Structural and Elastic Properties of DNA and Chromatin

Dissertation
zur Erlangung des Grades

“Doktor der Naturwissenschaften”

am Fachbereich Physik
der Johannes Gutenberg-Universität
in Mainz

Boris Mergell
geb. in Neustadt an der Weinstraße

Tag der mündlichen Prüfung: 6.10.2003

Abstract

In this thesis the elastic properties of DNA are investigated on various length scales using MC and scaling arguments. Three different system are considered: a generic model for double-stranded polymers, a mesoscopic model for DNA at the base-pair level, and a crossed-linker model of the chromatin fiber.

We investigate a chain consisting of two coupled worm-like chains with constant distance between the strands. The effects due to double-strandedness of the chain are studied. In a previous analytical study of this system an intrinsic twist-stretch coupling and a tendency of kinking is predicted. Even though a local twist structure is observed the predicted features are not recovered.

A new model for DNA at the base-pair level is presented. The base-pairs are treated as flat rigid ellipsoids and the sugar-phosphate backbones are represented as stiff harmonic springs. The base-pair stacking interaction is modeled by a variant of the Gay-Berne potential. This simple mesoscopic model of DNA combines ideas of the stack-of-plates model and the worm-like chain model. The resulting helical conformations are studied using energy minimization and Monte-Carlo simulations. It is shown by systematic coarse-graining how the elastic constants of a worm-like chain are related to the local fluctuations of the base-pair step parameters. Even though a lot of microscopic details of the base-pair geometry is neglected the model can be optimized to obtain a B-DNA conformation as ground state and reasonable elastic properties. The anisotropy of the bending angles is just a consequence of the plate-like shape of the base-pairs and the twist-stretch coupling is the result of the preferred stacking of neighboring base-pairs and the rigid backbones. Therefore these properties are included in the model. Moreover the model allows to simulate much larger length scales than it is possible with atomistic simulations due to the simplification of the force-field and in particular due to the possibility of non-local Monte-Carlo moves. As a first application the behavior under stretching is investigated. In agreement with micromanipulation experiments on single DNA molecules one observes a force-plateau in the force-extension curves corresponding to an overstretching transition from B-DNA to a so-called S-DNA state. The model suggests a structure for S-DNA with highly inclined base-pairs in order to enable at least partial base-pair stacking.

Finally a simple model for chromatin is introduced to study its structural and elastic properties. The underlying geometry of the modeled fiber is based on a crossed-linker model. The chromatosomes are treated as disk-like objects. Excluded volume and short range nucleosomal interaction are taken into account by a variant of the Gay-Berne potential. Under consideration is the influence of the nucleosomal interactions on elastic and structural properties of the fiber. It is found that the bending rigidity and the stretching modulus of the fiber increase with more compact fibers. For a reasonable parameterization of the fiber for physiological conditions and sufficiently high attraction between the nucleosomes a force-extension curve is found similar to stretching experiments on single chromatin fibers. For very small stretching forces a kinked fiber forming a loop is observed. If larger forces are applied the loop formation is stretched out and a decondensation of the fiber takes place.

Zusammenfassung

In dieser Arbeit werden elastische und strukturelle Eigenschaften von DNA auf verschiedenen Längenskalen untersucht. Dafür werden drei verschiedene Systeme mit Hilfe von Monte-Carlo Simulationen und Skalenargumenten näher betrachtet: ein generisches Modell für Doppelstrangmoleküle, ein mesoscopisches Modell für DNA und ein 'Crossed-Linker'-Modell für Chromatin.

Auswirkungen der Doppelstrangstruktur einer Kette, die aus zwei gekoppelten wurmartigen Ketten mit konstantem Abstand besteht, werden ermittelt. Obwohl lokal helikale Strukturen auftreten, können die in einer früheren analytischen Arbeit vorausgesagten Eigenschaften wie eine Kopplung zwischen Torsion und Streckung oder eine Tendenz zu Konformationen mit Knicken nicht bestätigt werden.

Es wird ein neues Modell für DNA auf der Basenpaarebene vorgestellt. Die Basenpaare werden als flache, steife Ellipsoide behandelt. Das Zuckerphosphatrückgrat ist durch steife harmonische Federn repräsentiert. Die Basenpaarstapelwechselwirkungen werden mit einem modifizierten Gay-Berne-Potential modelliert. Dieses einfache mesoskopische Modell verbindet die Ideen so genannter 'Stack-of-Plates'-Modelle und des Modells wurmartiger Ketten. Mit Hilfe von Energieminimierung und Monte-Carlo-Simulationen werden die gewonnen helikalen Strukturen untersucht. Durch systematisches 'Coarse-Graining' werden die elastischen Konstanten des Modells wurmartiger Ketten als Funktion der lokalen Fluktuationen der Basenpaarschrittparameter berechnet. Obwohl einige mikroskopische Details der Basenpaargeometrie vernachlässigt werden, kann das Modell so optimiert werden, daß man eine B-DNA-Konformation als Grundzustand und vernünftige elastische Eigenschaften erhält. Andererseits sind beispielsweise mikroskopische Details wie die Anisotropie der Biegewinkel, die eine Folge der oblaten Form der Basenpaare ist, oder die Verdrehungs-Strekkungs-Kopplung, bedingt durch das bevorzugte Stapeln benachbarter Basenpaare und das steife Rückgrat, in dem Modell enthalten. Verglichen mit atomistischen Computersimulationen ergeben sich folgende Vorteile. Erstens ist das involvierte Kraftfeld einfacher und die Energieberechung somit weniger zeitaufwendig, und zweitens ist man nicht auf lokale Monte-Carlo-Schritte angewiesen. Daraus ergibt sich die Möglichkeit, wesentlich längere Ketten zu simulieren. Als eine erste Anwendung des Modells wird das Verhalten der DNA unter Spannung untersucht. In Übereinstimmung mit Streckexperimenten an einzelnen DNA-Molekülen beobachtet man ein Kraftplateau in der Spannungs-Dehnungs-Kurve, bei dem sich die DNA einem strukturellen Übergang von B-DNA zur so genannten S-DNA unterzieht. Innerhalb des Modells findet man eine S-DNA-Struktur mit stark geneigten Basenpaaren, so daß die Basenpaare wenigsten teilweise übereinander liegen können.

Schließlich wird ein einfaches Modell für die Chromatinfaser eingeführt. Die zugrundeliegende Geometrie der modellierten Faser basiert auf einem so genannten 'Crossed-Linker'-Modell. Das Chromatosom wird als Zylinder behandelt. Volumenausschluß und kurzreichweitige Nukleosomenwechselwirkung sind durch ein modifiziertes Gay-Berne-Potential berücksichtigt. Untersucht wird der Einfluß der Nukleosomenwechselwirkung auf elastische und strukturelle Eigenschaften der Chromatinfaser. Dabei stellt man fest, daß die Biegesteifigkeit und der Streckmodul der Faser wächst, je kompakter die Faser ist. Für vernünftige Modellparameter der Faser, mit denen wichtige experimentell bestimmte Observablen in physiologischen Bedingungen wiedergeben werden können, und ausreichend starker Anziehung der Chromatosomen erhält man eine globale Strukturänderung, wenn an der Faser gezogen wird. Für sehr kleine Zugkräfte wird eine Schleifenkonformation beobachtet, bei der die Faser geknickt ist. Für stärkere Kräfte öffnet sich diese Struktur in einem ersten Schritt gefolgt von einer Dekondensation der Faser.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

DNA

Since the discovery of the double-stranded helical structure of DNA by Watson and Crick [WC53] enormous progress was made in synthesis, sequencing, and manipulation techniques. Nowadays it is possible to synthesize special DNA sequences, it is known how to determine the base-pair sequence of a particular DNA chain, and how to insert genes. Physicist were actively involved in these developments. Besides its biological role it is interesting to work on DNA since it comes closest to the theorist's notion of a polymer chain.

DNA structure

DNA is a polymer composed of only four types of subunits. These are the deoxyribonucleotides containing the bases adenine (A), cytosine (C), guanine (G), and thymine (T), a sugar (deoxyribose), and a negatively charged phosphate group. The nucleotides are linked together by covalent phosphodiester bonds [ABL+94]. 1953 Watson and Crick [WC53] discovered the spatial structure of DNA analyzing X-ray diffraction patterns of DNA fibers. It turned out that DNA is a double-stranded helix where the bases are on the inside of the helix and the sugar phosphates are on the outside. This requires specific base-pairing via hydrogen bonds between a large purine base (A or C) on one strand and a small pyrimidine base (T or C) on the other chain. The base-pair sequence is usually referred to as the primary structure of DNA.

Depending on the sequence, i.e. details of the chemical structure, and the external conditions

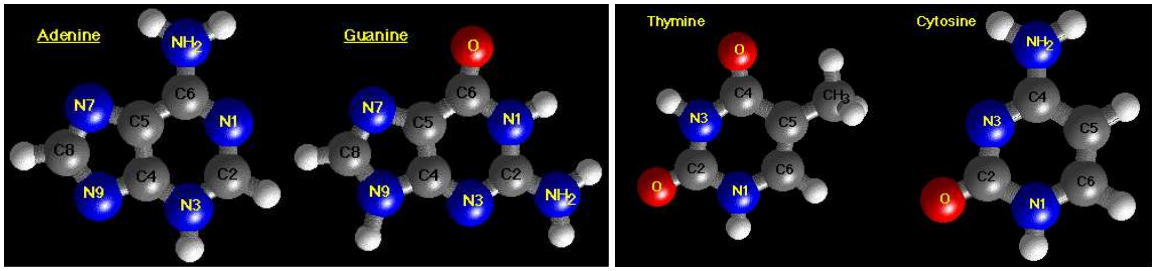


Figure 1.1: Chemical structure of the four DNA bases. Taken from [Hal].

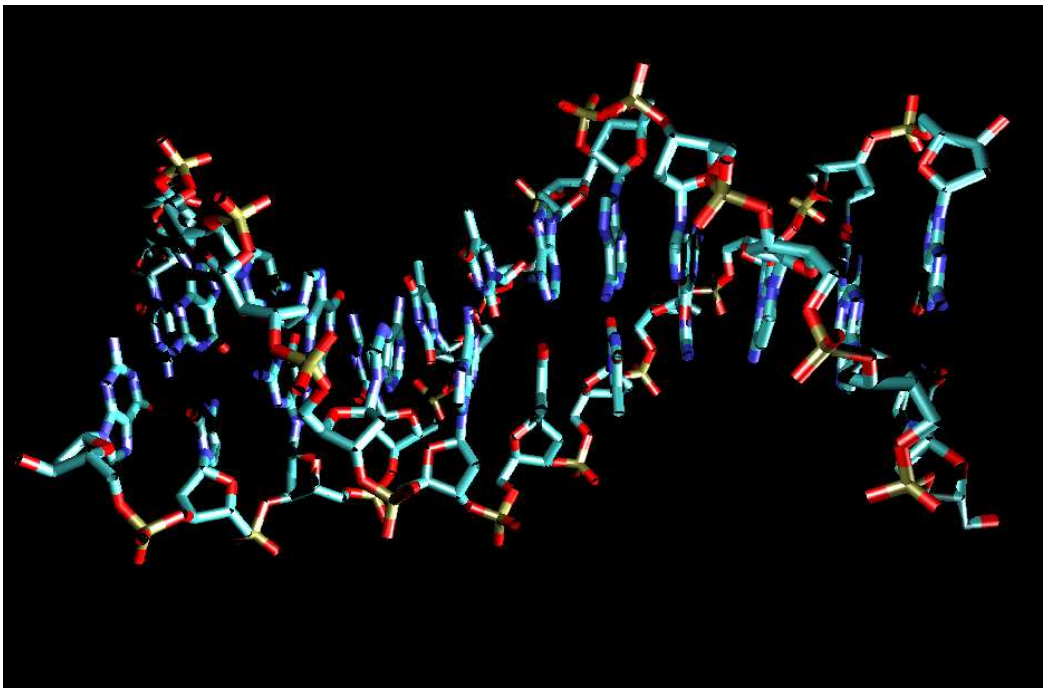


Figure 1.2: Structure of a B-DNA configuration.

DNA can adopt different conformations which vary in helical repeat length and handedness of the corresponding helix (this is the so-called secondary structure). Under physiological conditions the DNA is found in the so-called B-form, a right-handed double-stranded helix with a helical repeat length of about 10.4bp/34nm, a helix diameter of about 2nm and a base(-pair) thickness of about 0.34nm.

Imposing a torsion on the DNA the double-stranded helix can adopt particular (tertiary) structures. Either the helix forms a toroidal structure called solenoidal super-helix or it forms an interwound structure called plectonemic super-helix. This interwound structure is well-known as an annoyance in everyday life such as the snarling of the telephone cords.

What is interesting about DNA?

DNA carries the genetic code which is read by protein complexes. During transcription the DNA molecule gets highly deformed due to mechanical forces exerted by these enzymes. Therefore the mechanical properties of DNA play an important role for its functionality in the cell. Besides that the DNA (about 1m) is compacted via several organization steps in order to fit into the cell nucleus (with a diameter of about μm).

DNA is a rather stiff polymer with a persistence length of about 50nm in contrast to most synthetic polymers with persistence lengths of the order of 1nm. The elastic properties of DNA, that is the resistance to bending and twisting, as well as the base-pair sequence are of biological interest. The wrapping of the DNA around the histone core (protein complex composed of eight histone octamers), which is the first organization step of the DNA folding into the cell nucleus, is a competition of the energy that must be paid to bend the DNA and the energy that is gained due to electrostatic

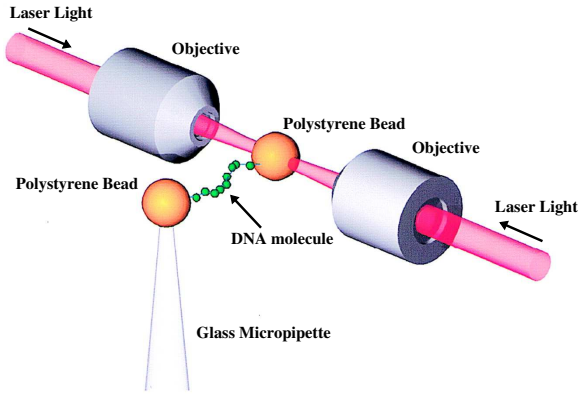
attraction of the negatively charged sugar phosphate backbone and positively charged histone tails. During cellular processes such as replication or transcription the DNA is locally highly deformed by enzymes in order to open the double-stranded helix and to read the genetic information. It is known that the genetic information in DNA determines not only the amino acid sequences of encoded proteins and RNA but also the geometry and deformability of DNA at the base-pair level. For example, there is evidence that the TATA-box sequence, important in the initiation of DNA transcription, may achieve this function by virtue of its enhanced flexibility [JCL+96, DMK99].

Investigating the structure and elasticity of DNA on various length scales

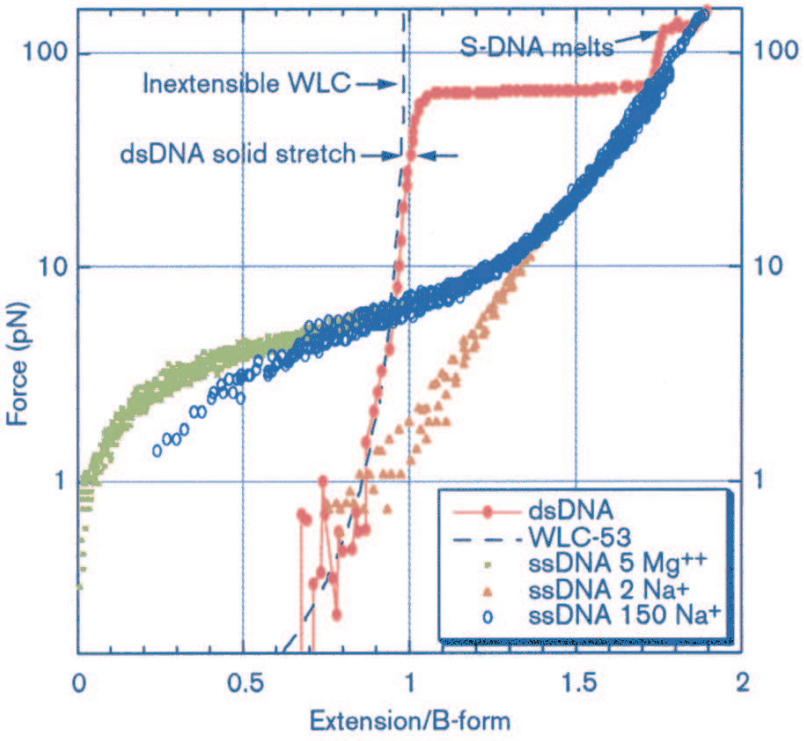
In recent years a lot of experimental techniques have been developed for the investigation of the elasticity and structure of DNA on various length scales. X-ray diffraction patterns of fibers or single crystals of DNA oligomers have led to a detailed picture of possible DNA conformations with atomistic resolution [Dic92, DDC+82]. It has turned out to be useful to discuss these conformations in terms of the relative position and orientation of adjacent bases and base-pairs [BPO94, HC95, EHL97, LS88] according to certain guidelines [DBC+89]. This provides a mechanical interpretation of the biological function of particular sequences [CD99].

Various optical methods such as time-dependent fluorescence depolarization [MRZ82, HCFS96] or electron microscopy [BWC90, BFK+95] helped to quantify the bending and torsional persistence length of DNA. In time-dependent fluorescence depolarization studies the decay of the fluorescence depolarization anisotropy of a DNA-ethidium-bromide complex is investigated. Ethidium bromide is a hydrophobic molecule of roughly the same size as the base-pairs that fluoresces green. It is used as intercalating agent since it likes to slip between two base-pairs. The connection between the fluorescence depolarization experiments and the torsion and bending dynamics of DNA has been analyzed in detail by Barkley and Zimm [BZ79]. Cryo-electron [BFK+95] microscopy makes it possible to directly observe the path of DNA reportedly without any adsorption and drying artifacts known from classical electron microscopy (EM). The reconstructed contour of several DNA fragments can then be used to calculate the bending persistence length. EM studies are also used to deduce an average structure for supercoiled circular DNA in solution [BWC90]. Furthermore cyclization experiments [VV02, PMSV00] which measure the probability that short DNA segments form loops [MMK96, SY84] can be used as a probe for the local shape, bending and torsional rigidity of the DNA.

An interesting development of the last decade are nanomechanical manipulation techniques on single DNA molecules [BSLS00, LLA+02, BBS03]. These methods offer the opportunity to study the mechanical response of a single DNA molecule to external forces and torques, to measure the forces generated in biochemical reactions such as the DNA-polymerase reaction [WSY+00] and to measure the typical forces necessary to pull two DNA strands apart [ERBH97, BERH97, BERH98]. They have also been used to measure the relaxation dynamics of a stretched molecule by hydrodynamic flow when the flow stopped [PSLC95, PSC97]. To manipulate a single molecule the molecule needs to be caught by its extremities. Different strategies have been used: magnetic beads [SFB92, SAB+96, SABC98, SCB98, SAB+99, SBC99], glass needles [SFB92, SCB96, ERBH97, BERH97, BERH98], optical traps [BMSS94], and AFM [CSRTG00]. External forces and/or torques can deform DNA and induce new structures [CLH+96, ABLC98]. The force-extension curves of Ref. [SCB96] show that under a critical stretching force of about 65pN torsionally relaxed double-stranded DNA (dsDNA) molecules undergo a structural transition to an overstretched S-DNA conformation. The increase in length amounts to 1.8 of its B-form. Strick *et al.* [SAB+96] developed



(a) Sketch of an experimental micromanipulation set-up where one latex bead is caught in an optical trap (optical tweezer). Taken out of [CB00].



(b) Force-extension measurements on ssDNA and dsDNA. Taken out of [BSLS00].

Figure 1.3: Micromanipulation set-up and measured stress-strain curves.

a technique which allows to constrain the DNA torsionally. They found for positively supercoiled DNA stretched by 3 pN a new highly twisted structure called P-DNA. The structure of S-DNA and P-DNA is still under discussion [BSLS00, LLA$^+$02]. Essevaz-Roulet *et al.* [ERBH97] have separately attached the two complementary strands of a DNA double helix to a glass slide and a glass micro-needle. Displacing the slide away from the needle opens the double helix. They measured forces of about 10pN to open an AT sequence and 15pN to open a CG sequence. In principle it is possible to determine the sequence of the DNA chain by the use of this method. But it turned out that due to thermal fluctuations only variations of the AT versus GC content on a scale of 100 bases could be detected. An overview over the development of micromanipulation techniques and the results can be found in [BBS03].

Interpretation of the experimental results

The experiments are usually rationalized within the framework of two models. (i) So-called stack-of-plates models are used to compactify the X-ray diffraction data. The bases are treated as plates. Attaching a coordinate frame to each base the configurations can be analyzed in terms of base-pair and base-pair step variables [DBC$^+$89, BO94, BPO94, EHL97, HC97a]. In this way a detailed picture of possible DNA configuration of specific DNA sequences is obtained. O'Hern *et al.* [OKLN98] developed an elasticity theory in terms of small deviations of the translational and rotational degrees of freedom of the base-pairs and calculated the twist-stretch coupling for three different ground states. Manning and Maddocks [MMK96] investigated sequence effects on cyclization rates of short DNA pieces using the unstressed shape of the underlying sequence and its elastic constants. (ii) On large length scales, that is in the low stretching force regime, DNA behaves effectively as a rigid rod. Therefore it is convenient to treat the DNA as a worm-like chain with a certain resistance to bending and torsion [MS94, MS95c, MS95a, MS95b]. The theory is in good agreement with experimental force-extension data up to forces of 20pN. Later extensions of the worm-like chain model accounted for effects due to double-strandedness [LGK98, EBK95], for the anisotropy of the bending angles [PR00a, PR00b], for twist-stretch coupling [KLNO97, Mar97, Nel98] and for topological effects [PR01] to study the elastic response of the chain depending on the linking number. Zhou *et al.* [HYZc99, YHZc00, ZZOY00] proposed a model based on bending and base-stacking interactions in an attempt to provide a unified framework to understand DNA elasticity for small as well as for large stretching forces beyond the overstretching transition [ZL01]. Their results agree quite well with experiments up to very large forces. Marko [Mar97, Mar98] incorporated the overstretching transition by including extra terms into the worm-like chain Hamiltonian which take into account the high cooperativity of the transition [CLH$^+$96] and define a stable S-DNA state. Storm and Nelson [SN02b, SN02a] recently proposed a two state model for the elasticity of nicked, double-stranded DNA. It allows coexistence of two conformational states (B- and S-DNA) of the DNA, each with its

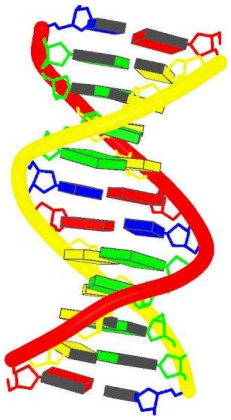


Figure 1.4: B-DNA configuration. Bases are represented as color-coded plates. Taken from [GAL].

own bending and stretching elastic constants. Fitting the model to an overstretched data set yields values for the persistence length and the stretching modulus of each state as well as the energy which is necessary to flip a B-DNA segment into the S-state.

Atomistic computer simulations have been used to predict possible DNA conformations for overstretched S-DNA [CLH$^+$96, SAB$^+$99, LL99] or stretched and overwound P-DNA [ABLC98] and to investigate electrostatic effects on DNA structure [HMIWW00, HWW02, HSH03]. Furthermore they allow for the calculation of all internal coupling constants of bases and base-pairs [LSHL00]. The formulation of a mesoscopic model raises the question which atoms should be combined to one monomer unit. To address this question the fluctuations of interatomic distances were measured in atomistic computer simulations [BFLG99]. If the mean squared fluctuations of the distance between two atoms are smaller than a certain threshold value it is assumed that they belong to the same rigid sub-unit. Depending on the threshold value a complete hierarchy of possible models can be formulated, e.g. treating the bases or the base-pairs and the backbone as rigid entities. Existing mesoscopic models on the base-pair level are based on a harmonic energy functional that incorporates the bending anisotropy and experimentally known correlations of base-pair step parameters such as twist, roll and slide [OZ00]. Simulation methods on larger length scales are essentially discretizations of the worm-like chain model with additional excluded volume and electrostatic interactions. Sequence effects and details of the chemical structure such as the bending anisotropy and the twist-stretch coupling are ignored [OZ00]. These low-resolution models are used to account for the topological properties of supercoiled DNA [SO92, CL94, CL96, VC95].

Chromatin

Chromatin structure

The cell nucleus contains DNA with a total contour length of about $1m$. Without any compaction the DNA chain would form a swollen coil of approximately $100\mu m$ diameter [Sch03] which is a factor of 100 larger than the cell diameter. In order to fit into the cell nucleus the DNA is compacted via several organization steps. The first structural level of this process is the formation of an array of nucleosomes. Eight histone octamers (two copies of H2A, H2B, H3 and H4) make a particle around which the DNA wraps in $1\frac{3}{4}$ turns [LMR$^+$97] due to electrostatic interactions between negatively charged DNA backbone and positively charged histone octamer units. Often the in- and outcoming linker DNA is glued together by so-called linker histones (H1 and H5). The nucleosome is the fundamental unit of the chromatin fiber. A nucleosome without linker DNA is called chromatosome. The nucleosomes in turn organize further into the chromatin fiber probably mediated by positively charged lysin rich histone tails [LMR$^+$97]. Bennink *et al.* [BPL$^+$01] studied the compaction rate of a DNA strand under stress in the presence of histones depending on the pulling force. They find an 8-fold compaction for moderate forces (1 pN). The folding hierarchy of DNA into the chromosome is illustrated in Fig. 1.5.

While the structure of individual core particles is established with atomistic resolution [LMR$^+$97] much less is known about the chromatin structure. In low salt concentration the fiber is in a

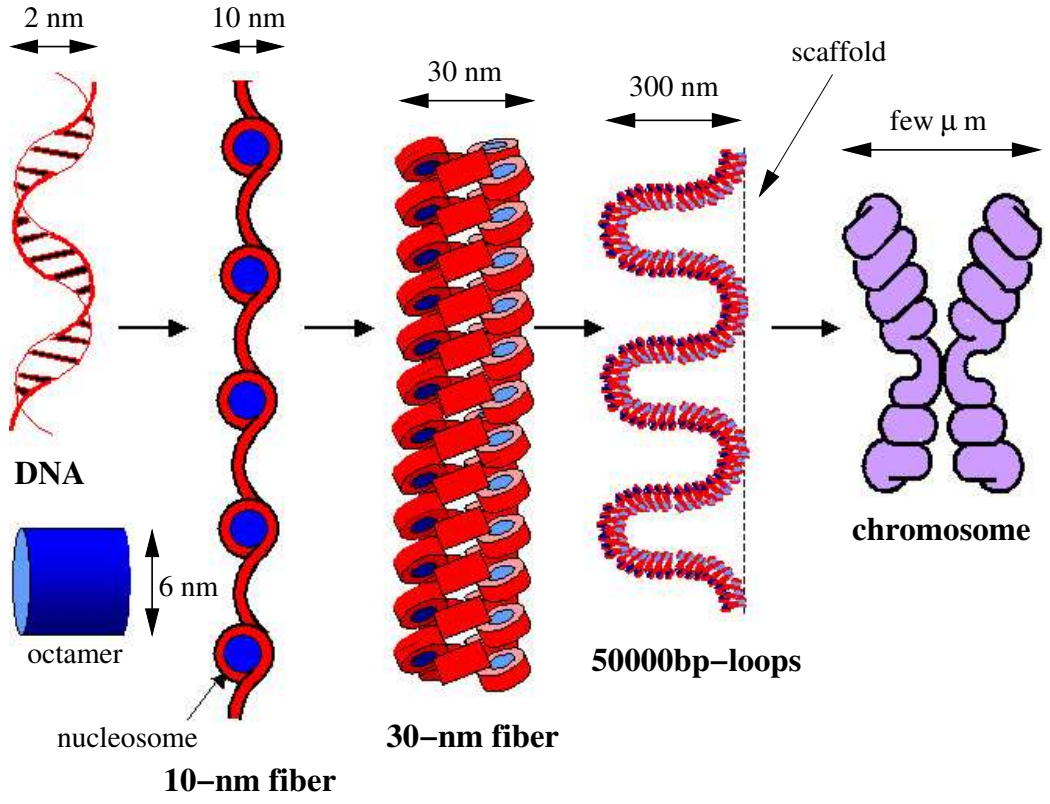


Figure 1.5: Organization steps of the DNA into the chromosome. Taken out of [Sch03].

swollen state forming a 'bead-on-a-string' structure referred to as the 10nm fiber [TKK79]. With increasing salt concentration the chromatin fiber starts to condense into more compact structures [GR87, BHDW98, BHG+98]. For very high salt concentrations close to physiological conditions the fiber forms a very dense fiber with a diameter of about 30nm.

There is still controversy about the structure of this so-called 30nm fiber [HZ95, HZ96]. The two competing classes of models are the solenoid models [FK76] and the crossed-linker models [WGHW93, SGB01]. In the solenoid models the position of successive nucleosomes form a helix and the linker DNA is assumed to be bent being able to connect neighboring nucleosomes. In the crossed-linker model the linker DNA is straight and connects nucleosomes on opposite sides of the fiber building up a three-dimensional zig-zag pattern. In principle it should be possible to distinguish between the two geometries using cryo-electron microscopy [BHDW98, BHG+98]. In fact, for low and intermediate salt concentrations zig-zag-like patterns are found supporting the crossed-linker models. But these methods cannot resolve the linker DNA geometry in physiological conditions such that the solenoid model cannot be excluded [HZ96].

Chromatin elasticity

The elastic properties of the chromatin fiber strongly depend on the underlying geometry. Stretching a solenoidal structure with a small helical pitch will cost much less energy than stretching a rather straight zig-zag structure. Furthermore the elastic properties of the chromatin fiber will be influenced by the state of compaction due to nucleosomal interactions. Nanomechanical manipulation techniques helped to gain deeper insights into the structural and elastic properties of the chromatin fiber. Cui and Bustamante [CB00] measured force-extension curves of single chromatin fibers in different salt concentration. They found a purely extensible worm-like chain behavior for small salt concentrations whereas for higher salt concentrations a force plateau occurred at roughly 5pN indicating a structural transition of the fiber. This force plateau is probably due to short-ranged nucleosome-nucleosome interactions which are not present in the more loose structures in low salt concentrations. Stretching the fiber with forces up to 25pN results in the release of the histones [BLL+01, BTSY+02].

The low salt concentration force-extension curves can be compared with elastic theories based on the solenoid and crossed-linker geometry. Schiessel *et al.* [SGB01, Sch03] and Ben-Haim *et al.* [BHLV01, BHLV02] calculated the mechanical response of the fiber to small stretching forces around the undistorted ground state assuming a crossed-linker geometry. They derived a relationship between the elastic constants of the fiber, the elastic constants of the linker DNA and the underlying geometry. When comparing their predictions to the experimental results it seems that the crossed-linker model is supported.

Computer simulations help to investigate effects due to excluded volume and short-ranged attractive interactions between the nucleosomes. Katritch *et al.* [KBO00] performed simulations including the linker DNA elasticity, excluded volume of the chromatosomes represented as spherical beads and a short-ranged attractive potential between chromatosomes. They investigated the mechanical response of the simulated fiber for varying linker length, entry-exit angle of the linker DNA and the bead diameter and compared their results to the experimental data of Cui and Bustamante [CB00]. It is shown that for a given fiber geometry a force plateau appears if the well depth of the attractive potential is increased. Wedemann and Langowski [WL02] introduced a model which treats the chromatosomes as disks with a radius of 10nm and a height of 6nm. The excluded volume and short-ranged attractive interactions of the nucleosomes are taken into account by a Gay-Berne

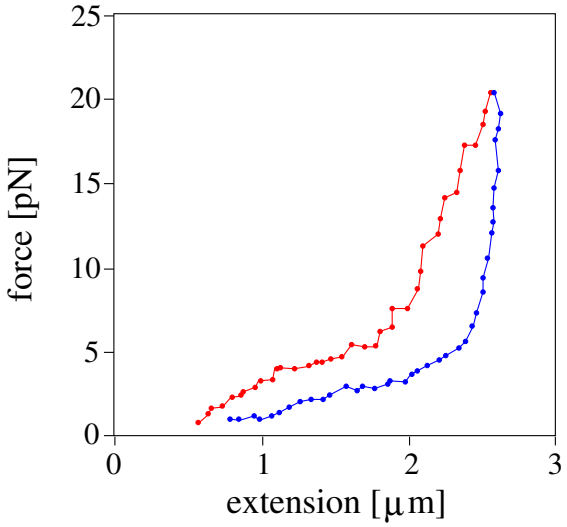


Figure 1.6: Force-extension curves for chicken erythrocyte chromatin fibers in 40mM NaCl. The red curve corresponds to the stretch cycle while the blue curve refers to the release cycle. Taken out of [CB00].

potential. Moreover the elastic and electrostatic energy of the DNA is included into the Hamiltonian. Structural properties such as the linear mass density and the persistence length of the fiber are studied. For model parameters corresponding to physiological conditions they found a linear mass density of about 6 nucleosomes$/11$nm in good agreement with experimental data [GR87, BHG$^+$98] and persistence lengths of about 200nm. Beard and Schlick [BS01] were interested in the folding mechanism of oligonucleosomes depending on the salt concentration. Therefore they included all 253 charges distributed over the surface of the chromatosome. The linker DNA is treated as a negatively charged elastic rod. For low salt concentrations the modeled trinucleosomes adopt disordered zig-zag conformations with rather straight linker DNA. High salt concentrations lead to condensation.

Outline of the thesis

This thesis is organized as follows. In chapter 2 we discuss structural and elastic properties of a system of two coupled worm-like chains. In particular we argue that there is neither a preference for kinking, nor a twist-stretch coupling, nor oscillations in the autocorrelation function of the bond-director field (pointing from one strand to the other) inherent in the system as predicted by an earlier study [LGK98]. In chapter 3 a generic model of DNA at the base-pair level is presented. We discuss structural and elastic properties as functions of the model parameters. To parameterize the model we demand that macroscopic variables such as the persistence length must be reproduced. The undistorted ground-state should match a B-DNA conformation. For the final parameterization of the model we perform simulations with an external stretching force. We also find an overstretching regime in the force-extension curves. The resulting S-DNA shows a skew ladder conformation. Moreover structural as well as elastic properties of the simulated chain below, at and beyond the critical force are investigated. Some preliminary ideas are presented in chapter 4 about how measured coupling constants from atomistic simulations can be mapped onto the effective coupling constants of a quadratic Hamiltonian on a larger length scale. In chapter 5 we study structural and elastic properties of chromatin. We calculate a structural phase diagram characterizing allowed and forbidden structures depending on the model parameters. We discuss the influence of the excluded volume interaction of the nucleosomes on the persistence length of the fiber. Furthermore we present some results on simulations performed with an additional stretching force.

Chapter 2

Elastic and structural properties of two coupled worm-like chains

We use computer simulations and scaling arguments to investigate statistical and structural properties of a semiflexible ribbon composed of isosceles triangles. We study two different models, one where the bending energy is calculated from the angles between the normal vectors of adjacent triangles, the second where the edges are viewed as semiflexible polymers so that the bending energy is related to the angles between the tangent vectors of next-nearest neighbor triangles. The first model can be solved exactly whereas the second is more involved. It was recently introduced by Liverpool and Golestanian [LGK98, GL00] as a model for double-stranded biopolymers such as DNA. Comparing observables such as the autocorrelation functions of the tangent vectors and the bond-director field, the probability distribution functions of the end-to-end distance, and the mean squared twist we confirm the existence of local twist correlation, but find no indications for other predicted features such as twist-stretch coupling, kinks, or oscillations in the autocorrelation function of the bond-director field [MEE02].

2.1 Introduction

A characteristic feature of many biopolymers is their high bending stiffness. Contour lengths of the order of μm and persistence lengths of the order of 50nm in the case of DNA even allow microscopy techniques to be used to directly observe their structure and dynamics [PSLC95, PSC97]. The model mostly used to interpret recent experimental data of micromechanical manipulations of single DNA chains [BMSS94, PSLC95, PSC97, SAB$^+$96, SCB98, CLH$^+$96] is that of the Kratky-Porod worm-like chain in which the polymer flexibility is determined by a single length, the persistence length l_p. Generalizations account for the chain helicity and coupling terms between bending, stretching, and twisting allowed by symmetry [Yam97, MS95c, MS94, MS95b, MS95a, MN97, KLNO97, PR01, PR00a, PR00b, KM97, MMK96]. All these continuum models of DNA neglect the double-stranded structure of DNA and one may ask, if this feature could not cause qualitatively different behavior.

The bending stiffness of single- and double-stranded DNA, for example, differs by a factor of 25 [FK90]. The simplest model which takes the double-strandedness into account is the railway-track model [EBK95] where two worm-like chains are coupled with harmonic springs. In two dimensions one finds drastical consequences: the bending fluctuations in the plane of the ribbon are strongly suppressed. The molecule becomes effectively stiffer on larger length scales. But the relevant

question is: what are the effects in three dimensions? Liverpool *et al.* [LGK98, GL00] investigated a version of the railway-track model in three dimensions where bending in the plane of the ribbon is forbidden by a constraint. Using analytical and simulation techniques they predict the existence of a low temperature regime where ribbons adopt a kink-rod structure due to a spontaneously appearing short-range twist structure resulting in an oscillatory behavior of the autocorrelation function of the bond-director field. Furthermore a twist-stretch coupling is predicted.

We study the discretized version of the simulation model of Liverpool *et al.* [LGK98, GL00] in the low temperature regime with the help of scaling arguments and MC simulations. In order to understand and to quantify the effects arising from the local twist structure of the Liverpool model we compare it with an analytically more tractable model where the bending stiffness is defined via the interaction of the normal vectors so that there is no tendency to form helical structures. Furthermore, we perform several MC simulation runs with an additional external force in order to test if the preferred buckling mechanism occurs via kinks.

2.2 Theoretical background

2.2.1 The worm-like chain model

The nanomechanical stretching experiments on single DNA molecules led to a revival of interest in the worm-like chain (WLC) model [SFB92, SCB96, MS95c]. The predicted stress-strain relations agree quite well with the experimental data of double-stranded DNA.

The elastic energy of a WLC with constant contour length L is given by the line integral:

$$\mathcal{H}_{WLC} = \frac{E}{2}\int_0^L ds \left(\frac{d^2\mathbf{r}(s)}{ds^2}\right)^2 = \frac{E}{2}\int_0^L ds \left(\frac{d\mathbf{t}(s)}{ds}\right)^2 = \frac{E}{2}\int_0^L ds\kappa(s)^2, \tag{2.1}$$

where s denotes arc length, $\mathbf{t}(s) = \frac{d}{ds}\mathbf{r}(s)$ denotes the unit tangent vector field, $\kappa(s)$ corresponds to the curvature of the chain, E is the bending modulus. Eq. (2.1) describes the resistance of the chain to bending. The tangent vector correlation function $\langle \mathbf{t}(s)\cdot\mathbf{t}(s')\rangle$ decays exponentially with correlation length l_p which is referred to as the bending persistence length:

$$\langle \mathbf{t}(s)\cdot\mathbf{t}(s')\rangle = \exp\left(-\frac{|s-s'|}{l_p}\right). \tag{2.2}$$

The bending modulus E is connected to the persistence length l_p via $l_p = E/k_BT$. The mean square end-to-end distance R_E^2 is given by [FK73, HH66]:

$$\begin{aligned} R_E^2 &= \langle(\mathbf{R}(L)-\mathbf{R}(0))^2\rangle \\ &= \langle \int_0^L ds\mathbf{t}(s)\cdot\int_0^L ds'\mathbf{t}(s')\rangle \\ &= \int_0^L ds\int_0^L ds'\langle\mathbf{t}(s)\cdot\mathbf{t}(s')\rangle \\ &= 2Ll_p - 2l_p^2\left(1-\exp\left(-\frac{L}{l_p}\right)\right). \end{aligned} \tag{2.3}$$

The two limiting case are:

$$R_E^2 = \begin{cases} 2Ll_p, & \frac{L}{l_p}\gg 1 \quad \text{random walk limit} \\ L^2, & \frac{L}{l_p}\ll 1 \quad \text{rigid rod limit} \end{cases}. \tag{2.4}$$

A generalized theory does not only treat the resistance of the chain to bending but also to twisting. This is done within the framework of rod theory.

2.2.2 Rod theory

A configuration of a rod is a parameterized space curve $\mathbf{r}(s)$ along with a parameterized family of right-handed orthonormal triads $\mathbf{d}_i(s)$, that is three unit vectors which satisfy for each s the constraints

$$\mathbf{d}_i(s) \cdot \mathbf{d}_j(s) = \delta_{ij} \tag{2.5}$$

$$\mathbf{d}_3(s) = \mathbf{d}_1(s) \times \mathbf{d}_2(s) \tag{2.6}$$

with δ_{ij} being the Kronecker delta. Note that the orthonormal basis set $\{\mathbf{d}_i(s)\}$ contains additional information to the space curve $\mathbf{r}(s)$. It is not just defined by $\mathbf{r}(s)$ and its derivatives as it is the case in the Frenet-Serret frame. One should think of $\mathbf{r}(s)$ as the centerline of the rod (the midcurve defined by the two sugar-phosphate backbones in the case of DNA) while $\{\mathbf{d}_i(s)\}$ specifies the orientation of each cross-section of the rod (some smooth interpolation of the orientation of each set of base-pair in the case of DNA) [Mad00].

With no loss of generality the vector $\mathbf{v}(s)$ can be defined as

$$\mathbf{v}(s) = \frac{d}{ds}\mathbf{r}(s) \tag{2.7}$$

where the components of $\mathbf{v}(s)$ with respect to the triad $\{\mathbf{d}_i(s)\}$ is given by

$$v_i(s) = \mathbf{v}(s) \cdot \mathbf{d}_i(s). \tag{2.8}$$

As the triad is an orthonormal basis set they satisfy kinematic equations of the form

$$\frac{d}{ds}\mathbf{d}_i(s) = \mathbf{u}(s) \times \mathbf{d}_i(s) \tag{2.9}$$

$$\frac{d}{ds}\mathbf{d}_i(s) = \epsilon_{jik} u_j(s) \mathbf{d}_k(s) \tag{2.10}$$

with $\mathbf{u}(s) = u_i(s)\mathbf{d}_i(s)$. This can be shown in the following way. Since $\{\mathbf{d}_i(s)\}$ is a basis set it exists a matrix such that

$$\Omega_{ij} = \left(\frac{d}{ds}\mathbf{d}_i\right) \cdot \mathbf{d}_j. \tag{2.11}$$

Differentiating the orthonormality relation $\mathbf{d}_i(s) \cdot \mathbf{d}_j(s) = \delta_{ij}$ yields $\Omega_{ij} = -\Omega_{ji}$. Multiplying Eq. (2.9) with $\mathbf{d}_m$ gives

$$\Omega_{jm} = \epsilon_{jim} u_i(s). \tag{2.12}$$

Multiplying both sides by ϵ_{jpm} and using the identity $\epsilon_{jpm}\epsilon_{jim} = 2\delta_{pi}$ results in the following unique solution for $\mathbf{u}(s)$:

$$u_p(s) = \frac{1}{2}\epsilon_{jpm}\Omega_{jm} \tag{2.13}$$

$$\mathbf{u}(s) = u_i(s)\mathbf{d}_i(s) = \frac{1}{2}\epsilon_{jim}\Omega_{jm}\mathbf{d}_i(s) = \frac{1}{2}\epsilon_{jim}\left(\left(\frac{d}{ds}\mathbf{d}_j(s)\right) \cdot \mathbf{d}_m(s)\right)\mathbf{d}_i(s). \tag{2.14}$$

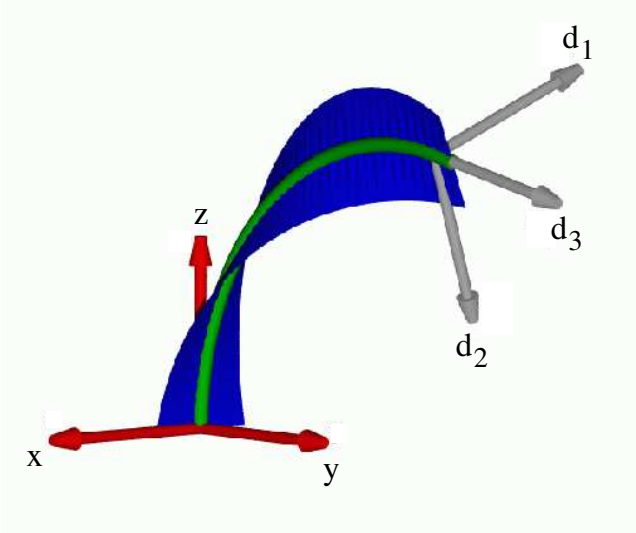


Figure 2.1: Illustration of the used frame. Taken from [Mad00].

Thus the so-called Darboux vector $\mathbf{u}(s)$ exists and is given by

$$u_i = \frac{1}{2}\epsilon_{jim}\left(\frac{d}{ds}\mathbf{d}_j(s)\right)\cdot\mathbf{d}_m(s). \tag{2.15}$$

Eq. (2.9) can be regarded as the generalized Frenet equations.

The components $u_i(s)$ determine the frame $\{\mathbf{d}_i(s)\}$ through integration of the differential equations of Eq. (2.9). Once the frame $\{\mathbf{d}_i(s)\}$ is known the components of $v_i(s)$ determine the centerline $\mathbf{r}(s)$ through integration of Eq. (2.7). Therefore the six scalar functions $u_i(s)$ and $v_i(s)$ characterize the configuration of the rod and are called a set of strains. Together with specified reference strains $\hat{u}_i(s)$ and $\hat{v}_i(s)$ where the reference state is assumed to be the minimum energy or unstressed configuration the strains can be related to shear, stretch, bending, and twist. If s is chosen to be arclength along the centerline of the reference curve $\hat{\mathbf{r}}(s)$ so that

$$|\frac{d}{ds}\hat{\mathbf{r}}(s)| = |\hat{\mathbf{v}}(s)| = 1 \tag{2.16}$$

$|\mathbf{v}(s)| - 1$ is a measure of the local extension and compression respectively. Furthermore the reference unit vector field $\mathbf{d}_3(s)$ can be chosen such that it points into the direction of the tangent vector field of the reference curve $\hat{\mathbf{r}}(s)$:

$$\frac{d}{ds}\hat{\mathbf{r}}(s) = \mathbf{d}_3(s). \tag{2.17}$$

In this case v_1 and v_2 corresponds to shear strains and v_3 can be regarded as a stretch or compression. u_1 and u_2 describe bending strains while u_3 correspond to a twist strain [Mad00].

2.2.3 Parameterization of a ribbon

A ribbon is an inextensible, unshearable rod which can be parameterized by the arclength s. To each point s a triad of unit vectors $\{\mathbf{d}_i(s)\}$ is attached. The vectors $\mathbf{d}_1(s)$ and $\mathbf{d}_2(s)$ are directed along

the two principle axis of the cross section, the vector $\mathbf{d}_3(s)$ is the tangent vector. As the triad is an orthonormal basis set they satisfy the kinematic equations of Eq. (2.9) with $u_j(s)$ representing bend ($u_1(s)$ out-of-plane, and $u_2(s)$ in-plane) and twist strains ($u_3(s)$) respectively. One can find a relation between the ordinary Frenet equations containing only two parameters, the curvature $\kappa(s)$ and the torsion $\tau(s)$

$$\frac{d\mathbf{t}(s)}{ds} = \kappa(s)\mathbf{n}(s) \tag{2.18}$$

$$\frac{d\mathbf{b}(s)}{ds} = -\tau(s)\mathbf{n}(s) \tag{2.19}$$

$$\frac{d\mathbf{n}(s)}{ds} = \tau(s)\mathbf{b}(s) - \kappa(s)\mathbf{t}(s) \tag{2.20}$$

and Eqs. (2.9) by fixing $\mathbf{d}_3(s) = \mathbf{t}(s)$ so that $\mathbf{d}_1(s)$ and $\mathbf{d}_2(s)$ are given by a rotation around $\mathbf{t}(s)$ with angle $\Psi(s)$

$$\mathbf{d}_1(s) = \cos(\Psi(s))\mathbf{b}(s) - \sin(\Psi(s))\mathbf{n}(s) \tag{2.21}$$

$$\mathbf{d}_2(s) = \cos(\Psi(s))\mathbf{n}(s) + \sin(\Psi(s))\mathbf{b}(s) \tag{2.22}$$

In this context $\Psi(s)$ can be seen as the twist angle [PR00a, Mag01]. A straightforward calculation (see Eq. (2.15)) gives for the generalized torsions:

$$\begin{aligned} u_1(s) &= \frac{d\mathbf{d}_3(s)}{ds} \cdot \mathbf{d}_2(s) = \kappa(s)\mathbf{n}(s) \cdot (\cos(\Psi(s))\mathbf{n}(s) + \sin(\Psi(s))\mathbf{b}(s)) \\ &= \kappa(s)\cos\Psi(s) \end{aligned} \tag{2.23}$$

$$\begin{aligned} u_2(s) &= -\frac{d\mathbf{d}_3(s)}{ds} \cdot \mathbf{d}_1(s) = -\kappa(s)\mathbf{n}(s) \cdot (\cos(\Psi(s))\mathbf{b}(s) - \sin(\Psi(s))\mathbf{n}(s)) \\ &= \kappa(s)\sin\Psi(s) \end{aligned} \tag{2.24}$$

$$\begin{aligned} u_3(s) &= -\frac{d\mathbf{d}_1(s)}{ds} \cdot \mathbf{d}_2(s) \\ &= \frac{d(\cos(\Psi(s))\mathbf{b}(s) - \sin(\Psi(s))\mathbf{n}(s))}{ds} \cdot (\cos(\Psi(s))\mathbf{n}(s) + \sin(\Psi(s))\mathbf{b}(s)) \\ &= \left(\sin(\Psi(s))\mathbf{b}(s)\left(\frac{\Psi(s)}{ds} + \tau(s)\right) + \cos(\Psi(s))\mathbf{n}(s)\left(\frac{\Psi(s)}{ds} + \tau(s)\right)\right) \\ &\quad \cdot (\cos(\Psi(s))\mathbf{n}(s) + \sin(\Psi(s))\mathbf{b}(s)) \\ &= \tau(s) + \frac{d\Psi(s)}{ds}. \end{aligned} \tag{2.25}$$

The total twist Tw of a ribbon is thus given by the integration of the local twist $u_3(s)$ along the contour normalized by the factor 2π

$$\mathrm{Tw} = \frac{1}{2\pi}\int_0^L u_3(s)ds \tag{2.26}$$

with L being the contour length. Together with the parameter set $\hat{u}_i(s)$, which determines whether the stress-free reference configuration includes spontaneous curvature and twist, the elastic part of the Hamiltonian is usually defined by quadratic terms in $u_i(s) - \hat{u}_i(s)$ [PR01, PR00a, PR00b, MS94, MS95b, MS95a, MN97, KLNO97, KM97, MMK96, NSJK96].

The inextensibility and unshearability condition requires that the strains $v_i(s)$ in any configuration equal the strains $\hat{v}_i(s)$ in the reference configuration such that

$$v_i(s) \equiv \hat{v}_i(s) \tag{2.27}$$

and

$$v_1(s) = v_2(s) = 0 \tag{2.28}$$
$$v_3(s) = 1. \tag{2.29}$$

2.3 Continuous description of two coupled semiflexible chains

It is an interesting question to which extent this generic description applies to more microscopic models of DNA [ZL01]. The simplest case is that of a "railway track" or ladder model consisting of two (or more) semiflexible chains

$$\mathcal{H}_{tt} = \frac{k}{2}\int_0^L ds \left\{ \left(\frac{d^2\mathbf{r}_1(s)}{ds^2}\right)^2 + \left(\frac{d^2\mathbf{r}_2(s)}{ds^2}\right)^2 \right\}, \tag{2.30}$$

plus a coupling between opposite points on different chains [EBK95]. Liverpool *et al.* [LGK98, GL00] considered the limit where the distance a between the coupling points (i.e. the width of the ribbon) is imposed as a rigid constraint which prevents bending in the plane of the ribbon: $\frac{d\mathbf{t}(s)}{ds}\cdot\mathbf{b}(s) = 0$ where $\mathbf{t}(s) = \frac{d\mathbf{r}(s)}{ds}$ is the tangent vector to the mid-curve $\mathbf{r}(s)$ with

$$\mathbf{r}(s) = \mathbf{r}_1(s) - \frac{a\mathbf{b}(s)}{2} = \mathbf{r}_2(s) + \frac{a\mathbf{b}(s)}{2} = \frac{\mathbf{r}_1(s)+\mathbf{r}_2(s)}{2} \tag{2.31}$$

and $\mathbf{b}(s)$ is the bond-director pointing from one strand to the other. Note, that the constraint is equivalent to $\Psi(s) = 0$, that is

$$u_1(s) = \kappa(s) \tag{2.32}$$
$$u_2(s) = 0 \tag{2.33}$$
$$u_3(s) = \tau(s). \tag{2.34}$$

Rewriting Eq. (2.30) in terms of ribbon variables they found

$$\begin{aligned}\mathcal{H}_{tt} &= \frac{k}{2}\int_0^L ds \left\{ \left(\frac{d^2\mathbf{r}(s)}{ds^2} + \frac{a}{2}\frac{d^2\mathbf{b}(s)}{ds^2}\right)^2 + \left(\frac{d^2\mathbf{r}(s)}{ds^2} - \frac{a}{2}\frac{d^2\mathbf{b}(s)}{ds^2}\right)^2 \right\} \\ &= \frac{k}{2}\int_0^L ds \left\{ 2\left(\frac{d^2\mathbf{r}(s)}{ds^2}\right)^2 + \frac{a^2}{2}\left(\frac{d^2\mathbf{b}(s)}{ds^2}\right)^2 \right\}\end{aligned} \tag{2.35}$$

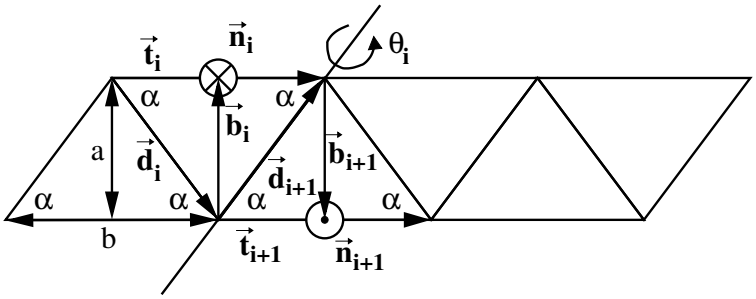


Figure 2.2: Illustration of the used variables. The length of each triangle $|\mathbf{t_i}|$ corresponds to the bond length b and the height $a = |\mathbf{b_i}| = \frac{1}{2} b \tan(\alpha)$ defines the strand separation length. $\{\theta_i\}$ terms the folding angles.

which can also be expressed as

$$\left(\frac{d\mathbf{t}(s)}{ds}\right)^2 = \kappa(s)^2 \tag{2.36}$$

$$\begin{aligned}\left(\frac{d^2\mathbf{b}(s)}{ds^2}\right)^2 &= \left(\mathbf{b}(s)\left(u_1(s)u_2(s) - \frac{du_3(s)}{ds}\right) - \mathbf{n}(s)\left(u_1(s)^2 + u_3(s)^2\right)\right. \\ &\quad \left. + \mathbf{t}(s)\left(u_2(s)u_3(s) + \frac{du_1(s)}{ds}\right)\right)^2 \\ &= \left(\frac{du_1(s)}{ds}\right)^2 + \left(u_1(s)^2 + u_3(s)^2\right)^2 + \left(\frac{du_3(s)}{ds}\right)^2 \\ &= \left(\frac{d\kappa(s)}{ds}\right)^2 + \left(\frac{d\tau(s)}{ds}\right)^2 + \left(\kappa(s)^2 + \tau(s)^2\right)^2 . \end{aligned} \tag{2.37}$$

Thus the Hamiltonian is not just a quadratic function in the strains $u_i(s)$ but derivatives as well as higher order terms of the strains are involved.

Note, that henceforth we use $\mathbf{b}(s)$ as the bond-director and $\mathbf{n}(s)$ as the normal vector to the ribbon plane.

