Close encounters with DNA



2014 J. Phys.: Condens. Matter 26 413101

(http://iopscience.iop.org/0953-8984/26/41/413101)

IOP Publishing Journal of Physics: Condensed Matter
J. Phys.: Condens. Matter **26** (2014) 413101 (36pp) doi:10.1088/0953-8984/26/41/413101

## Topical Review

# Close encounters with DNA

**C Maffeo[1,3], J Yoo[1,3], J Comer[1], D B Wells[1], B Luan[2] and A Aksimentiev[1]**

[1] Department of Physics, University of Illinois, Urbana, IL, USA
[2] IBM T J Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY, USA

E-mail: aksiment@illinois.edu

Received 19 May 2014, revised 25 July 2014
Accepted for publication 1 August 2014
Published 19 September 2014

**Abstract**

Over the past ten years, the all-atom molecular dynamics method has grown in the scale of both systems and processes amenable to it and in its ability to make quantitative predictions about the behavior of experimental systems. The field of computational DNA research is no exception, witnessing a dramatic increase in the size of systems simulated with atomic resolution, the duration of individual simulations and the realism of the simulation outcomes. In this topical review, we describe the hallmark physical properties of DNA from the perspective of all-atom simulations. We demonstrate the amazing ability of such simulations to reveal the microscopic physical origins of experimentally observed phenomena. We also discuss the frustrating limitations associated with imperfections of present atomic force fields and inadequate sampling. The review is focused on the following four physical properties of DNA: effective electric charge, response to an external mechanical force, interaction with other DNA molecules and behavior in an external electric field.

Keywords: molecular dynamics, nucleic acids, effective charge, nanopore, electrophoresis, charge inversion, counterions

(Some figures may appear in colour only in the online journal)

## 1. Introduction

After water and oxygen, DNA is, very likely, the most famous molecule of life known to humankind. This is not surprising, as we all know that an eye-catching, double-helical molecule of DNA carries instructions to manufacture and assemble all the components of a living organism. The wealth of information encoded in a DNA molecule often overshadows its unusual physical properties. For example, the force–extension dependence of double-stranded DNA has a well-defined plasticity plateau that is associated with melting or conformational change of its two strands. Despite being highly negatively charged, DNA molecules can attract one another and form a condensed state. The direction of DNA motion in an external electric field can reverse upon changing the concentration of the surrounding electrolyte. The nucleotide sequence of DNA is usually determined using methods that rely on the electrophoretic motion of DNA, a physical process of little direct biological relevance. Even the biological role of DNA as storage for genetic information is affected by its sequence-specific physical properties [1–3].

Despite the large number of theoretical and experimental studies, the nature of the microscopic processes that give rise to the above phenomena remain highly debated. With the advent of massively parallel supercomputers it has become possible to characterize these processes directly, through all-atom molecular dynamics (MD) simulations. In this topical review, we present an up-close perspective of the major physical properties of DNA. Because the all-atom MD method explicitly describes the trajectory of every atom in the system with femtosecond resolution, it has the potential to give unparalleled insight into an experimental system. The primary use of the MD method is to suggest a physically plausible explanation or justification of an experimental measurement by animating an equivalent system *in silico*. Equipped with a physically correct description of interatomic interactions and adequate computational power, the MD method should be able to predict the physical behavior of any biological system. Despite ever-increasing availability of massive parallel computing platforms, making quantitative predictions using MD remains challenging,

[3] These authors contributed equally.

 

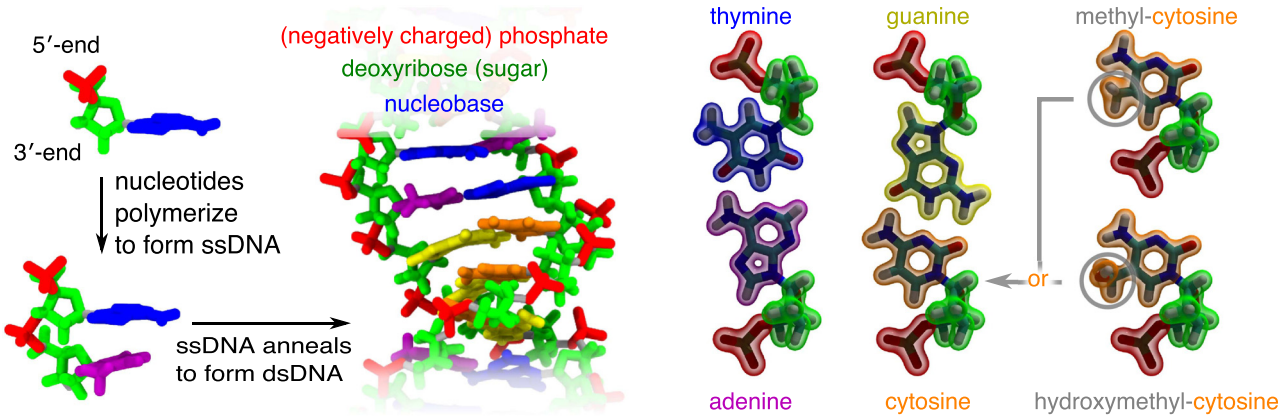


**Figure 1.** Chemical model of DNA. DNA is a polymer composed of nucleotides, each having a negatively charged phosphate, a deoxyribose sugar ring and one of the four nucleobases: adenine, thymine, guanine, cytosine. Two single DNA strands form a double helix held together through non-covalent interactions. In addition to the four types of DNA nucleotides, chemical modification of DNA occurs frequently and includes methylation and hydroxymethylation.

in part due to imperfections of the inter-atom interaction models.

Before we proceed, let's review the basic chemical structure of DNA, figure 1. A molecule of DNA is a polymer made up of many DNA nucleotides linearly arranged into a polymer chain. Single-stranded DNA (ssDNA) is made of one such chain, whereas in double-stranded DNA, two ssDNA molecules are arranged into a DNA double helix through non-covalent interactions. The basic unit of DNA structure—a DNA nucleotide—has three major groups: backbone, sugar and base. The backbone is negatively charged under physiological conditions and has a direction (5′-to-3′) determined by the order of the atoms forming the backbone. The sugar group links the backbone to the base. The chemical difference between DNA and RNA is the presence of an extra hydroxyl (OH) moiety in the sugar group, which strongly alters the properties of the molecule. The DNA base carries genetic information and typically comes in the one following four types: adenine (A), cytosine (C), guanine (G) and thymine (T). The complementary hydrogen bond paring of A with T and G with C governs the nucleotide sequence-specific assembly of two single strands into a double helix. The most familiar conformation of a DNA duplex is the so-called B-form duplex shown in figure 1, but DNA can also adopt a similar but more compact conformation known as an A-form duplex. Except where specified, discussion about double-stranded DNA pertains to B-form DNA.

Chemical modification of DNA bases is common. The most abundant variation is the addition of a methyl group to a cytosine, transforming it into a methylated cytosine—a carrier of epigenetic information (derived from the history of an organism). Modifications are possible, including hydroxymethylation and other derivatives [4]. Under physiological conditions, each DNA nucleotide carries a charge of one electron. A single DNA strand is much more flexible than a double helix. The following discussion implicitly assumes that DNA is in an aqueous environment at pH 7.0.

The rest of the review is organized as follows. After a brief description of the history and the present state of all-atom modeling of DNA, we describe the application of the MD method to the study of equilibrium properties of DNA, including its charge and mechanical properties, the behavior of DNA under an external force, the interaction between DNA molecules and finally the motion of DNA in an external electric field. Where possible, we begin our description of the physical phenomena with a brief review of conventional theoretic models.

## 2. Historical perspective on DNA modeling and simulation

The year 1952 witnessed the publication of the first solid evidence that DNA is the genetic material [5]. Just one year later in 1953, the base-paired double-helical structure of DNA was proposed [6]. In that same year, Hermann Staudinger received a Nobel Prize for his work in 1922 demonstrating that polymers are composed of long chains of covalently bonded atoms, and Paul Flory—later a Nobel laureate for his work in theoretical polymer chemistry—published his seminal book *Principles of Polymer Chemistry*. Thus, a theoretical framework for discussing polymers was available for the study of DNA, including the widely employed freely jointed chain (FJC) and worm-like chain (WLC) models [7–9], see section 5.2. No time was wasted; in that same year, the WLC model was used to analyze the angular dependence of the intensity of scattered light to suggest a surprisingly accurate measure of the flexibility of DNA [10, 11]. Other early theoretical studies of DNA focused on the thermodynamic transition of denaturation using a variety of approaches [12–15]. One perhaps iconic approach was an adaptation of the one-dimensional Ising model (originally developed to describe magnetization), whereby the likelihood of a given base pair's unpairing depends on the paired status of its nearest neighbors in a DNA helix [12].

By 1975, it was known that DNA was a rather rigid polymer, yet it could form compact structures like chromatin [16], sparking a debate about kinked versus smoothly bent DNA. In the earliest atomic-modeling studies that employed computation, models of kinked and bent DNA were produced by performing least-squares searches of toy potentials linking rigid groups of atoms [17, 18]. Soon after, Levitt borrowed atomic

interaction potentials from protein and RNA refinement studies to relax the full set of atomic coordinates of a DNA molecule that was bent and twisted by varying amounts [19]. Although both solvent and electrostatics were absent in the calculations and the calculations probed only the local energy minima, the study correctly determined the 10.5 base pairs-per-turn pitch of a DNA helix in solution, departing from the 10 base pairs-per-turn pitch observed in the solid-state Watson–Crick structure.

The first room-temperature computations involving DNA were performed by Clementi and Corongiu in 1979 and 1980 [20, 21]. These Monte Carlo (MC) simulations, performed using a model optimized to reproduce energies of *ab initio* calculations, revealed the structure of water molecules around fixed DNA helices and nucleotides. In the MC method, atomic coordinates were propagated according to Boltzmann statistics, so equilibrium, but not dynamical, properties of the system could be studied. In the beginning of 1983, Levitt, borrowing again the methods developed for the study of proteins, reported an MD simulation that showed asymmetric bending and twisting motions of duplex DNA in the absence of solvent [22]. This was followed only a few months later by a similar MD report by the Karplus group [23]. Finally in 1985, the Kollman group performed the first MD simulation of a DNA fragment in electrolyte solution [24].

Beveridge *et al* comprehensively reviewed the following decade of MD studies of DNA [25, 26]. At that time, simulations were rapidly approaching the nanosecond timescale [27]. Nevertheless, most simulations from that era described DNA using implicit solvent and truncated electrostatics. In 1995, it was noted that the outcome of a simulation could qualitatively depend on the method used to describe the dielectric environment of a charged polymer [28]. The Kollman group (1995) demonstrated that simulations of DNA using the particle mesh Ewald (PME) method, which efficiently calculates long-range electrostatic interactions in Fourier space, are more accurate than the simulations performed using truncated electrostatics [29]. Although accurate representation of electrostatics is essential for modeling of a highly charged molecule such as DNA, the community was at first slow to adopt the PME method due to the high computational cost [30]. At that time, several 'second-generation' all-atom force fields with explicit solvent were released, including AMBER-94 and CHARMM22 [31]. Around the same time, National Science Foundation supercomputers became more accessible to researchers and the major MD codes AMBER and CHARMM were parallelized.

Prompted by these methodological advances, longer and more accurate studies of DNA structure and dynamics were performed. Of particular note, Cheatham and Kollman (1996) observed a spontaneous transition of DNA from A-form to a more stable B-form during the first multi-nanosecond simulation of explicitly solvated DNA, indicating good accuracy for the AMBER force field [32]. Shortly thereafter Young, Ravishanker and Beveridge reported the first 5-nanosecond trajectory of B-DNA [33], which revealed substantial fluctuation of the DNA structure, in agreement with x-ray crystallography and NMR. The latter study was one among several that found that MD simulations employing the CHARMM or AMBER force field could reproduce the crystal and NMR structures [30], which was an important test of the force fields. The sequence-specificity of DNA structures began to be investigated including comparisons of twist, roll, and tilt obtained for different base pair stacks [34]. More dramatically, the intrinsic curvature of A-tract DNA was observed [33] (see section 5.2 for details). Back in 1986, it had been recognized from crystal structures that water may considerably affect the conformation of DNA (A-form versus B-form) [35]. Feig and Pettitt (1998) reviewed MD simulations that investigated the structure and properties of water surrounding DNA [36]. Another question of outstanding importance was the location and dynamics of ions, which were believed to bind DNA tightly [37]. A number of studies investigated whether and where the ions would bind to DNA [38] and began to characterize the ion atmosphere [26, 39, 40]. Around the same time, the generalized Born method [41] was introduced to implicit solvent simulations with AMBER [42] and CHARMM [43] parameters for more accurate estimation of the electrostatic solvation energy.

From 1995 to 2000, the era of quantitative MD simulations of DNA began to emerge, starting with free energy perturbation simulations of ligand binding [30, 34]. The structural and intramolecular character of DNA was investigated through free energy simulations of base pairing, stacking, and DNA stretching [30]. In particular, MacKerell and Lee (1999) used umbrella sampling MD simulations and atomic force microscopy to study stretch-induced melting of short DNA fragments [44]. Near-quantitative agreement was obtained between simulation and experiment, indicating good overall performance of the CHARMM22 force field. It must be cautioned that the AFM results were not in complete agreement with the seminal study of stretching long (48.5 kbp) λ-phage DNA using an optical trap [45], perhaps because different DNA constructs were employed [44]. Nevertheless, the simulation represented the first, to our knowledge, quantitative comparison of mechanical DNA properties observed in experiment and in simulation.

In the new millennium, there have been a few general reviews of DNA simulations [53–55]. However, the field has grown to such an extent that there have been very many special-topic reviews on a wide range of subjects. For example, the base stacking interactions were reviewed from a QM perspective [46]. Another recent review that included QM and MD descriptions of DNA focused on the backbone rather than bases [56]. Work continued toward understanding the relationship between DNA and the solvent [57, 58], however, it was only recently found that the standard parametrization of cation–phosphate interaction had considerable problems [59]. Advances in computing permitted the mechanical properties of DNA to be examined in more detail [60–63]. Finally, there were many investigations of DNA association with other molecules, including small molecules such as anti-cancer drugs [64, 65], sequence-specific DNA binding proteins [64, 66], non-specific DNA binding proteins such as the nucleosome [67] and synthetic structures such as carbon nanotubes [68] and silicon nitride nanopores [69].

In recent years, MD simulations of DNA have moved toward large systems and long durations. The first microsecond

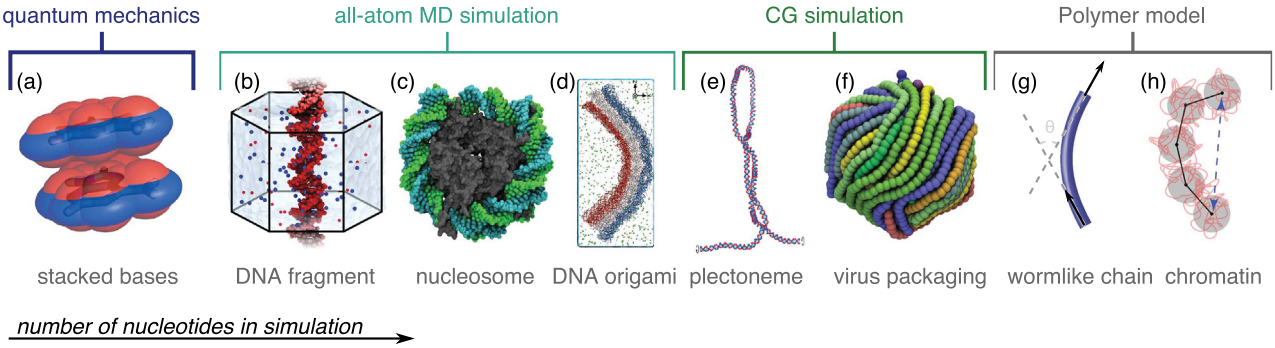


**Figure 2.** Nucleic acid systems that represent the range of scales amenable to various computational and theoretical methods. From left to right, generally increasing in number of nucleotides: (*a*) base-stacking interactions can be studied through quantum mechanics (QM) calculations, and reveal asymmetric van der Waals (vdW) radii [46]. The all-atom MD method with explicit or implicit solvent provides a balance between computational speed and descriptive detail. This method can be used to study systems ranging from a few nucleotides to (*b*) several turns of DNA [47] to (*c*) hundreds of nucleotides as in the nucleosome [48]. (*d*) A large multi-layer, curved DNA origami construct described by MD simulation with explicit solvent [49]. Coarse-grained simulations can describe DNA with a wide range of detail, from near atomic (e.g. to describe (*e*) a plectoneme [50]) to many nucleotides per site (e.g. to describe (*f*) the packaging of a virus [51]). The two dominant polymer models of DNA are (*g*) the WLC model [9], and the FJC model [7, 8]. The WLC model describes DNA under tension better than FJC, but FJC is convenient for very coarse descriptions of DNA, (*h*) such as at the level of chromatin [52]. Panels adapted with permission from (*a*) [46], (*b*) [47], (*d*) [49], (*e*) [50], (*f*) [51] and (*h*) [52].

simulation of B-DNA was performed by the Orozco group in 2007 [70], which revealed the limitations of the AMBER-99 force field and a new parametrization of the all-atom model that addressed the problem. Meanwhile, a number of impressive developments have happened in the related field of RNA systems simulations. An all-atom model of one of the smallest viruses—the satellite tobacco mosaic virus—was simulated for ∼50 ns, which included 60 copies of the coat protein, a 1063-nucleotide single-stranded RNA molecule and the electrolyte solution, about 1 million atoms in total [71]. Even larger RNA systems have since been simulated. Atomic simulations of the ribosome, an enormous protein synthesis factory predominantly composed of RNA, were recently reviewed by Sanbomatsu [72]. The simulations examined the conformational changes in the ribosome, the effects of point mutations and quantified the kinetics and free energy barriers of conformational transitions using a 3 million atom model and the aggregate simulation time of ∼2 $\mu$s.

Large-scale MD simulations of DNA systems have only recently matched the scale of the largest RNA simulations, perhaps because of the lack of atomic-scale structures of large DNA assemblies. Matching the set-up of DNA array experiments of Rau and Parsegian [73], Yoo and Aksimentiev simulated the structure, dynamics, ionic atmosphere and intermolecular forces of a DNA array [59]. The outcome of these simulations revealed the limitations of the standard ion–DNA interactions and a method to fine-tune the parameters to achieve quantitative agreement with experiments (see section 3.3). Another example was a simulation of end-to-end self-assembly of a large number (∼450) of duplex DNA [74]. One of the current frontiers of DNA simulations are DNA origami and related self-assembled nanostructures [75–77]. The first atomic-resolution simulation of DNA origami was reported very recently [49], revealing the *situ* structure of several DNA origami designs as well as their local and global mechanical properties.

The above discussion has focused predominantly on fully-atomic simulation of DNA. By employing a less detailed, 'coarse-grained' (CG) model the timescale accessible to simulations of DNA can be significantly extended [50, 78, 79]. In the past five years, many CG models of DNA have been developed to reproduce various properties of double-stranded DNA [80–90]. Most available CG models employ a few interaction sites to represent each nucleotide, but CG models can span an enormous range of scales, from the near-atomic [89] to the meso and macroscopic [51]. An excellent review of available CG DNA models can be found in the recent article from the de Pablo group [80].

Parameters describing the interactions of CG models are usually obtained by fitting against experimental data, such as the nearest neighbor DNA melting parameters [91], or by reproducing structural parameters from atomic resolution simulations. In analogy to the atomic force fields, which are optimized against a mix of finer quantum calculations and experimental data, it is our belief that the most accurate general purpose CG models of DNA will be obtained through a combination of all-atom simulations and single molecule experiment. A recent application of such an approach is a CG model of single-stranded thymine homopolymers specifically developed to match both the all-atom simulations and single-molecule fluorescence resonance energy transfer measurements [92].

Finally, the coarsest biologically relevant description of DNA ignores its physical form altogether, representing each nucleotide with one of four letters, A-C-G-T. In the field of genomics, computers play an instrumental role in organizing, sorting, search and comparing the genomes of different species and individuals. Since this field lies largely outside of our expertise, we merely mention computational genomics as an important and exciting area of computational modeling of the biological function of DNA.

## 3. All-atom force fields for DNA simulations

The MD method approximates the QM of chemical bonds using classical equations of motion. Such an approximation

requires a set of functions and parameters—the molecular force field—to describe the interactions between the atoms. In general, biomolecular force fields can be categorized according to the level of description of the surrounding media (gas phase, implicit solvent, explicit solvent), functional form of the force field, and crucially the method used to obtain the parameters.

Prominent early efforts in force field development for DNA were undertaken in the Karplus [23] and Kollman [24, 93, 94] groups, eventually leading to the CHARMM and AMBER parameter sets. The functional forms of those force fields are similar to those found today: harmonic terms on the bonds, angles and improper dihedral angles; a periodic term for each dihedral angle; and Lennard-Jones (LJ) and Coulomb terms between non-bonded atoms. Below we briefly review the formulations of the all-atom MD method and describe recent advances in this area.

### 3.1. General formulation of the all-atom MD method

In the MD method, point particles represent atoms. The atomic force field (discussed below) provides all atoms with predefined physical properties such as partial charges and vdW interaction parameters. The connectivity (or covalent bonds) among atoms is given *a priori*. For computational efficiency, the physical parameters remain the same during a given simulation, with a few exceptions such as alchemical free energy simulations [95, 96]. Next-generation force fields may include a dynamic description of molecular properties such as polarizability [97, 98]. For example, a polarizable CHARMM force field based on the classical Drude oscillator model with a complete description of DNA, water and ions was published recently [99, 100]. This model was optimized using several different levels of computations: quantum calculations for base–cation interactions and backbone torsions; thermodynamic osmotic pressure calculation for cation–phosphate interactions; hydration free energy calculation for DNA–water interactions.

Covalently bonded atoms interact with each other through bonded potentials, while the other atom pairs interact through non-bonded potentials:

$$U(\mathbf{r}^N) = U_{\text{bonded}}(\mathbf{r}^N) + U_{\text{nonbonded}}(\mathbf{r}^N), \tag{1}$$

where $\mathbf{r}^N$ denotes the coordinates of $N$ atoms in the system. Bonded interactions model quantum mechanical behavior of covalently connected atoms by means of harmonic bond, angle and improper dihedral angle restraints, and periodic dihedral angle potentials:

$$U_{\text{bonded}} = \sum_{\text{bonds}} K_b(b-b_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \sum_{\text{torsions}} K_\chi\{1+\cos(n\chi-\delta)\} + \sum_{\text{impropers}} K_\phi(\phi-\phi_0)^2, \tag{2}$$

where $K_b$ and $b_0$ are the force constant and equilibrium distance of the bond, respectively; $K_\theta$ and $\theta_0$, the force constant and equilibrium value of the angle, respectively; $K_\chi$, $n$ and $\delta$, the force constant, multiplicity and phase of the dihedral, respectively; $K_\phi$ and $\phi_0$, the force constant and equilibrium value of the improper dihedral angle [101]. The non-bonded potential usually applies to atom pairs separated by more than two covalent bonds and consists of LJ potential for vdW interactions and Coulomb potential for electrostatic interactions:

$$U_{\text{nonbonded}} = \sum_{\text{nonbond}} \left\{ 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{\epsilon r_{ij}} \right\}, \tag{3}$$

where $\epsilon_{ij}$ is the well depth; $\sigma_{ij}$, the finite distance at which the LJ potential is zero; $r_{ij}$, the interatomic distance; $q_{i,j}$, atomic charges for atom $i$ and $j$. The bonded parameters are empirically calibrated based on the quantum mechanical calculations of small molecules, whereas the non-bonded parameters are mainly derived from quantum chemistry calculations (e.g. partial charges) and empirical matching of thermodynamic data (e.g. hydration free energy).

Intramolecular and intermolecular non-bonded forces dominate the tertiary structure of a molecule and molecular binding, respectively. Thus, accurate description of the non-bonded forces is essential. In the 1980s and 1990s, the Coulomb interactions, when included, were truncated at a relatively short distance (~1 nm). Since solvent was prohibitively expensive to treat explicitly, the early Karplus and Kollman force fields treated it implicitly with the approximation that the dielectric constant depended on the distance between two charged atoms [102, 103]. This practice occurred apparently without physical justification, but nevertheless provided reasonable results [26]. Usually the charge of the DNA was significantly reduced to represent the effect of counterions in accordance with Manning counterion condensation theory (see section 4.1.2 for details) [26, 103].

Presently, the most practical and accurate treatment of long-range interactions takes advantage of the periodic boundary conditions employed in most modern MD simulations by using the Ewald summation method, which efficiently computes both short- and long-range electrostatic interactions in crystal systems. Cheatham *et al* have demonstrated the importance of using the Ewald method for DNA simulations by showing that B-form DNA conformations were stable only when the Ewald method was used [29]. Optimized versions of the Ewald method [104] permit highly parallelized MD simulations to be efficiently performed.

