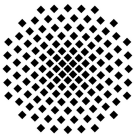


Universität Stuttgart

Simulations of DNA translocation through nanopores

Von der Fakultät für Mathematik und Physik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Stefan Kesselheim

aus Würselen

Hauptberichter:	Prof. Dr. Christian Holm
Mitberichter:	Prof. Dr. Johannes Kästner
	Prof. Dr. Roland Netz
Tag der mündlichen Prüfung:	29. Juli 2014

Institut für Computerphysik der Universität Stuttgart

2015

Contents

1 Zusammenfassung

Nanoporen sind vielversprechende Kandidaten für hochempfindliche, neuartige Sensoren. Mit ihnen kann der Transport winziger Mengen eines Materials, bis hin zu einzelnen Molekülen, nachgewiesen werden. Das grundsätzliche Messprinzip ist verblüffend einfach. Eine Membran mit einer einzigen Pore teilt einen Probenraum, der mit einer wässrigen Lösung gefüllt ist. Durch Anlegen einer Spannung zwischen beiden Teilen fließt ein elektrischer Strom durch die Pore. Verändert sich die Zusammensetzung des Porenraums kann dies als Modulation des elektrischen Stroms nachgewiesen werden. Dies kann etwa beim Durchqueren der Pore von Biomolekülen, z. B. DNA geschehen. Der Transport solcher Moleküle durch eine Nanopore wird als *Translokation* bezeichnet. Bereits frühe Experimente [1] zeigten, dass sich anhand verschiedener Modulationen des Stroms DNA-Moleküle mit unterschiedlichen Basensequenzen unterscheiden lassen. Die wohl wichtigste Vision der Forscherinnen und Forscher auf diesem Gebiet ist die Möglichkeit, Nanoporen zum Auslesen des Genoms zu nutzen. Eine günstige und schnelle Technologie für diese Aufgabe würde die Tür zu individualisierter medizinischer Behandlung öffnen, bei der Therapien anhand des individuellen genetischen Profils zugeschnitten werden könnten. Kommerzielle Sensoren sind bereits angekündigt [2], sind aber noch nicht verfügbar.

Die Ladung der DNA hat eine besondere Stellung. Ein hochgeladenes DNA-Molekül ist umgeben von einer Gegenionenwolke, die experimentell typischerweise aus Kaliumionen besteht. Wird eine Spannung angelegt, kommt es zu komplexen elektrohydrodynamischen Phänomenen wie Elektrophorese und Elektroosmose. Ein vollständiges Verständnis dieser Mechanismen ist anspruchsvoll. Daher wurden in dieser Dissertation Computersimulationsmethoden genutzt um ein besseres Verständnis zu erzielen. Diese ermöglichen es, sehr grundsätzliche, bisher nicht geklärte Phänomene zu verstehen. Die große Herausforderung besteht darin, die dabei auftretenden Transportprozesse realistisch im Computer abzubilden. Die Zeit- und Längenskalen, die dafür notwendig sind übersteigen das, was mit aktuellen Rechnern mit atomarer Auflösung simulierbar ist, um einige Größenordnungen. Daher wurden

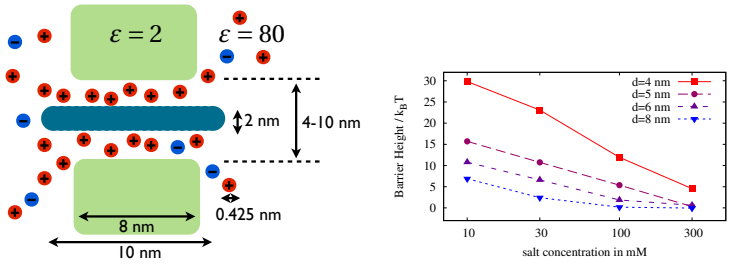
Methoden weiterentwickelt, mit denen eine vergröberte Modellierung mit einer reduzierten Anzahl an Freiheitsgraden möglich ist. Durch Simulationen auf atomarer Ebene wurde weiterhin untersucht, wie realistisch die genutzten vergröberten Modelle sind. So konnte gezeigt werden, unter welchen Bedingungen kompatible Ergebnisse zwischen den genutzten Modellskalen zu erzielen sind.

Zwei Hauptaspekte wurden untersucht. Einerseits wurde betrachtet, inwiefern elektrostatische Barrieren der Translokation von DNA entgegenwirken. Dies beinhaltet nicht nur die Rolle der Gegenionenwolke, die durch die Pore deformiert wird, sondern auch mögliche dielektrische Randkräfte hervorgerufen durch den Unterschied der dielektrischen Konstante von Wasser und Membran. Andererseits wurde die Frage untersucht, welche Mechanismen die beobachteten Modulationen des elektrischen Stroms aufgrund der Anwesenheit von DNA-Molekülen hervorrufen. Dabei wurde durch den Vergleich von Modellen auf drei Auflösungsniveaus, nämlich Kontinuum, vergröberte Partikelsimulation und atomistische Simulation, analysiert, inwiefern die Mechanismen quantitativ korrekt auf vergröbertem Niveau wiedergegeben werden.

Elektrostatische Barrieren

Eine der zentralen experimentellen Fragen besteht darin, wie man die Geschwindigkeit des DNA-Transports verringern kann, damit mehr Zeit besteht, das translokierende Molekül zu identifizieren. Zum besseren Verständnis der auftretenden Transportraten und -geschwindigkeiten haben wir Barrieren der freien Energie betrachtet, die dem Transport entgegenwirken bzw. ihn verlangsamen. Für diese Untersuchung wurde eine vergröberte molekulardynamische Simulation verwendet, bei der das Wasser implizit anhand seiner dielektrischen Konstante und die Ionen explizit als Partikel repräsentiert wurden.

Bisher ungeklärt war insbesondere die Frage, wie der Unterschied in der dielektrischen Konstante zwischen Wasser und Membran den Transport, insbesondere die auftretenden Barrieren, beeinflusst. In der gewählten vergröberten Darstellung kann der dielektrische Kontrast mit dem ICC$\star$ (induced charge computation)-Algorithmus miteinbezogen werden [3]. Dabei wird die Grenzflächen zwischen beiden Medien durch induzierte Ladungen dargestellt, die iterativ bestimmt werden. Hierzu ist lediglich die Kenntnis des elektrischen Feldes auf der Oberfläche notwendig, und insofern kann jeder übliche Elektrostatik-Löser leicht um diesen Aspekt erweitert werden. Das Skalierungsverhalten und die Periodizität der Lösung wird dabei von dem genutzten Löser übernommen. Eine grundlegende Implementierung des ICC$\star$-Algorithmus im



(a) (b)

Abbildung 1.1: (a) Skizze des Simulationsaufbaus. Eine Nanopore wird als zylindrische Kavität in einer ebenen Membran mit einer dielektrischen Konstante von $\varepsilon = 2$ und Wasser als dielektrisches Kontinuumm mit $\varepsilon = 80$ modelliert. Ein 10 nm langes Fragment doppelsträngiger DNA wird dabei durch viele überlappende Kugeln mit einem Durchmesser von 2 nm repräsentiert. Ionen werden dargestellt als geladene Kugeln mit 0.425 nm Radius. Aus dem Potential der mittleren Kraft bestimmen wir die Barrierenhöhe (b), die hier für verschiedene Porendurchmesser als Funktion der Salzkonzentration aufgetragen ist.

Simulationspaket ESPResSo lag bereits vor. Diese wurde parallelisiert und optimiert, und für leitende Oberflächen erweitert.

Abb. 1.1a zeigt eine Skizze der genutzen Simulationen. Eine generische Nanopore wird als zylindrischer Hohlraum in einer flachen Membran dargestellt. Ein kurzes Fragment doppelsträngiger DNA wird als 10 nm langer, geladener Zylinder und die umgebenden Ionen als sphärische Partikel mit 0,425 nm Radius, entsprechend dem *restricted primitive model* [4], mit einer Elementarladung repräsentiert. Die Oberfläche der Membran wird diskretisiert und die induzierten Ladungen in jedem Molekulardynamikschritt iterativ bestimmt. Es wurden je nach gewählten Bedingungen sehr große Barrieren in der freien Energie festgestellt. Abb. 1.1b zeigt eine Zusammenfassung der Ergebnisse. Für geringe Konzentrationen des umgebenden Elektrolyten werden große Translokationsbarrieren gefunden, da hier die Ladung der DNA nicht vollständig abgeschirmt wird. Steigt die Konzentration, so nimmt die Abschirmung zu und die Barriere sinkt. Entscheidend ist dabei das Verhältnis aus Porenradius und Debye-Länge. Die dielektrischen Randkräfte führen bei der gewählten Porenlänge etwa zu einer Verdopplung der auftretenden Barrieren.

Weiterhin wurden Translokationsbarrieren für ein vereinfachtes Modell ein-

zelsträngiger DNA-Moleküle bestimmt. Die dabei vorhandene Flexibilität der DNA führt zusätzlich zu einer entropischen Barriere, die in guter Näherung unabhängig von der Salzkonzentration ist. Die auftretenden elektrostatischen Barrieren erhöhen sich durch die Flexibilität deutlich, da sich immer Teile des DNA-Moleküls in der Nähe der Porenwand aufhalten. Eine wichtige Voraussetzung für diese Untersuchung war die Entwicklung einer geeigneten Reaktionskoordinate, die es ermöglicht den vollständigen Transportprozess eines Moleküls aus der Lösung in die Pore und wieder hinaus zu beschreiben. So konnten wir zeigen, dass die Barriere vor allem in dem Bereich auftritt, wo sich ein erstes Stück DNA in die Pore hineinbewegt.

Eine Alternative zu der vergröberten teilchenbasierten Darstellung ist die Wahl einer Kontinuumsbeschreibung. Im thermodynamischen Gleichgewicht wird die Ionenverteilung und das elektrostatische Potential in guter Näherung durch die Poisson-Boltzmann (PB)-Gleichung, eine nichtlineare partielle Differentialgleichung beschrieben. Es wurde eine sehr einfache Modellierung mit der PB-Gleichung durchgeführt. Es wurde dabei angenommen, dass das DNA-Fragment als sehr langer, geladener Zylinder zentriert in einem gleichermaßen langen ungeladenen Zylinder dargestellt werden kann. Das resultierende Problem ist eindimensional und kann mit einfachen Mitteln numerisch gelöst werden. Aus dieser Lösung lässt sich die Barriere pro Länge bestimmen. Multipliziert mit der Länge der Pore ergab sich eine hervorragende Korrelation zu den Ergebnissen, die molekulardynamisch unter Einbeziehung des dielektrischen Kontrasts ergeben hatten.

Die Berücksichtigung von dielektrischen Grenzflächen in der PB-Gleichung ist dabei mit der iPBS (Iterated Poisson-Boltzmann Solution) möglich. Diese Methode basiert auf einer Lösung der Poisson-Boltzmann-Gleichung mit der Finite-Elemente-Methode. Die gestellten Randbedingungen werden iteriert, und dabei die induzierten Ladungen auf dielektrischen Grenzflächen bestimmt. Eine Implementation dieser Methode wurde im Rahmen dieser Arbeit betreut. Sie wurde genutzt, um Translokationsbarrieren für doppelsträngige DNA-Fragmente endlicher Länge zu berechnen. Es konnte gezeigt werden, dass im Limes sehr langer Poren, unabhängig von der dielektrischen Konstante des Membranmaterials, dieselbe Barriere wie in der einfachen eindimensionalen Darstellung erhalten wird. Wird der dielektrische Kontrast berücksichtigt, erhöht sich die Barriere schneller bei längeren Poren, da der dielektrische Kontrast eine Neutralisierung des Inneren erzwingt. Der Unterschied ist besonders ausgeprägt bei der in den Simulationen zuvor gewählten Porenlänge. Dies erklärt die gute Übereinstimmung der eindimensionalem PB-Modellierung mit den MD-Simulationen mit dielektrischen Kontrast.

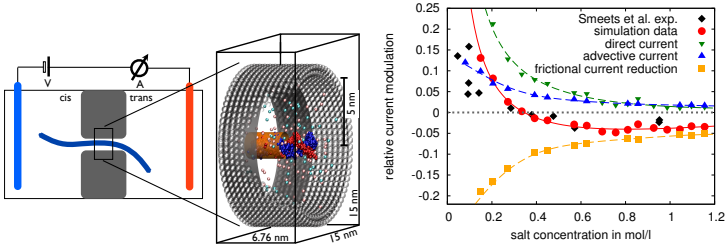
Modulation der Nanonporenleitfähigkeit durch DNA

Die wichtigste experimentell zugängliche Observable ist der gemessene Ionenstrom. Seine Modulation ermöglicht es, Rückschlüssen auf die Art des transportierten Biomoleküls zu ziehen. Interessanterweise war bisher nicht quantitativ geklärt, wie die Modulation des Stroms durch DNA zustande kommt. Experimentell wurde z. B. gefunden, dass bei hohen Salzkonzentrationen ein DNA-Molekül die Leitfähigkeit herabsetzt, bei niedrigen Salzkonzentrationen es aber die Porenleitfähigkeit erhöht [5]. Um dies genauer zu untersuchen, wurden hybride Gitter-Boltzmann/Molekulardynamik (LB/MD)-Simulationen und atomistische Molekulardynamiksimulationen durchgeführt.

Mithilfe der LB/MD-Simulationen konnten in Übereinstimmung mit der Literatur vier verschiedene Effekte identifiziert werden, die die Leitfähigkeit wesentlich beeinflussen. Zunächst hängt die Leitfähigkeit von der Anzahl an Ionen ab, die sich in der Pore befinden. Ein DNA-Molekül blockiert einen Teil des verfügbaren Volumens und verringert damit die Anzahl der den Strom tragenden Ionen. Da ein solches Molekül jedoch geladen ist, bringt es auch eine gewisse Anzahl von Gegenionen in die Pore ein, die die Leitfähigkeit wiederum erhöhen können. Darüber hinaus muss der an der DNA auftretende elektroosmotische Fluss berücksichtigt werden. Zuletzt ist es wichtig, dass Ionen in Wandnähe weniger mobil sind als in freier Lösung, was wiederum zu einer Verringerung der Leitfähigkeit führt. Das Zusammenspiel all dieser Effekte ist verantwortlich für die experimentellen Beobachtungen.

Eine quantitativ korrekte Vorhersage für all diese Größen wurde von vergröberten Simulationen nicht geliefert. Wir haben daher atomistische Simulationen eines möglichst einfachen Systems durchgeführt, das all diese Effekte abdeckt [6]. Dieses System, bestehend aus einem in einer Pore gehaltenen DNA-Fragment, wird in Abb. 1.2a dargestellt. Sowohl die Pore als auch die DNA reichen bis an den Rand des Simulationsvolumens, wo periodische Randbedingungen dafür sorgen, dass Partikel, die das System auf einer Seite verlassen, es auf der anderen Seite wieder betreten. Dies ermöglicht es, das bereits oben erwähnte Zylinder-im-Zylinder-Modell mit atomarer Auflösung zu studieren. Die Leitfähigkeit des Systems unter Einfluss eines externen elektrischen Feldes wurde für zahlreiche Salzkonzentrationen untersucht.

In Abb. 1.2b zeigen wir die Ergebnisse und den Vergleich mit experimentellen Daten. Unsere Simulationen reproduzieren die experimentellen Beobachtungen hervorragend. Ein Vergleich mit einem einfachen Kontinuumsmodell zeigt: die Oberflächenreibung der Ionen mit der DNA ist verantwortlich für die Stromblockaden. Nur in einem Modell, in dem diese quantitativ korrekt



(a) (b)

Abbildung 1.2: (a) Skizze der atomistischen Simulationen. Ein quaderförmiges Simulationsvolumen von 6.76 nm Länge enthält ein Fragment doppelsträngiger DNA, sowie eine zylindrische Porenwand mit 5 nm Radius. Zusätzlich enthält die Pore Wassermoleküle sowie Ionen. Der durch Anlegen eines elektrischen Feldes hervorgerufene Strom wird gemessen. (b) Vergleich der Stromänderung zwischen dem Fall mit und ohne DNA zu experimentellen Daten [5] bei verschiedenen Saltzkonzentrationen. Oberhalb von 350 mM Salzkonzentration verringert die Anwesenheit der DNA in der Pore die Leitfähigkeit, darunter ist sie durch die DNA erhöht. Die Erhöhung ist darauf zurückzuführen, dass die DNA zusätzliche Gegenionen in die Pore bringt, die zum Stromleitungsprozess beitragen können. Für die Verringerung essentiell ist neben dem von der DNA ausgefüllten Volumen der in dieser Arbeit zum ersten Mal quantitativ bestimmte Reibungseffekt zwischen Ionen und DNA.

wiedergegeben werden, kann auch quantitative Übereinstimmung mit dem Experiment erreicht werden.

Schlussfolgerungen

Aus experimenteller Perspektive lässt sich vor allem eine Schlussfolgerung ziehen: Werden wider Erwarten keine Translokationsvorgänge beobachtet, so sollte der Salzgehalt der Lösung erhöht werden. Damit lassen sich elektrostatische Barrieren verringern. Dies ist vermutlich einer der Gründe, warum experimentell typischerweise hohe Salzkonzentrationen vorgezogen werden. Aus unseren Untersuchung zur Porenleitfähigkeit ergibt sich, dass der Ladung der DNA eine sehr wichtige Rolle zukommt, da sie sowohl die Anzahl der zur Verfügung stehenden Ladungsträger, als auch die Stärke des elektroosmotischer Flusses

festlegt. Bei Biomolekülen, deren Ladung sich z. B. durch Änderung des pH-Wertes verändern lässt, kann daher eine andere Einstellung der Ladung die Chance einer eindeutigen Detektion verbessern. Auch die beobachtete Rate von Translokationsvorgängen kann auf diese Weise beeinflusst werden.

Auf theoretischer Ebene sind vielfältige Schlüsse möglich. Es wurde festgestellt, dass die Modellierung doppelsträngiger DNA als geladener Zylinder eine vernünftige Näherung ist, solange nur Regionen weiter als 1.5 nm von der Achse entfernt, betrachtet werden. Die Ionenverteilung ist realistisch und dies rechtfertigt die zur Bestimmung der Translokationsbarrieren genutzten Modelle. Hierbei ist die partikelbasierte Darstellung in der Anwesenheit von einwertigen Gegenionen aus unserer Sicht der PB-Modellierung nicht überlegen. Die Unterschiede in der Vorhersage beider Methoden ist vermutlich in vielen Fällen deutlich geringer als die Ungewissheit genauer experimenteller Parameter. Dabei ist die Kontinuumsmethode basierend auf der Lösung der PB-Gleichung deutlich effizienter. Die vergröberte molekulardynamische Simulation kann ihre Stärken bei flexiblen Molekülen wie die von uns untersuchte einzelsträngige DNA ausspielen, da hier die Entropie des Kettenmoleküls explizit betrachtet werden muss. Dies ist in einem Kontinuumsansatz nur schwer möglich.

Diese Näherung der doppelsträngigen DNA als geladener Zylinder ist auch außerhalb des thermodynamischen Gleichgewichts vernünftig, wenn etwa die Stärke des elektroosmotischen Flusses betrachtet werden soll. Zu unserer Überraschung ist die Vorhersage der Stärke des elektroosmotischen Flusses gut, und nur wenn quantitativ besonders präzise Aussagen notwendig sind, ist eine atomistische Simulation notwendig. Es wurde gezeigt, dass in Form der hybriden LB/MD-Methode eine verlässliches quantitatives Werkzeug zur Verfügung steht. Hier gilt allerdings wiederum, dass diese Methode ihr volles Potential besonders dann entwickeln kann, wenn flexible Strukturen untersucht werden sollen. Die vergröberten Modelle lassen jedoch keinen genauen Schluss auf die Leitfähigkeitsmodulation zu. Hierfür ist die genaue Mikrostruktur der DNA, einschließlich der betrachteten Basen und die genaue Wechselwirkung mit dem Ionen von großer Wichtigkeit. Der beobachtete Oberflächenreibungseffekt erscheint von sehr grundlegender Natur zu sein, und eine genauere Untersuchung wäre daher wünschenswert.

Diese Arbeit zeigt insgesamt, dass trotz der enormen Zunahme der Leistungsfähigkeit von Computern nicht alle Probleme mit atomistischen Simulationen bewältigt werden können. Die Größe notwendiger Systeme und die große Zeitdauer der zu untersuchenden Prozesse macht es notwendig, eine Vergröberung durchzuführen, die alle für die Fragestellung relevanten Aspekte korrekt reproduziert. Die Frage welche Aspekte relevant sind, muss jedoch im Einzelfall

aufgrund physikalischer Einsicht und Vergleich mit der feineren Skala entschieden werden. Darüber hinaus liefern die dabei entwickelten vergröberten Beschreibungen das Vokabular, mit dem die physikalische Einsicht formuliert wird. Insofern sind sie für die Wissenschaft eine essentielle Grundlage.

2 Introduction

Since 20 years nanopores offer scientists interesting insights into the world of single biomolecules. Conceptually, a nanopore is just a very small hole with a diameter of only a few nanometers in a thin membrane. The fascinating feature of nanopores is that they can be used to probe single molecules with an astonishingly simple measurement principle, the so-called the *resistive pulse technique* [7]. When a nanopore is immersed in an electrolyte and a voltage is applied, the dissolved ions create an electric current through the pore. The extension of the pore is so small, that a single biomolecule inside the pore is enough to produce a measurable perturbation of the current. The shape and magnitude of the pulses reveal properties of the molecule traversing the pore and can be used to probe single molecules. In 2000, Meller et al. showed that DNA with different sequences can be discriminated by their resistive pulses [8]. After this seminal finding the idea of using nanopores as read-out sensors for DNA was born. Twelve years later, in early 2012, the company Oxford Nanopore Technologies indeed announced two nanopore-based sensors for DNA sequencing applications [2]. Their sensors were promised to be commercially available within 2012, but so far (June 2014) Oxford Nanopore's sensors are not yet being sold, and the company has remained silent about the sensors' future [9]. Even though not all scientists in the field share the optimism that nanopores are a promising technology for DNA sequencing, it is undisputed that nanopores make versatile sensors. This and the sheer amount of different proposed applications make it a safe bet to say that we will see commercially applied nanopore technology in the future.

This thesis discusses two fundamental questions regarding DNA transport through nanopores. First, we investigate the rate with which molecules enter a nanopore, and what physical mechanisms hinder their passage. Second, we study the electric conduction of a nanopore, and how the conduction is altered by the presence of a macromolecule. In both cases, we encounter a complex interplay of several different physical effects. By using models with different degrees of detail and realism we can disentangle the contributing factors and their role for macromolecular transport and conduction. Our insights are a further step towards a better understanding of the details of macromolecular

transport through nanopores, and therefore also towards technological applications. In the following, we will outline briefly the questions we investigated and the methods we applied.

Already in one of the first publications regarding nanopores by Sung and Park [10], an entropic barrier for polymers crossing a membrane through a small hole was postulated, as the polymer has to assume a configuration compatible with the shape of the membrane. Works of Zhang and Shklovskii [11] and Kumar and Muthukumar [12] demonstrated that electrostatic interactions also create free energy barriers, but so far they had not been investigated systematically. Especially the role of the difference in dielectric permittivity between water and the membrane material had not been investigated. With the ICC$\star$ algorithm (induced charge computation, the star can stand for various Coulomb solvers [3]) a new tool to take this effect into account in molecular dynamics (MD) simulations has become available. In a past investigation, a significant free energy barrier in the absence of salt ions, hence at very low electrolyte concentrations, had been found [13]. We investigated the free energy barrier of short, stiff segments of double-stranded DNA (dsDNA) in electrolyte solutions of different concentration. Here the cases of neglecting and taking into account the dielectric contrast were compared. We then investigated the free energy barrier for a flexible chain, a model of single-stranded DNA (ssDNA). A very important step making this investigation possible was the formulation of a reaction coordinate that is capable of describing the translocation progress all the way from approach to exit.

Furthermore, the free energy barriers for the dsDNA segment were investigated by means of the Poisson-Boltzmann (PB) theory. Initially, this was performed without high ambitions: the simple numerical solver of the PB equation in one dimension written for the second major part was reused and applied to calculate the work necessary to embed an infinite charged cylinder into a cylindrical cavity. Interestingly, the results agreed well with the simulations in which the dielectric contrast had been taken into account, even though the model did not contain any notion of dielectric permittivity. We could explain this observation by the insight that both models, with and without dielectric contrast, converge to the same free energy barrier in the limiting case of long pores. The rate of convergence, however, is faster in the presence of dielectric contrast, and therefore the agreement was better for pores of finite length in that case. This convergence was explicitly shown using the IPBS algorithm proposed by S. Tyagi, which was implemented in a Diploma thesis to which the author was an advisor [14]. The algorithm is a numerical method to solve the PB equation in arbitrary geometries with different regions of arbitrary

dielectric constant and could be directly applied to investigate the pores of finite length as we did before with molecular dynamics simulations.

In the second part presented in Chapter 5, a very fundamental feature regarding the resistive pulse technique is addressed, the sign and magnitude of the pulses. We observed that the simple one-dimension model based on the electrokinetic equations suggested in Ref. [15] is not capable of predicting the DNA-induced reduction of conductance in the nanopore correctly. Instead, the model predicts a conductivity increase, up to very high electrolyte concentrations, a feature that in experiments is only observed at low concentrations. To adress this issue, we developed two different models in analogy to the model of van Dorp et al. [15], but each based on different simulation techniques. We used a hybrid molecular dynamics / Lattice Boltzmann (MD/LB) simulation method as well as all-atom simulations (AA). We did so using geometries reflecting the original model as closely as possible. The deviations between the prediction of the MD/LB simulations and the electrokinetic theory could be clarified in terms of well-known physical effects, but the expected conductivity reduction was only found for very high electrolyte concentrations. The all-atom model, however, did reproduce the current reduction in agreement with the experiments. This turned out to be due to the fact that ions near the DNA are less mobile as they experience a surface friction effect. This is not taken into account in the other models, and hence they fail in predicting the reduction.

Throughout this thesis, different physical effects are disentangled by using *coarse-grained* descriptions. This means, it is attempted to find models in which many degrees of freedom, e.g. the positions of all atoms in a system, are replace by fewer degrees of freedom. Our aim of coarse-graining is not solely simplifying the considered system for computational convenience. Instead, we use the simplified models as a tool for analysis. If it is possible to construct a set of models, in which certain effects can simply be switched on and off, this is a very powerful tool to gain physical insight. Then, the importance of individual effects can unambiguously assessed by comparing predictions with and without this effect. It is important to note, that coarse-grained descriptions are not only tool for investigation, but also provide the concepts used to formulate qualitative findings. Take, for example, one of the key statements of Chapter 4: "Charged objects in water are repelled from interfaces with most other media". The notions of object, media an interface are coarse representations of many atomic degrees of freedom.

Unfortunately, in modelling, this ideal situation is rare. Physical models are not Lego bricks, that can be simply be put together into very complex struc-

tures. Constructing consistent and meaningful physical models is a delicate task. Comparing the outcome of different models requires a fair amount of knowledge about the expected agreement with experiments and other simulation results, and where differences can be anticipated. For this reason it is important to know along which lines these models and simulation methods have been constructed. We therefore give a relatively detailed account of theories and computational methods in Chapter 3. Knowing the guiding principles is an important prerequisite for understanding what degree of accuracy can be expected.

This thesis is organised as follows. After a short summary in German, we describe the general strategy and setting of the thesis in this introduction. In the third chapter, important background information about DNA, nanopores and our modelling strategy. In particular, we give condensed information about the structure of DNA, characteristics of nanopores and an overview over experiments performed with both (Sec. 3.1), the statistical mechanics necessary for understanding of the applied models and algorithms (Sec. 3.2), molecular dynamics simulations (Sec. 3.3), and continuum models in context of charged soft matter systems (Sec. 3.4). In the following two extensive chapters the result regarding the two main research directions are presented. Chapter 4 is devoted to the role of electrostatic interactions for barriers opposing the translocation of DNA molecules through nanopores with special emphasis on the role of dielectric permittivity. It consists of an introduction to the theory of dielectric media and computational algorithms to take interfaces into account (Sec. 4.1), two example systems for which we study the role of dielectric interfaces (Sec. 4.2), molecular dynamics simulations regarding the height of electrostatic barriers (Sec. 4.3), and finally two Poisson-Boltzmann equation based models of DNA translocation barriers (Sec. 4.4). In Chapter 5 the modulation of the resistance of a nanopore by the presence of a DNA is investigated where we make use of three different models. In Sec. 5.1 we discuss general aspects of the electrical conductance of a nanopore, for which we present the prediction of a simple continuum model. In Sec. 5.2 we apply a mesoscopic molecular dynamics simulation method and in Sec. 5.3 an all-atom simulation is used to address the same question. In each step the results are compared, leading to a clear picture of the accuracy of each method. Both the findings of Chapters 4 and 5 are summarized at the end of each. We conclude in Chapter 6 with a general summary, a discussion of the applied models and suggest future research.

Although this thesis has only one author, the actual research was performed in a team. Therefore, I use “we” as the pronoun of the first person voice use throughout the scientific discussion. My starting point was the work of

M. Süzen. In the time during his PhD, the question of electrostatic barriers and the role of dielectric interfaces had already been raised. I benefitted strongly from his preliminary studies, as well as from a working and tested implementation of the ICC$\star$ algorithm in ESPResSo. With these important steps taken, it was possible to prepare the manuscripts of two articles within the first year of my PhD [16, 17], of which one was published as a full peer-reviewed paper [18]. Additionally, I advised four theses: Preliminary work on the Lattice-Boltzmann method was performed by G. Rempfer during his B.Sc. thesis [19], and F. Zeller [20] performed simulations on similar systems. A. Schlaich implemented the iPBS algorithm applied in Sec. 4.4 [14]. W. Müller prepared the simulation methods used in Sec. 5.3 and obtained preliminary results while preparing his Diploma thesis [21]. I enjoyed working with them an I am grateful for their contributions to the work presented here.

Published Work

Most of the work presented in this thesis is published. In total 5 peer-reviewed journal articles and 2 book chapters were produced.

Journal articles:

- S. Kesselheim, M. Sega, C. Holm, *Applying ICC* to DNA translocation: Effect of dielectric boundaries*, Comp. Phys. Comm., **182**, 33 (2011).
- S. Kesselheim, M. Sega, C. Holm, *Effects of dielectric mismatch and chain flexibility on the translocation barriers of charged macromolecules through solid state nanopores*, Soft Matter, **8**, 36 (2012).
- A. Arnold, K. Breitsprecher, F. Fahrenberger, S. Kesselheim, O. Lenz, C. Holm, *Efficient algorithms for electrostatic interactions including dielectric contrasts*, Entropy **15**, 11(2013).
- S. Kesselheim, W. Müller, C. Holm, *Origin of current blockades in nanopore translocation experiments*, Phys. Rev. Lett. **122**, 018101 (2014).
- T. Ertl, M. Krone, K. Scharnowski, S. Kesselheim, C. Holm, *Visual analysis for space-time aggregation of biomolecular simulations*, Faraday Discussions, accepted manuscript, (2014).

Peer-reviewed book chapters:

- A. Arnold, O. Lenz, S. Kesselheim, R. Weeber, F. Fahrenberger, D. Roehm, P. Kosovan, C. Holm, *ESPResSo 3.1 - Molecular Dynamics Software for Coarse-Grained Models*, in *Lecture Notes in Computational Science and Engineering 89*, M. Griebel, M.A. Schweitzer (Eds.), Springer, (2012).
- S. Kesselheim, C. Holm, *Modeling DNA in Nanopores*, in *Electrostatics of Soft and Disordered Media*, R. Podgornik, D.S. Dean, A. Naji, J. Dobnikar (Eds.), Pan Stanford (2013).

Other projects

Beyond the nanopores project I was involved in two other projects, that have lead to publications:

- **Tracer diffusion in crowded channels** In cooperation with Rajarshi Chakrabarti, a postdoc at the ICP, the diffusive transport of small tracer particles in a cylindrial pore, grafted with polymers was studied. We investigated how the transport of material along a channel is influenced by the crowding with polymers with repulsive interactions as well as with attractive interactions. I was involved in all phases of this work: I assisted in preparing simulation scripts and software infrastructure to perform parameter studies, and contributed to the publication. The results were published as
  R. Chakrabarti, S. Kesselheim, P. Kosovan, C. Holm, *Tracer diffusion in a crowded cylindrical channel*, Phys. Rev. E 87 (2013).
  This project was intended as a preliminary study towards a more realistic model for transport through the molecular pore complex, a complex structure which gates transport into and out of the cell nucleus.
- **Growth of colloidal crystals** In collaboration with A. Arnold and D. Roehm, both at the ICP, the growth of a crystal consisting of charge stabilized colloids was investigated. We applied the Lattice-Boltzmann method to show that hydrodynamic interactions lead to a significant slowdown of crystal growth. This interesting result disputes the analogy of colloidal system with smaller systems, specifically atomic or molecular systems near the freezing point. The results were published as
  D. Roehm, S. Kesselheim, A. Arnold, *Hydrodynamic interactions slow down crystallization of soft colloids*, advance article in Soft Matter (2014).

3 Experimental and theoretical framework

In the four sections of this chapter we describe the framework in which this thesis is set. The first section contains a brief overview over DNA and nanopore-based experiments. In the next section the necessary statistical mechanics background is presented. It addresses mainly the question of how physical laws from a finer model can be transferred to a coarser one, as this theme is found in many different aspects throughout this thesis. We then give an overview over different conceptual and algorithmic aspects of molecular dynamics (MD) simulations. Here, the focus lies on coarse-grained MD simulations, and especially on the software ESPResSo which is developed mainly at the ICP and was used to perform all coarse-grained simulations that are part of this thesis. Finally, we introduce two continuum models useful for modeling charged soft matter systems, namely the Poisson-Boltzmann equation for static properties and the electrokinetic equations for dynamic properties.

3.1 DNA and nanopores

Sensing single DNA molecules has arguably been the most important application for nanopores. The thrive for a next-generation DNA sequencing method is highly stimulating for the nanopore field in science. In this thesis we investigate two fundamental questions regarding how DNA can be sensed. We start off by giving background information about DNA molecules and a brief account of state of the art in nanopore experiments.

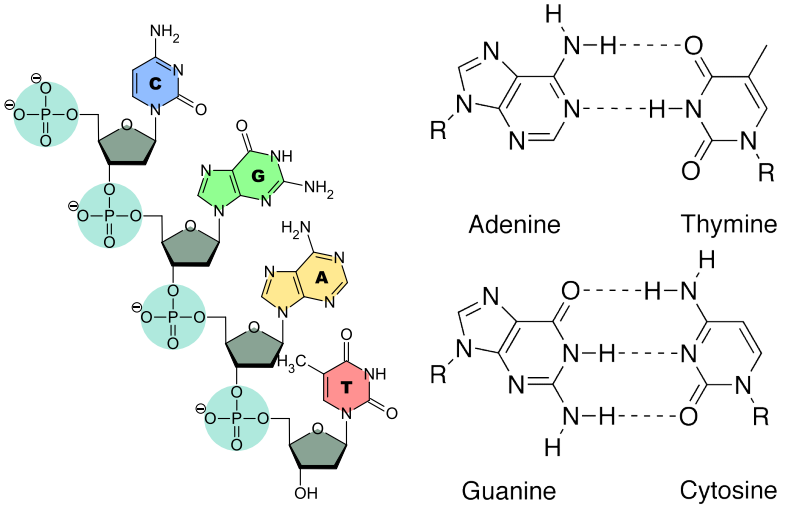
3.1.1 DNA: The molecule of life

The genetic information of a living organism is encoded in only a few molecules: Each of the 23 human chromosomes contains a single deoxyribonucleic acid (DNA) molecule carrying the genome. DNA is a very long

chain molecule. In a human cell its length is in total about 1.5 m [22]. The building blocks were identified already in 1919 by Levene [23]. Each of the four *nucleobases*, or short *bases*, Adenine, Cytosine, Guanine and Thymine is connected to a sugar group, which in turn is bonded to a phosphate group. Together they form a *nucleotide*. The chemical structures are shown in Fig. 3.1. Many phosphate groups are covalently bonded when forming the backbone of the DNA. The phosphate group is a strong acid, which gives rise to the name of the structures formed by nucleotides: *nucleic acid* [24]. The sequence of the bases in the DNA strand encodes the genetic information. Determining the DNA sequence is a challenging, costly and time–consuming process. Impressive progress has been made in the last decade [25, 26], bringing down the price to currently $5000 for sequencing the genome of a human [9]. Even easier access to genetic sequences would open the door for personlized medicine, in which treatments can be adjusted to individual genetic profiles. Nanopore-based technology is considered a promising candidate for a technology that allows for significantly cheaper DNA sequencing.

James D. Watson and Francis H. Crick discovered a very important aspect of the DNA structure, the base pairing mechanism depicted in 3.1b. This mechanism was identified in 1953 [27] based on X-ray diffraction images taken by Rosalind Franklin et al. [28]. The nucleobases Adenine and Thymine can form stable pairs by forming two hydrogen bonds at the same time. Cytosine and Guanine can form three hydrogen bonds and accordingly, their paired configuration is slightly more stable [29]. All other possible pairings are significantly less stable. Due to this specific pairing mechanism in living cells, two strands of DNA are wound around a common axis, forming a double helix. This famous structure is shown in Fig. 3.2, as a wire model by Watson and Crick [28], and rendered snapshots from the simulation presented in chapter 5. The two strands are complementary, thus every nucleotide in one strand is opposed by the matching partner on the other strand. In the course of cell replication the two strands are separated and new complementary strands are constructed base by base using each of the strands as blueprints. In vitro the two strands can be separated by elevating the temperature, the DNA *melts* [30]. In many nanopore experiments the resulting single-stranded DNA (ssDNA) is used, and not the double stranded DNA (dsDNA), in which the genetic information is typically stored.

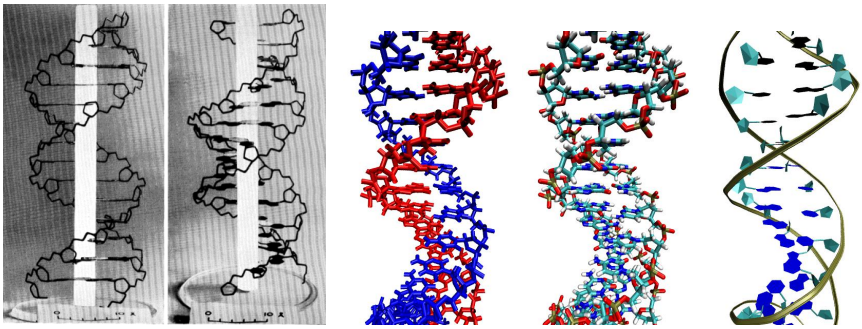
The acidity of the backbone is directly responsible for the DNA's charge. In aqueous solutions with moderate pH the acidic groups are fully dissociated and every phosphate group carries one elementary charge. In this thesis effects related to its charge are the main features under investigation.



(a) (b)

Figure 3.1: (a) The primary structure of a single DNA strand. Each of the four nucleobases (blue, green, yellow and red) is covalently bond to a dexyribose–phospate group. (b) The specific hydrogen bonds between different nucleobases. Within an Adenine-Thymine (AT) base pair two hydrogen bonds form. In a Cytosine-Guanine (CG) three hydrogen bonds are present (CG). Illustrations from the Wikimedia Commons (`DNA-Nucleobases.svg`, `AT_base_pair_jypx3.png`, `GC_base_pair_jypx3.png`).

The mechanism with which living cells replicate the genetic material to prepare for cell division can be imitated in vitro. The enzyme that constructs complementary strands, the *DNA polymerase*, also works in the test tube. In order to achieve rapid replication special temperature-resistant variants extracted from bacteria living in high temperature environments are necessary [32, 33]. By sequentially heating and cooling a solution containing a DNA sample, DNA polymerase and a supply of nucleotides, a single DNA strand can be copied [34]. Many cycles of this procedure are executed, so that the number of DNA molecules is doubled in every step, a process called *polymerase chain reaction* (PCR). This mechanism is an important workhorse in many medical and biological labs. Mullis was awarded the Nobel Prize in Chemistry in 1993 for this discovery [35].

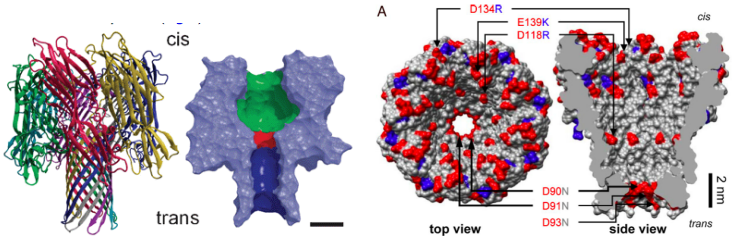


(a) (b)

Figure 3.2: The helical structure of DNA. (a) A wire model published by Watson and Crick in 1954 [28]. (b) The DNA molecule simulated in section 5.3 rendered using three different rendering methods in VMD [31].

In 1977 Sanger et al. published the first genetic sequence: the sequence of the bacterium Φ X174 [36]. Their sequencing method, now termed the Sanger method [37], is based on the PCR. In a slightly more modern version it works as follows [33, 38, 39]. First, a DNA sample is enhanced by PCR and melted. Then four DNA replication experiments are performed, each with a different additive. In each of the experiments one of the nucelotides is added also in a modified version. This modified nucleotide, the dideoxynucleotide is built into the chain by the polymerase like the original nucleotides but interrupts the polymerization. The construction of a complementary strand is aborted. In one of the reactions e.g. a dideoxy-adenine is added, and at every position where the complementary strand has a thymine base, there is a small probability that the chain reaction is interrupted. Then the DNA is melted and the resulting mixture of short oligomers created as complementary strands is size separated by gel electrophoresis. From the length distribution of the short strands in each of the sample the base type at each location can be deduced. This *chain termination* method can be fully automated, but still is a laborious process when applied to a full genome.

The Sanger method was the gold standard of genetic sequencing for many years, but during the last decade several next-generation methods have turned into viable alternatives [39]. The Archon X prize, however, still remains to be claimed: $ 10 Mio. are awarded to the first team that can sequence “100 human genomes in 30 days or less, for no more than $10,000 USD per genome sequenced” [40]. Nanopores might be the technology that can achieve this



(a) (b)

Figure 3.3: Pore proteins used as nanopores. (a) The α-Hemolysine protein. Left: the seven identical units are colored differently. Right: colors indicate the three different regions of the pore. (from [43]) (b) The MspA pore protein. The amino acids replaced in by genetic engineering are marked read and blue (from [44]).

goal, and in text books they already appear as candidates for next-generation sequencing [39].

3.1.2 Nanopore translocation experiments

The first successful nanopore experiments used single biological pore proteins embedded in a lipid bilayer membrane [41]. Membrane proteins called α-Hemolysine, extracted from the bacterium *staphylococcus aureus*, were embedded into a lipid membrane separating two water compartments. Fig. 3.3a depicts this setup. This protein is hollow, so applying a voltage between the two compartments results in a small ion current through the pore. Adding ssDNA molecules to the chamber containing the negative electrode alters the current. Short events appear, in which the current is largely blocked. These events are interpreted as individual DNA molecules passing the pore in the lipid membrane, a process called *translocation*. A few years later, Meller et al. could discriminate between a homopolymer consisting only of cytosines and a homopolymer of adenines from the duration of the translocation events and the current blockade level [8]. Since then an impressive number of similar experiments have been conducted. The Nanopore Site [42] currently lists more than 60 groups working on nanopore experiments.

Experiments with biological nanopores have significantly broadened in tech-

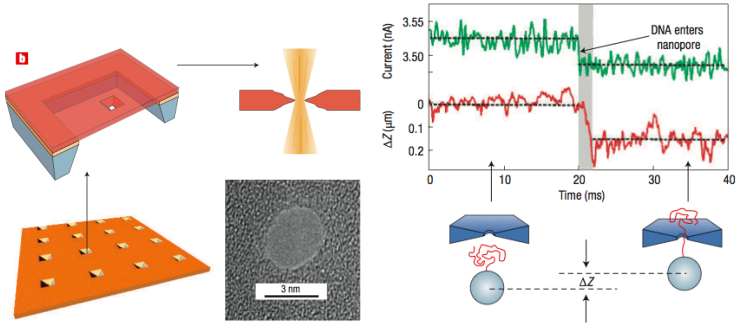


Figure 3.4: (a) Fabrication of a solid-state nanopore. In several surface deposition and etching steps, a thin free-standing membrane is created. By means of an ion or electron beam a thin hole is milled into the membrane (from [50]). (b) Experiments in Ref. [51]. A DNA molecule is attached to a colloid held in a laser tweezer in front of a nanopore. When a voltage is applied, it is dragged into the nanopore. This can be deduced consistently from the change in current (green line) and the colloid position (red line).

niques and scope. Relatively recently, the MspA protein extracted from a myobacterium [44, 45] shown in Fig.3.3b has been proposed as an alternative to the α-Hemolysine pore protein. In a genetic engineering process negative charged residues were removed from this protein, and DNA translocation could be observed. Oxford Nanopore Technology's sensor is based on this protein. By genetic engineering also the properties of the α-Hemolysine pore can be altered. In 2001 Howorka et al. attached a short single DNA strand inside the pore and were able to detect specific DNA sequences [46]. By replacing certain amino acids other interesting features were realized, such as a binding site for metal ions [47] or a receptor for the explosive TNT [48]. More very fascinating sensing applications can be found in the reviews by Venkatesan and Bashir [49] as well as Howorka and Siwy [43].

Only a few years after the first experiments with biological pores, conceptually identical experiments were performed with synthetic pores. The process to create them is sketched in Fig. 3.4a. Originally, a group in Harvard applied a technique called *ion beam sculpting* to create a thin hole in a free-standing silicon nitride membrane [52]. Using an ion beam, a pore was drilled into this

membrane and by varying the intensity and energy of the ion beam the pore could be reduced in size. Many different groups developed similar nanofabrication techniques to create single pores in different silicon–based substrates (e.g. [53, 54]). This class of pores is called *solid–state* nanopores. A good overview can be obtained from the reviews of Dekker [50] and also Howorka and Siwy [43], and the very recent review of Miles et al. [55]. As the substrates are solid, these nanopores are easier to handle than biological nanopores. By means of lithographic techniques they can eventually be produced in large arrays, which makes them very attractive for high-throughput sensing applications.

The surface of solid-state nanopores can be modified by different techniques. With atomic layer deposition, thin layers of e.g. alumina [56,57] or gold [58] can be used to fine-tune the size and surface properties of solid-state nanopores. Alternatively, different polymer coatings can be used. By silanization, different coatings can be bound covalently to the nanopore surface, which allows e.g. for the construction of pH-sensitive conduction characteristics [59]. Gold-coated surface can be functionalized with self-assembled monolayers [60], a process with which a broad range of materials can be bound to the nanopore. Kowalczyk et al. covered a solid-state nanopore with nucleoporins, reminiscent of the biological nuclear pore complex [61]. A relatively recent trend is the use of graphene or boron nitride as membrane materials. Both can be produced as very thin material, down to single-atomic layers. These materials are used to cover solid state nanopores and the beam of a transmission electron microscope is used to stamp a hole into these extremely thin membranes [62–64]. In these setup, the in-plane electric conductivity of the graphene can be used to infer information about molecules in its vicinity [65]. Similar effects can also be achieved with nanowires placed across nanopores [66]. In all these cases, the field effect due to the DNA charge is used as sensing mechanism. Other groups were able to include gold electrodes over the pores to enable the detection of molecules by means of the tunneling current [67, 68]. This tunnel current also contains valuable information about the translocating molecule and can be used to identify individual molecules.

An entirely different, fascinatingly simple nanopore fabrication technique was developed in the lab of U. Keyser. In a so-called pipette-puller, glass micropipettes are heated with a laser and pulled apart [69]. Using the right protocol, opening diameters down to around 15 nm can be achieved. As a micropipette is a product of mass-fabrication and very cheap, this technique is especially well suited for commercial sensing applications. By using transmission electron microscope beam, the opening can even be shrunk to only a few nanometers [70]. The transparency of the nanocapillary also allows for

direct imaging of molecules inside of the capillary. With this technique, the translocation of fluorescently labelled dsDNA could be observed in a light microscope [71].

A novel, very interesting technique relies on track-etching. In this case a polymeric membrane material such as polyethylene terephthalate (PET) is targeted by a heavy ion beam from a suitable source. After this irradiation, an etching agent is added to one side of the membrane which dissolves the polymer partly. As the ion beam has altered the structure of the membrane material, the etching occurs mostly along the track the ions took through the material. The other compartment is filled with a neutralization agent, which stops the etching process once the two compartments are connected. This process results in highly asymmetric nanopores with a conical shape and a small opening down to a few nanometers. The track-etching technique creates a highly charged surfaces with deprotonated carboxyl groups at a density around 1 e/nm^2 [72]. These high surface charges, in conjunction with the asymmetric shape of the nanopore can lead to interesting nonlinear electrokinetic effects such as current rectifications similar to the effects in a semiconductor diode [73, 74]. Also very interesting relationships between an applied electric field and the wetting properties of the surface have been observed [72].

Both the glass nanocapillary and the track-etched PET sensors are probably too coarse for DNA sequencing. Their strength lies in the robustness, simplicity and tunability, and this makes them interesting candidates for more qualitative sensing applications. They bridge the length scale towards micro- and nanofluidics. In micro- and nanofluidics, the idea of electric circuits on silicon wafers is transferred to applications with liquids. Very small amounts of liquids are controlled very precisely to be able to miniaturize analytics as far as possible. Very complex processes are performed in microfluidic environments, where the ultimate goal is integrating a whole a lab to a single chip. A very interesting overview over this fast-growing field can be found in the reviews [75–77]. Nanopores can fill an important role in these settings due to their capabilities as sensors for single biomolecules.

While the detection of the ion currents might already be sufficient for technological applications, other observables have proved helpful for scientific insight. An interesting experiments conducted in the lab of C. Dekker is shown in Fig. 3.4b. A double-stranded DNA molecule is attached to a colloidal particle fixed by a laser tweezer. The colloid can be positioned in front of a solid-state nanopore, an by switching the voltage on, the DNA can reversibly be inserted into the nanopore [51]. This setup allows to study the same DNA molecule during many insertion processes and at the same time measure the

force acting on the DNA molecule. The same technique, which is dubbed *single-molecule force spectroscopy*, was also realized with a glass micropipette nanopore [78].

Recently a very innovative nanopore approach was presented, termed DNA origami nanopores. DNA origami [79] is a technique in which three-dimensional structures are created by self-assembly of DNA. The structures are held together by *Holliday junctions*, where two adjacent dsDNAs are interconnected because one strand of both cross each other and pair with the opposite strand. Hollow structures from DNA have been designed and inserted either into lipid bilayer membranes [80], solid-state nanopores [81] or into glass micropipette nanopores [82]. With this technique, the atom-precise control over the material known from biological nanopores can be combines with the control and tunability of synthetic nanopores. The coarse-grained simulation methods developed in this thesis might facilitate fundamental insights into how the DNA origami technique can be refined.

Despite their complexity, we reduce nanopores to structureless holes in walls, throughout this work. This assumption is sufficient to answer our basic, yet fundamental questions about DNA and ion transport.

3.2 Statistical physics in the soft matter regime

DNA and nanopores are set in a field known as soft condensed matter. On the map of science, this field is located at the interface between biology, chemistry, and physics. What is the unifying feature of soft matter? Daoud and Williams state in the introduction of their book that a good example soft matter is "a molecular system giving a strong response to a very weak command signal" [83]. What is a weak and what is a strong command signal is, however, not clarified.

In the American community a different term is more common: complex fluids [84]. Complexity here means that the spectrum of properties, effects, and structures is vast and that properties can change upon small applied forces. The term *fluid* refers to the fact that a material does not sustain stress, but responds by deformation. This mechanical notion is clearly also the origin of *soft* in the term soft matter [85]. The fluidity also stresses the importance of the thermal energy: Liquids are liquid because the intermolecular interaction energies are comparable to the thermal energy [86], and the entropy gain associated with the positional and orientational disorder dominates over the energy loss of a phase change to a state with long-range order. This allows us to identify the energy scale on which stimuli are sufficient to induce a response in a soft matter system: the thermal energy $k_B T$.

We here give a few typical soft matter system, each with a characteristic trait that highlights in what sense they are soft.

- Liquid crystals: A material made out of highly anisotropic molecules or particles that can exhibit crystal order in one spatial direction and liquid–like structure in another direction.

- Colloids: small particles or droplets immersed in a solvent. They can form, depending on their mutual interactions, solid, liquid and gas phase and these phase transitions are sometimes investigated as models for phase transitions in other systems.

- Polymers: Long chain molecules, whose behaviour is largely driven by configurational entropy. We give a short introduction in Sec. 3.2.7.

- Gels: crosslinked networks of polymers that take up a lots of liquid. In a good solvent they can swell to many times their original size.

- Room temperature ionic liquids: a substance of large ions that is liquid at room temperature. They are promising alternative solvents in chemical engineering, as their vapour pressure is negligible.
- Magnetic fluids and gels: magnetic particles immersed in a liquid phase, sometimes connected with polymers. They can be manipulated and controlled by external magnetic fields.
- Most biological matter: Many biological structures from cell walls to proteins are manipulated by small amounts of energy in living organisms, and structural and functional changes can be induced with e.g. a small change in temperature.

All these systems share the property that they are soft with respect to deformation and that small temperature changes can induce large macroscopic changes, sometimes phase transitions. Typically they are investigated under room temperature conditions and ambient pressure.

The importance of the thermal energy implies two things: First, these systems are governed by a balance of energy and entropy. Small changes of the temperature can induce significant structural changes. Second, understanding intrinsic thermal fluctuations is of crucial importance. Fluctuations dictate the dynamics of all processes. Consequently, understanding soft matter systems typically requires considering not only the scale of single atoms, and the continuum scale of e.g. classical thermodynamics, but at least one length scale in between, a domain sometimes called the mesoscale. For DNA translocation experiments this mesoscale is the scale of DNA molecules: much larger than atoms, but still small enough that thermal fluctuations remain important. The fascinating property of nanopores is that in these experiments the mesoscale is directly accessible.

3.2.1 The multiscale view

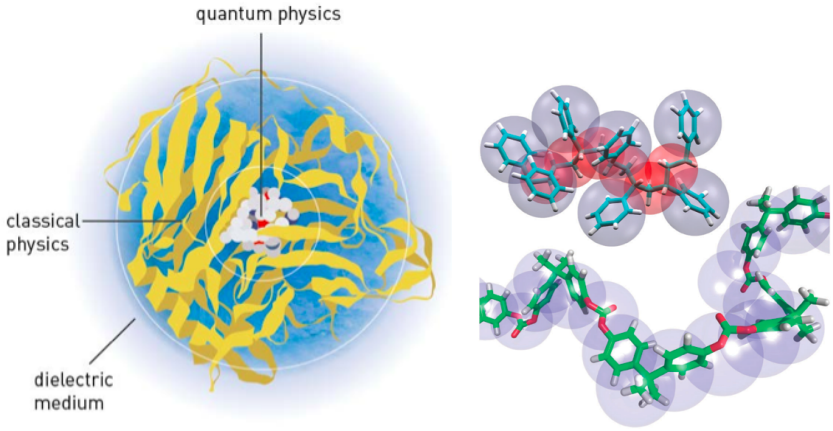
The 2013 Nobel Prize in Chemistry was awarded to Karplus, Warshel and Lewitt "for the development of multiscale models for complex chemical systems". In the announcement three very influential works by these scientists were mentioned [87]. In 1972, Karplus and Warshel studied electronic properties of a planar molecule which was partly described by quantum mechanics and partly by classical mechanics [88]. In 1976, Warshel and Lewitt studied the catalytic activity of a protein where the active part was again modelled by quantum mechanics, and the rest of the protein was described by classical

mechanics [89]. Additionally, in 1975, Warshel and Lewitt suggested a simulation method, where a protein was represented only by two "superatoms" instead of many individual atoms [90].

In the first two works physical theories of two sorts are combined into a single simulation. An example of a modern quantum mechanics/molecular mechanics (QM/MM) simulation is shown in fig 3.5a. Here an enzyme and the surrounding water is modelled by means of classical mechanics, and the active center of the enzyme is modelled by means of quantum mechanics. The complexity in these methods lies in providing a physically justified, consistent and numerically stable coupling scheme between the two physical theories. The suggestion and development of this class of QM/MM approaches was a breakthrough, that enabled scientists to study quantum effects, e.g. chemical reactions, in systems which are too large for a purely quantum-mechanical model. Today, some forty years later, the available amount of computer power has increased tremendously, but is still far from being abundant. Therefore QM/MM approaches are still under active development, and applied to a wide variety of systems [91, 92].

The strategy of the third work mentioned in the Nobel Prize announcement is the formulation of a method with fewer degrees of freedom than the original model, where the physical laws follow also classical mechanics. Here the conceptual difficulty lies in constructing a coarser model, in which important properties of the fine model are represented correctly. The loss of information should be minimal. Two approaches are possible: In systematic approaches both models are considered explicitly, and strategies and algorithms are devised to construct a coarse model that reproduces certain target properties as well as possible. Alternatively, in an axiomatic approach, coarse models are assumed based on fundamental physical laws, combined with scientific intuition. The connection to a microscopic model is not made explicitly. The emergent properties typically resemble the original system qualitatively, although sometimes the coarse-grained model can be constructed in such a way, that certain properties are exactly reproduced.

Constructing coarse-grained models in a systematic and consistent way is a very active field of research [86, 95]. The systematic approach is sketched in Fig. 3.5b. Several microscopic degrees of freedom, e.g. the coordinates of a group of atoms, are replaced by a single degree of freedom, the position of a *superatom*, or sometimes just a *bead*. Besides the question of how to obtain a coarse-grained model, it can can never fully represent the original model in all its properties, but only a subset. Using the iterative Boltzmann inversion technique [96] it is e.g. possible to create a pair potential that reproduces a



(a) (b)

Figure 3.5: (a) Conceptual sketch of a QM/MM simulation. A region modelled by classical mechanics, in this case an enzyme, hosts a small region treated by means of quantum mechanics (from [93]). (b) Mapping scheme in the systematic coarse-graining of polystyere (top) and polycarbonate (bottom) (from [94]). In both schemes a set of atoms is represented as a single sphere in a coarse-grained model.

given radial distribution function exactly. This, however, is not a sufficient condition for reproducing other structural properties correctly [97]. For example, a coarse-grained representation of water molecules can at the same time perfectly reproduce the radial distribution function while the tetrahedral packing matches poorly [98]. This clearly shows that the construction of coarse-grained models is delicate and has to be performed with lots of care.

The idea of coarse-graining can be combined with the hybrid approach used in QM/MM calculations. The vision of adaptive resolution simulations is to create a single model where the resolution can locally be adjusted as necessary [99]. A fully consistent method would allow scientists to simulate very large systems where many degrees of freedom have been eliminated. Only certain interesting spots are tackled with high, typically atomistic, resolution. Interesting examples are e.g. biomolecular systems, where the functional regions are clearly located, the surrounding solvent, however, is mostly not of interest. All adaptive-resolution MD schemes are currently at a relatively early

development stage: The first approach [100] was is not even a decade old. Adaptive resolution simulations might eventually evolve into what all-atom MD simulations are now: standard tools that with relatively little training.