2.4 Geometry of triangulated ribbons

Following Liverpool *et al.* [LGK98, GL00] we consider ribbons discretized by triangulation. In order to extract some fundamental properties of double-stranded semiflexible polymers we consider a ribbon-like system composed of isosceles triangles as shown in Fig. 2.2. The orientation of each triangle is given by $N-1$ rotations around the edges of the triangles with folding angles $\{\theta_i\}$. N is the number of triangles characterized by a set of trihedrons $\{\mathbf{t_i}, \mathbf{b_i}, \mathbf{n_i}\}$ where $\mathbf{t_i}$ is the tangent vector of the ith triangle, $\mathbf{b_i}$ is the bond-director, and $\mathbf{n_i}$ is the normal vector. Note that $\{\mathbf{t_i}, \mathbf{b_i}, \mathbf{n_i}\}$ is normalized. Going from one set of trihedrons $\{\mathbf{t_i}, \mathbf{b_i}, \mathbf{n_i}\}$ to the neighbor set $\{\mathbf{t_{i+1}}, \mathbf{b_{i+1}}, \mathbf{n_{i+1}}\}$ implies a rotation $\mathcal{R}_i$ around the edge between the respective triangles with angle θ_i and a reflection

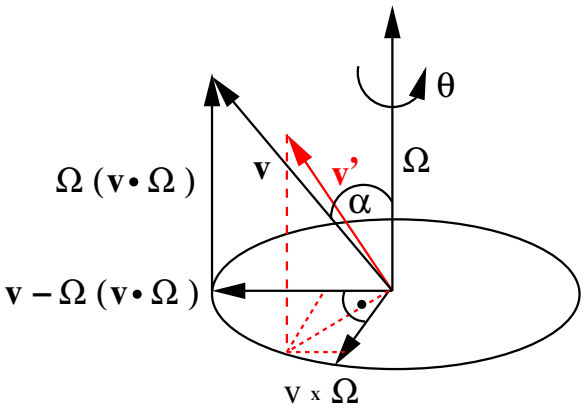


Figure 2.3: Geometrical derivation of Eq. 2.41. Decompose the vector $\mathbf{v}$ which we want to rotate around the unit vector $\mathbf{\Omega}$ with angle θ into a parallel component $\mathbf{v}_{\parallel} = (\mathbf{v}\cdot\mathbf{\Omega})\mathbf{\Omega}$ and two perpendicular components $\mathbf{v}_{\perp,1} = \mathbf{v} - \mathbf{v}_{\parallel}$ and $\mathbf{v}_{\perp,2} = \mathbf{v}\times\mathbf{\Omega}$ such that $\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp,1} + \mathbf{v}_{\perp,2}$. The two perpendicular components has the same length: $|\mathbf{v}_{\perp,1}| = \sqrt{(\mathbf{v}-(\mathbf{v}\cdot\mathbf{\Omega})\mathbf{\Omega})^2} = \sqrt{v^2-(\mathbf{v}\cdot\mathbf{\Omega})^2} = \sqrt{v^2 - v^2\cos(\alpha)^2} = v\sin(\alpha) = |\mathbf{v}\times\mathbf{\Omega}| = |\mathbf{v}_{\perp,2}|$ where α is the angle between $\mathbf{v}$ and $\mathbf{\Omega}$. Before the rotation the component in $\mathbf{v}_{\perp,2}$-direction is zero. The rotated vector $\mathbf{v}'$ is given by $\mathbf{v}' = \mathbf{v}'_{\parallel} + \mathbf{v}'_{\perp,1} + \mathbf{v}'_{\perp,2}$ where the parallel component $\mathbf{v}'_{\parallel} = \mathbf{v}_{\parallel}$ is unchanged by the rotation and one obtains by projecting the rotated vector into the $\mathbf{v}_{\perp,1} - \mathbf{v}_{\perp,2}$-plane for the perpendicular components $\mathbf{v}'_{\perp,1} = \cos(\theta)\mathbf{v}_{\perp,1}$ and $\mathbf{v}'_{\perp,2} = \sin(\theta)\mathbf{v}_{\perp,2}$. Thus it yields in the end: $\mathbf{v}' = (\mathbf{v}\cdot\mathbf{\Omega})\mathbf{\Omega} + \cos(\theta)(\mathbf{v}-(\mathbf{v}\cdot\mathbf{\Omega})\mathbf{\Omega}) + \sin(\theta)(\mathbf{v}\times\mathbf{\Omega})$.

of $\mathbf{b_i}$ and $\mathbf{n_i}$, i.e.

$$\begin{pmatrix}\mathbf{t_{i+1}}\\ \mathbf{b_{i+1}}\\ \mathbf{n_{i+1}}\end{pmatrix} = \mathcal{T}\mathcal{R}_i \begin{pmatrix}\mathbf{t_i}\\ \mathbf{b_i}\\ \mathbf{n_i}\end{pmatrix} \tag{2.38}$$

with

$$\mathcal{T} = \begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix} \tag{2.39}$$

$$\mathcal{R}_i = \begin{pmatrix}\mathbf{t_i}\cdot\mathbf{t_{i+1}} & \mathbf{t_i}\cdot\mathbf{b_{i+1}} & \mathbf{t_i}\cdot\mathbf{n_{i+1}}\\ \mathbf{b_i}\cdot\mathbf{t_{i+1}} & \mathbf{b_i}\cdot\mathbf{b_{i+1}} & \mathbf{b_i}\cdot\mathbf{n_{i+1}}\\ \mathbf{n_i}\cdot\mathbf{t_{i+1}} & \mathbf{n_i}\cdot\mathbf{b_{i+1}} & \mathbf{n_i}\cdot\mathbf{n_{i+1}}\end{pmatrix}. \tag{2.40}$$

The matrix product $\mathcal{T}\mathcal{R}_i$ can be viewed as a transfer matrix. To evaluate the scalar products of $\mathcal{R}_i$ the neighbor set of trihedrons $\{\mathbf{t_{i+1}}, \mathbf{b_{i+1}}, \mathbf{n_{i+1}}\}$ has to be determined. In general a vector $\vec{v}$ which is rotated around an axis $\vec{\Omega}$ with angle θ is given by [Gol91] (see Fig. 2.3)

$$\vec{v}\,' = \vec{v}\cos(\theta) + \left(\frac{\vec{\Omega}}{|\vec{\Omega}|}\times\vec{v}\right)\sin(\theta) + \frac{\vec{\Omega}}{|\vec{\Omega}|}\left(\frac{\vec{\Omega}}{|\vec{\Omega}|}\cdot\vec{v}\right)(1-\cos(\theta))\,. \tag{2.41}$$

Using Eq. 2.41 one obtains for $\mathbf{t_{i+1}}$

$$\mathbf{t_{i+1}} = \mathbf{t_i}\cos(\theta_i) + \left(\frac{\vec{d}_{i+1}}{|\vec{d}_{i+1}|} \times \mathbf{t_i}\right)\sin(\theta_i) + \frac{\vec{d}_{i+1}}{|\vec{d}_{i+1}|}\left(\frac{\vec{d}_{i+1}}{|\vec{d}_{i+1}|} \cdot \mathbf{t_i}\right)(1-\cos(\theta_i)) \tag{2.42}$$

with

$$\vec{d}_{i+1} = \frac{\vec{t}_i}{2} + \vec{b}_i. \tag{2.43}$$

The evaluation of the cross product and the scalar product yields

$$\begin{aligned}\mathbf{t_{i+1}} &= \mathbf{t_i}\cos(\theta_i) + \left(\frac{\frac{\vec{t}_i}{2}+\vec{b}_i}{|\frac{\vec{t}_i}{2}+\vec{b}_i|} \times \mathbf{t_i}\right)\sin(\theta_i) + \frac{\frac{\vec{t}_i}{2}+\vec{b}_i}{|\frac{\vec{t}_i}{2}+\vec{b}_i|}\left(\frac{\frac{\vec{t}_i}{2}+\vec{b}_i}{|\frac{\vec{t}_i}{2}+\vec{b}_i|} \cdot \mathbf{t_i}\right)(1-\cos(\theta_i)) \\ &= \mathbf{t_i}\cos(\theta_i) + \frac{|\vec{b}_i|}{|\vec{d}_{i+1}|}\mathbf{n_i}\sin(\theta_i) + \frac{\frac{\vec{t}_i}{2}+\vec{b}_i}{|\vec{d}_{i+1}|}\frac{|\vec{t}_i|}{2|\vec{d}_{i+1}|}(1-\cos(\theta_i))\,.\end{aligned} \tag{2.44}$$

where $\frac{|\vec{t}_i|}{2|\vec{d}_{i+1}|} = \cos(\alpha)$ and $\frac{|\vec{b}_i|}{|\vec{d}_{i+1}|} = \sin(\alpha)$ for geometric reasons (see Fig. 2.2). Thus the following results for the 3 scalar products is obtained:

$$\mathcal{R}_{i,11} = \mathbf{t_i} \cdot \mathbf{t_{i+1}} = \cos(\theta_i) + \sin(\alpha)^2(1-\cos(\theta_i)) \tag{2.45}$$
$$\mathcal{R}_{i,21} = \mathbf{b_i} \cdot \mathbf{t_{i+1}} = \sin(\alpha)\cos(\alpha)(1-\cos(\theta_i)) \tag{2.46}$$
$$\mathcal{R}_{i,31} = \mathbf{n_i} \cdot \mathbf{t_{i+1}} = \sin(\alpha)\sin(\theta_i). \tag{2.47}$$

The remaining terms of the rotation matrix $\mathcal{R}_i$ can be calculated in an analogous fashion resulting in

$$\begin{aligned}\mathcal{R}_{i,11} &= \cos(\theta_i) + \cos(\alpha)^2(1-\cos(\theta_i)) \\ \mathcal{R}_{i,12} &= -\cos(\alpha)\sin(\alpha)(1-\cos(\theta_i)) \\ \mathcal{R}_{i,13} &= -\sin(\alpha)\sin(\theta_i) \\ \mathcal{R}_{i,21} &= \cos(\alpha)\sin(\alpha)(1-\cos(\theta_i)) \\ \mathcal{R}_{i,22} &= \cos(\theta_i) + \sin(\alpha)^2(1-\cos(\theta_i)) \\ \mathcal{R}_{i,23} &= -\cos(\alpha)\sin(\theta_i) \\ \mathcal{R}_{i,31} &= \sin(\alpha)\sin(\theta_i) \\ \mathcal{R}_{i,32} &= \cos(\alpha)\sin(\theta_i) \\ \mathcal{R}_{i,33} &= \cos(\theta_i).\end{aligned} \tag{2.48}$$

In order to quantify properties such as bending and twisting within the given discretization we study the relation between the folding angles θ_i and these quantities which is illustrated in Fig. 2.4. One recognizes that the chain is not bent in case of $\theta_i - \theta_{i+1} = \delta\theta_i = 0$ and that purely twisted structures correspond to $\theta_i \equiv const$. The chain is untwisted but bent if $\delta\theta_i = 2\theta_i$. In case of $\theta_i \neq \pm\theta_{i+1}$ and $\theta_i \neq 0$ the chain is bent and twisted simultaneously resulting in solenoidal/torsional structures as is illustrated in Fig. 2.4(f). A kink is characterized by unlike twists meeting at an edge as it is shown in Fig. 2.4(d).

Due to the triangulation of the ribbon three triangles has to be considered to calculate the discretized expressions for the out-of-plane bending strain

$$u_1(s) = -\left(\frac{d}{ds}\mathbf{t}(s)\right)\cdot\mathbf{n}(s) \approx -\frac{\mathbf{t}(s+\Delta s)-\mathbf{t}(s)}{\Delta s}\cdot\mathbf{n}(s) = -\frac{1}{\Delta s}\mathbf{t}(s+\Delta s)\cdot\mathbf{n}(s) \tag{2.49}$$

and the twist strain

$$u_3(s) = \left(\frac{d}{ds}\mathbf{b}(s)\right)\cdot\mathbf{n}(s) \approx \frac{\mathbf{b}(s+\Delta s)-\mathbf{b}(s)}{\Delta s}\cdot\mathbf{n}(s) = \frac{1}{\Delta s}\mathbf{b}(s+\Delta s)\cdot\mathbf{n}(s) \tag{2.50}$$

which we call κ_i and τ_i respectively. The local curvature κ_i and the local twist rate τ_i between triangle i and $i+2$ are therefore given by

$$\kappa_i \equiv -\frac{1}{b}\sum_{j=i}^{i+1}\mathbf{n}_j\cdot\mathbf{t}_{j+1} \approx \frac{\sin(\alpha)}{b}\delta\theta_i \tag{2.51}$$

$$\tau_i \equiv \frac{1}{b}\sum_{j=i}^{i+1}\mathbf{n}_j\cdot\mathbf{b}_{j+1} \approx \frac{\cos(\alpha)}{b}(\theta_i+\theta_{i+1}), \tag{2.52}$$

where the accuracy of the right-hand side expressions only depends on the refinement of the discretization, i.e. on the values of b and α. Hence a spontaneous bending can be introduced via an additional term to the Hamiltonian with

$$\mathcal{H}_{curv} = k_{curv}\sum_i\left(\sum_{j=i}^{i+1}\mathbf{n}_j\cdot\mathbf{t}_{j+1}-\delta\theta_{sp,i}\right)^2 \tag{2.53}$$

and a spontaneous twist can be introduced by an additional term

$$\mathcal{H}_{Tw} = k_{Tw}\sum_i\left(\sum_{j=i}^{i+1}\mathbf{n}_j\cdot\mathbf{b}_{j+1}-\theta_{sp,i}\right)^2. \tag{2.54}$$

Note, that the total twist Tw is given by $\mathrm{Tw} = 1/(2\pi)\sum_i\tau_i$.

2.5 Model description

The bending stiffness within the given discretization can be taken into account by various interactions. One possible definition of a bending stiffness, which makes the problem analytically tractable, is a nearest neighbor interaction (plaquette stiffness) between the normal vectors $\{\mathbf{n_i}\}$ in analogy to the triangulation of vesicles [KG92] which results in the following Hamiltonian

$$\frac{\mathcal{H}_{nn}}{k_BT} = k\sum_{i=1}^{N-1}\left(1+\mathbf{n_i}\cdot\mathbf{n_{i+1}}\right). \tag{2.55}$$

In contrast Liverpool *et al.* [LGK98, GL00] were interested in the statistical mechanics of coupled worm-like chains and therefore chose a next-nearest neighbor interaction (edge stiffness) between the tangent vectors $\{\mathbf{t_i}\}$ with rigidity k so that the Hamiltonian is given by

$$\frac{\mathcal{H}_{tt}}{k_BT} = k\sum_{i=1}^{N-2}\left(1-\mathbf{t_i}\cdot\mathbf{t_{i+2}}\right). \tag{2.56}$$

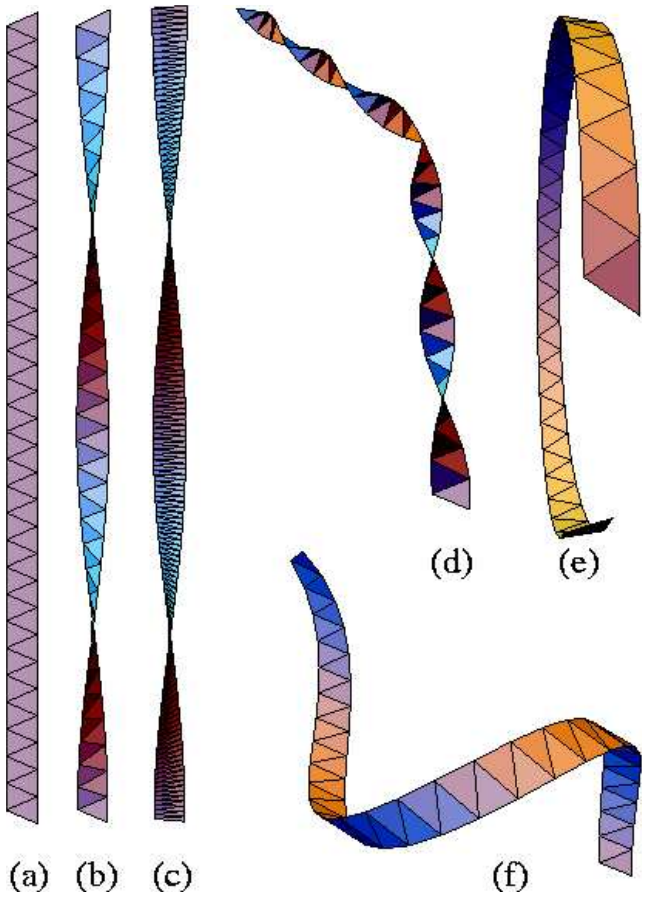


Figure 2.4: Illustration of bending, twisting, and kinking. (a) A flat ribbon as ground state conformation. (b) A twisted structure (c) The same twisted structure obtained with a smoother discretization. (d) Unlike twists meeting at the center resulting in a kink with θ_i positive for $i < N/2$, negative for $i \geq N/2$, and $|\theta_i| = |\theta_{i+1}|$, i.e. $\delta\theta_i = 0$, $\forall i \neq N/2$ and $\delta\theta_{N/2} = 2\theta_{N/2}$. (e) A bent structure. (f) A mixture of bent and twist resembling a solenoidal structure.

Both definitions lead to a flat ribbon as the ground state conformation for zero temperatures $T = 0$.

The above defined interactions lead to very distinct conformational features of the ribbon which can be understood by building up the ribbon just by adding successively the triangles in the absence of thermal fluctuations. Assuming that $\theta_1 \neq 0$ all subsequent angles θ_i with $i > 1$ vanish in the case of the nearest neighbor interaction ($\mathcal{H}_{nn}$). In contrast the tangent-tangent interaction ($\mathcal{H}_{tt}$) leads to the formation of a helix with $\theta_i = \theta_{i+1}$ as a result of the enforced alignment of the tangent vectors. This suggests a correlation of the folding angles $\{\theta_i\}$ which entails at least locally helical structures.

Assuming that the chains are rather stiff (continuum limit), i.e. small folding angles θ_i, one can expand the Hamiltonians with respect to θ_i. Since $\mathcal{H}_{nn}$ is diagonal in θ_i, it is sufficient to consider terms up to second order. $\mathcal{H}_{tt}$ contains coupling terms between θ_i and θ_{i+1} which makes it necessary

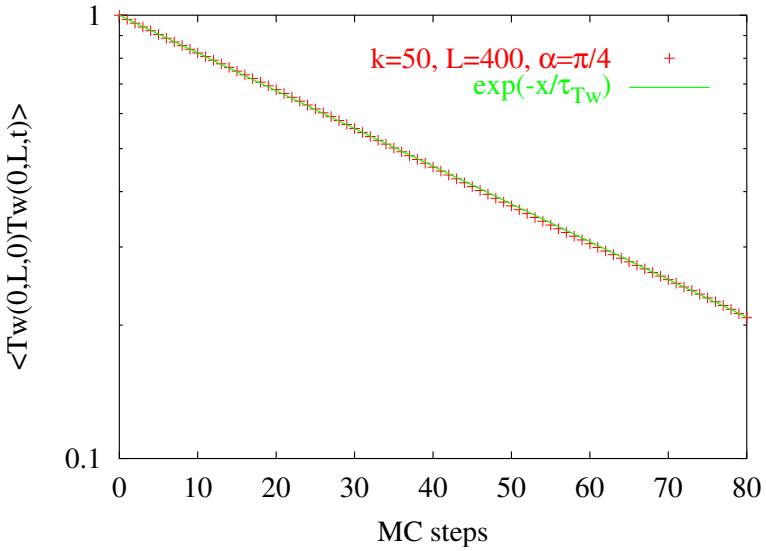


Figure 2.5: Measurement of the 'time' autocorrelation function of the total twist of the chain. The longest correlation time we observed was $\tau_{Tw} = 51$ MC steps for the total twist $\mathrm{Tw}(0, L)$ of the chain.

to keep terms up to fourth order in the analysis:

$$\frac{\mathcal{H}_{nn}}{k_B T} \approx \frac{k}{2} \sum_{i=1}^{N-1} \theta_i^2 \tag{2.57}$$

$$\frac{\mathcal{H}_{tt}}{k_B T} \approx \frac{k}{2} \sum_{i=1}^{N-2} \left\{ \sin(\alpha)^2 \delta\theta_i^2 \left(1 - \frac{1}{12}\delta\theta_i^2\right) + \sin(\alpha)^2 \cos(\alpha)^2 \theta_i^2 \theta_{i+1}^2 \right\} \tag{2.58}$$

with $\delta\theta_i = \theta_i - \theta_{i+1}$.

2.6 MC Simulation

Both models have local interactions and can be studied conveniently using a dynamic MC scheme. Trial moves consist of small random changes of the folding angles by a small amplitude $1/\sqrt{k}$, where k is the bending stiffness, and are accepted or rejected according to the Metropolis scheme [MRR$^+$53]. In the simulations we always use the full Hamiltonians Eq. (2.55) and (2.56). MC moves changing the folding angles correspond to the well-known Pivot algorithm [LB00]. The conformations are subsequently recalculated from Eqs. (2.38)-(2.48) and analyzed. Each simulation run comprises 10^5 MC-moves where one MC move corresponds to $N - 1$ trials with N being the number of triangles. The longest correlation time we observed was on the order of 50 MC moves for the total twist of the chain. In order to check if equilibrium is reached we compared simulation runs with a flat initial conformation, i.e $\theta_i = 0$, with simulation runs with crumpled conformations corresponding to equally distributed angles θ_i out of the interval $[-1/\sqrt{k}; 1/\sqrt{k}]$. Both runs yield the same results for the calculated observables.

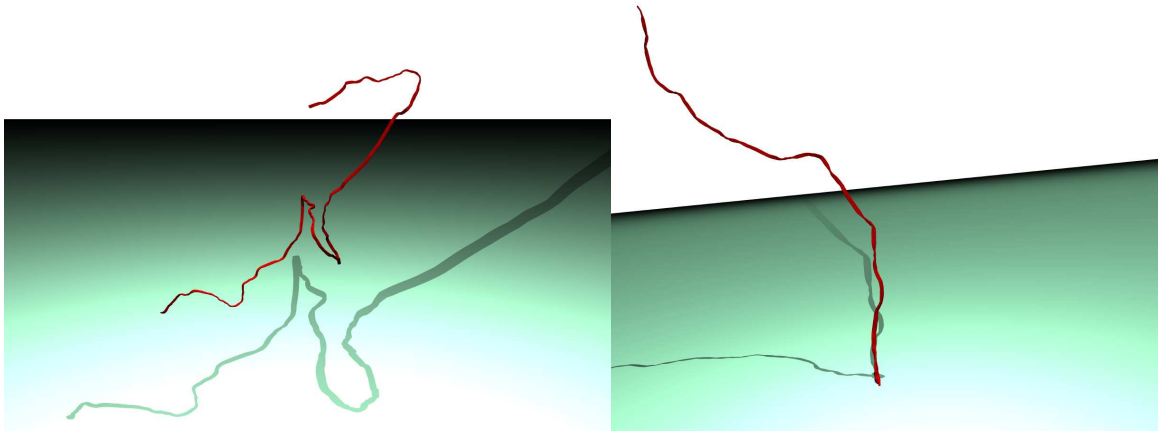


Figure 2.6: Snapshots for $\mathcal{H}_{nn}$ (left) and $\mathcal{H}_{tt}$ (right) with $k = 50$, $N = 800$, $\alpha = \pi/3$. The figure illustrates that just by watching conformational snapshots it is not possible to decide if kinks are present due to projection of a three-dimensional object onto a two-dimensional plane (see [LGK98]).

2.7 Plaquette Stiffness

Since the Hamiltonian $\mathcal{H}_{nn}$ of Eq. (2.57) is quadratic and diagonal in θ_i the solution in angle space is trivial. As a consequence of the independence of successive folding angles it yields $\langle\theta_i\theta_j\rangle = \frac{1}{k}\delta_{ij}$ and $\langle\mathcal{A}\rangle = \langle\prod_{k=i}^{j}(\mathcal{T}\mathcal{R}_k)\rangle = \langle\mathcal{T}\mathcal{R}_k\rangle^{j-i}$ where the matrix product is carried out in the eigenvector basis of $\langle\mathcal{T}\mathcal{R}_k\rangle$ (the eigenvectors depend only on the geometry of the triangles). The diagonal elements of $\langle\mathcal{A}\rangle$ are the correlation functions of $\langle\mathbf{t_i}\cdot\mathbf{t_j}\rangle, \langle\mathbf{b_i}\cdot\mathbf{b_j}\rangle, \langle\mathbf{n_i}\cdot\mathbf{n_j}\rangle$. Thus one calculates $\langle\mathcal{T}\mathcal{R}_k\rangle$

$$\begin{aligned}\langle\mathcal{T}\mathcal{R}_k\rangle = \mathcal{T}\langle\mathcal{R}_k\rangle &= \frac{\mathcal{T}\int\mathcal{D}[\theta]\mathcal{R}_k\exp(-\frac{\mathcal{H}_{nn}}{k_BT})}{\int\mathcal{D}[\theta]\exp(-\frac{\mathcal{H}_{nn}}{k_BT})} \\ &= \begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}\begin{pmatrix}\frac{-1+4k+\cos(2\alpha)}{4k} & -\frac{\cos(\alpha)\sin(\alpha)}{2k} & 0\\ \frac{\cos(\alpha)\sin(\alpha)}{2k} & \frac{-1+2k+\sin(\alpha)^2}{2k} & 0\\ 0 & 0 & -1+\frac{1}{2k}\end{pmatrix},\end{aligned} \tag{2.59}$$

diagonalizes it such that $\langle\mathcal{T}\mathcal{R}_k\rangle = \mathcal{S}\mathcal{D}\mathcal{S}^T$ with $\mathcal{S}$ being the eigenvector matrix and $\mathcal{D}$ being the diagonalized matrix and raises $\mathcal{D}$ to the power of $j-i$. Transforming back results in $\langle\mathcal{A}\rangle = \langle\mathcal{T}\mathcal{R}_k\rangle^{j-i} = \mathcal{S}\mathcal{D}^{j-i}\mathcal{S}^T$. Finally the continuum chain limit can be performed with $s = (j-i)b$, $l_p = bk/\sin(\alpha)^2$, $a = \frac{1}{2}b\tan(\alpha)$, $(j-i)\to\infty$, $b\to 0$, i.e. $a\to 0$, where l_p is the persistence length, a is the strand separation, b is the Kuhn segment length, $0 < s < L$ is the arclength, and L is the contour length. Note that within this model α is a fixed parameter that determines bending characteristics of the ribbon. Since the calculation is rather demanding and gives very complicated expressions we used a Mathematica script to evaluate the correlation matrix. In the end an exact expression for the

autocorrelation functions is obtained:

$$\langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle = \exp\left(-\frac{s}{l_p}\right) \tag{2.60}$$

$$\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle = \exp\left(-\frac{s}{l_p \tan(\alpha)^2}\right) \tag{2.61}$$

$$\langle \mathbf{n}(0) \cdot \mathbf{n}(s) \rangle = \exp\left(-\frac{s}{l_p \sin(\alpha)^2}\right). \tag{2.62}$$

For $\alpha = \pi/2$ the triangles degenerate to rectangles such that bending can only occur out of the plane of the ribbon and the usual worm-like chain result for two dimensions is recovered. All cross-correlation functions (the off-diagonal elements of $\langle \mathcal{A} \rangle$) vanish. Eqs. (2.61), (2.62) represent the persistence length $l_{p,in} = l_p \tan(\alpha)^2$ for bending within the plane of the ribbon and the persistence length $l_{p,out} = l_p \sin(\alpha)^2$ for bending out of the plane of the ribbon respectively [NSJK96]. This model was recently treated as a twisted zig-zag fiber within the framework of a two-angle model for studying structural properties of chromatin [SGB01].

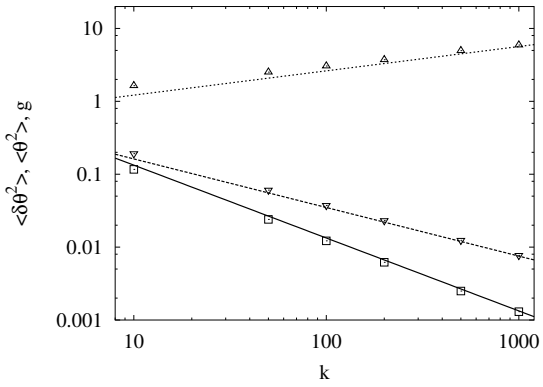
From the tangent-tangent correlation function the mean squared end-to-end distance can be calculated:

$$\begin{aligned} R_E^2 = \langle (\mathbf{R}(L) - \mathbf{R}(0))^2 \rangle &= \int_0^L ds_1 \int_0^L ds_2 \langle \mathbf{t}(s_1) \cdot \mathbf{t}(s_2) \rangle \\ &= 2Ll_p - 2l_p^2 \left(1 - \exp\left(-\frac{L}{l_p}\right)\right). \end{aligned} \tag{2.63}$$

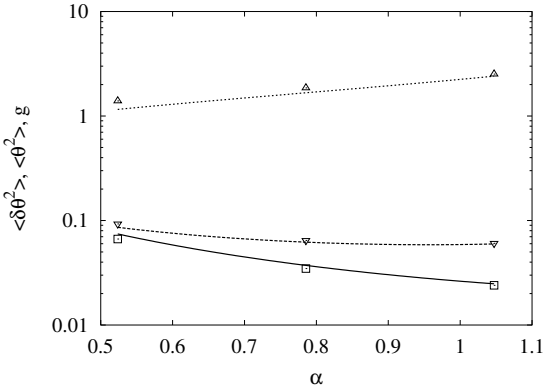
Eqs. (2.60) and (2.63) are identical to results for single worm-like chains [DE86]. Eq. (2.63) interpolates between the limiting behaviors of random coils ($2Ll_p$) for $L \gg l_p$ and rigid rods (L^2) for $L \ll l_p$.

2.8 Edge Stiffness

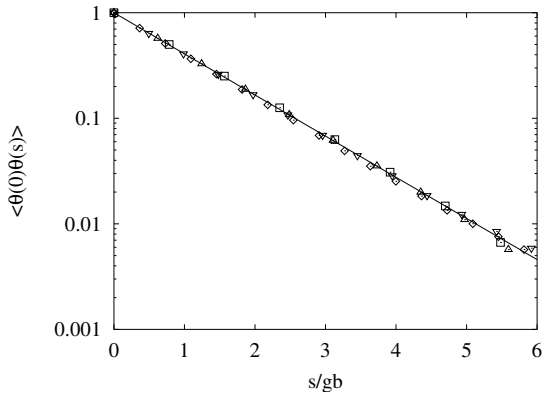
In the following we present a simple scaling argument which allows us to rationalize the behavior of the Liverpool model. Consider first the $\delta\theta$ part of Eq. (2.58). In the absence of other terms the folding angles would perform a simple random walk with step length $\langle \delta\theta_i^2 \rangle = \frac{1}{k \sin(\alpha)^2}$. The leading term limiting the fluctuations of the folding angles around zero is of order $\mathcal{O}(\theta_i^4)$. The behavior of the coupled system can be inferred from scaling arguments similar to those used for polymer adsorption. Consider a vanishing folding angle and follow the chain in either direction. Up to a characteristic number of steps g the folding angles will show simple diffusion. As a consequence the mean-squared folding angle averaged over this short segment is $\langle \theta_i^2 \rangle = g \langle \delta\theta_i^2 \rangle$ corresponding to a potential energy $\frac{E_{ex}}{k_B T} \sim g \langle \theta_i^4 \rangle \sim 3g \langle \theta_i^2 \rangle^2 \sim 3g^3 \langle \delta\theta_i^2 \rangle$. Note that for a Gaussian distribution function P(x) the 4th and 2nd moment are related by $\langle x^4 \rangle = 3 \langle x^2 \rangle^2$. The free diffusion of the



(a) $\langle\theta_i^2\rangle$ (downward triangles), $\langle\delta\theta_i^2\rangle$ (squares), and g (upward triangles) with $\alpha = \pi/3$ as a function of k.



(b) $\langle\theta_i^2\rangle$ (downward triangles), $\langle\delta\theta_i^2\rangle$ (squares), and g (upward triangles) with $k = 50$ as a function of α.



(c) Numerical evidence for the derived expression of $\langle\theta_i\theta_j\rangle$. The data refer to $k = 50$ (squares), $k = 100$ (circles), $k = 200$ (upward triangles), $k = 500$ (downward triangles) and $\alpha = \pi/4$.

Figure 2.7: Simulation versus scaling. We determine the correct prefactor $g_{fit} = 0.56 \pm 0.05$ of g from the numerical data of the folding angle correlation function $\langle\theta_i\theta_j\rangle$ which is our solely free parameter and use it for all following comparisons between scaling analysis and numerical results.

folding angles has to stop when this potential energy is of order k_BT resulting in a condition for g:

$$\begin{aligned}
\sin(\alpha)^2\cos(\alpha)^2 k\, g\langle\theta^4\rangle &= 1 \\
3\sin(\alpha)^2\cos(\alpha)^2 k\, g\langle\theta^2\rangle^2 &= 1 \\
3\sin(\alpha)^2\cos(\alpha)^2 k\, g^3\langle\delta\theta^2\rangle^2 &= 1 \\
g &= \left(\frac{1}{3}\frac{k^2\sin(\alpha)^4}{\sin(\alpha)^2\cos(\alpha)^2 k}\right)^{\frac{1}{3}} \\
g &= \left(\frac{k\tan(\alpha)^2}{3}\right)^{\frac{1}{3}}. \qquad (2.64)
\end{aligned}$$

As already mentioned in section 2.5 the folding angles θ_i are correlated. The previous scaling considerations suggest that the correlation length is on the order of g. This is confirmed by the following calculation. The mean square displacement of the folding angles is given by

$$\langle(\theta_i-\theta_j)^2\rangle = 2\langle\theta_i^2\rangle - 2\langle\theta_i\theta_j\rangle = \begin{cases} 2|i-j|\langle\delta\theta_i^2\rangle, & |i-j| \ll g, \\ 2\langle\theta_i^2\rangle, & |i-j| \gg g. \end{cases} \tag{2.65}$$

Thus the correlation function $\langle\theta_i\theta_j\rangle$ can be expressed as

$$\frac{\langle\theta_i\theta_j\rangle}{\langle\theta_i^2\rangle} = \begin{cases} 1-\frac{|i-j|}{g}, & |i-j| \ll g \\ 0, & |i-j| \gg g \end{cases} \approx \begin{cases} \exp\left(-\frac{|i-j|}{g}\right), & |i-j| \ll g \\ 0, & |i-j| \gg g \end{cases}. \tag{2.66}$$

Altogether we obtain

$$\langle\delta\theta_i^2\rangle = \frac{1}{k\sin(\alpha)^2} \tag{2.67}$$

$$g \sim \left(\frac{k\tan(\alpha)^2}{3}\right)^{\frac{1}{3}} \tag{2.68}$$

$$\langle\theta_i^2\rangle = g\langle\delta\theta_i^2\rangle \tag{2.69}$$

$$\frac{\langle\theta_i\theta_j\rangle}{\langle\theta_i^2\rangle} = \exp\left(-\frac{|j-i|}{g}\right). \tag{2.70}$$

Fig. 2.7 shows that these arguments are fully supported by the results of our MC simulations with $g = (0.56 \pm 0.05)\left(\frac{k\tan(\alpha)^2}{3}\right)^{\frac{1}{3}}$.

Using again the transfer matrix ansatz and considering only terms on the order of $\mathcal{O}(\theta_i^2)$ the following expression for $\mathbf{t_i}\cdot\mathbf{t_j}$, $\mathbf{b_i}\cdot\mathbf{b_j}$ and $\mathbf{n_i}\cdot\mathbf{n_j}$ is obtained in the low temperature limit:

$$\mathbf{t_i}\cdot\mathbf{t_j} = 1 - \frac{\sin(\alpha)^2}{2}\left(\sum_{k=i/2}^{j/2}\delta\theta_{2k}\right)^2 \tag{2.71}$$

$$\mathbf{b_i}\cdot\mathbf{b_j} = 1 - \frac{\cos(\alpha)^2}{2}\left(\sum_{k=i}^{j-1}\theta_k^2 + 2\sum_{k=i}^{j-1}\sum_{k'=k+1}^{j-1}\theta_k\theta_{k'}\right) \tag{2.72}$$

$$\mathbf{n_i}\cdot\mathbf{n_j} = 1 - \frac{1}{2}\sum_{k=i/2}^{j/2}\delta\theta_{2k}^2 + \cos(\alpha)^2\sum_{k=i}^{j-1}\sum_{k'=k+1}^{j-1}\theta_k\theta_{k'}. \tag{2.73}$$

Note that i, j are either odd or even depending on which strand is under consideration. Without loss of generality we choose i, j to be even. First of all we use that $\langle\mathbf{t}(0)\cdot\mathbf{t}(s)\rangle$ has to interpolate between 1 for $s = 0$ and 0 for $s \to \infty$ and that the right hand side of Eq. (2.71) is the Taylor expansion up to first order of the exponential function $\exp\left(\frac{\sin(\alpha)^2}{2}\left(\sum_{k=i/2}^{j/2}\delta\theta_{2k}\right)^2\right)$. Substituting then $s = 2|j-i|b$ and $l_p = 4bk$, performing the continuum chain limit with $b \to 0$ and $\alpha \to \pi/2$ respectively, i.e. keeping the strand separation a constant, yields the following expression for the autocorrelation function of the tangent vectors:

$$\langle\mathbf{t}(0)\cdot\mathbf{t}(s)\rangle = \exp\left(-\frac{s}{l_p}\right). \tag{2.74}$$

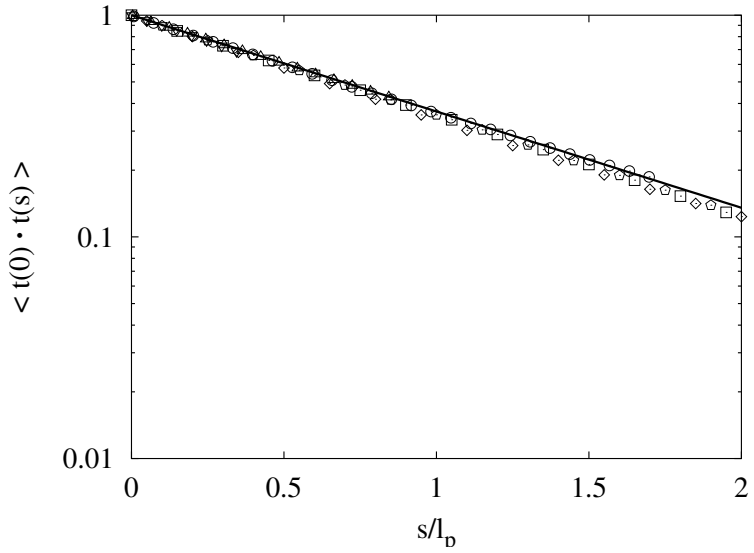


Figure 2.8: Comparison of MC data and analytical results (solid line) for the autocorrelation function of the tangent vectors with $k = 50$ (squares), $k = 100$ (circles), $k = 200$ (upward triangles), $k = 500$ (downward triangles) and $\alpha = \pi/3$, and $k = 50$ and $\alpha = \pi/4$ (diamonds), $\alpha = \pi/6$ (pentagons).

Thus the mean squared end-to-end distance R_E^2 becomes identical to Eq. (2.63). Eq. (2.74) is confirmed by our MC simulation data shown in Fig. 2.8.

To get an idea of the structural properties characterized by the autocorrelation function of the bond-directors $\langle \mathbf{b_i} \cdot \mathbf{b_j} \rangle$ we calculate the mean squared twist $\langle \mathrm{Tw}(i,j)^2 \rangle$ of the ribbon. Following the definition of the local twist rate τ_i of Eq. (2.52) the total twist between two triangles of index i and j is just the sum of the local twist angles determined by the projections of the normal vector of the ith triangle onto the bond-director of the $(i+1)$th triangle, that is

$$\mathrm{Tw}(i,j) = \frac{1}{2\pi} \sum_{k=i}^{j-1} \mathbf{n_i} \cdot \mathbf{b_{i+1}} = \frac{\cos(\alpha)}{2\pi} \sum_{k=i}^{j-1} \theta_k. \tag{2.75}$$

Comparing Eq. (2.72) and (2.75) we find for small twist angles

$$\langle \mathbf{b_i} \cdot \mathbf{b_j} \rangle = 1 - 2\pi^2 \langle \mathrm{Tw}(i,j)^2 \rangle. \tag{2.76}$$

Hence the autocorrelation function of the bond-directors can be seen as a measure for the local twist structure of the ribbon.

In contrast to the plaquette stiffness model, the angles θ_i in the edge stiffness model are correlated (see Eq. (2.70)). Therefore the double summation over $\langle \theta_i \theta_j \rangle$ in Eq. (2.72) proceeds along the

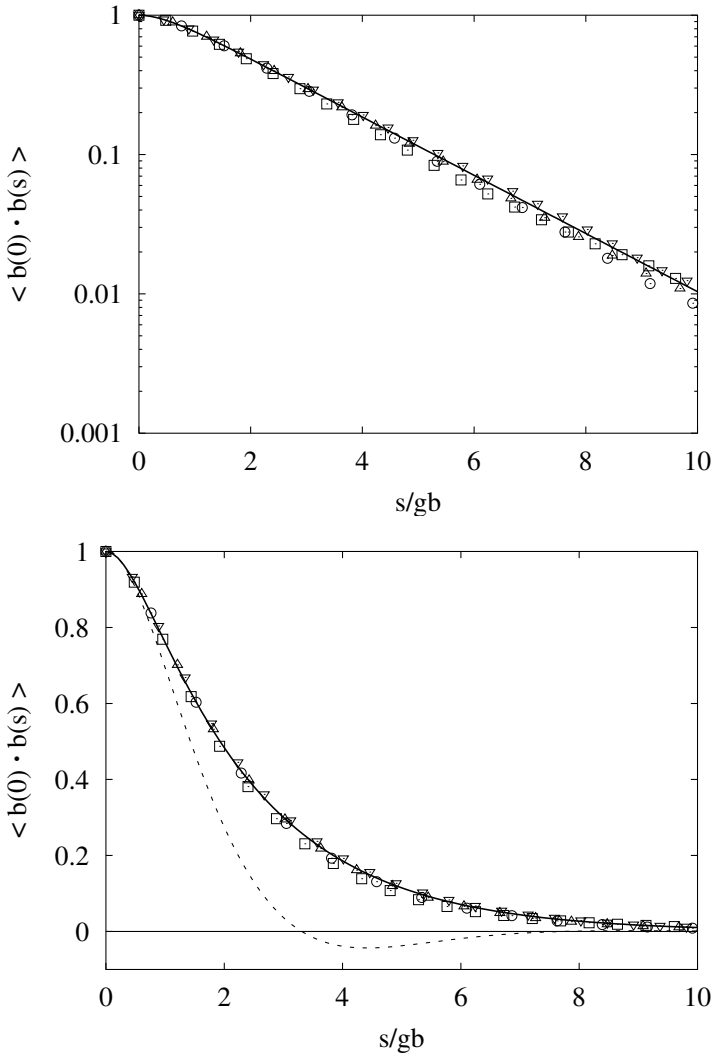


Figure 2.9: (a) Autocorrelation function of the bond-directors with $k = 50$ (squares), $k = 100$ (circles), $k = 200$ (upward triangles), $k = 500$ (downward triangles) and $\alpha = \pi/3$. The data show the predicted functional form (solid line) for $\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle$ of Eq. (2.79). In order to check the scaling argument of Eq. (2.79) we determined the correct prefactor $g_{fit} = 0.56 \pm 0.05$ of g with the help of the numerical data of $\langle \theta_i \theta_j \rangle$ (see Fig. 2.7) and inserted it into Eq. (2.79). The agreement is excellent. (b) Comparison of our simulation data with the analytical result of Liverpool *et al.* [LGK98, GL00] (dashed line). The predicted oscillation and resultant pitch is not recovered. But we find the same scaling behavior of the helical persistence length with $l_b = gb \sim l_p^{\frac{1}{3}} a^{\frac{2}{3}}$. It is also striking that the predicted functional form of Liverpool *et al.* is in very good agreement with our numerical data within one helical persistence length l_b.

lines of the calculation of the mean squared end-to-end distance of the worm-like chain model

$$\begin{aligned} \langle \mathrm{Tw}(i,j)^2 \rangle &= \frac{\cos(\alpha)^2}{(2\pi)^2} \sum_{k=i}^{j} \sum_{k'=i}^{j} \langle \theta_k \theta_{k'} \rangle \\ &= \frac{\cos(\alpha)^2}{(2\pi)^2} \left(\sum_{k} \langle \theta_k^2 \rangle + 2 \sum_{k=i}^{j} \sum_{k'=k+1}^{j} \langle \theta_k \theta_{k'} \rangle \right) \end{aligned} \tag{2.77}$$

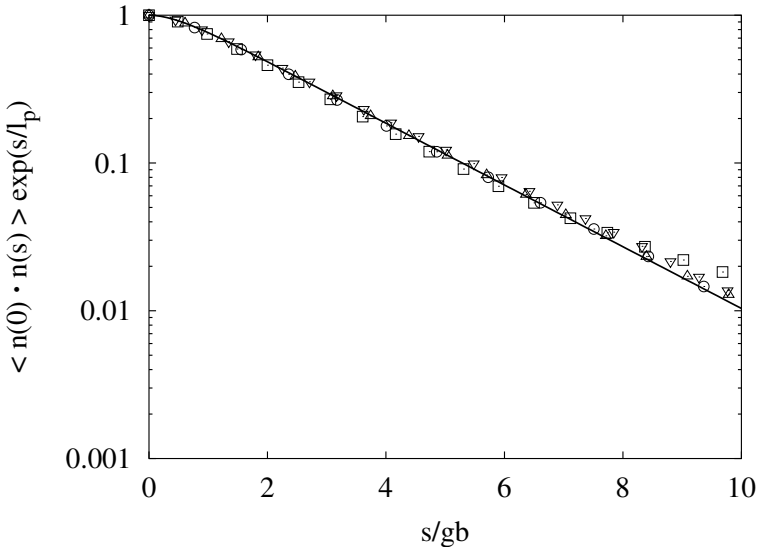


Figure 2.10: Autocorrelation function of the normal vectors with $k = 50$ (squares), $k = 100$ (circles), $k = 200$ (upward triangles), $k = 500$ (downward triangles) and $\alpha = \pi/3$. The ratio of the normal vector correlation function and the tangent correlation function $\langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle$ (see Eq. (2.81)) is shown (solid line) so that the same exponential decay as for $\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle$ should be regained. This is in agreement with the numerical data.

Using the scaling expressions of Eqs. (2.68) and (2.69), the same substitutions as in Eq. (2.74), and performing the continuum chain limit the following relationship for the mean-squared twist is obtained

$$\begin{aligned}
\langle \mathrm{Tw}(0,s)^2 \rangle &= \lim_{\substack{b\to 0 \\ \alpha \to \pi/2}} \left\{ \frac{\cos(\alpha)^2}{(2\pi)^2} \langle \theta_i^2 \rangle \left[\frac{s}{b} + 2g\frac{s}{b} - 4g^2 \left(1 - \exp\left(-\frac{s}{2gb} \right) \right) \right] \right\} \\
&= \lim_{\substack{b\to 0 \\ \alpha \to \pi/2}} \left\{ \frac{2g^3 \cos(\alpha)^2}{(2\pi)^2} \langle \delta\theta_i^2 \rangle \left[\frac{s}{2g^2 b} + \frac{s}{gb} - 2 \left(1 - \exp\left(-\frac{s}{2gb} \right) \right) \right] \right\} \\
&= \frac{1}{6\pi^2} \left(\frac{s}{gb} - 2 \left(1 - \exp\left(-\frac{s}{2gb} \right) \right) \right),
\end{aligned} \tag{2.78}$$

and the autocorrelation function of the bond-directors

$$\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle = \exp\left(-2\pi^2 \langle \mathrm{Tw}(0,s)^2 \rangle \right) \tag{2.79}$$

with

$$\begin{aligned}
gb &= g_{fit} \left(\frac{k \tan(\alpha)^2}{3} \right)^{\frac{1}{3}} b = g_{fit} \left(\frac{l_p \tan(\alpha)^2}{12b} \right)^{\frac{1}{3}} b \\
&= g_{fit} \frac{l_p^{\frac{1}{3}} \tan(\alpha)^{\frac{2}{3}} b^{\frac{2}{3}}}{12^{\frac{1}{3}}} = \frac{g_{fit}}{3^{\frac{1}{3}}} l_p^{\frac{1}{3}} a^{\frac{2}{3}},
\end{aligned} \tag{2.80}$$

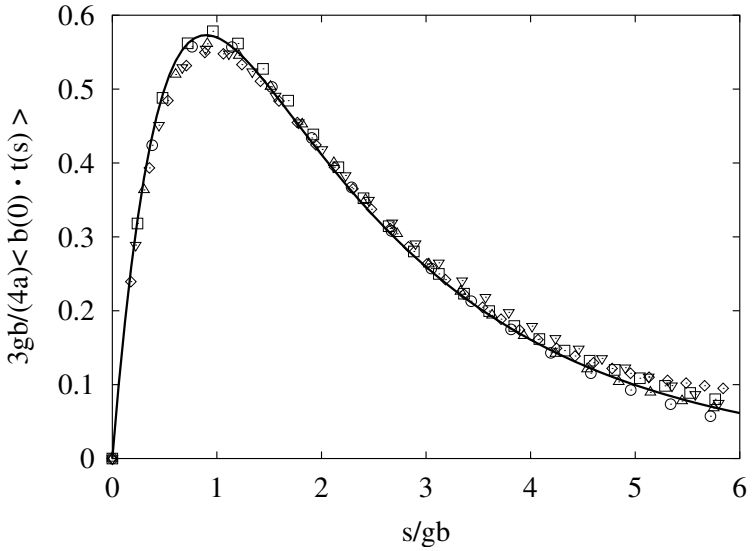


Figure 2.11: Crosscorrelation function of the bond-directors and the tangent vectors with $k = 50$ (squares), $k = 100$ (circles), $k = 200$ (upward triangles), $k = 500$ (downward triangles), $k = 1000$ (diamonds) and $\alpha = \pi/3$. The data validate the predicted functional form (solid line) for $\langle \mathbf{b}(0) \cdot \mathbf{t}(s) \rangle$ of Eq. (2.82).

where $g_{fit} = 0.56 \pm 0.05$ is the fitted prefactor for the scaling function g. a represents the strand separation of the ribbon which is given by $a = |\mathbf{b_i}| = \frac{1}{2} b \tan(\alpha)$. Hence we observe two length scales influencing the local twist structure of the ribbon: on the one hand the single strand persistence length l_p and on the other hand the strand separation a. The predicted scaling behavior of $\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle$ can be observed in the simulation data as it is shown in Fig. (2.9). Note that $\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle$ as well as all other calculated observables within this model is independent of the geometry of the triangles in contrast to the previous model where α was a fixed parameter which influenced the bending properties of the ribbon.

Eq. (2.73) can be evaluated in the same manner resulting in

$$\begin{aligned} \langle \mathbf{n}(0) \cdot \mathbf{n}(s) \rangle &= \langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle \langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle \\ &= \exp\left(-\frac{s}{l_p} - 2\pi^2 \langle \mathrm{Tw}(0,s)^2 \rangle \right). \end{aligned} \tag{2.81}$$

Eq. (2.81) shows that the autocorrelation function of the normal vectors is the product of $\langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle$ and $\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle$. For very stiff chains, the tangent correlation function gives just small corrections to the normal vector correlation function. Therefore Eq. (2.79) can be interpreted as the rigid rod limit of Eq. (2.81).

Other important structural properties of the ribbon can be extracted out of the crosscorrelation functions. $\langle \mathbf{n}(0) \cdot \mathbf{t}(s) \rangle$ and $\langle \mathbf{n}(0) \cdot \mathbf{b}(s) \rangle$ describe the mean curvature and mean twist respectively and vanish in both models for symmetry reasons. For $\langle \mathbf{b}(0) \cdot \mathbf{t}(s) \rangle$ we empirically observe the following relationship:

$$\langle \mathbf{b}(0) \cdot \mathbf{t}(s) \rangle = (2\pi)^2 a \frac{d}{ds} \langle \mathrm{Tw}(0,s)^2 \rangle \exp\left(-(2\pi)^2 \langle \mathrm{Tw}(0,s)^2 \rangle \right). \tag{2.82}$$

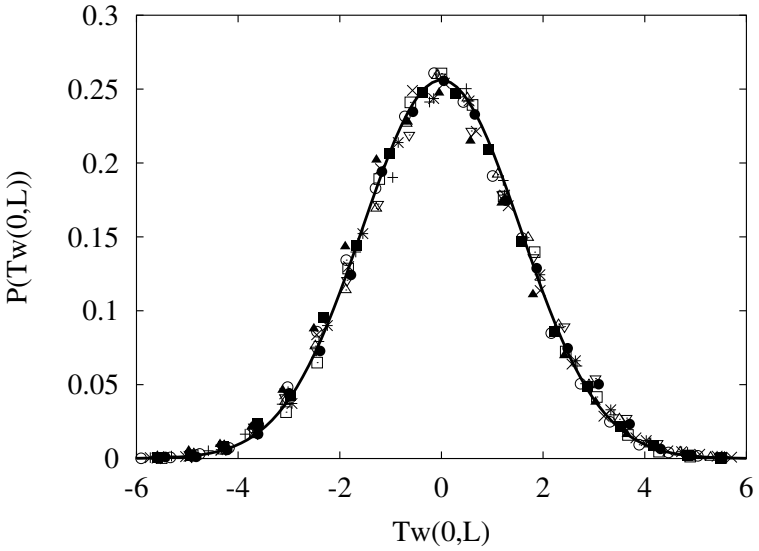


Figure 2.12: Comparison of the probability distribution functions of the total twist of the ribbon with $f = f = \{0, 0.01, 0.02, ..., 0.09\}$ and $l_p = L = 400$ with the scaling analysis for $f = 0$. The same Gaussian shape is recovered for all values of f.

Eq. (2.82) can be understood qualitatively in the following way. Due to the anisotropic rigidity of the ribbon the scalar product $\mathbf{b}(0) \cdot \mathbf{t}(s)$ is only non-zero if the chain is bent and twisted simultaneously. In case the ribbon is either solely bent or solely twisted the bond-directors are always perpendicular to the tangent vectors and the scalar product $\mathbf{b}(0) \cdot \mathbf{t}(s)$ vanishes for all s. The rate of mean twist of one helical persistence length $l_b = gb$ which defines the size of the locally existing helical structures can be calculated with Eq. (2.78) yielding

$$\sqrt{\langle \mathrm{Tw}(0, bg)^2 \rangle} = \frac{1}{3\pi^2} \left[1 - 2 \left(1 - \exp\left(-\frac{1}{2} \right) \right) \right] \approx \frac{1}{16}. \tag{2.83}$$

This corresponds to a typical twist angle of $\Psi = \frac{\pi}{8}$ using $\mathrm{Tw} = 2\pi\Psi$. Within l_b the twist rate is determined by the derivative of the mean squared twist $\frac{d}{ds}\langle \mathrm{Tw}(0, s)^2 \rangle$ which gives rise to the increasing correlation function $\langle \mathbf{b}(0) \cdot \mathbf{t}(s) \rangle$ up to the maximum value at $l_b = gb$. For larger internal distances of the chain the rate of mean twist is a random sequence of $\pm\frac{1}{16}$ so that the crosscorrelation function has to vanish and therefore decreases exponentially with $\exp\left(-(2\pi)^2 \langle \mathrm{Tw}(0, s)^2 \rangle\right)$. Fig. 2.11 compares Eq. (2.82) with our numerical data. It fully supports our argument.

2.9 Behavior under compression: Euler Buckling vs. Kinks

As discussed in section 2.4 the edge stiffness model includes local twist correlations at least on small length scales as a consequence of the correlation of the folding angles $\{\theta_i\}$. In order to understand and to quantify the effects arising from the local twist we measured the probability distribution functions of the folding angles, of the twist, and of the end-to-end-distance for different rigidities and compared the latter with the usual worm-like chain model to see which differences occur.

If there is a preference for kinking this property can be enforced by applying an additional constant force $\mathbf{F}_{buck} = f\mathbf{R}_E/R_E$ which compresses the ribbon. In addition the change in the end-to-end distance R_E caused by the buckling force should affect the twist distribution function $P(\mathrm{Tw})$ if R_E and Tw are coupled.

For small forces we calculate the change of twist $\langle \Delta \mathrm{Tw}(0,L)^2 \rangle$ under the influence of the external force $F_{buck} = f$ within the framework of linear response theory. Expanding $\langle \mathrm{Tw}(0,L)^2 \rangle$ around $f = 0$ yields

$$\langle \mathrm{Tw}(0,L)^2 \rangle = \langle \mathrm{Tw}(0,L)^2 \rangle_{f=0} + \left(\frac{d}{df} \langle \mathrm{Tw}(0,L)^2 \rangle \right)_{f=0} f + \mathcal{O}(f^2). \tag{2.84}$$

For small forces higher order terms in f can be neglected. The change of twist can be calculated by straightforward differentiation:

$$\begin{aligned}
&\langle \Delta \mathrm{Tw}(0,L)^2 \rangle = \langle \mathrm{Tw}(0,L)^2 \rangle - \langle \mathrm{Tw}(0,L)^2 \rangle_{f=0} \\
&= \left(\frac{d}{df} \langle \mathrm{Tw}(0,L)^2 \rangle \right)_{f=0} f \\
&= \left\{ \frac{d}{df} \left(\frac{\int \mathcal{D}[\theta] \mathrm{Tw}(0,L)^2 \exp(-\beta(\mathcal{H}_{tt} + fR_E))}{\int \mathcal{D}[\theta] \exp(-\beta(\mathcal{H}_{tt} + fR_E))} \right) \right\}_{f=0} f \\
&= \left\{ \frac{\int \mathcal{D}[\theta](-\beta R_E) \mathrm{Tw}(0,L)^2 \exp(-\beta(\mathcal{H}_{tt} + fR_E)) \int \mathcal{D}[\theta] \exp(-\beta(\mathcal{H}_{tt} + fR_E))}{\left(\int \mathcal{D}[\theta] \exp(-\beta(\mathcal{H}_{tt} + fR_E)) \right)^2} \right. \\
&\left. - \frac{\int \mathcal{D}[\theta] \mathrm{Tw}(0,L)^2 \exp(-\beta(\mathcal{H}_{tt} + fR_E)) \int \mathcal{D}[\theta](-\beta R_E) \exp(-\beta(\mathcal{H}_{tt} + fR_E))}{\left(\int \mathcal{D}[\theta] \exp(-\beta(\mathcal{H}_{tt} + fR_E)) \right)^2} \right\}_{f=0} f \\
&= -\beta f \left(\langle R_E \mathrm{Tw}(0,L)^2 \rangle_{f=0} - \langle R_E \rangle_{f=0} \langle \mathrm{Tw}(0,L)^2 \rangle_{f=0} \right)
\end{aligned} \tag{2.85}$$

with $\beta = 1/k_B T$. This predicts a change of the mean squared twist of the chain if a twist-stretch coupling determined by $\langle R_E \mathrm{Tw}(0,L)^2 \rangle_{f=0}$ exists. Note that $\langle R_E \mathrm{Tw}(0,L) \rangle_{f=0}$ vanishes due to symmetry reasons. The evaluation of our numerical data yields that $\langle R_E \mathrm{Tw}(0,L)^2 \rangle$ is uncorrelated, too. To quantify if higher order terms in f contribute to a change of $\langle \mathrm{Tw}(0,L)^2 \rangle$ we carried out several simulation runs with varying force strengths $f = \{0, 0.01, 0.02, ..., 0.09\}$ corresponding to $\frac{R_E(f)}{R_E(0)} = \{1, 0.95, 0.87, 0.71, 0.51, 0.36, 0.26, 0.21, 0.17, 0.15\}$.

Fig. 2.12 shows the same Gaussian shape for all measured probability distribution functions of the total twist of the ribbon $P(\mathrm{Tw}(0,L), f)$. This implicates that there is no twist-stretch coupling inherent in the system. The same is valid for the distribution function of the folding angles.

Moreover we measured the probability distribution function $P(R_E, f)$ of the end-to-end distance R_E for all applied forces f. Using the multiple histogram method developed by Ferrenberg and Swendsen [FS88] all measured histograms can then be recombined with a reweighting procedure to a single probability distribution function $P(R_E)$ with overall very good statistics. Fig. 2.13 shows $P(R_E)$ for $\mathcal{H}_{tt}$ and the worm-like chain model. Quite contrary to a shift to noticeably shorter end-to-end distances R_E as one would expect for the above described phenomena of kinks just the usual worm-like chain behavior is recovered. This indicates that the ribbon just bends under the external force in contradiction to a kink-rod structure. Another quantity which is sensitive to the presence of kinks is a three-point correlation function of the end-to-end distance R_E and the twist

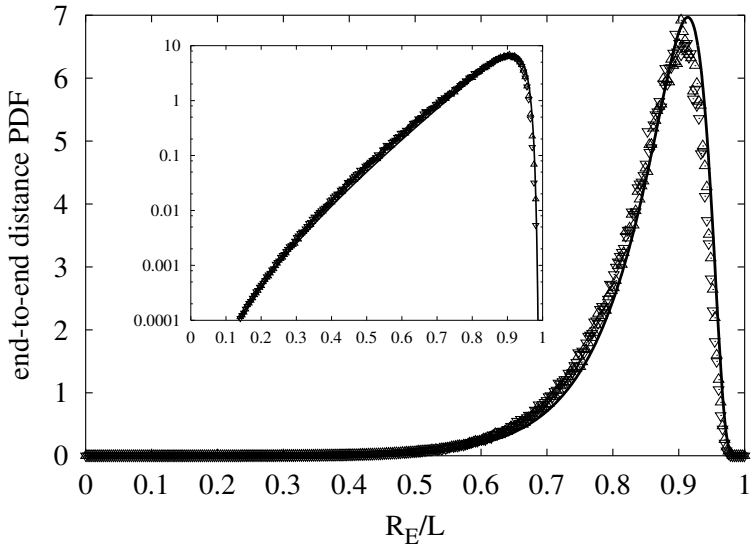


Figure 2.13: Probability distribution functions of the end-to-end distance of the edge stiffness model for different discretizations ($N = 800$ upward triangles, $N = 600$ downward triangles) with $\alpha = \frac{\pi}{4}$ and $l_p = L = 400$ calculated with the help of the multiple histogram method [FS88] and the usual worm-like chain model (solid line). The PDF of the worm-like chain model is calculated with the derived analytical expression of Wilhelm and Frey [WF96] which is valid in the limit of very stiff chains.

to the left $\mathrm{Tw}(0, \frac{L}{2})$, and to the right $\mathrm{Tw}(\frac{L}{2}, L)$ of the center of the chain. Due to the buckling force the center of the chain is labeled which means that a kink is detected if the end-to-end distances with $\mathrm{Tw}(0, \frac{L}{2})\mathrm{Tw}(\frac{L}{2}, L) < 0$ (unlike twists meeting at the center) are smaller than the end-to-end distances with $\mathrm{Tw}(0, \frac{L}{2})\mathrm{Tw}(\frac{L}{2}, L) > 0$ (like twists meeting at the center). Fig. 2.14 shows the mean end-to-end distance depending on the value of $\mathrm{Tw}(0, \frac{L}{2})\mathrm{Tw}(\frac{L}{2}, L)$ for $l_p = 200$, $L = 400$, and $f = 0$, $f = 0.03$, $f = 0.06$. We do not find an asymmetry between like and unlike twists meeting at the center as it would support the prediction of kinks made by Liverpool *et al.* [LGK98, GL00].

