].

In the majority of these models the attraction is possible in the presence of multivalent cations, but theoretically the sharing of monovalent associated cations can also lead to attraction [82]. The non-pair-wise-additive rod-rod counterion correlation force has been shown to be important for formation of DNA bundles [80] and for description of (collective) fluctuations in dense DNA arrays [81].

However, in all above-mentioned models DNAs interact as homogeneously-charged rods, no effect *helical symmetry* is considered. The models however are based on the correlations of "*artificial*" non-uniform patterns of associated cations on nearest rods. Computer simulations however show that the DNA electrostatic potential [83] and distribution of cations near the DNA surface [17] follow the symmetry and periodicity of the helix itself.

Kornyshev-Leikin theory: Recently the *first* theory of interaction of helical macromolecules has been developed, which takes *explicitly* into account the *helical distribution of DNA phosphates and adsorbed cations* [84]. The problem of potential distribution in the system of two helical molecules on the basis of the *linear PB equation* has been solved *exactly*. This theory is constructed for *zero temperature* in a sense that no fluctuations of adsorbed cations and correlations between them are considered. Nevertheless, the theory gives a *quantitative* electrostatic explanation of DNA condensation at high neutralization of DNA charge [85].

In this Chapter the DNA condensation with Mn^{2+} at elevated temperatures is investigated on the basis of the Kornyshev-Leikin theory [86]. In the next Section the model and the basic idea is discussed. In Section 1.4 the basic results and predictions of the Kornyshev-Leikin are summarized, as well as their experimental verifications. In Section 1.5 the free energy is constructed which takes into account the non-fixed distribution of adsorbed cations between DNA grooves. In Section 1.5.1 we solve the problem of potential distribution in the cell model on the basis of linearized PB equation. In Section 1.6.1 the basic equations for cations redistribution are derived, which are used for calculation of the interaction pressure in Section 1.6.2. In Section 1.7.2 the effect of model parameters on the force and the entropy curves is discussed. In Section 1.7.4 the other possible contribution to temperature-induced aggregation are discussed. In Section 1.8 the auxiliary results of the model and some results beyond the model are presented. In section 1.9 the special case of azimuthally frustrated DNA lattices is investigated in the framework of the Kornyshev-Leikin theory.

1.2 Specificity of Mn^{2+}-binding to DNA

Divalent metal ions plays an important, and in many cases essential, role in chemistry of nucleic acids and their structural and dynamic properties. For example, Mn^{2+} can, like Mg^{2+}, stabilize secondary and tetrary DNA structure and can replace Mg^{2+} in a variety of enzymatic reactions involving DNA and RNA. In particular, Mn^{2+} substitution for Mg^{2+} is known to increase the number of errors in replication.

$Mn^{2+}-DNA$ interaction: It was shown that in aqueous solutions of some di-valent cations DNAs self-assemble spontaneously into dense aggregates (Mn^{2+} and Cd^{2+} [37, 49]), whereas in

solutions of other cations does not (Ca^{2+} and Mg^{2+} [37]). It means that the capability of ions to condense DNA depends not only on their valences but also on where and how strong the cations bind to DNA. A large number of experimental works have been devoted to interaction of Mn^{2+} with DNA. It was argued that Mn^{2+} as well as some other divalent transition metal ions induce DNA condensation due to their ability to bind in the major DNA groove, particularly at N7 atom of guanine [14, 87].

The proton relaxation enhancement technique [88] has also shown that Mn^{2+} possesses two types of binding. The first, weak, electrostatic binding of cation with two contiguous phosphates, whereas the second, strong, binding appears to vary almost proportionally to GC content of DNA. It suggests that at small concentration Mn^{2+} binds preferentially to phosphates, whereas at moderate and large concentration – to bases, decreasing DNA melting temperature [89] (see Fig. 1.18). Binding of Mn to bases increases with GC content of DNA, as well as the effect of Mn on DNA condensation upon heating in osmotic stress experiments [49]. Many adsorption isotherms which take into account the specificity Mn^{2+}-binding to DNA have been suggested [14]

Double-helix of B-DNA: Such specificity of Mn^{2+}-binding cannot be taken into account within the framework of the DNA-rod-model. Indeed, *B-DNA is not a homogeneously-charged cylinder*, but has the periodic helical distribution of the phosphates (Fig. 1.1). *B*-DNA has two grooves, the ratio of their heights is nearly 2/3. Different cations may adsorb in different places on the DNA surface: in the minor groove (majority of alkali cations), in the major groove (multivalent cations, some divalent cations), or on the phosphate strings (as commonly believed, Ca and Mg).

Such preference in cations adsorption has been *explicitly* taken into account in the theory of electrostatic interaction of helical macromolecules, developed recently by A.A.Kornyshev and S.Leikin [84, 85] (it is summarized in Section 1.4). This theory shows that adsorption of cations in the major groove stimulates DNA-DNA attraction [85] due to *natural* charge separation along the DNA axis. That is why the multivalent cations adsorbing in the major groove cause DNA precipitation, whereas the majority of mono- and di-valent cations adsorbing in the minor groove do not [85]. *This theory is the basic tool of investigation of DNA phenomena in the present work.*

1.2.1 Puzzle of Mn-DNA condensation

However, several aspects of DNA condensation are still not completely understood. One such feature is the spontaneous DNA aggregation at elevated temperatures in the presence of Mn^{2+}. It has been investigated in the osmotic stress experiments coupled with X-ray diffraction [49]. This technique has been proved to be an extremely exact tool of measuring of force between various biopolymers in the last 20-50Å surface-to-surface separation in solution of various agents [90–93]. It is successfully used for measurement of DNA-DNA forces for a long time [62].

Phenomenon: The osmotic stress measurements reveal an abrupt compression of DNA columnar hexagonal aggregate upon heating in the presence of *only* $MnCl_2$ and $Mn(ClO_4)_2$ in multimillimolar range, in 10mM NaCl buffer solution [49]. The equilibrium separation between DNA axis is $28-30$Å. Similar effect was detected also for Cd^{2+}-salts [49]. With temperature DNA-DNA attraction becomes stronger, DNA-DNA force at large separation ($R \simeq 35$Å) decreases, and the transition pressure decreases as well, Fig. 1.11b. The force at very short distances ($R < 26$Å) is almost temperature-independent. *The larger the GC DNA content, the stronger the effect of* Mn^{2+} was detected [49], that indicates Mn binding in the major

groove. With increase of Mn^{2+} concentration the attraction also become stronger. The entropy increases nearly exponentially with compression of the DNA array, the typical value of entropy change is k_B/bp [66] (see Section 1.6). In this Chapter, on the basis of the Kornyshev-Leikin theory [84], we construct a model for description of this phenomenon [86].

1.2.2 Condensation of Mn-DNA upon heating is entropy-driven

Counter-intuitively, temperature-favored aggregation into ordered assemblies is quite common for biopolymers. It was observed not only with Mn^{2+}-DNA but also with collagen [94], light chains of myosin [95], hydroxypropylcellulose [93] and a number of other macromolecules. Thermodynamic laws dictate that temperature-favored processes must somehow increase the overall entropy of the system. Because formation of an ordered assembly obviously decreases the translational and configurational entropy of macromolecules, there must be some other source of the entropy increase that drives the temperature-favored assembly. Potential sources of entropy increase upon formation of Mn^{2+}-DNA aggregates are: (i) release of water molecules structured at DNA surface or in the solvation of Mn^{2+} ions, (ii) reorganization of Mn^{2+} ions bound to DNA [49], explored below, (iii) relaxation and partial melting of DNA backbone, resulting in higher conformational freedom of DNA bases. The question is, which (if any) of these mechanisms is the dominating one?

The release of structured water is a traditional explanation of temperature-favored aggregation of biological macromolecules [96]. Water release indeed governs the polymerization of tobacco mosaic virus [96], gelation of methylcellulose with temperature [96], and, most probably, the temperature-induced aggregation of collagen [94]. It was also proposed as the mechanism of temperature effect in Mn^{2+}-induced DNA condensation [49, 66]. However, no experimental evidence directly supporting this interpretation for Mn-DNA was found. Moreover, comparison of strongly temperature-dependent intermolecular forces in Mn^{2+}-DNA aggregates [49] and temperature-independent, but otherwise similar forces in $Co(NH_3)_6^{3+}$-aggregates [64] suggests that the temperature dependence may be a peculiar feature of Mn^{2+}.

1.3 Distribution of cations on DNA

1.3.1 The model of Mn-DNA aggregation

Here we consider the potential role of re-arrangement of bound Mn^{2+} in the entropy increase observed upon DNA condensation. We base our analysis on the following. It was argued previously that Mn^{2+} as well as some other divalent transition metal ions induce DNA condensation due to their ability to bind in the major groove, particularly at N7 atom of GC base pairs. (This is in contrast to Ca^{2+} and Mg^{2+} ions which have much higher binding affinity to DNA phosphates, do condense single-stranded DNA, but does not condense double-stranded DNA.) Such ion specificity can be explained by taking into account specific helical symmetry features of the DNA charge pattern [85]. Binding of counterions in the major groove induces a substantial charge separation resulting in areas of positive and negative charge alternating with 34Å helical periodicity along the molecular axis. Such charge separation allows alignment of molecules with oppositely charged areas facing each other in a zipper-like fashion, producing an electrostatic attraction sufficiently strong to condense DNA [85], Section 1.4.

Binding entropy: One expects cations to have much *larger translational freedom* in the wider major groove compared to ions bound to phosphates in the minor groove, Section 1.7. A shift in

the equilibrium towards binding in the major groove upon DNA aggregation may, thus, result in a substantial increase in the system entropy and lead to temperature-favored condensation. Thus, we formulate theoretical predictions for this mechanism and compare them with available experimental data. We also discuss the other possible mechanism where the possible entropy increase may come from increasing conformational freedom of DNA backbone (Sections 1.8.5 and 1.8.6).

We base our model on the theory of electrostatic interaction between helical macromolecules [84], Section 1.4, that accounts for realistic surface charge pattern of DNA backbone and distinguishes between different binding sites for counterions [85]. For the purposes of the present work we assume that Mn can bind only in minor or in major DNA groove.

1.3.2 Two-state adsorption model for Mn

Below a two-state adsorption model is constructed, where cations can reside on two types of adsorption sites, inside minor and major grooves of DNA. Each type of sites is characterized by adsorption energy and maximal number of sites. The electrostatic problem of adsorption on single DNA is treated within the framework of Debye-Bjerrum approximation. Adsorbed cations take part in chemical bonds with the corresponding sites in the grooves. The rest cations and anions are delocalized around DNA in the electric field of the molecule with renormalized charge. The introduced surface model leads to two coupled Frumkin-type adsorption isotherms for groove coverage. In the main text we use similar model for Mn^{2+}-binding to DNA in the aggregate.

Surface Free Energy (per unit length of DNA) of adsorbed cations can be written in the model of two-state lattice-gas [97] as

$$F_s = -k_{\rm B}T \ln \left\{ \prod_{i=1,2} \frac{\overline{N}_i!}{N_i!(\overline{N}_i - N_i)!} \mathrm{Exp}\left[\frac{e_0 z \Psi_i + E_i}{k_{\rm B}T}\right]^{N_i} \mathrm{Exp}\left[\frac{m}{2}\frac{N_i}{\overline{N}_i}\frac{X_i}{k_{\rm B}T}\right]^{N_i} \right\}, \qquad (1.3)$$

where subscript $i = 1, 2$ labels DNA minor and major grooves, respectively. Here E_i is the chemisorption energy, Ψ_i is the electrostatic potential in the grooves, $\overline{N}_i$ is the maximal number of adsorption sites, N_i is the number of adsorbed cations, m is the number of nearest neighbors in the pattern of adsorbed cations ($m = 2$ for linear cations patterns in the grooves), X_i is the interaction energy between the adsorbed cations. No inter-groove interaction between the adsorbed cations is considered. The partition function of the adsorbed cations is unity.

Applying the Stirling formula we get

$$F_s = \sum_{i=1,2} \left\{ N_i[ze_0\Psi_i + E_i] + \frac{N_i^2}{\overline{N}_i} X_i + k_{\rm B}T \left[N_i \ln \frac{N_i}{\overline{N}_i} + (\overline{N}_i - N_i) \ln \frac{\overline{N}_i - N_i}{\overline{N}_i} \right] \right\}. \qquad (1.4)$$

The terms in (1.4) correspond to the electrostatic and chemical energy of ion adsorption, intra-groove interaction, and lattice-gas mixing-entropy over the adsorption sites, respectively.

The potential in the grooves, Ψ_i, is assumed further to be equal to *uniform* electrostatic potential on the surface of the molecule. This simplification is not crucial for the purposes of this section since below we consider only the dependence of DNA charge compensation on E_i, which anyway are poorly known. They appear in Eq. (1.4) as a sum with electrostatic potential in the grooves.

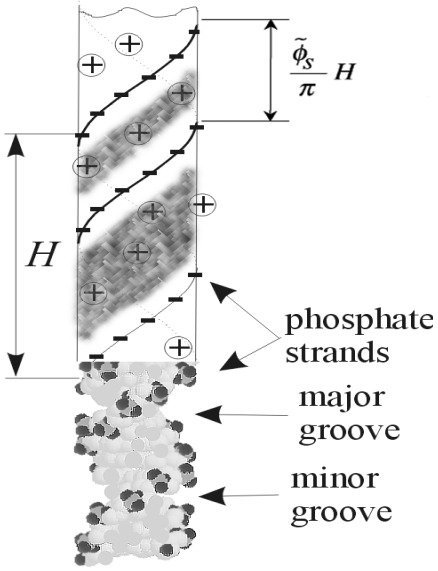

Figure 1.1: *Sketch of B-DNA structure.* H is DNA pitch, $\tilde{\phi}_s$ is the azimuthal half-width of the minor groove. Grey bands depict surfaces available for partitioning of adsorbed cations ($\oplus$).

Single molecule: To calculate this potential we consider DNA as infinite uniformly charged cylinder of the radius a in the cell model [22], in the presence of $z_+ : z_-$ salt with bulk cation concentration n_0. The solvent is modeled by a medium with dielectric constant $\varepsilon_{\rm w} \sim 80$ at all distances from the molecule. The distribution of dimensionless electrostatic potential $\Phi(r)=\frac{e_0\Psi_i}{k_{\rm B}T}$ around the cylinder in the cell model is described by Poisson-Boltzmann equation

$$\frac{d^2\Phi(r)}{dr^2} + \frac{1}{r}\frac{d\Phi(r)}{dr} = -4\pi l_B z_+ n_0 \left[e^{-z_+\Phi(r)} - e^{z_-\Phi(r)}\right] \tag{1.5}$$

with boundary conditions at the surface of the molecule and on the cell boundary

$$a\left.\frac{d\Phi}{dr}\right|_{r=a} = 2\xi(1-\theta), \quad \left.\frac{d\Phi}{dr}\right|_{r=R_s} = 0. \tag{1.6}$$

Here the total DNA fraction of compensated DNA charge is $\theta = (N_1 + N_2)/N$, where $1/N$ is the linear charge density of DNA phosphates. From these equations we need to find only the value of surface potential.

Below we linearize PB equation (1.5), assuming that $|\Phi(r)| \ll 1$. Having found the real DNA charge density, $\xi(1-\theta)$, we have to verify our assumption. Obviously, it will work for highly neutralized molecule and we need to set sufficiently large adsorption energies.

After linearization of Eq. (1.5), its solution, satisfying the boundary conditions (1.6), is

$$\Phi(r) = \frac{2\xi(1-\theta)}{\kappa a}\frac{I_0(\kappa r)K_1(\kappa R) + K_0(\kappa r)I_1(\kappa R)}{I_1(\kappa a)K_1(\kappa R) - I_1(\kappa R)K_1(\kappa a)}, \tag{1.7}$$

where $(I,K)_n(x)$ are the modified Bessel functions of nth order, and $\kappa = \sqrt{4\pi l_B z_+(z_+ + z_-)n_0}$ is the reciprocal Debye screening length.

Adsorption isotherm: The groove coverages, $\theta_i = N_i/\overline{N}_i$, are obtained from equalization of

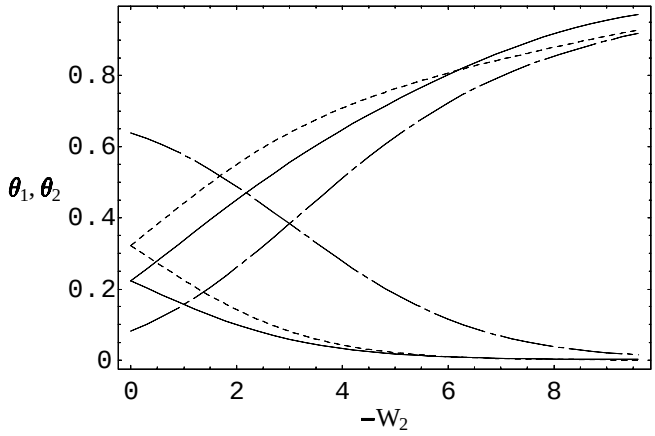

Figure 1.2: *Coverage of the major groove θ_2 increases with adsorption energy in the major groove, $-W_2$.* Coverage of the minor groove decreases. Parameters: $\xi = 4.2$, $a = 9.5$Å and $z = 3$, $n_0 = 0.1M$, $W_1 = 0$ (solid curves), $n_0 = 0.01M$, $W_1 = 0$ (dashed curves), $n_0 = 0.01M$, $W_1 = -3$ (dashed-dotted curves). θ_2 approaches unity for $-W_2 \sim 10 \div 12$. Higher adsorption energies may lead to a charge overcompensation since we have chosen that $\theta = 1$ under full occupation of *each* groove ($\overline{N}_i = N$). For $z = 4$ charge compensation is more pronounced due to the stronger contribution to binding constant form $\exp[-z_+\Phi(a)]$. At $n_0 = 0.1$ M and $\theta = 0.7$, $\Phi(a) \sim -0.92$ and the Debye-Bjerrum approximation can indeed be used. At smaller θ it works not so well.
Amount of associated cations on bulk cation concentration has been analyzed in another electrostatic model in [98], where competitive binding of several species on DNA has also been considered. The improvement of Manning counterion condensation theory for the case of tight ligand binding has been suggested and various binding regimes have been analyzed.

electrochemical potential of cation in the bulk and in the grooves, that leads to two coupled Frumkin-type adsorption isotherms

$$\frac{\theta_i}{1-\theta_i} = \frac{n_0}{n_w}\exp\{-z_+\Phi(a) - W_i - 2\chi_i\theta_i\}. \tag{1.8}$$

Here $n_w \approx 55.5M$; $W_i = (E_i - \mu^0)/(k_{\rm B}T)$, $\chi_i = X_i/(k_{\rm B}T)$ are the dimensionless chemisorption and intra-groove interaction energies. We use Eq. (1.7) to find $\Phi(a,\theta)$ and solve system (1.8) with respect to θ_i. The dependence of groove coverages on the concentration of adsorbing salt and adsorption energies are trivial (Fig. 1.2 for $z = 3$) and are only useful when the values of $\overline{N}$, W, and χ are known.[6]

We use similar two-state model for description of adsorption on DNA duplexes in hexagonal lattice, Section 1.5, and for analysis of concentrational dependence of pressure curves, Sec. 1.8.1.

[6] The determination of $\overline{N}_i$, W_i, and χ_i for DNA in a salt solution is however a complicated problem which requires a complex all-atom calculations. These parameters depend on the type of cation, DNA GC-content, bp sequence, and DNA properties in a given solution. For example, negative values of χ_i correspond to additional indirect attraction between adsorbed cations and 3-valued solution may take place [99]. The positive χ_i prevent accumulation of cations in the grooves. As we have insufficient information on cation adsorption patterns, we have plotted the results for $\chi_i = 0$.

1.4 Kornyshev-Leikin theory of electrostatic interaction of helical macromolecules

Below we briefly summarize the results of the theory of *electrostatic* interaction of two long *ideal* helical macromolecules, including DNA duplexes, developed recently by A.A.Kornyshev and S.Leikin [84, 85]. We start with used assumptions, then consider equations for electrostatic potential, and the final expression for interaction energy. Then we consider how this theory explains some puzzling features of interactions measured between DNA, collagen, and four-stranded guanosine helices.

The idea is to build an analytical theory which connects the molecular structure and intermolecular interaction. *The starting assumptions* are: (i) the dielectric response of water is linear and local, (ii) the molecules have straight, cylindrical, water-impermeable inner cores which are parallel to each other, (iii) the electric field in the water layer separating macromolecules can be described in the framework of the linear Debye-Hückel theory. The latter assumption does not capture the counterion condensation in the region of high electric field near DNA surface. However, it is possible to *combine* the Debye-Hückel field equation with the Bjerrum theory of ion pair formation [100, 101].

The strategy is to (i) solve the Debye-Hückel model for an electric field created by two opposing helical molecules with helical distribution of fixed surface charges *and* condensed counterions (*charges are treated explicitly)*; (ii) calculate the resulting electrostatic energy and the interaction energy [84].

1.4.1 Debye-Hückel-Bjerrum approximation

Interaction in electrolyte solution: Many biological macromolecules, DNA in particular, have high density of surface charges. Under physiological conditions, they are surrounded by an ionic atmosphere with the Debye length $\lambda_D \approx 7$Å. It is natural to expect here a nonlinear screening of fixed surface charges by electrolyte ions which is often treated within the mean-field, nonlinear Poisson–Boltzmann theory [5, 71, 102]. However, most of the counterions which contribute to the nonlinear screening lie within a narrow layer around each molecule. We refer to them as condensed counterions. The thickness of this layer can be estimated as $d_c \lesssim A/(4\pi l_B)$ [102], where A is the average area per elementary charge on the molecular surface and l_B is the Bjerrum length. For most biological macromolecules A$\lesssim$100Å^2 and $d_c < 2$Å. Mean-field description of an electric field inside such a thin layer is inappropriate.

We replace the Poisson–Boltzmann approximation by explicit treatment of condensed counterions. We describe the nonlinear screening layer (as well as the molecular-core/water interface) as infinitesimally thin surface that may have an arbitrary, inhomogeneous charge density. This surface contains fixed surface charges, chemisorbed ions, and mobile, condensed counterions. We treat the diffuse ionic atmosphere outside this surface within the Debye-Hückel theory.

This approach is similar to the two-state model (condensed counterions and free ions) commonly used in polyelectrolyte theory at nonvanishing salt concentrations. It is also similar to the Debye-Hückel-Bjerrum model which has proved to be quite successful in the theory of concentrated electrolyte solutions, including the theory of Coulomb criticality [100, 101, 103]. This model is accurate when the ratio of d_c to all other characteristic lengths in the system (the Debye length, the surface-to-surface distance between the molecules, the helical pitch, etc.) is small [102].

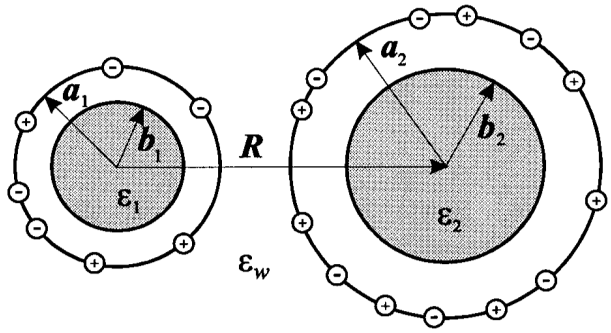

Figure 1.3: Horizontal cross section of two interacting parallel rod-like molecules. The molecules have water-impermeable cylindrical cores (shaded circles). Discrete charges are located on the coaxial cylindrical surfaces of, generally, larger radii. The charges may form any surface patterns.

1.4.2 The field equations and their solution

Several different models were proposed for electrostatic and non-electrostatic short-range force fields created by polar macromolecules in an aqueous solution [59, 104]. Here we use a generalized mean-field formalism [59, 105]. Within this formalism, the free energy of the system of two interacting DNA duplexes is expressed through the potential, φ, of a mean field as

$$E = \frac{1}{2}\int d^3\mathbf{r}\rho\left(\mathbf{r}\right)\varphi\left(\mathbf{r}\right), \tag{1.9}$$

where $d^3\mathbf{r}$ indicates integration over the system volume.

The volume density, $\rho\left(\mathbf{r}\right)$, of fixed surface charges (including adsorbed ions) is related to the surface charge density as $\rho_\nu\left(\mathbf{r}\right) = \sigma_n\left(z,\phi\right)\delta\left(R_\nu - a_\nu\right)$, where $\delta\left(x\right)$ is the Dirac $\delta-$function, $(z,\ \varphi,\ R_\nu)$ is the cylindrical coordinate system for the molecule ν, and α_ν are the radii of the outer cylinders where the charges are located, Fig. 1.3.

Inside the inner cores, the potential satisfies the Poisson equation,

$$\Delta\varphi^{core}\left(\mathbf{r}\right) = 0, \tag{1.10}$$

outside the cores, it obeys the linear Poisson–Boltzmann equation

$$\Delta\varphi^{out}\left(\mathbf{r}\right) - \kappa^2\varphi^{out}\left(\mathbf{r}\right) = 4\pi\alpha\rho\left(\mathbf{r}\right). \tag{1.11}$$

The continuity conditions,

$$\varphi^{core}\left(b_\nu\right) = \varphi^{out}\left(b_\nu\right) \tag{1.12}$$

and

$$\left(\nabla_n\varphi^{out}\left(\mathbf{r}\right)\right)_{r=b_\nu} = \gamma_\nu\left(\nabla_n\varphi^{core}\left(\mathbf{r}\right)\right)_{r=b_\nu} \tag{1.13}$$

relate φ^{out} and φ^{core} at the core/water interfaces. Here b_ν is the radius of the core of the νth molecule and $\nabla_n\varphi$ is the gradient component normal to the core surface.

For the electrostatic interaction,

$$\alpha = \varepsilon_w^{-1},\ \kappa = \lambda_D^{-1},\ \gamma_\nu = \varepsilon_\nu/\varepsilon_w, \tag{1.14}$$

$\rho\left(\mathbf{r}\right)$ is the electrostatic charge density; where ε_w and ε_ν are the macroscopic dielectric constants

of water and of the inner cores, respectively.

Linear vs. nonlinear: Despite its well-known limitations, the linearized Poisson–Boltzmann equation captures the qualitative force features between not too highly-charged objects and in many cases yields reasonable quantitative estimates. Many biological helices, like collagen, have an approximately equal number of positively and negatively charged groups or have a major fraction of the charge neutralized by adsorbed counterions, like spontaneously condensed DNAs. Then the *nonlinear screening* is pronounced only in the immediate vicinity of discrete charges so that the linear model may give accurate results at large and intermediate distances.

We neglect dielectric saturation of the solvent in immediate vicinity of ions since we will not apply our results at such short distances. Furthermore, we do not account for nonlocal dielectric response of water and neglect hydration forces that may be caused by specific water ordering near molecular groups. [7]

1.4.3 Interaction free energy

Potential distribution. The general procedure of calculation of electrostatic potential around two helical macromolecules with discrete charge distributions is presented in Ref. [84]. Here we summarize the results.

The potential created by the charge density $\sigma_\nu(z,\phi)$ on cylindrical surfaces of the radius $r_\nu > b_\nu$ around each core is [84, 85]

$$\widetilde{\varphi}_\nu(r,q,n) = \sum_{\nu,\mu=1}^{2} \sum_{m=-\infty}^{\infty} \tau_{n,m}^{\nu,\mu}(q,R)\,\widetilde{\sigma}_\nu(q,m), \tag{1.15}$$

where

$$\widetilde{\varphi}_\nu(r,q,n) = \frac{1}{2\pi}\int_0^{2\pi} d\phi \int_{-\infty}^{\infty} dz \varphi_\nu(r_\nu,q,n)\, e^{in\phi} e^{iqz} \tag{1.16}$$

is the Fourier transform of the potential near the core of the molecule ν in cylindrical coordinates (r,z,ϕ) associated with the molecular axis, $\widetilde{\sigma}_\nu(q,m)$ is a similar transform of $\sigma_\nu(z,\phi)$, and R is DNA-DNA interaxial separation.

At $R > 2b + \lambda_D$ the following approximation holds

$$\tau_{n,m}^{\nu,\mu}(q,R) - \tau_{n,m}^{\nu,\mu}(q,\infty) \approx \frac{r_\mu}{b^2} Q_{n,m}^{\nu,\mu}(q,R)\,\zeta_n^\nu(q,r_\nu)\,\zeta_m^\mu(q,r_\mu). \tag{1.17}$$

Here

$$\zeta_n^\nu(q,r_\nu) = \frac{K_n'(\widetilde{\kappa}b)\,I_n(\widetilde{\kappa}r_\nu) - K_n(\widetilde{\kappa}r_\nu)\,I_n'(\widetilde{\kappa}b) + \frac{\varepsilon_c q I_n'(qb)}{\varepsilon_w \widetilde{\kappa} I_n(qb)}\left[K_n(\widetilde{\kappa}r_\nu)\,I_n(\widetilde{\kappa}b) - K_n(\widetilde{\kappa}b)\,I_n(\widetilde{\kappa}r_\nu)\right]}{\left[K_n'(\widetilde{\kappa}b)\,I_n(\widetilde{\kappa}b) - K_n(\widetilde{\kappa}b)\,I_n'(\widetilde{\kappa}b)\right]\left(1 - \frac{\varepsilon_c q K_n(\widetilde{\kappa}b) I_n'(qb)}{\varepsilon_w \widetilde{\kappa} K_n'(\widetilde{\kappa}b) I_n(qb)}\right)}, \tag{1.18}$$

where

$$\widetilde{\kappa} = \sqrt{\kappa^2 + q^2}, \quad \kappa = \lambda_D^{-1}. \tag{1.19}$$

[7] Note that a phenomenological model of hydration forces [106] also uses the same field Eq. (1.11) with $\kappa = \lambda_w^{-1}$, $\gamma_\nu = 0$, where $\rho(\mathbf{r})$ is the density of hydration field sources, α is a phenomenological parameter, and $\lambda_w \approx 4$Å is a correlation length in water [59]. This is a purely empirical approach relating observed short-range, exponential interactions between polar surfaces in water to a phenomenological hydration field. This model gives a consistent description of a variety of existing experimental data [59].

The electrostatic propagators $Q_{n,m}^{\nu,\mu}(q,R)$ in Eq. (1.17) are in general very cumbersome [84], but for the case of interaction of two similar molecules ($b_\nu \equiv b$) with fixed surface charges lying directly on inner core/water interface ($a_\nu = b_\nu$), they can be reduced to [58, 85]

$$Q_{n,m}^{\nu,\mu}(q,R) = \frac{4\pi(-1)^{\mu n-\nu m} K_{n-m}(\widetilde{\kappa}R)}{\varepsilon_w \widetilde{\kappa}^2 K_n'(\widetilde{\kappa}b) K_m'(\widetilde{\kappa}b)} \propto e^{-\widetilde{\kappa}R},\ \nu \neq \mu,\quad \nu,\mu = 1,2 \tag{1.20}$$

$$Q_{n,m}^{\nu,\nu}(q,R) = -\frac{4\pi(-1)^{\nu n-\nu m} \Omega_{n,m}(\widetilde{\kappa}R,\widetilde{\kappa}b)}{\varepsilon_w \widetilde{\kappa}^2 K_n'(\widetilde{\kappa}b) K_m'(\widetilde{\kappa}b)} \propto e^{-2\widetilde{\kappa}R}, \tag{1.21}$$

where

$$\Omega_{n,m}(x,y) = \sum_{j=-\infty}^{j=+\infty} K_{n-j}(x) K_{j-m}(x) \frac{I_j'(y)}{K_j'(y)}. \tag{1.22}$$

Since $\varepsilon_c/\varepsilon_w \ll 1$, $\zeta_n^\nu \sim 1$ can be used for $(r_\nu - b)/b \ll 1$ [85].

Interaction energy: With the help these expressions the interaction energy between two rigid infinitely long DNAs for a given distribution of charges on their surfaces,

$$E_{int} = E(R) - E(\infty), \tag{1.23}$$

can be written as [58, 84, 85] (per unit length)

$$E_{int} = \frac{1}{2}\sum_{\nu,\mu=1}^{2}\sum_{n,m=1}^{\infty}\int_{-\infty}^{+\infty} dq Q_{n,m}^{\nu,\mu}(q,R)\, s_{n,m}^{\nu,\mu}(q), \tag{1.24}$$

where

$$s_{n,m}^{\nu,\mu}(q) = \lim_{L\to\infty}\left[\frac{\widetilde{\sigma}_\nu^{eff}(q,n)\,\widetilde{\sigma}_\mu^{eff}(-q,-m) + \widetilde{\sigma}_\nu^{eff}(-q,-n)\,\widetilde{\sigma}_\mu^{eff}(q,m)}{2L}\right], \tag{1.25}$$

$$\widetilde{\sigma}_\nu^{eff}(q,n) = \int_a^B \widetilde{\rho}_\nu(r_\nu,q,n)\,\zeta_n^\nu(q,r_\nu)\frac{r_\nu dr_\nu}{b_\nu}, \tag{1.26}$$

$$\widetilde{\rho}_\nu(r,q,n) = \frac{1}{2\pi}\int_0^{2\pi} d\phi \int_{-L/2}^{L/2} dz \rho_\nu(r_\nu,q,n)\, e^{in\phi} e^{iqz}. \tag{1.27}$$

It was assumed that the adsorbed and condensed counterions lie within the non-overlapping layers around DNA cores ($b \le r \le B,\ R > 2B$). Here $\widetilde{\rho}(r,q,n)$ is the Fourier transform of the charge density, and L is the length of the molecules.

Equations above define the *interaction Hamiltonian* for given $s_{n,m}^{\nu,\mu}(q)$. To account for fluctuations in $\widetilde{\rho}(r,q,n)$, one can add the energy of isolated DNA and chemical interaction of ions with DNA (Section 1.6) and calculate the *partition function*. The latter however is a complicated problem and it is not yet done. Thus, everywhere below we consider the consequences of the interaction Hamiltonian, but not of the free energy.

1.4.4 DNA-DNA interaction

The theory of counterion condensation [6] and the majority of the models of attraction between polyelectrolytes [5,71,73,74] (see Section 1.1.2 for details) presume that associated counterions are freely mobile. Such an assumption may hold for alkali metal ions. It is doubtful or breaks down for divalent alkali-earth and multivalent cations, since they possess strong chemical affinity to specific sites on the DNA surface. Therefore, while heuristically useful, theories of this type are not applicable to DNA condensation caused by these cations.

B-DNAs: Based on the experimental evidences, we assume that DNA-condensing counterions are chemisorbed and form a rigid, R–independent, $s_{n,m}^{\nu,\mu}(q)$ pattern. Then Eq. (1.24) gives the free energy of interaction between the molecules. Below we explicitly describe *phosphate strands as two helical lines of charges* and approximate various patterns of chemisorbed counterions by a *three-state model*, so that for adsorption pattern we can write

$$s_{n,m}^{\nu,\mu}(q) = 2\pi\overline{\sigma}^2\delta_{n,m}\delta(q+ng)\cos\left[q\Delta z\left(1-\delta_{\nu,\mu}\right)\right]\times \left[f_1\theta+(-1)^n f_2\theta-(1-f_3\theta)\cos\left(n\widetilde{\phi}_s\right)\right]. \quad (1.28)$$

Here $\overline{\sigma}=16.8\mu C/\mathrm{cm}^2$ and $-\theta\overline{\sigma}$ are the average effective surface charge densities of phosphates and of adsorbed counterions, respectively; $\delta_{x,y}$ and $\delta(x)$ are the Kronecker's and Dirac's deltas;

$$\tilde{\phi}_s \approx 0.4\pi \quad (1.29)$$

is the azimuthal half-width of the minor groove of B-DNA [2] (Fig. 1.1, 2.1); $g=2\pi/H$, $H\approx 34\text{Å}$ is the DNA pitch; f_i are the fractions of ions in the middle of the minor (f_1) and major (f_2) grooves and on the strands (f_3); $f_1+f_2+f_3=1$. DNA alignment is described by the axial shift $0\leq\Delta z\leq H$, Fig. 1.20c.

After substitution of Eq. (1.28) into Eqs. (1.24) and (1.25), we find the *free energy of interaction* of two B-DNAs

$$u_{int}=u_0\sum_{n=-\infty}^{+\infty}\left[f_1\theta+(-1)^n f_2\theta-(1-f_3\theta)\cos\left(n\widetilde{\phi}_s\right)\right]^2\times \frac{(-1)^n K_0(\kappa_n R)\cos[ng\Delta z]-\Omega_{n,n}(\kappa_n R,\kappa_n a)}{\left[(\kappa_n/\kappa)K_n'(\kappa_n a)\right]^2}, \quad (1.30)$$

where

$$u_0=8\pi^2\overline{\sigma}^2/\left(\varepsilon\kappa^2\right),\ \kappa_n=\sqrt{\kappa^2+n^2g^2}. \quad (1.31)$$

Expression (1.30) is the basic expression for DNA-DNA electrostatic interaction used in the present work for incorporation of non-fixed distribution of adsorbed cations treated in this Chapter, as well as non-ideality and torsional softness of DNA backbone, considered in the second Chapter.

1.4.5 Predictions of the theory and their experimental confirmation

Let us consider some puzzles of DNA-DNA interaction in the light of this theory.

DNA condensation: The different ability of ions to condense DNA can be addressed via binding of cations to different sites on DNA. Indeed, multivalent cations condensing DNA (spermine, spermidine, $Co(NH_3)_6^{3+}$) prefer to bind into major DNA groove [16]. According to the theory, it causes sharp charge separation along the axis of the molecule and facilitate DNA-DNA zipper-like attraction [85]. As shown in [85], the energy profit of adsorption in the major groove may be as huge as $50k_BT$ per DNA persistence length (≈ 500Å in physiological solution [45]). Analogously, ability of Mn^{2+} to condense DNA [49] follows from its preferential binding into the major groove [14]. Ca^{2+}, in spite of its much strong affinity to DNA phosphates, does not condense double-stranded DNA [37,49]: binding to the phosphates virtually neutralizes their charge and does not produce strong attraction.