According to the correspondence principle, classical mechanics emerge from quantum mechanics in the limit a of vanishing Planck's constant. Hence the laws of classical mechanics are a special case of quantum mechanics, in a certain asymptotic limit. In this limit the field quantities of quantum mechanics can be replaced by discrete degrees of freedom, by classical particles. In classical continuum theories, such as continuum mechanics, the opposite approach is applied: A continuum theory is constructed from a classical particle model. In this case a high-dimensional configuration space is replaced by a small set of functions smoothly varying in space. Formulating a continuum theory of a particle system clearly is a coarse-graining procedure. Interestingly, in many cases, the emergent continuum theories are asymptotically *exact*, thus the obtained physical laws hold exactly if the considered time and length scales are large and the applied forces are small enough. The microscopic properties are propagated to the macroscale in terms of transport coefficients, parameters which describe how a system responds to certain macroscopic stimuli. These can e.g. be obtained from molecular dynamics simulations [101]. Important examples considered in this thesis are the theory of electrostatics in dielectric media, and continuum hydrodynamics, and its generalization to immersed ions and applied electric fields, called electrokinetics.

The question of chemical kinetics is: at what rate does a chemical reaction happen? To answer this question it is necessary to have a measure of how far a chemical reaction has progressed. Macroscopically the amount of the reaction product is a very good measure for that. Microscopically, on the level of single molecules, this is expressed in terms of a *reaction coordinate*. For a dimerization reaction it is clear: If the to constituents are close enough and have the right orientation, they form a complex. If they are far apart, they are not bound. In this case, the progress of the reaction can be measured with a geometric coordinate. Clearly the rate of the chemical reaction is closely related to the dynamics in this single reaction coordinate, as we will describe in more detail in Sec. 3.2.6. The choice of the reaction coordinate, or a set of coordinates is not always obvious. A good choice is crucial for the success of dynamic models. In Sec. 4.3.3 we will describe a new reaction coordinate for the translocation of flexible polymer chains. It is important to note that, in an abstract sense, a describing a process in terms of a reaction coordinate is also a coarse-graining method: the complex motion of many molecules is projected onto a single or a few axes.

Throughout this work, all of these coarse-graining strategies are going to be applied. In the following chapter, we study the free energy of a charged chain molecule when passing through a nanopore. We want to understand if the frequency of successful translocations is reduced by electrostatic effects. We construct a reaction coordinate for this process that allows us to measure barriers, which oppose the translocation process. These simulations are performed using ad-hoc constructed models of DNA molecules and pores, where a model with implicit water and explicit ions is used. The water is coarse-grained out and we rely on a dielectric continuum description of the water and the membrane material, in which the pore is located. In addition, we perform similar investigations using a full continuum description based on the Poisson-Boltzmann theory. In the next chapter, we present results where ion transport is studied with models of different resolution. To do so , we apply a continuum model, based on the electrokinetic equations, an implicit-water-explicit-ions model, where the Lattice-Boltzmann method is applied to replace the effect of water on dynamics, and a fully explicit all-atom model. The comparison of these models allows us to determine on which modelling level which effects are reproduced to what degree of accuracy.

3.2.2 Equilibrium statistical mechanics

The material appearing in the following sections is a compilation from different textbook sources. It is a collection of important theoretic concepts and an overview where these concepts appear in the field of statistical mechanics. The simulation models and the applied simulation methodology are not derived rigorously in this section, but we attempt to make them plausible, and to identify the most important approximations. The central theme of this section is the question of how to convert a description of a (classical) system with many degrees of freedom into a model with fewer degrees of freedom.

The most important sources were the books of Landau and Lifshitz [102], Balescu [103], Reichl [104], and Evans and Morris [105], which cover equilibrium and non-equilibrium statistical mechanics. Other helpful sources were the books by Risken [106], Liboff [107], and Röpke [108] and the more classical books by Kubo [109] and Zwanzig [110]. Further background knowledge was obtained from Hansen and McDonald [111] and the physical chemistry book of Atkins. [112]. Finally we mention the more simulation-oriented textbooks of Frenkel and Smit [101], and Allen and Tildesley [113], and different textbooks on the science of colloids [114–116], which also contain a very readable perspective on statistical mechanics. The information about non-equilibrium

statistical mechanics in Sec. 3.2.3 and Sec.3.2.4 are also compiled from these sources.

In the following we will assume that the system under consideration can be described as a classical system. Quantum effects are neglected. The system may be described by the Euclidian coordinates $\boldsymbol{r}_i$ of n atoms, and their velocities $\boldsymbol{v}_i$. The main goal of this section is to formulate laws for a set of "coarse–grained" coordinates and velocities $\boldsymbol{R}_i$ and $\boldsymbol{V}_i$ that reflect the original system in some respect. For brevity, we refer to the column vectors containing all positions as $\mathfrak{r}$ and $\mathfrak{R}$ respectively, and the vectors containing all velocities as $\mathfrak{v}$ and $\mathfrak{V}$. The space spanned by $\mathfrak{r}, \mathfrak{v}$ and $\mathfrak{R}$, and $\mathfrak{V}$ is denoted as phase space.

The "fine–grained" system has a total energy H, where

$$H = \sum_{i=1}^{n} \frac{m_i}{2} \boldsymbol{v}_i^2 + V\left(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n\right). \tag{3.1}$$

Here m_i is the mass of the ith particle and V is the potential energy of the system, which we will not specify further here. Restricting ourselves to Euclidian coordinates allows us to use terms from Newtonian and Hamiltonian mechanics interchangeably: the canonical momenta $\boldsymbol{p}$ are trivially related to the velocities by $\boldsymbol{p}_i = m_i \boldsymbol{v}_i$. We hence can refer to the total energy as the Hamiltonian $H\left(\mathfrak{r}, \mathfrak{v}\right)$.

The Newtonian equations of motions read as

$$\begin{aligned} \dot{\boldsymbol{r}}_i &= \boldsymbol{v}_i \\ m_i \dot{\boldsymbol{v}}_i &= -\nabla_{\boldsymbol{r}_i} V\left(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n\right). \end{aligned} \tag{3.2}$$

The dot denotes the time derivative, and $\nabla_{\boldsymbol{r}_i}$ indicates that the gradient is to be taken with respect to the spacial coordinate of the ith particle. This gradient corresponds to the negative of the force F_i acting on particle i. In the shorthand notation we write:

$$\begin{aligned} \dot{\mathfrak{r}} &= \mathfrak{v} \\ \mathfrak{m} \dot{\mathfrak{v}} &= -\nabla_{\mathfrak{r}} V\left(\mathfrak{r}\right), \end{aligned} \tag{3.3}$$

where $\mathfrak{m}$ is a diagonal tensor containing the masses of all particles.

The change of paradigm between classical mechanics and statistical mechanics is that the entity under consideration is not a single trajectory but many trajectories are investigated by means of statistics. The most general quantity

to describe these statistics of trajectories is the phase space density $f(\mathfrak{r}, \mathfrak{v}, t)$, which is the probability density of finding the particles at positions $\boldsymbol{r}_i$ with velocities $\boldsymbol{v}_i$ at time t. It is normalized to unity, hence

$$\int \mathrm{d}^{3n}\mathfrak{r}\, \mathrm{d}^{3n}\mathfrak{v}\, f(\mathfrak{r}, \mathfrak{v}, t) = 1. \tag{3.4}$$

The time evolution of f can be constructed from the time evolution of single trajectories that follow Newton's equation of motion. This leads to the Liouville equation:

$$\frac{\partial}{\partial t} f + \mathfrak{v} \cdot \nabla_{\mathfrak{r}} f - \nabla_{\mathfrak{r}} V \cdot \nabla_{\mathfrak{v}} f = 0. \tag{3.5}$$

By integrating this expression over phase space, and inserting Newton's equation of motion, it can be seen that the normalization in 3.12 remains valid for all times.

Typically one is interested in the statistical mean $\langle B \rangle$ of a quantity B, that is a function of the configuration of the system. It is obtained by integration over the phase space:

$$\langle B(t) \rangle = \int \mathrm{d}^{3n}\mathfrak{r}\, \mathrm{d}^{3n}\mathfrak{v}\, B(\mathfrak{r}, \mathfrak{v})\, f(\mathfrak{r}, \mathfrak{v}, t)\,. \tag{3.6}$$

The angular brackets $\langle \cdot \rangle$ are sometimes are left out, if it is clear that we are referring to averages. Alternatively two-time quantities can be of interest. Then a statistical statement is made about the time evolution of the system. A typical case is the correlation function of two quantities A and B, which is defined as

$$C_{AB}(t, t') = \int \mathrm{d}^{3n}\mathfrak{r}\, \mathrm{d}^{3n}\mathfrak{v} \int \mathrm{d}^{3n}\mathfrak{r}\, \mathrm{d}^{3n}\mathfrak{v}\, A(\mathfrak{r}, \mathfrak{v})\, B(\mathfrak{r}', \mathfrak{v}')\, f(\mathfrak{r}, \mathfrak{v}, t)\, f(\mathfrak{r}', \mathfrak{v}', t')\,. \tag{3.7}$$

Quantities of interest are not necessarily scalars, but can be defined as function in space. The particle number density ρ can e.g. be defined as

$$\rho(\boldsymbol{r}, t) = \int \mathrm{d}^{3n}\mathfrak{r}\, \mathrm{d}^{3n}\mathfrak{v}\, f(\mathfrak{r}, \mathfrak{v}, t) \sum_i \delta(\boldsymbol{r}_i - \boldsymbol{r})\,, \tag{3.8}$$

Here $\delta(\cdot)$ is Dirac's delta function[1]. If it is given with a vector valued-

[1] For simplicity we ignore the distribution character of the delta functions completely in this work.

argument, it is to be interpreted as a product of delta functions for all components.

The initial conditions specified for Newton's equations of motion are transferred into specifying the initial conditions of the phase space density $f(\mathfrak{r}, \mathfrak{v}, 0)$. A steady state, or stationary state, expresses that f is independent of time. Restricting all considerations to steady states is a great simplification. Despite this simplification, only few general statements about steady states are known, and formulating general theorems about these states is ongoing research (see e.g. [117]).

The concept of thermodynamic equilibrium however is much more restrictive, but yields the very powerful apparatus of equilibrium statistical mechanics. For an isolated system, the concept of thermodynamic equilibrium can easily be introduced. In an isolated system, the energy E is by definition conserved. In that case thermal equilibrium is characterized by the fact that all state points $(\mathfrak{r}, \mathfrak{v})$ with that energy have the same probability. The entropy is maximal. For an isolated system, the entropy is defined as the logarithm of the phase space volume compatible with the energy E:

$$S(E) = k_B \log \int \mathrm{d}^{3n}\mathfrak{r}\mathrm{d}^{3n}\mathfrak{v}\delta\left(H(\mathfrak{r}, \mathfrak{v}) - E\right), \tag{3.9}$$

where k_B is the Boltzmann constant. This naturally leads to the definition of the temperature and the thermodynamic beta, as the derivative of the entropy with respect to the energy:

$$\frac{1}{k_B}\frac{\partial S}{\partial E} = \beta = \frac{1}{k_B T}, \tag{3.10}$$

where T is the absolute temperature.

For a system in equilibrium that can exchange heat with a reservoir at a fixed temperature the phase space density is given by the Boltzmann distribution:

$$f(\mathfrak{r}, \mathfrak{v}) = \frac{1}{Y} e^{-\beta H(\mathfrak{r}, \mathfrak{v})}. \tag{3.11}$$

Here Y is used to obtain a proper normalization. It is defined as

$$Y = \int \mathrm{d}^{3n}\mathfrak{r}\,\mathrm{d}^{3n}\mathfrak{v}\, e^{-\beta H(\mathfrak{r}, \mathfrak{v})}, \tag{3.12}$$

and thus is identical to the partition function up to a prefactor related to the indistinguishability of the particles. For a Hamiltonian which is the sum of

the kinetic Energy E_{kin} and the potential energy V, Y can be written as a product of a velocity part X and a configurational part Z, where

$$\begin{aligned} X &= \int \mathrm{d}^{3n}\mathfrak{v} e^{-\beta E_{\text{kin}}(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)} \quad \text{and} \\ Z &= \int \mathrm{d}^{3n}\mathfrak{r} e^{-\beta V(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_n)} \end{aligned} \tag{3.13}$$

The Boltzmann distribution in Eq. 3.11 can be factorized into $n+1$ independent factors

$$f(\mathfrak{r},\mathfrak{v}) = f(\mathfrak{r}) \prod f_i(v_i). \tag{3.14}$$

The velocity distribution, is given by

$$f_i(v_i) = \left(\frac{m_i}{2\pi k_B T}\right)^{3/2} \exp\left(\frac{m_i \boldsymbol{v}_i^2}{2k_B T}\right), \tag{3.15}$$

the Maxwell-Boltzmann distribution All information about the interaction of particles is contained in the configuration partition function Z. The probability density in configuration space is given by

$$f(\mathfrak{r}) = \frac{1}{Z} e^{-\beta V(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_n)}, \tag{3.16}$$

the $n+1$st factor of the factorization of Eq. 3.11. With the Boltzmann distribution it is possible to immediately obtain the probability to find the system in a state characterized by coarse–grained coordinate $\mathfrak{R}$, that is a function of the coordinates $\mathfrak{R}(\mathfrak{r})$:

$$p\left(\mathfrak{R}'\right) = \int \mathrm{d}^{3n}\mathfrak{r} e^{-\beta V(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_n)} \delta\left(\mathfrak{R}' - \mathfrak{R}(\mathfrak{r})\right). \tag{3.17}$$

As an example we might consider a coarse-grained space, which contains the first $N < n$ positions of the configuration space of the original system. Then the probability distribution can be written as

$$p(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_N) = \int \mathrm{d}\boldsymbol{r}_{N+1}^3 \cdots \mathrm{d}_n^3\mathfrak{r}\, e^{-\beta V(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_n)}. \tag{3.18}$$

This expression is named the reduced probability function. The above form of the Boltzmann distribution is the reason why this process of eliminating variables is dubbed *integrating out* degrees of freedom. The free energy landscape

$\mathcal{F}(\mathfrak{r})$ is defined as

$$\mathcal{F}(\mathfrak{R}) = -\frac{1}{\beta} \log p(\mathfrak{R}). \tag{3.19}$$

The word landscape is introduced here because the probability distribution exactly reflects the probability distribution of a single particle in an N-dimensional landscape where the altitude is given, up to appropriate constants, by the value of the free energy at the respective position. By taking the gradient of the expression above, a useful expression can be obtained:

$$\nabla_{\mathfrak{R}} \mathcal{F}(\mathfrak{R}) = \int \mathrm{d}^{3n}\mathfrak{r}\, p(\mathfrak{r})\, \nabla_{\mathfrak{R}} H(\mathfrak{r})\, \delta(\mathfrak{R} - \mathfrak{R}(\mathfrak{r})). \tag{3.20}$$

This can be read as follows: The derivative of the free energy landscape with respect to a coarse-grained coordinate $\mathfrak{R}$, is the mean of the (generalized) force, where the average is taken at all points $\mathfrak{r}$ compatible with the particular coarse-grained configuration $\mathfrak{R}$. Therefore free energy landscape is also called the potential of mean force. Both will be used interchangeably in this thesis.

The difficulty of systematic coarse–graining lies in finding a good model for $\mathcal{F}(\mathfrak{R})$. Even if $V(\mathfrak{r})$ is a sum of pair potentials, the coarse-grained potential will include many-body terms and can be relatively complicated. It is possible to approximate it by pair potentials, e.g. with the iterative Boltzmann inversion procedure (the name was coined by [96], the method is older), which a tabulated pair potential is constructed iteratively that exactly, meaning to arbitrary accuracy, reproduces a radial distribution function. The existence of a unique potential is guaranteed by the Henderson theorem [118].

In equilibrium it is possible to algorithmically derive coarse-grained representations of a system reflecting the fine system in some respect. Assuming Newtonian equations of motion in this effective many-body potential does not create realistic dynamics, as frequent collisions between particles, that have been coarse-grained out, lead a much faster relaxation of momentum towards equilibrium. Formal access to this aspect of coarse-graining can be gained using projection operator techniques. The theoretical foundation for these techniques was developed by Zwanzig [119] Mori [120], and therefore they are nowadays known as the Mori-Zwanzig formalism.

Before studying the details we will give an example of a coarse-grained model for dynamics, Brownian motion. It is so well known, because it was one of the first examples, where the statistical nature of matter was directly observable in experiments.

3.2.3 Brownian motion: A simple case of coarse-grained dynamics

In 1827 the Botanist R. Brown observed that pollen grains in water performed an disordered, stochastic motion [108, 121]. This motion originates from collisions of the water molecules with the immersed particle. With Einsteins seminal work [122], as Renn points out [123], it was not only possible to understand this behaviour qualitatively, but to also use the observation of Brownian motion to measure a quantity that links the micro- to the macroscale: the Avogadro number. Alternatively, by observing Brownian motion one can connect the energy scale of the macro- and the microscale. The thermal energy $k_B T$ can be inferred from measurements of the diffusion constant and the radius of particles and the viscosity of the surrounding medium.

Einstein's contribution to the understanding of Brownian motion was twofold [122]: First, he found that the mean square displacement of a particle is closely related to the self diffusion coefficient of these particles. It equals twice the slope of mean square displacement

$$D = \frac{1}{2} \lim_{\tau \to \infty} \frac{\mathrm{d}}{\mathrm{d}\tau} \langle [x(t) - x(t+\tau)]^2 \rangle_t, \tag{3.21}$$

in the limit of infinite τ.

Second, Einstein discovered the fact that the diffusion coefficient is related to the response of the particle to small applied forces. The formula he derived relates the diffusion coefficient D to the gas constant R, the Avogadro number N_A, the viscosity η and the particle radius a:

$$D = \frac{RT}{N_A} \frac{1}{6\pi\eta a}. \tag{3.22}$$

His finding is based on the very simple argument that the response of the particle to a random force caused by the solvent must be identical to the response to an external force. This is the insight on which the more general fluctuation dissipation theorem discussed below is founded.

A spherical particle immersed in a medium with a viscosity η moving at velocity $\boldsymbol{v}$ experiences, according to Stokes' law, a frictional force $\boldsymbol{F}_D$ given by

$$\boldsymbol{F}_D = -6\pi\eta a \boldsymbol{v} \tag{3.23}$$

The factor $6\pi\eta a$ is defined as the friction coefficient γ of the particle. From the assumption above, and the identification that forces on a particle originate

from gradients of the osmotic pressure, he derived Eq. 3.22.

This finding motivated Paul Langevin to suggest the following effective equation of motion for Brownian particles [124], here given in one dimension. He assumed a friction force $F_D = -\gamma v$ acting on moving particles. In addition he introduced a stochastic force F_R [108] due to collisions with solvent molecules with zero mean and a finite variance. In addition, an external force F_C can be present. Including the forces into Newton's equation of motion leads to the Langevin equation:

$$m\ddot{r} = -\gamma\dot{r} + F_R + F_C. \tag{3.24}$$

The random force is assumed to have zero mean and to be uncorrelated in time. Its strength is characterized by its variance, which we denote by A. This reads as

$$\begin{aligned} \langle F_R \rangle &= 0, \\ \langle F_R(t) F_R(t') \rangle &= A\delta(t - t'). \end{aligned} \tag{3.25}$$

Eq. 3.24 is a stochastic equation of motion. To obtain meaningful results, it is necessary to calculated averages over all realizations of the stochastic process F_R. Two important observations can be made: First, in the absence of external forces the mean square velocity of the particle fulfills:

$$\langle v^2 \rangle = \frac{A}{m}. \tag{3.26}$$

Second, if the external force F can be written as the gradient of a potential V that diverges for $x \to \pm\infty$, in the long time limit, the probability density of finding a particle at position x converges to:

$$p(x) \propto e^{\frac{-2V(x)}{A}}. \tag{3.27}$$

Both suggest to identify the variance of the random force A as twice the thermal energy $k_B T$.

This means that by applying a velocity-dependent frictional force, and a random force of the right magnitude to a particle, it is possible to generate configurations in according to the Boltzmann distribution. Hence, if Newton's equation of motion are exchanged for the Langevin equation in a computer simulation, it is possible to prescribe the temperature and obtain a series of configurations with probabilities according to the Boltzmann distribution. Observable averages can be calculated as time averages of the trajectories, and they approximate the ensemble average in the canonical ensemble. Frictional

and random forces mimic the action of a heat bath fixing the temperature of the system. This thermostatting mechanism is widely applied especially in implicit solvent molecular dynamics simulations, including the simulations in Chapter 4.

The fundamental solution to the Langevin equation as presented in Eq. 3.24 is an exponentially decaying velocity. The characteristic timescale of the decay γ/m is in many cases very small. If this time scale is much smaller than the other time scales of the system the system is *overdamped*. In this limit is possible to treat the dynamics without considering the velocities explicitly. We will discuss this issue below, when introducing hydrodynamic interactions and their algorithmic treatment.

For the Brownian motion of a spherical particle much larger than the solvent particles, a theory which is exact in the limit of large space and time scales can be formulated. Only based on the assumption of the validity of Stokes' law, the diffusion constant can be predicted, given that the Boltzmann constant is known. This prediction is possible by virtue of the fluctuation-dissipation theorem, which links the dissipative frictional force to the speed of the diffusion process. The more general framework is discussed in the following section.

3.2.4 Theory of linear response

In the following section we provide a short introduction to the theory how thermodynamic systems react to small applied forces, the theory of linear response. It is a generalization of what was found for the case of Brownian motion, the relation between the diffusion coefficient and the mobility to general, time-dependent forces and equilibrium fluctuations.

If we assume that the system under consideration is subject to some force $\boldsymbol{F}$, then it is often[2] justified to expand the response $\boldsymbol{j}$ of the system (both will be defined below) in a Taylor expansion

$$j_i = \sum_j \chi_{ij} F_j + \cdots . \tag{3.28}$$

The matrix χ_{ij} is the response matrix, admittance or susceptibility of the system. In general the response of the system is not instantaneous, but needs

[2] In some cases, e.g. two-dimensional hydrodynamics, the theory of linear response fails, as the obtained response coefficients diverge.

a certain time to develop. This is expressed by a time-dependent susceptibility:

$$j_i(t) = \sum_j \int_{-\infty}^{\infty} \mathrm{d}t' \chi_{ij}(t - t') F_j(t') + \cdots . \tag{3.29}$$

Causality demands that $\chi_{ij}(\tau) = 0$ for $\tau < 0$. The Fourier transform of this equation leads to the same form:

$$j_i(\omega) = \sum_j \chi_{ij}(\omega) F_j(\omega), \tag{3.30}$$

and Eq. 3.28 is the low frequency ($\omega \to 0$) limit of this equation. The time-dependent response coefficient or its Fourier transformed frequency-dependent response can be equivalently used to describe the response to a time-dependent force.

So far we have introduced the response functions without defining the corresponding forces or fluxes. Two types of forces can be identified [109]: *mechanical forces*, which can be taken into account directly in the equations of motion, and *thermal forces*, which require a macroscopic collective picture. According to Kubo et al. [109], the Hamiltonian corresponding to mechanical forces must be of the form

$$\mathcal{H} = -\boldsymbol{F} \cdot \boldsymbol{X}. \tag{3.31}$$

Then the corresponding flux can be identified as the time derivatives of the displacement $\boldsymbol{X}$. According to Evans and Morris [105], thermal forces must be connected to their respective fluxes by the form of the entropy production rate

$$T\frac{\mathrm{d}S}{\mathrm{d}t} = \boldsymbol{F} \cdot \boldsymbol{j}. \tag{3.32}$$

$\boldsymbol{F} \cdot \dot{\boldsymbol{X}}$ is the amount of work the external force performs on the system per unit time (the power). This work is converted into heat by dissipation. It is thus plausible that both definition of forces and fluxes for mechanical and thermal forces are in agreement.

Onsager found that the response matrix must be symmetric, and the real and imaginary parts of the admission coefficients are related by Kramers-Kronig relations. A sufficient condition for this is the time-reversibly of the underlying equations of motion.

There is an intimate relationship between the decay of equilibrium fluctuations and the response matrix: the fluctuation-dissipation theorem (FDT). The FDT links the correlation function C_{XY} of two variable X and Y in equi-

librium to the value of the admittance χ_{XY}. It reflects the fact that small applied forces must effect the system in the same way as stochastic fluctuations. For our consideration it is not necessary to formulate the full FDT. We need only one of its immediate consequences, the Green-Kubo formula. It links the Fourier-Transform of the correlation function of the time derivate of two variables $C_{\dot{X}\dot{Y}}$ to the susceptibility by

$$\chi_{XY}(\omega) = \frac{1}{k_B T} \int \mathrm{d}\tau e^{-i\omega\tau} C_{\dot{X}\dot{Y}}(\tau), \tag{3.33}$$

where the correlation function is given by:

$$C_{\dot{X}\dot{Y}} = \langle X(t)\, Y(t+\tau)\rangle_t. \tag{3.34}$$

Here the time average is taken in a system which is not perturbed by the external force, hence an equilibrium system. In context of Brownian motion, we are e.g. interested in the mobility μ, which thus is related to the Force f applied to a particle along the coordinate x. In the limit $\omega \to 0$ we obtain:

$$\mu = \frac{1}{k_B T} \int_0^\infty \mathrm{d}\tau \langle v(t)\, v(t+\tau)\rangle_t, \tag{3.35}$$

where v denotes the velocity of the particle. The equivalence with Eq. 3.21 can be shown by integration by parts [101].

The Green-Kubo relationship is a very useful tool for practical computation, as it allows us to perform a simulation in equilibrium and extract all response functions from a single trajectory. On the other hand, it is a requirement for a thermodynamically consistent simulation. When constructing an algorithm for the study of soft matter systems, it is thus of paramount importance, to make sure it fulfills this requirement. If it does not, it is not possible to obtain unique transport coefficients. We will show in Sec. 5.2 that this prerequisite is not always fulfilled for the Lattice-Boltzmann algorithm.

3.2.5 Hydrodynamics: The continuum limit

The hydrodynamic picture of fluids is based on the assumption of a local thermodynamic equilibrium. With this assumption, it is possible to locally use the statistical mechanics framework developed before, and construct equations based on conservation laws which describe a system on a global scale. Locally, the system can fluctuate, but the subsystems are chosen so large,

that fluctuations are weak or negligible. Only conserved quantities need to be considered, as they can not be created or annihilated, and therefore affect a system on a global scale. Evans and Morris characterize this perspective as follows: "At the hydrodynamic level we are interested in the macroscopic evolution of densities of conserved extensive variables such as mass, energy and momentum. Because these quantities are conserved, their respective densities can only change by a process of redistribution. [...] this means that the relaxation of these densities is slow, and therefore the relaxation plays a macroscopic role." Hydrodynamic theories thus are exact on large time and length scales. Lower bounds to the length scale in a gas is e.g. given by the mean free path, and to the timescale by the collision frequency [105]. In this section, we perform this analysis for a fluid under the effect of forces and pressure differences. This leads to the Navier-Stokes equations. Finally we describe how hydrodynamic interactions can be derived from it.

We change the perspective from individual particles to fields, e.g. the mass density density or the momentum density. Microscopically, they can be defined in analogy to the number density in Eq. 3.8. For every conserved quantity a continuity equation must hold:

$$\frac{\partial q}{\partial t} + \nabla \cdot \tilde{\boldsymbol{j}}_q = 0. \tag{3.36}$$

Here $\tilde{\boldsymbol{j}}_q$ is a flux, which we do not specify further for now. The key idea is to distinguish between a flux related to the motion of the underlying medium, the convective transport, and a flux that originates from local deviations from equilibrium, the diffusive transport. Which type of transport is more important depends on the specific application. In soft matter applications, e.g. heat conduction can almost always be considered a process that is much faster than convective heat transport, and the latter may be neglected for that reason.

If one assumes that a medium is in motion, with a certain position dependent velocity $\boldsymbol{u}$, then the flux of an extensive quantity q, measured per mass density of the medium, can be written as

$$\tilde{\boldsymbol{j}}_q = \rho q \boldsymbol{u} + \boldsymbol{j}_q \tag{3.37}$$

where ρ denotes the local mass density, and $\boldsymbol{j}$ is a flux resulting from a local imbalance of thermodynamic or mechanic forces. If the flux of mass itself is considered, q is set to 1 (the mass density per mass density). In a medium with one component, no mass fluxes that are not associated with a corresponding velocity field need to be considered, and one immediately obtains the mass

continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \boldsymbol{u} = 0. \tag{3.38}$$

To consider the momentum density, it is important to realize, that actually three quantities are considered at the same time, namely the three components of the momentum. Hence momentum fluxes are tensor valued, where the component j_{ij} measures the flux density of i-momentum in j-direction. It can be shown generally that the momentum flux tensor must be symmetric, as a result of the conservation of angular momentum [105].

Now, we consider the mechanical forces acting on a volume element. If the pressure in the surrounding is not constant, a volume element experiences a force proportional to the local pressure gradient:

$$\boldsymbol{f} = \nabla p. \tag{3.39}$$

In the present form this appears as a source term, but we can recast it in the form of the divergence of a tensor, in order to obtain the form of Eq. 3.36. We interpret the tensor

$$p_H = \begin{pmatrix} p/3 & 0 & 0 \\ 0 & p/3 & 0 \\ 0 & 0 & p/3 \end{pmatrix} \tag{3.40}$$

as the tensor describing the momentum flux in a co-moving reference frame due to a local imbalance in the pressure.

These hydrostatic forces are solely caused by a local deviation from mechanical equilibrium. They are conservative, in the sense that one volume element can perform reversible work on another volume element without dissipation, and this work can be recovered by expansion. This is true e.g. for sound waves, which can propagate through a medium with only small losses.

A second class of forces arises from the local local deformation of the medium. The tensor $\nabla \boldsymbol{u}$, the rate of deformation, is a measure for this. It measures how strongly a volume element is deformed per unit time. This deformation leads to a deviation from local thermal equilibrium. If we choose a reference point $\boldsymbol{r}_0$ in space, then for a point $\boldsymbol{r}$ in its vicinity the velocity is given by:

$$\boldsymbol{u}(r) = \boldsymbol{u}(\boldsymbol{r}_0) + \nabla \boldsymbol{u} \cdot (\boldsymbol{r} - \boldsymbol{r}_0). \tag{3.41}$$

The deviation from local equilibrium can be expanded in orders of the rate of deformation. The viscosity is obtained as the linear response coefficient in the limit of zero frequency of the deformation. For an isotropic medium only two

independent response functions are possible: forces opposing a compression, and forces opposing shear. This can be seen from the following decomposition of a tensor A [104]:

$$A = \underbrace{\frac{1}{3}\mathrm{Tr}A\,\mathbb{I}}_{\text{diagonal}} + \underbrace{\frac{1}{2}\left(A + A^T - \frac{2}{3}\mathrm{Tr}A\,\mathbb{I}\right)}_{\text{symmetric traceless}} + \underbrace{\frac{1}{2}\left(A - A^T\right)}_{\text{antisymmetric}} \tag{3.42}$$

$$\tag{3.43}$$

Applying this to the rate of shear tensor ∇u, one obtains:

$$\nabla \boldsymbol{u} = \underbrace{\frac{1}{3}\nabla \cdot \boldsymbol{u}\mathbb{I}}_{\varepsilon_d} + \underbrace{\frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T - \frac{2}{3}\nabla \cdot \boldsymbol{u}\right)}_{\varepsilon_s} + \underbrace{\frac{1}{2}\left(\nabla \boldsymbol{u} - (\nabla \boldsymbol{u})^T\right)}_{\varepsilon_s}, \tag{3.44}$$

with the diagonal ϵ_d, a traceless symmetric ϵ_s, an antisymmetric ϵ_a component of the strain rate tensor. For simple liquids no momentum flux arises from the antisymmetric part of the deformation. Then the viscous momentum flux tensor can be written as

$$\boldsymbol{p}_v = \eta_v \varepsilon_d + \eta \varepsilon_s. \tag{3.45}$$

Here the bulk viscosity, hence the dissipative force counteracting compression η_v and the dynamic viscosity η appear, as the only two independent parameters possible in an isotropic fluid. Inserting this into the conservation equation leads to the (compressible) Navier-Stokes equations:

$$\partial_t \rho \boldsymbol{u} + \nabla \cdot \rho \boldsymbol{u}\boldsymbol{u} = -\nabla p + \eta \nabla^2 \boldsymbol{u} - \left(\eta_v + \frac{1}{3}\eta\right) \nabla \left(\nabla \cdot \boldsymbol{u}\right) \tag{3.46}$$

The last conserved quantity that could be considered is the energy. This leads to an equation where heat conductivity is considered explicitly. In most soft matter applications it is however to assume that temperature is transported much faster through the system than particles. Therefore temperature does not have to be considered explicitly.

If the velocities are much smaller than the speed of sound, the flow is in good approximation incompressible. Then it fulfills

$$\nabla \cdot \boldsymbol{u} = 0. \tag{3.47}$$

The Navier-Stokes equation then can be simplified to:

$$\rho\partial_t \boldsymbol{u} + \rho\boldsymbol{u}\cdot\nabla\boldsymbol{u} = -\nabla p + \eta\nabla^2\boldsymbol{u}. \tag{3.48}$$

The character of the hydrostatic pressure changes when performing this approximation. In the compressible Navier-Stokes equations it was related to the local density by the equation of state $p(\rho)$. Now the density is constant, and the pressure appears purely as a degree of freedom that needs to be chosen such that the flow field can fulfill Eq. 3.47 and Eq. 3.50 simultaneously.

By a dimensional analysis of the Navier-Stokes equation, a parameter that relates the viscous forces to the inertia effects the can be identified, the Reynolds Number

$$\mathrm{Re} = \frac{\rho u_0 l}{\eta}. \tag{3.49}$$

Here u_0 is the scale of the velocity and l is a characteristic length scale of the system. In virtually all soft matter applications the velocities as well as the length scales are sufficiently small to make the small Reynolds number approximation. This leads to the Stokes equation, where the convective term in the Navier-Stokes equation is neglected:

$$\rho\partial_t \boldsymbol{u} = -\nabla p + \eta\nabla^2\boldsymbol{u} + \boldsymbol{f}. \tag{3.50}$$

Here we have added a *body force density* $\boldsymbol{f}$, which can e.g. correspond to gravity or to forces related to electric fields, as will be seen below. In the stationary case, where $\partial_t \boldsymbol{u} = 0$, the Stokes equation, plus the continuity equation are a linear, coupled set of PDEs, which can be analyzed in a similar fashion as the Poisson equation for electrostatics.

Stokes' law can be obtained by solving the Stokes equation with the boundary conditions $\boldsymbol{u} = \boldsymbol{v}$ on a spherical surface, and $\boldsymbol{u} = 0$ at infinite distance from the sphere. The resulting flow field is asymptotically equal to a Stokeslet multiplied with the force applied to the particle. Only for the near field, corrections are necessary. The boundary condition $\boldsymbol{u} = \boldsymbol{v}$ is called the *no-slip* boundary condition. Its validity has been under heavy debate, as it can not be derived from first principles. Until the advent of microfluidics, however, no evidence was found that it might be violated. Examples, for which slip effect has been related to dissolved gas or hydrophobic interactions [125]. Still there is no convergence of the discussion. It has e.g. been shown, that surface roughness has a similar effect as slip [126, 127]. The no-slip boundary condition remains a somewhat mysterious but very efficient tool in the study of hydrodynamic effects.

In a three-dimensional unbound system a point force $\boldsymbol{F}\delta(\boldsymbol{r}-\boldsymbol{r}_0)$ leads to a flow field

$$\boldsymbol{u}(\boldsymbol{r}) = \frac{1}{8\pi\eta\,|\boldsymbol{r}-\boldsymbol{r}_0|}\left(1+\frac{(\boldsymbol{r}-\boldsymbol{r}_0)(\boldsymbol{r}-\boldsymbol{r}_0)}{|\boldsymbol{r}-\boldsymbol{r}_0|^2}\right)\cdot\boldsymbol{F}. \tag{3.51}$$

This expression is the fundamental solution of the Stokes equation, and corresponds formally to the Coulomb potential of electrostatics. Every solution of the Stokes equation can be written as a superposition of expressions of the type 3.51. The fundamental solution is referred to as the Stokeslet.

Assume we want to study the motion of n spherical particles immersed in a liquid on which the forces $\boldsymbol{F}_i$ act. We want to know their velocities $\boldsymbol{v}_i$ in the stationary state. This problem can be solved approximately by using a pairwise additive approximation. We will show the route to the simplest of these approximation, the Oseen form of the mobility matrix, sometimes called the Oseen tensor. In the stationary state, the force on each particle, the applied force $\boldsymbol{F}_i$, is right away transferred to the fluid. In terms of the Stokeslet, the flow field can be written as

$$\boldsymbol{u}(\boldsymbol{r}) = \sum_i \frac{1}{8\pi\eta\,|\boldsymbol{r}-\boldsymbol{r}_i|}\left(1+\frac{(\boldsymbol{r}-\boldsymbol{r}_i)(\boldsymbol{r}-\boldsymbol{r}_i)}{|\boldsymbol{r}-\boldsymbol{r}_i|^2}\right)\cdot\boldsymbol{F}_i. \tag{3.52}$$

By virtue of Stokes' law, the hydroynamic force on the particles can be evaluated as $\boldsymbol{F}_{i,h} = 6\pi\eta a\,(\boldsymbol{u}(\boldsymbol{r}_i)-\boldsymbol{v}_i)$. The hydrodynamic force in turn must be the negative applied force. This expression can solved for the velocities $\boldsymbol{v}_i$ and we obtain:

$$\boldsymbol{v}_i = \frac{1}{6\pi\eta a}\boldsymbol{F}_i + \sum_{i\neq j}\frac{1}{8\pi\eta\,|\boldsymbol{r}_{ij}|}\left(1+\hat{\boldsymbol{r}}_{ij}\hat{\boldsymbol{r}}_{ij}\right)\boldsymbol{F}_j, \tag{3.53}$$

where we use a shorthand notation $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$ and $\hat{\boldsymbol{r}}_{ij} = \boldsymbol{r}_{ij}/|\boldsymbol{r}_{ij}|$ for the particle distances.

The expression above can be interpreted as a mobility matrix μ. The Oseen approximation is only valid for particle much further apart than their radius. Rotne and Prager gave an improved expression, which they obtained by minimizing the dissipation in the system [128]. In general, their expression is preferred over the Oseen form, as the mobility matrix in this case is positive definite. Even better expressions can be obtained by incorporating *lubrication corrections* to improve the short-range behaviour of the mobility matrix [129].

Taking into account diffusion, the particle motion can be expressed in terms of a Fokker-Planck equation [130] for the many particle probability density $p(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n)$,

$$\partial_t p = \sum_{i,j} \nabla_{\boldsymbol{r}_i} \cdot D_{ij} \left[\nabla_{\boldsymbol{r}_j} p + \boldsymbol{F}_j p \right]. \tag{3.54}$$

This expression is suitable for theoretical analysis. Examples are the calculations of the diffusion coefficients in suspension of colloidal particles [131–133] and the diffusion of polymers [134]. To perform computer simulations with this type of interaction tensor, Ermak and McCammon derived a simulation scheme, where the corresponding stochastic equation of motion is used to create trajectories that include the hydrodynamic interactions [130], which will briefly be describe below.

3.2.6 Theory of the transition state: Dynamics in 1D

The theory of collective dynamics in a single coordinate is probably oldest branch of statistical physics out of thermal equilibrium, as it describes the dynamics of chemical reactions. We give a short introduction following the review of Hänggi and Borkovec [135].

In the late 19th century van't Hoff [136] and Arrhenius [137] both found an exponential dependence of reaction rates with the negative inverse temperature, now known as the Arrhenius law:

$$k = k_0 \exp - \frac{E_A}{k_B T} \tag{3.55}$$

Here k is the rate at which the reaction occurs, E_A is the activation energy and k_0 the so-called pre-exponential factor. At their time neither the atomistic nature of matter was certain nor the quantitative value of the Boltzmann constant or Avogadro number was known, and hence only the constant E_A/k_B could be determined for each reaction. The interpretation as a microscopic activation energy could only be given later.

The key to giving an explanation of this behaviour was the interpretation of E_A as the energy difference between the educt and some intermediate structure that forms in course of the chemical reaction, the transition state. In this simplification, only three states are considered: the educt state, the transition state and the product state. To deduce the rate at which the reaction occurs, we first make the following consideration: If there were no product

state, the probability of finding the system in the transition state is would be proportional to the Boltzmann factor:

$$p(T) = \frac{e^{\frac{-E_A}{k_B T}}}{1 + e^{\frac{-E_A}{k_B T}}} \approx e^{\frac{-E_A}{k_B T}} \tag{3.56}$$

The actual rate at which a reaction from the transition state to the product occurs is proportional to the probability of finding the system in the transition state. With these assumption the Arrhenius equation is immediately obtained.

The work of Kramers [138] provides the connection to Brownian motion in one dimension. His assumption was that a chemical reaction and Brownian motion are linked such that every molecule (or group of molecule) undergoes a Langevin-type motion in a certain coordinate, the reaction coordinate. The potential energy contains all information about the chemical environment. The product and the educt state correspond to local minima in the potential energy. Chemical kinetics can then be derived by investigating the statistics of the corresponding trajectories, from the analysis of the Fokker-Planck equation. From our consideration in Sec. 3.2.2, it is obvious that the potential energy that needs to be inserted in the Langevin equation is actually the free energy, or potential of mean force expressed in the reaction coordinate.

Two important limits of Kramer's theory can be identified: In the low friction limit, the Langevin equation reflects Newton's equation of motion, valid for equations in the gas phase [139]. The high friction limit is appropriate for typical dense soft matter systems. Here ballistic motion is largely suppressed due to frictional forces. Interestingly in both, the high and the low friction limit, rates of the van't-Hoff-Arrhenius form with different pre-exponential factors arise.

These considerations emphasize the importance of the role of the transition state for the dynamics of all processes, also in the field of soft matter. One example is the folding of biomolecules. In this case it is very interesting to identify transition states, i.e. states which a system is likely to pass, when moving between two stable states. Modifying the transition state can dramatically influence the dynamics of the corresponding process.

In Ch. 4 we consider the dynamics of transport of a DNA molecule through a pore. We calculate the free energy barrier in terms of a certain reaction coordinate. We find a free energy barrier that strongly depends on conditions like the salt concentration and the diameter of the pore. In terms of the theory

of the transition state this suggests, that also the transport rate depends significantly on the free energy barrier.

An important assumption for these consideration is that the dynamics can be described in terms of Brownian motion in one dimension. This is only justified if all other processes happen faster than the considered process. In the following section we will discuss DNA from the point of view of polymer science. This will lead to a simple, but very interesting point. For long polymer lengths the DNA dynamics in a nanopore can not be described in terms of a single reaction coordinate, because the assumption that only one slow degree of freedom exists is invalid. The chain lengths considered in Sec. 4.3.4 should be, however, short enough to be on the safe side.

3.2.7 DNA from the point of view of polymer science

DNA's importance for biology can hardly be underestimated. From a somewhat coarser perspective, the DNA molecule is, however, a polymer that follows the same laws a synthetic polymer, namely the statistical physics of chain molecules. We give a very brief introduction to a few theoretical concepts that will be needed to justify our coarse-grained DNA models. The information is compiled from various text books [140–144].

Starting with the work of Debye, Flory, Kramers, Kuhn and others scientists in the early 20th century a new branch of physics was founded: the physics of long chain–like molecules. A particularly successful perspective was pioneered by de Gennes: the scaling approach to polymer science. The key observation was, that polymer properties follow scaling laws in the limit of long lengths. In this scaling limit the chemical details are not important. These universality concepts had been developed in the context of critical phenomena, so that polymer science could benefit from the "vast knowledge" [140] that had been accumulated in the field. This universality allows polymer scientists to formulate different models and use any of them to answer a particular question. The answer can always be transferred to other models by a simple rescaling procedure. This also facilitates the use of computers. A model does not have to be chosen to achieve detailed agreement to the physical system of interest, but can be chosen by the convenience of its computational evaluation. We illustrate this by explaining a few chain models. We follow the textbook of Grosberg et al. [141]:

In the *freely jointed chain model* a chain is represented by a series of N points (*monomers*) with a mutual distance l in space. No further restrictions are

imposed, e.g. the chain can bend sharply and cross itself. Alternatively the chain can described by its starting point and $N-1$ bond vectors $\boldsymbol{u}_i$ of length l. The bond vectors are uncorrelated, which can be expressed as

$$\langle \boldsymbol{u}_i \boldsymbol{u}_j \rangle = l^2 \delta_{ij}, \tag{3.57}$$

where $\delta_{ij}$is the Kronecker delta:

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j, \\ 0 \text{ else.} \end{cases} \tag{3.58}$$

The end-to-end vector $\boldsymbol{R}$ is the sum of all bond vectors and the end-to-end distance R_e is its length. The total length of the chain is called the contour length $L_c = (N-1)\,l$. It is easy to show that the end-to-end distance grows with the square root of the number of monomers N:

$$\langle R_e \rangle = \sqrt{N-1} l. \tag{3.59}$$

If we now consider a real chain molecule, say polyethylen, where the repeat unit is CH_2, the vectors that connect carbon atoms clearly can not fluctuate freely. On the contrary, the bond angle will stay close to the tetrahedral angle of 110 °. This, of course, reduces the flexibility. In other polymers other mechanisms creating stiffness are present. In our model for single-stranded DNA shown in Sec. 4.3.4, for example, electrostatic repulsion favors locally straight conformations. The role of stiffness can be expressed by a single parameter: the persistence length. It is the characteristic decay length of orientational correlation of monomers. Another measure of stiffness is the effective segment length, the Kuhn length l_K. It is defined such that the end-to-end distance for a stiff chain exactly fulfills

$$\langle R_e \rangle = l_K \sqrt{N}. \tag{3.60}$$

Both stiffness measures are consistent, thus both serve for rescaling. Their numerical value does depend on the underlying polymer model. An equivalent model is the random walk on a lattice. It can be considered a special case of the freely jointed chain where only bond angles compatible with a certain lattice (e.g. square or cubic lattice) are accepted. Also in this case the end-to-end distance increases with the square root of the chain length N:

$$\langle R_e \rangle = \frac{1}{\sqrt{d}} a \sqrt{N-1}. \tag{3.61}$$

Here d denotes the dimension and a the lattice spacing, and square or cubic lattice is assumed. It is obvious that the Kuhn length is different from the bond length. Physical quantities, can be calculated in both models in completely equivalent ways.

The fundamental universality in the behaviour above is expressed in scaling laws. The equation above would be expressed as

$$R_e \sim N^{1/2}, \tag{3.62}$$

which reads as follows: R_e scales as N to the power of $1/2$. It means that the end-to-end distance asymptotically approaches the functional form $A\sqrt{N}$, where the numerical prefactor A, the length scale, is dropped. This scaling law as such holds for all polymers of the according universality class, and for a particular model the prefactor must be determined individually.

The chain models we have presented so far are *ideal*, which means that interactions between the monomers are neglected. In real systems the interactions can only be neglected under very special conditions, so called θ–conditions. Under these conditions, repulsive interactions (*excluded volume interactions*) are counterbalanced by attractive interactions. Under all other conditions, at least one extra parameter is necessary for the relative importance of excluded volume interactions: the *second virial coefficient*. The underlying idea is borrowed from the theory of dense gases and dilute solutions. For low densities the effect of interactions can be described solely by taking into account particle pairs. The second virial coefficient is defined as

$$B = \int_0^\infty \left(1 - exp\left(-\frac{V(\boldsymbol{r})}{kT}\right)\right) \mathrm{d}^3 r, \tag{3.63}$$

where $V(\boldsymbol{r})$ is the interaction potential of two monomers. It thus characterizes the balance of attractive and repulsive forces, weighted by the Boltzmann factor. For high temperatures the second virial coefficient is positive due to repulsive "excluded volume" interactions. For low temperature in presence of attractive forces it is negative. The temperature for which the second virial coefficient is zero is called the θ-temperature. The condition $T > \theta$ is sometimes called *good solvent condition* as the potential u in the equation above refers to the effective pair potential, where the solvent molecules have been integrated out. Under good solvent conditions the interaction between chain and solvent is more favourable that the direct chain–chain interaction. The solvent condition separates three solvent different classes, with different scaling behaviour. (a) Under good solvent condition the end-to-end distance

scales like $R_e \sim N^{\nu}$, where ν is the Flory exponent. It equals 0.588 in three spatial dimension . (b) Under θ-conditions the ideal behaviour is recovered: $R_e \sim N^{1/2}$ because attractive and repulsive forces cancel. (c) For $B < 0$, under bad solvent conditions, attractive forces dominate and the polymer chain collapses to a *globule*, where $R_e \sim N^{1/3}$.

The typical computer model for polymers under good solvent condition is the bead-spring model [145] where monomer interactions are modelled by the Weeks-Chandler-Anderson potential and finite-extensible non-linear elastic (FENE) bond. The WCA potential is a truncated and shifted variant of the Lennard-Jones potential:

$$\frac{U_{\mathrm{LJ}}(r)}{k_B T} = \begin{cases} 4\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] + 1 & \text{if } r < 2^{1/6}\sigma \\ 0 & \text{otherwise.} \end{cases} \tag{3.64}$$

It can be considered an MD variant of a hard sphere interaction potential. The FENE bond potential reads as

$$U_{\mathrm{FENE}}(r) = -\frac{K r_{\max}^2}{2} \log\left[1 - (r/r_{\max})^2\right]. \tag{3.65}$$

It diverges at $r = r_{\max}$, which represents the finite extensibility of polymer bonds. This model, plus the Langevin equation of motion is probably the most widely used computer model for polymers.

When addressing the dynamics of a single chain, or a dilute solution in a solvent, two competing models have been developed. In the Rouse model hydrodynamic interactions are neglected, and each monomer undergoes an overdamped Brownian motion. In this case, the time of a polymer to relax to its equilibrium state scales as $N^{1+2\nu}$. In the Zimm model, hydrodynamics are taken into account, and correlation between monomer motion accelerate the relaxation, and the relaxation time scales as as $N^{3\nu}$. These relaxation times will be important below.

A quite simple question has attracted much interest since the advent of DNA translocation experiments and yet remains not fully solved: How does the mean translocation time τ of a chain molecule depend on its number of monomers N?

Inserting a polymer into a pore is associated with a decrease in chain entropy as a large number of conformations is disallowed by steric hindrance. Assuming that s monomers are on one side of the pore and $N - s$ monomers are on

the other side of the pore, and assuming the pore is very narrow and short, the entropic cost can be expressed as the free energy penalty of tethering two polymer chains of size s and $N - s$ to a wall. According to [146] and [10] this is associated with and entropic free energy barrier:

$$\mathcal{A}(s)/k_BT = \gamma(\log s + \log(N - s)), \tag{3.66}$$

where γ is a prefactor of order 1 that depends on the chain model (0.69 for self–avoiding chains in 3D). The definition of the free energy requires thermodynamic equilibrium, which means that all other degrees of freedom have sufficient time relax to thermodynamic equilibrium while changes of the reaction coordinate s happen on a longer timescale. In other words: all degrees of freedom of the system change much more rapidly than the reaction coordinate s. The chain is in *quasiequilibrium* for every value of s. Under these conditions, the translocation process can be described as a diffusion process of s in an external potential. This would imply a scaling of $\tau \sim N^2$, just as for free diffusional processes the square distance travelled is linear in time. A simple but very interesting observation brought forward in [147] is, that the relaxation time of the chain ends scales as $N^{1+2\nu}$ in the Rouse model and for long enough chains will be larger than the N^2 scaling of the translocation time. The Zimm model is invalid in the presence of walls as they screen hydrodynamic interactions. Thus, for long enough chains the relaxation time will be longer than the translocation time itself, and thus the quasiequilibrium assumption must be invalid.

A particularly clear account of this intrinsic non-equilibrium effect was given by Gauthier et al. [148], who showed with molecular dynamics simulations that the escape of a chain from a nanopore changes dramatically if interrupted half–way and given time to relax to its equilibrium. De Haan et al. were able to show, that a whole range of exponents can be obtained in computer experiments by systematically varying the solvent viscosity and the friction in the pore [149]. This explains why so many different apparently contradicting values had been found before. Only very recently the discussion appears to converge, at least for the case of unforced translocation, after Sakaue introduced the concept of chain tension propagation [150], which was extended e.g. by Rowghanian and Grosberg [151]. The simple question however has proved to have a quite complicated answer and a full understanding is not yet achieved.

3.3 Simulation approaches to soft matter

The "softness" property of soft matter is a very important reason why computer simulation models have been applied to soft matter systems already very early. Processes that require only a small timespan, or a small perturbation, are computationally much more feasible than those where materials withstand strong external forces. On the contrary, in solid state materials for example, the fastest vibrational time scales and other time scales, e.g. dislocation dynamics, are much further separated, and therefore simulations require much more computational effort. Some of the first computer simulations were performed to calculate the thermodynamic properties of a hard disc fluid [152]. This was only possible, because liquids are soft and therefore explore their phase space in a relatively small amount of time.

We would like to stress here in what sense the term *simulation* is applied in this thesis: We refer to procedures relying on the concept of classical particles, where a potential energy function, and possibly a propagation mechanism specified, and particle configurations are generated successively by a computer code. These simulation, also termed *molecular* simulations are hence a classical example of bottom-up modelling: properties of small constituents are assumed, and collective properties can be obtained from the calculations.

3.3.1 Monte-Carlo simulations

Two main directions can be identified in molecular simulations [153]. In Monte–Carlo approaches a stochastic algorithm is used to generate a series of configurations $\mathbf{r}$ of the system according to a given distribution function f. The method to generate these configurations is chosen solely by convenience and efficiency considerations. In Molecular Dynamics approaches, the mechanism to generate configurations is part of the modeling process. The equations of motion chosen are usually Newton's equation of motion, possibly with modifications to incorporate a heat bath, or the Langevin equation.

Monte-Carlo methods are mainly suited to measure static quantities Q, that can be expressed as averages over configuration space:

$$Q = \langle Q(\mathbf{r}) \rangle_f . \tag{3.67}$$

The "dynamics", thus the sequence of configurations that is obtained is not direct physical meaning. The most famous method to generate configurations

is the Metropolis algorithm [152]. In this algorithm initially a random starting configuration $\mathfrak{r}_0$ is generated. New configurations are generated in subsequent moves consisting of three steps: (1) a small displacements $\boldsymbol{\delta}\mathfrak{r}$ is chosen randomly. (2) the potential energy of the configuration is evaluated, (3) the move is accepted with a probability

$$p = \begin{cases} 1 & \text{if } V(\mathfrak{r}+\delta\mathfrak{r}) < V(\mathfrak{r}) \\ \exp\left(-\dfrac{V(\mathfrak{r}+\delta\mathfrak{r}) - V(\mathfrak{r})}{k_B T}\right) & \text{if } V(\mathfrak{r}+\delta\mathfrak{r}) < V(\mathfrak{r}), \end{cases} \tag{3.68}$$

where T is the desired temperature. The probability distribution of the resulting Markov chain of states converges to the Boltzmann distribution

$$p(\mathfrak{r}) \propto \exp -\frac{V(\mathfrak{r})}{k_B T}, \tag{3.69}$$

if the probability distribution of random displacements fulfills certain criteria. Then quantities Q can be calculated as averages over the sampled states.

Monte-Carlo methods offer a lot of algorithmic freedom, as the propagation method can be chosen in accordance with the problem that is studied. On the other hand, they have been somewhat on the decline in the last decades, because typically they are difficult to implement efficiently on parallel computer architectures. The Metropolis algorithm requires for efficient execution that every move changes the potential energy by $\lesssim k_B T$, so the elementary steps are very small. Typically, only single particles are displaced, requiring only few computational operations. This not a promising scenario for parallel computers. Molecular dynamics simulations are much better suited for parallel execution, because all particles are displaced, which requires a calculation of the forces on all particles in every step.

3.3.2 Molecular dynamics simulations

The idea of molecular dynamics simulation very simple. One constructs a particle model with conservative interactions and computes its time evolution by numerically integrating Newton's equations of motion. The interactions typically conserve energy, and hence are called potentials. They depend only on the positions of the other particles. Newton's equation of motion for particle i with mass m reads as

$$m_i \frac{\partial^2}{\partial t^2} \boldsymbol{r}_i = -\nabla_{\boldsymbol{r}_i} V(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n). \tag{3.70}$$

Here n denotes the number of particles present in the system, and V the potential energy of the system according to the model. In three dimension one obtains thus a set of $3n$ coupled ordinary differential equations (ODEs) that can be integrated with standard methods [154]. The timestep δt is chosen time independent. Explicit integration schemes are in typically preferred over implicit schemes as the number of degrees of freedom can be very large. A frequently applied integration scheme is the Velocity-Verlet (VV) scheme which consists of four substeps. Let denote $\boldsymbol{F}_i(t)$ be the force that acts on particle i based on the positions $\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n$ evaluated at time t and $\boldsymbol{v}_i$ the velocity of particle i . Then the four steps of the VV integration scheme read as

$$\begin{aligned}
&(1) \quad \boldsymbol{v}_i(t+\delta t/2) = \boldsymbol{v}_i(t) + \frac{\delta t}{2}\frac{1}{m_i}\boldsymbol{F}_i(t) \\
&(2) \quad \boldsymbol{r}_i(t+\delta t) = \boldsymbol{r}_i(t) + \delta t\, \boldsymbol{v}_i(t+\delta t/2) \\
&(3) \quad \boldsymbol{F}_i(t+\delta t) = -\nabla_{\boldsymbol{r}_i} V(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_n) \\
&(4) \quad \boldsymbol{v}_i(t+\delta t) = \boldsymbol{v}_i(t+\delta t/2) + \frac{\delta t}{2}\frac{1}{m_i}\boldsymbol{F}_i(t+\delta t)
\end{aligned} \tag{3.71}$$

The VV scheme is a second order propagation scheme that is time–reversible. The time–reversibility condition is important because it is a good indicator for the stability. The VV scheme can be considered the gold standard for MD applications. The expensive force evaluation only has to be performed once per step. Time–reversibility ensures stability and the computer code is almost as simple as for a first–order integration scheme.

It is tempting to apply the same integration scheme also for the integration of the Langevin equation. This is done in ESPResSo as well, but according to [155] this is not a good choice under all conditions. It must be noted, that the Langevin equation itself is not time–reversible and thus this feature of VV scheme is in vain. In other words, the frictional force and the noise in the Langevin ensure that trajectories are *not* stable in the sense of classical trajectories, but are in (in the continuous limit) nowhere differentiable. The current implementation in ESPResSo uses indeed the positions at full timesteps and the velocities at half timesteps for the calculation of the conservative forces and the dissipative forces, which already indicates a certain inconsistency. In the limit of weak frictional forces, however, Newton's equations of motion are recovered and in this limit it is possibly useful to apply the VV scheme also to the Langevin equations of motion.

Apart from using the Langevin equations of motion, different mechanisms can be applied to allow for heat exchange with a hypothetical reservoir. These

thermostats typically are modifications to the equations of motion, and allow the user to prescribe a certain temperature. In the simulations in Sec. 5.3 a stochastic velocity rescaling mechanism is used, that is based on the Berendsen thermostat [156]. In regular intervals the kinetic energy is calculated. If it smaller than the mean kinetic energy expected from the Boltzmann distribution, the velocities are scaled up, and if it is lower, they are scaled down. By adding a stochastic contribution to this rescaling mechanism, it can be assured that the correct distribution of the kinetic energy and hence a proper canonic ensemble can be obtained [157].

If the interactions are of finite range or have a finite cut-off r_c, the time necessary for an update can be reduced to $\mathcal{O}(n)$. This is done using the concept illustrated in Fig. 3.6. The simulation box is partitioned into cells with an edge length $\gtrsim r_c$. The interactions of a particle in a particular cell only need to be calculated with particles from neighbouring cells. The lists containing the particles are updated in every MD step. The resulting scheme is known as the *linked cell algorithm*. Cell lists are particularly well suited for a parallelization based on a spatial decomposition of the simulation domain. In this case, every process is responsible for calculating the forces in a well-defined region in space. At the boundary of this region, the process maintains copies of the particles in cells surrounding this region. The forces and positions on these copies, the so-called *ghost particles*, must be communicated in every MD step. This can lead to a significant dominant communication overhead. To keep the overhead small compared to the actual computational work, the number of particles per process must be sufficiently large.