2.10 Summary

We have reinvestigated the mechanical properties of the model introduced by Liverpool *et al.* [LGK98, GL00] of a double-stranded semiflexible polymer and rationalized the results of our MC simulations with the help of a simple scaling argument. We recover the predicted simple exponential decay of the tangent-tangent correlation function with the single strand persistence length l_p and that $\langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle$ is independent of the separation a of the two strands, which is in addition to l_p the other relevant length scale in the problem. Also in agreement with Ref. [LGK98, GL00] we find that the helical persistence length l_b and the helical pitch P scale with $l_p^{1/3} a^{2/3}$. Qualitatively, one would expect to see oscillations in the bond-director correlation function, if $P \leq l_b$. This can be understood by calculating the rate of mean twist within $l_b = gb$, i.e. $\sqrt{\langle \mathrm{Tw}(0, gb)^2 \rangle}$. If the mean twist exceeds π an oscillatory behavior has to be observed. But our calculation gives a twist rate within $l_b = gb$

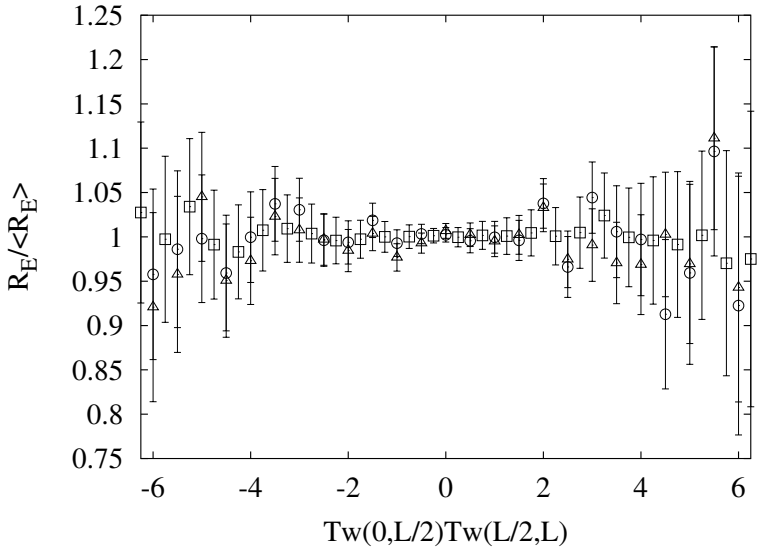


Figure 2.14: End-to-end distance R_E as a function of the product of the twist left and right of the center of the chain Tw$(0, \frac{L}{2})$Tw$(\frac{L}{2}, L)$, which is a measure for unlike (negative sign) and like (positive sign) twists meeting at the center, and as a function of the applied buckling force with $f = 0$ (squares), 0.03 (circles), 0.06 (triangles) and $l_p = 0.5L = 200$. R_E refers to the average of one interval of Tw$(0, \frac{L}{2})$Tw$(\frac{L}{2}, L)$ and $\langle R_E \rangle$ refers to the mean value of all sampled end-to-end distances. One does not find an asymmetry between end-to-end distances for like and unlike twists meeting at the center. The larger fluctuations for larger values of Tw$(0, \frac{L}{2})$Tw$(\frac{L}{2}, L)$ are the result of a poorer sampling rate.

of approximately $\pm 1/16$. For larger distances of the chain the rate of mean twist is just given by a random sequence of $\pm 1/16$ and thus cannot account for an oscillatory behavior of $\langle \mathbf{b}(0) \cdot \mathbf{b}(s) \rangle$. Liverpool *et al.* predict $P = l_b$, while our analysis indicates $P = 16 l_b$ as it is demonstrated in Fig. 2.9 (b). The authors claimed support from their own simulations, but failed to provide a quantitative comparison between their numerical and analytical results. In fact the presented oscillations seem to be ordinary fluctuations within the statistical errors. But as can be seen in Fig. 2.9 (b) the predicted functional form for the bond-director autocorrelation function is in very good agreement with our numerical data as well as with our scaling results within one helical persistence length $l_b = gb$.

Moreover our simulation results with applied constant buckling forces do not provide any evidence of the predicted tendency of kinking or the claimed twist-stretch coupling. Thus contrary to the claim made in Ref. [LGK98, GL00] the local twist structure does not suffice to explain experimental observations such as the twist-stretch coupling [SAB$^+$96, SABC98] and the kink-rod structures [KSBS93] of helical double-stranded molecules. These features require the inclusion of a spontaneous twist incorporated by an additional term in the Hamiltonian, e.g. $\mathcal{H}_{Tw} = k_{Tw} \sum_i \left(\sum_{j=i}^{i+1} \mathbf{n}_j \cdot \mathbf{b}_{j+1} - \theta_{sp,i} \right)^2$, [PR01, PR00a, PR00b, Mar97, KLNO97, YHZc00, ZZOY00, HYZc99].

Chapter 3

Simulating DNA at the base-pair level

We present a generic model for DNA at the base-pair level. We use a variant of the Gay-Berne potential to represent the stacking energy between neighboring base-pairs. The sugar-phosphate backbones are taken into account by semi-rigid harmonic springs with a non-zero spring length. The competition of these two interactions and the introduction of a simple geometrical constraint lead to a stacked right-handed B-DNA-like conformation. The mapping of the presented model to the Marko-Siggia and the Stack-of-Plates model enables us to optimize the free model parameters in order to reproduce the experimentally known observables such as persistence lengths, mean and mean squared base-pair step parameters. For the optimized model parameters we measured the critical force where the transition from B- to S-DNA occurs to be approximately 140pN. We recover an overstretched S-DNA conformation with highly inclined bases that partially preserves the stacking of successive base-pairs [MEE03].

3.1 Introduction

Following the discovery of the double helix by Watson and Crick [WC53], the structure and elasticity of DNA has been investigated on various length scales. X-ray diffraction studies of single crystals of DNA oligomers have led to a detailed picture of possible DNA conformations [DDC$^+$82, Dic92] with atomistic resolution. Information on the behavior of DNA on larger scales is accessible through NMR [Jam95] and various optical methods [MRZ82, SS86], video [PQSC94] and electron microscopy [BWC90]. An interesting development of the last decade are nanomechanical experiments with *individual* DNA molecules [SFB92, SCB96, CLH$^+$96, ERBH97, ABLC98] which, for example, reveal the intricate interplay of supercoiling on large length scales and local denaturation of the double-helical structure.

Experimental results are usually rationalized in the framework of two types of models: base-pair steps and variants of the continuum elastic worm-like chain. The first, more local, approach describes the relative location and orientation of neighboring base-pairs in terms of intuitive parameters such as twist, rise, slide, roll etc. [CD84, DBC$^+$89, LO99, OBB$^+$01]. In particular, it provides a mechanical interpretation of the biological function of particular sequences [CD99]. The second approach models DNA on length scales beyond the helical pitch as a worm-like chain (WLC) with empirical parameters describing the resistance to bending, twisting and stretching [MS94, MS95c]. The results are in remarkable agreement with the nanomechanical experiments mentioned above [PSLC95]. WLC models are commonly used in order to address biologically important phenomena

such as supercoiling [CW90, SO92, CL94] or the wrapping of DNA around histones [SWBG01]. In principle, the two descriptions of DNA are linked by a systematic coarse-graining procedure: From given (average) values of rise, twist, slide etc. the shape of the corresponding helix on large scales [CD84, HC97b, CD99] can be reconstructed. Similarly, the elastic constant characterizing the continuum model are related to the local elastic energies in a stack-of-plates model [OKLN98].

Difficulties are encountered in situations which cannot be described by a linear response analysis around the undisturbed (B-DNA) ground state. This situation arises routinely during cellular processes and is therefore of considerable biological interest [CD99]. A characteristic feature, observed in many nanomechanical experiments, is the occurrence of plateaus in force-elongation curves [SCB96, CLH$^+$96, ABLC98]. These plateaus are interpreted as structural transitions between microscopically distinct states. While atomistic simulations have played an important role in identifying possible local structures such as S- and P-DNA [CLH$^+$96, ABLC98], this approach is limited to relatively short DNA segments containing several dozen base-pairs. The behavior of longer chains is interpreted on the basis of stack-of-plates models with step-type dependent parameters and free energy penalties for non-B steps. Realistic force-elongation are obtained by a suitable choice of parameters and as the consequence of constraints for the total extension and twist (or their conjugate forces) [ALCM01]. Similar models describing the non-linear response of B-DNA to stretching [HYZc99] or untwisting [BCP99, CM99] predict stability thresholds for B-DNA due to a combination of more realistic, short-range interaction potentials for rise with twist-rise coupling enforced by the sugar-phosphate backbones.

Clearly, the agreement with experimental data will increase with the amount of details which is faithfully represented in a DNA model. However, there is strong evidence both from atomistic simulations [BFLG99] as well as from the analysis of oligomer crystal structures [HC97a] that the base-pair level provides a sensible compromise between conceptual simplicity, computational cost and degree of reality. While Lavery et al. [BFLG99] have shown that the base-pairs effectively behave as rigid entities, the results of El Hassan and Calladine [HC97a] and of Hunter et al. [HL97, Hun93] suggest that the dinucleotide parameters observed in oligomer crystals can be understood as a consequence of van-der-Waals and electrostatic interactions between the neighboring base-pairs and constraints imposed by the sugar-phosphate backbone.

The purpose of the present chapter is the introduction of a class of “DNA-like”-molecules with simplified interactions resolved at the base or base-pair level. In order to represent the stacking interactions between neighboring bases (base-pairs) we use a variant [EE03] of the Gay-Berne potential [GB81] used in studies of discotic liquid crystals. The sugar-phosphate backbones are reduced to semi-rigid springs connecting the edges of the disks/ellipsoids. Using Monte-Carlo simulations we explore the local stacking and the global helical properties as a function of the model parameters. In particular, we measure the effective parameters needed to describe our systems in terms of stack-of-plates (SOP) and worm-like chain models respectively. This allows us to construct models of our systems which faithfully represent the equilibrium structure, fluctuations and linear response of DNA. At the same time we preserve the possibility of local structural transitions, e.g. in response to external forces.

This chapter is organized as follows. In the second section we introduce the base-pair parameters to discuss the helix geometry in terms of these variables. Furthermore we discuss how to translate the base-pair parameters in macroscopic variables such as bending and torsional rigidity. In the third section we introduce the model and discuss the methods (MC simulation, energy minimization) that we use to explore its behavior. In the fourth section we present the resulting equilibrium structures, the persistence lengths as a function of the model parameters, and the behavior under stretching.

3.2 Theoretical Background

3.2.1 The worm-like chain model revisited: WLC with stretching modulus and WLC under tension

We already introduced in section 2.2.1 the Hamiltonian of the WLC. Let us examine the discretized version of Eq. (2.1) with an additional stretching term accounting for fluctuations of the bond length of the segment i along the $\mathbf{t}_i$-direction:

$$\frac{\mathcal{H}_{WLC}}{k_BT} = \frac{k}{2}\sum_{i=1}^{N-1}\mathbf{t}_i\cdot\mathbf{t}_{i+1} + \frac{\gamma}{2}\sum_{i=1}^{N-1}\frac{(b_i-b_0)^2}{b_0} \tag{3.1}$$

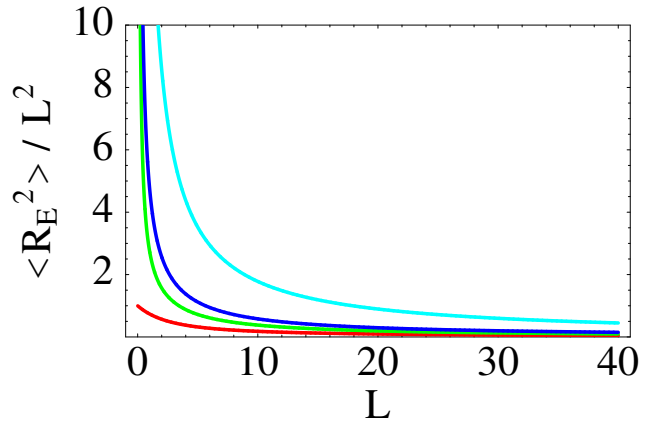
with $\vec{b}_i = \vec{r}_{i+1} - \vec{r}_i = b_i\mathbf{t}_i$ being the ith bond vector with length b_i, $\vec{r}_i$ representing the position vectors, and b_0 being the length around which the segments fluctuate. Thus we only allow for one-dimensional longitudinal fluctuations (along the $\mathbf{t}_i$-direction) in the segment length b_i. In the following we are interested in the change of the mean-square end-to-end distance $\langle R_E^2\rangle$ compared to the incompressible WLC model. Under the assumption that b_i and $\mathbf{t}_i$ are uncorrelated it yields

$$\begin{aligned}
\langle R_E^2\rangle &= \sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\langle\vec{b}_i\cdot\vec{b}_j\rangle \\
&= \sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\langle b_i\, b_j\,\mathbf{t}_i\cdot\mathbf{t}_j\rangle \\
&= \sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\langle b_i\, b_j\rangle\langle\mathbf{t}_i\cdot\mathbf{t}_j\rangle \\
&= \sum_{i=1}^{N-1}\langle b_i^2\rangle + \sum_{i=1}^{N-1}\sum_{j\neq 1}^{N-1}\langle b_i\rangle\langle b_j\rangle\langle\mathbf{t}_i\cdot\mathbf{t}_j\rangle \\
&= \sum_{i=1}^{N-1}(\langle b_i^2\rangle - \langle b_i\rangle^2) + \sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\langle b_i\rangle^2\langle\mathbf{t}_i\cdot\mathbf{t}_j\rangle \\
&= N(\langle b_i^2\rangle - \langle b_i\rangle^2) + \sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\langle b_i\rangle^2\langle\mathbf{t}_i\cdot\mathbf{t}_j\rangle
\end{aligned} \tag{3.2}$$

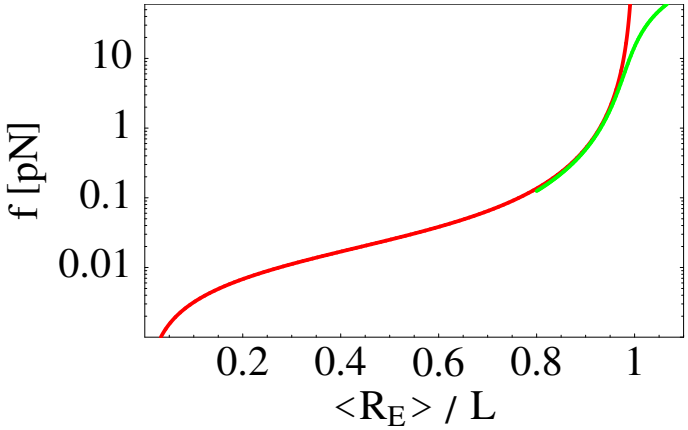
where the first term is the additional contribution from the stretching modulus of the segments and the second part corresponds to the usual worm-like chain result. The first and second moment of the bond vectors is given by

$$\langle b_i\rangle = \frac{\int db\, b_i \exp\left(-\frac{\gamma}{2k_BT}\sum_{i=1}^{N-1}\frac{(b_i-b_0)^2}{b_0}\right)}{\int db\, \exp\left(-\frac{\gamma}{2k_BT}\sum_{i=1}^{N-1}\frac{(b_i-b_0)^2}{b_0}\right)} = b_0 \tag{3.3}$$

$$\langle b_i^2\rangle = \frac{\int db\, b_i^2 \exp\left(-\frac{\gamma}{2k_BT}\sum_{i=1}^{N-1}\frac{(b_i-b_0)^2}{b_0}\right)}{\int db\, \exp\left(-\frac{\gamma}{2k_BT}\sum_{i=1}^{N-1}\frac{(b_i-b_0)^2}{b_0}\right)} = b_0\left(b_0+\frac{1}{\gamma}\right). \tag{3.4}$$



(a) Mean squared end-to-end distance for the inextensible (red) and extensible WLC model ($1/(\gamma l_p) = 0.5$: green, $1/(\gamma l_p) = 0.25$: blue, $1/(\gamma l_p) = 0.0625$: cyan).



(b) Illustration of the calculated stress-strain relations for the WLC model (red) and the WLC model extended by a linear stretching term (green).

Figure 3.1: Stress-strain relations and mean squared end-to-end distance for the inextensible and extensible WLC model.

In the limit of long chains the following relationship for the mean-square end-to-end distance is obtained

$$\langle R_E^2 \rangle = 2Ll_p - 2l_p^2 \left(1 - \exp\left(-\frac{L}{l_p}\right)\right) + \frac{L}{\gamma} \tag{3.5}$$

with $L = Nb_0$ being the contour length and $l_p = kb_0$ denoting the bending persistence length. There is just an additional term $\frac{L}{\gamma}$ to the usual worm-like chain result of Eq. (2.3). Note that γ is

determined by

$$\gamma = \frac{b_0}{\langle (b_i - b_0)^2 \rangle}. \tag{3.6}$$

The energy of a worm-like chain with bending persistence length l_p in the presence of an external stretching force f in $\mathbf{z}$-direction is given by

$$\frac{\mathcal{H}_{WLC,stretch}}{k_B T} = \frac{l_p}{2} \int_0^L ds \left(\frac{d\mathbf{t}(s)}{ds} \right)^2 - f \int_0^L \mathbf{z} \cdot \mathbf{t}(s). \tag{3.7}$$

If the force f in Eq. (3.7) is used as a Lagrange multiplier to fix the end-to-end extension $R_E = \mathbf{z} \cdot (\vec{r}(L) - \vec{r}(0))$ of the chain the free energy $\mathcal{F}$ corresponds to the quantum-mechanical ground state energy of a dipolar rotator with moment of inertia l_p subject to an electric field f [BMSS94, MS95c]. The force extension relation can be solved numerically

$$\frac{\langle R_E \rangle}{L} = \frac{k_B T}{L} \frac{d}{df} \ln(\mathcal{Z}) = -\frac{d}{df} \frac{\mathcal{F}}{L}, \tag{3.8}$$

where $\mathcal{Z}$ is the partition function given by the path integral

$$\mathcal{Z} = \mathcal{Z}(\mathbf{t}_0, 0; \mathbf{t}_1, L) = \int \mathcal{D}\mathbf{t} \, \exp\left(\frac{\mathcal{H}_{WLC,stretch}}{k_B T} \right). \tag{3.9}$$

The large and small stretching force regime admit analytical asymptotic solutions [MS95c]

$$\frac{\langle R_E \rangle}{L} = \begin{cases} \frac{2 f l_p}{3 k_B T} & f l_p \ll 1 \\ 1 - \left(\frac{k_B T}{4 f l_p} \right)^{\frac{1}{2}} & f l_p \gg 1 \end{cases} \tag{3.10}$$

which can be combined by the approximative interpolation formula

$$\frac{f l_p}{k_B T} = \frac{1}{4} \left(1 - \frac{\langle R_E \rangle}{L} \right)^{-2} + \frac{\langle R_E \rangle}{L} - \frac{1}{4}. \tag{3.11}$$

Since for large stretching forces the longitudinal component (along the stretching direction) of the tangent vector $\mathbf{t}_\perp$ does not contribute to the elastic energy the bending persistence length l_p of a worm-like chain under tension can be evaluated as a function of the applied force f and the average of $\mathbf{t}_\perp^2$ [MS95c]:

$$l_p = \frac{1}{f \langle \mathbf{t}_\perp^2 \rangle^2}. \tag{3.12}$$

Experimental force-extension data show that for forces larger than 15pN the measured extensions exceed the contour length of the DNA [SCB96, CLH$^+$96]. These deviations from the inextensible worm-like chain case indicate that the constraint of fixed chain length has to be released by adding a linear stretching term with stretch modulus γ into the energy (see Eq. (3.1)). By doing so one obtains in the asymptotic stretching regime with $f l_p \gg 1$ the following stress-strain relation [Odi95, SN02b]

$$\frac{\langle R_E \rangle}{L} = 1 - \left(\frac{k_B T}{4 f l_p} \right)^{\frac{1}{2}} + \frac{f}{\gamma}. \tag{3.13}$$

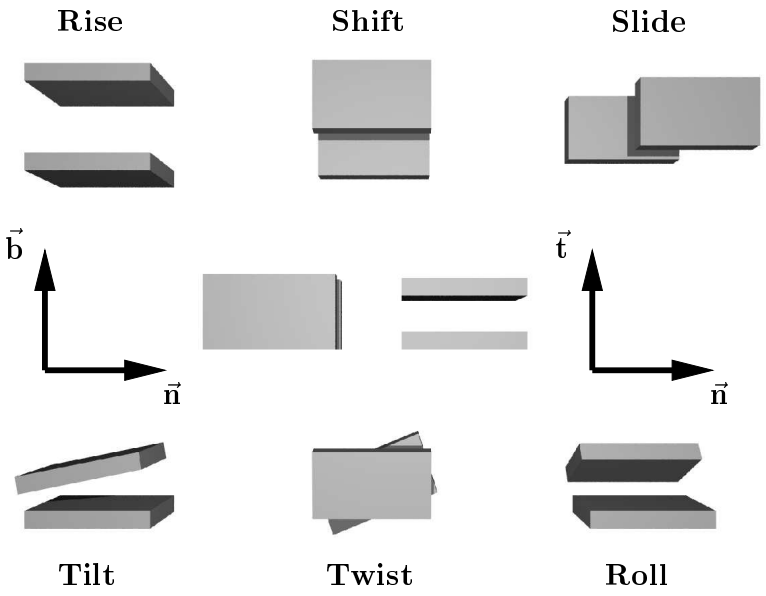


Figure 3.2: Illustration of all six base-pair step parameters and the corresponding coordinate system. The involved mathematics to calculate the step parameters is discussed in appendix B.

3.2.2 Helix geometry

To resolve and interpret X-ray diffraction studies on DNA oligomers the relative position and orientation of successive base-pairs are analyzed in terms of Rise (Ri), Slide (Sl), Shift (Sh), Twist (Tw), Roll (Ro), and Tilt (Ti) [BPO94] (see Fig. 3.2). In order to illustrate the relation between these local parameters and the overall shape of the resulting helix we discuss a simple geometrical model where DNA is viewed as a twisted ladder where all bars lie in one plane. For vanishing bending angles with Ro $=$ Ti $= 0$ each step is characterized by four parameters: Ri, Sl, Sh, and Tw [CD99]. Within the given geometry a base-pair can be characterized by its position $\mathbf{r}$ and the angle of its main axis with the $\mathbf{n}$/$\mathbf{b}$-axis ($\mathbf{n}$ points into the direction of the large axis, $\mathbf{b}$ points into the direction of the small axis, and $\mathbf{t}$, representing the tangent vector of the resulting helix, is perpendicular to the $\mathbf{n}$-$\mathbf{b}$- plane as it is illustrated in Fig. 3.2). At each step the center points are displaced by a distance $\sqrt{\text{Sl}^2 + \text{Sh}^2}$ in the $\mathbf{n} - \mathbf{b}-$plane. The angle between successive steps is equal to the twist angle and the center points are located on a helix with radius $r = \sqrt{\text{Sl}^2 + \text{Sh}^2}/(2\sin(\text{Tw}/2))$.

In the following we study the consequences of imposing a simple constraint on the bond lengths l_1 and l_2 representing the two sugar phosphate backbones (the rigid bonds connect the right and left edges of the bars along the $\mathbf{n}$-axis respectively). Ri is the typical height of a step which we will try to impose on the grounds that it represents the preferred stacking distance of neighboring base pairs. We choose Ri $= 3.3$Å corresponding to the B-DNA value. One possibility to fulfill the constraint $l_1 = l_2 = l = 6$Å is pure twist. In this case a relationship of the twist angle and the width of the

base-pairs d, the backbone length l and the imposed rise Ri is obtained:

$$\mathrm{Tw} = \arccos\left(\frac{d^2 - 2l^2 + 2\mathrm{Ri}^2}{d^2}\right). \tag{3.14}$$

Another possibility is to keep the rotational orientation of the base pair (Tw $= 0$), but to displace its center in the **n**-**b**-plane, in which case $\mathrm{Ri}^2 + \mathrm{Sl}^2 + \mathrm{Sh}^2 \equiv l^2$. With Sh $= 0$, it results in a skewed ladder with skew angle $\arcsin(\mathrm{Sl}/l)/\pi$ [CD99].

The general case can be solved as well. In a first step a general condition is obtained that needs to be fulfilled by any combination of Sh, Sl, and Tw independently of Ri. For non-vanishing Tw this yields a relation between Sh and Sl:

$$\tan(\mathrm{Tw}) = \frac{\mathrm{Sh}}{\mathrm{Sl}}. \tag{3.15}$$

Using Eq. (3.15) the general equation can finally be solved:

$$\mathrm{Sl} = \frac{1}{\sqrt{2}}\left[\cos(\frac{\mathrm{Tw}}{2})^2\sqrt{\sec(\frac{\mathrm{Tw}}{2})^2(2l^2 - d^2 - \mathrm{Ri}^2)}\right]. \tag{3.16}$$

Eq. (3.16) is a result of the mechanical coupling of slide and shift respectively and twist due to the backbones. Treating the rise again as a constraint the twist is reduced for increasing slide or shift motion. The center-center distance c of two neighboring base-pairs is given by

$$c = \sqrt{\mathrm{Ri}^2 + \mathrm{Sl}^2\,(1 + \tan(\mathrm{Tw})^2)}. \tag{3.17}$$

For Tw $= 0$ and a given value of Ri the center-center distance is equal to the backbone length l and for $\mathrm{Tw} = \arccos\left((d^2 - 2l^2 + 2\mathrm{Ri}^2)/d^2\right)$ one obtains $c = \mathrm{Ri}$.

3.2.3 Thermal fluctuations

In this section we discuss how to calculate the effective coupling constants of a harmonic system valid within linear response theory describing the couplings of the base-pair step parameters along the chain. Furthermore we show how to translate measured mean and mean squared values of the 6 microscopic base-pair step parameters into macroscopic observables such as bending and torsional persistence length. This provides the linkage between the two descriptions: WLC (worm-like chain) versus SOP (stack-of-plates) model.

Within linear response theory it should be possible to map our model onto a Gaussian system where all translational and rotational degrees of freedom are harmonically coupled. We refer to this model as the stack-of-plates (SOP) model [OKLN98]. The effective coupling constants are given by the second derivatives of the free energy in terms of base-pair step variables around the equilibrium configuration. This yields 6×6 matrices $\mathcal{K}^{nm}$ describing the couplings of the base-pair step parameters of neighboring base-pairs along the chain:

$$\mathcal{K}^{nm} = \frac{\partial^2 \mathcal{F}}{\partial x_i^n \partial x_j^m}. \tag{3.18}$$

Therefor one can calculate the $(N-1) \times (N-1)$ correlation matrix $\mathcal{C}$ in terms of base-pair step parameters. N is thereby the number of base-pairs.

$$\langle \mathcal{C} \rangle = \begin{pmatrix} \mathcal{K}^{11} & \mathcal{K}^{12} & \mathcal{K}^{13} & \mathcal{K}^{14} & \dots \\ \mathcal{K}^{12} & \mathcal{K}^{22} & \mathcal{K}^{23} & \mathcal{K}^{24} & \dots \\ & & \ddots & & \end{pmatrix}^{-1}. \tag{3.19}$$

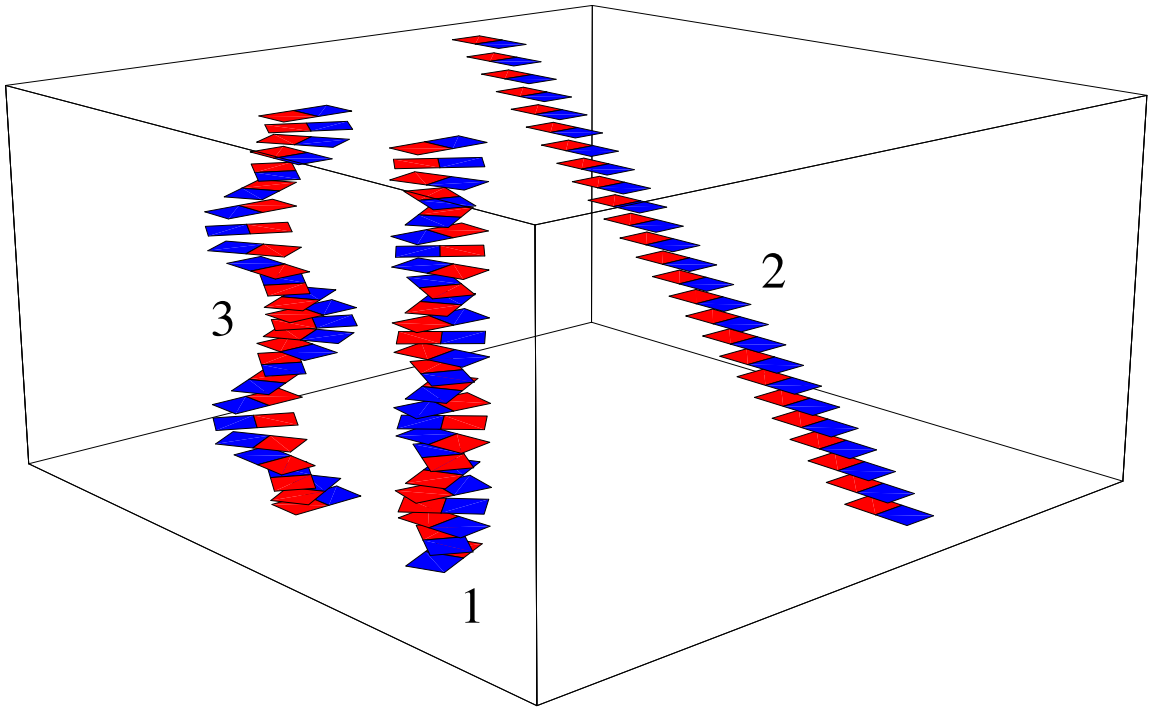


Figure 3.3: Illustration of DNA geometry for a diameter of $d = 16$Å: (1) Twisted ladder with Sl $=$ Sh $= 0$, Ri $= 3.3$Å, Tw $\approx 2\pi/10$, (2) Skewed ladder with Tw $=$ Sh $= 0$, Ri $= 3.4$Å, Sl ≈ 5.0Å, (3) Helix with Tw $= 2\pi/12$, Ri $= 3.4$Å, Sl ≈ 2.7Å, Sh ≈ 1.6Å.

The inversion of $\mathcal{C}$ results in a generalized connectivity matrix with effective coupling constants as entries.

The following considerations are based on the assumption that one only deals with nearest-neighbor interactions. Then successive base-pair steps are independent of each other and the calculation of the orientational correlation matrix becomes feasible. In the absence of spontaneous displacements (Sl $=$ Sh $= 0$) and spontaneous bending angles (Ti $=$ Ro $= 0$) as it is the case for B-DNA going from one base-pair to the neighboring implies three operations. In order to be independent of the reference base-pair one first rotates the respective base-pair into the mid-frame with $\mathcal{R}(Tw_{sp}/2)$ ($\mathcal{R}$ is a rotation matrix, Tw_{sp} denotes the spontaneous twist), followed by a subsequent overall rotation in the mid-frame which takes the thermal motion of Ro, Ti and Tw into account

$$\mathcal{A} = \begin{pmatrix} \mathbf{t_i} \cdot \mathbf{t_{i+1}} & \mathbf{t_i} \cdot \mathbf{b_{i+1}} & \mathbf{t_i} \cdot \mathbf{n_{i+1}} \\ \mathbf{b_i} \cdot \mathbf{t_{i+1}} & \mathbf{b_i} \cdot \mathbf{b_{i+1}} & \mathbf{b_i} \cdot \mathbf{n_{i+1}} \\ \mathbf{n_i} \cdot \mathbf{t_{i+1}} & \mathbf{n_i} \cdot \mathbf{b_{i+1}} & \mathbf{n_i} \cdot \mathbf{n_{i+1}} \end{pmatrix} \tag{3.20}$$

with

$$
\begin{aligned}
\mathcal{A}_{11} &= \cos(\Omega) + (1-\cos(\Omega))\frac{\mathrm{Tw}^2}{\Omega^2} \\
\mathcal{A}_{12} &= (1-\cos(\Omega))\frac{\mathrm{Tw}\,\mathrm{Ti}}{\Omega^2} - \frac{\mathrm{Ro}}{\Omega}\sin(\Omega) \\
\mathcal{A}_{13} &= (1-\cos(\Omega))\frac{\mathrm{Tw}\,\mathrm{Ro}}{\Omega^2} + \frac{\mathrm{Ti}}{\Omega}\sin(\Omega) \\
\mathcal{A}_{21} &= (1-\cos(\Omega))\frac{\mathrm{Tw}\,\mathrm{Ti}}{\Omega^2} + \frac{\mathrm{Ro}}{\Omega}\sin(\Omega) \\
\mathcal{A}_{22} &= \cos(\Omega) + (1-\cos(\Omega))\frac{\mathrm{Ti}^2}{\Omega^2} \\
\mathcal{A}_{23} &= (1-\cos(\Omega))\frac{\mathrm{Ti}\,\mathrm{Ro}}{\Omega^2} - \frac{\mathrm{Tw}}{\Omega}\sin(\Omega) \\
\mathcal{A}_{31} &= (1-\cos(\Omega))\frac{\mathrm{Tw}\,\mathrm{Ro}}{\Omega^2} - \frac{\mathrm{Ti}}{\Omega}\sin(\Omega) \\
\mathcal{A}_{32} &= (1-\cos(\Omega))\frac{\mathrm{Ti}\,\mathrm{Ro}}{\Omega^2} + \frac{\mathrm{Tw}}{\Omega}\sin(\Omega) \\
\mathcal{A}_{33} &= \cos(\Omega) + (1-\cos(\Omega))\frac{\mathrm{Ro}^2}{\Omega^2}
\end{aligned}
\tag{3.21}
$$

and a final rotation due to the spontaneous twist $\mathcal{R}(Tw_{sp}/2)$. Note that $\Omega = \sqrt{\mathrm{Tw}^2 + \mathrm{Ti}^2 + \mathrm{Ro}^2}$ and that $\mathcal{R}(Tw_{sp}/2)$ is given by

$$
\mathcal{R}(Tw_{sp}/2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\left(Tw_{sp}/2\right) & -\sin\left(Tw_{sp}/2\right) \\ 0 & \sin\left(Tw_{sp}/2\right) & \cos\left(Tw_{sp}/2\right) \end{pmatrix}. \tag{3.22}
$$

The orientational correlation matrix between two neighboring base pairs can be written as

$$
\langle \mathcal{O}_{i\,i+1} \rangle = \mathcal{R}(Tw_{sp}/2)\,\langle \mathcal{A} \rangle\,\mathcal{R}(Tw_{sp}/2) \tag{3.23}
$$

$\mathcal{A}$ describes the fluctuations around the mean values. Under the assumption of small angles Ro, Ti and Tw it yields

$$
\begin{aligned}
\langle \mathcal{O}_{i\,i+1} \rangle_{11} &= 1 - \langle \mathrm{Ti}^2 \rangle - \langle \mathrm{Ro}^2 \rangle \\
\langle \mathcal{O}_{i\,i+1} \rangle_{12} &= 0 \\
\langle \mathcal{O}_{i\,i+1} \rangle_{13} &= 0 \\
\langle \mathcal{O}_{i\,i+1} \rangle_{21} &= 0 \\
\langle \mathcal{O}_{i\,i+1} \rangle_{22} &= \frac{1}{2}\left(\langle \mathrm{Ti}^2 \rangle - \langle \mathrm{Ro}^2 \rangle - \cos(\mathrm{Tw}_{sp})(2\langle \mathrm{Tw}^2 \rangle + \langle \mathrm{Ti}^2 \rangle + \langle \mathrm{Ro}^2 \rangle - 2)\right) \\
\langle \mathcal{O}_{i\,i+1} \rangle_{23} &= \frac{1}{2}\left(2\langle \mathrm{Tw}^2 \rangle + \langle \mathrm{Ti}^2 \rangle + \langle \mathrm{Ro}^2 \rangle - 2\right)\sin(\mathrm{Tw}_{sp}) \\
\langle \mathcal{O}_{i\,i+1} \rangle_{31} &= 0 \\
\langle \mathcal{O}_{i\,i+1} \rangle_{32} &= -\frac{1}{2}\left(2\langle \mathrm{Tw}^2 \rangle + \langle \mathrm{Ti}^2 \rangle + \langle \mathrm{Ro}^2 \rangle - 2\right)\sin(\mathrm{Tw}_{sp}) \\
\langle \mathcal{O}_{i\,i+1} \rangle_{33} &= \frac{1}{2}\left(\langle \mathrm{Ro}^2 \rangle - \langle \mathrm{Ti}^2 \rangle - \cos(\mathrm{Tw}_{sp})(2\langle \mathrm{Tw}^2 \rangle + \langle \mathrm{Ti}^2 \rangle + \langle \mathrm{Ro}^2 \rangle - 2)\right).
\end{aligned}
\tag{3.24}
$$

As a consequence of the independence of successive base-pair step parameters one finds $\langle \mathcal{O}_{i\,j} \rangle = (\mathcal{R}(Tw_{sp}/2)\,\langle \mathcal{A} \rangle\,\mathcal{R}(Tw_{sp}/2))^{j-i}$ where the matrix product is carried out in the eigenvector basis of $\mathcal{R}(Tw_{sp}/2)\,\langle \mathcal{A} \rangle\,\mathcal{R}(Tw_{sp}/2)$. In the end a relationship of the mean and mean squared local base-pair step parameters and the bending and torsional persistence length is obtained. The calculation yields an exponentially decaying tangent-tangent correlation function

$$\langle \mathbf{t}(0) \cdot \mathbf{t}(s) \rangle = \exp(-s/l_p) \tag{3.25}$$

with a bending persistence length

$$l_p = \frac{2\langle \mathrm{Ri} \rangle}{(\langle \mathrm{Ti}^2 \rangle + \langle \mathrm{Ro}^2 \rangle)}. \tag{3.26}$$

In the following we will calculate the torsional persistence length. Making use of a simple relationship between the local twist and the base-pair orientations turns out to be more convenient than the transfer matrix approach.

The (bi)normal-(bi)normal correlation function is an exponentially decaying function with an oscillating term depending on the helical repeat length $h = p\langle \mathrm{Ri} \rangle$ and the helical pitch $p = 2\pi/\langle \mathrm{Tw} \rangle$ respectively, namely

$$\langle \mathbf{n}(0) \cdot \mathbf{n}(s) \rangle = \exp(-s/l_n)\cos(2\pi\, s/h). \tag{3.27}$$

The torsional persistence length $l_n = l_b$ can be calculated then in the following way. It can be shown that the twist angle Tw of two successive base-pairs is related to the orientations $\{\mathbf{t}, \mathbf{b}, \mathbf{n}\}$ and $\{\mathbf{t}', \mathbf{b}', \mathbf{n}'\}$ through

$$\cos(\mathrm{Tw}) = \frac{\mathbf{n} \cdot \mathbf{n}' + \mathbf{b} \cdot \mathbf{b}'}{1 + \mathbf{t} \cdot \mathbf{t}'} \tag{3.28}$$

(see appendix C). Taking the mean and using the fact that the orientational correlation functions and twist correlation function decay exponentially

$$\exp(-1/l_{Tw}) = \frac{2\exp(-1/l_n)}{1 + \exp(-1/l_p)} \tag{3.29}$$

yields in the case of stiff filaments a simple expression of l_n depending on l_p and l_{Tw}:

$$\frac{l_n}{2} = \frac{l_b}{2} = \left(\frac{2}{l_{Tw}} + \frac{1}{l_p} \right)^{-1}, \tag{3.30}$$

where the twist persistence length is defined as

$$l_{Tw} = \frac{\langle \mathrm{Ri} \rangle}{\langle \mathrm{Tw}^2 \rangle}. \tag{3.31}$$

3.3 Model and methods

Qualitatively the geometrical considerations suggest a B-DNA like ground state and the transition to a skewed ladder conformation under the influence of a sufficiently high stretching force, because this provides the possibility to lengthen the chain and to partially conserve stacking. Quantitative modeling requires the specification of a Hamiltonian.

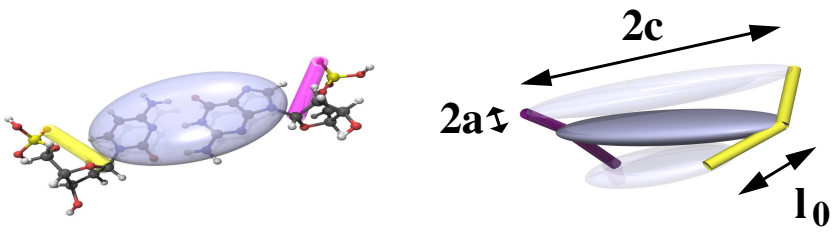


Figure 3.4: (left) Illustration of the underlying idea. The base-pairs are represented as rigid ellipsoids. The sugar-phosphate backbone is treated as semi-rigid springs connecting the edges of the ellipsoid. (right) Introduced interactions lead to a right-handed twisted structure.

3.3.1 Introduction of the Hamiltonian

The observed conformation of a dinucleotide base-pair step represents a compromise between (i) the base stacking interactions (bases are hydrophobic and the base-pairs can exclude water by closing the gap in between them) and (ii) the preferred backbone conformation (the equilibrium backbone length restricts the conformational space accessible to the base-pairs) [PH98]. Packer and Hunter [PH98] have shown that roll, tilt and rise are backbone-independent parameters. They depend mainly on the stacking interaction of successive base-pairs. In contrast twist is solely controlled by the constraints imposed by a rigid backbone. Slide and shift are sequence-dependent. While it is possible to introduce sequence dependant effects into our model, they are ignored in the present work.

We propose a generic model for DNA where the molecule is described as a stack of thin, rigid ellipsoids representing the base-pairs (Fig. 3.4). The shape of the ellipsoids is given by three radii a, b, c of the main axes in the body frames which can be used to define a structure matrix

$$\mathcal{S} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}. \tag{3.32}$$

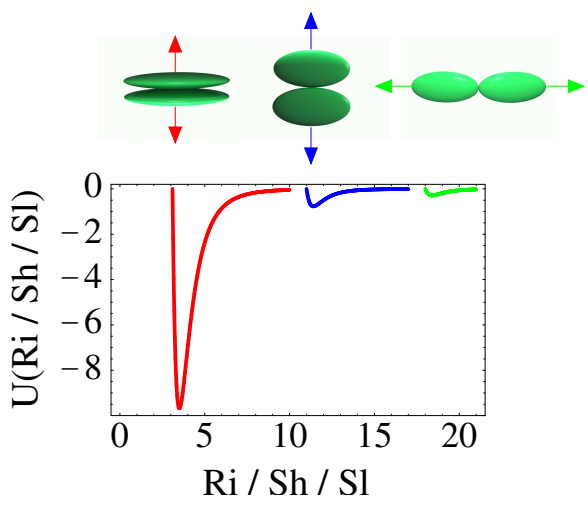
$2a$ corresponds to the thickness, $2b$ to the depth which is a free parameter in the model, and $2c = 18$Å to the width of the ellipsoid which is fixed to the diameter of a B-DNA helix. The thickness $2a$ will be chosen in such a way that the minimum center-center distance for perfect stacking reproduces the experimentally known value of 3.3Å.

The attraction and the excluded volume between the base-pairs is modeled by a variant of the Gay-Berne potential [EE03] for ellipsoids of arbitrary shape $\mathcal{S}_i$, relative position $\vec{r}_{12}$ and orientation $\mathbf{A}_i$. The potential can be written as a product of three terms:

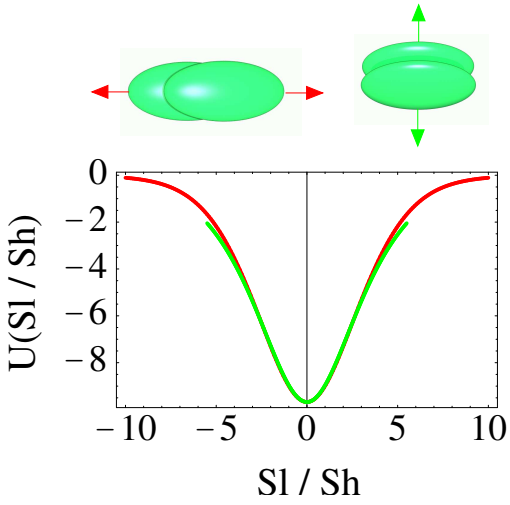
$$U(\mathbf{A}_1, \mathbf{A}_2, \vec{r}_{12}) = U_{\mathrm{r}}(\mathbf{A}_1, \mathbf{A}_2, \vec{r}_{12}) \eta_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12})\, \chi_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12}). \tag{3.33}$$

The first term controls the distance dependence of the interaction and has the form of a simple LJ potential

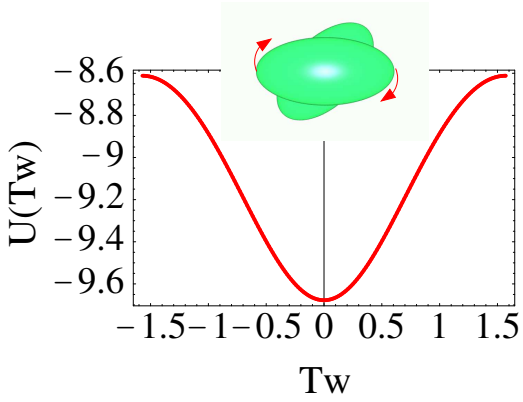
$$U_{\mathrm{r}} = 4\epsilon_{GB} \left(\left(\frac{\sigma}{h + \gamma\sigma} \right)^{12} - \left(\frac{\sigma}{h + \gamma\sigma} \right)^{6} \right) \tag{3.34}$$



(a) Distance dependence of the potential energy for all possible pole contacts.



(b) Dependence of the potential energy on slide (green) and shift (blue) motion around the stacked conformation (Ri=3.5, Sl=Sh=Tw=Ti=Ro=0.0).



(c) Twist angular dependence of the potential energy for two stacked ellipsoids

Figure 3.5: Distance and angular dependence of the Gay-Berne potential between two similar oblate ellipsoids with semi-axes (3.3,10,18) and $\gamma = 1.0$, $\sigma = 3.3$.

where the interparticle distance r is replaced by the distance h of closest approach between the two bodies:

$$h \equiv \min(|\vec{r}_i - \vec{r}_j|) \; \forall (i,j) \tag{3.35}$$

with $i \in Body\,1$ and $j \in Body\,2$. The range of interaction is controlled by an atomistic length scale $\sigma = 3.3$Å, representing the effective diameter of a base-pair.

In general, the calculation of h is non-trivial. We use the following approximative calculation

scheme which is usually employed in connection with the Gay-Berne potential:

$$h(\mathbf{A}_1, \mathbf{A}_2, \vec{r}_{12}) = r_{12} - \sigma_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12}) \tag{3.36}$$

$$\sigma_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12}) = [\frac{1}{2}\hat{r}_{12}^T \mathbf{G}_{12}^{-1}(\mathbf{A}_1, \mathbf{A}_2)\hat{r}_{12}]^{-1/2} \tag{3.37}$$

$$\mathbf{G}_{12}(\mathbf{A}_1, \mathbf{A}_2) = \mathbf{A}_1^T \mathbf{S}_1^2 \mathbf{A}_1 + \mathbf{A}_2^T \mathbf{S}_2^2 \mathbf{A}_2. \tag{3.38}$$

In the present case of oblate objects with rather perfect stacking behavior Eq. (3.36) produces only small deviations from the exact solution of Eq. (3.35) [PWLW94, PW85, PRPL96].

The other two terms in Eq. (3.33) control the interaction strength as a function of the relative orientation $\mathbf{A}_1^t \mathbf{A}_2$ and position $\vec{r}_{12}$ of interacting ellipsoids:

$$\eta_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12}) = \frac{\det[\mathbf{S}_1]/\sigma_1^2 + \det[\mathbf{S}_2]/\sigma_2^2}{(\det[\mathbf{H}_{12}]/(\sigma_1 + \sigma_2))^{1/2}} \tag{3.39}$$

$$\mathbf{H}_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12}) = \frac{1}{\sigma_1}\mathbf{A}_1^T \mathbf{S}_1^2 \mathbf{A}_1 + \frac{1}{\sigma_2}\mathbf{A}_2^T \mathbf{S}_2^2 \mathbf{A}_2 \tag{3.40}$$

$$\sigma_i(\mathbf{A}_i, \hat{r}_{12}) \equiv \left(\hat{r}_{12}^T \;\; \mathbf{A}_1^T \mathbf{S}_i^{-2} \mathbf{A}_1 \;\; \hat{r}_{12}\right)^{-1/2} \tag{3.41}$$

and

$$\chi_{12}(\mathbf{A}_1, \mathbf{A}_2, \hat{r}_{12}) = [2\hat{r}_{12}^T \;\; \mathbf{B}_{12}^{-1}(\mathbf{A}_1, \mathbf{A}_2)\; \hat{r}_{12}] \tag{3.42}$$

$$\mathbf{B}_{12}(\mathbf{A}_1, \mathbf{A}_2) = \mathbf{A}_1^T \mathbf{E}_1 \mathbf{A}_1 + \mathbf{A}_2^T \mathbf{E}_2 \mathbf{A}_2 \tag{3.43}$$

with

$$\mathbf{E}_i = \sigma \begin{pmatrix} \frac{a_i}{b_i\, c_i} & 0 & 0 \\ 0 & \frac{b_i}{a_i\, c_i} & 0 \\ 0 & 0 & \frac{c_i}{a_i\, b_i} \end{pmatrix} = \frac{\sigma}{\det[\mathbf{S}_i]}\mathbf{S}_i^2. \tag{3.44}$$

In Fig. 3.5 the distance dependence of the Gay-Berne potential for all possible pole contacts as well as the dependence on Tw, Sh, and Sl around the stacked (on top) conformation is shown.

We neglect electrostatic interactions between neighboring base-pairs since at physiological conditions the stacking interaction dominates [Hun93, CD99].

At this point we have to find appropriate values for the thickness $2a$ and the parameter γ of Eq. (3.34). Both parameters influence the minimum of the Gay-Berne potential. There are essentially two possible procedures. One way is to make use of the parameterization result of Everaers and Ejtehadi [EE03], i.e. $\gamma = 2^{1/6} - 30^{-1/6}$, and to choose a value of $a \approx 0.7$ that yields the minimum center-center distance of 3.3Å for perfect stacking. Unfortunately it turns out that the fluctuations of the bending angles strongly depend on the flatness of the ellipsoids. The more flat the ellipsoids are the smaller are the fluctuations of the bending angles so that one ends up with extremely stiff filaments with a persistence length of a few thousand base-pairs. This can be seen clearly for the extreme case of two perfectly stacked plates: each bending move leads then to an immediate overlap of the plates. That is why we choose the second possibility. We keep γ as a free parameter that is used in the end to shift the potential minimum to the desired value and fix the width of the ellipsoids to be approximately half the known rise value $a = 1.55$Å. This requires $\gamma = 1.07$ close to the standard choice $\gamma = 1$ employed in the literature [BFZ98].

The sugar phosphate backbone is known to be nearly inextensible. The distance between adjacent sugars varies from 5.5Å to 6.5Å [CD99]. This is taken into account by two stiff harmonic

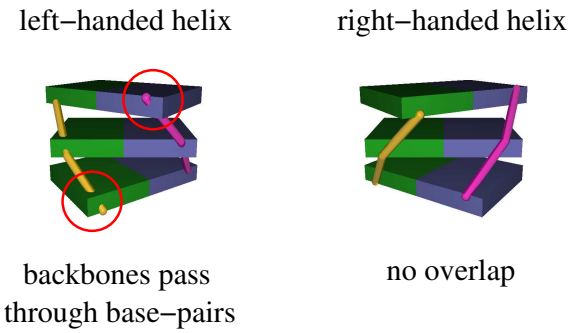


Figure 3.6: Consider two plates that are connected by two rigid bonds. The right bond is anchored at the upper right corner whereas the left bond is anchored at the lower left corner. In case of a right-handed helix there is no overlap of the backbone with the base-pairs whereas in case of a left-handed twist the backbones pass through the plate. Such moves are rejected in the Monte Carlo procedure.

springs with length $l_1 = l_2 = 6.0$Å connecting neighboring ellipsoids (see Fig. 3.4). The anchor points are situated along the centerline in $\vec{n}$-direction (compare Fig. 3.2 and Fig. 3.4) with a distance of ± 8Å from the center of mass. The backbone is thus represented by an elastic spring with non-zero spring length $l_0 = 6$Å

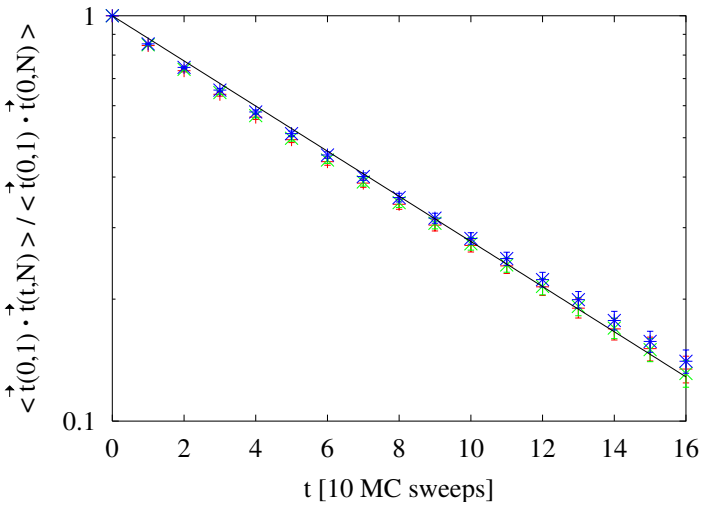
$$\mathcal{H}_{el} = \frac{k}{2}\left[(|\mathbf{r}_{1,i+1} - \mathbf{r}_{1,i}| - l_0)^2 + (|\mathbf{r}_{2,i+1} - \mathbf{r}_{2,i}| - l_0)^2\right]. \tag{3.45}$$

The competition between the GB potential that forces the ellipsoids to maximize the contact area and the harmonic springs with non-zero spring length that does not like to be compressed leads to a twist in either direction of the order of $\pm\pi/5$. The right-handedness of the DNA helix is due to excluded volume interactions between the bases and the backbone which we do not represent explicitly (Fig. 3.6). Rather we break the symmetry by rejecting moves which lead to local twist smaller than $-\pi/18$.

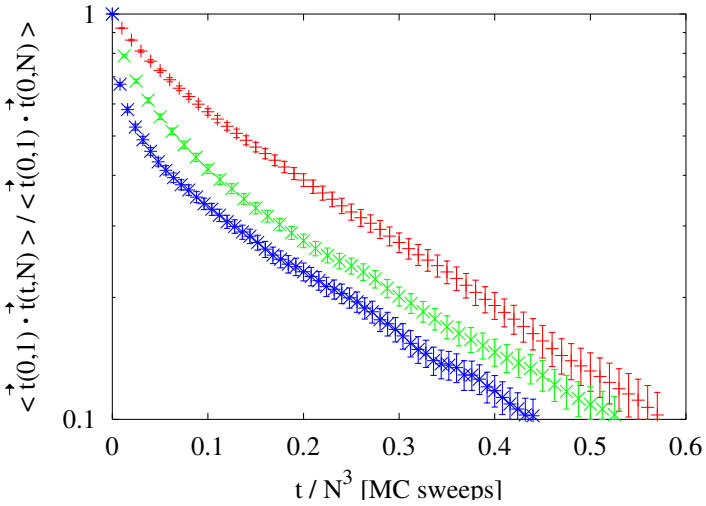
Thus we are left with three free parameters in our model, the GB energy depth $\epsilon = \min(U)$ which controls the stacking interaction, the spring constant k which controls the torsional rigidity, and the depth b of the ellipsoids which influences mainly the fluctuations of the bending angles. All other parameters such as the width and the height of the ellipsoids, or the range of interaction $\sigma = 3.3$Å which determines the width of the GB potential are fixed in order to reproduce the experimental values for B-DNA.

3.3.2 MC simulation

In our model all interactions are local and it can therefore conveniently be studied using a MC scheme. In addition to trial moves consisting of local displacements and rotations of one ellipsoid by a small amplitude, it is possible to employ global moves which modify the position and the orientation of large parts of the chain. The moves are analogous of (i) the well-known pivot move [LB00], and (ii) a crankshaft move where two randomly chosen points along the chain define the axis of rotation around which the inner part of the chain is rotated. The moves are accepted or rejected according to the Metropolis scheme [MRR$^+$53].



(a) Time correlation function for local moves.



(b) Time correlation function for global moves.

Figure 3.7: Time correlation functions of the scalar product of the tangent vectors of the first and the last monomer $\tau = \vec{t}(0,1) \cdot \vec{t}(t,N)$ with $N = 10$ (red), $N = 20$ (green), $N = 50$ (blue) and $2b = 11$Å, $\epsilon = 20k_BT$, $k = 64k_BT/$Å^2 for (a) global and (b) local moves. It is observed that τ_{global} is independent of the chain length N whereas τ_{local} scales as N^3. The 'time' is measured in units of sweeps where one MC sweep corresponds to N trials. The CPU-time for one sweep scales as N^2 in case of global moves and as N in case of local moves. Thus the simulation time t scales as $t_{local} \propto N^4$ and $t_{global} \propto N^2$.

Fig. 3.7 shows that these global moves significantly improve the efficiency of the simulation. We measured the correlation time τ of the scalar product of the tangent vectors of the first and the last monomer of 200 independent simulation runs with $N = 10, 20, 50$ monomers using (i) only local moves and (ii) local and global moves (ratio 1:1). The correlation time of the global moves is independent of the chain length with $\tau_{global} \approx 78$ sweeps whereas τ_{local} scales as N^3.

Each simulation run comprises 10^6 MC sweeps where one MC sweep corresponds to $2N$ trials (one rotational and one translational move per base-pair) with N denoting the number of monomers. The amplitude is chosen such that the acceptance rate equals approximately to 50%. Every 1000 sweeps we store a snapshot of the DNA conformation. We measured the 'time' correlation functions of the end-to-end distance, the rise of one base-pair inside the chain and all three orientational angles of the first and the last monomer and of two neighboring monomers inside the chain in order to extract the longest relaxation time τ_{max}. We observe $\tau_{max} < 1000$ for all simulation runs.