### 3.2. Recent updates to the second-generation force fields

An all-atom empirical force field for DNA is a set of bonded (equation (2)) and non-bonded (equation (3)) parameters optimized utilizing QM calculations and thermodynamics data. Although there exist several all-atom force fields, the CHARMM [105] and AMBER [106] force fields are two most popular choices for the simulations of DNA. Steady advances in cluster computing continue to increase the size and duration of MD simulations, outpacing force field development. More often than not, simulations today greatly exceed the duration and complexity of the simulations that were originally used to develop and validate the force fields. It is, therefore, perhaps unsurprising that the force fields require frequent updates to

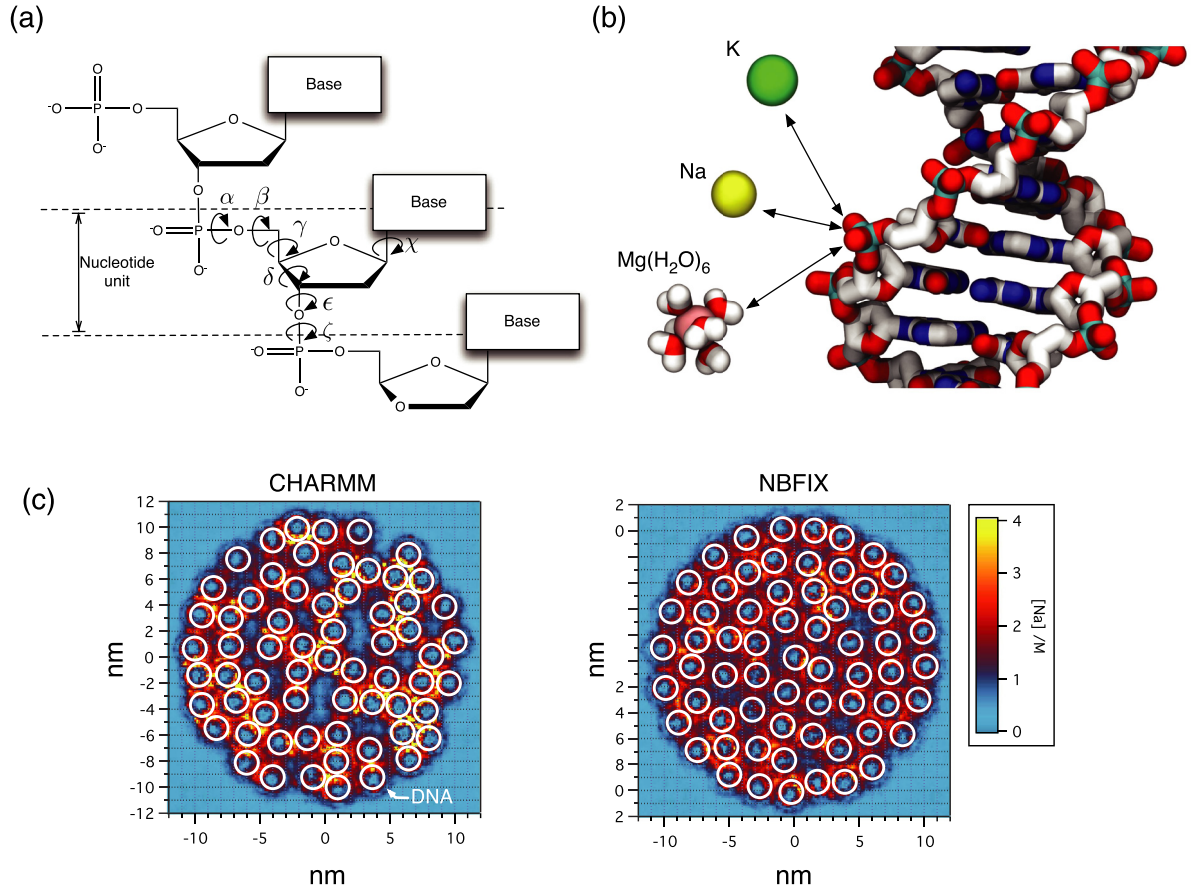


**Figure 3.** Intra-DNA and DNA–ion interactions. (*a*) Schematic illustration of the chemical structure of DNA. Seven torsion angles are required to describe the conformation of a nucleotide unit. (*b*) DNA–cation interactions. The phosphate groups of the DNA backbone strongly interact with cations because of electrostatic attraction. The accuracy of these opposite-charge interactions is essential for simulations of DNA–ion and ion-mediated DNA–DNA interactions [59]. Magnesium is shown in a $Mg(H_2O)_6$ complex to emphasize its ability to coordinate water molecules. See section 4.2.2 for a detailed discussion of $Mg^{2+}$ parametrization. Bulk water and DNA hydrogen atoms are omitted for clarity. (*c*) Custom NBFIX corrections improve accuracy of DNA array simulations. The panels show representative configurations of 64 dsDNA molecules (white circles) confined in a cylindrical volume (not shown). The color indicates the local density of cations. (Left) the artificially strong attractions between cations and DNA phosphate results in erroneous clustering of DNA. (Right) using custom parameters to describe cation–DNA phosphate interactions recovers hexagonal packing of DNA helices [59], in agreement with experiments [73].

keep up with the opportunities offered by modern supercomputing systems.

*3.2.1. Backbone.* Since the original development of the 'second generation' DNA force fields about 20 years ago [106, 105], both the AMBER and CHARMM force fields have been updated several times. The majority of the updates focused on refinements of the bonded parameters, in particular the torsional potentials, using QM calculations to improve the description of intra-helical conformations [55, 107]. The large number of updates reflects the complexity of the backbone torsional motions in DNA, see figure 3(*a*). Specific updates of the torsional potentials include refinements of $\chi$ and $\gamma$ to improve the balance of A- and B-form DNA populations [108], $\alpha$ and $\gamma$ to improve the stability of B-form DNA conformations in long simulations (>10 ns) [109, 110], and $\epsilon$ and $\zeta$ to improve sequence-dependent distributions of B-DNA subpopulations [111]. Thus, six torsion angles along the DNA backbone are required to describe the conformation of a nucleotide unit, whereas a peptide bond requires only two. The structure and fluctuations of the backbone are essential for proper description of protein recognition [56].

Compared to the canonical helical conformations of ds-DNA, non-canonical conformations (e.g. loops, bulges, and kinks or single-stranded DNA) are harder to validate due to significantly larger torsional spaces [112]. It has been suggested recently that conventional AMBER and CHARMM force fields cannot properly describe non-canonical conformations of DNA [113]. This criticism prompted development of AMBER force field variants, e.g. ff11 and ff12 [114, 115]. Readers interested in the detailed history of DNA force field development are referred to recent reviews on this subject [55, 107].

*3.2.2. Ions.* It has been a common practice to use the AMBER force field together with parameters for inorganic ions, e.g. Na, K, Mg, Ca, developed by Åqvist *et al* about 25 years ago [116]. Similarly, ion parameters developed more than 20 years ago are commonly used with the CHARMM force field [117].

While parametrized to match the ion solvation energy, the default sets of ion parameters in both AMBER and CHARMM overestimate the interaction energy of cation–anion pairs as well as the energy of select ion–charged group interactions. When used in a simulation, such inaccurate parametrization leads to the formation of artificial ion clusters [118, 119]. Updates to the default ion parameters were put forward by the Cheatham group for the AMBER force field (Cheatham ion parameters) [120, 121] and the Roux group for the CHARMM force field (Roux ion parameters) [119]. The ion parameters were developed to work with the TIP3P water model, so in principle one could employ either the Cheatham or Roux ion parameters in simulations using either force field.

Even the updated parameters were found to produce considerable artifacts in simulations of dense DNA systems [59]. Such high density DNA systems are commonplace in biological systems and in nanotechnological applications of DNA. For example, in a fully packed bacteriophage (a virus infecting bacteria), the inter-DNA distance is $<30$ Å [122, 123] and hence the molar concentration of the bare DNA charge is higher than 1M. The DNA charges must be neutralized by a similar amount of opposite charge counterions, which include both inorganic, such as those shown in figure 3(*b*), and more complex organic compounds, such as biogenic polyamines or histone tails. For proper description of high density DNA systems, an accurate parametrization of inter-DNA and DNA–ion interactions is as important as the parametrization of intra-DNA interactions.

Figure 3(*c*) illustrates some of the problems with the standard ion parameter sets. When 64 dsDNA molecules were simulated in a cylindrical well permeable to ions, the DNA molecules formed clusters mediated by exaggerated binding of ions to the DNA phosphate groups [59]. The pressure required to keep the DNA molecules confined to the cylindrical well was an order of magnitude less than in experiment [73]. Yoo and Aksimentiev introduced custom corrections to the vdW parameters (NBFIX) describing the interaction of a specific cation (Li, Na, K or Mg) with a phosphate oxygen to bring the simulated osmotic pressure of the binary mixture into accordance with the experiments of acetate (a proxy for phosphate)–cation solutions [124]. Using such corrections considerably improved the realism of DNA array simulations. Specifically, the simulations could reproduce the hexagonal packing of the DNA helices; the average inter-DNA distance, and the dependence of the internal DNA pressure on the inter-DNA distance, all in agreement with the corresponding experimental measurements [73]. Additional validation of the NBFIX corrections was obtained by simulating competitive binding of two cation species to DNA [125], which quantitatively reproduced ion counting experiments [126] (see section 4.2).

A detailed discussion of the parametrization strategies for accurate description of divalent ions in the context of nucleic acid simulations is given in section 4.2.2.

*3.2.3. Base-stacking.* As the computational capabilities of modern computer systems keeps on increasing, so does the appetite for simulations of ever more realistic biological systems involving DNA, which often feature non-canonical conformations of DNA and DNA–protein assemblies. One of the major problems in describing such conformations is the strength of base stacking interactions of DNA nucleotides. For example, a recent study has quantitatively investigated the end-to-end association of short duplex DNA fragments using free energy and brute-force simulations [74], see section 6.4. The magnitude of the end-to-end attraction was found to be considerably larger than the corresponding experimental estimates [127, 128].

Base-stacking interactions have been the subject of extensive QM studies [46]. Acceptable agreement was found for the quantum gas-phase enthalpies and the AMBER-99 force field, but the atomic radii in the force field were noted to be larger than expected from state-of-the-art QM calculations (in the direction normal to the base). It is important to note that the latter was a recent discovery enabled by simulations employing a basis set with a higher level description of electron correlations than previously possible. The results of the advanced QM calculations were taken into account by the Garcia group in a recent modification of the AMBER-99 force field that eliminated unrealistic aggregation of nucleotides at low concentrations [129]. Using this sequence-independent correction to the base-stacking interaction permitted a more realistic simulation of folding of one RNA molecule [129]. However, the correction has not yet been validated for a variety of RNA and DNA structures.

*3.2.4. Best practices.* At the time of writing this review, the latest and probably the most recommended force field choices for the simulations of canonical B-DNA duplexes [55, 107] is AMBER ff99bsc0 [110] and CHARMM36 [111]. Simulations of ssDNA are best performed using the CHARMM36 force field, as the AMBER parametrization was found to enforce the helical conformation of dsDNA on single DNA strands [112]. We strongly recommend using the NBFIX corrections to describe ion–phosphate interactions [59]. Additional improvement of accuracy could be expected from the base-stacking corrections [129], however such corrections have not yet been validated in combination with other recent force field updates. A combination of several recent force field updates might be required for accurate simulations of non-canonical nucleic acids, such as Holliday junctions, ribozyme, DNA/RNA hairpins or synthetic DNA nanostructures. We strongly urge early adopters to always test new force field modifications before employing them in long term and computationally expensive projects.

## 4. Electrostatics and ion atmosphere

### 4.1. Theoretical background

What is the charge of DNA? While a simple answer might be the charge of an electron per nucleotide under standard conditions, we often speak of effective charges due to the complex electrostatics of the molecule in solution. There are various ways to measure the charge of DNA and therefore various ways interpret its value. Perhaps the most basic way to measure charge is to place a charged probe at some distance from the molecule and measure the force. If we make the charge of the probe small enough, then we can make the measurement with only negligible effects on the subject of the measurement,

allowing us to calculate the force on a probe of charge $q$ from the electric potential $\phi$ by $\mathbf{F} = -q\nabla\phi$. For measuring DNA–ion interactions, the charge of the probe may need to be much smaller than $e$, the charge of a proton, but for the moment let's not consider the complication of actually making experimental measurements of the charge this way. We will soon find that determining the effective charge of DNA from the potential at some distance away from it or vice versa is complex even in theory.

In biological organisms, DNA is most commonly found in its stiff double-helical B-form, which is about 2.5 nm in diameter and has a persistence length about 20 times greater [130]. Therefore, DNA is often modeled as an infinitely long rod, with varying levels of detail in its charge distribution. Below we model DNA as a uniformly charged cylinder, allowing for easy calculation, but neglecting the complexity of the ion distribution near the DNA. In the B-form double helix, each base pair typically represents a 0.34 nm step along the helical axis, although the length of the step varies a bit with the DNA sequence. Therefore, DNA has a standard charge density of $Q/L = -2e/0.34$ nm. To compute the electrostatic potential, we can employ Gauss's law from elementary electrostatics, which tells us that the charge inside any closed surface can be related to the potential at the surface by

$$Q_{\text{enclosed}} = -\epsilon_0 \oint_{\text{closed surface}} \mathrm{d}\mathbf{a} \cdot \nabla\phi. \tag{4}$$

Computing $\phi$ from Gauss's law for a uniformly charged cylinder is a standard textbook problem, and is calculated by integrating over a closed surface with the form of a 'can'—a capped cylinder of radius $r$ sharing its axis with the axis of the uniformly charged cylinder. We obtain the solution

$$\phi(r) = -\frac{1}{2\pi\epsilon_0}\frac{Q}{L}\ln(r/r_0), \tag{5}$$

where $r$ is the distance between the uniformly charged cylinder and the point at which we are evaluating $\phi$, $\epsilon_0$ is the electric constant, $Q/L$ is the linear charge density of the DNA, and $r_0$ is an arbitrary distance where $\phi(r_0) = 0$.

This calculation, however, was performed in vacuum. In water, the $Q_{\text{enclosed}}$ of Gauss's law must include also the charges of the water molecules. Although water molecules are neutral on the whole, they have an electric dipole moment whose orientation becomes biased by an electric field. When placing our imaginary 'can' into water, the surface of the can severs the dipoles, leaving unbalanced charges within the surface. At low fields, water can be treated as a homogeneous isotropic linearly polarizable medium, and the effect of water polarization can be captured by a scalar dielectric constant, so that equation (5) becomes

$$\phi(r) = -\frac{1}{2\pi\epsilon_0\epsilon_{\text{r}}}\frac{Q}{L}\ln(r/r_0), \tag{6}$$

where $\epsilon_{\text{r}}$ is the relative permittivity, or dielectric constant, of water. Note that we must be careful with the assumption of a uniform dielectric constant since high electric fields and water structuring can modify the dielectric constant of water near ions or DNA [131].

Here is an example where we can illustrate the idea of an effective charge. Suppose we let $Q_{\text{eff}} = Q/\epsilon_{\text{r}}$. This substitution is not often made, but it serves as an example. The advantage of this substitution is that the equation for the potential in water regains the simple form it had for the vacuum. However, we must keep in mind that the effective charge is not an intrinsic property of DNA, but depends also on the solvent. Moreover, under the conditions where the dielectric properties of water can no longer be described by a single scalar value, $Q_{\text{eff}}$ will also cease to be constant. Introducing an effective charge can be useful for simplifying a complex theory; however, in all cases, as in this case, we must be wary about its range of validity.

DNA is invariably suspended in an aqueous electrolyte solution containing many small mobile ions, such as $Na^+$, $K^+$, $Cl^-$, or ions of higher valencies such as $Mg^{2+}$. The high charge of DNA therefore causes a complex rearrangement of these mobile ions, making the form of $\phi$ calculated above invalid. A large amount of literature has been focused on understanding the interaction between nucleic acids and ions due to the importance of the phenomenon in processes necessary for life including protein–nucleic acid binding, gene expression, and chromosome packaging as well as in biotechnology applications. Here, we briefly describe some pertinent research in the field to examine the use of effective charges for DNA.

*4.1.1. Poisson–Boltzmann theory.* A number of analytical and computational techniques exist in the literature to study the interaction between DNA and ions, each with varying levels of complexity and range of applicability. Due to its simplicity, Poisson–Boltzmann (PB) theory is almost invariably the starting point for such studies. A good description of PB and related mean field theories is given by Grosberg *et al* [132]. In a simple derivation, PB theory begins with the Poisson equation with a scalar permittivity, $\nabla^2\phi = -\rho/(\epsilon_0\epsilon_{\text{r}})$, where $\phi$ is the electric potential, $\rho$ is the charge density, $\epsilon_0$ is the electric constant, and $\epsilon_{\text{r}}$ is the dielectric constant (relative permittivity) of the medium. We then assume that small mobile ions rearrange themselves in accord with Boltzmann statistics. We treat the ions as point particles so that their energy is simply $E = q\phi$, where $q$ is the charge of an ion. In doing this, we have considered interactions of the ions only through the average potential $\phi$ and are neglecting interactions such as steric repulsion between nearby ions and hydration effects. Let's suppose that we have a net-neutral univalent electrolyte like KCl in an aqueous solution. Then the PB equation is

$$\nabla^2\phi = -\frac{ec_0}{\epsilon_0\epsilon_r}\left[\exp(-e\phi\beta) - \exp(e\phi\beta)\right], \tag{7}$$

where $\beta = 1/k_{\text{B}}T$ is the inverse of the thermal energy, $c_0$ is the bulk concentration of the ions (the concentration where $\phi = 0$) and the first and second exponential terms represent the charge densities of the cations and anions, respectively. Despite the fact that the PB equation neglects ion correlations (being a mean field theory) and steric repulsion between ions at small separations, it contains much of the physics necessary to understand the interaction between DNA and ions. Recent updates to the PB formalism allows the inclusion of steric interactions [133, 134].

If we are working in a region where the potential is small ($e\phi\beta \ll 1$), then we can linearize the PB equation by expanding the exponentials by $\exp(x) \sim 1+x$. With the linear PB equation the following potential is obtained for a uniformly charged cylinder [135],

$$\phi(r) = \frac{2\xi r_{\mathrm{DH}}}{e\beta a}\frac{K_0(r/r_{\mathrm{DH}})}{K_1(a/r_{\mathrm{DH}})}, \tag{8}$$

where $K_0$ and $K_1$ are modified Bessel functions [136], $a$ is the radius of the DNA, $r_{\mathrm{DH}} = \sqrt{\epsilon_0\epsilon_{\mathrm{r}}/(8\pi e^2 c_0\beta)}$ is the Debye–Hückel screening length, and $\xi = |Q/L|\ e\beta/(4\pi\epsilon_0\epsilon_{\mathrm{r}})$.

*4.1.2. Ion condensation.* Equation (8) might be the end of the story for describing interactions between ions and DNA for dilute electrolytes (the $c_0 \to 0$ limit). However, when the charge density of the uniformly charged cylinder $Q/L$ becomes sufficiently high so that $\xi > 1$, non-linear effects in the PB equation lead to a change in character of the solutions [132]. A certain fraction of the counterions, $1 - 1/\xi$, becomes associated with the uniformly charged cylinder—in a phenomenon referred to as counterion condensation [37]. Note that here the word 'condensation' refers to a loose association and does not imply physical contact between the ions and DNA. This phenomenon is especially interesting because it applies to the B-form DNA double helix in water at room temperature: using the formulas above we find that $\xi = 4.2$ and that 76% of the counterions are condensed.

What is the nature of this condensation? Originally, Manning demonstrated that the free energy of a point charge interacting with a line charge is infinite if the line charge density exceeds a certain threshold. The critical charge density turns out to be ~25% of the DNA charge for monovalent electrolyte [37]. The divergence of the free energy only occurs for the model where the point charge can approach the line charge arbitrarily close. Nevertheless, something akin to counterion condensation emerges simply from the non-linear PB equation as detailed above for more realistic systems like a finite-sized cylinder. However, in realistic models of DNA, counterion condensation cannot be thought of as binding of ions to specific sites on the DNA [126], as x-ray scattering shows no discrete structure. Manning [137] speaks of the condensed ions as a liquid and emphasizes that the number of condensed ions is determined only by electrostatics and thermodynamics [138, 139].

Mathematically, condensation can be manifested as counterions that remain at finite distances from the DNA as the bulk ion concentration is reduced to zero. Beyond the Debye–Hückel screening length $r_{\mathrm{DH}}$, interactions between charged bodies decrease rapidly; therefore, $r_{\mathrm{DH}}$ can be thought as the characteristic length scale for electrostatic interaction in the solution. As $c_0$ is reduced, $r_{\mathrm{DH}}$ grows; however, the radius that contains the condensed fraction of ions grows more slowly, $R_{\mathrm{cond}} \sim \sqrt{r_{\mathrm{DH}}}$ [135]. In a sense, therefore, the DNA holds the condensed ions more and more tightly as the bulk ion concentration is reduced, since the radius enclosing the condensed ions relative to the characteristic length scale decreases: $R_{\mathrm{cond}}/r_{\mathrm{DH}} \sim 1/\sqrt{r_{\mathrm{DH}}}$.

Here is where introducing an effective charge, $Q_{\mathrm{eff}}$, can be useful. Let the effective charge of the DNA be its bare charge minus the charge of the condensed counterions, $Q_{\mathrm{eff}} = Q - Q(1 - 1/\xi) = Q/\xi$. At distances beyond the radius confining the condensed ions, $r \gg R_{\mathrm{cond}} \approx \sqrt{2ar_{\mathrm{DH}}}\exp[(\xi-2)/(2\xi-2)]$, we can approximate the potential in a dilute electrolyte [135] by

$$\phi(r) = \frac{2\xi_{\mathrm{eff}} r_{\mathrm{DH}}}{e\beta a}\frac{K_0(r/r_{\mathrm{DH}})}{K_1(a/r_{\mathrm{DH}})}. \tag{9}$$

Equation (9) is nearly identical to the formula obtained from the linearized PB equation (equation (8)). The only difference is that $\xi$ is replaced by $\xi_{\mathrm{eff}}$, which differs from $\xi$ only in that the bare charge $Q$ has been replaced with the effective charge $Q_{\mathrm{eff}}$. Thus, under restricted conditions, we can think of the DNA and its condensed counterions as a single entity with a charge density of $Q_{\mathrm{eff}}/L$.

We have continually used an infinite uniformly charged cylinder as a model for DNA. We might worry that the theory of counterion condensation would fail for a more realistic DNA model, for example, having a molecule of finite length or using a helical charge distribution. As it turns out, the condensed fraction $1 - 1/\xi$ and the critical value of the charge density where condensation occurs is fairly universal. Caused by the long-range electrostatic effects, the condensed fraction and critical charge density do not depend on the local details of the DNA's charge distribution—the same results have been found for discrete charges in a double helical arrangement [140]. As for the assumption of infinite length, the calculations presented above hold when the DNA is significantly longer than $r_{\mathrm{DH}}$ [141]. Note, that we must be careful about this criterion, since we are working the dilute ($r_{\mathrm{DH}} \to \infty$) limit. Most importantly, a number of experiments and simulations show evidence of a sharp change in a variety of observables near the critical value $\xi = 1$ [142].

Thus far, we have considered the dilute limit. Can the above considerations be extended to sizeable ion concentrations? Gauss's law (equation (4)) relates the potential of a closed surface to the quantity of charge inside the surface. Suppose that you place a cylinder around a DNA double helix and measure the charge of the mobile ions inside as function of the cylinder's diameter. The charge of the DNA should be completely screened at a sufficiently large distance; therefore, for a large diameter cylinder the charge of the mobile ions within should exactly cancel the DNA charge. A cylinder of a sufficiently small diameter contains no ions. For intermediate diameters, one might expect the charge of the ions within the cylinder to monotonically increase, approaching the total charge of the DNA, which is exactly what PB theory predicts. This prediction, however, does not always hold. At high bulk ion concentrations (2.5M), MC simulations using explicit ions and implicit solvent showed a rise in the ion charge beyond that needed to neutralize the DNA [143]. This effect is known as overneutralization or charge inversion and could be produced by several physical mechanisms [132, 144].

### 4.2. Simulations of ion atmosphere around DNA helices

DNA is a polyanion with a large local charge density; when we assume a cylinder of 1 nm radius for a DNA double helix, the

local DNA charge density is about 1.5M. For charge neutrality, a similar number of counterions is expected around DNA [37], and this diffuse layer of counterions is called the 'ion atmosphere' [145]. The physiologically relevant ions (e.g. $Na^+$, $K^+$, $Mg^{2+}$, $Ca^{2+}$, and polyamines) bind DNA competitively, playing an important role in DNA stability, structure and function [146]. In the following subsections, we discuss the recent experimental and simulation studies on the ion atmosphere and technical issues related to the simulations of divalent cations.

*4.2.1. Experimental and simulation studies of the ion atmosphere.* Experimental ion detection is still technically challenging and details of the counterion distributions around DNA remain elusive.

Currently, small-angle x-ray scattering (SAXS) is the most widely used experimental method for studying the counterion atmosphere. Using this method, Pollack and co-workers directly demonstrated the existence of the counterion atmosphere around DNA [147] and the competition of cations with different valences [148, 149]. However, the SAXS technique can only provide a qualitative description of the counterion distribution. Recently, Herschlag and co-workers developed a novel technique, buffer equilibration and atomic emission spectroscopy (BE–AES), that enables direct counting of ionic species condensed to DNA [126]. The quantitative ion counts from the BE–AES method can complement the qualitative spatial distribution data from SAXS, providing a more detailed view of the ion atmosphere.

The microscopic structure and dynamics of the ion atmosphere around DNA can be characterized using the MD method. Pioneering studies utilizing explicit solvent and the Ewald summation technique for long-range electrostatics have been performed since the mid-1990s. Young and co-workers reported that counterions could be found in high concentration within the DNA grooves, but they were not 'stuck' to the DNA [38]. Ordered $Na^+$ sites were observed in the major and minor grooves of DNA [40]. Early simulations suggested that the minor groove of AT DNA narrowed upon $Na^+$ binding, but later studies disagreed [150, 151]. We caution that minor groove binding sites may be artifacts of the force fields employed at the time. The ion-binding sites on the DNA required ~50 ns to be fully sampled, while the overall ion atmosphere and DNA conformation relaxed much more quickly around ~5 ns [152]. Because water also plays an important role for DNA structure in aqueous solutions [26], the orientation and density of water molecules around DNA, especially the minor groove, were scrutinized in the above studies.

Figure 4 illustrates the distributions of ions around a poly(dA)·poly(dT) DNA duplex in 170 mM and 320 mM NaCl electrolytes obtained in previously reported MD simulations [47]. $Na^+$ ions are found in both grooves of the DNA with the highest concentration just outside the DNA phosphate groups. Further away, the concentration of $Na^+$ ions rapidly decays while that of $Cl^-$ ions rises to the bulk concentration at about 30 Å from the DNA helical axis. The ions were found to be highly mobile and only loosely associated with the DNA. The Manning radius, defined to be the radius at which 75%

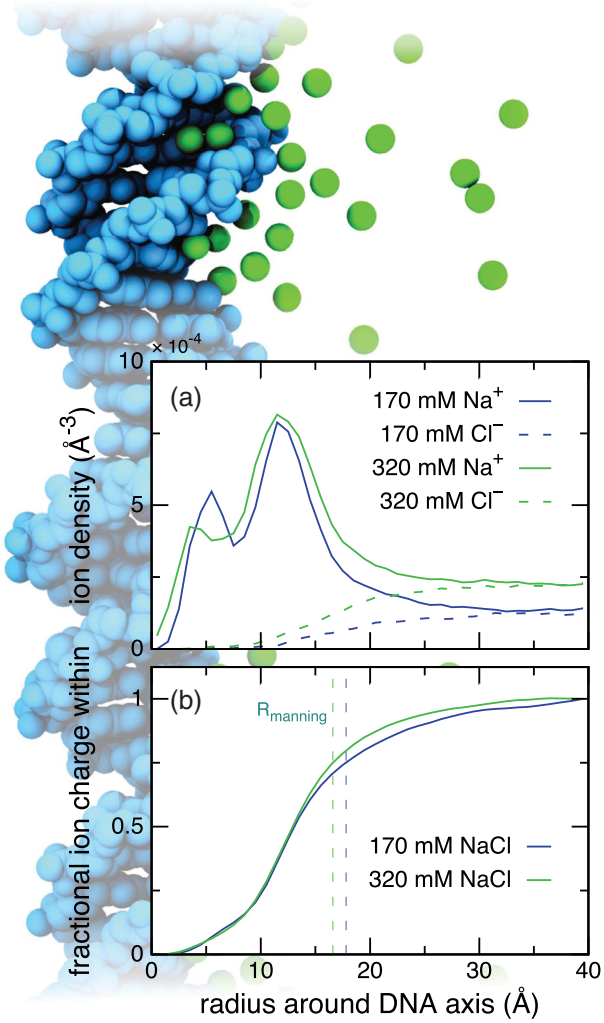


**Figure 4.** Typical distribution of monovalent ions around duplex DNA. (*a*) Density of $Na^+$ (solid) and $Cl^-$ (dashed) ions as a function of distance from the axis of a DNA molecule corresponding to bulk ion concentrations of 170 mM (blue) and 320 mM (green). (*b*) Charge of ions within the specified distance from the DNA axis, normalized by the DNA charge. Colors are as in (*a*). The Manning radius is indicated for each ion concentration as a vertical dashed line. The ion distributions were taken from trajectories reported in [47]. The background shows the DNA molecule (cyan) surrounded by ions (green) superimposed from several snapshots of an MD trajectory. The molecular image is arranged to scale with the data so that the axis of the DNA corresponds to the zero mark of the graph's $x$-axis.

charge neutralization occurred, was found to be about 16–17 Å and only weakly depended on the ion concentration in the range of 170–320 mM. However, the radius was found to be significantly larger for low ion concentration solutions; at 30 mM the radius was about 33 Å [47].