Electrostatic interactions can not be treated with this method, as their long-range nature forbids cutting them off at a certain distance. In the following section we discuss how they can be calculated efficiently.

3.3.3 Electrostatic interactions

In a molecular dynamics simulation, calculating the electrostatic forces is a challenging issue. As the electrostatic potential decays only as $1/r$, it is long-ranged. The occurring complication can be already seen by the following simple examples adapted from [158]. For a short-ranged interaction potential it is possible to perform a density expansion of the pressure. The first term in the expansion is $\propto c$, where c is the number density. The prefactor is the

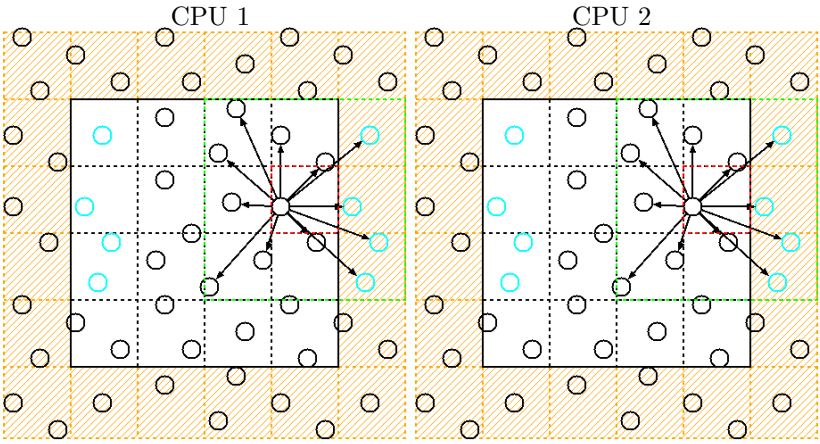


Figure 3.6: Illustration of the domain decomposition parallelization scheme of a molecular dynamics algorithm using cell lists. The computational domain is decomposed into subdomains according to the number of CPUs. Every CPU stores the particles inside its computational domain and a ghost region surrounding its computational domain. When particle cross a subdomain border, they move from CPU to CPU. Each subdomain is decomposed into cells. The cell size is chosen such, that only a small number of neighbouring cells, typically only direct neighbours, need to be considered in the calculation of interactions. This reduces the computational complexity of molecular dynamics simulations to $\mathcal{O}(n)$. Illustration from `http://en.wikipedia.org/wiki/Cell_lists`.

second virial coefficient:

$$B = 2\pi \int_0^\infty \mathrm{d}r\, r^2 V(r)\,. \tag{3.72}$$

If V is the Coulomb potential $\propto 1/r$, the integral diverges, not only for short distances (this could be repaired by introducing a repulsive pair interaction) but also for large r. Indeed, the pressure for a charged system can be expanded in orders of $\sqrt{c}$, so that the $p(c)$ has an infinite slope at $c = 0$, reflecting the divergence in Eq. 3.72.

In some systems, a truncation of electrostatic interactions is possible because the long range nature is compensated for by screening effects [159]. But even then the cutoff radius typically must be relatively large. As the computational

effort strongly depends cubically on the cut-off distance, this is undesired. A notable number of algorithms has been developed to circumvent this problem. The first method was developed by Ewald [160] to calculate the Madelung constant. More advanced variants of this Ewald summation method are applied in most common MD codes, including ESPResSo. Therefore, a brief introduction is given here. Before that, we would like to make two remarks. First, we introduce Poisson's equation and the applied unit system, and second, we briefly introduce the Fourier transformation. Different text book sources (e.g. [101]) and original research articles [161–163] were used as sources. As improvement of electrostatic algorithms is an active area of research in the Holm group and the SFB 716, some parts of the following section may also stem from personal communication with experts.

We apply a unit system in which the thermal energy k_BT is the scale of energy and the elementary charge e is the scale of charge. In this unit system the Bjerrum length l_B emerges naturally. It is the distance at which the potential energy of two elementary charges equals k_BT. It is given by

$$l_B = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r k_B T}, \tag{3.73}$$

where ε_r is the relative permittivity of the considered medium. In this unit system the Poisson equation for the electrostatic potential Φ, measured in k_BT per elementary charge, reads as

$$\Delta\Phi = -4\pi l_B \rho, \tag{3.74}$$

where ρ is the charge density in elementary charges per unit length cubed. The electric field is the negative gradient of the electric field $E = -\nabla\Phi$ and is measured in k_BT per elementary charge and unit length. In Chapter 4, we will discuss extensively how to introduce an inhomogeneous dielectric constant.

We additionally introduce the Fourier transformation. We define the continuous Fourier transformation $\hat{f}$ of a function f as

$$\hat{f}(\boldsymbol{k}) = \int e^{-i\boldsymbol{k}\boldsymbol{x}} f(\boldsymbol{x})\, \mathrm{d}^3\boldsymbol{x}. \tag{3.75}$$

Here $\boldsymbol{x}$ is a 3-dimensional vector. The prefactor for the proper normalization are defined as part of the back-transformation:

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^3} \int e^{i\boldsymbol{k}\boldsymbol{x}} \hat{f}(\boldsymbol{k})\, \mathrm{d}^3\boldsymbol{k}. \tag{3.76}$$

For a periodic function only a discrete set of Fourier coefficients is different from zero, namely those that are compatible with the length of the unit cell. In the x direction with unit cell size L_x these are:

$$k_x = 2\pi n / L_x, \tag{3.77}$$

where n is an integer number. Clearly, if a function f with a finite support is given, the Fourier transformation of its periodic continuation (assuming a certain unit cell) can easily be constructed, as this implies only discarding the Fourier modes incompatible with the assumed unit cell size. The Ewald summation makes use of this fact. Periodic boundary conditions automatically appear due to the application of Fourier transformations.

The key idea of the Ewald summation is to formally split the charge density into two, calculate the electric energy and the electric field separately with different methods and add them up again at the end. According to their characteristics, one is called the short-range (sr) and the other one the long-range (lr) part of the charge density, electrostatic potential and force respectively. The short range part is treated in real space and the long range part in Fourier space. A full treatment in Fourier space is impractical because for small distances the potential can become very steep, and thus many Fourier modes would have to be taken into account. Therefore, before solving the Poisson equation in Fourier space a smoothing operation of the charge density is performed. This is done by a convolution with a smoothing kernel S. Formally we write

$$\rho_{lr}(r) = \int \mathrm{d}^3\boldsymbol{r}' \rho(\boldsymbol{r}') S(\boldsymbol{r} - \boldsymbol{r}'). \tag{3.78}$$

In case of the Ewald summation a convolution with a Gaussian smoothing kernel is used. We have

$$S(r) = \left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \exp -\alpha^2 r^2. \tag{3.79}$$

The parameter α is called the splitting parameter and its inverse is the characteristic smoothing length.

The short range charge distribution is constructed as the complement:

$$\rho_{sr}(r) = \int \mathrm{d}^3 r' \rho(r') (1 - S(r - r')). \tag{3.80}$$

In the short range charge distribution, every peak in the original charge distribution is thus surrounded by a smoothed complementary Gaussian distri-

bution of opposite sign. Conceptually this is shown in Fig. 3.7a. Accordingly the electrostatic potential and field are each split into a component due to the short and long range charge distribution ρ_{sl} and ρ_{lr}. We denote them by Φ_{lr} and Φ_{sr}, and $\boldsymbol{E}_{lr}$ and $\boldsymbol{E}_{sr}$. The total electrostatic energy can be calculated from

$$U_{\mathrm{el}} = \frac{1}{2}\int \mathrm{d}V \rho \Phi_{sr} + \frac{1}{2}\int \mathrm{d}V \rho \Phi_{lr}. \tag{3.81}$$

The first sum can easily be evaluated in real space, as the short-range contribution to the electrostatic potential caused by by each particle decays on the length scale $1/\alpha$ as shown in Fig. 3.7b. The second sum is evaluated in Fourier space. This leads to an expression that quickly converges in Fourier space:

$$\int \mathrm{d}V \rho \Phi_{lr} = \frac{1}{2L^3}\sum_{\boldsymbol{k}\neq 0} \frac{4\pi l_B}{k^2} e^{-k^2/4\alpha^2} \left|\hat{\rho}\left(\boldsymbol{k}\right)\right|^2. \tag{3.82}$$

Here $\hat{\rho}\left(\boldsymbol{k}\right)$ is the Fourier transform of the charge density. The long range electrostatic potential can be identified as $-k^2 e^{-k^2/4\alpha^2}\hat{\rho}\left(\boldsymbol{k}\right)$, as the convolution in Eq. 3.78 in Fourier space reduces to a simple product, and the Laplace operator is just a multiplication with $-k^2$.

The forces are obtained analogously. The short range contribution to the forces is obtained directly from differentiation of the potential energy contribution. The long range on particle i part can be expressed in the following form:

$$\boldsymbol{F}_i = q_i \sum_k e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} \frac{4\pi l_B \boldsymbol{k}}{k^2} e^{-k^2/4\alpha^2} \hat{\rho}\left(\boldsymbol{k}\right), \tag{3.83}$$

hence also as a function decaying quickly in Fourier space. The computational effort of the Ewald summation of n charges scales as $n^{3/2}$, if the splitting parameter is chosen optimally [164, 165].

The class of mesh-Ewald summation algorithms makes use of the fact that the computational effort for the Fourier transformation can be substantially reduced by the use of the Fast Fourier Transformation (FFT). The FFT is a computational algorithm for the calculation of the discrete Fourier transformation. The discrete Fourier Transformation is very similar to the Fourier Transformation shown before but for a function defined on an equidistant grid of points. It has a favourable scaling of $\mathcal{O}\left(N \log N\right)$ where N is the number of grid points. As it is very widely used very efficient computational libraries as the FFTW [166] are available. For the application of a mesh-Ewald summation, first a suitable grid-representation for the point charges must be constructed. This process is called charge assignment. After the

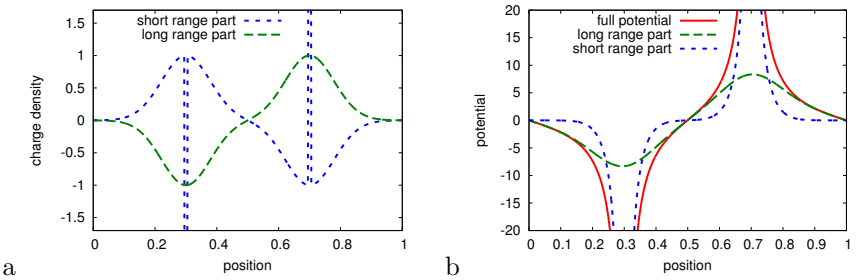


Figure 3.7: The Ewald decomposition of the (a) charge density and (b) the electrostatic potential. Two charges -1 and +1 are positioned at x=0.3 and x=0.7 of a periodic box with edge length 1. For the calculation of the short range potential, the source charges are surrounded by Gaussian screening charges. The long range charge density is the negative of the screening charge density, which corresponds to a smeared variant of the original charge density. The splitting parameter is in this case set to $10/l$, where l denotes the box length, thus the decay length of the Gaussians is 0.1. The electrostatic potential is decomposed into a rapidly decaying short range contribution and a long wavelength long ranged contribution, which is calculated in Fourier space. This illustration was made using ESPResSo, where the P^3M, thus a mesh–based variant of the Ewald summation was used for the calculation.

application of the FFT, operations in Fourier Space and application of the inverse FFT, the electrostatic potential or electric field need to be evaluated at the continuous particle positions. This is done by a interpolation. The choice of functions for charge assignment and interpolation is one key ingredient. In the P^3M method cardinal B-splines are chose, while in other variants different functions are chosen. The favourable scaling of the FFT transfers to the mesh-Ewald algorithms. Given n charged particles the algorithm scales as $n \log n$.

One drawback of the Ewald summation and its descendants is the fact that periodic boundary conditions are automatically applied in all spacial directions. When simulating planar systems, it is, however, desired to have periodic images only in two directions. An alternative approach was for 2D periodic systems was developed by Arnold and Holm, the MMM2D algorithm [167–169]. We give a brief overview here, as its extension to flat dielectric and conducting

interfaces is discussed in the following chapter.

We assume a system that is periodic in x and y direction and not periodic in z direction. We perform a Fourier transformation of the Poisson equation only in x and y direction, denoting the x-wave number by p and the y-wave number by q. Then the Poisson equation reads as

$$\left(-p^2 - q^2 + \frac{\partial^2}{\partial z^2}\right) \Phi = -4\pi\rho\left(p, q, z\right). \tag{3.84}$$

Thus an ordinary differential equation in z obtained. For $p = q = 0$ it easily solved by

$$\Phi\left(0, 0, z\right) = 4\pi l_B \int \mathrm{d}z' \left|z - z'\right| \rho\left(0, 0, z\right). \tag{3.85}$$

In physical terms this can be translated as: For the lowest Fourier mode, the electrostatic potential can be calculated by assuming all charges to be smeared out over planes, and using the well-known solution of the potential around a charged plane. A solution for higher Fourier modes can be obtained by the well known solution of the PDE in Eq. 3.86:

$$\Phi\left(p, q, z\right) = \frac{2\pi l_B}{\sqrt{p^2 + q^2}} \int \mathrm{d}z' \exp\left(-\sqrt{p^2 + q^2}\left|z - z'\right|\right) \rho\left(p, q, z'\right). \tag{3.86}$$

The sum that appears in the back–transformation into real space has favorable convergence properties only if the distance between charges in the distance in z-direction is large. In the simulation box in the MMM2D algorithm is subdivided into slabs in the z-direction. Eq. 3.86 is used to calculate the interaction between non-adjacent slabs. For neighbouring slabs, a different expression is applied: the MMM2D near field formula. This formula has the complementary property of being well convergent for if distances in z direction are small.

3.3.4 Hydrodynamics

Hydrodynamic effects can be incorporated into simulations in different ways. The first algorithm was suggested by Ermak and McCammon [130]. Its underlying dynamics is an overdamped Langevin equation. In every step displacements for all particles are calculated. The underlying physical assumption is, that the correlation time of velocities are much smaller than the time step. Mechanical forces lead to deterministic displacements and random forces lead to

random displacements. Apart from the force calculation two time-consuming steps must be mentioned: (a) calculating the diffusion matrix according to Eq. 3.53 or a different approximation (b) a Cholesky decomposition of the diffusion matrix. The second step is necessary to obtain correlated displacements with a correlation matrix related to the diffusion tensor,

$$\Delta x_i \Delta x_j = 2D_{ij}\Delta t. \tag{3.87}$$

The Cholesky decomposition scales like $\mathcal{O}\left(n^3\right)$, and originally was the limiting factor. Alternatives of $\mathcal{O}\left(n^{2.5}\right)$ [170] and recently even of $\mathcal{O}\left(n^2\right)$ [171] have been proposed.

The Ermak-McCammon scheme was called *Brownian Dynamics*, as purely Brownian, diffusive processes could be studied. Mainly in the group of Brady, it has been generalized to situations including inertia, who coined the term *Stokesian Dynamics* [129]. The further development included an Ewald-Type summation for the diffusion tensor, and the inclusion of particle rotation and shear flow. In addition they added short range corrections of the diffusion matrix, *lubrication corrections*. Further improvements were the acceleration of the algorithm by using Fast Fourier transformations, with a method in the spirit of the $\mathrm{P^3M}$ algorithm [172]. This brings the computational cost down to $\mathcal{O}\left(n \log n\right)$.

In the last two decades, several alternatives, sometimes called *mesoscopic* methods, have emerged. The idea here is that the solvent is replaced by a medium that resembles a Newtonian fluid "well enough" without being as costly as e.g. an explicit water model. One example is the dissipative particle dynamics (DPD) method. Originally suggested by Hoogerbrugge and Koelman [173] it features the idea of combining soft particles with a momentum conserving thermostat [174]. The general idea is to thermalize particle pairs, rather than particles, and construct their dissipative interaction so that momentum conservation is obeyed. Asymptotically an isothermal Newtonian fluid is obtained. From molecular dynamics methods, it inherits the $\mathcal{O}\left(n\right)$ scaling, and a well defined temperature is obtained automatically. Its flexibility has been enhanced by generalizing the dissipative particle interaction [175], and also the connection to the smooth particle hydrodynamics method has been made [176]. The Lattice Boltzmann Method, another method of this class, is described in detail in Sec. 5.2.

3.3.5 The software package ESPResSo

When one aims at simulating coarse-grained models, it is clear that there is not a unique typical simulation situation. For example, when simulating an ideal chain (without excluded volume interactions) with harmonic bonds, a decomposition of the computational domain is not helpful as no non-bonded interactions have to be calculated. However, when an explicit solvent model is considered, short-ranged non-bonded interactions appear where a domain decomposition is much more efficient. In atomistic biomolecular simulations the situation is different. Here the density of atoms and the short range interaction is always similar. In this case, performance questions are dominated by how the water models are most efficiently simulated. The necessary flexibility is much lower. When the decision was made by the Holm Group in 1998 to develop an own code for simulations two main ideas were followed [177]. First, no other available code offered the flexibility necessary. Second, the usage of a program language (TCL) as control mechanism for the simulation provides more flexibility and extensibility than any other available software.

The second point is most important from a user's point of view. Typically, lengthy input files need to be written to control what a simulation code does. In these input files, parameters that control the simulation are set. In ESPResSo the user writes a TCL program, where the TCL idiom is extended with commands that perform MD-specific task, like adjusting the system or switching algorithms on and off. The same approach is used in the visualization software VMD [31]. This variant is much more powerful as it allows the user to explicitly formulate the logic of his or her computational task. As an example, we would like to show the TCL script that was used to calculate the electrostatic potential for the system with two charges in Fig. 3.7:

```
1  setmd box_l 1. 1. 1.
2  setmd time_step 0.01
3  setmd skin 0.01
4  thermostat off
5
6  set dist 0.4
7  set alpha 10.
8
9  part 0 pos [ expr 0.5-$dist/2 -$d / 2 ] 0 0 q 1
10 part 1 pos [ expr 0.5+$dist/2 +$d / 2 ] 0 0 q -1
11
12 inter coulomb 1. p3m 0.25 16 4 $alpha
13
14 set fs [ list ]
15 set q_test 0.0001
16
```

```
17  set n 1000
18  set d [ expr 1./$n ]
19
20  set p 0
21  for { set i 0 } { $i < $n } { incr i } {
22    set p [ expr $p + $d ]
23    part 2 pos [ expr $p ] 0 0 q $q_test
24    integrate 0
25    set f [ expr [ lindex [ part 2 print f ] 0 ]
    / $q_test ]
26    lappend fs $f
27  }
```

Here is a brief overview over the program. In the first lines, the system geometry and basic MD parameters necessary for the program to run are set up. In line 9 and 10, the two source particles are placed and their charge is assigned. In line 12, the P^3M algorithm is switched on with the desired parameters. Finally, in the loop following line 21, a test charge is moved through the system to measure the electrostatic field. The logic of this task is not one that fits into typical MD procedures. After every force calculation step the system is modified by an external influence: The user wants the test charge to be displaced. The equations of motion are not integrated at all. As the user controls the structure of the process and not only single parameters, the above task can easily be solved using ESPResSo.

The simulation package ESPResSo is actively developed by a group of core developers who is largely located in Stuttgart. SK has been one of the core developers during the preparation of this thesis, and has made major contributions to several parts. All contributions are integral part of the 3.1 version of ESPResSo [178] or prepared for later versions. We hereby give a brief overview

- The ICC$\star$ algorithm which is applied in Chapter 4 was improved. Initially a purely serial version was available. The implementation was carefully checked, and parallelized for execution on distributed memory computers. It was extended for the use with homogeneous external fields. The implementation is explained in detail below.

- A module for the efficient calculation of correlation–type quantities was implemented. It is a generic framework to calculate quantities that depend on the state of the simulation at two different times, like correlations and mean-square-displacements (MSD). It can be considered a special case of the algorithm presented in [101], where the lag time is not equidistantly spaced, but where larger lag times can have larger

time spacings. This is especially suited for correlation–type data as they often have features distributed over several orders of magnitude in time.

- The present implementation of the Lattice-Boltzmann algorithm was carefully checked and the code was cleaned up. It was extended for the use of arbitrarily shaped boundaries using the application of the link bounce back method, including boundaries with nonzero velocity. The user interface was improved. This included the calculation of the pressure tensor field and functions to calculate the average fluid field, also in a cylindrical coordinate system.
- An implementation of tabulated external potentials. Based on a potential energy map given on a regular grid the forces particle forces are calculated. The potential was used to include the result of a finite element calculation of the electrostatic potential into an MD simulation.

3.4 Continuum theories of charged soft matter systems

In this section we briefly introduce an approach in soft matter science where the particle concept is abandoned in favour of a continuum descriptions. In these approach the particle entropy is automatically contained and only the averages of fields are considered. Therefore, continuum approaches allow for a very efficient modelling, if they are applicable.

3.4.1 Ions in equilibrium: Poisson–Boltzmann theory

The interaction of charged objects in aqueous solution is modified not only by the polarizability of the surrounding water, but also by present ions. Even in completely desalinated water the autoprotolysis of water molecules leads to a concentration of H_3O^+ and OH^- ions of at least 10^{-7} mol/l. Ions which are free to move, will accumulate near objects bearing a charge of the opposite sign. This reduces the magnitude of electric fields and the field created by a charged object is negligible if it is enclosed in a sufficiently large electrolyte compartment. In the presence of large electrolyte concentrations this screening effect is so strong that the long range-nature of electrostatics often can be entirely neglected. The distribution of ions around charged objects is governed by the interplay of energy, i.e. the attraction of counterions, and entropy, the tendency of ions to distribute homogeneously. The theory proposed by Debye and Hückel published in 1923 [179] describes this interplay exactly, in a certain limit that will be discussed below. It is generalized by a theory that, depending on the scientific context, is called Poisson-Boltzmann theory or Gouy-Chapman theory. We will present a short derivation in the following lines, and apply the name of Poisson-Boltzmann theory throughout this work.

We reformulate here the derivation of Russel et al. [116]. For simplicity we assume the presence of only two univalent ion species + and −. In the PB theory the chemical potential of both ion species is assumed to be

$$\mu_\pm = \log c_\pm/c_0 \pm \Phi. \tag{3.88}$$

Her c denotes the local ion concentration of both species. The chemical potential μ and the electrostatic potential Φ are measured in units of the thermal energy k_BT and in k_BT per elementary charge. Eq. 3.88 corresponds to the

ideal gas chemical potential in the reduced external potential Φ. The reference state is chosen in a hypothetical reservoir, where the zero point of the electrostatic potential is fixed. The reservoir is characterized by its concentration c_0. In equilibrium, the free energy is minimal if the chemical potential is constant in space. With this assumption a Boltzmann distribution of the ions in the electrostatic potential is obtained:

$$c^{\pm} = c_0 \exp \mp \Phi. \tag{3.89}$$

To derive the PB equation, the mean electrostatic potential created by the charge densities c_+ and c_i is inserted as the external potential Φ into this equation. This Boltzmann distribution is inserted inserted into the Poisson equation, leading to

$$\begin{aligned} \Delta\Phi &= -4\pi l_B \left[c_0 \exp\left(-\Phi\right) - c_0 \exp\left(\Phi\right)\right] \\ &= 8\pi l_B c_0 \sinh \Phi. \end{aligned} \tag{3.90}$$

The prefactor $8\pi l_B c_0$ is identified as the inverse square of a characteristic length: the Debye length $l_D = \sqrt{1/8\pi l_B c_0}$. For low surface potentials solutions of the form $\exp\left(-d/l_D\right)$ are obtained, where d is the distance to a charged object. Thus the Debye length is the characteristic length scale on which the electrostatic potential decays.

At two points approximations were made to construct this form of the PB equation: First, the chemical potential was constructed based on the ideal gas chemical potential. Second, the mean electrostatic potential was assumed to be acting on the ions. Both assumptions are quite severe, however the PB equation is very successful at describing ion behaviour in univalent salt under ambient conditions in water (see e.g. [180]).

The free energy of a system in the Poisson-Boltzmann theory can be calculated from the following expression [181, 182]:

$$F = \int_V \frac{1}{2} \boldsymbol{E} \cdot \boldsymbol{E} + c_0 \left(\phi \sinh \phi - \cosh \phi + 1\right) \tag{3.91}$$

This expression is very convenient as it can be used to directly calculate the free energy difference between two states. It will be used in Sec. 4.4.4 to calculate the free energy difference between two positions of a model DNA molecule.

3.4.2 The electrokinetic equations

The phenomenon of electrophoresis seems not very spectacular. When a charged object is placed in an aqueous phase and exposed to an electric field, it starts to accelerate until it reaches a stationary state. However, a remarkable difference to the application of a mechanical force is apparent when considering the counterion cloud of the object. It is charged oppositely and thus accelerates the solvent in the opposite direction. When one considers the total force that is applied onto a region large enough so that the counterion cloud is fully inside, it must be zero. Therefore, also the momentum flux out of the region must be zero in the stationary state. Noting that the Navier-Stokes equation is a conservation equation for the momentum it appears obvious that e.g. the resulting flow field must be quite different.

In rest of this section we present a brief account of the electrokinetic equations, which a set of partial differential equations sometime called the "standard model of electrokinetics" [183–185]. It can be considered a generalization of the PB equation for out-of-equilibrium situations. In equilibrium, the PB equations are recovered. In Sec. 3.4 of this thesis, they will be used to construct a continuum model for charge transport in nanopores, and to our surprise this simple continuum model yields very good agreement with our atomistic simulation even on the sub-nanometer scale. It is formulated for two univalent ion species + and − to keep the formulas simple but expressive.

In the standard model of electrokinetics the ions follow a convection-diffusion equation. Their number flux density $\boldsymbol{j}_\pm$ thus obeys

$$\boldsymbol{j}_\pm = -D\left(\nabla c_\pm \pm c_\pm \nabla\Phi\right) + c_\pm \boldsymbol{u}. \tag{3.92}$$

Here, D is the diffusion coefficient and $\boldsymbol{u}$ the solvent velocity. We have assumed D to be identical for both ion species. The solvent is described by a Stokes equation that is driven by a gradient in the hydrodynamic pressure p, the osmotic pressure p_{osm} and by the electrostatic potential that exerts a force density on the solvent. Then the Stokes equation reads as follows:

$$\eta\Delta u = -\nabla p - \nabla p_{\mathrm{osm}} + \left(c_+ - c_-\right)\nabla\Phi. \tag{3.93}$$

Here η denotes the dynamic viscosity of the solvent. The osmotic pressure in our unit system is just the sum of the two concentrations: $p_{\mathrm{osm}} = c_+ + c_-$. It can be noted that for the solution of the PB equation the osmotic pressure and the electrostatic force exactly cancel. Therefore the solvent is not accelerated in this case. Finally the electrostatic potential Φ is described by the Poisson

equation

$$\Delta\Phi = -4\pi l_B \left(c_+ - c_-\right). \tag{3.94}$$

The equations are closed by continuity equations which, in a stationary state, read as

$$\nabla \cdot \boldsymbol{j}_\pm = 0, \qquad \nabla \cdot \boldsymbol{u} = 0 \tag{3.95}$$

The whole set of equations thus consists of 13 nonlinear first order (Diffusion-related) and second order (Poisson and Stokes) partial differential equations. Directly applying the continuity equation for the ion numbers leaves 7 PDEs of second order.

An interesting observation can be made: When the ion velocities $\boldsymbol{v}_\pm = \boldsymbol{j}_\pm / c_\pm$ are introduced, the following form of the Stokes equation can immediately obtained:

$$\eta\Delta u = -\nabla p - \frac{1}{D}\left(c_+ \left(\boldsymbol{v}_+ - \boldsymbol{u}\right) + c_- \left(\boldsymbol{v}_+ - \boldsymbol{u}\right)\right). \tag{3.96}$$

The physical interpretation of this form is the appearance of a drag force between ions ans solvent:

$$\boldsymbol{f}_{\text{drag}} = \gamma c \left(\boldsymbol{v} - \boldsymbol{u}\right), \tag{3.97}$$

with a drag coefficient $\gamma = 1/D$. Note that the usually appearing factor $k_B T$ is absorbed into our definition of the energy scale.

The diffusion-convection equation is based on the free energy of an ideal gas. If the fluxes and the flow field are zero, the Boltzmann distribution of an ideal gas in an external potential is recovered. Therefore, in the case of vanishing ion fluxes, the Poisson-Boltzmann equation is recovered as a special case of the electrokinetic equations. The reservoir concentration then is given explicitly as a boundary condition.

Under certain conditions, the effect of solvent motion on the ions is negligible. Then the convective term in Eq. 3.92 can be dropped. The resulting set of diffusion equations and the Poisson equation is known as the Poisson-Nernst-Planck model of ion transport. Formally they can e.g. be obtained as the limit of infinite viscosity. These equations are used for the description e.g. of ion migration through biological pore [186]. In the nanopore community, the have e.g. been applied to understand the current rectification properties of nanopores with an asymmetric charge distribution [74].

It is very common to simplify the electrokinetic equations by linearization. The means, the solution is expanded around the solution of the system which

is not driven by external forces, the solution of the PB equation. Formally this corresponds to a first order perturbation theory [187]. In case of the electrokinetic equations, the driving force is included in the boundary conditions (BCs). To perform the perturbation expansion, the solution to the electrokinetic equation is written as power series in ϵ:

$$\Phi = \Phi^0 + \epsilon\Phi^1 + \mathcal{O}\left(\epsilon^2\right), \tag{3.98}$$

$$c_\pm = c_\pm^0 + \epsilon c_\pm^1 + \mathcal{O}\left(\epsilon^2\right), \tag{3.99}$$

$$\boldsymbol{u} = \boldsymbol{u}^0 + \epsilon\boldsymbol{u}^1 + \mathcal{O}\left(\epsilon^2\right). \tag{3.100}$$

$$\tag{3.101}$$

As the unperturbed state (0) is the undriven case $V(r) = 0$, the solution can be obtained from the PB equation. Then, the force density on the fluid is zero, and hence $\boldsymbol{u}^0 =$. By inserting these formal expansions into the electrokinetic equation, we can derive equations for Φ^n, $c_\pm^n$ and $\boldsymbol{u}^n$. We the expressions by order of ϵ, and on every order of ϵ the equations must be valid independently. This way, we construct a hierarchy of equations, each linear in the highest appearing order of ϵ. The equations corresponding to first order of ϵ are given here:

$$\begin{aligned}
\Delta\Phi^1 &= -4\pi l_B\left(c_+^1 - c_-^1\right), \\
0 &= \nabla\cdot\left(-D\nabla c_\pm^1 \mp D\nabla c_\pm^1\nabla\Phi_0 \mp D\nabla c_\pm^0\nabla\Phi^1 + c_\pm^0\boldsymbol{u}^1\right), \\
\eta\Delta u^1 &= -\nabla p^1 + \left(c_+^1 - c_-^1\right)\nabla\Phi^0 + \left(c_+^0 - c_-^0\right)\nabla\Phi^1.
\end{aligned} \tag{3.102}$$

Solving this linear set of equations is significantly easier than solving the full problem. For many physical questions, the first order expansion is sufficient. Other effects, for example, induced charge electroosmotic flow [188, 189] or ion current rectification in nanopores [74], are second order effects, where the system of equations in second order of ϵ must be solved.

The electrokinetic equations are a simple and well-understood continuum model, where many effects can be modelled accurately, at least in the limit of low surface potentials and low salt concentrations, where the theory is exact. However, the lack of knowledge of the right boundary conditions is in practice a severe limitation. In Chapter 5, we construct a model of nanopore conduction from this model. There, DNA is included as a charged cylinder. By comparison to atomistic simulations, we can show that this assumption works surprisingly well. However, we find that a surface friction effect is very important for understanding the conduction modification caused by DNA. This effect is conceptually not in included in the electrokinetic equations. Identify-

ing this effect is, however, only possible, because the electrokinetic equations describe the general behaviour well, and the system simulated with atomistic resolution deviates only in a certain region close to the DNA.

4 Translocation barriers

In this chapter we present our approach to one of two main questions we investigated in course of the preparation of this thesis: Is there a free energy barrier for DNA that opposes DNA molecules from entering a nanopore? Free energy barriers greatly influence the expected rate of translocation and the ratio of successful to unsuccessful translocation processes. A thorough understanding of them therefore is an important ingredient to understanding translocation dynamics.

We investigated free energy barriers with special emphasis on the role of electrostatics. DNA is a highly charged molecule, and potentially interacts directly and through its surrounding counterion cloud with obstacles and boundaries. We applied two different modelling approaches, a coarse-grained molecular dynamics simulation approach and an approach based on the Poisson-Boltzmann equation. One of the challenges of this chapter lied in obtaining compatible results with both methods.

Electrostatic barriers can, on the one hand, be present, if the pore is smaller than the extension of the counterion cloud of the molecular. In this case, the Debye layer must deform, which is associated with a free energy cost. On the other hand, typical membrane materials have a dielectric constant significantly smaller than water. This dielectric contrast gives rise to *dielectric boundary forces* (DBFs) which potentially increase a free energy barrier. In most investigations, these DBFs have so far been neglected. This is probably due to the fact that they are computationally difficult to take into account correctly. In coarse-grained MD simulations they could, until recently, only be taken into account with method of computational complexity $\mathcal{O}\left(n^2\right)$ [190,191] if surfaces of arbitrary shape were to be considered. Recently, an algorithm proposed by S. Tyagi [3] and developed in the group of C. Holm has made it possible to take these effects into account more efficiently, the ICC$\star$ algorithm. In the Poisson-Boltzmann equation, typically the dielectric constant is not taken explicitly into account, but only in terms of boundary conditions corresponding to a infinitely low or infinitely high dielectric constant inside the boundary. We were able to take the dielectric constant of the pore material explicitly

into account using the iPBS algorithm. The idea is very similar to the ICC$\star$ algorithm, but not only the induced charges, but also the boundary conditions to the PB equation are iterated. The PB equation itself is solved by means of a Finite Element Method. Originally it has also been proposed by S. Tyagi, and has been re-implemented using the DUNE framework by A. Schlaich in course of the work of his diploma thesis [14], to which SK was an advisor.

The project of investigating translocation barriers for short fragments of ds-DNA with coarse-grained MD simulations had already been started, as a part of the PhD thesis of M. Süzen [13]. However only first example studies had been performed, indicating that under certain conditions the DBFs strongly increase electrostatic free energy barriers. We extended theses studies to a broad range of conditions, pore sizes and obtained a good understanding of the expected barriers. We also developed a method to calculate the free energy barrier of a highly flexible polymer resembling a longer piece of single-stranded DNA. We could show that in this scenario, due to the flexibility of the DNA molecule, electrostatic effects, especially dielectric effects are even more important.

This chapter is organized as follows. We first give a textbook level introduction to the physics of dielectric media. Then we present algorithmic approaches to take dielectric interfaces into account in coarse grained MD simulations. This includes a description of the ICC$\star$ algorithm, but also a brief account of the ICMMM2D method [192]. In the following section we show results for a simple setting, which already contains many aspects of the subsequently applied model for DNA, the distribution of ions dissolved in water in the vicinity of a dielectric interface. Then we investigate the role of DBFs for DNA translocation: For models of double-stranded DNA and single-stranded DNA the occurring translocation barriers are assessed. In the last section we present the results we obtained from the Poisson-Boltzmann theory, and compare them to the results of the MD simulations. Finally, we briefly summarize the finding of this chapter.

The content of this chapter was published in in total four articles [18,178,193, 194]. In the first three of these articles SK was the main author, performed all simulations, suggested the structure and content of the articles and prepared the figures. Several parts of these articles are reproduced literally in this thesis. Other parts are paraphrased, or otherwise adapted. To the fourth article different authors contributed significantly. Here only data produced by simulations of SK are displayed, and most text parts have completely been rewritten. Only Sec. 4.1.4 is reproduced almost verbally, as it fulfills the same standards of being originally authored by SK as the other three articles. The

artwork presented from this article was partly performed by F. Fahrenberger based on drafts by SK. His support is highly appreciated. The reproduction of the content in this thesis is in agreement with the coauthors of the articles. The content of Sec. 4.4 is currently in preparation for publication.

4.1 Polarization effects in simulations

Implicit solvent models are largely based on a continuum approximation of the solvent. One strategy is describing the solvent as a dielectric continuum, an approach pioneered by Born [195]. Reducing a solvent to its dielectric properties is very appealing as the number of degrees of freedom necessary to consider can be significantly lowered. It is hence a coarse-graining strategy. In this section we outline the theory of continuum electrostatics, and how its equations can be used to approximate the forces acting on charged particles near interfaces. This leads to the formulation of two algorithms, the ICC$\star$ algorithm (Induced Charge Calculation, the star can stand for different Coulomb solvers), and the Image Charge MMM2D (ICMMM2D) algorithm. we apply the ICC$\star$ algorithm in the next sections to the calculation of ion distributions near interfaces and free energy barriers DNA molecules have to overcome when crossing a dielectric pore.

Standard textbooks on electrostatics [112, 196–199] and on the theory of the liquid state [111, 200] are the main sources of the following section. No single source is strictly followed, as the theory is relatively common. We apply SI units in this section, as the reduced units require the concept of a dielectric constant.

While preparing this thesis also a different system not directly related to nanopores and DNA was brought to the author's attention: Ionic-liquid based supercapacitors with graphene electrodes [201–204]. In this system ions approach the interface between vacuum and graphene, a material with a high electric susceptibility, which can, potentially, be approximated as a conducting medium. It appeared natural to investigate if and how the considerations for dielectric media can be extended to conducting media. This will be discussed in Sec. 4.1.4.

4.1.1 Electrostatics in polarizable media

When electric fields are applied to a material, the charge distribution in the material changes: it becomes is *polarized*. The most common occurrence of this effect in every day life is the refraction of light at interfaces with different dielectric constants. The deflection of the electromagnetic rays is caused by the difference in polarizability of the two media. In this work we focus on how polarization affects the role of electrostatic effects in different soft matter systems. Thus we are interested its impact on static electric fields rather than on electromagnetic waves.

The microscopic origin of the dielectric effect is a distortion of the charge distribution in the material when an electric field is applied, the polarization effect. In water the dielectric effect is very large. In an electric field the water dipoles align antiparallel to the field which cancels 79/80 of the external field and, of course, dramatically changes the role and strength of electrostatic interactions. Therefore this effect needs to be considered in every charged system.

Particles with a nonzero charge density embedded in dielectric media polarize the surrounding medium. Molecules for which the charge distribution is relevant are called *polar*, and the highest non-vanishing multipole moment is used to characterize them e.g. as dipolar, or quadrupolar. The term "monopolar" is uncommon, and ions are referred to as being charged. In the Born model of solvation, the free energy of solvation of a molecule is estimated from the energy difference between a state where the molecule is in vacuum, and where the molecule is embedded in a spherical cavity in a homogeneous dielectric medium with the same bulk dielectric constant as the solvent [195]. Following this line of thought, near the interface of two dielectric media forces must be expected that drag polar objects towards the phase with the higher dielectric constant. Polar objects immersed in water thus would be mostly repelled from interfaces since water is better polarizable than most other media. As the forces acting on molecules close to interfaces are of electrostatic origin they can be long ranged. Their far field can be approximated in terms of continuum electrostatic theory. We now systematically develop the necessary theory.

When applying an electric field $\boldsymbol{E}$ to an isolated molecule, its electronic structure, and also the position of its nuclei is distorted, it is polarized. For small applied fields, this effect is linear in the field strength. The perturbation of the charge density, ρ_m, can be expanded in multipoles, and the term connecting the dipole moment of the perturbed charge density to the applied field is

defined as the polarizability of an isolated molecule α:

$$\alpha = \lim_{E \to 0} \frac{1}{E} \int \mathrm{d}^3 \boldsymbol{r} \rho_m(\boldsymbol{r}) \boldsymbol{r}.$$

In general α is a symmetric tensor of rank two, as the polarizability is not necessarily isotropic. Recently, a lot of effort has been made to devise interactions in molecular systems that incorporate electronic polarizability, without explicitly considering the quantum mechanical structure of the electrons in a molecule. From these polarizable force fields (see e.g. [205, 206]) a significant improvement of the accuracy of molecular simulations is expected.

In a dense medium, the collective response can be expressed in terms of a dipole density $\boldsymbol{P}$. The proportionality constant between $\boldsymbol{P}$ and the electric field, is called the dielectric susceptibility χ, which is tensorial as well. Finding a general relationship between the polarizability of atoms or molecules and the susceptibility is a nontrivial issue. Mean-field theories can be constructed, analogous to the construction of the macroscopic magnetization from the density of magnetic moments in a paramagnet, leading to the Clausius-Mossotti relationship [207]. In general, however, the collective response can be qualitatively different from the response of individual molecules due to strong mutual interaction and material anisotropy. This can lead to interesting phenomena such as ferroelectricity and the piezoelectric effect. It is important to note that the timescale on which the electron shell responds is below a femtosecond and hence can be considered instantaneous for soft matter applications.

In liquid phases, also the orientation of polar molecules needs to be considered: In water for example, the orientational polarization is more than a factor of 50 stronger than electronic polarizability. The dipole density in a medium with a number density n of molecules with a dipole moment D is given by

$$\boldsymbol{P} = nD\langle \cos\theta \rangle \hat{\boldsymbol{E}},$$

where θ is the angle between the dipole vector of a molecule, and the applied field, and $\hat{\boldsymbol{E}}$ is a unit vector pointing in the direction of the applied electric field. Note that, due to the isotropy of the liquid phase, no off-diagonal terms in the response can appear. The susceptibility must be a scalar.

At this point it becomes obvious that the theory of polarization is in general nonlocal. Expressing e.g. the potential energy of a molecule in an electric field by $\boldsymbol{P} \cdot \boldsymbol{E}$ is only a good approximation if $\boldsymbol{E}$ is approximately constant on the length scale of the molecular extension. If this is not the case, terms of higher order must be considered. The same is valid for the response: Applied

external fields not only create local dipole moments, but also multipoles of higher order that need to be considered. The non-locality can e.g. considered by means of integral equation theories [208, 209]. This lies, however, beyond the scope of this thesis.

If only the response on large length scales is of interest, it is sufficient to consider it in terms of the dipole density, as only dipole fields are long-ranged. In this limit the response of an isotropic medium can be expressed solely by its dielectric susceptibility χ, which relates the mean local electric field $\boldsymbol{E}$ to the mean dipole density $\boldsymbol{P}$:

$$\boldsymbol{P}(\boldsymbol{r}) = \chi \boldsymbol{E}(\boldsymbol{r}). \tag{4.1}$$

The potential created by the dipole density $\boldsymbol{P}(\boldsymbol{r})$ can be expressed as:

$$\Phi_P = \frac{1}{4\pi\varepsilon_0} \int \mathrm{d}^3\boldsymbol{r}' \, \frac{\boldsymbol{P}(\boldsymbol{r}') \cdot (\boldsymbol{r} - \boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|^3}.$$

In two simple steps this can be rewritten as:

$$\begin{aligned} \Phi_P &= -\frac{1}{4\pi\varepsilon_0} \int \mathrm{d}^3\boldsymbol{r}' \, \boldsymbol{P}(\boldsymbol{r}') \cdot \nabla_r \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \\ &= -\frac{1}{4\pi\varepsilon_0} \int \mathrm{d}^3\boldsymbol{r}' \, \nabla_{r'} \cdot \boldsymbol{P}(\boldsymbol{r}') \, \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|}. \end{aligned} \tag{4.2}$$

This expression is has the same form as the potential created by a charge density $-\nabla \cdot \boldsymbol{P}$. Therefore $-\nabla \cdot \boldsymbol{P}$ is called the *induced charge density* and denoted by ρ_{ind}. By taking the divergence of Eq. 4.1 we can derive an expression for it:

$$\rho_{\mathrm{ind}} = -\boldsymbol{E} \cdot \nabla\chi - \chi \nabla \cdot \boldsymbol{E}. \tag{4.3}$$

The induced charge density thus is nonzero only where the susceptibility changes or where the electric field has sources.

To derive the Poisson equation in a dielectric medium we assume a density ρ_f of fixed (contrary to induced) charges. Sources of the electric field $\boldsymbol{E}$ are the fixed and the induced charges:

$$\nabla \cdot \boldsymbol{E} = \frac{1}{\varepsilon_0} (\rho_f + \rho_{\mathrm{ind}}). \tag{4.4}$$

By inserting Eq. 4.3, it can be rewritten as:

$$\nabla \cdot \left(1 + \frac{\chi}{\varepsilon_0}\right) \boldsymbol{E} = \rho_f / \varepsilon_0. \tag{4.5}$$

The relative permittivity ε_r is defined as $\varepsilon_r = (1 + \chi)$. With this definition we we can write down the Poisson equation in a dielectric medium:

$$\nabla \cdot (\varepsilon_r \nabla \Phi) = -4\pi \rho_f / \varepsilon_0. \tag{4.6}$$

An important point can be noticed. If a homogeneous relative permittivity ε_r is assumed, it can be brought to the right hand side, and appears as a factor $1/\varepsilon_r$ in front of the fixed charge density. This means, all charges appear scaled down by this factor, and hence, if the polarizability of a medium is large, electrostatics is significally weakened.

We have now laid out the necessary equations to present the ICC$\star$ algorithm. Before going on, we introduce a concept that is very useful for the interpretation of the role of dielectrics, the concept of image charges. With this concept we can discuss qualitative aspects of the resulting forces.

4.1.2 The image charge method

A very simple though mathematically entirely correct picture of dielectric effects can be drawn using the concept of image charges. This technique is standard in textbooks considering dielectric interfaces. We develop it for planar interfaces, although it is also suited for other interfaces if they are sufficiently symmetric, e.g. dielectric spheres [210–212]. The key idea behind the picture of image charges is determining a charge distribution inside the dielectric that creates the same electrostatic field outside as obtained from a full solution of the Poisson equation including a dielectric interface. Assuming a dielectric interface at $x = 0$ between ε_1 in the right half–space $x > 0$ and ε_2 in the left half–space $x < 0$ the solution of the full Poisson equation is recovered by assuming an image charge density

$$\rho_{\mathrm{im}}(x, y, z) = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \rho_f(-x, y, z). \tag{4.7}$$

in the left half–space, and assuming a homogeneous permittivity ε_1. Two limits can be identified. If $\varepsilon_1 \ll \varepsilon_2$ the image charge distribution is the negative of the original charge distribution. This limit corresponds to conducting

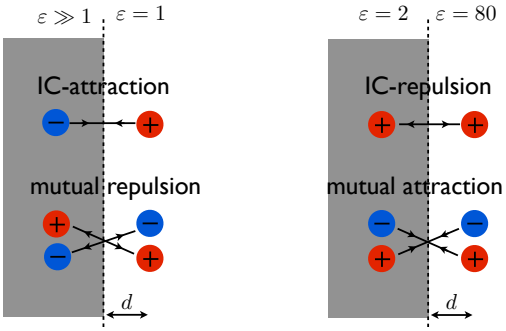


Figure 4.1: Image charges of a single charge (upper half) and a pair of charges (lower half). We consider the limits $\varepsilon_1 \ll \varepsilon_2$ (left) and $\varepsilon_1 \gg \varepsilon_2$ (right). Individual charges are attracted to interfaces with a medium with a higher dielectric constant, and repelled by interfaces with a medium with a lower dielectric constant. Pairs of charges attract each other less in front of a wall of high permittivity, and their mutual attraction is increased in front of boundaries with media with a lower permittivity.

medium and yields a constant electrostatic potential on the dielectric interface. The opposite limit $\varepsilon_1 \gg \varepsilon_2$ corresponds to a situation often encountered in aqueous systems: The dielectric constant of water at ambient conditions is around 80, which is much larger than most other materials. The image charge distribution ρ_{im} is identical to the charge distribution in the water phase with the same sign. In Fig. 4.1 we qualitatively depict the influence of a dielectric interface of both types on a single point charge and a pair of oppositely charged point charges.

For a single elementary charge the image charge picture allows us to deduce potential energy V (in units of the thermal energy) of a single elementary charge as a function of the distance to the dielectric interface d:

$$V(d) = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \frac{l_B}{2d}.$$

We use reduced units again, and the Bjerrum length is defined in medium 1, in which the particle is immersed. It is important to note the sign of the potential: For $\varepsilon_2 > \varepsilon_1$ it is attractive; a charge is attracted by its own image, while for $\varepsilon_2 < \varepsilon_1$ it is repulsive. At maximum dielectric contrast, the image

charge is of equal magnitude as the real charge, and the interaction strength depends on the permittivity of medium 1. For a pair of particles of opposite sign the situation is as follows: In the case $\varepsilon_2 > \varepsilon_1$ each particle is repelled by the image of the other particle, which weakens the mutual attraction. For $\varepsilon_2 < \varepsilon_1$ they attract each other stronger as each particle is attracted by the image of the other one. The general character of these statements remain the same also for dielectric interfaces of other shapes.

We will repeat these two qualitative statements for $\varepsilon_2 < \varepsilon_1$ again in other words, as it is the very important for the qualitative understanding of our results. We assume that the medium in which the charges are immersed is water, and assume a boundary to a medium of significantly lower dielectric permittivity. First, charged objects are repelled by the boundary, if they are not neutralized by other charges. Second, the mutual attraction of unlike-charged object is enhanced in the vicinity of the boundary. This leads to a higher tendency of neutralizing charged objects.

If a second dielectric interface is present, the situation becomes more complicated: At every dielectric interface also the image charges due to all other interfaces need to be reflected. In the two limits of large dielectric contrast, this leads to an infinite series of image charges, which does not decrease with increasing number of reflections. The occurring sum is only conditionally convergent, and the interpretation of this sum lead to extensive discussions in our institute. We will present our interpretation in Sec. 4.1.4. For every finite value of ε_r no convergence problems arise. The magnitude of the image charges decays exponentially with the number of reflections. An efficient way to take these image charges into account is the summation method used in the ICMMM2D method [192], which is based on a Fourier decomposition of the electric field in the parallel direction.

We will now turn to the presentation of the ICC$\star$ algorithm that allows for dielectric interfaces of arbitrary shape.

4.1.3 The ICC$\star$ algorithm

We have seen above, that the formulation of electrostatics in dielectric media in terms of a polarization density and in terms of an induced charge density are equivalent. The idea of the ICC$\star$ algorithm is to determine the induced charge distribution in an iterative procedure. This is possible because the equation we derive for it below has a fix-point structure. A simple iterative scheme can be applied to determine the charges, and only the electric field at the

boundary is necessary for the iteration. For that reason, the iterative scheme can be used in conjunction with any Coulomb solver for point charges.

The Poisson equation given in Eq. 4.6 can be rewritten as:

$$\Delta\Phi = -\frac{1}{\varepsilon_r \varepsilon_0}\rho_f - \frac{1}{\varepsilon_r}\nabla\varepsilon_r \cdot \nabla\Phi. \tag{4.8}$$

The first term of the right-hand side of this equation contains the screened fixed charges, and the second term describes charge density induced in regions where the dielectric permittivity varies. We assume a discontinuous jump of the dielectric constant between two regions from ε_1 to ε_2.

In this case the induced charge density ρ_{ind} can be replaced by a induced surface charge density σ_{ind}. By evaluating the second source term in the equation above it can be identified as

$$\sigma_{\text{ind}} = 2\varepsilon_0 \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \boldsymbol{E} \cdot \boldsymbol{n}, \tag{4.9}$$

where $\boldsymbol{n}$ is the local normal vector on the surface. This equation couples the magnitude of an induced charge density to the local magnitude of the electric field. If hence the local field strength was known, the induced surface charge density could be determined from Eq. 4.9. As the local field strength, however, depends on the magnitude of the induced charge density everywhere else in the system, this is not possible. But Eq. 4.9 can be used to construct a surface charge density iteratively. From an initial guess of the induced charge density the electric field at the boundary can be calculated, leading to an improved estimate of the induced charge density. This procedure can be iterated until self-consistency is achieved. This is the idea of the ICC$\star$ algorithm.

To go into more detail, let us introduce the Green's function G for the Laplace operator. In an unbounded 3D geometry it is just $\dfrac{1}{4\pi\,|r - r'|}$, but it can also include, for example, periodic boundary conditions. The Laplace operator can be eliminated from Eq. 4.8 and we obtain

$$\Phi = \frac{1}{\varepsilon_0}\int_V G\left(\boldsymbol{r}, \boldsymbol{r}'\right) \frac{\rho_f\left(\boldsymbol{r}'\right)}{\varepsilon\left(\boldsymbol{r}'\right)}\,\mathrm{d}^3\boldsymbol{r}' + \frac{1}{\varepsilon_0}\int_A G\left(\boldsymbol{r}, \boldsymbol{r}'\right)\sigma_{\text{ind}}\left(\boldsymbol{r}'\right)\,\mathrm{d}A'. \tag{4.10}$$

The volume integral extends over the whole volume and the surface integral extends over all dielectric interfaces. The potential is now expressed in terms of the Green's function of a homogeneous dielectric, yet the induced charge

density σ_{ind} is still unknown. By taking the gradient and inserting this expression in the Eq. 4.9, we obtain the following integral equation:

$$\sigma_{\text{ind}} = -2\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2}\left(\int_V \nabla_{\boldsymbol{r}} G\left(\boldsymbol{r}, \boldsymbol{r}'\right)\frac{\rho\left(\boldsymbol{r}'\right)}{\varepsilon_1}\mathrm{d}^3\boldsymbol{r}' + \int_A \nabla_{\boldsymbol{r}} G\left(\boldsymbol{r}, \boldsymbol{r}'\right)\sigma\left(\boldsymbol{r}'\right)\mathrm{d}A'\right). \tag{4.11}$$

This result is easily generalized to multiple regions with different permittivities. We assume a discretized surface of m point charges on the dielectric interface the equation above for discretization point k can be written as

$$q_k = -A_k\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2}\boldsymbol{n}_k \cdot \left[\sum_{i=1}^{n}\frac{q_i}{\varepsilon\varepsilon_0}\nabla_{r_k} G\left(\boldsymbol{r}_k, \boldsymbol{r}_i\right) + \sum_{j=1, j\neq k}^{m}\frac{q_j}{\varepsilon_0}\nabla_{r_k} G\left(\boldsymbol{r}_k, \boldsymbol{r}_j\right)\right],$$

where A_k is the surface area and $\boldsymbol{n}_k$ the normal vector of the surface element k. The term in square brackets is just the electric field acting at the position of point k assuming a homogeneous dielectric constant ε_1 in the system, created by conventional (not induced) charges. Any standard Coulomb solver can thus be used to calculate the sum in the expression above. The desired solution of all q_k is the fix point of the following iteration:

$$q_k^{l+1} = \left(1 - \omega\right)q_k^l + \omega A_k\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2}\boldsymbol{n}_k \cdot \boldsymbol{E}\left(\left[q_i\right], \left[q_j^l\right]\right).$$

This iteration belongs to the class of Successive Overrelaxation (SOR) Schemes. It turned out that this iteration is very stable when performed for typical soft matter system. With a choice of $\omega \approx 0.7$ no stability issues occur. In every MD step only 1-3 iterations are necessary as the particle positions change only slightly.

An important advantage of this algorithm is, that the computationally most costly part, the evaluation of the electric field, can be done with *any* usual electrostatics solver without modifications. Thus not only the computational efficiency, but also the periodicity is inherited from the underlying Coulomb solver. The complexity of the algorithm remains unchanged by the presence of induced surface charges. However, the number of particles can increase considerably. We found it sufficient to discretize the surface with mutual particle distances equal to the distance of closest approach. For the system shown in section 4.2.2 this would mean in total 1600 surface charges per wall, compared to less than 100 ions in the system.

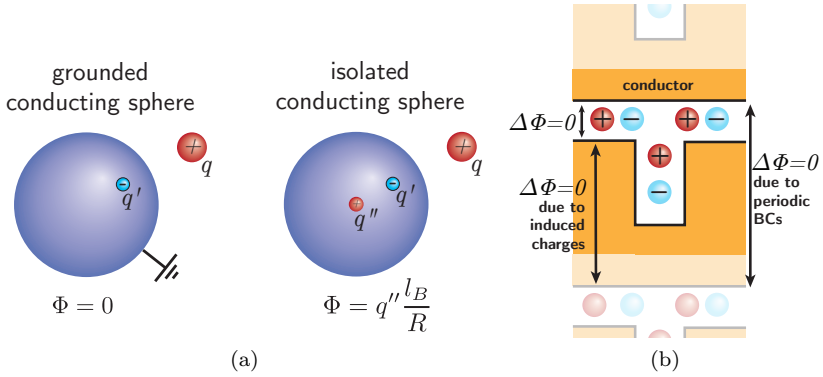


Figure 4.2: (a) Illustration of the dielectric boundary problem of single charge q outside of a grounded conducting sphere. The problem can be solved by assuming an image charge q' inside the sphere, leading to zero potential Φ on its surface. If the sphere is assumed to be conducting, but isolated, the excess charge q' has to be cancelled by adding a second charge q'' in the center of the sphere, which leads to a constant surface potential Φ. (adapted from [196]). (b) A more complex geometry with an upper and lower electrode (yellow). The electrodes are treated with the ICC$\star$ algorithm. If a Coulomb solver with periodic boundary conditions (BCs) in vertical direction is applied, the potential difference between both electrodes is automatically zero. This is because the periodicity yields zero potential difference between an electrode and its periodic image and the ICC$\star$ algorithm ensures that the two electrodes connected over periodic BCs are on equal potential.

4.1.4 Conducting interfaces

Metallic boundary conditions are the $\varepsilon_r \to \infty$ limit of Equations 4.6, where ε_r denotes the dielectric constant of the surrounding medium. The corresponding dielectric contrast becomes -1, so that the field automatically vanishes in the conductor. In the case of two interfaces, this leads to constant potentials on each of the two interfaces, but the potentials are not necessarily the same. It is, however, also possible to fix the electrostatic potential difference for two surfaces in a system, which is especially interesting for the simulation of batteries with conducting electrodes. In the following, a brief account on how this can be achieved with image-charge based methods as well as the ICC$\star$ method is given.

The starting point is the textbook example of a single point charge outside of a conducting sphere as depicted in Fig. 4.2a. A metallic sphere brought into an electric field can either be isolated, i.e. the charge on the sphere is constant, or on a constant electrostatic potential, typically grounded. Again, the boundary problem can be solved by adding an image charge opposite to the source charge in the sphere. This ensures, that the surface potential of the sphere does not vary, but is constant. A second image charge can be placed at the center, which dictates the electrostatic potential at the surface of the sphere: for an isolated sphere with zero net charge, the image charge at the center must be of the same magnitude as the other image charge, but with opposite sign. For a grounded sphere, it is simply zero. In the following, we will show how the same conceptual idea can be transferred to a system in a slitlike geometry. This way we can treat isolated as well as grounded boundary conditions both with the image charge and induced charge methods.

In the MMM2D method a solution of the Poisson equation periodic in x and y direction and non-periodic in z direction is constructed. To do this it is helpful to construct a partially Fourier-transformed charge distribution $\rho\left(k_x, k_y, z\right)$. For two conducting planes at $z = 0$ and $z = L_z$ the series of images can easily be constructed: $\rho^n = (-1)^n \rho\left(k_x, k_y, z + nL_z\right)$. Here n denotes the index of the image, $n \in \mathbb{Z}\backslash\{0\}$. This sum can efficiently be evaluated by ICMMM2D, an extension of the MMM2D method [192]. In the conducting limit, all Fourier modes $k_x \neq 0, k_y \neq 0$ serve to hold the potential constant on each conducting plates individually.