An estimate for the CPU time required for one sweep for chains of length $N = 100$ on a AMD Athlon MP 2000+ processor results in $0.026s$ which is equivalent to $1.33 \times 10^{-4}s$ per move.

3.3.3 Energy minimization

We complemented the simulation study by zero temperature considerations that help to discuss the geometric structure that is obtained by the introduced interactions and to rationalize the MC simulation data. Furthermore zero temperature considerations can be used to obtain an estimate of the critical force f_{crit} that must be applied to enable the structural transition from B-DNA to the overstretched S-DNA configuration as a function of the model parameters $\{\epsilon, k, b\}$.

3.4 Results

In the following we will try to motivate an appropriate parameter set $\{\epsilon,\ k,\ b\}$ that can be used for further investigations within the framework of the presented model. Therefor we explore the parameter dependence of experimental observables such as the bending persistence length of B-DNA $l_p \approx 150$bp, the torsional persistence length $l_t \approx 260$bp [SBC99], the mean values and correlations of all six base-pair parameters and the critical pulling force $f_{crit} \approx 65$pN [CLH$^+$96, LL99, LLA$^+$02, BSLS00] that must be applied to enable the structural transition from B-DNA to the overstretched S-DNA configuration. In fact, static and dynamic contributions to the bending persistence length l_p of DNA are still under discussion. It is known that l_p depends on both the intrinsic curvature of the double helix due to spontaneous bending of particular base-pair sequences and the thermal fluctuations of the bending angles. Bensimon *et al.* introduced disorder into the WLC model by an additional set of preferred random orientation between successive segments [BDM98]. They derived a relationship of the pure persistence length l_{pure}, i.e. without disorder, the effective persistence length l_{eff}, and the probability distribution function of a set of random orientation $P(\{\Psi_i\})$ and compared their result to MC simulations. Assuming a Gaussian distribution function for $P(\{\Psi_i\})$ one finds a simple renormalization formula in the limit of large and small disorder

$$\frac{l_{eff}}{l_{pure}} = \begin{cases} 1 - \frac{1}{2}\sqrt{\frac{l_{pure}}{l_{disorder}}} & , \quad \frac{l_{pure}}{l_{disorder}} \ll 1 \\ 2\,\frac{l_{disorder}}{l_{pure}} & , \quad \frac{l_{pure}}{l_{disorder}} \gg 1 \end{cases} . \tag{3.46}$$

Since we are dealing with intrinsically straight filaments with $1/l_{disorder} = 0$, we measure l_{pure}. Recent estimates of $l_{disorder}$ range between 430 [BFK$^+$95] and 4800 [VV02] base-pairs using cryo-

T	⟨Ri⟩	⟨Sh⟩	⟨Sl⟩	⟨Tw⟩	⟨Ti⟩	⟨Ro⟩	$\langle c \rangle$	l_p
0	3.26	0.0	0.0	0.64	0.0	0.0	3.26	∞
1	3.37	0.01	-0.01	0.62	0.0	0.0	3.47	172.8
2	3.76	-0.01	-0.03	0.47	0.0	0.0	4.41	25.3
3	4.10	-0.01	0.01	0.34	0.0	-0.01	5.07	14.4
5	4.30	0.03	-0.02	0.27	0.0	0.01	5.39	13.6

Table 3.1: Dependence of mean values of all six step parameters and of the mean center-center distance $\langle c \rangle$ on the temperature for $2b = 11$Å, $\epsilon = 20 k_B T$, $k = 64 k_B T/$Å^2. ⟨Ri⟩, ⟨Sh⟩, ⟨Sl⟩ and $\langle c \rangle$ are measured in [$\mathring{A}$], l_p in base-pairs.

electron microscopy and cyclization experiments respectively implicating values between 105 and 140 base-pairs for l_{pure} if it is assumed that the angles Ψ_i are small.

3.4.1 Equilibrium structure

As a first step we study the equilibrium structure of our chains as a function of the model parameters. To investigate the ground state conformation we rationalize the MC simulation results with the help of the geometrical considerations and minimum energy calculations. In the end we will choose parameters for which our model reproduces the experimental values of B-DNA [CD99]:

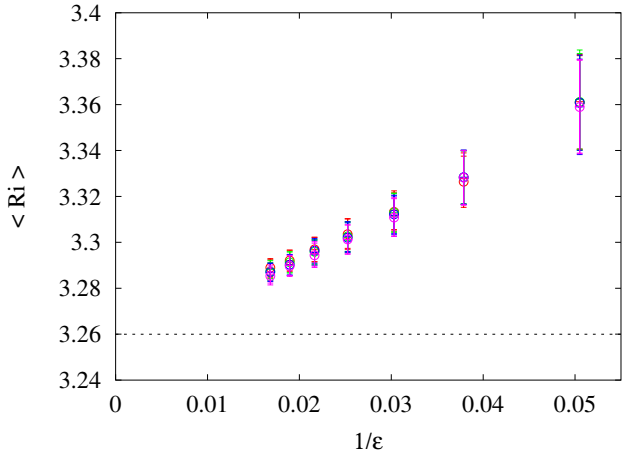
$$\begin{aligned}
\langle \text{Ri} \rangle &= 3.3 - 3.4\text{Å} \\
\langle \text{Sl} \rangle &= 0\text{Å} \\
\langle \text{Sh} \rangle &= 0\text{Å} \\
\langle \text{Tw} \rangle &= 2\pi/10.5 - 2\pi/10 \\
\langle \text{Ti} \rangle &= 0 \\
\langle \text{Ro} \rangle &= 0.
\end{aligned}$$

We use the following reduced units in our calculations. The energy is measured in units of $k_B T$, lengths in units of Å, forces in units of $k_B T \text{Å}^{-1} \approx 40$pN.

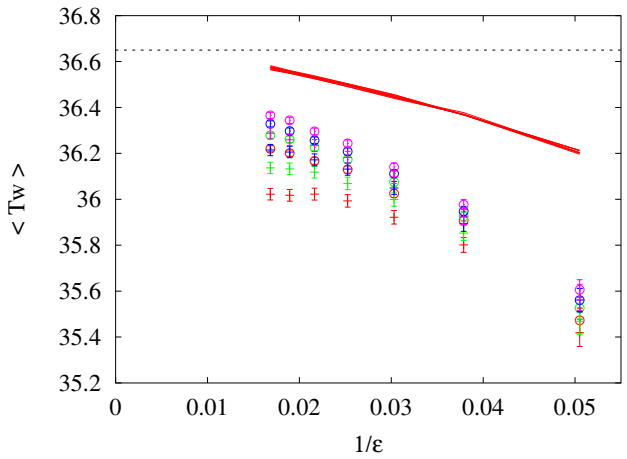
We start by minimizing the energy for the various conformations shown in Fig. 3.3 to verify that our model Hamiltonian indeed prefers the B-Form. Since we have only local (nearest neighbor) interactions we can restrict the calculations to two base-pairs. There are three local minima which have to be considered: (i) a stacked, twisted conformation with Ri $= 3.3$, Sl, Sh, Ti, Ro $= 0$, Tw $= \pi/10$, (ii) a skewed ladder with Ri $= 3.3$, Sl $= 5.0$, Sh, Tw, Ti, Ro $= 0$, and (iii) an unwound helix with Ri $= 6.0$, Sl, Sh, Ti, Ro $= 0$, Tw $= 0$. Without an external pulling force the global minimum is found to be the stacked twisted conformation.

We investigated the dependence of Ri and Tw on the GB energy depth ϵ that controls the stacking energy for different spring constants k. Ri depends neither on ϵ nor on k nor on b. It shows a constant value of Ri ≈ 3.3Å for all parameter sets $\{\epsilon, k, b\}$. The resulting Tw of the minimum energy calculation coincides with the geometrically determined value under the assumption of fixed Ri up to a critical ϵ. Up to that value the springs behave effectively as rigid rods. The critical ϵ is determined by the torque $\tau(k, \epsilon)$ that has to be applied to open the twisted structure for a given value of Ri.

Using MC simulations we can study the effects arising from thermal fluctuations. Plotting ⟨Ri⟩, and ⟨Tw⟩ as a function of the GB energy depth ϵ one recognizes that in general ⟨Ri⟩ is larger than



(a) Rise as a function of ϵ. $\langle \mathrm{Ri} \rangle$ depends only on ϵ. In the limit of $\epsilon \to \infty$ the minimum energy value is reached.



(b) Twist as a function of ϵ. In addition to the MC data and the minimum energy calculation we calculated the twist with Eq. (3.14) using the measured mean rise values of (a). It can be observed that $\langle \mathrm{Tw} \rangle$ changes with all three model parameters. Increasing y and k decreases especially the fluctuations of Tw and Sh so that $\langle \mathrm{Tw} \rangle$ increases as a result of the mechanical coupling of the shift and twist motion. In the limit of $\epsilon, k \to \infty$ the minimum energy value is reached.

Figure 3.8: Rise [Å] and twist [^{o}] as a function of ϵ [k_BT] for $2b = 8,\ 9,\ 10,\ 11$Å (red, green, blue, purple). For every b there are two data sets for $k = 32$ (plus), 64 (circles) [$k_BT/\text{Å}^2$]. The dotted line corresponds to the minimum energy value.

Ri$(T = 0)$. It converges only for large values of ϵ to the minimum energy values. This can be understood as follows. Without fluctuations the two base-pairs are perfectly stacked taking the minimum energy configuration Ri $= 3.3$Å, Sl, Sh, Ti, Ro $= 0$, and Tw $= \pi/10$. As the temperature is increased the fluctuations can only occur to larger Ri values due to the repulsion of neighboring base-pairs. A decrease of Ri would cause the base-pairs to intersect. Increasing the stacking energy reduces the fluctuations in the direction of the tangent vector and leads to smaller $\langle$Ri$\rangle$ value. In the limit $\epsilon \to \infty$ it should reach the minimum energy value which is observed from the simulation data. In turn the increase of the mean value of rise results in a smaller twist angle $\langle$Tw$\rangle$. We can calculate with the help of Eq. (3.14) the expected twist using the measured mean values of $\langle$Ri$\rangle$. Fig. 3.8 shows that there is no agreement. The deviations are due to fluctuations in Sl and Sh which cause the base-pairs to untwist. This is the mechanical coupling of Sl, Sh, and Tw due to the backbones already mentioned in section 3.2.2. It is observed that a stiffer spring k and a larger depth of the ellipsoids b result in larger mean twist values. Increasing the spring constant k means decreasing the fluctuations of the twist and, due to the mechanical coupling, of the shift motion around the mean values which explains the larger mean twist values. An increase of the ellipsoidal depth b in turn decreases the fluctuations of the bending angles. The coupling of the tilt fluctuations with the shift fluctuations leads to larger values for $\langle$Tw$\rangle$. The corresponding limit where $\langle$Tw$\rangle \to$ Tw(T $= 0$) is given by $k, \epsilon \to \infty$.

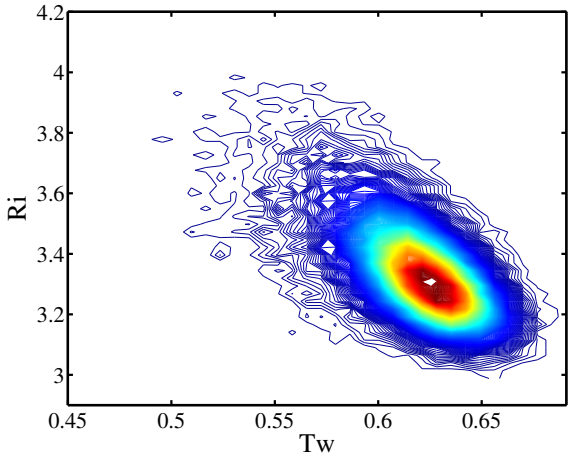
The measurement of the mean values of all six base-pair step parameters for different temperatures is shown in Table 3.4.1. One can see that with increasing temperature the twist angles decrease while the mean value of rise increase. The increase of the center-center distance is not only due to fluctuations in Ri but also due to fluctuations in Sl and Sh. That is why there are strong deviations of $\langle c \rangle$ from $\langle$Ri$\rangle$ even though the mean values of Sl and Sh vanish. Note that the mean backbone length $\langle l \rangle$ always amounts to about 6Å.

The calculation of the probability distribution functions of all six base-pair parameters shows that especially the rise and twist motion do not follow a Gaussian behavior. The deviation of the distribution functions from the Gaussian shape depends mainly on the stacking energy determined by ϵ. For smaller values of ϵ one observes larger deviations than for large ϵ values.

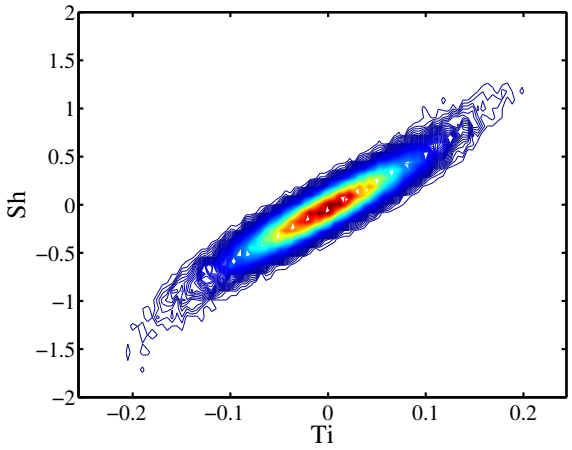
It is worthwhile to mention that there are mainly two correlations between the base-pair parameters. The first is a microscopic twist-stretch coupling determined by a correlation of Ri and Tw, i.e. an untwisting of the helix implicates larger rise values. A twist-stretch coupling was introduced in earlier rod models [KLNO97, Mar97, Nel98] motivated by experiments with torsionally constrained DNA [SAB$^+$96] which allow for the determination of this constant. Here it is the result of the preferred stacking of neighboring base-pairs and the rigid backbones. The second correlation is due to constrained tilt motion. If we return to our geometrical ladder model we recognize immediately that a tilt motion alone will always violate the constraint of fixed backbone length l. Even though we allow for backbone fluctuations in the simulation the bonds are very rigid which makes tilting energetically unfavorable. To circumvent this constraint tilting always involves a directed shift

	Ri	Sh	Sl	Tw	Ti	Ro
$\langle x \rangle$	3.359	0.0	0.0	0.621	0.0	0.0
$\langle x^2 \rangle - \langle x \rangle^2$	0.020	0.115	0.423	0.001	0.003	0.009

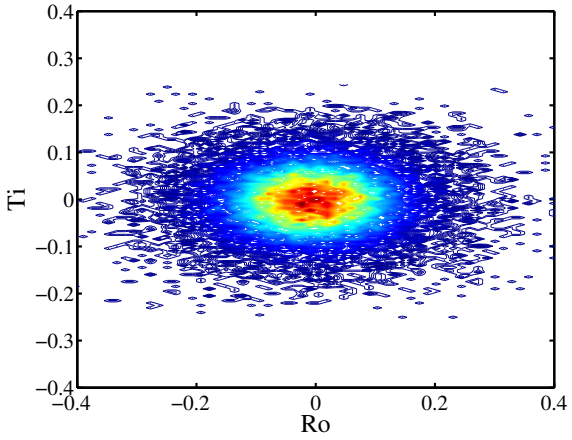
Table 3.2: Mean values and mean squared fluctuations of all step parameters for $\epsilon = 20 k_B T$, $k = 64 k_B T/$Å^2, $2b = 11$Å.



(a) Contour plot of measured clouds for rise-twist.



(b) Contour plot of measured clouds for shift-tilt.



(c) Contour plot of measured clouds for roll-tilt.

Figure 3.9: Contour plots of measured clouds for rise-twist, shift-tilt, and roll-tilt to demonstrate internal couplings and the anisotropy of the bending angles ($2b = 11$Å, $\epsilon = 20k_BT$, $k = 64k_BT/$Å^2).

motion.

Fig. 3.9 shows that we recover the anisotropy of the bending angles Ro and Ti as a result of the spatial dimensions of the ellipsoids. Since the overlap of successive ellipsoids is larger in case of rolling it is more favorable to roll than to tilt.

The correlations can be quantified by calculating the correlation matrix $\mathcal{C}$ of Eq. (3.19). Inverting $\mathcal{C}$ yields the effective coupling constants of the SOP model $\mathcal{K} = \mathcal{C}^{-1}$. Due to the local interactions it suffices to calculate mean and mean squared values of Ri, Sl, Sh, Tw, Ro, and Ti characterizing

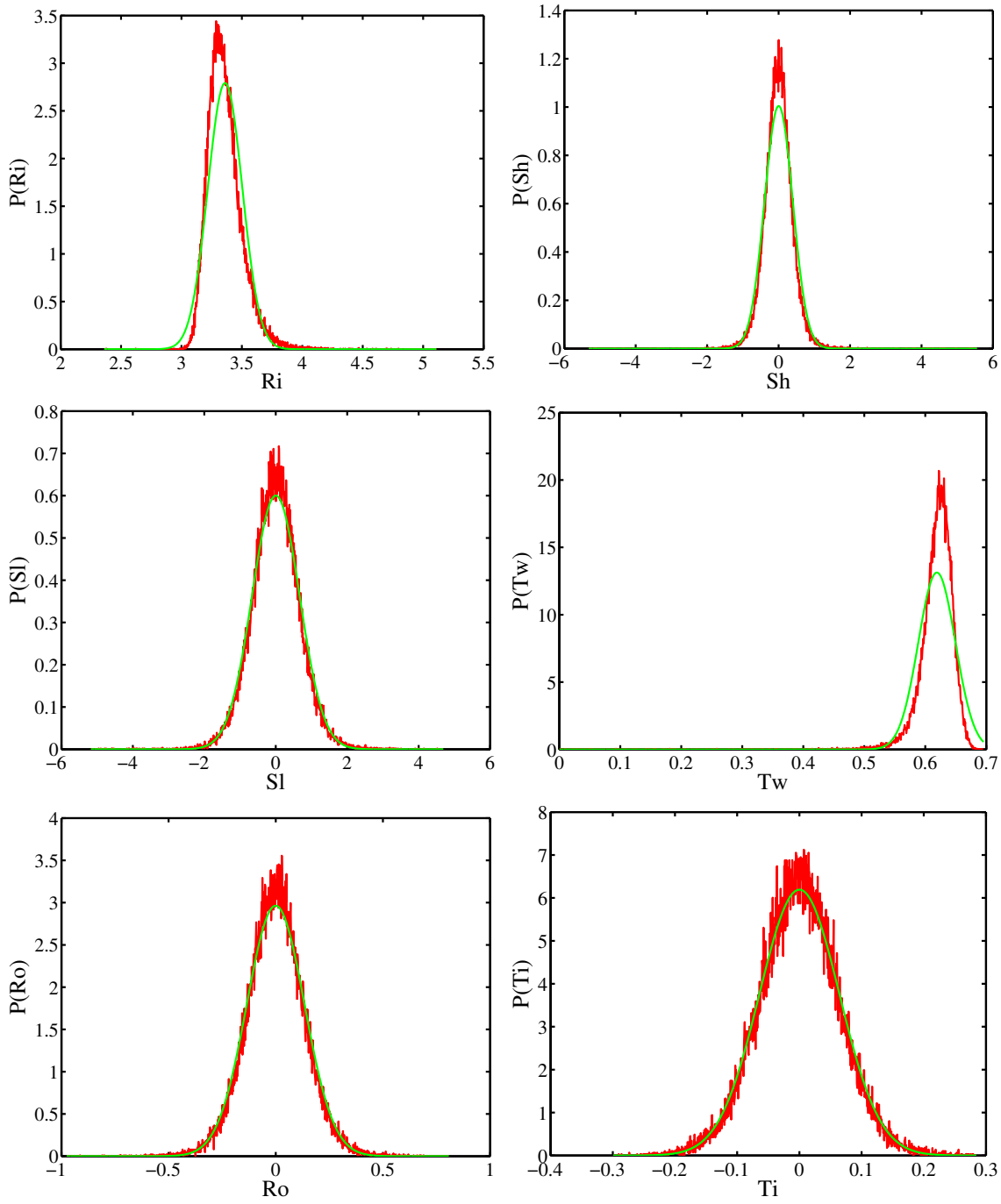


Figure 3.10: Comparison of probability distribution functions of all base-pair parameters for $\epsilon = 20k_BT$, $k = 64k_BT/\text{Å}^2$, $2b = 8\text{Å}$. The Gaussians are plotted with the measured mean and mean squared values of the MC simulation.

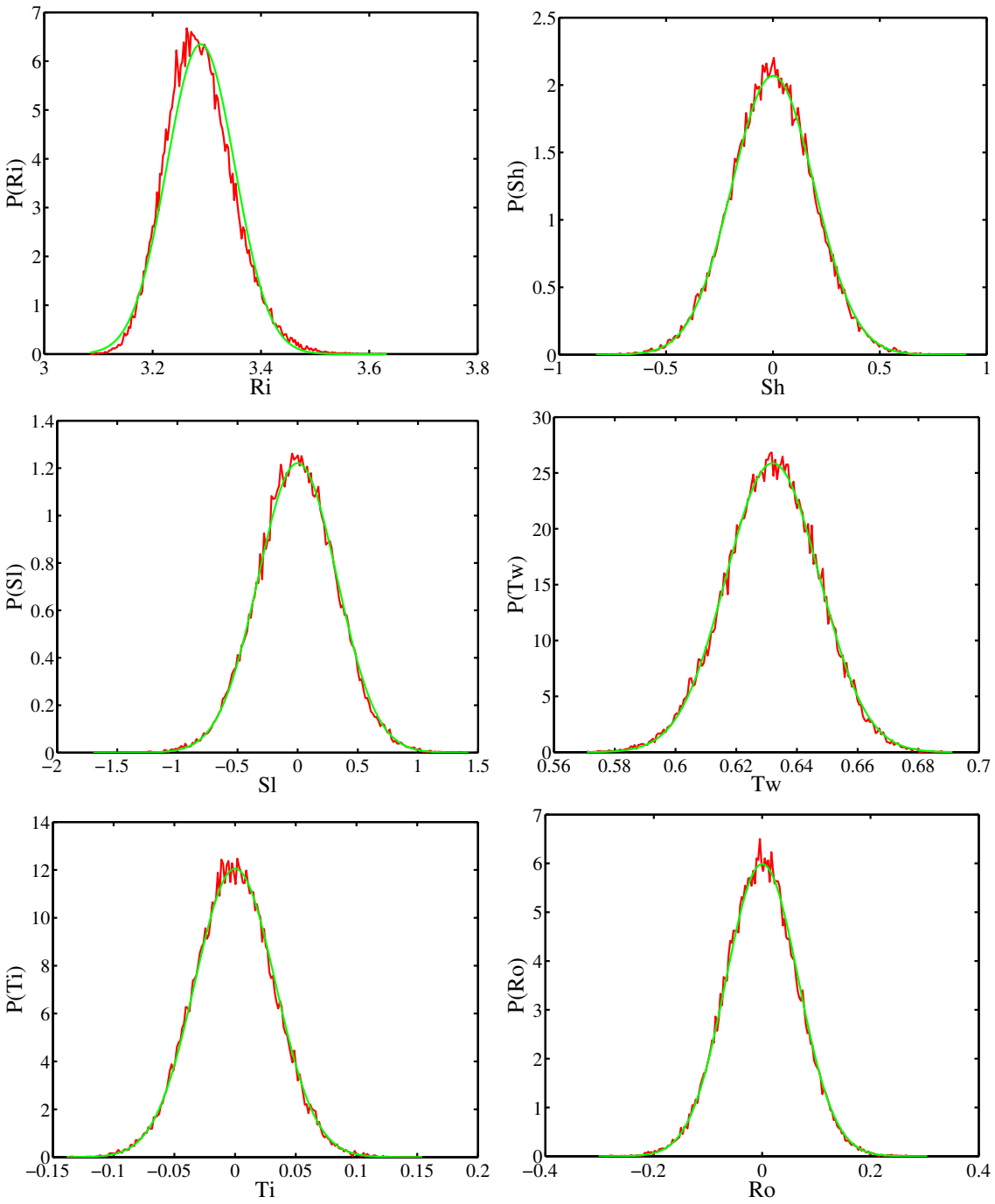
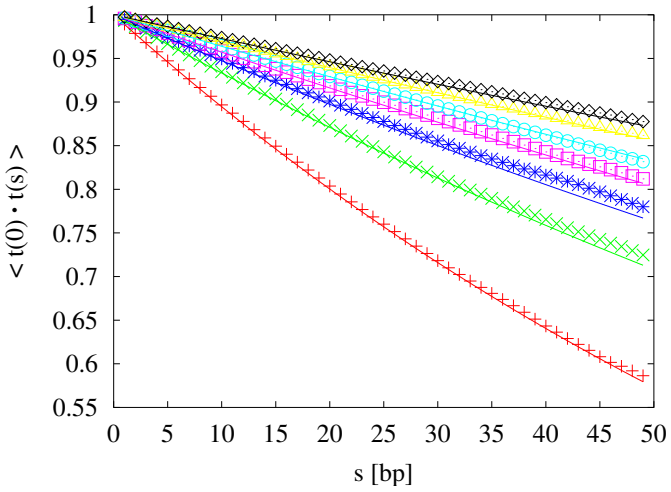
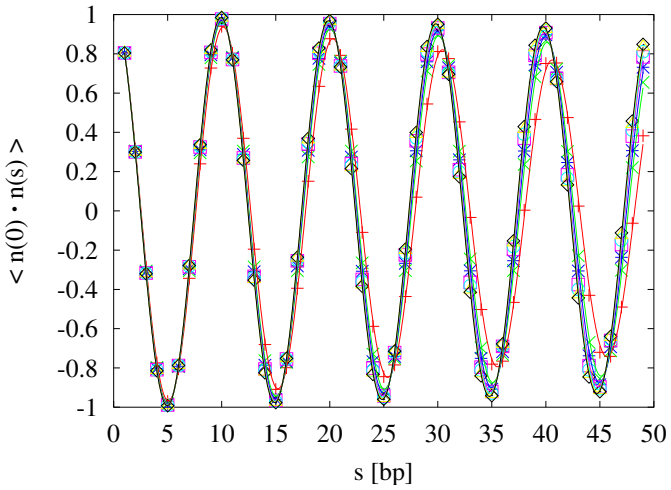


Figure 3.11: Comparison of probability distribution functions of all base-pair parameters for $\epsilon = 60k_BT$, $k = 64k_BT/\text{Å}^2$, $2b = 8\text{Å}$. The Gaussians are plotted with the measured mean and mean squared values of the MC simulation. One can see that for large ϵ values the distribution functions are essentially Gaussian.



(a) Tangent vector correlation function $\langle \mathbf{t}_i \cdot \mathbf{t}_j \rangle$.



(b) Normal vector correlation function $\langle \mathbf{n}_i \cdot \mathbf{n}_j \rangle$.

Figure 3.12: Comparison of analytical expressions Eqs. (3.26) and (3.30) for l_p and l_n (solid lines) with numerically calculated orientational correlation functions (data points) for $2b = 8$Å, $k = 64 k_B T/\text{Å}^2$, and $\epsilon = 20, \ldots, 60$ $[k_B T]$ (from bottom to top).

the 'internal' couplings of the base-pairs steps:

$$\mathcal{C} = (\sigma)_{ij}, \forall i, j \in \{1, \ldots, 6\} \tag{3.47}$$

with $\sigma_{x,y} = \langle xy\rangle - \langle x\rangle\langle y\rangle$. One obtains for $\epsilon = 20k_BT$, $k = 64k_BT/\text{Å}^2$, $2b = 11\text{Å}$

$$\mathcal{K} = \mathcal{C}^{-1} = \begin{pmatrix} 81\text{Å}^{-2} & -5\text{Å}^{-2} & 0 & 267\text{Å}^{-1} & 29\text{Å}^{-1} & 0 \\ -5\text{Å}^{-2} & 60\text{Å}^{-2} & 0 & -48\text{Å}^{-1} & -340\text{Å}^{-1} & 0 \\ 0 & 0 & 2\text{Å}^{-2} & -1\text{Å}^{-1} & 0 & 0 \\ 267\text{Å}^{-1} & -48\text{Å}^{-1} & -1\text{Å}^{-1} & 2211 & 261 & -2 \\ 29\text{Å}^{-1} & -340\text{Å}^{-1} & 0 & 261 & 2244 & -1 \\ 0 & 0 & 0 & -2 & -1 & 117 \end{pmatrix}. \tag{3.48}$$

Thus a significant twist-stretch, shift-tilt and twist-tilt coupling is recovered.

3.4.2 Bending and torsional rigidity and stretching modulus

The correlation matrix of Eq. (3.47) can also be used to check Eqs. (3.26) and (3.30). Therefore we measured the orientational correlation functions $\langle \mathbf{t}_i \cdot \mathbf{t}_j\rangle$, $\langle \mathbf{n}_i \cdot \mathbf{n}_j\rangle$, $\langle \mathbf{b}_i \cdot \mathbf{b}_j\rangle$ and compared the results to the analytical expressions as it is illustrated in Fig. 3.12. The agreement is excellent.

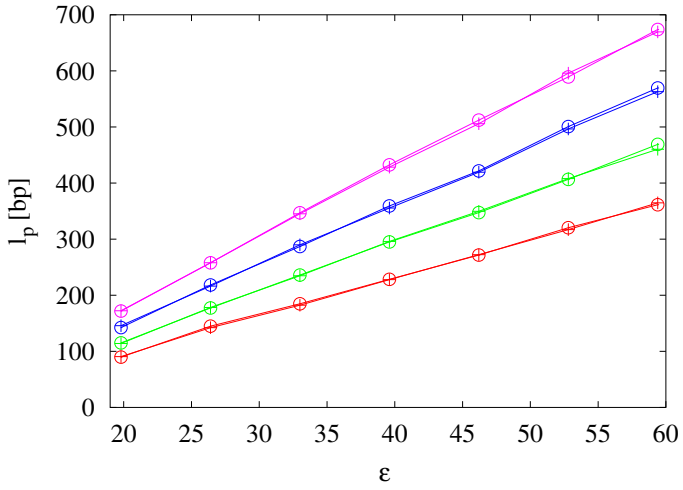
The simulation data show that the bending persistence length does not depend on the spring constant k. But it strongly depends on ϵ being responsible for the energy that must be paid to tilt or roll two respective base-pairs. Since a change of twist for constant Ri is proportional to a change in bond length the bond energy contributes to the twist persistence length explaining the dependence of l_{Tw} on k (compare Fig. 3.13).

We also measured the mean-square end-to-end distance $\langle R_E^2\rangle$ and find that $\langle R_E^2\rangle$ deviates from the usual WLC chain result due to the compressibility of the chain. So as to investigate the origin of the compressibility we calculate $\langle R_E^2\rangle$ for the following geometry. We consider two base-pairs without spontaneous bending angles such that the end-to-end vector $\vec{R}_E$ can be expressed as

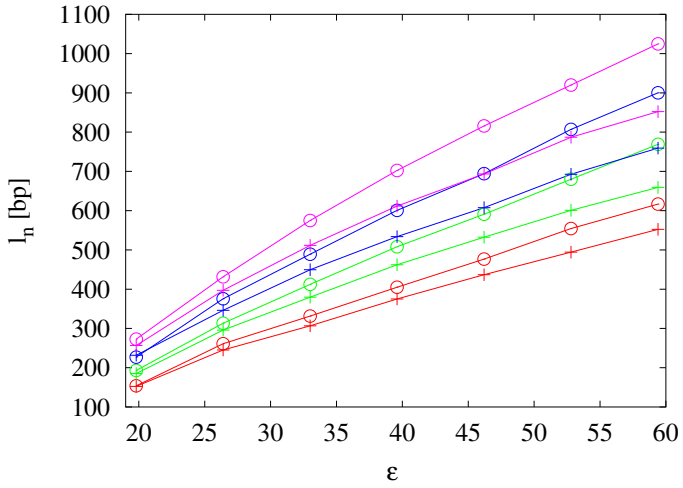
$$\vec{R}_E = \sum_i \vec{c}_i = \sum_i (\text{Ri}\,\mathbf{t}_i + \text{Sh}\,\mathbf{b}_i + \text{Sl}\,\mathbf{n}_i). \tag{3.49}$$

The coordinate system $\{\mathbf{t}_i, \mathbf{b}_i, \mathbf{n}_i\}$ is illustrated in Fig. 3.2. $\vec{c}_i$ denotes the center-center distance of two neighboring base-pairs. Since successive base-pair step parameters are independent of each other, and the translational step parameters Ri and Sh and Sl are uncorrelated the mean-square end-to-end distance $\langle R_E^2\rangle$ is given by

$$\begin{aligned} \langle R_E^2\rangle &= \sum_i \sum_j \langle \vec{c}_i \cdot \vec{c}_j\rangle \\ &= \sum_i \sum_j (\langle \text{Ri}_i\text{Ri}_j\rangle \langle \mathbf{t}_i \cdot \mathbf{t}_j\rangle + \langle \text{Sh}_i\text{Sh}_j\rangle \langle \mathbf{b}_i \cdot \mathbf{t}_j\rangle + \langle \text{Sl}_i\text{Sl}_j\rangle \langle \mathbf{n}_i \cdot \mathbf{t}_j\rangle) \\ &\Downarrow \langle \text{Ri}_i\text{Ri}_j\rangle = \langle \text{Ri}_i\rangle\langle \text{Ri}_j\rangle = \langle \text{Ri}\rangle^2, \text{ etc., } \forall i \neq j,\ \langle \text{Sl}\rangle = \langle \text{Sh}\rangle = 0 \\ &= \sum_i (\langle \text{Ri}^2\rangle + \langle \text{Sh}^2\rangle + \langle \text{Sl}^2\rangle) + \sum_i \sum_{j\neq i} \langle \text{Ri}\rangle^2 \langle \mathbf{t}_i \cdot \mathbf{t}_j\rangle \\ &= \sum_i (\langle c_i^2\rangle - \langle \text{Ri}\rangle^2) + \sum_i \sum_j \langle \text{Ri}\rangle^2 \langle \mathbf{t}_i \cdot \mathbf{t}_j\rangle \\ &= \frac{N\langle \text{Ri}\rangle}{\gamma} + 2N\langle \text{Ri}\rangle l_p - 2l_p^2 \left(1 - \exp\left(-\frac{N\langle \text{Ri}\rangle}{l_p}\right)\right). \end{aligned} \tag{3.50}$$



(a) Bending persistence length l_p.



(b) Torsional persistence length l_n.

Figure 3.13: Dependency of bending and torsional persistence length on the spring constant k, the width of the ellipsoids b and the energy depth ϵ. We measured the persistence lengths for varying width sizes $2b = 8, 9, 10, 11$Å (red, green, blue, purple) and for two different spring constants $k = 32$ (plus), 64 (circles) $[k_BT/\text{Å}^2]$. The bending persistence length depends solely on b and ϵ. It gets larger for larger ϵ and b values. But it does not depend on k (the curves for different k values corresponding to the same width b lie one upon the other). The torsional persistence length in turn depends on k, since a change of twist for constant Ri is proportional to a change in bond length.

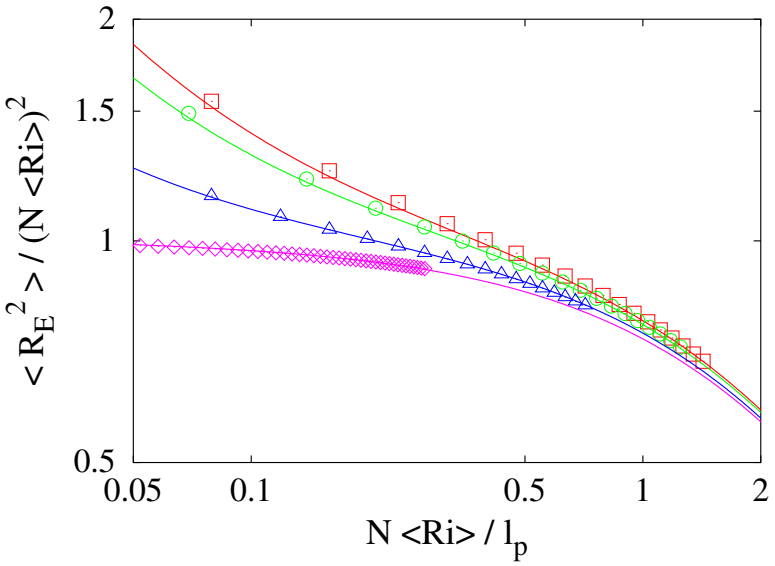


Figure 3.14: Comparison of the simulation data with $\epsilon = 20 k_B T$, $k = 64 k_B T/\text{Å}^2$, $2b = 11\text{Å}$, and $T = 1, 2, 3, 5$ (from bottom to top) to Eqs. (3.26), (3.50) and (3.51) (solid lines). Using the measured bending persistence lengths and the stretching moduli we find a good agreement with the predicted behavior. For $T = 1$ we obtain $\gamma = 6.02\text{Å}^{-1}$.

N denotes the number of base-pairs. Note that $\langle \text{Sl} \rangle$ and $\langle \text{Sh} \rangle$ vanish. Using $\langle c_i^2 \rangle = \langle \text{Ri}^2 \rangle + \langle \text{Sh}^2 \rangle + \langle \text{Sl}^2 \rangle$ the stretching modulus γ is simply given by

$$\gamma = \frac{\langle \text{Ri} \rangle}{(\langle \text{Ri}^2 \rangle - \langle \text{Ri} \rangle^2) + \langle \text{Sh}^2 \rangle + \langle \text{Sl}^2 \rangle}. \tag{3.51}$$

We compared the data for different temperatures T to Eq. (3.50) using the measured bending persistence lengths l_p and stretching moduli γ (see Fig. 3.14). The agreement is excellent. This indicates that *transverse* slide and shift fluctuations contribute to the *longitudinal* stretching modulus of the chain.

3.4.3 Stretching

Extension experiments on double-stranded B-DNA have shown that the overstretching transition occurs when the molecule is subjected to stretching forces of 65pN or more [BSLS00]. The DNA molecule thereby increases in length by a factor of 1.8 times the normal contour length. This overstretched DNA conformation is called S-DNA. The structure of S-DNA is still under discussion. First evidence of possible S-DNA conformations were provided by Lavery *et al.* [CLH$^+$96, LL99, LLA$^+$02] using atomistic computer simulations.

In principle one can imagine two possible scenarios how the transition from B-DNA to S-DNA occurs within our model. Either the chain untwists and unstacks resulting in an untwisted ladder with approximately 1.8 times the equilibrium length, or the chain untwists and the base-pairs slide against each other resulting in a skewed ladder with the same S-DNA length. The second scenario should be

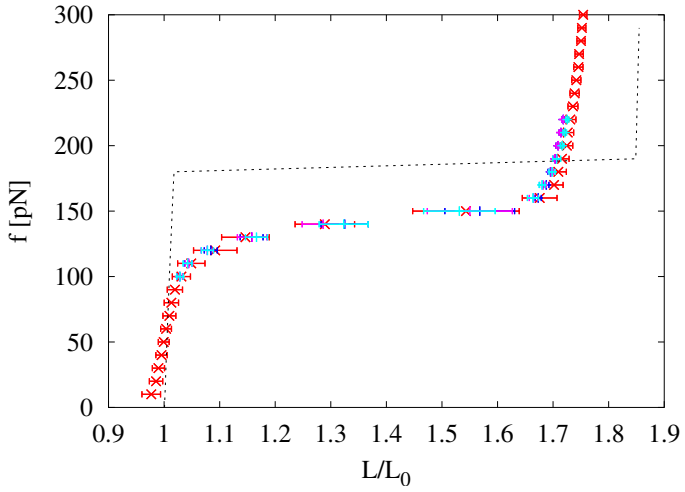
energetically favorable since it provides a possibility to partially conserve the stacking of successive base-pairs. In fact molecular modeling of the DNA stretching process [CLH$^+$96, LL99, LLA$^+$02] yielded both a conformation with strong inclination of base-pairs and an unwound ribbon depending on which strand one pulls.

We expect that the critical force f_{crit} where the structural transition from B-DNA to overstretched S-DNA occurs depends only on the GB energy depth ϵ controlling the stacking energy. So as a first step to find an appropriate value of ϵ as input parameter for the MC simulation we minimize the Hamiltonian with an additional stretching energy $E_{pull} = fc_{i,i+1}$, where the stretching force acts along the center-of-mass axis, with respect to Ri, Sl and Tw for a given pulling force f. Fig. 3.15 shows the resulting stress-strain curve. First the pulling force acts solely against the stacking energy up to the critical force where a jump from $L(f_{crit-})/L_0 \approx 1.05$ to $L(f_{crit+})/L_0 = \sqrt{\text{Ri}^2 + \text{Sl}^2}/\text{Ri} \approx 1.8$ occurs, followed by another slow increase of the length caused by overstretching the bonds. $L_0 = L(F = 0) = \text{Ri}$ denotes the stress-free center-of-mass distance. As already mentioned three local minima are obtained: (i) a stacked, twisted conformation, (ii) a skewed ladder, and (iii) an unwound helix. The strength of the applied stretching force determines which of the local minima becomes the global one. The global minimum for small stretching forces is determined to be the stacked, twisted conformation and the global minima for stretching forces larger than f_{crit} is found to be the skewed ladder. Therefore the broadness of the force plateau depends solely on the ratio of l/Ri determined by the geometry of the base-pairs S and the bond length $l = 6.0$Å. A linear relationship between the critical force and the stacking energy ϵ is obtained so that it is possible to extrapolate to smaller ϵ values to extract the ϵ value that reproduces the experimental value of $f_{crit} \approx 65$pN. This suggests a value of $\epsilon \approx 7$.

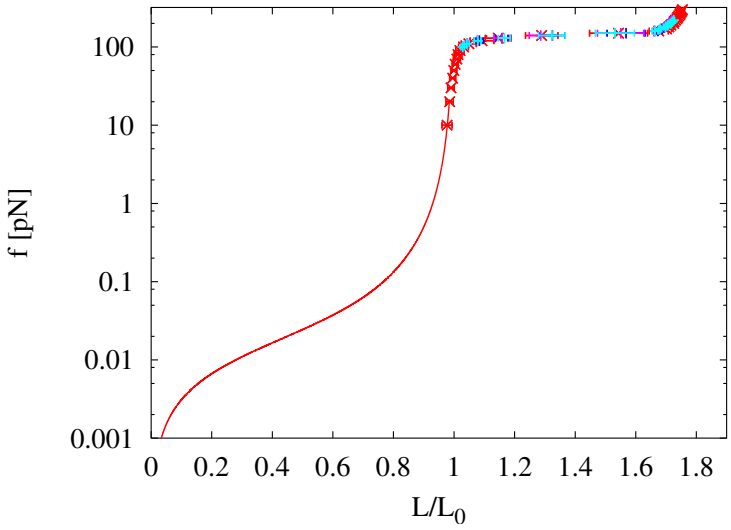
The simulation results of the previous sections show several problems when this value of ϵ is chosen. First of all the correct persistence lengths cannot be reproduced, the chain is far to flexible. Secondly the undistorted ground state is not a B-DNA anymore. The thermal fluctuations suffice to unstack and untwist the chain locally. That is why one has to choose larger ϵ values even though the critical force is going to be overestimated.

Therefore we choose the following way to fix the parameter set $\{b, \epsilon, k\}$. First of all we choose a value for the stacking energy that reproduces correctly the persistence length. Afterwards the torsional persistence length is fixed to the experimentally known values by choosing an appropriate spring constant k. The depth of the base-pairs has also an influence on the persistence lengths of the chain. If the depth b is decreased larger fluctuations for all three rotational parameters are gained such that the persistence lengths get smaller. Furthermore the geometric structure and the behavior under pulling is very sensitive to b. Too small values provoke non-B-DNA conformations or unphysical S-DNA conformations. We choose for b a value of 11Å for those reasons. For $\epsilon = 20$ and $k = 64$ a bending stiffness of $l_p = 170$bp and a torsional stiffness of $l_n = 270$bp are obtained close to the experimental values. We use this parameter set to simulate the corresponding stress-strain relation.

The simulated stress-strain curves for 50 base-pairs show three different regimes (see Fig. 3.15). (i) For small stretching forces the WLC behavior of the DNA in addition with linear stretching elasticity of the backbones is recovered. This regime is completely determined by the chain length N. Due to the coarse-graining procedure that provides analytic expressions of the persistence lengths depending on the base-pair parameters it is not necessary to simulate a chain of a few thousand base-pairs. The stress-strain relation of the entropic and WLC stretching regime (small relative extensions L/L_0 and small forces) is known analytically [MS95c, Odi95]. (ii) Around the critical force



(a) Stress strain relation calculated by minimum energy calculation (black) and obtained by MC simulation (red: $N = 50$ and blue, cyan, purple: $N = 500$). The deviation in the critical force is due to entropic contributions.



(b) Logarithmic representation with WLC result.

Figure 3.15: Force-extension curve for final parameterization. We simulated $N = 50$ and $N = 500$ monomers, so that the WLC behavior (red solid line in (b)) is recovered for small stretching forces followed by a structural transition. The elastic response of S-DNA is controlled by overstretching the bonds.

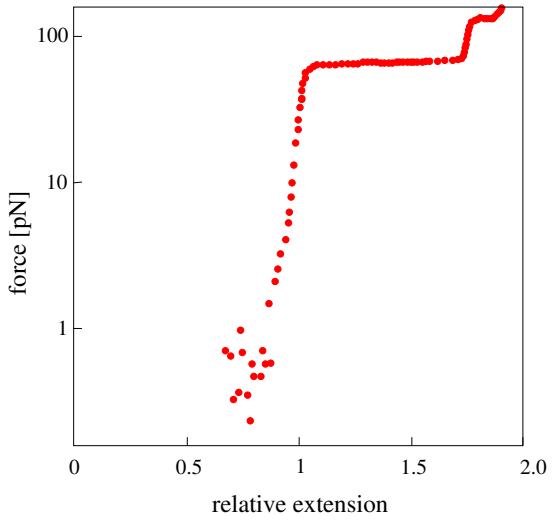
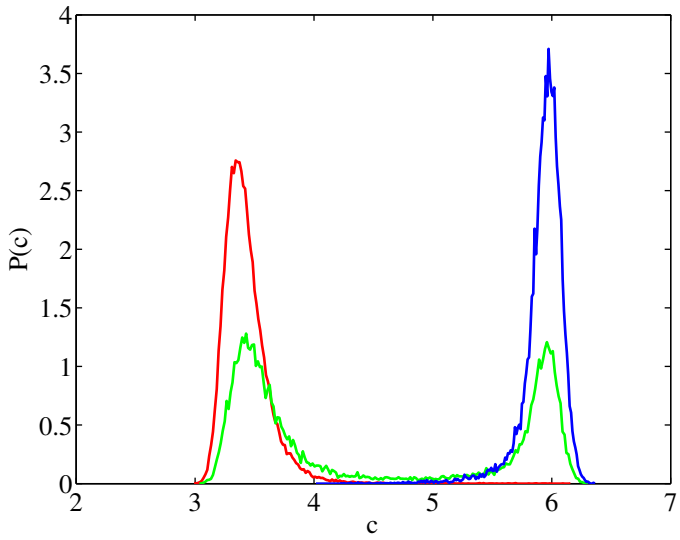


Figure 3.16: Experimentally measured stress-strain curves by [BSLS00].

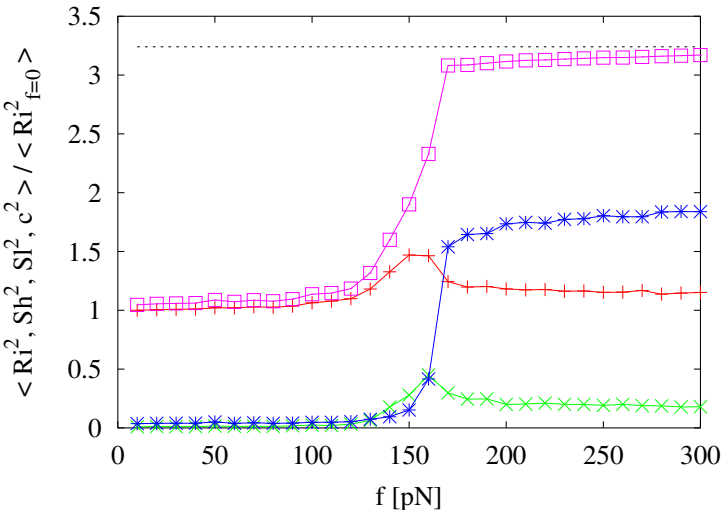
$f_{crit} \approx 140$pN which is mainly determined by the stacking energy of the base-pairs the structural transition from B-DNA to S-DNA occurs. (iii) For larger forces the bonds become overstretched. Our MC simulations suggest a critical force $f_{crit} \approx 140$pN which is slightly smaller than the value $f_{crit} \approx 180$pN calculated by minimizing the energy. This is due to entropic contributions.

In order to further characterize the B-to-S-transition we measured the mean, mean-square values and the probability distribution functions of rise, slide, shift, etc. as a function of the applied forces. The evaluation of the MC data shows that the mean values of shift, roll and tilt are completely independent of the applied stretching force and vanish for all f. Rise increases at the critical force from the undisturbed value of 3.3Å to approximately 4.0Å and decays subsequently to the undisturbed value. Quite interestingly, the mean value of slide jumps from its undisturbed value of 0 to ± 5Å (no direction is favored) and the twist changes at the critical force from $\pi/10$ to 0. The calculation of the distribution function of the center-center distance c of two neighboring base-pairs for $f = 140$pN yields a double-peaked distribution (see Fig. 3.17) indicating that part of the chain is in the B-form and part of the chain in the S-form. The contribution of the three translational degrees of freedom to the center-center distance c is shown in Fig. 3.17. The S-DNA conformation is characterized by Ri $= 3.3$Å, Sl $= \pm 5$Å and Tw $= 0$. In agreement with Refs. [CLH$^+$96, LL99] we obtain a conformation with highly inclined base-pairs still allowing for partial stacking of successive base-pairs.

The evolution of the S-DNA conformation depending on the applied stretching force is illustrated in Figs. 3.20-3.22. We measured the probability distribution functions of all six step variables around the critical force for $f = 130$pN and $f = 140$pN, and beyond f_{crit} for $f = 250$pN. Figs. 3.20 and 3.21 stress the fact that around the critical force part of the chain is in the B-form and part of the chain is in the S-form giving rise for large fluctuations in Ri, Sl and Tw. Note that the twist distribution function shows a cutoff at $\pi/16$ which represent the geometrical constraint in order to enforce a right-handed helical structure as ground state conformation. Moreover it can be observed



(a) PDF of the center-center distance of successive base-pairs for $f = 0$ (red), 140 (green), 200 (blue) pN.



(b) Mean squared values of rise (red), shift (green), slide (blue), and center-of-mass distance (purple) for neighboring base-pairs as a function of the stretching force f. The dashed line corresponds to the S-DNA center-of-mass distance.

Figure 3.17: Probability distribution function of the center-center distance of successive base-pairs and contributions of the translational degrees of freedom to $\langle c^2 \rangle$. Note that $\langle \mathrm{Tw} \rangle$ of the resulting S-DNA conformation vanishes as predicted by Eq. (3.16).

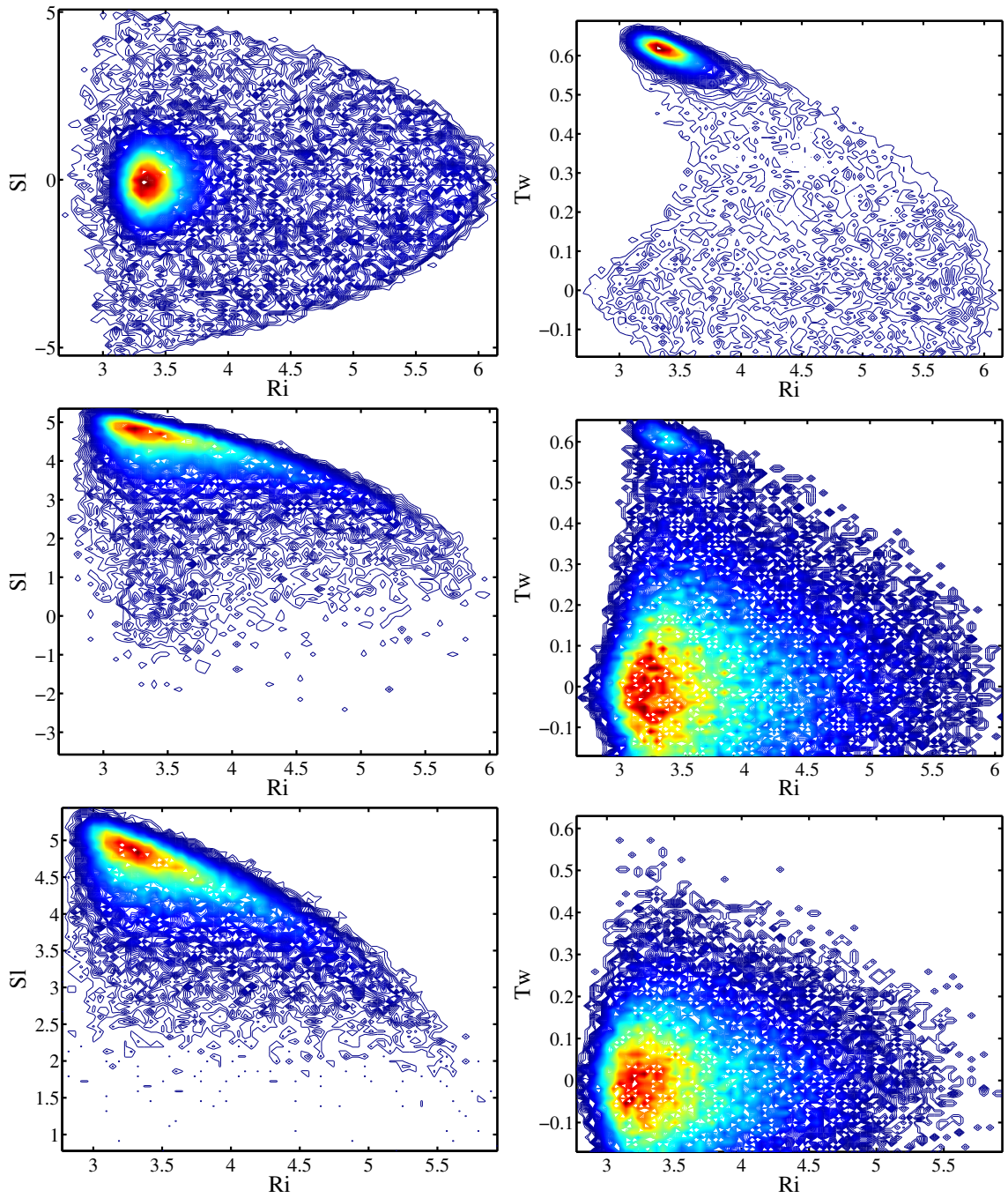


Figure 3.18: Contour plot of rise versus slide and versus twist respectively for $f = 130\text{pN}, 140\text{pN}, 250\text{pN}$. This illustrates how the S-DNA structure builds up with increasing stretching force f.

that for the S-DNA conformation the shift, twist and roll fluctuations are much larger than in the B-DNA conformation. In the S-DNA state the contact area of two successive base-pairs is much

(a) zero stretching force

(b) at the critical force

(c) beyond the critical force

Figure 3.19: Snapshots of chains with length $N = 50$ for the final parameterization of Eq. (3.52). One can clearly see the right-handed B-DNA conformation (a) and the skewed ladder structure of the S-DNA conformation (c).

smaller than in the B-DNA state such that the energy loss due to fluctuations around the S-DNA ground state is strongly reduced. In addition twist fluctuations lead to much smaller fluctuations in the backbone length.

3.5 Discussion

We have introduced a simple model Hamiltonian describing double-stranded DNA on the base-pair level. Due to the simplification of the force-field and, in particular, the possibility of non-local MC moves our model provides access to much larger length scales than atomistic simulations. For example, $4h$ on a AMD Athlon MP 2000+ processor are sufficient in order to generate 1000 independent conformations for chains consisting of $N = 100$ base-pairs.

In the data analysis, the main emphasis was on deriving the elastic constants on the elastic rod level from the analysis of thermal fluctuations of base-pair step parameters. Assuming a twisted ladder as ground state conformation one can provide an analytical relationship between the persistence lengths and the local elastic constants given by eqs. (3.26), (3.30) [1]. Future work has to show, if it is

[1]The general case where the ground state is characterized by spontaneous rotations as well as spontaneous displacements as in the A-DNA conformation is more involved. This is the subject of ongoing work.

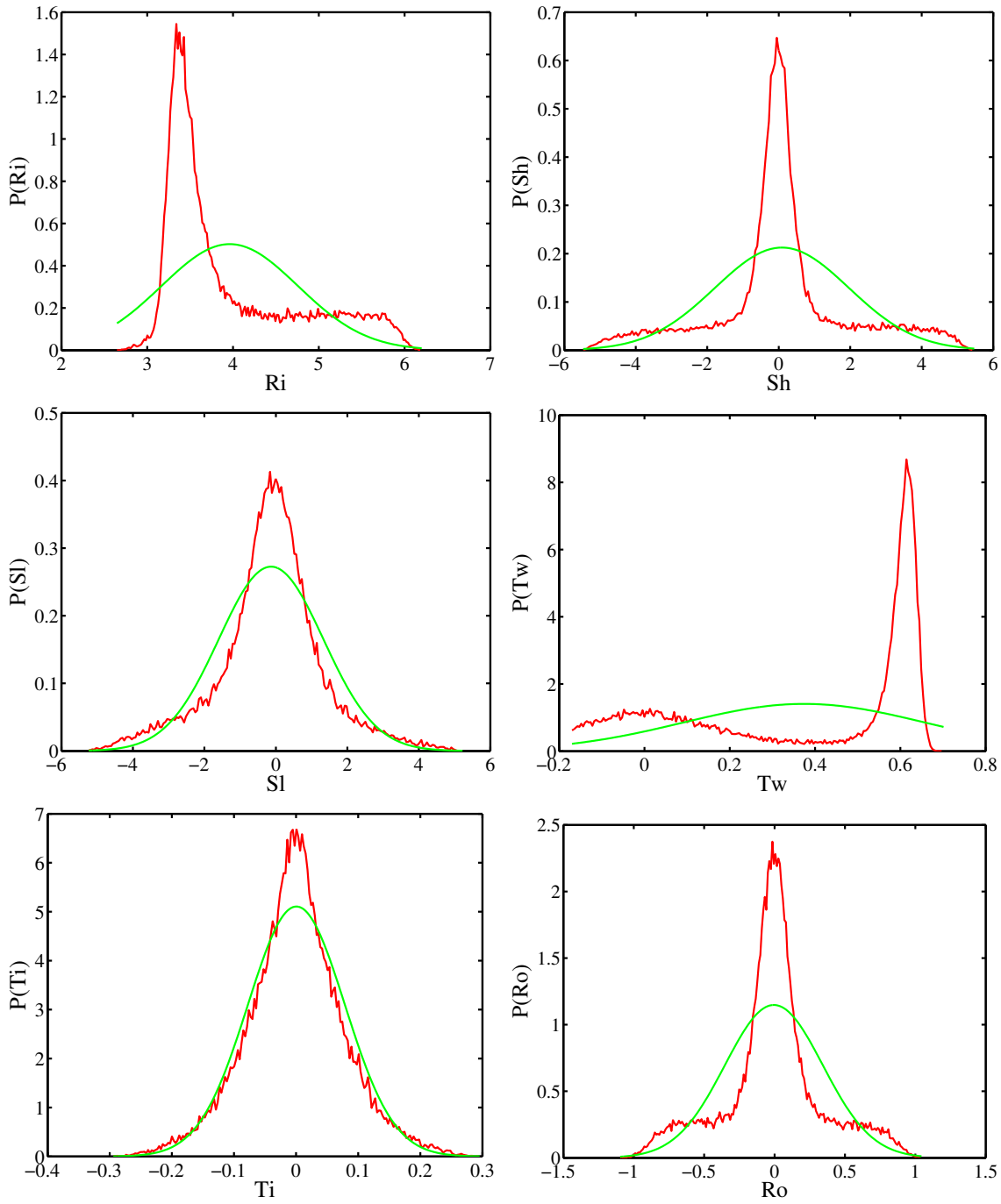


Figure 3.20: Comparison of probability distribution functions of all base-pair parameters for the final parameterization with $\epsilon = 20$, $k = 64$, $2b = 11$ and a stretching force of $f = 130$pN, i.e. slightly below the critical force. The Gaussians are plotted with the measured mean and mean squared values of the MC simulation with 50 monomers.