DNA overwinding in fibers: Regardless of the distribution of cations in condensed layer, the theory predicts that interaction between DNA helices is more energetically favorable at integer number of bp per helical turn. In this case DNAs can align so that the axial phosphate ridges of one molecule face axial grooves separating the same ridge on opposite molecule, minimizing unfavorable interaction of the phosphates. The estimated energy gain from optimizing of phosphate-phosphate interaction [107] is sufficient to overcome the torsional DNA overwinding from 10.5 bp/turn in solution [108] to *nearly* 10 bp/turn in dense fibers [109].

Interaction between guanosine four-stranded helices. Another quantitative confirmation of the theory come from the recent measurements of the forces between helices formed by stacked guanosine monophosphate tetramers [110, 111]. At small concentration of KCl the DLVO repulsion is observed in this system. At high concentrations (up to molar KCl) a short-range exponential repulsion force is observed. Its decay length is ~ 0.7Å and *almost* independent on salt concentration. At even higher concentration (up to 4M), KCl facilitates the attraction with 1.3Å decay length and causes the spontaneous aggregation of guanosine helices [110,111].

The Kornyshev-Leikin theory predicts two dominant non-DLVO components of electrostatic force at high degree of surface charge neutralization by adsorbed cations. The first one is the image-charge repulsion of charge of one helix from dielectric core of another helix, Eq. (1.21). The decay length of this repulsion depends on the number of helical strands, N, on helical pitch, H, and on the Debye screening length of the solution, λ_D, as

$$\lambda = \frac{\lambda_D}{2\sqrt{1+\lambda_D^2\frac{4\pi^2N^2}{H^2}}}. \tag{1.32}$$

This expression gives for four-stranded guanosine helices the decay length $\lambda = 0.7$Å, almost independent on univalent salt concentration up to 2M. The second force is the exponential attraction with two time longer decay length, Eq. (1.20), which is also in agreement with experiments [111].

Predictions: The Kornyshev-Leikin theory also predicts B-to-A DNA transition triggered by adsorption of cations and DNA mesomorphism in dense assemblies [112], symmetry laws for interaction of duplexes [107], when the discreteness of DNA charges is taken into account explicitly. The Kornyshev-Leikin theory with conjunction of the model of cations re-distribution between the grooves, described in the next Section, give a quantitative explanation of spontaneous condensation of Mn-DNA assembly [86]. The generalization of the theory for the case of interaction of DNAs at all interaxial angles gives a qualitative [102] and quantitative [113] agreement for the cholesteric pitch of DNA liquid-crystalline phase.

1.5 The free energy of DNA assembly

In this Section we construct the total free energy of DNA columnar hexagonal assembly, as a sum of electrostatic interaction energy of all nearest neighbors and the free energy of adsorbed cations. Total free energy minimum determines the optimal distribution of cations on DNA surface for each value of DNA-DNA separation and temperature. It suggests a description of the observed temperature dependence of DNA-DNA force and the entropy increase with compression of DNA assembly [49].

DNA Structure and Charge Pattern. Following [84,85] we assume that bare DNA is a dielectric cylinder with two negatively charged phosphate strands described as spiral line-charges on the cylinder surface (Fig. 1.1). We distinguish two types of counterions: (a) bound, treated as surface charges, and (b) free, treated within the linear Debye-Hückel theory. We assume that bound counterions follow the basic double helical motif of DNA. For simplicity, we distinguish only minor and major groove binding and describe the two populations of counterions (with f and $1-f$ are the fractions of cations, respectively) as positively charged *spiral lines* located in the middle of the grooves .

We calculate the free energy of a hexagonally packed DNA assembly as a function of separation between helices and average occupation numbers for the *two* types of binding sites (the charge density of the corresponding spiral lines). This is the simplest model that accounts for helical symmetry of surface charge pattern on DNA and possible changes in this pattern upon DNA aggregation. It can be adopted to include other binding sites [85] and generalized to describe each phosphate or counterion as a discrete charge [107], but this should not change main qualitative features of intermolecular forces in Mn-DNA assembly.

Free Energy. Similar to interaction between two DNA helices analyzed previously [85], it is convenient to represent the free energy per molecule in DNA hexagonal aggregate in the following form

$$F\left(R,\theta,f\right)=F_{cyl}(R,\theta)+F_{helix}\left(R,\theta,f\right)+F_{ion}\left(R,\theta,f\right), \tag{1.33}$$

where F is the free energy normalized per unit length of one DNA molecule, R is the interaxial distance between molecules in the aggregate.

The components of the free energy are (i) F_{cyl}, the interaction free energy associated with the average net charge of DNA (free energy of an aggregate of homogeneously charged cylinders whose surface charge density is equal to $(1-\theta)\overline{\sigma}$, where $\overline{\sigma}$ is the average surface charge density of bare DNA); (ii) F_{helix}, the interaction free energy associated with inhomogeneous, helical pattern of DNA surface charges; and (iii) F_{ion}, the free energy of counterion binding to an isolated DNA molecule. All these free energy terms are considered in next sections.

1.5.1 Homogeneously-charged rods in the cell model

Homogeneously Charged Cylinders. Calculation of F_{cyl} is a traditional problem which was solved with the help of several approximations (see, e.g., [24, 114]). Since the differences between the results are not substantial and not critical for understanding the temperature sensitivity of Mn^{2+}-DNA forces, we use the simplest cell model [23, 24, 115], where a molecule is considered to be surrounded by a cylindrical cell of the radius R_s (R_s is chosen so that the volume of the cell is equal to the volume per polyelectrolyte molecule in solution).

We calculate the distribution of the electrostatic potential $\Psi(r)$ inside the cell using the Poisson-Boltzmann equation linearized around $\Psi(R_s)$ rather than around the potential in the bulk electrolyte and calculate $\Psi(R_s)$ self-consistently. This leads to renormalized screening

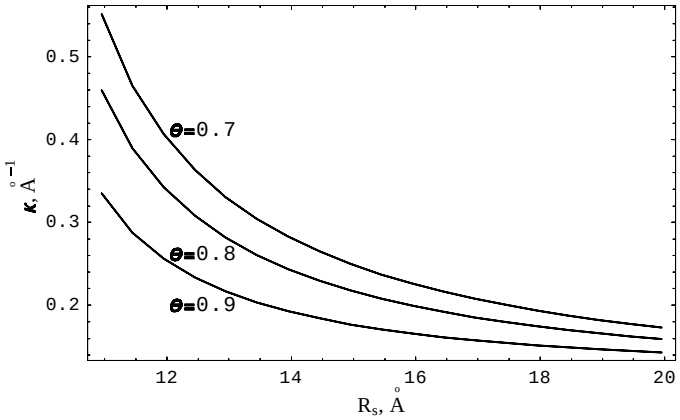

Figure 1.4: *Effective Debye screening length* κ^{-1} *in the cell model decreases substantially with increasing aggregate density* ($a = 9.5$ Å, $n_0 = 0.05$ M, and $T = 20$ °C). This length is calculated from PB equation for uniformly-charged cylinders with B-DNA parameters, Eq. (1.39).
Further this length is used as a Debye length for screening of interaction of *helical* charge distributions on DNAs (harmonics with $n > 0$ in Eq. (1.30)). Thus, in our model the interaction of uniformly-charged DNA rods affects the interaction of the *spirals* of phosphates and of the cations via κ.

length inside the aggregate, κ^{-1} (Fig. 1.4), which reflects Donnan equilibrium (higher concentration of counterions and shorter screening length compared to bulk electrolyte). We assume that θ is independent of R and of absolute temperature T (possible effects of $\theta\,(R_s)$ −dependence on DNA-DNA force are discussed in Section 1.8.1)

Note that the variation of the potential inside the cell $\Delta\Psi$ does not exceed $k_{\rm B}T/e_0$, where e_0 is the elementary charge, $k_{\rm B}$ is the Boltzmann constant (Fig. 1.7). Consistent with this, detailed comparison shows little difference between our simplified approximation and full numerical solution of the nonlinear Poisson-Boltzmann equation inside the cell, Fig. 1.7. [8]

Cell Model. The potential in a hexagonal lattice of parallel, homogeneously charged, cylindrical molecules can be approximated by a potential created by molecule inside a cylindrical cell of the radius $R_s = R\sqrt{\frac{\sqrt{3}}{2\pi}}$, where R is the interaxial distance in the lattice [22]. Within the standard cell model, the dimensionless electrostatic potential $\Phi(r) = \frac{e_0\Psi(r)}{k_{\rm B}T}$ is described by the nonlinear Poisson-Boltzmann (PB) equation

$$\frac{d^2\Phi(r)}{dr^2} + \frac{1}{r}\frac{d\Phi(r)}{dr} = -4\pi l_B n_0 z_+ \left(e^{-z_+\Phi(r)} - e^{z_-\Phi(r)}\right), \tag{1.34}$$

where a is the molecular radius, z_+ and z_- are the valences of electrolyte ions, n_0 is the bulk cation concentration, r is the radial distance from the axis of the cell, l_B is the Bjerrum length, and ε is the temperature-dependent dielectric constant of water [116]. No dependence of ε on salt concentration in the aggregate is considered, since (i) this dependence may not be the same as in the bulk [117], (ii) this may not allow us to get the final results in analytical form. No effects of dielectric saturation of the solvent and the nonlocal dielectric response are taken into account as well.

The following boundary conditions should be satisfied at the molecular surface ($r = a$) and

[8] In pressure calculation we use the solution of *linearized* PB equation, since the interaction of helical charge distributions is also considered within the linear PB theory, Eq. (1.30).

at the cell surface ($r = R_s$)

$$a\frac{\partial\Phi}{\partial r}\bigg|_{r=a} = 2\xi(1-\theta), \quad \frac{\partial\Phi}{\partial r}\bigg|_{r=R_s} = 0, \tag{1.35}$$

where $\xi = l_B/b$, and b is the mean distance between unit negative charges along the axis of the molecule.

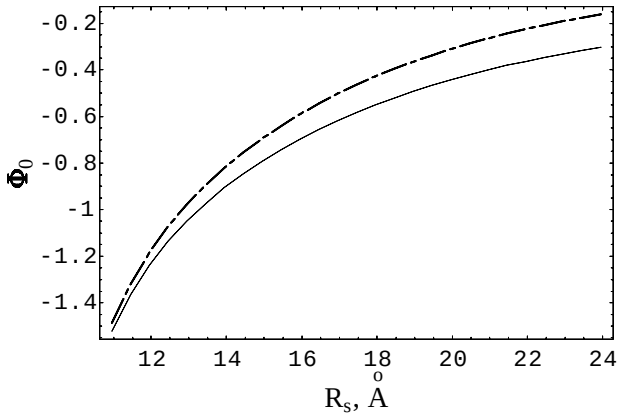

Figure 1.5: *The Donnan potential in the cell, Φ_{Don}, (solid) and the mid-point potential as the solution of linearized PB equation, Eq.* (1.38), (*dashed curve*) *are very close in dense assemblies.* Donnan potential is always more negative than $\Phi(R_s)$ due to the finite potential variation in the cell. Parameters: $z_+ = z_- = 2$, $n_0 = 0.05M$, $\theta = 0.8$, a=9.5Å.

Potential Distribution. For the purpose of the present work, we linearize Eq. (1.34) near the potential on the cell boundary $\Phi(R_s)$, i.e. we assume that

$$\Phi(r) = \Phi(R_s) + \delta\Phi(r), \tag{1.36}$$

where $|\delta\Phi(r)| \ll 1$. Eq. (1.34) then takes the form

$$\frac{d^2\delta\Phi(r)}{dr^2} + \frac{1}{r}\frac{d\delta\Phi(r)}{dr} = -B(-\delta\Phi(r) + A/B), \tag{1.37}$$

where $A = 4\pi l_B n_0 z_+ \left(e^{-z_+\Phi(R_s)} - e^{z_-\Phi(R_s)}\right)$, $B = 4\pi l_B n_0 z_+ \left(z_+ e^{-z_+\Phi(R_s)} + z_- e^{z_-\Phi(R_s)}\right)$. This procedure is similar to the so-called Debye-Bjerrum approximation and valid when potential variation inside the cell is small.

The solution of Eqs. (1.37), (1.35) is

$$\Phi(r,\theta) = \Phi(R_s) + \frac{e^{-z_+\Phi(R_s)} - e^{z_-\Phi(R_s)}}{z_+ e^{-z_+\Phi(R_s)} + z_- e^{z_-\Phi(R_s)}} + \frac{2\xi(1-\theta)}{\kappa a}\frac{I_0(\kappa r)K_1(\kappa R_s) + K_0(\kappa r)I_1(\kappa R_s)}{I_1(\kappa a)K_1(\kappa R_s) - I_1(\kappa R_s)K_1(\kappa a)}, \tag{1.38}$$

where

$$\kappa^2 = B = 4\pi l_B n_0 z_+ \left(z_+ e^{-z_+\Phi(R_s)} + z_- e^{z_-\Phi(R_s)}\right) \tag{1.39}$$

is an effective screening length inside the cell, Fig. 1.4. At $r = R_s$, this expression transforms into a transcendental equation that defines the value of $\Phi(R_s)$. Note that the variation of the

potential decreases and $\Phi(R_s)$ approaches the Donnan potential with decreasing R_s, Fig. 1.5.[9]

1.5.2 Linearized and nonlinear Poisson-Boltzmann equations

Linearized solution of PB equation always underestimates concentration of counterions near a highly charged surface and, therefore, overestimates the surface potential compared to the full solution. However, under the conditions relevant for the present work (strong adsorption of counterions on molecular surface and substantially reduced net surface charge density), the difference between the linearized solution and full numerical solution of nonlinear PB equation is small and can be neglected (Fig. 1.7). As we will see later, the effect of this difference on total interaction pressure in the assembly is much smaller than the effect of counterion re-partitioning between DNA grooves, Fig. 1.11.

Interaction Free Energy and Pressure. Repulsive pressure in the lattice of *homogeneously* charged cylinders, p_{cyl}, can be expressed via concentrations of ions at the cell boundary ($n_+(R_s)$ and $n_-(R_s)$) and in the bulk ($n_+(\infty)$ and $n_-(\infty)$) as follows [23, 118, 119][10]

$$p_{cyl} = k_B T \left[n_+(R_s) + n_-(R_s) - n_+(\infty) - n_-(\infty) \right] = n_0 k_{\mathrm{B}} T \left(e^{-z_+ \Phi(R_s)} - 1 + \frac{z_+}{z_-} e^{z_- \Phi(R_s)} - \frac{z_+}{z_-} \right). \quad (1.40)$$

For 2:1 electrolyte this expression turns into Eq. (1.53), used further for $MnCl_2$. Similar expression has been recently used to rationalize the measured osmotic properties of columnar DNA assembly in solutions of simple electrolytes [114]. The dependence of the mid-point potential and p_{cyl} on θ are presented in Fig. 1.6. The free energy of interaction between uniformly charged cylinders can then be found as

$$F_{cyl}(R) = -\int_{\infty}^{R} R' \sqrt{3} p_{cyl}(R')\, dR'. \quad (1.41)$$

Note that the electrostatic free energy of the cell, $F_{cyl}(R)$, and corresponding homogeneous electrostatic *repulsive pressure of overlapping diffuse double layers* $p_{cyl}(R)$, can also be found via a *charging process* as follows [118]

$$F_{cyl}(R) = \frac{1}{L} \int_0^1 \frac{d\lambda}{\lambda} \int_V \rho(r)_\lambda \Psi(r)_\lambda d^3 r, \quad p_{cyl}(R) = -\left(\frac{2 \partial F_{cyl}(R)}{\partial R^2 \sqrt{3}} \right)_{T,\sigma}, \quad (1.42)$$

[9] For DNA assembly with cell radius R_s in the presence of $z:z$ electrolyte with bulk concentration n_0, the Donnan potential, Φ_{Don}, is calculated from the electroneutrality condition of the DNA assembly [22] as $\Phi_{Don} = -z^{-1} \operatorname{arcsin h}[\zeta]$, where $\zeta = \frac{(1-\theta)}{2\pi z n_0 b (R_s^2 - a^2)}$. The *average* concentrations of positive (+) and negative (−) ions inside the cell are $n_\pm = n_0 e^{\mp z \Phi_{Don}} = n_0 \left(\pm\zeta + \sqrt{\zeta^2 + 1} \right)$. The *average* reciprocal Debye screening length is then $\kappa_{Don} = \sqrt{8\pi l_B z^2 n_0 \sqrt{\zeta^2 + 1}}$. For dense DNA assemblies the averaged quantities are close to determined from the linearized PB equation owing to a small potential variation in the cell.

[10] This pressure does not contain the contribution of all structural forces, viz. hydration, entropic, and depletion forces, as well as the Debye-Hückel correction to ideal pressure due to interaction of ions in electrolyte. Water is considered as structureless dielectric continuum.

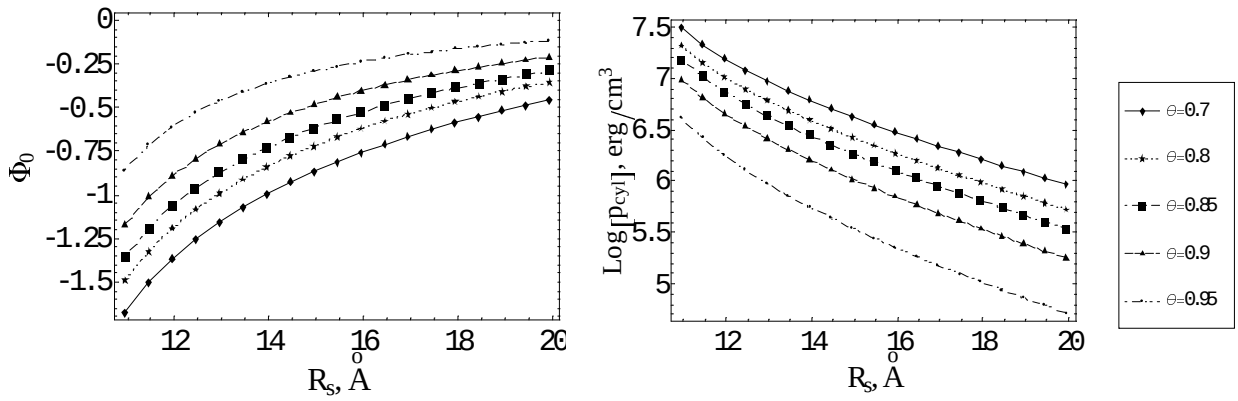

Figure 1.6: *The mid-point potential $\Phi_0 = \Phi(R_s)$ and corresponding homogeneous repulsive pressure, p_{cyl}, Eq. (1.53), decrease with increase of DNA charge neutralization, θ.* All curves are plotted at room temperature for 2:1 electrolyte with $n_0 = 0.05M$.

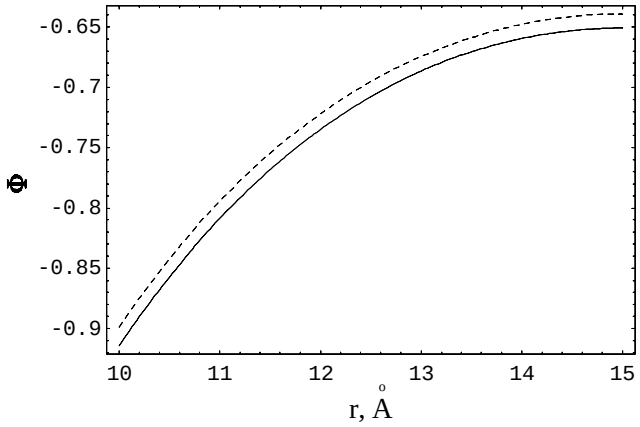

Figure 1.7: *Potential distributions in the cell obtained from linearized PB equation, solution (1.38) (solid curve) and full numerical solution of the nonlinear PB equation (dashed curve) are very close.* Parameters: $n_0 = 0.05M$ (2:1 electrolyte), $\theta = 0.85$, $R_s = 15$ Å, and $T = 20$ °C. The potential drop between the cylinder surface and cell boundary remains small (< 1) and the agreement of the two solutions remains reasonably good in the whole range of R_s and temperatures used in the present Chapter.

Note that in spite of the fact that PB equation together with boundary conditions (1.35) contains the electroneutrality of the cell, after linearization the approximate solution satisfies this condition within several percent accuracy. The simplicity of used numerical scheme also results in a mistake of the same order.

Numerical solution: To find numerical solution of nonlinear PB equation (1.34) we start with the derivative of electrostatic potential on the cylinder surface, Eq. (1.35) and some value of surface potential. The simplest finite-difference scheme shows that at some separation from the cylinder surface the potential derivative turns to zero. Then, to compare with approximate solution of linearized PB equation at proper cell radius, Eq. (1.38), we choose such value of surface potential for numerical scheme, that the calculated potential has the turning point at the same cell radius R_s, as analytical solution. Doing so, fortunately, we reduce the problem of solving of nonlinear equation with boundary conditions at $r = a$ and at $r = R_s$ for potential derivatives to the problem with initial conditions on one end of the interval, that is much simpler.

where ρ is the charge density of ions, λ is the charging parameter, σ is the surface charge density of the cylinder, $V = LR^2\sqrt{3}/2$ is the cell volume. For cylindrical geometry the identity of this pressure with Eq. (1.40) can be proved analogously to the planar case [118]. Note that the calculation of p_{cyl} by charging cycle is more complicated if $\theta = \theta(R_s)$.

In the limit $\kappa R_s \gg 1$ the mid-point potential and the uniform pressure are given by simple formula (in $z:z$ electrolyte)

$$\Phi_0 \equiv \Phi(R_s) = -\frac{2\xi(1-\theta)}{\kappa_0 a K_1(\kappa_0 a)}\sqrt{\frac{2\pi}{\kappa_0 R_s}}e^{-\kappa_0 R_s}, \quad p_{cyl} = z^2 n_0 k_B T\left[\frac{2\xi(1-\theta)}{\kappa_0 a K_1(\kappa_0 a)}\right]^2 \frac{2\pi}{\kappa_0 R_s}e^{-2\kappa_0 R_s}, \tag{1.43}$$

where $\kappa_0 = \sqrt{8\pi l_B z^2 n_0}$. This expression reproduces the exact results for reasonable θ and n_0 at $R_s \gtrsim 18$Å.

1.6 Total interaction pressure

1.6.1 The minimum of the free energy

Assumptions and approximations: Following [85] we assume that all molecules in Mn-DNA aggregate are aligned parallel to each other *without any axial shifts*. Such alignment is energetically favorable at large separations or when a majority of counterions is bound in the major groove. For simplicity, for description of temperature-dependent forces below we neglect the possibility of other, frustrated alignments (azimuthal alignment of DNAs in hexagonal lattice is considered in Section 1.9).

We take into account only the interaction between *nearest-neighbor* molecules because of the rapid, exponential decay of $\alpha_n(R)$. [11] We assume that considered electrostatic interaction is pair-wise additive. [12] We assume also that all DNAs in the aggregate have identical fraction of cations bound in the major/minor grooves and the same fraction of charge compensation. [13]

Helices. We approximate the free energy associated with the helical distribution of charges on DNA by the sum of pair interaction free energies for nearest neighbors derived in [85], (see Eq. (1.30))

$$F_{helix}(R,\theta,f) = \frac{48\pi^2\overline{\sigma}^2}{\varepsilon\kappa^2}\sum_{n=1}^{\infty}\alpha_n(R)[f\theta + (-1)^n(1-f)\theta - \cos(n\tilde{\phi}_s)]^2, \tag{1.44}$$

where $\varepsilon = \varepsilon_w$ is the macroscopic dielectric constant of water and

$$\alpha_n(R) = \frac{(-1)^n K_0(\kappa_n R) - \sum\limits_{j=-\infty}^{j=+\infty}[K_{j-n}(\kappa_n R)]^2\,\frac{I_j'(\kappa_n a)}{K_j'(\kappa_n a)}}{[(\kappa_n/\kappa)\,K_n'(\kappa_n a)]^2}. \tag{1.45}$$

[11] We expect that the main conclusions of the model will not change substantially if the interaction beyond nearest-neighbor are taken into account. The effect of DNA-DNA-interaction on counterions re-distribution can become even stronger due to a larger amount of interaction pairs which from each molecule in the assembly.

[12] This approximation is valid for direct electrostatic interaction, but is questionable for fluctuation-induced forces. For instance, non-pairwise-additive effects of (collective) fluctuations of associated cations was shown to be important for dense assemblies of charged DNA *rods* [81].

[13] Thus, only symmetric solutions are considered, some results of behavior of DNAs with non-symmetric distributions on the lattice are considered in Section 1.8.3.

Here $K_n(x)$, $K_n'(x)$ and $I_n(x)$, $I_n'(x)$ are the modified Bessel functions of nth order and their derivatives, $\kappa_n = \sqrt{\kappa^2 + 4\pi^2 n^2/H^2}$, H is the DNA pitch, a is the DNA radius, and $\tilde{\phi}_s$ is the azimuthal half-width of the minor groove (see Fig. 1.1, 2.1).

Counterion Binding. We neglect short-range interactions between bound cations in the grooves and approximate the free energy of counterion binding by

$$F_{ion}(R,\theta,f) = N_1(E_1 - TS_1) + N_2(E_2 - TS_2) - TS_{conf}. \tag{1.46}$$

Here E_i is the binding energy at a given site in minor ($i=1$) or major ($i=2$) grooves. Note that it may have both electrostatic and non-electrostatic components. S_i is the binding entropy. The axial densities (number per unit length of DNA) of counterions bound in each groove are

$$N_1(R) = \frac{f\theta}{z_+ b}, \quad N_2(R) = \frac{(1-f)\,\theta}{z_+ b}, \tag{1.47}$$

where b^{-1} is the axial density of charged phosphate groups on bare DNA ($b = 1.7$Å), z_+ is the counterion charge number ($z_+ = 2$ for Mn-DNA). Under the *assumption* that the adsorbed ion is confined between four nearest neighbor phosphates, the linear density of binding sites in each groove is $\overline{N} = 1/(2b)$.[14]

To find the configurational entropy of the distribution of bound counterions among these sites, we use a two-state ideal lattice-gas model [28,97] which gives (Section 1.3.2)

$$S_{conf}(R) = k_{\rm B} \sum_{i=1,2} \left[N_i \ln \frac{\overline{N}}{N_i} + (\overline{N} - N_i) \ln \frac{\overline{N}}{\overline{N} - N_i} \right]. \tag{1.48}$$

We assume that $\Delta E_{ads} = E_2 - E_1 > 0$ *because of closer proximity of* Mn^{2+} *to negatively charged phosphates. We also assume that*

$$\Delta S_{ads} = S_2 - S_1 > 0, \tag{1.49}$$

e.g., because of larger space per binding site available in the major groove.

Counterion Partitioning. Eq. (1.33) defines the free energy functional with respect to the minor groove fraction f. Neglecting fluctuations, minimization of Eq. (1.33) yields an equation on f [15]

$$f = \frac{1}{2} + \frac{1}{2} \left\{ \frac{(1+K)}{\theta(1-K)} - \sqrt{\left[\frac{(1+K)}{\theta(1-K)} - 1 \right]^2 + \frac{4K}{\theta(1-K)}} \right\}. \tag{1.50}$$

Here the partition coefficient K depends on f and R

[14] Note again that the number of adsorption sites in each groove depends on many factors, GC DNA content is one of them. In principle, we can account to GC-factor choosing or, if the experimental data are sufficient, calculating the density of adsorption sites in each groove and the adsorption energies. The calculation of surfaces available for cation binding, becomes then however more problematic. Thus, in the main text we have simplified the model assuming equal number of adsorption sites in both grooves, $\overline{N}_1 = \overline{N}_2$. We vary the groove coverages thus only by adsorption energies.

[15] This solution follows from the free energy extremum: $\frac{N_2}{N-N_2} \frac{\overline{N}_1 - (N-N_2)}{\overline{N}_2 - N_2} = K$, where, in general, $\overline{N}_1 \neq \overline{N}_2$.

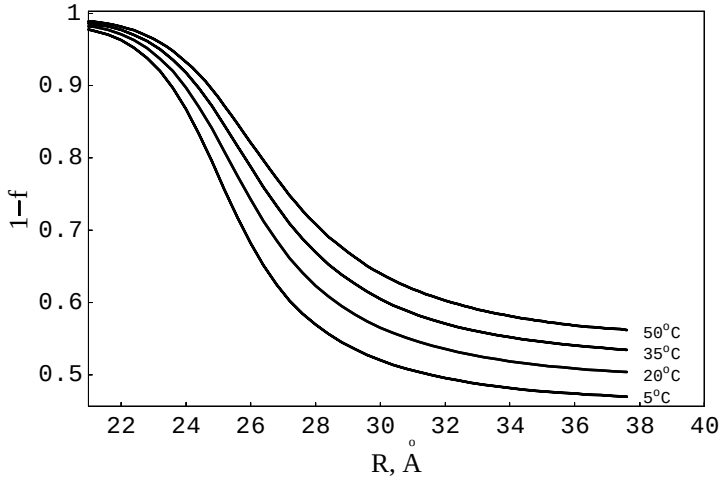

Figure 1.8: *The change in the fraction of cations bound in the major groove,* $1-f$*, upon aggregate compression at* $n_0 = 0.05$ *M,* $\theta = 0.85$*,* $\Delta E_{ads} = 4k_BT$*, and* $\Delta S_{ads} = 4k_B$*.* Under these conditions, 4-5 cations per helical pitch of DNA move from minor to major groove upon aggregate compression.

$$K(f,R) = K\left(f,\infty\right)\exp\left\{-\frac{192\pi^2 b\overline{\sigma}^2}{\varepsilon\kappa^2 k_{\rm B}T}\sum_{n=1}^{\infty}\alpha_n(R)\left[(-1)^n-1\right]\left[f\theta+(-1)^n(1-f)\theta-\cos(n\tilde{\phi}_s)\right]\right\}, \tag{1.51}$$

and $K\left(f,\infty\right) = \exp\left[-\left(\Delta E_{ads} - T\Delta S_{ads}\right)/\left(k_{\rm B}T\right)\right]$. The solution of Eq. (1.50) with compression of the aggregate is presented in Fig. 1.8.

1.6.2 Intermolecular forces

Intermolecular Forces in Mn^{2+}-DNA aggregates were experimentally studied by the osmotic stress technique [59, 62] which directly measures an interaction pressure

$$\Pi(R) = -\frac{dF(R)}{dA} = -\left[\frac{\partial\left(F_{cyl}(R,\theta)+F_{helix}\left(R,\theta,f\right)\right)}{\sqrt{3}R\partial R}\right]_{f,\theta} \equiv p_{cyl}(R,\theta)+p_{helix}(R,\theta,f), \tag{1.52}$$

where $A = \sqrt{3}R^2/2$ is the cross-section area per molecule in the aggregate. Here we took into account that $\partial F/\partial f = 0$ and $\partial F_{ion}/\partial A = 0$. In expression (1.52) we did not take the derivative $\frac{\partial F}{\partial\kappa(R)}\frac{\partial\kappa(R)}{\partial R}$ into account.

The two contributions to the interaction pressure are given by Eq. (1.40)

$$p_{cyl} = n_0 k_{\rm B}T\left(e^{-\frac{2e_0\Psi(R_s,\theta)}{k_{\rm B}T}}+2e^{\frac{e_0\Psi(R_s,\theta)}{k_{\rm B}T}}-3\right), \tag{1.53}$$

where n_0 is the concentration of $MnCl_2$ in the bulk solution outside the DNA aggregate; and

according to Eq. (1.44) [85]

$$p_{helix} = \frac{16\sqrt{3}\pi^2\overline{\sigma}^2}{\varepsilon R}\sum_{n=1}^{\infty}\frac{(-1)^n K_1(\kappa_n R) + 2\sum\limits_{j=-\infty}^{j=+\infty}\frac{K_{j-n}(\kappa_n R)K'_{j-n}(\kappa_n R)I'_j(\kappa_n a)}{K'_j(\kappa_n a)}}{\kappa_n[K'_n(\kappa_n a)]^2}\times$$
$$[f\theta + (-1)^n(1-f)\theta - \cos(n\tilde{\phi}_s)]^2. \quad (1.54)$$

1.7 Entropy change and the force

Entropy. The model accounts for the following sources of the entropy change with variation of R: (a) the entropy of bound counterions (Fig. 1.9), (b) the entropy of free electrolyte ions and counterions and (c) the entropy of water polarization fluctuations (at each temperature, we use the observed value of $\varepsilon(T)$) (Fig. 1.10).

With compression of DNA assembly and with temperature increase adsorbed cations move into the major groove, that make the patterns of adsorbed cations more complimentary and stimulate DNA-DNA attraction. Within this model, the dominating contribution to the entropy change is associated with repartitioning of bound counterions, Fig. 1.9,

$$\Delta S(R) = \Delta N\,(R)\,\Delta S_{ads} + \Delta S_{conf}\,(R)\,, \quad (1.55)$$

where $\Delta N\,(R) = N_2\,(R) - N_2(\infty)$ is the axial density of cations transferred to the major groove with compression and $\Delta S_{conf}\,(R) = S_{conf}\,(R) - S_{conf}\,(\infty)$ in the difference in configurational entropy of distribution of cations between the grooves, see Eq. (1.48). The latter entropy increases with compression of the aggregate when in the final state the cations have more possibilities to arrange themselves over the available adsorption sites in both grooves.

1.7.1 Partition function of adsorbed cation

Partition function of adsorbed cation in ith DNA groove, q_i, is approximated in the main text by its *translational* part only. I.e., ΔS_{ads} is determined by translational freedom of adsorbed cations in the adsorption sites in the grooves, along DNA surface. In general, due to the difference in the strength of adsorption bonds, cations in the grooves may have different vibrational entropy of motion perpendicular to DNA surface, along the adsorption bond. We explore it below but neglect in the main text. [16]

The partition function q_i, can be written as $q_i = q_i^{tr} q_i^{vibr}$ [28, 120], where q_i^{tr} is the *translational* partition function which depends on available surfaces for translational motion of a cation, q_i^{vibr} is the *vibrational* partition function which describes the vibration of cation *perpendicular* to DNA surface. q_i^{tr} can be calculated as the partition function of point particle with

[16] We neglect also by the change in *hydration* of cations and DNA upon binding. Below we want to explore *separately* the electrostatic consequences of transfer of cations between the grooves. We realize that the changes in hydration can be very important, since the entropy release per one water molecule transferred from the adsorbed state into the bulk can be $\sim k_B T$. However the *quantitative* model, which would describe the specificity of Mn^{2+} compared to cobalt hexammine, requires much more experimental data on DNA hydration in the dense arrays in the presence of these cations.