The $k_x = k_y = 0$ term is related to the potential difference ΔV between the two planes. It is given by.

$$\Delta V = \frac{1}{\varepsilon_r \varepsilon_0} \int_0^{L_z} \rho\left(k_x = 0, k_y = 0, z\right) z \, \mathrm{d}z. \tag{4.12}$$

In terms of the z-component of the dipole moment of the system, $P_z = \sum_i q_i z_i$, this equals

$$V = -\frac{1}{\varepsilon_r L_x L_y} P_z. \tag{4.13}$$

In order to cancel this potential difference, one has to apply a constant external field $\boldsymbol{E} = (0, 0, \frac{1}{\varepsilon_r L_x L_y L_z} P_z)$ in every MD step. This additional field corresponds to that created by the central charge in the spherical image charge picture, which puts the surfaces to the desired potential. In the planar limit, it, however, shifts to a position infinitely far away from the charge distribu-

tion, which obfuscates their possible presence. To obtain a nonzero voltage ΔV, in addition a field $\Delta V/L_z$ can be applied in z direction.

Systems treated with ICC$\star$ require no special measures if the surface on constant potential is connected within the simulation box. However some attention is required when considering electrodes at the boundary of the simulation box as depicted in figure 4.2b. If an electrostatics method is applied which is not periodic in the respective direction, this will result in two electrically unconnected surfaces. To obtain surfaces on the same electrostatic potential, such as two electrically connected plates, it is sufficient to use a solver periodic in the respective direction and leave a gap to the periodic images. This can be seen from the simple example depicted in figure 4.2b. If a solver with periodic boundary conditions is used, for example the Ewald summation, the difference between the electrostatic potential at a given position and its nearest periodic image is necessarily zero due to periodicity. But since the induced charges create a constant potential in both sections of the conductor, these must be the same throughout the whole, periodically connected conductor.

To obtain a nonzero electrostatic potential difference, the solution of the Poisson equation with zero potential can be superimposed with a solution of the empty simulation box with nonzero potential on the surfaces. This requires a solution of the Laplace equation that is then applied as an external field, just as in the simple case of parallel plates. To do that, we have implemented reading in tabulated external potentials, which are applied to charged particles weighted with the according charge into the simulation package ESPResSo. It also takes care that the external potential is not applied to theICC$\star$ discretization charges. The solution of the Laplace equation has to be obtained externally.

In order to illustrate the considerations above we show its importance on a simple model system. We chose a planar geometry because in this geometry it is possible to use both the image charge method as well as the induced charge method. Thus our system consists of a set of charges confined by two parallel conducting planes. In the following, we will show that isolated plates can be simulated either by using the ICMMM2D algorithm without correction, or by using the ICC$\star$ method with a Coulomb solver which is not periodic in the direction of the planes' normal vectors. Connected plates at zero potential difference can either be simulated with using the ICMMM2D method with the correction derived above, or using the ICC$\star$ method with a Coulomb solver which is periodic in the normal direction. As solver for the fully periodic case, we use the P^3M algorithm [161, 162], as solver for the partially periodic case MMM2D [213, 214].

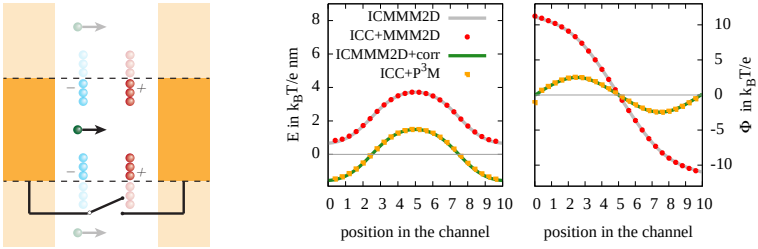


Figure 4.3: (a) Sketch of the model system that was used to probe the influence of grounded and isolated metallic boundary conditions. The two possible setups are depicted by adding a switch to electrically connect the two plates. (b) The resulting electric field E and the electrostatic potential Φ between the plates calculated with different algorithms for disconnected plates (red/grey) and connected plates (yellow/green).

For simplicity we construct a constant, spatially fixed charge distribution with a net dipole moment and probe the electrostatic field with a small test charge $q = 10^{-9}e$ that is moved through the system. Metallic boundary conditions are created at $z = 0$ and $z = L_z = 10\,\text{nm}$ by using the four algorithms described above. The dielectric permittivity between the surrounding metal plates is assumed to be $\varepsilon_W = 80$ for bulk water. The charge distribution we chose is depicted in figure 4.3a: fixed, charged particles form two oppositely charged plates at $z = 0.25L_z$ and $0.75L_z$. In one of the two directions parallel to the surrounding metal plates a void of length 0.5 L_x is left, so that the test charge can be moved through the gaps without getting close to the charged plates. We measure the electric field by performing a force calculation with the respective algorithm and dividing the obtained force by the small value of the test charge. In figure 4.3b we report the measured electric field and the potential obtained from integration of the electric field. We observe the expected behavior: The shape of the electric field is identical in all cases up to a constant. In the cases where the algorithms are supposed to simulate two connected metal electrodes the electric field is shifted downwards, so that the integral is zero, and both electrodes have the same potential.

4.2 Electrolytes near dielectric interfaces

In the following we report two examples, where the dielectric contrast between an electrolyte solution and a second material are important. Both are not directly related to an experimental system, but are interesting examples, where many important aspects can be already discussed. They are based on the primitive model of electrolytes, which we will introduce first.

4.2.1 The primitive model of electrolytes

The natural generalization of hard sphere systems to electrostatic interactions is the *primitive model* of electrolytes. In this simple model, the repulsive interactions between spherical particles is extended by a Coulomb interaction. This model as such is suited for salt melts [111], but can be extended to solutions, by taking into account the solvent polarizability in terms of a scalar dielectric constant ε_r. In general arbitrary mixtures of ion species can, of course, be considered. Electroneutrality is, however, necessary. The *restricted primitive model* corresponds to the special case of the primitive model, where two ion species of equal valency and diameter are considered.

For the considerations in Sec. 3.2 it is apparent, that the hard sphere plus Coulomb interaction is an approximation of the ion-ion potential of mean force. In water, a first shell of water molecules, the first solvation shell is bound very tightly to any small ion. This is taken into account in the primitive model in terms of an effective ion radius. There is, however, no unique definition of an ion radius. Several possible definitions exist, e.g. the Born radius describing the free energy of solvation, or the hydrodynamic radius which will be discussed in Sec. 5.2. For the primitive model of electrolytes composed of small ions an effective ion radius of 0.425 nm is very common [4, 215]. The first appearance of this number we found was the publication of Card and Valleau from 1970 [216].

A generalization of the primitive model can be done in terms of a softer potential. In Hansen/McDonald [111], a repulsive potential of the form

$$V(r) = \varepsilon_r \left(\frac{\sigma}{r}\right)^n$$

is proposed, which yields the hard sphere interaction in the limit $n \to \infty$. In our simulations we applied a Weeks-Chandler-Anderson potential, a purely

repulsive truncated and shifted variant of the Lennard-Jones interaction:

$$\frac{U(r)}{k_BT} = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + 1 & \text{if } r < 2^{1/6}\sigma \\ 0 & \text{otherwise.} \end{cases} \tag{4.14}$$

Here the energy parameter ε was set to k_BT and the canonical ion radius of 0.425 nm was applied. This variant has the advantage of being continuously differentiable everywhere, which is helpful for the numerical integration of the equations of motion. At the same time, the range in which it is nonzero is relatively small, as $2^{1/6} \approx 1.122$.

We however expect that the details of the ion-ion interaction is only of minor importance for the results as the considered ion densities are relatively small. In Ref. [180] it was e.g. shown that the Poisson-Boltzmann theory can well describe the ion distribution around DNA, although in PB theory ions are treated as point-like.

4.2.2 Salt distribution in a slit pore

As an illustrative example, we investigated the distribution of a 3:1 electrolyte in a two-dimensional slit, in which the walls have a different dielectric permittivity as that of the solvent.

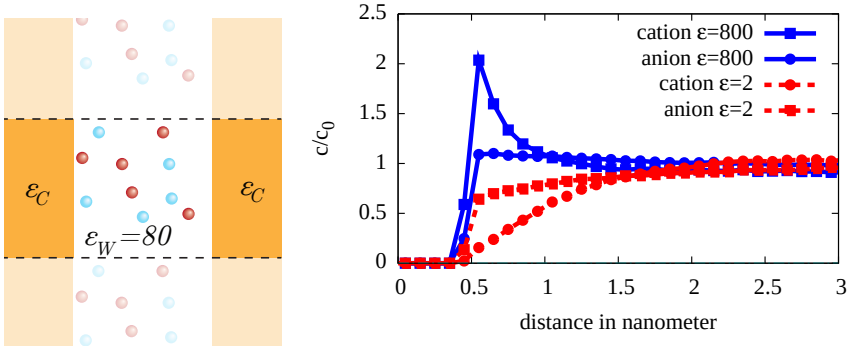
The simulation setup is sketched in Fig. 4.4a. We use a box of 20x20x11 nm^3 in which walls are created at z=0.5 nm and z = 10.5 nm. This is done using a potential, of the WCA-form, except only the coordinate normal to the wall $z - z_0$ is inserted as distance. z_0 is chosen such that the potential equals k_BT at the desired position. 23 trivalent cations and 69 univalent anions were added to the system, corresponding to a concentration of 10 mM. The expected Debye length is about 1 nm, and hence we expect the interactions between the walls to be small.

In the calculations the ELCIC (electrostatic layer correction with image charges) algorithm was applied [217]. In this algorithm, the infinite series of image charges is summed up as in the ICMMM2D. The electrostatic forces of the particles which are not images, however, is calculated differently. The result of a 3D-periodic electrostatics method, in this case P^3M, is corrected such that effectively a system only periodic in the directions parallel to the slab is considered. Image charges are constructed by reflections at the box boundaries in z-direction. The parameter d determining the approximate distance

of closest approach of ions to the wall hence is set to 0.5 nm. We investigated a wall of high permittivity and a wall of low permittivity in the same simulation. One wall has a permittivity of $\varepsilon_r = 2$, which is typical for biological membranes. Water is assumed have $\varepsilon_r = 80$. The prefactor in Eq. 4.7 then is already at a value of 0.95, hence the strength of the image charges is almost as large as the original charges. On the other wall we assumed $\varepsilon_r = 800$, which is close to a conducting boundary.

In Fig. 4.4b the density distribution is reported. Both walls are shown in one graph, and each concentration is divided by c_0=10 mM. Near the high permittivity wall both ion species accumulate. The accumulation of the trivalent cation is much stronger than the univalent anion because the image charges are larger by a factor of 9. This creates a net charge in the first layer, which is then compensated for by an extended layer of anion excess. Image charges lead to a repulsion of the ions on the side with the low dielectric medium. Here again the strength is larger for the cation, leading to a negatively charged boundary layer and a positive second layer.

The range of the accumulation and depletion layers is similar to the Debye length. We will find this consistently throughout all applications: Dielectric effects, as all electrostatic effects are screened on the length scale of the Debye length, and hence only need to be considered if the Debye length is comparable to other length scales in the system.



(a) (b)

Figure 4.4: (a) Sketch of our simulation setup. A 3:1 electrolyte, between two walls with a dielectric constant different from that of water. The slab is periodically replicated parallel to the walls, vertical in the sketch.
(b) Density distribution of univalent anions and trivalent cations near the dielectric interface. The dielectric interfaces is placed at $x = 0$, and a repulsive potential maintains a minimum distance of 0.5 nanometers of all ions. A a high permittivity medium with $\varepsilon_r = 800$, strongly attracts cations, while anions are less attracted. A low permittivity $\varepsilon_r = 2$, repels both ion species. Image charges are stronger for the trivalent cations, leading to a charged boundary layer with opposite sign in both cases.

4.2.3 Salt distribution in a cylindrical pore

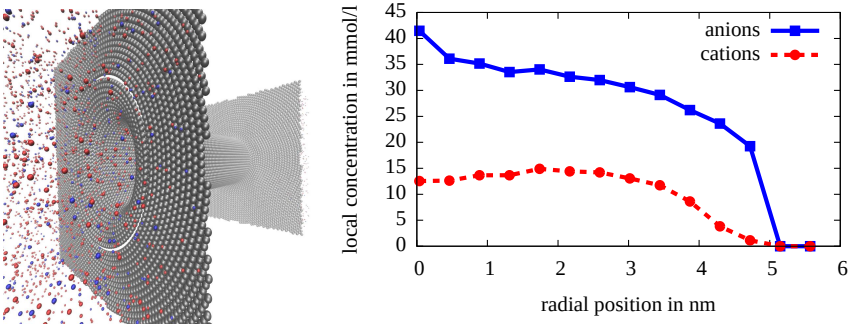
As a second example we investigate how the dielectric permittivity of a membrane affects the distribution of the same electrolyte in a cylindrical pore. We expect the qualitative findings to be similar. With this example we can demonstrate, that the ICC⋆-algorithm is applicable to more complex shapes.

We created a discretized representation of a cylindrical pore in a flat membrane with smooth openings. We devised a procedure where the user specifies the desired pore length, radius and the average distance between discretization points. Then the particles are placed automatically. This was implemented in the `TCL` language and is shipped with ESPResSo versions > 3.1. Conical shapes are possible, but will not be used in this thesis. The normal vectors and areas for each discretization point are stored in `TCL` variables that can be handed over to the ICC⋆ function.

The pore is incorporated in the MD simulations as what is called a *constraint* in ESPResSo. The key ingredient is the *signed distance function*, a mathematical function $d(\boldsymbol{r})$, that maps from every point in the simulation box to a real number. It describes the distance to the closest point on the surface of the object. A position with a negative distance to a constraint is considered to be inside the wall. This distance is inserted the into interaction potential. Forces are calculated as the negative gradient of this expression. We apply the WCA interaction with $\sigma = 1$nm and $\epsilon = 1 k_B T$. The shape parameters of the distance function are chosen so that the interaction potential equals 1 $k_B T$ for the given size parameters. An implementation of the signed distance function did exist in ESPResSo. We only adapted it for conical pores. This implementation, however, was error-prone and tedious, and we consider it a very helpful extension of ESPResSo if these procedures could be partly automated. In Sec. 6.4 we will comment further on this issue.

In Fig. 4.5a a snapshot of the system is shown. A box of size 22x22x70nm^3, was filled with 649 cations and 1947 anions. This number had been calculated from the volume in the box that was not occupied by membrane material assuming an electrolyte concentration of 10 mmol/l. A pore of diameter 10 nm and length 40 nm was created. The potential acting on the particles is again a WCA potential, where the distance function was used as argument to the WCA potential. The discretization points of the surface were shifted by 0.5 nm into the position of the boundary. Periodic boundary conditions were applied in all directions of the system with the P^{3}M algorithm. The dielectric constant of the wall material was set to $\varepsilon_r = 2$.

We measured the distribution in the central half of the cylindrical pore, using the new routines for analysis on the fly described in Sec. 3.3.5. Qualitatively our findings are consistent with those for the flat wall. Both ion species are repelled from the pore boundary, and only after a few Debye lengths, the concentration profile is flat again.



(a) (b)

Figure 4.5: (a) The ICC⋆ example system: positive and negative ions are displayed as red and blue spheres, the ICC⋆ discretization points by grey spheres. (b) Ion density of both species in the pore measured in the center of the pore.

4.3 Translocation Barriers for DNA molecules

In theoretical approaches [218] the translocation process is usually subdivided into three different phases, namely: (I) The approach of a DNA molecule to the pore opening, where the molecule undergoes a drift/diffusion motion; (II) The capture process where the first piece of a DNA enters the pore; (III) The translocation phase, where the contour of the chain molecule moves through the pore. In phase II, chain-like molecules have to partly un-coil and enter the pore with either one end or with a loop. This is also called the nucleation process as the molecule has to overcome a free energy barrier in this phase. In this section we investigate the free energy barrier for short, stiff segments of double-stranded DNA molecules and long, flexibility single-stranded DNA molecules. After some remarks on the experimental situations, we present our simulation model and its prediction for the dsDNA molecules. Then we show our results for ssDNA, after having introduced a new reaction coordinate especially well suited for the capture process of flexible chain molecules.

4.3.1 General considerations

In the capture phase, a DNA molecule has to overcome a free energy barrier that is entropic in nature since the confinement reduces the number of allowed configurations. In general, this contribution will depend on the flexibility of the macromolecule. However, a charged macromolecules such as DNA it is surrounded by a cloud of counterions, and as it enters the confined geometry of the pore, the counterion cloud has to follow. This can enhance the barrier, since under most conditions the confinement also perturbs the distribution of counterions. Kumar et al. showed [12] that the free energy of a flexible polyelectrolyte confined into a spherical cavity is increased by 6-9 k_BT when one of its ends is anchored to the cavity. This anchoring can be assumed to be comparable to the capture process as one end of the chain is confined in space.

For much smaller pores, very similar questions have been asked e.g. by Parsegian [219] and Levitt [220,221]. They considered the ion conduction through thin biological pores. They found electrostatic interactions, especially related to the permittivity to be of major importance. The electrostatic barrier for ssDNA in a biological α-hemolysin pore was calculated in a work by Zhang and Shklovskii [11]. They found a value of more than 10 k_BT. In these calculations the barrier is governed by the dielectric mismatch effect between the surrounding water and the membrane material that is less polarizable. Findings based on simulations that take into account the dielectric contrast indicated that a dielectric contrast increases the tendency to neutralize charges in the pore [222] and compresses the counterion cloud [223]. Previous simulation works on this effect were however restricted to generic pore models with small diameters not much larger than a nanometer and small simulation cells, because the employed scheme to solve Poisson equation in the presence of dielectric discontinuities was computationally very demanding. The ICC$\star$ algorithm is a very viable alternative, and we could already show that it allows us to simulation large systems.

Experimentally a high ($10k_BT$ and more) barrier that depends on the salt concentration is well supported: Brun et al. investigated the voltage dependence of the capture rate of dextran sulfonate, a polyelectrolyte [224]. From exponential fits of the translocation frequency they estimate the free energy barrier to be $\sim 10k_BT$ for high salt concentration (1 M) while being higher ($\sim 14k_BT$) for lower salt concentration (0.1 M). The experiments in the group of Meller [225] indicated that ssDNA translocating through α-Hemolysine crosses a free energy barrier that strongly depends on the buffer

solution. Their explanation seems to omit the effect of dielectric contrast, which could provide a good physical justification. Oukhaled et al. also showed that for α-Hemolysine pores the frequency of capture events is inversely proportional to the Debye length [226].

4.3.2 The translocation barrier for a double-stranded DNA molecule

To study the influence of dielectric interfaces on DNA translocation it is necessary to construct a coarse–grained representation of a pore and a DNA molecule. Most importantly we want to study the question if dielectric boundary forces are responsible for a significant electrostatic barrier that opposes the translocation process. We wish to apply a coarse–grained picture where an implicit solvent approach is justified. This is clearly not the case for protein–based nanopores. Solid-state nanopores have diameters of more than 5 nm in most cases. Here coarse-grained representations can be expected to work much better.

Our pore representation is a cylindrical hole in a flat membrane of 8 nm thickness that is smoothed at the openings to avoid sharp edges. In comparison to experiments this pore is relatively short. For example in the experiments of Keyser et al. , a length of around 40 nm is reported [5]. In the MD simulations it is modelled exactly as the pore by its distance function and a WCA potential as described in Sec. 4.2.3.

To evaluate the dielectric boundary forces we require a model that reasonably well reproduces the electric field on the nanometer length scale. Most importantly this means to reproduce the DNA's charge density correctly. Our dsDNA model is motivated by the fact that the persistence length of B-DNA is between 40 and 60 nm [227]. In other words it is rather stiff on the length scale of the size of small synthetic pores. This allows us to approximate a 10 nm (=30 base pairs) long piece of DNA as a stiff rod. As its backbone's acidic groups are fully dissociated we assume it to be charged with $2e$/bp and approximate it as a rigid object composed of 58 overlapping spheres of diameter 2 nm with a mutual distance of 0.167 nm each carrying one elementary charge. This rigid cylinder model is found in the literature in many places, and is sometimes dubbed the *primitive model of DNA* [228, 229], reminiscent of the primitive electrolyte model introduced above.

The rationale to use such a crude and simple model is – beyond the ease of computations – the insight, that the exact interaction between ions and

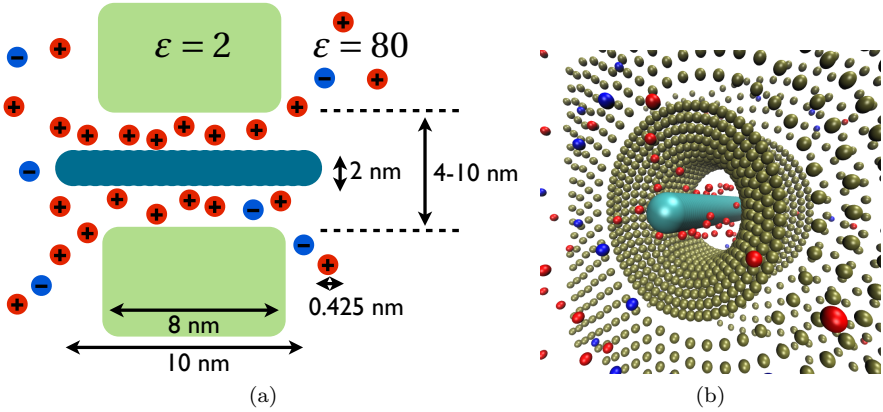
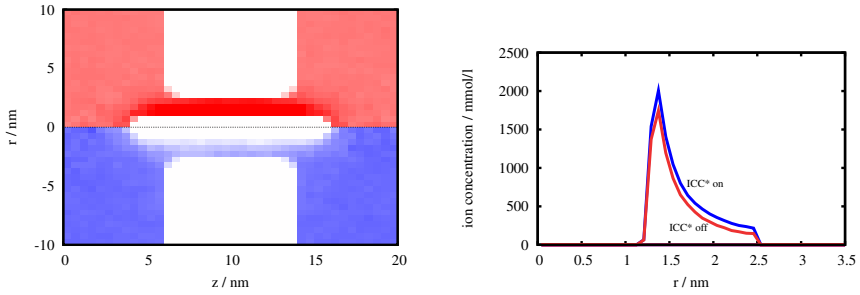


Figure 4.6: (a) Sketch of the simulation setup applied for the translocation of a rigid model dsDNA. A nanopore is represented as a cylindrical cavity of length 8 nm and varying diameter in a flat membrane. A fragment of dsDNA of 10 nm length is represented by 58 overlapping spheres with 2 nm diameter and fixed on the pore axis. Ions are represented as spheres with a diameter of 0.425 nm water, a dielectric medium with $\varepsilon = 80$. The dielectric constant of the membrane can be set to $\varepsilon = 2$ by using the ICC$\star$ algorithm. (b) Simulation snapshot. Anions and cations are displayed as blue and red spheres, and the DNA beads as cyan spheres. The discretization points of the ICC$\star$ algorithm, the discrete induced charges are shown as golden spheres.

DNA is not of importance. If the DNA does not approach a wall to shorter distances that the length scale of its microstructure, the exact shape will appear smeared out. Creating a representation which has the right asymptotic electrostatic potential is sufficient to yield a correct free energy barrier. A necessary condition for that is, of course, repulsive interactions between DNA and wall. If attraction dominated, the DNA would stick to the wall and it would be necessary to create a more detailed model. In chapter 5.3 this model is compared to a full atomistic model of dsDNA. There it will be shown, that the charged-rod model applied in this section is surprisingly accurate.

In Fig. 4.6a the most important simulation parameters are sketched. Fig. 4.6b shows a snapshot of the system. We applied a time step of $\tau = 0.005\sqrt{m\sigma^2/k_BT}$, and the Langevin equation (Eq. 3.24) to thermalize the system with a friction coefficient of $0.5/\tau$. The simulations were performed using an earlier version of ESPResSo, and the induced charges were recom-



(a) (b)

Figure 4.7: Coion and counterion densities for a salt concentration of 100 mM. (a) The ion density as a function of the position z along the pore axis and of the distance r from it. A 10 nm long DNA fragment is centered in the pore. Density of counterions (upper part, red) and coions (lower part, blue). Darker colors indicate higher density. The increase in density of counterions and the depletion of coions close to the DNA can easily be seen. (b) The counterion density in the vicinity of the DNA measured in the center of the pore as a function of the distance from the DNA. Upper curve: dielectric contrast is taken into account. Lower curve: Dielectric contrast is neglected.

puted every 10 MD steps. All ion distributions remained independent of this parameter. All simulations were performed by adding 58 counterions, and a certain number of salt ions to the system. The number was calculated from the available volume in the box, multiplied with the electrolyte concentration to be considered. We investigated a salt-free case, and concentrations of 10 mM, 30 mM, 100 mM, and 300 mM. Each simulation was run with and without taking the dielectric contrast between the pore and the water into account.

In Fig. 4.7 we show the ion distributions obtained for the 10 mM electrolyte case. A Debye layer forming around the DNA can clearly be observed. Near the DNA the local ion concentration goes up to around 2 M. It decays within the pore to slightly more than 100 mM, indicating that the electrostatic potential has almost decayed to zero. The counterion density is increased when taking into account the dielectric contrast. This reflects the enhanced attraction between unlike-charges, or the tendency to neutralize charged objects as qualitatively found in Sec. 4.1.2.

To assess the free energy barrier for translocation we made a simplification

regarding the geometry. In principle the position and orientation of the rod is given by five independent parameters, e.g. the position of the center of mass, and two angles. We however further simplify the geometry, by assuming that the DNA position is constraint to the pore axis and aligned parallel to the pore. This is justified as the shape of the pore requires an alignment of the rod. On the other hand we dramatically reduce the configuration space necessary to be sampled. The whole process can be expressed in terms of one number only, the axial (z) component of the center of mass (COM) position of the rod. The free energy can hence be obtained as the potential of the mean total force in z direction. This allowed us to perform a completely straightforward calculation. We fix the DNA at equidistant positions z_i with a mutual distance d. Fixing means simply that the particle positions are not updated in integration step of the MD run. The force necessary to constrain the DNA at that particular z-position trivially is the negative of the force acting in every MD step. The free energy profile $\mathcal{F}(z)$ then is obtained from a numerical integration of the mean constraint force F

$$\mathcal{F}(z_n) = -\sum_{i=1}^{n} \langle F(z_i) \rangle d. \tag{4.15}$$

As the system setup is symmetric with respect to reflection at the center plane of the membrane, we symmetrized the free energy in the graphs shown below.

Free energy profiles for five parameter sets are displayed for a 5 nm pore in Fig. 4.8. It is shown for three salt concentrations, 10 mM, 30 mM, and 100 mM, in which the dielectric contrast is taken into account and 10 mM and 30 mM, in which the dielectric contrast was neglected. Data for other concentration was calculated but omitted here as they show similar behaviour. All datasets exhibit a significant free energy barrier. This indicates that the Debye layers are compressed, which is assiociated with a free energy loss. In all cases the free energy barrier is larger with dielectric contrast, and strongly depends on the salt concentration. The height of the barriers is significant: In the 10 mM with dielectric contrast more than 10 k_BT is observed.

We further investigated the dependence of the barriers on the salt concentration and the width of the pore, as we expect that the ratio of the Debye and pore radius governs the height of the barrier. The results are shown in Fig. 4.9. The barrier height decreases strongly with the electrolyte concentration. For c=300 mM the barrier is negligible even for the smallest pore size. In this case the Debye length is about 0.5 nm, which is two times smaller than the

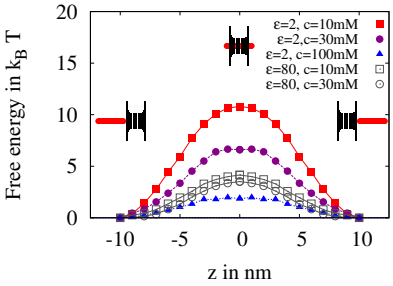


Figure 4.8: Free energy as a function of the z component of the COM position of the DNA. The potential is set to zero at the position where the DNA rod is centered between two periodic replicas of the pore. The salt concentrations 10 mM, 30 mM and 100 mM are shown as filled squares, circles and triangles for cases in which dielectric contrast is taken into account. The salt concentrations 10 mM and 30 mM are shown as open squares and circles for the case where the dielectric contrast was neglected.

shortest distance between the DNA and the pore wall. The barriers including dielectric contrast are typically a factor 2-3 larger than the barriers obtained without dielectric contrast.

Next we want to consider the free energy barrier for flexible molecules reflecting single-stranded DNA. One important question considered was, what is an appropriate reaction coordinate for the process. The COM position of the chain is a possible choice. However, we have seen, that the free energy profile is very steep at the point where only a short piece of the DNA enters the pore. We devised an alternative reaction coordinate for chain molecules, that can capture this aspect well. It is presented in the next section.

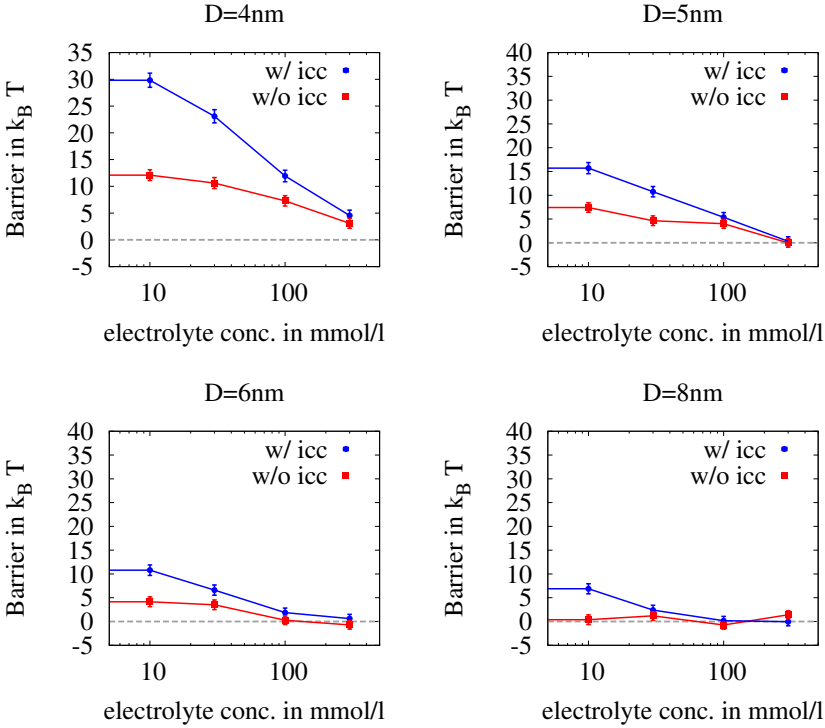


Figure 4.9: The translocation barriers obtained for pore diameters between 4 and 8 nm as a function of the electrolyte concentration. Results obtained with dielectric contrast are shown as circles, and without dielectric contrast as squares.

4.3.3 A new reaction coordinate for the translocation of chain molecules

Our aim is to compute the free energy landscape along a reaction coordinate which is suitable to characterize the progress of translocation over all three different translocation phases. It is necessary that this coordinate is chosen so that it can be unambiguously determined from a given configuration and that it is a continuous function of the configuration. It is also helpful for practical purposes if the reaction coordinate is simple enough so that generalized forces can be calculated easily from configurations. In the early article of Muthukumar [146], the suggested reaction coordinate is the index of the monomer centered in the pore at a given instant in time. This, however, can neither describe the approach of the macromolecule to the pore, nor is it well defined if more than one monomer is centered in the pore. In general, the use of this reaction coordinate presents difficulties if both ends of the chain are on identical sides of the membrane. This corresponds "the approach" of the macromolecule towards the pore, thus part of phase I. In phase I the desired reaction coordinate, however, should describe the distance of the chain from the membrane. Our following suggestion allows to combine both approaches and correctly capture the symmetry of the pore and our model chain.

We assume that the polymer configuration is described by a parametric function $\boldsymbol{r}(l)$ where $-l_C/2 < l < l_C/2$ spans the contour length l_C. We then distinguish two cases, namely: (a) Both ends are the same side of the membrane. The reaction coordinate R takes the value $\pm(d + l_C/2)$, where d is the Euclidean distance to the middle of the pore of the polymer end closest to the pore itself. The positive and negative sign distinguish between the cis and the trans compartment. (b) Both ends are on different sides of the membrane. In this case we assign to the reaction coordinate R the value l at which the contour $\boldsymbol{r}(l)$ intersects with the separation plane. If more than one intersection is found, since their number has to be odd, the central one is taken. Hence, our reaction coordinate spans the interval from $-\infty$ to $+\infty$ and is able to describe the whole translocation process as a continuous functional of the polymer contour.

Differently from other possible reaction coordinates, it covers the fact that the insertion of only a few monomers into the pore is potentially associated with a steep increase of the free energy. The generalized forces along this reaction coordinates are just the usual Cartesian forces on the chain bead centered in the pore or on the head bead, respectively. Also, the force is continuous at the transition point of the two parts, as the head bead coincides with the bead

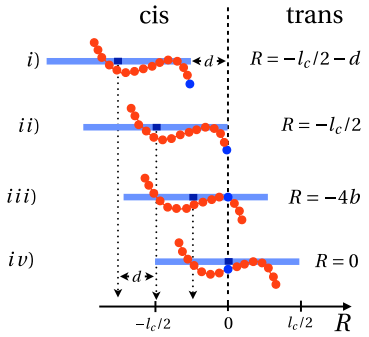


Figure 4.10: The reaction coordinate R of a stiff chain and a flexible chain. We distinguish two cases: (i) both ends of the chain are on the same side of the mid- plane through the membrane and (iii, iv) both ends of the chain are on different sides. (ii) is the transition point. For the stiff chain, R is the axial position of the center of mass of the chain. For the flexible chain in case (i) the distance of the end monomer plus half the contour length l_c is taken for R. In (iii) and (iv) the distance between the position of the monomer centered in the pore and the position of the central monomer of the chain taken along the polymer contour is chosen. The sign indicates if the chain is mainly in the cis or trans compartment. The definition for the stiff chain is a special case of the more general definition for flexible chains.

centered in the pore at the transition between case (a) and (b). For a single, stiff chain confined to the pore axis the reaction coordinate reduces to the axial position of its center in a cylindrical coordinate system centered in the middle of the pore.

4.3.4 The translocation barrier for a single-stranded DNA molecule

To model ssDNA, we applied a model inspired by common coarse-grained polyelectrolytes models, a bead spring model [230, 231]. In our model, spherical beads are connected by harmonic springs, with a potential

$$V(r) = \frac{1}{2} k (r - r_0)^2 \tag{4.16}$$

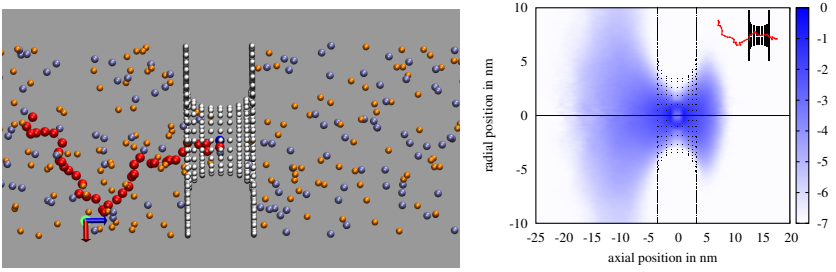


Figure 4.11: (a) Simulation snapshot of a flexible ssDNA molecule (represented by red beads) through a pore. The particles used for the discretize the pore surface for the ICC$\star$ algorithm are shown in white, cations in orange and anions in blue. (b) Density of beads found in simulation with the bead with index 10 fixed in the pore. Darker colors correspond to a higher probability of finding a bead at the corresponding position. A representative configuration is displayed in the upper right corner.

with an equilibrium bond length of r_0 =1 nm and a $k = 30k_BT$/nm. The beads mutually interact with a WCA potential with σ=1 nm and $\varepsilon = k_BT$. The line charge density of ssDNA is slightly less than 3 charges per nanometer, and hence each bead was assigned three elementary charges. The resulting mean bond length was around 1.1 nm, and hence the line charge density of ssDNA was reproduced with sufficient accuracy. We considered polymers with $N = 40$ beads, corresponding to 120 bases, which resulted in a contour length of 44 nm. We chose a box length of 20x20x60 nm^3, sufficient to host the ssDNA molecule. Fortunately an improvement of the P^{3}M implementation to non-cubic boxed allowed us to perform the simulation without extending the box size in x and y direction. This would have dramatically increased the number of ions to consider. We found a discretization of the pore for the calculation of induced charges with a resolution of 1 nm sufficient to obtain accurate forces, and hence we could significantly reduce the number of necessary surface charges. A pore of 10 nm diameter was chosen, as we expected the force in general to be higher due to closer approach of the DNA to the pore wall. All other parameters were chosen in accordance with the simulations of dsDNA presented above. In Fig. 4.11a a snapshot of a simulation is shown. In addition in Fig. 4.11b we display a density map of the beads for a simulation in which the monomer with index 10 was fixed in the center of the pore. The equilibrium extension of the polymer can be estimated from this graph.

We calculated the radius of gyration of the chain in electrolytes in different salt concentrations. It was found in the range of 5.5 and 6 nm depending on the electrolyte concentration. This was significantly larger than for an uncharged chain, as the electrostatic between adjacent beads is known lead to an apparent stiffening of the chain. By employing also an angular potential of the form

$$V(\phi) = \frac{1}{2} k_\phi \phi^2$$

to an uncharged chain, with k_ϕ=$10 k_B T$, a similar radius of gyration was obtained. Furthermore, we stiffened the charged polymers by applying and angular potential with $k_\phi = 30 k_B T$. This way we can minimize the entropic contribution to the free energy barrier. The two additional models, the uncharged polymer with a similar radius of gyration and the artificially stiffened charged polymer, are used for reference simulations. With the uncharged polymer we can probe the entropic contribution to the free energy barrier that will help us interpreting the simulation results. With the stiffened polymer we can investigate the free energy barrier with a much smaller entropic contribution. This allows us to proble electrostatic and entropic effects separately.

For the free energy calculation with the reaction coordinate for chains introduced above we made some assumptions. The contour of the polymer is assumed to be a piecewise linear function, connecting the bead positions. For configurations where both ends of the polymer were on different sides of the symmetry plane we only evaluated it at positions, where one bead was exactly placed at the center of the pore. The fact that the contour length on both sides could slightly fluctuate was neglected. Furthermore the reaction space was restricted only to configurations where the bead by which the reaction coordinate was determined is placed on the axis. These approximations allow us to constrain the system to a certain value of the reaction coordinate by fixing one bead of the polymer chain at a certain position.

The free energy is calculated as follows. The system in the region (a) of the definition of the reaction coordinate, where the polymer has not yet entered the pore, reaction coordinate is constraint to equidistant z_i with mutual distances d by placing one end bead at $(0, 0, z_i + L_C/2)$. We only consider negative reaction coordinates here. For positive values the + sign must be exchanged for a − sign. The force acting on the end bead was measured, and the free energy was calculated from Eq. 4.15. In the region (b) where the chain crosses the membrane, the reaction coordinate is constraint by fixing the bead with index $z_i/b + N/2$, where b is the mean bond length is fixed in the center of the pore. The mean force acting on that bead is calculated in a simulation,

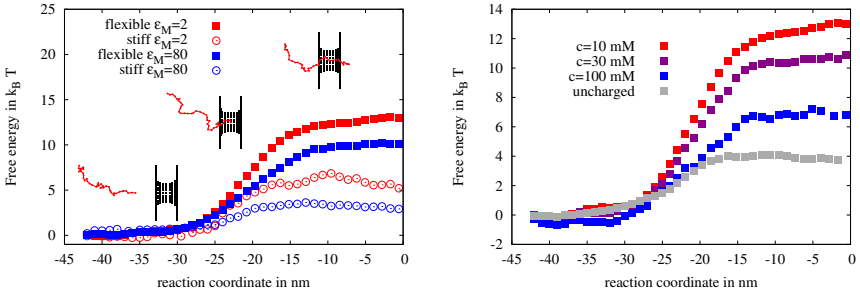


Figure 4.12: (a) Potential of mean force for a flexible ssDNA segment at a salt concentration of 10 mM. The free energy changes steeply as the first 10 nm of the ssDNA-molecule enter the pore. (b) The potential of mean force for a ssDNA at different salt concentrations with $\varepsilon_M = 2$. For comparison the free energy profile of an uncharged polymer with a comparable radius of gyration is shown in grey.

and the free energy is again calculated from Eq. 4.15, except the distance d is exchanged by the mean bond length b.

In Fig. 4.12a the free energy for four different configurations, all in 10 mM salt concentration, is shown. We compare again the cases with and without dielectric contrast. We do this for a fully flexible, and an artificially stiffened DNA. For a flexible DNA, the barriers are general higher, and taking into account the dielectric contrast also increases the barriers by about 4 k_BT in both cases. The flexibility seems to be of major importance. It enhances the barriers by around 8 k_BT in both cases. Two contributions may be important. (a) There is an entropic barrier directly associated with the chain flexibility, and (b) due to the flexibility electrostatic effects are magnified, as the beads can get closer to the wall. For comparison we performed simulations also with an uncharged polymer of which we had matched the radius of gyration as described above. It is shown in Fig. 4.12b. It experiences a barrier of about $4k_BT$, which is of purely entropic origin.

These results suggest the following approximate decomposition of the free energy barrier for a ssDNA molecule in 10 mM salt: 4 k_BT due the entropy loss in the pore, and 7-12 k_BT due to electrostatic effects. Taking into account the dielectric contrast between membrane and DNA again approximately doubles

the height of the free energy barrier, consistent with our findings for double-stranded DNA.

For comparison we also performed calculations in different salt concentrations. The barriers are displayed in Fig. 4.12b. Again a significant decrease of the barrier height with increasing salt concentration is observed. Also for 100 mM the barrier remains clearly above the barrier measured for the purely entropic barrier.

We conclude that the considered 50 nm long ssDNA strand is subject to considerable free energy barrier in the 10 nm wide pore. The entropic contribution accounts for around 4 k_BT. Under an electrolyte concentration of 10 mM we measure a total barrier is around 14 k_BT, of which we attribute 10 k_BT to electrostatic effects. Neglecting the dielectric contrast, the electrostatic barrier is considerably smaller. The electrostatic contribution to the free energy barrier, however, rapidly decays when increasing the salt concentration. At 100 mM it is reduced to only 2-3 k_BT. This finding is consistent with those made above for dsDNA. The dielectric constant approximately double the free energy barrier. When comparing the absolute numbers, it is, however, important to note, than in this investigation the DNA charge density was significantly smaller than in the investigations for dsDNA and at the same time the pore was much larger in diameter. This demonstrates that the flexibility enhances the role of electrostatic barriers significantly.

4.4 Poisson-Boltzmann models for the translocation barrier

In the following section we present the translocation barrier of a charged cylinder, a model of double-stranded DNA, calculated by means of the Poisson-Boltzmann (PB) equation. First, we will show results obtained with a very simple one-dimensional model which agrees reasonably well with the data we obtained from MD simulation described in the previous section. Interestingly, only the MD data including dielectric contrast was especially well reproduced. This is very surprising as the simple PB model does not even contain any notion of dielectric permittivity at all. We were originally very surprised by this fact, however we found a convincing explanation for this behaviour. Second, we investigated the translocation barrier estimated for a cylinder of finite length with the iPBS method, suggested by S. Tyagi, and worked out in detail by A. Schlaich in his diploma thesis [14]. In the iPBS, method a finite element

solution of the PB equation is combined with the idea of the ICC⋆ algorithm. The boundary condition to the FEM solution are iterated. With this model we can support our more qualitative discussion of the one-dimensional model.

4.4.1 A simple one-dimensional Poisson-Boltzmann model for the translocation barrier

One effect observed in the calculations above, is that inside the pore the degree of counterions was enhanced by the dielectric contrast. This can also be understood in terms of a modified electrostatic pair interaction. In the image charge picture in the vicinity of a low permittivity dielectric, the attraction of unlike-charged objects in enhanced. Applying this to the DNA-electrolyte system, we expect, that the effective attraction between cations and DNA is higher when the DNA is inside a pore. On the other hand, if the DNA is completely neutralized, we expect no enhanced attraction, as an object surrounded by a fully developed counterion cloud does not induce net charge in an adjacent interface.

The expected degree of neutralization will depend on the pore length: A long pore will lead to a strong energetic penalty if the DNA is not neutralized, while for a short pore a non-neutral DNA molecule will lead to a much smaller energy penalty.

We performed simulations with different pore lengths between 4 nm and 20 nm at 10 mM salt and measured the degree, to which the DNA charge in the center of a pore was neutralized. This degree was defined as the net ion charge in a thin slab at the pore center divided by the DNA charge in the same slab. A neutralization degree of 1 thus means that each DNA charge is exactly compensated for by counterions.

The results are depicted in Fig. 4.13a. The dielectric contrast enhances the tendency that the interior of the pore is electrostatically neutral at much lower pore lengths than in the situation without dielectric contrast. In both cases, full neutralization is reached in the limit of an infinitely long pore. This is due to the fact that the self–energy of an infinitely long charged rod diverges and therefore must be compensated by counterions. In the limit of zero salt reservoir concentration as treated in [232], this divergence leads to the following: Even when decreasing the ion reservoir concentration to zero a finite number of ions stays next to the charged cylinder. This behaviour is known as Manning condensation. The effect of charge neutralization allows us to make the following consideration. If the interior of the pore is fully

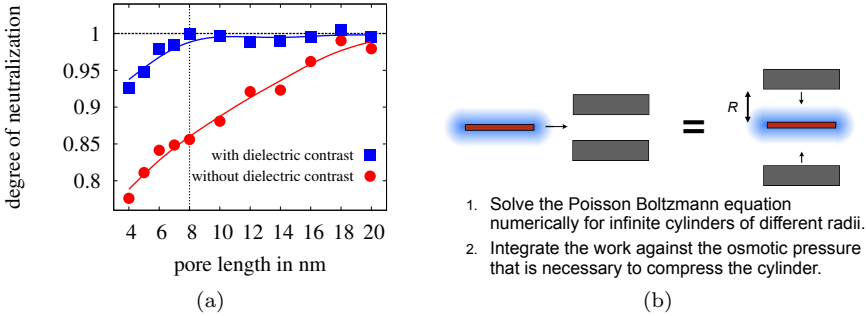


Figure 4.13: (a) The degree of neutralization as a function of the pore length. The vertical dotted line corresponds to the length for which free energy barriers are reported. (b) Sketch depicting the idea of the PB model applied to calculate free energy barriers.

neutralized, no image charges are to be taken into account. Then the effect of the pore confinement is only a compression of the counterion cloud into a radius r. This can be modelled with the Poisson–Boltzmann equation as described in the following.

We assume DNA to be a cylinder of radius 1 nm and one elementary charge per 0.167 nm and the pore to be a cylinder of radius R without surface charges. The boundary conditions thus are the following:

$$\left.\frac{\partial \Phi}{\partial r}\right|_{r=1\mathrm{nm}} = 4\pi l_B \sigma_{\mathrm{DNA}}, \tag{4.17}$$

$$\left.\frac{\partial \Phi}{\partial r}\right|_{r=R} = 0, \tag{4.18}$$

where $\sigma_{\mathrm{DNA}} = 0.94e/\mathrm{nm}^2$. The second boundary condition corresponds to the requirement of charge neutrality of the system.

The work to change the volume of the pore is determined by the osmotic pressure on the cylinder (see the sketch in Fig. 4.13b). In the PB approximation this is just the ideal gas pressure of the ions at the boundary of the cell. However, we do not want the compress the whole ion cloud, but allow particle and volume exchange with the reservoir. Under these conditions the excess

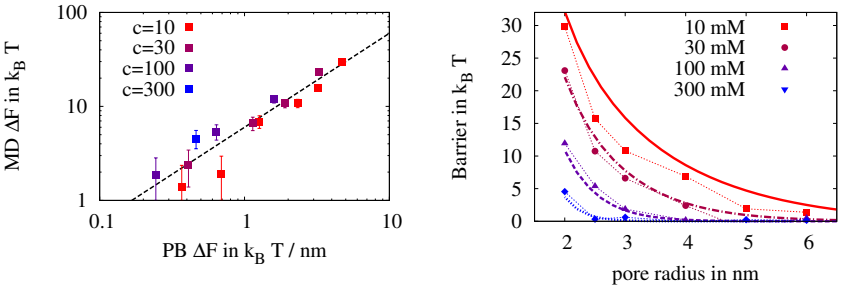


Figure 4.14: (a) Correlation plot: The free energy barrier per pore length computed from Poisson-Boltzmann theory (horizontal axis) correlated with the result from molecule dynamics simulations of Sec. 4.3.2 (vertical axis). The slope of the dashed line is 6 nm, the best choice of an effective pore length. (b) The free energy barrier calculated from MD simulations (symbols connected with thin lines), and PB model (thick lines).

osmotic pressure, meaning the osmotic pressure minus the osmotic pressure in the reservoir. It reads as:

$$p^{\mathrm{e}}(R) = k_B T \left(c_+ + c_- - 2c_0\right)\Big|_{r=R} \tag{4.19}$$

$$= 2k_B T c_0 \left(\cosh \Phi(R) - 1\right). \tag{4.20}$$

This allows us to calculate the work needed to shrink the pore to the desired size. by integrating $p^e \mathrm{d}V$ from an infinitely large cylinder to a finite cylinder of radius R. In our case, we obtain the work per unit length as the integral:

$$\mathcal{W}/l = \int_{\infty}^{R} p^{\mathrm{e}}(R')\, 2\pi R' \mathrm{d}R'.$$

We determine this work from numerical solutions of the PB equation and a simple Euler integration scheme. We apply a finite difference scheme implemented in the programming language PYTHON using the NUMPY numeric library.

In Fig. 4.14 we report the results obtained with the PB model in comparison to the results of the coarse grained MD model including dielectric contrast. First, we report the correlation of the PB barriers with the MD data. In

this graph, the MD barrier is reported as a function of the PB barrier for the same pore radius and salt concentration. The correlation is very good. This is surprising given the simplicity of the PB model. From the slope of the plot it is possible to extract an effective pore length. The slope of the dashed line is 6 nm, which agrees reasonably well with the 8 nm pore we have used in the MD simulations.

In Fig. 4.14b we report the free energy barrier as a function of the pore radius for different salt concentration. The PB barriers are scaled to an effective pore length of 6 nm. The agreement is again surprisingly good: The deviations are within a few k_BT. This indicates that the general physical effects are captured correctly by the PB model. Although our considerations make it plausible that the PB model is appropriate, it remains surprising to us, that the complex effect of dielectric boundary forces can be boiled down to a model that does not contain dielectrics at all. It encourages us to believe, that other simple models in physics chosen with care and good physical intuition sometimes can explain more than expected at first sight.

As a further investigation step we calculated the free energy barriers obtained for a cylindrical DNA molecule of finite length with the iPBS method, which is described below.

4.4.2 The iPBS algorithm

In this section we give a brief overview over the *iterative Poisson-Boltzmann solver* (iPBS) method and apply it to the question of Sec. 4.3.2, the translocation barrier of a stiff dsDNA fragment. It is based on a finite element solution of the PB equation, for which boundary conditions are determined iteratively until self-consistence is fulfilled. This iterative strategy was suggested by S. Tyagi, who also wrote a prototype solution in MATLAB. Later on, it was implemented DUNE-PDELAB by A. Schlaich. SK was advisor to the Diploma thesis of A. Schlaich, where it was implemented and tested on other benchmark problems.

In many "classic" calculations involving the PB equation the electric potential is only evaluated in the region, where ions can enter. Non-ionic charges are introduced in terms of boundary conditions to the electrostatic potential (Dirichlet-type) or to the normal component of the electric field (Neumann-type), as in text books of classical electrodynamics [196]. Physically, the situation is more complicated. In general the Poisson equation must be solved in the whole space where the field can enter. Applying Dirichlet and Neumann boundary conditions are only approximations but not always an appropriate model for physical interfaces. The iPBS algorithm is a method to tackle general systems, where the geometry of the domains can be arbitrarily complicated, and where dielectric interfaces can be introduced naturally.

One of the reasons why the approximate Dirichlet and Neumann boundary conditions are so abundant, is that the more general situation is more difficult to treat from a numerical point of view. The boundary value problem on some spatial region $\mathcal{A}$

$$\begin{aligned} \Delta\Phi &= \kappa^2 \sinh\Phi \quad \text{on } \mathcal{A}, \\ \frac{\partial\Phi}{\partial n} &= f(r) \quad \text{on } \partial\mathcal{A} \end{aligned} \tag{4.21}$$

with Neumann BCs can be solved with standard finite element methods for elliptic equations. We call a problem of this kind a *restricted problem*. In general, the situation is more complicated, because multiple regions with different permittivities are to be considered. The strategy of iPBS lies in determining the boundary condition f iteratively from solutions of the restricted problem only. This can be achieved by performing an explicit Coulomb integration.

In Fig. 4.15 we depict the geometry we have in mind: A system $\mathcal{S}$ consists of a region $\mathcal{A}$ filled with an electrolyte and a region $\mathcal{B}$, composed of several

regions $\mathcal{B}_i$ without ions. The permittivity in region $\mathcal{A}$ is given by ε_A and in every region $\mathcal{B}_i$ by ε_i. The surfaces of the regions $\mathcal{B}_i$ can be charged. With the iPBS method, it is also possible to consider cases, in which the charge density depends on the local electrostatic potential as the result of a chemical dissociation reaction. We neglect this possibility in this thesis.

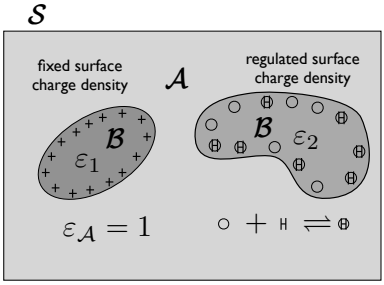


Figure 4.15: Schematic sketch of the systems considered with the iPBS algorithm. A system $\mathcal{S}$ consists of a region $\mathcal{A}$ where ions are described by the PB equation and several regions $\mathcal{B}_i$ that do not contain ions. These regions may contain charges on their surfaces that are either fixed, or regulated by some regulation mechanism. The dielectric permittivity is constant in $\mathcal{A}$ and each of the $\mathcal{B}_i$, but may jump at interfaces.

The physical problem that we want to solve can be expressed in terms of a combination of the Poisson equation (eq.4.6) and the Poisson-Boltzmann equation:

$$\nabla \cdot (\varepsilon \nabla \Phi) = \begin{cases} 8\pi l_B c_0 \sinh \Phi & \text{on } \mathcal{A}, \\ 0 & \text{on } \mathcal{B}_i, \\ -4\pi l_B \sigma_{\mathrm{f}} & \text{on } \partial\mathcal{A}. \end{cases} \tag{4.22}$$

Here σ_{f} denotes a fixed, spatially varying charge density on the interface $\partial\mathcal{A}$. We apply reduced units, which allows us to keep the notation shorthand. We explicitly keep the dielectric constant in all media, and assume that the Bjerrum length l_B is defined in some reference medium (e.g. vacuum) relative to which the permittivity is defined.

At dielectric interfaces, the electric field is discontinuous. In the presence of

a surface charge σ_f the discontinuity is given by

$$\varepsilon_A \boldsymbol{E}_A \cdot \boldsymbol{n} - \varepsilon_{B_i} \boldsymbol{E}_{B_i} \cdot \boldsymbol{n} = 4\pi\sigma_f, \tag{4.23}$$

where $\boldsymbol{n}$ denotes the normal vector of the surface an $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ the electric field on both sides of the discontinuity. Eq. 4.23 can be rewritten as

$$\boldsymbol{E}_A \cdot \boldsymbol{n} = \frac{4\pi l_B}{\varepsilon_A + \varepsilon_B}\sigma_f + \frac{\varepsilon_B}{\varepsilon_B + \varepsilon_A}\left(\boldsymbol{E}_A + \boldsymbol{E}_B\right) \cdot \boldsymbol{n}. \tag{4.24}$$

Obviously, $\boldsymbol{E}_A \cdot \boldsymbol{n}$ is exactly the boundary value f in the boundary condition of Eq. 4.21 we seek to determine. The sum $(\boldsymbol{E}_A + \boldsymbol{E}_B)$ is independent of the charge the particular boundary segment and we will derive an expression for it in the following.

Analogously to the derivation of the ICC$\star$ algorithm in Sec. 4.1.3 we rewrite the full PB equation and introduce the induced charge density σ_{ind}:

$$\Delta\Phi = \begin{cases} 8\pi l_B c_0/\varepsilon_A \sinh\Phi & \text{on } \mathcal{A}, \\ -\dfrac{4\pi l_B}{\varepsilon_A}\left(\dfrac{2\varepsilon_A}{\varepsilon_A+\varepsilon_B}\sigma_{\text{f}} + \underbrace{\dfrac{\varepsilon_A}{2\pi l_B}\dfrac{\varepsilon_A-\varepsilon_i}{\varepsilon_A+\varepsilon_i}\nabla\Phi\cdot\boldsymbol{n}}_{\sigma_{\text{ind}}}\right) & \text{on } \partial\mathcal{B}_i. \end{cases} \tag{4.25}$$

To obtain the electric field in the whole system $\mathcal{S}$ without solving the Poisson-Boltzmann equation everywhere, we use the Greens function G of the Laplace operator in $\mathcal{S}$. In this way, we can express the electric field in terms of integrals over $\mathcal{A}$ and the interfaces $\partial\mathcal{B}_i$:

$$\begin{aligned} \boldsymbol{E}_C(\boldsymbol{r}) =& 2l_B/\varepsilon_A \int_{\mathcal{A}} \nabla_{\boldsymbol{r}} G(\boldsymbol{r},\boldsymbol{r}')\, c_0 \sinh\Phi(\boldsymbol{r}')\, \mathrm{d}V' \\ & - l_B/\varepsilon_A \sum_i \int_{\partial\mathcal{B}_i} \nabla_{\boldsymbol{r}} G(\boldsymbol{r},\boldsymbol{r}') \left[\frac{2\varepsilon_A}{\varepsilon_A+\varepsilon_B}\sigma_f(\boldsymbol{r}') + \sigma_{\text{ind}}(\boldsymbol{r}')\right] \mathrm{d}A' \end{aligned} \tag{4.26}$$

The so far unknown expression $\boldsymbol{E}_A + \boldsymbol{E}_B$ in Eq. 4.24 can be identified as twice the field created by all charges in the system, except for the surface segment that is currently considered. Hence we can write

$$f(r) = \frac{4\pi l_B}{\varepsilon_A + \varepsilon_B}\sigma_f + \frac{2\varepsilon_B}{\varepsilon_B + \varepsilon_A}\boldsymbol{E}_C \cdot \boldsymbol{n} \tag{4.27}$$

This equation is the heart of the iPBS algorithm. If a function f can be found that fulfills Eq. 4.27, the solution of the restricted problem with this BC coincides with the solution of the full problem. A further complication lies in the fact, that the induced charge density σ_{ind} is also unknown. The fix point structure of this equation, however, make it possible to formulate a simple iterative procedure to construct both functions at the same time.

The iteration is constructed analogously to the ICC$\star$ algorithm. Initially we choose

$$f^0 = 4\pi l_B/\varepsilon_A \sigma_f \tag{4.28}$$

and induced charges

$$\sigma^0_{\text{ind}} = 0 \tag{4.29}$$

and solve the restricted problem with these boundary conditions. Then we construct subsequent improvements to both f and σ by the update rules

$$f^{i+1} = (1-\alpha_f)\, f^i + \alpha_f \frac{4\pi l_B}{\varepsilon_A + \varepsilon_B}\sigma_f + \frac{2\varepsilon_B}{\varepsilon_B + \varepsilon_A}\boldsymbol{E}_C \cdot \boldsymbol{n} \tag{4.30}$$

and

$$\sigma^{i+1}_{\text{ind}} = (1-\alpha_\sigma)\, \sigma^i_{\text{ind}} + \alpha_\sigma \frac{1}{8\pi l_B}\frac{1-\varepsilon}{1+\varepsilon}\boldsymbol{E}^i \cdot \boldsymbol{n}. \tag{4.31}$$

The field at the boundary is evaluated from Eq. 4.26. As this expression depends on the potential Φ, it is is necessary to solve the restricted problem in every iteration step. In our investigations the above iterations converged without convergence problems if $\alpha_f = 0.7$ and $\alpha_\sigma = 0.2$ were used. These iterations again are part of the class of successive overrelaxation (SOR) methods.