*4.2.2. The roles of divalent ions.* Divalent cations, such as $Mg^{2+}$ and $Ca^{2+}$, can have a dramatic effect on the structure and function of DNA and RNA [153]. According to their function, they can be described as 'coordinating' and 'diffuse' cations. The 'coordinating' divalent cations are found at active sites of enzymes where they catalyze chemical reactions [153]. Such divalent cations are essential for the activities of many tRNA

and ribozyme molecules [153–155], as well as ATP-dependent biological motors, and can be often resolved in crystal structures [153]. The function of 'diffuse' divalent cations is to reduce the internal stress in DNA or RNA structures by partially neutralizing the negative charge of the phosphate groups [156–158].

Developing a single set of parameters capable of describing the two functional roles of divalent cations remains a major challenge. The catalytic ability of divalent ions is largely due to their +2 charge, which can impart strong electrostatic forces on nearby negatively charged moieties such as water oxygen or anions. Simulations employing a polarizable force field demonstrated that the dipole moments of waters in the first solvation shell of $Ca^{2+}$ and $Mg^{2+}$ were respectively increased by 20% and 40% [159], an effect that is difficult to capture using non-polarizable force fields. The physical size of a divalent cation dramatically effects its properties. For example, the water exchange rates in the first and second solvation shells are significantly different for $Mg^{2+}$ ($\sim$10 $\mu$s) and $Ca^{2+}$ (<ns) [160, 161]. The difference cannot be explained by considering the individual interactions between the ions and the water molecules. Rather, water packing in the first solvation shell determines the slow kinetics for $Mg^{2+}$, which has exactly six water molecules (versus 7–8 for $Ca^{2+}$) that fit perfectly in a stable octahedral conformation that prevents other water molecules from intervening [160].

Thus, accurate description of divalent ions by MD requires realistic treatment of the polarizability and kinetics of surrounding water. The most straightforward approach is to employ a polarizable force field to allow the atomic dipole moments respond to the electric field of a divalent cation during a simulation. Several groups have already used development versions of polarizable force fields and obtained significantly improved agreement with experiment [159, 160, 162, 163]. However, a general purpose polarizable force field capable of describing large nucleic acid systems and long time scale (>100 ns) processes is presently not available.

Several approaches have been proposed to improve the realism of divalent cation simulations within the framework of standard CHARMM and AMBER force fields. Allnér *et al* optimized the LJ parameters of $Mg^{2+}$ and $Ca^{2+}$ so that they can be used directly with the CHARMM and AMBER force fields [164]. By performing a series of binding free energy calculations between $Mg^{2+}$/$Ca^{2+}$ and water/phosphate molecules, they demonstrated that the new parameter sets can reproduce water kinetics in the first solvation shell better than the standard force fields. Saxena and Sept proposed a novel multi-site model of $Mg^{2+}$/$Ca^{2+}$, in which these divalent cations are modeled as six dummy atoms arranged in an octahedron formation [165]. These two recent studies primarily focused on reproducing the structure and kinetics of the first solvation shell of $Mg^{2+}$ and $Ca^{2+}$ ions. However, proper coordination of the first solvation shell does not guarantee correct description of long-ranged interactions between $Mg^{2+}$ or $Ca^{2+}$ and other molecules such as DNA. Furthermore, the very long exchange time of water in the first solvation shell of $Mg^{2+}$ ($\sim$ 10$\mu$s [160]) suggests that the coordination chemistry of a typical $Mg^{2+}$ ion should remain unchanged during a typical MD simulation. For example, if a coordinating $Mg^{2+}$ ion is initially in direct contact with a DNA phosphate oxygen, it will remain bound to the DNA for the duration of the simulation. Thus, the coordination state of each crystallographically resolved $Mg^{2+}$ ion must be carefully considered at the set-up stage of an MD simulation [153].

When a fully solvated $Mg^{2+}$ ion approaches DNA, it is unclear whether it will retain its six coordinating water molecules or some of them will be exchanged for phosphate oxygens. Several lines of evidence suggest that the vast majority of $Mg^{2+}$ ions will remain forming a $Mg^{2+}$-hexahydrate complex ($Mg^{2+}(H_2O)_6$) shown in figure 3(*b*). First, measurements using fluorescence and thermal melting methods indicated that tertiary structure of an RNA pseudoknot does not have directly bound $Mg^{2+}$ [166]. Second, measurements of DNA pressure and inter-DNA spacing in a DNA array system in the presence of $Mg^{2+}$ and $Ca^{2+}$ have shown that $Ca^{2+}$ lowers the DNA array pressure slightly more efficiently than $Mg^{2+}$ does [73]. If (meta)-stable $Mg^{2+}$–DNA complexation were possible, $Mg^{2+}$ would bridge neighboring DNA helices better than $Ca^{2+}$, which would result in a lower internal pressure compared to $Ca^{2+}$. Third, if a $Mg^{2+}$ ion directly bridges the phosphate groups of two DNA helices, the inter-helical distance would be about 22–24 Å, which is inconsistent with experimentally measured distances of at least 25 Å, even at extremely high external pressure ($\sim$ 50 bar) [73]. Fourth, if specific direct binding of $Mg^{2+}$ to DNA phosphates were energetically favorable they would be crystallographically resolved. Finally, a $MgCl_2$ solution does not contain $Mg^{2+}$–$Cl^-$ pairs according to Raman spectroscopy [160].

Yoo and Aksimentiev proposed a model, in which a $Mg(H_2O)_6$ complex is treated as a stable molecular complex [59]. Such an approach is reminiscent of using cobalt hexamine, $Co(NH_3)_6$, a molecule having the same geometry as $Mg(H_2O)_6$, as an analog for $Mg^{2+}$ in single-molecule experiments [167]. Achieving quantitative agreement between simulated [59] and measured [73] DNA array data was only possible by increasing the dipole moments of water molecules forming the $Mg(H_2O)_6$ complex to the values observed using a polarizable model [159]. The biggest limitation of the magnesium hexahydrate model is that it requires, *a priori*, a different treatment of chemical moieties that bind $Mg^{2+}$ directly or indirectly. Extension of the model to $Ca^{2+}$ is not straightforward, as the first solvation shell of $Ca^{2+}$ is not as stable as $Mg^{2+}$, and hence, restraining 7–8 water molecules around $Ca^{2+}$ may cause artifacts. One possible solution to the above problems is to dynamically reassign parameters for water molecules as they move in and out of the first solvation shell of a divalent cation.

*4.2.3. Competitive binding.* Physiological ionic solutions usually contain various cation species. Of the monovalent ions, $K^+$ is more prevalent in cells than $Na^+$ and interacts with DNA differently; $K^+$ does not enter the minor groove of DNA [151, 168, 169] and $K^+$ ions diffuse about the DNA surface significantly more quickly than $Na^+$ [168, 169].

Early simulations employing high concentrations of $Mg^{2+}$ ions demonstrated that the DNA conformation shifted towards the A-form [170]. Later simulations showed that $Mg^{2+}$ binds

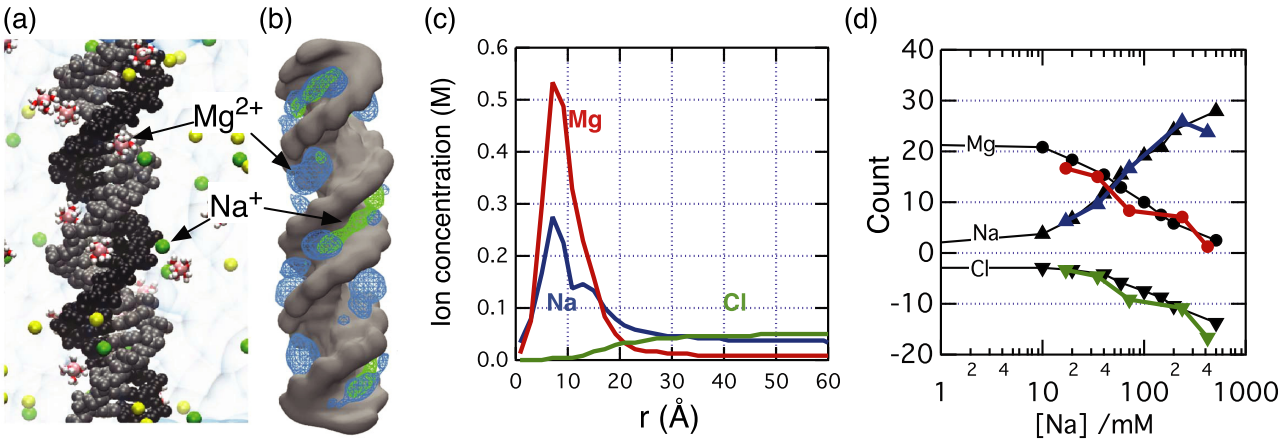


**Figure 5.** Competitive binding of $Mg^{2+}$ and $Na^+$ to dsDNA. (*a*) A representative configuration of ions near a 24 bp duplex. The two DNA strands are shown in black and gray; $Na^+$ and $Cl^-$ ions are shown as yellow and green spheres. The first solvation shell of each $Mg^{2+}$ (pink) ion is explicitly shown. (*b*) Volumetric representations of DNA (gray), $Mg^{2+}$ (blue) and $Na^+$ (green) obtained by averaging a ∼100 ns trajectory. The density maps are shown as isosurfaces of 0.5 and 0.02 $Å^{-3}$ atom number density for DNA and ions, respectively. (*c*) Ion concentration as a function of the radial distance from the DNA axis. Data in (*a*)-(*c*) correspond to bulk ion concentrations of ∼5mM $Mg^{2+}$, ∼40mM $Na^+$ and ∼50mM $Cl^-$. (*d*) Simulated excess number of $Mg^{2+}$ (red), $Na^+$ (blue) and $Cl^-$ (green) ions as a function of $Na^+$ concentration at the background of 5mM $Mg^{2+}$. For comparison, experimental data [126] are shown in black. Reprinted with permission from [125]. Copyright 2012 American Chemical Society.

to the DNA backbone and G·C base pairs preferentially, and that a DNA molecule was more rigid in $Mg^{2+}$ electrolyte than in $Na^+$ [171].

By utilizing the new NBFIX-enabled parametrization, Yoo and Aksimentiev could quantitatively describe the ionic atmosphere of DNA surrounded by two competing cation species [125]. Figure 5 illustrates the ion atmosphere in one such system containing a mixture of $Na^+$ and $Mg^{2+}$ ions. In a 100 ns MD trajectory, $Na^+$was found to bind predominantly to the minor grooves while $Mg^{2+}$ was found to bind to the minor and major grooves with equal affinity, figures 5(*a*) and (*b*). Due to the stronger binding affinity of $Mg^{2+}$, the local concentration of $Mg^{2+}$ near DNA is about twice that of $Na^+$, although the bulk concentration of $Na^+$ is about an order of magnitude higher than that of $Mg^{2+}$, figure 5(*c*). Chloride is completely excluded near the DNA due to electrostatic repulsions. Overall, all ion concentrations converge to their bulk values at around 30 Å from the center of the DNA, figure 5(*c*). By integrating the ion concentration profiles in figure 5(*c*), the researchers could evaluate the excess ion counts for each species, figure 5(*d*). The comparison between the computed and experimental ion count data shows quantitative agreements, indicating that the NBFIX parameters provide a realistic description of ion atmosphere [125]. For the competition between $Na^+$ and $K^+$, it was found that $K^+$ prefers the major groove while $Na^+$ prefers the minor groove. For the $Na^+$ and $Li^+$, both cations prefer binding to the minor groove. For all data set, quantitative reproduction of the experimental data [126] was demonstrated.

*4.2.4. Multivalent electrolytes.* Multivalent (valence of 3 or greater) polyamines, including spermine, spermidine and histone tails, play an essential role in genome compaction [172, 173]. Spermine (+4) was found to compete with $Na^+$ for binding DNA in the minor groove [174–177]. Other polymanines, putrescine (+2) and spermidine (+3) were found to bind DNA much like spermine with putrescine also binding somewhat in the major groove [175]. All three of these natural polyamines displaced water in the first hydration layer of the DNA [174, 176].

Theoretical developments beyond mean field treatments have resulted in the prediction of charge inversion of DNA. The strongly correlated liquid theory of Shklovskii [178, 179] predicted that higher valency counterions could result in DNA of net positive charge due to counterion condensation. MD simulations have been applied to investigate such a possibility [177]. An effectively infinite fragment of DNA was placed at the center of a nanochannel and surrounded by electrolyte containing various concentrations of $Na^+$, $Mg^{2+}$, spermidine (3+) and spermine (4+) cations. The systems were simulated using the MD method until the ion distributions around the DNA reached equilibrium. The possibility of charge inversion was examined using Gauss's law, equation (4). The total charge of DNA and the surrounding electrolyte enclosed within a cylinder co-axial with the DNA was plotted against the radius of the cylinder. For monovalent electrolytes, the effective electric charge of DNA and the surrounding cations monotonically increased reaching zero at large distances, indicating the absence of charge inversion. For some multivalent electrolytes, the charge became positive, indicating overscreening or charge inversions. The study further investigated the relationship of such electric charge inversion with the inversion of the electrophoretic mobility of DNA, finding that the two phenomena have different physical origins as discussed in section 7.3.4.

## 5. DNA mechanics

### 5.1. Equilibrium conformations

Why do two complementary DNA strands form a double helix? An obvious answer is the complementary hydrogen bonding

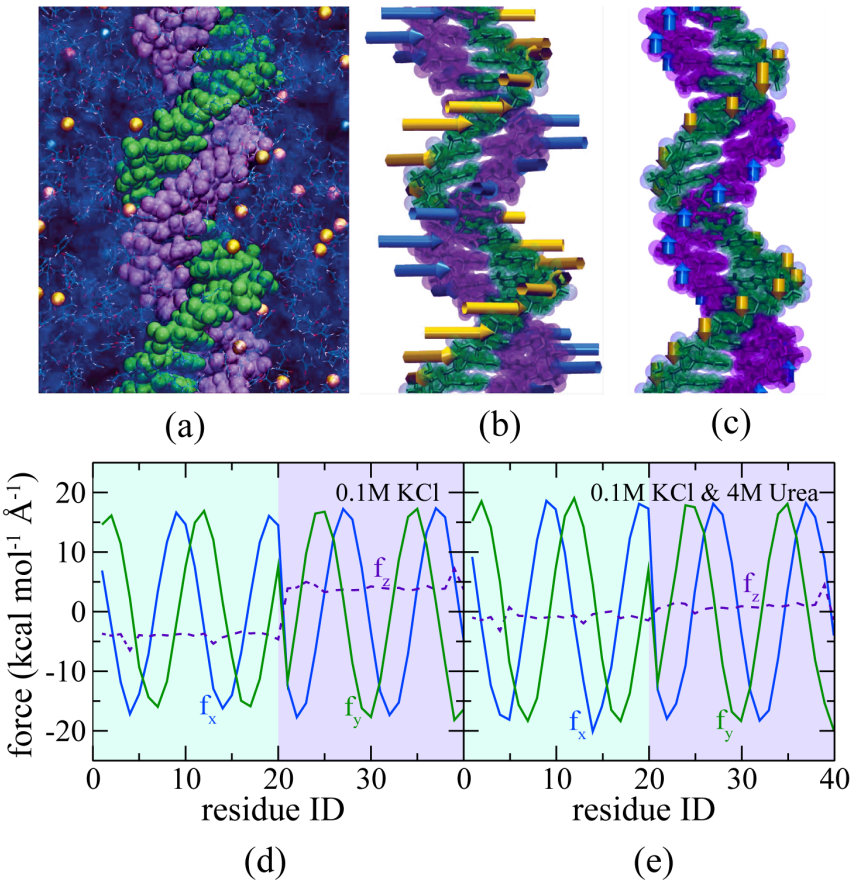


**Figure 6.** MD simulations of the solvent force on dsDNA. (*a*) Simulation set-up. A fragment of dsDNA (green and purple) is surrounded by an aqueous solution of $K^+$ (gold) and $Cl^-$ (cyan) ions. Bonds connecting atoms in each urea molecule are shown as lines; water molecules are not shown. (*b*), (*c*) Schematic representation of the lateral (panel (*b*)) and vertical (panel (*c*)) mechanical forces (represented by arrows) on each nucleotide in 0.1M KCl. (*d*) Cartesian components of the force acting on each nucleotide of the DNA at 0.1M KCl. (*e*) Same as in panel (*d*), but in the presence of 4M urea. In panels (*d*) and (*e*), nucleotides that belong to the one strand of dsDNA are numbered from 1 to 20 and to the other strand from 21 to 40; the force along the *z*-axis is shown using dashed lines.

of the DNA nucleotides from the two strands. However, DNA hybridization is a balancing act, requiring all the forces involved (hydrogen bonding, base-stacking, electrostatics and solvation) to have the right magnitude.

All-atom MD simulations should be able, in principle, to answer exactly that question, as the magnitude of the force applied to every atom of the simulation system is known at every step of a simulation. Nevertheless, studies reporting the average forces acting on individual components of biomolecules are rare as they require averaging the instantaneous forces as frequently as every simulation time step to provide statistically meaningful estimates of the average force [180]. Even when such averaging is done, extracting the contributions from forces of different origins is difficult, in particular when the electrostatic interactions are not calculated in pairwise fashion, which is typically the case when the PME method is used.

The force experienced by DNA nucleotides in a double helix from the solvent (water and ions) surrounding the DNA can be surprisingly strong. Figure 6 illustrates the forces experienced by individual nucleotides from the solvent in MD simulations performed by Luan and Aksimentiev. In these simulations, a 20 base-pair fragment of dsDNA (poly(dA)·poly(dT)) molecule was solvated in 0.1M KCl electrolyte with and without 4M urea added, figure 6(*a*). The 5′-end and the 3′-end of each DNA strand were covalently linked through the periodic boundary of the simulation system, forming an effectively infinite molecule. The simulations were performed using an older AMBER parameter set [106] and a custom version of NAMD [180]. The force between each nucleotide in the DNA and the entire electrolyte was calculated at every time step. For each atom of a nucleotide, the vdW and electrostatic forces exerted by the atoms of the electrolyte within 12 Å of that atom were recorded. The long-range Coulomb forces from the electrolyte atoms outside the 12 Å cutoff were extracted from the PME calculation. The total force on a nucleotide was then computed as a sum of the forces on the individual atoms of that nucleotide. The forces shown in figure 6 were averaged over MD trajectories tens of nanoseconds in duration. To verify the accuracy of the force calculation, the total force on a nucleotide from all other nucleotides of the DNA fragment was calculated. The latter force balanced the force on the same nucleotide from the surrounding electrolyte, as expected.

Figures 6(*b*) and (*c*) respectively illustrate the lateral (normal to the DNA axis) and vertical (along the DNA axis) forces from the solvent on individual nucleotides of dsDNA.

Each arrow points along the direction of the force and the length of the arrows represents the magnitude of the force. Figures 6(*d*) and (*e*) plot the Cartesian components of the forces. It is interesting to notice that the lateral force on each nucleotide is directed tangential to the DNA helix, applying a torque that intertwines the two DNA strands into a double helix. The vertical forces on both strands are in the opposite directions and push the two strands towards each other. The addition of 4M urea to the 0.1M KCl electrolyte considerably reduced the vertical forces, causing a wider minor groove of dsDNA. MD simulations of that kind might provide a mechanical account of urea's action as a denaturant of dsDNA.

In the above example, the DNA was linked to itself and hence was unable to change its conformation. However, if the bonds across the boundaries were removed, the DNA would lose its rod-like appearance and exhibit some degree of bending. Depending on the nucleotide sequence, fragments of dsDNA may exhibit different conformations, such as B-DNA or A-DNA. While interactions between nucleotides forming a DNA double helix are expected to affect the conformation, the interactions between dsDNA and the surrounding electrolyte also contribute. Melting-temperature measurements suggest that the dsDNA structure is more stable in high ion strength electrolytes.

The local bending of dsDNA is highly dependent on the DNA sequence. For example, dsDNA is more likely to bend at the AT step. The network of hydrogen bonds formed by water molecules at the surface of DNA varies in accord with the DNA sequence. For a poly(dA)·poly(dT) fragment, a water 'spine' is formed in the minor groove of the dsDNA, bridging the N3 and O2 atoms in neighboring base pairs. When a sequence contains 4–6 dA nucleotides repeated with helical periodicity, the DNA curves on average toward the minor groove by ∼10–20° per turn. These so-called A-tracts have been recently reviewed in outstanding detail [181]. The molecular mechanisms that underlie A-tract curvature are a challenge to determine because there are many possible causes, including solvent interactions, ion-binding, anomalous stacking or base-pairing. It is tempting to attempt to describe the A-tract curvature using a nearest neighbor model, such as the model that accurately describes the free energy of DNA hybridization as a sum of energies associated with pairs of stacked base pairs [91]. However, it was shown using MD simulation that the conformation of a CG base pair stack depended on the sequence context in the DNA [182]. Thus, it appears that the sequence-dependent attributes of DNA mechanics cannot be fully described by a nearest neighbor model, at least if we assume that the MD result was not due to artifacts in the force field.

At the length scale greatly exceeding the helical pitch of dsDNA, polymer models can provide an adequate description of dsDNA's equilibrium conformations. The WLC model mathematically represents a polymer of length $l$ by a continuous set of unit-length tangent vectors $\vec{t}(s)$ parametrized by the distance along the contour $s \in (0, l)$. The polymer is assumed to behave like a cylindrically symmetric beam with a mechanical rigidity described by an elastic modulus. Thus local bending of the polymer requires an energy d$u$ proportional to $|\frac{\partial \vec{t}}{\partial s}|^2$. In thermodynamic equilibrium $\frac{\partial \vec{t}(s)}{\partial s}$ should sample from the Boltzmann distribution and it can be shown by integrating over all possible polymer configurations that the expected value for $\langle \vec{t}(s) \cdot \vec{t}(0) \rangle = \mathrm{e}^{-s/L_p}$ [9]. Thus, the tangent vector of a polymer is expected to be correlated for a characteristic contour length of $L_\mathrm{p}$, the persistence length. For DNA, a range of values for $L_\mathrm{p}$ have been obtained, centered around 45–50 nm by fitting the WLC model to a wide variety of experimental measurements [183]. Finally, it is worth noting that the WLC model has been extended to incorporate elastic stretching as well as twisting [184–186].

### 5.2. Mechanical stretching

The mechanical properties of dsDNA have been the subject of extensive study [45, 62, 189–194] because of their fundamental importance to gene regulation processes in biological cells. The first single-molecule measurements of the elastic properties of DNA were performed nearly 20 years ago and revealed that dsDNA can be mechanically melted to form ssDNA [45]. At low forces (below 20 pN) stretching of dsDNA is accurately described by the WLC model, which captures the entropy-driven contraction of dsDNA.

Single-molecule manipulation experiments revealed a characteristic plateau in the force–extension curve of dsDNA that signifies a highly cooperative transition from a canonical B-DNA structure to an overstretched S-DNA conformation, the so-called B-to-S transition [188]. Despite extensive experimental and theoretical studies, the nature of the B-to-S transition remains controversial [189]. MD simulations revealed that stretching DNA with a force transforms the canonical B-DNA structure into a ladder-like conformation [188, 190]. However, subsequent thermodynamic analysis of DNA stretching suggested that the 'B–S' plateau indicates a melting transition, i.e. separation of the two DNA strands that occurs at the beginning of (and throughout) the plateau [191]. Furthermore, both theory [192] and experiment [193] suggest that depending on the twist of the DNA helix, i.e. the number of base pairs per turn of the helix, several DNA conformations may coexist during the transition.

A single strand of DNA behaves differently under applied force. Typically, the force–extension curve of ssDNA does not exhibit an overstretching plateau [194]. However, a plateau reminiscent of that observed in stretching dsDNA can be observed for homopolymeric (such as poly(dA)) single-stranded DNA [195] and RNA [196] (for example, poly(A)), suggesting that base-stacking in such molecules may enforce a helical secondary structure. In general, single-molecule stretching of ssDNA can be challenging to interpret because of a large number of secondary structures that a long random-sequence ssDNA may form during the measurement [92].

*5.2.1. Steered molecular dynamics.* Complementing experimental studies of DNA mechanics, stretching of DNA has been studied using MD simulations. Perhaps the most straightforward way to model single-molecule stretching is to apply a constant force to one end of the DNA while holding the other end fixed. Despite its simplicity, this method is seldom used.

Although a large enough force is guaranteed to produce pronounced stretching, it is almost always irreversible and requires force magnitudes that greatly exceed those applied in experiment. A small force, in contrast, will produce only modest elastic deformation. For an activated process, which typically describes forced unfolding of a biomolecule, the likelihood of observing structural transition is an exponential function of the applied force. Hence, it is very difficult to 'dial in' the right magnitude of the force to observe the stretching process that maximizes the realism of the simulation given the limited duration of an MD trajectory.

In contrast, the steered molecular dynamics (SMD) approach [197], guarantees the observation of a deformation process of a prescribed magnitude within the available simulation time. Figure 7(*a*) illustrates this method using a fragment of ssDNA. In this simulation, a homopolymer of 20 adenine nucleotides was stretched in a 0.1M NaCl solution. The coordinates of the O3′ atom at the 3′-end were fixed whereas the O5′ atom at the 5′-end was subject to a harmonic potential (spring constant $k = 10\,\mathrm{pN\,Å^{-1}}$) mimicking the action of a virtual spring. The other end of the virtual spring was pulled at a constant velocity of $2\,\mathrm{Å\,ns^{-1}}$. At any time during the simulation, the force applied to the 5′-end of the DNA could be computed from the extension of the spring. The extension of the ssDNA was obtained as the distance between the terminal O5′ and O3′ atoms.