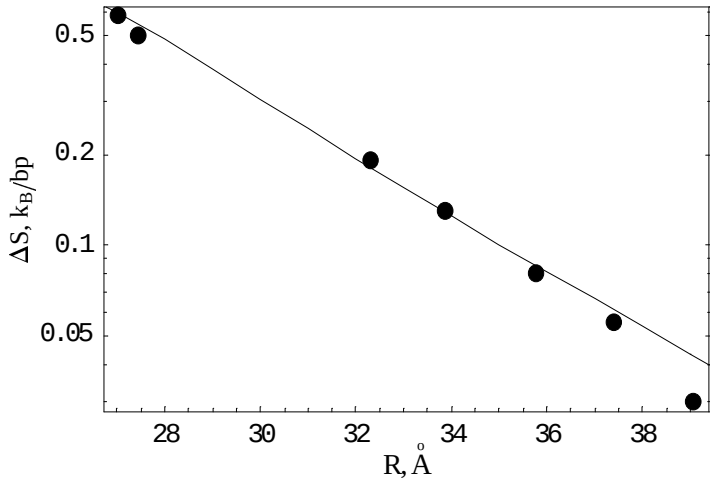

Figure 1.9: *The calculated (solid curve) change in the entropy of adsorbed cations and measured (filled circles, data from [66]) change in the total entropy upon compression of DNA aggregate at* $n_0 = 0.025$ *M and* $T = 35$ ^{o}C. The same θ, ΔE_{ads}, and ΔS_{ads} as in Fig. 1.8 were used for the calculation. Note that the value of the entropy change depends on ΔS_{ads}, but the *decay length* is almost independent of ΔS_{ads}. This length appears in the model naturally and depends on the decay lengths of Kornyshev-Leikin interaction energy, (1.30), and on the density-dependent Debye screening length $\kappa\,(R_s)$, determined from Eq. (1.39).

mass m in square box of the size a_i [28]

$$q_i^{tr} = \frac{2\pi m k_B T}{h^2} a_i^2 \equiv \frac{a_i^2}{\lambda_{Br}^2}, \tag{1.56}$$

where λ_{Br} is the thermal de Broglie length of Mn^{2+}. Considering each cation as harmonic oscillator of frequency ν_i, we get

$$q_i^{vibr}\,(T) = e^{-\frac{h\nu_i}{2k_{\rm B}T}} \Big/ \left(1 - e^{-\frac{h\nu_i}{k_{\rm B}T}}\right). \tag{1.57}$$

Then $K \sim \exp\left\{-\frac{F_2 - F_1}{k_{\rm B}T}\right\} = \frac{s_2}{s_1}\frac{q_2^{vibr}}{q_1^{vibr}}$, where $s_i \simeq a_i^2$ are the surface area per one adsorption site in each groove. The entropy difference upon transfer of ΔN cations (per unit length of DNA) from the minor to major groove is given by

$$\Delta S = [\Delta N(\Delta S^{tr} + \Delta S^{vibr}) + \Delta S_{conf}]. \tag{1.58}$$

Here the difference of translational and vibrational entropy per cation is, respectively,

$$\Delta S^{tr} = S_2^{tr} - S_1^{tr} = k_{\rm B} \ln\left(\frac{s_2}{s_1}\right) \tag{1.59}$$

and

$$\Delta S^{vibr} = S_2^{vibr} - S_1^{vibr}, \quad S_i^{vibr} = k_{\rm B}\left[\frac{w_i}{e^{w_i} - 1} - \ln\left(1 - e^{-w_i}\right)\right], \quad w_i = \frac{h\nu_i}{k_{\rm B}T} \tag{1.60}$$

For strong adsorption $w_i \gg 1$. Adsorption energies were assumed to be temperature-independent

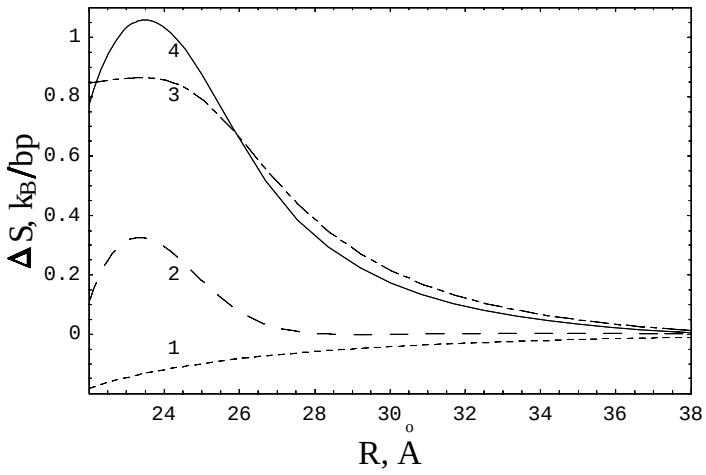

Figure 1.10: *Entropy change with interaxial separation upon DNA aggregate compression.* Following Eq. (1.33) we distinguish three different contributions to entropy: (1) $S_{cyl} = -\partial F_{cyl}/\partial T$, (2) $S_{helix} = -\partial F_{helix}/\partial T$, and (3) $S_{ion} = -\partial F_{ion}/\partial T$. The total entropy change (4) is dominated by the entropy of repartitioning of ions (S_{ion}). The values of S_{ion} were calculated from Eqs. (1.46), (1.48), (1.55). The values of S_{cyl} and S_{helix} were obtained by numerical differentiation of F_{cyl} and F_{helix} where we took into account the measured [116] temperature dependence of water dielectric constant $\varepsilon\,(T)$.
Note that the entropy change due to the complex derivatives, $-\frac{\partial F_{cyl,\,helix}}{\partial \varepsilon(T)}\frac{\partial \varepsilon(T)}{\partial T}$, comes from re-arrangement of water molecules associated with interaction upon the change in DNA-DNA interaxial separation. Like for two point like-charged particles [29], for homogeneously charged cylinders this contribution is negative (curve 1). The dependence of F_{helix} on temperature come primarily from dependence of pre-exponential factor $u_0\,(\varepsilon\,(T))$, Eqs. (1.30), (1.44), and may be either positive or negative.

in Eq. (1.58). Further we neglect by vibrational degree of freedom of cations, $q_i^{vibr} = 1$. Thus $\Delta S_{ads} \equiv \Delta S^{tr}$ and entropy change is determined by different available adsorption surfaces ($s_2 > s_1$) and by configurational entropy of cation distributions, ΔS_{conf}, Eq. (1.55).

1.7.2 Parameters of the model

The overall entropy change and, thus, the temperature dependence of intermolecular forces depend primarily on ΔS_{ads}. It is the most important adjustable parameter of the model, which, however, cannot be accurately calculated, based on present knowledge. Indeed, the value of ΔS_{ads} may depend on the strength of counterion binding in each groove, space available within each binding site, counterion and binding site hydration, etc. It may depend both on the type of counterion and base pair composition of DNA. The two other adjustable parameters, θ and ΔE_{ads}, also depend on some of these factors and may vary upon a change in experimental conditions. Our goal here is to see whether the observed force curves and entropy changes can be explained at reasonable values of ΔS_{ads}, θ and ΔE_{ads}. For the calculations, we use the macroscopic water dielectric constant $\varepsilon(T)$ [116] and the following values of other parameters: $a \approx 9.5$Å, $H \approx 34$Å, $\tilde{\phi}_s \approx 0.4\pi$, $\overline{\sigma} \approx 16.8\mu C/\mathrm{cm}^2$.

Forces and Entropy. Figs. 1.9 and 1.11 compare the entropy and the osmotic stress vs. separation curves measured in columnar aggregates of Mn^{2+}-DNA [49], and the corresponding curves calculated within our model. Fig. 1.10 compare the other contributions to entropy change with the entropy of adsorbed cations.

The model reproduces experimental observations at reasonable values of the adjustable parameters: $\theta = 0.85 - 0.9$, $\Delta E_{ads} = 3 - 5\ \ k_{\mathrm{B}}T$, and $\Delta S_{ads} = 2 - 5\ \ k_{\mathrm{B}}$. Indeed, $\theta = 0.85 - 0.9$

agrees with previous estimates for Mn^{2+}-DNA based on the available experimental data. The range of ΔE_{ads} is consistent with our electrostatic estimate of $\Delta E_{ads} \sim 2 - 8\ k_{\rm B}T$ (assuming $\varepsilon \sim$ 20-to-80 inside DNA grooves [122]); ΔS_{ads} is *expected* to be of the order of several $k_{\rm B}$. For instance, the geometrically available surface areas per counterion are $A_1 \sim 10\text{Å}^2$ and $A_2 \sim 50\text{Å}^2$ in the minor and major grooves; so that just the change in the translational entropy upon counterion translocation may already contribute

$$\Delta S_{ads} \sim k_{\rm B} \ln (A_2/A_1) \sim 1.5 k_{\rm B}. \tag{1.61}$$

Unfortunately, we do not know the contribution of specific (chemical) interactions of bound Mn^{2+} with DNA bases and adjacent phosphates to ΔE_{ads} and ΔS_{ads}. Nor do we know the energetic and entropic cost of the change in solvation of DNA grooves and $Mn^{2+}-$ions upon the ion translocation from minor to major groove. These unknown factors may change the values (and even the sign) of ΔE_{ads} and ΔS_{ads}. Still, interpretation of the observed data in terms of the described model is plausible.

1.7.3 Calculated pressure curve and cations re-distribution

Within such interpretation, the net force between DNA helices results from a balance between three main components: (i) repulsion associated with the net, average charge of the molecules, (ii) attraction associated with alignment of negatively charged phosphate strands opposite to positively charged major grooves on adjacent DNA, and (iii) image-charge repulsion of phosphate strands and bound counterions from dielectric cores of adjacent DNA. The average-charge repulsion decreases as $\exp(-\kappa R)$, the attraction decays as $\exp\left(-\sqrt{\kappa^2 + 4\pi^2/H^2}R\right)$, and the image-charge repulsion as $\exp\left(-2\sqrt{\kappa^2 + 4\pi^2/H^2}R\right)$ [84]. At larger separations between DNA – prior to the transition to a less hydrated, condensed state – the dominating intermolecular force is the average-charge repulsion. The magnitude of the net force is, however, reduced by the attraction.

Elevated temperature favors higher fraction of counterions in the major groove (due to higher binding entropy). This results in stronger attraction and smaller absolute value of the net force. At closer distances, the attraction (which increases with decreasing R faster than the average-charge repulsion) becomes more important and the net force becomes less repulsive or even attractive. This leads to a jump-wise decrease in the separation under osmotic stress and the transition to the condensed state. In the condensed state, the image-charge repulsion – which has the shortest decay length and the largest amplitude of the three forces – prevents further dehydration. It is responsible for the repulsive branch of the net force observed at close distances.

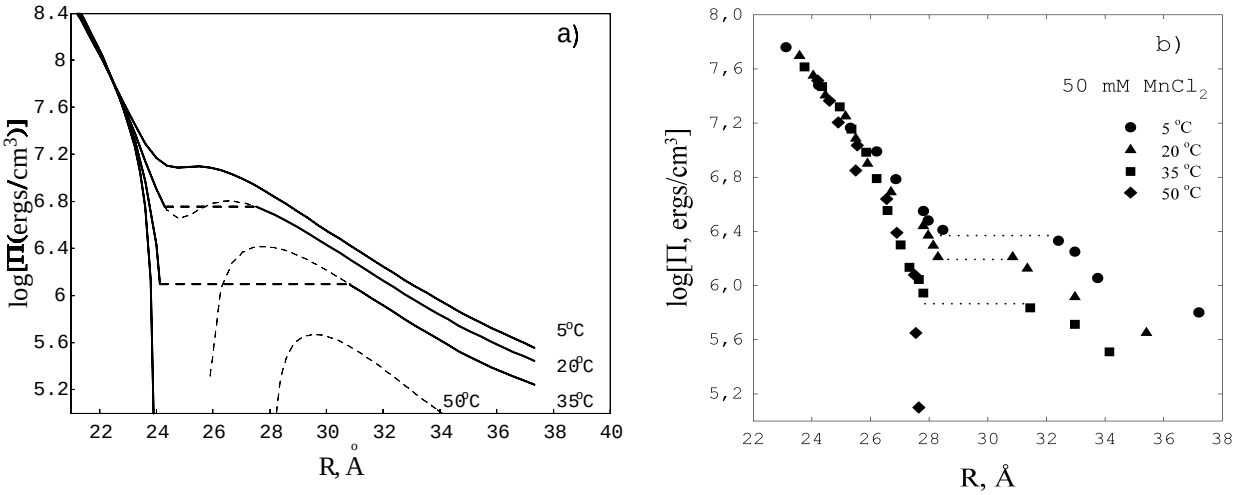

Figure 1.11: *The calculated (a) and measured (b) (data from [49]) pressure,* Π*, in DNA lattice vs. interaxial separation R at different temperatures are very similar.* Same parameters as in Fig. 1.8 were used. The solid lines show the equilibrium pressure-distance curves that exhibit jumpwise transitions from a more hydrated to a less hydrated state. The transitions are indicated by dashed lines. The pressure-distance curve branches that are metastable or unstable at constant pressure (osmotic stress) conditions are depicted by dotted lines. We have used Maxwell relation of "equal surfaces" [121] to calculate the transition pressure.
Note that since the Debye length becomes shorter with compression, the region of attraction appears for much smaller surface-to-surface DNA separation, than it was predicted in physiological solution [85].
The lattice separation of predicted repulsion-attraction transition are by 3-4Å smaller than measured [49]. This may come from 9.5Å used for DNA radius. The first shell of tightly bound on DNA water molecules [15], increases effectively the "radius" of DNA by ~ 3Å.

Counterion Partitioning and Specificity. Because counterion translocation from minor to major groove increases intermolecular attraction and, thus, reduces the free energy of DNA aggregate, compression of the aggregate by osmotic stress induces such translocation (Fig. 1.8). Most of counterion repartitioning occurs during the jump-wise transition to the condensed state. Overall, we expect 3-5 bound counterions per helical pitch of DNA (10 base pairs) to move from minor to major groove with compression. As we have pointed out, the present model associates the observed entropy increase upon DNA condensation and aggregate compression with this repartitioning.

Because it predicts strong temperature dependence of intermolecular forces only in a fairly narrow range of ΔE_{ads} and ΔS_{ads}, the present model may explain the observed counterion specificity. For instance, assuming that ΔE_{ads} and ΔS_{ads} for Mn^{2+} are within this range, it is natural to expect ΔE_{ads} and ΔS_{ads} for $Co(NH_3)_6^{3+}$ to be substantially different and, therefore, outside the range. Thus, even though the dependencies of intermolecular forces on concentrations of these ions are qualitatively similar [64], the absence of strong (if any) temperature dependence for $Co(NH_3)_6^{3+}$ is not surprising.

Note in the end that no effect of anions on DNA-DNA force has been investigated in the present model, since the anion concentration inside DNA phase is much lower than in the bulk (Donnan effect). Thus, we have neglected by the influence of anion type and properties on temperature-induced DNA condensation. It was however observed [49] that attraction in Mn-DNA assembly is enhanced in the presence of Mn perchlorate compared to cloride. This says

in favor of water release mechanism, since ClO_4^- is known to increase the entropy of bulk water (bond breaker). Thus, a water molecules coming from DNA phase upon aggregate compression, where it is either adsorbed on DNA or confined between the molecules, can win more entropy. It would be interesting to investigate the magnitude of this effect compared to predicted cations re-distribution.

Conclusions: In the present Chapter we focused on potential contribution of counterion redistribution on DNA surface to the temperature-induced DNA aggregation [49]. The corresponding model may potentially explain the existing observation at reasonable assumptions about the energetic (ΔE_{ads}) and entropic (ΔS_{ads}) cost of the redistribution. The existing data do not allow one to distinguish unambiguously this and other possible mechanisms of the temperature dependence. Still, the present work offers a strategy for further experiments. Ideally, one would measure ΔE_{ads} and ΔS_{ads} or partitioning of counterions between DNA grooves and compare the results with our predictions. In reality, direct experiments of this type are difficult to do, but the present theory may also be helpful in evaluation of indirect data. We hope that future studies will eventually resolve this puzzle.

1.7.4 Other interpretations of experimental data

The interpretation of reported observations in terms of our model is consistent with the available experimental data. However, the data are not sufficient for excluding other interpretations.

Destabilization of DNA backbone and DNA melting: In particular, it was proposed that temperature-induced aggregation of DNA in the presence of Mn^{2+} may be caused by cross-linking of melted sections of DNA mediated by the ions [123–125,138]. However, Mn^{2+}-induced DNA melting was observed only above 60°C and at fairly low pH ($\simeq 4$ [123]), i.e. when DNA is substantially less stable than under the conditions of osmotic stress force measurements. Furthermore, such interpretation is at odds with x-ray diffraction patterns observed from Mn^{2+}-DNA aggregates in the latter experiments. Even though it is difficult to exclude that a fraction of the sample could be melted, the osmotic stress/x-ray method reveals only the forces between DNA helices that exhibit normal *B*-DNA scattering pattern. We however explore some consequences of a model of (collective) DNA melting and cross-linking in Section 1.8.5. While DNA melting and base pair cross-linking via transition metal ions could be the mechanism of temperature-induced aggregation at low pH and high temperature observed in [123], it is unlikely to explain the temperature-dependent forces and aggregation observed at neutral pH and much lower temperatures in [49,66]. Within the framework of cross-linking model it is still unclear why Mn and Cd cause DNA condensation whereas Ni and Zn do not [49], in spite of the fact that all these cations are known to destabilize DNA structure, Fig. 1.18.

Still, partial *destabilization* of DNA backbone by Mn^{2+} may be important even in the latter case. Indeed, aggregation of real *non-ideally helical* DNA sequences requires torsional deformation of DNA and DNA unwinding from 10.5 bp/turn in solution [108] to 10 bp/turn in dense fibers [109] is observed. As previously suggested, this torsional deformation may be needed to make the opposition of negatively charged phosphate strands and positively charged grooves possible over entire length of the molecules [126]. Otherwise it would be disrupted by sequence-dependent variations in the helical pitch (see Chapter 2). Partial destabilization of DNA backbone by Mn^{2+} may reduce torsional rigidity and make the deformation easier. Furthermore, it may increase thermal motions of the backbone. The latter may help to adjust two similar but not completely complementary charge patterns on surfaces of two opposing DNA for a better match (see [127] and discussion in Section 1.8.6).

Hydration forces and water-release mechanism: Here we have rationalized the intermolecular forces in terms of an electrostatic model. However, it was proposed that *hydration* forces arising from modulation of water structure near DNA surface may give a substantial contribution to the net DNA-DNA interaction [59].

The interaction between almost neutral *collagen triple helices* for example is indeed governed by hydration force [90,128] associated with energetic cost of re-arrangement of hydrogen-bond-network in the water between molecules. The temperature-dependence collagen condensation [94], very similar to Mn-DNA condensation, can be described on the basis of release of structured water. The predicted 0.7Å decay length of hydration forces depends both on water correlation length ($\sim$ 4Å) and on the periodicity of surface hydration pattern. It is excellent agreement with experiment. The decay length of electrostatic force is too long to fit the data [107].

DNA vs. collagen: The surface *hydration* pattern of collagen is regular (determined primarily by helical backbone) in contrast to surface *charge* pattern, which is *irregular*. It distinguishes collagen from other biopolymers. For *DNA duplexes* both surface *charge* patterns and surface *hydration* patterns follow the helical symmetry of phosphate residues [58]. Thus, both experimentally and theoretically, it is difficult to distinguish the electrostatic and hydration contribution to DNA-DNA forces [107]. Osmotic stress experiments in very concentrated NaCl solutions have been done to exclude the electrostatic contribution to the force, but the analysis of data is somewhat ambiguous. Theoretically, both forces have very similar functional dependence on DNA charge periodicity and the results depend on a number of poorly known factors (water dielectric response, re-arrangement of hydrogen bond network, etc.) that makes the comparison of force amplitudes very difficult.

It was proposed that the entropy increase associated with release of structured water upon compression may be responsible for the observed temperature dependence of Mn-DNA forces [49]. Comparison of Mn^{2+}-DNA that does exhibit strong temperature dependence and $Co(NH_3)_6^{3+}$-DNA that does not (despite quantitatively similar forces) suggests that the temperature dependence is unlikely to be related to the release of structured water. The entropy effect of such a release should be similar in both cases. However, the release of water structured around counterions, either upon their binding or upon dehydration of DNA aggregate, may be different. So that this mechanism cannot be excluded as well.

1.8 Auxiliary results of the model and beyond the model

1.8.1 Adsorption isotherm and its effect on interaction pressure

In the calculations above we assumed that θ is independent of the density of DNA assembly. This assumption was based on the experimental observation that the change in the number of adsorbed cations upon compression of DNA aggregate is $\sim$0.01/bp [49] (and, therefore, the change in θ should also be $\sim$0.01, i.e. small). [17]

To test whether such change is consistent with the parameters used in the present model, we calculated an expected change in θ upon variation in the cell radius based on a simplified form of adsorption isotherm [28,120], which results from equalization of electrochemical potential of

[17] For $Co(NH_3)_6^{3+}$, contrary to Mn^{2+}, 0.2 cation/bp are additionally bound during the shrinking transition [64], that causes substantial change in θ.

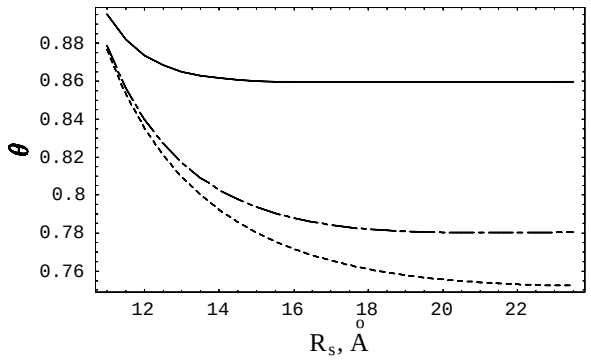

Figure 1.12: DNA charge compensation θ changes only slightly with compression of the cell, Eq. (1.62). Parameters: $n_0 = 0.05M$, $F_{ads}/(k_BT) = -3$ (dotted), -6 (dotted-dashed), -9 (solid curves). θ increases with compression since the value of surface potential increases. The larger the value of θ, the smaller its variation with compression.

cations on molecular surface and in the bulk solution

$$\frac{\theta}{1-\theta} = \frac{n_0}{n_w}\exp\left[-\frac{F_{ads}}{k_{\rm B}T}\right]\exp\left[-z\Phi\left(a,\theta\right)\right], \tag{1.62}$$

where $n_w = 55.5M$ is the water molarity, F_{ads} is the average adsorption free energy on DNA surface, $\Phi\left(a,\theta\right)$ is the surface potential of the rod in the cell model, Eq. (1.38).

The dependence of θ on R_s calculated from Eqs. (1.62), (1.38) at different values of F_{ads} is plotted in Fig. 1.12. We find that the change in θ is indeed $\sim$0.01 when $\theta\left(\infty\right)$ exceeds 0.8-0.85 which is the range of θ required for Mn^{2+}-induced condensation within our model and in experiments. The value of adsorption free energy that produces such θ ($F_{ads} \sim 10\ k_{\rm B}T$) has the same order of magnitude but is slightly larger than the change in the adsorption free energy upon Mn^{2+} transition from minor to major groove. In other words, it also appears to be consistent with model assumptions required for explaining the observed temperature dependence.

Adsorption isotherm (1.62) results in θ increase only in the region $R \lesssim 24$Å, where the image-charge repulsion between DNAs dominates. Therefore, including of adsorption isotherm in the model of Mn-DNA aggregation can change calculated pressure curves only slightly. Thus, we did not include $\theta\left(R\right)$-effect in the main part, but present $\Pi\left(\theta\right)$ −dependences in Fig. 1.13. As follows from the Kornyshev-Leikin theory, the appearance of attraction between DNAs depends on the threshold value of θ, when favorable strand–groove register overcome the repulsion of non-compensated DNA charge, Fig. 1.13. This value increases with n_0 according to (1.62), that implies the appearance of attraction at $R \sim 28$Å (cobalt hexammine [64], Mn^{2+} and Cd^{2+} [49]).

The specificity of binding of each type of cations determine the values of θ and f on DNA. These θ and f can satisfy theoretically predicted DNA condensation condition. For example, many alkali cations adsorb in the minor groove [9] (large f), that requires much higher θ for DNA condensation, Fig. 1.13c. The multivalent cations are bound usually in the major groove [16] (small f) and much stronger (larger θ), that warrants easier DNA condensation Fig. 1.13a.

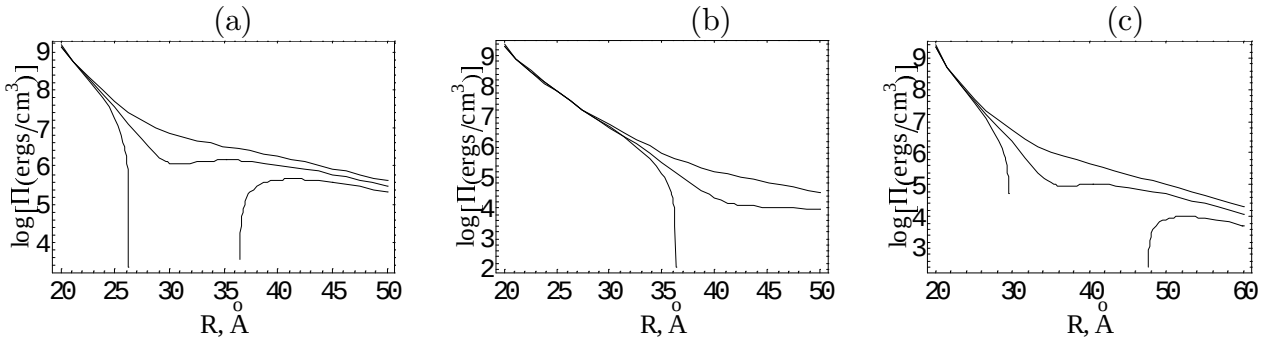

Figure 1.13: *DNA-DNA attraction is triggered by small variation of the pattern of adsorbed cations.* Parameters: $\kappa_D^{-1} = 7$Å $=$Const (R), fixed $R-$*independent* pattern of adsorbed cations. Longe-ranged DNA repulsion decreases with θ : (a) $f = 0$, $\theta = 0.5,\ 0.55,\ 0.6$; (b) $f = 0.5$, $\theta = 0.85,\ 0.9,\ 0.95$; (c) $f = 1$, $\theta = 0.75,\ 0.8,\ 0.85$. The situation with 50/50 occupation of the grooves ($f = 0.5$) is the most unfavorable for condensation since it requires θ close to unity. It is so since charge separation along the helix decreases.

1.8.2 Abrupt transitions of cations between the grooves

In Fig. 1.8 we have presented the case when the occupation of the major groove increases *smoothly* with compression of DNA assembly. It was so because the *proper* distribution of the adsorbed cations between the grooves on single DNA has been chosen. Here we present the solution of Eq. (1.50) when *minor* groove is occupied with compression, or groove occupation changes jump-wise. We propose also a phenomenological DNA surface model which may account for such behavior.

Idea. If minor groove is highly occupied on single DNA, it corresponds to a deep well of DNA surface energy at small f. The interaction energy is minimal (at all distances!) when all cations are adsorbed in the major groove. There can be a critical separation when two minima of total energy equalize, and an abrupt transition in groove occupation takes place.

Strong adsorption in the minor groove. The surface free energy, $F_s(N_2)$, has a deep minimum at small N_2. The Kornyshev-Leikin interaction energy, $F_{int}(N_2)$, has two minima at the edges. When R decreases both minima of F_{int} become deeper relative to the energy in maximum, Fig. 1.14. The total free energy per one DNA in assembly, $F = F_s + 3F_{int}$, has only one minimum for all R at small f. With compression, f decrease, Fig. 1.15a.

At strong adsorption in the major groove both F_{int} and F have a minimum at large N_2. When R decreases, this minimum becomes relatively deeper for larger N_2, Fig. 1.15b. Only this case was considered in the main text.

Intermediate case. When there is a slight preference for adsorption in minor grooves, F may have *two minima*, their depth equalizes at a certain R. When R decreases, the minimum at higher N_2 is lower, and a transition when N_2 jumps from smaller to higher value takes place, Fig. 1.15c.

A phenomenological model proposed below can result in such abrupt re-arrangement of adsorbed cations. In the spirit of Landau order-parameter-theory, we *suppose* that the surface free energy of ith DNA has the form

$$F_{si} = A\frac{\varphi_i^2}{2} + B\frac{\varphi_i^4}{4} + P\varphi_i. \tag{1.63}$$

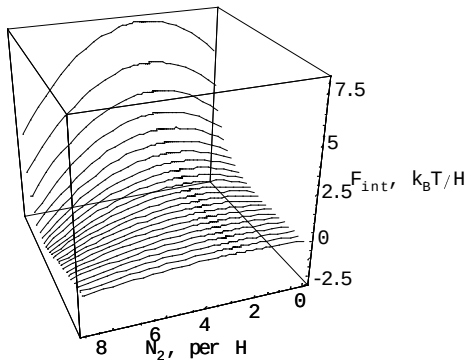

Figure 1.14: *The pair DNA-DNA interaction energy, Eq. (1.30) has two minima (at $f = 0$ and $f = 1$) which become relatively deeper with approach of the molecules* (nearest side of the cube $R = 40$Å, the most far side – $R = 20$Å). Parameters: $\Delta z = 0$, $\theta = 0.85, n_0 = 0.05M, T = 300K$.
Note that the possibility of abrupt transitions should survive also for *non-zero mutual shifts*, since the interaction energy surface still has minima at the edges: two minima at $f = 1$ and $\Delta z \neq 0$, and one minimum at $f = 0$ and $\Delta z = 0$, see Fig. 2b in [85]. The investigation of non-zero shift in the aggregate is however itself a non–trivial problem (Section 1.9).

Here $\varphi_i = f_i - \overline{f}_i$, where f_i and $\overline{f}$ $(0 \leq f_i, \overline{f} \leq 1)$, are the fractions of cations in the minor groove at finite and at infinite separation, respectively. The coefficients $B > 0$, A, P should reproduce the specificity of binding of considered cations. If $A < 0$, F_s has a two-well shape.[18]

Idea: negative P favors larger f on single DNA. With compression due to the minimum of F_{int} at $f = 0$, the optimal f may change abruptly, from larger to smaller value. It could lead to jump in F_{int} and in the force.

Model: Since coefficients α_n, Eq. (1.45), decrease quickly with n, only first two f-dependent terms in F_{int} are important for the present model. Then F reduces to Landau Hamiltonian with two bilinearly-coupled order parameters [129]

$$U = A(\varphi_1^2 + \varphi_2^2) + \frac{B}{2}(\varphi_1^4 + \varphi_2^4) + 2C\varphi_1\varphi_2 - D(\varphi_1 + \varphi_2) + G, \tag{1.64}$$

where $C = 4u_0\theta^2\alpha_1$, $D = -P - 4u_0\theta E\alpha_1$, $J = 2\theta\overline{f} - \theta - \cos(\tilde{\phi}_s)$,
$G = u_0\left\{2\alpha_2[\theta - \cos(2\tilde{\phi}_s)]^2 + 2E^2\alpha_1\right\}$. The minimization of (1.64) with respect to φ_1, φ_2, Δz at fixed $\overline{f}$ and θ yields to Euler equations

$$\left\{\begin{array}{c} (\varphi_1 - \varphi_2)[A - C + B(\varphi_1^2 + \varphi_2^2 + \varphi_1\varphi_2)] = 0,\ (\varphi_1 + \varphi_2)[A + C + B(\varphi_1^2 + \varphi_2^2 - \varphi_1\varphi_2)] = D, \\ \sin(g\Delta z)\left\{A_1[4\theta^2\varphi_1\varphi_2 + 2\theta J(\varphi_1 + \varphi_2) + \lambda^2] - 4A_2\cos(g\Delta z)[\theta - \cos(2\tilde{\phi}_s)]^2\right\} = 0, \end{array}\right. \tag{1.65}$$

where $A_n = K_0(\kappa_n R)[\kappa_n K_n'(\kappa_n b)]^{-2}$.

[18] The model gives also a hint for accounting of temperature-dependent DNA condensation if $A = p\,(n_0)\,\frac{T^\star - T}{T^\star}$, $T^\star\,(n_0)$ is the critical temperature when DNAs start to condense [44, 49]; p is the adjustable parameter.

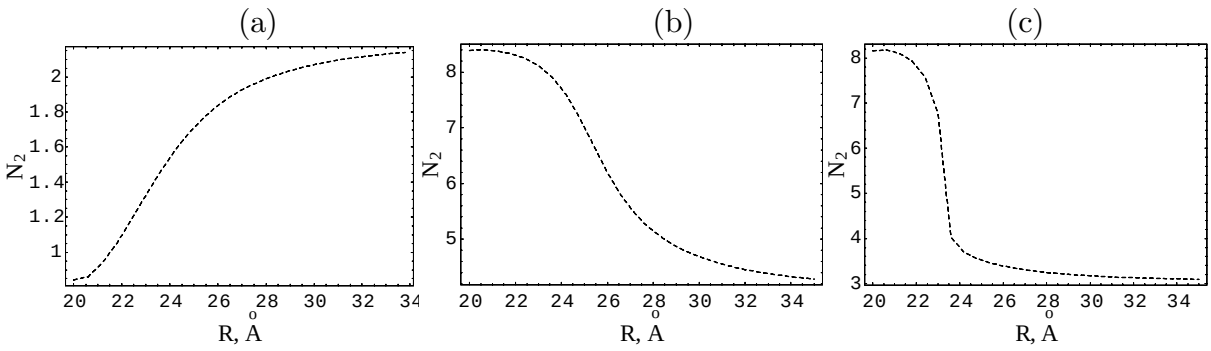

Figure 1.15: *The number of cations in the major grove per pitch, Eq. (1.50), at strong occupation of the minor (a) and major (b) groove, and in intermediate case (c).* Parameters as in Fig. 1.14 and $\eta = 0.4$, $s_2/s_2 = 9, W_1 = 0, W_2 = (a) - 3,\ (b) - 5, (c) - 4$. In this figure $N_2 = 10\theta\,(1-f)\,,\ \eta$ is the fraction of adsorption sites in the minor groove.

Further we consider only symmetric solutions so that $\varphi_1 = \varphi_2 = \varphi$. At $\Delta z \equiv 0$ Eqs. (1.65) reduce to cubic equation

$$2B\varphi^3 + (2A + 8u_0\theta^2\alpha_1)\varphi + (P + 4u_0\theta J\alpha_1) = 0. \tag{1.66}$$

Its solutions can describe the abrupt transitions from large to small f with compression of DNA lattice. For $\Delta z \not\equiv 0$ we also get a cubic equation for f.

We do not speculate about the physical meaning and value of coefficients A, B, P, since the present experimental data do not allow to say unambiguously whether this model has any physical relevance to real DNA. It is however possible that at short R the electrostatic interaction is strong enough to make such transformation of DNA backbone, that new adsorption sites appear in the (major) grooves, that could facilitate DNA-DNA attraction.

1.8.3 Behavior of DNAs of two types on the lattice

Below we account for *different* distributions of cations on DNAs and consider how the interconvertible DNAs of two types behave on hexagonal lattice with its compression. We use the Bragg-Williams approximation to calculate the number of corresponding interacting pairs. Only nearest-neighbor-interaction are considered. The Kornyshev-Leikin pair interaction energy is used as interaction energy of DNAs on the lattice [85], Eq. (1.30). Debye screening length and patterns of adsorbed cations are constant.

Idea: Since the interaction energy of DNAs with highly occupied major groove is lower than that with highly occupied minor groove, the larger amount of interaction pairs with highly occupied major groove is energetically favored. Entropically ($T \neq 0$) however 50/50 distribution of molecules over two possible types is favored. Below we show that the number of DNAs with highly-occupied major groove always *increases* with compression.

Model: Let only f_1 and $f_2 \equiv 1 - f_1$ occupations of the minor groove are possible (f_1 corresponds to highly occupied minor groove (say, 80%)). Let M_i molecules of the length L have f_i cation distribution, $i = 1, 2$. Then $f = \frac{M_1}{M_1+M_2}$ is the fraction of DNAs of the first type on the lattice. This quantity we want to find.

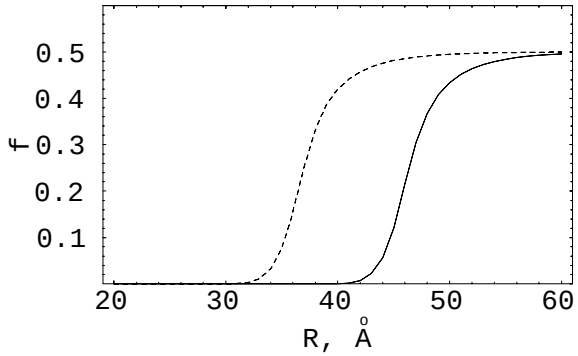

Figure 1.16: *The fraction of molecules with high occupation of the minor groove decreases with compression.* Parameters: $\theta = 0.8$, $f_1 = 0.75$, $\kappa_D^{-1} = 7$Å, $L = 10H$ (solid curve) and $L = H$ (dashed curve).

There are three possible interacting pairs: $f_1 : f_1$, $f_1 : f_2$, $f_2 : f_2$, interaction energy of each of them is, Eq. (1.30), (per length L)

$$F\left(f_i, f_j\right) = F_{ij} = u_0 L \sum_{n=1}^{\infty} 2\alpha_n \left[f_i\theta + (-1)^n \left(1 - f_i\right)\theta - \cos\left(n\widetilde{\phi}_s\right)\right] \times \left[f_j\theta + (-1)^n \left(1 - f_j\right)\theta - \cos\left(n\widetilde{\phi}_s\right)\right],$$

where $\alpha_n = \alpha_n\left(R, \kappa, a\right)$ are known functions [85], Eq. (1.30), $\theta = \text{Const}$, $z_i = 2$. If the molecules are distributed randomly on the lattice, their numbers are $\frac{zM_1^2}{2(M_1+M_2)}$, $\frac{2zM_1M_2}{2(M_1+M_2)}$, and $\frac{zM_2^2}{2(M_1+M_2)}$, respectively. Here z is the lattice coordination number, $z = 6$ for hexagonal lattice.

By construction the surface free energies of all DNAs are equal, $F_{s1} = F_{s2}$. Then the *free* energy of $M = M_1 + M_2$ interacting molecules is

$$F = M\left\{f F_{s1} + \left(1 - f\right) F_{s2} + k_{\mathrm{B}}T\left[f \ln f + \left(1 - f\right)\ln\left(1 - f\right)\right] + z\left[f^2 F_{11} + 2f\left(1 - f\right)F_{12} + \left(1 - f\right)^2 F_{22}\right]/2\right\}. \quad (1.67)$$

Results: Extremum of F with respect to M_1 at fixed M gives the equation on f (Fig. 1.16)

$$\frac{f}{1 - f} = A\left(R\right)\exp\left[f B\left(R\right)\right], \quad (1.68)$$

where $A\left(R\right) = \exp\left[z\left(F_{12} - F_{22}\right)/\left(k_{\mathrm{B}}T\right)\right]$, $B\left(R\right) = z\left(2F_{12} - F_{11} - F_{22}\right)/\left(k_{\mathrm{B}}T\right)$. DNA length gives the energetic scale where the "conversion" effect becomes to be significant. We have observed that f decreases always monotonically with compression and no phase transition occurs. Considered model suggests that if there is a limitation on the maximal number of adsorption sites in the minor groove (repulsion between adsorbed cations, etc.), the occupation of the major groove will increase with compression.

1.8.4 Ions of the finite size and Poisson-Boltzmann equation

Motivation: Below we present the approximate analytical solution for potential distributions near one and two charged surfaces with taking into account finite size of ions in solution. For highly-neutralized DNA the saturation effect considered below is of no significant importance. In general, however, it is useful to have an analytical solution, at least for planar case.

The Poisson-Boltzmann (PB) mean-field theory considers ions as point-like particles and neglects by statistical correlations between them. The first assumption leads to overestimation of ion concentration near highly-charged surface, especially in the case of multivalent ions. The calculated concentration of ions near highly-charged surface (with charge density of bare B-DNA) may be as large as 50M. Physically, however, the maximal-available concentration of ions is their close-packing concentration, $\overline{n} \sim a^{-3}$, where a is the size of an ion, possibly hydrated.