4.4.3 The Finite Element solution of the PB equation

We applied a finite element method (FEM) for the solution of the restricted problem. In the following we give a brief introduction to the applied method. Comparing to the vast literature and modern developments our method is classical. It is not much more sophisticated than that formulated by Courant int 1943 [233], which is often quoted as the birth of the finite element method [234]. Using a FEM was initially chosen because its resolution can be spacially adjusted. In systems governed by the PB equation this is very helpful as the Debye length can be much smaller than the system size, and thus a fine resolution next to charged surfaces and a coarse resolution far away from

charged surfaces is desired to keep the computational effort low. It was technically realized using the DUNE framework [235], which will be described below. We apply the notation and concepts from the DUNE-PDElab Howto document [236].

The key idea in the finite element method is to approximate the solution of a PDE with a sum of functions which are nonzero only in a small region of space, and which can be described by a finite set of coefficients. In boundary value problems an algebraic set of equation for the coefficients is derived, which is subsequently solved numerically. Here, we sketch the FEM with Lagrangian triangular elements with polynomial basis functions of degree k, so called Pk elements. P stands for "pyramid" in this abbreviation. In all cases investigated with IPBS the cylindrical symmetry was used to decrease the dimension of the system to two. Three-dimensional systems were only investigated to prove the functionality of our implementation.

The main ingredient for an FE based solution of a PDE are the elements. The nomenclature here is somewhat ambiguous. Speaking loosely, the term element refers an element of the decomposition of the domain in which the PDE is solved. In mathematical terms [237], an element is a triple composed of (a) a domain $\hat{\Omega}$, (b) a set of linear independent basis functions $\{\psi_i\}$ defined on $\hat{\Omega}$, and (c) a set of functionals that are defined on the space spanned by the functions ψ_i. These two notions can usually be clearly distinguished by context.

The computational domain is decomposed into n regions with simple shapes, the triangulation. It typically consists of triangles or rectangles in two dimensions, or of pyramids or cuboids in three dimensions. All elements can be mapped to a *reference element* by an affine transformation μ_e, the element transformation. Here e is the index of the corresponding element. This reference element corresponds to the domain $\hat{\Omega}$ of the mathematical definition of an element. For simplicity we will assume the same reference element for all elements. We additionally assume a *conformal* triangulation: all vertices of the triangulation must be a vertex to all adjacent triangles.

The function space is chosen to be the space of all polynomials of degree k on $\hat{\Omega}$. One important trick now lies in a good representation. One possible choice is a Lagrangian basis, that is also used in Lagrangian interpolation. One the reference element certain points are chosen, so called nodal points, on which all functions of the basis set are zero, but one, which equals unity. The concept of a nodal basis is sketched in fig. 4.16. The nodal points are now chosen such that, under reverse application of the finite element transformation, every image of a nodal point that is at an element boundary is not only a nodal

point on one element but on all adjacent elements. This is always possible if the triangulation is conformal. The corners of the triangle are typical nodal points, and e.g. for P2 elements also the midpoints of all intersections. In this nodal representation we tag the basis entry by the index of the nodal point where it is nonzero, and denote it by ψ_l. The number of the necessary nodal points depends on the order of the polynomials taken as basis. In two dimensions, for polynomials of order one, three points are necessary, for P2 elements six, and for order three polynomials ten.

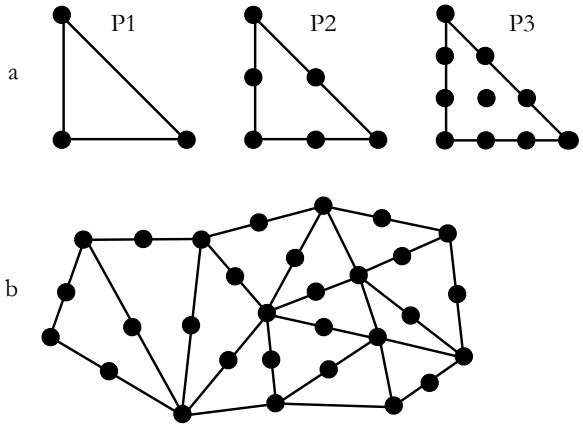


Figure 4.16: (a) Triangular Lagrangrian Elements of order 1, 2 and 3. The reference elements are depicted as triangles, and the nodal points a circles. (b) Conformal triangulation of an arbitrary region, including nodal points for P2 elements. Nodal points of adjacent elements fall on top of each other. Every function of the discrete function space in which solutions are searched for is uniquely characterized by its value on the nodal points.

A global function space is constructed by transforming every point in $\mathcal{A}$ to a point in $\hat{\Omega}$ using the appropriate element transformation. In a nodal representation, continuity of the global functions can be ensured by using the same coefficient for those basis functions, which share the nodal point in global space. We denote the set of pairs of elements and basis function indices (e, l) which share the global nodal point with index j by $\mathcal{L}(j)$. Then the global

basis function j can be written as:

$$\Psi_j(x) = \sum_{(e,l)\in L(j)} \psi\left(\mu_e^{-1}(x)\right)\chi_e(x). \tag{4.32}$$

Here χ_e is the characteristic function of element e, which ensures that every basis function can only contribute if evaluated within $\hat{\Omega}$. Thus finally we have constructed a basis of continuous functions in $\mathcal{A}$, where every basis function has only a small support. Different elements, now meant in the mathematical sense, allow for the construction of a different global basis. With this method it is e.g. also possible to construct a basis continuous in the first or higher derivatives, but we stick to the most basic case.

Every global nodal point corresponds to one degree of freedom. In the following section we will show how it is possible to derive a set of algebraic equation from the PB equation. The solution of this set of equations thus yield the coefficients needed for the global basis functions given in Eq. 4.32.

The starting point to solve the PB equation with the FEM is writing down its weak form. Let u and v be a square integrable functions. A function u is called weak solution of the PB equation if:

$$\int_{\mathcal{A}} v\Delta u - v\kappa^2 \sinh u = 0 \tag{4.33}$$

for all functions v. Functions u for which this condition might hold are called trial functions, functions v are called test functions. For practical purposes the vector space of admitted functions $\mathcal{H}$ is restricted to those constructed before. In the FEM this is done not only for the space of trial functions, but also for the test function space. Both are restricted to the span of the global basis. A function is called weak solution with respect to $\mathcal{H}$ if Eq. 4.33 holds for all test functions $v \in \mathcal{H}$. We can thus construct n independent equations, by setting $v = \Psi_j$, for all $0 < j < n$. Eq. 4.33 can be reformulated by means of Green's first identity:

$$-\int_{\mathcal{A}} \left(\nabla\Psi_j \cdot \nabla u + \Psi_j \kappa^2 \sinh u\right) \mathrm{d}V + \int_{\partial\mathcal{A}} v\nabla u \cdot n \,\mathrm{d}A = 0. \tag{4.34}$$

Noting that the support of Ψ_j consists of only a few elements, it is obvious that the integral in Eq. 4.34 can also be restricted to only a few elements. Neumann BCs are incorporated easily in the above formulation by replacing $\nabla u = f$ in the surface integral. We expand u in the global basis, with coefficients u_i

and thus find a nonlinear set of equations for u_i. For a given set (vector) of coefficients $/u$ the RHS of all n equations above is called the residual $\mathcal{R}(\boldsymbol{u})$. Finding the solution can be formally written as:

$$\mathcal{R}(\boldsymbol{u}) = 0. \tag{4.35}$$

Only for linear PDEs the residual is linear in $\boldsymbol{u}$. In general, including the PB equation, a nonlinear set of algebraic equations must be solved. We apply a Newton method to solve this problem. The Hessian of $\mathcal{R}$ is calculated by numerical differentiation. In every step of the Newton method, a sparse linear system must be solved. This can easily be exchanged. We obtained the fastest convergence with a biconjugate gradient method with a successive-overrelaxation preconditioner.

The realization of iPBS in DUNE

The DUNE software framework was written to assist in solving problems similar to that given above. This supports not only FE methods, but also Finite Volume methods and Discontinuous Galerkin methods. It consists of the following submodules:

- dune-grid [238, 239]: Management of triangulations. It provides an abstract interface for different *grid managers*, other software that stores the actual data. The grid managers used in IPBS are UG [240] and ALU [241, 242]. The grid managers can be used interchangeably. Dune-grid supports a domain decomposition between processes and provides an interface for MPI communication of data.

- dune-localfunctions: Provides the elements in the mathematical sense. This includes Lagrangian elements of arbitrary order in two and three dimensions. Also more complex elements are available.

- dune-istl [243, 244]: A library of iterative linear solvers and interface to third-party linear solvers. Built-in solvers are (bi-) conjugate gradient solvers, and multigrid-solvers and with different preconditioners. External solver that can be used are SUPERLU, and PETSc.

- dune-pdelab [245]: a module that connects those mentioned above so that the solution of discrete problems is convenient. The software concepts follow very closely the mathematical constructs.

In the IPBS program, the following steps are performed:

1. Reading in a configuration file.
2. Reading in a mesh file.
3. Solution of the restricted problem, using Lagrangian elements of arbitrary order.
4. Calculation of the electric field using Eq. 4.25.
5. Update of regulated and induced charges and the flux, using Eqs. 4.31 and 4.30.
6. Analysis.

Steps 3-5 are iterated until the relative change of each quantity is below a certain threshold, typically 10^{-4}.

4.4.4 Application to the free energy barrier of dsDNA

The method outlined above was used to calculate the free energy barrier that a DNA molecule has to overcome when entering a pore. The geometry we used was constructed as closely as possible to that used in the MD simulations: a cylindrical pore with smoothed openings in a flat membrane. The DNA molecule was modelled as a charged cylinder of 1 nm radius. The surface charge density was calculated from the linear charge density used in the MD simulations. The cylinder symmetry of the problem could be used to reduce the problem to an effective two-dimensional problem. The surface of the pore was iterated to take into account different values of the dielectric permittivity.

For each parameter set, two calculations were performed to obtain a free energy barrier: one in which the DNA is far away from the pore, and one where it is inside. The free energy of both states is calculated from Eq. 3.91, and the difference corresponds to the free energy barrier opposing the translocation process. In Fig. 4.17, typical meshes we used the calculations are shown. The density of triangles can be increased in the vicinity of the DNA and the pore, and near the box boundary the resolution is decreased.

The pore radius, salt concentration and pore length were varied. The DNA molecule was always kept 2 nm longer than the length of the pore. For each parameter set two types of simulations were performed: A pair of simulations in which the dielectric constant of the membrane was set to $\varepsilon = 2$ and one in which the dielectric constant was set to $\varepsilon = 80$. With all this data we can access the whole parameter space investigated with the MD simulations, plus

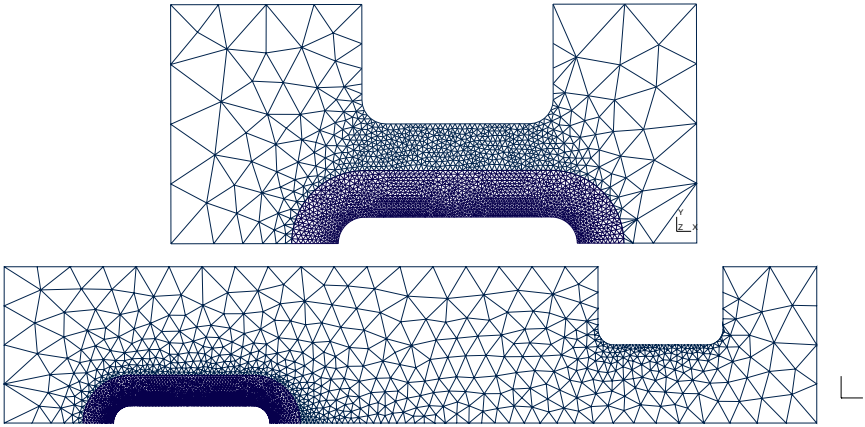


Figure 4.17: The meshed used to calculate the free energy barrier of a 10nm DNA fragment in a pore of length 8 nm. The free energy of both configurations was evaluated; the difference corresponds to the free energy barrier of the DNA.

we can consider different pore lengths. The calculations did not take longer than a day on a single CPU.

The results are depicted in Fig. 4.18a and 4.18b. In Fig. 4.18a we display the free energy barrier obtained for a 10 nm DNA strand in an 8 nm pore as a function of the pore radius 10 mM and 100 mM electrolyte concentrations. We obtain the same behaviour as in the MD simulations. The barrier drops rapidly from more than 10 k_BT when the pore radius increases from two to four to six nanometers. The decay for 10 mM is slower than for 100 mM reflecting the longer Debye length of around 3 nm compared to around 1 nm. When assuming a dielectric constant of $\varepsilon = 80$ of the membrane material, the barriers are almost by a factor of two lower than for $\varepsilon = 2$.

In Fig. 4.18b we display the free energy barrier we obtained for pores of different lengths, and DNA molecules of accordingly slightly larger length. The barriers are divided by the pore length. The free energy barriers for 10 mM and 30 mM electrolyte concentration are shown, but all other concentrations investigated exhibit the same trends. The free energy barrier per unit length increases with increasing pore length and converges to the value calculated from the one-dimensional theory in Sec. 4.4 for the respective concentration. The convergence is accelerated by a dielectric contrast between the membrane

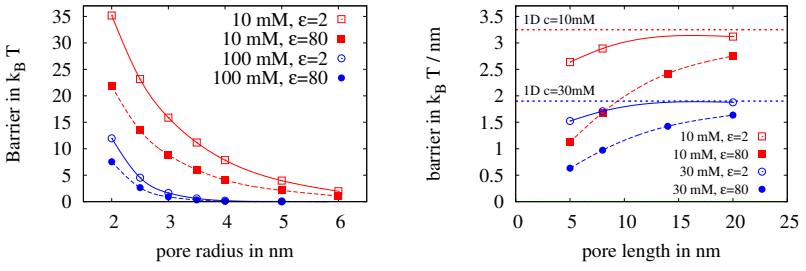


Figure 4.18: The free energy barrier obtained for a DNA molecule. (a) The free energy barrier for a 10nm strand of dsDNA in a 5nm pore of length 8nm for two electrolyte concentrations. 10 mM is displayed as squares, and 100 mM is displayed as circles. We compare the situations between a pore permittivity of $\varepsilon = 2$ to $\varepsilon = 80$. (b) The free energy barrier per unit length of DNA molecule which is 2nm longer than the pore. Symbols are chosen as before, expect for the dataset for c =100mM is replaced by the dataset of c =30mM. In the limit of long pores, both converge against the result of the one-dimensional PB model (Sec. 4.4), which shown as horizontal lines.

material and the surrounding water as neutralization is increased. For the 8 nm pore considered above without dielectric contrast around 50% of the limiting value is reached, while the dielectric contrast the fraction is around 80%.

These results conform with our arguments presented above. In the limit of an infinite pore, the free energy barrier is given by the simple one-dimensional model of Sec. 4.4. The question for which pore length this model is adequate depends on the dielectric properties of the membrane materials. Low-ε media reduce this necessary length so that for the 8 nm long pores considered in the MD simulations the one-dimensional model is applicable. Without dielectric contrast the necessary length is much longer, around 20 nm. This finding is in agreement with the neutralization data from the MD simulations shown in Fig. 4.13a.

4.5 Summary

In this chapter have investigated an important questions regarding DNA translocation: Under which conditions do electrostatic interactions cause significant free energy barriers? We have presented the ICC$\star$ algorithm, an algorithm suited for including dielectric boundary forces into coarse grained molecular dynamics simulations. This algorithm is an extension to an existing Coulomb solver, where additional charges represent the dielectric boundaries. Their magnitude is determined self-consistently in every simulation step. In course of this work, this algorithm was improved and extended and can now be used in parallel molecular dynamics simulations for surfaces of arbitrary shape. The dominant additional computational cost stems from the increased number of particles with electrostatic interactions. The complexity of the underlying electrostatics algorithm remains unchanged when taking into account induced surface charges.

With this method, we investigated the free energy barrier for 10 nm-long, stiff rods, resembling a fragment of double-stranded DNA. We could show that if the pore radius is comparable to or smaller than the Debye length, significant free energy barriers appear. This can be as high as several k_BT per nanometer pore length, and hence can suppress translocation even at high applied voltages. When taking into account dielectric boundary forces, these barriers are significantly enhanced, typically by a factor of two. We were able to show that this barrier only partly originates from repulsive dielectric boundary forces, but also happens as a consequence of entropic loss of the counterion cloud. Especially in the presence of dielectric contrast, the system minimizes its free energy by dragging more counterions into the pore. This enhances the screening of the bare DNA charge, but at the expense of counterion entropy.

For flexible molecules, very similar effects are apparent, but somewhat increased in strength. We performed simulations with a flexible polyelectrolyte chain resembling single-stranded DNA with a contour length of 45 nm. With a pore of twice the radius and a flexible molecule with only half the charge density of the stiff molecule considered before, we obtained similar free energy barriers. This results from the chain entropy which drives a longer section of the contour into the pore and closer to the pore boundary than in the case of a stiff rod. Again, with increasing salt concentration, the barrier height decreases. Dielectric contrast increases the height of the barrier, again with contributions due to counterion entropy and direct interaction.

For the stiff dsDNA piece, we could show, that alternatively to a particle-based

model, a simple Poisson-Boltzmann model is surprisingly accurate. The very simple one-dimensional cylinder-in-cylinder model leads to almost identical free energy barriers. The deviations to the particle simulation models are not larger than errors caused by the uncertainty of experimental conditions. Interestingly the agreement with simulations including dielectric contrast was much better than with that without dielectric contrast. From the data of particle-based simulation we could show that for the simulation with dielectric contrast the degree of neutralization of the interior of the pore was larger. As the neutrality of the pore interior is a fundamental assumption of the PB model, this explains why the agreement with simulations with dielectric contrast is better.

Using the iPBS algorithm, we could support this argument. In the iPBS algorithm the Poisson-Boltzmann equation is solved by means of the Finite Element Method. The boundary conditions are iterated in a fashion similar to ICC$\star$. This allowed us to compute free energy barriers in essentially the same geometry as we used for particle simulations. We could show that, in the limit of an infinitely long pore, the free energy per pore length converges against the prediction of the one-dimensional PB model. For the pore length considered in the simulations before, the result with and without dielectric contrast deviate. As the dielectric contrast enhances the neutralization, the barrier measured with dielectric contrast converges faster to the limiting value.

Two important questions naturally appear at this point: (a) "How accurate are the models we have applied?" and (b) "What are the implications of our findings for the dynamics of DNA translocations". The next chapter contains interesting insight especially about point (a), and therefore we comment on these question in detail in the concluding chapter 6.

5 Nanopore conductance in the presence of DNA

Most nanopore-based experiments are based on measuring the ion current through the pore. A macromolecule crossing the pore is observed as a modulation. These current modulations they contain valuable information about the molecules transported through the pore. But what is the origin of the modulation? For the Coulter Counter, developed in the 50s to count red blood cells, the answer is simple. Cells obstruct a certain fraction of the pore cross section, and this increases the resistance of the circuit. The Coulter principle states that the relative change of the current current is roughly proportional to ratio of the cell volume to the pore volume [246]. Qualitatively, the same observation is made in most nanopore experiments. In this chapter we will see that in the case of charged biomolecules, such as DNA, the situation is more complicated. In addition, electrokinetic aspects must be taken into account. An interesting observation is, for example, that when a low electrolyte buffer concentration is used, pulses of conductivity increase can be observed [5, 247, 248]. This already indicates that the situation is more complex than the Coulter Counter.

So far the understanding of the current modulations was not quantitative. This unsatisfactory situation stimulated our work on understanding the relevant mechanisms. In this chapter, we focus on understanding the experimental data of Smeets et al. [5]. They systematically investigated the role of the salt concentration in translocation experiments with double-stranded DNA (dsDNA) in solid state nanopores. For concentrations above around 350 mM a current reduction was observed, and below this concentration a current increase was observed. We consider a quantitative understanding of this crossover point an important benchmark, because this point should be insensitive to many details, like the exact geometry of the pore.

We started our investigations by applying the model suggested by van Dorp et al. [15] to nanopore conductivity. It has been used successfully to predict the mechanical force necessary to stall a DNA molecule in a nanopore while

it is subject to an electric field. In this model an infinite charged cylinder representing the DNA is embedded in an infinite uncharged cylinder, the pore. The electrokinetic equations are solved in this simple one-dimensional geometry. We refer to it as *the electrokinetic infinite cylinder model* to distinguish it from our other models. The electrokinetic infinite cylinder model does not predict the conductance crossover in agreement with the experiments, but predicts a crossover at much higher concentrations. As this model fails to describe the behaviour quantitatively, we chose to investigate the system with a hybrid of molecular dynamics (MD) simulations and the Lattice Boltzmann (LB) method. The geometry we applied was chosen to represent exactly the electrokinetic cylinder model. The challenge is to develop this well-established method for qualitative insight towards a quantitatively reliable tool. As a third step, we constructed a model with atomistic resolution. The geometry was also designed to resemble the electrokinetic cylinder model as closely as possible. We thus gradually increase the level of detail of the representation. By taking into account the atomic details of the DNA we can, indeed, understand the experimental data, although the coarse-grained modelling approaches turn out to be surprisingly accurate.

This chapter is organized as follows. First, in Sec. 5.1, we give a more detailed description of the experiments by Smeets et al. [5] and van Dorp et al.. In the first set of experiments the conductance crossover effect was investigated in detail. In course of the second experiments the electrokinetic infinite cylinder model has been developed. We present a detailed account of this model and its prediction for the conductance. In Sec. 5.2 we describe the MD/LB model, including several benchmark tests we have performed. In Sec. 5.3 we describe the all-atom model, and the results obtained for the conductance. This model is also used to study DNA molecules with different base pair compositions, where we make use of a visualization tool that was developed in a joint project with the visualization department within the SFB 716.

5.1 Current blockade or enhancement

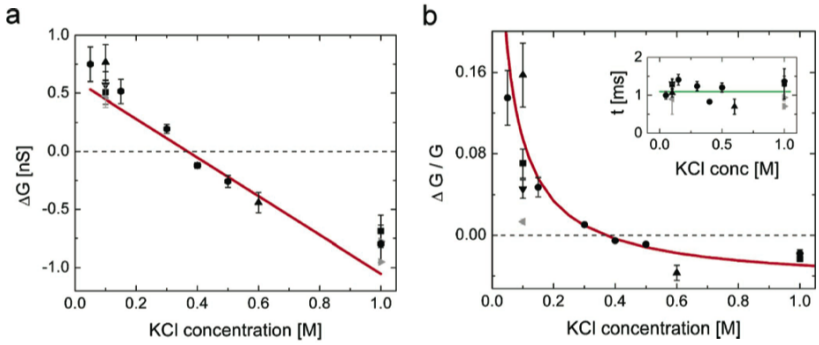
In this section we present the starting point of our project. We discuss the experiments of Smeets et al. [5] where the crossover between current blockade and current enhancement has been investigated systematically for the first time. They already gave the main qualitative arguments, but our simulations presented in the following sections allow us to discuss their validity quantitatively. In context of the experiments of van Dorp et al. [15], the electrokinetic

infinite cylinder model was presented. While originally applied to calculate forces acting on DNA molecules in nanopores we apply it to calculate the current modulation caused by a DNA molecule. We discuss its prediction in some detail as in the following all simulation results will be compared to it.

5.1.1 The reference experiments

In their experiments Smeets et al. investigated systematically the role of electrolyte concentration for the current magnitude [5]. They performed translocation experiments with double-stranded DNA (dsDNA) under different electrolyte conditions. The buffer concentration was changed between very low salt conditions, around 0.05 M, and 1 M. They used solid-state nanopores of around 10 nanometer diameter which were drilled with a transmission electron microscope into a free-standing silicon nitride membrane. Short resistance pulses were observed as in other experiments, indicating DNA molecules traversing the pore. For high salt concentrations, they observed a reduction of the current by a few percent. Under low salt concentration conditions, the pulses inverted their sign. Then the current was enhanced by the presence of a DNA molecule. Their findings are depicted in Fig. 5.1. The nanopore conductance corresponds to conductance changes between -1 and +1 nS depending on the salt concentration. The crossover point between current enhancement and current reduction is between 300 and 400 mM. If normalized by the open pore current without DNA, the modulations are between +16% and -4%, depending on the salt concentration.

In Ref. [5], this observation was explained by the competition of two effects. First, the volume occupied by the DNA molecules can not be occupied by ions. This leads to a steric exclusion of ions, and hence a reduction of the current. Second, DNA is a highly charged object. The electrostatic interactions guide additional ions into the pore, its counterion cloud. The number of ions displaced from the pore should be roughly proportional to the salt concentration, while the number of additional counterions is approximately constant. With these assumptions, Smeets et al. obtained a good fit of their data by assuming a cylindrical geometry of 10 nm diameter and assuming that the DNA obstructs circular cross section of 2.2 nm diameter. In their analysis, the counterion number was used as adjustable parameter. They obtained a value of 0.58 K^+ ions per base pair. This is only 29% of the "bare" charge of DNA. This discrepancy is interpreted as the DNA's effective charge. They noted that assuming two K^+ per base pair and but a reduction of the mobility to 29% of its bulk value leads to the same conductance modulation. Thus,



(a) (b)

Figure 5.1: (a) Absolute and (b) relative change in conductance ΔG during DNA translocation events as a function of the salt concentration of the solution. At low salt concentration, the conduction increases by almost 1 nS or 15 %, while it drops at high salt concentration by -1 nS or around -4 %. From Ref. [5].

the reduced counterion mobility is a different interpretation of the parameters obtained from the fit.

Significant effort has been taken to make more physical quantities accessible to direct observation in these experiments. An interesting example is the approach taken by Keyser et al. in the Dekker lab in Delft. The setup is sketched in Fig. 5.2. In these experiments dsDNA was attached to a colloidal particle, which could be controlled by a laser tweezer. The colloidal particle was moved to the vicinity of the nanopore. On applying an electric field the DNA moved into the nanopore. This insertion process was measured by monitoring the electric current. When the particle was moved away from the pore the current exhibited an abrupt drop to the baseline current; the DNA was pulled out of the pore.

From the displacement of the colloid in the laser trap the force acting on the colloid, and hence the forces acting on the DNA can be deduced. The observed forces were dependent on the radius of the pore. For pores of around ten nanometers in diameter forces around 220 pN/V were observed, while for larger pores (up to 80 nm) the force decreased to around 100 pN/V. A naive model where the force is assumed to be the DNAs bare charge density mul-

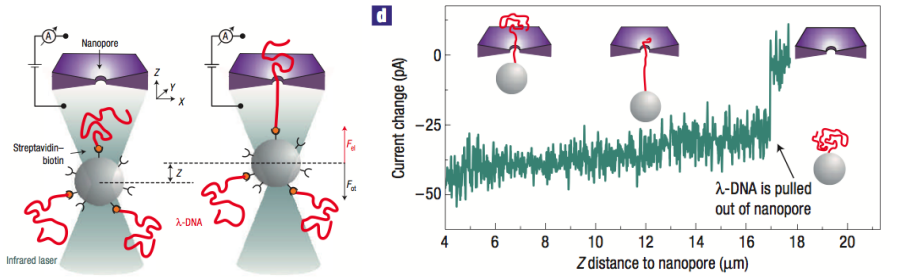


Figure 5.2: (a) Experimental setup of the force measurements in Ref. [249]. DsDNA is attached to colloidal particle in laser tweezers. By placement of the particle in front of a nanopore and application of an electric fied, the DNA can be reversibly inserted into the pore. (b) Conductance measurement in the same setup. As the particle is moved away from the membrane, the current decreases and eventually DNA leaves the pore. Both illustrations from [249].

tiplied with the applied voltage, however, yields a much higher force of more than 900 pN/V. The model described in the following section was constructed to understand this behaviour quantitatively.

5.1.2 The electrokinetic infinite cylinder model

Before turning to the question of modulations in the resistance we briefly recapitulate the nanopore conductance model for cylindrical nanopores of finite size presented in [250]. Using this simple model we can argue, that for the considerations below it is sufficient to assume a cylindrical pore, where the entrance and exit regions are neglected.

Fig. 5.3 displays the geometry and the equivalent circuit. A cylindrical pore of length L and radius R connects two reservoirs containing electrodes between which the voltage is applied. The total resistance of the system can be expressed in terms of a series of access resistances R_{ac} on both sides of the membrane, and the resistance of the cylindrical pore R_{cyl}. In 1968 [251], Hille considered the conductance of a single pore in an infinitely thing membrane. He approximated the surrounding electrolyte as a homogeneously conducting medium with a conductivity σ, and a pore as a hole in an infinitesimally thin sheet. He approximated the equipotential planes as hemispheres, and this

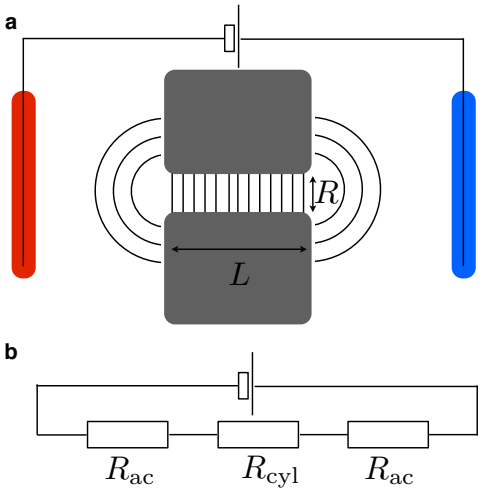


Figure 5.3: (a) Sketch of the considered geometry. A planar membrane of width L contains a cylindrical pore of radius R. Two electrodes are inserted on the different sides of the membrane. Near the pore, the equipotential planes (shown as black lines) are hemi-spherical and inside they are planar. (b) Equivalent circuit of nanopore conductance. The applied voltage drops over a series of an access resistance R_{ac}, a resistance of the pore R_{cyl} and an identical access resistance.

lead him to the following expression for the conductance of an infinitely thin circular pore.

$$G = \sigma \frac{\pi}{4R}. \tag{5.1}$$

Hall later realized that the exact result of the capacity of a charged disc could be applied to improve this result [252]. The differential form of Ohm's law

$$\nabla \cdot \sigma \boldsymbol{E} = 0 \tag{5.2}$$

that describes a conducting medium is formally identical with the Laplace equation of electrostatics in the absence of further charges. By symmetry, the symmetry plane of the system must be on constant potential, and hence the potential in the conductance must follow the same form as the potential around a disc on constant potential. The normal derivative of the potential is, in one case, proportional to the current density through the midplane, and

in the other case proportional to the charge density. This consideration leads to the following expression of the conductance of an infinitely thin pore of:

$$G = \sigma \frac{1}{2R}. \tag{5.3}$$

The conductance of a cylinder of radius R and length L is:

$$G_{\text{cyl}} = \sigma \frac{\pi R^2}{2L}. \tag{5.4}$$

The access resistance of the cylindrical pore is given by half of the resistance of an infinitely thin pore. Using Hall's formula for the access resistance, the conductance of the whole G system can be written as:

$$G^{-1} = \frac{1}{\sigma}\left(\frac{L}{\pi R^2} + \frac{1}{2R}\right) \tag{5.5}$$

$$= \frac{1}{\sigma}\frac{L}{\pi R^2}\left(1 + \frac{\pi R}{2L}\right). \tag{5.6}$$

From this expression it can be seen that if for pores with a much smaller radius than length the access resistance in negligible. In the experiments of Smeets et al. , the radius-to-length ratio of the pores is around 0.12, thus the correction due to a finite pore length is around 20%. In the following considerations, we do not investigate the absolute current, but only how it changes relative to the open pore current. By neglecting the access resistance, we will clearly overestimate the current modulations by a similar percentage, as access resistance will be influenced less by the presence of a DNA molecule. This approximation, however, simplifies the discussion and the crossover point, where the open pore current and the current with DNA are identical, is obviously insensitive to the magnitude of the pore resistance.

In Ref. [15], the authors explain the forces acting on a colloidal particle which was used to hold a DNA molecule in a nanopore also measured by Keyser et al. [51, 249]. They used the electrokinetic equations to model the flow of ions and water and chose a particularly simple geometry. An infinitely long impermeable charged cylinder on the z-axis of the coordinate system with a radius of 1.1 nm representing the DNA molecule is assumed. The pore is represented as an infinitely long impermeable cylindrical tube also centered in the coordinate system. Originally, the radius R of the pore is varied. We, however, choose a fixed radius of 5 nm in correspondence to the experimental data of Ref. [5]. On the surface of both cylinders, no-slip boundary conditions

$\boldsymbol{u} = 0$ for the water are assumed. No pressure and concentration gradients are present, so the solution of the electrokinetic equations is independent of z. No currents and no fluid flow in radial directions can occur, as the cylinders are impermeable ($j^{\pm} = 0$). The solution of this model is very simple: First the model can be solved for the concentrations in radial direction. This can be done by solving the Poisson-Boltzmann equation

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}}{\mathrm{d}r}\Phi\right) = \frac{1}{l_D^2}\sinh\Phi, \tag{5.7}$$

in the radial coordinate r. Here the Laplace operator in cylindrical coordinates has been inserted. l_D is the Debye length. Neumann boundary conditions are assumed at the two cylinder surfaces:

$$\left.\frac{\mathrm{d}}{\mathrm{d}r}\Phi\right|_{r=1.1\mathrm{nm}} = 4\pi l_B\sigma,$$
$$\left.\frac{\mathrm{d}}{\mathrm{d}r}\Phi\right|_{r=5\mathrm{nm}} = 0.$$

Here the cylinder charge density σ is set to the value expected from distributing a line charge density of 5.9 e/nm homogeneously on a cylinder of 1.1 nm radius. This yields a surface charge density of 0.85 e/nm^2. We solved this boundary value problem numerically by means of finite differences using PYTHON.

The concentrations of ions can be obtained from the electrostatic potential simply by applying the Boltzmann distribution:

$$c^{\pm}(r) = c_0 \exp\mp\Phi. \tag{5.8}$$

The body force on the water acts only in axial (z) direction and can be written as:

$$f_z = \left(c^+ - c^-\right)E, \tag{5.9}$$

where E denotes the strength of the externally applied electric field. Assuming a long but finite pore, it is related to the applied voltage V by V/l, where l is the length of the pore.

The Stokes equation reduces to an ordinary differential equation in the z component of the fluid velocity u_z,

$$\eta\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\frac{\mathrm{d}}{\mathrm{d}r}u_z\right) + \left(c^+ - c^-\right)E = 0. \tag{5.10}$$

It can be solved by a double integration. Here, two integration constants appear, that are chosen to fulfill no-slip boundary conditions on both walls.

The force on the counterions is partly transferred into the DNA molecule and can be calculated from the value of the viscous stress

$$\eta \frac{\mathrm{d}}{\mathrm{d}r} u_z \tag{5.11}$$

at the surface of the DNA molecule. Interestingly, the force acting on a DNA is independent of the pore length for a given voltage. Shorter pores lead to higher electric fields, and therefore identical forces if end effects are negligible.

In the simple one-dimensional cylinder-in-cylinder geometry the ion velocities are given by

$$v_z^{\pm} = u_z \pm \mu^{\pm} E. \tag{5.12}$$

In accordance with the experimental values for KCl, we assume $\mu^+ = \mu^-$. In Fig. 5.4 we present the density and velocity profiles obtained from the electrokinetic infinite cylinder model for two different salt concentrations, 100 mM and 800 mM. Qualitatively, both datasets have very similar features. The counterion density decays steeply from a very high value at the DNA to a its bulk value. The decay for 800 mM is more rapid than for 100 mM, reflecting a Debye length of around 0.35 nm and 1 nm respectively. The ion motion is much faster than the water motion which indicates that convection due to electroosmotic flow is only a small correction to the conductivity.

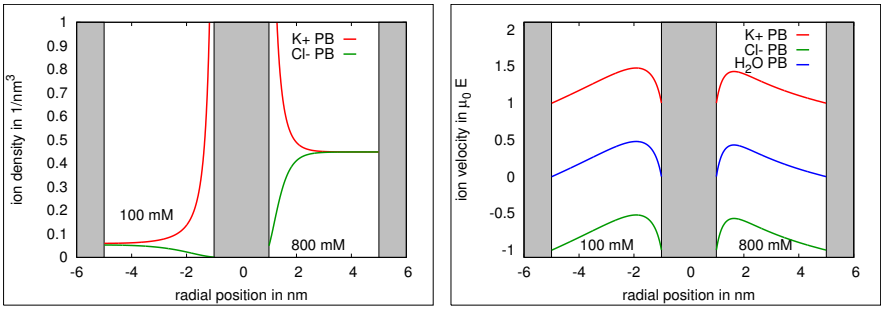
Interestingly, in this simple geometry, the decoupling of the equations can be used, to obtain a form of the electrokinetic equations with only one free parameter. All other constants can be absorbed into scaling factors. Physically the remaining constant can be interpreted the effective hydrodynamic radius of the ions, which is given by

$$r_h = \frac{k_B T}{6\pi\eta D}, \tag{5.13}$$

hence the diameter of a sphere with the same diffusion constant in the same solvent. The ion velocity can be brought to the following form:

$$v = \mu E \left(1 + \frac{r_h}{a} u^\star\right), \tag{5.14}$$

where a denotes a characteristic length of the system and $u^\star$ is the water velocity divided by μE. The hydrodynamic radius thus characterizes the relative importance of the water motion for the ion current. We distinguish the



(a) (b)

Figure 5.4: (a) Ion density profile and (b) ion velocity and water velocity profile as predicted by the electrokinetic infinite cylinder model. We show the results for 100 mM (left half) and an 800 mM buffer concentration (right half). The velocities are normalized by $\mu_0 E$, where μ_0 is the infinite dilution mobility and E the electric field along the pore axis.

two contributions to the ion motion as *direct* velocity and *convective* velocity, and their respective contributions to the total current. For ions with a very small hydrodynamic radius, the convective contribution is negligible, and the transport is dominated by direct motion. Large ions couple stronger to the water, and the convective transport is more important. Matching the hydrodynamic radius of ions is therefore a necessary condition for a quantitatively correct model. This will be very important in the following sections.

To compare the current in the absence of a DNA molecule we apply the same model, just without the inner cylinder representing DNA. As the pore wall is assumed to be uncharged, the solution is trivial. The current is just given by

$$I_{\text{w/o DNA}} = \pi R^2 2 c_0 \mu E. \tag{5.15}$$

The ion velocity profiles can readily be integrated yielding the total current I. We define the following decomposition of the total current: The current density is decomposed into a direct current I_D, which excludes the water motion and an convective current I_W solely due to water motion. The corresponding

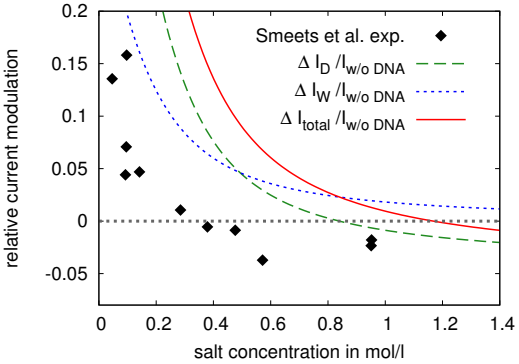


Figure 5.5: Relative modulation of the electric current. Experimental data by Smeets et al. [5] is shown as diamonds and the prediction of the electrokinetic model presented in Ref. [15] is displayed as a solid line. In addition, the prediction of the electrokinetic model is divided into the two contributions defined in the text: a direct contribution caused purely by the change of the number of ions in the pore and a contribution due to ion convection, shown as dashed and dotted line, respectively.

densities j_D and j_W read as:

$$j_D = \left(c^+ + c^-\right) \mu E \tag{5.16}$$

$$j_W = \left(c^+ - c^-\right) u_z. \tag{5.17}$$

I_D and I_W are defined as the corresponding integrals over the pore cross section. In absence of a DNA molecule $I_W = 0$ as no water motion is observed.

For comparison with the experimental data of Ref. [5], we plot the results of this model in Fig. 5.5. We present the relative current modulation

$$\frac{I_{\text{with DNA}} - I_{\text{w/o DNA}}}{I_{\text{w/o DNA}}} \tag{5.18}$$

as a function of the salt concentration c_0 in the bulk solution. It can clearly be seen that the model predicts a crossover between current enhancement and current reduction for $c_0 \approx 1.2\,\text{M}$, in clear quantitative disagreement with the experiments. The direct current modulation reflects the change in the total number of ions upon DNA insertion. At around $c_0 > 0.85\,\text{M}$ the direct current

contribution changes its sign. For larger concentrations the number of ions displaced from the pore due to the volume occupied by the DNA cylinder exceeds the number of counterions brought into the pore by the DNA. The convection always enhances the current, as it makes counterion transport more efficient and only appears in presence of the DNA.

The observed crossover concentrations are significantly higher than those measured experimentally. This matches the observation of Smeets et al., that only by assuming reduced, effective charge of the DNA leads to a consistent prediction of the crossover. The same "trick" would, of course work here. Inserting a smaller charge of the cylinder decrease the conductance in the presence of DNA. We are, however, interested in the physical mechanism, that leads to the decrease of conductance. Therefore, our next step was developing a model which incorporates "more" physics, a model based on molecular dynamics simulations. It is presented in the following section.

5.2 Lattice-Boltzmann/MD simulations

In this section we will present a model of ion transport through nanopores based on coarse-grained simulations. The modeling strategy relies on treating all charged objects as explicit particles, and replace the water molecules by a continuum. As dynamic quantities are considered, a simple dielectric picture will not be enough. The water, as any solvent, assumes an active role in the dynamics. Two very important effects of the presence of water have already been discussed before: electroosmotic flow and hydrodynamic interactions. Electroosmotic flow occurs as a collective phenomenon, where the concerted motion of many ions induces water flow. Its relevance for the conduction process has already been addressed in context of the electrokinetic infinite cylinder model of the previous section. Hydrodynamic interactions are also relevant in ion conduction phenomena. We will show below that they cause, for example, a decrease of the specific conductivity of an electrolyte with increasing concentration. To take both effects into account, we apply the Lattice Boltzmann method (LBM), a lattice-based solvent representation to which particles are coupled. The ion motion is taken into account by Molecular Dynamics (MD). This concept was pioneered by Ladd [253] and Ahlrichs and Dünweg [254]. Before going into the details, we give a more thorough motivation why this method was used. It contains key ideas from Ref. [255].

In liquid systems, momentum is subject to a diffusive motion. Momentum flux occurs in the direction of a gradient of the momentum density. In the

time-dependent Navier-Stokes equation, the kinematic viscosity ν, related to the dynamic viscosity η and the fluid density ρ, by $\nu = \eta/\rho$ takes the role of a diffusion constant. Particles immersed in the liquid are subject to Brownian motion. By taking the ratio of the particle diffusion coefficient D and the kinematic viscosity it is possible to quantify, how much faster momentum is transported through the system than particles. This ratio is defined as the Schmidt number

$$\mathrm{Sc} = \frac{\nu}{D}. \tag{5.19}$$

The hydrodynamic interactions described in Sec. 3.2.5 are derived assuming instantaneous momentum transfer between particles. They are derived in the infinite Schmidt number limit. This assumption leads to a non-local coupling of particles, where the motion of all particles is linked by a many-body hydrodynamic interaction. Under virtually all conditions in liquid systems the Schmidt number is very high, at least in the order of 10^3. Its magnitude will be discussed in more detail below. The non-locality of of hydrodynamic interactions, and the intrinsic structure of a conservation law causes hydrodynamic interactions to be long ranged. In general they can not be cut beyond a certain distance without introducing systematic errors. Therefore typically algorithms scaling $\mathcal{O}\left(n^2\right)$, where n is the number of particles, must be applied to take them into account. This unfortunate scaling behaviour can be avoided by using a novel class of simulation algorithms, *mesoscopic* simulation methods.

The general idea of mesoscopic simulation methods lies in replacing the "real" mechanism of momentum transport by a "synthetic" mechanism that is computationally more convenient. This means that fluid parameters, which are unimportant for the result can be chosen at will from a certain range. The Stokes equation is the low-Reynolds number limit of the Navier-Stokes equation. Simulating a fluid with a small but finite Reynolds number will yield equivalent results, but possibly at lower computational costs. This robustness of the physics laws against details allows for a great freedom in the choice of computational methods. It is possible to reformulate hydrodynamics entirely in a local way. A good example for this is the *dissipative particle dynamics* (DPD) approach [173,174]. In DPD, real molecules are replaced by soft particles which do not only possess a conservative interaction, but also a dissipative interaction based on a pairwise friction. All interactions conserve momentum and therefore Navier-Stokes-like behaviour is recovered. The interactions are short-ranged and hence $\mathcal{O}\left(n\right)$ algorithms for particle methods can be applied. The resulting fluid is highly compressible due to the soft particle interaction. The compressibility can be much larger than the compressibility of the fluid

that is actually considered, as the velocities are sufficiently low that compressibility effects are anyways negligible. The soft potential allows for a larger time step and thus reduces the computational effort. Therefore also the Schmidt number in DPD is typically also much smaller than in reality. Under which conditions this is problematic, is, however, not entirely clear.

We have seen above that for the system consisting of ions and water the hydrodynamic radius of the particles must be matched. This is the key prerequisite to obtain the correct ratio of convective and direct ion transport and therefore must be matched in our mesoscopic approach. The numerical value of the viscosity, on the contrary, is not important. By using appropriate scaling factors we can a posteriori deduce the physical values of all observables.

A particularly successful mesoscopic modelling strategy is the Lattice-Boltzmann method (LBM). It originates from the field of Lattice Gas Cellular Automata (LGCAs). In LGCAs particles are confined to discrete positions on a lattice, and can only assume a velocity which connect lattice positions in one time step [256]. With this crude form of dynamics, also hydrodynamic behaviour can be recovered, if momentum conservation is assured. Entirely local collision rules lead to $\mathcal{O}(n)$ complexity. The LBM is a generalization of LGCAs. Instead of boolean populations, i.e. a certain lattice site is occupied with a particle or not, floating point numbers are introduced which contain the number of particles with a certain velocity at a certain lattice site [257]. This statistical perspective allows to draw the connection to the Boltzmann equation. In the following section 5.2 we give a derivation of the LBM as presented by He and Luo [258]. In their derivation the Boltzmann equation, and not LCGAs, is used as starting point. The nomenclature Lattice-Boltzmann method (LBM) and Lattice-Boltzmann equation (LBE) are used interchangeably. When referring to the LBE typically the mathematical structure of the method is addressed, while the term LBM refers more to the general method and philosophy. We sometimes also apply the term LB fluid because we refer to it by its function in simulations which also contain particles.

Two main strategies have been proposed to couple particles to an LB fluid, the approaches suggested by Ladd [253] and by Ahlrichs and Dünweg [254]. In Ladd's method particles treated in the same way as other boundaries to the LB fluid, with the exception that they can move. In Ahlrichs' and Dünweg's method particle couple to the fluid by momentum exchanged based on mutual friction. As the coupling occurs at a specific spatial point, we refer to this approach also as the point coupling scheme. We will apply the it for the ions. Our representation of the DNA molecule is inspired by the Ladd scheme, with the simplification that the DNA is immobile. Especially the point coupling

has so far rarely been applied as tool with numerical accuracy. We therefore have performed a series of benchmark tests to investigate what accuracy we can expect from it. This is presented in Sec. 5.2.2. One important aspect is investigating the properties of an electrolyte modelled with this model.

Finally we apply the model presented before to the problem of nanopore conduction. We chose a system that resembles closely the electrokinetic infinite cylinder model described before. We systematically compare its predictions to that of electrokinetic infinite cylinder model, and work out how good the agreement is, and where differences occur.

5.2.1 The Lattice Boltzmann Method

Two possible routes can be taken to construct the LBE: First, it is possible to consider the LBE a generalization of LCGAs, as described in [256]. A second approach is the route of kinetic theory, in which the LBE is constructed, as a discretization of the linearized Boltzmann equation. We will take the second route here, as it appears natural, in context of MD simulations to take the route where classical particles, and not lattice particles are involved. We follow the original description by He and Luo [258,259] and that of Shan et al. [260]. We restrict ourselves to the Lattice-Bhatnagar-Gross-Krook (LBGK) form of the LBM.

The Boltzmann equation describes the time-evolution of a dilute gaseous system, where the mean free path is not negligible in comparison to other length scales. The considered quantity is the one-particle distribution function $f(\boldsymbol{r}, \boldsymbol{v})$, the number of particles in an infinitesimal phase space volume around $\boldsymbol{r}$ and $\boldsymbol{v}$. Analogous to the particle number density in Eq. 3.8 it can be defined as

$$f(\boldsymbol{r}, \boldsymbol{v}, t) = \int \mathrm{d}^{3n}\mathfrak{r}\, \mathrm{d}^{3n}\mathfrak{v}\, \tilde{f}(\mathfrak{r}, \mathfrak{v}, t) \sum_i \delta(\boldsymbol{r}_i - \boldsymbol{r})\, \delta(\boldsymbol{v}_i - \boldsymbol{v}). \tag{5.20}$$

Here we have added a tilde to the n-particle distribution function. For non-interacting particles in a potential $V(\boldsymbol{r})$, the time-evolution is given by:

$$\partial_t f + \boldsymbol{v} \cdot \nabla_r f - \nabla_r \boldsymbol{V}(\boldsymbol{r}) \cdot \nabla_v f = 0. \tag{5.21}$$

For free particles, it hence looks exactly like the Liouville equation, except that the arguments are only 3-dimensional and not $3N$-dimensional.

Deviations from free propagation can be incorporated on the right hand side of the equation. Boltzmann suggested a term in which correlations between particles are neglected, appropriate for a dilute gas. With his *molecular chaos assumption* the probability of finding a collision pair with velocities $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is assumed to be proportional to the product $f(\boldsymbol{r}, \boldsymbol{v}_1) f(\boldsymbol{r}, \boldsymbol{v}_2)$. In this approximation, a closed expression for the time evolution of f can be given. The necessary input is the differential cross section of the particles, which describes the outcome of collisions as a function of the relative orientation of the velocity vectors of the colliding particles, that depends on the pair potential.

With Boltzmann's molecular chaos assumption it is possible to close the Boltzmann equation, and connect the microscale of single particles to the macroscale where many collisions are averaged over. A different closure is the one suggested by Bhatnagar, Gross and Krook [261] (BGK). It is an effective theory where no particular underlying mechanism is assumed. The central assumption is that the relaxation towards local thermodynamic equilibrium can be described solely by a linear relaxation process. This is expressed as

$$\partial_t f(\boldsymbol{r}, \boldsymbol{v}, t) = \frac{1}{\tau_r} (f - f_{\mathrm{eq}}) . \tag{5.22}$$

τ_r is called the relaxation time. Obviously, the distribution function converges against the equilibrium distribution f_{eq}, if the system is not driven by external forces. For the equilibrium distribution a Boltzmann distribution

$$f_{\mathrm{eq}} = \frac{\rho}{(2\pi k_B T)^{3/2}} \exp\left(\frac{(\boldsymbol{v} - \boldsymbol{u})^2}{2 m k_B T} \right) \tag{5.23}$$

shifted by the local mean velocity $\boldsymbol{u}$ is assumed. The BGK equation is momentum conserving, but energy is not conserved.

The BGK model fulfills the criterion given by Evans and Morris quoted in Sec. 3.2.5. The collision process is is a completely generic mechanism, but as it conserves momentum it leads to a well defined asymptotic behaviour. It is possible to derive the Navier-Stokes equation from this assumption for large time and length scales. The transport coefficients depend solely on the choice of τ_r.

To obtain the Lattice-Boltzmann equation, we consider again the full Boltzmann equation. Assuming a constant potential, and disregarding the collisions, Eq. 5.21 can be integrated along a line $\boldsymbol{r} + \boldsymbol{v}t$ over a time interval τ.

This leads to

$$f\left(\boldsymbol{r}+\boldsymbol{v}\tau,\boldsymbol{v},t+\tau\right)=f\left(\boldsymbol{r},\boldsymbol{v},t\right). \tag{5.24}$$

This expresses the fact that all particles at $(\boldsymbol{r},\boldsymbol{v})$ travel to $(\boldsymbol{r}+\boldsymbol{v}\tau,\boldsymbol{v})$ in the time interval. For sufficiently small time intervals τ also BGK collision expression can be rewritten as

$$f\left(\boldsymbol{r},\boldsymbol{v},t+\tau\right)-f\left(\boldsymbol{r},\boldsymbol{v},t\right)=\frac{1}{\tau_r}\left(f-f_{\mathrm{eq}}\right). \tag{5.25}$$

To obtain the Lattice-Boltzmann equation this equation is discretized by choosing a cubic lattice with discrete lattice points $\boldsymbol{r}_{mno}=am\boldsymbol{e}_x+an\boldsymbol{e}_y+ao\boldsymbol{e}_z$, with a lattice constant a. The vectors $\boldsymbol{e}_x,\boldsymbol{e}_y,\boldsymbol{e}_z$ denote the unit vectors in the three spatial directions. The velocity space is discretized accordingly into discrete velocities $\boldsymbol{v}_{ijk}$, where the spacing is given by the lattice constant divided by the time step,

$$\boldsymbol{v}_{ijk}=\frac{a}{\tau}i\boldsymbol{e}_x+\frac{a}{\tau}j\boldsymbol{e}_y+\frac{a}{\tau}k\boldsymbol{e}_z \tag{5.26}$$

The choice of a and τ so far appears arbitrary. They are, however, chosen in a particular relation to the mass of the particles and temperature of the system. This allows for a calculation of certain integrals based on the Gauss-Hermite quadrature.

In the Gauss-Hermite quadrature, integrals of the form

$$I=\int_{\infty}^{\infty}P\left(x\right)e^{-x^2} \tag{5.27}$$

are considered. If P a polynomial of degree of not more than $\leq 2n-1$, the integral I can be expressed exactly by a sum of n terms

$$I=\sum_{i=1}^{n}w_iP\left(x_i\right), \tag{5.28}$$

if both the weight w_i and the abscissae x_i are be chosen appropriately. If a non-polynomial function is used the sum in Eq. 5.28 is only an approximation. The correct choice of the x_i in one dimension are the n roots of the Hermite polynomial of degree n. For the abscissae and weights, tables are available [262]. In multiple dimensions corresponding integration formulas can be constructed, where a corresponding subset of the space of polynomials of several arguments can be integrated exactly. The choice of the considered polynomial degree results in the number of integration points that have to be

used in Eq. 5.28.

Calculating hydrodynamic fields involves integrals in velocity space. They can be identified as the moments of the distribution function. The density e.g. is given by

$$\rho(\boldsymbol{r},t) = \int \mathrm{d}v\, f(\boldsymbol{r},\boldsymbol{v},t)\,, \tag{5.29}$$

and the momentum density $\boldsymbol{\rho u}$ is given by

$$\rho u(\boldsymbol{r},t) = \int \mathrm{d}v\, \boldsymbol{v} f(\boldsymbol{r},\boldsymbol{v},t)\,. \tag{5.30}$$

By using the Gauss-Hermite quadrature these integrals can be calculated approximately, if the distribution does not deviate too much from a Gaussian. This is only the case for applied forces. An arbitrary distribution function that deviates only weakly from a Gaussian can be described in terms of its moments. This can be expressed in terms of a Taylor expansion in three dimensions,

$$f(\boldsymbol{r},\boldsymbol{v},t) \approx (a(\boldsymbol{r},t) + \boldsymbol{b}(\boldsymbol{r},t)\cdot\boldsymbol{v} + c(\boldsymbol{r},t) : \boldsymbol{vv} + \cdots)\, e^{-\boldsymbol{m}v^2/2k_BT}. \tag{5.31}$$

where : denotes the double tensor contraction. The constants a, b, c are determined from matching the moments of f. When considering isothermal flow, it is sufficient to match moments up to second order [255]. For $n = 2$ the roots of the Hermite polynomials are 0 and $\pm\sqrt{3/2}$. The velocity is rescaled to $\boldsymbol{x} = \sqrt{mv/2k_BT}\boldsymbol{v}$, so that e.g. the density integral is of the form

$$\rho = \sqrt{\frac{m}{2k_BT}} \int \mathrm{d}x g(\boldsymbol{r},\boldsymbol{x},t)\, e^{-x^2} \tag{5.32}$$

in which the Gauss-Hermite quadrature can be applied. Here g is the polynomial approximation of Eq. 5.31. The Maxwell-Boltzmann distribution is brought to the same form by expanding it around small velocities $\boldsymbol{u}$,

$$f_{\mathrm{eq}}(\boldsymbol{v}) = \frac{\rho}{(2\pi k_BT)^{3/2}} \left(1 + \frac{m\boldsymbol{v}\cdot\boldsymbol{u}}{3k_BT} + \frac{m^2(\boldsymbol{v}\cdot\boldsymbol{u})^2}{18k_B^2T^2} - \frac{m^2\boldsymbol{v}^2}{9k_B^2T^2}\right) e^{-m\boldsymbol{v}^2/2k_BT}. \tag{5.33}$$

With this method it is therefore possible to express the local particle distribution function in terms of a few discrete quantities, namely the moments of the distribution.

The connection to the discretization of the Boltzmann equation is made such

that exactly the points at which f is evaluated, the abscissae, correspond to the discretization points of time and space. Every quadrature point corresponds exactly to the discretization points in Eq. 5.26. The naming scheme introduced by Qian [263] characterizes every LBM by the spacial dimension and the number of abscissae in the Gauss-Hermite quadrature. In two spatial dimensions the most prominent example is the D2Q9 scheme, where in the nine integration points $(0,0)$, $(\pm a/\tau, 0)$, $(0, \pm a/\tau)$, and $(\pm a/\tau, \pm a/\tau)$ are considered. Throughout this thesis, the D3Q19 variant will be applied. Its 19 velocities are depicted in Fig. 5.6. They correspond to the vectors

$$\begin{aligned} \boldsymbol{v}_0 &= (0,0,0), \\ \boldsymbol{v}_1, \cdots, \boldsymbol{v}_6 &= (\pm a/\tau, 0, 0), (0, \pm a/\tau, 0), (0, 0, \pm a/\tau), \text{and} \\ \boldsymbol{v}_7, \cdots, \boldsymbol{v}_{18} &= (\pm a/\tau, \pm a/\tau, 0), (0, \pm a/\tau, \pm a/\tau), (\pm a/\tau, 0, \pm a/\tau), \end{aligned} \tag{5.34}$$

and a/τ is identified as $\sqrt{m/3k_BT}$. This condition links the choice of the spatial and temporal discretization to the mass of the particles and the temperature of the system. The associated weights in the Gauss-Hermite quadrature are

$$w_0 = \frac{1}{3}, \qquad w_1, \cdots, w_6 = \frac{1}{18}, \qquad \text{and } w_7, \cdots, w_{18} = \frac{1}{36}. \tag{5.35}$$

It is important to note that the information about the full distribution function f has been reduced to its value at fixed positions $\boldsymbol{v}_i$ in velocity space, the populations. All formulas look as if the velocity space had been discretized in the first place, but they are merely the abscissae of the Gauss-Hermite quadrature.