Starting from the initial conformation of a B-form DNA (figure 7(*b*)), the DNA strand was stretched by about twofold in 30 ns. In apparent agreement with experiment [195], the simulation revealed a force–extension plateau, figure 7(*b*). In the plateau region, the base stacking was observed to have breaks, as illustrated by the inset to figure 7(*b*) (at $L = 1.5L_0$). At that stage of the stretching process, ssDNA lost its helical structure, however many consecutive bases remained stacked. When overstretched ($L = 1.9L_0$), the backbone of the ssDNA was nearly straight. The stacking of neighboring bases in this stretched ssDNA was weak and could transiently form and break.

While it is tempting to describe the above simulation as successful, it actually illustrates several potential pitfalls of the MD method. First of all, the force field used for the above simulation (AMBER-94) poorly describes the conformation of ssDNA [112], which probably explains helical base stacking in partially stretched ssDNA. Second, the force plateau could also be observed in simulated stretching of poly(dT) strands, which was not observed in experiment [195]. Finally, the magnitude of the forces required to stretch ssDNA is significantly greater than the magnitude of the forces applied in experiment. This is perhaps the largest drawback of the SMD method and is a direct consequence of the short (in comparison to experiment) simulation timescale.

In general, ssDNA exhibits dynamics on timescales that are long compared to the all-atom MD timescale. Thus, the timescale of end-to-end collisions for a 20 nt dT fragment was experimentally measured to be 800 ns [198]. It is therefore difficult to sample ssDNA conformations adequately in atomic resolution simulations, especially when the solvent is represented explicitly. For many applications, the need to enhance sampling efficiency warrants the use of CG representations of DNA. Most CG DNA models have been optimized to represent the properties of double-stranded DNA [80–90], with only a few models capable of reproducing physical properties of ssDNA [80, 92, 199].

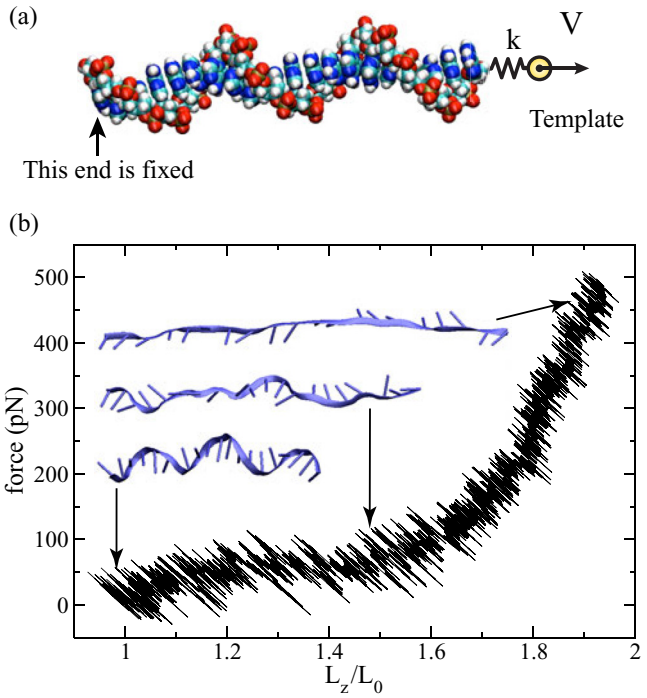


**Figure 7.** SMD simulation of ssDNA stretching. (*a*) Set-up of the simulations. A single DNA strand (vdW spheres) is placed in electrolyte solution (not shown). One end of the DNA is fixed in space; the other end is coupled to a template particle through a harmonic potential ($k$ is the spring constant of the potential). Moving the template with a constant velocity $V$ stretches the DNA. The difference between the coordinates of the template and the end of the DNA reports the applied force. (*b*) The force–extension dependence of a poly$(dA)_{20}$ fragment simulated using SMD. The inset shows three representative conformations of ssDNA corresponding to different amounts of stretching ($L = L_0$, $1.5L_0$ and $1.9L_0$), respectively. $L_0$ denotes the extension of a single DNA strand in a DNA double helix.

*5.2.2. Anisotropic pressure control.* As an alternative to SMD-type simulations of short dsDNA fragments [152, 188, 190, 200, 201], Luan and Aksimentiev suggested an anisotropic pressure-control method to stretch an effectively infinite DNA molecule [202]. Figure 8(*a*) illustrates a typical set-up of such a simulation. A fragment of dsDNA, containing two helical turns, is submerged in 1M KCl electrolyte. In each strand of the DNA helix, the backbone of the first nucleotide is covalently linked to the backbone of the last nucleotide over the periodic boundary of the system. This simulation set-up avoids the effect of free ends and is particularly suited for simulations of dsDNA stretching at a constant twist.

To stretch DNA, simulations are performed under anisotropic pressure conditions maintained using the Nosé–Hoover Langevin piston pressure control [203]. While the $P_{xx}$ and $P_{yy}$ components of the pressure tensor are kept at 1 bar, the $P_{zz}$ component is set to a negative value, stretching the

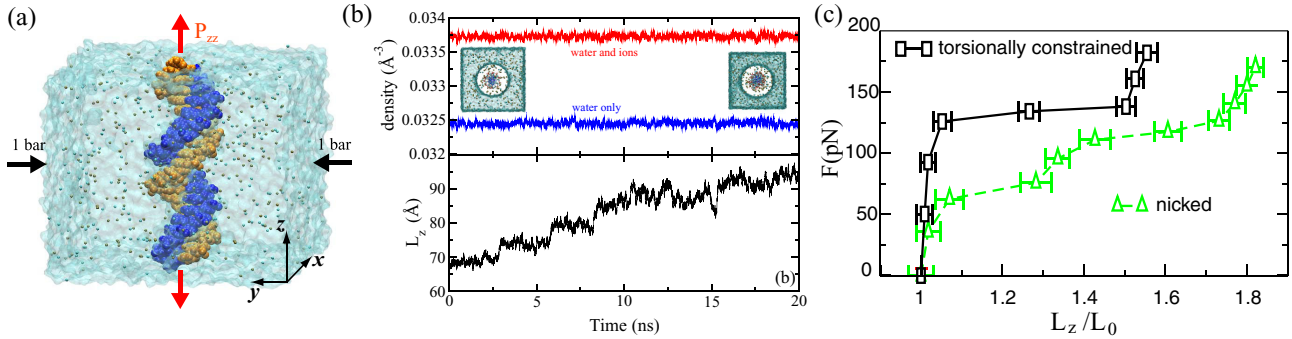


**Figure 8.** Stretching DNA using anisotropic pressure control. (*a*) Simulation set-up. A fragment of a DNA helix (orange and blue) is placed in a rectangular box of water (semi-transparent surface) and surrounded by potassium (tan) and chloride (cyan) ions. The strands of DNA are linked to themselves across the periodic boundary. The lateral pressures ($P_{xx}$,$P_{yy}$) are maintained at 1 bar, while the longitudinal (along the DNA helical axis) pressure ($P_{zz}$) is set to a negative value to stretch DNA. (*b*) A DNA overstretching process, simulated using anisotropic pressure control. (Top) densities of the bulk water and electrolyte away from DNA versus simulation time. (Bottom) the length of the DNA fragment versus simulation time. (*c*) Simulated force–extension dependences of torsionally constrained and nicked dsDNA molecules. The figures are adapted from [202], with permission. Copyright 2008 by the American Physical Society.

system in the $z$ direction until the total internal stress in the simulated system balances the applied pressure $P_{zz}$. During this process, the cross-sectional area of the simulation system decreases as the water density remains constant, figure 8(*b*). A typical process of dsDNA overstretching is illustrated at the bottom panel of figure 8(*b*). Repeating such simulations at different negative values of $P_{zz}$ yields the dependence of stress versus strain.

Assuming the pressure in bulk water is isotropic, the $zz$ component of the water stress tensor $P_{zz}^{\text{water}} \approx P_{xx} \approx 1$ bar. From the balance of the applied external pressure $P_{zz}$ and the internal stress in DNA and water, the tensile force inside DNA can be written as

$$F = -(P_{zz}S - P_{zz}^{\text{water}}(S - S_{\text{DNA}})), \tag{10}$$

where $S$ and $S_{\text{DNA}}$ are the cross-sectional areas of the simulation system and DNA, respectively, in the $x$–$y$ plane. Because the product of $P_{zz}^{\text{water}} S_{\text{DNA}}$ is less than 0.5 pN, equation (10) to compute the tensile force in DNA can be simplified as

$$F = -(P_{\text{zz}} - 1) \times S. \tag{11}$$

Thus, the force–extension dependence of a DNA molecule can be computed from the stress–strain relation of the entire system.

Figure 8(*c*) shows the force–extension curves obtained using the anisotropic pressure control method for torsionally constrained poly($dA_{10}dC_{11}$)·poly($dT_{10}dG_{11}$) and nicked poly($dC_{21}$)·poly($dG_{21}$) dsDNA. Each force–extension curve shown in figure 8(*c*) has three characteristic regions: elastic, transition and overstretched. The force required to initiate the overstretching transition is in good qualitative agreement with experiment. When torsionally unrestrained, λ-DNA exhibits an overstretching plateau at a tension of about 70 pN and that the tension slightly increases in the plateau region [188, 194]. For torsionally constrained DNA, experiment has shown a significantly higher yield force of about 110 pN [193]. Direct quantitative comparison, however, is difficult because of the heterogeneous sequence of λ-DNA. A detailed analysis of the sequence effects can be found in [202].

### 5.3. Twisting dsDNA

A DNA molecule can be bent, stretched along its helical axis, or twisted about this axis. Measurements of the extension of a double-stranded DNA molecule under low applied force (<10 pN) probes the mechanical bending of DNA [204]. At higher force (but still below the overstretching transition), the same measurements can characterize elastic stretching along the helical axis [204]. Such experiments are usually performed using optical traps which allow the DNA to relax torsionally. In biological cells, torsional stress in DNA was shown to regulate gene expression [205, 206].

Usually, the term flexibility refers to a polymer's ability to bend, which is quantified by its persistence length. For a generic dsDNA molecule, the persistence length is ∼45–50 nm [183, 62] or about 20 times its diameter (2.5 nm), which makes DNA fairly stiff for a natural polymer. In direct analogy to the persistence length, one can define a torsional persistence length for dsDNA that represents the characteristic length scale along which a DNA molecule's twist is correlated. This length is similar to the persistence length, but varies considerably in the literature from 36–75 nm [183, 207, 208].

Analysis of the crystal structures suggests that the intrinsic twist of a DNA helix depends on its nucleotide sequence and can range from 31° for an AG step to 40° for a TA step [209, 210]. Here we follow the usual convention that the first letter is on the 5′-end of one strand, and we note that some sequences are 'palindromes' (e.g. AA with TT). The standard deviation of the twist is similar for different base pair steps, ranging from 3.9° for an AA step to 6.5° for a CA step, suggesting a mild dependence of the torsional persistence length on the DNA sequence. The sequence-dependent average twist and twist moduli were obtained from MD simulations of DNA fragments [207] using the AMBER force field [106] without the parmbsc0 refinement. The results of the MD simulations were in overall agreement with the results of the crystal structure survey [210] with some discrepancies, primarily for GG steps. CG steps were found to be the most conformationally flexible, and AT steps the least. Overall, the study demonstrated significant

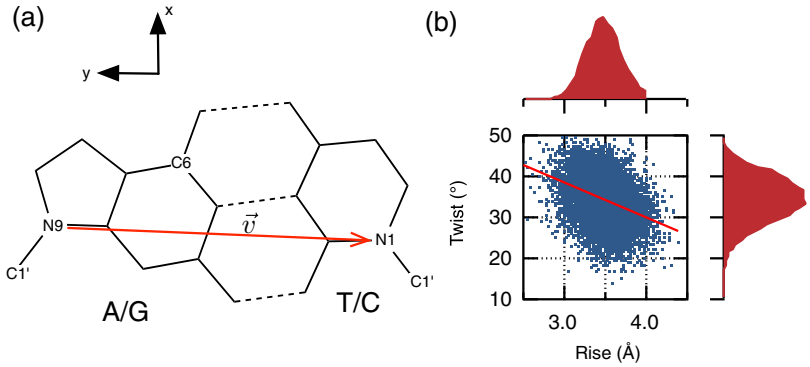


**Figure 9.** MD simulation of twist-stretching coupling. (*a*) Definition of a base pair geometry. The orientation of a base pair is characterized using a plane passing through atoms N9, N1 and C6, and a vector, $\vec{v}$, connecting atoms N9 and N1. (*b*) Scatter plot of instantaneous twist versus instantaneous rise from a 50 ns simulation of dsDNA. The red line shows a linear fit to the scatter plot, indicating slight ($R^2 = 0.12$) anticorrelation of twist and rise.

sequence-specific coupling between various coordinates used to describe base pair steps.

A more recent simulation study employed the newer AMBER-99 force field and a much larger, systematic set of sequences to examine the flexibility of specific base steps [211]. Despite reporting on overall flexibility, which included contributions from both twist and bending, the results of the study were consistent with previous observations, assessing the TA, CG and CA (but not AT, GC and AC) steps as the most flexible. The authors emphasized a correlation between the flexibility of a step and the absence of water molecules in the minor groove that bridge the base pairs forming the step.

Simulations of longer DNA molecules have also been performed. Mazur has studied many DNA properties through simulations employing an all-atom MD method that permits the integration time step to be increased from 1–2 to 10 fs by treating groups of DNA atoms rigidly [212]. That method enabled long-timescale (>160 ns) simulations of short (<2 turns) DNA fragments that revealed a torsional persistence length of 120 nm for AMBER-98 [213]. In contrast however, Noy and Golestanian [214] more recently found the torsional persistence length to be 80 nm from simulations of relatively long (3–5 turns) random sequence DNA molecules using the AMBER-99 force field with the parmbsc0 correction. The disagreement resulted in a debate between the two groups [215, 216], which may partially be attributed to the rather different simulation protocols employed.

A number of elastic models were developed that incorporated twisting and bending of DNA [184–186]. In general, elastic deformations of DNA may be coupled and some elastic responses may be asymmetric. In particular, twist–bend coupling and asymmetric bending were included in an elastic model that ignored stretching [217]. By contrast, single-molecule experiments demonstrated that stretching and twisting are counterintuitively coupled at low force (under 30 pN) with the DNA twisting *more* when stretched [218]. This led to a refinement of the extensible WLC model that incorporates twist and twist–stretch coupling [219].

Presently, MD simulations cannot reproduce the experimentally observed stretch–twist coupling. Figure 9 illustrates our analysis of the MD simulations reported in [125]. In those simulations, performed using the CHARMM force field [220], a 24 bp dsDNA was free to both rotate and extend. To compute the twist, we defined the orientation of each base pair, $\vec{v}$, according to the standard reference frame [221], figure 9(*a*). For two consecutive base pairs $n$ and $n+1$, the twist is $\arccos\left(\frac{\vec{v}_n \cdot \vec{v}_{n+1}}{|\vec{v}_n||\vec{v}_{n+1}|}\right)$. The rise is computed as the average distance between the plane of N9, N1 and C6 atoms of the base pair $n$ and those three atoms of the base pair $n+1$. Figure 9 shows a scatter plot of the instantaneous twist versus instantaneous raise computed from a 50 ns MD trajectory at a mixture of 50 mM NaCl and 50 mM KCl. A linear regression fit to the scatter plot reveals a weak anticorrelation of the twist with extension, which is opposite to the dependence observed in single-molecule experiments [218, 222, 223].

When DNA is twisted enough, linear elastic models fail to describe its properties. This was demonstrated in MD simulations where a three-turn DNA molecule was twisted by −0.22 to +0.39 turns [224]. Initially, the DNA had uniform twist, but by the ends of a 10 ns simulation it was slightly undertwisted or extremely overtwisted; the DNA conformation included a mixture of B-DNA and locally denatured DNA accommodating the overall twist. The MD simulations also indicated that counterions were able to access the minor groove more easily in overwound than underwound DNA.

### *5.4. DNA looping*

Because cells can suppress expression of genes by looping DNA, the behavior of DNA in small loops has long been of interest. Under thermal fluctuations or applied force, DNA bends away from its equilibrium curvature. For moderate bending, the WLC model describes this bending very well with a 45–50 nm persistence length. Early simulations were not long enough to allow accurate determination of the persistence length of DNA. However, the studies from Mazur [213] and from Noy and Golestanian [214] reported persistence lengths of 65-85 nm for AMBER-98 and 43–51 nm for AMBER-99bsc0, respectively.

The likelihood of DNA having complementary single-stranded overhangs to form a loop [225], or to cyclize, depends

on both the bending and twisting properties of DNA. As a function of DNA length, this likelihood has a peak around 400 bp, where the DNA is neither too long (so that the ends cannot find each other), nor too short (so that the DNA is unable to bend into a loop). This assay has been used to unambiguously distinguish intrinsic sequence-dependent curvature from isotropic flexibility [225]. For DNA molecules shorter than 500 bp, the likelihood of DNA cyclization was found to fluctuate periodically with respect to the length of the DNA fragment [62].

Under high curvature, the WLC model must break down, however, there is much disagreement with respect to the critical curvature where the model loses its validity. It was experimentally observed that very short DNA fragments (∼100 bp) could form circles more readily than predicted by the elastic models of DNA [208], suggesting strong bending or kinking. Similar results were observed from atomic force microscope images [1] and from the interference of x-ray scattering off gold nanoparticles tethered to either end of a short DNA molecule [226]. Other experiments rigorously defended the validity of the WLC model for short DNA [62, 63, 227, 228]. The controversy continues as DNA was seen to be highly flexible in recent melting experiments of curved DNA [229] and in single-molecule experiments where the looping of 100 bp DNA fragments was monitored by fluorescence [3]. It is interesting to note that most experimental studies that reported sharp bending of DNA employed either divalent electrolytes or high concentrations of monovalent electrolyte [3, 229, 1]. In the nucleus, counterions with valence $> 2$ strongly screen electrostatic interactions, so studies of DNA bending at elevated concentrations of monovalent electrolyte may still be biologically relevant.

All-atom simulations demonstrated spontaneous kinking of initially well-formed 94 bp DNA minicircles with two different linking numbers that were expected to flank the torsionally relaxed state [230]. The kinks broke stacking interactions with bending of 90–120° towards the minor groove. A more recent study involved shorter simulations (10–30 ns) of homopolymeric DNA minicircles under various degrees of torsional strain [231]. In that study, kinks and duplex melting were observed in torsionally strained DNA, but not in the simulations of relaxed DNA, perhaps because of the short durations of these simulations (10 ns).

Given that the atomic models employed (AMBER-94/99) significantly overestimate the base-stacking interaction free-energy [129], the above studies appear to suggest that DNA is likely to be kinked in 94 bp circles. However, caution should be taken in jumping to this conclusion; the force field likely enhances the free energy required to form a kink, but it also may enhance the free energy needed to form continuous curves. Whether a DNA circle kinks in MD simulation depends on the relative magnitude of these free energies. This discussion then begs the question of whether the AMBER force field overestimates the bending rigidity of unkinked DNA. Unfortunately, both experiments and all-atom MD simulation are locked in debates that pertain to the likelihood of looping of short DNA sequences.

### *5.5. Nucleosomes*

Sharply bent DNA structures naturally occur in nucleosomes—the storage units of eukaryotic genomes. In a typical nucleosome, a 150-bp-long DNA segment wraps almost twice around a protein core called a histone octamer complex. The interaction of DNA with the histone proteins is thought to play a critical role in gene regulation [232]. When the existence of nucleosomes was first established, the DNA in nucleosomes was expected to be kinked (e.g. [16]). However, subsequent crystallographic studies revealed that DNA can wrap around a histone octamer without disrupting its base-pairing pattern [232]. Interested readers are referred to a recent review on this subject [232].

For a eukaryotic gene to be transcribed, nucleosomal DNA must be unwrapped from the histone core. Not surprisingly, a number of single-molecule techniques have been applied to study the internal mechanics of nucleosomes [233–236]. Despite extensive efforts, the structural details of the unwrapping process remain elusive. Several research groups have used computer simulations to study the structure and dynamics of nucleosomes. Until recently, all-atom MD studies of nucleosomes have been tempered by the relatively large size of the simulation system. By far a more serious obstacle is the slow dynamics of unstructured parts of the histone proteins (so-called histone tails), which strongly interact with DNA because of their polybasic nature. Consequently, CG approaches have historically prevailed [237–239]. However, several all-atom MD simulations of nucleosomes have been reported already [48, 240]. For example, Biswas *et al* reported a 100 ns simulation that characterized interactions of histone tails with dsDNA [48].

A potentially interesting approach is to use the SMD method to simulate forced unwrapping of nucleosomal DNA [240]. If performed properly, such a simulation could provide a structural interpretation of the single-molecule pulling experiments [233–236]. In the simulation featured in figure 10, unwrapping of dsDNA was produced by increasing the length of a virtual spring whose ends were anchored at the centers of mass of the terminal 10 bp fragments of the DNA. Initially, the equilibrium length of the spring was increased at a rate of 1 nm ns$^{-1}$, which was reduced to 0.2 nm ns$^{-1}$ after 18 ns. While pulling the ends of the nucleosomal DNA apart, we observed symmetric unwrapping of the DNA conformation until approximately 40 ns, figure 10(*d*). After 44 ns, local melting of dsDNA was initiated, figure 10(*e*). Thereafter, only the propagation of DNA melting, not DNA unwrapping, was observed. Such a disappointing outcome is perhaps not surprising, as the SMD force exceeded 100 pN at the moment of the DNA melting, figure 10(*f*). This simulation indicates that a significantly slower SMD pulling will be required to properly reproduce experimental unwrapping of nucleosomes.

## 6. DNA–DNA interactions

Direct, physical interactions between DNA molecules are central to many fundamental biological processes and underpin emerging DNA nanotechnology. Below we describe the surprising range of DNA–DNA forces that sensitively depend on

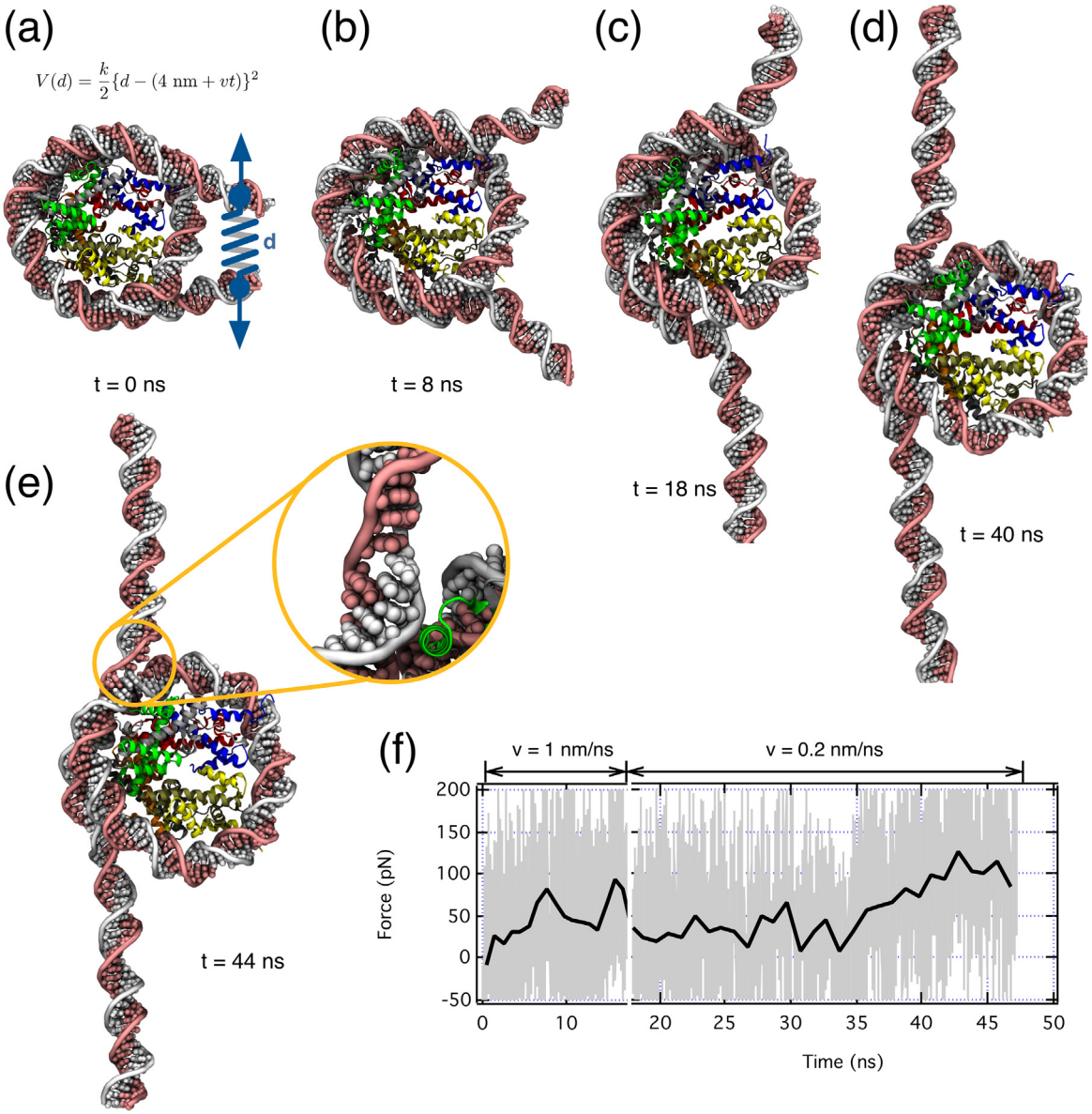


**Figure 10.** SMD simulation of nucleosomal DNA unwrapping. (*a*)–(*e*) Representative conformations of a partially unwrapped nucleosome at several instances of the SMD trajectory. To realize the SMD protocol, a harmonic spring potential was applied between the centers of mass of the terminal fragments (10 bp) of DNA. The equilibrium length of the spring was increased at a rate of $v$ to unwrap the nucleosome. The inset in panel (*e*) illustrates local melting of DNA produced by the excessive SMD force (∼100 pN). (*f*) SMD force as a function of the simulation time. Note that the pulling rate was reduced by a factor of 5 at $t \sim 18$ ns to facilitate structural relaxation.

the properties of the interacting molecules, e.g. their mutual orientation, as well as the properties of the environment, e.g. electrolyte concentration. The discussion below does not cover DNA–DNA forces mediated by proteins.