Modified Nonlinear PB (MNPB) equation. The idea of incorporation of the finite size of ions into the PB equation is very old [130], but it became known again several years ago [131]. To account for the finite size of ions, usually, the lattice-gas free energy (for one type of ions)

$$F_{lat} = k_{\mathrm{B}}T \int_V \left\{ n\left(r\right) \ln\left[\frac{n\left(r\right)}{\overline{n}}\right] + \left[\overline{n} - n\left(r\right)\right] \ln\left[\frac{\overline{n} - n\left(r\right)}{\overline{n}}\right] \right\} d^3r, \tag{1.69}$$

is used for volume free energy of ions, instead of the ideal-gas free energy,

$$F_{id} \sim k_{\mathrm{B}}T \int_V n\left(r\right) \ln n\left(r\right) d^3r. \tag{1.70}$$

It modifies PB equation for dimensionless electrostatic potential, $\Phi\left(r\right) = e\psi\left(r\right) / \left(k_{\mathrm{B}}T\right)$. In solution of $z-$valent cations *only* with bulk concentration n_0 it gives (in planar geometry)

$$\frac{d^2\Phi\left(x\right)}{dx^2} = -\kappa^2 \frac{e^{-z\Phi(x)}}{1 - \frac{n_0}{\overline{n}} + \frac{n_0}{\overline{n}} e^{-z\Phi(x)}}, \tag{1.71}$$

where $\kappa^2 = 4\pi l_B z n_0$, l_B is the Bjerrum length, $\overline{n} = a^{-3}$ is the close-packing concentration of ions on cubic lattice with edge a. For negatively charged surface(s) $\Phi\left(x\right) < 0$. At $\overline{n} \to \infty$, Eq. (1.71) turns into usual PB equation.

Single charged surface in salt solution. After first integration of Eq. (1.71) from the surface ($x = 0$) to infinity, and using the Gauß theorem, $\Phi'_s \equiv \Phi'|_{x=0} = -\frac{4\pi l_B \sigma}{e} > 0$, we get the exact expression for surface potential as a function of surface charge density σ [131],

$$\Phi_s = -z^{-1} \left[\ln\left(e^{\zeta} - (1-\alpha)\right) - \ln\alpha \right], \tag{1.72}$$

where $\zeta = 2\pi a^3 \sigma^2 / \left(\varepsilon k_{\mathrm{B}}T\right)$, $\alpha \equiv n_0/\overline{n}$.

Far from the surface, $\kappa_D x \gg 1$, the potential is small, $|\Phi| \ll 1$, we may expand exponents and arrive at Debye-Hückel (DH) equation, $\frac{d^2\Phi(x)}{dx^2} \approx -\kappa^2 \left[1 - z\left(1-\alpha\right)\Phi\left(x\right)\right]$, with exponentially decaying solution

$$\Phi\left(x\right) \equiv \Phi_2\left(x\right) = -\frac{Be^{-\kappa_D x}}{z\left(1-\alpha\right)}, \tag{1.73}$$

where $\kappa_D^2 = z\left(1-\alpha\right)\kappa^2$.

Near the highly-charged surface the potential may be so large that $\frac{n_0}{\overline{n}} e^{-z\Phi(r)} \gg 1$. Then a *saturation layer* appears near the surface, where the concentration of ions approaches $\overline{n}$. Eq. (1.71) can thus be simplified as $\frac{d^2\Phi(x)}{dx^2} \approx -\frac{\kappa^2}{\alpha}$. After two integrations, we get a parabolic

approximation for potential near the surface,

$$\Phi(x) \equiv \Phi_1(x) = \Phi_s + \Phi_s' x - \frac{\kappa^2}{2\alpha} x^2. \tag{1.74}$$

The thickness of saturation layer, l^* [131], is estimated as the point of zero derivative of potential (1.74)

$$l^* \cong \frac{\sigma a^3}{ze}. \tag{1.75}$$

It is proportional to the surface charge density and ion's volume. Below we find analytically the approximate solution of Eq. (1.71). At $\kappa_D x \gg 1$, we use DH solution (1.73), whereas near the surface, at $\kappa x \ll \kappa l^* = \frac{\sigma a^3 \kappa}{ze}$, we use solution (1.74). If these two regions overlap, matching the value of potential and its derivative in some point, x_m, we find integration constant B and parameter x_m as

$$x_m = \left(\frac{\alpha z (1-\alpha) \Phi_s'}{\kappa_D^2} - \frac{1}{\kappa_D} \right) + \sqrt{\frac{1}{\kappa_D^2} + \frac{\alpha z^2 (1-\alpha)^2 \Phi_s'^2}{\kappa_D^4} + \frac{2\alpha z (1-\alpha) \Phi_s}{\kappa_D^2}},$$

$$B = \left(\Phi_s' - \frac{\kappa_D^2 x_m}{\alpha z (1-\alpha)} \right) \frac{z(1-\alpha)}{\kappa_D} e^{\kappa_D x_m}. \tag{1.76}$$

The final distribution of potential differs *only* slightly from the numerical solution of Eq. (1.71) [131], Fig. 1.17.

Two charged surfaces. We find the potential distribution between two charged surfaces on the distance D, with counterions between them to ensure the *electroneutrality* of the system. Like for PB equation between two surfaces [29], in MNPB case we need to find the cation concentration in the mid-point between the surfaces self-consistently for each surface-to-surface separation.

Since only the potential difference is important, we set potential in the mid-point to zero, $\Phi(x = 0) = 0$. (Cation concentration in the mid-point is not zero, of course, and varies with R and σ.) Then DH region of small potential always exists near the middle point, where we write the expansion

$$\Phi(x) \equiv \Phi_2(x) = \frac{1 - \cosh(\kappa_D x)}{z(1-\alpha)}, \tag{1.77}$$

where the condition of zero potential derivative in the mid-point was used. Near the surfaces we use solution (1.74)

$$\Phi(x) \equiv \Phi_1(x) = \Phi_s + \Phi_s'(x + D/2) - \frac{\kappa^2}{2\alpha}(x + D/2)^2 \tag{1.78}$$

with only difference, that now the mid-point concentration n_0 is not fixed, but is determined self-consistently.

Then we match the potential and its derivative at some distance x_m from the surface,

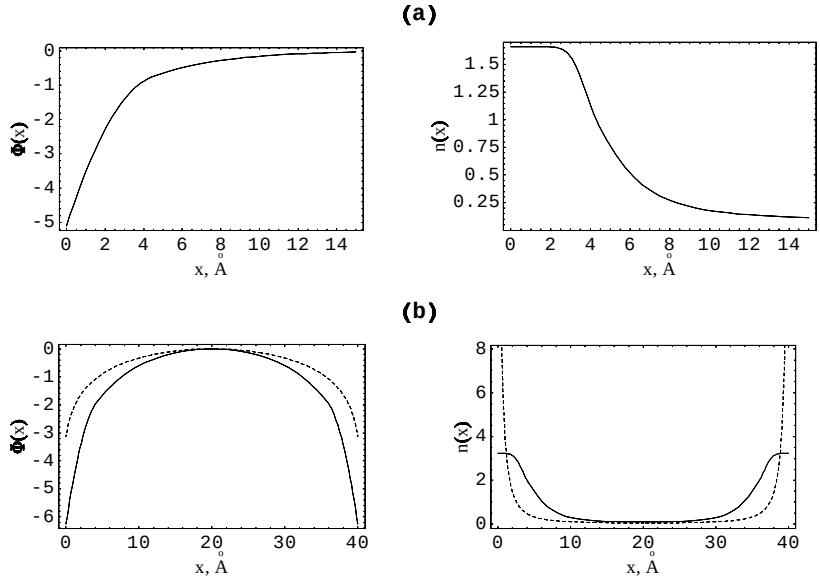

Figure 1.17: *Distribution of potential and of finite-size cations near one (a) and two (b) negatively charged surfaces.*
(a): Parameters: $a = 10\text{Å},\ z = 4,\ n_0 = 0.1M,\ \sigma = 1e/\left(50\text{Å}^2\right)$. For these parameters: $\overline{n} = 1.66M,\ l^* = 5\text{Å},\ \kappa_D^{-1} = 3.5\text{Å},\ x_m = 4.4\text{Å}$. The value of potential in sewing point is -0.76, that justifies used DH approximation. The approximate solution for potential is very close to full numenrical solution [131].
(b) Parameters: $D = 40\text{Å},\ z = 2,\ a = 8\text{Å},\ \sigma = 1e/\left(50\text{Å}^2\right)$. Then $\overline{n} = 3.2M,\ l^* = 5.1\text{Å},\ \kappa_D^{-1} = 3.4\text{Å},\ x_m = 3.5\text{Å}$, i.e. $l^* > x_m > \kappa_D^{-1}$ and there is a region of overlap of two used approximations. The value of potential in matching point is -1.4. This value can be decreased if larger ions or larger σ are considered. Dotted curves – the solution of NPB equation, $\{z\Phi_{NPB}(x) = \ln\left[\cos^2\left[Kx\right]\right]$, where the mid-point concentration is found from equation $\sigma = K^{-1}zn_0\tan\left[KD/2\right]$, where $K^2 = 2\pi l_B n_0 z^2\}$. Solid curves – is the approximate solution of MNPB equation, Eqs. (1.77), (1.78). Expression (1.72) for surface potential is used in calculations for two plates.
The cation profile reveals a plateau region near the surfaces. In addition to considered *bulk* close-packing effect, the finite-size-ions cannot come closer than $\sim a/2$ to the surface. When this effect is also considered, the cation concentration near the surface decreases substantially [132], and the maximum at the distance $\sim a/2$ from the surface can appear.

unknown in advance, and found n_0 and x_m from matching conditions,

$$\Phi_s + \Phi_s' x_m - \frac{\kappa_D^2 x_m^2}{2\alpha z\left(1-\alpha\right)} = \frac{1-\cosh\left[\kappa_D\left(-D/2+x_m\right)\right]}{z\left(1-\alpha\right)},$$
$$\Phi_s' - \frac{\kappa_D^2 x_m}{\alpha z\left(1-\alpha\right)} = \frac{\kappa_D \sinh\left[\kappa_D\left(D/2-x_m\right)\right]}{z\left(1-\alpha\right)}. \quad (1.79)$$

Unfortunately, these coupled equations can be treated only numerically. Note that second equation in (1.79) is equivalent to approximate electroneutrality condition. Indeed, we obtain the constructed approximate solution satisfies electroneutrality condition within 3-5% accuracy, that confirms used approximations.

1.8.5 Are DNA melting and aggregation coupled?

As we noted in Sec. 1.6, the proposed mechanism of entropically-driven DNA aggregation in the presence of Mn^{2+} does not exclude another interpretations because of insufficient experimental

data. Below we explore temperature-induced DNA aggregation which arises from temperature- and cation-induced DNA melting and cross-linking.

Helix-to-coil DNA transition. Since two phases cannot coexist in 1D system [121], the majority models of DNA melting result in alternating coil-helix sections along the molecule. The basic tendency is that the combinatorial entropy of melted/helix sections favors the shorter length of the sequences, the energy profit promotes the larger amount of bp in helical state and longer sequences. Most of the works on helix-coil transition of bio-molecules are based on the nearest-neighbors Ising models [133, 134]. Many factors, affecting DNA stability in solutions have also been considered theoretically (solvent composition, pH and ionic strength, specifically adsorbing cations, force-induced melting [133], DNA electrostatic charge [135], etc.).

Motivation. Melting of Mn-DNA in assembly can differ from melting of single DNA. Some earth-transition cations (for example, Mn^{2+}, Cu^{+}, Cd^{2+}, etc.) are known to form a chelation complexes with DNA (N7-(cation)-O6, N7-(cation)-phosphate, N3C-(cation)-phosphate), destabilizing the helix and at moderate concentrations decreasing DNA melting temperature T_m [14, 89, 136] (see Fig. 1.18). Such re-construction of helix may lead to appearance of *new* additional sites for cation binding on DNA (N3A, N1A, etc. [137]). The possible cross-linking via a cation bridging between these sites may "glue" DNAs together that was suggested as a mechanism of temperature-induced DNA aggregation in the presence of backbone destabilizing cations [123, 125].

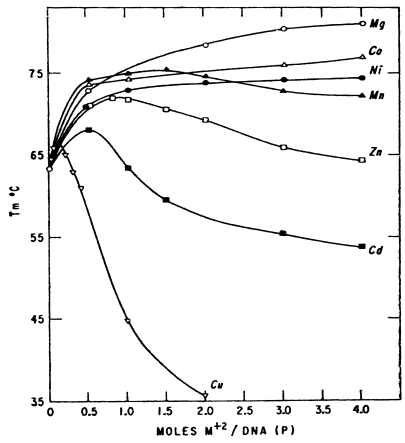

Figure 1.18: *Transition-earth cations at moderate concentrations decreases DNA melting temperature, T_m* (data from [89]). More aggressive Cu^{+} melts DNA even at room temperature.

Raman spectroscopy experiments indeed show that perturbations of structural DNA groups and DNA aggregation occur simultaneously [123, 136, 138]. In spite of the insufficient accuracy and ambiguous interpretation of Raman spectroscopy data (pH=4-5 [123]), below we present a simple model of collective DNA melting.

Mn-DNA is not a unique system, where X-linking can occur. In solutions of polyamines, for example, the measured equilibrium spacing of DNA lattice is correlated with the length of condensing agent [47], supporting X-link-driven DNA condensation.

The theory of DNA melting and its modification. We consider the melting theory of homopolymer DNA of N bp long in the solution of agent, that binds with different affinities to melted and helical bp. Let the number of bp in melted ($i = 1$) and helix ($i = 2$) sections are N_i, $i = 1, 2$, the number of bound cations are k_i, n is the number of helix-coil boundaries, F_s is

the free energy of creation of one such boundary. The free energy then takes the form [134]

$$F = nF_s + N_1F_1 + N_2F_2 + k_1\psi_1 + k_2\psi_2 \\ - k_{\rm B}T\ln\left[\frac{N_1!}{n!\,(N_1-n)!}\frac{N_2!}{n!\,(N_2-n)!}\right] - k_{\rm B}T\ln\left[\frac{N_1!}{k_1!\,(N_1-k_1)!}\frac{N_2!}{k_2!\,(N_2-k_2)!}\right]. \quad (1.80)$$

Here F_i are the free energies of melted and helix bp, ψ_i are the binding free energies of a cation to corresponding bp. Adsorption of cations is assumed to be non-cooperative and adsorbed cations do not interact. Helix-coil transition occurs cooperatively due to finite energy F_s.

Basic equations: Applying the Stirling's formula, extremum F over $n,\ N_1,\ k_{1,2}$ at $N = N_1 + N_2 = \text{Const}$ can be written as

$$\left(\frac{N_1}{n}-1\right)\left(\frac{N_2}{n}-1\right) = \sigma^{-1},\ \frac{1-n/N_2}{1-n/N_1} = s\frac{1-k_1/N_1}{1-k_2/N_2},\ \frac{k_i/N_i}{1-k_i/N_i} = n_0K_i, \quad (1.81)$$

where $\sigma = \exp\left[-\frac{F_s}{k_{\rm B}T}\right]$ is the so-called DNA cooperativity factor, $\sigma \sim 10^{-4}$ [134], $s = \exp\left[-\frac{F_2-F_1}{k_{\rm B}T}\right] \equiv \exp\left[-\frac{\Delta F}{k_{\rm B}T}\right]$, $K_i = \frac{1}{n_w}\exp\left[-\frac{\psi_i-\mu_0}{k_{\rm B}T}\right]$ are the binding constants, $i = 1, 2$. If $n_0 = 0$, we arrive at two coupled quadratic equations for the fraction of the molecule in helical state, ϑ, and the number of boundaries, n, with well-known solutions

$$\vartheta = \frac{N_2}{N} = \frac{1}{2}\left[1 + \frac{s-1}{\sqrt{(s-1)^2+4s\sigma}}\right],\ \frac{n}{N} = -\frac{\sigma}{1-\sigma}\left[1 - \frac{s+1}{\sqrt{(s-1)^2+4s\sigma}}\right]. \quad (1.82)$$

According to Eq. (1.82), the smaller σ, the sharper the transition [133]. The number of sections n increases with T and increases sharply with decrease of F_s. At small T DNA is in helix-state, $\Delta F < 0$. If $k_1/N_1 \to 1$, the r.h.s. of Eq. (1.81) tends to zero, and $N_1 \to N$. I.e., adsorption of cations on melted bp stimulates DNA melting.

It is reasonable to assume however that in order to bind to two single DNA strands and connect them together, some energy, P_s, must be paid to overcome the entropy of ss-sections. We assume that $P_s = \text{Const}$ and does not depend on the length of ss-section. The free energy of one DNA then is $F_{sur} = F + mP_s - k_{\rm B}T\ln\left[\frac{k_1!}{m!(k_1-m)!}\frac{(N_1-k_1)!}{m!(N_1-k_1-m)!}\right]$, where the new term is the mixing entropy of the number of combinations to form m regions from k_1 bound cations, and of m regions without bound cations from $N_1 - k_1$ remaining sites. It appears that the number of cations bound to ss-sections is correlated with helix-coil transition; T_m increases with P_s.

Cooperative DNA melting and DNA condensation. We assume below that the melted sections on two juxtaposed DNAs are distributed randomly, i.e. the probability to find a melted-melted contact is $p_X = (N_1/N)^2$. Then the number of sites for X-linking is $N_X = N_1^2/(N_1+N_2)$. Electrostatic interaction between DNAs is considered *below* in the model of uniformly charged rods, $F_{int} = hN\frac{8\pi^2\overline{\sigma}^2}{\varepsilon\kappa^2}\frac{K_0(\kappa R)}{(K_1(\kappa a))^2}\left(1 - \frac{z_i(k_1+k_2)}{2N}\right)^2$ (energies are measured per N bp). DNAs also interact via X-links, the corresponding free energy of two DNAs with n_X X-links is: $F_X = n_XE_X - k_{\rm B}T\ln\left[\frac{N_X!}{n_X!(N_X-n_X)!}\right]$.

The extremum of the total free energy, $F_{tot} = F_X + F_{int} + 2F_{sur}$, with respect to $n_X,\ N_1,\ n,\ k_{1,2}$ gives

$$\frac{Nn_X/N_1^2}{1-Nn_X/N_1^2} = e^{-\frac{E_X}{k_BT}}, \quad \frac{1-n/N_2}{1-n/N_1} = e^{-\frac{F_2-F_1}{k_BT}} \frac{1-k_1/N_1}{1-k_2/N_2} \left(1-\frac{Nn_X}{N_1^2}\right)^{N_1/N}, \tag{1.83a}$$

$$\left(\frac{N_1}{n}-1\right)\left(\frac{N_2}{n}-1\right) = e^{\frac{F_s}{k_BT}}, \tag{1.83b}$$

$$\frac{k_i/N_i}{1-k_i/N_i} = \frac{n_0}{n_w} \exp\left[-\frac{\psi_i-\mu_0}{k_BT}\right] \exp\left[h\frac{8\pi^2\overline{\sigma}^2}{\varepsilon\kappa^2}\frac{z_i}{2}\frac{K_0(\kappa R)}{[K_1(\kappa a)]^2}\left(1-\frac{z_i(k_1+k_2)}{N}\right)\right]. \tag{1.83c}$$

Two last equations are obtained by equalization of electrochemical potential of cation adsorbed on helical and melted sites of each DNA, to bulk electrochemical potential. $\theta = z_i(k_1+k_2)/N$, is not a constant, but changes with R according to adsorption isotherms (1.83c).

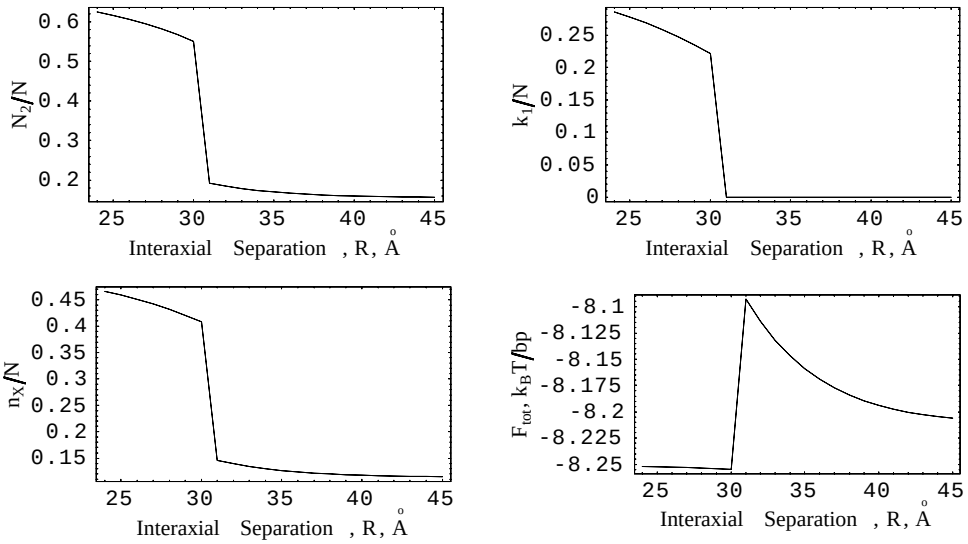

Figure 1.19: *Attraction of partially melted DNAs can be triggered by their collective melting.* The fraction of melted bp, N_2/N, of zipped bp, n_X/N, of adsorbed cations, k_1/N, increase abruptly at $R = R_x = 30$Å. The total free energy, F_{tot}, also decreases abruptly due to X-linking. Parameters: $\Delta F = -2k_BT$, $F_s = 2k_BT$, $E_X = -1k_BT$, $\psi_1 = -1k_BT$, $\psi_2 = -8k_BT$, $n_0 = 0.1M$, $\kappa_D^{-1} = 7$Å. The plausible values have been chosen for *adjustable* parameters of the model, E_X, R^*, P_s, and ψ_i.

Results: If $k_i = 0$ (no melting ions) and E_X is large, $n_X \to N_1^2/N$ and the R.H.S. of second Eq. (1.83a) tends to zero, $N_1 \to N$, i.e. almost all bp are melted and zipped together via X-links. Physically, a possibility to crosslink may appear at DNA-DNA separations smaller than some critical R_x. If it happens abruptly, an abrupt change in all quantities takes place, Fig. 1.19. When $R < R_x$ and DNAs are further pushed together, the number of adsorbed cations increases in order to compensate DNA charge, according to (1.83c). The repulsion becomes weaker compared to a fixed number of adsorbed cations.

Water release upon melting: B-DNA is highly-hydrated molecule, 10-30 water molecules are ordered per nucleotide [2,15]. From the point of view of accessible surfaces, ds-DNA should be much more hydrated than ss-DNA, and *upon melting* water should release from DNA. With compression DNAs melt and water releases progressively into the bulk, increasing the entropy of DNA assembly. It again says in favor of water release mechanism of DNA condensation with Mn. Still, we consider the *electrostatic* model of DNA aggregation with Mn, Section 1.6, as the most reasonable and straightforward one, its results are in good agreement with experimental

data.

1.8.6 Aggregation of non-ideal duplexes, the role of bp fluctuations

In Section 1.6 we have considered the temperature-induced aggregation of *ideal* DNA duplexes in the present of Mn^{2+}. Real DNA is however not an ideal staircase and such imperfect duplexes interact differently than ideal [126]. Below we discuss how the destabilization of DNA backbone allows *non-ideal non-rigid* helices recognize and attract each other.

The concept of temperature-induced complementarity has been put forward in Ref. [127], where the important role of fluctuations for biomolecular recognition was emphasized and the recognition code for hydration forces between complimentary patterns of polar groups has been suggested.

Idea: Consider two opposing surfaces with distributed polar groups. Let they attract each other if there is a mutual complementarity of "charge" distributions. Let laterally-rigid, incommensurate surfaces always repel each other (it corresponds to low temperature when charges do not fluctuate). At finite temperature fluctuations grow and charges acquire more freedom on the substrate.[19] Thermal agitation allows charges to arrange on the surfaces better, the patterns become more complimentary with a possibility of net attraction forces. At higher temperature, the fluctuations increase and can destroy the favorable arrangement of charges, interaction may again become repulsive. When surfaces come closer to each other, interaction and adjustment of charges become stronger. It necessary leads to increase of the number of configurations for surface groups and serve as a source of entropy increase.

Role of Mn: In assembly of non-ideally-helical rigid DNAs *the sequence-dependent twist* leads to disruption of strand-groove register, and long rigid DNA fragments always repel each other [126], Section 2.3. Finite *DNA torsional rigidity* allows molecules to relax sequence-dependent mismatches of the twist (see Section 2.4) and they can attract each other. Since Mn is known to destabilize DNA backbone, this cation can decrease DNA torsional rigidity and allow molecules to attract each other easier.

Role of fluctuations: At ambient temperature the fluctuations of azimuthal angle may *diminish* the incommensuration of DNA charge patterns caused by sequence-dependent twist. It may help to restore strand-groove register and facilitate DNA-DNA attraction. However torsional bp fluctuations *smooth* charge distributions on DNAs, that weakens DNA-DNA attraction, according to the Kornyshev-Leikin theory. At high temperature bp fluctuations disrupt the complementarity of charge patterns, diminishing the attraction. [20] With compression DNAs interact stronger, that must increase the entropy of the backbone adjustment.

Note also that simultaneous consideration of DNA sequence-dependent twist and bp fluctuations is also important for calculation of the optimal number of bp/turn on imperfect interacting DNAs. This result should correct the 10bp/turn favored by electrostatics on ideal helices [107].

[19] Note that the correlation between associated cations may also strengthen the attraction with temperature: for instance, configuration $\left\{\begin{smallmatrix}+&-&+&-&+&-\\-&+&-&+&-&+\end{smallmatrix}\right\}$ of cations and polycation charges, minimizing the electrostatic energy at $T=0$, produces weaker attraction than configuration $\left\{\begin{smallmatrix}+&+&+&-&-&-\\-&-&-&+&+&+\end{smallmatrix}\right\}$ at $T\neq 0$ [139].

[20] The additional length-scale, the fluctuation coherence length, $\lambda_f(T)$, appears in this model. If $\lambda_f \gg \lambda_c$, the fluctuations are too small to help system to reach a commensurate charge patterns. If $\lambda_f \sim \lambda_c$, i.e. the thermal bp fluctuations are of the order of random twist fluctuations, the fluctuation can improve DNA non-ideality and facilitate attraction. If $\lambda_f \ll \lambda_c$, the fluctuations are too large and disrupt strand-groove register.

1.9 Azimuthally frustrated DNA lattices

Motivation: The main topic of Chapter 1 is to describe the temperature-induced DNA condensation in the presence of Mn^{2+}. We have considered however only the simplest alignment of DNAs on hexagonal lattice, without any axial shifts between the molecules, $\Delta z \equiv 0$, see p. 30. It is realistic for large DNA-DNA separations or high occupation of the major groove by adsorbed cations [84] (occupation of the major groove in our model is indeed strong, Fig. 1.8).

Pair DNA-DNA interaction energy, Eq. (1.30), results in the *nonzero optimal shift* (nonzero azimuthal orientation angle) between two DNAs at close separations [84], Fig. 1.20b,c. On the lattice, where each molecule takes part in many interactions, the optimization of all angles is constrained by packing symmetry and may involve many-body effects. It is not clear how the azimuthal part of interaction energy is optimized on a given lattice and we address this point below.

We consider electrostatic interaction between DNAs in columnar hexagonal lattice on the basis of the Kornyshev-Leikin theory [85]. We consider as before only nearest-neighbor interaction and consider them to be pair-wise additive. We calculate the optimal azimuthal orientational angles which minimize the orientation of DNAs in elementary triangles on the lattice. The lattice constructed from such triangles is frustrated.

1.9.1 Experimental observations and puzzles

Experiment: It has recently been shown that positional order within the hexatic DNA phase under the external osmotic stress is more *liquid-like* the more DNA density is increased. The measured correlation length goes from about five neighbors at 24.0Å DNA interaxial spacing to about eight neighbors at R =27.5 Å, Fig. 1.20a. It is surprising and counter-intuitive, since we would expect that DNA array becomes more positionally-organized when we approach the crystalline phase. It was suggested [140] that progressive disordering of DNA packing at higher densities may be due to increasing frustration of the molecules as they try to satisfy both the positional and the angular constraints imposed by electrostatic interaction potential, Eq. (1.30), [85].

1.9.2 Optimal azimuthal angles on hexagonal lattice

Let pair DNA-DNA interaction energy (per unit length) depends on mutual azimuthal orientation angle of two DNA, $\phi \equiv \delta\phi = \Phi_1 - \Phi_2$, as

$$E(\phi) = -a_1(R)\cos[\phi] + a_2(R)\cos[2\phi], \quad a_1, a_2 > 0, \tag{1.84}$$

where $a_1(R)$ and $a_2(R)$ are known functions of DNA-DNA separation R, the parameters of solution and of DNA intrinsic parameters. For electrostatic interaction these coefficients have been calculated in [85] (see Section 1.4 and Appendix A). [21] We calculate only the *ground-state configuration* of the molecules: no influence of temperature either on interaction energy or packing symmetry is considered. As was shown in [84,85], energy (1.84) is minimal at $\phi_* = 0$ at $a_1 > 4a_2$ ($\gamma < 1/2$, $\gamma \equiv 2a_2/a_1$) and at a nonzero angle, $\phi_* = \pm\arccos[a_1/(4a_2)]$, at $\gamma > 1/2$.

[21] Below however we do not restrict ourselves to only electrostatic interaction and first calculated the properties of packing of the molecules with general form pair interaction energy, Eq. (1.84). We do so since the patterns of water adsorbed on DNA follow DNA helical symmetry [15] and the hydration forces [59] *should* also depend on DNA mutual azimuthal orientation, as it was indeed predicted [58].

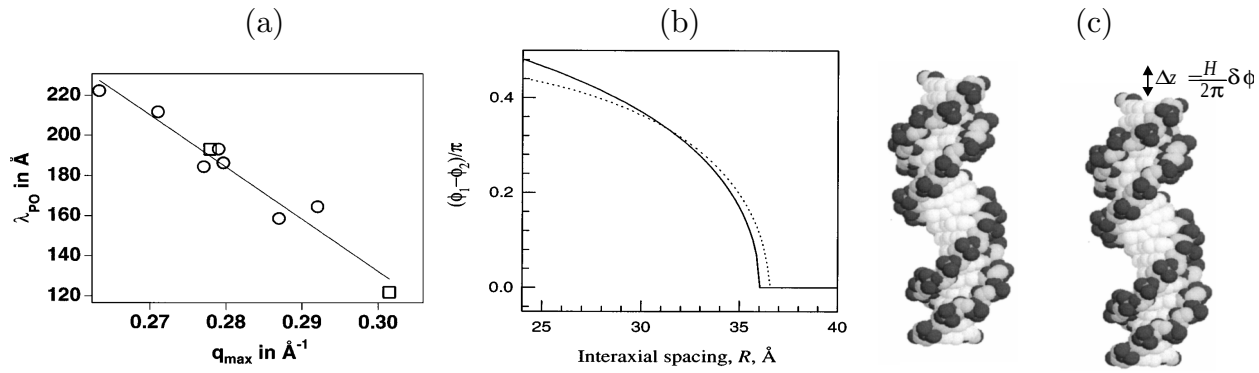

Figure 1.20: *The measured positional correlation length (a) decreases when DNA lattice is compressed. The optimal azimuthal angle between two DNA is nonzero at close DNA-DNA separation,* as follows from Kornyshev-Leikin theory (b,c). Data from [84] and [140]. The interaxial DNA-DNA separation is recalculated from inverse lattice spacing as $R = 4\pi/\left(\sqrt{3}q_{\max}\right)$.
The recent theoretical predictions of frustrated DNA assemblies, based both on the phenomenolocal theory [141] and computer simulations of DNA lattice with exact DNA-DNA pair interaction potential [142], suggest a variety of possible azimuthal "spin" structures. However, in first work no exact DNA-DNA interaction potentail was used, whereas in the second work the predicted spin structures do not garantee the minimum of the system.

General consideration. The interaction energy of νth DNA molecule of the length L on hexagonal lattice is

$$E\left(\phi_\nu,\phi_\mu\right)=\frac{1}{2L}\int\limits_0^L dz\sum_{\mu=1}^{6}\left\{-a_1(R)\cos\left(\phi_\nu-\phi_\mu\right)+a_{_2}(R)\cos\left[2\left(\phi_\nu-\phi_\mu\right)\right]\right\}, \tag{1.85}$$

where index μ labels 6 nearest-neighbors. From the symmetry of the ground state, $\phi_{\nu,\mu+2}=\phi_{\nu,\mu}$, and (1.85) can be written in terms of new variables

$$x=\phi_1-\left(\phi_2+\phi_3\right)/2,\ y=\phi_2-\phi_3 \tag{1.86}$$

as the interaction energy density of three molecules in elementary equilateral triangle,

$$E_\Delta\left(x,y\right)=-a_1\cos\left[y\right]+a_2\cos\left[2y\right]-2a_1\cos\left[x\right]\cos\left[y/2\right]+2a_2\cos\left[y\right]\cos\left[2x\right]. \tag{1.87}$$

Energy minimum: The extremum of this energy, $\partial E_\Delta/\partial x=\partial E_\Delta/\partial y=0$, leads to multiple solutions for optimal values of variables x and y. These solutions correspond to energy maxima and minima, we choose the solutions of the energy minima. The angle differences between the molecules in such "optimal triangle" are constructed from optimal x and y as

$$\Delta_1=\phi_2-\phi_1=y/2-x,\ \Delta_2=\phi_3-\phi_2=-y,\ \Delta_3=\Delta_1+\Delta_2. \tag{1.88}$$

Analysis of energy extremum shows that energy E_Δ is minimal at (Fig. 1.21)

$$\Delta=\Delta_1=\pm\arccos\left[\frac{1}{4}+\frac{1}{4}\sqrt{1+\frac{2a_1}{a_2}}\right]. \tag{1.89}$$

We found also that in such triangle two differences of azimuthal angles between the molecules

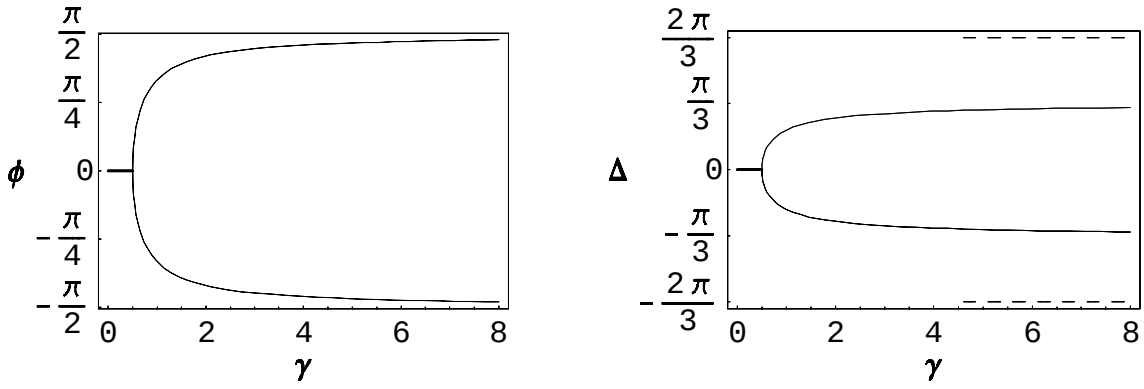

Figure 1.21: *The optimal azimuthal angle between two DNAs (left) and angle differences in elementary triangles:* Δ_1*– solid,* Δ_2 *– dashed curves.* For $\gamma \gg 1$, $\phi_* \to \pi/2$, $\Delta_1 \to \pm\pi/3$, $2\Delta_1 \to \Delta_2$, and the energies of two optimal triangles equalize.

are equal to Δ_1 and the third difference is two times larger. The energy of such triangle, Eq. (1.87), is

$$E_{\Delta_1} = -\frac{a_2 a_1^2 - 2a_1 a_2^2 + 10 a_2^3 + 2\left[2a_1/a_2 + 1\right]^{3/2}}{8a_2^2}. \tag{1.90}$$

When $a_2 \gtrsim 2.3a_1$, the second minimal triangle appears, with

$$\Delta_2 = \pm 2\pi/3. \tag{1.91}$$

This triangle also has $\Delta_2, \Delta_2, 2\Delta_2$ angle differences, but its energy is higher,

$$E_{\Delta_2} = 3\left(a_1 - a_2\right)/2. \tag{1.92}$$

Electrostatic interaction of B-DNAs: The assumption of hexagonal DNA lattice is realistic for $R\sim$ 25÷34Å. For B-DNA-B-DNA *electrostatic* interaction the expression for coefficients $a_1(R)$, $a_2(R)$ are [85], see Eq. (1.30)

$$a_1(R) = \frac{16\pi^2\overline{\sigma}^2}{\varepsilon\kappa_1^2}\frac{\left[2f\theta - \theta - \cos\left[0.4\ \pi\right]\right]^2 K_0(\kappa_1 R)}{\left[K_1'(\kappa_1 a)\right]},\ a_2(R) = \frac{16\pi^2\overline{\sigma}^2}{\varepsilon\kappa_2^2}\frac{\left[\theta - \cos\left[0.8\ \pi\right]\right]^2 K_0(\kappa_2 R)}{\left[K_2'(\kappa_2 a)\right]}, \tag{1.93}$$

where $\kappa_n = \sqrt{\kappa^2 + n^2\frac{4\pi^2}{H^2}}$; a is the DNA radius, θ is the DNA charge compensation fraction, f is the fraction of cations in the minor groove, κ_D^{-1} is the Debye screening length of the solution.