On the level of ρ and $\boldsymbol{u}$ as defined in eqs. 5.29 and 5.30 hydrodynamic behaviour is recovered asymptotically. The dynamic viscosity, the only relevant transport coefficient for low Reynolds number flow, is obtained from the value of the relaxation time τ_r by a multiple time scale expansion. This approach is known as the Chapman-Enskog analysis. It is found that for the LBE in the BGK form the viscosity is related to the relaxation time by

$$\eta = \frac{k_BT\tau}{a^3}\frac{1+\tau/\tau_r}{1-\tau/\tau_r}. \tag{5.36}$$

In typical soft matter applications diffusion is very important. If diffusive mechanisms are considered, the fluctuations of a medium can, of course, not be neglected. Fluctuations can be incorporated into the LBM by adding a stochastic expression to the collision step. It was shown by Dünweg, Schiller

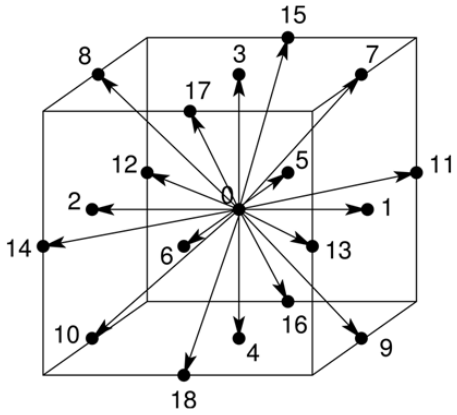


Figure 5.6: The 19 velocities of the D3Q19 Lattice-Boltzmann equation. One population is at rest, six move to nearest neighbours and twelve to next-nearest neighbours. From [264].

and Ladd that this can be expressed best in terms of the framework of multiple relaxation time (MRT) Lattice Boltzmann Method [265–268].

In the LBGK equation all populations relax towards the local equilibrium with a single relaxation time constant. In general, a linear collision operator can have multiple characteristic relaxation times. The expression of Eq. 5.25 can be generalized to

$$f_i^\star = f_{i,\mathrm{eq}} + \sum_j \mathcal{C}_{ij} \left(f_j - f_{j,\mathrm{eq}} \right), \tag{5.37}$$

which reduces to Eq. 5.25 in the case $\mathcal{C}_{ij} = \dfrac{1}{\tau_r}\delta_{ij}$. The idea of the multiple relaxation time formulation, lies in generalizing this equation such that the collision operator $\mathcal{C}_{ij}$ conserves mass and momentum, but otherwise is formulated with maximum freedom. This can be done by constructing new orthonormal basis set of the space spanned by the local populations, the mode space.

For the D3Q19 model it is possible to construct the following classes of modes:

- the density
- the three components of the momentum density

- a bulk mode describing local compression or expansion
- five modes related to shear stress
- nine modes which are not directly related to physical quantities, so called *ghost* or *kinetic* modes.

The density and momentum density are conserved, and hence the corresponding eigenvalue of the collision matrix is unity. The eigenvalues of the bulk and stress modes control the bulk and shear viscosity of the fluid. The relaxation of ghost modes does not have an immediate physical interpretation. For D3Q19 two independent relaxation times for the ghost modes can be chosen. A proper choice of the ghost mode relaxation time can be used to improve the stability of the LBM with certain boundary conditions [269].

The collision process leads to a relaxation of these modes towards their equilibrium value, if they are not conserved. In the same spirit as in the Langevin equation, under conditions with nonzero temperature, random fluctuations must be added to them. It was shown that these fluctuations need to be added to all modes to obtain a thermodynamically consistent system in equilibrium [155, 268]. This is most conveniently done in the multiple-relaxation-time (MRT) formulation of the LBM. The implementation in ESPResSo follows this scheme [266].

The most widely used method to include boundaries to an LB fluid was already applied for LGCAs, the *bounce back method*. Originally, discrete particles propagated to wall nodes were reflected to where they came with inverted velocities [270]. In terms of populations this means, all populations streamed into a boundary node are reflected back to the node where they came from with inverted velocities. This does not cause overwriting of relevant data, because exactly this population would have been streamed to that node from a boundary node. With this method in the LBM a zero-velocity boundary condition half-way between to nodes is produced [271, 272]. This statement only applies if the viscosity is not too high. In general the position of the boundary will depend on the viscosity [273]. Interestingly we have observed that this leads to cases where refining the lattice and keeping the viscosity constant worsens the agreement with the expected result as an apparent slip appears. A thorough discussion of possible boundary implementations is beyond the scope of our project, as we consider the LBM mostly as a tool to mediate hydrodynamic effects.

5.2.2 Coupling small particles to the LB fluid

In many applications of the LBM, it is necessary to couple the actual soft matter to the solvent the that is modelled by the LBM. The first model for particle coupling was suggested by Ladd et al. [274]. In this coupling scheme, particles are represented as bounce-back boundaries to the LB fluid. By summing up the transferred linear and angular momentum during a time step, the force and torques on the particles are calculated and their equations of motion are integrated. If necessary, the shape of the boundary is updated. In Ladd's work, a stochastic collision operator was used for the first time. Due to the fluctuating stress, Brownian motion of the particles can be observed, as desired for soft matter systems.

A different coupling method was introduced by Ahlrichs and Dünweg [254]. The general idea is that small particles are coupled by their expected frictional momentum exchange with the fluid. In this model it is assumed that a particle moving with velocity $\boldsymbol{v}$ at position $\boldsymbol{r}$ experiences a frictional force

$$\boldsymbol{F}_F = -\gamma \left(\boldsymbol{v} - u\left(\boldsymbol{r}\right) \right). \tag{5.38}$$

The coefficient γ is a friction coefficient, which we will discuss below. To obtain the fluid velocity at a continuous particle position, the fluid velocity is interpolated from the surrounding fluid nodes. The same force with opposite sign is applied to the fluid to conserve momentum in this coupling process. In the present implementation in ESPResSo, a first and a second order interpolation scheme can be employed, and the interpolation weights are also used to spread the force on the surrounding lattice nodes. In the following section we will discuss several aspects of this coupling scheme. Similar analyses were performed by Nash et al. and Ollila et al. [275, 276].

We note that the considered coupling mechanism is not overdamped, but resolves the inertial time scale. It is helpful to quantify this scale for an illustrative case. Assuming a colloid of $R = 1$ μm radius with the density of water immersed in water, the characteristic time scale of the decay of the velocity is

$$m/\gamma = \rho \frac{4\pi}{3} R^3 / 6\pi\eta R \approx 200\,\mathrm{ns}. \tag{5.39}$$

In comparison, the time to diffuse over its own radius is

$$r^2/Dk_BTR^2/6\pi\eta R \approx 5\,\mathrm{s}. \tag{5.40}$$

This time scale is a factor of approximately 10^7 larger than the inertial time

scale. Accordingly, the Schmidt number is 10^7. Matching both numbers independently is thus very difficult, as the number of time steps must be at least equal to the ratio of the two. Using the Lattice-Boltzmann algorithm for colloidal particles can only be practical following the idea of *telescoping of time scales* [277]. If two time scales in a physical system are sufficiently separated, a simulation method can be applied where the time scales are closer together but still sufficiently separated. Practically a much smaller Schmidt number may be as good as the correct one, which allows to significantly reduce the number of time steps. The clear advantage of using an algorithm where the inertial time scale is resolved, is that hydrodynamic interactions can be achieved by purely local operations.

Interestingly the friction coefficient γ in Eq. 5.38 is not identical to the inverse mobility of a particle. When a force $\boldsymbol{F}_{\text{ext}}$ is applied to a particle also the underlying fluid starts to move, as it is accelerated by the friction force. A stationary state is reached, when the applied forces and the friction force cancel, and then the applied force is completely transferred to the fluid. The particle velocity will then be given by

$$\boldsymbol{v} = \boldsymbol{F}_{\text{ext}}/\gamma + \boldsymbol{u}\left(\boldsymbol{r}\right). \tag{5.41}$$

Note, that due to the finite resolution of the grid, the fluid velocity at the particle position remains finite, in contrast to the divergent Stokeslet in the continuum case. Hence, if the resolution of the LB fluid is raised, the velocity at the position of the particle can be expected to be proportional to the resolution, as the flow field, decaying like $1/r$, is resolved finer. Furthermore in the Stokes equation the fluid velocity is proportional to the inverse of the viscosity. These dimensional arguments lead to a particle mobility of the form

$$\boldsymbol{\mu} = \frac{1}{\gamma} + \frac{g}{\eta a}, \tag{5.42}$$

which was suggested by Ahlrichs and Dünweg [278]. g is a numerical constant, which is supposedly independent of all other parameters. We performed the same check of this formula as Ahlrichs and Dünweg, and could confirm this behaviour for sufficiently high viscosities. As the LB method only fulfills the Stokes equation in the limit of large lengths, the field on short distances may deviate from the $1/\eta a$ scaling discussed above. Therefore it is legitimate that the numerical value of g also depends weakly on the viscosity. For the D3Q19 method with the first order coupling, g is around 0.04.

We define the hydrodynamic radius r_H of a particle coupled by this method as the radius of a sphere with the same mobility in the same medium. This

leads to the following expression

$$\frac{1}{r_H} = \frac{6\pi\eta}{\gamma} + \frac{6\pi g}{a}. \tag{5.43}$$

Thus for all choices of η and γ, the hydrodynamic radius can be at maximum

$$r_H \leq \frac{a}{6\pi g} \approx 0.75a. \tag{5.44}$$

This point is very important from the practical perspective. Making the lattice of the LBM finer does not leave the physics unchanged. On the contrary, it weakens the hydrodynamic interactions, as the hydrodynamic radius is decreased.

In addition, the mobility of a particle in periodic boundary conditions is system size dependent. This issue was first discussed by Hasimoto [279]. His actual physical question was the permeability of a periodic array of spheres. In this context he analyzed the response of a fluid subject to a delta-shaped force replicated periodically in space on a cubic lattice. He applied an technique inspired by the Ewald sum, where automatically a homogeneous counter-force density is assumed, because otherwise the result would be divergent. His result leads to the following form of the mobility of a spherical particle of radius a in a cubic box of length L:

$$\mu = \frac{1}{6\pi\eta a}\left(1 - 2.81\frac{r_h}{L}\right). \tag{5.45}$$

This expression is asymptotically exact for large a/L.

In order to check the agreement between Hasimoto's prediction and the LBM, we performed computer simulations, in which we applied a force to an isolated particle in periodic boundary conditions. A homogeneous counterforce was applied to the fluid, in agreement with Hasimoto's calculation. These calculations were performed with boxes of different sizes and checked the validity of the formula. The results are depicted in Fig. 5.7. Here we report only one data set, but the agreement between the Hasimoto prediction and all simulation results was very good. Only relatively large boxes (corresponding to $> 20a$) were investigated, and therefore higher order corrections to the Hasimoto formula are negligible.

The finite size correction of the mobility can be interpreted as a result of the hydrodynamic interaction between a particle and its periodic images. The good agreement with Hasimoto's formula underlines that HIs are modelled

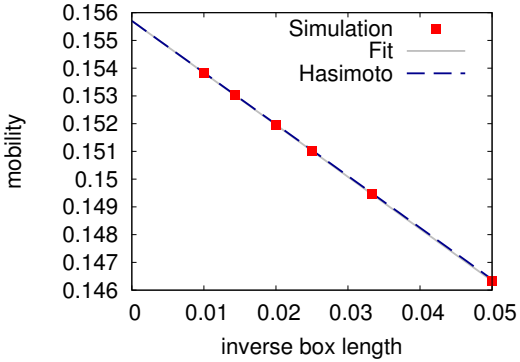
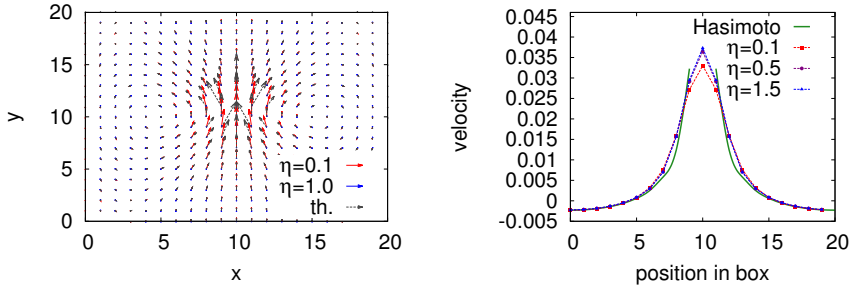


Figure 5.7: Finite size dependence on the mobility of a single particle coupled with the point coupling method. The simulation results (points) are coincide perfectly with the theory by Hasimoto [279]. This indicates that the hydrodynamic interaction with periodic images in the simulation method supports the notion of a hydrodynamic radius of particle coupled with the point coupling method.

asymptotically correctly. Thus, unique hydrodynamic radius is a justified concept for this coupling method. The same effective radius which describes the particle mobility also determines the strength of the hydrodynamic interactions.

Additionally we wanted to measure the hydrodynamic interactions of a pair of particles. However, this quantity is also influenced by finite size effects, and it was difficult to correct for this effect. Alternatively, we used the flow field created by the particles and the counterforce in the simulations above and compared it to the solution of Stokes equation subject to a point force in periodic boundary condition. This can be expressed as a infinite sum of Stokeslets. For the summation an Ewald summation technique can be applied. This calculation was first performed by Beenakker [280], and we implemented the formulas from [281].

In Fig. 5.8 we present the results. We find a good agreement of the analytic solution and the LB results in the percent range at distances of around 3-5 lattice sites between particles. Therefore we can also expect hydrodynamic interactions to be in agreement with the Oseen-Tensor level hydrodynamic interactions.



(a) (b)

Figure 5.8: Flow field in a box subject to a point force at the center. (a) Scaled flow field in the plane where the force is applied. LB results are indicated as solid arrows and the analytic solution as dashed arrow. (b) The vertical component of the scaled velocity along the horizontal centerline of graph (a). The simulations are displayed as dashed lines with symbols at the lattice nodes, and the analytic solution as solid line. The velocity units are arbitrary, as only the linear response is considered.

As a further test we addressed the question if the coupling method fulfills the fluctuation dissipation theorem, or, more precisely, the Green-Kubo formula. Schiller and Dünweg have found a proof [255], but we will show below, that not in all cases of practical relevance the conditions of their proof is fulfilled.

To define a time dependent mobility $\mu(\tau)$, we assume a particle at rest that is subject to step force that is switched on at $\tau = 0$. The particle will assume a velocity $v(\tau)$. We define $\mu(\tau)$ as the velocity after time τ, divided by the magnitude of the applied force F. This quantity can be extracted easily from the same simulation we have used to calculate the long time mobility. It is straightforward to show that this quantity is the inverse Fourier transformed frequency-dependent mobility. By means of the Green-Kubo formula it is related to the integral of the velocity autocorrelation function by

$$\mu(\tau) = \frac{1}{k_B T} \int_0^\tau \mathrm{d}\tau' \left\langle v(t)\, v(t+\tau') \right\rangle_t . \tag{5.46}$$

This is a generalization of Eq. 3.35; the "normal" mobility is retained for $\tau \to \infty$. We denote the time-derivative of the mobility as the response function g. It is accessible directly by simulations by using a delta-function-shaped

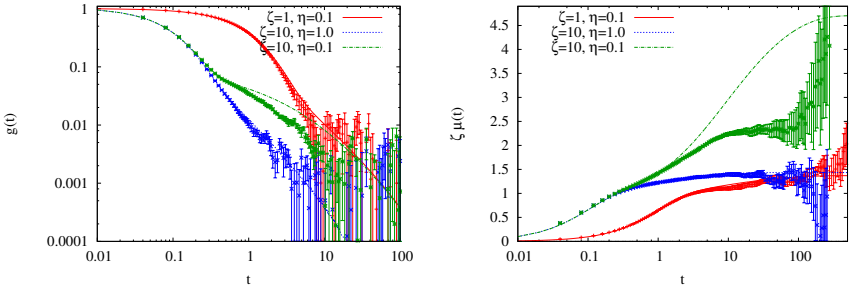


Figure 5.9: Test of the fluctuation dissipation theorem. (a) Velocity autocorrelation function (symbols) vs. response to a short force pulse (lines). (b) Integral of the the velocity autocorrelation function vs. response to a step force. Friction constant and viscosity are give in simulation units.

force in time, a kick, and measuring the velocity of as a function in time. It is the fundamental solution of the corresponding equation of motion, and can be therefore also called the Green's function of the particle motion. The velocity response to any force $F(t)$, can be obtained from its convolution with g. By virtue of the FDT, it must be identical to the velocity autocorrelation function up to a constant $1/k_BT$.

We also calculated the velocity autocorrelation function for a single particle in periodic boundary conditions for different values of the viscosity η and friction coefficient γ, with a lattice constant of 1. The results are presented in Fig. 5.9. It observed that the expected agreement is only found, if the viscosity and friction coefficient are sufficiently large. The long time limit of the mobility measured from fluctuations and from pulling the particle differs up to a factor of two under the considered conditions. We found that this behaviour is insensitive to the LB time step, and therefore is unlikely to be related to LB internals. It however appears to be an intrinsic property of the coupling mechanism.

So far we have not been able to systematically track the origin of the deviations. Lower viscosities increase the deviations, but it still remains unclear, what exactly causes the deviations. In the proof in Ref. [255], Schiller and Ladd argue only on the level of hydrodynamic fields, and hence assume that the fluid is in local equilibrium on the time scale of the coupling. By decreas-

ing the viscosity, the time scale on which the hydrodynamic fields develop increases. Therefore, it is plausible that the time scales for establishing a viscous flow and of the relaxation of particle momentum are not sufficiently well separated to apply the proof.

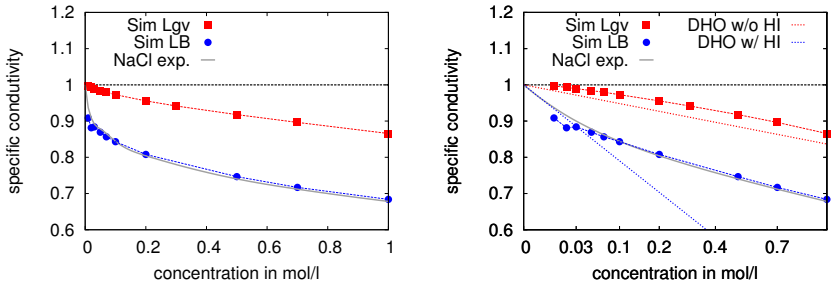
As a further test, we investigated the conductance of an electrolyte. Onsager [282], based on work of Debye and Hückel, derived a model how the conductance of an electrolyte depends on the concentration. His theory applies for sufficiently small concentrations, and contains two effects, *relaxation* and *retardation*. First, the radial distribution of ions is skewed by the applied electric field, the relaxation effect. This leads to a reduction of the field strength at the position of the ion, and hence a reduction of the conductance. Second, ions of opposite charge moving in opposite directions are subject to hydrodynamic interactions which also slows them down, the retardation effect. In Onsager's calculation, hydrodynamic interactions are taken into account on the Oseen level. The results of this calculation is known as the Debye-Hückel-Onsager limiting law and is contained e.g. in many physical chemistry books. For a 1:1 electrolyte. it can be brought to a particularly simple form, where it reads as

$$\sigma = 2c\mu_0 \left(1 - 0.07\frac{l_B}{l_D} - \frac{r_h}{l_D}\right). \tag{5.47}$$

It is apparent, that the retardation term vanishes for vanishing hydrodynamic radii.

For small c, the Debye-Hückel-Onsager law is exact. Interestingly, this means that the conductivity can not be expanded in a Taylor series in c, as the slope at $c = 0$ is infinite, but must be expanded in terms of $\sqrt{c}$. Different empirical extensions have been proposed to extend the range of applicability of such a formula to higher concentrations.

We performed simulations of a univalent electrolyte at room temperature represented by spherical particles interacting with the WCA potential with $\sigma = 0.45$ nm and $\epsilon = k_B T$. The hydrodynamic radius chosen in agreement with that of K and Cl ions to 0.15 nm. This was done by using a lattice constant of 0.34 nm. This length was chosen as the unit of length in the simulations. The mass of both cations and anions was chosen as unity. Using the mean experimental mass this fixes the unit of time of our simulations to 1.5 ps. We chose a viscosity of 0.8 in simulation units, corresponding to 0.11 mPa s. This is a factor of 10 smaller than real water, but the diffusion constant accordingly is with 1.3 10^{-8}m$^/$s also factor of 10 larger. The Schmidt number is even smaller by a factor of 100, as we have reduced the viscosity and enhanced the diffusion constant. The box shape is cubic with an edge length of



(a) (b)

Figure 5.10: Specific conductance of an electrolyte simulated with Langevin dynamics (Lgv) and the LB method. (a) Plotted with a linear concentration axis, (b) axis scaled by the square root of the concentration. Simulation data with Langevin dynamics are shown as squares, and data obtained with the LB method are shown as circles. Experimental data, interpolated with splines, is shown as grey lines. The result of the Debye-Hückel-Onsager (DHO) formula is added in the right plot as dashed lines, both with (blue) and without the hydrodynamic contribution (red).

20 simulation units, or 6.68 nm. The choice the parameters convenient from a practical perspective, as the diffusion constant is relatively high, and therefore simulated systems exhibit a high signal-to-noise ratio. On the other hand, the FDT is fulfilled sufficiently precisely to derived unique transport coefficients. It is, however, not unique, and from our perspective further investigations of the role of the parameters are necessary.

In Fig. 5.10 the result is depicted and compared to experimental results to NaCl from [283]. Furthermore we compare the results to simulations without hydrodynamic interactions using Langevin dynamics. The friction coefficient in the Langevin dynamics simulations are chosen such that the mobility of an isolated particle is matched. The conductance of the experimental data is normalized by its value at infinite dilution, and the simulation data is normalized by the mobility of isolated particles.

The agreement with the experimental data is very good. The simulation data captures the experimental curve almost to the statistical accuracy of the simulation data. The relaxation effect is weaker than the retardation effect, but constitutes an important contribution. For small electrolyte concentrations

we find considerable deviations, that are caused by the finite size of our simulation box. For the smallest considered salt concentration only four ion pairs were used. Such a good agreement between simulation and experiment is very surprising from our perspective, but very encouraging. Our simulation data not only captures low concentration as the Debye-Hückel-Onsager law, but can describe the experimental data over the entire concentration range.

5.2.3 Nanopore conduction with the LB model

The electrolyte model described above was applied to calculate the conductance of a nanopore in the presence and in the absence of a DNA molecule. Inspired by the electrokinetic continuum model, the DNA molecule was represented as a cylinder and the nanopore as a cylindrical channel. We considered a system where both the pore and the DNA periodically repeated across the boundary. This way we can simulate a finite piece of pore and DNA that imitates an infinite system. The DNA cylinder was charged by placing particles with one elementary charge at a mutual distance of 0.16 nm along the cylinder axis. The DNA-ion interaction was modelled by a WCA interaction with $\epsilon = k_BT$ and σ =0.425 nm, and the distance was shifted by 0.575 nm to obtain a steep potential and a distance of closest approach of around 1 nm. The Bjerrum length of 0.71 nm reflects a temperature of 298 K and an aqueous solvent. A box length of 25 nm was chosen and again all lengths were expressed in units of the LB lattice constant of 0.34 nm.

The total charge of the cylinder consisted of 144 elementary charges. The system was neutralized by adding 144 positive counterions, and a varying number of extra ions was added to the system. In this geometry it is impossible to use an ion reservoir. Our strategy was therefore adding a different numbers of salt ions, perform the simulation and determine the concentration of a hypothetically attached reservoir a posteriori. This leads to a somewhat different salt concentration than one obtains by dividing the number of ions by the pore volume. We determined the salt concentrations by calculating the local concentration as a function of the distance to the pore axis, and fitting a constant to the mean of the concentration of positive and negative ions in a region between 3 and 4.5 nm away from the pore axis. If the Poisson-Boltzmann equation held exactly, this method would yield exact results given that the electrostatic potentials are smaller than k_BT/e. This is the case in the regions used for averaging. The same method was also applied for the atomistic simulations described below, where we performed further simulations to justify it.

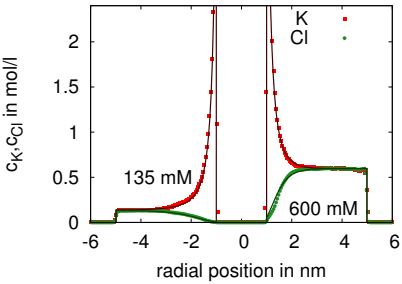


Figure 5.11: Radial density profile of ions. K^+ ions are shown as squares, and Cl^- ions as circles as obtained from the simulations. Corresponding results of the electrostatic continuum model are shown as lines. The left half of the plot is corresponds to a reservoir concentration of 135 mM and the right half to 600 mM.

In Fig. 5.11 we display the ion densities we obtained from our simulations for two different data sets, corresponding to 101 and 505 ion pairs added to the system. The corresponding reservoir concentrations were determined to be 135 mM and 600 mM. The agreement between the electrokinetic infinite cylinder model and the simulations is good. Especially for the higher salt concentration case the decay of the counterion cloud towards the bulk value is somewhat faster. This indicates that correlation effects are present that are not contained in the PB theory. The divergence between the two models is weak. We therefore can say, consistent with previous findings [284], that the ion density and hence also the electrostatic potential around a charged rod with the line charge density of DNA is described well by the Poisson-Boltzmann equation, and that correlation effects are not a dominant source of error.

The position-dependent water and ion velocities are depicted in Fig. 5.12a. The water velocities obtained in the simulations are significantly in the simulations than in the electrokinetic model. This encloses both datasets shown, although the deviations for higher salt concentrations are larger. Qualitatively the fast decay of the counterion cloud in the simulations will lead to a smaller water flow. This can be understood by considering the extreme case of ions sitting right at the boundary. The force acting on them is immediately transferred into the DNA and they do not constitute electroomotic flow. The effect can be considered analogous to a lever. The larger the distance of the

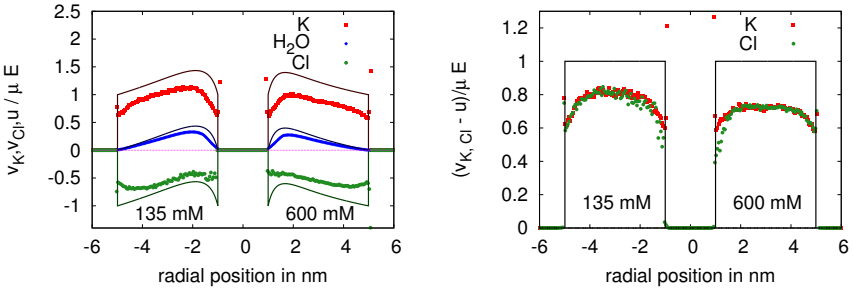


Figure 5.12: (a) Ion velocity as a function of the distance to pore axis. K^+ and Cl^- ions as squares and circles resp. The velocity of the LB fluid replacing the water is displayed as diamonds. The corresponding continuum model prediction are shown as lines. (b) Ion mobility as a function of the distance from the pore axis pore axis.

counterions from the DNA, the larger is the EOF they create. This qualitative finding is also the reason why the EOF is smaller when the Debye length gets smaller. A second aspect probably is related to the coupling scheme of the LB particles. Due to the multilinear assignment scheme forces in a certain range are directly transferred into the DNA without causing an acceleration of the water. This could be a second source of deviation, which we have not completely ruled out. Finally it is interesting to note that the position of the hydrodynamic boundary is found exactly at the desired distance from the pore axis, although the cylinder is represented on a lattice. The ion velocities we obtain are in good qualitative agreement between the values obtained from the continuum model, although significantly smaller. This is not surprising as the conductance decreases with increasing concentration in the coarse-grained simulation model.

In the electrokinetic continuum model the ion velocity is just the water velocity plus a constant offset. Inspired by this notion we define an effective position-dependent mobility as the difference between the local ion velocity and the water velocity, as

$$\mu^{\pm} = \pm \left(v_z^{\pm}(r) - u_z^{\pm}(r) \right). \tag{5.48}$$

It is shown in Fig. 5.12b for both ion species, again for the same two datasets as used before. We normalize their values by the mobility of an isolated particle.

At least for the higher concentration data set the mobility is almost constant in the center between the DNA and the pore wall. This constant however is significantly smaller than unity, which reflects the concentration dependence of the conductance. Towards the walls detail below. The mobility of counterions is somewhat larger than the mobility of the coions. For this behaviour we have so far not found a satisfactory explanation, but it is found throughout our data.

On the macroscopic scale the mobility of particles in a solvent is reduced near a wall. In analogy with the image charge method this effect can be explained in term of image particles moving in the opposite direction. A particle interacts hydrodynamically with this image particle, leading to a decrease in mobility. For a single planar wall this can be expressed in terms of the Blake tensor [285], that replaces the Oseen tensor near walls, just as Coulombs law can be corrected by introducing an image charge on the other side of the wall. Asymptotically the correction behaves as equation

$$\mu\left(d\right)_{\parallel} = \mu_0 \left(1 - \frac{8}{9}\frac{a}{d} + \cdots\right), \tag{5.49}$$

where d is the distance from the wall. Perkins and Jones derived a quasi-exact expression by solving the Stokes equation in this geometry [286]. We applied their result, as stated in [287], to construct an expression for the position-and concentration dependent mobility. We assumed that the mobility is corrected by three factors, a factor depending on the local concentration , and the influence of two walls. The concentration factor was determined from an interpolation of the experimental data, and we used the local concentration of ions of the opposite charge, as the opposite species is responsible for the relaxation and retardation effects. The wall factors were constructed by inserting the distance to both walls into the Perkins-Jones formula. For simplicity we only state its asymptotic form

$$\mu\left(r\right) = \mu_0 \times f_{\text{conc}} \times \left(1 - \frac{8}{9}\frac{a}{r - r_{\text{DNA}}}\right) \times \left(1 - \frac{8}{9}\frac{a}{r - r_{\text{wall}}}\right). \tag{5.50}$$

It is depicted in Fig.5.12b as solid lines. Even though the agreement to the simulation data is not perfect we reproduce the trends in the simulation data well. It appears plausible that these two factors must be taken into account, even though they conceptually not included in the electrokinetic equations.

Finally we investigate the current predictions of the coarse grained simulation model. The current is measured from the displacement of all ions during

the production part of the simulation, divided by the applied field and the simulation time. We write the total current as product the product

$$I = ni\mu_0 E \tag{5.51}$$

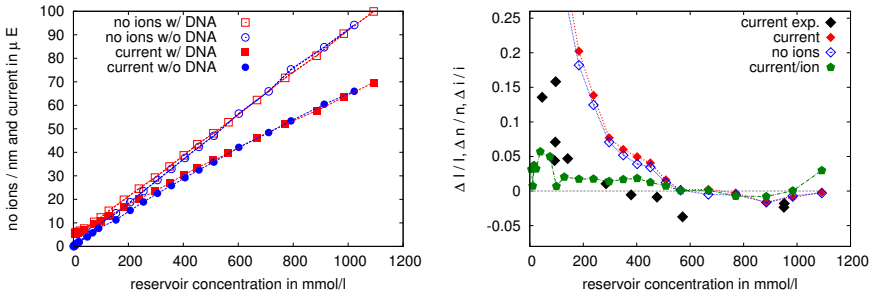
of the number of ions per unit length $n = \left(N^+ + N^-\right)/L$, multiplied by the product of the velocity of an isolated particle $\mu_0 E$ and a dimensionless current per particle i, which contains all deviations from the behaviour of isolated particles. This way we can distinguish between two influences, how a DNA affects the number of ions present in the pore, and how easily these ions can move. If all ions moved at the velocity of an isolated particle, i would equal unity. This way it is possible to express the conductivity of the system in "equivalent ions per nanometer" N/Li, thus in the same units as the number of ions in the system.

In Fig. 5.13a we present the results obtained both for the number of ions the pore contains per unit length and the reduced measured current. Both are shown as a function of the effective reservoir concentration which was calculated a posteriori for every dataset. We performed all simulations also for a DNA-free system, where the same a posteriori analysis procedure was applied. We find that for low concentrations the number of ions in the pore and the current is significantly larger when a DNA is present in the pore, reflecting the fact that the conductivity is dominated by the counterion cloud surrounding the DNA. For all datasets the current is considerable smaller than the number of ions per unit length, thus i is smaller than unity and decreases with increasing concentration. This reflects the decrease of electrolyte conductance with increasing concentration.

For concentrations above around 600 mM, both the number of ions and the current exhibit a crossover. Above this concentration, the conductivity is higher without a DNA. Although the relative statistical error is small, this crossover is not much larger than the error. To investigate this further, we fitted a third order polynomial to the data without DNA, and investigated the relative difference in current I, number of ions n, and current per ion i. The validity of the third order polynomial fit was verified by a χ^2-test. We define the relative modulation of the current as

$$\Delta I/I = \left(I_{\text{with DNA}} - I_{\text{without DNA}}\right)/I_{\text{without DNA}}, \tag{5.52}$$

and the relative modulations of n and i analogously. The results are shown in Fig. 5.13b, and compared to the experimental data of Smeets et al. . Here, the crossover around 700 mM in the simulation data is clearly visible. Interestingly



(a) (b)

Figure 5.13: (a) Comparison of the number of ions n in the pore (open symbols) and the reduced ion current I (solid symbols). Data with a DNA present is depicted as squares, and data without DNA as circles. (b) Relative modulation of the current $\Delta I/I$, the number of ions in the pore $\Delta n/n$ and the current per ion $\Delta i/i$. Experimental data for of Smeets et al. [5] is shown as black diamonds, the modulation of the current as smaller red solid diamonds, the modulation of the number of ions as open diamonds and the modulation of the current per ion as solid pentagons. The size of the statistical errors is comparable to the symbol size.

it occurs both in $\Delta n/n$ and $\Delta i/i$ at a similar concentration, but in both cases the reduction is with $\lesssim 1\%$ very weak.

We interpret the data as follows. The relative modulation of the number of ions per unit length exhibits the same crossover behaviour as in electrokinetic infinite cylinder model. The crossover concentration appears to be somewhat smaller than the 800 mM determined in the continuum model. This can, however, be explained by the occurrence of correlation effects which are neglected in the continuum model. In all cases, the current per ion, which is equivalent to the electroosmotic contribution to the current in the continuum model, is smaller not only in absolute numbers due to the concentration dependent conductivity, but also smaller in relative terms. Two sources of this effect can be identified. First, the electoosmotic flow occuring in the simulation is smaller than in the continuum model. Second, the mobility of ions near the DNA is smaller due to the hydrodynamic image particle interaction.

To this point the data obtained from the simulations is only partly satisfactory. We have identified four potentially important effects, (a) the excluded volume

of the DNA which ions can not enter, (b) the counterion cloud of the DNA which brings extra ions into the pore, (c) the electroosmotic flow and (d) the mobility reduction near boundaries. The last aspect is a result of momentum conservation and particle number conservation, it is very likely to occur also for real DNA. On the other hand, the models we applied here are clearly too coarse for a quantitative analysis. The magnitude of the effect only a few percent and on that level the exact shape, position of charges, surface structure and other microscopic details are likely to be important. To achieve a better understanding of these effects, we decided to perform also simulation with atomistic resolution, where these aspects are represented much more realistically. They will be described in the following section.

5.3 Atomistic model of nanopore conductance

In this section we present our atomistic resolution model of a DNA in a nanopore. From the previous models we have learned what are the key effects that determine how much the nanopore current is modulated by the presence of a DNA molecule: (a) the excluded volume of the DNA, (b) the DNA charge and the structure of the counterion cloud (c) electroosmotic flow, and (d) friction effects near the DNA. The magnitude of each effect however remains obscure. All of them are significantly affected by the microstructure of the DNA. Therefore we studied these quantities with a model with atomic resolution to achieve a more realistic picture.

The guiding principle of atomistic simulations is that every atom is represented by a particle. At present, atomistic computer simulations of biomolecules, such as proteins and nucleic acids have reached a point where they can be applied without many years of experience, but after a relatively short training period. Different software packages, such als GROMACS (GROningen MAchine for Chemical Simulations) [288–292], AMBER (Assisted Model Building with Energy Refinement) [293, 294], CHARMM (Chemistry at HARvard Macromolecular Mechanics) [295, 296] and NAMD (Not (just) Another Molecular Dynamics program) [297, 298] are installed and ready to use on many supercomputers. Using them requires only relatively little knowledge about the internals, and a vast amount of computer simulation results is published in a broad range of scientific journals. From crystallographic X-ray measurements the structure of an impressive number of biomolecules is available. Downloading one of them from the Protein Database (PDB), and studying it in simulation can be learned in a few days. The possibility to zoom into details

which are experimentally inaccessible allow scientists to gain an entirely new perspective on microscopic mechanisms, that, for example, be e.g. biologically relevant.

In Fig. 5.14 we outline our simulation setup. Just as in our previous simulations, we investigate the central part of a cylindrical nanopore. We apply periodic boundary conditions to mimic an infinite system. The system length is chosen in agreement with the symmetry of a DNA molecule. We use a simulation box size of 6.76 nm, which contains exactly two full turns of the double helix. We choose a DNA homopolymer which consists only of CG base pairs. In Sec. 5.3.4, we compare the results also to a DNA containing only AT base pairs. The DNA is held in place by harmonic springs attached to the phosphorous atoms of the backbone, so that it can deform and fluctuate but not move away from the initial position. Water and ions are added to the system. An electric field is applied parallel to the pore axis and the ion motion is monitored and analyzed. For comparison, identical simulations without a DNA molecule were performed. In principle, this system is identical with the electrokinetic infinite cylinder model, and the LB/MD model. The only difference is that all degrees of freedom are taken into account explicitly in terms of particles. Similar simulation studies have been conducted by Luan et. al. [299], but with a focus on measuring the force acting on the DNA.

Despite the relative ease of performing simulations, the question remains how reliable the results are. Crucial is the choice of simulation parameters, most importantly the choice of interaction potentials between atoms. The dataset containing all interaction potentials is called a *force field*. Determining these parameters from experimental data is an extreme complex inverse problem. Due to the complexity of this problem, there are several well-established sets of parameters, where somewhat different philosophies, and also different underlying data guided the development. Each force field was developed in conjunction with a particular set of parameters for water molecules, the water model. Choosing a force field is a particularly difficult problem and requires expert knowledge. For our simulations it was unfortunately not possible to use one of the most popular combinations.

For the simuluations, we used GROMACS 4.5.5 [289, 291, 300]. This simulation software is very efficient and highly optimized. Especially the implementation of common water models, for which virtually all atomistic biomolecular simulations spend most of the computation time, is fast, due hand-written assembler kernels with a minimum number of floating point operations. Its relatively narrow focus on biomolecular systems compared e.g. to ESPResSo makes these optimizations very useful.

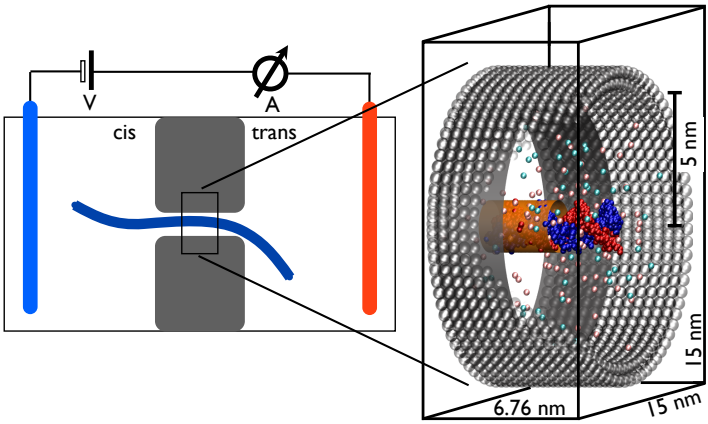


Figure 5.14: The simulation setup used for the atomistic simulations. We focus on the central, cylindrical part of a long nanopore. 20 base pairs, corresponding to two full turns of the helix are simulated in peridioc boundary conditions. The pore is formed by Lennard-Jones bead forming a cylinder of radius 5 nm. A box of $6.76 \times 15 \times 15\,\mathrm{nm}^3$ is used, and an electric field is applied along the pore axis.

The somewhat difficult choice of a combination of a force field, a water model, and interaction parameters of KCl will be discussed in the following section. Then we will present the details of the applied simulation protocol. In the following section, we present the simulation results we obtained for the conductivity of a nanopore in the presence of a DNA molecule. They are compared to the prediction of the electrokinetic infinite cylinder model. Finally, we investigate detailed features of the obtained data by means of visual analytics, where we also discuss sequence-specific effects.

5.3.1 Choice of the force field

An important ingredient for our simulations is the choice of the force field. Our simulation is not too far from standard tasks in biolomolecular simulations and therefore we choose one of the classical, (non-polarizable) force fields. On the market there are four families of force fields for biomolecules: AMBER, CHARMM, GROMOS and OPLS. The first three have been developed in connection with simulation packages with the same name. Only the first

three are regularly applied for DNA. OPLS is therefore excluded from the further discussion. Our focus lies on the dynamic properties of ions, and this introduces certain complications to the discussion. Our rationale will be presented in this paragraph accompanied by some background information on each force field.

The AMBER and CHARMM force fields are the most widely used force fields for DNA [301]. Both are parametrized with the TIP3P water model [302]. The first nucleic acid simulations with AMBER were performed with the AMBER94 parameter set of Cornell et al. [303]. The force field parameters were derived by a combination of quantum-chemical calculation for determining the partial charges and experimental values of thermodynamic properties. The AMBER99 parameter set [304] still is the reference for simulations of DNA, with the exception that unrestrained DNA deforms in very long simulation runs [305]. The same group therefore suggested a modification of angular and dihedral potentials, the ParmBSC0 parametrization, which results in a somewhat stiffer double-stranded DNA and appears to be the state of the art in the simulation of DNA. In the AMBER03 force field [306] torsional potentials and partial charges were reparametrized. Interestingly, in the parameter set recommended in the manual for the newest version 12 of the AMBER software, *ff12sb*, is based on the 99 version including the ParmBSC0 modifications of the force field again and not on the AMBER03 parameter set. The CHARMM force field was initially used for nucleic acids in 1995 [307]. It was updated to be compatible with the CHARM27 version of the force field and also new ion parameters were included [308]. The resulting DNA conformations, however, are somewhere between the A and B form [309]. In the GROMOS force field hydrogen atoms bound to non-polar atoms are included into a "super-atom". This slight simplification appears not to affect the overall quality of the force field. The GROMOS force field is somewhat less common in DNA simulations [301]. The most mature parameter set for nuculeic acids, 45A4 [310], unfortunately is currently not available in GROMACS. GROMOS is parametrized with the SPC water model [311].

In our simulations the conformations of the DNA molecule will only very weakly depend on the force field. On the one hand, we fix the phosphorous atoms of its backbone and therefore it cannot bend. On the other hand it has not ends where it can start unzipping. Only fine details as the size of the groves and the exact surface exposed to the solvent can somewhat depend on the applied force field. Therefore the accurate reproduction of a certain geometry is not the most worrying point in the choice of a force field. Our main focus lies on the ions. In general it can be stated that the quality of the ion parameters has not been not considered a very important point in

the force field development. For example, the only place where the standard GROMOS parameters for sodium and chloride are documented, is the software manual. Potassium is even not available at all. We are not aware of any independent study which assesses the quality of the predictions obtained with these ions. A second example is related to the default ion parameters in the AMBER force field. Only relatively recently it has been found that small stable salt clusters form already at moderate concentrations when the alkali ion parameters by Aqvist [312] are used. They are not encouraged to be used any more, [313,314]. Noy even states: "The use of AMBER-standard combination of LB-adapted [Lorentz-Bertelot-mixing rule adapted] Aqvist parameters for cations and Dang's parameters for Cl^- should be disregarded as a potential source of artefacts in simulations at medium or large ionic strengths.". The source of trouble originates from the mixing rule used to calculate the Lennard-Jones interaction parameters between unequal particles. AMBER uses the Lorentz-Bertholot mixing rule, i.e. the arithmetic mean of σ and ϵ of the individual atoms, while the Aquist parameter set requires a geometric mixing rule [315]. In the last couple of years a trend has developed to systematically improve the quality of ionic force fields, e.g. [301,316,317]. This process is not converged, but has created an even larger variety of ion parameter sets with so far still partly unknown properties.

In addition to the cluster formation, two main points are to be worried about when making a choice of ion parameters, (a) the DNA/ion interaction and (b) the hydrodynamic radius of the ions. For the first point, it is very difficult to make an appropriate statement. Experimental data about ion-DNA binding is in principle availabe, e.g. [318, 319], and was also used to improve ion parametrizations [301] but the evidence for or against a particular choice appeared vague to us. The second point, the hydrodynamic radius is much simpler to evaluate from our point of view. We define the hydrodynamic radius r_h as the radius of a spherical particle in the same solvent of viscosity η, for which the Stokes-Einstein formula predicts the same diffusion coefficient D. Then r_h is given by

$$r_h = \frac{k_B T}{6\pi\eta D}. \tag{5.53}$$

Its importance has been stressed in the previous sections. Only if the hydrodynamic radius is reproduced correctly, the contribution of convection to the overall current is correct. Noy et al. showed [320] that essentially every combination of popular ion parameters with every common water model in combination with the ParmBSC0 variant of the AMBER force field leads to very similar DNA conformations [320]. Therefore we chose to apply the AMBER force field, and select a suitable set of ion parameters and water model.

Ion	TIP3P		SPC/E		exp.	
	D	r_h	D	r_h	D	r_h
K^+	2.22	0.31	1.2	0.25	1.957	0.112
Na^+	3.30	0.21	2.05	0.154	1.334	0.167
Cl^+	3.08	0.22	1.74	0.148	2.032	0.108
Viscosity	0.321		0.729		1.00	

Table 5.1: Diffusion constants in $10^{-9}m^2/s$ as reported in [324] and hydrodynamic radii in nanometer for the Smith-Dang-Garrett ion parameters. In addition, we report the literature values for the viscosity of water in mPa s.

This way we can also evade the Aqvist ion parameters. One popular choice is using the parameters of Smith, Dang and Garrett [321–323] (abbreviated by Dang, as he is the only author of all articles), especially as the Dang chloride ion is the default of AMBER anyways. From the simulation results of Joung et al. [324], we calculated the hydrodynamic radii of sodium, potassium and chloride ions in SPC/E and TIP3P water. The results, including experimental values also from [324] are summarized in table 5.1. We note that the viscosity of both TIP3P water and SPC/E water are significantly smaller than the experimental values. The viscosity for TIP3P is around 1/3 and that of SPC/E is around 2/3 the viscosity of real water. Correspondingly, the diffusion constants of all ions are higher in the simulations than in reality. The hydrodynamic radius in all cases is larger than the experimental values, which therefore must lead to somewhat large convective current contributions. Hence, that the Dang-SPC/E combination is a better choice than the combination than choosing TIP3P water.

Originally we had performed the full set of simulations using the GROMOS parameter set. Also in these simulations the hydrodynamic radius of ions was significantly too large. We did not determine it exactly, but from investigating the ion velocity profiles we found that the magnitude of the EOF compared to the ion velocity was larger than we had expected. Indeed this even spoiled observing the conductivity crossover. We therefore have chosen to use the AMBER-SPC/E-Dang combination which is not canonical, but tested for long time stability by Noy et al. and where the hydrodynamic radii are acceptable. While experimentally also NaCl is used in experiments, we focus on KCl as in the experiments of Smeets et al. [5].

Our force field choice was justified, given the available data, but not unique. It would be very helpful to compare the simulation results presented here to the prediction of other force fields. However, this goes beyond the scope of

this thesis. In the conclusions section, we will discuss how this work could be continued in this respect.

5.3.2 Simulation setup and analysis procedure

We briefly summarize the procedure we perform in course of a simulation.

The DNA strands are generated using the Nucleic Acid Builder (NAB) from the AMBER toolbox [294]. This tool provides a powerful interface to automatically generate complex nucleic acid structures. Although we studied only very simple examples, we chose it because it provides the flexibility to solve also more complex tasks. Changing from a poly-CG DNA molecule to a poly-AT could e.g. be achieved with minimal effort. It generates a PDB (protein database) file which then is converted into a GROMACS structure using the GROMACS tool `pdb2gmx`. From solely atomic coordinates and residue names, this tool creates a *topology* containing interaction parameters, including bonds, angular and dihedral potentials. In this process also the desired force field must be chosen.

The resulting topology has to be modified to be compatible with periodic boundary conditions. This was done with a tool written in course of W. Müller's Diploma thesis. NAB automatically creates special terminal groups at the ends of the DNA strand. Our tool chops these terminal groups off and automatically creates bonds between the new ends of the chain. Then harmonic springs, *restraints* in GROMACS terminology, are created which keep the phosphorous atoms in place. Finally the pore atoms are created and a also fixed in space by harmonic restraints. A pore of length 6.76 nm and radius 5 nm is created by placing two layers of Lennard-Jones particles with $\sigma = 0.5\,\mathrm{nm}$ at mutual distances of 0.5 nm on a cylinder in a simulation box of 15 x 15 x 6.76 nm^3. Adding water and ions is a somewhat more complicated process described below.

Before the start of a simulation, several annealing steps are taken that ensure the system reaches a configuration with a small potential energy. First, we minimize the energy of the system in vacuum with a gradient method, then we add water molecules. The water molecules are created as periodic replications of an equilibrated box containing 216 water molecules. This box is provided by GROMACS and ensures that the systems is not far away from an equilibrium configuration when starting a simulation. Water molecules overlapping with the DNA and the pore pseudoatoms are removed. For this procedure a distance criterion based on the van-der-Waals radii of the atoms

is used. The applied radii of the pore atoms are adjusted to yield the bulk density of water in the region between the wall and the DNA. The energy of the resulting hydrated DNA molecule is minimized again to ensure that energetically unfavorable configurations are eliminated and the first MD steps can be be performed without stability problems. Then the production run is initiated. To make sure the system is in a stationary state, a certain portion, in most simulations the first nanosecond is discarded. During this period, the stationary ion distribution develops.

During the production run, the bonds involving hydrogen atoms are constraint in length with the P-LINCS algorithm [325], because this allows for a larger time step, here to two femtoseconds. A stochastic velocity rescaling thermostat [157] keeps the temperature of the system constant. In principle this thermostat is not momentum-conserving and therefore does not produce correct hydrodynamics. We use a relatively weak coupling constant of 5 ps where, on the length scale of the considered box size, hydrodynamics are reproduced with good accuracy. The restraints are sinks of the momentum anyway, and therefore the system is not extremely sensitive to the violation of momentum conservation as long as the coupling is sufficiently weak. We also performed simulations with different coupling times up to 50 ps, hence an even weaker coupling, and the results were indistinguishable.

Electrostatic interactions are calculated using the SPME method [326] with a grid constant of 0.125 nm. For short range interactions a 1.4 nm cut-off is used, in accordance with the specification of the AMBER force field. An electric field of 0.2 V/nm is applied along the axis of the system. For every salt concentration ten independent runs of five nanoseconds were performed. For analysis we hence obtained around 40 ns of simulation trajectories for each salt concentration. In most simulation runs, we used the GROMACS shared memory parallelization using one node of the BW-Grid cluster with eight processor cores. The performance of around five nanoseconds per day could have been increased by using a higher number of processors and distributed memory parallelization. As several hundred independent calculations had to be performed, this degree of parallelization was sufficient and the parallelization overhead was minimal.

Once per 200 time steps all particle positions are stored to disc. This creates a relatively large amount of data, around 3 gigabyte per five nanosecond run. We chose such a high output frequency as we want to resolve the velocity autocorrelation time of the ions to maximize the statistical accuracy. After each run, in an automatic procedure, the ion trajectories and a short piece of the water trajectories are extracted from the compressed GROMACS output

format and analyzed by a binning procedure. In this binning procedure for every ion the distance from the cylinder axis is calculated and the density in every bin is calculated by accumulating the weights of a linear interpolation for each bin. The flux density is calculated by multiplying the displacement in the each interval with the same weights and accumulating it per bin. Finally the result is divided by the volume of each bin, the number of snapshots and the timespan of each interval. The mean local velocity is calculated as the ratio of flux density and density. For bins with a very low density, this can lead to high statistical errors and therefore if the density was below a threshold 0.001 ions per cubic nanometer the velocity was set to zero. To obtain a three-dimensional density and velocity fields, we performed a three-dimensional binning procedure we describe in more detail below.

In our simulations we could not consider a reservoir explicitly, as it is not compatible with our geometry. Our strategy for obtaining the reservoir concentrations is adding a certain number of ion pairs to the simulation box and determining the effective concentration of a hypothetically attached reservoir afterwards. Our method works as follows: In the Poisson-Boltzmann (PB) theory, the densities $c_\pm$ of univalent ions with signs + and - is given by

$$c_\pm = c_0 \exp \mp \Phi. \tag{5.54}$$

Here Φ denotes the electrostatic potential measured in $k_B T$ per elementary charge and c_0 the reservoir concentration. Then the mean of the + and - concentration can be expanded in the potential, and we obtain

$$\frac{1}{2}\left(c_+ + c_-\right) = c_0 + \mathcal{O}\left(\Phi^2\right). \tag{5.55}$$

Thus, for sufficiently low electrostatic potentials, the mean of both ion concentrations equals the reservoir concentration. We observed in our simulations, that for distances between 3 and 4 nm away from DNA, the local ion concentrations of both species are very similar for all considered datasets. This is shown in detail below. Therefore, the mean concentration of both species in this region was assumed to be the reservoir concentration.

In order to check that this method of determining the reservoir concentration is valid we performed further simulations in a box without the constraining wall particles with a square cross section, which was fully filled with water and where we applied periodic boundary conditions in all direction. Since in this case the box is much larger, it can serve as a reservoir, and at the border the DNA's presence is fully screened. We will briefly elucidate this by discussing another quantity, the mean ion density, in the total simulation volume.

In the simulations with a pore wall, a volume

$$V_1 = \pi R^2 L_z, \tag{5.56}$$

is accessible to DNA, water and ions, where R is the pore radius (5 nm) and L_z the box length in axial direction (6.76 nm). In the simulations without pore a volume

$$V_2 = L_x L_y L_z \tag{5.57}$$

is available. Here L_x and L_y are the box dimensions in x and y direction (15 nm). The available volume V_2 is approximately four times larger than V_1. We performed extra simulations in the larger system where we wanted to reproduce the ion distribution for a certain dataset of the original simulations with pore. Let N_1 be the number of ion pairs which we had used in the original simulations and c_0 the effective reservoir concentration we have determined from the original simulations. The additional number of ions pairs to add to the box was then determined from the reservoir concentration c_0 we had determined from the method described above:

$$N_2 = N_1 + (V_2 - V_1)c_0. \tag{5.58}$$

This was justified as the pore originally had already been large enough to provide an almost complete screening of the DNA's charge. We chose the original datasets with $N_1 = 16$ (lowest concentration), $N_1 = 64$ (close to the crossover) and $N_1 = 256$ (a high salt concentration of almost 1 mol/l). Eq. 5.58 gave $N_2 = 64$, $N_2 = 241$ and $N_2 = 842$, respectively. We compared the ion distributions explicitly for all data sets in Fig. 5.16. As can be seen, for all datasets the ion distribution around the DNA of the original data and the new simulations are in agreement, and the ion concentration in the new datasets converges as anticipated to the reservoir c_0 concentration we had determined before. An important point however was, that the relaxation time of the ion density was significantly larger in these simulations and the first four nanoseconds of every run had to be discarded.

In addition, we investigate how the ratio of ion pairs per volume N/V (expressed in mol/l here) compares to the reservoir concentration c_0 we had determined from the original simulations. In the limit of an infinite box they should be identical. For the now much larger box, we expect them to be closer. The results are shown in table 5.2. Indeed, the difference between N/V and c_0 decrease with the larger accessible volume. The factor of four in volume is reflected in a comparable reduction of the difference.

We have shown that this method allows us to determine the effective reser-

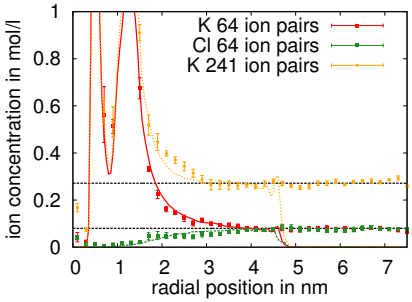


Figure 5.15: Local ion concentration around the DNA. The data of the original setup with pore are shown as lines, and the data of the setup without pore are shown as symbols with error bars. The K concentration (solid line) and Cl concentration (dashed line) of the original run with 16 ion pairs (dotted line) are compared to simulation in a cuboid box where the number of ion pairs we had to add was determined from eq. 5.58. The good agreement between the ion distributions indicates the method that we used to determine c_0 was correct.

voir concentration *a posteriori* without extra computational effort. The same analysis mechanism was also used for the simulations without DNA to make sure that boundary effects caused by the pore are correspondingly taken into account.

5.3.3 Simulation results

In the following we perform a detailed investigation of the static and dynamic properties of the system. It is split into two parts: First, the density profiles $c^{\pm}(r)$ and velocity profiles $v^{\pm}(r), u(r)$ are compared to the continuum model. Second we determine the electric current through the pore and compare this to the experimental results of [5].

The static ion density profile reflects the typical behavior known from the literature (e.g. [299,327,328]). Fig. 5.16 shows, as an example, the simulation with 64 added ion pairs. The counterion density exhibits two significant peaks. One peak occurs at r=0.5 nm (full height not shown), and a second peak is observed at r=1.25 nm. The first peak is related to counterions entering into the major and minor grove of the DNA. This peak compensates for around

N_1	N_1/V_1	N_2	N_2/V_2	c_0
16	0.050 mol/l	64	0.070 mol/l	0.080 mol/l
80	0.28 mol/l	278	0.26 mol/l	0.25 mol/l
256	0.80 mol/l	842	0.92 mol/l	0.98 mol/l

Table 5.2: Comparison of the mean ion density N/V and the concentration of a hypothetical reservoir c_0, from simulations in the cylindrical pore geometry (1) and in the cuboid box without constraining the pore pseudoatoms (2). In the second set of simulations the deviation between mean ion density and c_0 is approximately a factor of four smaller, reflecting the volume increase by a factor of four.

1/3 of the DNA's bare charge at all considered salt concentration. The second peak indicates the effective radius of the DNA backbone. For larger distances than 1.25 nm, the ion distribution is well described by the continuum model. In the region between $r = 3\,\text{nm}$ and $r = 4\,\text{nm}$ an almost complete screening of the DNA charge was observed for all salt concentrations.

In the right panel of Fig. 5.16 we display the ion velocities $v^{\pm}$ of both ion species and the water velocity u as a function of the radial position r for the same dataset. The electrokinetic infinite cylinder model (dotted curves) and the atomistic simulation data are qualitatively similar. An electroosmotic flow of water is observed in the migration direction of the K^+ ions. Both ion species move significantly faster than the water whose velocity tends to zero in the atomistic model approximately at the inner boundary of the continuum model. The agreement of the hydrodynamic no-slip boundary between both models up to the size of a water molecule is remarkable.

Motivated by the simple decomposition of the ion velocity $v^{\pm} = \mu^{\pm} E + u$ in the continuum model, we define a position dependent mobility of the ions by computing their actual velocity minus the local water velocity, $\mu(r)^{\pm} = \left(v^{\pm} - u\right)/E$. In the region farthest away from both the pore wall and the DNA the mobility is constant, and we denote this constant value by μ_D, the free mobility. We find that its value is concentration dependent, and the trends of the experimental data in bulk electrolyte [283] are well reproduced (not shown). This was, however, not investigated in more detail. In the left panel of Fig. 5.17 we show the position dependent K^+ mobility for electrolyte concentrations between 0.15M and 1.2M, normalized by μ_D. All curves collapse onto a single master curve. The range of the mobility reduction is comparable with the effect found in Ref. [329]. The different geometries, however, make a direct comparison difficult.