### *6.1. DNA hybridization*

The most fundamental DNA–DNA interaction is hybridization, which is the formation of a double helix by non-covalent self-assembly of two ssDNA molecules carrying complementary nucleotide sequence [6]. Despite tremendous interest in this fundamental problem, fully atomistic simulations of DNA hybridization remains a challenging problem because sampling the conformational space of ssDNA is beyond current computational power. For example, single-molecule experiments have shown that hybridization of a ∼10 bp DNA molecule at mM concentration occurs on the millisecond-timescale [241]. Instead, all-atom MD simulations have focused on the observation of single base binding–unbinding events [242] or folding of short hairpins (about 4 bps) [129, 245–248]. Currently, the most practical solution for simulations of DNA/RNA hairpin folding is the replica exchange MD (REMD) technique, which enhances sampling of a thermodynamic ensemble by accelerating energy barrier crossing using coordinate-swapping among ∼100 replicas each held at different temperature [129, 244, 246].

### 6.2. Side-by-side repulsion of duplex DNA

DNA molecules carry a high negative charge and hence electrostatically repel one another unless the DNA charge is compensated by polycationic condensing agents, such as cobalt hexamine (4+), spermidine (3+), spermine (4+), polylysine or polyarginine [73, 172, 247, 248, 249]. While the effective charge of a macromolecule can be theoretically determined from the force between a point charge and the macromolecule, in practice such forces are difficult to measure. Computational methods offer this capability (section 4.2.4), but their validation requires connection to experimental observables.

A related, experimentally attainable approach is to determine the effective force between charged macromolecules. Quantitative measurement of the force between DNA molecules as a function of inter-DNA distance was not possible until Parsegian and co-workers developed, about 30 years ago, a method to condense DNA using osmotic pressure of polyethylene glycol (PEG) polymer [73, 250]. In that seminal work, the authors found that condensed DNA arranged in a hexagonal lattice pattern. The study established that inter-DNA forces are generally repulsive in the presence of monovalent and divalent cations. The only two exceptions are manganese ($Mn^{2+}$) and cadmium ($Cd^{2+}$) ions that can induce DNA condensation in a specific temperature range [251], see section 6.3. Furthermore, they observed that the inter-DNA repulsion decays exponentially at short distances regardless of cation types and argued that 'hydration forces', not electrostatic forces, are at the origin of the cation-independent repulsion [73].

In principle, the force between two DNA molecules can be estimated from the non-linear Poisson–Boltzmann (NLPB) equation (section 4.1.1). The PB equation is non-linear so the ionic densities surrounding two DNA molecules in isolation do not add trivially when the molecules are brought into proximity. Therefore, one can expect that the effective charge of DNA determined by measuring the interaction with a point charge may not describe the interaction between two or more DNA molecules. Nevertheless, it is useful to consider an effective charge for the purpose of modeling these interactions. Additionally, an effective charge highlights deviations between the behavior of real DNA and that predicted by idealized models. As we shall see, the effective charge of DNA as determined through DNA–DNA interactions can depend on the ionic strength and the DNA–DNA separation. One should, however, keep in mind that the NLPB equation neglects many effects including hydration, finite ion size, correlations in the ion atmosphere and local dielectric response of the medium [252–254].

Prior to the availability of the ion–phosphate NBFIX parameters, several all-atom MD simulations quantified the repulsion between parallel dsDNA molecules [47, 255, 256]. The studies demonstrated significant repulsion with forces around 5–10 pN/turn between parallel DNA molecules having center of mass separations between 25 and 30 Å around physiological KCl or NaCl concentrations. Replacing potassium ions with sodium ions was found to decrease the repulsion [255], in qualitative agreement with experimental measurements of DNA condensation under osmotic pressure [257].

Using the NBFIX corrections for ion–phosphate interactions (described in section 3.2.2), all-atom MD simulations could quantitatively reproduce the experimental inter-DNA distance and the osmotic pressure in DNA arrays [59]. In an MD simulation, various ionic species of the same valence differ from one another solely by their mass, the vdW radius and the well depth of the vdW potential. A custom NBFIX correction for a particular ion–phosphate pair selectively increases the vdW radius *just for that pair interaction*. For monovalent ions, the adjustment needed to bring simulation into agreement with experiment was larger for smaller ions. Potassium ions, which have the largest vdW radius of common ions, needed the smallest adjustment. Hence the results from the previous studies performed without the NBFIX corrections are likely to be the most accurate in the case of KCl electrolytes.

In the dilute regime, repulsive DNA–DNA interactions can be experimentally studied by several means. For example, small angle x-ray scattering experiments can provide an average measure of the pairwise interaction potential between particles in a solution, and demonstrate overall repulsion of long DNA fragments in mono- and divalent electrolyte. Similarly, forces involved in packaging viral capsids have been related to DNA–DNA repulsion [258]. Another system for the study pairwise DNA interactions can be created by twisting a long DNA molecule tethered to a magnetic bead and a stationary surface, whereupon the DNA will buckle forming a twisted loop called a plectonemic supercoil, figure 11. The two halves of the plectonemic loop are wrapped closely around one another even when DNA–DNA interaction is repulsive. As the DNA molecule continues to be twisted, DNA is transferred from the tether region to the plectoneme, which grows in an ion-concentration dependent manner. Given a functional form, the strength of the electrostatic interaction can be inferred from the rate of plectonemic growth.

In a collaborative study of experimentalists and modelers, DNA–DNA interactions within a plectoneme were modeled as that of uniformly charged cylinders interacting through a potential obtained from the linearized PB equation scaled by a charge reduction factor (CRF) [47]. The CRF was used as an adjustable parameter to fit theoretical predictions to experimental data, yielding a value of 0.42 for the CRF in 1 : 1 electrolyte over a wide range of concentrations. All-atom MD simulations verified that a CRF of 0.42 improved the linearized PB prediction of the force between parallel dsDNA molecules. This CRF contrasts with that from Manning counterion condensation theory, demonstrating that the effective charge of DNA is a process-specific quantity.

### 6.3. Side-by-side attraction

In all kingdoms of life, the genome is highly condensed for efficient and safe storage and gene regulation. The simplest life forms where such compaction is observed are DNA viruses such as bacteriophages. When the genome of a bacteriophage is fully packed inside its capsid, the inter-DNA distance is less than 30 Å and the nucleotide concentration easily exceeds 1M. Eukaryotic genomes are also densely packed inside the nucleus. For example, the 3 billion base pair long human

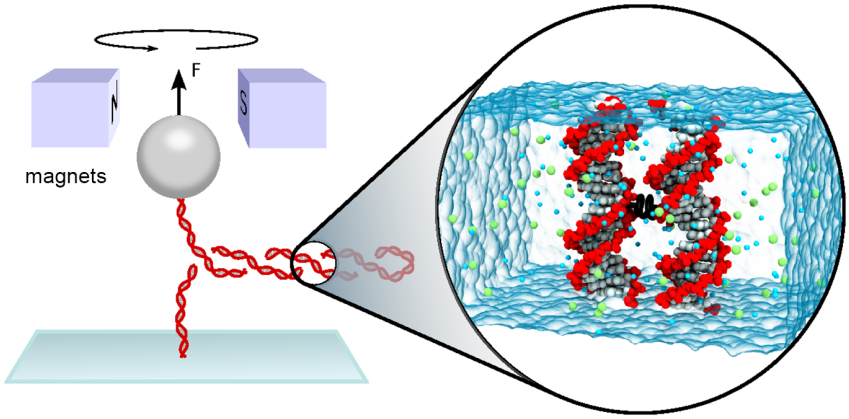


**Figure 11.** Schematic representing the experimental protocol used to produce plectonemic supercoils [47]. DNA is tethered to a stationary surface and a magnetic bead. A constant tension is applied to the DNA as it is twisted. At some critical number of twists, the DNA buckles to form a plectoneme (red curve). MD simulations quantified the interactions between the two halves of the plectoneme by averaging the force needed to restrain two effectively infinite DNA fragments about various separations. A typical simulation system is shown below the drawing of the experimental set-up. Water is shown a transparent molecular surface. Ions (green and cyan) and DNA atoms (gray, red backbone) are shown as vdW spheres. A small spring is drawn between the DNA, representing the method used to obtain the force of the interaction.

genome is packed inside a nucleus roughly 10 $\mu$m in diameter. Simple estimation of nucleotide concentration, assuming only DNA exists inside the nucleus, gives about 10 mM concentration; considering non-DNA biomolecules such as histone and transcription factor proteins, one can easily imagine that the effective concentration of DNA inside the nucleus significantly exceeds 10 mM. Therefore, DNA condensation has been the focus of intense research for several decades owing to its immense importance in biological and biomedical processes.

Polyvalent counterions (or polycations) play a critical role in stabilizing DNA condensates, providing electrostatic screening for the highly charged naked DNA and even facilitating attraction of like-charged DNA under certain conditions. A notable example is polybasic peptides in eukaryote cells that condense the eukaryotic genome. Histone tails and, especially, H1 histones are rich in basic residues such as lysine and arginine so that they can stabilize highly condensed chromatin fibers. Protamine, which is a relatively short polybasic polypeptide, is also known to condense genome in sperm cells. Moreover, non-peptide biogenic polyamines such as putrescine, spermidine and spermine are abundant in eukaryotic cells, indicating potential roles in genome condensation [259, 260].

Inspired by the biologically abundant polyamines and their potential importance in biological functions, *in vitro* studies of condensed DNA phases by biogenic polyamines have been performed by several groups. Long strands of duplex DNA (more than 400 bp) pack into dense toroids or rod-like structures within which they are assembled into a highly organized hexagonal lattice [261]. Short strands of DNA form liquid crystalline aggregates in the presence of multivalent electrolytes such as spermine and spermidine [248, 262]. Most *in vitro* studies have shown that the electrostatic forces between DNA chains in bulk solutions containing mono- or divalent electrolytes are purely repulsive [73, 263]. To induce condensation, a counterion valence of $Z \geqslant 3$ is required [262, 172]. Recent experiments involving small-angle x-ray scattering demonstrated that short duplexes of DNA containing 20 base pairs could attract each other in divalent solutions exceeding a threshold value of counterion concentration [127]. However, the same group revealed that the DNA association observed in the presence of divalent cations was caused by end-to-end stacking of short DNA oligomers [127, 128]. Subsequent experiments demonstrated that such association was inhibited by the presence of short $T_4$ loops at the ends of the DNA [128]. Thus, experimental data convincingly indicates that DNA condensation at physiological temperature can occur only in the presence of polycations with at least +3 valence. The only two DNA-condensing divalent cations are $Mn^{2+}$ and $Cd^{2+}$ [251]. Because $Mn^{2+}$ or $Cd^{2+}$ can induce DNA condensation only in a certain temperature range (30–60 °C), the $Mn^{2+}$- or -$Cd^{2+}$ + induced condensation is usually attributed to an entropic contribution of water [251]. However, the detailed mechanism is not yet known.

The microscopic mechanism of DNA condensation is a subject of a debate. The two main forces implicated in the mechanism arise from hydration and electrostatics. Experimental studies by Rau *et al* [247] suggested that the reconfiguration of water between the macromolecular surfaces as a possible mechanism for DNA compaction. The concept was developed from direct measurement of the osmotic stress as a function of the separation between parallel DNA helices. Multivalent ions bound to DNA appeared to reconfigure the water between the DNA surfaces to create long range attractive hydration forces. The authors observed that the B-form DNA molecules condensed by multivalent cations assembled into well-defined lattices of lattice spacing greater than the diameter of a double helix. Measurements of the forces between dsDNA that were pushed together from a self-assembled state revealed an exponentially increasing repulsion of a 1.3 to 1.5 Å decay length. The decay length was found to be insensitive to the ionic strength and to the valence of the cationic species. These findings could not be explained by direct bridging of neighboring DNA molecules by cationic

ligands or by the electrostatic double layer theory. Hence the authors conjectured that an alternate mechanism involving a disruption of the water structure between neighboring DNA molecules was at play in DNA condensation.

The other competing theory of DNA condensation is based on counterion fluctuations and correlation effects due to multivalent counterions. Bloomfield *et al* [264] proposed that the phenomenon of DNA lateral condensation originates from correlations in the counterion environment of homogeneously charged cylinders (representing DNA in this theory). When the counterion distribution has a quasi-two-dimensional character, i.e. the counterions are localized within a thin layer close to the cylinder surface, the problem can be reduced to the interaction between layers of adsorbed but mobile counterions neutralizing surfaces of similar charge density. The electrostatic repulsion between counterions can produce layers of alternating positive and negative charge near the surface. Two such opposing surfaces adjust complementarily to one another to minimize their total energy, resulting in a mutual attraction of the surfaces.

A similar theory proposed by Shklovski [265] suggested that counterions at the surface of the cylinders can exhibit quasi-crystalline ordering. Since dsDNA has a rigid rod-like structure, a bundle of such parallel rods can serve as a background for counterions to form a Wigner crystal. The binding energy of the bundle originates from the correlations of counterion motion. This simple picture, however ignores the discrete nature of the DNA charge. The same year Kornyshev and Leikin [254] proposed an electrostatic zipper model for DNA aggregation. The authors showed that binding of counterions to the grooves of DNA creates an axial separation of positive and negative charges. The attraction between a negatively charged strand and the positively charged grooves of the neighboring molecule creates an electrostatic zipper running along the whole length of the molecule. Divalent cations such as $Ca^{2+}$ and $Mg^{2+}$, which have high affinity to phosphates, do not induce DNA condensation within the electrostatic zipper model because of weak charge separation. However, ions that preferentially bind to the grooves and not to the DNA backbone, such as $Mn^{2+}$, enhance DNA condensation.

While the implications of these theories cannot be directly tested by experiments, all-atom MD simulations might provide an insight into the molecular mechanism of attraction. Dai *et al* [266] investigated the interaction between two parallel duplex DNA molecules in the presence of multivalent counterions, such as putrescine, spermidine, spermine and cobalt hexamine. The inter-DNA potentials obtained from the umbrella sampling technique showed that with increasing valence of the counterions and smaller ligand size, the interaction becomes more attractive. The attractive force was associated with the formation of short-lived ion-bridges, i.e. multivalent ions simultaneously bound to two DNA molecules. There was no evidence for long-range persistent two-dimensional ordering of counterions at the surface of the DNA as suggested by the counterion correlation theory. Using all-atom MD simulations, Luan and Aksimentiev demonstrated that monovalent and divalent electrolytes can include attraction between two parallel DNA fragments [256]. The simulations suggested that, although a mutual attraction between DNA double helices in NaCl electrolyte is possible, the attractive potential well is too shallow to allow for condensation. Although these two studies elucidated general aspects on cation-induced DNA–DNA interactions, they could not reproduce the experimental data quantitatively due to artificially strong cation–DNA interactions as described in section 3.2.2. For example, in both reports, the local free energy minima were at ∼23–24 Å, which is about 5 Å shorter than that in spermine-induced DNA condensate.

Using the improved parametrization of cation–phosphate interactions, one can compute the interaction free energy between two parallel DNA helices, similar to the method used in [266, 256]. The outcome of such simulations is much more consistent with the experimental data [73, 172]. Figure 12(*a*) illustrates the simulation set-up with two parallel DNA helices submerged in a solution of spermine(4+) and NaCl. Figure 12(*b*) shows the interaction free energy computed as a function of DNA–DNA distance using the conventional umbrella sampling technique. Consistent with experiment [73], inter-DNA forces are always repulsive at all ranges of DNA–DNA distance in the presence of $Na^+$ and $Mg^{2+}$ although $Mg^{2+}$ lowers the inter-DNA repulsion significantly. At sub-mM spermine concentrations, long-range inter-DNA attraction develops, indicated by the free energy minimum near 28 Å, which is quantitatively consistent with experiment [172].

### 6.4. End-to-end attraction of duplex DNA

Recently, the Clark and Bellini groups studied liquid crystal phases formed by dense solutions of short (8 bp) duplex DNA fragments [267–269]. Surprisingly, the researchers observed particular kinds of liquid crystal phases that could be formed only by elongated molecules characterized by energetically favorable axial alignment. The only explanation for the results was that the short duplex DNA fragments were aggregating end-to-end into long rod-like structures. From the concentration dependence of the phase transitions, the authors estimated a standard binding free energy of the end-to-end interaction to be between $-2.3$ and $-4.7$ kcal mol$^{-1}$ [267]. Experimental evidence supporting end-to-end attraction was independently obtained through small-angle x-ray scattering [128]. End-to-end attraction of DNA was an unexpected discovery that may have implications for a variety of systems, as outlined in figure 13(*c*). In particular, end-to-end attraction of duplex DNA could play an important role in guiding the initial steps in the repair of double-stranded DNA breakage, a serious form of DNA damage [270].

The magnitude and mechanism of end-to-end attraction was investigated using a variety of all-atom MD methods [74]. In 40 out of 40 simulations, two axially aligned duplex DNA fragments rapidly collapsed to an end-to-end bound complex. Strikingly, DNA with terminal phosphates were found to stack with a continuous conformation of the backbone that resembled B-DNA. The bound complexes were stable in subsequent simulations (without axial restraints) for hundreds of nanoseconds. Using SMD, select complexes were ruptured and an upper bound of the end-to-end free energy was estimated. Maffeo and co-workers also employed umbrella sampling simulations,

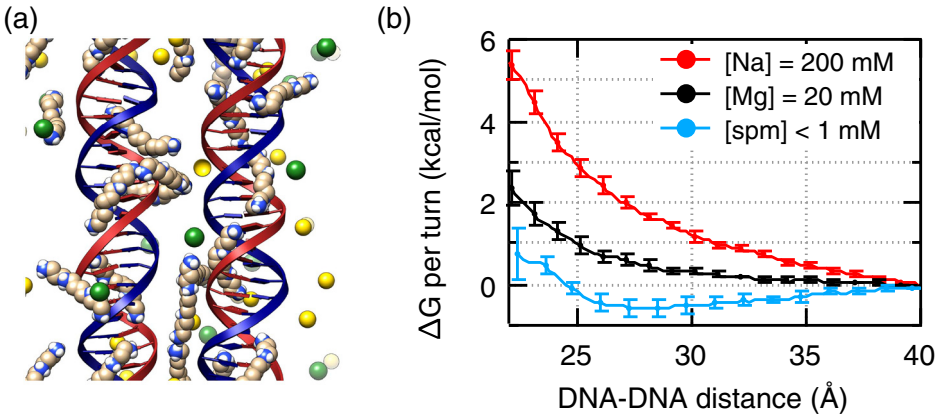


**Figure 12.** Interaction free energy of two parallel DNA helices. (*a*) Simulation set-up contains two parallel DNA helices ($G_{20}$·$C_{20}$), which are effectively infinite under periodic boundary condition. The DNA helices are shown using a cartoon representation; $Na^+$, $Cl^-$, and spermine ions are shown as vdW spheres. Water is not shown for clarity. (*b*) The free energy of two DNA helices versus the DNA–DNA distance. The simulations were performed for the following three ionic conditions: $[Na^+] = 200$mM (red); $[Na^+] = 200$mM, $[Mg^{2+}] = 20$mM (black); $[Na^+] = 200$mM, [spermine] <1mM (blue). Consistent with experimental data [73, 172], attraction is observed only in the presence of polyamine. Data are taken from [280].

the gold standard of methods for obtaining free energy, within thermodynamic cycle adapted from Woo and Roux [271] to estimate the standard binding free energy of end-to-end assembly at $-5.4 \pm 1$ kcal mol$^{-1}$. To our knowledge, this was the largest system to which the method had been applied.

Since the magnitude of the free energy was quite large compared to experimental estimates [267], the researchers verified the free-energy calculations through a brute-force simulation of a massive system containing ∼450 initially separated DNA fragments, see figure 13(*a*). The DNA fragments spontaneously associated end-to-end forming long chains during the 250 ns simulation, see figure 13(*b*). The trajectory allowed the association rate $k_{on}$ to be evaluated. During the simulation, only one end-to-end junction ruptured, allowing for a very rough estimation of the dissociation rate $k_{off}$. The logarithm of the ratio of those rates is the standard binding free energy, which was found to lie between $-4.4$ and $-7.6$ kcal mol$^{-1}$, consistent with the prior estimate. Since that study was performed, the Garcia group published a correction to the force field that reduces the strength of the stacking interaction [129], so the magnitude of the free energy obtained in the above study could have been overestimated.

More recently, the Sciortino group employed CG simulations to validate a theoretical model that, given a stacking free energy, provides a link between the concentration of short DNA duplexes and the degree of the end-to-end polymerization, which can in turn be related to the liquid crystal phase of DNA solution [272]. The theory was then used to estimate the range of end-to-end stacking free energies compatible with the experimentally observed liquid crystal phase transitions, finding the end-to-end association free energy in the range of $-0.4$ and $-2.4$ kcal mol$^{-1}$.

## 7. DNA in electric field

As a highly charged molecule, DNA moves in an external electric field. This phenomenon, electrophoresis, has many important applications in biotechnology and forensics. For example, gel or capillary electrophoresis is a key step of the Sanger's method of sequencing DNA [273]. The electrophoretic motion of DNA through nanopores and nanochannels is essential to many emerging methods of DNA characterization, in particular the so-called nanopore sequencing [274–276]. From a purely scientific point of view, the motion of DNA in electric field is a particularly interesting subject as it permits validation and further improvement of various theoretical models of polyelectrolyte solutions.

### 7.1. Free-solution electrophoresis

The movement of an object, such as DNA, under the influence of an electric field is described by its electrophoretic mobility $\mu$:

$$\mu = \frac{v}{E}, \tag{12}$$

where $E$ is the electric field and $v$ is the measured velocity. In free-solution electrophoresis, however, the behavior of $\mu$ for DNA can be puzzling. In particular, $\mu$ of DNA, which might reasonably be expected to be proportional to the number of bases $N$, is actually independent of $N$ for large $N$ [277, 278]. Moreover, under certain conditions $\mu(N)$ displays a *maximum* before lowering slightly to its large-$N$ value [277].

How can this counterintuitive result be understood? Authors sometimes turn to the Einstein relation between $\mu$ and the diffusion constant $D$:

$$\mu = \frac{QD}{k_B T}, \tag{13}$$

where $Q$ is the electric charge of DNA. However, for a long polymer the diffusion constant scales as $D(N) \sim 1/N^{\nu}$, where $\nu$ is 1/2 or 3/5, depending upon whether excluded volume effects are important. For shorter DNA ($N < 250$), $\nu$ was found to be approximately 0.68 [279]. Since $Q \sim N$ and $\mu$ is independent of $N$, equation (13) cannot be correct [279].

Combined with measurements of the diffusivity of DNA [278, 279], one may be tempted to salvage equation (13)

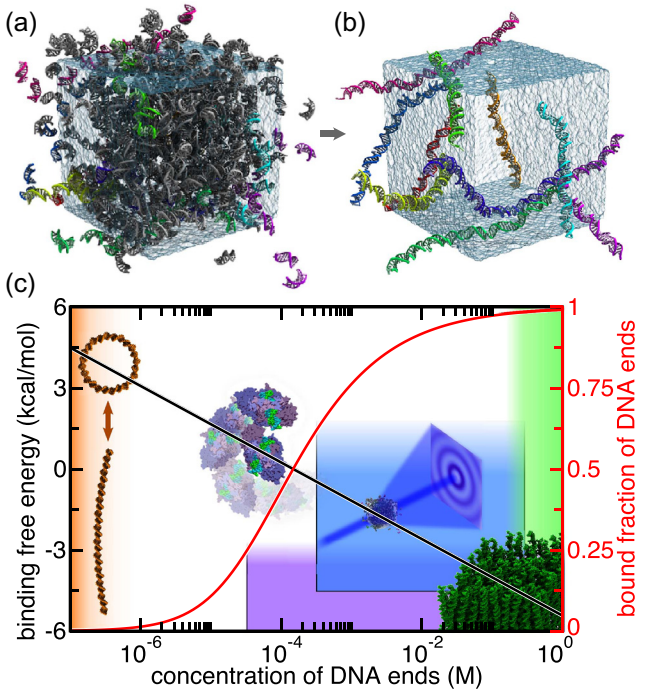


**Figure 13.** End-to-end association of duplex DNA fragments. (*a*) Simulation system containing a solution of 458 DNA fragments near the density of the experimentally determined isotropic–nematic phase transition [267]. Most DNA fragments are shown in gray. Those fragments that formed the ten largest end-to-end chains at the end of the simulation are shown in color. The simulation used periodic boundary conditions, lasted 260 ns and included about 1 500 000 atoms. Water in the simulation unit cell is shown as a semitransparent molecular surface. (*b*) The ten largest end-to-end chains at the end of the simulation. The kinetic association and dissociation rates enabled estimation of the standard binding free energy of the end-to-end interaction. (*c*) End-to-end attraction in different DNA systems. The effective binding free energy (black) and the fraction of bound DNA ends (red) are plotted against the reference concentration of DNA ends. Images illustrate four DNA systems in which the end-to-end attraction may or may not play a role. From top left to bottom right: blunt-ended DNA circles (orange); repair of DNA during non-homologous end joining [270] (purple); small-angle x-ray scattering experiments [127] (blue); DNA aggregation into liquid crystal phases [267] (green). Figures adapted with permission from [74].

by assigning DNA an *effective* charge [281]. Based on the above scaling argument, such a construction would scale as $Q_{\rm eff}(N) \sim N^{1-\nu}$. However, this obscures the reasons for the failure of equation (13). As pointed out by Hoagland *et al* [277], the Einstein relation neglects the very important effect of counterions surrounding the DNA. An applied electric field acts on these counterions as well; since they are by definition of the opposite charge of DNA, an electro-osmotic flow is set up which acts against the motion of DNA.

The effects of the electro-osmotic flow are significant. It leads not only to a lower $Q_{\rm eff}$ of DNA, but qualitatively changes the behavior of DNA motion in an electric field. The counterion motion effectively screens hydrodynamic interactions between monomers [282]. Thus, the fact that a long strand of DNA is a single, covalently bonded polymer is irrelevant to its motion, and the mobility $\mu$ is independent of the polymer length. The DNA coil is transparent to the motion of counterions and is hence described as 'free-draining' [283]. For diffusive motion, on the other hand, hydrodynamic forces between monomers are not screened, leading to the failure of equation (13) [279].

The independence of DNA mobility on $N$ led to the so-called *local force picture*. In the local force picture, DNA is thought of as chain of segments, each with an effective charge $q_{\rm k}$ and friction coefficient $\xi_{\rm k}$. The segments do not exert force on each other, and thus the velocity of the DNA is given by $v = N_{\rm k} q_{\rm k} E / N_{\rm k} \xi_{\rm k} = q_{\rm k} E / \xi_{\rm k}$, independent of the number of segments $N_{\rm k}$. However, this picture can be misleading: as pointed out by Vivoy [284], the local force picture implicitly assumes that the background fluid is at rest.

### 7.2. Sieved electrophoresis

Since in free-solution electrophoresis the DNA mobility $\mu$ is independent of the DNA length, free-solution electrophoresis cannot generally be used for mass fractionation, and therefore is not useful for DNA sequencing applications [285]. In the presence of a *sieving matrix*, however, DNA mobility once again depends on the DNA length. The sieving matrix may be a gel such as polyacrylamide or agarose, or a polymer such as polydimethylacyrylamide (for an exhaustive review of sieving media, see [286]).