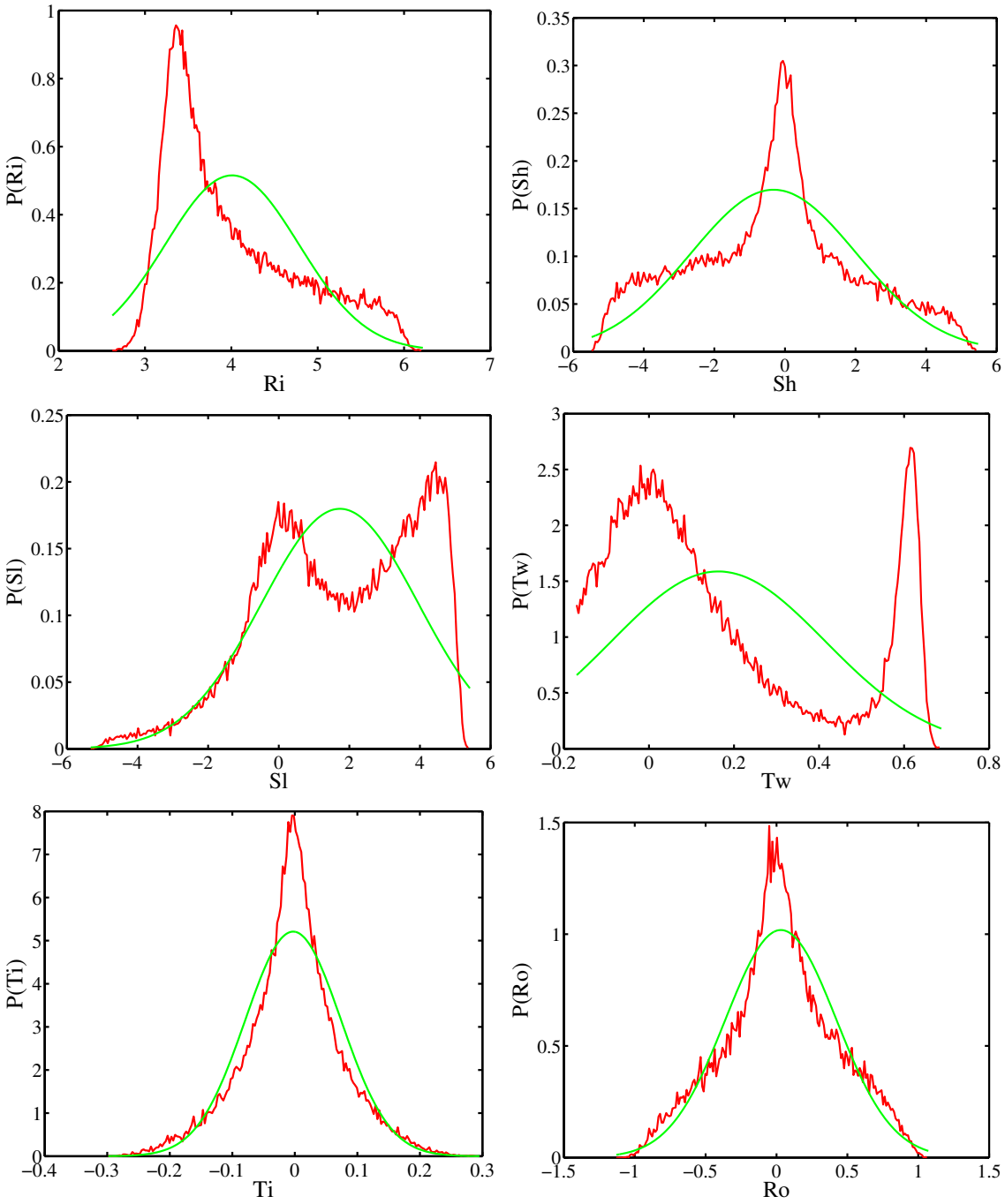


Figure 3.21: Comparison of probability distribution functions of all base-pair parameters for the final parameterization with $\epsilon = 20$, $k = 64$, $2b = 11$ and a stretching force of $f = 140$pN, i.e. at the critical force. The Gaussians are plotted with the measured mean and mean squared values of the MC simulation with 50 monomers.

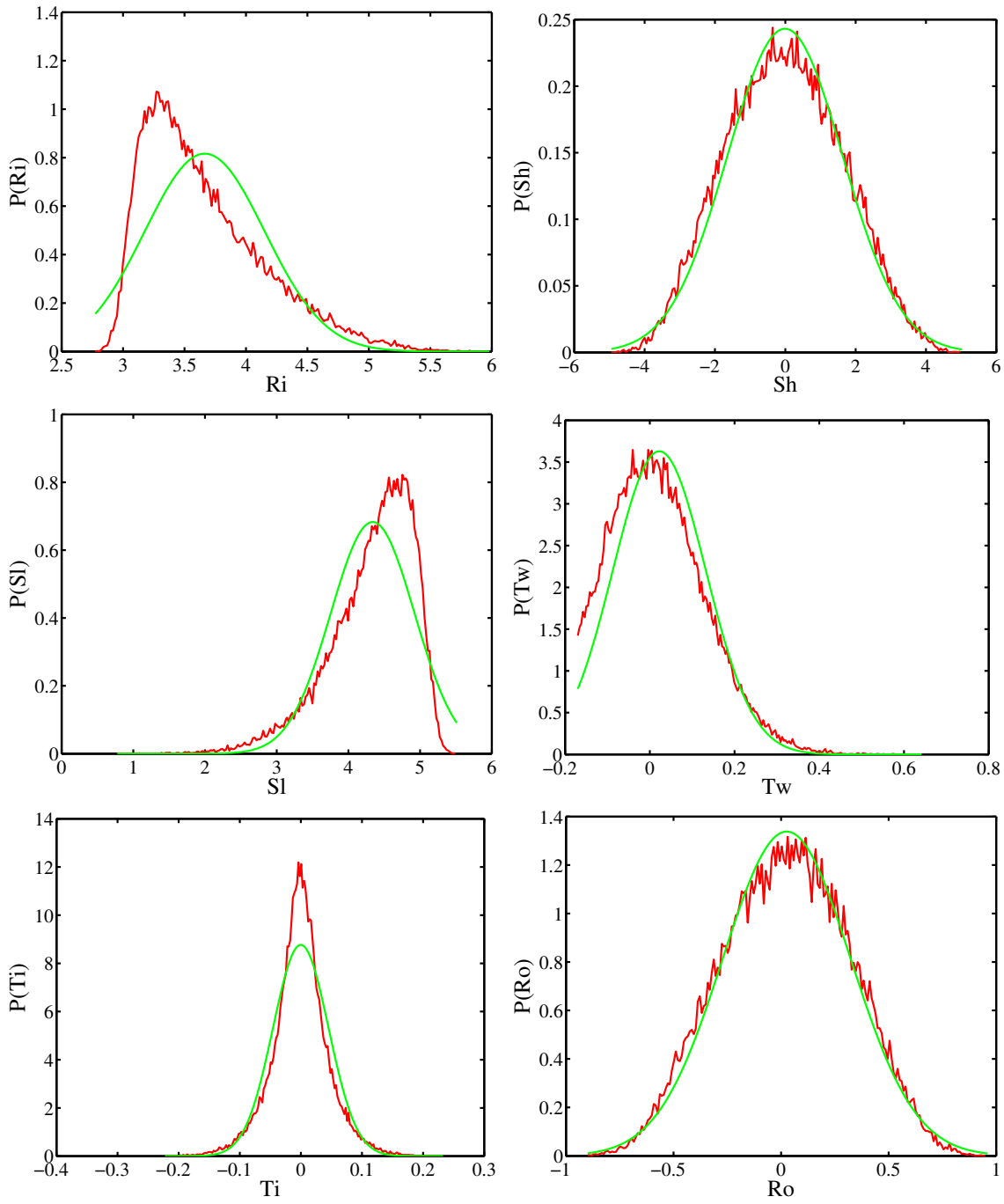


Figure 3.22: Comparison of probability distribution functions of all base-pair parameters for the final parameterization with $\epsilon = 20$, $k = 64$, $2b = 11$ and a stretching force of $f = 250$pN, i.e. beyond the critical force. The Gaussians are plotted with the measured mean and mean squared values of the MC simulation with 50 monomers.

possible to obtain suitable parameters for our mesoscopic model from a corresponding analysis of atomistic simulations [LL00a] or quantum-chemical calculations [GB99]. In the present paper, we have chosen a top-down approach, i.e. we try to reproduce the experimentally measured behavior of DNA on length scales *beyond* the base diameter. The analysis of the persistence lengths, the mean and mean squared values of all six base-pair parameters and the critical force, where the structural transition from B-DNA to S-DNA takes place, as a function of the model parameters $\{b, k, \epsilon\}$ and the applied stretching force f suggests the following parameter set:

$$\begin{aligned} 2b &= 11\text{Å} \\ \epsilon &= 20 k_B T \\ k &= 64 k_B T/\text{Å}^2. \end{aligned} \tag{3.52}$$

It reproduces the correct persistence lengths for B-DNA and entails the correct mean values of the base-pair step parameters known by X-ray diffraction studies. While the present model does not include the distinction between the minor and major groove and suppresses all internal degrees of freedom of the base-pairs such as propellor twist, it nevertheless reproduces some experimentally observed features on the base-pair level. For example, the anisotropy of the bending angles (rolling is easier than tilting) is just a consequence of the plate-like shape of the base-pairs and the twist-stretch coupling is the result of the preferred stacking of neighboring base-pairs and the rigid backbones.

The measured critical force is overestimated by a factor of 2 and cannot be improved further by fine-tuning of the three free model parameters $\{b, k, \epsilon\}$. f_{crit} depends solely on the stacking energy value ϵ that cannot be reduced further. Otherwise neither the correct equilibrium structure of B-DNA nor the correct persistence lengths would be reproduced. Our model suggests a structure for S-DNA with highly inclined base-pairs so as to enable at least partial base-pair stacking. This is in good agreement with results of atomistic B-DNA simulations by Lavery *et al.* [CLH$^+$96, LL99]. They found a force plateau of 140pN for freely rotating ends [CLH$^+$96]. The mapping to the SOP model yields the following twist-stretch (Ri-Tw) coupling constant $k_{Ri,Tw} = (\mathcal{C}^{-1})_{Ri,Tw} = 267/\text{Å}$. $k_{Ri,Tw}$ is the microscopic coupling of rise and twist describing the untwisting of the chain due to an increase of rise (compare also Fig. 3.9).

Possible applications of the present model include the investigation of (i) the charge renormalization of the WLC elastic constants [PHP00], (ii) the microscopic origins of the cooperativity of the B-to-S transition [SN02a], and (iii) the influence of nicks in the sugar-phosphate backbone on force-elongation curves. In particular, our model provides a physically sensible framework to study the intercalation of certain drugs or of ethidium bromide between base pairs. The latter is a hydrophobic molecule of roughly the same size as the base-pairs that fluoresces green and likes to slip between two base-pairs forming an DNA-ethidium-bromide complex. The fluorescence properties allow to measure the persistence lengths of DNA [SS86]. It was also used to argue that the force plateau is the result of a DNA conformational transition [CLH$^+$96].

In the future, we plan to generalize our approach to a description on the base level which includes the possibility of hydrogen-bond breaking between complementary bases along the lines of Ref. [BCP99, CM99]. A suitably parameterized model allows a more detailed investigation of DNA unzipping experiments [BERH97] as well as a direct comparison between the two mechanism currently discussed for the B-to-S transition: the formation of skewed ladder conformations (as in the present paper) versus local denaturation [WWRB01, RB01a, RB01b]. Clearly, it is possible to study

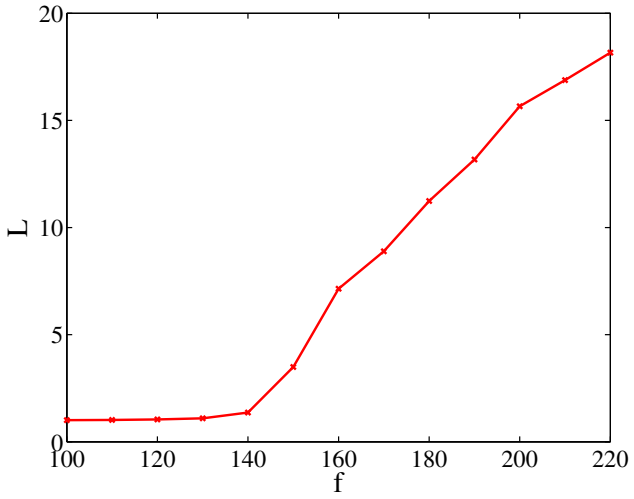
sequence-effects and even more refined models of DNA. For example, it is possible to mimic minor and major groove by bringing the backbones closer to one side of the ellipsoids without observing non-B-DNA like ground states. The relaxation of the internal degrees of freedom of the base-pairs characterized by another set of parameters (propeller twist, stagger, etc.) should help to reduce artifacts which are due to the ellipsoidal shape of the base-pairs. Sequence effects enter via the strength of the hydrogen bonds ($E_{GC} = 2.9 k_B T$ versus $E_{AT} = 1.3 k_B T$) as well as via base dependent stacking interactions [Hun93]. For example, one finds for guanine a concentration of negative charge on the major-groove edge whereas for cytosine one finds a concentration of positive charge on the major-groove edge. For adanine and thymine instead there is no strong joint concentration of partial charges [CD99]. It is known that in a solution of water and ethanol where the hydrophobic effect is less dominant these partial charges cause GG/CC steps to adopt A- or C-forms [FSH99] by a negative slide and positive roll motion and a positive slide motion respectively. Thus by varying the ratio of the strengths of the stacking versus the electrostatic energy it should be possible to study the transition from B-DNA to A-DNA and C-DNA respectively.

3.6 Summary

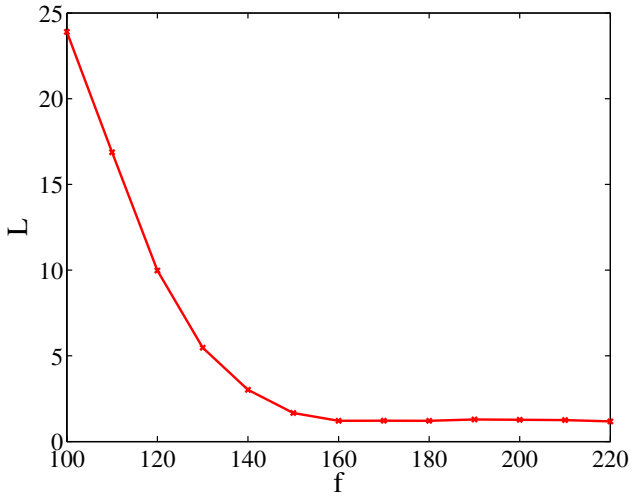
Inspired by the results of El Hassan and Calladine [HC97a] and of Hunter et al. [HL97, Hun93] we have put forward the idea of constructing simplified DNA models on the base(-pair) level where discotic ellipsoids (whose stacking interactions are modeled via coarse-grained potentials [EE03, GB81]) are linked to each other in such a way as to preserve the DNA geometry, its major mechanical degrees of freedom and the physical driving forces for the structure formation [CD99].

In the present paper, we have used energy minimization and Monte Carlo simulations to study a simple representative of this class of DNA models with non-separable base-pairs. For a suitable choice of parameters we obtained a B-DNA like ground state as well as realistic values for the bend and twist persistence lengths. The latter were obtained by analyzing the thermal fluctuations of long filaments as well as by a systematic coarse-graining from the stack-of-plates to the elastic rod level. In studying the response of DNA to external forces or torques, models of the present type are not restricted to the regime of small local deformations. Rather by specifying a physically motivated Hamiltonian for *arbitrary* base-(step) parameters, our ansatz allows for realistic local structural transitions. For the simple case of a stretching force we observed a transition from a twisted helix to a skewed ladder conformation. While our results suggest a similar structure for S-DNA as atomistic simulations [CLH$^+$96], the DNA model studied in this paper can, of course, not be used to rule out the alternate possibility of local strand separations [WWRB01, RB01a, RB01b].

In our opinion, the base(-pair) level provides a sensible compromise between conceptual simplicity, computational cost and degree of reality. Besides providing access to much larger scales than atomistic simulations, the derivation of such models from more microscopic considerations provides considerable insight. At the same time, they may serve to validate and unify analytical approaches aiming at (averaged) properties on larger scales [ALCM01, HYZc99, BCP99, CM99, SN02a]. Finally we note that the applicability of linked-ellipsoid models is not restricted to the base-pair level of DNA as the same techniques can, for example, also be used to study chromatin [WL02, KBO00, MSE]. This is the subject of chapter 5.



(a) Typical domain length of S-DNA.



(b) Typical domain length of B-DNA.

Figure 3.23: Typical domain length L [bp] of (a) S-DNA, (b) B-DNA as a function of the stretching force f [pN]. The parameter set of Eq. (3.52) is used. The simulated chains consist of 500 base-pairs. We averaged over three independent runs. There are two S-DNA states corresponding to positive (+) and negative (-) slide. None of the states should be preferred which is indicated by the simulation data.

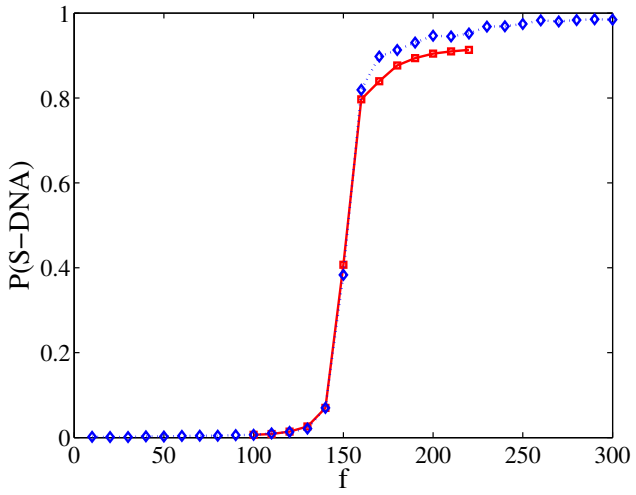
3.7 Outlook

Recently Storm and Nelson [SN02b, SN02a] introduced a two-state model for the elasticity of double-stranded DNA that could be solved analytically. The energy contains an Ising part taking into account the free energy cost of converting a single segment from B- to S-state and the energy creating a B-to-S-interface, an elastic part taking into account the bending stiffness of B- and S-DNA and the bending rigidity between a S- and B-segment. The full energy functional reads

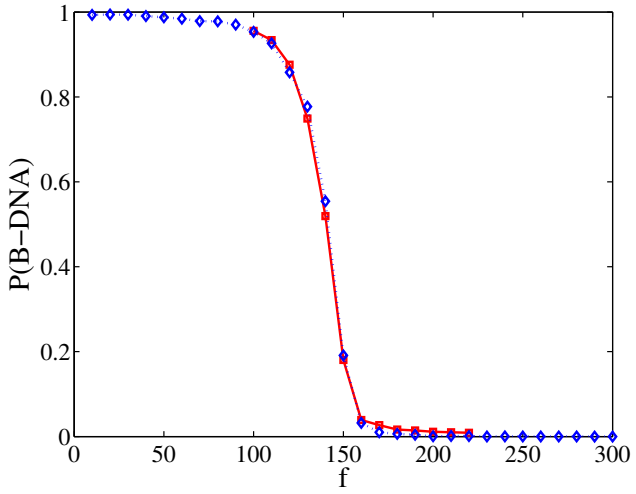
$$\begin{aligned}\frac{\mathcal{E}}{k_BT} = -\sum_{i=1}^{N-1}\Big[&\frac{\alpha}{2}(\sigma_i+\sigma_{i+1})+\gamma(\sigma_i\sigma_{i+1}-1)\\ &+\frac{fb}{2k_BT}\left(\frac{1}{2}((\sigma_i+1)+(\sigma_i-1)\zeta)\mathbf{t}_i\cdot\mathbf{z}+\frac{1}{2}((\sigma_{i+1}+1)+(\sigma_{i+1}-1)\zeta)\mathbf{t}_{i+1}\cdot\mathbf{z}\right)\\ &-\frac{A}{2b}\left(\frac{\beta}{4}(1-\sigma_i)(1-\sigma_{i+1})+|\sigma_i-\sigma_{i+1}|\eta+\frac{1}{4}(1+\sigma_i)(1+\sigma_{i+1})\right)\theta_{i,i+1}^2\Big].\end{aligned} \tag{3.53}$$

σ_i takes into account in which state segment i is found. $\sigma_i = 1$ corresponds to the B-state and $\sigma_i = -1$ to the S-state. $2\alpha k_BT$ denotes the B-to-S conversion energy, $2\gamma k_BT$ the B-to-S interfacial energy. b refers to the B-DNA segment length, ζb to the S-DNA segment length. The bending stiffness of B-DNA (S-DNA) is denoted by A ($\beta\zeta A$). The bending rigidity of a B-DNA and a S-DNA segment is termed ηA. f refers to the applied stretching force and $\mathbf{z}$ is a normalized vector pointing into the stretching direction. $\mathbf{t}_i$ denotes the tangent vector of the ith segment, θ_i corresponds to the bending angle between segment i and $i+1$. Using transfer matrix techniques they fitted their theory to the force-extension data of Refs. [CLH$^+$96, SCB96]. The stretching moduli for the B- (E^B) and S-DNA (E^S) is implemented to first order by replacing f with $f(1+\frac{f}{2E^{S,B}})$ for the two respective states in the transfer matrix. With the obtained values for $2\alpha k_BT$, $2\gamma k_BT$, ζ, E^B, E^S, and A they calculated the typical length of a S-domain and the relative population of the S-state versus the applied stretching force. It is possible to apply their theory to our simple but realistic DNA model. With our model it is possible to measure these quantities directly and to compare it to the fitting results of Storm and Nelson [SN02b, SN02a]. One major difference is that it is not possible to define a pure B- or S-state as it is the case in the theory. Moreover our model allows for more than just two states. Open ladder conformations are found, too, and the S-state is divided in a S(+)- and S(-)-state with positive and negative slide respectively. None of the S-state is preferred.

We analyze data of simulations with applied stretching force and chain lengths of $N = 50$ and $N = 500$. We use the final parameterization of Eq. (3.52). Without stretching force the simulated chain fluctuates around the B-DNA form with $\langle\Delta\text{Ri}^2\rangle = 0.02$, $\langle\Delta\text{Sh}^2\rangle = 0.12$, $\langle\Delta\text{Sl}^2\rangle = 0.42$, $\langle\Delta\text{Tw}^2\rangle = 0.001$, $\langle\Delta\text{Ro}^2\rangle = 0.010$, $\langle\Delta\text{Ti}^2\rangle = 0.003$, whereas beyond the critical force at 210pN the fluctuations around the S-form amount to $\langle\Delta\text{Ri}^2\rangle = 0.21$, $\langle\Delta\text{Sh}^2\rangle = 2.32$, $\langle\Delta\text{Sl}^2\rangle = 0.28$, $\langle\Delta\text{Tw}^2\rangle = 0.011$, $\langle\Delta\text{Ro}^2\rangle = 0.08$, $\langle\Delta\text{Ti}^2\rangle = 0.002$. That is why we count conformations with Sl > 3.0 as S(+)-state and with Sl < -3.0 as S(-)-state respectively. A B-DNA conformation is counted if Tw $> \pi/6$. For the calculation of the compartment length of a S- or a B-domain we determine the mean number of base-pairs in a row belonging to one domain. Eq. (3.12) can be used to measure the bending persistence length of S-DNA. We find the effective bending persistence length of the S-state to be about 60Å, which is order of magnitudes smaller than the value found for the B-state. It is about three times as large as it is expected for two single DNA strands within our model. Note that the Kuhn segment length corresponds to two times the backbone length such



(a) Relative population of S-DNA counting both S-state.



(b) Relative population of B-DNA.

Figure 3.24: Relative population P of S-DNA and B-DNA as a function of applied strain f [pN] for chains with 500 (blue) and 50 (red) base-pairs. The parameter set of Eq. (3.52) is used. The measured population of the S-state looks very similar to the result calculated by [SN02b, SN02a].

that $l_{p,ssDNA} = 12$Å. Thus the partial stacking of the base-pairs in the skewed ladder conformation contributes to the bending stiffness of the chain. Storm and Nelson [SN02b] predict persistence lengths of about $72 - 120$Å from the fitting of the experimental data of Refs. [CLH$^+$96] and [SCB96] respectively. The value obtained from stretching experiments on ssDNA amounts to about 7.5Å [CSRTG00].

Figs. 3.23, 3.24 show measured relative populations and typical domain lengths of S-DNA and B-DNA as a function of applied strain. In contrast to the model of Storm and Nelson [SN02b] the S-DNA state is divided into a S(+) state with positive slide and a S(-) state with negative slide. Neither should be preferred. In fact the analysis of three different runs with 500 base-pairs indicate that the relative population of the S(+) and S(-) state beyond the critical force amounts to about 0.5. It has to be noted that the diffusion time of the interface seems to be quite large such that one has to average over a sufficiently large number of independent runs. The measured functions of the S-state, where the S(+)- as well as S(-)-state are counted, are quite similar to the result obtained by Storm and Nelson [SN02b]. We measure an asymptotic slope of the increase of domain length of S-DNA of about $0.25\,\frac{\mathrm{bp}}{\mathrm{pN}}$ which is roughly 40 times smaller than the value determined by Storm and Nelson. A first analysis of the interface between the S(+) and S(-) state indicate that the segments in between unstack to form an open ladder conformation. For large stretching forces essentially no B-DNA states are left.

Chapter 4

Renormalization of coupling constants

We presented in the previous chapter a systematic way of coarse-graining. For a pure B-DNA ground state, i.e. there is only a spontaneous twist involved, and under the assumption of local interactions (no next-nearest neighbor interactions) we calculated the bending and torsional persistence length as a function of the mean-square fluctuations of the rotational degrees of freedom. For more complex ground state geometries, i.e. with spontaneous displacements and spontaneous bending angles, the derived relationships are not valid anymore. In what follows we discuss how to rescale the effective coupling constants using renormalization techniques if one goes to larger length scales.

Since in the presented model only local interactions are involved one can determine the renormalized coupling constants with a simple sampling procedure using the measured coupling constants of section 3.4.1. In such a way one can produce a long sequence of base-pairs which can subsequently be analyzed for neighbors, next-nearest neighbors and so forth. Snapshots of the simple sample method for B- and A-DNA conformations are illustrated in Fig. 4.2. If the model contains long-range interactions one has to use Monte-Carlo renormalization methods [LB00].

Renormalization always implies that certain degrees of freedom will be integrated out. In the present case there are six degrees of freedom per base-pair (3 translations and 3 rotations) which can be expressed by six step parameters or by fluctuations of positions and orientations of each base-pair. We will discuss two general strategies to renormalize the measured elastic constants for Gaussian systems on larger length scales. One strategy involves a blocking procedure where a certain number of variables is combined in one 'super'-variable whereas in the other strategy the renormalized coupling constants between variable i and $i + n_{renorm}$ are calculated where $n_{renorm} > 1$ is an integer. As an example consider the renormalization of Ising spin systems. Blocking corresponds to summing up a certain amount of spins to one 'super' spin whereas decimation corresponds to choosing a single spin out of each renormalized cell [LB00]. In general one can express the transformation by an operator $\mathfrak{R}$ acting on the Hamiltonian to be renormalized. $\mathfrak{R}$ reduces the number of degrees of freedom by a factor of n_{renorm}. The free energy remains unchanged by the renormalization process since it is only expressed in terms of new variables. Note that due to the renormalization procedure one can encounter interactions between monomers separated by an arbitrarily large distance.

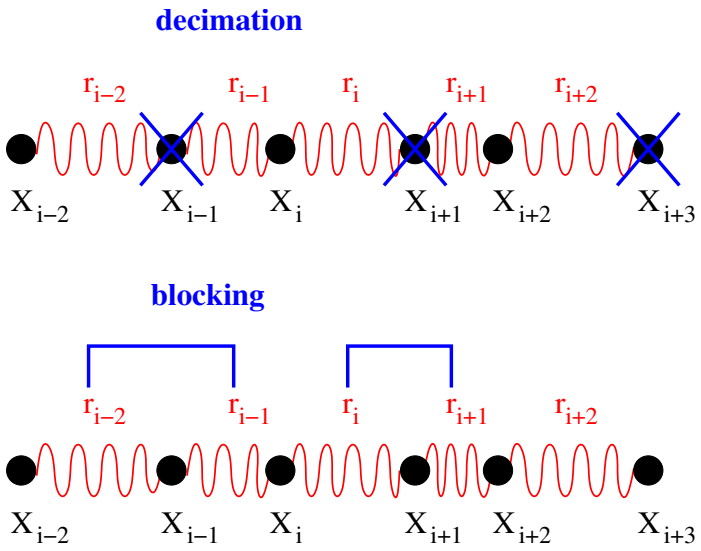


Figure 4.1: Illustration of blocking and decimation for a one dimensional Gaussian chain. The chain variables are either X_i representing the position vectors of each monomer or the center-center distances of neighboring monomers $r_i = X_{i+1} - X_i$. Decimation of every second position vector X_i is equivalent to blocking of r_i and r_{i+1}.

4.1 Renormalization of Gaussian systems with additive variables

In general the Hamiltonian of a Gaussian system can be expressed as

$$\begin{aligned} \frac{\mathcal{H}}{k_B T} &= (\underline{X} - \underline{X}_0)^T \, \underline{\underline{K}} \, (\underline{X} - \underline{X}_0) \\ &= \underline{X}^T \, \underline{\underline{K}} \, \underline{X} - 2\underline{X}_0^T \, \underline{\underline{K}} \, \underline{X} + \underline{X}_0^T \, \underline{\underline{K}} \, \underline{X}_0. \end{aligned} \tag{4.1}$$

$\underline{\underline{K}}$ denotes the elastic matrix with $d\,N \times d\,N$ dimensions where N is the number of monomers and d is the number of degrees of freedom. Its entries are the coupling constants along the chain analogously to the connectivity matrix of a polymer network. Note that $\underline{\underline{K}}$ is symmetric. $\underline{X}$ is a dN-dimensional vector. $\underline{X}_0$ refers to spontaneous values around which $\underline{X}$ fluctuates. The free energy is given by $\mathcal{F} = k_B T \ln(\mathcal{Z}) = \mathcal{H}$ where the partition function $\mathcal{Z}$ is given by

$$\mathcal{Z} = \int \mathcal{D}[\underline{X}] \exp\left(-\frac{\mathcal{H}}{k_B T}\right). \tag{4.2}$$

In the following we consider two renormalization procedures of a Gaussian system (as illustrated in Fig. 4.1 for a 1d Gaussian chain) with a generalized Hamiltonian of the form

$$\frac{\mathcal{H}_g}{k_B T} = \underline{X}^T \, \underline{\underline{K}} \, \underline{X} - 2\underline{k}^T \, \underline{X} + C \tag{4.3}$$

where C is a constant. We calculate the scaling of the effective interactions of the remaining variables under the assumption that the variables $\underline{X}$ are *additive*. We split the Hamiltonian into two parts corresponding to the variables $\underline{X}_1$ we want to keep and the variables $\underline{X}_2$ we integrate out. The

Hamiltonian can be rewritten in terms of the new variables and the matrices that connect them as follows [Mag85]:

$$\begin{aligned}\frac{\mathcal{H}_g}{k_BT} &= \underline{X}^T\,\underline{\underline{K}}\,\underline{X} + 2\underline{k}^T\,\underline{X} + C \\ &= \begin{pmatrix}\underline{X}_1^T & \underline{X}_2^T\end{pmatrix}\begin{pmatrix}\underline{\underline{K}}_{11} & \underline{\underline{K}}_{12} \\ \underline{\underline{K}}_{21} & \underline{\underline{K}}_{22}\end{pmatrix}\begin{pmatrix}\underline{X}_1 \\ \underline{X}_2\end{pmatrix} + 2\begin{pmatrix}\underline{k}_1^T & \underline{k}_2^T\end{pmatrix}\begin{pmatrix}\underline{X}_1 \\ \underline{X}_2\end{pmatrix} + C \\ &= \underline{X}_1^T\,\underline{\underline{K}}_{11}\,\underline{X}_1 + \underline{X}_1^T\,(\underline{\underline{K}}_{12} + \underline{\underline{K}}_{21}^T)\,\underline{X}_2 + \underline{X}_2^T\,\underline{\underline{K}}_{22}\,\underline{X}_2 + 2\underline{k}_1^T\,\underline{X}_1 + 2\underline{k}_2^T\,\underline{X}_2 + C.\end{aligned} \tag{4.4}$$

4.1.1 Decimation

Decimation implies keeping every second, or every third, etc. monomer and integrating out all other variables. The integration over the variables $\underline{X}_2$ can be performed by completing the square:

$$\int \mathcal{D}[\underline{Y}]\,\exp(-\underline{Y}^T\,\underline{\underline{A}}\,\underline{Y} - \underline{b}^T\,\underline{Y}) = \frac{\exp(\underline{b}^T\,\underline{\underline{A}}^{-1}\,\underline{b})}{\sqrt{\det(\mathcal{A}_{\mathcal{J}})}} \tag{4.5}$$

with

$$(\mathcal{A}_{\mathcal{J}})_{ij} = \frac{\partial^2(\underline{Y}^T\,\underline{\underline{A}}\,\underline{Y})}{\partial y_i \partial y_j}. \tag{4.6}$$

This results in the following equations for the scaling of the effective coupling constants and the spontaneous values contained in $\underline{k} = \underline{X}_0^T\underline{\underline{K}}$ and $C = \underline{X}_0^T\underline{\underline{K}}\,\underline{X}_0$:

$$\begin{aligned}\underline{\underline{K}}' &= \underline{\underline{K}}_{11} - \frac{1}{4}(\underline{\underline{K}}_{12}^T + \underline{\underline{K}}_{21})^T\,\underline{\underline{K}}_{22}^{-1}(\underline{\underline{K}}_{12}^T + \underline{\underline{K}}_{21}) \\ \underline{k}' &= \underline{k}_1 - \frac{1}{2}(\underline{\underline{K}}_{12}^T + \underline{\underline{K}}_{21})\,\underline{\underline{K}}_{22}^{-1}\underline{k}_2 \\ C' &= C - \underline{k}_2^T\,\underline{\underline{K}}_{22}^{-1}\,\underline{k}_2 + \frac{1}{2}\ln(\det(\mathcal{K}_{\mathcal{J},22}))\end{aligned} \tag{4.7}$$

Note that C scales in order to conserve the absolute value of the free energy after each decimation step, i.e. $\mathcal{F}_g' = \mathcal{F}_g'' = \ldots = \mathcal{F}_g$. The renormalized partition function is thus given by

$$\mathcal{Z}_g' = \int \mathcal{D}[\underline{X}_1]\exp\left(-\frac{\mathcal{H}_g'}{k_BT}\right)\det(\mathcal{K}_{\mathcal{J},22})^{-\frac{1}{2}} \tag{4.8}$$

with

$$\frac{\mathcal{H}_g'}{k_BT} = \underline{X}_1^T\,\underline{\underline{K}}'\,\underline{X}_1 + \underline{k}'^T\,\underline{X}_1 + C'. \tag{4.9}$$

4.1.2 Blocking

The calculation of the renormalization formulas for the blocking procedure is slightly more involved. In general the integration scheme can be written as

$$\int \prod_j \mathcal{D}[\underline{X}_j]\exp(\frac{\mathcal{H}_g}{k_BT})\delta(\underline{Y} - \sum_{i=1}^{k}\underline{X}_i). \tag{4.10}$$

The vectors $\underline{X}_i$ are chosen such that each block contains k variables. For the sake of simplicity we discuss the case where each block consists of two variables. Blocking more than two variables can be done iteratively.

Starting again from Eq. (4.4) one has to perform two integrations where in the first step (1) one of the variables, say $\underline{X}_1$, is replaced by $\underline{Y} - \underline{X}_2$. The second integration step (2) yields the renormalized Hamiltonian expressed in terms of the new variables $\underline{Y}$:

$$\int \mathcal{D}[\underline{X}_1]\,\mathcal{D}[\underline{X}_2]\exp(\frac{\mathcal{H}_g}{k_BT})\delta\left(\underline{Y} - (\underline{X}_1 + \underline{X}_2)\right)$$

$$\stackrel{(1)}{\Longrightarrow}$$

$$\underline{X}_2^T(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})\underline{X}_2 + \left[\underline{Y}^T(\underline{\underline{K}}_{12} + \underline{\underline{K}}_{21}^T - 2\underline{\underline{K}}_{11}) + 2(\underline{k}_2 - \underline{k}_1)^T\right]\underline{X}_2 + \underline{Y}^T\,\underline{\underline{K}}_{11}\,\underline{Y} + 2\underline{k}_1^T\underline{Y} + C)$$

$$\stackrel{(2)}{\Longrightarrow} \frac{\mathcal{H}'_g}{k_BT} =$$

$$\underline{Y}^T\left[\underline{\underline{K}}_{11} - \frac{1}{4}(\underline{\underline{K}}_{12} + \underline{\underline{K}}_{21}^T - 2\underline{\underline{K}}_{11})(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})^{-1}(\underline{\underline{K}}_{12}^T + \underline{\underline{K}}_{21} - 2\underline{\underline{K}}_{11})\right]\underline{Y}$$
$$+ 2\left[\underline{k}_1 - \frac{1}{2}(\underline{\underline{K}}_{12} + \underline{\underline{K}}_{21}^T - 2\underline{\underline{K}}_{11})(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})^{-1}(\underline{k}_2 - \underline{k}_1)\right]^T\underline{Y}$$
$$+ C - (\underline{k}_2 - \underline{k}_1)^T(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})^{-1}(\underline{k}_2 - \underline{k}_1). \tag{4.11}$$

Hence we obtain the following scaling relations for the coupling constants and the spontaneous values:

$$\begin{aligned}
\underline{\underline{K}}' &= \underline{\underline{K}}_{11} - \frac{1}{4}(\underline{\underline{K}}_{12} + \underline{\underline{K}}_{21}^T - 2\underline{\underline{K}}_{11})(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})^{-1}(\underline{\underline{K}}_{12}^T + \underline{\underline{K}}_{21} - 2\underline{\underline{K}}_{11}) \\
\underline{k}' &= \underline{k}_1 - \frac{1}{2}(\underline{\underline{K}}_{12} + \underline{\underline{K}}_{21}^T - 2\underline{\underline{K}}_{11})(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})^{-1}(\underline{k}_2 - \underline{k}_1) \\
C' &= C - (\underline{k}_2 - \underline{k}_1)^T(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})^{-1}(\underline{k}_2 - \underline{k}_1) + \frac{1}{2}\ln(\det(\mathcal{K}_\mathcal{J}))
\end{aligned} \tag{4.12}$$

where the matrix $\mathcal{K}_\mathcal{J}$ is given by

$$(\mathcal{K}_\mathcal{J})_{ij} = \frac{\partial^2\left(\underline{X}_2^T(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12} - \underline{\underline{K}}_{21}^T + \underline{\underline{K}}_{22})\underline{X}_2\right)}{\partial x_{2,i}\partial x_{2,j}}. \tag{4.13}$$

Note that the discussed decimation and blocking strategy are special cases of a general renor-

malization scheme [Mag85] given by

$$\exp\left(-\frac{\underline{X}'^T\,\underline{\underline{K}}'\,\underline{X}'}{k_BT}\right)=\int\mathcal{D}[\underline{X}]\,\exp\left(-\frac{\underline{X}^T\,\underline{\underline{K}}\,\underline{X}}{k_BT}\right)\delta\left(\underline{X}'-\underline{\underline{\Gamma}}\,\underline{X}\right). \tag{4.14}$$

Because of the special nature of Gaussian integrals the renormalized Hamiltonians will always be quadratic in the variables $\underline{X}'$.

The derived scaling relations are only valid for the renormalization of *Gaussian* systems with *additive* variables. Interestingly, even though the distribution functions are non-Gaussian one can use the renormalization procedure to calculate the rescaled coupling constants within linear response theory, since it yields

$$\begin{aligned}\langle\underline{X}'_1\underline{X}'^T_2\rangle&=\int\mathcal{D}[\underline{X}']\,\underline{X}'_1\underline{X}'^T_2\exp\left(-\frac{\mathcal{H}'(\underline{X}')}{k_BT}\right)\\&=\int\mathcal{D}[\underline{X}']\,\underline{X}'_1\underline{X}'^T_2\int\mathcal{D}[\underline{X}]\,\exp\left(-\frac{\mathcal{H}(\underline{X})}{k_BT}\right)\delta\left(\underline{X}'-\underline{\underline{\Gamma}}\underline{X}\right)\\&=\int\mathcal{D}[\underline{X}]\,(\underline{\underline{\Gamma}}\,\underline{X}_1)(\underline{\underline{\Gamma}}\,\underline{X}_2)^T\exp\left(-\frac{\mathcal{H}(\underline{X})}{k_BT}\right)\\&=\langle\underline{\underline{\Gamma}}\,\underline{X}_1\underline{X}_2^T\,\underline{\underline{\Gamma}}^T\rangle\\&=\underline{\underline{\Gamma}}\,\langle\underline{X}_1\underline{X}_2^T\rangle\,\underline{\underline{\Gamma}}^T.\end{aligned} \tag{4.15}$$

4.1.3 Comparison of derived scaling relations with brute force integration

We derived a calculation scheme for the rescaled variables of a Gaussian system for decimation and blocking. In order to test whether Eqs. (4.7) and (4.12) are correct we constructed a 24×24 large symmetric matrix $\underline{\underline{K}}$ and compared rescaling results of our matrix manipulation scheme using MatLab with brute force integration using Mathematica, that is

$$\int\prod_{i=1}^{6}\prod_{i=13}^{18}dx_i\exp\left(-\underline{X}^T\,\underline{\underline{K}}\,\underline{X}\right) \tag{4.16}$$

in case of decimation and

$$\int\prod_{i=1}^{24}dx_i\exp\left(-\underline{X}^T\,\underline{\underline{K}}\,\underline{X}\right)\prod_{j=1}^{6}\delta\left(y_j-(x_j+x_{j+6})\right)\,\delta\left(y_j-(x_{j+12}+x_{j+18})\right) \tag{4.17}$$

in case of blocking. The resulting rescaled 12×12 large matrices are identical (data not shown).

4.2 Discussion

We already know that both the rotational and the translational base-pair step parameters are in general *non-additive*. The angular parameters between base-pair i and j can be calculated by the scheme discussed in appendix B where the rotation matrix $\mathcal{R}_{i,j}$ is given by

$$\mathcal{R}_{i,j}=\mathcal{A}_i^T\mathcal{A}_j \tag{4.18}$$

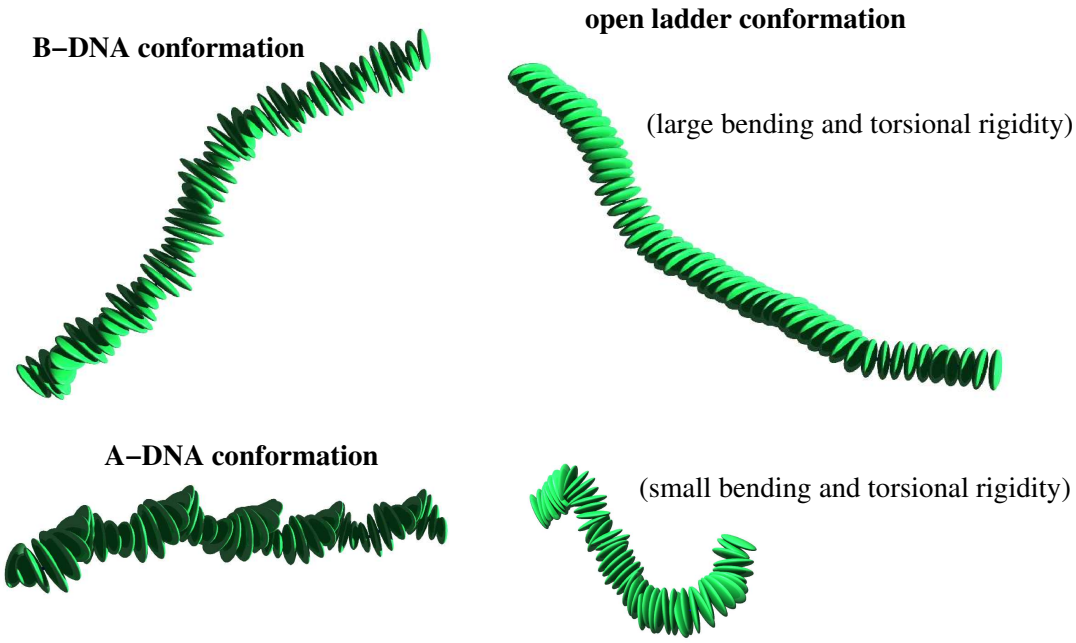


Figure 4.2: Simple sampling snapshots for a B-DNA conformation with coupled twist and rise motion, a A-DNA conformation with negative slide (-2.0Å) and positive roll (0.1^o), and straight conformations with no spontaneous rotations or displacements.

with $\mathcal{A}_i = \prod_{j=1}^{i} \mathcal{R}_{j,j-1}$ and $\mathcal{R}_{i,i-1} = \mathcal{R}(\mathrm{Tw}_{i,i-1}, \mathrm{Ti}_{i,i-1}, \mathrm{Ro}_{i,i-1})$. Eq. (4.18) defines how the rotational step parameters sum up along the chain. Only for small rotations one finds that the angular parameters such as Tw, Ti, Ro are additive while even in that case the translational degrees of freedom are not. Going from one base-pair to the next-nearest neighbor implies then for the rotational base-pair step parameters

$$\begin{aligned} \mathrm{Tw}_{i-1,i+1} &= \mathrm{Tw}_{i-1,i} + \mathrm{Tw}_{i,i+1} \\ \mathrm{Ti}_{i-1,i+1} &= \mathrm{Ti}_{i-1,i} + \mathrm{Ti}_{i,i+1} \\ \mathrm{Ro}_{i-1,i+1} &= \mathrm{Ro}_{i-1,i} + \mathrm{Ro}_{i,i+1}. \end{aligned} \tag{4.19}$$

On the other hand the calculation of the translational step parameters is more complicated. The translational vector $\mathcal{T}_{i-1,i}$ between two neighboring base-pairs i and $i-1$ expressed in terms of $\mathrm{Ri}_{i-1,i}$, $\mathrm{Sh}_{i-1,i}$, and $\mathrm{Sl}_{i-1,i}$ (see Eq. (B.15)) can be written as

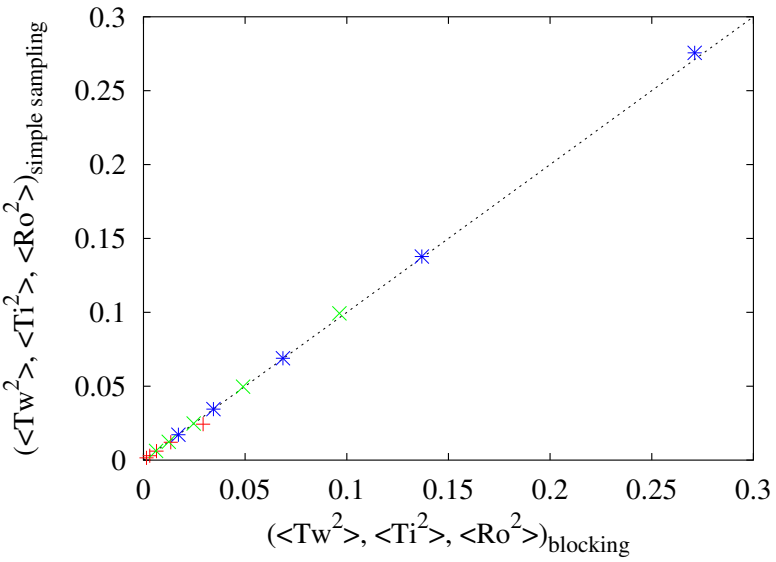
$$\mathcal{T}_{i-1,i} = \mathcal{R}_{i-1,i}^{-\frac{1}{2}} \mathcal{A}_i^{-1} (\vec{c}_i - \vec{c}_{i-1}) \tag{4.20}$$

with $\vec{c}_i$ being the position vector of the center-of-mass of base-pair i (see appendix B). Using iteratively Eq. (4.20) and

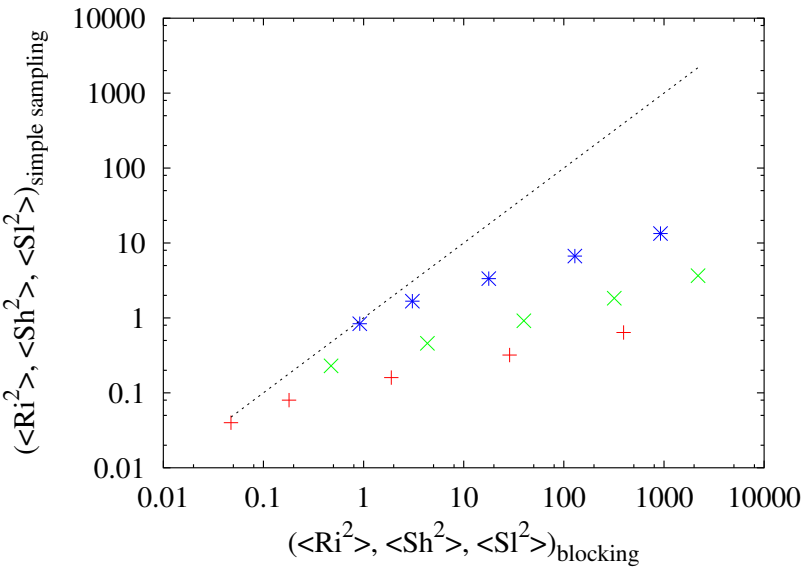
$$\vec{c}_i = \vec{c}_{i-1} + \mathcal{A}_i \mathcal{R}_{i-1,i}^{\frac{1}{2}} \mathcal{T}_{i-1,i} \tag{4.21}$$

one can calculate the translational vector between next-nearest neighbors:

$$\mathcal{T}_{i-1,i+1} = \mathcal{R}_{i-1,i+1}^{-\frac{1}{2}} \mathcal{A}_{i+1}^{-1} \mathcal{A}_i \mathcal{R}_{i-1,i}^{\frac{1}{2}} \mathcal{T}_{i-1,i} + \mathcal{R}_{i-1,i+1}^{-\frac{1}{2}} \mathcal{R}_{i,i+1}^{\frac{1}{2}} \mathcal{T}_{i,i+1}. \tag{4.22}$$



(a) $\langle Tw^2 \rangle$ (red), $\langle Ti^2 \rangle$ (green), $\langle Ro^2 \rangle$ (blue).



(b) $\langle Ri^2 \rangle$ (red), $\langle Sh^2 \rangle$ (green), $\langle Sl^2 \rangle$ (blue).

Figure 4.3: Comparison of simple sampling results for large bending and torsional stiffness with rescaling procedure for 2, 4, 8, 16, and 32 blocked variables

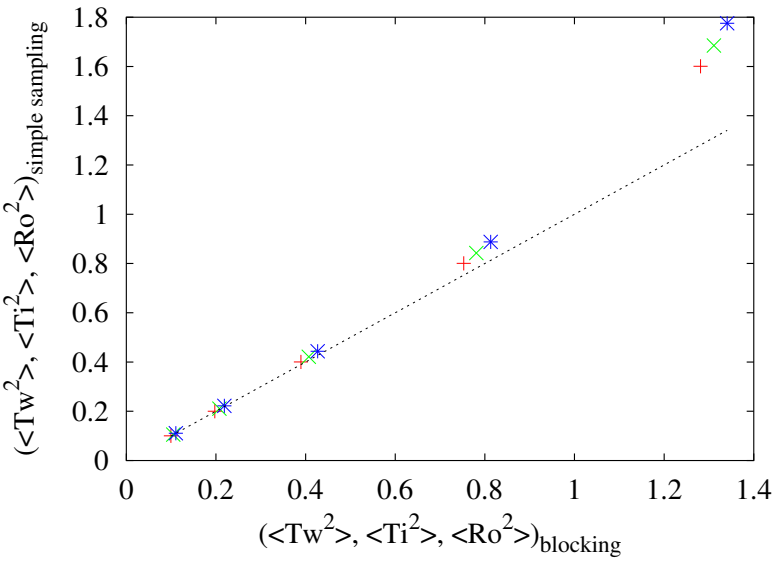


Figure 4.4: Comparison of simple sampling for small bending and torsional stiffness results with rescaling procedure for 2, 4, 8, 16, and 32 blocked variables.

This illustrates the *non-additive* behavior of the translational degrees of freedom under blocking. Thus one cannot expect the rescaling formula of Eq. (4.12) to work. Moreover it is very unlikely that one can find a rescaling relation in case of blocking using Eqs. (4.22) and (4.18) since the integrals to be carried out are not Gaussian anymore.

To check this further we determined renormalized coupling constants using simple sampling of a stack-of-plates model. First of all we consider a system where no spontaneous displacements and no spontaneous rotations are involved. Furthermore there are no internal couplings such as a twist-stretch coupling. Every step is sampled using

$$x = x_{sp} + \frac{1}{\sqrt{k_x}} \cdot n_{rand}\left([-1;1]\right) \tag{4.23}$$

where x is a step parameter, k_x refers to the corresponding coupling constant, x_{sp} denotes a spontaneous value, and $n_{rand}\left([-1;1]\right)$ is a Gaussian random number between 1 and -1. 20000 chain conformations are produced and subsequently analyzed. For the measurement of a decimated chain with renormalization number n_{renorm} we sampled $20000 \cdot n_{renorm}$ to have the same statistics. The elastic constants are chosen such that only small angles are involved. The resulting renormalized coupling constants can be compared with the result obtained by Eq. (4.12) using the original elastic matrix $\mathcal{K}$. Therefore we build a 12×12 matrix

$$\begin{pmatrix} \mathcal{K} & 0 \\ 0 & \mathcal{K} \end{pmatrix} \tag{4.24}$$

which is blocked subsequently. The same is done with the resulting renormalized 6×6 matrix in order to go to larger renormalization numbers with $n_{renorm} = 2^i$ ($i > 1$ is an integer). Note that decimation in the simple sampling routine corresponds to blocking base-pair step parameters. The

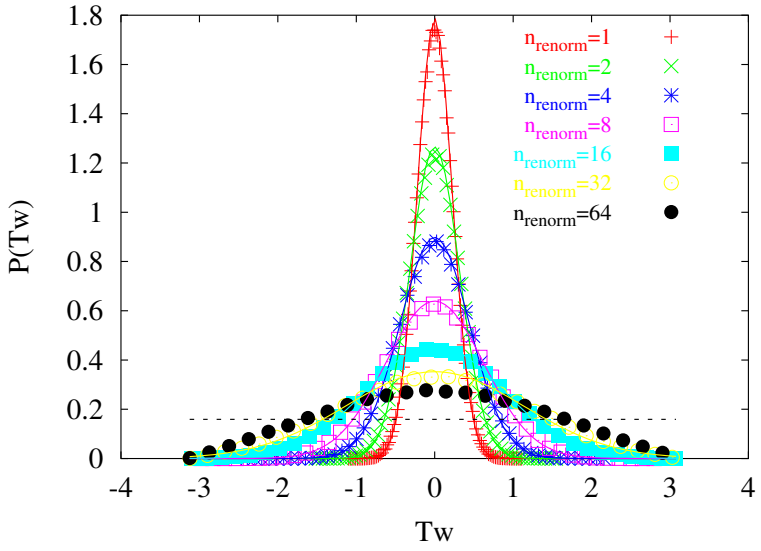


Figure 4.5: Twist distribution function for different renormalization steps measured by simple sampling. The solid lines correspond to the corresponding Gaussians functions calculated with the measured mean squared values. For large n_{renorm} one recovers deviations from the Gaussian shape. The distribution function tends towards the Gaussian limit (uniform distribution, black dotted line) for large renormalization numbers.

results are illustrated in Fig. 4.3. Indeed the rotational part of the elastic matrix can be renormalized using the derived scaling relation for the blocking procedure since the rotational parameters are *additive* for small rotations. Of course, for even larger n_{renorm} the additivity of the angular parameters will break down, since they are not small anymore. For $n_{renorm} \to \infty$ one recovers the Gaussian chain limit where the plates can freely rotate. In Fig. 4.5 the twist distribution for various degrees of renormalization is shown. One observes that $P(\mathrm{Tw})$ converges to a uniform distribution for $n_{renorm} \to \infty$ which corresponds to the freely rotating chain limit. For the translational part one recovers quite substantial deviations of the blocking procedure from the simple sampling results, especially for large renormalization numbers. It underlines the *non-additivity* of the base-pair step parameters displayed in Eq. (4.22) which is not taken into account in the renormalization calculation of Eq. (4.12).

The situation gets even worse if one introduces a geometrical coupling between the rise and twist degree of freedom. For the sampling procedure we used coupling constants k_x with $k_x = (\langle x^2 \rangle - \langle x \rangle^2)^{-1}$ where the mean and mean squared values of the step parameters x were measured for the final parameterization set of Eq. (3.52). The twist-rise coupling is introduced via Eq. (3.14). We use $d = 16$Å for the plate diameter, $l = 6$Å for the length of the respective rods connecting the edges of the plates and we impose a spontaneous rise with Ri $= 3.3$Å. The renormalized twist coupling constant does not coincide with the one measured by simple sampling due to the introduced coupling. The renormalized elastic matrix measured by simple sampling is illustrated in Fig. 4.6.