At separations larger than the separation of spontaneous symmetry breakdown, $R > R_*$,all angle differences between the molecules in elementary triangle are zero (R_* is found from equation $4a_2(R_*)/a_1(R_*) = 1$). At $R < R_*$ these differences are nonzero, Eqs. (1.89), (1.91), and all DNAs in triangle are rotated (shifted) with respect to each other, Fig. 1.20c. We found that for strong compensation of DNA charge (large θ) and strong occupation of major groove (small f) the only optimal triangle with Δ_1 is realized, Fig. 1.22. For relatively small θ and large f the second optimal triangle can appear, Fig. 1.23.

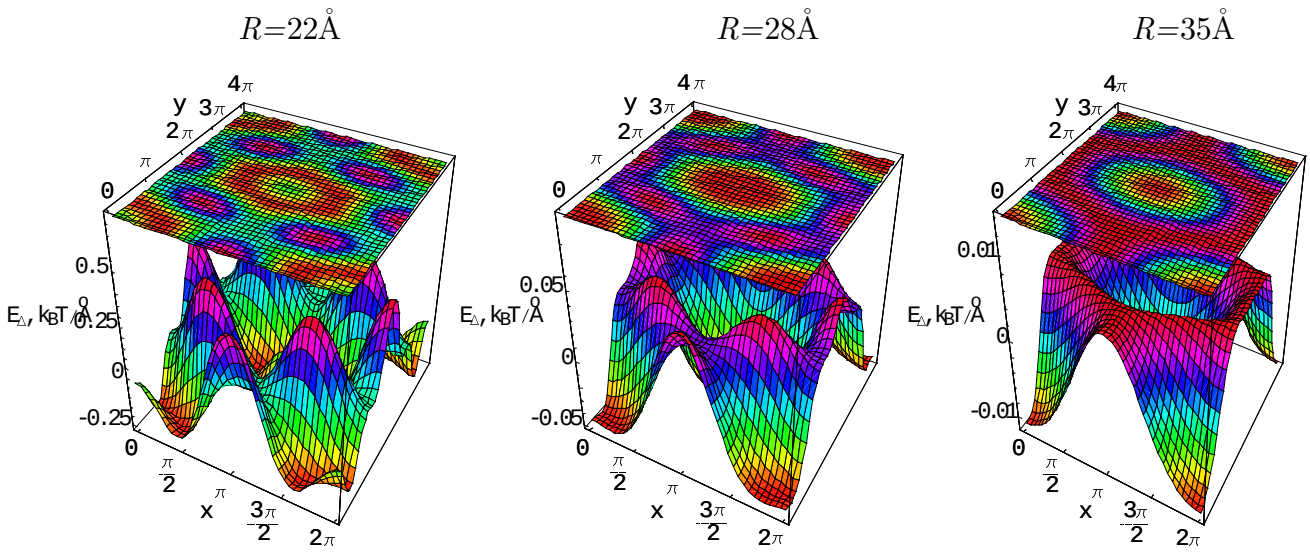

Figure 1.22: *At close separations the energy of elementary triangle, $E_\Delta(x,y)$, is minimal at nonzero azimuthal angles between the molecules (*at $R \lesssim 30$Å*).* Parameters: $1/\kappa_D = 7$Å, $f = 0.3$, $\theta = 0.8$. At large R the mutual azimuthal angle between all DNAs is zero ($R = 35$Å). Predicted nonzero azimuthal angles, *which are consistent with hexagonal lattice,* may be the physical reason why the DNA lattices at high densities are frustrated [144]. Due to $R-\phi-$coupling in the interaction potentail, the increase of azimuthal frustrations can affect (disrupt) the positional order on the lattice [140].
We find that even at $R = 22$Å the permutation minima of the interaction energy of triangle are separated by small energy barrier, $\sim 0.01 k_B T/$Å ($\sim 5 k_B T$ for 500Å long DNAs). Frustrations of the mutual angle within the valley from one minimum to another costs little energy. This weakens biaxial correlations at ambient temperature and lead to suppression of chiral interactions and, probably, to a cholesteric-hexagonal transition [113].

Beyond nearest-neighbor interactions: The lattice constructed from optimal triangles has minimal energy; noninteracting triangles are distributed randomly on the lattice. Let now the molecules on the distance $R\sqrt{3}$ interact and let it may affect the mutual orientation of triangles on the lattice, but cannot change the optimal angles in triangle, $\Delta_{1,2}$. It is reasonable since electrostatic interaction decreases nearly exponentially with separation, Eq. (1.93).

Only zero and $3\Delta_1$ angle differences are possible between the molecules on the distance $R\sqrt{3}$. The interaction energy $E(3\Delta_1) = \left[(a_1{+}4a_2)^2 \sqrt{a_2\,(a_2{+}2a_1)} + a_1\,(a_2{-}a_1)\,(8a_2 - a_1)\right] / [16a_2^2]$ is always lower than $E(0) = a_2 - a_1$. Thus, at $T = 0$ the triangles are arranged so that *all* angle differences between DNAs on $R\sqrt{3}$-distance are zero (such lattice can be constructed). With temperature increase the lattice of minimal triangles *start to disorder* and zero differences between diagonal molecules will appear.

Finite temperature: phase diagram of DNA assembly. The calculation of the full phase diagram of *dense* DNA assembly at finite temperature is a complicated problem. For hexagonal lattice it is similar 2D XY models in the theory of magnetism and Ising model of *order-disorder transitions*, but the $R-\phi$-coupling in interaction energy makes the problem even more non-trivial. At high densities however the interaction energy is much larger than the thermal energy, the low-temperature expansion can be used [143]. The full analysis however is not yet done.

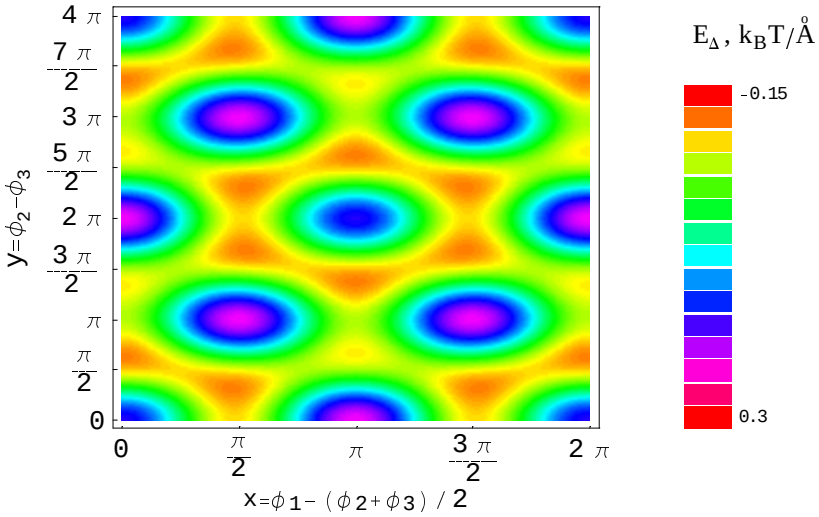

Figure 1.23: *At relatively small θ and high occupation of the minor groove the optimal triangle with angle difference Δ_2 appears.* Six permutation minima corresponding to the first optimal triangle, Δ_1, are depicted as red spots, whereas those for second triangle are seen as yellow spots. Parameters: $R = 22$Å, $f = 1$, $\theta = 0.55$, $\kappa_D^{-1} = 7$Å. Each color in the column on the right corresponds to $0.01 k_B T/$Å rise of E_Δ.

1.9.3 Observed lattice structures: Hexagonal *vs.* orthorhombic

Experimental facts: Analysis of X-ray-diffraction pictures has shown that natural DNA in fibers in the presence of alkali cations can adopt A, B, and C configurations [144–146]. The particular DNA-form is largely determined by the type of cations in solution and relative humidity around the fibre. These parameters also determine the symmetry of DNA lattices (*hexagonal, orthorhombic, etc.*). On each lattice DNA molecules are forces to pack with definite vertical shifts with respect to each other.

For example, Li-DNAs at 66% relative humidity are in the B-form (10bp/turn) and packed into *orthorhombic* crystalline lattice with $\pm\frac{H}{6}$ or $\pm\frac{H}{8}$ shifts between nearest molecules [144] (DNA pitch is $H = 33.7$Å) [22], see Fig. 1.24a. Li-, Na-, K-, and Rb- B-DNAs at 92% humidity adopt *hexagonal* phase with $\left(0, +\frac{H}{6}, -\frac{H}{6}\right)$ shifts between the molecules in elementary triangle [144]. It is consistent with $(\Delta, \Delta, 2\Delta)$ angle differences, predicted above from the Kornyshev-Leikin theory, Eqs. (1.89), (1.91).

At 44% humidities Li-DNA adopt C-form ($9\frac{1}{3}$bp/turn!), Fig. 1.24, but in no case A-form, which is however observed in fibers of Na-, Rb-, and K-DNAs at 75% humidity. [23] Li-DNA adopt hexagonal or orthorhombic lattice with nonzero axial shifts, in contrast to Na- and K-DNAs, which are packed in *monoclinic* lattice without any axial shifts [144, 146].

Unfortunately, all experiments have been done for very dense DNA lattices, where the edge-groove interlocking may take place. Still, some puzzles remain. For example, best-fitting

[22] The molecules are distributed over two types of sheets (1,2), parallel to DNA axis. Molecules m_1 are displaced along DNA axis by 11.0Å relative to molecules m_2. The sheet of contiguous molecules is displaced randomly either by $H/3$ or -$H/3$ along DNA axes relative to the neighboring sheet.

[23] It suggests that in the osmotic stress experiments, which are usually performed at 10mM of NaCl, B-to-A transition in dense DNA assemblies can take place. It rearranges water structure around DNAs and can give rise to entropy increase with compression.

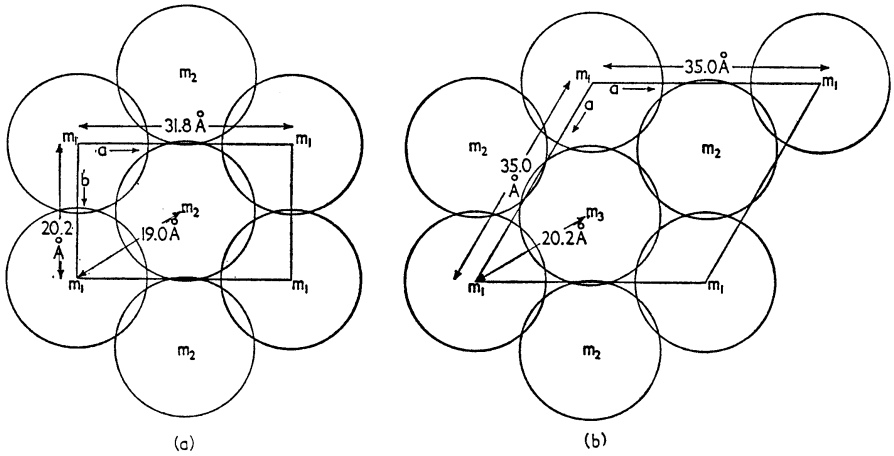

Figure 1.24: *Arrangement of C-DNAs in the unit cell: (a) orthorhombic packing, (b) hexagonal packing* (data from [146]). *B*-DNAs adopt orthorhombic lattice similar to (a), but with larger lattice constants, 31.2Å, 22.7Å, and 19.3Å instead of 31.8Å, 20.2Å, and 19.0Å, respectively, for C-DNA.
(a) vertical shifts $\pm\frac{H}{3}$ and $\pm\frac{2H}{3}$ were observed between nearest molecules m_1 and m_2 [145]. Nearest DNAs can also be translated randomly along the axis by $\pm\frac{H}{2}$.

geometrical arrangement has never been observed for *C*-DNA. It was suggested that it is not favored electrostatically [145]. Below we explore the consequences of DNA-DNA electrostatic interaction, Eq. (1.30), on DNA packing symmetry on non-equidistant lattice. [24]

Idea: As was put forward by Parsegian *et al.* [140], "the orthorhombic phase solves the angular frustration problem by distorting the hexagonal equilateral into the isosceles triangles." On the basis of this idea the first *phenomenological* model of distorted hexagonal lattice has recently been suggested [141].

Model: We base our analysis on Eq. (1.84) for *exact DNA-DNA pair interaction energy.* We calculate the energy E_Δ of elementary isosceles triangle with angle $\alpha \neq \pi/3$ near one vortex. We want to find the optimal value of α, if the density of the lattice is held constant, and the optimal mutual shift between the molecules.

Extremum of E_Δ however does not result in an analytical expression for the optimal x, y. Analysis of the energy surface shows that E_Δ loses the permutation minima, specific for hexagonal lattice, and has a shallow minimum for x, y, which satisfy the condition $y \approx \pi - 2x$. The optimal angle differences between DNAs in the triangle are

$$\Delta_1 \approx \pi/2 - 2x,\ \Delta_2 \approx 2x - \pi,\ \Delta_3 \approx -\pi/2. \tag{1.94}$$

I.e., two closest molecules are shifted by $\frac{\pi}{2}\frac{H}{2\pi} = \frac{H}{4}$, whereas the angles between two other molecules, which are far apart, can take any value subject to condition $y + 2x \approx \pi$. The value of E_Δ in the minimum decrease when α decreases. Experiments indeed show [144] that in orthorhombic lattice some molecules interlock into each other, whereas the other molecules are several Å apart. Thus, electrostatic interaction stimulates packing in non-equidistant lattice, where the *nearest* molecules adopt the most favorable mutual orientation.

[24] This interaction energy may be inaccurate for such tightly packed lattices because (i) the linear PB equation does not work (potential exceeds k_BT/e_0), (ii) the dielectric constant of water is not more macroscopic [147] and non-local screening may take place [148]. We hope however that it may strongly affect the absolute value of interaction energy, but not its *azimuthal* dependence.

Chapter 2

Interaction and recognition of non-ideal DNA duplexes

The basic molecular mechanisms that govern the search of DNA homology and subsequent homologous pairing during synapsis are not completely understood. The complementarity of single DNA strands and subsequent homologous recombination mediated by RecA-family-proteins are commonly considered as the basic mechanism of homologous recognition in pro- and eukaryotic cells. Recent experiments however suggest that there can be also RecA-independent mechanisms of DNA-DNA recognition, which involve *direct* DNA-DNA interaction. In this chapter we explore one such mechanism: the recognition of homologous DNA sequences mediated by their *electrostatic* interaction.

2.1 DNA-DNA recognition: the role and description

2.1.1 Homologous recombination and the length of homology

The essential problem in genetic recombination is how homologous DNA sequences recognize each other in gene shuffling reaction or DNA repair. These reactions represent the fundamental metabolic activity of life. *In vivo*, homologous recombination proceeds in many steps and involves many different molecules. It is often a response to single- or double-stranded damage of DNA duplex in bacteria and fungi [149, 150]. In some yeast and mammalian cells the chromosomal DNA double-strand breaks are also efficiently repaired by homologous and illegitimate recombination, respectively [151]. Without these mechanisms a re-arrangement of genetic information or loss of chromosomal segment would have been unavoidable. *In vitro*, the specificity of homologous sequences to recognize each other on a distance is used for gene manipulation [152] and gene targeting [153, 154]. Thus, the understanding of the mechanisms of DNA pairing is extremely important.

The dependence of recombination frequency on the length of DNA homology has been measured for many DNAs in different conditions. For example, minimum 30 and 50 base-pairs (bp) long DNA homology were shown to be required for recombination between the cII cistrons of bacteriophage T4 [155] and between plasmid and phage of *Escherichia coli* [156, 157]. It was also shown that in both cases the recombination frequency reveals two different regions in dependence on the sequence length L, for short and long sequences. For long sequences in all experiments the recombination frequency increases linearly with the sequence length, for short sequences it increases $\sim L^n$ with $n > 1$. Similar tendencies but with longer required homology

length (150-200 bp) were detected for homologous recombination of mammalian cells [158]. It was shown [156] that mismatches in DNA homology substantially decrease the recombination frequency (in *E.coli* for instance reduction of homology from 100% to 90% decreases the frequency over 40-fold [157]). It suggests that DNA-DNA recognition depends on the complementarity of DNA sequences over the whole length of homology.

The complementarity of single DNA strands has been considered as a mechanism of recognition of nucleic acids during synapsis [159]. Random-walk theoretical models of branch migration during homologous recombination have been developed for description the above-described dependences of recombination frequency [160, 161]. Although these models result in reasonable linear dependence of recombination frequency for long sequences, and L^3-dependence for short sequences [160], they do not answer the question how the homologous sequences recognize each other *before* the onset of recombination process.

2.1.2 Electrostatic mechanism of homology recognition

Recent investigations on pairing of homologous chromosomes of budding yeast [162] suggest that *recognition may proceed via pairing of intact double-stranded DNA fragments*. Since reported recognition is independent of specific recombination proteins (RecA), it was suggested that "identification of homology at the chromosomal level during somatic pairing is determined by processes other than those involved in searching for homology at the DNA level during recombination repair". "In budding yeast, just prior to entering the meiotic program, homologs are paired via multiple interstitial interactions between chemically intact chromosomes" [163]. It was argued that pairing contact should be unstable and dynamic, that may be initiated by *direct* interaction between homologous DNA fragments [162].

Recently it was shown that DNA-DNA *electrostatic* interaction crucially depends on DNA surface charge patterns [85] (Section 1.4). From the theory follows that two juxtaposed *ideal* helices can attract each other under a favorable mutual azimuthal alignment. This effect gave rise to a concept of "an electrostatic zipper motif for DNA aggregation" [85]. Modification of this theory to include the sequence-dependent twist between adjacent DNA monomers [126] has shown that interaction between two uncorrelated DNA sequences differs dramatically from the interaction between two homologous sequences. It was argued that this kind of interaction may be responsible for a snap-shot *electrostatic recognition of homologous DNA sequences* on a distance [126].

Basic Idea: Qualitatively, this result may be explained without a complicated algebra. Indeed, DNA is not an ideal staircase. Step angles are slightly distorted for each step, and the pattern of these distortions correlates with the text of the sequence [164, 165].

Two *homologous* duplexes in parallel juxtaposition will have almost identical patterns of distortions of the steps, and they can be aligned in such a way, that the motifs of positive and negative charges will stay in register along the whole length of the sequence. This strengthens attraction between the duplexes and allows homologous sequences to recognize each other on a distance. They can come into a closer juxtaposition from the solution that is necessary for subsequent recombination process.

On the contrary, two *non-homologous* sequences have random relative to each other texts and the related patterns of distortions. Their quasi-helical charge distributions can be positioned in a register over a certain length, but they will inevitably lose register for longer sequences. The attraction between them will therefore be much weaker or even turn into repulsion [126]. The characteristic length over which two juxtaposed duplexes with uncorrelated

texts completely lose register was found to be equal to [126]

$$\lambda_c = h/\Delta\Omega^2, \tag{2.1}$$

where h is the vertical rise between DNA base pairs and $\Delta\Omega$ is the root mean square variation of the twist angle for each sequence; λ_c was called the helical coherence length. For B-DNA $h \approx 3.4\,\text{Å}$, $\Delta\Omega \approx 0.07 - 0.1\,\text{rad}$ [166–169], and thus $\lambda_c \approx 300 - 700\text{Å}$. On the length-scales larger than λ_c the mismatch accumulates according to the law of random walk.

The difference between the interaction energy of *rigid* non-homologous and homologous duplexes of the same length is positive and is called recognition energy. It is typically several $k_{\rm B}T$ for sequences $L \sim$100bp long at interaxial separations of $R = 30\text{Å}$ [85]. It was also shown that the interaction energy increases nearly exponentially with approach of the duplexes and the recognition energy grows. The latter is pronounced for long sequences, and is larger the longer the sequence. This kind of electrostatic, 'snap-shot' recognition mechanism may explain the puzzle of homologous recombination: the observed growth of the frequency of recombination events with the length of DNA homology [155, 156, 158]. The electrostatic recognition of sequences as a whole could be a source of selective screening in a primary search in which DNA need not unzip [170].

2.1.3 Recognition energy and DNA torsional softness

These conclusions were derived under assumption that the juxtaposed duplexes are torsionally rigid. However DNAs have a finite *torsional* persistent length, λ_t [126]. In the model this length depend on DNA torsional rigidity modulus, DNA-DNA interaxial separation, and the patterns of adsorbed charges, $\lambda_t \sim 200 - 700\text{Å}$ [126]. For sequences longer than the torsional length, $L > \lambda_t$ the effects of torsional elasticity may not be ignored, although if $\lambda_t \gg \lambda_c \sim L$, the account of torsional elasticity could give only a small correction to the recognition energy. Since typically $\lambda_t \sim \lambda_c \sim L$, the problem needs special analysis.

Torsional elasticity will somehow allow DNA sequences to relax the accumulating mismatch, at the cost of the energy of torsional deformation. This will diminish the corresponding recognition energy. The effect will be the stronger the torsionally softer the molecules. But how will it be realized and how much will it change the recognition energy dependence on the sequence length is most interesting to know. The basic equation for approaching this problem was suggested in Ref. [126], but it was solved there only in the limit of $(\lambda_t/\lambda_c) \to \infty$. In this Chapter we treat this problem also for very soft sequences, $(\lambda_t/\lambda_c) \ll 1$, and in intermediate case, $(\lambda_t/\lambda_c) \sim 1$.

There is a number of experimental indications on the importance of DNA torsional deformations. It is well established that DNA structure changes in dense aggregates subject to external conditions. The impetus for the observed DNA overwinding from 10.4−10.6 bp per DNA helical pitch in solutions [108, 171] to nearly 10 bp/pitch in hydrated fibers [109] as well as B-to-A DNA transition in dense aggregates at low humidity [144, 146], was recently explained by the gain in corresponding electrostatic interaction energy [107, 112]. We are unaware of experimentally detected laws of how torsional unwinding-overwinding relaxes the energy, accumulated due to intrinsic distortions mismatch. However, the correlations between B-DNA twisting and base-pair morphology were repeatedly discussed [166].

The theory predicts that sequence-dependent DNA twist prevents attraction between rigid DNA duplexes with uncorrelated texts [126]. It is known, however, that *in vitro* some ions cause aggregation of even random DNAs [37]. Some of these cations, in particular cobalt hex-

ammine, are known to destabilize DNA backbone [45], that may help DNA to condense. It was established that identical sequences aggregate without significant backbone deformations [168], contrary to random fragments. This may be considered as another indication that torsional deformations are important for DNA-DNA recognition. In addition, the value of DNA torsional rigidity [172, 173] may strongly depend on external conditions (specifically adsorbing cations, properties of solution, temperature, etc.). Thus, if the effect of finite torsional rigidity on the recognition energy were strong, one could expect a nontrivial influence of external conditions on the identification of homology.

Below we explore the effect of finite DNA torsional rigidity on electrostatic interaction of helically imperfect DNA duplexes [174]. In Section 2.2 we derive the basic equations in the simplest approximation for DNA-DNA interaction energy. In Sections 2.3 and 2.4 we find approximate solutions of these equations for rigid and soft helices, and calculate the energy of the system. In Section 2.5 we use the parabolic approximation to estimate the energy of interacting duplexes with B-DNA parameters. In Section 2.6 we discuss the consequences of the model when a more accurate expression for interaction energy is used. In Section 2.8 we derive kink-like solutions of Euler equation for ideal helices and discuss their importance.

2.2 The free energy functional

The interaction free energy of ideal DNA duplexes of the length L at interaxial separation R has been calculated in Refs. [84, 85]. It can be written in the form

$$E_{\text{int}} \approx L\left[a_0\left(R\right) - a_1\left(R\right)\cos\left[\delta\phi\right] + a_2\left(R\right)\cos\left[2\delta\phi\right]\right], \tag{2.2}$$

with the coefficients $a_i\left(R\right)$ decaying nearly exponentially with R (when $R \gg \lambda_D$) (see Appendix A). They depend crucially on DNA surface charge pattern. Since coefficients a_i quickly decrease with index i, the first two $\delta\phi$-dependent terms usually reproduce basic dependence of interaction energy on mutual azimuthal orientation of the molecules [84] (see footnote in Appendix A).

Here mutual orientational angle is the difference of azimuthal orientations of the middle of the minor groove of two molecules, Fig. 2.1,

$$\delta\phi = \Phi_1 - \Phi_2. \tag{2.3}$$

For ideal helices this angle does not change along the molecules in juxtaposition, i.e. along z-axis.

However DNA is not an ideal helix, and $\delta\phi$ changes along axis z. Within one molecule each combination of adjacent base pairs has a preferred twist angle $\Omega = \langle\Omega\rangle \pm \Delta\Omega$, $\langle\Omega\rangle = 34^\circ - 35^\circ$ and $\Delta\Omega = 4^\circ - 6^\circ$ [167–169]. The distribution of the preferred twist angle, $\Omega\left(z\right)$, is a sequence "fingerprint", Fig. 2.2.

Using expression (2.2) for a local density of interaction energy, which is valid for long ($L \gg H$) molecules with small twit variation ($|\Delta\Omega| \ll \langle\Omega\rangle$), we may write

$$E_{\text{int}}\left(L\right) \approx \int_0^L dz\left\{a_0\left(R\right) - a_1\left(R\right)\cos\left[\delta\phi\left(z\right)\right] + a_2\left(R\right)\cos\left[2\delta\phi\left(z\right)\right]\right\}. \tag{2.4}$$

For each of the molecules in juxtaposition the actual twist angle may differ from $\Omega\left(z\right)$ because intermolecular interaction (and thermal fluctuations) may cause torsional deformation

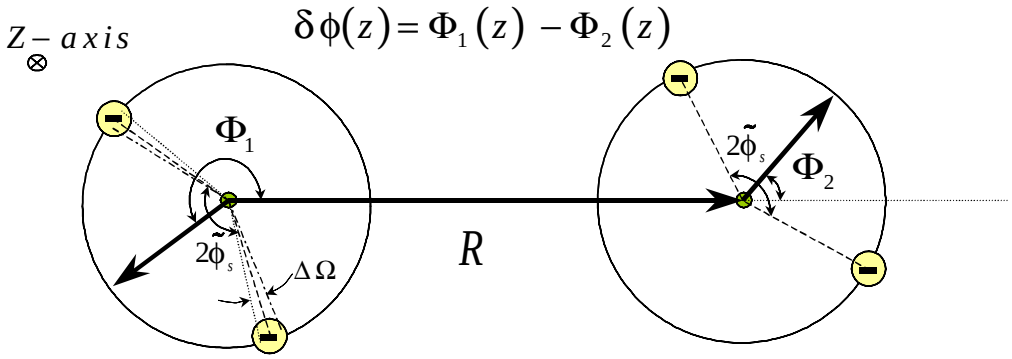

Figure 2.1: *A plane perpendicular to two non-ideal B-DNA helices on the distance R.* Cross sections of the phosphate strands are depicted as bright circles. We describe the orientation of bp on each molecule at axial position z by the azimuthal angle $\Phi_{1,2}(z)$ of the middle of the minor groove. Each combination of adjacent bp has a preferred twist angle $\Omega_{1,2}(z) = \langle\Omega\rangle \pm \Delta\Omega(z)$, where $\Delta\Omega = 4^\circ - 6^\circ$ [167].

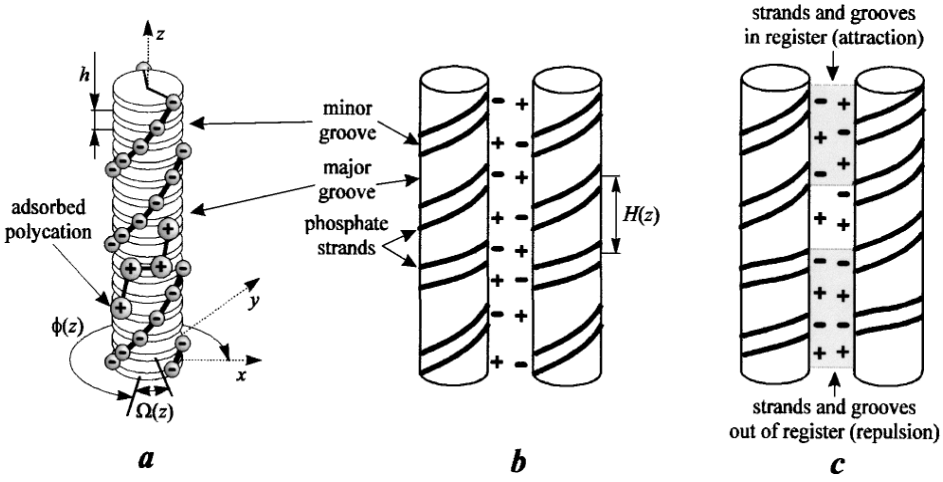

Figure 2.2: *Homologous sequences are in register (b), whereas for non-homologous sequences the sequence-dependent twist destroyes the register (c)* (data from [126]). (a) *B*-DNA schematically drawn as a stack of base pairs. Each base pair has two negatively charged phosphate groups. Sequence-dependent twist variation $\Omega(z)$ leads to local variation of DNA helical pitch $H(z)$.

of DNA. To find this angle we write the torsional energy for each molecule in the form

$$E_{t,i} = \frac{C}{2}\int_0^L dz \left(\frac{d\Phi_i(z)}{dz} - \Omega_i(z)\right)^2, \tag{2.5}$$

where C is DNA torsional rigidity modulus, and index $i = 1, 2$ labels the molecules. In calculations below we will use the value of $C = 3 \times 10^{-19}$ erg cm [172]. [1]

[1]The latter of course depends on DNA bp-sequence, the presence of cations and environmental conditions. However at ambient physiological conditions 10^{-19} erg cm order of magnitude is certain. For instance, the values of [172] $C \approx 3.4 \times 10^{-19}$ erg cm and [175] $C \approx 2.4 \times 10^{-19}$ erg cm have been reported, on the basis of experiments on cyclization kinetics. The Monte-Carlo simulations of ring closure probabilities coupled with statistical mechanics of DNA chains give a value very close to the former estimate [172]. Measurements of DNA supercoiling free energy [176] treated by means of the statistical mechanics of supercoils give $C \approx 2.9 \times$

The sum of interaction and torsional free energies gives the total free energy as a functional of $\delta\phi(z)$. The minimum of the functional provides the most favorable *local* azimuthal alignment of torsionally flexible DNAs. In the main part of this Chapter we consider a simplified form of interaction energy, in which the a_2- term is neglected in Eq. (2.4). This is appropriate for not too close separations between the duplexes and for high occupation of major grooves by adsorbed cations [85]. The results for $a_2 \neq 0$ are more sophisticated and briefly discussed in Section 2.7.

Then the total energy of two DNA duplexes in juxtaposition reads (see Appendix A)

$$E(L) \simeq h \int_0^{L/h} \left[a_0 - a_1 \cos\left[\delta\phi(Z)\right] + \frac{C}{4h^2} \left(\frac{d\delta\phi(Z)}{dZ} - \delta\Omega(Z) \right)^2 \right] dZ, \tag{2.6}$$

where $Z = z/h$ is the corresponding lengths scaled to the vertical rise;

$$\delta\Omega(Z) = \Omega_1(Z) - \Omega_2(Z). \tag{2.7}$$

For homologous sequences $\delta\Omega(Z) \equiv 0$, whereas for non-homologous, random sequences $\delta\Omega(Z)$ can be approximated by a delta-correlated random Gaussian field with zero average, $\langle \delta\Omega(Z) \rangle = 0$,

$$\langle \delta\Omega(Z)\, \delta\Omega(Z') \rangle = 2\Delta\Omega^2 \delta(Z - Z'), \tag{2.8}$$

where $\Delta\Omega$ is the root-mean-square variation of the twist on each sequence. Minimizing energy functional (2.6), we obtain the Euler equation on the running optimal mutual azimuthal angle $\delta\phi(Z)$,

$$\frac{d^2\delta\phi}{dZ^2} - \kappa^2 \sin\left[\delta\phi\right] = \frac{d\delta\Omega}{dZ}. \tag{2.9}$$

Here $\kappa = h/\lambda_t$, where

$$\lambda_t = \sqrt{\frac{C}{2a_1(R)}} \tag{2.10}$$

is the characteristic torsional length. It decreases with approach of the duplexes, Fig. 2.6, and with their softening.

Eq. (2.9) is the Sine-Gordon equation in a random field. Many non-trivial solutions of this equation have been investigated (Brownian motion in periodic potential [177], kink-dynamics in random media [178–181], soliton pinning by impurities [182–185], Frenkel-Kontorova model [186] and non-linear oscillations [187], etc.). For parameters of real DNA in standard conditions solution of Eq. (2.9) may not involve a coupling of disorder and non-linearity and can be investigated in simple harmonic approximation, Sections 2.4 and 2.5.

Below we solve Eq. (2.9) in this regime, insert the solution into Eq. (2.6), average the energy over the realizations of $\delta\Omega(z)$, and thus get the mean-field system energy as a function of model parameters.

10^{-19} erg cm [173], if DNA is modeled as an elastic isotropic rod.

2.3 Rigid DNA duplexes

Consider the case of absolutely rigid chains: $\kappa = 0$. It was studied in Ref. [126], and we essentially reproduce its results for the simplified form of the interaction energy, Eq. (2.6). We also extend them for duplexes of finite length.

Since $\kappa = 0$, the second term in the r.h.s. of Eq. (2.9) is absent. Then, $\delta\phi(Z)$ is not affected by the interaction between duplexes, but is predetermined by the intrinsic pattern of twist angle $\delta\Omega(Z)$ (the one they would have had, if they did not interact with each other):

$$\frac{d\delta\phi(Z)}{dZ} = \delta\Omega(Z). \tag{2.11}$$

(the constant of integration is of no importance for the calculation and is put to zero). The solution of Eq. (2.9), satisfying the boundary condition

$$\delta\phi(Z)|_{Z=0} = 0 \tag{2.12}$$

is

$$\delta\phi(Z) = \int_0^Z \delta\Omega(Z')dZ'. \tag{2.13}$$

Using the Gaussian-statistics we find

$$\langle\cos[\delta\phi(z)]\rangle = e^{-\frac{1}{2}\langle\delta\phi^2(z)\rangle} = e^{-z/\lambda_c}, \tag{2.14}$$

i.e. the angular correlations decrease exponentially along the helices; the decay length is equal to the coherence length. The larger the deviation of the twist angle from its mean value, the faster the decay of correlations along the molecules. For long sequences $\delta\phi(Z)$ obeys the law of random walk: its mean square displacement increases $\propto L$, whereas short sequences ($L \lesssim \lambda_c$) remain correlated due to condition (2.12).

The average energy reads

$$\langle E(L)\rangle = a_0 L - a_1\lambda_c\left(1 - e^{-L/\lambda_c}\right). \tag{2.15}$$

Since $a_{0,1} > 0$, $\langle E(L)\rangle$ *increases* with L linearly for long sequences ($L \gg \lambda_c$), Fig. 2.3. First correction in $\kappa \ll 1$ reveals the same tendency (Appendix B). For short sequences ($L \lesssim \lambda_c$), $\langle E(L)\rangle$ decreases since $\delta\phi(Z)$ is situated near the minimum of interaction energy, Eq. (2.14). Thus, the helical non-ideality of rigid DNAs disrupts strand-groove register [85] and long rigid DNA fragments with uncorrelated texts always *repel* each other [126].

The recognition energy, ΔE, is the difference between the interaction energy of uncorrelated sequences and the interaction energy of identical sequences

$$\langle E(L)\rangle = (a_0 - a_1)L. \tag{2.16}$$

It has the form

$$\Delta E(L) = a_1\lambda_c\left[L/\lambda_c + e^{-L/\lambda_c} - 1\right]. \tag{2.17}$$

This energy increases with L *linearly* for long sequences ($L \gg \lambda_c$),

$$\Delta E \approx a_1 L, \tag{2.18}$$

but it vanishes quadratically for short sequences ($L \ll \lambda_c$),

$$\Delta E \approx a_1 L^2 / (2\lambda_c) . \tag{2.19}$$

Note that the used approximation is strictly valid for DNA duplexes of finite rigidity it they are much shorter than λ_t.

Azimuthally free sequences: The case considered above corresponds to the situation when the two duplexes are fixed on one end and have second end free. But how do two 'azimuthally free' DNA duplexes align? Let them be free to decide at which point Z_* they choose $\delta\phi(Z_*) = 0$. Integrating Eq. (2.11), we get $\delta\phi(Z) = \int_{Z_*}^{Z} \delta\Omega(Z')\, dZ'$. The energy minimization over Z_* results in $Z_* = L/(2h)$, i.e. correlations persist from the center of the chains to their ends. The average energy,

$$\langle E(L) \rangle = a_0 L - 2a_1\lambda_c \left(1 - e^{-L/(2\lambda_c)}\right), \tag{2.20}$$

is thus lower than energy (2.15), Fig. 2.4. The recognition energy,

$$\Delta E = a_1\lambda_c \left[L/\lambda_c + 2e^{-L/(2\lambda_c)} - 2\right], \tag{2.21}$$

coincides with Eq. (2.17) in the limit of long L, as it should be, but for short duplexes it increases with L^2 two times slower, $\Delta E \approx a_1 L^2/(4\lambda_c)$.

2.4 Soft DNA duplexes

2.4.1 Infinite sequences

Below we consider the case when the interaction energy of two DNAs is much larger than their torsional energy. It is the case of small C and large a_1, i.e. short λ_t. This case, opposite to the one considered in the previous section, will be called the limit of *soft* fragments. In this limit, the duplexes relax practically after each torsional mismatch, keeping $\delta\phi$ close to zero.