A sieving matrix can be thought of as a network of pores of typical size $b$. DNA with a radius of gyration $R_{\rm g} \gg b$ must radically deform to electrophorese through such a network, forming (at low fields) a random chain. The motion of such DNA is called 'reptation', akin to the motion of an entangled polymer [287], and is described mathematically by various reptation models, such as biased reptation (BRM) and biased reptation with fluctuations (BRF) models [284, 288]. In these models, DNA is confined to a tube (made of connected pores) and undergoes biased diffusional motion in that tube.

Why does sieving work? Because it breaks the scaling between friction and driving forces. In its 'tortured' (snake-like) conformation, the effective force on DNA $F_{\rm eff} \sim R_x$, where $R_x$ is the vector between the two DNA ends projected in the field direction. In weak fields, DNA behaves as a Gaussian chain, and therefore $R_x \sim N^{1/2}$ [284, 288]. Meanwhile, the friction still scales as $\xi \sim N$, and so the velocity in the tube $v_{\rm curv} = F_{\rm eff}/\xi \sim N^{-1/2}$. Projecting this velocity along the field direction yields the familiar result $v \sim N^{-1}$.

At low field, the DNA indeed forms a Gaussian chain, to a good approximation. As the field increases, however, the DNA tends to orient along the field direction. In a strong field, the DNA is fully oriented, and mass fractionation is no longer possible because the effective force on DNA no longer scales as $N^{1/2}$, but as $N$, thus making the mobility $\mu$ independent of $N$. The same breakdown occurs as $N$ increases, limiting the effective read length of sieving-based DNA sequencing techniques to $\sim$ 1000 bp.

### 7.3. Nanopore electrophoresis

Over the last decade, DNA electrophoresis through a nanometer-sized pore (nanopore) has been extensively studied [289–291]. The device used in experiment usually contains

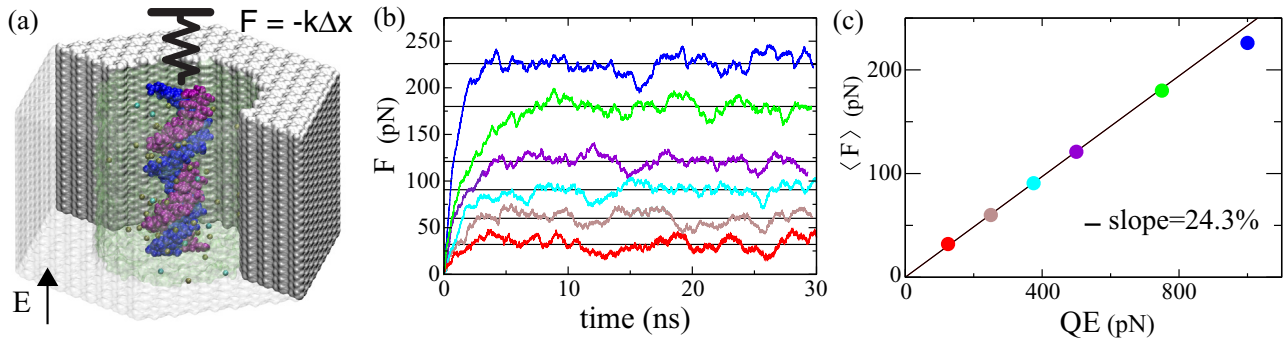


**Figure 14.** MD simulation of the effective force on DNA in a nanopore. (*a*) Simulation set-up. Two strands of DNA are colored in purple and blue; $K^+$ (tan) and $Cl^-$ (cyan) ions are shown as spheres; water (green) is shown as a semitransparent surface. A mechanical tether force is applied to the DNA via a harmonic potential (a virtual spring). One end of the spring is fixed in space whereas the other end is attached to the center of mass of the DNA fragment. A uniform external electric field $E$ is applied to the whole system. The DNA molecule is made effectively infinite by connecting the backbone to its image across the periodic boundary. (*b*) The restraining force versus the simulation time for several values of the applied electric field. The color of the lines corresponds to the color of the symbols in panel (*c*). (*c*) The average effective force acting on DNA versus the nominal force of the electrostatic field $QE$, where $Q$ is the nominal electrical charge of DNA. Figures were adapted from [310] with permission. Copyright 2008 by the American Physical Society.

two fluid chambers separated by a thin (tens of nanometers) membrane containing a nanopore. The membrane and nanopore may be biological [289, 292], such as a protein pore in a lipid membrane, or synthetic [293–295], such as a solid-state nanopore drilled in a thin silicon-nitride or silicon-oxide film. The confined geometry of a nanopore facilitates detection and identification of individual biological molecules [296]. Nanopores have been suggested as promising systems for realizing low-cost and high-throughput sequencing of DNA [274, 275, 297].

In nanopore electrophoresis, DNA is transported from one side of the membrane to the other by a transmembrane electric potential. The translocation events are detected and characterized by measuring the blockades of transmembrane ionic current [298, 299]. If the size of the pore constriction is comparable to the diameter of the DNA molecule, the ionic current can be sensitive to the DNA's nucleotide sequence [300, 301]. Other mechanisms for DNA sequence readout using nanopores have been suggested, including electron tunneling current and local electrostatic potential measurements [302–305]. In general, DNA translocation through nanopores is found to occur more slowly than one might naïvely expect based on its charge and the magnitude of the electric field [306]. Thus, quantitative characterization of the force experienced by DNA in a nanopore is critical to understanding the microscopic mechanics of the DNA transport.

*7.3.1. Effective driving force on DNA in a nanopore.* The effective force on DNA in a nanopore was measured directly by trapping one end of the DNA molecule with optical tweezers while the other end of the molecule was subject to an electric field in a nanopore [307]. The measurements revealed scaling of the effective force $F$ with the electrostatic field $E$, i.e. $F = q_{\text{eff}}E$, where the scaling factor $q_{\text{eff}}$, also known as the DNA's effective charge, was found to be about 25% of the DNA's bare charge $Q$. However, the interpretation of these measurements was ambiguous. On the one hand, the measurements are in perfect agreement with the Manning condensation theory [37] that predicts a 76% reduction of the DNA charge arising from the electrostatic field of counterions that condense near the DNA surface. Such interpretation neglects solvent as a possible source of the effective screening and assumes that the fraction $q^*$ of the DNA charge that is not screened by the electrostatic field of the condensed counterions is the effective charge that determines the effective force. On the other hand, theoretical and experimental studies of DNA electrophoresis suggested that hydrodynamic drag is an important factor influencing DNA transport [282, 279]. The importance of hydrodynamic interactions between DNA and the solvent *inside* a solid-state nanopore was pointed out by Ghosal [308, 309].

To determine the microscopic origin of the effective driving force acting on DNA in a nanopore, Luan and Aksimentiev performed all-atom MD simulations of a mechanically tethered DNA molecule subject to an electric field in a nanopore [310]. Figure 14(*a*) illustrates a typical set-up of the MD simulation: a DNA fragment subject to an electric field is confined in a cylindrical channel, cut from a $Si_3N_4$ crystal. Mimicking the action of optical tweezers, a harmonic potential (a virtual spring of a spring constant $k$) is attached to the center of mass of the DNA fragment. Subject to the applied electric field, the DNA initially drifts opposite the field direction, figure 14(*b*), until the force of the restraining potential balances the effective driving force of the electric field. Repeating the simulations at several values of the applied potential reveals the dependence of the effective electrophoretic force on the electric field strength, figure 14(*c*). For the particular nanopore geometry considered in the study, the simulations revealed a linear dependence of the effective forces $F$ on the applied electric fields $E$. The effective DNA charge determined as the ratio of $F$ to $E$ was found to be ~25% of the nominal DNA charge, in apparent agreement with the Manning condensation theory. As shown below, such an agreement was coincidental.

Recent works [309–311] have clearly shown the prominent role of the electro-osmotic flow in determining the effective force on DNA in a nanopore. Experimentally,

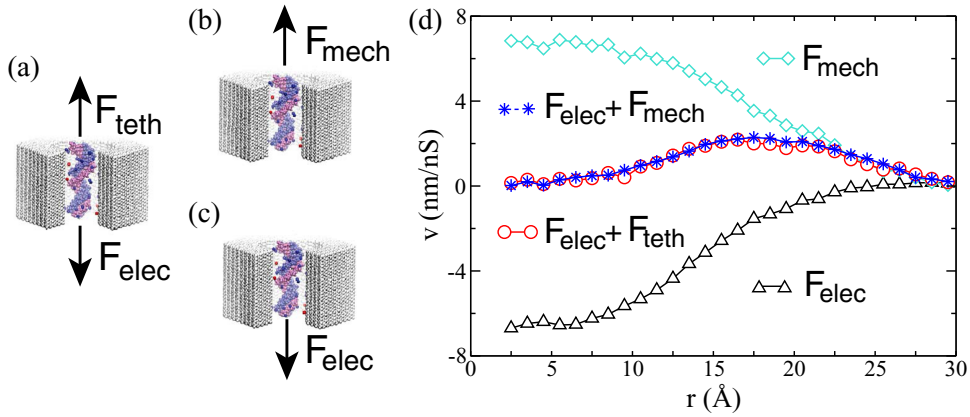


**Figure 15.** Electro-osmotic screening of the DNA charge. (*a*) A stationary DNA fragment is subject to a restraining force $F_{\text{teth}}$ and an electrophoretic force $F_{\text{elec}}$. (*b*) The DNA fragment is displaced through the nanopore with a constant velocity by a mechanical pulling force $F_{\text{mech}}$. (*c*) The DNA fragment is displaced through the nanopore by an electrophoretic force $F_{\text{elec}}$. (*d*) Water velocity as a function of radial position, calculated from all-atom MD simulations corresponding to the set-ups shown in panels (*a*)–(*c*). The DNA surface is located at approximately 11 Å; the surface of the nanochannel is located at ∼30 Å. A superposition (blue stars) of the flow profiles observed in mechanical pulling (cyan diamonds; panel (*b*)) and electrophoresis (black triangles; panel (*c*)) simulations reproduces the flow profile observed in the simulations of the effective electrophoretic force (red circles; panel (*a*)). Positive values of the water velocity corresponds to the upward direction in *a*)–(*c*). The data were taken from [310].

measurements of the force on DNA in nanopores of different size showed that the effective charge is actually a decreasing function of pore radius [311]. Thus, the effective charge of DNA in a pore should also decrease as the pore radius increases. The following relationship was predicted theoretically by Ghosal [309] by solving the coupled Poisson and Stokes equations:

$$Q_{\text{eff}} = F_{\text{mech}}/E = 2\pi\epsilon L\frac{\Phi(a) - \Phi(R)}{\ln(R/a)}, \qquad (14)$$

where $L$ is the membrane thickness, $\epsilon$ is the dielectric constant of the electrolyte, $a$ and $R$ are respectively the radii of the DNA and nanopore, and $\Phi(a)$ and $\Phi(R)$ are $\zeta$-potentials at DNA and pore surfaces, respectively. Finding potentials $\Phi(a)$ and $\Phi(R)$ numerically by solving the PB equation, the experimental results [311] could be quantitatively reproduced using equation (14) with a constant offset.

Complementing the theoretical and experimental studies, MD simulations have elucidated the contribution of the electro-osmotic flow to the effective force experienced by DNA in a nanopore [310]. Figure 15 illustrates the local velocity of water in a 3 nm radius nanopore observed in three MD simulations. The first simulation, figure 15(*a*), directly measures the effective electrophoretic force $F_{\text{elec}}$ by balancing it with a tether force $F_{\text{teth}}$, just like in figure 14(*a*). Although the DNA fragment remains stationary, ionic current flows through the nanopore, dragging water molecules along. The resulting water flow has a parabolic profile, reaching zero at the surface of the DNA and the nanopore, figure 15(*d*) (circles). In this particular simulation, the tether (spring) force is 118 pN and the applied electric field is 8 mV Å$^{-1}$. In the absence of electric field, figure 15(*b*), DNA is pulled by a mechanical force of the tether with a velocity of 7 nm ns$^{-1}$; the pulling force (∼118 pN) is balanced by the hydrodynamic friction. The flow profile (diamonds in figure 15(*d*)) has a maximum of ∼7 nm ns$^{-1}$ at the DNA surface, implying no-slip flow. Subject to the electric field of 8 mV Å$^{-1}$ and in the absence of the restraining force of the tether, figure 15(*c*), the DNA moves in the direction opposite to the direction of the field with an average velocity of −7 nm ns$^{-1}$. The solvent flow (triangles in figure 15(*d*)) has the same velocity near the DNA surface as the DNA fragment.

The results of the above simulations indicate that a stationary DNA subject to simultaneous action of mechanical and electric forces can be considered as a superposition of the electrophoretic motion ($v = \mu E$) and mechanical pulling ($v = F/\xi$). Indeed, the superposition of the flow profiles produced by the electrically and mechanically driven motions of DNA (stars in figure 15(*d*)) reproduces the flow profile observed when DNA is held stationary by the simultaneous action of the electrophoretic and tether forces. From the relation of superposition

$$F = \xi\mu E, \qquad (15)$$

where $\xi$ is the friction coefficient and $\mu$ is the electrophoretic mobility [310]. This relation was directly verified by MD simulations of the DNA's electrophoretic and mechanical motion through nanochannels of different radii and different surface roughnesses [310]. Note that the above expression was previously derived [282] for DNA electrophoresis in a gel. Importantly, it allows for a more satisfying and illuminating expression for the effective charge of DNA:

$$Q_{\text{eff}} = \xi\mu. \qquad (16)$$

The two seemingly different expressions for the effective charge of DNA, equations (14) and (16), are consistent with each other, given that the friction coefficient $\xi$ and the electrophoretic mobility $\mu$, derived from electrohydrodynamics equations [309, 311], are

$$\xi = \frac{2\pi\eta L}{\ln(R/a)}, \qquad (17)$$

$$\mu = \epsilon\frac{\Phi(a) - \Phi(R)}{\eta}, \qquad (18)$$

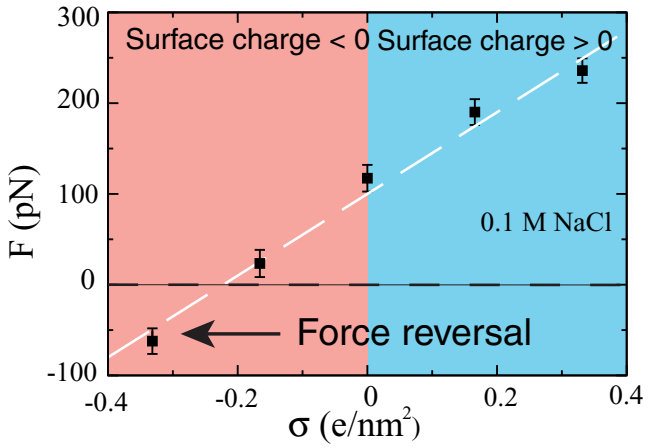


**Figure 16.** The effective force acting on a DNA fragment in a solid-state nanochannel of surface charge density $\sigma$. The force was measured using a setup similar to that shown in figure 14(*a*). The data were taken from [312].

where $\eta$ is the viscosity of the electrolyte. With explicit expressions for $\xi$ and $\mu$, the equivalence of equations (14) and (16) becomes obvious.

*7.3.2. DNA translocation through charged nanopore.* The effect of the hydrodynamic flow on the effective charge of DNA is even more dramatic in a charged nanopore because of the additional electro-osmotic flow that develops near the charged nanopore surface. Luan and Aksimentiev have showed that the surface charge density can regulate the electrophoretic mobility of DNA [312]: lowering the surface charge reduces the effective electrophoretic force and hence the electrophoretic mobility. When the surface charge density is positive, the induced electro-osmotic flow near the pore surface moves in the same direction as DNA, increasing the effective force on and the translocation velocity of DNA, in comparison to an uncharged pore, see figure 16. When the surface charge density is negative, the induced electro-osmotic flow near the pore surface acts against the electrophoretic force, reducing the DNA translocation speed. Further decreasing the surface charge density can even reverse the direction of the electrophoretic force, figure 16. Consequently, the effective electrophoretic charge of DNA is also a function of the nanopore surface charge density [312]. Because the nanopore surface serves as a boundary of the hydrodynamic flow, other surface properties, such as roughness and hydrophobicity, can affect the flow profile and thereby the effective charge of DNA in a nanopore.

*7.3.3. Slowing down DNA transport in LiCl.* DNA translocation experiments in the Cees Dekker group at TU Delft revealed an unexpected phenomenon: the size, not the charge, of cations was found to affect the DNA translocation velocity through a nanopore [313]. Specifically, the transport of DNA was found to be slower when DNA was surrounded my smaller monovalent cations (at the same cation concentration). Such a phenomenon could not be explained by continuum modeling as the size of cations is not taken into account by the theory.

MD simulations illustrated in figure 17 determined the microscopic origin of this phenomenon [313]. At the same bulk ion concentration, the instantaneous number of $Li^+$, $Na^+$, or $K^+$ ions bound to the surface of DNA was found to be the same. Consequently the total charge of the surrounding solvent is the same regardless of the ion type and hence the force driving the electro-osmotic flow. However, the duration of the transient bound state between cations and DNA was found to sensitively depend on the type of cations: the bound state formed by $Li^+$ lasted, on average, longer than the bound state formed by $Na^+$, and the latter were found to be longer than the bound state formed by $K^+$. Using a simplified model, figure 17(*b*), it was demonstrated that such a change in bond strength could indeed affect the force transmitted from ion to DNA, producing better effective neutralization of the DNA charge and hence reducing the velocity of DNA transport through a nanopore.

*7.3.4. Electrophoretic charge inversion.* Motion of a charged object in electric field in vacuum can serve as a measure of the object's electric charge. In the case of DNA in electrolyte solution, interpretation of such measurements is not straightforward. At the very least, measurements of the direction of the electrophoretic motion can be used to detect the inversion of the DNA's effective charge. Indeed, by measuring the electrophoretic mobility of DNA, Besteman *et al* concluded that charge inversion could occur when DNA was surrounded by spermine ($[C_{10}N_4H_{30}]^{4+}$) and cobalt sepulchrate ($[CoC_{12}H_{30}N_8]^{3+}$) counterions, and not by two other types of trivalent cations [314]. Strictly speaking, the above experiments have only shown the inversion of the *effective electrophoretic* charge of DNA.

Using the MD method, Luan and Aksimentiev [177] demonstrated inversion of the electrophoretic mobility of ds-DNA in the presence of spermine(4+) and spermidine(3+) counterions, figure 18. These simulations were already described in section 4.2.4 when inversion of the DNA's electrical charge was discussed. When the amount of spermidine molecules was just enough to neutralize the DNA charge, the DNA was observed to move in the direction opposite to the direction of the electric field, figure 18. By increasing the number of spermidine cations the direction of the electrophoretic motion could be reversed. The study has also unequivocally shown that inversion of the DNA's electric charge (measured using the Gauss law, section 4.2.4) is not equivalent to inversion of the DNA's electrophoretic charge (indicated by the direction of the DNA's motion) [177]. The electric charge inversion is related to the ionic screening within the plane perpendicular to the DNA duplex and thus is insensitive to the action of an external electric field directed parallel to the DNA helical axis. Inversion of the DNA's electrophoretic charge is conditioned by both direct binding of counterions to DNA and the electro-osmotic effect, which in turn depends on the boundary condition for the electro-osmotic flow.

## 8. Concluding remarks

The steady advancement of scientific computing leaves no doubt that atomistic MD simulations will play a major role

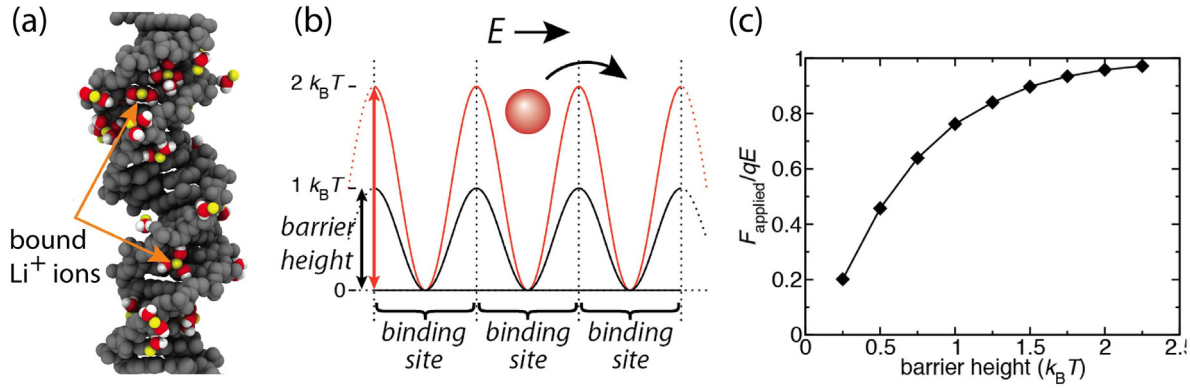


**Figure 17.** The type of monovalent ions affect the translocation velocity of DNA. (*a*) A snapshot from an MD simulation illustrating binding of $Li^+$ ions to dsDNA. DNA is shown in gray, lithium is shown in yellow, and water is shown in red and white. Only those water molecules involved in the lithium–DNA bonds are shown. (*b*) A toy model describing ion binding to DNA. Red and black lines schematically represent 1D potentials describing affinity of two different cations to DNA. The depth of the free-energy minima are the same and so is the instantaneous number of ions bound to DNA. The barrier for hopping between the adjacent sites along DNA varies with the ion size and so is the strength of the bond. (*c*) The barrier height (the strength of the cation–DNA bond) affects the effective charge of DNA. The effect of water flow is neglected in this model. Adapted with permission from [313]. Copyright 2012 American Chemical Society.

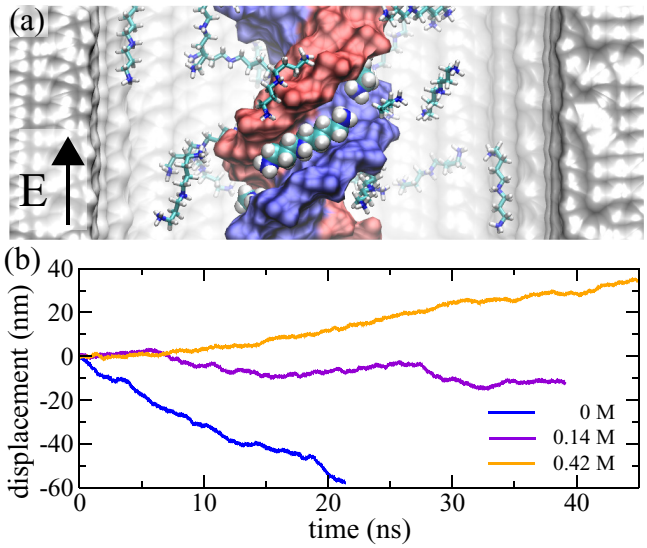


**Figure 18.** Reversal of the electrophoretic motion of DNA in a multivalent electrolyte. (*a*) Setup of MD simulations. The two strands of a DNA duplex are shown using a molecular surface representation, spermidine molecules directly bound to the groove of DNA are shown using vdW spheres, spermidine molecules dissolved in solution are shown using a ball-and-stick model. Subject to electric field, the DNA molecule moves through the nanochannel (gray molecular surface). (*b*) Center-of-mass displacement of the DNA duplex versus simulation time at three different concentrations of spermidine counterions. The data were taken from [177].

in the future development of the nucleic acids research field. Although several challenges lie ahead with regard to improvement of the molecular force field and sampling of the conformational space, it has already become clear that MD simulations can not only compliment experiment but make testable predictions and lead development in the field. Exciting opportunities await in combining single-molecule experiment with atomistic simulations to uncover the physical mechanisms that biological cells use to store, transcribe, duplicate and repair DNA. Atomistic simulations will allow exploration of DNA nanostructures for a variety of applications, from advanced materials to drug delivery. For the physics community, atomistic simulations of DNA are expected to firmly link theory to experiment, leaving no room for interpretation.

## Acknowledgments

This work was supported by the grants from the National Science Foundation (DMR-0955959 and PHY-1430124) and the National Institutes of Health (R01-HG005115 and R01-HG007406). The authors gladly acknowledge supercomputer time provided through XSEDE Allocation Grant MCA05S028 and the Blue Waters Petascale Supercomputer facility of the University of Illinois at Urbana-Champaign.