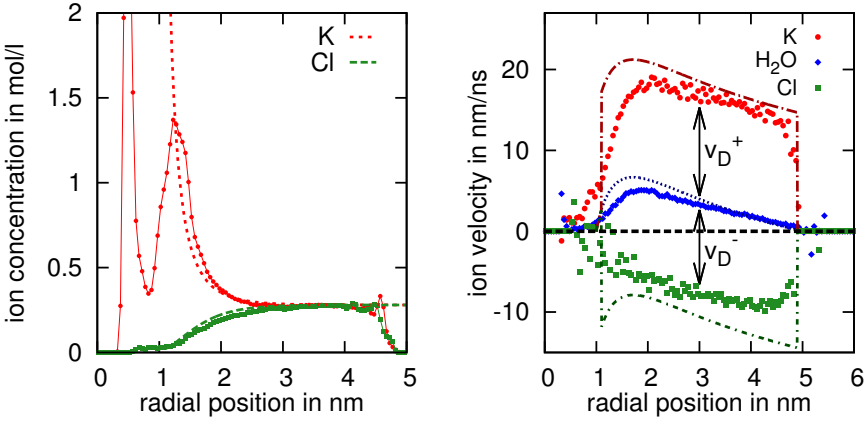
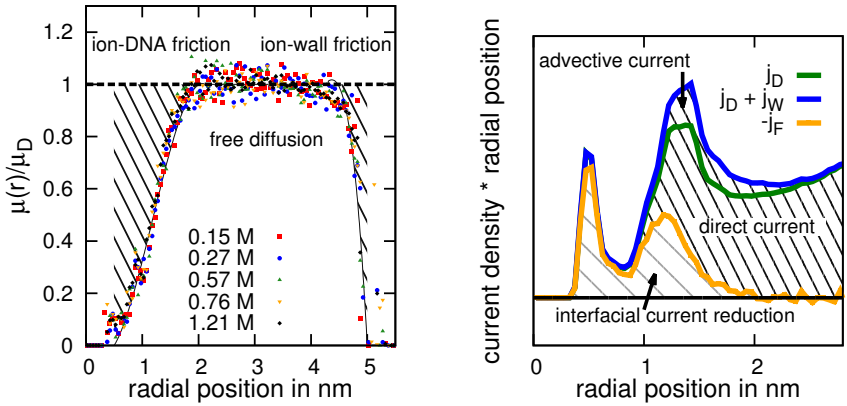


Figure 5.16: (a) Comparison of the ion concentration as a function of the distance r from the pore axis at around 0.3 mol/l of the atomistic data (points) to the continuum model (broken lines). The long distance decay of the second peak at r=1.25 nm is well described by the continuum model. (b) Comparison of the measured ion and water velocities as a function of r.

Zhang and Shklovskii e.g. introduced a friction coefficient between ions and DNA and [11], which they calculate based on hydrodynamic arguments, similar to those we brought forward in Sec. 5.2.3. In Comer et al. measured the diffusion coefficient of ions in the vicinity of the different nucleotides in atomistic simulations [329]. They found that the diffusion coefficient changes by more than a factor of two in the first nanometer from the nucleotide, independent of the specific base. We interpret the mobility reduction at the boundary as being caused by three factors. First, the major and minor groove produce a microscopically rough DNA surface, and ions trapped inside are virtually immobile. Second, the charge pattern on the DNA causes electrofriction [330, 331] that can extend beyond the DNA backbone. Finally, hydrodynamic effects will slow down the motion of particles near a hydrodynamic boundary, as Comer et al. have shown.

In order to quantify how the mobility reduction next to the boundary affects the current, we split the current into three contributions: (1) the current I_{D} that would be expected if the mobility was constant and the water was



(a) (b)

Figure 5.17: (a) Position dependent ion mobility normalized by the bulk mobility $(v-u)/\mu_D E$ for five different salt concentration. Near the DNA and the wall a significant reduction is observed. (b) Observed current densities as a function of r. We distinguish the direct current j_D, the convective current j_W and the (negative) frictional current j_F. The direct current near the DNA is largely cancelled by friction.

immobile, (2) the convective current I_{W} due to the water flow, and (3) the (negative) current contribution I_{F} that is suppressed by interfacial friction. We express all of these in terms of current densities which we define as the direct current density $j_{\mathrm{D}} = c^+\mu_D^+ E - c^-\mu_D^- E$, the convective current density (as in [332]) $j_{\mathrm{W}} = c^+u - c^-u$, and the interfacial current reduction density $j_{\mathrm{F}} = j - j_{\mathrm{D}} - j_{\mathrm{W}}$, where j is the total current density. The direct current depends, apart from the concentration dependence of the mobility, only on the number of ions of both species inside the pore. Therefore it contains both, the extra counterions that the DNA brings into the pore, and the ions expelled from the pore due to the DNA's finite volume.

In the right panel of Fig. 5.17 we report the direct and convective current density, and the density of interfacial current reduction. They are multiplied by the radius, so that the area under the graph corresponds to the magnitude of the associated current. The area between the j_F curve (yellow) and the sum of j_D and j_W (blue) represents the total current. The shape of the

direct current follows exactly the ion concentration reported in Fig. 5.16. The convective current $j_W = (c_+ - c_-)\,u$ is nonzero only in the regions with net charge, namely the Debye layer. It only occurs if a DNA is in the pore, and it is a positive modulation of the total current. In the proximity of the DNA, the conduction is largely suppressed. The fact that ions in the DNA's grooves are virtually immobile is reflected in j_D and j_F being almost equal. This however applies only to 1/3 of the counterions. Ions further away from the pore axis, also at $r > 1\,\mathrm{nm}$, are considerably reduced in mobility, resulting in an larger frictional current reduction.

We performed identical simulations also for a DNA-free pore. We investigate the relative current modulation $(I_{\text{with DNA}} - I_{\text{without DNA}})\,/I_{\text{without DNA}}$ as this quantity depends neither on the length of the investigated pore nor on the absolute value of the mobility of the ions. The total current was calculated from the displacement of the ions during a simulation run. In Fig. 5.18 the results are displayed together with the experimental data. The lines are only guides to the eye. The current obtained in the DNA-free simulations is interpolated by a third order polynomial to get a continuous function to determine the relative current alteration for every salt concentration. We decompose our data into a direct, an convective and a frictional component. For all salt concentrations the current modulations from the simulations are in good agreement with the experimental data. For high salt conditions a current reduction of around 5% is observed. The experiments suggest a somewhat lower reduction, however with uncertainties similar to the deviations between experiments and simulation. For low salt concentrations the predicted enhancement is larger than experimentally observed. This can be justified by noting that the experimentally used nanopores are not very long (around 40 nm) and have an hourglass shape. Only the radius at the constriction corresponds to the 10 nm used here. When the pore is wider the relative enhancement or blockade can be expected to be smaller. The crossover point between enhancement and blockade occurs at an electrolyte concentration of 0.3 mol/l, agreeing with the experiments.

The observed direct current is always larger in presence of the DNA since the number of ions in the pore at equal reservoir concentration is always larger with DNA in the investigated concentration range up to 1.2 mol/l. The continuum model predicts this to be the case up to 0.85 mol/l. Thus the obstruction of the cross section due to the DNA's finite volume is always overcompensated by the presence of the Debye layer. The convective contribution to the current only appears in the presence of a DNA strand and is always positive. The effect that is responsible for the crossover effect, however, can be clearly identified. Without the mobility reduction at the surface of the DNA no current

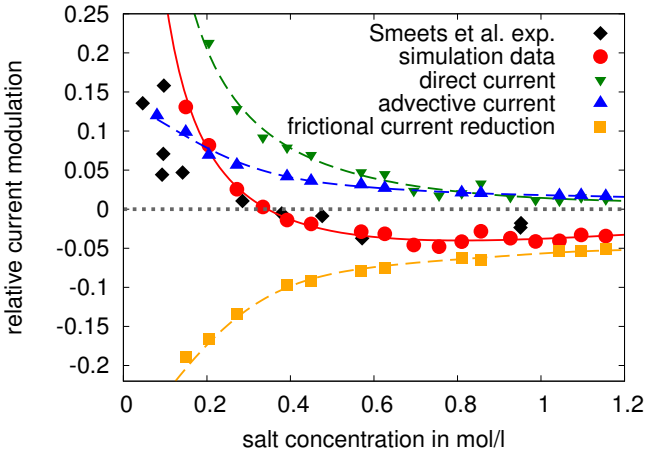


Figure 5.18: Relative current modulation caused by the DNA as a function of salt concentration for the experiments (diamonds) and simulations (circles). The quantitative agreement between simulations and experiments is good, especially the location of the crossover point between enhancement and reduction. We decomposed the total current into direct current (downward triangles), convective current (upward triangles), and friction (squares) and consider their contribution to the modulation independently. The current reduction due to friction is responsible for observing an overall current reduction at large salt concentrations.

blockades would be observed.

5.3.4 Visual analysis of sequence-specific effects

In the one-dimensional radial projection of the ion densities and velocities, very interesting information is lost. So far we could only speculate about the mechanisms slowing down ions in the vicinity of the DNA molecule. To understand the mechanisms in more detail, it is useful to investigate the data in all spatial dimensions, especially to identify how the molecular structure of the DNA affects the conduction process. After all, if DNA sequencing from measuring the ion current is possible at all, the current modulation must somehow depend on the DNA sequence in the pore. Even though double-

stranded DNA is less well suited for sequencing we decided, as a first step, to compare the current modulation of a poly-CG and a poly-AT DNA.

The analysis was performed in cooperation with the Visualisation Research Centre of the University in Stuttgart (VISUS). The goal of our joint project was to analyze our data with regard to correspondences between the ion dynamics and the DNA structure. From the visualization point of view, this problem is tackled with a *visual analytics* approach [333–335]. The key ingredient in visual analytics is using the human's perceptual and cognitive capability and extending it by suitable visualization. In a tight interaction of visualization development and the user, tailored techniques of data visualization, abstraction and filtering are developed. This approach stems from the problem that available datasets are growing in size and complexity. The human is put "in the loop" of the process of representation of the data, visual interaction and computational analysis. In this incremental process of insight and refinement, researchers can work out the "important" aspects of a dataset.

An important necessity for visual analytics are visualization techniques which allow for interaction with the data. This requires the use of parallel computer architectures and efficient algorithms. In our case this is realized by using the software framework `MegaMol` [336], which is developed at VISUS for interactive visualization of large and complex datasets. Our effort resulted in a software application where data aggregated from many independent simulation is displayed in conjunction with a particles-based representation of the DNA. By using modern rendering techniques on state-of-the-art hardware this allows for visualizations of our datasets with at least 30 frames per second. All visualization parameters can be adjusted on the fly. The user can interactively rotate the representation, adjust thresholds, add filters, make elements transparent, add and remove different visualization elements, and explore the data for interesting features.

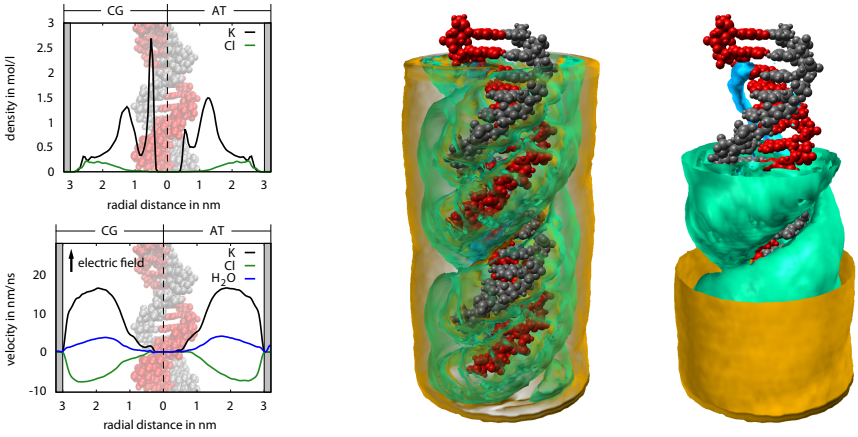
Different visualization techniques were combined in this application, which are briefly outlined below. Before that, we briefly describe the applied aggregation technique.

A challenging aspect of a three-dimensional data representation is the noise of the data due to statistical uncertainties. To maximize the statistical accuracy we reduced the pore radius to three nanometers and ran in total 100 simulations of eight nanoseconds, with otherwise identical simulation parameters as above. With a C program, a three-dimensional binning of the density and the flux density was performed. Each ion was distributed over the eight neighbouring bins with weights according to the multilinear expression

which is also used for the interpolation of Lattice-Boltzmann velocities. In this binning method we made use of the intrinsic symmetry of the data. By construction, the system is, except for artifacts due to the cubic periodicity of the simulation box, invariant under a translation of 0.338 nm and rotation by 36°, as all base pairs are identical. We therefore counted each ion twenty times at coordinates translated by $n{\cdot}0.338\,\mathrm{nm}$ and rotated by $n \cdot 36°$. This minimizes the statistical errors, but the resulting dataset has the same shape as the simulation box, and is therefore easy to use in the visualization software. We applied a file format from the `VTK` toolbox [337], so, in principle, the data can also be visualized with `VTK`-based tools as `ParaView` [338].

As means to display the ion densities, we apply *isosurfaces*, a very common and intuitive technique. An isosurface is defined as the level-set of a scalar function in space [339]. They are often visualized using either a mesh generated by method like Marching Cubes [340] or direct volume rendering. In our application, we use GPU-based ray marching [341], a direct volume rendering method that can be used to render a smooth isosurface. The basic idea of ray marching is to traverse the scene along a ray that starts at the camera, thereby sampling the scalar field in a given interval. If the scalar value crosses the given isovalue between two sample points (which are obtained using trilinear interpolation), the location of the isosurface along the ray is approximated using linear interpolation. Isosurface ray marching yields high visual quality and runs very fast on current graphics hardware [341]. It has the additional inherent benefit that it renders semi-transparent surfaces correctly. Further information can be added to the isosurface by a color coding.

Originally we planned to apply stream lines to visualize the velocity vector field. Stream lines are the trajectories of massless particles where the force on the particle corresponds to the value of the vector field at each position. Further information can be encoded in the local color and thickness of the line. Visual clutter can be avoided by limiting the number of stream lines. By graphical insepction, it however turned out that stream lines are only well suited for divergence-free fields, such as flow fields and electrostatic fields. The velocity field is not divergence-free, but only the flux density field, and therefore the results appeared implausible to the observer. Alternative to stream lines we implemented a glyph-based representation of the velocity field. In this case we chose arrows as glyphs, indicating the direction of the velocity field. The length was scaled by the velocity and with different color codings extra information could be added. The DNA atoms are represented by spherical glyphs, where the color coding is an adjustable mix of molecule index (in this case in index of the DNA strand it belongs to) and the atom type. Also



(a) (b)

Figure 5.19: (a) One-dimensional projection of the ion density (top) and the ion velocity (bottom) for both CG- and AT-DNA. The right parts of the diagrams show the curves for CG-DNA, while the left parts show the curves for AT-DNA. The DNA strand and the pore walls are given for reference. (b) Two screenshots taken from our application showing three ion density isosurfaces for the CG-DNA molecule. The innermost, blue isosurface encapsulate ions inside the major groove of the DNA and strictly follows the helix shape. The outermost, yellow isosurface exhibits an almost perfectly circular cross section with a radius of ∼1.5 nm, indicating that for larger distances a rod-like model of DNA is absolutely adequate.

other representation, such a *stick* representation, where each chemical bond is represented as a stick connecting the atom centers, are available.

First we compare the ion densities of the CG- and AT-DNA in the well-known one-dimensional projection in the upper left left panel of Fig. 5.19a. It is noted that the very high first peak in the potassium ion concentration observed already in the CG-Simulations before is much less pronounced for the AT-DNA. In the latter case the fraction of charge of the DNA already compensated for within the grooves is smaller. Correspondingly the second peak, corresponding to the onset of the Debye-type screening is slightly higher. The chloride ion distribution in both cases is indistinguishable. At the pore boundary a surprisingly pronounced layering behaviour can be observed.

In Fig. 5.19b we display isosurfaces of the ion densities around the CG-DNA in two variants. The two snapshots were rendered with the raycasting method described above, where in the left variant the isosurfaces were rendered with a high degree of transparency, and as opaque surfaces in the right variant. The iso-levels of the three surfaces correspond to around 0.8 M, 1.6 M and 3.5 M respectively. The outermost yellow surfaces is virtually cylindric with a radius of around 1.5 nm. This indicates that already at this short distance the structure of the DNA is entirely smeared out. The ion distribution is indistinguishable from that of a structureless rod. Close to the DNA more structure can be observed. The intermediate isosurface is almost cylindrical, but exhibits a gap in the major groove of the DNA. The isosurface with the highest density forms a tube-like structure within the major groove, indicating a region with a high binding affinity for the ions.

In Fig. 5.20 we have highlighted the qualitative differences of the regions with the highest ion concentrations of the CG and AT simulations. We combine a transparent isosurface rendering with a stick representation of the DNA molecule, which is cropped to a region of interest at a particular base pair. In the CG dataset we observe a tube structure of highest ion concentration which winds around the DNA axis. It is located directly at the N and O groups of the C or G base. At positions between two base pairs, the O atom of the next pair is also close, and this effect makes the tube structure continuous. Even for the highest iso-values no isolated spots are observed. In case of the AT dataset, two binding regions can be identified. On the one hand, ions accumulate in the major grove in direct vicinity of the hydrogen bond-forming atoms of the nucleobase. This effect is, however, significantly lower than for the CG homopolymer, and the binding spots form a pearl-necklace structure. The regions of highest potassium density are located adjacent to the phosphate groups of the backbone. Each of the two oxygen atoms of the phosphate groups bear a charge of -0.7761e, hence this binding behaviour likely to be caused by electrostatic interactions.

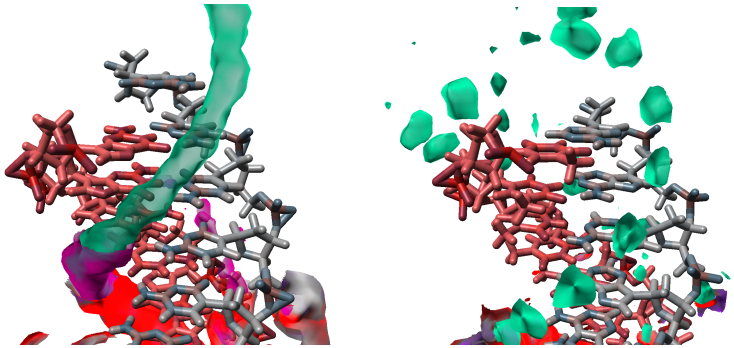


Figure 5.20: Ion binding to a CG-DNA (left) and an AT-DNA (right). For the CG-DNA the innermost layer of ions binds very tightly to the major groove. The ions prefer close binding to the nitrogen and oxygen atoms in the major grove which are involved in the base pairing. In case of AT-DNA the major groove is far less attractive. The ions also bind there, but much stronger to the negatively charged oxygen atoms at the backbone.

Furthermore we investigated the ion velocity field. In Fig. 5.21 we present a combined rendering of stream lines and arrow glyphs of the velocity field around the CG DNA. Additionally, the highest density isosurface is color-coded from red to blue indicating an outflow or inflow, respectively. To our surprise this rendering was not very helpful as the streamlines exhibit very unstructured behaviour, and some arrows point upward and some point downward. But we could make a very interesting observation. When we reduced the sampling rate with which the velocity field was calculated, we found the general trend that the magnitude of the velocity field reduced significantly (right panel). Reducing the sampling rate in this case means that the ion velocities are not calculated from their displacement between subsequent snapshots, but from snapshots with a higher distance along the time axis. This results in a smoothening of the trajectories, similar to a low-pass filter. This result is remarkable. The arrows pointing in seemingly random directions are not a result of statistical noise, because this would have increased when lowering the available number of independent snapshots. On the contrary, the ions must perform some high frequency vibrational motion that does not average out.

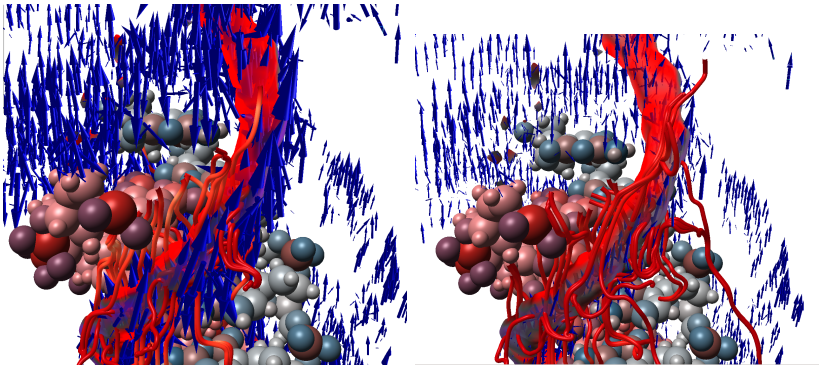


Figure 5.21: Ion motion near the CG-DNA taken with two different sampling rates: 5/ps (left) and 0.1/ps (right). The color coding indicates the normal component of the local ion velocity, and the ion velocity is additionally displayed as arrows. Stream lines support the perception of the vector field. On short time scales a rapid, ballistic ion motion is apparent. If the trajectories are smoothened by using a lower sampling rates, the short time motion is averaged out. On long time scales, the motion of the ions is largely suppressed, reflecting the friction effect near the surface.

We adapted our visualisations to investigate this phenomenon in more detail. Fig. 5.22 shows the arrow glyph visualisation of the ion velocity vector field inside the grooves (regions with high ion density, green isosurface). The arrows are colour-coded according to their direction: Red arrows indicate velocity vectors that point in the direction of the electric current. This is the anticipated direction of ion motion. There is, however, a region in the major groove where the ions move in the opposite direction. Intuitively, this should not happen, since the force of the applied current is stronger than other competing effects (e.g. negatively charged atoms in the DNA). By reducing the sampling rate, which leads to a smoothing the individual trajectories, this effect is weakened, and entirely disappears for even lower sampling rates. Ultimately, the up- and down-motion of the ions within the major groove cancel almost entirely, as seen in the one-dimensional projection of Fig. 5.19a. The origin of the behaviour, however, remains obscure to us. The phenomenon is not understood, and we can only speculate about the underlying mechanism. The close interaction with the visualization department, allowed us to clearly and unambiguously identify this effect. We suggest to study it further in a coarse-grained scenario, where the parameters are under better control and

the computational effort is smaller.

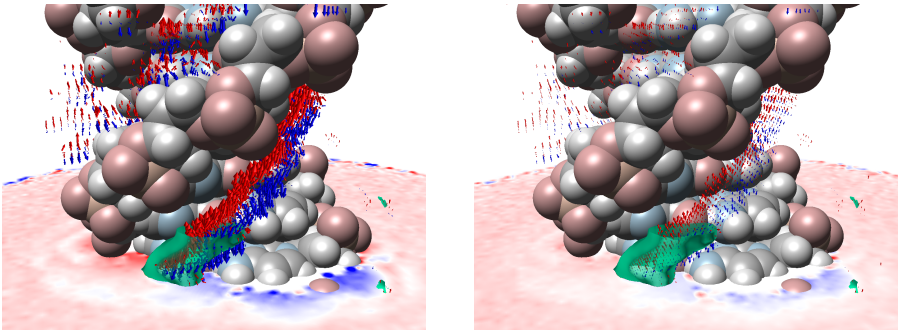


Figure 5.22: Ion motion inside the major and minor groove of CG-DNA taken with two different sampling rates: 5/ps (left) and 0.1/ps (right). The green isosurface encircles the region with high ion density, which is following the the major groove. In this region, ion motion is partially in compliance with the electric current (upwards, red arrow glyphs). There is, however, a distinct region where the direction of the ion motion opposes the direction of the current (blue arrow glyphs). The atoms of the DNA are colour-coded by their chemical element using the common CPK colouring scheme.

5.4 Summary

We have investigated the question what mechanisms cause ion current modulations in nanopore translocation experiments. More specifically, have we investigated why a DNA in a nanopore can cause current enhancements at low salt concentrations, and current reductions at high salt concentrations of the buffer solution. Four important effects and their qualitative influence on the current modulation could be identified: (a) Excluded volume. Ions are displaced from the volume occupied by a DNA molecule. This leads to a current reduction. (b) DNA charge. As the DNA is charged it is surrounded by a counterion cloud. Under most experimental conditions this counterion cloud can enter the pore and the counterion presence enhances the conductivity. (c) Electroosmostic flow. The DNA creates an electroosmotic flow of water molecules inside the pore. This leads to an convective transport of ions and hence an increase in conductance. (d) Surface friction. Ions near the DNA surface are less mobile than ions far away from the pore. This decreases the conductance. The relative importance of these mechanisms depends strongly on the details of the considered system.

As a first step, we applied a very simple model where the DNA is modelled as a infinite charged cylinder in a cylindrical pore and where the ion motion is described by the electrokinetic equations. In this model the experimental conductivity modulations can not be observed in agreement with the experiments. For concentrations up to 1.2 M the conductance with DNA is higher than without DNA. By replacing the continuum description of ions and water by a explicit particles and an LB fluid, the crossover concentration between enhancement and reduction moved down to around 0.7 M, which is in better agreement with the experiments. The MD/LB simulations contain hydrodynamic interactions of ions with interfaces, which reduced the mobility at the boundary. This friction effect is conceptually not contained in the electrokinetic equations. In a third step, we performed simulations of the same system with atomistic resolution, including an all-atom model of the DNA molecule. In this model the friction of ions with the DNA molecule is more pronounced, and a crossover behaviour in agreement with the experiments is found.

The continuum model based on the electrokinetic equations had been successfully applied to predict the electrohydrodynamic forces acting on DNA molecules in pores [15]. Its most important drawback is that conceptually no friction effects are included. In principle, it is possible to add this effect by using a position-dependent diffusion coefficient. A second effect missing in the electrokinetic equation is the concentration-dependent conductance of

electrolytes. This effect is of electrohydrodynamic origin and, to our surprise, is included in the LB/MD model almost in quantitative agreement with experimental data. It however only weakly affects the question of the magnitude of the current modulation in a nanopore caused by a DNA molecule. In the LB/MD model we observe a friction effect between ions and surfaces caused by hydrodynamic interactions. As the DNA surface in our model is only represented in a very primitive way we can, however, not expect quantitative agreement.

The LB/MD method we used so far had only rarely been used as a tool for quantitative predictions. We have shown that by choosing the right parameters a well defined hydrodynamic radius of the particles is obtained. The maximum hydrodynamic radius that can be obtained with the nearest-neighbour coupling suggested by Ahlrichs and Duenweg [278] is around half of the applied lattice constant. Choosing too small viscosities can lead to a violation of the fluctuation-dissipation relation and is therefore discouraged. We have shown, that quantitative predictions in agreement with the electrokinetic equations plus the two extra effects contained in the simulation method are possible. Our analysis of the methodology is, however, not complete. We will comment on that discussing further work.

Quantitative agreement with the experimental data could be achieved by rebuilding the same geometry in all-atom simulations. We found a surface friction effect that is significantly more pronounced than in the LB/MD simulations. It leads to an reduced effective conductivity of the counterion cloud. The general agreement between the all-atom model and the coarser models presented above is surprisingly good. For distances larger than around 1.5 nm the ion density measured in atomistic simulation is virtually indistinguishable from that of a charged cylinder. Also the magnitude of the electroosmotic flow and the position of the effective hydrodynamic boundary in both models agree remarkably well. In a detailed three-dimensional visual analysis of the ion densities around a DNA molecule determined from atomistic simulations we could clearly distinguish features of a poly-AT and a poly-CG DNA molecule. For the poly-CG a significant amount of ions accumulates in the major groove, which is almost absent of the poly-AT DNA. In this case the ion densities in vicinity of the phosphate group of the backbone, where the DNA charge is located, is somewhat higher. The visual analysis also revealed a very interesting behaviour. In the tube of highest ion concentration within the major groove of the CG-DNA in one region the ions move in the direction of the applied field, and in some regions in the opposite directions. Both currents largely cancel out. So far, this behaviour is not understood, but it appears related to a high frequency ballistic motion of the ions.

Choosing a suitable set of interaction parameters for the all-atom simulations turned out to be more difficult than anticipated. No standard set of parameters could yield hydrodynamic ion radii in agreement with experiments. We applied a combination of the AMBER force field, the SPC/E water model, and an ion model suggested by Dang. This was the only combination we found in the literature that had been tested for DNA simulations and where the hydrodynamic radii are not much larger than the experimental values of potassium and chloride ions. The role of the force field for the simulation results remains an open question. Below we will discuss how these investigations can be extended with regard to this question.

6 Conclusions

In this concluding chapter I will review the work that was done, discuss the modelling strategies and formulate ideas on how to proceed with this work. Except for the summary, it is written in the first person singular voice, as the judgements and views put forward are mostly my personal opinion. At the point of writing this thesis the application of an extension of this project, together with the Special Research Center (Sonderforschungsbereich) it is situated in, are prepared to be extended. Also, a long-term collaboration with the experimental group of U. Keyser, which possibly includes myself, is being prepared. Thus, many plans and different strategies have been discussed in the last months. This chapter is a good opportunity to formulate my judgement on the present findings and recommendations for future work. Therefore, this chapter has become quite extensive.

6.1 Summary

In this thesis we studied nanopores, one of the most promising biosensor technologies. The measurement principle is astonishingly simple. A membrane separates two water-filled compartments and only a single nanometer-sized hole connects them. By applying a voltage, biomolecules can be driven through the pore. The voltage also causes a current of ions, e.g. KCl, through the pore, which can be measured. From ion current measurements, detailed information about the transported biomolecule can be obtained. The most attractive application is the sequencing of DNA, which can be potentially made much cheaper and faster than the current state of the art. The versatility of nanopores, however, makes them also very attractive for other scientific and commercial applications.

We studied two fundamental aspects of nanopore-based sensors: (a) Are there free energy barriers of electrostatic origin which prevent charged molecules

from crossing pores? (b) What is the physical origin of the ion current modulations caused by DNA molecules transported through a nanopore? Both questions were answered with a combination of particle-based molecular-dynamics simulations and the application of a continuum theory.

We investigated the question regarding electrostatic free energy barriers with simple coarse-grained models for double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA). Short dsDNA fragments were modelled as stiff rods of 2 nm diameter and ssDNA was modelled by a flexible bead-spring model. In both cases we chose the charge density in agreement with the experimental charge density of about 6 e/nm and 3 e/nm. A univalent electrolyte was modelled as charged spheres in an implicit solvent model, where water was assumed to act as a homogeneous dielectric medium with a permittivity of 80. We investigated especially the role of the dielectric permittivity of the membrane material. This was done by means of the ICC$\star$ algorithm, where induced charges on a dielectricm interface are determined iteratively in every simulation step. This algorithm was improved considerably, parallelized and is available as part of the simulation package ESPResSo since version 3.1. In our simulation the pore was modelled as a homogeneous dielectric medium with a permittivity of 2, reflecting the sharp dielectric contrast between typical membrane materials and the surrounding water. It was assumed to be an eight nanometers long, cylindrical cavity in a flat wall. We compared the cases with and without taking the dielectric contrast into account.

We determined the free energy profile for the transport of DNA molecules by calculating the potential of mean force. For dsDNA we used the axial coordinate of the center of mass of the DNA as a reaction coordinate and additionally assumed that the DNA is constraint to the pore axis. We measured the free energy barrier for a 10 nm dsDNA fragment in pores of different diameters at different electrolyte concentrations. The barrier height we determined were as large as $20\,k_BT$, corresponding to around 50 kJ/mol for small pores and low salt concentrations. With increasing salt concentration and pore radius the barrier sharply decreases as more and more of the DNA charge is screened due the formation of a counterion cloud. The relevant length scale for screening is the Debye length which is between around 3 nm at 0.01 mol/l and 0.3 nm at 1 mol/l. Its ratio to the pore radius governs the magnitude of the barrier height. The dielectric contrast increases the barrier height. As a rule of thumb, we found that it approximately doubles under the considered conditions. This is, however, only valid for the pore length we considered.

For single-stranded DNA we suggested a new reaction coordinate which allows us to describe exactly both the approach of one end of the DNA towards

the pore opening and the transport process. This is achieved by using the Euclidean distance of the DNA end to the pore center in if the DNA does not cross the pore, and by using the fraction of the DNA that has already translocated if the DNA crosses the pore. By stitching these two regions together, a continuous function is obtained that describes over the whole process how far the translocation has proceeded. With this reaction coordinate we calculated the free energy barrier for a 40 nm ssDNA molecule in a 10 nm wide pore. We found it to be between 15 and 5 k_BT depending on the salt concentration. Even at high salt concentrations an barrier of a few k_BT remains, reflecting the entropic loss of the inner degrees of freedom of the DNA, as it has to assume a shape compatible with the membrane. The dielectric permittivity here also increases the barrier significantly.

Interestingly we found that the free energy barriers calculated for dsDNA correlated well with the free energy barrier calculated from a very simple model based on the Poisson-Boltzmann theory. In the Poisson-Boltzmann theory, the ion density in the electrostatic potential is assumed to be governed by the Boltzmann distribution, while the electrostatic potential is generated by the ions themselves. It is a simple mean-field theory. In the applied model, DNA is assumed to be an infinite charged cylinder which is centered in an infinite uncharged cylinder. The free energy barrier is calculated from solving the resulting one-dimensional differential equation numerically in pores of different radii and calculating the work necessary to shrink the pore to the desired size. The dielectric permittivity does not appear on this level of description as the pore must be entirely neutralized and no induced charges can occur. We could show that, for the pore length considered in the MD simulations, the DNA is only partly screened but that the central region of a very long pore must be neutral as otherwise the self-energy diverges. This divergence is also responsible for the phenomenon of Manning condensation. When we consider longer pores in the simulations, indeed the central regions of the pore is entirely neutralized regardless if dielectric contrast is considered or not. If dielectric contrast is considered, the screening is stronger and therefore we find such a good correlation between the barriers from the Poisson-Boltzmann model and the simulations including dielectric contrast.

We could support this finding by calculating the free energy barrier with the iPBS method. With the iPBS method, the Poisson-Boltzmann equation can be solved in arbitrary geometries including dielectric contrast between different regions. This is done with a combination of the Finite Element Method and an induced charge iteration similar to the ICC$\star$ algorithm. We applied the iPBS method to calculate the free energy barrier of DNA molecules in a nanopore. In agreement with the MD simulations, DNA was approximated as

a charged cylinder on the axis of a cylindrical pore. We could show that, in the limit of infinite pore length and DNA length, the same free energy barriers as in the one-dimensional model is obtained. This is valid regardless of the question if dielectric contrast between the membrane material and the water is considered. For pores of only 8 nm length, as considered in the MD simulations, we, however, found that the barriers are significantly higher if dielectric contrast is considered.

Summarizing our findings regarding electrostatic free energy barriers we find: Electrostatic interactions are subject to screening due to the presence of salt ions. Under all relevant experimental conditions, a significant amount of salt is present. The charge can still lead to considerable free energy barriers, if the Debye length is comparable to the pore radius. The dielectric contrast does not change this picture qualitatively. It increases the tendency that charged objects in front of dielectric interfaces are neutralized. For long molecules in long pores, the dielectric contrast does not have to be taken into account, as the self-energy which increases with increasing length also requires neutralization.

The second question we investigated was the physical origin of ion current modulations in DNA translocation experiments. In virtually all translocation experiments a voltage is applied over the nanopore and the molecules to be sensed are dragged into the pore. During the experiment, a steady electric current formed by ions, typically KCl, develops. When DNA molecules, enter the pore, their presence alters the electric current. Four important effects contribute to these modulations modulation: (1) the excluded volume of the DNA, (2) the DNA charge, (3) electroosmotic flow and (4) ion friction near the DNA surface. Each of these effects is suggested in the literature, but so far they have not been quantified in detail.

Our goal was to improve the quantitative understanding of the interplay of these effects. We performed a modelling procedure composed of three stages: We started off with a continuum model based on the electrokinetic equations. Then we constructed a coarse grained simulation model. Finally we performed simulations with atomistic resolution. The challenge in this process was obtaining mutually compatible results, or a consistent hierarchy of models where only more details but not systematic deviations were introduced on each level.

First, we applied the model developed by van Dorp et al. [15] to predict the ion current modulation caused by a DNA molecule. It was the inspiration for the simple one-dimensional continuum model discussed above and therefore is very similar. A cylinder bearing the line charge density of double-stranded

DNA is embedded in an uncharged cylinder. The motion of ions and water are described by the electrokinetic equations. If no field was applied this model would reduce to the Poisson-Boltzmann equation, and hence exactly to the infinite cylinder model of the first section. An electric field is applied along the pore axis, which leads to the transport of ions and water along the pore. From the integral of the current density over the pore cross section the current is inferred. This model contains three of the four effects mentioned above. The cylinder radius governs the excluded volume of the DNA. It reduces the cross section and hence the number of ions available. The solution of the Poisson-Boltzmann equation takes into account the charge of the DNA, and how many counterions enter and how many coions leave the pore in equilibrium. Finally the solution of the Stokes equation leads to a convective contribution to the ion transport, because counterions travelling in the direction of the developing electroosmotic flow can move easier.

The experimental data we wanted to understand are, among others, the experiments of Smeets et al. [5]. They observed, that in translocation experiments with dsDNA through solid-state nanopores the sign of the current modulations depends on the concentration of the electrolyte that is used. For low salt concentrations, they observed positive modulations; the ion current is increased by the presence of a DNA molecule inside the pore. For high salt concentrations they observed the behaviour found in most experiments, a current reduction. The crossover between current increase and reduction was determined at concentrations between 300 mM and 400 mM. The model of van Dorp et al. had originally been developed to predict the force necessary to stall a DNA molecule in the pore in a very similar experiment setup. Therefore, we were optimistic that it would be successful in describing the ion current as well. However, we found that the crossover is predicted at concentrations larger than 1 M. Even by omitting the convective current contribution the crossover still is above 800 mM. Based on an even simpler model, Smeets et al. had argued that is was justified to assume an "effective" DNA charge significantly smaller than the bare charge could be assumed to explain the data. They speculated that this charge reduction might be related to a reduced mobility of ions. On this level of modelling this effect is conceptually not included and we therefore continued to more complex models of the same system.

We constructed a conceptually identical model based on a hybrid Molecular Dynamics/ Lattice Boltzmann (MD/LB) simulation. This approach is an extension of conventional implicit solvent molecular dynamics simulations, where the dynamic properties of the solvent are taken into account in terms of a very efficient lattice-based fluid representation. Compared to an entirely particle-

based method the computational effort is much lower, and the remarkable computational capabilities of graphics cards can be used efficiently. Particles are coupled to the LB fluid with an frictional force. From a dimensional analysis of the electrokinetic equations we could show that the effective hydrodynamic radius of the ions must be matched to obtain the correct ratio of ion velocity and magnitude of the electrokinetic flow. We could show that for the MD/LB method a unique hydrodynamic particle radius can be identified. By choosing the right particle radius the concentration dependence of the specific conductivity of an electrolyte is predicted in quantitative agreement with experimental data. To study the nanopore conductance properties, we constructed a model reflecting the infinite cylinder model as exactly as possible. We created a cylinder which is periodically repeated through periodic boundary conditions and which bears the same line charge. It was embedded it into a cylinder. The DNA charge was neutralized and a varying number of ions was added to the system. The effective reservoir concentration was determined, in accordance with the electrokinetic model, as the mean concentration near the pore boundary. The radial distribution of ions and the occurring ion and water velocities were in agreement with the predictions of the electrokinetic equations. We found, however, that the absolute conductance of the pore was significantly reduced (up to around 30%) due to the concentration dependent electrolyte conductance. Additionally we found, that the mobility of particles near the walls, including the DNA, was lowered. This is in qualitative agreement with macroscopic hydrodynamics, but the magnitude of the slowdown did not follow any well-established form. This effect, however, shifted the crossover concentration to around 600 mM. As the mobility reduction somewhat arbitrary, we decided to study the same effect with a model of finer resolution.

As a third step, we constructed a very similar model using atomistic simulations. A dsDNA molecule was placed in a periodic simulation box such that it was closed over periodic boundaries. This pseudo-infinite DNA strand of 20 base pairs was fixed in a cylindrical pore represented by Lennard-Jones beads. We chose a combination of ion force field and water model with realistic hydrodynamic radii and used the only atomic force field with which this combination has been tested in the literature. An electric field was applied along the pore axis and we again determined the position-dependent ion densities and ion and water velocities. The ion distributions was in good agreement with the predictions of the electrokinetic model for distance from the DNA axes larger than 1.5 nm. For shorter distances specific binding inside the DNA's major and minor groves occurred. To our surprise especially the electroosmotic flow predicted by the continuum model was almost qualitatively correct. The de-

viations e.g. of the effective hydrodynamic boundary was smaller than the size of a water molecule. The ion velocity is described very well by a position dependent mobility, for which we find a concentration-independent shape, if we correct for the concentration- dependent conductance. Furthermore, we find that this mobility modulation near the DNA surface is responsible for the current blockades in the experiments. We found a crossover concentration in quantitative agreement with the experiments and the modulations we found were somewhat larger than those found experimentally. This, however, can be explained by the fact that we neglect the access resistance of both sides of the pore, which is less affected by the presence of the DNA.

These investigations were performed with a DNA molecule consisting only of cytosine-guanine (CG) base pairs. In addition, we performed simulations with a DNA consisting of only adenine-thymine (AT) base pairs. In a cooperation with the Visualisation Research Centre at the University of Stuttgart we developed a Visual Analytics tool to investigate the detailed features of the simulation results. With this tool we could identify the structures attracting the potassium ions. For the CG molecule we found the N-H-O structure in the major grove as binding site. The continuity of the tube structure is possible, because the next N-H-O region is close as well. For the AT, the amine group in the major groove and the oxygen atom in the minor grove are the sites where the highest potassium concentration is found. The origin of these differences is so far not fully understood. A very interesting observation was made regarding the motions of ions in the tube of highest ion concentration of the CG-DNA. It can be separated into a region where the ions travel parallel to the electric field and a region where the move antiparallel. This behaviour must be related to a high frequency vibration as it disappears when the ion trajectory are sampled with a lower sampling rate. It is so far also not understood.

6.2 Discussion of models and methods

In the present work we applied models based on three different approaches, each with a different resolution. We applied all-atom simulations, implicit solvent simulations, and continuum models. Each of these three approaches has clear strengths and drawbacks. The computational effort and therefore the time and length scales possible to access increase significantly with the resolution. In orders of magnitude, we spent CPU-years for all-atom simulations,

CPU-months for coarse-grained simulations and CPU-days for continuum-level calculations. On the contrary the length scales were nanometers for the atomistic simulations, tens of nanometers for the coarse-grained simulations and hundreds of nanometers on the continuum level. These numbers already indicate that all of these approaches are complementary. A consistent connection between these scales, however, is necessary, and we will summarize our findings regarding the consistency here. We will start the discussion in terms of properties regarding thermodynamic equilibrium and then proceed to dynamic properties.

The electrokinetic equations are derived from very general assumptions and can be expected to be exact in the limit of large length scales and low ion concentrations and low electrostatic surface potentials. While this limit may be unrealistic for many applications we believe that this a helpful starting point. In equilibrium, i.e. in the absence of driving forces like voltages or pressure or concentration gradients, they reduce to a single equation, the Poisson-Boltzmann equation. In this equation only a single free parameter, the Debye length governs the physics. It is the length scale on which electrostatic interactions are screened and it depends on the electrolyte concentration and the dielectric constant of the solvent. Ion correlations are neglected in this mean-field-theory. By comparison to simulations with explicit ions, we found that in aqueous systems containing a cylinder with the line charge density of DNA this approximation is not severe. This is in agreement with the literature. By comparison to atomistic simulations, we found that describing the ions by means of Poisson-Boltzmann theory around a cylinder of around 1 nm radius is absolutely adequate, as long as the far field more the 1.5 nm away from the DNA axis is of interest. This finding also makes us confident that the free energy barriers calculated from coarse-grained simulations are realistic. For distances smaller than around 1.5 nm, effects specific to the DNA type appear important. Our simulations indicate that around one third of the DNA charge is compensated for by ions entering the DNA grooves. This fraction does depend on the DNA type that is considered. The far field, however, remains virtually unaffected by the DNA composition.

The of role dielectric interfaces is twofold. In the context of an interface between water and a medium with a lower dielectric constant, their effect is twofold. On the single ion level, they are repulsive and therefore prevent ions from being close. Furthermore, they increase correlations. This is only visible at conditions under which ion correlations are important at all. In a univalent aqueous electrolyte, the effect is negligible, but in trivalent salt, some effect could be detected. On the Poisson-Boltzmann mean-field-level of description, as correlation effects, they are not contained. In the presence

of charged macroions, e.g. DNA, the interface repels the charged objects. In addition, it increase the electrostatic attraction between counterions and the charged objects. This effect is screened on the Debye length scale. Applying a particle-based description for the dsDNA translocation barriers was by no means better than a PB-based model. The differences between the prediction of both was smaller than the uncertainty e.g. of the geometry from the experiments. Correlation effects were of no importance for the considered question and therefore the PB equation is absolutely adequate. With the iPBS algorithm, dielectric interfaces can be treated satisfactorily, and even the simple one-dimensional continuum model yielded acceptable results.

In case of the flexible model for single stranded DNA, the coarse-grained simulations played their full potential. We were able to model a DNA strand of forty nanometer length in a simulation box, which was even larger. In this case the high configurational degree of freedom is difficult to take into account correctly on the continuum level. For the all-atom-model the computational effort rises to enormous amounts. On the continuum scale the geometric simplicity, which made the modelling with double-stranded DNA so easy, is lost. In other terms, the conformational space of the DNA and the ions is sampled at the same time in the coarse-grained model. However, it is fair to say that our effort in parametrizing the ssDNA was small. The physics of single-stranded DNA is much more complex, especially as it can form complex structures, for example so-called hairpins, by interactions with itself. A more realistic model must take these into account in more detail.

For dynamic effects, the electrokinetic equations contain only one further parameter, the hydrodynamic radius of the ions. The hydrodynamic ion radius can be measured in terms of the diffusion constant of the ions and the solvent viscosity. The physical interpretation of this parameter is as follows: Particles with small hydrodynamic radii can slip though the solvent easily. Their motion is only weakly affected by fluid flow. If the hydrodynamic radius of the particles is large, water flow and particle transport are tightly coupled. Consistent with this statement is the fact that the strength hydrodynamic interactions also only depends on this quantity. The Debye-Hückel-Onsager expression for the electrolyte concentration contains the hydrodynamic radius for the same reason. Any model containing more microscopic structure must match the ratio of this length scale to other length scales in the system to be in asymptotic agreement with the electrokinetic equations.

We could show, that the hybrid Lattice-Boltzmann/Molecular Dynamics simulations, the hydrodynamic radius of the particles is well-defined and can be adjusted to the desired number. Hydrodynamic interactions are reproduced

sufficiently accurately so that the experimental electrolyte conductance can be reproduced almost numerically. In this sense, the LB/MD method is an extension to the electrokinetic equations, that can reproduce their behaviour asymptotically. It does contain hydrodynamic interactions as a correlation effect. Near boundaries, the particle mobility is reduced. From the atomistic simulations of DNA, we could show, that this effect is physically relevant, but further work is necessary to understand its magnitude better. In this sense the LB/MD method is not predictive, and its accuracy for different conditions still must be explored better.

On the all-atom level we proved the existence and relevance of a surface friction effect for ions. We believe that this effect is independent of the applied force field, but caused by a combination of a microscopic surface roughness, electrofriction due to charge patterns on the surface and a water-mediated hydrodynamic frictional force. We are, however, somewhat suspicious that the findings may quantitatively dependent of the applied force field. This is because ion parameters mostly have been adjusted to reproduce bulk electrolyte properties well, but the interaction with other materials, e.g. DNA, has not been considered in the parametrization. The ion parameters used in canonically with the different popular force fields have been used more traditionally than because their behaviour is well tested. To obtain ions with a realistic hydrodynamic radius, we had to rely on a non-standard combination of ion parameters and water model, which was not very satisfactory. We suggest studying these effects further by systematically checking further combinations of force fields, ion parameters and water models.

Furthermore, a very interesting open question is, how reliable a discontinuous dielectric constant can model the behaviour of ions or also dipolar molecules in the vicinity of the interfaces. This is especially important if the membrane material has a higher dielectric constant than the solution. This can happen e.g. near graphene sheets, because in plane graphene is a conductor. By assuming an infinite dielectric constant, ions are strongly attracted to the surface. Even if this is realistic, an open question is where to position the dielectric interface relative to the atoms. In our simulations the distance between the dielectric interface and the plane, where repulsive interactions set is a model parameter, but our findings were not changed much by altering that distance, because the image charge effects were repulsive. In case of attractive dielectric boundary forces potential with comparisons to quantum mechanical calculations would be advisable.

6.3 Future work

In Chapter 4, we investigated free energy profiles of DNA fragments. This investigation aimed at learning about the translocation rate. Our investigation, however, did indeed not address the translocation rate directly. Two aspects are still missing when considering the translocation rate in experiments. First, knowledge about the translocation rate requires knowledge about the pre-exponential factor of the Arrhenius law. This factor, also known as the attempt rate, is proportional to the diffusion constant and the DNA concentration. It also involves a factor related to the exact geometry and definition of the reaction coordinate. Second, it is necessary to take the driving voltage into account.

The charged rod system described in Sec. 4.3.2 would be a good starting point for this investigation. Dielectric boundary forces could be neglected to minimize the computational effort. As a first step it would be interesting to study the translocation rate of a charged rod which is confined to the pore axis but can move along the axis without applied voltage. In this system it is straightforward to apply Kramers theory. By adding an appropriate amount of salt ions, or choosing a wider pore, the free energy barrier can be reduced, so that translocation happens sufficiently often and rates can be measured directly and compared to Kramers theory. No further complications are expected and therefore this system can act as a benchmark system for the understanding of these processes.

For tighter pores and lower salt concentration, the spontaneous translocation rate will drop significantly, and methods of rare event sampling will be necessary to study them. As techniques like Umbrella sampling still rely on Kramer's theory, we suggest to apply the Forward Flux sampling technique [342, 343]. In this method, unbiased molecular dynamics trajectories along a reaction coordinate can be constructed. A series of points along this reaction coordinates serves as milestones, an each of them defines a hypersurface in phase space. Trajectories crossing a hypersurface are forked into many by using different series of random number for thermostatting. This way it can be assured for every hypersurface, that some trajectories will cross the next hypersurface, without applying a biasing force. From a statistical analysis of the trajectory fragments, an overall rate of the process can be obtained. As no complications e.g. from slow internal degrees of freedom are expected, agreement of conventional simulations, Kramers' theory and Forward Flux sampling should be obtained.

For biased system is important to recognize that the driving force is not trivial to determine. For sufficiently small voltages, the electric field is described by a Laplace equation with zero field boundary condition on the pore. The force acting on the DNA, however, has to be obtained by also considering electroosmotic flow. In Ref. [344] this is e.g. done by solving the electrokinetic equations in a very similar geometry as in 4.4.2. The mean force necessary to stall the DNA at various positions along the axis is calculated by solving the electrokinetic equations with the Finite Element Method. The integral results is then used as the effective potential including the driving force for Kramers' theory. This procedure appears sound, but our system would allow us to verify this procedure without additional assumptions. For this the methodological process with the LB method presented in this thesis is helpful. As the Forward Flux Sampling method also allows for nonequilibrium conditions, it would be possible to perform this investigation also for small driving forces and high free energy barriers. We believe it is important to show that this process can not be described by a concept of an effective charge, and that the out-of-equilibrium conditions requires special care.

Moving from a DNA fragment confined to the pore axis to a fragment with full orientational and positional freedom makes the considerations above more difficult. Brute-Force simulations will require significantly longer simulation time for DNA molecules to find the pore entrance. The reaction coordinate used in Sec. 4.3.2 is probably not well suited for this enterprise. It would be important to also take into account the distance in parallel direction, so that the reaction coordinate can distinguish between DNA molecule in the flat region of the membrane and in the pore. It could, for example, be possible to use the geometrical distance from the disc corresponding to the pore opening. An alternative possibility could be inspired by the solution of the electrostatic problem of a pore with electrodes placed far away from the pore openings. Evaluating this value at the position of the center of mass of the DNA fragment also could form a viable alternative. In all of these cases, an entropic contribution to the free energy, the Fixman potential [345], needs to be taken into account. With these modifications, the same machinery as for the constraint case can be used. With all these intermediate steps it should be possible to obtain a full picture of the translocation of stiff DNA fragments. All steps can be performed largely without new implementations, since already a well tested framework for the Forward Flux Sampling Method is available [346, 347]. This would make this project very well suited for a Master thesis.

Furthermore, the information gained on the atomistic level could be transferred back to the coarser scales. The good general agreement of the LB/MD

simulation both with the electrokinetic theory and the atomistic simulations make it plausible that only small modifications could lead to a representation that is suited to study more complex questions, e.g. regarding more complex geometries or nonlinear responses. In the literature, a plethora of model with several "super-atoms" per nucleotide or base pair is available [348–355]. Only few of the were though designed to be used in combination with an explicit ion representation. The first step in this project would be to find a DNA model which reproduces the ion distribution correctly. It could be attempted to represent the different base pairs differently, to reproduce the significant differences in ion distribution we found by the visual analysis. The parameter space of such a problem should not be underestimated. Different numbers of beads per base, with different sizes, and possibly anisotropic shapes and attractive interactions with different strength must be tested. From first attempts, I have learned that approximate results for the ion densities can be obtained from only minutes of simulation time. A GPU implementation of electrostatics could be very helpful to achieve a workflow which is not too tedious. As a second step, the DNA could be made mobile, but it is probably very helpful to have working model for a fixed DNA as a starting point.

A second detail that must be transported to the coarser scale is the position-dependent mobility. To our knowledge, this effect has so far only been taken into account Comer and Aksimentiev [329]. In their atomic resolution Brownian dynamics, ions are treated explicitly and DNA is taken into account including all atomic detail. The water has been incorporated into potentials of Mean Force between the ions and all four nucleotides that have been deduced from extensive simulations. Beyond, the diffusion coefficient of the ions depends on the distance from the nucleotide surface. For the LB/MD hybrid method it could be possible to include an extra friction term in the spirit of the Dissipative Particle Dynamics method could be considered. The pairwise particle friction of DPD could pose a viable way to include the extra friction effect thermodynamically consistently, and including momentum conservation. The dynamic properties of the particles coupled to the LB fluid is only partly well understood. We have shown, that the method can be used as a quantitative tool for simulating electrokinetic systems. Our simulation results indicate, that the hydrodynamic particle radius is unique, and determines indeed the strength of hydrodynamic interactions. Yet, we have shown, that the fluctuation-dissipation theorem for particle mobility and diffusion is only valid in a part of the parameters space. These deviations should be investigated more systematically to understand why this can happen.

Interesting structures that could be studied with a coarse grained model of DNA are DNA Origami nanopores. In these pores, DNA Origami [79] is used

to create nanopores. [80–82]. The key advantage lies in the accuracy of the DNA self-assemblies. First computer simulations of DNA origami have been published, [356], but the size of experimentally used DNA origami nanopores makes the all-atom simulations very demanding. A coarse-grained DNA representation with implicit water and realistic ion dynamics would allow for realistic and computationally tractable models.

With sufficient computational resources, it would be relatively easy to investigate further combinations of force fields, ion parameters and water models in the same simulation setup we used in Sec. 5.3. In principle, all combinations should yield similar ion distributions around the DNA. The workflow is already adapted for the use of different force fields and water models, and only ion parameters must be added. The visual analysis tool developed in this thesis will then be especially helpful, as qualitative features of the ion distributions must be compared.

In most translocation experiments, single-stranded DNA is used. The main motivation, is that, of course, DNA sequencing can only with work if a single DNA strand passes the constriction. Therefore, it appears very rewarding to extend our investigation to single-stranded DNA. At least one important parameter will then appear, which has so far not been relevant, the chain tension. Single-stranded DNA will fluctuate much more, and therefore the interplay of entropic forces within the chain and external forces from the chain ends can significantly affect the chain geometry. The analysis will also turn out to be more difficult, as fixing parts of the DNA will lead to unrealistic results. Therefore, new measures must be found to characterize the ion distribution and position dependent friction effects. Thus, studying ssDNA will require conceptually new ideas and is therefore much more difficult than the study of dsDNA presented here.

6.4 Concluding remarks

Finally, we would like to review our work from a somewhat more philosophical perspective. I would like to comment especially the way we apply the concept of a model. In other fields of physics, typically a single theory that describes all observations is searched for (see e.g. [357]). In our case, we are, however not striving for *the* model, which entirely explains the experimental data. On the contrary, for all we know, the soft matter regime is satisfactorily explained in terms of quantum mechanics. So why are we discussing models? One very plausible point is, of course, the computational feasibility. We have

e.g. shown, that we can estimate free energy barriers reasonably by solving a one-dimensional ordinary differential equation. This is convenient, but not the only reason.

In our case, the problem setting is somewhat different. The physical theory capable of describing all the phenomena considered is well known. Quantum mechanics are known to accurately describe the effects of relevant to soft matter systems. The generality of QM is also its fate: For our purposes We are in no need to find a unique and highly accurate representation, which can "stand for" the material world [358]. We, on the one hand, are in the need models which are hand computationally feasible. Only if available computational resources allow us to evaluate the model and make predictions, a model is useful. On the other hand, we need models, as a tool to analyze and interpret the data we have obtained from simulations. In this sense our understanding is model-based itself because we rely on the using simpler models to understand the prediction of a complex model. In our *model-based understanding*, we deliberately leave out details and realism leads to scientific insight. Yi describes the strategy as follows: Identifying the features of a model, and what a model can do – with and without extra assumptions – helps us to shape our physical intuitions of the model [359]. Precision and generality of models of complex systems are sometimes considered contradictory [360], and the tradeoff between the two we choose, is typically on the side of generality. Interestingly, this approach is a "novelty [...] in the philosophy of science" [361].

When projecting the current developement of computational power into the future, it becomes plausible, that we are only observing an early stage of *in silico* modelling. In a few decades, atomistic simulations with billions of particles might be standard. I believe that then it is even more important to have an appropriate language of formulating scientific insight. Possibly in this future, computer simulation methods will be much easier to perform and much more accessible for amateurs. But at that point, it is very important to possess a language in which expected behaviour can be separated from unexpected behaviour. With expected, I refer to that behaviour, which could have been inferred from simple considerations of very basic scientific laws. Especially then, it is necessary to have a sharp intuition and good coarse-grained representations.

6.5 Acknowledgements

This work was conducted at the Institute for Computational Physics (ICP) at the University of Stuttgart. At the ICP, I experienced a positive, productive and open atmosphere. I enjoyed helpful support, lively discussions and an enormous input. Beneath a pleasant working environment, I made friends and I am sure these friendships will survive the separation over significant distances, and personal development in very different directions. The productive and inspiring surrounding was largely created by my supervisor, Prof. Christian Holm. I share his believe that the best science is done when people have fun and enjoy their work as a team. This meant also, that attending conferences, and working in international teams, was appreciated. Therefore, I could attend not only fascinating talks, but also meet most inspiring people from all over the world. Christian gave me a lot of freedom, and a lot of responsibility which I both highly appreciate.

Most of the simulation work presented here was performed with the software package ESPResSo. I highly appreciate the help and support of the community of ESPResSo developers and users. Being part of the development team is a very pleasant, useful and inspiring experience. Scientific software development is necessarily a compromise between rapid development for personal scientific use, and fundamental improvements to the core. The motivating discussions, and the productive atmosphere in the ESPResSo team will forever influence my relationship to software and its development, and therefore will be a very important point for my professional and scientific future.

The work presented in this thesis was funded by the German Science Foundation (Deutsch Forschungsgemeinschaft, DFG), as project C5 “Macromolecular transport through nanoscale pores” in the Collaborative Research Center (Sonderforschungsbereich, SFB) “Dynamic simulation of systems with large particle numbers”. The generous funding allowed me to take part in numerous excellent, international conferences. I greatly benefited from stimulating discussion with scientist from different fields, sometimes with very different perspectives.