Thus the presented renormalization procedure can only be used to calculate the renormalized angular coupling constants (in case of rather stiff filaments) for small n_{renorm} in a system where translational and rotational motion is not coupled. Otherwise one has to use the described sim-

$$\mathcal{K} = \begin{pmatrix} 18876 & 0 & 0 & 216838 & 0 & 0 \\ 0 & 9 & 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 216838 & -2 & 0 & 2497586 & 0 & -1 \\ 0 & 0 & 0 & 0 & 322 & 0 \\ 0 & 0 & 0 & -1 & 0 & 116 \end{pmatrix}$$

$$\Downarrow n_{renorm} = 2$$

$$\begin{pmatrix} 104 & 0 & 0 & 1127 & 0 & 0 \\ 0 & 3 & 0 & 0 & -2 & 8 \\ 0 & 0 & 2 & 0 & -2 & -5 \\ 1127 & 0 & 0 & 15316 & 1 & -1 \\ 0 & -2 & -2 & 0 & 131 & -4 \\ 0 & 8 & -5 & -1 & -4 & 84 \end{pmatrix}$$

$$\Downarrow n_{renorm} = 4$$

$$\begin{pmatrix} 12 & 0 & 0 & 89 & 0 & 0 \\ 0 & 1 & 0 & 0 & -3 & 2 \\ 0 & 0 & 1 & 0 & -2 & -4 \\ 89 & 0 & 0 & 1910 & 0 & 0 \\ 0 & -3 & -2 & 0 & 41 & 1 \\ 0 & 2 & -4 & 0 & 1 & 50 \end{pmatrix}$$

$$\Downarrow n_{renorm} = 8$$

$$\begin{pmatrix} 1 & 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & -3 & 3 \\ 10 & 0 & 0 & 675 & 0 & 0 \\ 0 & 2 & -3 & 0 & 99 & -2 \\ 0 & 4 & 3 & 0 & -2 & 126 \end{pmatrix}$$

Figure 4.6: Renormalization results by simple sampling. The renormalized elastic matrices for $n_{renorm} = 2,\ 4,\ 8$ are shown. $\mathcal{K}$ corresponds to the original system.

ple sampling method to create long sequences which can subsequently be used to calculate the renormalized elastic matrix for an elastic rod model. The necessary coupling constants which are needed for the sampling procedure can be extracted out of atomistic molecular dynamics simulations. Depending on the underlying sequence and on solvent conditions the effective constants and the spontaneous values will vary from B-DNA to A- and C-DNA values. The parameterized rod model can subsequently be used to investigate structural and elastic properties of linear and circular DNA [KML97, MRKL98, FMM00], cyclization rates [MMK96], and effects on supercoiling phenomena [CL96] on much larger length scales. Instead of including empirically some coupling terms in the worm-like chain model this is a systematic way of calculating the present coupling terms from the underlying base-pair sequence. The introduction of geometrical couplings, e.g. between shift and tilt, can be quite demanding. One has to find analytical relationships as the ones discussed in section 3.2.2. In order to circumvent this problem the following strategy can be embarked for the

simple sampling procedure:

- calculation of the elastic matrix by inverting the correlation matrix
- diagonalization of the elastic matrix
- independent random numbers can be drawn for the eigen-modes
- the conformation is obtained by transforming back.

If long-ranged interactions are involved certainly the simple sampling method cannot be used anymore. Rather one has to implement Monte-Carlo renormalization techniques to measure the renormalized coupling constants along the chain.

It is even more challenging to keep the non-linear aspects of the DNA such as the overstretching transition during renormalization. In principle one could renormalize separately the elastic matrix of the B-DNA and the S-DNA provided that it is possible to write down an effective Hamiltonian such that the critical force where the structural transition occurs is an invariant under the renormalization group operator. An Ising-like model as it is introduced by Storm and Nelson [SN02b, SN02a] could be a good starting point.

Chapter 5

A computer simulation study of the influence of nucleosomal interactions on chromatin structure and elasticity

We introduce a simple model for the chromatin fiber to study its structural and elastic properties. The underlying geometry (of the modeled fiber) is based on the two-angle model [SGB01], which belongs to the class of crossed-linker models. The chromatosomes are treated as disk-like objects. Excluded volume and short range nucleosomal attraction is taken into account by a variant of the Gay-Berne potential [GB81]. We investigate the influence of the nucleosomal interactions on elastic properties of the fiber, such as the bending stiffness and the stretching modulus, and on structural properties, such as the mass density. We find that the bending rigidity and the stretching modulus of the fiber increase with larger nucleosomal disk sizes. Moreover, we apply an external stretching force to the system and measure the resulting force-extension curves. For a reasonable parameterization of the chain for physiological conditions and sufficiently high attraction we find a force-plateau in agreement with experiments [CB00], where the plateau corresponds to a structural transition of the fiber. In a first step the overall fiber geometry – a loop conformation – is stretched out, followed by a decondensation of the fiber as the second step.

5.1 Introduction

It is known that inside the cells of all procaryotic organisms DNA is wrapped around so-called histone octamers. This complex is called nucleosome. The nucleosomes linked together by DNA segments organize further into the chromatin fiber (see section 1). Electrostatic interactions between the nucleosomes (probably mediated by lysin rich histone tails [LMR$^+$97]) give rise to higher order structures of the chromatin fiber [HZ96, BHG$^+$98]. For low salt concentrations a 'beads-on-a-string' structure is observed, sometimes referred to as the 10-nm fiber. For higher salt concentrations (> 40 mM) the fiber appears to thicken folding into a condensed structure with a diameter of roughly 30 nm. Furthermore, linker histones strongly influence the higher order structure of chromatin. They glue the DNA strands entering and exiting the histone core particle together by forming a stem structure [BHG$^+$98]. In the absence of linker histones the entry-exit angle of the in- and outcoming DNA is larger, leading to more open structures. While the structure of the nucleosome is known with atomistic resolution [LMR$^+$97] there is still considerable controversy about the structure of the 30-

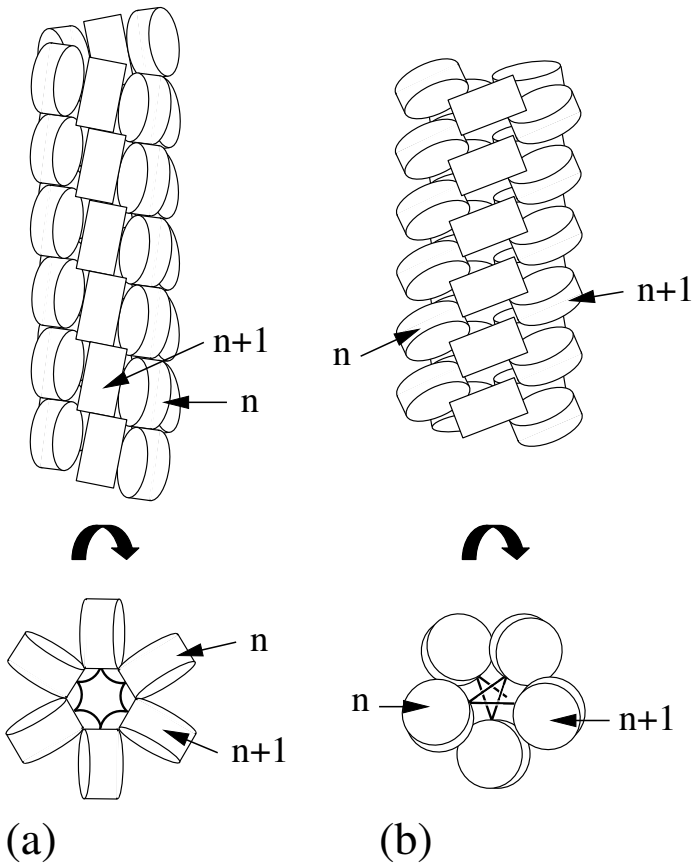


Figure 5.1: Top and side view of (a) the solenoidal and (b) the crossed-linker model. Taken from [Sch03]. Note that the chromatin structure is not a regular helix. Thermal fluctuations [KBO00, WL02] as well as varying linker lengths lead to irregularities of the overall structure [WGHW93].

nm chromatin fiber. Essentially there are two classes of models: (i) the solenoid models [FK76] and (ii) the crossed-linker models [WGHW93, SGB01]. In the solenoid models one assumes that the successive nucleosomes form a helix where the normal vector of the nucleosomes is roughly perpendicular to the solenoidal axis. The entry-exit side faces inward towards the solenoidal axis and the linker DNA must bend in order to connect neighboring nucleosomes. In the crossed-linker model the linker DNA is straight and connects nucleosomes on opposite sides of the fiber (compare Fig. 5.1).

The higher order structure of chromatin for various salt concentration has been studied experimentally for oligonucleosomes using electron cryo-microscopy [BHDW98, BHG$^+$98], neutron scattering and scanning transmission electron microscopy [GR87, HZ95]. Structural parameters such as the mass density (number of nucleosomes per 11nm) and the linker entry-exit angle are measured to characterize the state of compaction. All these studies show an open zig-zag like fiber structure for low salt concentration. Furthermore, studies on di- and trinucleosomes indicate that increasing the salt concentration towards physiological conditions does not induce a bending of the linker DNA, in contradiction to the solenoidal model [GR87, HZ96, HZ95, BHDW98, BHG$^+$98]. Rather

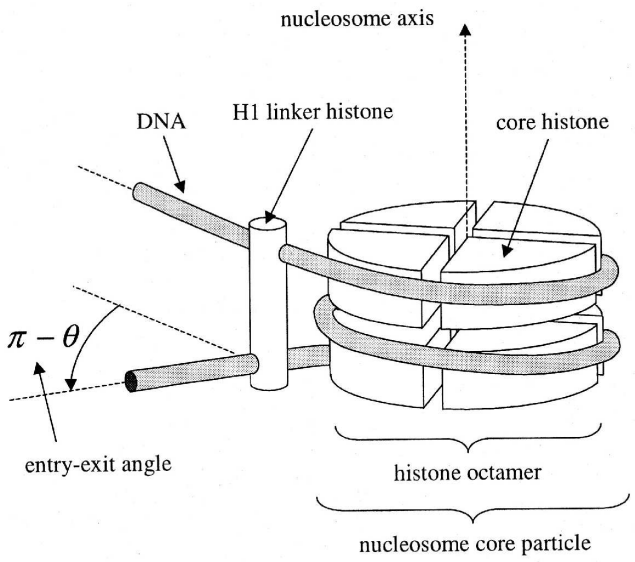


Figure 5.2: Schematic representation of nucleosome. The DNA is wound in $1\frac{3}{4}$ turns around the histone octamer. Taken from [SGB01].

nucleosome-nucleosome interactions, the existence of the linker histone and the salt dependence of the entry-exit angle [WGHW93, HZ96, BHDW98, BHG$^+$98, Sch02] seem to be the relevant ingredients to account for the different folding states. Another parameter that controls the compaction of the fiber is the helical twist of the linker DNA between two nucleosomes. Fluctuations of 1bp linker length would lead to a change in twist of $\pm 2\pi/10$. However, it has been pointed out by Widom and coworkers [Wid92, YLW93] that such changes in twist lead to very large free energy penalties. Therefore this possibility is very unlikely and can be excluded. Unfortunately none of these techniques permit to resolve the internal structure in a compacted fiber, such as the conformation of the linker DNA. All these experimental observations still cannot exclude the possibility of solenoidal structures at high ionic strengths.

The development of micromanipulation techniques on single molecules permits to gain new insight into the structural and elastic properties of the chromatin fiber. Cui and Bustamante [CB00] measured force-extension curves of single chromatin fibers under different ionic conditions. For low salt concentration the fiber behaves like an extensible WLC, whereas for high salt concentrations one finds a force plateau at around 5 pN where the fiber starts to get longer with little increase in tension. This transition is identified with the termination of nucleosomal-nucleosomal attraction. The comparison of the predictions made by computer simulations [KBO00, BS01, WL02] and of analytical approaches [SGB01, BHLV01, BHLV02] with the stretching experiments seems to support the crossed-linker models.

In the following we will discuss possible structures that one obtains for a regular two-angle model [SGB01], which belongs to the class of crossed-linker models. Accounting for the geometry, the elasticity of the linker DNA, and the twist-stretch coupling of the fiber, one can calculate within linear response the effective bending and torsional persistence length, stretching modulus and twist stretch coupling, and the resulting stress-strain curves. In these analytical considerations excluded volume effects of the nucleosomes as well as the attractive interaction between nucleosomes are not

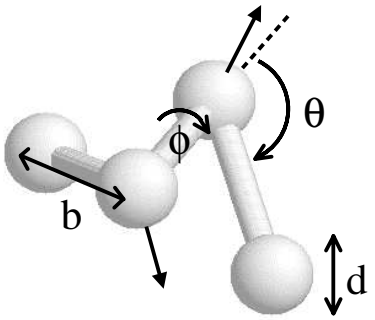


Figure 5.3: Illustration of the two-angle model. $\pi - \theta$ corresponds to the entry-exit angle, ϕ is the rotational angle, b is the linker length and d the nucleosomal diameter. The arrows denote the axis determined by the superhelix of the nucleosomal DNA. Taken from [SGB01].

taken into account. This is reasonable for low salt concentration where one finds open zig-zag-like structures with large internucleosomal distances. However, under physiological conditions the fiber is closely packed and excluded volume as well as nucleosomal attraction is important. These effects can be addressed with the help of computer simulations.

In section 5.2 we recapitulate earlier results of the two-angle model concerning the fiber geometry and concerning the elastic properties as a function of the underlying geometry and linker DNA elasticity. The chromatin model is introduced in section 5.3. We use MC simulations to investigate the structural and elastic properties of the modeled fiber. Details on the MC moves and the correlation time are summarized in section 5.5. In section 5.4 we study the phase diagram which classifies allowed and forbidden structures in the two-angle plane as a function of nucleosomal size and linker DNA length. The influence of nucleosome-nucleosome interactions on the elastic properties of the fiber is discussed in section 5.6. In section 5.6.3 the response of the fiber to stretching is investigated.

5.2 The two-angle fiber

The two-angle model assumes that the geometric structure of the 30-nm fiber can be derived from the single-nucleosome structure [SGB01]. Consider for example the wrapping of the DNA around the histone core. It is found that DNA does not wind an integral number of turns around the histone core. Rather, only about $1\frac{3}{4}$ turns are completed which implies a non-zero entry-exit angle of the in- and outcoming DNA (see Fig 5.2). In the presence of linker histones which glue the two strands together one obtains stem-like structures [BHG$^+$98]. Even though the exact value of the entry-exit angle $\pi - \theta$ depends on salt concentration and on the presence or absence of linker histones, one can nevertheless assume that θ is determined at the single-nucleosome level. In addition there is a rotational angle ϕ involved (see Fig. 5.3) which determines the orientation of the nucleosomes along the string. The rotational angle is a periodic function on the length of the linker DNA b connecting two nucleosomes. Thus ϕ is given by multiples of $2\pi/10$, the twist angle associated with individual base-pair steps.

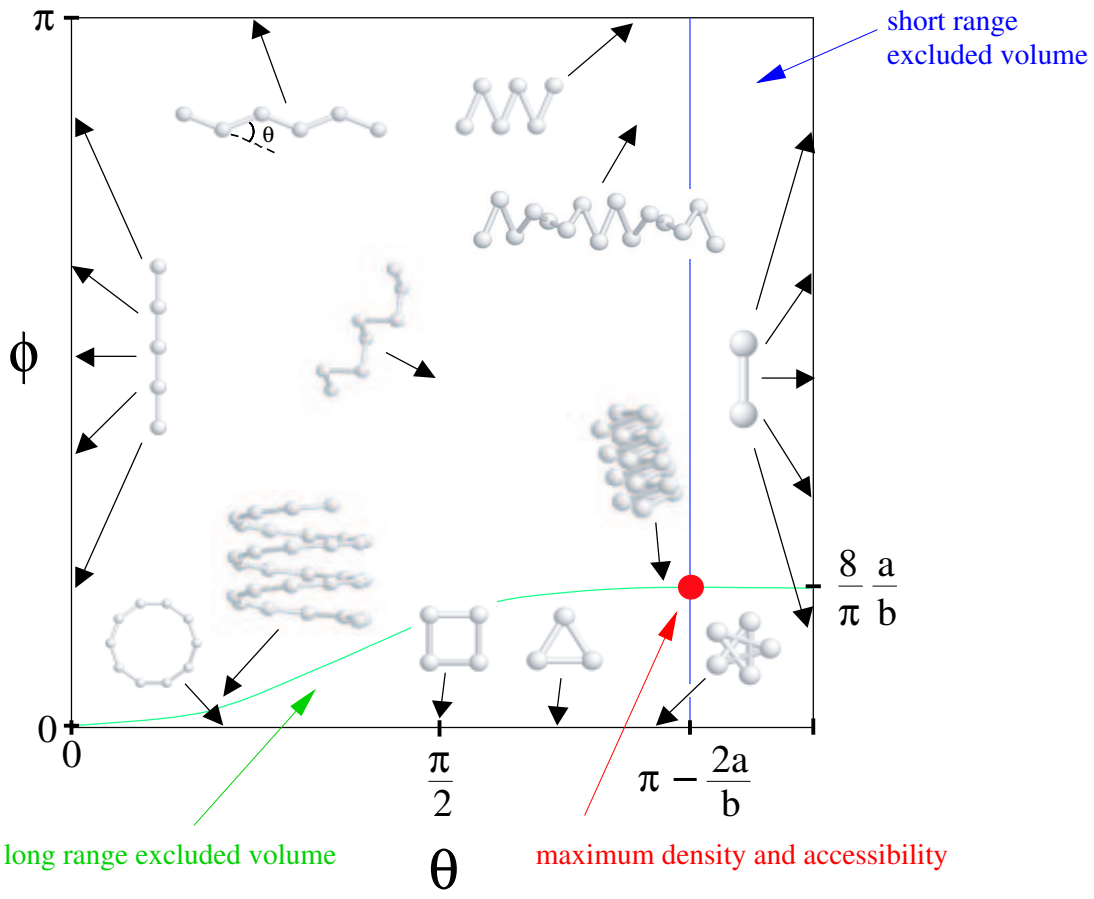


Figure 5.4: Overview of resulting conformations in the θ-ϕ-plane. The drawn green and blue line show the boundaries to forbidden structures due to long range excluded volume effects (small ϕ values), and due to short range excluded volume effects (small θ values) respectively. Taken from [SGB01].

5.2.1 Geometrical considerations

The two-angle model characterizes the fiber structure via three parameters: the linker length b, the entry-exit angle $\pi - \theta$ and the twist angle ϕ. It assumes that the linker DNA is straight, that the chromatin fiber is regular, and that the nucleosomes are point-like situated at the joints of the linker DNA. Exclude volume effects are not considered.

In Fig. 5.4 different possible structures are depicted. If either one of the angles is 0 or π the resulting structure is planar. In case of $\phi = 0$ one obtains planar structures varying from $\frac{2\pi}{\theta}$-polygons for $\theta = \frac{2\pi}{n}$ with n being an integer to star-like structures. In the special case of $\theta = \pi/2$ one finds a square. For $\theta = \frac{\pi(n-1)}{n}$ closed star-like structures are encountered. In particular $n = 5$ corresponds to the regular pentagram. In case of $\theta = 0$ a straight line is recovered. $\phi = \pi$ yields planar zig-zag structures and $\theta = \pi$ produces 'dumbbell' conformations [SGB01].

Non-zero values of θ and ϕ lead to three-dimensional fibers. They can be further subdivided into three classes. Structures with small θ and ϕ values resemble solenoids. Structures where ϕ is

still small but θ takes large values form fibers with crossed linkers. Finally, fibers with ϕ close to π lead to twisted zig-zag structures.

The overall fiber geometry can be described as a function of the underlying three local geometric parameters $\{b, \theta, \phi\}$. One can construct a spiral characterized by a radius R and a pitch angle ψ such that the nucleosomes (but not necessarily the linker chain) are situated along this spiral. The distance b of successive nucleosomes along the spiral is fixed. It is possible to derive a relationship of R and ψ and the local parameters $\{b, \theta, \phi\}$ [Sch03]:

$$R = \frac{b \sin\left(\frac{\theta}{2}\right)}{2 - 2\cos\left(\frac{\theta}{2}\right)^2 \cos\left(\frac{\phi}{2}\right)^2} \tag{5.1}$$

$$\cot(\psi) = \frac{\tan\left(\frac{\theta}{2}\right) \arccos\left(2\cos\left(\frac{\theta}{2}\right)^2 \cos\left(\frac{\phi}{2}\right)^2 - 1\right)}{2\sin\left(\frac{\phi}{2}\right)\sqrt{1 - \cos\left(\frac{\theta}{2}\right)^2 \cos\left(\frac{\phi}{2}\right)^2}} \tag{5.2}$$

Furthermore one can calculate the distance s_0 of two successive nucleosomes along the fiber axis

$$s_0 = \frac{b \sin\left(\frac{\phi}{2}\right)}{\sqrt{\sec\left(\frac{\theta}{2}\right)^2 - \cos\left(\frac{\phi}{2}\right)^2}}. \tag{5.3}$$

Thus the mass density λ (the number of nucleosomes per 11 nm) is given by

$$\lambda = \frac{11\text{nm}}{s_0}. \tag{5.4}$$

The contour length L of the fiber consisting of N nucleosomes is determined by $L = N s_0$.

Up to now we neglected excluded volume interactions. However, certain areas of the θ-ϕ phase diagram are forbidden due to overlapping nucleosomes. In the following considerations we assume that the nucleosomes are spherical with a radius a and that their centers are located at the joints of the linker DNA. One can distinguish between two types of interactions: (i) short range excluded volume interactions between monomer i and $i \pm 2$. This requires that the entry-exit angle has to be sufficiently large with

$$\begin{aligned} b \sin\left(\frac{\pi - \theta}{2}\right) &> a \\ \theta &< 2\arccos\left(\frac{a}{b}\right) \end{aligned} \tag{5.5}$$

(compare blue line in Fig. 5.4). (ii) Long range excluded volume interactions in case of very small ϕ values. This becomes clear if one considers planar structures with $\phi = 0$ that run into themselves. Circular structures with $\theta = \frac{2\pi}{n}$ will have an overlap after n nucleosomal steps if

$$n s_0 < 2a, \tag{5.6}$$

where $n s_0 = \frac{2\pi}{\psi} s_0$ corresponds to one helical pitch. Eq. (5.6) can be solved for ϕ, yielding a condition that must be fulfilled to avoid a steric clash after n steps [Sch03] (compare green line

in Fig. 5.4). The examination of a circular structure provides an upper bound estimation for the angle ϕ. Structures that deviate sightly from the n-polygon case lead to a fine structure of possible ϕ values. We will discuss this in more detail in section 5.4.

It is interesting to consider the implications of this geometrical model for the compaction and accessibility of DNA in chromatin. Due to the very large ratio of DNA length to cell nucleus diameter it is clear that inactive regions with DNA sequences that do not carry genetic information should be packed as densely as possible. On the other hand active regions must be also accessible to protein complexes that bind to the DNA to read a specific sequence. This leads to the question where in the phase diagram the point is situated that maximizes both the compaction and the accessibility.

Maximum compaction is reached if the bulk density (the number of nucleosomes per unit volume of the master solenoid)

$$\rho = \frac{1}{\pi R^2 s_0} \tag{5.7}$$

is maximized. Clearly this will be the case for structures with internal linker DNA. In particular, the highest density that satisfies the condition of excluded volume is obtained for the largest possible value of θ and the smallest possible value of ϕ. That corresponds to a structure where the nucleosomes are in closest contact with its $i \pm 2$ neighbors and the neighbors after/before one helical turn.

To achieve maximum accessibility for highly compacted structures with a given entry-exit angle $\pi - \theta$ Schiessel *et al* [SGB01] searched for structures which give the maximum reduction of the line density ρ_L with

$$\rho_L = \frac{1}{s_0} \tag{5.8}$$

for a small change of the angle θ. This is identical with looking for the maximum of $\frac{d\rho_L}{d\theta}$ which was termed accessibility. Interestingly this analysis results in the same pair of angles as it is obtained by the calculation of the maximum bulk density (compare red point in Fig. 5.4).

5.2.2 Elastic properties

The stretching experiments of Cui and Bustamante indicate that the chromatin fiber is a highly flexible chain with a large amount of twistable and bendable linker DNA. Of course, the elastic properties also depend on the geometry of the fiber. A zig-zag fiber with a small θ value will have a larger stretching modulus than a compact star-like structure.

In order to calculate the mechanical parameters of the chromatin fiber Ben-Haim *et al* [BHLV01, BHLV02] describe the two-angle fiber as an extensible WLC with additional terms accounting for the twist rigidity of the fiber and the twist-stretch coupling. Within linear response theory one can calculate the elastic free energy of the fiber in terms of the local relative extension $u(s)$ of the fiber, its local twist $\Omega(s)$, and its local curvature $\rho(s)$

$$\frac{\mathcal{F}_{fiber}}{k_B T} = \int ds \left(\frac{A}{2}\rho(s)^2 + \frac{C}{2}\Omega(s)^2 + \frac{\gamma}{2}u(s)^2 + D\Omega(s)u(s) \right), \tag{5.9}$$

where A denotes the bending stiffness of the fiber, C corresponds to the twist rigidity, γ denotes the stretching modulus, and D refers to the twist-stretch coupling. This description has recently been used to discuss the force-extension relations of single DNA molecules [Mar97, Mar98]. Using the fact that the elastic energy stored in the solenoid is nothing but the sum of elastic energies stored

in its linkers, i.e. $\mathcal{F}_{fiber} = \mathcal{F}_{linker}$ it is possible to derive a relationship between the fiber elastic constants, the DNA elastic constants and the parameters describing the relaxed ($T = 0$) geometry of the fiber. Treating the DNA as an inextensible WLC with bending and torsional rigidity Ben-Haim *et al* [BHLV01, BHLV02] calculated the elastic energies stored in the linker DNA corresponding to the twisting and bending degrees of freedom and compared this result to the energy expression of the fiber. This results in the following equations [Sch03]:

$$\gamma = \frac{s_0}{k_B T\, b_{linker}} \frac{C + \Delta S \cos(z)^2}{R^2 \cos\left(\frac{\eta}{2}\right)^2} f(\eta, z) \tag{5.10}$$

$$A = \frac{l_p s_0}{b_{linker}} \frac{2 l_{Tw}}{l_p + l_{Tw} - \Delta S \cos(z)^2} \tag{5.11}$$

$$C = \frac{s_0}{b_{linker}} \left(\frac{l_{Tw}}{3} \tan\left(\frac{\eta}{2}\right)^2 + l_p - \Delta S \cos(z)^2 \right) f(\eta, z) \tag{5.12}$$

$$D = -\frac{s_0}{k_B T\, b_{linker}} \frac{\Delta S \cos(z) \sin(z)}{R \cos\left(\frac{\eta}{2}\right) f(\eta, z)} \tag{5.13}$$

where

$$f(\eta, z) = \frac{3 l_p}{3 l_p + \tan\left(\frac{\eta}{2}\right)^2 (l_{Tw} + \Delta S \cos(z)^2)} \tag{5.14}$$

and the fiber radius R and s_0 are given by Eq. (5.1) and (5.3). z denotes the angle between the fiber axis and the linker:

$$z = \arccos\left(\frac{s_0}{b_{linker}}\right). \tag{5.15}$$

η refers to the angle between neighboring nucleosomes, as seen when viewed down the fiber axis, and can thus be identified with the twist angle of the solenoid given by

$$\eta = \cot(\psi)\frac{s_0}{R} = \arccos\left(2\cos\left(\frac{\theta}{2}\right)^2 \cos\left(\frac{\phi}{2}\right)^2 - 1\right) \tag{5.16}$$

where the result of ψ of Eq. (5.2) has been inserted. $\frac{\eta}{s_0}$ corresponds to the twist rate of the unperturbed fiber. l_p and l_{Tw} describe the bending and twist rigidity of the linker DNA and ΔS is given by their difference $l_p - l_{Tw}$. b_{linker} denotes the linker length.

Eqs. (5.10) and (5.11) can be compared to results obtained by fitting the extensible WLC model to stress-strain curves of measurements in low salt concentration [CB00]. One obtains a value of about 30nm for the fiber bending persistence length A and 1.25nm for the stretching modulus γ. It should be stressed that these results are based on the release part of the force-extension curves at low ionic strength, i.e. open chromatin fibers, where one does not find nucleosome-nucleosome interactions such that the underlying theory is valid. Using a value of $\theta = 50^\circ$ one obtains with Eq. (5.11) and (5.10) similar result for linker lengths $n = 31 \pm k\,p$ base-pairs where k is an integer and $p = 10$ corresponds to the DNA pitch [BHLV02].

As already mentioned, in high salt concentration, where the fiber adopts very compact structures there are also nucleosome-nucleosome interactions involved which are neglected in the present model. These attractive interaction can be mediated by the lysin-rich core histone tails and cause

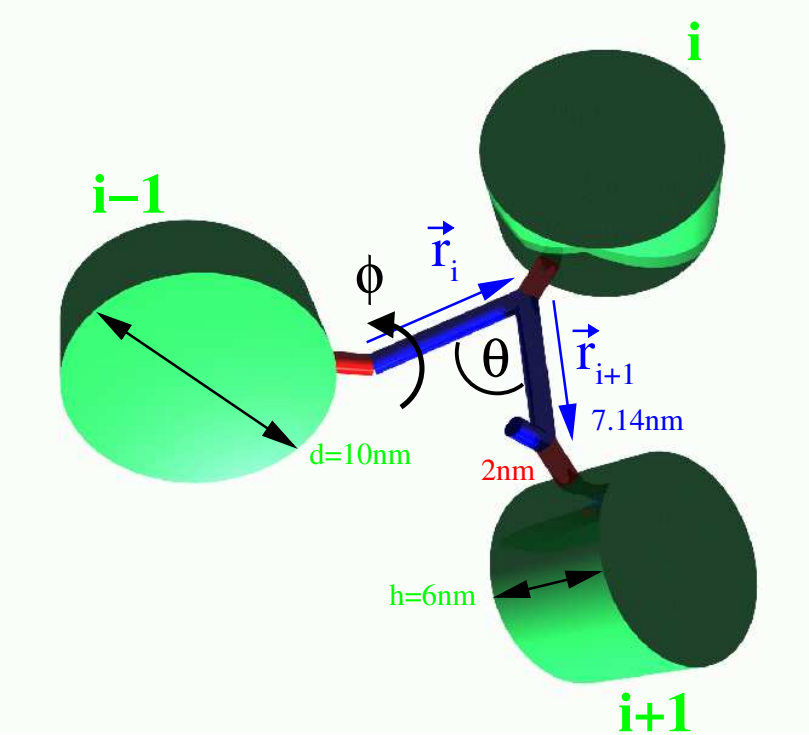


Figure 5.5: Illustration of the modeled fiber. The red cylinders correspond to the stem, the blue ones to the linker DNA.

the fiber to contract. Liquid crystal studies on core particles have indeed shown fiber-like columnar structures, presumably due to attractive nucleosome-nucleosome interactions [LL97, LL00b].

Wedemann *et al* [WL02] performed MC simulations of the 30-nm fiber where they modeled the nucleosomes as disks with a radius of 5nm and a height of 6nm. Excluded volume and attractive nucleosome-nucleosome interaction were included via a Gay-Berne potential. The underlying geometry and elasticity of the fiber was closely related to the two-angle model. Electrostatic and hydrodynamic interactions were also taken into account. They found for $\theta = 154°$, $\phi = 110°$, $b_{linker} = 11$bp a persistence length of $A = 260$nm which is about 20 times larger than the corresponding value of 13nm calculated with Eq. (5.11).

We want to study in more detail the nucleosomal effects due to excluded volume and due to nucleosome-nucleosome attraction on the elastic properties of the fiber and on the behavior under stretching. For this we use a model which is based on the previously discussed two-angle-fiber with additional nucleosomal interactions. We model the nucleosomal interactions with a variant of the Gay-Berne potential [EE03] similarly to Wedemann *et al* [WL02]. The geometry of the fiber is slightly modified in order to take the stem structure of the nucleosomes in the presence of linker histones into account.

5.3 The chromatin model

Concerning the geometry of the modeled fiber we assume that the in- and outcoming DNA is glued together by a linker histone forming a stem [BHG$^+$98]. This will be of importance when we calculate the effective length of the linker DNA. The DNA length per nucleosome varies from tissue to tissue. In case of chicken erythrocyte the DNA length amounts to about 212 base-pairs. We will use this value in the following considerations. Later in this chapter the effect of linker length variations on the fiber geometry and elasticity will be discussed in more detail. About 146 base-pairs wrap in $1\frac{3}{4}$ turns around the histone core particle. Thus 2 turns correspond to 177 base-pairs. The stem ends at about 7nm from the center of the core particle. Assuming that the linker histones are situated directly at the core particles, 2 times 7 base-pairs are involved in forming the stem. This means for chicken erythrocyte chromatin fibers that there are roughly 21 base-pairs of linker DNA left. Fig. 5.5 shows a sketch of the chromatin model. The core particle with its DNA is treated as ellipsoidal disks with a diameter of 10nm and a height of 6nm corresponding to the experimental values [LL97, LMR$^+$97]. The stem (red cylinders) has a length of 2nm and the linker DNA (blue cylinders) is 7.14nm long.

The linker DNA is discretized in four segments in order to allow for bending and torsional deformations. Each segment i represents about 5bp and is labeled by a set of basis vectors $\{\mathbf{t}_i, \mathbf{n}_i, \mathbf{b}_i\}$ where $\mathbf{t}_i$ denotes the tangent vector, $\mathbf{n}_i$ the normal and $\mathbf{b}_i$ the binormal vector. The elastic energy of the linkers is thus described by

$$\frac{\mathcal{H}_{el}}{k_B T} = \frac{l_p}{2b} \sum_{i=1}^{4(N-1)+2} (\beta_i - \beta_{sp})^2 + \frac{l_{Tw}}{2b} \sum_{i=1}^{4(N-1)+2} (\tau_i - \tau_{sp})^2 \tag{5.17}$$

where l_p and l_{Tw} are the bending and twist rigidity respectively, N is the number of nucleosomes and b denotes the segment length. $\beta_i = \arccos(\mathbf{t}_i \cdot \mathbf{t}_{i+1})$ refers to the bending angle between two neighboring segments and τ_i denotes the twist angle given by Eq. (B.11). The spontaneous bending angle $\beta_{sp} = \theta$ takes the kink of in- and outcoming linker DNA into account and is only non-zero for those segment pairs that are connected to a stem. $\tau_{sp} = \phi/3$ enforces the right-handed helicity of the DNA which in turn gives rise for the fiber twist angle ϕ. Note that the linker length and the fiber twist angle are actually coupled. This is so because the DNA has to face with its minor groove the binding sites on the histone spool. Thus an increase of 1bp causes an increase of twist of $2\pi/10$. A calculated linker length of 21bp corresponds to $\phi = 2\pi/10$. However, for simplicity we keep ϕ as a free parameter which controls the fiber geometry while the linker length variations amount at most to $\pm$1nm and can thus be neglected.

To account for the orientation of the chromatosomes the disks are also labeled by a set of three orthonormal basis vectors $\{\mathbf{T}_i, \mathbf{N}_i, \mathbf{B}_i\}$. $\mathbf{N}_i$ is the normal vector perpendicular to the disk plane given by

$$\mathbf{N}_i = \frac{\vec{r}_i \times \vec{r}_{i+1}}{|\vec{r}_i \times \vec{r}_{i+1}|}. \tag{5.18}$$

The vectors $\{\vec{r}_i\}$ connect the stems of neighboring chromatosomes (see Fig. 5.5). $\mathbf{B}_i$ points from the joint of the linker DNA towards the disk resulting in

$$\mathbf{B}_i = \frac{\vec{r}_i - \vec{r}_{i+1}}{|\vec{r}_i - \vec{r}_{i+1}|}. \tag{5.19}$$

$\mathbf{T}_i$ is obtained by orthonormality of the basis set:

$$\mathbf{T}_i = \mathbf{N}_i \times \mathbf{B}_i. \tag{5.20}$$

As already mentioned we use a variant of the Gay-Berne potential [EE03] (compare section 3.3.1) to model the excluded volume interactions and the attractive interactions due to bridging of lysin-rich tails between the chromatosomes. We choose $\gamma = 1.0$ of Eq. (3.34) and the effective diameter of the nucleosomal disks is chosen to be $\sigma = h$ where h is the height of the chromatosome. The structure matrix is given by [LMR$^+$97]

$$\mathcal{S} = \begin{pmatrix} h/2 & 0 & 0 \\ 0 & d/2 & 0 \\ 0 & 0 & d/2 \end{pmatrix} \tag{5.21}$$

where $d = 1.67h$ is the diameter of the chromatosome. Using the experimentally determined spatial dimensions of the nucleosome core particle [LMR$^+$97] this leads to

$$\mathcal{S} = \begin{pmatrix} 3.0 & 0 & 0 \\ 0 & 5.0 & 0 \\ 0 & 0 & 5.0 \end{pmatrix} \text{ nm}. \tag{5.22}$$

The parameterization of the GB potential entails a lateral spacing of ≈ 7.0 nm and a vertical spacing of ≈ 11.0 nm of the disks in good agreement with experimental results of Livolant *et al* [LL97]. It should be noted that the diameter of the chromatosome represented by $\mathcal{S}$ is composed of two times the DNA diameter and the histone core diameter. The GB parameter ϵ_{GB} which determines the energy well depth ϵ will be chosen such that ϵ amounts to about 1-4k_BT close to the experimental value found by stretching a chromatin fiber [CB00]. The effect of ϵ on the stress-strain relation is discussed in more detail in section 5.6.3.

The parameters characterizing the underlying geometry such as the entry-exit angle $\pi - \theta$, the rotational angle ϕ and the linker length b_{linker} will be varied in order to study their influence on the elastic and structural properties of the fiber.

5.4 Investigation of the fine-structure of the two-angle phase diagram

Before we perform simulations to study the elastic properties of the modeled chromatin fiber we first investigate the structural properties within the described geometry of Fig. 5.5. We measured the two-angle diagram and the energy landscape depending on $\{\theta, \phi\}$ for a given linker length b_{linker}. The nucleosomal disk size is fixed to $d = 10$nm and $h = 6$nm. To calculate the phase diagram we start with a configuration given by $\theta = \theta_{initial}$ and $\phi = \pi/2$ for a fixed linker length b_{linker}. We add successively nucleosomes building up the fiber and check if there is an overlap or not. In case of overlap we update the rotational angle by $\phi_{new} = \phi_{old} + \frac{\phi_{old}}{2}$. Otherwise, if the building procedure satisfies the condition $L > \frac{2\pi}{\psi} s_0$, i.e. the actual fiber length L exceeds one helical pitch without producing an overlap, the rotational angle is updated by $\phi_{new} = \phi_{old} - \frac{\phi_{old}}{2}$. This results in a phase boundary between allowed and forbidden structures determined by the excluded volume condition which is shown for various linker lengths in Fig. 5.8. One can observe that with increasing linker length b_{linker} the fine structure of the phase boundary is increased. This can be understood as follows. Imagine a situation where the nucleosomes are closely packed on a solenoid for a certain linker length. Increasing the linker length entails more space between nucleosome i

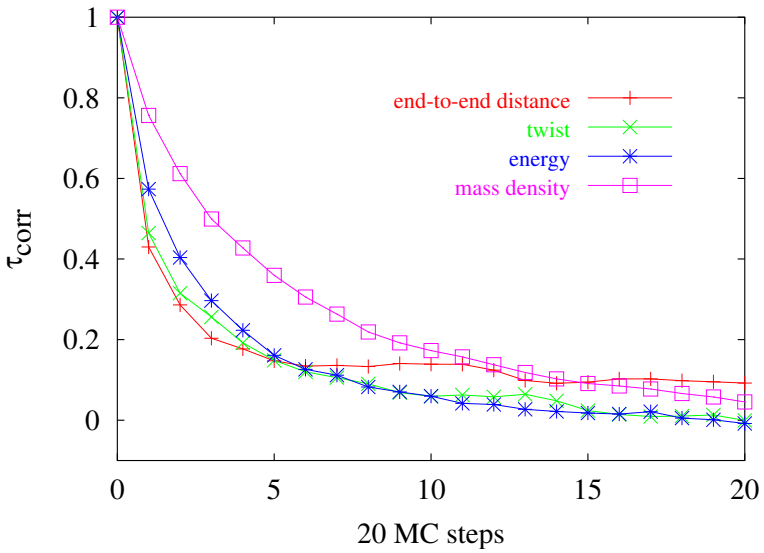


Figure 5.6: Illustration of the measured correlation time of the energy, the mass density the end-to-end distance and twist for a fiber with 100 nucleosomes, and $b = 7.14$nm, $\theta = 145°$, $\phi = 100°$, $\epsilon = 1k_BT$. It is shown that the longest relaxation time amounts to approximately 100 MC steps.

and $i + 1$ such that after one helical turn it is possible (depending on b_{linker}) that a nucleosome fits in. This creates an additional bump into the phase boundary. Fig. 5.8 shows the calculated energy landscape for straight linkers of length $b_{linker} = 7.14$ nm, i.e. there is no bending energy involved. The energetically favored conformations can be found close to the phase boundary. Note that the boundary due to short range excluded volume interactions is smeared out to smaller values of the rotational angle ϕ due to the stem structure of the nucleosomes. Without a stem the boundary would be a straight line following Eq. (5.2.1). The previously introduced method for the calculation of the phase diagram cannot resolve this overhang.

In fact, the rotational angle ϕ and the length of the linker DNA b_{linker} are coupled. An additional linker length of one base-pair corresponds to an increase in ϕ of 36^o. In Fig. 5.7 the phase boundary between allowed and forbidden structures concerning nucleosomal clashes is shown for coupled b_{linker} and ϕ, and for two different nucleosomal disk sizes. Starting with $b_{linker} = 0.1$nm we increase successively the linker length for a given angle of θ and check if the respective structure shows an overlap of nucleosomes or not. Each time a boundary between overlap and no overlap or vice versa is passed we write out the corresponding angle space coordinate. Note that the rotational angle ϕ is updated for every step by $\phi = b_{linker}/0.34\text{nm} \cdot 36^o$. All structures belonging to the area right of the phase boundary in Fig. 5.7 are forbidden.

5.5 Monte-Carlo simulations

We use a Monte-Carlo scheme to simulate the chromatin fiber which relies on three moves: (i) a local move where one chooses randomly one nucleosome which is rotated around an axis determined by two points on the in- and out-coming linker DNA by a small random angle, (ii) a non-local pivot

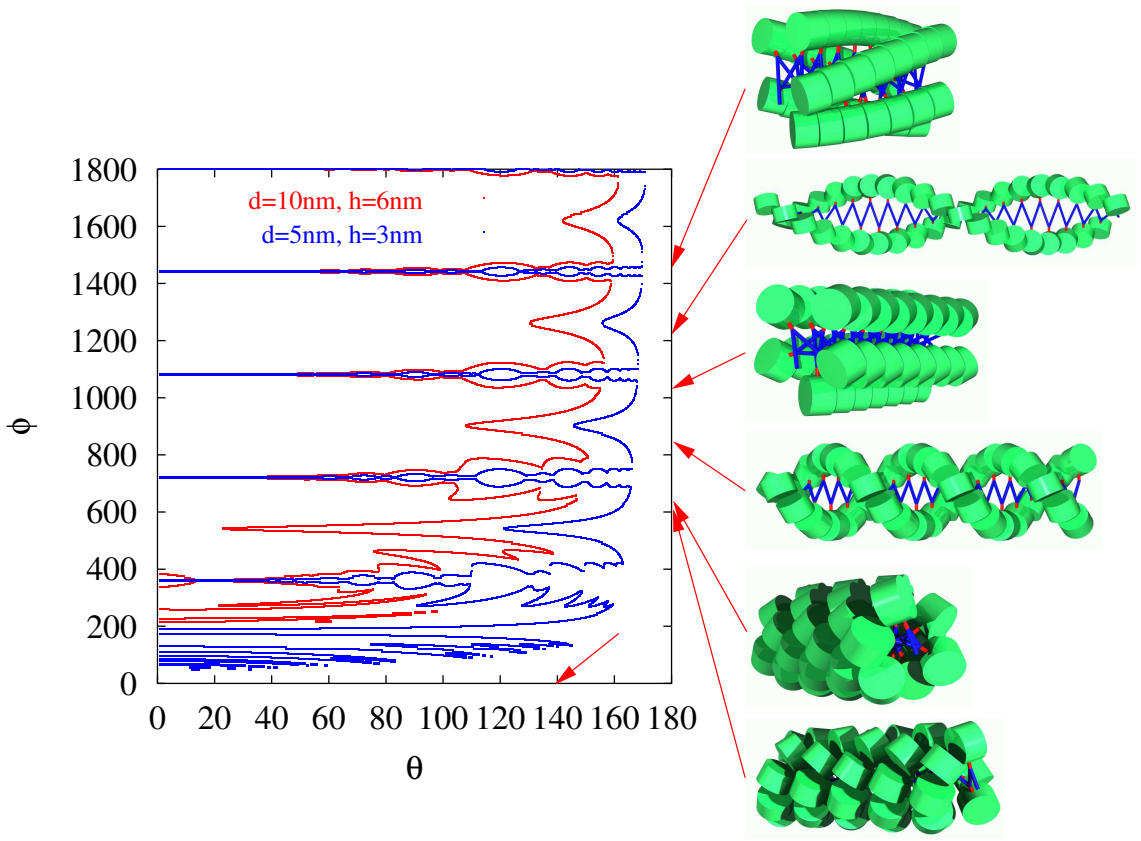


Figure 5.7: Phase diagram for θ and ϕ where the linker length b_{linker} is coupled to the rotational angle ϕ via $b_{linker} = \phi/36^o \cdot 0.34$nm. The snapshots correspond to structures lying directly on the boundaries for $\theta = 140^o$ and the respective ϕ values indicated by the arrows. One observes an increase of the fiber diameter with increasing linker length. The red data points refer to a nucleosomal disk size of $d = 10$nm and $h = 6$nm, whereas the blue data set corresponds to half the disk size.

move where a random segment point is chosen around which the shorter part of the chain is rotated around a random axis by a random angle, (iii) a non-local crankshaft move where two random points along the DNA segments define the axis of rotation around which the inner part of the chain is rotated. The moves are accepted or rejected according to the Metropolis scheme [MRR$^+$53].

Each simulation run consists of 200000 MC sweeps where one trial corresponds to N_{DNA} trials with N_{DNA} being the number of DNA segments. We simulated either fibers with $N = 50$ (simulations with applied stretching force) or $N = 100$ nucleosomes such that the linker DNA segments amounts to $N_{DNA} = 202$ and $N_{DNA} = 404$ respectively. The amplitudes are chosen such that the acceptance rate equals approximately 50%. Every 20 moves we save a configuration. As initial conformation we used the relaxed ($T = 0$) fiber structure. In order to determine the longest relaxation time τ_{corr} of the system we measured the 'time' correlation functions of the energy, the mass density, the end-to-end distance and twist (see Fig. 5.6). It is found that $\tau_{corr} \approx 100$ MC sweeps.

Note that the entry-exit angle θ is not varied during the simulation.

We use the following reduced units: lengths are measured in [nm] and forces are measured in

$[k_BT/\text{nm} = 4\text{pN}]$.

5.6 Structural and elastic properties of the simulated fiber

5.6.1 Influence of the nucleosomal interaction on the structure of the fiber

In section 5.2 we discussed the implications of entry-exit angle, rotational angle and linker length on the structure of the ground state ($T = 0$). In the absence of excluded volume interactions thermal motion leads to rather distorted conformations due to twist and bend fluctuations of the linker DNA (see Fig. 5.9). If the excluded volume of the nucleosomes is taken into account these fluctuations are reduced resulting in more ordered structures. The fiber fluctuates only slightly around the ground state conformation since the accessible space for bending and twist fluctuations is strongly reduced due to the spatial extensions of the nucleosomes. In case of $\theta = 145°$, $\phi = 100°$ a helical structure can be observed. If additionally nucleosomal attraction is included, the fiber can form loops. The occurrence of loops only depends on the ratio of the attractive well of the Gay-Berne potential and the contour length of the fiber. For a given attractive well and sufficiently large chains a loop formation is observed. The occurrence of loops has strong effects on the behavior of the fiber under stretching. This point will be discussed in more detail in section 5.6.3.

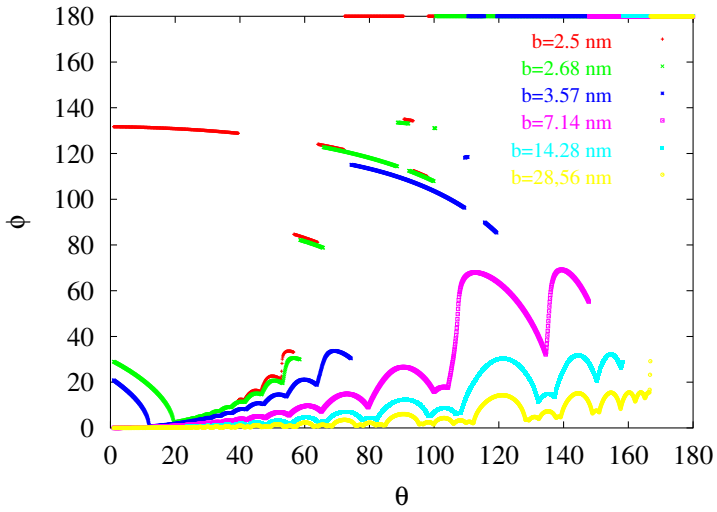
5.6.2 Influence of the nucleosomal interaction on the elastic properties of the fiber

Rough estimate of excluded volume effects

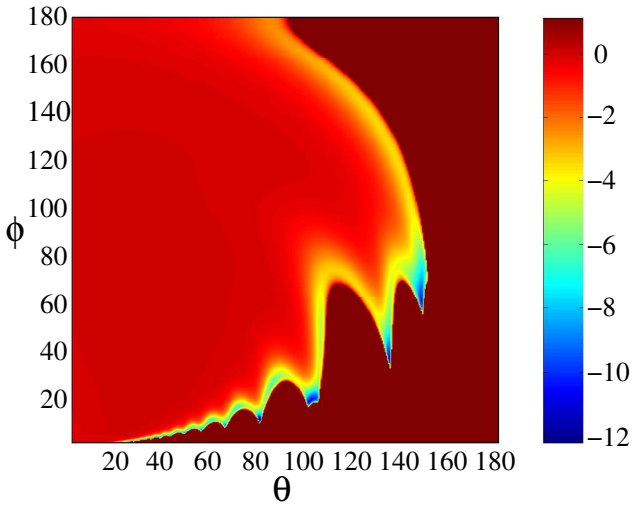
For a given pair of angles one can also estimate the maximum possible size of the nucleosomes for which no nucleosomal overlap is found. This can also be done the other way around using the phase diagram of Fig. 5.8. Here the disk size is kept constant. The most compact structure for a given value of θ is found to be the smallest possible value of ϕ that does not lead to an overlap. This is nearly identical to the energy minimum for the imposed θ value. For such a configuration there is no space left for bending motions such that the bending persistence length is infinitely large. Scaling down the disk size enables bending fluctuations, the bending persistence length becomes smaller. This explains qualitatively the results shown in Fig. 5.14 (see below). A more quantitative analysis can be carried out for the geometry shown in Fig. 5.10, where we consider for the sake of simplicity a bending motion in two dimensions. We assume that the nucleosomes are spheres with radius R_0 and consider a fiber geometry where the nucleosomes are stacked on top of each other with a center-center distance c forming columns which are parallel to the fiber axis. Bending such a fiber by an angle β leads to a curvature of the fiber axis with $\kappa = \frac{\beta}{b_{pitch}}$ and radius $R_\kappa = \frac{1}{\kappa}$. b_{pitch} refers to the helical pitch of the chromatin fiber. The distance c between the centers of two neighboring nucleosomes (within one column) is given by $c \approx (R_\kappa - s - R_0)\beta$ where s denotes the stem length. Thus the available space between two spheres results in $c - 2R_0$. This yields a condition for the maximum possible bending angle β_{max} that does not lead to an overlap of the nucleosomes for given values of R_0, s and b_{pitch}:

$$\beta_{max} = \frac{b_{pitch} - 2R_0}{s + R_0}. \tag{5.23}$$

Note that the diameter of the nucleosomes has to be smaller than or equal to b_{pitch}. The bending persistence length A_{sphere} is related to the mean squared fluctuations of β. Assuming that the

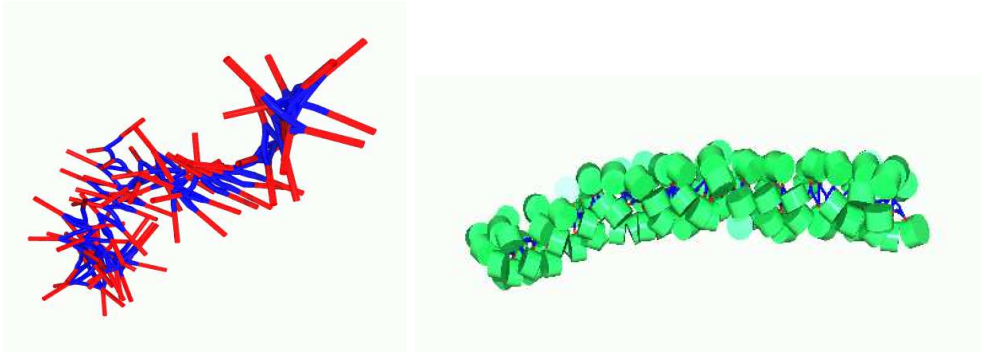


(a) We calculated the boundary between forbidden and allowed structures for different linker lengths corresponding to 7.3, 7.9, 10.5, 21, 42 and 84 bp. The nucleosomal disk size is fixed to $d = 10$nm and $h = 6$nm. All structures below the data points are forbidden due to excluded volume interactions of the nucleosomes.



(b) Overlap is found in the dark red area. The energy minima are located at the blue spots (compare the color code of the energy scale).

Figure 5.8: Illustration of fine structure of the two-angle phase diagram (using hard core repulsion) and contour plot of the energy surface for $b_{linker} = 7.14$ nm (using soft core attraction).



(a) without nucleosomal interactions (b) with nucleosomal interactions

Figure 5.9: Snapshots of fiber with (b) and without (a) nucleosomal interactions for $\theta = 145°$, $\phi = 100°$ and $b_{linker} = 7.14$nm, $d = 10$nm, $h = 6$nm.

bending fluctuations are suppressed by the accessible space between the nucleosomes one can use $A_{sphere} \approx \frac{b_{pitch}}{\beta_{max}^2}$ as a rough estimate of the resulting persistence length. Fig. 5.11 shows the bending persistence length A_{sphere} as a function of the radius of the spheres. It is found that A_{sphere} increases rapidly as $2R_0$ comes close to b_{pitch}. For $2R_0 = b_{pitch}$, i.e. $\theta_{max} = 0$, A_{sphere} diverges. The overall behavior of the estimated bending persistence length A_{sphere} is very similar to the simulation result (compare Fig. 5.11). Deviations are encountered for several reasons: (i) we neglect the disk-like

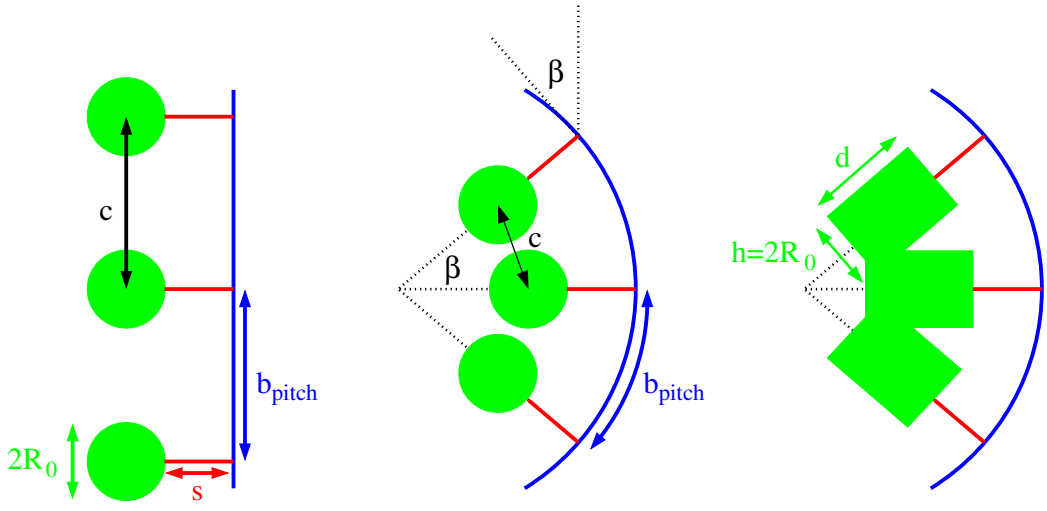


Figure 5.10: Illustration (side view) of fiber geometry with nucleosomal spheres of radius R_0 to estimate the bending persistence length as a function of sphere size for given linker length b_{linker} and stem s. On the right one can see the bent fiber where we substituted the spheres with disks of size $h = 2R_0$ and $d = 1.6h$. Here an overlap is found for the same bending angle β.

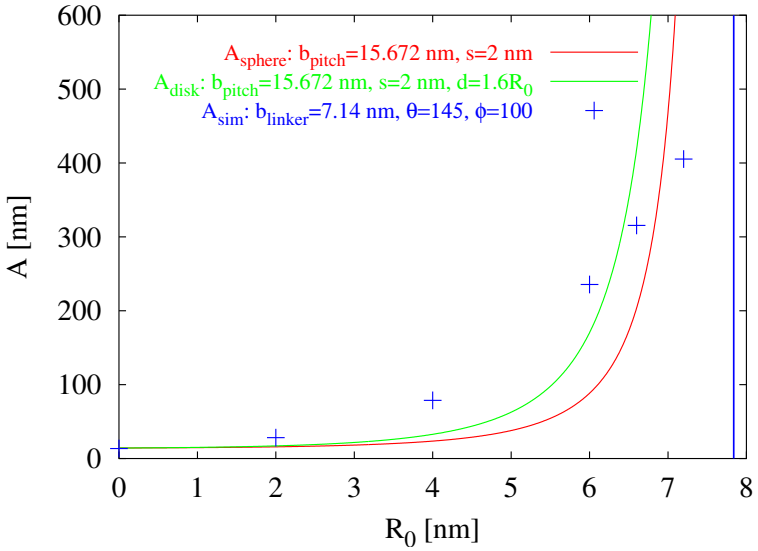


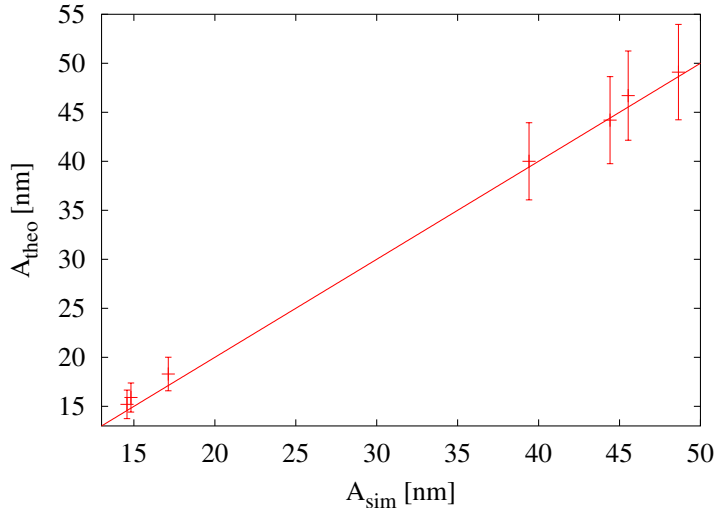
Figure 5.11: Comparison of the estimated bending persistence length of the fiber (A_{sphere}: red line, A_{disk}: green line) using the geometry illustrated in Fig. 5.10 and the simulation data (blue data points) with $\theta = 145°$ and $\phi = 100°$. The helical pitch is fixed to $b_{pitch} = 15.67$ nm in order to enforce a divergence of the persistence length for a radius of 7.8 nm as it is found in the simulation. The length of the stem amounts to $s = 2$ nm. Deviations are mainly due to the fact that the nucleosomes of the simulated fiber do not stack perfectly in columns which is one of our assumptions for an estimate of A_{sphere} and A_{disk}.

structure of the nucleosomes. The diameter of the spheres corresponds to the height of the disks so that we underestimate A depending on the 3-d structure under consideration. An analogous calculation of the distance between the edges of the disks results in $\beta_{max} = \frac{b_{pitch}-2R_0}{s+d}$ where d is the diameter of the disk, i.e. $A_{sphere} < A_{disk}$. (ii) We only treat bending in 2 dimensions. Bending out of the plane becomes important for zig-zag-like structures. In this case we overestimate the bending persistence length. Only for very compact structures (large θ small ϕ) where one encounters isotropic bending in all directions our estimate becomes comparable to the simulation results.