## References

[1] Wiggins P A, van der Heijden T, Moreno-Herrero F, Spakowitz A, Phillips R, Widom J, Dekker C and Nelson P C 2006 High flexibility of DNA on short length scales probed by atomic force microscopy *Nature Nanotechnol.* **1** 137–41

[2] Kaplan N, Moore I K, Fondufe-Mittendorf Y, Gossett A J, Tillo D, Field Y, LeProust E M, Hughes T R, Lieb J D, Widom J and Segal E 2009 The DNA-encoded nucleosome organization of a eukaryotic genome *Nature* **458** 362–6

[3] Vafabakhsh R and Ha T 2012 Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization *Science* **337** 1097–101

[4] Ito S, Shen L, Dai Q, Wu S C, Collins L B, Swenberg J A, He C and Zhang Y 2011 Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine *Science* **333** 1300–3

[5] Hershey A D and Chase M 1952 Independent functions of viral protein and nucleic acid in growth of bacteriophage *J. Gen. Physiol.* **36** 39–56

[6] Watson J D and Crick F H C 1953 A structure for deoxyribose nucleic acids *Nature* **171** 737–8

[7] Kuhn W 1936 Beziehungen zwischen Molekülgröße, statistischer Molekülgestalt und elastischen Eigenschaften hochpolymerer Stoffe *Colloid Polym. Sci.* **76** 258–71

[8] Kramers H A 1946 The behavior of macromolecules in inhomogeneous flow *J. Chem. Phys.* **14** 415

[9] Kratky O and Porod G 1949 Röntgenuntersuchung gelöster Fadenmoleküle *Recl. Trav. Chim. Pay.* B **68** 1106–22

[10] Peterlin V A 1953 Lichtzerstreuung an ziemlich gestreckten Fadenmolekülen *Makromol. Chem.* **9** 244–68

[11] Doty P 1957 The physical chemistry of deoxyribonucleic acids *J. Cell. Comparat. Physiol.* **49** 27–57

[12] Hill T L 1959 Generalization of the one-dimensional Ising model applicable to helix transitions in nucleic acids and proteins *J. Chem. Phys.* **30** 383–7

[13] Gibbs J H and DiMarzio E A 1959 Statistical mechanics of helix–coil transitions in biological macromolecules *J. Chem. Phys.* **30** 271

[14] Zimm B H 1960 Theory of 'melting' of the helical form in double chains of the DNA type *J. Chem. Phys.* **33** 1349–56

[15] de Gennes P G 1969 Some conformation problems for long macromolecules *Rep. Prog. Phys.* **32** 187–205

[16] Crick F H C and Klug A 1975 Kinky helix *Nature* **255** 530–3

[17] Sobell H M, Tsai C C, Jain S C and Gilbert S G 1977 Visualization of drug–nucleic acid interactions at atomic resolution: III. Unifying structural concepts in understanding drug–DNA interactions and their broader implications in understanding protein–DNA interactions *J. Mol. Biol.* **114** 333–65

[18] Sussman J L and Trifonov E N 1978 Possibility of nonkinked packing of DNA in chromatin *Proc. Natl Acad. Sci. USA* **75** 103–7

[19] Levitt M 1978 How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations *Proc. Natl Acad. Sci. USA* **75** 640–4

[20] Clementi E and Corongiu G 1979 Interaction of water with DNA single and double helix in the B conformation *Int. J. Quantum Chem.* **16** 897–915

[21] Clementi E and Corongiu G 1980 A theoretical study on the water structure for nucleic acids bases and base pairs in solution at $T = 300$ K *J. Chem. Phys.* **72** 3979–92

[22] Levitt M 1983 Computer simulation of DNA double-helix dynamics *Symp. on Quantitative Biology (Cold Spring Harbor, NY)* vol 47, pp 251–62

[23] Tidor B, Irikura K K, Brooks B R and Karplus M 1983 Dynamics of DNA oligomers *J. Biomol. Struct. Dyn.* **1** 231–52

[24] Seibel G L, Singh U C and Kollman P A 1985 A molecular dynamics simulation of double-helical B-DNA including counterions and water *Proc. Natl Acad. Sci. USA* **82** 6537–40

[25] Beveridge D L, Swaminathan S, Ravishanker G, Withka J M, Srinivasan J, Prevost C, Louise-May S, Langley D R, DiCapua F M and Bolton P H 1993 Molecular dynamics simulation on hydration, structure and motions of DNA oligomers *Water and Biological Macromolecules* ed E Westhof (Boca Raton, FL: CRC Press) chapter 6, p 165

[26] Jayaram B and Beveridge D 1996 Modelling DNA in aqueous solutions—theoretical and computer simulation studies on the ion atmosphere of DNA *Annu. Rev. Biophys. Biomol. Struct.* **25** 367–94

[27] Beveridge D L and Ravishanker G 1994 Molecular dynamics studies of DNA *Curr. Opin. Struct. Biol.* **4** 246–55

[28] Anderson C F and Record M T 1995 Salt–nucleic acid interactions *Annu. Rev. Phys. Chem.* **46** 657–700

[29] Cheatham T E III, Miller J L, Fox T, Darden T A and Kollman P A 1995 Molecular dynamics simulations on solvated biomolecular systems: the particle mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins *J. Am. Chem. Soc.* **117** 4193–4

[30] Cheatham T E and Kollman P A 2000 Molecular dynamics simulation of nucleic acids *Annu. Rev. Phys. Chem.* **51** 435–71

[31] Beveridge D L and McConnell K J 2000 Nucleic acids: theory and computer simulation, Y2K *Curr. Opin. Struct. Biol.* **10** 182–96

[32] Cheatham T E and Kollman P A 1996 Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution *J. Mol. Biol.* **259** 434–44

[33] Young M A, Ravishanker G and Beveridge D L 1997 A 5-nanosecond molecular dynamics trajectory for B-DNA: analysis of structure, motion and solvation *Biophys. J.* **73** 2313–36

[34] Cheatham T E and Young M A 2000 Molecular dynamics simulation of nucleic acids: successes, limitations, and promise *Biopolymers* **56** 232–56

[35] Saenger W, Hunter W N and Kennard O 1986 DNA conformation is determined by economics in the hydration of phosphate groups *Nature* **324** 385–8

[36] Feig M and Pettitt B M 1998 A molecular simulation picture of DNA hydration around A- and B-DNA *Biopolymers* **48** 199–209

[37] Manning G S 1969 Limiting laws and counterion condensation in polyelectrolyte solutions: I. Colligative properties *J. Chem. Phys.* **51** 924

[38] Young M A, Jayaram B and Beveridge D L 1997 Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: fractional occupancy of electronegative pockets *J. Am. Chem. Soc.* **119** 59–69

[39] Lyubartsev A P and Nordenskiöld L 1995 Monte Carlo simulation study of ion distribution and osmotic pressure in hexagonally oriented DNA *J. Phys. Chem.* **99** 10373–82

[40] Feig M and Pettitt B M 1999 Sodium and chlorine ions as part of the DNA solvation shell *Biophys. J.* **77** 1769–81

[41] Still W C, Tempczyk A, Hawley R C and Hendrickson T 1990 Semianalytical treatment of solvation for molecular mechanics and dynamics *J. Am. Chem. Soc.* **112** 6127–9

[42] Jayaram B, Sprous D and Beveridge D L 1998 Solvation free energy of biomacromolecules: parameters for a modified generalized Born model consistent with the AMBER force field *J. Phys. Chem.* B **102** 9571–6

[43] Dominy B N and Brooks C L III 1999 Development of a generalized Born model parametrization for proteins and nucleic acids *J. Phys. Chem.* B **103** 3765–73

[44] MacKerell A D and Lee G U 1999 Structure, force, and energy of a double-stranded DNA oligonucleotide under tensile loads *Eur. Biophys. J.* **28** 415–26

[45] Smith S B, Cui Y and Bustamante C 1996 Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules *Science* **271** 795–9

[46] Šponer J, Šponer J E, Mládek A, Jurečka P, Banáš P and Otyepka M 2013 Nature and magnitude of aromatic base stacking in DNA and RNA: quantum chemistry, molecular mechanics, and experiment *Biopolymers* **99** 978–88

[47] Maffeo C, Schöpflin R, Brutzer H, Stehr R, Aksimentiev A, Wedemann G and Seidel R 2010 DNA–DNA interactions in tight supercoils are described by a small effective charge density *Phys. Rev. Lett.* **105** 158101

[48] Biswas M, Voltz K, Smith J C and Langowski J 2011 Role of histone tails in structural stability of the nucleosome *PLoS Comput. Biol.* **7** e1002279

[49] Yoo J and Aksimentiev A 2013 *In situ* structure and dynamics of DNA origami determined through molecular dynamics simulations *Proc. Natl Acad. Sci. USA* **110** 20099–104

[50] Doye J P K *et al* 2013 Coarse-graining DNA for simulations of DNA nanotechnology *Phys. Chem. Chem. Phys.* **15** 20395–414

[51] Rollins G C, Petrov A S and Harvey S C 2008 The role of DNA twist in the packaging of viral genomes *Biophys. J.* **94** L38–40

[52] Tark-Dame M, van Driel R and Heermann D W 2011 Chromatin folding–from biology to polymer models and back *J. Cell Sci.* **124** 839–45

[53] Norberg J and Nilsson L 2002 Molecular dynamics applied to nucleic acids *Acc. Chem. Res.* **35** 465–72

[54] Cheatham T E 2004 Simulation and modeling of nucleic acid structure, dynamics and interactions *Curr. Opin. Struct. Biol.* **14** 360–7

[55] Pérez A, Luque F J and Orozco M 2012 Frontiers in molecular dynamics simulations of DNA *Acc. Chem. Res.* **45** 196–205

[56] Šponer J, Mládek A, Šponer J E, Svozil D, Zgarbová M, Banáš P, Jurečka P and Otyepka M 2012 The DNA and RNA sugar–phosphate backbone emerges as the key player. An overview of quantum-chemical, structural biology and simulation studies *Phys. Chem. Chem. Phys.* **14** 15257–77

[57] Egli M 2002 DNA–cation interactions: quo vadis? *Chem. Biol.* **9** 277–86

[58] Mocci F and Laaksonen F 2012 Insight into nucleic acid counterion interactions from inside molecular dynamics simulations is 'worth its salt' *Soft Matter* **8** 9268

[59] Yoo J and Aksimentiev A 2012 Improved parametrization of $Li^+$, $Na^+$, $K^+$, and $Mg^{2+}$ ions for all-atom molecular dynamics simulations of nucleic acid systems *J. Phys. Chem. Lett.* **3** 45–50

[60] Beveridge D L, Dixit S B, Barreiro G and Thayer K M 2004 Molecular dynamics simulations of DNA curvature and flexibility: helix phasing and premelting *Biopolymers* **73** 380–403

[61] Prévost C, Takahashi M and Lavery R 2009 Deforming DNA: from physics to biology *ChemPhysChem* **10** 1399–404

[62] Peters J P and Maher L J 2010 DNA curvature and flexibility *in vitro* and *in vivo Quart. Rev. Biophys.* **43** 23–63

[63] Vologodskii A and Frank-Kamenetskii M D 2013 Strong bending of the DNA double helix *Nucleic Acids Res.* **41** 6785–92

[64] Giudice E and Lavery R 2002 Simulations of nucleic acids and their complexes *Acc. Chem. Res.* **35** 350–7

[65] Spiegel K and Magistrato A 2006 Modeling anticancer drug–DNA interactions via mixed QM/MM molecular dynamics simulations *Org. Biomol. Chem.* **4** 2507–17

[66] Wang H and Laughton C A 2007 Molecular modelling methods for prediction of sequence-selectivity in DNA recognition *Methods* **42** 196–203

[67] Sharma S, Ding F and Dokholyan N V 2007 Multiscale modeling of nucleosome dynamics *Biophys. J.* **92** 1457–70

[68] Gao H and Kong Y 2004 Simulations of DNA-nanotube interactions *Annu. Rev. Mater. Res.* **34** 129–52

[69] Aksimentiev A, Heng J B, Timp G and Schulten K 2004 Microscopic kinetics of DNA translocation through synthetic nanopores *Biophys. J.* **87** 2086–97

[70] Pérez A, Luque F J and Orozco M 2007 Dynamics of B-DNA on the microsecond time scale *J. Am. Chem. Soc.* **129** 14739–45

[71] Freddolino P L, Arkhipov A S, Larson S B, McPherson A and Schulten K 2006 Molecular dynamics simulations of the complete satellite tobacco mosaic virus *Structure* **14** 437–49

[72] Sanbonmatsu K Y 2012 Computational studies of molecular machines: the ribosome *Curr. Opin. Struct. Biol.* **22** 168–74

[73] Rau D C, Lee B and Parsegian V A 1984 Measurement of the repulsive force between polyelectrolyte molecules in ionic solution: hydration forces between parallel DNA double helices *Proc. Natl Acad. Sci. USA* **81** 2621–5

[74] Maffeo C, Luan B and Aksimentiev A 2012 End-to-end attraction of duplex DNA *Nucleic Acids Res.* **40** 3812–21

[75] Rothemund P W K 2006 Folding DNA to create nanoscale shapes and patterns *Nature* **440** 297–302

[76] Dietz H, Douglas S M and Shih W M 2009 Folding DNA into twisted and curved nanoscale shapes *Science* **325** 725–30

[77] Ke Y, Douglas S M, Liu M, Sharma J, Cheng A, Leung A, Liu Y, Shih W M and Yan H 2009 Multilayer DNA origami packed on a square lattice *J. Am. Chem. Soc.* **131** 15903–8

[78] Saunders M G and Voth G A 2013 Coarse-graining methods for computational biology *Annu. Rev. Biophys.* **42** 73–93

[79] Potoyan D A, Savelyev A and Papoian G A 2013 Recent successes in coarse-grained modeling of DNA *WIREs Comput. Mol. Sci.* **3** 69–83

[80] Hinckley D M, Freeman G S, Whitmer J K and de Pablo J J 2013 An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: structure, thermodynamics, and dynamics of hybridization *J. Chem. Phys.* **139** 144903

[81] Šulc P, Romano F, Ouldridge T E, Rovigatti L, Doye J P K and Louis A A 2012 Sequence-dependent thermodynamics of a coarse-grained DNA model *J. Chem. Phys.* **137** 135101

[82] De Biase P M, Solano C J F, Markosyan S, Czapla L and Noskov S Yu 2012 BROMOC-D: Brownian dynamics/Monte-Carlo program suite to study ion and DNA permeation in nanopores *J. Chem. Theory Comput.* **8** 2540–51

[83] Savelyev A and Papoian G A 2010 Chemically accurate coarse graining of double-stranded DNA *Proc. Natl Acad. Sci. USA* **107** 20340–5

[84] He Y, Maciejczyk M, Ołdziej S, Scheraga H A and Liwo A 2013 Mean-field interactions between nucleic-acid-base dipoles can drive the formation of a double helix *Phys. Rev. Lett.* **110** 098101

[85] Hsu C W, Fyta M, Lakatos G, Melchionna S and Kaxiras E 2012 *Ab initio* determination of coarse-grained interactions in double-stranded DNA *J. Chem. Phys.* **137** 105102

[86] Morriss-Andrews A, Rottler J and Plotkin S S 2010 A systematically coarse-grained model for DNA and its predictions for persistence length, stacking, twist, and chirality *J. Chem. Phys.* **132** 035105

[87] Dans P D, Zeida A, Machado M R and Pantano S 2010 A coarse grained model for atomic-detailed DNA simulations with explicit electrostatics *J. Chem. Theory Comput.* **6** 1711–25

[88] Maciejczyk M, Spasic A, Liwo A and Scheraga H A 2010 Coarse-grained model of nucleic acid bases *J. Comput. Chem.* **31** 1644–55

[89] Gopal S M, Mukherjee S, Cheng Y-M and Feig M 2010 PRIMO/PRIMONA: a coarse-grained model for proteins and nucleic acids that preserves near-atomistic accuracy *Proteins: Struct., Funct., Bioinf.* **78** 1266–81

[90] DeMille R C, Cheatham T E and Molinero V 2011 A coarse-grained model of DNA with explicit solvation by water and ions *J. Phys. Chem.* B **115** 132–42

[91] SantaLucia J Jr 1998 A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics *Proc. Natl Acad. Sci. USA* **95** 1460–5

[92] Maffeo C, Ngo T T M, Ha T and Aksimentiev A 2014 A coarse-grained model of unstructured single-stranded DNA derived from atomistic simulation and single-molecule experiment *J. Chem. Theory Comput.* **10** 2891–6

[93] Nuss M E and Kollman P A 1979 Electrostatic potentials of deoxydinucleoside monophosphates: I. Deoxydinucleoside monophosphates and actinomycin chromophore interactions *J. Med. Chem.* **22** 1517–24

[94] Kollman P A, Weiner P K and Dearing A 1981 Studies of nucleotide conformations and interactions. The relative stabilities of double-helical B-DNA sequence isomers *Biopolymers* **20** 2583–621

[95] Straatsma T P and McCammon J A 1991 Multiconfiguration thermodynamic integration *J. Chem. Phys.* **95** 1175

[96] Straatsma T P and McCammon J A 1992 Computational alchemy *Annu. Rev. Phys. Chem.* **43** 407–35

[97] Baker C M, Lopes P E M, Zhu X, Roux B and MacKerell A D 2010 Accurate calculation of hydration free energies using pair-specific Lennard-Jones parameters in the CHARMM Drude polarizable force field *J. Chem. Theory Comput.* **6** 1181–98

[98] Ponder J W *et al* 2010 Current status of the AMOEBA polarizable force field *J. Phys. Chem.* B **114** 2549–64

[99] Savelyev A and MacKerell A D 2014 All-atom polarizable force field for DNA based on the classical Drude oscillator model *J. Comput. Chem.* **35** 1219–39

[100] Savelyev A and MacKerell A D 2014 Balancing the interactions of ions, water, and DNA in the Drude polarizable force field *J. Phys. Chem.* B **118** 6742–57

[101] Mackinnon R 2004 Potassium channels and the atomic basis of selective ion conduction (nobel lecture) *Angew. Chem. Int. Edn* **43** 4265–77

[102] van Gunsteren W F and Berendsen H J C 1990 Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry *Angew. Chem. Int. Edn Engl.* **29** 992–1023

[103] Nilsson L and Karplus M 1986 Empirical energy functions for energy minimization and dynamics of nucleic acids *J. Comput. Chem.* **7** 591–616

[104] Darden T, York D and Pedersen L 1993 Particle mesh Ewald: an $n$ log($n$) method for Ewald sums in large systems *J. Chem. Phys.* **98** 10089–92

[105] Foloppe N and MacKerrell A D Jr 2000 All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data *J. Comput. Chem.* **21** 86–104

[106] Cornell W D, Cieplak P, Bayly C I, Gould I R, Merz K M, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W and Kollman P A 1995 A second generation force field for the simulation of proteins, nucleic acids, and organic molecules *J. Am. Chem. Soc.* **117** 5179–97

[107] Cheatham T E and Case D A 2013 Twenty-five years of nucleic acid simulations *Biopolymers* **99** 969–77

[108] Cheatham T E, Cieplak P and Kollman P A 1999 A modified version of the Cornell *et al* force field with improved sugar pucker phases and helical repeat *J. Biomol. Struct. Dyn.* **16** 845–62

[109] Várnai P, Djuranovic D, Lavery R and Hartmann B 2002 $\alpha/\gamma$ transitions in the B-DNA backbone *Nucleic Acids Res.* **30** 5398–406

[110] Perez A, Marchan I, Svozil D, Sponer J, Cheatham T E, Laughton C A and Orozco M 2007 Refinement of the AMBER force field for nucleic acids: improving the description of $\alpha/\gamma$ conformers *Biophys. J.* **92** 3817–29

[111] Hart K, Foloppe N, Baker C M, Denning E J, Nilsson L and MacKerell A D 2012 Optimization of the CHARMM additive force field for DNA: improved treatment of the BI/BII conformational equilibrium *J. Chem. Theory Comput.* **8** 348–62

[112] Guy A T, Piggot T J and Khalid S 2012 Single-stranded DNA within nanopores: conformational dynamics and implications for sequencing; a molecular dynamics simulation study *Biophys. J.* **103** 1028–36

[113] Ditzler M A, Otyepka M, Sponer J and Walter N G 2010 Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in *Acc. Chem. Res.* **43** 40–7

[114] Banás P, Hollas D, Zgarbová M, Jurecka P, Orozco M, Cheatham T E, Sponer J and Otyepka M 2010 Performance of molecular mechanics force fields for RNA simulations: stability of UUCG and GNRA hairpins *J. Chem. Theory Comput.* **6** 3836–49

[115] Zgarbová M, Jurečka P, Banáš P, Otyepka M, Šponer J E, Leontis N B, Zirbel C L and Šponer J 2011 Noncanonical hydrogen bonding in nucleic acids. benchmark evaluation of key base phosphate interactions in folded RNA molecules using quantum-chemical calculations and molecular dynamics simulations *J. Phys. Chem.* A **115** 11277–92

[116] Åqvist J 1990 Ion water interaction potentials derived from free-energy perturbation simulations *J. Phys. Chem.* **94** 8021–4

[117] Beglov D and Roux B 1994 Finite representation of an infinite bulk system: solvent boundary potential for computer simulations *J. Chem. Phys.* **100** 9050–63

[118] Auffinger P, Cheatham T E and Vaiana A C 2007 Spontaneous formation of KCl aggregates in biomolecular simulations: a force field issue? *J. Chem. Theory Comput.* **3** 1851–9

[119] Luo Y and Roux B 2009 Simulation of osmotic pressure in concentrated aqueous salt solutions *J. Phys. Chem. Lett.* **1** 183–9

[120] Joung I S and Cheatham T E 2008 Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations *J. Phys. Chem.* B **112** 9020–41

[121] Joung I S and Cheatham T E 2009 Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters *J. Phys. Chem.* B **113** 13279–90

[122] Knobler C M and Gelbart W M 2009 Physical chemistry of DNA viruses *Annu. Rev. Phys. Chem.* **60** 367–83

[123] Gelbart W M and Knobler C M 2009 Virology pressurized viruses *Science* **323** 1682–3

[124] Luo Y, Egwolf B, Walters D E and Roux B 2010 Ion selectivity of $\alpha$-hemolysin with a $\beta$-cyclodextrin adapter: I. Single ion potential of mean force and diffusion coefficient *J. Phys. Chem.* B **144** 952–8

[125] Yoo J and Aksimentiev A 2012 Competitive binding of cations to duplex DNA revealed through molecular dynamics simulations *J. Phys. Chem.* B **116** 12946–54

[126] Bai Y, Greenfeld M, Travers K J, Chu V B, Lipfert J, Doniach S and Herschlag D 2007 Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids *J. Am. Chem. Soc.* **129** 14981–88

[127] Qiu X, Andresen K, Kwok L W, Lamb J S, Park H Y and Pollack L 2007 Inter-DNA attraction mediated by divalent counterions *Phys. Rev. Lett.* **99** 38104

[128] Li L, Pabit S A, Lamb J S, Park H Y and Pollack L 2008 Closing the lid on DNA end-to-end stacking interactions *Appl. Phys. Lett.* **92** 223901

[129] Chen A A and García A E 2013 High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations *Proc. Natl Acad. Sci. USA* **110** 16820–5

[130] Baumann S G, Smith S B, Bloomfield V A and Bustamante C 1997 Ionic effects on the elasticity of single DNA molecules *Proc. Natl Acad. Sci. USA* **94** 6185–90

[131] Young M A, Jayaram B and Beveridge D L 1998 Local dielectric environment of B-DNA in solution: results from a 14 ns molecular dynamics trajectory *J. Phys. Chem.* B **102** 7666–9

[132] Grosberg A Yu, Nguyen T T and Shklovskii B I 2002 Colloquium: the physics of charge inversion in chemical and biological systems *Rev. Mod. Phys.* **74** 329–45

[133] Kilic M S, Bazant M Z and Ajdari A 2007 Steric effects in the dynamics of electrolytes at large applied voltages: II. Modified Poisson–Nernst–Planck equations *Phys. Rev.* E **75** 021503

[134] Chaudhry J H, Comer J, Aksimentiev A and Olson L N 2014 A stabilized finite element method for modified Poisson–Nernst–Planck equations to determine ion flow through a nanopore *Commun. Comput. Phys* **15** 93–125

[135] Le Bret M and Zimm B H 1984 Distribution of counterions around a cylindrical polyelectrolyte and Manning's condensation theory *Biopolymers* **23** 287–312

[136] Abramowitz M and Stegun I A 1965 *Handbook of Mathematical Functions* (New York: Dover)

[137] Manning G 1981 Limiting laws and counterion condensation in polyelectrolyte solutions: VII. Electrophoretic mobility and conductance *J. Phys. Chem.* **85** 1506

[138] Manning G S 1996 The critical onset of counterion condensation: a survey of its experimental and theoretical basis *Ber. Bunsenges. Phys. Chem.* **100** 909–22

[139] Manning G S 2008 Approximate solutions to some problems in polyelectrolyte theory involving nonuniform charge distributions *Macromolecules* **41** 6217–27

[140] Manning G S 2001 Counterion condensation on a helical charge lattice *Macromolecules* **34** 4650–5

[141] Manning G S and Mohanty U 1997 Counterion condensation on ionic oligomers *Physica* A **247** 196–204

[142] Manning R S, Maddocks J H and Kahn J D 1996 A continuum rod model of sequence-dependent DNA structure *J. Chem. Phys.* **105** 5626–46

[143] Montoro J C G and Abascal J L F 1995 Ionic distribution around simple DNA models: I. Cylindrically averaged properties *J. Chem. Phys.* **103** 8273–84

[144] Gavryushov S and Zielenkiewicz P 1998 Electrostatic potential of B-DNA: effect of interionic correlations *Biophys. J.* **75** 2732–42

[145] Manning G S 1978 The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides *Quart. Rev. Biophys.* **2** 179–246

[146] Hud N V and Plavec J 2003 A unified model for the origin of DNA sequence-directed curvature *Biopolymers* **69** 144–58

[147] Das R, Mills T T, Kwok L W, Maskel G S, Millett I S, Doniach S, Finkelstein K D, Herschlag D and Pollack L 2003 Counterion distribution around DNA probed by solution x-ray scattering *Phys. Rev. Lett.* **90** 188103

[148] Andresen K, Das R, Park H Y, Smith H, Kwok L W, Lamb J S, Kirkland E J, Herschlag D, Finkelstein K D and Pollack L 2004 Spatial distribution of competing ions around DNA in solution *Phys. Rev. Lett.* **93** 248103

[149] Andresen K, Qiu X, Pabit S A, Lamb J S, Park H Y, Kwok L W and Pollack L 2008 Mono- and trivalent ions around DNA: a small-angle scattering study of competition and interactions *Biophys. J.* **95** 287–95

[150] Rueda M, Cubero E, Laughton C A and Orozco M 2004 Exploring the counterion atmosphere around DNA: what can be learned from molecular dynamics simulations? *Biophys. J.* **87** 800–11

[151] Várnai P and Zakrzewska K 2004 DNA and its counterions: a molecular dynamics study *Nucleic Acids Res.* **32** 4269–80

[152] Ponomarev S, Thayer K M and Beveridge D L 2004 Ion motions in molecular dynamics simulations on DNA *Proc. Natl Acad. Sci. USA* **101** 14771–5

[153] Draper D E, Grilley D and Soto A M 2005 Ions and RNA folding *Annu. Rev. Biophys. Biomol. Struct.* **34** 221–43

[154] Walter F, Murchie A I H and Lilley D M J 1998 Folding of the four-way RNA junction of the hairpin ribozyme *Biochemistry* **37** 17629–36

[155] Cate J H, Gooding A R, Podell E, Zhou K, Golden B L, Kundrot C E, Cech T R and Doudna J A 1996 Crystal structure of a group i ribozyme domain: principles of RNA packing *Science* **273** 1678–85

[156] McKinney S A, Déclais A-C C, Lilley D M J and Ha T 2003 Structural dynamics of individual Holliday junctions *Nature Struct. Biol.* **10** 93–7

[157] Hohng S, Zhou R, Nahas M K, Yu J, Schulten K, Lilley D M J and Ha T 2007 Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the Holliday junction *Science* **318** 279–83

[158] Soto A M, Misra V and Draper D E 2007 Tertiary structure of an RNA pseudoknot is stabilized by 'diffuse' $Mg^{2+}$ ions *Biochemistry* **46** 2973–83

[159] Jiao D, King C, Grossfield A, Darden T A and Ren P 2006 Simulation of $Ca^{2+}$ and $Mg^{2+}$ solvation using polarizable atomic multipole potential *J. Phys. Chem.* B **110** 18553–9

[160] Callahan K M, Casillas-Ituarte N N, Roeselová M, Allen H C and Tobias D J 2010 Solvation of magnesium dication: molecular dynamics simulation and vibrational spectroscopic study of magnesium chloride in aqueous solutions *J. Phys. Chem.* A **114** 5141–8

[161] Bleuzen A, Pittet P-A, Helm L and Merbach A E 1997 Water exchange on magnesium(II) in aqueous solution: a variable temperature and pressure 17o NMR study *Magn. Res. Chem.* **35** 765–73

[162] Piquemal J-P, Perera L, Cisneros G A, Ren P, Pedersen L G and Darden T A 2006 Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field: from energetics to structure *J. Chem. Phys.* **125** 054511

[163] Mamatkulov S, Fyta M and Netz R R 2013 Force fields for divalent cations based on single-ion and ion-pair properties *J. Chem. Phys.* **138** 024505