Thus, we may replace $\sin[\delta\phi(Z)]$ by $\delta\phi(Z)$ in Eq. (2.9) (harmonic approximation), which results in linear differential equation

$$\frac{d^2\delta\phi}{dZ^2} - \kappa^2\delta\phi = \frac{d\delta\Omega}{dZ}. \tag{2.22}$$

The solution of this equation for infinite chains reads

$$\delta\phi(Z) = -\frac{1}{2\kappa}\int_{-\infty}^{+\infty} dZ' e^{-\kappa|Z-Z'|}\frac{d\delta\Omega(Z')}{dZ'} = \frac{e^{-\kappa Z}}{2}\int_{-\infty}^{Z} dZ' e^{\kappa Z'}\delta\Omega(Z') - \frac{e^{\kappa Z}}{2}\int_{Z}^{\infty} dZ' e^{-\kappa Z'}\delta\Omega(Z') . \tag{2.23}$$

It is only valid if

$$\left\langle \delta\phi^2(Z) \right\rangle = \frac{\lambda_t}{2\lambda_c} \ll 1, \tag{2.24}$$

i.e., for soft chains (short λ_t) or for small $\Delta\Omega$ (long λ_c). Note that the average in Eq. (2.24)

does not depend on the coordinate along the chains, since we have neglected by boundary conditions in Eq. (2.23). Compared to rigid fragments, there will be a long-range order in the mutual alignment of the duplexes: correlations persist along the whole length of the fragments.

Since $\left\langle \left(d\delta\phi/dZ - \delta\Omega\right)^2 \right\rangle = \kappa^2 \left\langle \delta\phi^2 \left(Z\right) \right\rangle$, the average energy for solution (2.23) is calculated from Eq. (2.6) as

$$\left\langle E\left(L\right)\right\rangle = a_0 L - a_1 L \left[e^{-\lambda_t/\left(4\lambda_c\right)} - \lambda_t/\left(4\lambda_c\right)\right]. \tag{2.25}$$

However, in order to stay within the accuracy of linear approximation, we should expand the exponential, i.e.

$$\left\langle E\left(L\right)\right\rangle = a_0 L - a_1 L \left[1 - \lambda_t/\left(2\lambda_c\right)\right]. \tag{2.26}$$

Thus, the total energy of soft chains is proportional to the sequence length L (dotted-dashed line in Fig. 2.3b), because each azimuthal mismatch is corrected by local torsional deformation of the backbone. The recognition energy for soft duplexes is proportional to L for all L

$$\Delta E = a_1 L \frac{\lambda_t}{2\lambda_c}. \tag{2.27}$$

As follows from Eq. (2.25), one half of ΔE comes from the energy of torsional adjustment, and another half – is the difference of interaction energy of homologous ($\delta\phi = 0$) and uncorrelated chains. Note, that solution (2.23) must not satisfy condition (2.12).

2.4.2 Finite sequences

The general solution of Eq. (2.22) is

$$\delta\phi\left(Z\right) = A_1 e^{-\kappa Z} + A_2 e^{\kappa Z} + \\ \frac{e^{-\kappa Z}}{2}\int\limits_{Z_1}^{Z} \delta\Omega\left(\xi\right) e^{\kappa\xi} d\xi - \frac{e^{\kappa Z}}{2}\int\limits_{Z}^{Z_2} \delta\Omega\left(\xi\right) e^{-\kappa\xi} d\xi + \frac{\delta\Omega\left(Z_1\right) e^{\kappa Z_1} e^{-\kappa Z} - \delta\Omega\left(Z_2\right) e^{-\kappa Z_2} e^{\kappa Z}}{2\kappa}, \tag{2.28}$$

where the limits of integration $Z_{1,2}$ and the constants $A_{1,2}$ are arbitrary. Their values should be obtained from the boundary conditions. Without loss of generality we put $Z_1 = 0$, $Z_2 = L$ and

$$\delta\Omega\left(0\right) = \delta\Omega\left(L\right) = 0. \tag{2.29}$$

Since we expect that solution (2.28) is valid also only for soft sequences, we again use the harmonic approximation expanding $\left\langle\cos\left[\delta\phi\left(Z\right)\right]\right\rangle \approx 1 - \left\langle\delta\phi\left(Z\right)^2\right\rangle/2$ in interaction energy. It simplifies the expression for average energy, the minimization of which with respect to $A_{1,2}$ gives $A_1 = A_2 = 0$. Thus, the average

$$\left\langle \delta\phi^2\left(z\right)\right\rangle = \frac{\lambda_t}{2\lambda_c}\left[1 - \frac{e^{-2z/\lambda_t} + e^{-2(L-z)/\lambda_t}}{2}\right] \tag{2.30}$$

decays with the length $\lambda_t/2$ to the value calculated for infinite soft chains, Eq. (2.24). The third term in the calculated average energy,

$$\left\langle E\left(L\right)\right\rangle = a_0 L - a_1 L\left(1 - \frac{\lambda_t}{2\lambda_c}\right) - \frac{a_1\lambda_t^2}{4\lambda_c}\left(1 - \exp\left[-\frac{2L}{\lambda_t}\right]\right), \tag{2.31}$$

is the correction to Eq. (2.26), which takes into account the finite length of the duplexes (see Fig. 2.3). This term is negative, since soft finite chains are more correlated on the ends, Eq. (2.30), due to condition (2.29).

The calculated recognition energy (Fig. 2.5),

$$\Delta E = a_1 L \frac{\lambda_t}{2\lambda_c} - \frac{a_1 \lambda_t^2}{4\lambda_c} \left(1 - \exp\left[-2L/\lambda_t\right]\right), \tag{2.32}$$

increases linearly for long ($L \gg \lambda_t$) sequences,

$$\Delta E \approx a_1 L \frac{\lambda_t}{2\lambda_c}. \tag{2.33}$$

It increases $\lambda_t/(2\lambda_c)$-times slower with L than the recognition energy of long rigid chains, Eq. (2.18). ΔE increases quadratically for short ($L \ll \lambda_t$) sequences, similarly to the recognition energy of rigid chains, Eq. (2.17), $\Delta E \approx a_1 L^2/(2\lambda_c)$. This is understandable, since both approximations consider the sequences much shorter than DNA torsional length, λ_t.

2.5 Approximate solution in parabolic approximation

Below we find approximate solution of Eq. (2.9) which allows us to calculate the energy of much more rigid chains and of the chains with larger twist variation, than in the limit of soft chains. It appears than the interaction of random B-DNA sequences and the adjustment of their twist patterns can be considered by this method.

The main difficulty in solving Eq. (2.9) is in its nonlinearity. However, a fairly good approximation to it could be obtained if the $-\cos[\delta\phi]$ in Eq. (2.6) is replaced by two parabolas with a minimum at $\delta\phi = 0$ and a maximum at $\delta\phi = \pi$. The parabolas, which intercept $\delta\phi$-axis at $-\pi/2$, $\pi/2$ and $\pi/2$, $3\pi/2$, are given by the functions $U_1(\delta\phi) = -a_1\left(1 - \frac{8}{\pi^2}\frac{\delta\phi^2}{2}\right)$ and $U_2(\delta\phi) = a_1\left(1 - \frac{8}{\pi^2}\frac{(\delta\phi-\pi)^2}{2}\right)$, in these intervals, respectively. [2] We model the potential $-a_1\cos[\delta\phi]$ on whole axis by 2π -periodic translations of these parabolas.

Then, in the interval of $-\pi/2 \le \delta\phi \le 3\pi/2$, the corresponding Euler equations are

$$\frac{d^2\delta\phi_1}{dZ^2} - \kappa_0^2\delta\phi_1 = \frac{d\delta\Omega}{dZ}, \quad -\pi/2 \le \delta\phi_1 \le \pi/2, \tag{2.34}$$

$$\frac{d^2\delta\phi_2}{dZ^2} - \kappa_0^2\delta\phi_2 = \frac{d\delta\Omega}{dZ}, \quad \pi/2 \le \delta\phi_2 \le 3\pi/2, \tag{2.35}$$

where $\kappa_0 = \kappa\sqrt{8}/\pi$. The solution for $\delta\phi_1(Z)$ is constructed similarly to solution (2.28). In considered region the solution of Eqs. (2.34), (2.35) must be matched in the point where $\delta\phi_1 = \pi/2$. The matching procedure is, however, ambiguous in view of the random functions under the integrals. We therefore use an *approximate* condition on the mean squared angle, assuming that Eq. (2.34) is valid when

$$\sqrt{\left\langle \delta\phi_1^2(Z)\right\rangle} \le \pi/2. \tag{2.36}$$

[2] Such approximation makes soft chains softer and gives lower estimation for their energy.

Comparing this condition with condition (2.24) makes it clear that the range of validity of $\delta\phi_1$ is much larger than of soft chains solution, Eq. (2.23). It makes $\delta\phi_1$ valid for more rigid chains and for larger $\Delta\Omega$. It appears that for the typical values of parameters of B-DNA this approximate solution is quite representative for any length of the sequence.

Thus, we get that

$$\left\langle \delta\phi^2\left(z\right)\right\rangle = \frac{\lambda_t \pi}{\lambda_c 8\sqrt{2}} \left\{ 2 - \exp\left[-\frac{\sqrt{8}}{\pi}\frac{2z}{\lambda_t} \right] - \exp\left[-\frac{\sqrt{8}}{\pi}\frac{2L}{\lambda_t} \right] \exp\left[\frac{\sqrt{8}}{\pi}\frac{2z}{\lambda_t} \right] \right\} \tag{2.37}$$

and the average energy (Fig. 2.3)

$$\left\langle E\left(L\right)\right\rangle = a_0 L - a_1 L \left(1 - \frac{\lambda_t}{2\lambda_c}\frac{\sqrt{8}}{\pi} \right) - \frac{a_1 \lambda_t^2}{4\lambda_c} \left(1 - \exp\left[-\frac{\sqrt{8}}{\pi}\frac{2L}{\lambda_t} \right] \right). \tag{2.38}$$

According to condition (2.36), the average angle remains within the first parabola when

$$\lambda_t / \lambda_c \leq \sqrt{2}\pi. \tag{2.39}$$

This determines the region of parameters where $\delta\phi_1$ is valid and no $\delta\phi_2$ solution exists. This constraint is much weaker than the inequality (2.24), that is the basic advantage of parabolic approximation.

Criterion (2.39) however fails (i) for large $\Delta\Omega$ (small λ_c: mismatch accumulated very quickly), (ii) for very rigid chains (at large C: chains are not soft enough to correct the angle mismatches), (iii) when DNAs are far apart form each other (the case of small a_1 and thus large λ_t, when electrostatic forces are not strong enough to keep $\delta\phi$ near zero). Eventually, for B-DNAs at $R = 30$Å interaxial separation, $\lambda_t = 200 \div 800$Å (at physiological conditions), $\lambda_c = 200 \div 600$Å [126], and that warrants the validity of Eq. (2.38).

2.6 How to affect the recognition energy?

Energy comparing: rigid and soft duplexes. The averaged interaction energies for all the discussed approximations are displayed in Fig. 2.3. The basic conclusions are: (i) the energy of long absolutely rigid duplexes *increases* with the sequence length L and they repel each other; (ii) the energy of soft duplexes *decreases* $\sim L$, and the duplexes can attract each other always acquiring the optimal azimuthal configuration. The finite torsional rigidity thus allows DNAs to correct the accumulating mismatch and turn repulsion into attraction. The attraction-repulsion behavior of unrelated B-DNA sequences depends on their groove occupation, Fig. (2.3).

Possible mechanism of homologous recombination: Homologous recombination is commonly believed to be initiated by breaks in DNA duplex. Specialized RecA proteins are thought to coat the 3' single-stranded tails produced from such breaks. This complex promotes the association of an intact partner double helix and catalyzes the exchange with homologous sequence [162, 188]. This is the basic mechanism of homologous recombination in *E.coli* [156]. Corresponding RecA homologs can also be involved in homologous recombination of eukaryotic cells [188]. However, RecA-independent homologous recombination in pairing of chromosomes in yeast [162] suggests that direct DNA-DNA interaction are important for alignment and exchange of homologous DNA sequences as well.

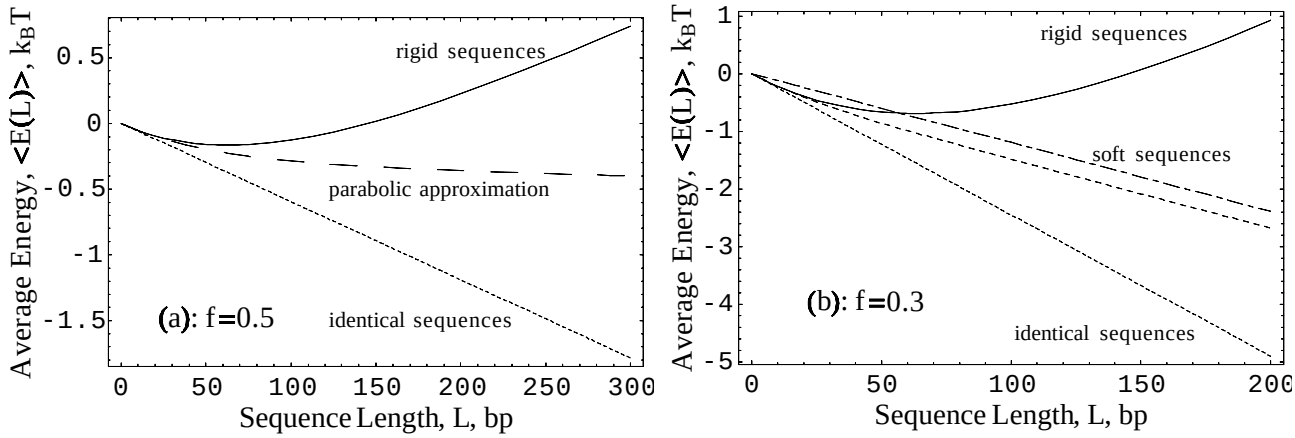

Figure 2.3: *Torsional softness decreases the energy of DNA sequences with uncorrelated texts.* Negative energy values indicate favorable juxtaposition compared to infinite separation between the duplexes. Parameters: $R = 30$Å, $a = 9$Å, $\theta = 0.8$, $1/\kappa_D = 7$Å, $\lambda_c = 310$Å; (a) $f = 0.5$, $\lambda_t = 324$Å and (b) $f = 0.3$, $\lambda_t = 158$Å. In case (a) DNAs are rigid and soft-chains-approximation does not work; in case (b) $\lambda_t/(2\lambda_c) \approx 1/4$ and the chains are soft. Compared to rigid sequences, the energy of soft chains decreases with the sequence length and such sequences can attract each other. Notations: the average energy of homologous sequences (dotted), random absolutely rigid sequences, Eq. (2.15) (solid), random soft sequences, Eq. (2.26) (dotted-dashed) and Eq. (2.31) (short-dashed), and random sequences within the parabolic approximation, Eq. (2.38) (long-dashed).

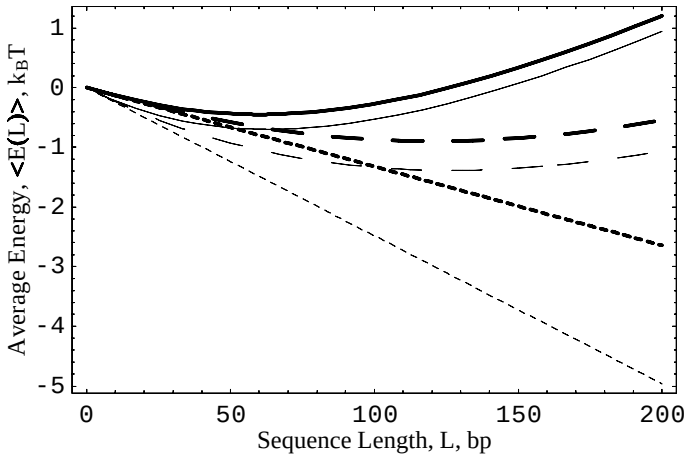

Figure 2.4: *Energy of optimally aligned rigid finite chains (long dashes) is lower than their energy when one end of chains is fixed (solid).* Thick curves: $a_2 \neq 0$ in (2.2), thin curves: $a_2 = 0$. Dotted curves: homologous sequences, Eqs. (2.16), (2.42); long-dashed curves: azimuthally free rigid sequences, Eqs. (2.20), (2.41); solid curves: rigid sequences with one end fixed, Eqs. (2.15), (2.40). Parameters of Fig. 2.3b.

It was shown [189] that in synaptic complex RecA pairs as short as 8 bp-long homology sequences. It seems to be too short both for avoiding of mistakes [190] and for carrying genome-wide search. The effective homologous recombination requires on the other hand 50-200 bp homology [155]. It suggests an alternative, two-step mechanism of homologous recombination [126]: the first step is a cross-grained alignment of 50-200 bp long DNA sequences, governed by direct DNA-DNA interactions, and the second step is the precise match of $\sim$ 10bp-long fragments, which may involve specific proteins. To test this hypothesis more advanced experiments should be done.

Recognition energy: In this Chapter we explore the role of DNA torsional rigidity on homology recognition of the first step of this scheme. We show that although recognition energy decreases for soft sequences, still it is $\sim k_{\rm B}T$ for 200bp-long homology, that warrants the efficient recognition of homologous DNA fragments.

The recognition energy of long duplexes increases proportionally their length, Fig. 2.5. Its slope depends on λ_t and λ_c, being maximal for rigid chains, Eq. (2.18), and minimal for soft chains, Eq. (2.27). For short sequences the recognition energy increases quadratically with the length of the sequence, Eq. (2.19). Similar nonlinear increase of recombination frequency for DNA homologous sequences shorter than 50bp was indeed observed in recombination of *E.coli* [157] and T4 phage, Fig. 2.5a.

Effect of parameters: Expressions (2.15), (2.26), (2.32), and (2.38) describe the dependence of the recognition energy on (i) DNA-DNA separation, the external parameters ((ii) temperature, (iii) Debye screening length of solution, etc.), and DNA intrinsic parameters ((iv) rms variation of twist angle, $\Delta\Omega$, and (v) counterion partitioning on DNA). Consider each effect separately.

(i) if molecules are forced into closer juxtaposition, $\lambda_t \sim a_1^{-1/2}$ decreases (Fig. 2.6) and DNAs adjust their sequence-dependent twist patterns better. The recognition energy for long sequences (both rigid and soft) is $\sim a_1$ and increases nearly exponentially when R decreases. Thus, the recognition energy increases $\sim a_1^{1/2}$.

(ii) Temperature-induced twist fluctuations, which were not taken into account in the present work, may also facilitate the complementarity of twist-angle-patterns in a proper temperature range (c.f. the mechanism of temperature-induced complementarity proposed in Ref. [127]). With temperature increase DNAs can make torsional deformations easier that decreases the recognition energy.

(iii) DNA-DNA interaction energy depends on Debye screening length via screening exponents and pre-exponential factor, Eqs. (A.4), (A.5). In particular, at higher ionic strength of solution DNA charges are screened better that decreases the recognition energy.

(iv) The effect of base-pair sequence is considered in the model through DNA coherence length λ_c. [3] It decreases with increase of root-mean-square variation of the twist angle, $\Delta\Omega$. It leads to disruption of strand-groove register on shorter lengths and increases the recognition energy, especially of long soft chains.

[3] The dependence of DNA twist rigidity on bp-sequence, predicted in computer simulations of real DNAs [191], can modify the results of the present model, being particularly important for kink-solitons, Section 2.7.

In addition, DNA is an extendable molecule [193]. The most known example of it is the variable 142-149 number of bp per turn of nucleosome [192]. DNA can be compressed or elongated in some bp sequences up to 1.5 times [193]. Such huge extension of the duplex, as well as its unwinding (18.5 bp per right-handed turn), is believed to be important for homologous recognition of RecA-ss-DNA and dsDNA fragments [191]. It can also be important for adjustment of sequence-dependent charge patterns on interacting DNAs in our model, since DNA backbone gets one more degree of freedom to relax twist mismatches.

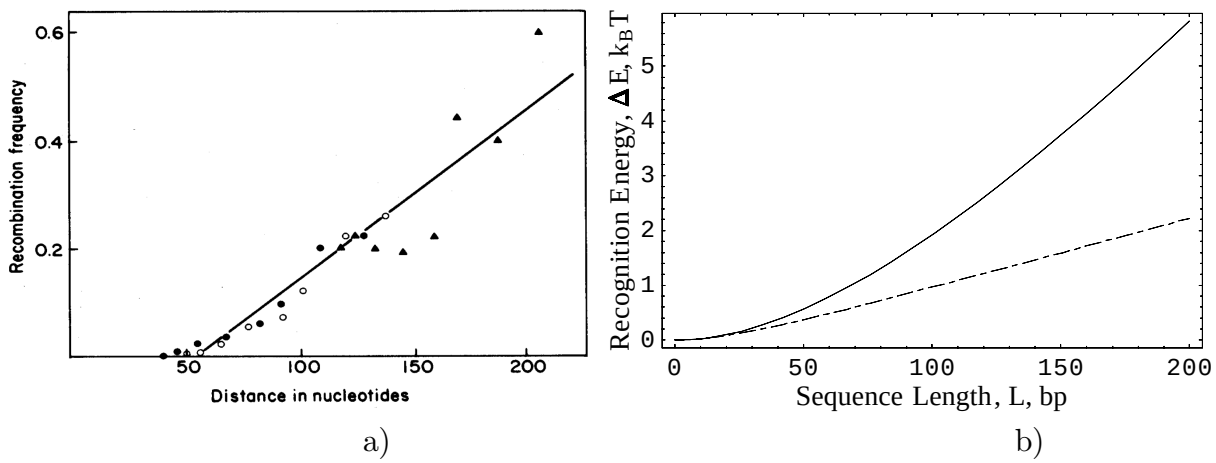

Figure 2.5: *Frequency of homologous recombination events measured by deletion-by-deletion crosses in T4 bacteriophage* [155] (a) and *the predicted electrostatic recognition energy* (b) *increase linearly with the leggnth of homology.* Parameters as in Fig. 2.3b, i.e. $R = 30$Å that is quite representative distance between homologous DNAs before and during recombination process. (a) The region of DNA homology is limited between the deletions in two phage DNAs. Triangles, empty and filled circles correspond to various location of deletions along DNA; the line is least square fit to data. 50bp is the minimal length of DNA homology required for primary pathway for recombination in T4. For shorter sequences the second mode of recombination with $\sim L^{n>1}$ -law was also detected [155]. In spite of the fact that the measured recombination frequency is independent of RecA-family-proteins, its similarity with calculated recognition energy should not be overestimated.

(v) Last but not least, coefficients a_i depend crucially on the patterns of adsorbed charges (see Appendix A). For example, strong occupation of the major groove decreases λ_t substantially (Fig. 2.6) that allows DNAs to make torsional adjustment easier and decreases the recognition energy (Fig. 2.5). [4] If the electrostatic mechanism is indeed involved in recombination process, this fact should affect the homology recognition.

All these parameters may control the efficiency of homologous recombination and it would have been extremely interesting to explore possible effects of this kind experimentally. For instance, the difference in *electrostatic* interaction of homologous and non-homologous sequences could be detected by osmotic stress experiments. [5]

Conclusions. We have calculated the interaction energy of torsionally non-rigid helically non-ideal DNA helices. We have shown that the recognition energy of torsionally soft duplexes decreases with chains softening. We have also found that for typical parameters of B-DNA the recognition, although diminished by torsional elasticity, is still possible. We have revealed a number of factors that can affect the recognition energy.

[4] This dramatic DNA softening suggests a hint why many cations, which bind in the major groove, cause precipitation of unrelated DNA sequences [37], whereas those bound in the minor groove, do not.

[5] Even if a force difference will be seen, the contribution of other forces must be somehow excluded (different sequences have different patterns of hydration, they interact differently with ions, and may have different bp-dependent torsional moduli). Only after this one can say unambiguously that the observed difference comes exclusively from the different laws of electrostatic interaction of homologous and non-homologous sequences. It is however difficult to do for DNA.

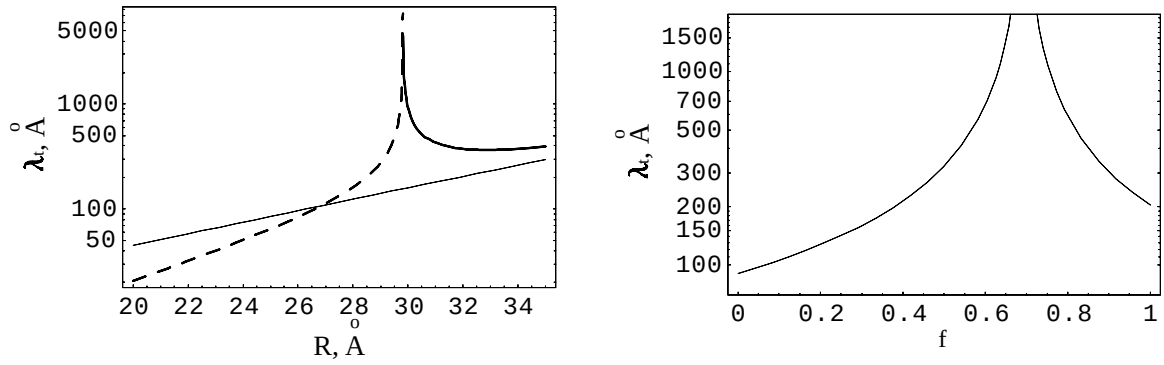

Figure 2.6: *The torsional lengths λ_t (thin), $\widetilde{\lambda}_t$, (thick solid) and $\widetilde{\widetilde{\lambda}}_t$ (thick dashed curve) describe the torsional softness of DNA duplexes in the model.* Parameters: $\theta = 0.8$, $\kappa_D^{-1} = 7$Å, (left) $f = 0.3$; (right): $R = 30$Å. λ_t decreases with approaching of the duplexes and with occupation of the major (an minor) groove. λ_t diverges when a_1 vanishes, $2f\theta - \theta - \cos[0.4\pi] = 0$.

2.7 Beyond cos(δφ) interaction potential: ready to kink

In this Section we explore the basic consequences of interaction energy (2.4) with $a_2 \neq 0$. Below we calculate the energy of the rigid chains, following [126], and use the harmonic approximation to calculate the energy of soft chains.

2.7.1 Rigid chains

Within the approximation of absolutely rigid chains, Eq. (2.13), the interaction energy of *random* sequences of the length L is [126]

$$\langle E(L)\rangle = a_0 L - a_1\lambda_c\left(1 - e^{-L/\lambda_c}\right) + \frac{a_2\lambda_c}{4}\left(1 - e^{-4L/\lambda_c}\right), \tag{2.40}$$

the energy of azimuthally *free unrelated* sequences at the most favorable alignment is lower

$$\langle E(L)\rangle = a_0 L - 2a_1\lambda_c\left(1 - e^{-L/\lambda_c}\right) + 2\frac{a_2\lambda_c}{4}\left(1 - e^{-2L/\lambda_c}\right), \tag{2.41}$$

the energy of *identical* sequences is the most favorable

$$\langle E(L)\rangle = (a_0 - a_1 + a_2)\, L. \tag{2.42}$$

All energies are shown in Fig. 2.4.

2.7.2 Soft chains: Two-well-approximation

The simplest $\cos[\delta\phi]$-approximation considered above works reasonably well at separations, larger than the separations of spontaneous symmetry breakdown $R_* \simeq 30 \div 35$Å in the interaction potential [84]. (R_* is determined from the condition $a_1(R_*)/[4a_2(R_*)] = 1$). At smaller separations we should also account for $\cos[2\delta\phi]$-term in $E_{\rm int}(\delta\phi)$, Eq. (2.2).

Large separations. At $R > R_*$ the mutual angle $\delta\phi = 0$ minimizes $E_{\rm int}$. In the harmonic

approximation ($\delta\phi \ll 1$), analogously to (2.26), for average energy we obtain

$$\langle E(L)\rangle = a_0 L - (a_1 - a_2) L + [a_1 - 4a_2]\,\tilde{\lambda}_t/(2\lambda_c), \tag{2.43}$$

where $\tilde{\lambda}_t = \sqrt{C/[2(a_1 - 4a_2)]}$ is a new DNA torsional length [85].

Small separations. At $R < R_*$ the interaction energy has a two-well shape, the angles $\delta\phi_* = \pm\arccos[a_1/(4a_2)]$ describes the minima of E_{int}. If rms twist variation $\Delta\Omega$ is small or DNA torsional rigidity is small, $\delta\phi$ will develop near the bottom of one energy well and no transition between the wells occurs. The criterion of validity of this approximation can be written, analogously to (2.36), in terms of mean squared angle, i.e. when

$$\sqrt{\left\langle (\delta\phi - \delta\phi_*)^2 \right\rangle} \lesssim \delta\phi_*, \tag{2.44}$$

or $\frac{\sqrt{Ca_2}}{\lambda_c\sqrt{2}\sqrt{(4a_2)^2 - a_1^2}} \lesssim \arccos^2\left[\frac{a_1}{4a_2}\right]$.

Then, using harmonic approximation for the shape of interaction energy near the minimum, we get

$$\langle E(L)\rangle = a_0 L - \left(\frac{a_1^2}{8a_2} + a_2\right) L + \left[4a_2 - \frac{a_1^2}{4a_2}\right]\frac{\tilde{\tilde{\lambda}}_t}{2\lambda_c}, \tag{2.45}$$

where $\tilde{\tilde{\lambda}}_t = \sqrt{C/[2(4a_2 - a_1^2/(4a_2))]}$ is new DNA torsional length.

Since in both cases we effectively consider soft chains, interaction energy (2.43) and (2.45) is the sum of the energy of homologous sequences plus a positive term proportional to DNA torsional rigidity modulus. Criterion (2.44) fails near R_*, as well as the harmonic approximation used for $R > R_*$, when the torsional lengths $\tilde{\lambda}_t$ and $\tilde{\tilde{\lambda}}_t$ diverge (Fig. 2.6).

2.7.3 The role of kink-solitons

Large $\Delta\Omega$ and large C may however cause the variation of angle $\sim \delta\phi_*$, and the approximation of single parabola, Eqs. (2.43), (2.45), fails. At $R < R_*$ a kink-like transition in $\delta\phi(Z)$ *may* take place, which switch the energy minima from one well to another (see small kink, Section 2.8). It can prevent the accumulation of large mismatches. At $R > R_*$ only $0 \to 2\pi$ transition can exist (big kink), that can require much larger mismatches and increases the energy of the system.

Biological importance: The torsional dynamic solitons on DNA may be responsible, in particular, for transport of specific proteins along the DNA [194]. It is suggested that proteins binding cause conformational distortion of DNA backbone, which may travel along the molecule without dispersion, carrying protein to corresponding gene (∼1kbp apart). It was shown that the static Sine-Gordon-like soliton placed within the promoter region develops along the DNA, whereas it remains static being placed outside this region [194].

Another indication of importance of undamped dynamic torsional deformation of the duplex comes from the observations of homologous recombination occurring in some yeast far (∼ 30 kbp!) from double-strand DNA breaks [150]. As one possible mechanism it was suggested that a "recombination machine" which consists of special enzymes and proteins, can entry at double-strand break and travel along DNA to a distant point before promoting recombination.

2.8 Azimuthal kinks on ideal non-rigid DNA duplexes

In previous Sections we have considered the effect of sequence-dependent DNA twist on the behavior of mutual azimuthal angle $\delta\phi\,(z)$ on rigid and soft chains. The approximations used did not however imply a kink-like solution for $\delta\phi\,(z)$. Here we consider the effect of DNA torsional softness on azimuthal angle of two *ideal* DNAs and investigate whether the interaction potential (2.46) implies a kink-like solution *along* interacting DNA duplexes.

As follows from Eq. (1.30), the pair *local* DNA-DNA interaction free energy has the form [85] (per unit length)

$$\varepsilon_{\rm int}\left(\delta\phi\right) = -a_1\left(R\right)\cos\left[\delta\phi\right] + 2a_2\left(R\right)\cos^2\left[\delta\phi\right], \tag{2.46}$$

where $a_1\left(R\right),\ a_2\left(R\right)$ are positively defined functions, exponentially decaying with DNA-DNA interaxial separation R (see Appendix A). These functions depend on DNA charge compensation θ, the fraction of counterions in the minor groove, f, the reciprocal Debye screening length of the solution, κ_D^{-1}. All these parameters are assumed below to be independent on R, possible effects of non-fixed patterns of adsorbed cation and increasing κ_D with approach of the molecules are considered in Chapter 1, [86].

As we noted, minimization of $\varepsilon_{\rm int}\left(\delta\phi\right)$ results in equation: $\sin\left[\delta\phi\right]\left(a_1 - 4a_2\cos\left[\delta\phi\right]\right) = 0$. Its solution is $\delta\phi = 0\,(\mathrm{mod}\,2\pi)$, when $R > R_*$, and

$$\delta\phi_* = \pm\arccos\left(\frac{a_1}{4a_2}\right), \tag{2.47}$$

when $R < R_*$. I.e., the spontaneous symmetry breakdown occurs and nonzero mutual azimuthal angle are optimal when two DNAs are closer than critical separation R_* [84], Fig. 2.9. R_* is found from the condition $\frac{a_1(R_*)}{4a_2(R_*)} = 1$ and is 30-35Å for typical DNA parameters. The value of interaction energy in the minima is $\varepsilon_{\rm int}\left(0\right) = 2a_2 - a_1$, when $R > R_*$, and $\varepsilon_{\rm int}\left(\delta\phi_*\right) = -\frac{a_1^2}{8a_2}$, when $R < R_*$.

The angles $\delta\phi = 0, 2\pi, 4\pi, ...$ and $\delta\phi = \delta\phi_*$, constant along the chains, describe the absolute energy minimum. The possible transitions between the angles in each set describe the states of the system with higher energy. We describe the properties of these excited states in the next Sections.

2.8.1 Free energy functional and basic equations

Below we consider DNA torsional elastic energy in the simplest model [126] with *macroscopic* rigidity modulus $C \simeq 3\times10^{-19}$ erg cm [172, 173], that leads to the following Hamiltonian of the system of two non-rigid *ideal* DNA molecules [195]

$$\mathcal{H} = \int_{-\infty}^{+\infty} dz\left\{\varepsilon_{\rm int}\left(\Phi_1 - \Phi_2\right) + \frac{C}{2}\left(\frac{d\Phi_1}{dz}\right)^2 + \frac{C}{2}\left(\frac{d\Phi_2}{dz}\right)^2\right\}. \tag{2.48}$$

The variation of $\mathcal{H}$ with respect to Φ_1 and Φ_2 leads to the following equation (see Appendix A)

$$2\frac{\partial\varepsilon_{\rm int}}{\partial\delta\phi} = C\frac{\partial^2\delta\phi}{\partial z^2}. \tag{2.49}$$

Multiplying both sides on $\frac{\partial\delta\phi}{\partial z}$ and integrating from z_0 to z, we get the Euler mean-filed equation

on the *local* mutual azimuthal angle

$$\pm\sqrt{\frac{4a_1}{C}}\,(z-z_0)=\int_{\delta\phi(z_0)}^{\delta\phi(z)}\frac{d\delta\phi}{\sqrt{\varepsilon_{\rm int}\,(\delta\phi)-\varepsilon_{\rm int}\,(\delta\phi_0)}}. \tag{2.50}$$

We *expect* this equation to have a kink-like solutions. It is useful thus to choose z_0 to be the position of the center of the kink. This can be done by choosing $\delta\phi\,(z_0)$ to be the value of $\delta\phi$ at the center of the kink. Like for the Sine-Gordon kink, we choose below $z_0=0$.

The energy of topological kink with the shape described by Eq. (2.50), is determined as

$$E=\sqrt{C}\int_{\delta\phi(z=-\infty)}^{\delta\phi(z=+\infty)}\sqrt{\varepsilon_{\rm int}\,(\delta\phi)-\varepsilon_{\rm int}\,(\delta\phi_0)}d\delta\phi. \tag{2.51}$$

2.8.2 Big kink

At $R{>}R_*$ angles $\delta\phi=0(\mathrm{mod}\ 2\pi)$ describe the minima of interaction energy of two DNAs. The kink solution derived below describe the transition along the molecule between two nearest minima of interaction energy, $\delta\phi=0$ and $\delta\phi=2\pi$. We will call below this kink-soliton "*big*" kink.

Thus, we put $\delta\phi\,(0)=\pi$. Then Eq. (2.50) turns into the following equation

$$\pm\sqrt{\frac{4a_1}{C}}z=\int_{\pi}^{\delta\phi(z)}\frac{d\delta\phi}{\sqrt{1-\cos\left[\delta\phi\right]-\gamma\left\{1-\cos^2\left[\delta\phi\right]\right\}}}, \tag{2.52}$$

where

$$\gamma\equiv 2a_2/a_1<1/2. \tag{2.53}$$

Direct integration of Eq. (2.52) gives

$$\pm\sqrt{\frac{4a_1}{C}}z=\pm\sqrt{\frac{2}{1-2\gamma}}\,\mathrm{arctan\,h}\left\{\frac{\sqrt{1-2\gamma}\cos\left[\delta\phi/2\right]}{\sqrt{\gamma\left(1+\cos\left[\delta\phi\right]\right)-1}}\right\}, \tag{2.54}$$

Resolving Eq. (2.54) with respect to $\delta\phi\,(z)$, we get the shape of the *big* kink

$$\delta\phi_b\,(z)=\begin{cases}\arccos\left\{\frac{\sinh^2[Z]-(1-2\gamma)}{\sinh^2[Z]+(1-2\gamma)}\right\}, & Z<0\\ 2\pi-\arccos\left\{\frac{\sinh^2[Z]-(1-2\gamma)}{\sinh^2[Z]+(1-2\gamma)}\right\}, & Z>0\end{cases}, \tag{2.55}$$

where the dimensionless length along the molecule, $Z=z\sqrt{1-2\gamma}\sqrt{\frac{2a_1}{C}}$, has been introduced. At $\gamma=0$ big kink turns into *Sine-Gordon* kink,

$$\delta\phi_{SG}\,(z)=4\arctan\left[\exp\left(\pm\frac{z}{\sqrt{C/2a_1}}\right)\right], \tag{2.56}$$

with energy $E_{SG}=8\lambda_t a_1=4\sqrt{2}\sqrt{Ca_1}$.