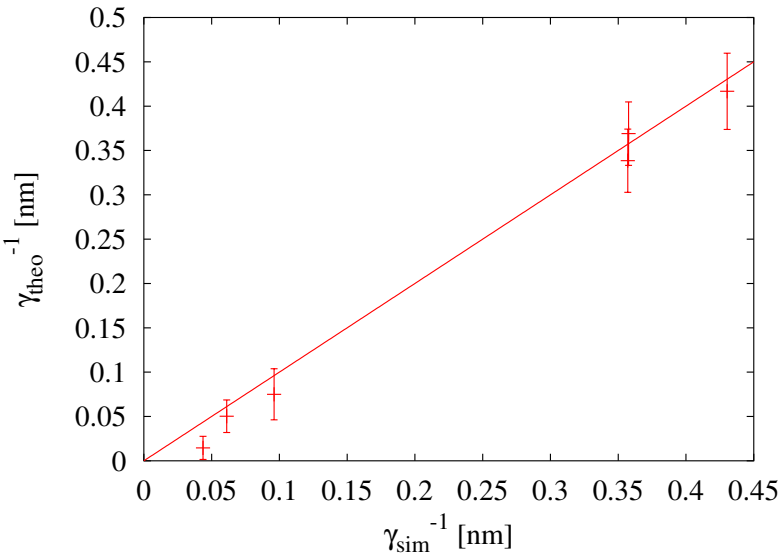
Simulation results

To test our simulation, we also measured the bending persistence length and the stretching modulus of the fiber for various combinations of θ and ϕ where we take nothing but the elasticity of the linker DNA into account. The calculation is done as follows: first of all we determine the fiber axis by calculating subsequently the centers of mass $\{\vec{c}_i\}$ defined by nucleosomes i up to $i + N_c$, $i + 2$ up to $i + 2 + N_c$, etc. N_c is chosen to match approximately one or two helical turns. Afterwards one can calculate the autocorrelation function of the tangent vectors defined by the fiber axis

$$\mathbf{t}_i = \frac{\vec{c}_{i+1} - \vec{c}_i}{|\vec{c}_{i+1} - \vec{c}_i|} \tag{5.24}$$



(a) Comparison of analytical results of the bending persistence length A_{theo} and the measured values A_{sim} of the simulation for different pairs of angles θ and ϕ. The agreement is excellent.



(b) Comparison of analytical results of the stretching modulus γ_{theo} and the measured values γ_{sim} of the simulation for different pairs of angles θ and ϕ.

Figure 5.12: Analytical and simulation results of the bending persistence length and the stretching modulus

in order to extract the persistence length of the fiber A from an exponential fit

$$\langle \mathbf{t}_i \cdot \mathbf{t}_j \rangle = \exp\left(-\frac{|i-j| b_{linker}}{A} \right). \tag{5.25}$$

The stretching modulus can be calculated via

$$\gamma = \frac{\langle L \rangle}{\langle \Delta L \rangle^2} \tag{5.26}$$

with $\Delta L = L - \langle L \rangle$ being the mean deviations around the average contour length of the fiber (length of the fiber axis). It should be noted that depending on N_c the estimated persistence length as well as the stretching modulus and the contour length of the fiber vary. In case of N_c being too large for example the stretching modulus is underestimated since bending fluctuations within i and $i + N_c$ contribute to γ. On the other hand values of N_c which are too small lead to a helicoidal fiber axis and the contour length of the fiber is overestimated. This entails a systematic error which must be minimized. Values of N_c corresponding to one or two helical turns appear to be reasonable.

Even though the analytical calculations determine the elastic constants of the fiber by a perturbation analysis around the $T = 0$ structure, the agreement between the simulation data and the analytical result for different pairs of angles is very good (see Fig. 5.12). In general the fluctuations lead to bending and twisting of the linker DNA away from the straight $T = 0$ conformation. This in turn leads to a smaller effective length per linker DNA and a subsequent change of the elastic constants. For rather extended zig-zag-like ground structures this effect is most pronounced and one finds the largest deviations from the analytical expressions.

Another possible method of analysis is the measurement of the mean squared nucleosomal distances R_E^2 along the chain. The data can subsequently be fitted to the extensible WLC model. Fit parameters are the contour length L, the persistence length A and the stretching modulus γ. We obtain:

- for $\theta = 145°$, $\phi = 100°$, $d = 11$nm, $h = 6.6$nm:
  - fitting result: $L/N = 1.92$nm, $A = 340$nm, $\gamma = 59\text{nm}^{-1}$, $\lambda = 5.8$
  - averaged contour method: $L/N = 1.90$nm, $A = 330$nm, $\gamma = 15\text{nm}^{-1}$, $\lambda = 5.6$
- for $\theta = 145°$, $\phi = 100°$, $d = 10$nm, $h = 6$nm:
  - fitting result $L/N = 1.84$nm, $A = 244$nm, $\gamma = 47\text{nm}^{-1}$, $\lambda = 6.0$
  - averaged contour method: $L/N = 1.82$nm, $A = 236$nm, $\gamma = 14\text{nm}^{-1}$, $\lambda = 6.1$.

The differences of L and A between both methods are smaller than 5%. But one should mention that some of the data for R_E^2 could not be fitted properly due to strong data scattering for small $j - i$, especially for the simulations with pure bending stiffness. That is also why one finds strong deviations for the stretching modulus γ between both methods, since γ is determined by the small distance region of R_E^2. The bending persistence length, on the other hand, is dominated by large distances where the data do not scatter strongly. In Fig. 5.13 we compare R_E^2 calculated with the raw data and with the average contour length data. One can see how the data set of R_E^2 is smoothed by the averaging procedure for small internal distances, while the large distance behavior does not change. The red line corresponds to the extensible WLC fit to the averaged data.

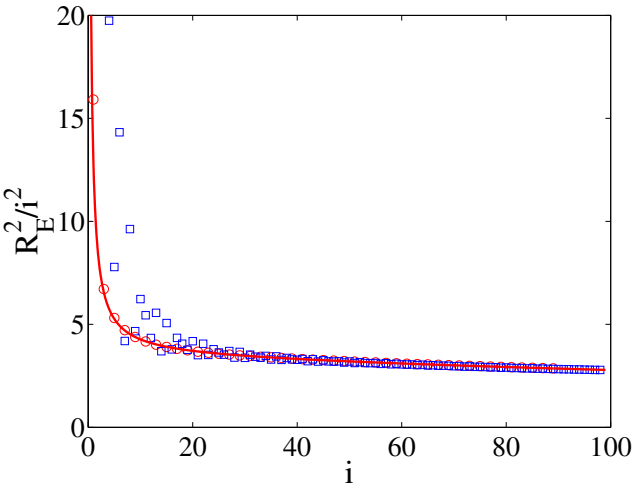


Figure 5.13: Mean squared internal distances of the nucleosomes R_E^2 measured with the center-of-mass of the nucleosomes (blue squares) and the smoothed contour (red circles) for $\theta = 145°$, $\phi = 100°$, $d = 10$nm, $h = 6$nm. i denotes the monomer index. Fitting result (red line): $L/N = 1.84$nm, $A = 244$nm, $\gamma = 47\text{nm}^{-1}$, $\lambda = 6.0$; averaged contour method: $L/N = 1.82$nm, $A = 236$nm, $\gamma = 14\text{nm}^{-1}$, $\lambda = 6.1$. The fitting procedure results in a $3 - 4$ times larger stretching modulus compared to the averaged contour method while the obtained contour length L, the persistence length A and the mass density λ are in good agreement.

We already discussed the limitations of the analytical approach due to the negligence of nucleosomal interactions. For low salt concentrations, that is for large entry-exit angles, one expects that the chromatin fiber is a highly flexible and extendable object. This behavior is captured by Eq. (5.11). However, for higher salt concentrations the structure becomes more compact such that excluded volume interactions play an important role. In order to investigate excluded volume effects on the bending persistence length of the fiber, we simulated various fibers with different nucleosome volumes. This is done without attractive forces. We simply check during the simulation if an overlap occurs and, if yes, such a move is then rejected. We keep the ratio of the nucleosomal height h to its diameter d constant with $\frac{h}{d} = 0.67$. For rather compact structures the bending fluctuations are impeded by the excluded volume of the disks. Thus one anticipates large bending persistence lengths of the fiber for large nucleosomal disk sizes. This is indeed confirmed by our simulations. Fig. 5.14 illustrates the dependence of the persistence length for given angles θ and ϕ on the nucleosomal disk size. We observe that the effect is smaller for less compact structures such as zig-zag fibers, since there is more room for bending fluctuations. The snapshots for different nucleosomal disk sizes in Fig. 5.14 nicely illustrate the stiffening of the fiber with increasing disk size.

In Fig. 5.15 the bending persistence length A of a fiber with and without nucleosomal attraction is shown. It is observed that the bending persistence length of a fiber with $\theta = 145°$, $\phi = 110°$ is reduced if nucleosomal attraction is present. Since neighboring nucleosomes come closer to each other if the fiber bends (see Fig. 5.10) the energy contribution due to nucleosomal attraction is enhanced. That is why bending is enhanced compared to the case of pure excluded volume interaction.

The stretching modulus γ of the fiber also increases with increasing disk size which is caused

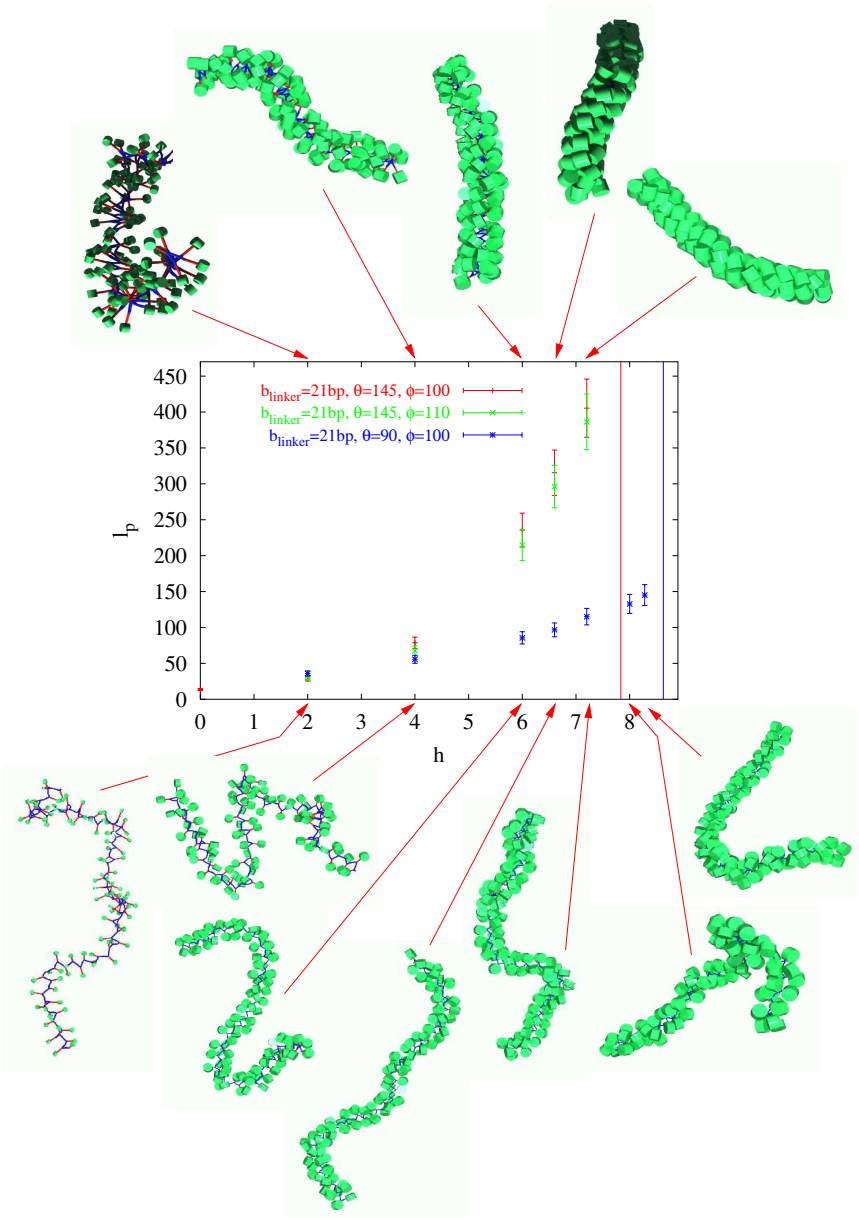


Figure 5.14: Effect of excluded volume interaction on the bending persistence length of the fiber for various pairs of angles θ and ϕ is shown (see inset). For a given linker length b_{linker} the fiber persistence length A [nm] grows if one increases the overall size of the nucleosomal disks while keeping the aspect ratio constant. The vertical lines show the limit where the ground structure $T = 0$ shows overlap of nucleosomes. Note that the divergence limits of the red and green data set are nearly identical. The persistence lengths of compact structures diverge for these values of h [nm]. For very small disk sizes the measured persistence lengths converge to the analytical values calculated with Eq. (5.11).

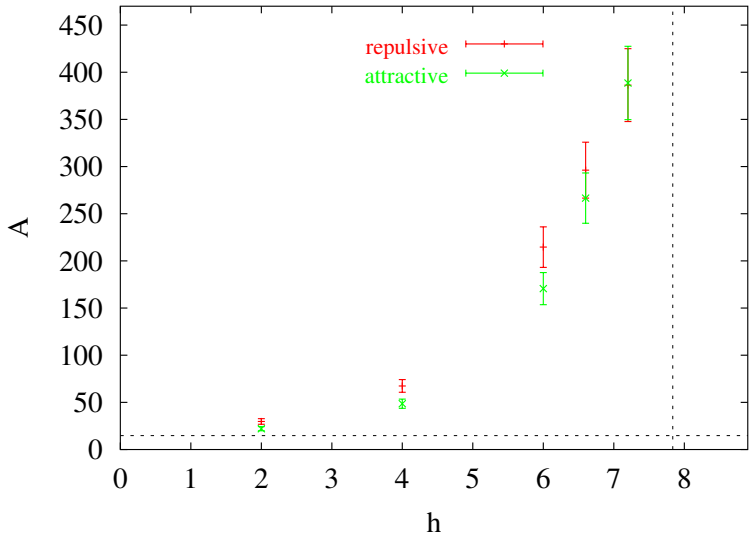


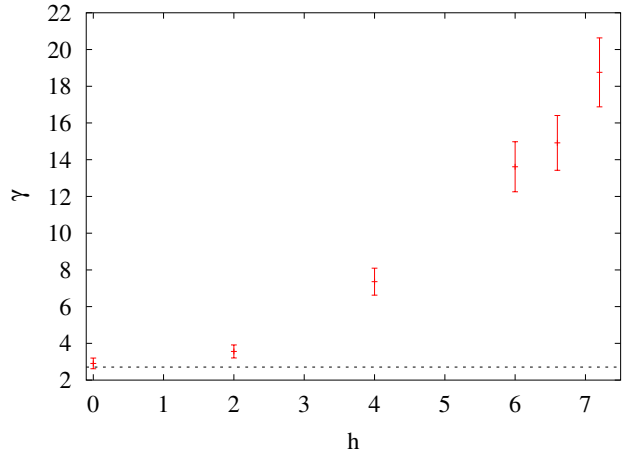
Figure 5.15: Persistence length A [nm] as a function of nucleosomal disk size (h [nm]) with $\theta = 145°$, $\phi = 110°$. We compare fibers with ($\epsilon = 1k_BT$) and without ($\epsilon = 0k_BT$) nucleosomal attraction where ϵ is the Gay-Berne energy depth for two stacked disks.

by the hindrance of longitudinal fluctuations due to the excluded volume of the nucleosomes. The mass density λ decreases with increasing disk size. This displays the already mentioned fact, that in general the fluctuations lead to bending and twisting of the linker DNA away from the straight $T = 0$ conformation. That in turn leads to a smaller effective length per linker DNA and a smaller contour length respectively. Since these bending and twisting fluctuations are reduced for larger nucleosomal disk sizes the mass density decreases. Fig. 5.16 illustrates the dependence of γ and λ on the nucleosomal disk size for fibers with $\theta = 145^o$, $\phi = 100^o$, and $b_{linker} = 7.14$nm. For $d = 10$nm, $h = 6$nm corresponding to the experimentally determined spatial dimension of the chromatosomes [LMR+97] we find a mass density of about $\lambda = 6.1$ very close to experimental values under physiological conditions [BHG+98] and a stretching modulus of about $\gamma = 13.6\text{nm}^{-1}$ which is roughly 10 times larger than the value determined by a fit of the release part of the force-extension curves at low ionic strength [CB00].

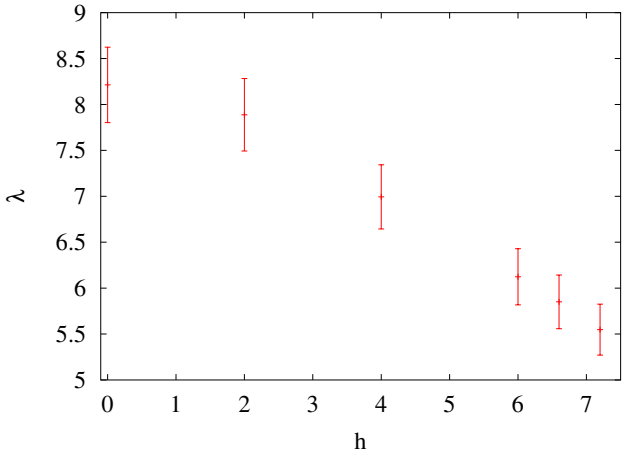
5.6.3 Stretching the fiber

In the stretching experiments of Cui and Bustamante [CB00] on chicken erythrocyte chromatin fibers in physiological conditions, i.e. for very compact fibers, a force-plateau at around 3-5 pN is found. The authors propose that the plateau corresponds to a condensation-decondensation transition due to nucleosome-nucleosome interactions.

Within our model the responsible parameter for a condensation-decondensation transition caused by short range attractive interaction between nucleosomes is the energy well depth ϵ of the GB potential. It determines the energy that must be paid to pull two nucleosomes apart. For sufficiently large values of ϵ one should find a force-plateau in the force-extension curve. On the other hand, one could also imagine that the fiber will get more and more compact until it prefers to form loop



(a) Stretching modulus γ vs. nucleosomal disk size (height h). The dotted line corresponds to the analytical result.



(b) Mass density λ vs. nucleosomal disk size (height h).

Figure 5.16: Stretching modulus γ [nm] and mass density λ [number of nucleosomes/11nm] vs. nucleosomal disk size (height h [nm]) for $\theta = 145^o$, $\phi = 100^o$, and $b_{linker} = 7.14$nm. For $d = 10$nm, $h = 6$nm corresponding to the experimental values [LMR$^+$97] we find a mass density of about $\lambda = 6.1$ very close to experimental values under physiological conditions [BHG$^+$98] and a stretching modulus of about $\gamma = 13.6\text{nm}^{-1}$ which is roughly 10 times larger than the value determined by a fit of the release part of the force-extension curves at low ionic strength [CB00].

structures with a kink near the center of the chain in order to maximize the contact area of the surface of the fiber. This will be discussed later on in more detail.

Since we are interested in the stretching behavior of a chromatin fiber under physiological con-

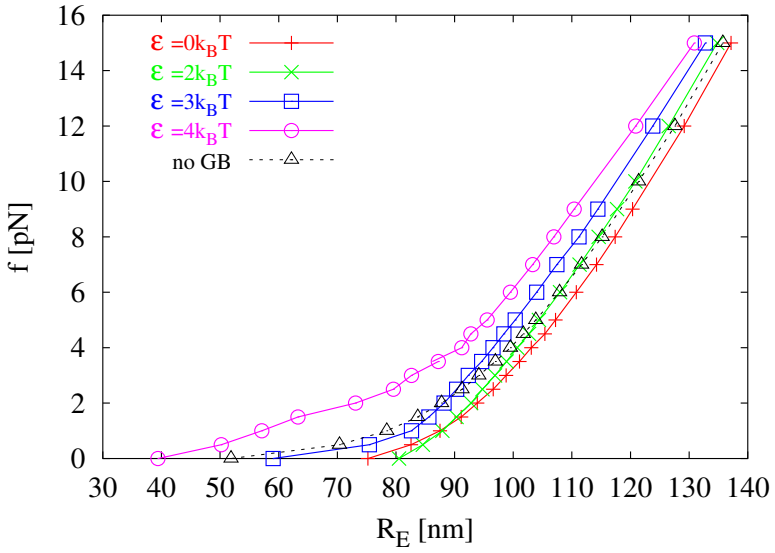


Figure 5.17: Force-extension curves for a fiber where only the linker DNA elasticity is taken into account and for fibers with hard-core repulsion ($\epsilon = 0$) and soft-core attraction between the nucleosomes with $\epsilon = 2, 3, 4k_BT$. A further increase of ϵ should shift the plateau to larger stretching forces f.

ditions we choose the following structural parameter set to simulate the fiber with external stretching force:

$$\theta = 145^o$$
$$\phi = 110^o$$
$$b_{linker} = 7.14\text{nm}.$$

$\pi - \theta$ is chosen in order to match the experimental value of 35^o measured in high salt concentration [BHG$^+$98]. b_{linker} and ϕ are fixed such that one obtains a reasonable value for the mass density λ of about 6.1 nucleosomes per 11nm measured with $\epsilon = 1k_BT$. Note that b_{linker} coincides with linker lengths of erythrocyte chromatin fibers in the presence of linker histones measured by digestion experiments [K. 89]. Even though the mass density will increase with increasing potential well depth ϵ it does not exceed a value of 7.5 nucleosomes per 11nm. This is still close to experimental data which suggest values of about $6 - 7$ nucleosomes per 11nm [GR87]. Certainly other pairs of $\{b_{linker}, \phi\}$ exist which produce a mass density of about 6 nucleosomes per 11nm, but the overall behavior of the stress-strain relations should not be influenced. Longer linker DNA segments will decrease the mass density and will increase the flexibility of the fiber for fixed ϕ. Therefore one will find the already mentioned kink instability for smaller values of ϵ.

We measure the force-extension relation for various values of ϵ. The stretching force acts along the center-to-center distance of the chain. The results are illustrated in Fig. 5.17. For small values of ϵ no force plateau is found. The force-extension curves follow the extensible WLC behavior. Note that since we simulate quite short chromatin fibers, finite size effects are present. For $\epsilon = 4k_BT$ we find a force-plateau at about 2pN which looks quite similar to the experimental data [CB00] (see Fig.

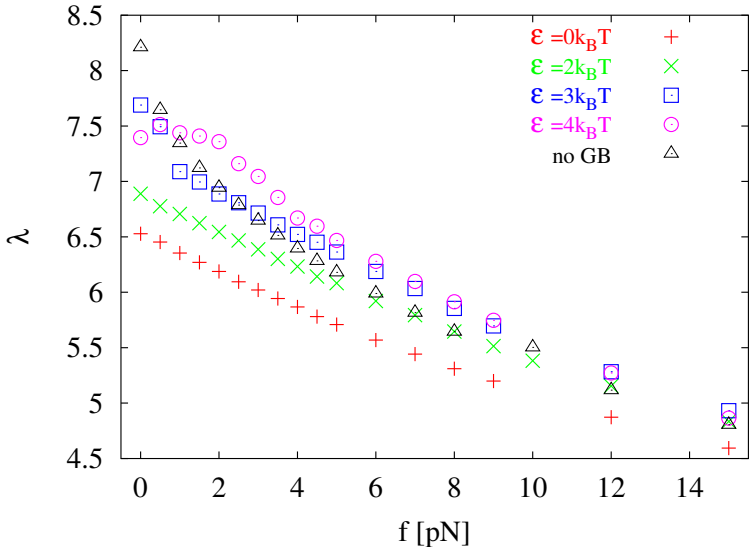


Figure 5.18: Stretching force f vs. mass density λ for a fiber with $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm where only the linker DNA elasticity is taken into account and for fibers with nucleosomal interactions with $\epsilon = 0,\ 2,\ 3,\ 4k_BT$. One can nicely see how the mass density stays constant for $\epsilon = 4k_BT$ up to 2pN. Up to these stretching forces only the loop formation of the fiber is pulled out.

5.21). A further increase of ϵ should shift the plateau to larger stretching forces f. The stress-strain curve is averaged over 10 different simulation runs. For the different runs the data scatter quite a lot in the low stretching force regime. This is due to a kinked loop formation of the fiber. Once the chain forms a kink at a certain position along the chain and the two parts of the chain are glued together the kink position seems to be quite immobile while the loop conformation can open due to thermal fluctuations. In principle the kink location can diffuse along the chain. But this diffusion process would lead to states of much higher energy. One possibility to overcome this problem is to run several simulations where the chains form kinks at different fiber positions. Subsequently one performs a kind of ensemble averaging over the different simulation runs.

In Fig. 5.22 we show the probability distributions of the end-to-end distance R_E for different GB energy well depths. One can clearly observe a shift to smaller values of R_E for $\epsilon \geq 3k_BT$ corresponding to the formation of kinks. By stretching the fiber the nucleosomal contacts of the two parts of the chain are pulled out in a first step (up to 2pN), followed by a decondensation transition of the fiber. Moreover, it is found that for large stretching forces (> 5pN) the probability distribution functions for the different ϵ values coincide. The measurement of the mean internal distances $\langle R_E \rangle$ underline the kink observation. For several runs $\langle R_E \rangle$ shows a parabolic shape where the maximum is found at the kink location (compare Fig. 5.19). Another possibility to detect kinked structures is the measurement of the contact matrix $\mathcal{M}_{contact}$ of the fiber for each snapshot. If a pair of nucleosomes i and j is in contact we count $\mathcal{M}^{(ij)}_{contact} = 1$ otherwise $\mathcal{M}^{(ij)}_{contact} = 0$. By adding up the contact matrices of each snapshot we obtain a two dimensional histogram as it is shown in Fig. 5.20. If no kinks are present and the nucleosomes of the fiber form a quite regular solenoidal

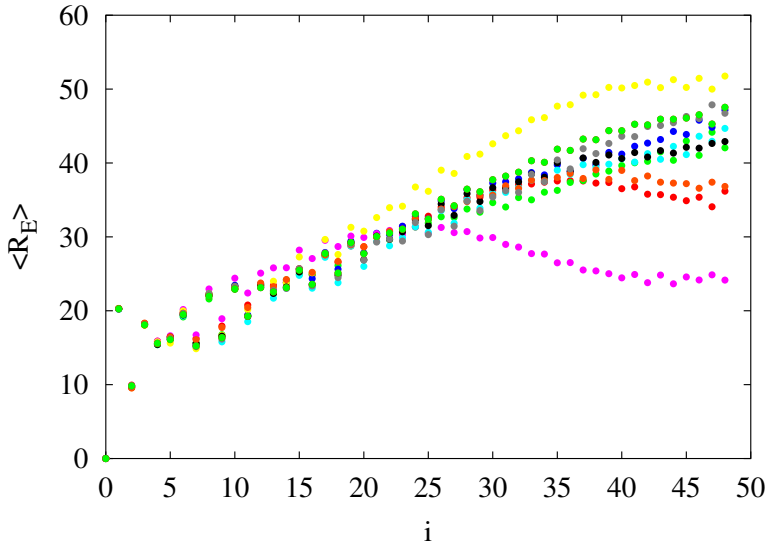


Figure 5.19: Mean internal distances $\langle R_E \rangle$ for the individual simulation runs (different colors correspond to different runs) with $\epsilon = 4k_BT$, $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm and no applied stretching force. i refers to the monomer index. Parabolic shapes are recovered where the maximum is found at the kink location.

structure two main stripes with some less pronounced side stripes parallel to the main diagonal are observed. In the presence of kinks some side branches perpendicular to the main diagonal, i.e. along the secondary diagonal, can be found. If loop conformations are present, the histograms show a cross-like pattern where the kink location is found where the two branches along the main and secondary diagonal respectively cross each other.

5.7 Discussion

We introduced a simple model based on the two-angle model by Schiessel *et al.* [SGB01] which additionally accounts for excluded volume and attractive interaction of the nucleosomes. The structural analysis of the two-angle phase diagram helped to identify the boundary between allowed and forbidden geometries. Moreover we investigated the influence of the nucleosome-nucleosome inter-

θ	ϕ	A	A_{el}
90	100	75	34
145	100	236	14
145	110	216	15

Table 5.1: Measured persistence lengths A in [nm] for various pairs of angles θ and ϕ in [°]. A_{el} denotes the analytical result without excluded volume, i.e. the contribution from the linker DNA. The spatial dimension of the nucleosomes is chosen to match the experimental values of $d = 10$nm and $h = 6$nm.

θ	ϕ	λ
145	79	7.0
145	100	6.3
145	110	6.1

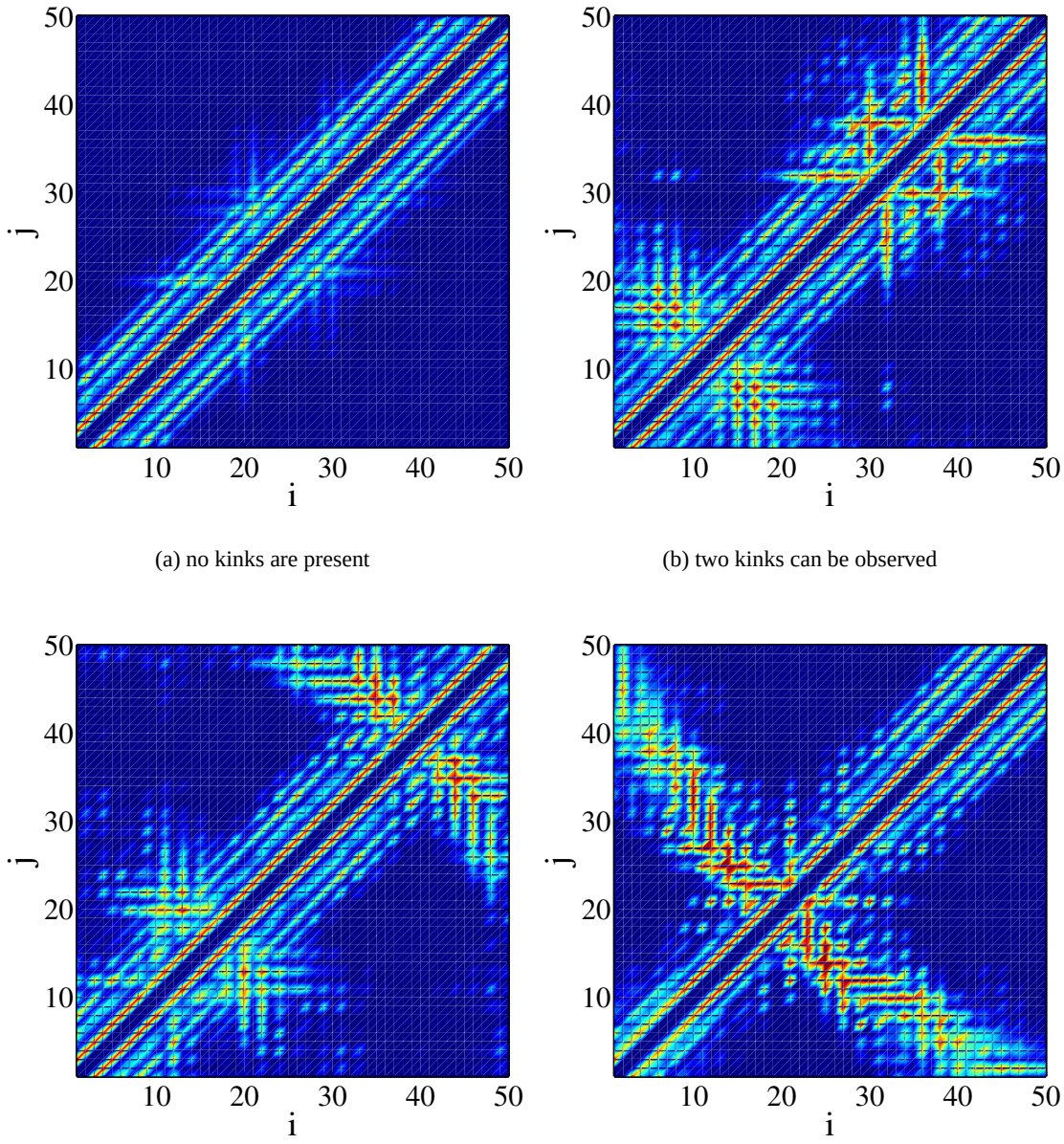
Table 5.2: Mass density λ [number of nucleosomes/11nm] for various values of ϕ with $\theta = 145^o$ and $\epsilon = 1k_BT$. For larger rotational angles ϕ the fiber gets less compact resulting in smaller mass densities λ.

action on structural as well as elastic properties of the chromatin fiber. We find a drastical increase in the bending persistence length of the fiber for increasing nucleosomal disk sizes. In particular for a disk size of $d = 10$nm and $h = 6$nm and an underlying geometry given by $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm where the entry-exit angle is fixed to the experimental value of 35° [BHG$^+$98] and the mass density of the fiber under physiological conditions [GR87] could be reproduced the bending persistence length amounts to 216nm close to the value 260nm found by earlier computer simulation studies [WL02]. Note that in Ref. [WL02] a linker length of $b_{linker} = 3.74$nm is chosen corresponding to 11 base-pairs which explains the slightly larger value. Since the density $\rho = (\pi R s_0)^{-1}$ of a fiber with given entry-exit angle $\pi - \theta$ and given linker length b_{linker} is decreased if one goes to larger values of the rotational angle ϕ there is more space left for bending fluctuations and one recovers smaller persistence lengths. On the other hand for fixed ϕ and b_{linker} one finds smaller densities for larger entry-exit angles which in turn leads to smaller persistence lengths (compare table 5.1).

We characterized the found transition for reasonable values of the model parameters θ, ϕ, b_{linker}, d, h to simulate a fiber under physiological conditions. The entry-exit angle is fixed to the experimental value of 35° [BHG$^+$98] and the mass density [GR87] could be reproduced.

For a rotational angle of $\phi = 100^o$ we recover a mass density of about $\lambda = 6.1$ very close to experimental values under physiological conditions [BHG$^+$98] and a stretching modulus of about $\gamma = 13.6\text{nm}^{-1}$ which is roughly 10 times larger than the value determined by the fitting procedure of Cui and Bustamante [CB00] (see Fig. 5.16). Remember that they used a fit of the release part of the force-extension curves at low ionic strength where the fiber is less compact and no excluded volume effects are present. The large stretching modulus is caused by the hindrance of longitudinal fluctuations due to the excluded volume of the nucleosomes. For larger rotational angles ϕ and fixed θ the fiber gets less compact resulting in smaller mass densities λ. This is displayed in table 5.2.

The simulations with applied stretching forces show for $\epsilon = 4k_BT$ a force-plateau in the stress-strain curve rather similar to the one observed in micromanipulation experiments on single chromatin fibers [CB00] (see Fig. 5.21). Due to the strong nucleosomal attraction the fiber forms kinks, and the two parts of the fiber glue together in order to maximize the contact area of the fiber surface. For small forces the chain is unglued, but kinks are still present. In contrast to DNA, the chromatin fiber does not undergo an instantaneous local structural transition if a critical stretching force is exceeded. Instead we observe a global structural change up to 2pN where the two parts of the loop are torn apart. For stretching forces larger than 2pN a decondensation of neighboring nucleosomes along the fiber takes place, but the overall structure of the fiber is not strongly distorted. For even larger forces the fiber is lengthened by increasing the angle formed by nucleosome $i-1$, i and $i+1$ which is realized by strong bending and by untwisting of the linker DNA. This stretching behavior of the fiber is additionally confirmed by the measurement of the mass density as a function of the stretching force f (see Fig. 5.18). In case of $\epsilon = 4k_BT$ one can clearly see a plateau up to forces



(a) no kinks are present

(b) two kinks can be observed

(c) two kinks and one loop can be observed

(d) one loop is present with kink location near the center of the fiber

Figure 5.20: Density plot of measured histograms of the contact matrix of the fiber with $\epsilon = 4k_BT$, $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm and with $f = 0$pN in case of (b), (c), and (d) and with $f = 3.5$pN in case of (a). i and j denote the nucleosome index.

$f = 2$pN. In this regime the loops are pulled out. For larger stretching forces the fiber starts to decondense by untwisting and strong bending of the linker DNA which results in smaller values for λ. For very large forces the fiber will from an extended string of nucleosomes. Some snapshots

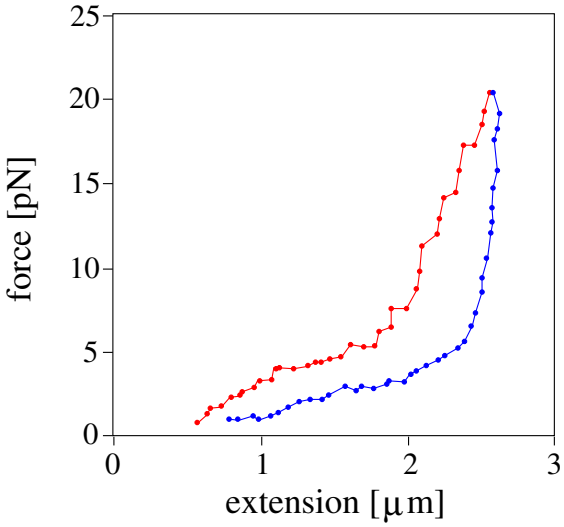


Figure 5.21: Force-extension curves for chicken erythrocyte chromatin fibers in 40mM NaCl. The red curve corresponds to the stretch cycle while the blue curve refers to the release cycle. Taken out of [CB00].

corresponding to different stretching forces for $\epsilon = 4k_BT$ are illustrated in Fig. 5.23. Even for quite large stretching forces f some of the nucleosomes are in contact such that there is no full decondensation. This appears also in the measured force-extension curves. In the high stretching regime one can observe a shift towards smaller end-to-end distances R_E as ϵ becomes larger. In case of longer linker DNA segments bending and untwisting the fiber after the decondensation transition would be easier such that neighboring nucleosomes are farther apart and a full decondensation can take place. As an additional result the mass density would decrease faster to the value corresponding to a fully extended fiber conformation. Moreover, the fiber will form more easily a kinked structure. For even larger values of ϵ one expects the force-plateau to be shifted to larger stretching forces.

Katritch *et al* [KBO00] used a similar model to simulate force-extension curves. The major difference consists in their assumption that ϕ is fully randomized. A quenched disorder of the rotational angle ϕ has rather strong effects on the structural as well as the elastic properties of the simulated fiber [KBO00]. Furthermore, spheres are used instead of disks. They investigated the dependence of linker length, entry-exit angle and rotational angle on the shape of the force-extension curves. The comparison of their simulation results to the relaxation data of the stretching experiments on chicken erythrocyte chromatin fibers [CB00] postulated values of $b_{linker} = 40$bp, $\theta = 130^o$, an effective nucleosomal diameter of $d_{eff} = 14$nm, and a random distribution of the rotational angle ϕ. While the determined entry-exit angle is close to the experimental value, $b_{linker} = 40$bp is in contradiction to digestion experiments yielding a value of about 20bp [K. 89], and $d_{eff} = 14$nm is quite large. For the optimized model they introduced an additional short ranged attractive potential between the nucleosomes. For large attraction they also recovered a force-plateau and a very strong reduction in the end-to-end distance without applied stretching force, but failed to provide a quantitative analysis of the structural changes during stretching.

We characterized the found transition for reasonable values of the model parameters θ, ϕ, b_{linker},

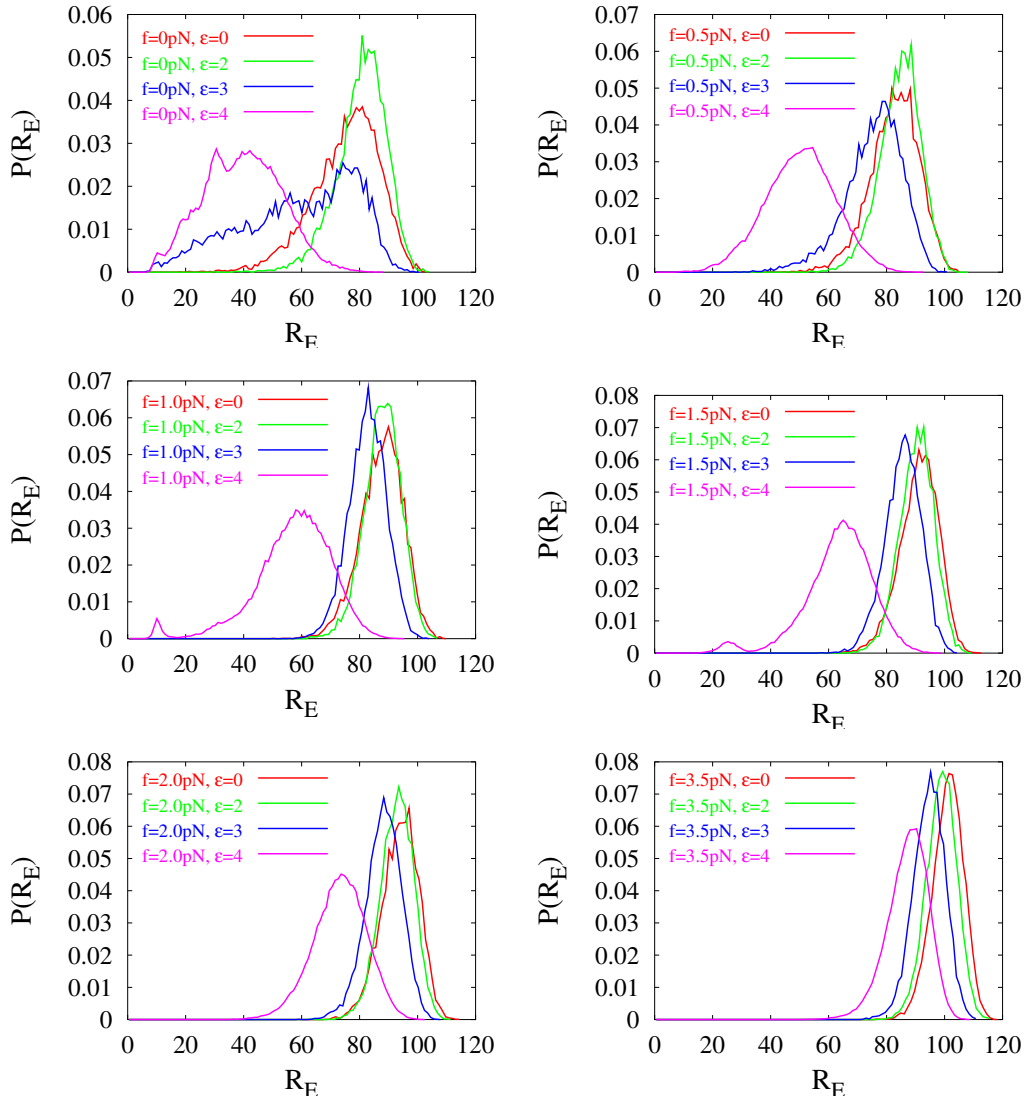


Figure 5.22: PDF of R_E of a fiber with $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm for various GB energy well depths and stretching forces (see legends). For $f = 0$pN and $\epsilon \geq 3k_BT$ a shift to smaller values of R_E is recovered corresponding to the formation of kinks. By stretching the fiber the nucleosomal contacts of the two parts of the chain, i.e. the loop formation, are stretched out.

d, and h to simulate a fiber under physiological conditions. The entry-exit angle is fixed to the experimental value of 35° [BHG$^+$98] and the mass density [GR87] could be reproduced. Within our model we predict a loop formation for strong nucleosomal attraction which is stretched out for forces up to 2pN. Whether or not kinks are present depends solely on the length of the fiber and on ϵ. For a given ϵ loop formations will emerge if the fiber length exceeds a critical value. We can measure the additional attractive energy due to the loop formation in the following way: we calculate the GB energy for a loop conformation and subtract the GB energy of a condensed but

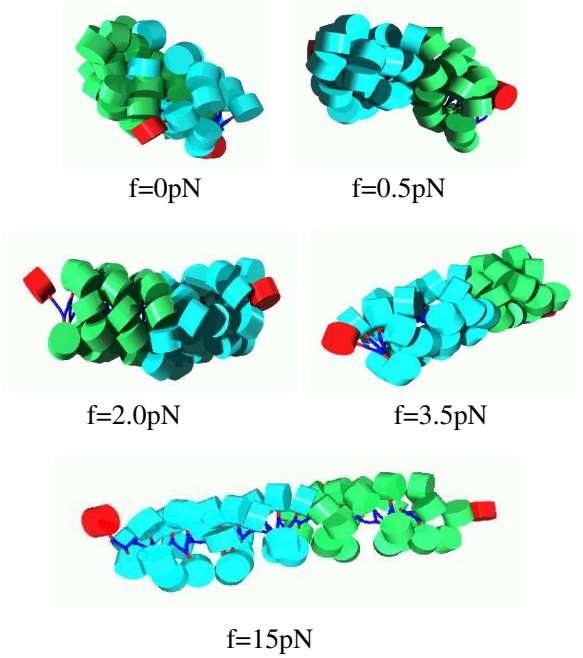


Figure 5.23: Snapshots of fibers with $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm corresponding to different stretching forces for $\epsilon = 4k_BT$. To facilitate the detection of a loop formation or a kink one half of the chain is dyed green while the other half is dyed cyan. The first and last nucleosomes are labeled red. For $f = 0$pN a kink is detected close to the center of the chain such that the fiber forms a loop in order to maximize the contact area of its surface. Up to $f = 2$pN the kink is still present but the loop gets pulled out. For $f = 15$pN the fiber is partially decondensated.

straight conformation for a given ϵ_{GB}. We obtain a value of $\mathcal{E}_{loop} = 3k_BT/4$nm for $\epsilon = 4k_BT$. If we assume a loop conformation which is composed of two straight chains connected by a semicircle we obtain an estimate for the radius of curvature with $R = \sqrt{\frac{l_p}{2}\mathcal{E}_{loop}}$. With $l_p \approx 200$nm the radius of curvature amounts to about 9nm which is about the size of the diameter of the nucleosomes $d = 10$nm.

Chapter 6

Summary of results and outlook on future work

We have investigated the structural and elastic properties of three different systems: (i) a generic model for double-stranded polymers consisting of two coupled worm-like chains, (ii) a mesoscopic model for DNA at the base-pair level, and (iii) a crossed-linker model of the chromatin fiber.

In a first step towards more microscopic models of DNA we studied the effects caused by the double-strandedness of DNA. We introduce the coupling between the two strands - taking into account the hydrogen bonding - as a rigid one such that bending in the plane of the ribbon is forbidden. We performed MC simulations and rationalized the results with the help of a simple scaling argument. Two relevant length scales are found in the problem: the single strand persistence length l_p and the separation length a of the two strands. We recover the predicted simple exponential decay of the tangent-tangent correlation function with the single strand persistence length. The tangent-tangent correlation function is independent of the separation of the two strands. Also in agreement with Ref. [LGK98, GL00] we find an inherent local twist structure with a helical persistence length l_b and a helical pitch P which scale with $l_p^{1/3}a^{2/3}$. But in contradiction to Ref. [LGK98, GL00] we do not observe an oscillatory behavior of the bond-director correlation function. Qualitatively, one would expect to see oscillations in the bond-director correlation function, if $P \leq l_b$. If the mean twist within one helical persistence length exceeds π an oscillatory behavior has to be observed. But our calculation gives a twist rate within $l_b = gb$ of approximately $\pm 1/16$. Moreover our simulation results with applied constant buckling forces do not provide any evidence of a tendency of kinking or an inherent twist-stretch coupling. A twist-stretch coupling has to be incorporated by the inclusion of an additional term in the Hamiltonian along the lines of Refs. [KLNO97, Mar97, Nel98]. We have dealt with these effects in chapter 3 where we have seen that the twist-stretch coupling is a natural result of the preferred stacking of neighboring base-pairs and the rigid backbones and must not be included by an extra term in the Hamiltonian. Interestingly, kinked structures are recovered in the modeled chromatin fiber.

In chapter 3 we have introduced a simple mesoscopic model of DNA on the base-pair level combining the ideas of SOP and WLC model. We studied the resulting helical conformation using energy minimization and Monte Carlo simulations. We showed by systematic coarse-graining that the elastic constants of the WLC model are related to the local fluctuations of the base-pair step parameters. Assuming a twisted ladder as ground state conformation an analytical relationship between the persistence lengths and the stretching modulus respectively, and the local elastic constants

can be provided given by Eqs. (3.26), (3.30), and (3.51). The general case where the ground state is characterized by spontaneous rotations as well as spontaneous displacements as in the A-DNA conformation is more involved. This is the subject of chapter 4.

For a suitable choice of parameters we obtained a B-DNA like ground state as well as realistic values for the bend and twist persistence lengths. The latter were obtained by analyzing the thermal fluctuations of long filaments as well as by a systematic coarse-graining from the stack-of-plates to the elastic rod level. In studying the response of DNA to external forces or torques, models of the present type are not restricted to the regime of small local deformations. The measured critical force is overestimated by a factor of 2 and cannot be improved further by fine-tuning of the three free model parameters $\{b, k, \epsilon\}$. f_{crit} depends solely on the stacking energy value ϵ that cannot be reduced further since it would reproduce neither the correct equilibrium structure of B-DNA nor the correct persistence lengths. It is an interesting question if the introduction of other stacking potentials can solve this problem (see for example Ref. [EE03]).

Our model suggests a structure for S-DNA with highly inclined base-pairs in order to enable at least partial base-pair stacking. This is in good agreement with results of atomistic B-DNA simulations by Lavery *et al.* [CLH$^+$96, LL99]. They found a force plateau of 140pN for freely rotating ends [CLH$^+$96]. The mapping to the SOP model yields the following twist-stretch (Ri-Tw) coupling constant $k_{Ri,Tw} = (\mathcal{C}^{-1})_{Ri,Tw} = 267/\text{Å}$ (see Eq. (3.48)). Here $k_{Ri,Tw}$ is the microscopic coupling of rise and twist describing the untwisting of the chain due to an increase of rise (compare also Fig. 3.9).

While our results suggest a similar structure for S-DNA as atomistic simulations [CLH$^+$96], the DNA model studied in this paper can, of course, not be used to rule out the alternate possibility of local strand separations [WWRB01, RB01a, RB01b]. Therefore one future project is the generalization of our approach to a description on the base level which includes the possibility of hydrogen-bond breaking between complementary bases along the lines of Ref. [BCP99, CM99]. Additionally a suitably parameterized model allows a more detailed investigation of DNA unzipping experiments [BERH97]. Clearly, it is possible to study sequence-effects and even more refined models of DNA. For example, it is possible to mimic minor and major groove by bringing the backbones closer to one side of the ellipsoids without observing non-B-DNA like ground states. The relaxation of the internal degrees of freedom of the base-pairs characterized by another set of parameters (propeller twist, stagger, etc.) should help to reduce artifacts which are due to the ellipsoidal shape of the base-pairs. Sequence effects enter via the strength of the hydrogen bonds ($E_{GC} = 2.9k_BT$ versus $E_{AT} = 1.3k_BT$) as well as via base dependent stacking interactions [Hun93]. For example, one finds for guanine a concentration of negative charge on the major-groove edge whereas for cytosine one finds a concentration of positive charge on the major-groove edge. For adanine and thymine instead there is no strong joint concentration of partial charges [CD99]. It is known that in a solution of water and ethanol where the hydrophobic effect is less dominant these partial charges cause GG/CC steps to adopt A- or C-forms [FSH99] by a negative slide and positive roll motion and a positive slide motion respectively. Thus by varying the ratio of the strengths of the stacking versus the electrostatic energy it should be possible to study the transition from B-DNA to A-DNA and C-DNA respectively. The introduction of electrostatic interactions in the model also provides the possibility to investigate the charge renormalization of the WLC elastic constants [PHP00]. As a last future perspective it should be mentioned that the presented model may serve to validate and unify analytical approaches aiming at (averaged) properties on larger scales [ALCM01, HYZc99, BCP99, CM99, SN02a]. We already started to quantify the B-to-S transition and the structural and elastic properties of the S-DNA state in more detail and to compare it to the fitting results of the two-state model of Storm and

Nelson [SN02b, SN02a].

In chapter 4 we discuss possible renormalization procedures to obtain the elastic constants of the WLC model. It is shown that in case of nearest-neighbor interaction we can work out a convenient simple sampling procedure to calculate the effective coupling constants on larger length scales. Therefor one has to implement the following steps: (i) calculation of the elastic matrix by inverting the correlation matrix, (ii) diagonalization of the elastic matrix, (iii) independent random numbers can be drawn for the eigen-modes, and (iv) the conformation is obtained by transforming back. If long-ranged interactions are involved certainly the simple sampling method cannot be used anymore. Rather one has to implement Monte-Carlo renormalization techniques to measure the renormalized coupling constants along the chain. These techniques are important if long-ranged electrostatic interactions are included in the DNA model. In the future one should make a thorough investigation of the renormalization of the elastic constants for different DNA ground states.

Another challenging project is to keep the non-linear aspects of the DNA such as the over-stretching transition during renormalization. In principle one could renormalize separately the elastic matrix of the B-DNA and the S-DNA provided that it is possible to write down an effective Hamiltonian such that the critical force where the structural transition occurs is an invariant under the renormalization group operator. An Ising-like model as it is introduced by Storm and Nelson [SN02b, SN02a] could be a good starting point.

One major advantage of linked-ellipsoid models as introduced in chapter 3 is that they are not restricted to the base-pair level of DNA as the same techniques can, for example, also be used to study chromatin. This is the subject of chapter 5. In the case of chromatin the ellipsoids represent chromatosomes. The underlying geometry of the linker DNA is based on the two-angle model by Schiessel *et al* [SGB01]. A variant of the Gay-Berne potential [EE03] takes the excluded volume interactions and the attractive interactions (probably due to bridging of lysin-rich tails between the chromatosomes) into account. Parameters characterizing the underlying geometry such as the entry-exit angle $\pi - \theta$, the rotational angle ϕ and the linker length b_{linker} were varied in order to study their influence on the elastic and structural properties of the fiber. The structural analysis of the two-angle phase diagram helped to identify the boundary between allowed and forbidden geometries. Taking in a first step only the excluded volume interactions into account we find a drastical increase in the bending persistence length of the fiber for increasing nucleosomal disk sizes. In particular for a disk size of $d = 10$nm and $h = 6$nm and an underlying geometry given by $\theta = 145^o$, $\phi = 110^o$, and $b_{linker} = 7.14$nm, where the entry-exit angle [BHG$^+$98] and the spatial dimension of the chromatosomes [LMR$^+$97] are fixed to their experimental values and the mass density of the fiber under physiological conditions [GR87] could be reproduced, the bending persistence length amounts to 216nm. The stretching modulus amounts to about $\gamma = 13.6\text{nm}^{-1}$. Both values are order of magnitudes larger than the values determined by the fitting procedure of Cui and Bustamante [CB00]. Note that they fitted the extensible WLC model to the release part of the force-extension curves at low ionic strength where the fiber is less compact and no excluded volume effects are present.

Interestingly, we observe kinked fiber structures in the presence of nucleosomal attraction. Whether kinks are present or not depends solely on the length of the fiber and ϵ. For a given ϵ loop formations will emerge if the fiber length exceeds a critical value in order to maximize the contact area of the surface of the fiber.

For a reasonable parameterization of the chain for physiological conditions and sufficiently high attraction we find a force-plateau corresponding to a structural transition of the fiber in agreement with experiments [CB00]. In a first step the overall fiber geometry - a loop conformation -

is stretched out, followed by a decondensation of the fiber as the second step.

Up to now we studied exclusively fibers with a regular geometry. It would be interesting to introduce irregularities in the linker length and the rotational angle respectively and to study its consequences on the elastic properties and the behavior under stretching of the fiber.

Another interesting project is the formulation of a new coarse-grained model on the next level of DNA compaction. It is known that the chromatin fiber further organizes into the chromosome. Unfortunately there are no reliable experimental data on this length scale which could provide access to structural or elastic properties. Maybe our chromatin model can help to measure these elastic properties and to come up with a possible ground structure on this length scale.

Appendix A

Multiple histogram method

The multiple histogram method of Ferrenberg and Swendsen [FS88] is a convenient tool to study the free energy barrier of two conformational states or phases. Depending on applied external fields one can enforce the system to stay in either state. They developed a tool which makes it possible to recombine measured histograms of different field strengths with a reweighting procedure to a single probability distribution function with overall very good statistics.

Consider the case where the unperturbed state is characterized by the Hamiltonian $\mathcal{H}_0$. The application of an external force determines which of the local minima of the system are preferentially sampled by the simulation. Thus one performs several simulations of systems i characterized by the Hamiltonian $\mathcal{H}_i = \mathcal{H}_0 + \mathcal{V}_i$ where $\mathcal{V}_i$ can be a function of a order parameter Q. In the case discussed in section 2.9 the external potential is given by $\mathcal{V}_i = fR_E$ where f is the buckling or stretching force and R_E corresponds to the end-to-end distance. Higher probabilities (compared to the usual wormlike chain result) of small end-to-end distances suggest the occurrence of kinks. Thus R_E is sensitive to the underlying structure. The parameter f should enhance the formation of a kink-rod structure in case with an intrinsic preference for such structures. In the case of a canonical ensemble the probability distribution function $P_i(Q)$ of the order parameter Q for a given external field $\mathcal{V}_i$ is given by

$$P_i(Q) = \frac{\int d\mathbf{r}^N \delta\left(Q - Q(\mathbf{r}^N)\right) \exp(-\beta\mathcal{H}_0 + \mathcal{V}_i)}{Z_i} \tag{A.1}$$

with Z_i being the partition function:

$$Z_i = \int d\mathbf{r}^N \exp(-\beta\mathcal{H}_0 + \mathcal{V}_i). \tag{A.2}$$

The unperturbed system is characterized by

$$P_0(Q) = \frac{\int d\mathbf{r}^N \delta\left(Q - Q(\mathbf{r}^N)\right) \exp(-\beta\mathcal{H}_0)}{Z_0}. \tag{A.3}$$

In the following we will address the question how $P_0(Q)$ can be estimated by the knowledge of $P_i(Q)$ (see e.g. [FS96, LB00]).