[164] Allnér O, Nilsson L and Villa A 2012 Magnesium ion–water coordination and exchange in biomolecular simulations *J. Chem. Theory Comput.* **8** 1493–502

[165] Saxena A and Sept D 2013 Multisite ion models that improve coordination and free energy calculations in molecular dynamics simulations *J. Chem. Theory Comput.* **9** 3538–42

[166] Sotomayor M, Vasquez V, Perozo E and Schulten K 2007 Ion conduction through MscS as determined by electrophysiology and simulation *Biophys. J.* **92** 886–902

[167] Joo C, McKinney S A, Lilley D M J and Ha T 2004 Exploring rare conformational species and ionic effects in DNA Holliday junctions using single-molecule spectroscopy *J. Mol. Biol.* **341** 739–51

[168] Auffinger P and Westhof E 2000 Water and ion binding around RNA and DNA (c,g) oligomers *J. Mol. Biol.* **300** 1113–31

[169] Cheng Y, Korolev N and Nordenskiöld L 2006 Similarities and differences in interaction of $K^+$ and $Na^+$ with condensed ordered DNA. A molecular dynamics computer simulation study *Nucleic Acids Res.* **34** 686–96

[170] MacKerell A D 1997 Influence of magnesium ions on duplex DNA structural, dynamic, and solvation properties *J. Phys. Chem.* B **101** 646–50

[171] Li W, Nordenskiöld L and Mu Y 2011 Sequence-specific $Mg^{2+}$–DNA interactions: a molecular dynamics simulation study *J. Phys. Chem.* B **115** 14713–20

[172] Todd B A, Parsegian V A, Shirahata A, Thomas T J and Rau D C 2008 Attractive forces between cation condensed DNA double helices *Biophys. J.* **94** 4775–82

[173] DeRouchey J, Hoover B and Rau D C 2013 A comparison of DNA compaction by arginine and lysine peptides: a physical basis for arginine rich protamines *Biochemistry* **52** 3000–9

[174] Korolev N, Lyubartsev A P, Laaksonen A and Nordenskiöld L 2002 On the competition between water, sodium ions, and spermine in binding to DNA: a molecular dynamics computer simulation study *Biophys. J.* **82** 2860–75

[175] Korolev N, Lyubartsev A P, Laaksonen A and Nordenskiöld L 2003 A molecular dynamics simulation study of oriented DNA with polyamine and sodium counterions: diffusion and averaged binding of water and cations *Nucleic Acids Res.* **31** 5971–81

[176] Korolev N, Lyubartsev A P, Laaksonen A and Nordenskiöld L 2004 Molecular dynamics simulation study of oriented polyamine- and Na-DNA: sequence specific interactions and effects on DNA structure *Biopolymers* **73** 542–55

[177] Luan B and Aksimentiev A 2010 Electric and electrophoretic inversion of the DNA charge in multivalent electrolytes *Soft Matter* **6** 243–6

[178] Shklovskii B I 1999 Screening of a macroion by multivalent ions: correlation-induced inversion of charge *Phys. Rev.* E **60** 5802–11

[179] Nguyen T T, Rouzina I and Shklovskii B I 2000 Reentrant condensation of DNA induced by multivalent counterions *J. Chem. Phys.* **112** 2562

[180] Zhao Q, Sigalov G, Dimitrov V, Dorvel B, Mirsaidov U, Sligar S, Aksimentiev A and Timp G 2007 Detecting SNPs using a synthetic nanopore *Nano Lett.* **7** 1680–5

[181] Haran T E and Mohanty U 2009 The unique structure of A-tracts and intrinsic DNA bending *Quart. Rev. Biophys.* **42** 41–81

[182] Beveridge D L *et al* 2004 Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides: I. Research design and results on d($c_p$g) steps *Biophys. J.* **87** 3799–813

[183] Hagerman P J 1988 Flexibility of DNA *Annu. Rev. Biophys. Biophys. Chem.* **17** 265–86

[184] Fuller F B 1971 The writhing number of a space curve *Proc. Natl Acad. Sci. USA* **68** 815–9

[185] Barkley M D and Zimm B H 1979 Theory of twisting and bending of chain macromolecules; analysis of the fluorescence depolarization of DNA *J. Chem. Phys.* **70** 2991–3007

[186] Shimada J and Yamakawa H 1984 Ring-closure probabilities for twisted wormlike chains: application to DNA *Macromolecules* **17** 689–98

[187] Bockelmann U 2004 Single-molecule manipulation of nucleic acids *Curr. Opin. Struct. Biol.* **14** 368–73

[188] Cluzel P, Lebrun A, Heller C, Lavery R, Viovy J-L, Chatenay D and Caron F 1996 DNA: an extensible molecule *Science* **271** 792–4

[189] Harris S A 2004 The physics of DNA stretching *Contemp. Phys.* **45** 11

[190] Konrad M W and Bolonick J I 1996 Molecular dynamics simulation of DNA stretching is consistent with the tension observed for extension and strand separation and predicts a novel ladder structure *J. Am. Chem. Soc.* **118** 10989

[191] Rouzina I and Bloomfield V 2001 Force-induced melting of the DNA double heix 1. Thermodynamic analysis *Biophys. J.* **80** 882–93

[192] Marko J F 1998 DNA under high tension: overstretching, undertwisting, and relaxation dynamics *Phys. Rev.* E **57** 2134–49

[193] Léger J F, Romano G, Sarkar A, Robert J, Bourdieu L, Chatenay D and Marko J F 1999 Structural transitions of a twisted and stretched DNA molecule *Phys. Rev. Lett.* **83** 1066–9

[194] Bustamante C, Bryant Z and Smith S B 2003 Ten years of tension: single-molecule DNA mechanics *Nature* **421** 423–7

[195] Ke C, Jiang Y, Rivera M, Clark R L and Marszalek P E 2007 Pulling geometry induced errors in single molecule force spectroscopy measurements *Biophys. J.* **92** 76–8

[196] Seol Y, Skinner G M and Visscher K 2007 Stretching of homopolymeric RNA reveals single-strand helices and base-stacking *Phys. Rev. Lett.* **98** 158103–6

[197] Isralewitz B, Izrailev S and Schulten K 1997 Binding pathway of retinal to bacterio-opsin: a prediction by molecular dynamics simulations *Biophys. J.* **73** 2972–9

[198] Uzawa T, Isoshima T, Ito Y, Ishimori K, Makarov D E and Plaxco K W 2013 Sequence and temperature dependence of the end-to-end collision dynamics of single-stranded DNA *Biophys. J.* **104** 2485–92

[199] Ouldridge T E, Louis A A and Doye J P K 2011 Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model *J. Chem. Phys.* **134** 085101

[200] Harris S A, Sands Z A and Laughton C A 2005 Molecular dynamics simulations of duplex stretching reveal the importance of entropy in determining the biomechanical properties of DNA *Biophys. J.* **88** 1684–91

[201] Feig M and Pettitt B M 1998 Structural equilibrium of DNA represented with different force fields *Biophys. J.* **75** 134–49

[202] Luan B and Aksimentiev A 2008 Strain softening in stretched DNA *Phys. Rev. Lett.* **101** 118101

[203] Martyna G J, Tobias D J and Klein M L 1994 Constant pressure molecular dynamics algorithms *J. Chem. Phys.* **101** 4177–89

[204] Bustamante C, Smith S B, Liphardt J and Smith D 2000 Single-molecule studies of DNA mechanics *Curr. Opin. Struct. Biol.* **10** 279–85

[205] Kouzine F, Sanford S, Elisha-Feil Z and Levens D 2008 The functional response of upstream DNA to dynamic supercoiling *in vivo*, *Nature Struct. Mol. Biol.* **15** 146–54

[206] Brooks T A and Hurley L H 2009 The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics *Nature Rev. Cancer* **9** 849–61

[207] Lankas F, Sponer J, Langowski J and Cheatham T E 2003 DNA basepair step deformability inferred from molecular dynamics simulations *Biophys. J.* **85** 2872–83

[208] Cloutier T E and Widom J 2004 Spontaneous sharp bending of double-stranded DNA *Mol. Cell* **14** 355–62

[209] Gorin A A, Zhurkin V B and Olson W K 1995 B-DNA twisting correlates with base-pair morphology *J. Mol. Biol.* **247** 34–48

[210] Olson W K, Gorin A A, Lu X-J, Hock L M and Zhurkin V B 1998 DNA sequence-dependent deformability deduced from protein-DNA crystal complexes *Proc. Natl Acad. Sci. USA* **95** 11163–8

[211] Yonetani Y and Kono H 2009 Sequence dependencies of DNA deformability and hydration in the minor groove *Biophys. J.* **97** 1138–47

[212] Mazur A K 1999 Symplectic integration of closed chain rigid body dynamics with internal coordinate equations of motion *J. Chem. Phys.* **111** 1407–14

[213] Mazur A K 2011 Local elasticity of strained DNA studied by all-atom simulations *Phys. Rev.* E **84** 021903

[214] Noy A and Golestanian R 2012 Length scale dependence of DNA mechanical properties *Phys. Rev. Lett.* **109** 228101

[215] Mazur A K 2013 Comment on Length scale dependence of DNA mechanical properties *Phys. Rev. Lett.* **111** 179801

[216] Noy A and Golestanian R 2013 Noy and Golestanian reply *Phys. Rev. Lett.* **111** 179802

[217] Marko J F and Siggia E D 1994 Bending and twisting elasticity of DNA *Macromolecules* **27** 981–8

[218] Gore J, Bryant Z, Nöllmann M, Le M U, Cozzarelli N R and Bustamante C 2006 DNA overwinds when stretched *Nature* **442** 836–9

[219] Gross P, Laurens N, Oddershede L B, Bockelmann U, Peterman E J G and Wuite G J L 2011 Quantifying how DNA stretches, melts and changes twist under tension *Nature Phys.* **7** 731–6

[220] MacKerell A D Jr and Banavali N K 2000 All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution *J. Comput. Chem.* **21** 105–20

[221] Olson W K *et al* 2001 A standard reference frame for the description of nucleic acid base-pair geometry *J. Mol. Biol.* **313** 229–37

[222] Lionnet T, Joubaud S, Lavery R, Bensimon D and Croquette V 2006 Wringing out DNA *Phys. Rev. Lett.* **96** 178102

[223] Sheinin M Y and Wang M D 2009 Twist-stretch coupling and phase transition during DNA supercoiling *Phys. Chem. Chem. Phys.* **11** 4800–3

[224] Randall G L, Zechiedrich L and Pettitt B M 2009 In the absence of writhe, DNA relieves torsional stress with localized, sequence-dependent structural failure to preserve B-form *Nucleic Acids Res.* **37** 5568–77

[225] Kahn J D and Crothers D M 1992 Protein-induced bending and DNA cyclization *Proc. Natl Acad. Sci. USA* **89** 6343–7

[226] Mathew-Fenn R S, Das R and Harbury P A B 2008 Remeasuring the double helix *Science* **322** 446–9

[227] Du Q, Smith C, Shiffeldrim N, Vologodskaia M and Vologodskii A 2005 Cyclization of short DNA fragments and bending fluctuations of the double helix *Proc. Natl Acad. Sci. USA* **102** 5397–402

[228] Mastroianni A J, Sivak D A, Geissler P L and Alivisatos A P 2009 Probing the conformational distributions of subpersistence length DNA *Biophys. J.* **97** 1408–17

[229] Qu H, Wang Y, Tseng C-Y and Zocchi G 2011 Critical torque for kink formation in double-stranded DNA *Phys. Rev.* X **1** 021008

[230] Lankas F, Lavery R and Maddocks J H 2006 Kinking occurs during molecular dynamics simulations of small DNA minicircles *Structure* **14** 1527–34

[231] Harris S A, Laughton C A and Liverpool T B 2008 Mapping the phase diagram of the writhe of DNA nanocircles using atomistic molecular dynamics simulations *Nucleic Acids Res.* **36** 21–9

[232] Andrews A J and Luger K 2011 Nucleosome structure(s) and stability: variations on a theme *Annu. Rev. Biophys.* **40** 99–117

[233] Brower-Toland B D, Smith C L, Yeh R C, Lis J T, Peterson C L and Wang M D 2002 Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA *Proc. Natl Acad. Sci. USA* **99** 1960–5

[234] Brower-Toland B, Wacker D A, Fulbright R M, Lis J T, Kraus W L and Wang M D 2005 Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes *J. Mol. Biol.* **346** 135–46

[235] Mihardja S, Spakowitz A J, Zhang Y and Bustamante C 2006 Effect of force on mononucleosomal dynamics *Proc. Natl Acad. Sci. USA* **103** 15871–6

[236] Hall M A, Shundrovsky A, Bai L, Fulbright R M, Lis J T and Wang M D 2009 High-resolution dynamic mapping of histone–DNA interactions in a nucleosome *Nature Struct. Mol. Biol.* **16** 124–9

[237] Arya G and Schlick T 2006 Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model *Proc. Natl Acad. Sci. USA* **103** 16236–41

[238] Voltz K, Trylska J, Calimet N, Smith J C and Langowski J 2012 Unwrapping of nucleosomal DNA ends: a multiscale molecular dynamics study *Biophys. J.* **102** 849–58

[239] Fan Y, Korolev N, Lyubartsev A P and Nordenskiöld L 2013 An advanced coarse-grained nucleosome core particle model for computer simulations of nucleosome–nucleosome interactions under varying ionic conditions *PLoS ONE* **8** e54228

[240] Ettig R, Kepper N, Stehr R, Wedemann G and Rippe K 2011 Dissecting DNA–histone interactions in the nucleosome by molecular dynamics simulations of DNA unwrapping *Biophys. J.* **101** 1999–2008

[241] Cisse I I, Kim H and Ha T 2012 A rule of seven in Watson–Crick base-pairing of mismatched sequences *Nature Struct. Mol. Biol.* **19** 623–7

[242] Hagan M F, Dinner A R, Chandler D and Chakraborty A K 2003 Atomistic understanding of kinetic pathways for single base-pair binding and unbinding in DNA *Proc. Natl Acad. Sci. USA* **100** 13922–7

[243] Sorin E J, Rhee Y M, Nakatani B J and Pande V S 2003 Insights into nucleic acid conformational dynamics from massively parallel stochastic simulations *Biophys. J.* **85** 790–803

[244] Garcia A E and Paschek D 2008 Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin *J. Am. Chem. Soc.* **130** 815–7

[245] Portella G and Orozco M 2010 Multiple routes to characterize the folding of a small DNA hairpin *Angew. Chem. Int. Edn Engl.* **49** 7673–6

[246] Kara M and Zacharias M 2014 Theoretical studies of nucleic acids folding *WIREs Comput. Mol. Sci.* **4** 116–26

[247] Rau D C and Parsegian V A 1992 Direct measurement of the intermolecular forces between counterion-condensed DNA double helices. Evidence for long range attractive hydration forces *Biophys. J.* **61** 246–59

[248] Pelta J, Livolant F and Sikorav J L 1996 DNA aggregation induced by polyamines and cobalthexamine *J. Biol. Chem.* **271** 5656–62

[249] Todd B A and Rau D C 2008 Interplay of ion binding and attraction in DNA condensed by multivalent cations *Nucleic Acids Res.* **36** 501–10

[250] Parsegian V A, Rand R P, Fuller N L and Rau D C 1986 Osmotic stress for the direct measurement of intermolecular forces *Methods Enzymol* **127** 400–16

[251] Rau D C and Parsegian V A 1992 Direct measurement of temperature-dependent solvation forces between DNA double helices *Biophys. J.* **61** 260–71

[252] Kornyshev A A and Leikin L 1997 Theory of interaction between helical molecules *J. Chem. Phys.* **107** 3656–74

[253] Kornyshev A A and Leikin S 1998 Electrostatic interaction between helical macromolecules in dense aggregates: an impetus for DNA poly- and meso-morphism *Proc. Natl Acad. Sci. USA* **95** 13579–84

[254] Kornyshev A A and Leikin S 1999 Electrostatic zipper motif for DNA aggregation *Phys. Rev. Lett.* **82** 4138–41

[255] Savelyev A and Papoian G A 2007 Inter-DNA electrostatics from explicit solvent molecular dynamics simulations *J. Am. Chem. Soc.* **129** 6060–1

[256] Luan B and Aksimentiev A 2008 DNA attraction in monovalent and divalent electrolytes *J. Am. Chem. Soc.* **130** 15754–5

[257] Podgornik R, Rau D C and Parsegian V A 1994 Parametrization of direct and soft steric-undulatory forces between DNA double helical polyelectrolytes in solutions of several different anions and cations *Biophys. J.* **66** 962–71

[258] Fuller D N, Rickgauer J P, Jardine P J, Grimes S, Anderson D L and Smith D E 2007 Ionic effects on viral DNA packaging and portal motor function in bacteriophage $\varphi$29 *Proc. Natl Acad. Sci. USA* **104** 11245–50

[259] Balasundaram D, Tabor C W and Tabor H 1991 Spermidine or spermine is essential for the aerobic growth of saccharomyces cerevisiae *Proc. Natl Acad. Sci. USA* **88** 5872–6

[260] Eisenberg T *et al* 2009 Induction of autophagy by spermidine promotes longevity *Nature Cell. Biol.* **11** 1305–14

[261] Hud N V and Vilfan I D 2005 Toroidal DNA condensates: unraveling the fine structure and the role of nucleation in determining size *Annu. Rev. Biophys. Biomol. Struct.* **34** 295–318

[262] Raspaud E, de la Cruz M O, Sikorav J L and Livolant F 1998 Precipitation of DNA by polyamines: a polyelectrolyte behavior *Biophys. J.* **74** 381–93

[263] Strey H H, Podgornik R, Rau D C and Parsegian V A 1998 DNA–DNA interactions *Curr. Opin. Struct. Biol.* **8** 309–13

[264] Rouzina I and Bloomfield V A 1996 Macroion attraction due to electrostatic correlation between screening counterions: I. Mobile surface-adsorbed ions and diffuse ion cloud *J. Phys. Chem.* **100** 9

[265] Shklovskii B I 1999 Wigner crystal model of counterion induced bundle formation of rod like polyelectrolytes *Phys. Rev. Lett.* **82** 3268–71

[266] Dai L, Mu Y, Nordenskiöld L and van der Maarel J R C 2008 Molecular dynamics simulation of multivalent-ion mediated attraction between DNA molecules *Phys. Rev. Lett.* **100** 118301

[267] Nakata M, Zanchetta G, Chapman B D, Jones C D, Cross J O, Pindak R, Bellini T and Clark N A 2007 End-to-end stacking and liquid crystal condensation of 6 to 20 base pair DNA duplexes *Science* **318** 1276

[268] Zanchetta G, Nakata M, Buscaglia M, Clark N A and Bellini T 2008 Liquid crystal ordering of DNA and RNA oligomers with partially overlapping sequences *J. Phys.: Condens. Matter* **20** 494214

[269] Rossi M, Zanchetta G, Klussmann S, Clark N and Bellini T 2013 Propagation of chirality in mixtures of natural and enantiomeric DNA oligomers *Phys. Rev. Lett.* **110** 107801

[270] Lieber M R, Ma Y, Pannicke U and Schwarz K 2003 Mechanism and regulation of human non-homologous DNA end-joining *Nature Rev. Mol. Cell Biol.* **4** 712–20

[271] Woo H-J and Roux B 2005 Calculation of absolute protein–ligand binding free energy from computer simulations *Proc. Natl Acad. Sci. USA* **102** 6825–30

[272] De Michele C, Rovigatti L, Bellini T and Sciortino F 2012 Self-assembly of short DNA duplexes: from a coarse-grained model to experiments through a theoretical link *Soft Matter* **8** 8388–98

[273] Sanger F and Coulson A R 1975 A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase *J. Mol. Biol.* **94** 441–8

[274] Branton D *et al* 2008 The potential and challenges of nanopore sequencing *Nature Biotechnol.* **26** 1146–53

[275] Venkatesan B M and Bashir R 2011 Nanopore sensors for nucleic acid analysis *Nature Nanotechnol.* **6** 615–24

[276] Timp W, Mirsaidov U M, Wang D, Comer J, Aksimentiev A and Timp G 2010 Nanopore sequencing: electrical measurements of the code of life *IEEE Trans. Nanotechnol.* **9** 281–94

[277] Hoagland D A, Arvanitidou E and Welch C 1999 Capillary electrophoresis measurements of the free solution mobility for several model polyelectrolyte systems *Macromolecules* **32** 6180–90

[278] Stellwagen E and Stellwagen N C 2002 Determining the electrophoretic mobility and translational diffusion coefficients of DNA molecules in free solution *Electrophoresis* **23** 2794–803

[279] Nkodo A E, Garnier J M, Tinland B, Ren H, Desruissaux C, McCormick L C, Drouin G and Slater G W 2001 Diffusion coefficient of DNA molecules during free solution electrophoresis *Electrophoresis* **22** 2424

[280] Yoo A and Aksimentiev A 2014 The physics of DNA in confinement, unpublished

[281] Stellwagen E, Lu Y and Stellwagen N C 2003 Unified description of electrophoresis and diffusion for DNA and other polyions *Biochemistry* **42** 11745–50

[282] Long D, Viovy J L and Ajdari A 1996 Simultaneous action of electric fields and nonelectric forces on a polyelectrolyte: motion and deformation *Phys. Rev. Lett.* **76** 3858–61

[283] Meagher R J, Won J-I, McCormick L C, Nedelcu S, Bertrand M M, Bertram J L, Drouin G, Barron A E and Slater G W 2005 End-labeled free-solution electrophoresis of DNA *Electrophoresis* **26** 331–50

[284] Viovy J L 2000 Electrophoresis of DNA and other polyelectrolytes: physical mechanisms *Rev. Mod. Phys.* **72** 813

[285] Sanger F, Nicklen S and Coulson A R 1977 DNA sequencing with chain-terminating inhibitors *Proc. Natl Acad. Sci. USA* **74** 5463

[286] Klepárník K and Boček P 2007 DNA diagnostics by capillary electrophoresis *Chem. Rev.* **107** 5279–317

[287] de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)

[288] Slater G W 2009 DNA gel electrophoresis: the reptation model(s) *Electrophoresis* **30** S181–7

[289] Kasianowicz J J, Brandin E, Branton D and Deamer D W 1996 Characterization of individual polynucleotide molecules using a membrane channel *Proc. Natl Acad. Sci. USA* **93** 13770–73

[290] Dekker C 2007 Solid-state nanopores *Nature Nanotechnol.* **2** 209–15

[291] Keyser U F 2011 Controlling molecular transport through nanopores *J. R. Soc. Interf.* **8** 1369–78

[292] Meller A, Nivon L and Branton D 2001 Voltage-driven DNA translocations through a nanopore *Phys. Rev. Lett.* **86** 3435–8

[293] Li J, Stein D, McMullan C, Branton D, Aziz M J and Golovchenko J A 2001 Ion-beam sculpting at nanometre length scales *Nature* **412** 166–9

[294] Storm A J, Chen J H, Ling X S, Zandergen H W and Dekker C 2003 Fabrication of solid-state nanopore with single-nanometre precision *Nature Mater.* **2** 537–40

[295] Heng J B, Ho C, Kim T, Timp R, Aksimentiev A, Grinkova Y V, Sligar S, Schulten K and Timp G 2004 Sizing DNA using a nanometer-diameter pore *Biophys. J.* **87** 2905–11

[296] Aksimentiev A 2010 Deciphering ionic current signatures of DNA transport through a nanopore *Nanoscale* **2** 468–83

[297] Deamer D W and Akeson M 2000 Nanopores and nucleic acids: prospects for ultrarapid sequencing *Trends Biotechnol.* **18** 147–51

[298] Smeets R M M, Keyser U F, Krapf D, Wu M-Y, Dekker N H and Dekker C 2006 Salt-dependence of ion transport and DNA translocation through solid-state nanopores *Nano Lett.* **6** 89–95

[299] Comer J, Dimitrov V, Zhao Q, Timp G and Aksimentiev A 2009 Microscopic mechanics of hairpin DNA translocation through synthetic nanopores *Biophys. J.* **96** 593–608

[300] Venta K, Shemer G, Puster M, Rodríguez-Manzo J A, Balan A, Rosenstein J K, Shepard K and Drndić M 2013 Differentiation of short, single-stranded DNA homopolymers in solid-state nanopores *ACS Nano* **7** 4629–36

[301] Comer J and Aksimentiev A 2012 Predicting the DNA sequence dependence of nanopore ion current using atomic-resolution Brownian dynamics *J. Phys. Chem.* C **116** 3376–93
[302] Gracheva M E, Aksimentiev A and Leburton J-P 2006 Electrical signatures of single-stranded DNA with single base mutations in a nanopore capacitor *Nanotechnology* **17** 3160–5
[303] Zwolak M and Di Ventra M 2008 Colloquium: physical approaches to DNA sequencing and detection *Rev. Mod. Phys.* **80** 141–65
[304] Huang S, He J, Chang S, Zhang P, Liang F, Li S, Tuchband M, Fuhrmann A, Ros R and Lindsay S 2010 Identifying single bases in a DNA oligomer with electron tunnelling *Nature Nanotechnol.* **5** 868–73
[305] Tsutsui M, Taniguchi M, Yokota K and Kawai T 2010 Identifying single nucleotides by tunnelling current *Nature Nanotechnol.* **5** 286–90
[306] Storm A J, Chen J H, Zandbergen H W and Dekker C 2005 Translocation of double-strand DNA through a silicon oxide nanopore *Phys. Rev.* E **71** 051903–13
[307] Keyser U, Koeleman B, Dorp S, Krapf D, Smeets R, Lemay S, Dekker N and Dekker C 2006 Direct force measurements on DNA in a solid-state nanopore *Nature Phys.* **2** 473–77
[308] Ghosal S 2006 Electrophoresis of a polyelectrolyte through a nanopore *Phys. Rev.* E **74** 041901
[309] Ghosal S 2007 Effect of salt concentration on the electrophoretic speed of a polyelectrolyte through a nanopore *Phys. Rev. Lett.* **98** 238104
[310] Luan B and Aksimentiev A 2008 Electro-osmotic screening of the DNA charge in a nanopore *Phys. Rev.* E **78** 021912
[311] van Dorp S, Keyser U F, Dekker N H, Dekker C and Lemay S G 2009 Origin of the electrophoretic force on DNA in solid-state nanopores *Nature Phys.* **5** 347–51
[312] Luan B and Aksimentiev A 2010 Control and reversal of the electrophoretic force on DNA in a charged nanopore *J. Phys.: Condens. Matter* **22** 454123
[313] Kowalczyk S W, Wells D B, Aksimentiev A and Dekker C 2012 Slowing down DNA translocation through a nanopore in lithium chloride *Nano Lett.* **12** 1038–44
[314] Besteman K, Van Eijk K and Lemay S G 2007 Charge inversion accompanies DNA condensation by multivalent ions *Nature Phys.* **3** 641