The half-width of $\delta\phi_b\,(z)$ is determined by $1/\left(\sqrt{1-2\gamma}\sqrt{\frac{2a_1}{C}}\right)=\lambda_t/\sqrt{1-2\gamma}$. The DNA torsional length is $\lambda_t\sim 200-500$Å for typical parameters, Fig. 2.6. The larger the DNA

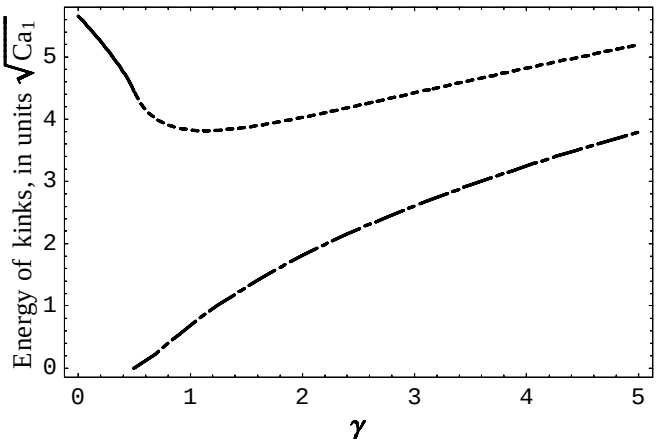

Figure 2.7: *Energy of kinks on ideal torsionally soft DNA duplexes.* Big kink, Eq. (2.57), (solid curve) exists at large separations $R > R_*$ ($\gamma < 1/2$) and coresponds to $0 \rightarrow 2\pi$ azimuthal rotation of mutual orientation angle. Small kink, Eq. (2.60), exists at $R < R_*$ and corresponds to $-\delta\phi_* \rightarrow +\delta\phi_*$ rotation, (dotted-dashed curve) and to $\delta\phi_* \rightarrow 2\pi - \delta\phi_*$ rotation (dashed curve). When $\gamma \gg 1$, the energies of the latter rotations equalize since $\delta\phi_* \rightarrow \pi/2$.

torsional rigidity, the wider the kink. When DNAs approach closer to each other, the coefficient $a_1(R)$ increases, that decreases the kink width. (stronger interaction can make more localized kink). In the point of spontaneous symmetry breakdown, at $\gamma = 1/2$, the width of the big kink diverges.

Kink-solution (2.55) describes $0 \rightarrow 2\pi$ torsional rotation along the molecule between the minima of interaction energy $\varepsilon_{\text{int}}(\delta\phi)$. The energy of big kink is calculated according to Eq. (2.51) as

$$E_b = \sqrt{C}\int_0^{2\pi}\sqrt{1-\cos[\delta\phi]-\gamma\{1-\cos^2[\delta\phi]\}}d\delta\phi$$
$$= 2\sqrt{2}\sqrt{Ca_1}\left\{\sqrt{1-2\gamma}+\frac{\arcsin\left[\sqrt{2\gamma}\right]}{\sqrt{2\gamma}}\right\}. \quad (2.57)$$

2.8.3 Small kink

For close DNA-DNA separations, $R < R_*$, the non-zero azimuthal angle $\delta\phi_* = \pm\arccos\left(\frac{1}{2\gamma}\right)$ minimizes the interaction energy (2.46). It appears that Euler equation (2.50) has a kink-like solution, below "*small*" kink, which describes the torsional rotation between the angles $\delta\phi_*$. We choose below $\delta\phi(0) = 0$.

From Eq. (2.50), the shape of the small kink is determined from equation

$$\pm\sqrt{\frac{4a_1}{C}}z = \int_0^{\delta\phi(z)}\frac{d\delta\phi}{\sqrt{\frac{1}{2\gamma}-\cos[\delta\phi]-\gamma\left(\frac{1}{(2\gamma)^2}-\cos^2[\delta\phi]\right)}} \quad (2.58)$$

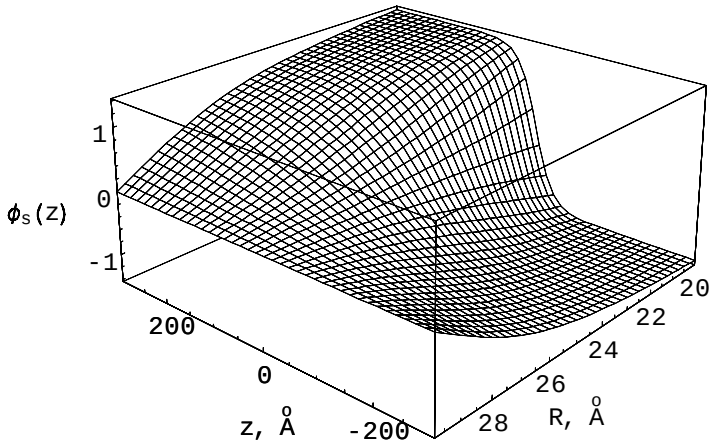

Figure 2.8: *The shape of the small kink*, Eq. (2.59). It exists at $R < R_*$ and corresponds to torsional rotation of DNA backbone between the angles $\pm\delta\phi_*$. With approach of DNAs to each other, kink width decreases and $\delta\phi_* \to \pm\pi$. Parameters: $\theta = 0.8,\ f = 0.3,\ \kappa_D^{-1} = 7$Å.

by direct integration as

$$\delta\phi_s(z) = 2\arctan\left\{\pm\sqrt{\frac{2\gamma-1}{2\gamma+1}}\tanh\left[z\sqrt{\frac{2a_1}{C}}\frac{1}{2}\sqrt{\frac{(2\gamma)^2-1}{2\gamma}}\right]\right\}. \tag{2.59}$$

Plus and minus in Eq. (2.59) corresponds to kink and anti-kink, respectively. Since $\delta\phi_* \to 0$ when $\gamma \to 1/2$, near the point of the symmetry break-down the energy of the small kink,

$$E_s = \sqrt{C}\int_{-\delta\phi_*}^{\delta\phi_*}\sqrt{\frac{1}{2\gamma}-\cos[\delta\phi]-\gamma\left(\frac{1}{(2\gamma)^2}-\cos^2[\delta\phi]\right)}d\delta\phi$$
$$= 2\sqrt{Ca_1\gamma}\left\{\sqrt{1-\frac{1}{(2\gamma)^2}}-\frac{\arccos\left[\frac{1}{2\gamma}\right]}{2\gamma}\right\}, \tag{2.60}$$

tends to zero. The half-width of the small kink *diverges* at $\gamma \to 1/2$.

At $R < R_*$ the second possibility for rotation of local azimuthal on interacting duplexes appears. It corresponds to azimuthal rotation from angle $\arccos\left(\frac{1}{2\gamma}\right)$ to angle $2\pi - \arccos\left(\frac{1}{2\gamma}\right)$. I.e., the angle difference of such rotation is always larger than π. Such rotation has the energy

$$E_s = \sqrt{C}\int_{\delta\phi_*}^{2\pi-\delta\phi_*}\sqrt{1-\cos[\delta\phi]-\gamma\{1-\cos^2[\delta\phi]\}}d\delta\phi =$$
$$= 2\sqrt{Ca_1\gamma}\left\{\sqrt{1-\frac{1}{(2\gamma)^2}}+\frac{\pi/2+\arcsin\left[\frac{1}{2\gamma}\right]}{2\gamma}\right\} \tag{2.61}$$

The shape of the small kink is depicted in Fig. 2.8, the energies of small and big kinks varying with DNA-DNA separation are depicted in Fig. 2.9.

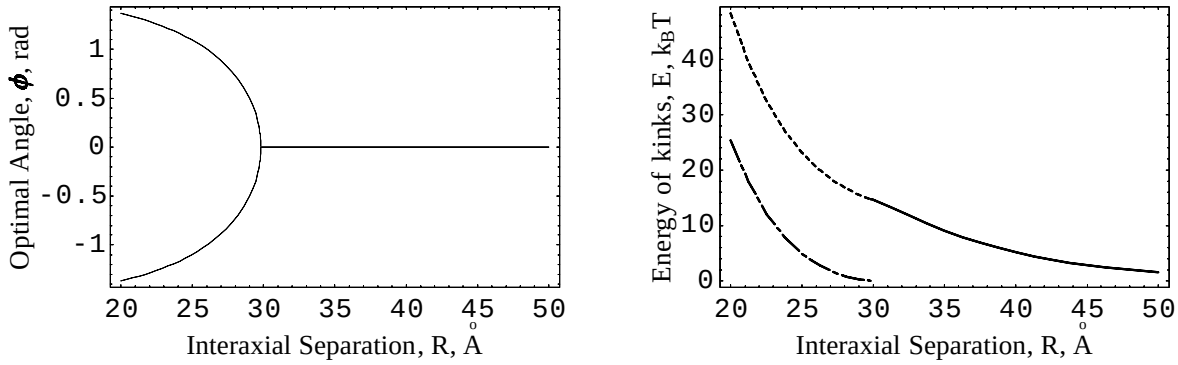

Figure 2.9: *The optimal azimuthal angle $\delta\phi_*$ between two ideal rigid DNA (left) and the energy of big, Eq. (2.57), and small, Eq. (2.60), kinks (right)* $\left(R_* \approx 30\text{Å}\right)$. The same notations for energies as in Fig. 2.7 as used, parameters as in Fig. 2.8. As expected, the energy of big kink is much higher than that of small kink. All energies increase nearly exponentially with compression.

2.8.4 Outlook

Concentration of kinks: The energy of two molecules with above-predicted kinks is higher than the energy of the system without kinks; one can consider kinks as *defects*. At zero temperature there will be no kinks, at finite temperature their density will be exponentially small. For *estimation* one can use the result for linear density of thermally-activated Sine-Gordon kinks [196]

$$n_{SG} = 2\sqrt{\frac{2}{\pi}}\frac{\sqrt{E_{SG}/\left(k_{\mathrm{B}}T\right)}}{d}\exp\left[-E_{SG}/\left(k_{\mathrm{B}}T\right)\right], \tag{2.62}$$

where E_{SG} is the energy of Sine-Gordon kink and d is its width.

For typical DNA parameters the energy of the big kink is $\sim 10-30k_{\mathrm{B}}T$, Fig. 2.9. According to Eq. (2.62), the concentration of thermally-activated big kinks is of the order of 1 kink per each $\sim 10^5 - 10^7$ bp. It is negligible concentration for relaxing of mismatches on DNAs with sequence-dependent twist variation (see previous Sections).

The concentration of small kinks can be much higher and their effect on interaction energy can be more appreciable, especially in the vicinity of the point of spontaneous symmetry breakdown, where the energy of small kink vanishes, Eq. (2.60). At $R \to R_*$ the half-width of the small kink however diverges, that diminishes its efficiency in relaxing of twist mismatches.

Order-disorder lattice: The equation, which governs the pinning of these kinks by random variation of DNA twist, is the Euler equation for free energy (A.3), (A.10) [126]

$$\frac{d^2\delta\phi}{dZ^2} - \tilde{\kappa}^2\sin\left(\delta\phi\right)\left[1 + \frac{8a_2}{a_1 - 4a_2}\sin^2\left(\delta\phi/2\right)\right] = \frac{d\delta\Omega}{dZ}, \tag{2.63}$$

where $\tilde{\kappa} = h/\tilde{\lambda}_t$.

Pinning of the kinks by random "impurities" of DNA twist pattern could (substantially) decrease the energy of kink nucleation. One could expect in this case the formation of a quasi-periodic wall of Frenkel-Kontorova-like kink dislocations [197, 198]. In this mechanism the long part of the chains is kept out of disorder whereas the accumulated mismatch is released in the

narrow region via a kink [6,7]. The effect of pinning of small and big kinks by random DNA twist patterns on the averaged energy requires additional investigation and is the subject of our future work.

[6] It is plausible at least for interaction of dinucleotide seqeuences, $\{^{\mathrm{AAAA...}}_{\mathrm{TTTT...}}\}$ with $\{^{\mathrm{CCCC...}}_{\mathrm{GGGG...}}\}$. (The sequences with large percentage of AT or GC bp is not a fiction, but they form block structure of DNA of higher organisms [134], they can be several kbp long.) The twist angle between nearest bp in such sequences is constant and equal to $\approx 35.6°$ and $\approx 33.6°$, respectively [167]). If such sequences are forced into juxtaposition, the constant incommensuration along the chains appears (first has 10.1bp/turn, the second – 10.5bp/turn [199]). The Kornyshev-Leikin-like DNA-DNA interaction is then approximated by pure Frenkel-Kontorova model. As a result, kink lattice of $\delta\phi$ with a constant period can appear, if the twist variation $\delta\Omega > 4\kappa/\pi \sim 1°$. The random twist variation will however disrupt this ideal lattice.

[7] The first estimation for fragments with random texts show that such structure can realize only on soft chains. If, for example, $\delta\phi$ does not respond to twist variation at all, only the chains with $\lambda_t^2 < h\lambda_c$ can profit from that (compare to criterion (2.24)).

On the other hand, we have performed *numerical simulations* of Eq. (2.9). We model DNA twist variation so that the twist angle of each bp differs from $\langle\Omega\rangle$ only by $\pm\Delta\Omega = \mathrm{Const}$. For very rigid chains, $\delta\phi\,(z)$ obeys random-walk motion along the chains. At $\kappa \sim 0.01$ the regions of correlation of $\delta\phi$ near 0, 2π, $4\pi, ...$ appear. Between these regions $\delta\phi$ *tracks* the twist variation. The relative width of correlated and uncorrelated regions depends on $\Delta\Omega$ and λ_t. The trajectory of $\delta\phi$ reminds a random walk motion *between the energy minima*, similar to trajectory of overdamped forced Sine-Gordon equation under perturbation by a thermal random force [200].

Conclusions

This thesis is focused on the *theory of electrostatic interaction of DNA duplexes*: the complementarity and recognition of DNA charge patterns, torsional adjustment of DNA backbone and DNA-DNA attraction are explored.

DNA helical symmetry determines to a large extend the properties of DNA *in vitro* and its functioning in cells. The exact pair interaction potential of ideal DNAs [84] provides a quantitative description of many of puzzles of DNA-DNA interaction. We have modified this potential to incorporate (i) the non-fixed, adjusting patterns of adsorbed cations and (ii) helical non-ideality and torsional flexibility of the DNA backbone. It allows to describe (i) the temperature-induced DNA condensation and (ii) interaction and recognition of non-ideal DNA duplexes.

(i) In the last 10-15Å surface-to-surface separation between DNAs, their helicity can give rise to a strong attraction. We have shown that DNA condensation is governed by temperature-induced adjustment of Mn^{2+} adsorption patterns that strengthens this attraction. The caused entropy increase upon aggregation is comparable with the traditional entropy of water release [86]. Advanced experiments are required to test the predicted features, including the predicted frustrated DNA lattices.

The model can be modified to describe the aggregation of *imperfect* DNA helices. The attraction in this case should be less pronounced, but temperature- and Mn-induced torsional DNA softening can help DNAs to condense. Thus, the entropy of the backbone, as well as the change of hydration upon adsorption of ions, are other plausible entropy sources, which can also contribute to aggregation of DNA and other biopolymers.

This study can help to elucidate *in vitro* DNA condensation in solutions of trivalent ions, where the specificity of binding is crucial as well. Incorporation of DNA bending rigidity in the model can also shed light on toroidal DNA condensation *in vivo* as well as on the properties of DNAs wrapped around histone cores in nucleosome.

(ii) Even slightly noticeable non-ideality of the DNA helical structure may have a huge effect on electrostatic interaction and, especially, recognition of DNA duplexes. Just that non-ideality allows homologous (identical) DNA fragments to recognize and attract each other [126]. This attraction can bring them into closer juxtaposition to ensure entering the recombination process. Non-homologous *rigid* sequences repel each other, and the stronger the longer the sequences.

We have shown that finite DNA torsional softness diminishes the recognition energy of homologous sequences, but the recognition still remains strong. The dependence of recognition energy on many factors and parameters has been calculated in the model [174]. Only systematic future experiments could prove whether they indeed affect DNA-DNA recognition. If they do, new controlling tools could be suggested for gene targeting and shuffling, as well as for DNA repair. This could help envisaging new methods for curing diseases, affected by erroneous recombination.

Acknowledgement

The author is very grateful first of all to his scientific supervisor, Prof. A.A.Kornyshev (Research Center Jülich, Germany; University of Düsseldorf, Germany; Imperial College, UK), and to Dr. S.Leikin (NIH, USA) for constant encouragement and help. Without their ideas and huge contribution as authors of both described above works, the present dissertation would not be possible.

The author is also thankful to Prof. A.M.Berezhkovskij (NIH, USA), Dr. P.L.Hansen (NIH, USA), Prof. H.Löwen (University of Düsseldorf, Germany), Dr. S.V.Malinin (Research Center Jülich), Prof. A.V.Parsegian (NIH, USA), Dr. D.C.Rau (NIH, USA), and Prof. J.M.Schurr (University of Washington, USA) for many stimulating discussions; and to Dr. E.Spohr for critical reading of several pieces of the manuscript.

Author wants to acknowledge the Research Center Jülich (Forschungszenrum Jülich GmbH, Germany) for the possibility to make this PhD work in Germany, and also Research Center Jülich, National Institute of Health (NIH, MD, USA), and Deutsche Forschungsgemeinschaft (Germany) for financial support of his trips to NIH, Bethesda, USA.

Appendix A

Non-rigid duplexes: The free energy functional

The free energy functional for two parallel *non-rigid non-ideal* DNA duplexes of the length L in juxtaposition at interaxial separation R can be written as a sum of their *local* torsional and electrostatic energy densities,

$$E\left(L\right)=\int_{0}^{L}dz\ \left[\varepsilon_{t}(z)+\varepsilon_{\mathrm{int}}(z)\right]. \tag{A.1}$$

Expressions for $\varepsilon_{t}\left(z\right)$ and $\varepsilon_{\mathrm{int}}\left(z\right)$ are written in terms of local twist angles, $\Phi_{1,2}(z)$, (Fig. 2.1)

$$\varepsilon_{t}(z)=\frac{C}{2}\left[\left(\frac{d\Phi_{1}}{dz}-\frac{\Omega_{1}(z)}{h}\right)^{2}+\left(\frac{d\Phi_{2}}{dz}-\frac{\Omega_{2}(z)}{h}\right)^{2}\right],\quad \varepsilon_{\mathrm{int}}(z)=\varepsilon_{\mathrm{int}}(\Phi_{1}(z)-\Phi_{2}(z)), \tag{A.2}$$

where subscript $i=1,2$ labels the molecules; h is DNA rise (axial phosphate-to-phosphate distance).

$\Omega_{i}(z)$ is the local azimuthal angle on isolated molecule, the preferred twist angle between adjacent base pairs. These angles take the values $\Omega_{i}=\langle\Omega\rangle\pm(4-6^{\circ})$, $\langle\Omega\rangle\approx34^{\circ}-35^{\circ}$ [167,169], that reflect the intrinsic non-ideality of helical DNA structure.

The modulus of DNA torsional rigidity C [172] determines to which extend the twist angles on interacting DNAs are close to their values on single molecule, $\Omega_{i}\left(z\right)$. In our model the *actual* twist angles, $h\left[d\Phi_{i}\left(z\right)/dz\right]$, differ from $\Omega_{i}\left(z\right)$, because electrostatic interaction may cause torsional DNA deformation.

Pair DNA-DNA *electrostatic* interaction free energy density, $\varepsilon_{\mathrm{int}}$, depends on $\delta\phi\left(z\right)=\Phi_{1}\left(z\right)-\Phi_{2}\left(z\right)$ (Fig. 2.1), and can be approximated by first two terms of sum (1.30) as [85] [1]

$$\varepsilon_{\mathrm{int}}\left(R,z\right)\approx a_{0}\left(R\right)-a_{1}\left(R\right)\cos\left[\delta\phi\left(z\right)\right]+a_{2}\left(R\right)\cos\left[2\delta\phi\left(z\right)\right], \tag{A.3}$$

[1] Note however that the situation is possible when factor $f\left(1,\theta\right)$ in a_{1} is small, whereas similar factors in $\mathrm{a}_{2,3}$ are not (a_{3} is described by Eq. (A.5) with $n=3$). For B-DNA with $\theta=0.8$ it occurs at $f\approx0.7$, when the inequality $a_{1}\ll a_{2}\ll a_{3}$ holds (a_{1} vanishes, see Fig. 2.6). I.e., for such occupation of the grooves the term with $\cos[3\delta\phi]$ must be taken into account in interaction energy (A.3) as well. Since the decay length of coefficients a_{n} decreases with n, Eq. (A.6), the contribution of a_{3}-term can be neglected already at $R\gtrsim40\mathring{\mathrm{A}}$.

where the coefficients for B-DNA have the form (see Eq. (1.30))

$$a_0(R) = \frac{8\pi^2\overline{\sigma}^2}{\varepsilon}\left\{\frac{(1-\theta)^2 K_0(\kappa_D R)}{\kappa_D^2 [K_1(\kappa_D a)]} - \sum_{n,j=-\infty}^{\infty}\frac{[f(n,\theta)]^2}{\kappa_n^2}\left[\frac{[K_{n-j}(\kappa_n R)]^2 I_j'(\kappa_n a)}{[K_n'(\kappa_n r)]^2 K_j'(\kappa_n a)}\right]\right\} > 0, \tag{A.4}$$

$$a_{n=1,2}(R) = \frac{16\pi^2\overline{\sigma}^2}{\varepsilon}\frac{[f(n,\theta)]^2}{\kappa_n^2}\frac{K_0(\kappa_n R)}{[K_n'(\kappa_n a)]^2} > 0, \tag{A.5}$$

and

$$f(n,\theta) = f\theta + (-1)^n(1-f)\theta - \cos[0.4\pi n], \qquad \kappa_n = \sqrt{\kappa_D^2 + \frac{4\pi^2 n^2}{H^2}}. \tag{A.6}$$

Here $a \approx 9$Å is the radius of the cylindrical surface formed by centers of phosphates, $H \approx 34$Å is the *average* value of B-DNA helical pitch [2]; $\overline{\sigma} \approx 16.8\mu C/\mathrm{cm}^2$ is the surface charge density of the phosphates; θ is the fraction of DNA charge neutralization by bound counterions; f and $1-f$ are the fractions of counterions bound in the minor and major groove, respectively (no binding of cations on strands is considered here); $\varepsilon \approx 80$ is the dielectric constant of water; κ_D is the Debye screening length ($1/\kappa_D \approx 7\mathring{A}$ in physiological solution); $I_n(x)$, $K_n(x)$, $I_n'(x)$ and $K_n'(x)$ are the modified Bessel functions and their derivatives, respectively.

For getting expression (A.3) both phosphates and adsorbed cations are approximated as uniformly charged helical strings on the distance a from DNA axis. The strings of cations run in the middle of the grooves. The inner core of DNA is approximated by a medium with low dielectric constant, $\varepsilon_{core} \ll \varepsilon$ (see [84] and Section 1.4 for more details).

Further, it is convenient to introduce new variables: $\delta\phi(z) = \Phi_1(z) - \Phi_2(z)$, $V(z) = \Phi_1(z) + \Phi_2(z)$, $\delta\Omega(z) = \Omega_1(z) - \Omega_2(z)$, $T(z) = \Omega_1(z) + \Omega_2(z)$, $P(z)^2 = \Omega_1(z)^2 + \Omega_2(z)^2$. In terms of these variables $\varepsilon_{\mathrm{int}}(z) = \varepsilon_{\mathrm{int}}(\delta\phi(z))$ and $\varepsilon_t(z) = \varepsilon_-(z) + \varepsilon_+(z)$, where

$$\varepsilon_-(z) = \frac{C}{2}\left[\frac{1}{2}\left(\frac{d\delta\phi}{dz}\right)^2 - \frac{d\delta\phi}{dz}\frac{\delta\Omega}{h}\right], \quad \varepsilon_+(z) = \frac{C}{2}\left[\frac{1}{2}\left(\frac{dV}{dz}\right)^2 - \frac{dV}{dz}\frac{T}{h} + \frac{P^2}{h^2}\right]. \tag{A.7}$$

The minimization of energy functional (A.1) over $\delta\phi(z)$ and $V(z)$ leads to Euler equations:

$$\frac{\partial\varepsilon_-}{\partial\delta\phi} + \frac{\partial\varepsilon_{\mathrm{int}}}{\partial\delta\phi} = 0, \quad \frac{\partial\varepsilon_+}{\partial V} = 0. \tag{A.8}$$

Introducing the dimensionless length variable, $Z = z/h$, we obtain the equations of the free energy extremum

$$\frac{d^2\delta\phi}{dZ^2} = \frac{d\delta\Omega}{dZ} + \frac{2h^2}{C}\frac{\partial\varepsilon_{\mathrm{int}}}{\partial\delta\phi}, \quad -\frac{d^2V}{dZ^2} + \frac{dT}{dZ} = 0. \tag{A.9}$$

[2] It is assumed that the seqeunce-dependent twist variation is so small that does not disrupt the DNA helical symmetry and does not change the average value of the pitch. It is also assumed, Eq. (A.3), that DNA interact locally is if they are infinitely long. The inaccuracy involved is hard to estimate in both cases. It is clear that if DNAs retain the helical symmetry in the model, no order-disirder transitions with variable pitch along the molecule are possible. The more accurate consideration of soft chains however may require to take into account the effects of $H(z)$ and $\langle H(R)\rangle$, since the unwinding of the chains can be favorable.

Using Eqs. (A.7), (A.9) we rewrite (A.1) in terms of $\delta\phi\,(Z)$ as

$$E\,(L) = h \int_0^{L/h} dZ \left[\varepsilon_{\mathrm{int}}\,(\delta\phi\,(Z)) + \frac{C}{4h^2} \left(\frac{d\delta\phi\,(Z)}{dZ} - \delta\Omega(Z) \right)^2 \right]. \tag{A.10}$$

Appendix B

Non-rigid duplexes: perturbation theory in small κ

Below we find correction term to energy of absolutely rigid chains, Eq. (2.15), in small κ. We substitute the solution as a sum of (2.13) and a small deviation, $\kappa^2\delta\Phi \ll 1$, into Eq. (2.6). The condition of applicability of this approximation is estimated below. Then the equation for $\delta\Phi$ takes the form

$$\frac{d^2\delta\Phi(Z)}{dZ^2} - \kappa^2 \cos\left[\delta\phi_r\left(Z\right)\right]\delta\Phi(Z) = \sin\left[\delta\phi_r\left(Z\right)\right], \tag{B.1}$$

where $\delta\phi_r\left(Z\right) \equiv \int_0^Z \delta\Omega(Z')dZ'$ is the solution for absolutely rigid chains, Eq. (2.13). For short sequences the accumulated mistake is statistically small, i.e. $\delta\phi_r\left(z \lesssim \lambda_c\right) \approx 0$. On such length scale we put $\cos\left[\delta\phi_r\left(z\right)\right] \approx 1$ in Eq. (B.1). For long sequences $\delta\phi_r\left(z\right)$ may take any values and this approximation fails.

The solution of resulting equation, which does not depend on the sequence length L, is

$$\delta\Phi(Z) = -\frac{1}{2\kappa}\int_0^\infty dZ' e^{-\kappa|Z-Z'|}\sin\left[\delta\phi_r\left(Z'\right)\right], \tag{B.2}$$

where we kept $\sin[\delta\phi_r]$-term that reflects the nonlinearity of Eq. (2.9). The average energy is

$$\begin{aligned}\left\langle E(L)\right\rangle &\approx h\int\limits_0^{L/h}\left(a_0 - a_1\left\langle\cos\delta\phi_r(Z)\right\rangle\right)dZ + ha_1\kappa^2\int\limits_0^{L/h}\left(\left\langle\delta\Phi(Z)\sin\delta\phi_r(Z)\right\rangle + \tfrac{1}{2}\left\langle\left(\delta\Phi'(Z)\right)^2\right\rangle\right)dZ \\ &\equiv \left\langle E_r(L)\right\rangle + \left\langle E_\kappa(L)\right\rangle,\end{aligned} \tag{B.3}$$

where we have used that $\kappa^2\delta\Phi \ll 1$, but do not expand $\sin[\delta\phi_r]$ and $\cos\left[\delta\phi_r\right]$. The first integral is the energy of infinitely rigid DNAs, $\langle E_r(L)\rangle$. The second integral, $\langle E_\kappa(L)\rangle$, is the correction to it due to finite DNA torsional rigidity. The first term in $\langle E_\kappa(L)\rangle$ is negative as energy gain of adjustment of torsionally non-rigid chains to minimum of $\varepsilon_{\rm int}$, whereas the second one is positive as the torsional energy of mismatch between $\delta\phi_r\left(Z\right)$ and corrected solution $\delta\phi\left(Z\right) = \delta\phi_r\left(Z\right) + \kappa^2\delta\Phi\left(Z,\kappa\right)$. In total $\langle E_\kappa(L)\rangle$ should be negative, since torsionally non-rigid chains have lower energy than the rigid ones.

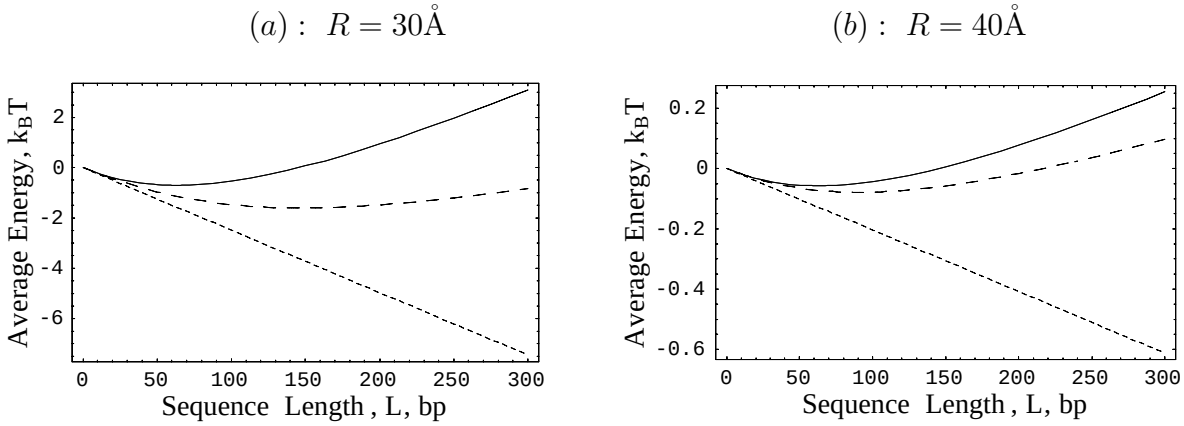

Figure B.1: *The average energy decreases when the chains are not infinitely rigid.* Long-dashed: $\langle E\left(L\right)\rangle$, solid: $\langle E_r\left(L\right)\rangle$, dotted: homologous sequences. Parameters of Fig. 2.3b. The approximation $\kappa^2\delta\Phi\left(Z\right) \ll 1$ is questionable at $R = 30$Å (a). At larger R, DNA-DNA interaction perturbs the twist pattern weaker, chains become effective more rigid and the approximation is valid (b).

Further we use "Gaussian averaging theorem", $\langle \exp\left[i(\delta\phi_r(Z) - \delta\phi_r(Z'))\right]\rangle = \exp\left[-\frac{1}{2}\left\langle\left(\delta\phi_r(Z) - \delta\phi_r(Z')\right)^2\right\rangle\right]$, since $\delta\phi_r(Z)$ can be represented as a sum of delta-correlated quantities. After cumbersome but straightforward integration we get

$$\langle E_\kappa(L)\rangle = -a_1 h\left\{\frac{l}{2(\tau+\kappa)} - \frac{\kappa+3\tau}{8\tau(\kappa+\tau)^2} + \frac{e^{-(\kappa+\tau)l}}{8\tau(\kappa+\tau)^2}\left[\kappa+3\tau-(\kappa^2-\tau^2)\frac{1-e^{(\kappa-3\tau)l}}{\kappa-3\tau}\right]\right\} + \frac{a_1 h\kappa^2}{16}\left\{\frac{2\gamma l}{\kappa(\kappa+\tau)^2} + \frac{-2\tau^3-10\tau^2\kappa-3\kappa^2\tau+\kappa^3}{2\kappa^2(\kappa+\tau)^2(\kappa+2\tau)}\left(1-e^{-(\kappa+\tau)l}\right) + \frac{e^{-(\kappa+\tau)l}}{2(\kappa-2\tau)}\left[\frac{1-e^{(\kappa-3\tau)l}}{\kappa-3\tau}\frac{\kappa-8\tau}{(\kappa+\tau)(\kappa+2\tau)} - \frac{1-e^{-(\kappa-\tau)l}}{\kappa-\tau}\frac{2\tau}{\kappa^2}\right]\right\}, \tag{B.4}$$

where $l = L/h$, $\tau = h/\lambda_c$.

For very short sequences ($L \ll \lambda_t, \lambda_c$), $\langle E_\kappa(L)\rangle \approx \frac{a_1\tau\kappa L(1-8\gamma l+3\kappa l)}{8(\kappa+\tau)(\kappa+2\tau)}$ can be positive. We suppose that this wrong result comes from approximations used for getting expression (B.3). It occurs for typical parameters at $L \lesssim 5$bp. For long sequences ($L \gg \lambda_t, \lambda_c$) the energy correction, $\langle E_\kappa(L)\rangle \approx -\frac{3a_1\kappa L}{8\tau}$. The proportionality coefficient is, as expected, the reciprocal torsional length, $\kappa = h/\lambda_t$.

Solution (B.4) is valid if $\left\langle\left(\kappa^2\delta\Phi\left(Z\right)\right)^2\right\rangle \ll 1$. The upper estimate, when unity is used for $\sin[\delta\phi_r]$ in (B.2), is $\left\langle\left(\kappa^2\delta\Phi\left(Z\right)\right)^2\right\rangle \simeq \left(2-e^{-\kappa Z}\right)^2/4$. The exact calculation gives a smaller value: for long sequences $\left\langle\left(\kappa^2\Phi\left(Z\right)\right)^2\right\rangle \approx \frac{\kappa}{4(\kappa+\tau)}$. Thus, for $\lambda_c = 300$Å which corresponds to real DNA, $\lambda_t \gtrsim 1000$Å satisfy this approximation, i.e. it is applicable for *very rigid* chains.

Although $\kappa^2\delta\Phi\left(Z\right) \ll 1$, $\langle E\left(L\right)\rangle$ may differ substantially from $\langle E_r\left(L\right)\rangle$, since this adjustment is integrated over the whole length of the sequence. Nevertheless, DNAs with long λ_t repel each other (Fig. B.1). Note that the approximation $\cos\left[\delta\phi_r\left(z\right)\right] \approx 1$ does *may not* work for long sequences, and we can use the calculated energy correction $\langle E_\kappa\left(L \gtrsim \lambda_c\right)\rangle$ only as an extrapolation.

Bibliography

[1] J.D. Watson and F.H.C. Crick, A structure of Deoxyribose nucleic acid, Nature, 171 737-738 (1953).

[2] W. Saenger, Principles of nucleic acid structure, Springer-Verlag, NY (1984).

[3] G.S. Manning, Limiting laws and counterion condensation in polyelectrolyte solutions, Biophys. Chem., 7 95-102 (1977)

[4] J. Ray and G.S. Manning, Counterion and coion distribution functions in the counterion condensation theory, Macromol., 32 4588-4595 (1999); Counterion condensation revisited, J. Biol. Str. Dyn., 16 461-476 (1998) and references therein.

[5] F. Oosawa, Polyelectrolytes, Marcel Dekker, NY (1971).

[6] G.S. Manning, The molecular theory of polyelectrolyte solution with applications to the properties of polynucleotides, Quart. Rev. Biophys., II 179-246 (1978).

[7] M.D. Frank-Kamenetskii, V.V. Anshelevich, and A.V. Lukashin, Polyelectrolyte model of DNA, Sov. Phys. Usp., 30 317-330 (1987).

[8] D. Stigter, Evaluation of the counterion condensation theory of polyelectrolytes, Biophys. J., 69, 380-388 (1995).

[9] L. McFail-Isom, C.C. Sines, and L.D. Williams, DNA structure: cations in charge?, Curr. Opin. Str. Biol., 9, 298-304 (1999).

[10] L. McFail-Isom, X.Shui, and L.D. Williams, Divalent cations stabilize unstacked conformations of DNA and RNA by interacting with base Π systems, Biochemistry, 37 17105-17111 (1998).

[11] X. Shui, C.C. Sines, L. McFail-Isom, D. VanDerveer, and L.D. Williams, Structure of the Potassium Form of CGCGAATTCGCG: DNA Deformation by Electrostatic Collapse around Inorganic Cations, Biochemistry, 37 16877-16887 (1998).

[12] X. Shui, L. McFail-Isom, G.G. Hu, and L.D. Williams, The *B*-DNA Dodecamer at High Resolution Reveals a Spine of Water on Sodium, Biochemistry, 37 8341-8355 (1998).

[13] K. Luger, A.M. Mäder, R.K. Richmond, D.F. Sangent, and T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8Å resolution, Nature, 389 251-260 (1997).

[14] J. Granot, J. Feigon, and D.R. Kearns, Interaction of DNA with divalent metal ions: I. ^{31}P-NMR studies, Biopolymers, 21 181-201 (1982); J. Granot and D.R. Kearns, II. Proton relaxation enchancement studies, 21 203-218 (1982); III. J. Granot and D.R. Kearns, II. Extent of metal

binding: experiment and theory, 21 219-232 (1982); IV. Competitive studies of Mn^{2+} binding to AT- and GC-rich DNAs, 21 873-883 (1982).