In computer simulations the probability distribution function $P_i(Q)$ can be computed by histograms $H_i(Q)$ which measure how often the system with energy $\mathcal{H}_0 + \mathcal{V}_i$ is found to take a value between Q and $Q + \Delta Q$ such that

$$P_i(Q)\Delta Q = \frac{\langle H_i(Q)\rangle}{N_i} \tag{A.4}$$

where N_i denotes the number of saved simulation steps and $\langle H_i(Q)\rangle$ denotes the average over different simulation runs. Certainly $\frac{\langle H_i(Q)\rangle}{N_i}$ will fluctuate around $P_i(Q)\Delta Q$ due to the finite length of a simulation run.

It can be shown [FS96, LB00] that $P_0(Q)$ can be reconstructed from every individual histogram $P_i(Q)$ (as long as there is an overlap of the sampled distribution functions) by

$$P_0(Q) = \exp(\beta\mathcal{V}_i)\frac{\mathcal{Z}_i}{\mathcal{Z}_0}P_i(Q). \tag{A.5}$$

This leads to the following ansatz where one reconstructs $P_0^{est}(Q)$ by a linear combination of the measured histograms $P_i(Q)$ corresponding to different external fields $\mathcal{V}_i$ [FS96]:

$$P_0(Q) = \sum_{i=1}^{n}\omega_i(Q)\exp(\beta\mathcal{V}_i)\frac{\mathcal{Z}_i}{\mathcal{Z}_0}P_i(Q). \tag{A.6}$$

where $\omega_i(Q)$ is a normalized weight function with

$$\sum_{i=1}^{n}\omega_i(Q) = 1 \tag{A.7}$$

which has to be determined. The weights are chosen in such a way that the variance of $P_0^{est}(Q)$, i.e. $\langle P_0^{est}(Q)^2\rangle - \langle P_0^{est}(Q)\rangle^2$, is minimized under the constraint of eq. (A.7) resulting in

$$P_0^{est}(Q) = \frac{\sum_{i=1}^{n}H_i(Q)}{\sum_{i=1}^{n}\exp(-\beta\mathcal{V}_i)N_i\frac{\mathcal{Z}_0}{\mathcal{Z}_i}} \tag{A.8}$$

where n denotes the number of histograms that should be recombined (details of the calculation can be found in [FS96]). Inserting eqs. (A.8),(A.5) in eq. (A.2) yields an implicit equation for $\mathcal{Z}_i$ that must be solved self-consistently:

$$\mathcal{Z}_i = \int dQ\exp(-\beta\mathcal{V}_i)\frac{\sum_{j=1}^{n}H_j(Q)}{\sum_{k=1}^{n}\exp(-\beta\mathcal{V}_k)\frac{N_k}{\mathcal{Z}_k}}. \tag{A.9}$$

Actually it is not possible to determine all absolute values of $\mathcal{Z}_i$. Rather one is interested in calculating their ratio $\mathcal{Z}_i/\mathcal{Z}_0$ in order to obtain the estimate of the probability distribution function $P_0^{est}(Q)$ (see eq. (A.8)). Thus we can fix one of the $\mathcal{Z}_i$, say $\mathcal{Z}_0$, at a constant value and solve the set of eqs. (A.9)) until self-consistency is reached, that is until a new iteration step fulfills the condition

$$\sum_{i=1}^{n}|\mathcal{Z}_i^{it,new} - \mathcal{Z}_i^{it,old}| < \epsilon \tag{A.10}$$

where ϵ denotes the desired accuracy of the iteration procedure.

Appendix B

Calculation of the base-pair step parameters

The following considerations are based on the base-pair step calculation scheme developed by Babcock *et al* [BPO94].

The definition of the base-pair step parameters has to ensure that the calculated rotational and translational step parameters should be independent of the direction of measurement, i.e. independent of the chain end one starts with. Only the signs of the parameters should change if one changes the direction of measurement. It is known that Euler angles do not satisfy this condition. Simultaneous rotations of a rigid body, on the other hand, fullfil the desired requirements. In general a vector $\vec{v}$ which is rotated around an axis $\vec{\Omega}$ with angle Ω is given by

$$\vec{v}\,' = \vec{v}\cos(\Omega) + \left(\frac{\vec{\Omega}}{|\vec{\Omega}|} \times \vec{v}\right)\sin(\Omega) + \frac{\vec{\Omega}}{|\vec{\Omega}|}\left(\frac{\vec{\Omega}}{|\vec{\Omega}|} \cdot \vec{v}\right)(1 - \cos(\Omega))\,. \tag{B.1}$$

as already discussed in section 2.4. This produces an overall rotation of the body about the origin along the vector $\vec{\Omega} = \{\Omega_x, \Omega_y, \Omega_z\}$ with

$$\Omega = |\vec{\Omega}| = \sqrt{\mathrm{Ti}^2 + \mathrm{Ro}^2 + \mathrm{Tw}^2} \tag{B.2}$$

(see Fig. B.1). This can be written as a rotation matrix $\mathcal{R}$ [JJ56]

$$\mathcal{R} = \begin{pmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} & \mathcal{R}_{13} \\ \mathcal{R}_{21} & \mathcal{R}_{22} & \mathcal{R}_{23} \\ \mathcal{R}_{31} & \mathcal{R}_{32} & \mathcal{R}_{33} \end{pmatrix} \tag{B.3}$$

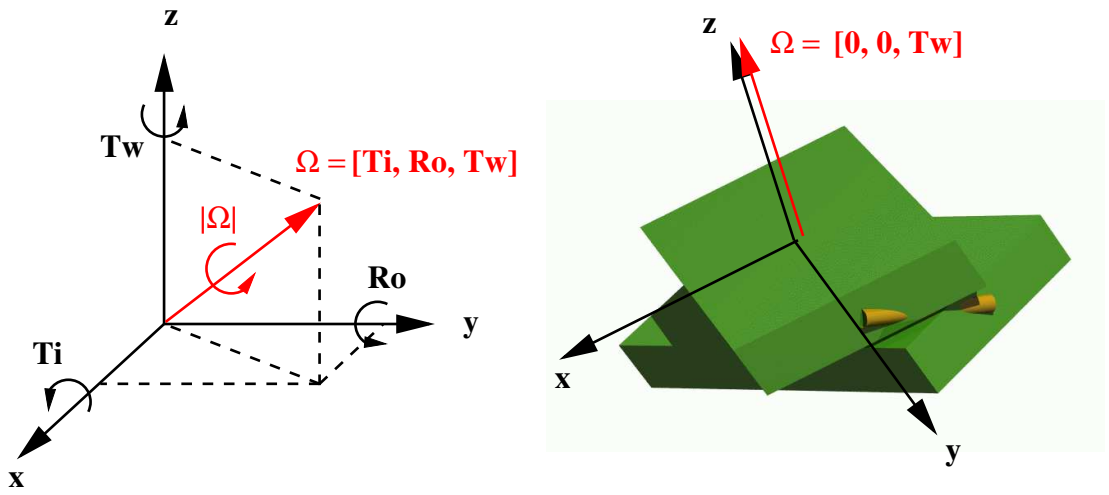


Figure B.1: Illustration of the rotation of a body about a vector $\vec{\Omega} = \{\Omega_x, \Omega_y, \Omega_z\}$.

with

$$
\begin{aligned}
\mathcal{R}_{11} &= \cos(\Omega) + (1-\cos(\Omega))\frac{\Omega_x^2}{\Omega^2} = \cos(\Omega) + (1-\cos(\Omega))\frac{\mathrm{Ti}^2}{\Omega^2} \\
\mathcal{R}_{12} &= (1-\cos(\Omega))\frac{\Omega_x\Omega_y}{\Omega^2} - \frac{\Omega_z}{\Omega}\sin(\Omega) = (1-\cos(\Omega))\frac{\mathrm{TiRo}}{\Omega^2} - \frac{\mathrm{Tw}}{\Omega}\sin(\Omega) \\
\mathcal{R}_{13} &= (1-\cos(\Omega))\frac{\Omega_x\Omega_z}{\Omega^2} + \frac{\Omega_y}{\Omega}\sin(\Omega) = (1-\cos(\Omega))\frac{\mathrm{TiTw}}{\Omega^2} + \frac{\mathrm{Ro}}{\Omega}\sin(\Omega) \\
\mathcal{R}_{21} &= (1-\cos(\Omega))\frac{\Omega_x\Omega_y}{\Omega^2} + \frac{\Omega_z}{\Omega}\sin(\Omega) = (1-\cos(\Omega))\frac{\mathrm{TiRo}}{\Omega^2} + \frac{\mathrm{Tw}}{\Omega}\sin(\Omega) \\
\mathcal{R}_{22} &= \cos(\Omega) + (1-\cos(\Omega))\frac{\Omega_y^2}{\Omega^2} = \cos(\Omega) + (1-\cos(\Omega))\frac{\mathrm{Ro}^2}{\Omega^2} \\
\mathcal{R}_{23} &= (1-\cos(\Omega))\frac{\Omega_y\Omega_z}{\Omega^2} - \frac{\Omega_x}{\Omega}\sin(\Omega) = (1-\cos(\Omega))\frac{\mathrm{RoTw}}{\Omega^2} - \frac{\mathrm{Ti}}{\Omega}\sin(\Omega) \\
\mathcal{R}_{31} &= (1-\cos(\Omega))\frac{\Omega_x\Omega_z}{\Omega^2} - \frac{\Omega_y}{\Omega}\sin(\Omega) = (1-\cos(\Omega))\frac{\mathrm{TiTw}}{\Omega^2} - \frac{\mathrm{Ro}}{\Omega}\sin(\Omega) \\
\mathcal{R}_{32} &= (1-\cos(\Omega))\frac{\Omega_y\Omega_z}{\Omega^2} + \frac{\Omega_x}{\Omega}\sin(\Omega) = (1-\cos(\Omega))\frac{\mathrm{RoTw}}{\Omega^2} + \frac{\mathrm{Ti}}{\Omega}\sin(\Omega) \\
\mathcal{R}_{33} &= \cos(\Omega) + (1-\cos(\Omega))\frac{\Omega_z^2}{\Omega^2} = \cos(\Omega) + (1-\cos(\Omega))\frac{\mathrm{Tw}^2}{\Omega^2}.
\end{aligned}
\tag{B.4}
$$

$\mathcal{R}$ is the rotation matrix which characterizes one base-pair step, i.e. it determines the rotation angles $\{\mathrm{Ti}, \mathrm{Ro}, \mathrm{Tw}\}$ which are involved to go from one base-pair to the neighboring one. Since only one rotation is involved changing the perspective from one coordinate frame to the other simply has the effect of reversing the direction of the rotation. For neighboring base-pair parameters changing the direction of measurement is equivalent to transposing the rotation matrix. Using Eq. (B.4) one obtains

$$\mathcal{R}(\mathrm{Ti}, \mathrm{Ro}, \mathrm{Tw})^T = \mathcal{R}(\text{-Ti}, \text{-Ro}, \text{-Tw}). \tag{B.5}$$

Thus changing the direction of measurement causes a change in signs of tilt, roll and twist while leaving their magnitudes unaltered.

The rotation angle Ω can be calculated by determining the trace of $\mathcal{R}$

$$tr(\mathcal{R}) = \sum_{i=1}^{3} \mathcal{R}_{ii} = 3\cos(\Omega) + (1 - \cos(\Omega))\frac{\mathrm{Ti}^2 + \mathrm{Ro}^2 + \mathrm{Tw}^2}{\Omega^2} = 2\cos(\Omega) + 1. \tag{B.6}$$

Thus one obtains for Ω

$$\cos(\Omega) = \frac{1}{2}(\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{33} - 1) \tag{B.7}$$

$$\Omega = \arccos\left(\frac{1}{2}(\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{33} - 1)\right). \tag{B.8}$$

Substituting Eq. B.7 into the diagonal terms of the rotation matrix $\mathcal{R}$ of Eq. B.4 results in a calculation scheme for $\{\mathrm{Ti}, \mathrm{Ro}, \mathrm{Tw}\}$:

$$|\mathrm{Ti}| = \Omega\sqrt{\frac{-\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{33} - 1}{\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{33} - 3}} \tag{B.9}$$

$$|\mathrm{Ro}| = \Omega\sqrt{\frac{\mathcal{R}_{11} - \mathcal{R}_{22} + \mathcal{R}_{33} - 1}{\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{33} - 3}} \tag{B.10}$$

$$|\mathrm{Tw}| = \Omega\sqrt{\frac{\mathcal{R}_{11} + \mathcal{R}_{22} - \mathcal{R}_{33} - 1}{\mathcal{R}_{11} + \mathcal{R}_{22} + \mathcal{R}_{33} - 3}}. \tag{B.11}$$

With Eqs. (B.8-B.11) the values of the rotational base-pair parameters are completely determined. In order to identify the signs of $\{\mathrm{Ti}, \mathrm{Ro}, \mathrm{Tw}\}$ one can make use of the signs and relative magnitudes of the off-diagonal elements of the rotation matrix $\mathcal{R}$. Eq. (B.8) restricts the value of Ω to lie in the interval of $[0, \pi]$. Thus $\sin(\Omega)$ must be non-negative. As a consequence one finds that $\mathrm{Ti} \geq 0$ in case of $\mathcal{R}_{32} \geq \mathcal{R}_{23}$. Otherwise Ti is negative. Furthermore if Ti is positive (negative) the sign of Ro is the same (opposite) as the sign of $\mathcal{R}_{21} + \mathcal{R}_{12}$ and the sign of Tw is the same (opposite) as the sign of $\mathcal{R}_{31} + \mathcal{R}_{13}$. These considerations lead to the following rules:

$$\begin{aligned}
&\text{if } \mathcal{R}_{32} \geq \mathcal{R}_{23} \text{ then} \\
&\quad \mathrm{Ti} = +|\mathrm{Ti}| \\
&\text{if } \mathcal{R}_{21} + \mathcal{R}_{12} \geq 0 \text{ then } \mathrm{Ro} = +|\mathrm{Ro}| \text{ else } \mathrm{Ro} = -|\mathrm{Ro}| \\
&\text{if } \mathcal{R}_{31} + \mathcal{R}_{13} \geq 0 \text{ then } \mathrm{Tw} = +|\mathrm{Tw}| \text{ else } \mathrm{Tw} = -|\mathrm{Tw}| \\
&\\
&\text{if } \mathcal{R}_{32} < \mathcal{R}_{23} \text{ then} \\
&\quad \mathrm{Ti} = -|\mathrm{Ti}| \\
&\text{if } \mathcal{R}_{21} + \mathcal{R}_{12} \geq 0 \text{ then } \mathrm{Ro} = -|\mathrm{Ro}| \text{ else } \mathrm{Ro} = +|\mathrm{Ro}| \\
&\text{if } \mathcal{R}_{31} + \mathcal{R}_{13} \geq 0 \text{ then } \mathrm{Tw} = -|\mathrm{Tw}| \text{ else } \mathrm{Tw} = +|\mathrm{Tw}|
\end{aligned} \tag{B.12}$$

For reasons of numerical stability these rules are only employed if $|\mathrm{Ti}| \geq |\mathrm{Ro}|$ and $|\mathrm{Ti}| \geq |\mathrm{Tw}|$. Other rules are used if $|\mathrm{Ro}|$ or $|\mathrm{Tw}|$ are largest in magnitude. For these two remaining cases one can

derive analogous sign rules [BPO94]: (i) $|\text{Ro}| \geq |\text{Ti}|$ and $|\text{Ro}| \geq |\text{Tw}|$:

$$\begin{aligned}
&\text{if } \mathcal{R}_{13} \geq \mathcal{R}_{31} \text{ then} \\
&\quad \text{Ro} = +|\text{Ro}| \\
&\text{if } \mathcal{R}_{21} + \mathcal{R}_{12} \geq 0 \text{ then Ti} = +|\text{Ti}| \text{ else Ti} = -|\text{Ti}| \\
&\text{if } \mathcal{R}_{32} + \mathcal{R}_{23} \geq 0 \text{ then Tw} = +|\text{Tw}| \text{ else Tw} = -|\text{Tw}| \\
\\
&\text{if } \mathcal{R}_{13} < \mathcal{R}_{31} \text{ then} \\
&\quad \text{Ro} = -|\text{Ro}| \\
&\text{if } \mathcal{R}_{21} + \mathcal{R}_{12} \geq 0 \text{ then Ti} = -|\text{Ti}| \text{ else Ti} = +|\text{Ti}| \\
&\text{if } \mathcal{R}_{32} + \mathcal{R}_{23} \geq 0 \text{ then Tw} = -|\text{Tw}| \text{ else Tw} = +|\text{Tw}|
\end{aligned} \tag{B.13}$$

and (ii) $|\text{Tw}| \geq |\text{Ti}|$ and $|\text{Tw}| \geq |\text{Ro}|$:

$$\begin{aligned}
&\text{if } \mathcal{R}_{21} \geq \mathcal{R}_{12} \text{ then} \\
&\quad \text{Tw} = +|\text{Tw}| \\
&\text{if } \mathcal{R}_{31} + \mathcal{R}_{13} \geq 0 \text{ then Ti} = +|\text{Ti}| \text{ else Ti} = -|\text{Ti}| \\
&\text{if } \mathcal{R}_{32} + \mathcal{R}_{23} \geq 0 \text{ then Ro} = +|\text{Ro}| \text{ else Ro} = -|\text{Ro}| \\
\\
&\text{if } \mathcal{R}_{21} < \mathcal{R}_{12} \text{ then} \\
&\quad \text{Tw} = -|\text{Tw}| \\
&\text{if } \mathcal{R}_{31} + \mathcal{R}_{13} \geq 0 \text{ then Ti} = -|\text{Ti}| \text{ else Ti} = +|\text{Ti}| \\
&\text{if } \mathcal{R}_{32} + \mathcal{R}_{23} \geq 0 \text{ then Ro} = -|\text{Ro}| \text{ else Ro} = +|\text{Ro}|.
\end{aligned} \tag{B.14}$$

Equations (B.8-B.14) provide the calculation scheme for the rotational base-pair step parameters.

With the knowledge of the rotational step parameters it is now possible to calculate the translational step parameters along the x-, y- and z- direction (shift, slide, rise). This vector is given by

$$\vec{\mathcal{T}}_{i,i+1} = \begin{pmatrix} \text{Sh} \\ \text{Sl} \\ \text{Ri} \end{pmatrix} = \mathcal{R}^{-\frac{1}{2}} \vec{r}_{i,i+1} \tag{B.15}$$

where the entries of $\vec{r}_{i,i+1}$ are the coordinates of the origin of frame $i+1$ as measured in frame i. The rotation matrix $\mathcal{R}^{-\frac{1}{2}}$ ensures that the translational step parameters are measured in a midway coordinate frame such that the magnitudes of $\{\text{Sh}, \text{Sl}, \text{Ri}\}$ are independent of the direction of measurement. The effect of measuring in the reverse direction is again a change in sign of $\vec{\mathcal{T}}_{i,i+1}$

$$\vec{\mathcal{T}}_{i,i+1} = -\vec{\mathcal{T}}_{i+1,i}. \tag{B.16}$$

In order to calculate the square-root of the rotation matrix $\mathcal{R}$ one can use the following general property of $\mathcal{R}$ of Eq. (B.4). For all values of k and m one can show [BPO94] that

$$\mathcal{R}(k\text{Ti}, k\text{Ro}, k\text{Tw})\mathcal{R}(m\text{Ti}, m\text{Ro}, m\text{Tw}) = \mathcal{R}((k+m)\text{Ti}, (k+m)\text{Ro}, (k+m)\text{Tw}) \tag{B.17}$$

Thus square-roots can be calculated by setting $k = m = \frac{1}{2}$. This implicates that we use the calculated values of $\{\text{Ti}, \text{Ro}, \text{Tw}\}$ divide them by two and make use of eqs (B.2) to (B.4) to generate the square-root matrix.

Appendix C

Relationship between twist/bending angles and the orientation of the base-pairs

Consider the following three systems $\mathcal{S}$, $\mathcal{S}'$, $\mathcal{S}''$, where system $\mathcal{S}'$ is randomly rotated with respect to $\mathcal{S}$. To find for example the twist angle τ between $\mathcal{S}$ and $\mathcal{S}'$ we rotate $\mathcal{S}'$ about a unit vector $\mathbf{z}$ with

$$\mathbf{z} = \frac{\mathbf{t} \times \mathbf{t}'}{|\mathbf{t} \times \mathbf{t}'|} = \frac{\mathbf{t} \times \mathbf{t}'}{\sin(\theta)} \tag{C.1}$$

which is perpendicular to the $\mathbf{t}$-$\mathbf{t}'$-plane. This undoes the bending and results in the $\mathcal{S}''$ frame (see Fig. C.1). The vector $\mathbf{t}''$ is thus parallel to $\mathbf{t}$ and the angle τ between $\mathbf{n}''$ and $\mathbf{n}$ and $\mathbf{b}''$ and $\mathbf{b}$

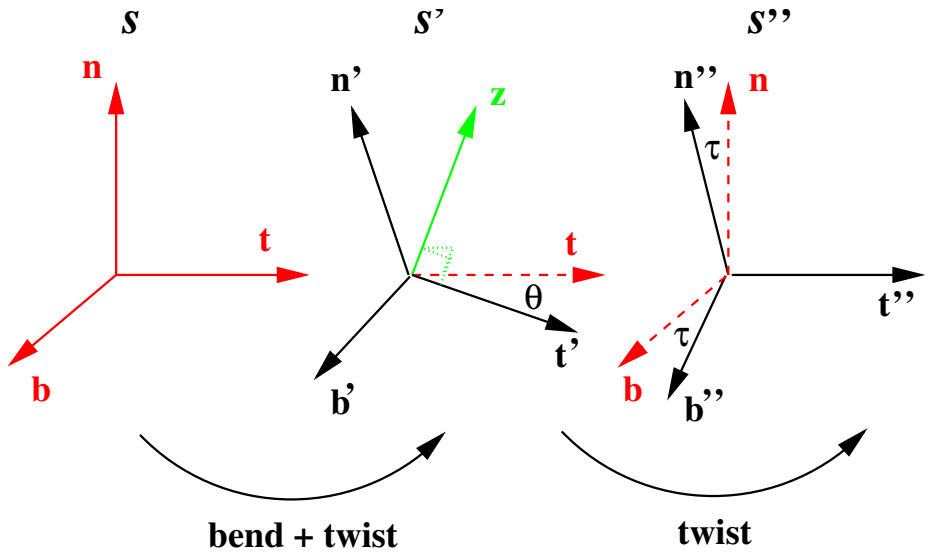


Figure C.1: Bending and twist angles in two rotated frames $\mathcal{S}$ and $\mathcal{S}'$.

respectively corresponds to the twist angle Tw, such that

$$\cos(\tau) = \cos(\mathrm{Tw}) = \mathbf{n}'' \cdot \mathbf{n} = \mathbf{b}'' \cdot \mathbf{b}. \tag{C.2}$$

Furthermore we can use Eq. (2.41) to calculate $\mathbf{n}''$ and $\mathbf{b}''$:

$$\mathbf{n}'' = \mathbf{n}' \cos(\theta) + \sin(\theta)\mathbf{z} \times \mathbf{n}' + (1 - \cos(\theta))(\mathbf{n}' \cdot \mathbf{z})\mathbf{z} \tag{C.3}$$
$$\mathbf{b}'' = \mathbf{b}' \cos(\theta) + \sin(\theta)\mathbf{z} \times \mathbf{b}' + (1 - \cos(\theta))(\mathbf{b}' \cdot \mathbf{z})\mathbf{z}. \tag{C.4}$$

Combining Eq. (C.2) and (C.3) we get

$$\cos(\mathrm{Tw}) = (\mathbf{n} \cdot \mathbf{n}') \cos(\theta) + \sin(\theta)\mathbf{n} \cdot (\mathbf{z} \times \mathbf{n}') + (1 - \cos(\theta))(\mathbf{n}' \cdot \mathbf{z})(\mathbf{n} \cdot \mathbf{z}). \tag{C.5}$$

We can make use of Eq. (C.1) to calculate the scalar and vector products of Eq. (C.5):

$$\mathbf{n}' \cdot \mathbf{z} = \frac{1}{\sin(\theta)}(\mathbf{t} \times \mathbf{t}') \cdot \mathbf{n}' = \frac{1}{\sin(\theta)}(\mathbf{n}' \times \mathbf{t}') \cdot \mathbf{t} = \frac{1}{\sin(\theta)}\mathbf{b}' \cdot \mathbf{t} \tag{C.6}$$

$$\mathbf{n} \cdot \mathbf{z} = \frac{1}{\sin(\theta)}(\mathbf{t} \times \mathbf{t}') \cdot \mathbf{n} = \frac{1}{\sin(\theta)}(\mathbf{n} \times \mathbf{t}) \cdot \mathbf{t}' = -\frac{1}{\sin(\theta)}\mathbf{b} \cdot \mathbf{t}' \tag{C.7}$$

$$\mathbf{z} \times \mathbf{n}' = \frac{1}{\sin(\theta)}(\mathbf{t} \times \mathbf{t}') \times \mathbf{n}' = \frac{1}{\sin(\theta)}(\mathbf{t}'(\mathbf{n}' \cdot \mathbf{t}) - \mathbf{t}(\mathbf{n}' \cdot \mathbf{t}')) = \frac{1}{\sin(\theta)}\mathbf{t}'(\mathbf{n}' \cdot \mathbf{t}). \tag{C.8}$$

Substituting these equations into Eq. (C.5) results in

$$\begin{aligned}
\cos(\mathrm{Tw}) &= (\mathbf{n} \cdot \mathbf{n}') \cos(\theta) - (\mathbf{n} \cdot \mathbf{t}')(\mathbf{n}' \cdot \mathbf{t}) - \frac{1 - \cos(\theta)}{\sin(\theta)^2}(\mathbf{b} \cdot \mathbf{t}')(\mathbf{b}' \cdot \mathbf{t}) \\
&= (\mathbf{n} \cdot \mathbf{n}')(\mathbf{t} \cdot \mathbf{t}') - (\mathbf{n} \cdot \mathbf{t}')(\mathbf{n}' \cdot \mathbf{t}) - \frac{1 - \mathbf{t} \cdot \mathbf{t}'}{(\mathbf{t} \times \mathbf{t}') \cdot (\mathbf{t} \times \mathbf{t}')}(\mathbf{b} \cdot \mathbf{t}')(\mathbf{b}' \cdot \mathbf{t}) \\
&= (\mathbf{n} \cdot \mathbf{n}')(\mathbf{t} \cdot \mathbf{t}') - (\mathbf{n} \cdot \mathbf{t}')(\mathbf{n}' \cdot \mathbf{t}) - \frac{1 - \mathbf{t} \cdot \mathbf{t}'}{(\mathbf{t} \cdot \mathbf{t})(\mathbf{t}' \cdot \mathbf{t}') - (\mathbf{t} \cdot \mathbf{t}')^2}(\mathbf{b} \cdot \mathbf{t}')(\mathbf{b}' \cdot \mathbf{t}) \\
&= (\mathbf{n} \cdot \mathbf{n}')(\mathbf{t} \cdot \mathbf{t}') - (\mathbf{n} \cdot \mathbf{t}')(\mathbf{n}' \cdot \mathbf{t}) - \frac{(\mathbf{b} \cdot \mathbf{t}')(\mathbf{b}' \cdot \mathbf{t})}{1 + (\mathbf{t} \cdot \mathbf{t}')}.
\end{aligned} \tag{C.9}$$

Using that

$$\begin{aligned}
(\mathbf{n} \cdot \mathbf{n}') &= (\mathbf{b} \times \mathbf{t}) \cdot (\mathbf{b}' \times \mathbf{t}') \\
&= (\mathbf{b}' \cdot \mathbf{b})(\mathbf{t}' \cdot \mathbf{t}) - (\mathbf{b}' \cdot \mathbf{t})(\mathbf{t}' \cdot \mathbf{b})
\end{aligned} \tag{C.10}$$

$$\begin{aligned}
(\mathbf{b} \cdot \mathbf{b}') &= (\mathbf{t} \times \mathbf{n}) \cdot (\mathbf{t}' \times \mathbf{n}') \\
&= (\mathbf{t}' \cdot \mathbf{t})(\mathbf{n}' \cdot \mathbf{n}) - (\mathbf{n}' \cdot \mathbf{t})(\mathbf{t}' \cdot \mathbf{n})
\end{aligned} \tag{C.11}$$

and

$$(\mathbf{b}' \cdot \mathbf{t})(\mathbf{t}' \cdot \mathbf{b}) = (\mathbf{b}' \cdot \mathbf{b})(\mathbf{t}' \cdot \mathbf{t}) - (\mathbf{n} \cdot \mathbf{n}') \tag{C.12}$$

$$(\mathbf{n}' \cdot \mathbf{t})(\mathbf{t}' \cdot \mathbf{n}) = (\mathbf{t}' \cdot \mathbf{t})(\mathbf{n}' \cdot \mathbf{n}) - (\mathbf{b} \cdot \mathbf{b}') \tag{C.13}$$

respectively one can rewrite Eq. (C.9) obtaining the following relationship between the twist angle and the base vectors of frame $\mathcal{S}$ and $\mathcal{S}'$:

$$\cos(\mathrm{Tw}) = \frac{\mathbf{n} \cdot \mathbf{n}' + \mathbf{b} \cdot \mathbf{b}'}{1 + \mathbf{t} \cdot \mathbf{t}'}. \tag{C.14}$$

The respective bending angles Ro (around $\mathbf{n}$-axis, compare Fig. 3.2) and Ti (around $\mathbf{b}$-axis) can be calculated in an analogous fashion. One finds that they are given by permutations of $\mathbf{t}$, $\mathbf{n}$, and $\mathbf{b}$ in Eq. (C.14):

$$\cos(\mathrm{Ti}) = \frac{\mathbf{n}\cdot\mathbf{n}' + \mathbf{t}\cdot\mathbf{t}'}{1+\mathbf{b}\cdot\mathbf{b}'} \tag{C.15}$$

$$\cos(\mathrm{Ro}) = \frac{\mathbf{t}\cdot\mathbf{t}' + \mathbf{b}\cdot\mathbf{b}'}{1+\mathbf{n}\cdot\mathbf{n}'}. \tag{C.16}$$

Bibliography

[ABL+94] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson. *Molecular Biology of the Cell*. Garland Publishing, Inc., New York and London, 1994.

[ABLC98] J.F. Allemand, D. Bensimon, R. Lavery, and V. Croquette. Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Natl. Acad. Sci. USA*, 95:14152–14157, 1998.

[ALCM01] A.Sarkar, J. F. Leger, D. Chatenay, and J. F. Marko. Structural transitions in DNA driven by external force and torque. *Phys. Rev. E*, 63:051903, 2001.

[BBS03] C. Bustamante, Z. Bryant, and S.B. Smith. Ten years of tension: single-molecule DNA mechanics. *Nature*, 421:423–427, 2003.

[BCP99] M. Barbi, S. Cocco, and M. Peyrard. Helicoidal model for DNA opening. *Physics Letters A*, 253:358–369, 1999.

[BDM98] D. Bensimon, D. Dohmi, and M. Mezard. Stretching a heteropolymer. *Europhys. Lett.*, 42(1):97–102, 1998.

[BERH97] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot. Molecular Stick-Slip Motion Revealed by Opening DNA with Piconewton Forces. *Phys. Rev. Lett.*, 79:4489–4492, 1997.

[BERH98] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot. DNA strand separation studied by single molecule force measurements. *Phys. Rev. E*, 58(2):2386–2394, 1998.

[BFK+95] J. Bednar, P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, and A. Stasiak. Determination of DNA Persistence Length by Cryo-electron Microscopy. Separation of the Static and Dynamic Contributions to the Apparent Persistence Length of DNA. *J. Mol. Biol.*, 254:579–594, 1995.

[BFLG99] N. Bruant, D. Flatters, R. Lavery, and D. Genest. From Atomic to Mesoscopic Descriptions of the Internal Dynamics of DNA. *Biophys. J.*, 77:2366–2376, 1999.

[BFZ98] R. Berardi, C. Fava, and C. Zannoni. A gay-berne potential for dissimilar biaxial particles. *Chem. Phys. Lett.*, 297:8–14, 1998.

[BHDW98] J. Bednar, R. A Horowitz, J. Dubochet, and C. L. Woodcock. Chromatin Conformation and Salt-induced Compaction: Three-dimensional Structural Information from Cryoelectron Microscopy. *J. Cell. Biol.*, 131:1365–1376, 1998.

[BHG+98] J. Bednar, R. A. Horowitz, S. A. Grigoryev, L. M. Carruthers, J. C. Hansen, A. J. Koster, and C. L. Woodcock. Nucleosomes, Linker DNA, and Linker Histones form a unique Structural Motif that Directs the higher-Order Folding and Compaction of Chromatin. *Proc. Natl. Acad. Sci. USA*, 95(24):14173–14178, 1998.

[BHLV01] E. Ben-Haim, A. Lesne, and J.-M. Victor. Chromatin: A tunable spring at work inside chromosomes. *Phys. Rev. E*, 64:051921, 2001.

[BHLV02] E. Ben-Haim, A. Lesne, and J.-M. Victor. Adaptive elastic properties of chromatin fiber. *Physica A*, 314:592–599, 2002.

[BLL+01] M. L. Bennink, S. H. Leuba, G. H. Leno, J. Zlatanova, B. G. de Grooth, and J. Greve. Unfolding individual nucleosomes by stretching single chromatin fibers with optical tweezers. *Nature Structural Biology*, 8(7):606–610, 2001.

[BMSS94] C. Bustamante, J. F. Marko, E.D. Siggia, and S. Smith. Entropic Elasticity of lambda-Phage DNA. *Science*, 265:1599–1600, 1994.

[BO94] M. S. Babcock and W. K. Olson. The Effect of Mathematics and Coordinate System on Comparability and 'Dependencies' of Nucleic Acid Structure Parameters. *J. Mol. Biol.*, 237:98–124, 1994.

[BPL+01] M. L. Bennink, L. H. Pope, S. H. Leuba, B. G. de Grooth, and J. Greve. Single Chromatin Fibre Assembly Using Optical Tweezers. *Single Mol.*, 2(2):91–97, 2001.

[BPO94] M. S. Babcock, E. P. D. Pednault, and W. K. Olson. Nucleic Acid Structure Analysis: Mathematics for Local Cartesian and Helical Structure Parameters That Are Truly Comparable Between Structures. *J. Mol. Biol.*, 237:125–156, 1994.

[BS01] D. A. Beard and T. Schlick. Computational Modeling Predicts the Structure and Dynamics of Chromatin Fiber. *Structure*, 9:105–114, 2001.

[BSLS00] C. Bustamante, S. B. Smith, J. Liphardt, and D. Smith. Single-molecule studies of DNA mechanics. *Current Opinion in Structural Biology*, 10:279–285, 2000.

[BTSY+02] B. D. Brower-Toland, C. L. Smith, R. C. Yeh, J. T. Lis, C. L. Peterson, and M. D. Wang. Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proc. Natl. Acad. Sci. USA*, 99(4):1960–1965, 2002.

[BWC90] T. C. Boles, J. H. White, and N. R. Cozzarelli. Structure of plectonemically supercoiled DNA. *J. Mol. Biol.*, 213:931–951, 1990.

[BZ79] M. D. Barkley and B. H. Zimm. Theory of twisting and bending of chain macromolecules; analysis of the fluorescence depolarization of DNA. *J. Chem. Phys.*, 70(6):2991–3007, 1979.

[CB00] Y. Cui and C. Bustamante. Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proc. Natl. Acad. Sci. USA*, 97(1):127–132, 2000.

[CD84] C. R. Calladine and H. R. Drew. A base-centered explanation of the B-to-A transition in DNA. *J. Mol. Biol.*, 178:773–782, 1984.

[CD99] C. R. Calladine and H. R. Drew. *Understanding DNA: The molecule and how it works*. Academic Press, 1999.

[CL94] G. Chirico and J. Langowski. Kinetics of DNA Supercoiling Studied by Brownian Dynamics Simulation. *Biopolymers*, 34:415–434, 1994.

[CL96] G. Chirico and J. Langowski. Brownian Dynamics Simulations of Supercoiled DNA with Bent Sequences. *Biophys. J.*, 71:955–971, 1996.

[CLH+96] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron. DNA: An Extensible Molecule. *Science*, 264:792–794, 1996.

[CM99] S. Cocco and R. Monasson. Statistical mechanics of torque induced denaturation of DNA. *Phys. Rev. Lett.*, 83:5178–5181, 1999.

[CSRTG00] H. Clausen-Schaumann, M. Rief, C. Tolksdorf, and H. E. Gaub. Mechanical Stability of Single DNA Molecules. *Biophys. J.*, 78:1997–2007, 2000.

[CW90] N. R. Cozzarelli and J. C. Wang. *DNA Topology and Its Biological Effects*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, 1990.

[DBC+89] R. E. Dickerson, M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, R. Lavery, H. C. Nelson, W. K. Olson, W. Saengar, Z. Shakked, H. Sklenar, D.M. Soumpasis, C.-S. Tung, E. von Kitzing, A. H.-C. Wang, and V. B. Zhurkin. Definitions and nomenclature of nucleic acid structure parameters. *EMBO Journal*, 8:1–4, 1989.

[DDC+82] R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini, and M. L. Kopka. The anatomy of A-, B- and Z-DNA. *Science*, 216:475–485, 1982.

[DE86] M. Doi and S.F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, 1986.

[Dic92] R. E. Dickerson. DNA structure from A to Z. *Methods in Enzymology*, 211:67–110, 1992.

[DMK99] N. A. Davis, S. S. Majee, and J. D. Kahn. TATA Box DNA Deformation with and without the TATA Box-binding Protein. *J. Mol. Biol.*, 291(2):249–265, 1999.

[EBK95] R. Everaers, R. Bundschuh, and K. Kremer. Fluctuations and Stiffness of Double-Stranded Polymers: Railway-Track Model. *Europhys. Lett.*, 29:263–268, 1995.

[EE03] R. Everaers and M. R. Ejtehadi. Interaction potentials of soft and hard ellipsoids. *Phys. Rev. E*, 67:041710, 2003.

[EHL97] M. A. Elhassan, C. A. Hunter, and X. J. Lu. Structure and Conformation of Helical Nucleic Acids: Analysis Programme (SCHNAaP). *J. Mol. Biol.*, 273:668–680, 1997.

[ERBH97] B. Essevaz-Roulet, U. Bockelmann, and F. Heslot. Mechanical separation of the complementary strands of DNA. *Proc. Natl. Acad. Sci. USA*, 94:11935–11940, 1997.

[FK73] M. Fixman and J. Kovac. Polymer conformational statistics. III. Modified Gaussian models of stiff chains. *J. Chem. Phys.*, 58(4):1564–1568, 1973.

[FK76] J. T. Finch and A. Klug. Solenoidal Model for Superstructure in Chromatin. *Proc. Natl. Acad. Sci. USA*, 73(6):1897–1901, 1976.

[FK90] M. D. Frank-Kamenetskii. Physical data for nucleic acids and their constituents. In W. Saenger, editor, *Numerical Data and Functional Relationships in Science and Technology, Subvolume c: Spectroscopic and Kinetic Data. Physical Data I*, VII/1c. Springer, Berlin, 1990.

[FMM00] P. B. Furrer, R. S. Manning, and J. H. Maddocks. DNA Rings with Multiple Energy Minima. *Biophys. J.*, 79:116–136, 2000.

[FS88] A. M. Ferrenberg and R. H. Swendsen. Optimized Monte Carlo Data Analysis. *Phys. Rev. Lett.*, 63:1195–1198, 1988.

[FS96] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, 1996.

[FSH99] Y. Fang, T. S. Spisz, and J. H. Hoh. Ethanol-induced structural transitions of DNA on mica. *Nucleic Acids Research*, 27:1943–1949, 1999.

[GAL] NDB ATLAS GALLERY. http://beta-ndb.rutgers.edu/atlas/xray/indexes/xray.b.dna.html.

[GB81] J. G. Gay and B. J. Berne. Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.*, 74(6):3316–3319, 1981.

[GB99] C. F. Guerra and F. M. Bickelhaupt. Charge transfer and environment effects responsible for characteristics of dna base pairing. *Angewandte Chemie-International Edition*, 38(19):2942–2945, 1999.

[GL00] R. Golestanian and T. B. Liverpool. Statistical mechanics of semiflexible ribbon polymers. *Phys. Rev. E*, 62(4):5488–5499, 2000.

[Gol91] H. Goldstein. *Klassische Mechanik*. Wiesbaden : Aula-Verlag, 1991.

[GR87] S. E. Gerchman and V. Ramakrishhnan. Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. *Proc. Natl. Acad. Sci. USA*, 84:7802–7806, 1987.

[Hal] R. B. Hallick. http://www.blc.arizona.edu/Molecular_Graphics/DNA_Structure/DNA_Tutorial.html.

[HC95] M. A. El Hassan and C. R. Calladine. The assessment of the geometry of dinucleotide steps in double-helical DNA; a new local calculation scheme. *J. Mol. Biol.*, 251:648–664, 1995.

[HC97a] M. A. El Hassan and C. R. Calladine. Conformational characteristics of DNA: empirical classifications and a hypothesis for the conformational behaviour of dinucleotide steps. *Phil. Trans. R. Soc. Lond. A*, 355:43–100, 1997.

[HC97b] M. A. El Hassan and C. R. Calladine. Curvature and torsion of infinitely repeating DNA sequences: a geometric investigation. *Proc. R. Soc. Lond. A*, 453:365–386, 1997.

[HCFS96] P. J. Heath, J. B. Clendenning, B. S. Fujimoto, and J. M. Schurr. Effect of Bending Strain on the Torsion Elastic Constant of DNA. *J. Mol. Biol.*, 260:718–730, 1996.

[HH66] R.A. Harris and J.E. Hearst. On Polymer Dynamics. *J. Chem. Phys.*, 44(7):2595–2602, 1966.

[HL97] C. A. Hunter and X.-J. Lu. DNA Base-stacking Interactions: A Comparison of Theoretical Calculations with Oligonucleotide X-ray Crystal Structures. *J. Mol. Biol.*, 265:603–619, 1997.

[HMIWW00] D. Hamelberg, L. McFail-Isom, L. D. Williams, and W. D. Wilson. Flexible structure of DNA: Ion dependence of minor-groove structure and dynamics. *J. Am. Chem. Soc.*, 122(43):10513–10520, 2000.

[HSH03] B. Hartmann, M. R. Sullivan, and L. F. Harris. Operator recognition by the phage 434 cl repressor: MD simulations of free and bound 50-bp DNA reveal important differences between the OR1 and OR2 sites. *Biopolymers*, 68(2):250–264, 2003.

[Hun93] C. A. Hunter. Sequence-dependent DNA Structure: The Role of Base Stacking Interactions. *J. Mol. Biol.*, 230:1025–1054, 1993.

[HWW02] D. Hamelberg, L. D. Williams, and W. D. Wilson. Effect of a neutralized phosphate backbone on the minor groove of B-DNA: molecular dynamics simulation studies. *Nucl. Acids. Res.*, 30(16):3615–3623, 2002.

[HYZc99] Z. Haijun, Z. Yang, and O.-Y. Zhong-can. Bending and Base-Stacking Interactions in Double-Stranded DNA. *Phys. Rev. Lett.*, 82(22):4560–4563, 1999.

[HZ95] K. Van Holde and J. Zlatanova. Chromatin Higher Order Structure: Chasing a Mirage? *J. Biol. Chem.*, 270(15):8373–8376, 1995.

[HZ96] K. Van Holde and J. Zlatanova. What determines the Folding of the Chromatin Fiber? *Proc. Natl. Acad. Sci. USA*, 93(20):10548–10555, 1996.

[Jam95] T. L. James. Nuclear Magnetic Resonance and Nucleic Acids. *Methods in Enzymology*, 261:1–640, 1995.

[JCL+96] Z. S. Juo, T. K. Chiu, P. M. Leiberman, I. Baikalov, A. J. Berk, and R. E. Dickerson. How proteins recognize the TATA box. *J. Mol. Biol.*, 261:239–254, 1996.

[JJ56] S. H. Jeffreys and B. Jeffreys. *Methods of Mathematical Physics*. Cambridge University Press, 1956.

[K. 89] K. E. van Holde. *Chromatin*. Springer, New York, 1989.

[KBO00] V. Katritch, C. Bustamante, and W. K. Olson. Pulling Chromatin Fibers: Computer Simulations of Direct Physical Micromanipulations. *J. Mol. Biol.*, 295(29-40), 2000.

[KG92] D. M. Kroll and G. Gompper. Scaling Behavior of Randomly Tiangulated Self-Avoiding Surfaces. *Phys. Rev. A*, 46(6):3119–3122, 1992.

[KLNO97] R. D. Kamien, T. C. Lubensky, P. Nelson, and C. S. O'Hern. Direct determination of DNA twist-stretch coupling. *Europhys. Lett.*, 38(3):237–242, 1997.

[KM97] S. Kehrbaum and J. H. Maddocks. Elastic rods, rigid bodies, quarternions and the last quadrature. *Philosophical Transactions Royal Society of London*, 355:2117–2136, 1997.

[KML97] K. Klenin, H. Merlitz, and J. Langowski. A Brownian Dynamics Program for the Simulation of Linear and Circular DNA and other Wormlike Chain Polyelectrolytes. *Biophys. J.*, 74:780–788, 1997.

[KSBS93] J. Käs, H. Strey, M. Bärmann, and E. Sackmann. Direct Measurement of the Wave-Vector-Dependent Bending Stiffness of Freely Flickering Actin-Filaments. *Europhys. Lett.*, 21:865–870, 1993.

[LB00] D. P. Landau and K. Binder. *Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, 2000.

[LGK98] T. B. Liverpool, R. Golestanian, and K. Kremer. Statistical Mechanics of Double-Stranded Semiflexible Polymers. *Phys. Rev. Lett.*, 80:405–408, 1998.

[LL97] A. Leforestier and F. Livolant. Liquid crystalline ordering of nucleosome core particles under macromolecular crowding conditions: evidence for a discotic columnar hexagonal phase. *Biophys. J.*, 73:1771–1776, 1997.

[LL99] R. Lavery and A. Lebrun. Modelling DNA stretching for physics and biology. *Genetica*, 106:75–84, 1999.

[LL00a] I. Lafontaine and R. Lavery. Optimization of nucleic acid sequences. *Biophys. J.*, 79:680–685, 2000.

[LL00b] F. Livolant and A. Leforestier. Chiral Discotic Columnar Germs of Nucleosome Core Particles. *Biophys. J.*, 78:2716–2729, 2000.

[LLA+02] R. Lavery, A. Lebrun, J.-F. Allemand, D. Bensimon, and V. Croquette. Structure and mechanics of single biomolecules: experiment and simulation. *J. Phys.: Condens. Matter*, 14:R383–R414, 2002.

[LMR+97] K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389:251–260, 1997.

[LO99] X. J. Lu and W. K. Olson. Resolving the discrepancies among nucleic acid conformational analysis. *J. Mol. Biol.*, 285:1563–1575, 1999.

[LS88] R. Lavery and H. Sklenar. The Definition of Generalised Helicoidal Parameters and of Axis Curvature for Irregular Nucleic Acids. *J. Biomol. Struct. Dynam.*, 6:63–91, 1988.

[LSHL00] F. Lankas, J. Sponer, P. Hobza, and J. Langowski. Sequence-dependent Elastic Properties of DNA. *J. Mol. Biol.*, 299:695–709, 2000.

[Mad00] J. H. Maddocks. The DNA modelling course Home Page. http://lcvmwww.epfl.ch/, 2000.

[Mag85] A. C. Maggs. *Critical dynamics near the percolation threshold.* PhD thesis, University of Oxford, 1985.

[Mag01] A. C. Maggs. Writhing geometry at finite temperature: Random walks and geometric phases for stiff polymers. *J. Chem. Phys.*, 114(13):5888–5896, 2001.

[Mar97] J. F. Marko. Stretching must twist DNA. *Europhys. Lett.*, 38(3):183–188, 1997.

[Mar98] J. H. Marko. DNA under high tension: Overstretching, undertwisting, and relaxation dynamics. *Phys. Rev. E*, 57(2):2134–2149, 1998.

[MEE02] B. Mergell, M. R. Ejtehadi, and R. Everaers. Statistical mechanics of triangulated ribbons. *Phys. Rev. E*, 66:011903, 2002.

[MEE03] B. Mergell, M. R. Ejtehadi, and R. Everaers. Generic model for DNA deformations at the base-pair level. *Phys. Rev. E*, 2003. submitted.

[MMK96] R. S. Manning, J. H. Maddocks, and J. D. Kahn. A continuum rod model of sequence-dependent DNA structure. *J. Chem. Phys.*, 105(13):5626–5646, 1996.

[MN97] J. D. Moroz and P. Nelson. Torsional Directed Walks, Entropic Elasticity, and DNA Twist Stiffness. *Proc. Natl. Acad. Sci. USA*, 94(26):14418–14422, 1997.

[MRKL98] H. Merlitz, K. Rippe, K. V. Klenin, and J. Langowski. Looping Dynamics of Linear DNA Molecules and the Effect of DNA Curvature: A Study by Brownian Dynamics Simulation. *Biophys. J.*, 74:773–779, 1998.

[MRR+53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. N. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.

[MRZ82] D. P. Millar, R. J. Robbins, and A. H. Zewail. Torsion and bending of nucleic-acids studied by sub-nanosecond time-resolved fluorescence depolarization of intercalated dyes. *J. Chem. Phys.*, 76:2080–2094, 1982.

[MS94] J. F. Marko and E. D. Siggia. Bending and twisting elasticity of DNA. *Macromolecules*, 27:981–988, 1994.

[MS95a] J. F. Marko and E. D. Siggia. Fluctuations and Supercoiling of DNA. *Science*, 265:506–508, 1995.

[MS95b] J. F. Marko and E. D. Siggia. Statistical mechanics of supercoiled DNA. *Phys. Rev. E*, 52(3):2912–2938, 1995.

[MS95c] J. F. Marko and E. D. Siggia. Stretching DNA. *Macromolecules*, 28:8759–8770, 1995.

[MSE] B. Mergell, H. Schiessel, and R. Everaers. in preparation.

[Nel98] P. Nelson. New Measurements of DNA Twist Elasticity. *Biophys. J.*, 74:2501–2503, 1998.

[NSJK96] I. A. Nyrkova, A. N. Semenov, J.-F. Joanny, and A. R. Khokhlov. Highly Anistropic Rigidity of "Ribbon-Like" Polymers: I. Chain Conformation in Dilute Solutions. *Journal de Physique II*, 6:1411–1428, 1996.

[OBB+01] W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X. J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C. S. Tung, E. Westhof, C. Wolberger, and H. M. Berman. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, 313:229–237, 2001.

[Odi95] T. Odijk. Stiff Chains and Filaments under Tension. *Macromolecules*, 28:7016–7018, 1995.

[OKLN98] C.S. O'Hern, R.D. Kamien, T.C. Lubensky, and P. Nelson. Elasticity Theory of a Twisted Stack of Plates. *Eur. Phys. J. B*, 1:95–102, 1998.

[OZ00] W. K. Olson and V. B. Zhurkin. Modeling DNA deformations. *Current Opinion in Structural Biology*, 10:286–297, 2000.

[PH98] M. J. Packer and C. A. Hunter. Sequence-dependent DNA Structure: The Role of the Sugar-phosphate Backbone. *J. Mol. Biol.*, 280:407–420, 1998.

[PHP00] R. Podgornik, P. L. Hansen, and V. A. Parsegian. Elastic moduli renormalization in self-interacting stretchable polyelectrolytes. *J. Chem. Phys.*, 113(30):9343–9350, 2000.

[PMSV00] A. A. Podtelezhnikov, C. Mao, N. C. Seeman, and A. Vologodskii. Multimerization-Cyclization of DNA Fragments as a Method of Conformational Analysis. *Biopolymers*, 79:2692–2704, 2000.

[PQSC94] T. T. Perkins, S.R. Quake, D.E. Smith, and S Chu. Relaxation of a Single DNA Molecule Observed by Optical Microscopy. *Science*, 264:822, 1994.

[PR00a] S. Panyukov and Y. Rabin. Fluctuating filaments: Statistical mechanics of helices. *Phys. Rev. E*, 62(5):7135–7146, 2000.

[PR00b] S. Panyukov and Y. Rabin. Thermal Fluctuations of Elastic Filaments with Spontaneous Curvature and Torsion. *Phys. Rev. Lett.*, 85(11):2404–2407, 2000.

[PR01] S. Panyukov and Y. Rabin. On the deformation of spontaneously twisted fluctuating ribbons. *cond-mat/0105116*, 2001.

[PRPL96] J. W. Perram, J. Rasmussen, E. Praestgaard, and J. L. Lebowitz. Ellipsoid contact potential: Theory and relation to overlap potentials. *Phys. Rev. E*, 54(6):6565–6572, 1996.

[PSC97] T. T. Perkins, D. E. Smith, and S. Chu. Single Polymer Dynamics in an Elongational Flow. *Science*, 276:2016–2020, 1997.

[PSLC95] T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu. Stretching of a Single Tethered Polymer in a Uniform Flow. *Science*, 268:83–87, 1995.

[PW85] J. W. Perram and M. S. Wertheim. Statistical Mechanics of hard ellipsoids: 1. Overlap Algorithm and the Contact Function. *J. Comput. Phys.*, 58(3):409–416, 1985.

[PWLW94] J. W. Perram, M. S. Wertheim, J. L. Lebowitz, and G. O. Williams. Monte-Carlo Simulation of Hard Spheroids. *Chem. Phys. Lett.*, 105(3):277–280, 1994.

[RB01a] I. Rouzina and V. A. Bloomfield. Force-Induced Melting of the DNA Double Helix: 1. Thermodynamic Analysis. *Biophys. J.*, 80:882–893, 2001.

[RB01b] I. Rouzina and V. A. Bloomfield. Force-Induced Melting of the DNA Double Helix: 2. Effect of Solution Conditions. *Biophys. J.*, 80:894–900, 2001.

[SAB+96] T. R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The Elasticity of a Single Supercoiled DNA Molecule . *Science*, 271:1835–1837, 1996.

[SAB+99] T. R. Strick, J.-F. Allemand, D. Bensimon, R. Lavery, and V. Croquette. Phase coexistence in a single DNA molecule. *Physica A*, 263:392–404, 1999.

[SABC98] T. R. Strick, J.-F. Allemand, D. Bensimon, and V. Croquette. Behaviour of Supercoiled DNA. *Biophys. J.*, 74:2016–2028, 1998.

[SBC99] T. R. Strick, D. Bensimon, and V. Croquette. Micro-mechanical measurement of the torsional modulus of DNA. *Genetica*, 106:57–62, 1999.

[SCB96] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules. *Science*, 271:795–799, 1996.

[SCB98] T. R. Strick, V. Croquette, and D. Bensimon. Homologous Pairing in Stretched Supercoiled DNA. *Proc. Natl. Acad. Sci. USA*, 95(18):10579–10583, 1998.

[Sch02] H. Schiessel. How short-ranged electrostatics controls the chromatin structure on much larger scales. *Europhys. Lett.*, 58(1):140–146, 2002.

[Sch03] H. Schiessel. The physics of chromatin. *J. Phys.: Condens. Matter*, 15(19):R699–R774, 2003.

[SFB92] S. B. Smith, L. Finzi, and C. Bustamante. Direkt Mechanical Measurements of the Elasticity of Single DNA-Molecules by using Magnetic Beads. *Science*, 258:1122–1126, 1992.

[SGB01] H. Schiessel, W. M. Gelbart, and R. Bruinsma. DNA Folding: Structural and Mechanical Properties of the Two-Angle Model for Chromatin. *Biophys. J.*, 80:1940–1956, 2001.

[SN02a] C. Storm and P. Nelson. The bend stiffness of S-DNA. *arXiv:physics/0212032*, 2002.

[SN02b] C. Storm and P. Nelson. Theory of High-Force DNA Stretching and Overstretching. *arXiv:physics/0206088*, 2002.

[SO92] T. Schlick and W. K. Olson. Supercoiled DNA Energetics and Dynamics by Computer Simulation. *J. Mol. Biol.*, 223:1089–1119, 1992.

[SS86] J. M. Schurr and K. S. Schmitz. Dynamic light-scattering-studies of biopolymers - effects of charge, shape, and flexibility. *Annual Review of Physical Chemistry*, 37:271–305, 1986.

[SWBG01] H. Schiessel, J. Widom, R. F. Bruinsma, and W. M. Gelbart. Polymer reptation and nucleosome repositioning. *Phys. Rev. Lett.*, 86:4414, 2001.

[SY84] J. Shimada and H. Yamakawa. Ring-Closure Probabilities for Twisted Wormlike Chains. *Macromolecules*, 17(4):689–698, 1984.

[TKK79] F. Thoma, T. Koller, and A. Klug. Involvement of the histone H1 in the organization of the nucleosome and of the salt-dependent super-structures of chromatin. *J. Cell. Biol.*, 83:404–427, 1979.

[VC95] A. V. Vologodskii and N. R. Cozzarelli. Modeling of Long-Range Electrostatic Interactions in DNA. *Biopolymers*, 35:289–296, 1995.

[VV02] M. Vologodskaia and A. Vologodskii. Contribution of the Intrinsic Curvature to Measured DNA Persistence Length. *J. Mol. Biol.*, 317:205–213, 2002.

[WC53] J. D. Watson and F. H. C. Crick. A structure for deoxyribose nucleic acid. *Nature*, 171:737–738, 1953.

[WF96] R. Wilhelm and E. Frey. Radial distribution function of semiflexible polymers. *Phys. Rev. Lett.*, 77(12):2581–2584, 1996.

[WGHW93] C. L. Woodcock, S. A. Grigoryev, R. A. Horowitz, and N. Whitaker. A Chromatin Folding Model that Incorporates Linker Variability Generates Fibers Resembling the Native Structures. *Proc. Natl. Acad. Sci. USA*, 90(19):9021–9025, 1993.

[Wid92] J. Widom. A relationship between helical twist of DNA and the ordered positioning of nucleosomes in all eukariotic cells. *Proc. Natl. Acad. Sci. USA*, 89(3):1095–1099, 1992.

[WL02] G. Wedemann and J. Langowski. Computer Simulation of the 30-Nanometer Chromatin Fiber. *Biophys. J.*, 82:2847–2859, 2002.

[WSY+00] G. J. L. Wuite, S. B. Smith, M. Young, D. Keller, and C. Bustamante. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature*, 404:103–106, 2000.

[WWRB01] M. C. Williams, J. R. Wenner, I. Rouzina, and V. A. Bloomfield. Effect of pH on the Overstretching Transition of Double-Stranded DNA: Evidence of Force-Induced DNA Melting. *Biophys. J.*, 80:874–881, 2001.

[Yam97] H. Yamakawa. *Helical Wormlike Chains in Polymer Solutions*. Springer, Berlin, 1997.

[YHZc00] Z. Yang, Z. Haijun, and O.-Y. Zhong-can. Monte-Carlo Implementation of Supercoiled Double-Stranded DNA. *Biophys. J.*, 78(4):1979–1987, 2000.

[YLW93] J. Yao, P. T. Lowary, and J. Widom. Twist constraint on linker DNA in the 30-nm chromatin fiber: Implications for nucleome phasing. *Proc. Natl. Acad. Sci. USA*, 90(20):9364–9368, 1993.

[ZL01] Z. Zhou and Pik-Yin Lai. On the consistency of two elastic models for double-stranded DNA. *Chem. Phys. Lett.*, 346:449–454, 2001.

[ZZOY00] H. Zhou, Y. Zhang, and Z. Ou-Yang. Elastic property of single double-stranded DNA molecules: Theoretical study and comparison with experiments. *Phys. Rev. E*, 62(1):1045–1058, 2000.