[15] H.R. Drew and R.E. Dickerson, Structure of *B*-DNA Dodecamer, J. Mol. Biol., 151 535-556 (1981).

[16] W.H. Braunlin and Q. Xu, Hexaamminecobalt (III) binding environments on double-helical DNA, Biopolymers, 32 1703-1711 (1992).

[17] J.C.G. Montoro and J.L.F. Abascal, Ionic distribution around simple *B*-DNA models II. Deviations from cylindrical symmetry, J. Chem .Phys., 109 6200-6210 (1998).

[18] N. Korolev, A.P. Lyubartsev, A. Rupperecht, and L. Nordenskiöld, Competitive Binding of Mg^{2+}, Ca^{2+}, Na^{+}, and K^{+} Ions to DNA in Oriented DNA Fibers: Experimental and Monte-Carlo Simulation Results, Biophys. J., 77 2736-2749 (1999).

[19] P.J. Heath and J.M. Schurr, Counterion condensation: Effects of site binding, fluctuations in nearest-neighbor interactions, and bending, Macromolecules, 25 4149-4159 (1992).

[20] H.H. Strey, V.A. Parsegian, and R. Podgornik, Equation of state for polymer liquid crystals: Theory and Experiment, Phys. Rev. E, 59 999-1008 (1999); S.M. Lindsay, S.A. Lee, J.W. Powell, T. Weidlich, C. Demarco, G.D. Lewen, N.J. Tao, A. Rupprecht, The origin of the A to B transition in DNA fibers and films, Biopolymers, 27 1015-1043 (1988).

[21] R. Podgornik, H.H. Strey, K. Gawrisch, D.C. Rau, A. Rupprecht, and V.A. Parsegian, Bond orientational order, molecular motion, and free energy of high-density DNA mesophases, Proc. Natl. Acad. Sci. U.S.A., 93 4261-4266 (1996).

[22] S.A. Rice, M. Nasagawa, and H. Morawetz, Polyelectrolyte Solutions, Academic Press, London-NY, 1961.

[23] A. Katchalsky, Polyelectrolytes, Pure Appl. Chem., 26 327-373 (1971).

[24] R.M. Fuoss, A. Katchalsky, and S. Lifson, The potential of an infinite rod-like molecule and the distribution of the counterions, Proc. Nat. Acad. Sci. USA, 37, 579-589 (1951); A. Katchalsky and S. Lifson, The electrostatic free energy of polyelectrolyte solution, J. Pol. Sci., XIII, 43-55 (1954).

[25] T. Alfrey, Jr., P.W. Berg, and H. Morawetz, The counterion distribution in solution of rod-like polyelectrolytes, J. Pol. Sci., VII, 543-547 (1951).

[26] M. Le Bret and B.H. Zimm, Monte-Carlo determination of the distribution of ions about a cylindrical polyelectrolyte, Biopolymers, 23 271-285, (1984); Distribution of counterions around a cylindrical polyelectrolyte and Manning condensation theory, Biopolymers, 23 287-312 (1984)

[27] M.D. Frank-Kamenetskii, How DNA helix breathes, Nature, 328 17-18 (1987).

[28] A.W. Adamson and A.P. Gast, Physical Chemistry of Surfaces, Sixth Edition, NY, Wiley-Interscience Publication, (1997).

[29] J.N. Israelachvili, Intermolecular and surface forces, Academic Press, NY, 1992.

[30] L.C. Gosule and J.A. Schellman, Compact form of DNA induced by spermidine, Nature, 259 333-335 (1976).

[31] J. Widom and R.L. Baldwin, Cation-induced toroidal condensation of DNA, J. Mol. Biol., 144 431-453 (1980).

[32] V.A. Bloomfield, Condensation of DNA by multivalent cations: consideration on mechanism, Biopolymers, 31 1471-1481 (1991).

[33] P.G. Arscott, J.R. Wenner, and V.A. Bloomfield, DNA condensation by cobalt hexaammine(III) in alkohol-water mixtures: dielectric constant and other solvent effects, Biopolymers, 35 345-364 (1995).

[34] G.E. Plum, P.G. Arscott, and V.A. Bloomfield, Condensation of DNA by trivalent cations. 2. Effect of cation structure, Biopolymers, 30 631-643 (1990).

[35] J.X. Tang and P.A. Janmey, The polyelectrolyte nature of F-actin and the mechanism of F-actin bundle formation, J. Biol. Chem., 271 8556-8563 (1996); J.X. Tang, T. Ito, T. Tao, P. Traub, and P.A. Janmey, Opposite effects of electrostatics and steric exclusion on bundle formation by F-actin and other filamentous polyelectrolytes, Biochemistry, 36 12600-12607 (1997), and references therein.

[36] J.X. Tang, S.E. Wong, P.T. Tran, and P.A. Jamney, Counterion-induced bundle formation of rodlike polyelectrolyte, Berichte der Bunsen-Gesellschaft- Phys.Chem.Chem.Phys., 100 796-806 (1996).

[37] V.A. Bloomfield, DNA condensation, Curr. Opin. Struct. Biol., 6 334-341 (1996).

[38] C. Ma and V.A. Bloomfield, Gel electrophoresis measurement of counterion condensation on DNA, Biopolymers, 35 211-216 (1995).

[39] J. Pelta, D. Durand, J. Doucet, and F. Livolant, DNA mesophases induced by spermidine: structural properties and biological implications, Biophys. J., 71 48-63 (1996).

[40] R.W. Wilson and V.A. Bloomfield, Counterion-induced condensation of DNA. A light-scattering study, Biochemistry, 18 2192-2196 (1979).

[41] D.K. Chattoraj, L.C. Gosule, and J.A. Schellman, DNA condensation with polyamines: 2. Electron-microscopic studies, J. Mol. Biol., 121 327-337 (1978).

[42] V.A. Bloomfield, DNA condensation by multivalent cations, Biopolymers, 98 269-282 (1998) and references therein.

[43] J. Pelta, F. Livolant, and J-L. Sikorav, DNA aggregation induced by polyamines and cobalthexammine, J. Biol. Chem., 271 5656-5662 (1996).

[44] E. Raspaund, M. Olvera de la Cruz, J.-L. Sikorav, and F. Livolant, Precipitation of DNA by polyamines: a polyelectrolyte behavior, Biophys. J., 74 381-393 (1998).

[45] C.G. Baumann, S.B. Smith, V.A. Bloomfield, and C. Bustamante, Ionic effects on the elasticity of single DNA molecules, Proc. Natl. Acad. Sci. U.S.A., 94 6185-6190 (1997).

[46] Y. Yoshikawa and K.Yoshikawa, Diaminoalkanes with as odd number of carbons atoms induce compaction of a single double-stranded DNA chain, FEBS Lett., 361 277-281 (1995).

[47] J.A. Schellman and N. Parthasarathy, X-ray diffraction studies on cations-collapsed DNA, J. Mol. Biol., 175 313-329 (1984).

[48] C. Ma and V.A. Bloomfield, Condensation of supercoiled DNA induced by $MnCl_2$, Biopys. J., 67 1678-1681 (1994).

[49] D.C. Rau and V.A. Parsegian, Direct measurement of temperature-dependence solvation forces between DNA double helices, Biophys. J., 61 260-271 (1992).

[50] S. Flock, R. Labarbe, and C. Houssier, Osmotic effectors and DNA structure: effect of glycine on precipitation of DNA by multivalent cations, J. Biomol. Str. Dyn., 13 87-102 (1995).

[51] N. Hud, F. Milanovich, and R. Balhorn, Evidence of novel secondary structure in DNA-bound protamine is revealed by Raman spectroscopy, Biochemistry, 33 7528-7535 (1994).

[52] H.A. Tajmir-Riahi, M. Naoui M, and R. Ahmad, The effects of cobalt hexammine and cobalt pentammine cations on the solution structure of calf-thymus DNA. DNA condensation and structural features studied by FT1R difference spectroscopy, J. Biomol. Struct. Dyn., 11 83-93.(1993).

[53] J.G. Duguid, J.M. Benevides, V.A. Bloomfield, and G.J. Thomas Jr., Raman Spectroscopy of DNA-metal complexes: I. Interactions and conformational effects of the divalent cations: Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Pd^{2+}, and Cd^{2+}, Biophys. J., 65 1916-1928 (1993).

[54] C. Ma, L. Sun, and V.A. Bloomfield, Condensation of plasmids enhanced by Z-DNA conformation of $d(CG)_n$ inserts, Biochemistry, 34 3521-3528 (1995).

[55] J.H. van de Sande, L.P. McIntosh, and T.M.Jovin, Mn^{2+} and other transitions metals at low concentrations induce the right-to-left helical transformation of poly[d(G-C)], EMBO J., 1 777-782 (1982) and references therein.

[56] J. Widom and R.L. Baldmin, Monomolecular condensation of λ-DNA by cobalt hexammine, Biopolymers, 22, 1595-1620 (1983).

[57] I. Koltover, K. Wagner, and C.R. Safinya, DNA condensation in two dimensions, Proc. Natl. Acad. Sci. USA, 97 14046-14051 (2000).

[58] S. Leikin, In "Hydration processes in biology", Ed.: M.-C. Bellisent-Funel, Theory of electrostatic and hydration interactions between helical macromolecules, pp.323-332, IOS-Press (1999).

[59] S. Leikin, R.P. Rand, D.C. Rau, and V.A. Parsegian, Hydration forces, Annu. Rev. Phys. Chem., 44 369-395 (1993).

[60] S. Leikin, In "Hydration processes in biology", Ed.: M.-C. Bellisent-Funel, Hydration forces between biological macromolecules: measurements, speculations, and established facts, pp.313-321, IOS-Press (1999).

[61] R. Podgornik, H.H. Strey D.C. Rau, and V.A. Parsegian, Watching molecules crowd: DNA double helices under osmotic stress, Biophys. Chem., 57 111-121 (1995).

[62] D.C. Rau, B.K. Lee, and V.A. Parsegian, Measurement of the repulsive force between polyelectrolyte molecules in ionic solution– hydration forces between parallel DNA double helices, Proc. Natl. Acad. Sci. U.S.A., 81 2621 -2625 (1984).

[63] H.H. Strey, V.A. Parsegian, and R. Podgornik, Equation of state for DNA liquid crystals: fluctuation enhanced electrostatic double layer repulsion, Phys. Rev. Lett., 78 895-898 (1997).

[64] D.C. Rau and V.A. Parsegian, Direct measurement of the intermolecular forces between counterion-condensed DNA double helices, Biophys. J., 61 246-259 (1992).

[65] A.P. Lyubartsev and L. Nordenskiöld, Monte-Carlo simulation study of ion correlation and osmotic pressure in hexagonally oriented DNA, J. Phys. Chem., 99 10373-10382 (1995).

[66] S. Leikin, D.C. Rau, and V.A. Parsegian, Measured entropy and enthalpy of hydration as a function of distance between DNA double helices, Phys. Rev. A, 44 5272-5278 (1991).

[67] I.M. Lifshitz, A.Y. Grosberg, and A.R. Khokhlov, Some problems of the statistical physics of polymer chains with volume interaction, Rev. Mod. Phys., 50 683-713 (1978); A.Y. Grosberg, and A.R. Khokhlov, Statistical Physics of Macromolecules, AIP, New York, (1994).

[68] L. Onsager, The effect of the shape on the interaction of colloidal particles, Ann. NY Acad. Sci., 627-659 (1949); A. Isihara, Theory of anisotropic colloidal solutions, J. Chem. Phys., 19 1142-1147 (1951).

[69] P.J. Flory, Statistical thermodynamics of semi-flexible chain molecules, Proc. Royal Soc. Lon. A Math., 234 60-73 (1956); Phase equilibria in solutions of rod-like particles, ibidem., 234 73-89 (1956).

[70] A.Y. Grosberg, and A.V. Zhestkov, On the compact form of linear duplex DNA: Globular states of the uniform elastic (persistent) macromolecule. J. Biomolec. Struct. Dynam. 3: 859-872 (1986).

[71] J.-L. Barrat and J.-F. Joanny, Theory of polyelectrolyte solutions, In Advances in Chemical Physics, Vol. XCIV. I. Prigogine and S. A. Rice, editors. John Wiley and Sons, New York, (1996).

[72] M. Olvera de la Cruz, L. Belloni, M. Delsanti, J.P. Dalbiez, O. Spalla, and M. Drifford, Precipitation of highly charged polyelectrolyte solutions in the presence of multivalent salts. J. Chem. Phys., 103 5781–5791 (1995).

[73] I. Rouzina and V.A. Bloomfield, Macroions attraction due to electrostatic correlation between screening counterions: I. Mobile surface-adsorbed ions and diffuse ion cloud, J. Phys. Chem., 100 9977-9989 (1996).

[74] N. Grønbech-Jensen, R.J. Mashl, R.F. Bruinsma, and W.M. Belbart, Counterion-induced attraction between rigid polyelectrolyte, Phys. Rev. Lett, 78 2477-2480 (1997).

[75] B.I. Shklovskii, Wigner crystal model of counterion induced bundle formation of rigid polyelectrolyte, Phys. Rev. Lett, 82 3268-3271 (1999).

[76] M.J. Stevens and M.O. Robins, Density functional theory of ionic screening: when do like charges attract?, Europhys. Lett., 12 91 (1990).

[77] S. Marčelja, Electrostatics of membrane adhesion, Biophys. J., 61 1117-1121 (1992); R. Kjellander, T.Akesson, B. Jonsson, and S. Marčelja, Double-layer interactions in monovalent and divalent electrolytes: A comparison of the anisotropic hypernetted chain theory and Monte-Carlo simulations, J. Chem. Phys., 97 1424-1431 (1992) and references therein.

[78] B.-Y. Ha and A.J. Liu, Counterion-mediated attraction between two like-charged rods, Phys. Rev. Lett., 79 1289-1292 (1997).

[79] B.-Y. Ha and A.J. Liu, Effect of Non-pairwise-additive interactions on bundles of rodlike polyelectrolytes, Phys. Rev. Lett., 81 1011-1014 (1998).

[80] B.-Y. Ha and A.J. Liu, Counterion-mediated, non-pairwise-additive attractions in bundles of like-charged rods, Phys. Rev. E, 60 803-813 (1999).

[81] R. Podgornik and V.A. Parsegian, Charge-Fluctuation Forces between Rodlike Polyelectrolytes: Pairwise Summability Reexamined, Phys. Rev. Lett., 80 1560-1564 (1998).

[82] J. Ray and G.S. Manning, An attractive force between two rodlike polyions mediated by the sharing of condensed counterions, Langmuir, 10 2450-2461 (1994).

[83] M.A. Young, B. Jayaram, and D.L. Beveridge, Intrusion of counterions into the spine of hydration of *B*-DNA: fractional occupancy of electronegative pockets, J. Am. Chem. Soc., 119 59-69 (1997).

[84] A.A. Kornyshev and S. Leikin, Theory of interaction between helical molecules, J. Chem. Phys., 107 3656-3674 (1997) ; Erratum: ibidem, 108 7035(E) (1998).

[85] A.A. Kornyshev and S. Leikin, Electrostatic zipper motif for DNA aggregation, Phys. Rev. Lett., 82 4138-4141 (1999).

[86] A.G. Cherstvy, A.A. Kornyshev, and S. Leikin, Temperature-dependent DNA condensation triggered by re-arrangement of adsorbed cations, submitted to J. Phys. Chem. B.

[87] R.M. Clement, J. Sturm, and M.P. Daune, Interaction of metallic cations with DNA: IV. Specific binding of Mg^{2+} and Mn^{2+}, Biopolymers, 12 405-421 (1981).

[88] R. van Steenwinkel, F. Campagnari, and M. Merlini, Interaction of Mn^{2+} with DNA as studied by proton-relaxation enchancement of solvent water, Biopolymers, 20 915-923 (1981).

[89] G.L. Eichorn and Y.Ae Shin, Interaction of metal ions with polynucleotides and related compounds: XII. the relative effect of various metal ions on DNA helicity, J. Am. Chem. Soc., 90 7323-7328 (1968).

[90] S. Leikin, D.C. Rau, and V.A. Parsegian, Temperature-favored assembly of collagen is driven by hydrophobic not hydrophilic interactions, Nature (Structural Biology), 2 205-210 (1995).

[91] S. Leikin, M.M. Kozlov, N.L. Fuller, R.P. Rand, Measured effects of diacylglycerol on structural and elastic properties of phospholipid membranes, Biophys. J., 71 2623-2632 (1996).

[92] D.C. Rau and V.A. Parsegian, Direct measurement of forces between linear polysaccharides xanthan and schizophyllan, Science, 249 1278-1281 (1990).

[93] C. Bonnet-Gonnet, S. Leikin, S. Chi, D.C. Rau, and V.A. Parsegian, Measurement of forces between hydroxypropylcellulose polymers: temperature favored assembly and salt exclusion, J. Phys. Chem., 105 1877-1886 (2001).

[94] S. Leikin, D.C. Rau, and V.A. Parsegian, Direct measurement of forces between self-assembled proteins: temperature dependent exponential forces between collagen triple helices, Proc. Nat. Acad. Sci. U.S.A., 91 276-280 (1994).

[95] T.A. Gill, J.K. Chan, K.F. Prochareon, and A.T. Paulson, Effect of salt concentration and temperature on heat-induced aggregation and gelation of fish myosin, Food Res. Intern., 25 333-341 (1992).

[96] M.A. Lauffer, Entropy-driven processes in biology, Springer, Berlin, 1975.

[97] Yu.Ya. Gurevich and Yu.I. Kharkats, Theory of the diffuse part of the double layer in solid electrolytes, Sov. Phys. Dokl., 21 642-645 (1976); A.A. Kornyshev and M.A. Vorotyntsev, Conductivity and space charge phenomena in solid electrolytes with one mobile charge carrier species, a review with original material, Electrochimica Acta, 26 303-323 (1981).

[98] I. Rouzina and V.A. Bloomfield, Competitive electrostatic binding of charged ligands to polyelectrolytes: planar and cylindrical geometries, J. Phys. Chem., 100 4292-4304 (1996).

[99] Yu.I. Kharkats, Bi-transitions in two-level disordered systems, Dokl. Akad. Nauk SSSR, 235 82-85 (1977) [Sov. Phys. Dokl., 22 385-387 (1977)].

[100] M.E. Fisher and Y. Levin, Criticality in Ionic Fluids: Debye-Hückel Theory, Bjerrum, and beyond, Phys. Rev. Lett., 71, 3826-3829 (1993).

[101] D. M. Zuckerman, M. E. Fisher, and B.P. Lee, Critique of primitive model electrolyte theories, Phys. Rev. E, 56 6569-6580 (1997).

[102] A.A. Kornyshev and S. Leikin, Electrostatic interaction between long, rigid helical macromolecules at all interaxial angles, Phys. Rev. E, 62 2576-2596 (2000) and references therein.

[103] B.P. Lee and M.E. Fisher, Density fluctuations in an electrolyte from generalized Debye-Hückel theory, Phys. Rev. Lett., 76 2906-2909 (1996); M. E. Fisher, The story of coulombic criticality, J. Stat. Phys., 75 1-36 (1994).

[104] G. Cevc, in Water and Biological Macromolecules, eds.: E. Westhof, Chemical Rubber, Boca Raton, FL, (1993), p. 338; R. Kjellander, On the image-charge model for the hydration force, J. Chem. Soc., Faraday Trans., 80 1323-1348 (1984); A.A. Kornyshev and S. Leikin, Fluctuation theory of hydration forces: the dramatic effect of inhomogeneous boundary conditions, Phys. Rev. A, 40 6431-6437 (1989); S. Leikin and A.A. Kornyshev, Theory of hydration forces – nonlocal electrostatic interaction of neutral surfaces, J. Chem. Phys., 92 6890-6898 (1990); S.Marcelja and N.Radic, Repulsion of interfaces due to boundary water, Chem. Phys. Lett., 42 129-130 (1976).

[105] S. Leikin and A.A. Kornyshev, The mean-field theory of dehydration transitions, Phys. Rev. A, 44 1156-1168 (1991).

[106] A.A. Kornyshev and S. Leikin, Fluctuation theory of hydration forces: the effect of inhomogeneous boundary conditions, Phys. Rev. A, 40 6431-6437 (1989).

[107] A.A. Kornyshev and S. Leikin, Symmetry laws for interaction between helical macromolecules, Biophys. J., 75 2513-2519 (1998).

[108] D. Rhodes and A. Klug, Helical periodicity of DNA determined by enzyme digestion, Nature, 286 573-579 (1980).

[109] S.B. Zimmerman and B.H. Pheiffer, Helical parameters of DNA do not change when DNA fibers are wetted: X-ray diffraction study, Proc. Natl. Acad. Sci. U.S.A., 76 2703-2709 (1979).

[110] P. Mariani and L. Saturni, Measurements of intercolumnar forces between parallel guanosine four-stranded helices, Biophys. J., 70 2867-2874 (1996).

[111] P. Mariani, F. Cluchi, and L. Saturni, Helix-specific interactions induce condensation of guanosine four-stranded helices in concentrated salt solutions, Biophys. J., 74 430-435 (1998).

[112] A.A. Kornyshev and S. Leikin, Electrostatic interaction between helical macromolecules in dense aggregates: An impetus for DNA poly- and meso-morphism, Proc. Natl. Acad. Sci. U.S.A., 95 13579-13584 (1998).

[113] A.A. Kornyshev, S. Leikin, S.V. Malinin, Chiral electrostatic interaction and cholesteric liquid crystals of DNA, Eur.Phys.J. E, 7 83-93 (2002).

[114] P.L. Hansen, R. Podgornik, and V.A. Parsegian, Osmotic properties of DNA, Phys. Rev. E, 64 0219071-0219074 (2001).

[115] S. Lifson, Improved approximation for the potential of spherical electrolyte molecules in solution, J. Chem. Phys., 27 700-701 (1957).

[116] CRC Handbook of Chemistry and Physics, 75th ed., ed. D.R.Lide, CRC Press, Boca Raton, 1994.

[117] R.A. Robinson and R.H. Stokes, Electrolyte solutions, London, (1959).

[118] E.J.W. Verwey and J.Th.G. Overbeek, Theory of the stability of lyophobic colloids, Elsevier (1948).

[119] B.V. Derjagin, N.B. Churaev, and V.M. Muller, Surface forces, Moskva, Nauka, (1987) [in Rissian].

[120] J. Lyklema, in "Adsorption from solution at the solid/liquid interface", Eds: G.D. Parfitt, C.H.Rochester, p.239, (1983).

[121] L.D. Landau and E.M. Lifshitz, Statistical Physics, Pergamon, New-York, (1978).

[122] G. Lamm and G.R. Pack, Calculation of dielectric constant near polyelectrolytes in solution, J. Phys. Chem., 101 959-965 (1997).

[123] J.G. Duguid, V.A. Bloomfield, J.M. Benevides, and G.J. Thomas Jr., Raman Spectroscopy of DNA-metal complexes. II. The thermal denaturation of DNA in the presence of Sr^{2+}, Ba^{2+},Ca^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, and Cd^{2+}, Biophys. J., 69 2623-2641 (1995).

[124] J.G. Duguid and V.A. Bloomfield, Aggregation of melted DNA by divalent metal ion-mediated cross-linking, Biophys. J., 69 2642-2648, (1995); D. Matulis, J. Rouzina, and V.A. Bloomfield, Thermodynamics of DNA binding and condensation, J. Mol. Biol., 296 1053-1063 (2000).

[125] J.H. Shibata and J.M. Schurr, A theory of aggregation in the thermal denaturation region of multistrand biopolymers, Biopolymers, 20 525-549 (1981).

[126] A.A. Kornyshev and S. Leikin, Sequence recognition in the pairing of DNA duplexes, Phys. Rev. Lett., 86 3666-3669 (2001).

[127] S. Leikin and V.A. Parsegian, Temperature-induced complementarity as a mechanism for biomolecular assembly, Proteins, 19 73-76 (1994).

[128] S. Leikin, V.A. Parsegian, W.H. Yang, and G.E. Walfaren, Raman spectral evidence for hydration forces between collagen triple helices, Proc. Natl. Acad. Sci., U.S.A., 94 11312-11317 (1997).

[129] A.A. Kornyshev, D.A. Kossakowski, and S. Leikin, Landau theory of a system with two bilinearly coupled order parameters in external field: exact mean-field solution, critical properties and isothermal susceptibility, Z. Naturforsch., 50a, 789-794 (1995).

[130] M. Eigen and E. Wicke, The thermodynamics of electrolytes at higher concentrations, J. Phys. Chem., 58 702-714 (1954).

[131] I. Borukhov, D. Andelman, and H. Orland, Steric effects in Electrolytes: a modified Poisson-Boltzmann equation, Phys. Rev. Lett., 79 435-438 (1997).

[132] J. Granot, Effect of finite ionic size on the solution of the Poisson-Boltzmann equation: application to the binding of divalent metal ions to DNA, Biopolymers, 22 1831-1841 (1983).

[133] D. Poland and H.A. Sheraga, Theory of helix-coil transition in biopolymers, Academic Press, NY-London, (1970).

[134] Yu.S. Lasurkin, M.D. Frank-Kamenetskii, and E.N. Trifonov, Melting of DNA: its study and application as a research method, Biopolymers, 9 1253-1306 (1970)

[135] B.H. Zimm and S.A.Rice, Helix-coil transition in charged macromolecules, Mol. Phys., 3 391 (1960).

[136] D.A. Knoll, M.G. Fried, and V.A. Bloomfield, "Structure and Expression: DNA and its Drug Complexes", Eds: M.H.Sarma, R.H.Sarma, Adenine Press, New York, v.2, pp. 123-146, (1988)

[137] H. Richard, J.P. Schreiber, and M. Daune, Interactions of metallic ions with DNA. V. DNA renaturation mechanism in the presence of Cu^{2+}, Biopolymers., 12 1-10 (1973).

[138] A.P. Yurgaitis and Yu.S. Lasurkin, Mechanism of DNA denaturation in the presence of manganese ions, Biopolymers, 20 967-975 (1981); Yu.P. Blagoi, V.A. Sorokin, and G.O. Gladchenko, Biopolymers, 22 1641-1656 (1983).

[139] J. Stilck, Y. Levin, and J. Arenzon, Thermodynamic properties of a simple model of like-charged attracting rods, J. Stat. Phys. 106, 287-299 (2002).

[140] 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 3105-3108 (2000).

[141] V. Lorman, R. Podgornik, and B. Žekš, Positional, reorientational, and bond orientational order in DNA mesophases, Phys. Rev. Lett., 87 218101-1–218101-4 (2001).

[142] H.M. Harreis, A.A. Kornyshev, C.N. Likos, H. Löwen, and G. Sutmann, Phase Behavior of Columnar DNA Assemblies, Phys. Rev. Lett., 89 018303-1–018303-4 (2002).

[143] V.L. Berezinskiĭ, Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group, Sov. Phys. JETP., 32 493-500 (1971); F. Wegner, Spin-ordering in a planar classical Heisenberg model, Z. Phys., 206 465-470 (1967).

[144] R. Langridge, H.R. Wilson, C.W. Hooper, M.H.F. Wilkins, and L.D.Hamilton The molecular configuration of deoxyribonucleic acid: I. X-ray diffraction study of a crystalline form of the lithium salt, J. Mol. Biol., 2 19-37 (1960).

[145] D.A. Marvin, M. Spencer, M.H.F. Wilkins, and L.D. Hamilton, The molecular configuration of deoxyribonucleic acid: I. X-ray diffraction study of the C form of the lithium salt, J. Mol. Biol., 3 547-565 (1961).

[146] W. Fuller, M.H.F. Wilkins, H.R. Wilson, and L.D. Hamilton, The molecular configuration of deoxyribonucleic acid: X-ray diffraction study of the A form, J. Mol. Biol.,12 60-80 (1965).

[147] M.L. Kopka, A. Fratani, H.R. Drew, and R.E. Dickerson, Ordered water-structure around a beta-DNA dodecamer: A quantitative study, J. Mol. Biol., 163, 129-146 (1983); J.Barciszewski, J. Jurczak, S. Porowski, T. Specht, V.A. Erdmann, The role of water structure in conformational changes of nucleic acids in ambient and high pressure conditions, Eur. J. Biochem., 260 293-307 (1999).

[148] A.A. Kornyshev, S. Leikin, and G. Sutmann, "Overscreening" in a polar liquid as a result of coupling between polarization and density fluctuations, Electrochim. Acta, 42 849-865 (1997).

[149] R. Holliday, A mechanism of gene conversion in fungi, Genet. Res., Camb., 5 282-304 (1964); R. Holliday, Genetic recombination in fungi, In Replication and Recombination of Genetic Material, Eds.: W.J. Peacock, R.D. Brock, Canberra, Australian Academy of Science, (1968).

[150] G.R. Smith, Homologous recombination near and far from DNA breaks: Alternative roles and contrasting views, Annu. Rev. Genet., 35 243-274 (2001).

[151] R.G. Sargent, M.A. Brenneman, and J.H. Wilson, Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination, Mol. Cell. Biol., 17 267-277 (1997) and references therein.

[152] K.M. Vasquez, K. Marburger, Z. Intody, and J.H. Wilson, Manipulating the mammalian genome by homologous recombination, Proc. Natl. Acad. Sci. U.S.A., 98 8403-8410 (2001).

[153] P. Hasty, J. Riveraperez, C. Chang, and A. Bregley, Target frequency and interaction pattern for insertion and replacement vectors in embryonic stem-cells, Mol. Cell. Biol., 11, 4509-4517 (1991); P. Hasty, J. Riveraperez, and A. Bregley, The length of homology required for gene targeting in embryonic stem-cells, Mol. Cell. Biol., 11 5586-5591 (1991).

[154] Y.S. Rong and K.G. Golic, Gene targeting by homologous recombination in Drosophila, Science, 288 2013-2018 (2000).

[155] B.S. Singer, L. Gold, P. Gauss, and D.H. Doherty, Determination of the amount of homology required for recombination of bacteriophage T4, Cell, 31 25-33 (1982).

[156] V.M. Watt, C.J. Ingles, M.S. Urdea, and W.J. Ritter, Homology requirements for recombination of Escherichia coli, Proc. Natl. Acad. Sci. U.S.A. 82, 4768-4772 (1985).

[157] P. Shen and H.V. Huang, Homologous recombination in Escherichia coli: dependence on substrate length and homology, Genetics, 112 441-457 (1986).

[158] J. Rubnitz and S. Subramani, The minimum amount of homology required for homologous recombination in mammalian cells, Mol. Cell. Biol., 4 2253-2258 (1984).

[159] B. Lewin, Genes IV, Oxford, 1997.

[160] Y. Fujitani and I. Kobayashi, Random walk model for homologous recombination, Phys. Rev. E, 52 6607-6622 (1995).

[161] Y. Fujitani, K. Yamamoto, and I. Kobayashi, Dependence of the frequency of homologous recombination on the homology length, Genetics, 140 797 (1995).

[162] S.M. Burgess, N. Kleckner, and B.M. Weiner, Somatic pairing of homologs in budding yeast: existence and modulation, Genes Dev., 13 1627-1641 (1999).

[163] B.M. Weiner and N. Kleckner, Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast, Cell, 77 977-991 (1994).

[164] R.R. Sinden, DNA structure and function, Academic Press, San-Diego, (1994).

[165] S. Neidle, DNA Structure and recognition, IRL Press at Oxford University Press, Oxford, (1994).

[166] A.A. Gorin, V.B. Zhurkin, and W.K. Olson, *B*-DNA twisting correlates with base-pair morphology, J. Mol. Biol., 247 34-48 (1995).

[167] W. Kabsch, C. Sander, and E.N. Trifonov, The 10 helical twist angles of *B*-DNA, Nucl. Acid. Res., 10 1097-1104 (1982).

[168] W.K. Olson, A.A. Gorin, X.-J. Lu, L.M. Hock, and V.B. Zhurkin, DNA sequence-dependent deformability deduced from protein–DNA crystal complexes, Proc. Natl. Acad. Sci. U.S.A., 95, 11163-11168 (1998).

[169] R.E. Dickerson and H.R. Drew, Structure of *B*-DNA dodecamer. II. Influence of base pair sequence on DNA structure, J. Mol. Biol., 149 761-786 (1981).

[170] DNA need not unzip, Phys. Rev. Focus (http://focus.aps.org/v7/st19.html).

[171] J.C. Wang, Helical repeat of DNA in solution, Proc. Natl. Acad. Sci. U.S.A., 76 200-203 (1979).

[172] D.M. Crothers, J. Drak, J.D. Kahn, and S.D .Levene, DNA bending flexibility, and helical repeat by cyclization kinetics, in "Methods in Enzymology", edited by D.M.J. Lilley and J.E. Dahlberg (Academic Press, San Diego, 1992), Vol. 212 B, p. 3, and references therein.

[173] M.D. Frank-Kamenetskii, A.V. Lukashin, V.V. Anshelevich, and A.V. Vologodskii, Torsional and bending rigidity of the double helix from data on small DNA rings, J. Biomol. Str. Dyn., 2 1005-1012 (1985).

[174] A.G. Cherstvy, A.A. Kornyshev, and S.Leikin, Effect of torsional elasticity on the interaction and recognition of DNA duplexes, in preparation.

[175] D. Shore and R. Baldwin, Energetics of DNA twisting. I. Relation between twist and cyclization probability, J. Mol. Biol., 170 957-981, (1983).

[176] D.S. Horowitz, J.C. Wang, Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling, J. Mol. Biol., 173 75-91 (1984).

[177] H. Risken, The Fokker-Planck Equation, Springer, Berlin, (1989).

[178] F.G. Bass, Yu.S. Kivshar, V.V. Konotop, and Yu.A. Sinitsyn, Dynamics of solitons under random perturbations, Phys. Rep., 157 63-181 (1988).

[179] F. Abdullaev, Theory of solitons in inhomogeneous media, Wiley, (1994).

[180] B.V. Chirikov, A universal instability of many-dimensional oscillator systems, Phys. Rep, 52 265-379 (1979);

[181] Yu.S. Kivshar and B.A. Malomed, Dynamics of solitons in nearly integrable systems, Rev. Mod. Phys., 61 763-915 (1989).

[182] D. Baeriswyl and A.R. Bishop, Phase fluctuations, disorder and nonlinearity in one dimension: I. The linear limit, J. Phys. C.: Solid St.Phys., 13 1403-1418 (1980).

[183] H. Reisinger and F. Schwalb, Defects in the Sine-Gordon model: Statics, Z. Phys. B, 52 151-170 (1983); H. Schmidt and F. Schwalb, Localized defects in a one-dimensional Ginzburg-Landau model, Z. Phys. B, 30 197-210 (1978).

[184] "Solitons in condensed matter", Eds.: A.R. Bishop and T.Schneider, Springer, 1978.

[185] A.Vl. Gurevich, S.L. Leikin, and R.G.Mints, Localization of nonlinear waves in randomly inhomogeneous media, Phys. Lett., 105A 31-33 (1984).

[186] O.M. Braun and Yu.S. Kivshar, Nonlinear dynamics of the Frenkel-Kontorova model, Phys. Rep., 306 1-108 (1998).

[187] G.M. Zaslavskii and V.A. Chirikov, Stochastic instability of non-linear oscillations, Sov. Phys. Usp., 14 548-568 (1972).

[188] J.E. Yansey-Wrona and R.D. Camerini-Otero, The search for DNA homology does not limit stable homologous pairing promoted by RecA protein, Curr. Biol., 5 1149-1158 (1995).

[189] P. Hsieh, C.S. Camerini-Otero, and R.D. Camerini-Otero, The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA, Proc. Natl. Acad. Sci., U.S.A., 89 6492-6496 (1992).

[190] R.D. Camerini-Otero and P. Hsieh, Homologous recombination proteins in prokaryotes and eukaryotes, Annu. Rev. Genet., 29 509–552 (1995).

[191] K.M. Kosikov, A.A. Gorin, V.B. Zhurkin, and W.K. Olson, DNA stretching and compression: large-scale simulations of double helical structures, J. Mol. Biol., 289 1301-1326 (1999).

[192] S.C. Satchwell, H.R. Drew, and A.A. Travers, Sequence periodicities in chicken nucleosome core DNA, J. Mol. Biol., 191 659-675 (1986).

[193] A. Stasiak, E. Di Capua, The helicity of DNA in complexes with RecA protein, Nature, 299 185-186 (1982); A. Stasiak, E. Di Capua, and T. Koller, Elongation of duplex DNA by RecA protein, J. Mol. Biol., 151 557-564 (1981).

[194] M. Salerno, Discrete model for DNA-protomer dynamics, Phys. Rev. A, 44 5292-5297 (1991).

[195] A.A. Kornyshev, unpublished results.

[196] A.R. Bishop, Nonlinear collective excitations in superionic conductors, J. Phys. C., 11 L329-L335 (1978); K.M. Leung and A.R. Bishop, Thermodynamics of the classical, easy-plane ferromagnetic chain with in-plane magnetic field, J. Phys. C., 16 5893-5898 (1978).

[197] T.A. Kontorova and Ya.I. Frenkel, On the theory of plastic deformation and twinning, Zh. Eksper. Teor.Fiz., 8 1340-1348 (1938). [in Russian]

[198] I.F. Lyuksyutov, A.G. Naumovets, and V.L. Pokrovsky, Two-Dimensional Crystals, Academic Press, Boston (1992).

[199] L.J. Peck and J.C. Wang, Sequence dependence of the helical repeat of DNA in solution, Science, 292 375-378 (1981); D. Rhodes and A. Klug, Sequence-dependent helical periodicity of DNA, Science, 292 378-380 (1981).

[200] M. Büttiker and R. Landauer, Nucleation theory of overdamped soliton motion, Phys. Rev. A, 23 1397-1410 (1981).