Ultimately I would like to thank my family for their continuous support, their curiosity and their positive attitude towards my life in Stuttgart. My parents offered me an inspiring learning environment in my childhood, where questions were answered, and curiosity was stimulated. I was taught that learning about math, science, music, languages, computers, and many other things is fun, and without enjoying science I had never finished this PhD. Teaming up with my

brother, was great fun, and still is. My family now also includes Dani, whose continuous support over the distance between Stuttgart and Aachen has made the time pleasant and joyful. She supported my work, joined my travels and helped me keep in touch with friends. After this work is finished, a new period of life will begin for us.

Bibliography

[1] L. N. A. Meller and D. Branton. Voltage-driven DNA translocations through a nanopore. *Physical Review Letters*, 86(15):3435, 2001.

[2] O. N. Technologies. Press release, 2012.

[3] C. Tyagi, M. Süzen, M. Sega, M. Barbosa, S. Kantorovich, and C. Holm. An iterative, fast, linear-scaling method for computing induced charges on arbitrary dielectric boundaries. *Journal of Chemical Physics*, 132:1154112, 2010.

[4] J. P. Valleau and L. K. Cohen. Primitive model electrolytes. I. Grand canonical monte carlo computations. *The Journal of Chemical Physics*, 72:5935, 1980.

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

[6] S. Kesselheim, W. Müller, and C. Holm. Origin of current blockades in nanopore translocation experiments. *Physical Review Letters*, 112:018101, 2014.

[7] R. DeBlois and C. Bean. Counting and sizing of submicron particles by the resistive pulse technique. *Review of Scientific Instruments*, 41(7):909–916, 1970.

[8] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, and D. Branton. Rapid nanopore discrimination between single polynucleotide molecules. *PNAS USA*, 97:1079–1084, 2000.

[9] J. Makoff. Cost of gene sequencing falls, raising hopes for medical advances. *The New York Times*, 2012.

[10] W. Sung and P. J. Park. Polymer translocation through a pore in a membrane. *Physical Review Letters*, 77(4):783, 1996.

[11] J. Zhang and B. Shklovskii. Effective charge and free energy of DNA

inside an ion channel. *Phys. Rev. E*, 75(2):21906, 2007.

[12] R. Kumar and M. Muthukumar. Origin of translocation barriers for polyelectrolyte chains. *Journal of Chemical Physics*, 131(19):194903, 2009.

[13] M. Süzen. *Induced charge computation.* PhD thesis, Goethe-Universität, Frankfut (Main), 2009.

[14] A. Schlaich. An iterative Poisson-Boltzmann solver for regions with dielectric mismatch. Diplomarbeit, University of Stuttgart, 2011.

[15] S. van Dorp, U. F. Keyser, N. H. Dekker, C. Dekker, and S. G. Lemay. Origin of the electrophoretic force on DNA in solid-state nanopores. *Nat Phys*, 5(5):347–351, 2009.

[16] S. Kesselheim, M. Sega, and C. Holm. The ICC* algorithm: A fast way to include dielectric boundary effects into molecular dynamics simulations. *arXiv preprint arXiv:1003.1271*, 2010.

[17] S. Kesselheim, M. Sega, and C. Holm. Influence of pore dielectric boundaries on the translocation barrier of DNA. *arXiv preprint arXiv:1002.2759*, 2010.

[18] S. Kesselheim, M. Sega, and C. Holm. Applying ICC* to DNA translocation. effect of dielectric boundaries. *Computer Physics Communications*, 182(1):33–35, 2011.

[19] G. Rempfer. Lattice-Boltzmann simulations in Complex Geometries. Bachelos thesis, University of Stuttgart, 2010.

[20] F. Zeller. Simulation of electroosmotic flow using the lattice boltzmann method. Bachelor thesis, University of Stuttgart, 2011.

[21] W. Müller. Full atomistic simulations of DNA in a nanopore. Diplomarbeit, University of Stuttgart, 2012.

[22] T. A. Waigh. *Applied biophysics: A molecular approach for physical scientists.* Wiley, Chichester, 2007.

[23] P. R. Cook. *Principles of nuclear structure and function.* Wiley-Liss, New York [u.a.], 2001.

[24] E. Sackmann and R. Merkel. *Lehrbuch der Biophysik.* Wiley-VCH, Weinheim, 2010.

[25] E. R. Mardis. A decade/'s perspective on DNA sequencing technology. *Nature*, 470(7333):198–203, 2011.

[26] K. Wetterstrand. DNA sequencing costs: Data from the NHGRI Genome Sequencing Program (GSP). `www.genome.gov/sequencingcosts`, accessed April 29 2013.

[27] J. D. Watson, F. H. Crick, et al. Molecular structure of nucleic acids. *Nature*, 171(4356):737–738, 1953.

[28] F. H. Crick and J. D. Watson. The complementary structure of deoxyribonucleic acid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 223(1152):80–96, 1954.

[29] H. F. Lodish. *Molecular cell biology*. Freeman, New York, NY, 2013.

[30] R. Cotterill. *Biophysik: eine Einführung*. Wiley-VCH, Weinheim, 2008.

[31] W. Humphrey, A. Dalke, and K. Schulten. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.

[32] R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239(4839):487–491, 1988.

[33] D. S. Nicholl. *Gentechnische Methoden*. Spektrum, Akad. Verl., Heidelberg, 2002.

[34] T. Reinard. *Molekularbiologische Methoden*. Ulmer, Stuttgart, 2010.

[35] J. W. Dale, M. v. Schantz, and N. Plant. *From genes to genomes: Concepts and applications of DNA technology*. Wiley-Blackwell, Chichester [u.a.], 2012.

[36] F. Sanger, G. Air, B. Barrell, N. Brown, A. Coulson, C. Fiddes, C. Hutchison, P. Slocombe, and M. Smith. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*, 265(5596):687, 1977.

[37] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12):5463–5467, 1977.

[38] A. J. F. Griffiths. *Introduction to genetic analysis*. Freeman, New York, NY, 2012.

[39] D. P. Clark and N. J. Pazdernik. *Molecular biology*. Elsevier/Academic Press, Amsterdam, 2013.

[40] A. Genomics. X-prize competition guidelines. `http://genomics.xprize.org/sites/genomics.xprize.org/files/docs/AGXP_Competition_Guidelines.pdf`, accessed April 29 2013, 2012.

[41] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences*, 93(24):13770–13773, 1996.

[42] The Nanopore Site. http://www.thenanoporesite.com/research-groups.html (accessed 29. Jan 2014).

[43] S. Howorka and Z. Siwy. Nanopore analytics: Sensing of single molecules. *Chemical Society Reviews*, 38(8):2360–2384, 2009.

[44] T. Z. Butler, M. Pavlenok, I. M. Derrington, M. Niederweis, and J. H. Gundlach. Single-molecule DNA detection with an engineered MspA protein nanopore. *Proceedings of the National Academy of Sciences*, 105(52):20647–20652, 2008.

[45] I. M. Derrington, T. Z. Butler, M. D. Collins, E. Manrao, M. Pavlenok, M. Niederweis, and J. H. Gundlach. Nanopore DNA sequencing with MspA. *Proceedings of the National Academy of Sciences*, 107(37):16060–16065, 2010.

[46] S. Howorka, S. Cheley, and H. Bayley. Sequence-specific detection of individual DNA strands using engineered nanopores. *Nature Biotechnology*, 19:636, 2001.

[47] O. Braha, B. Walker, S. Cheley, J. J. Kasianowicz, L. Song, J. E. Gouaux, and H. Bayley. Designed protein pores as components for biosensors. *Chemistry & biology*, 4(7):497–505, 1997.

[48] X. Guan, L.-Q. Gu, S. Cheley, O. Braha, and H. Bayley. Stochastic sensing of TNT with a genetically engineered pore. *ChemBioChem*, 6(10):1875–1881, 2005.

[49] B. Venkatesan and R. Bashir. Nanopore sensors for nucleic acid analysis. *Nature Nanotechnology*, 6(10):615–624, 2011.

[50] C. Dekker. Solid-state nanopores. *Nature Nanotech*, 2:209–215, 2007.

[51] U. Keyser, J. van der Does, C. Dekker, and N. Dekker. Optical tweez-

ers for force measurements on DNA in nanopores. *Review of Scientific Instruments*, 77:105105, 2006.

[52] J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. Golovchenko. Ion-beam sculpting at nanometre length scales. *Nature*, 412(6843):166–169, 2001.

[53] A. Storm, J. Chen, X. Ling, H. Zandbergen, and C. Dekker. Fabrication of solid-state nanopores with single-nanometre precision. *Nature Materials*, 2(8):537–540, 2003.

[54] S. E. Henrickson, M. Misakian, B. Robertson, and J. J. Kasianowicz. Driven DNA transport into an asymmetric nanometer-scale pore. *Physical Review Letters*, 85(14):3057, 2000.

[55] B. N. Miles, A. P. Ivanov, K. A. Wilson, F. Doğan, D. Japrung, and J. B. Edel. Single molecule sensing with solid-state nanopores: novel materials, methods, and applications. *Chemical Society Reviews*, 42(1):15–28, 2013.

[56] P. Chen, T. Mitsui, D. B. Farmer, J. Golovchenko, R. G. Gordon, and D. Branton. Atomic layer deposition to fine-tune the surface properties and diameters of fabricated nanopores. *Nano Letters*, 4(7):1333–1337, 2004.

[57] B. M. Venkatesan, A. B. Shah, J.-M. Zuo, and R. Bashir. Dna sensing using nanocrystalline surface-enhanced al2o3 nanopore sensors. *Advanced Functional Materials*, 20(8):1266–1275, 2010.

[58] R. Wei, D. Pedone, A. Zürner, M. Döblinger, and U. Rant. Fabrication of metallized nanopores in silicon nitride membranes for single-molecule sensing. *Small*, 6(13):1406–1414, 2010.

[59] M. Wanunu and A. Meller. Chemically modified solid-state nanopores. *Nano Letters*, 7(6):1580–1585, 2007.

[60] C. Vericat, M. Vela, G. Benitez, P. Carro, and R. Salvarezza. Self-assembled monolayers of thiols and dithiols on gold: new challenges for a well-known system. *Chemical Society Reviews*, 39(5):1805–1834, 2010.

[61] S. W. Kowalczyk, L. Kapinos, T. R. Blosser, T. Magalhaes, P. van Nies, R. Y. H. Lim, and C. Dekker. Single-molecule transport across an individual biomimetic nuclear pore complex. *Nature Nanotechnology*, 6:433, 2011.

[62] J. Zhou and F. Schmid. AC-field-induced polarization for uncharged colloids in salt solution: A dissipative particle dynamics simulation. *The European Physical Journal E*, 36(4):33, 2013.

[63] S. Garaj, W. Hubbard, A. Reina, J. Kong, D. Branton, and J. Golovchenko. Graphene as a subnanometre trans-electrode membrane. *Nature*, 467(7312):190–193, 2010.

[64] C. A. Merchant, K. Healy, M. Wanunu, V. Ray, N. Peterman, J. Bartel, M. D. Fischbein, K. Venta, Z. Luo, A. T. C. Johnson, and M. Drndiċ. DNA translocation through graphene nanopores. *Nano Letters*, 10(8):2915–2921, 2010.

[65] S. K. Min, W. Y. Kim, Y. Cho, and K. S. Kim. Fast DNA sequencing with a graphene-based nanochannel device. *Nature Nanotechnology*, 6(3):162–165, 2011.

[66] E. Sadki, S. Garaj, D. Vlassarev, J. A. Golovchenko, and D. Branton. Embedding a carbon nanotube across the diameter of a solid state nanopore. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, 29(5):053001–053001, 2011.

[67] M. Tsutsui, M. Taniguchi, K. Yokota, and T. Kawai. Identifying single nucleotides by tunnelling current. *Nature nanotechnology*, 5(4):286–290, 2010.

[68] A. Ivanov, E. Instuli, C. McGilvery, G. Baldwin, D. McComb, T. Albrecht, and J. Edel. DNA tunneling detector embedded in a nanopore. *Nano Letters*, 11(1):279, 2011.

[69] L. J. Steinbock, O. Otto, D. R. Skarstam, S. Jahn, C. Chimerel, J. L. Gornall, and U. F. Keyser. Probing DNA with micro- and nanocapillaries and optical tweezers. *Journal of Physics: Condensed Matter*, 22(45):454113, 2010.

[70] L. J. Steinbock, J. F. Steinbock, and A. Radenovic. Controllable shrinking and shaping of glass nanocapillaries under electron irradiation. *Nano Letters*, 13(4):1717–1723, 2013.

[71] V. V. Thacker, S. M. Hernandez-Ainsa, J. Gornall, L. J. Steinbock, and U. F. Keyser. Combining fluorescence imaging and ionic current detection in nanocapillaries. In *Nanopores for Bioanalytical Applications: Proceedings of the International Conference*, pages 43–45. The Royal Society of Chemistry, 2012.

[72] M. R. Powell, L. Cleary, M. Davenport, K. J. Shea, and Z. S. Siwy. Electric-field-induced wetting and dewetting in single hydrophobic nanopores. *Nature Nanotechnology*, 6(12):798–802, 2011.

[73] Z. Siwy and A. Fuliński. A nanodevice for rectification and pumping ions. *American Journal of Physics*, 72:567, 2004.

[74] I. Vlassiouk and Z. S. Siwy. Nanofluidic diode. *Nano Letters*, 7(3):552–556, 2007.

[75] T. Jain, R. J. S. Guerrero, C. A. Aguilar, and R. Karnik. Integration of solid-state nanopores in microfluidic networks via transfer printing of suspended membranes. *Analytical Chemistry*, 85(8):3871–3878, 2013.

[76] Z. Fekete, G. Huszka, A. Pongrácz, G. Jágerszki, R. Gyurcsányi, E. Vrouwe, and P. Fürjes. Integrated microfluidic environment for solid-state nanopore sensors. *Procedia Engineering*, 47:13–16, 2012.

[77] S. Choi, M. Goryll, L. Sin, P. Wong, and J. Chae. Microfluidic-based biosensors toward point-of-care detection of nucleic acids and proteins. *Microfluidics and Nanofluidics*, 10(2):231–247, 2011.

[78] N. Laohakunakorn, O. Otto, S. Sturm, K. Kroy, and U. F. Keyser. Dynamic single-molecule force spectroscopy using optical tweezers and nanopores. In *SPIE NanoScience+ Engineering*, page 88101F. International Society for Optics and Photonics, 2013.

[79] P. Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, 2006.

[80] R. Wei, T. G. Martin, U. Rant, and H. Dietz. DNA origami gatekeepers for solid-state nanopores. *Angewandte Chemie International Edition*, 51(20):4864–4867, 2012.

[81] N. A. W. Bell, C. R. Engst, M. Ablay, G. Divitini, C. Ducati, T. Liedl, and U. F. Keyser. DNA Origami nanopores. *Nano Letters*, 12(1):512–517, 2012.

[82] S. Hernàndez-Ainsa, N. A. W. Bell, V. V. Thacker, K. Göpfrich, K. Misiunas, M. E. Fuentes-Perez, F. Moreno-Herrero, and U. F. Keyser. DNA origami nanopores for controlling DNA translocation. *ACS Nano*, 7(7):6024–6030, 2013.

[83] M. Daoud and C. Williams. *Soft Matter Physics*. Springer, 1999.

[84] P.-G. de Gennes. Soft matter. Nobel Lecture, 1991.

[85] I. V. Hamley. *Introduction to Soft Matter*. Wiley, 2003.

[86] C. Peter and K. Kremer. Multiscale simulation of soft matter systems. *Faraday Discussions*, 144:9–24, 2010.

[87] Announcement of the 2013 Nobel Prize in chemistry. http://www.nobelprize.org/mediaplayer/index.php?id=1956.

[88] A. Warshel and M. Karplus. Calculation of ground and excited state potential surfaces of conjugated molecules. i. formulation and parametrization. *Journal of the American Chemical Society*, 94(16):5612–5625, 1972.

[89] A. Warshel and M. Levitt. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, 103(2):227 – 249, 1976.

[90] M. Levitt and A. Warshel. Computer simulation of protein folding. *Nature*, 253(5494):694–698, 1975.

[91] H. M. Senn and W. Thiel. QM/MM methods for biomolecular systems. *Angewandte Chemie International Edition*, 48(7):1198–1229, 2009.

[92] K. Meier, A. Choutko, J. Dolenc, A. P. Eichenberger, S. Riniker, and W. F. van Gunsteren. Biomolekulare simulationen mit mehreren auflösungsniveaus: ein überblick über methodische aspekte. *Angewandte Chemie*, 125(10):2888–2904, 2013.

[93] Nobelprize.org. Announcement of the 2013 Nobel Prize in Chemistry. http://www.nobelprize.org/mediaplayer/index.php?id=1956.

[94] D. Fritz, K. Koschke, V. A. Harmandaris, N. F. van der Vegt, and K. Kremer. Multiscale modeling of soft matter: scaling of dynamics. *Physical Chemistry Chemical Physics*, 13(22):10412–10420, 2011.

[95] C. Peter and K. Kremer. Multiscale simulation of soft matter systems - from the atomistic to the coarse-grained level and back. *Soft Matter*, 5(22):4357–4366, 2009.

[96] D. Reith, M. Pütz, and F. Müller-Plathe. Deriving effective mesoscale potentials from atomistic simulations. *Journal of Computational Chemistry*, 24(13):1624–1636, 2003.

[97] C. Junghans. *Between the scales: water from different perspectives*. PhD thesis, Johannes Gutenberg-Universität Mainz, 2010.

[98] H. Wang, C. Junghans, and K. Kremer. Comparative atomistic and coarse-grained study of water: What do we loose by coarse-graining? *European Physical Journal E: Soft Matter*, 28(2):221, 2009.

[99] S. O. Nielsen, P. B. Moore, and B. Ensing. Adaptive multiscale molecular dynamics of macromolecular fluids. *Physical Review Letters*, 105:237802, 2010.

[100] M. Praprotnik, L. D. Site, and K. Kremer. Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly. *Journal of Chemical Physics*, 123(22):224106, 2005.

[101] D. Frenkel. Colloidal systems: Playing tricks with designer "atoms". *Science*, 296:65, 2002.

[102] L. Landau and E. Lifshitz. Lehrbuch der theoretischen Physik, Band V: Statistische physik. *Akademie-Verlag, Berlin*, 1987.

[103] R. Balescu. Equilibrium and nonequilibrium statistical mechanics. *NASA STI/Recon Technical Report A*, 76:32809, 1975.

[104] L. E. Reichl. *A Modern Course in Statistical Physics*. Wiley-VCH, 2009.

[105] D. J. Evans and G. Morriss. *Statistical mechanics of nonequilibrium liquids*. Cambridge University Press, 1990.

[106] H. Risken. *The Fokker-Planck Equation*. Springer, Berlin, 1989.

[107] R. L. Liboff. *Kinetic theory: classical, quantum and relativistic descriptions*. Springer, New York, 2003.

[108] G. Röpke. *Statistische Mechanik für das Nichtgleichgewicht*. Deutscher Verlag der Wissenschaften, Berlin, 1987.

[109] R. Kubo, M. Toda, and H. N. *Statistical Physics II*. Springer, 1985.

[110] R. Zwanzig. *Nonequilibrium statistical mechanics*. Oxford University Press, 2001.

[111] J. P. Hansen and I. R. McDonald. *Theory of Simple Liquids*. Academic Press, London, 1986.

[112] P. W. Atkins. *Atkins Physical Chemistry, 6th edition*. Oxford University Publishing, 1998.

[113] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications. Clarendon Press, Oxford, 1987.

[114] R. J. Hunter and L. R. White. *Foundations of Colloid Science*. Oxford science publications. Clarendon Press, 1989.

[115] D. F. Evans and H. Wennerström. *The Colloidal Domain*. Wiley-VCH, New York, 1999.

[116] W. B. Russel, D. A. Saville, and W. R. Schowalter. *Colloidal Dispersions*. Cambridge University Press, Cambridge, UK, 1989.

[117] T. Speck and U. Seifert. Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state. *Europhysics Letters*, 74(3):391, 2006.

[118] R. Henderson. A uniqueness theorem for fluid pair correlation functions. *Physics Letters A*, 49(3):197–198, 1974.

[119] R. Zwanzig. Memory effects in irreversible thermodynamics. *Physical Review*, 124(4):983–992, 1961.

[120] H. Mori. Transport, collective motion, and Brownian motion. *Progress of Theoretical Physics*, 33(3):423–455, 1965.

[121] R. Brown. A brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine, or Annals of Chemistry, Mathematics, Astronomy, Natural History and General Science*, 4(21):161–173, 1828.

[122] A. Einstein. Über die von der molekularkinetischen Theorie der wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[123] J. Renn. Einstein's invention of Brownian motion. *Annalen der Physik*, 14(S1):23–37, 2005.

[124] P. Langevin. Sur la théorie du mouvement brownien. *CR Acad. Sci. Paris*, 146(530-533), 1908.

[125] E. Lauga, M. P. Brenner, and H. A. Stone. Microfluidics: the no-slip boundary condition. *arXiv preprint cond-mat/0501557*, 2005.

[126] C. Kunert and J. Harting. Roughness induced boundary slip in microchannel flows. *Physical Review Letters*, 99:176001, 2007.

[127] C. Kunert and J. Harting. Simulation of fluid flow in hydrophobic rough microchannels. In *International Journal of Computational Fluid Dy-*

namics, volume 22, page 475, 2008.

[128] J. Rotne and S. Prager. Variational treatment of hydrodynamic interaction in polymers. *The Journal of Chemical Physics*, 50:4831, 1969.

[129] J. F. Brady and G. Bossis. Stokesian dynamics. *Annual Review of Fluid Mechanics*, 20:111–157, 1988.

[130] D. L. Ermak and J. McCammon. Brownian dynamics with hydrodynamic interactions. *The Journal of Chemical Physics*, 69:1352, 1978.

[131] G. Batchelor. The stress system in a suspension of force-free particles. *J. Fluid Mech*, 41(3):545–570, 1970.

[132] G. K. Batchelor. Sedimentation in a dilute dispersion of spheres. *Journal of Fluid Mechanics*, 52:245–268, 1972.

[133] G. Batchelor. Brownian diffusion of particles with hydrodynamic interaction. *Journal of Fluid Mechanics*, 74(01):1–29, 1976.

[134] B. H. Zimm. Dynamics of polymer molecules in dilute solution: Viscoelasticity, flow birefringence and dielectric loss. *Journal of Chemical Physics*, 24(2):269–278, 1956.

[135] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after kramers. *Reviews of Modern Physics*, 62:251, 1990.

[136] J. Hoff and L. Dunsch. *Etudes de dynamique chimique, 1884*. Ostwalds Klassiker der exakten Wissenschaften. Akademische Verlagsgesellschaft, 1985.

[137] S. Arrhenius. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Zeitschrift für physikalische Chemie*, 4:226–248, 1889.

[138] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[139] M. R. Wright. *An Introduction to Chemical Kinetics*. Wiley, Hoboken, NJ, 2005.

[140] P. G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, NY, 1979.

[141] A. Y. Grosberg and A. R. Khokhlov. *Statistical Physics of Macromolecules*. AIP Press, 1994.

[142] M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Oxford Science Publications, 1986.

[143] M. Doi. *Introduction to Polymer Physics*. Clarendon Press, Oxford, 1996.

[144] G. Strobl. *The Physics of Polymers*. Sprigner, 2007.

[145] G. S. Grest and K. Kremer. Molecular dynamics simulation for polymers in the presence of a heat bath. *Phys. Rev. A*, 33(5):3628–31, 1986.

[146] M. Muthukumar. Polymer translocation through a hole. *Journal of Chemical Physics*, 111(22):10371, 1999.

[147] J. Chuang, Y. Kantor, and M. Kardar. Anomalous dynamics of translocation. *Phys. Rev. E*, 65:011802, 2001.

[148] M. Gauthier and G. Slater. Nondriven polymer translocation through a nanopore: Computational evidence that the escape and relaxation processes are coupled. *Phys. Rev. E*, 79(2):21802, 2009.

[149] H. de Haan and G. Slater. Memory effects during the unbiased translocation of a polymer through a nanopore. *The Journal of Chemical Physics*, 136:154903, 2012.

[150] T. Sakaue. Sucking genes into pores: Insight into driven translocation. *Phys. Rev. E*, 81:041808, 2010.

[151] P. Rowghanian and A. Grosberg. Force-driven polymer translocation through a nanopore: An old problem revisited. *The Journal of Physical Chemistry B*, 115(48):14127–14135, 2011.

[152] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087, 1953.

[153] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, San Diego, 2002.

[154] Bronstein and Semendjajev. *Taschenbuch der Mathematik*. Teubner, 1979.

[155] U. D. Schiller. An overview of integration schemes for molecular dynamics simulations. Technical report, MPI Mainz, 2008.

[156] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola,

and J. R. Haak. Molecular dynamics with coupling to a heat bath. *Journal of Chemical Physics*, 81:3684–3690, 1984.

[157] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *Journal of Chemical Physics*, 126(1):014101, 2007.

[158] J. K. Hansen and K. T. Forest. Type IV pilin structures: Insights on shared architecture, fiber assembly, receptor binding and type II secretion. *Journal of Molecular Microbiology and Biotechnology*, 11:192–207, 2006.

[159] D. Wolf, P. Keblinski, S. R. Phillpot, and J. Eggebrecht. Exact method for the simulation of coulombic systems by spherically truncated, pairwise r^{-1} summation. *Journal of Chemical Physics*, 110:29, 1999.

[160] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.

[161] M. Deserno and C. Holm. How to mesh up Ewald sums. I. A theoretical and numerical comparison of various particle mesh routines. *Journal of Chemical Physics*, 109:7678, 1998.

[162] M. Deserno and C. Holm. How to mesh up Ewald sums. II. An accurate error estimate for the Particle-Particle-Particle-Mesh algorithm. *Journal of Chemical Physics*, 109:7694, 1998.

[163] H. G. Petersen. Accuracy and efficiency of the particle mesh ewald method. *Journal of Chemical Physics*, 103(9):3668–3679, 1995.

[164] J. W. Perram, H. G. Petersen, and S. W. de Leeuw. An algorithm for the simulation of condensed matter which grows as the 3/2 power of the number of particles. *J. Mol. Phys.*, 65(4):875–893, 1988.

[165] N. Karasawa and W. A. Goddard III. Acceleration of convergence for lattice sums. *The Journal of Physical Chemistry*, 93(21):7320–7327, 1989.

[166] FFTW. Fastest fourier transform in the west – homepage, 2003.

[167] A. Arnold. Berechnung der elektrostatischen Wechselwirkung in 2d + h periodischen Systemen. Diploma thesis, Johannes Gutenberg–Universität, 2001.

[168] A. Arnold. *Computer Simulations of Charged Systems in Partially Periodic Geometries*. PhD thesis, Johannes Gutenberg-Universität, Mainz, Germany, 2004.

[169] A. Arnold, J. de Joannis, and C. Holm. Electrostatics in periodic slab geometries I. *Journal of Chemical Physics*, 117:2496–2502, 2002.

[170] M. Fixman. Construction of Langevin forces in the simulation of hydrodynamic interaction. *Macromolecules*, 19(4):1204–1207, 1986.

[171] T. Geyer and U. Winter. An $\mathcal{O}\left(n^2\right)$ approximation for hydrodynamic interactions in brownian dynamics simulations. *The Journal of Chemical Physics*, 130:114905, 2009.

[172] A. Sierou and J. F. Brady. Accelerated stokesian dynamics simulations. *Journal of Fluid Mechanics*, 448(1):115–146, 2001.

[173] P. J. Hoogerbrugge and J. M. V. A. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhysics Letters*, 19(3):155–160, 1992.

[174] P. Español and P. Warren. Statistical mechanics of Dissipative Particle Dynamics. *Europhysics Letters*, 30:191, 1995.

[175] C. Junghans, M. Praprotnik, and K. Kremer. Transport properties controlled by a thermostat: An extended dissipative particle dynamics thermostat. *Soft Matter*, 4(1):156–161, 2008.

[176] P. Español. Fluid particle dynamics: A synthesis of dissipative particle dynamics and smoothed particle dynamics. *Europhysics Letters*, 39(6):605, 1997.

[177] A. Arnold. personal communication, 2013.

[178] A. Arnold, O. Lenz, S. Kesselheim, R. Weeber, F. Fahrenberger, D. Röhm, P. Košovan, and C. Holm. ESPResSo 3.1 — molecular dynamics software for coarse-grained models. In M. Griebel and M. A. Schweitzer, editors, *Meshfree Methods for Partial Differential Equations VI*, volume 89 of *Lecture Notes in Computational Science and Engineering*, pages 1–23. Springer, 2013.

[179] P. Debye and E. Hückel. Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen. *Physikalische Zeitschrift*, 24(9):185, 1923.

[180] M. Deserno and C. Holm. Theory and simulations of rigid polyelectrolytes. *Molecular Physics*, 100(18):2941–2956, 2002.

[181] M. Krishnan. Electrostatic free energy for a confined nanoscale object in a fluid. *The Journal of Chemical Physics*, 138(11):–, 2013.

[182] M. Krishnan. Erratum: "electrostatic free energy for a confined nanoscale object in a fluid" [j. chem. phys.138, 114906 (2013)]. *The Journal of Chemical Physics*, 139(6), 2013.

[183] S. Dukhin, R. Zimmermann, and C. Werner. A concept for the generalization of the standard electrokinetic model. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 195(1):103–112, 2001.

[184] J. J. Lopez-Garcia, C. Grosse, and J. Horno. A new generalization of the standard electrokinetic model. *The Journal of Physical Chemistry B*, 111(30):8985–8992, 2007.

[185] M. Z. Bazant. Clean data with dirty surfaces in electrokinetics. *Physics*, 3:18, 2010.

[186] B. Eisenberg. Ions in fluctuating channels: transistors alive. *Fluctuation and Noise Letters*, 11(01), 2012.

[187] E. Hinch. *Perturbation Methods*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 1991.

[188] M. Z. Bazant, K. Thornton, and A. Ajdari. Diffuse-charge dynamics in electrochemical systems. *Phys. Rev. E*, 70:021506, 2004.

[189] M. Z. Bazant and T. M. Squires. Induced-charge electrokinetic phenomena. *Current Opinion in Colloid & Interface Science*, 15(3):203–213, 2010.

[190] D. Boda, D. Gillespie, B. Eisenberg, W. Nonner, and D. Henderson. Ionic soft matter: Modern trends in theory and applications. In D. Henderson, editor, *Ionic Soft Matter: Modern Trends in Theory and Applications*, pages 19–43. Springer, 2005.

[191] D. Boda, M. Valisko, B. Eisenberg, W. Nonner, and D. Gillespie. Computing induced charges in inhomogeneous dielectric media: Application in a Monte Carlo simulation of complex ionic system. *Journal of Chemical Physics*, 125(3):034901, 2006.

[192] S. Tyagi, A. Arnold, and C. Holm. ICMMM2D: An accurate method to include planar dielectric interfaces via image charge summation. *Journal of Chemical Physics*, 127:154723, 2007.

[193] S. Kesselheim, M. Sega, and C. Holm. Effects of dielectric mismatch and chain flexibility on the translocation barriers of charged macromolecules through solid state nanopores. *Soft Matter*, 8(36):9480–9486, 2012.

[194] S. Kesselheim and C. Holm. Modeling dna in nanopores. In R. Podgornik, D. Dean, A. Naji, and J. Dobnikar, editors, *Electrostatics of Soft and Disordered Media*. Pan Stanford, 2013.

[195] M. Born. Volumes and heats of hydration of ions. *Zeitschrift für Physik*, 1:45–48, 1920.

[196] J. D. Jackson. *Classical Electrodynamics*. Wiley, New York, 3rd edition, 1999.

[197] L. Landau and E. Lifshitz. Lehrbuch der theoretischen Physik, Band II: Klassische Feldtheorie, 1971.

[198] W. Demtröder. *Experimentalphysik 2: Elektrizität und Optik*. Springer, 2004.

[199] L. Landau and E. Lifshitz. Lehrbuch der theoretischen Physik, Band VIII: Elektrodynamik der Kontinua, 1977.

[200] R. B. B. Joseph O. Hirschfelder, Charles F. Curtiss. *The Molecular Theory of Gases and Liquids*. John Wiley & Sons, 1954.

[201] S. Kondrat and A. Kornyshev. Superionic state in double-layer capacitors with nanoporous electrodes. *Journal of Physics: Condensed Matter*, 23(2):022201, 2011.

[202] Z. Liu, T. Chen, A. T. Bell, and B. Smit. Improved united-atom force field for 1-alkyl-3-methylimidazolium chloride. *Journal of Physical Chemistry B*, 114(32):10692–10692, 2010.

[203] C. Merlet, M. Salanne, and B. Rotenberg. New coarse-grained models of imidazolium ionic liquids for bulk and interfacial molecular simulations. *The Journal of Physical Chemistry C*, 116(14):7687–7693, 2012.

[204] G. Feng, J. Zhang, and R. Qiao. Microstructure and capacitance of the electrical double layers at the interface of ionic liquids and planar electrodes. *The Journal of Physical Chemistry C*, 113(11):4549–4559, 2009.

[205] T. A. Halgren and W. Damm. Polarizable force fields. *Current Opinion in Structural Biology*, 11(2):236–242, 2001.

[206] O. Borodin. Polarizable force field development and molecular dynamics simulations of ionic liquids. *Journal of Physical Chemistry B*, 113(33):11463–11478, 2009.

[207] C. Kittel. *Einführung in die Festkörperphysik*. Oldenbourg, 2002.

[208] B. Roux, H. A. Yu, and M. Karplus. Molecular basis for the born model of ion solvation. *Journal of Physical Chemistry*, 94(11):4683–4688, 1990.

[209] D. Beglov and B. Roux. Solvation of complex molecules in a polar liquid: An integral equation theory. *The Journal of Chemical Physics*, 104(21):8678–8689, 1996.

[210] Z. Xu, Y. Liang, and X. Xing. Mellin transform and image charge method for dielectric sphere in an electrolyte. *arXiv preprint arXiv:1210.2180*, 2012.

[211] Z. Xu. A fast algorithm for treating dielectric discontinuities in charged spherical colloids. *Interdisciplinary Sciences: Computational Life Sciences*, 4(1):19–26, 2012.

[212] G. Dassios and J. C.-E. Sten. On the neumann function and the method of images in spherical and ellipsoidal geometry. *Mathematical Methods in the Applied Sciences*, 35(4):482–496, 2012.

[213] A. Arnold and C. Holm. MMM2D: A fast and accurate summation method for electrostatic interactions in 2D slab geometries. *Computer Physics Communications*, 148(3):327–348, 2002.

[214] A. Arnold and C. Holm. A novel method for calculating electrostatic interactions in 2D periodic slab geometries. *Chemical Physics Letters*, 354:324–330, 2002.

[215] K.-L. Yang, S. Yiacoumi, and C. Tsouris. Electrical double-layer formation. In *Dekker Encyclopedia of Nanoscience and Nanotechnology*, chapter 93, pages 1001–1014. CRC Press, 2007.

[216] D. Card and J. Valleau. Monte carlo study of the thermodynamics of electrolyte solutions. *The Journal of Chemical Physics*, 52:6232, 1970.

[217] S. Tyagi, A. Arnold, and C. Holm. Electrostatic layer correction with image charges: A linear scaling method to treat slab 2D + h systems with dielectric interfaces. *Journal of Chemical Physics*, 129(20):204102, 2008.

[218] M. Muthukumar. *Polymer Translocation*. Taylor and Francis, 2009.

[219] A. Parsegian. Energy of an ion crossing a low dielectric membrane: Solutions to four relevant electrostatic problems. *Nature*, 221:844, 1969.

[220] D. Levitt. Electrostatic calculations for an ion channel. I. Energy and potential profiles and interactions between ions. *Biophysical Journal*, 22(2):209–219, 1978.

[221] D. Levitt. Electrostatic calculations for an ion channel. II. kinetic behavior of the gramicidin A channel. *Biophysical Journal*, 22(2):221–248, 1978.

[222] Y. Rabin and M. Tanaka. DNA in nanopores: Counterion condensation and coion depletion. *Physical Review Letters*, 94:148103, 2005.

[223] L. Guo and E. Luijten. Structure of ions and water around a polyelectrolyte in a polarizable nanopore. *Intl. J. Mod. Phys. C*, 20:1485–1492, 2009.

[224] L. Brun, M. Pastoriza-Gallego, G. Oukhaled, J. Mathé, L. Bacri, L. Auvray, and J. Pelta. Dynamics of polyelectrolyte transport through a protein channel as a function of applied voltage. *Physical Review Letters*, 100:158302, 2008.

[225] D. J. Bonthuis, J. Zhang, B. Hornblower, J. Mathé, B. I. Shklovskii, and A. Meller. Self-energy-limited ion transport in subnanometer channels. *Physical Review Letters*, 97:128104, 2006.

[226] G. Oukhaled, L. Bacri, J. Mathé, J. Pelta, and L. Auvray. Effect of screening on the transport of polyelectrolytes through nanopores. *Europhysics Letters*, 82(4):48003, 2008.

[227] Z. Kam, N. Borochov, and H. Eisenberg. Dependence of laser light scattering of DNA on NaCl concentration. *Biopolymers*, 20(12):2671–2690, 2004.

[228] C. S. Murthy, R. J. Bacquet, and P. J. Rossky. Ionic distributions near polyelectrolytes. a comparison of theoretical approaches. *The Journal of Physical Chemistry*, 89(4):701–710, 1985.

[229] A. P. Lyubartsev and L. Nordenskiöld. Monte carlo simulation study of ion distribution and osmotic pressure in hexagonally oriented dna. *The Journal of Physical Chemistry*, 99(25):10373–10382, 1995.

[230] K. Grass, U. Böhme, U. Scheler, H. Cottet, and C. Holm. Importance of hydrodynamic shielding for the dynamic behavior of short polyelectrolyte chains. *Physical Review Letters*, 100:096104, 2008.

[231] G. W. Slater, C. Holm, M. V. Chubynsky, H. W. de Haan, A. Dubé,

K. Grass, O. A. Hickey, C. Kingsburry, D. Sean, T. N. Shendruk, and L. Zhan. Modeling the separation of macromolecules: A review of current computer simulation methods. *Electrophoresis*, 30(5):792–818, 2009.

[232] G. Manning. Limiting laws and counterion condensation in polyelectrolyte solutions I. colligative properties. *Journal of Chemical Physics*, 51:924–933, 1969.

[233] R. Courant. Variational methods for the solution of problems of equilibrium and vibrations. *Bull. Amer. Math. Soc*, 49(1):23, 1943.

[234] G. Pelosi. The finite-element method, Part I: R. L. Courant [historical corner]. *Antennas and Propagation Magazine, IEEE*, 49(2):180–182, 2007.

[235] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, S. Kuttanikkad, M. Ohlberger, and O. Sander. The distributed and unified numerics environment (DUNE). In *19th Symposium on Simulatin Technique in Hannover, September 12-14*, 2006.

[236] D. Project. Dune-pdelab howto. Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, 2011.

[237] P. G. Ciarlet. *The finite element method for elliptic problems*. Elsevier, 1978.

[238] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. part I: Abstract framework. *Computing*, 82(2–3):103–119, 2008.

[239] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. part II: Implementation and tests in DUNE. *Computing*, 82(2–3):121–138, 2008.

[240] P. Bastian, K. Birken, K. Johannsen, S. Lang, N. Neuss, H. Rentz-Reichert, and C. Wieners. UG – A flexible software toolbox for solving partial differential equations. *Computing and Visualization in Science*, 1:27–40, 1997.

[241] A. Dedner, C. Rohde, B. Schupp, and M. Wesenberg. A parallel, load-balanced MHD code on locally-adapted, unstructured grids in 3d. *Computing and Visualization in Science*, 7:79–96, 2004.

[242] A. Burri, A. Dedner, R. Klöfkorn, and M. Ohlberger. An efficient implementation of an adaptive and parallel grid in dune. In *Computational Science and High Performance Computing II*, pages 67–82. Springer, 2006.

[243] M. Blatt and P. Bastian. The iterative solver template library. In B. Kagstroem, E. Elmroth, J. Dongarra, and J. Wasniewski, editors, *Applied Parallel Computing. State of the Art in Scientific Computing*, volume 4699 of *Lecture Notes in Computer Science*, pages 666–675. Springer Berlin / Heidelberg, 2007.

[244] M. Blatt and P. Bastian. On the generic parallelisation of iterative solvers for the finite element method. *International Journal of Computational Science and Engineering*, 4(1):56–69, 2008.

[245] P. Bastian, F. Heimann, and S. Marnach. Generic implementation of finite element methods in the distributed and unified numerics environment (dune). *Kybernetika*, 46(2):294–315, 2010.

[246] H. E. Kubitschek. Electronic counting and sizing of bacteria. *Nature*, 182:234 – 235, 1958.

[247] H. Chang, F. Kosari, G. Andreadakis, M. A. Alam, G. Vasmatzis, and R. Bashir. Dna-mediated fluctuations in ionic current through silicon oxide nanopore channels. *Nano Letters*, 4(8):1551–1556, 2004.

[248] R. Fan, R. Kamik, M. Yue, D. Li, A. Majumdar, and P. Yang. DNA translocation in inorganic nanotubes. *Nano Letters*, 5:1633–1637, 2005.

[249] U. Keyser, B. Koeleman, S. van Dorp, D. Krapf, R. Smeets, S. Lemay, N. Dekker, and C. Dekker. Direct force measurements on DNA in a solid-state nanopore. *Nature Physics*, 2:473–477, 2006.

[250] S. W. Kowalczyk, A. Y. Grosberg, Y. Rabin, and C. Dekker. Modeling the conductance and dna blockade of solid-state nanopores. *Nanotechnology*, 22(31):315101, 2011.

[251] B. Hille. Pharmacological modifications of the sodium channels of frog nerve. *The Journal of general physiology*, 51(2):199–219, 1968.

[252] J. Hall. Access resistance of a small circular pore. *The Journal of general physiology*, 66(4):531–532, 1975.

[253] A. J. C. Ladd. Numerical simulations of particulate suspensions via a discretized boltzmann equation. part 1. theoretical foundation. *Journal*

of Fluid Mechanics, 271:285–309, 1994.

[254] P. Ahlrichs and B. Dünweg. Lattice-boltzmann simulation of polymer-solvent systems. *International Journal of Modern Physics C*, 9(8):1429–1438, 1998.

[255] B. Dünweg and A. J. C. Ladd. Lattice boltzmann simulations of soft matter systems. In *Advanced Computer Simulation Approaches for Soft Matter Sciences III*, volume 221 of *Advances in Polymer Science*, pages 89–166. Springer-Verlag Berlin, Berlin, Germany, 2009.

[256] S. Succi. *The lattice Boltzmann equation for fluid dynamics and beyond.* Oxford University Press, USA, 2001.

[257] R. Benzi, S. Succi, and M. Vergassola. The lattice boltzmann equation: theory and applications. *Physics Reports*, 222(3):145–197, 1992.

[258] X. He and L.-S. Luo. A priori derivation of the lattice boltzmann equation. *Physical Review E*, 55(6):R6333, 1997.

[259] X. He and L.-S. Luo. Theory of the lattice boltzmann method: From the boltzmann equation to the lattice boltzmann equation. *Physical Review E*, 56(6):6811, 1997.

[260] X. Shan, X.-F. Yuan, and H. Chen. Kinetic theory representation of hydrodynamics: a way beyond the navier-stokes equation. *Journal of Fluid Mechanics*, 550(1):413–441, 2006.

[261] P. L. Bhatnagar, E. P. Gross, and M. Krook. A model for collision processes in gases. i. small amplitude processes in charged and neutral one-component systems. *Physical Review*, 94(3):511, 1954.

[262] M. Abramowitz and I. Stegun. *Handbook of mathematical functions.* Dover Publications Inc., New York, 1970.

[263] Y. Qian, D. d'Humieres, and P. Lallemand. Lattice bgk models for navier-stokes equation. *Europhysics Letters*, 17:479, 1992.

[264] S. De, K. Nagendra, and K. Lakshmisha. Simulation of laminar flow in a three-dimensional lid-driven cavity by lattice boltzmann method. *International Journal of Numerical Methods for Heat & Fluid Flow*, 19(6):790–815, 2009.

[265] D. d'Humieres. Multiple–relaxation–time lattice boltzmann models in three dimensions. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*,

360(1792):437–451, 2002.

[266] B. Dünweg, U. Schiller, and A. J. C. Ladd. Statistical mechanics of the fluctuating lattice-boltzmann equation. *Phys. Rev. E*, 76:36704, 2007.

[267] B. Dünweg, U. D. Schiller, and A. J. C. Ladd. Progress in the understanding of the fluctuating lattice boltzmann equation. *Computer Physics Communications*, 180(4):605–608, 2009.

[268] U. D. Schiller. *Thermal fluctuations and boundary conditions in the lattice Boltzmann method*. PhD thesis, Johannes Gutenberg-Universitaet Mainz, Fachbereich 08: Physik, Mathematik und Informatik, 2008.

[269] D. d'Humières and I. Ginzburg. Viscosity independent numerical errors for lattice boltzmann models: from recurrence equations to "magic" collision numbers. *Computers & Mathematics with Applications*, 58(5):823–840, 2009.

[270] D. d'Humieres and P. Lallemand. Numerical simulations of hydrodynamics with lattice gas automata in two dimensions. *Complex Systems*, 1(4):599, 1987.

[271] D. P. Ziegler. Boundary conditions for lattice boltzmann simulations. *Journal of Statistical Physics*, 71:1171–1177, 1993.

[272] M. A. Gallivan, D. R. Noble, J. G. Georgiadis, and R. O. Buckius. An evaluation of the bounce-back boundary condition for lattice boltzmann simulations. *International Journal for Numerical Methods in Fluids*, 25(3):249–263, 1997.

[273] D. R. Noble, S. Chen, J. G. Georgiadis, and R. O. Buckius. A consistent hydrodynamic boundary condition for the lattice boltzmann method. *Physics of Fluids*, 7:203, 1995.

[274] A. J. C. Ladd. Dynamical simulations of sedimenting spheres. *Phys. Fluids A*, 5(2):299–310, 1993.

[275] R. W. Nash, R. Adhikari, and M. E. Cates. Singular forces and pointlike colloids in lattice boltzmann hydrodynamics. *Phys. Rev. E*, 77:026709, 2008.

[276] S. Ollila, C. Smith, T. Ala-Nissila, and C. Denniston. The hydrodynamic radius of particles in the hybrid Lattice Boltzmann–Molecular Dynamics method. *Multiscale Modeling & Simulation*, 11(1):213–243, 2013.

[277] J. T. Padding and A. A. Louis. Hydrodynamic interactions and brow-

nian forces in colloidal suspensions: Coarse-graining over time and length scales. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 74(3):031402, 2006.

[278] P. Ahlrichs and B. Dünweg. Simulation of a single polymer chain in solution by combining lattice boltzmann and molecular dynamics. *Journal of Chemical Physics*, 111(17):8225–8239, 1999.

[279] H. Hasimoto. On the periodic fundamental solutions of the stokes equations and their application to viscous flow past a cubic array of spheres. *Journal of Fluid Mechnics*, 5(02):317–328, 1959.

[280] C. Beenakker. Ewald sum of the rotne–prager tensor. *The Journal of Chemical Physics*, 85:1581, 1986.

[281] C. Pozrikidis. Computation of periodic green's functions of stokes flow. *Journal of Engineering Mathematics*, 30(1-2):79–96, 1996.

[282] L. Onsager. Theories of concentrated electrolytes. *Chemical Reviews*, 13(1):73–89, 1933.

[283] A. L. Horvath. *Handbook of Aqueous Electrolyte Solutions*. Ellis Horwood Ltd., Chichester, 1985.

[284] M. Deserno, C. Holm, J. Blaul, M. Ballauff, and M. Rehahn. The osmotic coefficient of rod-like polyelectrolytes: Computer simulation, analytical theory, and experiment. *European Physical Journal E: Soft Matter*, 5:97–103, 2001.

[285] J. Blake. A note on the image system for a Stokeslet in a no-slip boundary. *Proc. Camb. Phil. Soc*, 70:303–310, 1971.

[286] G. Perkins and R. Jones. Hydrodynamic interaction of a spherical particle with a planar boundary: Ii. hard wall. *Physica A: Statistical Mechanics and its Applications*, 189(3-4):447 – 477, 1992.

[287] Y. von Hansen, M. Hinczewski, and R. R. Netz. Hydrodynamic screening near planar boundaries: Effects on semiflexible polymer dynamics. *The Journal of Chemical Physics*, 134(23):–, 2011.

[288] H. Bekker, H. J. C. Berendsen, E. J. Dijkstra, S. Achterop, R. van Drunen, D. van der Spoel, A. Sijbers, H. Keegstra, B. Reitsma, and M. K. R. Renardus. Gromacs: A parallel computer for molecular dynamics simulations. In R. A. de Groot and J. Nadrchal, editors, *Physics Computing 92*. World Scientific, 1993.

[289] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, 91(1-3):43–56, 1995.

[290] D. van der Spoel, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, B. Hess, K. A. Feenstra, E. Lindahl, R. van Drunen, and H. J. C. Berendsen. *GROMACS: GROningen MAchine for Chemical Simulations – User Manual, Version 3.1*. Nijenborgh 4, 9747 AG Groningen, The Netherlands, 2001.

[291] D. V. D. Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16):1701–1718, 2005.

[292] D. van der Spoel, E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. Sijbers, K. A. Feenstra, R. van Drunen, and H. Berendsen. *GROMACS User Manual version 4.5*, 2005.

[293] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. R. Ross, I. T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman. AMber: Assisted model building with energy refinement – a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comp. Phys. Comm.*, 91:1–41, 1995.

[294] R. Salomon-Ferrer, D. A. Case, and R. C. Walker. An overview of the amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2012.

[295] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, S. Swaminathan, M. Karplus, et al. Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*, 4(2):187–217, 1983.

[296] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, et al. Charmm: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.

[297] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kale, R. D. Skeel, and K. Schulten. NAMD – a parallel, object-oriented molecular dynamics program. *Int. J. Supercomput. Ap.*, 10(4):251–268, 1996.

[298] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa,

C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[299] B. Luan and A. Aksimentiev. Electro-osmotic screening of the dna charge in a nanopore. *Physical Review E*, 78(2):021912, 2008.

[300] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, 2008.

[301] J. Yoo and A. Aksimentiev. Improved parametrization of li+, na+, k+, and mg2+ ions for all-atom molecular dynamics simulations of nucleic acid systems. *The Journal of Physical Chemistry Letters*, 3(1):45–50, 2011.

[302] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79(2):926–935, 1983.

[303] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz J.R, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117:5179–5197, 1995.

[304] T. E. Cheatham III, P. Cieplak, and P. A. Kollman. A modified version of the cornell et al. force field with improved sugar pucker phases and helical repeat. *Journal of Biomolecular Structure and Dynamics*, 16(4):845–862, 1999.

[305] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham III, C. A. Laughton, and M. Orozco. Refinement of the amber force field for nucleic acids: Improving the description of α γ conformers. *Biophysical journal*, 92(11):3817–3829, 2007.

[306] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry*, 24(16):1999–2012, 2003.

[307] A. D. MacKerell Jr, J. Wiorkiewicz-Kuczera, and M. Karplus. An all-

atom empirical energy function for the simulation of nucleic acids. *Journal of the American Chemical society*, 117(48):11946–11975, 1995.

[308] A. D. MacKerell, N. Banavali, and N. Foloppe. Development and current status of the charmm force field for nucleic acids. *Biopolymers*, 56(4):257–265, 2000.

[309] T. Schlick. *Molecular modeling and simulation: an interdisciplinary guide*. Springer, 2010.

[310] T. A. Soares, P. H. Hünenberger, M. A. Kastenholz, V. Kräutler, T. Lenz, R. D. Lins, C. Oostenbrink, and W. F. van Gunsteren. An improved nucleic acid parameter set for the gromos force field. *Journal of computational chemistry*, 26(7):725–737, 2005.

[311] H. J. C. Berendsen, J. P. M. Postma, W. F. Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. In B. Pullman, editor, *Intermolecular Forces*, volume 14 of *The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, pages 331–342. Springer Netherlands, 1981.

[312] J. Aqvist. Ion-water interaction potentials derived from free energy perturbation simulations. *The Journal of Physical Chemistry*, 94(21):8021–8024, 1990.

[313] P. Auffinger, T. E. Cheatham, and A. C. Vaiana. Spontaneous formation of kcl aggregates in biomolecular simulations: A force field issue? *Journal of Chemical Theory and Computation*, 3(5):1851–1859, 2007.

[314] A. A. Chen and R. V. Pappu. Parameters of monovalent ions in the amber-99 forcefield: Assessment of inaccuracies and proposed improvements. *The Journal of Physical Chemistry B*, 111(41):11884–11887, 2007.

[315] F. Mocci and A. Laaksonen. Insight into nucleic acid counterion interactions from inside molecular dynamics simulations is ???worth its salt??? *Soft Matter*, 8(36):9268–9284, 2012.

[316] P. Hunenberger and M. Reif. *Single-Ion Solvation*. RSC Theoretical and Computational Chemistry Series. The Royal Society of Chemistry, 2011.

[317] I. S. Joung and T. E. Cheatham III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *The journal of physical chemistry B*, 112(30):9020–9041,

2008.

[318] V. P. Denisov and B. Halle. Sequence-specific binding of counterions to b-dna. *Proceedings of the National Academy of Sciences*, 97(2):629–633, 2000.

[319] R. Owczarzy, B. G. Moreira, Y. You, M. A. Behlke, and J. A. Walder. Predicting stability of dna duplexes in solutions containing magnesium and monovalent cations. *Biochemistry*, 47(19):5336–5353, 2008.

[320] A. Noy, I. Soteras, F. J. Luque, and M. Orozco. The impact of monovalent ion force field model in nucleic acids simulations. *Physical Chemistry Chemical Physics*, 11(45):10596–10607, 2009.

[321] D. E. Smith and L. X. Dang. Computer simulations of nacl association in polarizable water. *Journal of Chemical Physics*, 100:3757, 1994.

[322] L. X. Dang and P. A. Kollman. Free energy of association of the k+: 18-crown-6 complex in water: a new molecular dynamics study. *The Journal of Physical Chemistry*, 99(1):55–58, 1995.

[323] L. X. Dang. Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: a molecular dynamics study. *Journal of the American Chemical Society*, 117(26):6954–6960, 1995.

[324] I. S. Joung and T. E. Cheatham III. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *The Journal of Physical Chemistry B*, 113(40):13279–13290, 2009.

[325] B. Hess. P–LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2008.

[326] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. Pedersen. A smooth Particle Mesh Ewald method. *Journal of Chemical Physics*, 103:8577, 1995.

[327] M. Feig and B. M. Pettitt. Sodium and chlorine ions as part of the dna solvation shell. *Biophysical Journal*, 77:1769???1781, 1999.

[328] C. Maffeo, R. Schöpflin, H. Brutzer, R. Stehr, A. Aksimentiev, G. Wedemann, and R. Seidel. Dna–dna interactions in tight supercoils are described by a small effective charge density. *Physical Review Letters*, 105(15):158101, 2010.

[329] J. Comer and A. Aksimentiev. Predicting the dna sequence dependence of nanopore ion current using atomic-resolution brownian dynamics. *The Journal of Physical Chemistry C*, 116(5):3376–3393, 2012.

[330] R. Netz. Electrofriction and dynamic Stern layers at planar charged surfaces. *Physical Review Letters*, 91(13):138101, 2003.

[331] Y. W. Kim and R. R. Netz. Electro-osmosis at inhomogeneous charged surfaces: Hydrodynamic versus electric friction. *The Journal of Chemical Physics*, 124:114709, 2006.

[332] P. Rowghanian and A. Y. Grosberg. Two cases of reciprocal relations for electric and hydrodynamic currents: A rigid polymer in a nano-channel and a polyelectrolyte gel. *The Journal of Chemical Physics*, 139(2):–, 2013.

[333] J. Thomas and K. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press, 2005.

[334] D. A. Keim, F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler. Visual analytics: Scope and challenges. In *Visual Data Mining*, pages 76–90. Springer, 2008.

[335] J. Kehrer and H. Hauser. Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):495–513, 2013.

[336] Megamol project website. http://www.vis.uni-stuttgart.de/megamol (last accessed: 2013-12-20).

[337] W. Schroeder, K. Martin, and B. Lorensen. *An Object-Oriented Approach To 3D Graphics*. Prentice Hall, 1997.

[338] A. Henderson, J. Ahrens, and C. Law. The ParaView Guide. Published by Kitware Inc., Clifton Park, 2004.

[339] T. Can, C.-I. Chen, and Y.-F. Wang. Efficient molecular surface generation using level-set methods. *Journal of Molecular Graphics and Modelling*, 25(4):442–454, 2006.

[340] W. E. Lorensen and H. E. Cline. Marching Cubes: A high resolution 3d surface construction algorithm. In *Proceedings of ACM SIGGRAPH Computer Graphics and Interactive Techniques*, pages 163–169, 1987.

[341] M. Hadwiger, P. Ljung, C. R. Salama, and T. Ropinski. Advanced illu-

mination techniques for GPU volume raycasting. In *ACM SIGGRAPH Asia 2008 courses*, pages 1–166, 2008.

[342] R. J. Allen, D. Frenkel, and P. R. ten Wolde. Simulating rare events in equilibrium or nonequilibrium stochastic systems. *Journal of Chemical Physics*, 124(2):024102, 2006.

[343] R. J. Allen, D. Frenkel, and P. R. ten Wolde. Forward flux sampling-type schemes for simulating rare events: Efficiency analysis. *Journal of Chemical Physics*, 124(19):194111, 2006.

[344] S. Getfert, T. Töws, and P. Reimann. Opposite translocation of long and short oligomers through a nanopore. *Phys. Rev. E*, 87:062710, 2013.

[345] M. Fixman. Simulation of polymer dynamics.I. general theory. *Journal of Chemical Physics*, 69:1527, 1978.

[346] Y. Dorozhko, K. Kratzer, Y. Yudin, A. Arnold, C. W. Glass, and M. Resch. Rare event sampling using the science experimental grid laboratory. In B. Topping and P. Iványi, editors, *Proceedings of the Fourteenth International Conference on Civil, Structural and Environmental Engineering Computing*, page Paper 207, Stirlingshire, UK, 2013. Civil-Comp Press.

[347] K. Kratzer and J. Berryman. Website of the Flexible Rare Event Sampling Harness System (FRESHS): www.freshs.org.

[348] E. Allahyarov and H. Löwen. Effective interaction between helical biomolecules. *Phys. Rev. E*, 62(4):5542–5556, 2000.

[349] E. Allahyarov, G. Gompper, and H. Löwen. Attraction between dna molecules mediated by multivalent ions. *Physical Review E*, 69(4):041904, 2004.

[350] E. Allahyarov, H. Löwen, and G. Gompper. Adsorption of monovalent and multivalent cations and anions on dna molecules. *Physical Review E*, 68(6):061903, 2003.

[351] T. A. Knotts IV, N. Rathore, D. C. Schwartz, and J. J. de Pablo. A coarse grain model for dna. *The Journal of Chemical Physics*, 126:084901, 2007.

[352] J. J. de Pablo. Coarse-grained simulations of macromolecules: from dna to nanocomposites. *Annual Review of Physical Chemistry*, 62:555–574, 2011.

[353] E. Sambriski, D. Schwartz, and J. de Pablo. A mesoscale model of DNA and its renaturation. *Biophysical Journal*, 96(5):1675 – 1690, 2009.

[354] C. W. Hsu, M. Fyta, G. Lakatos, S. Melchionna, and E. Kaxiras. Ab initio determination of coarse-grained interactions in double-stranded dna. *The Journal of Chemical Physics*, 137(10):–, 2012.

[355] J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Sulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J. Smith. Coarse-graining dna for simulations of dna nanotechnology. *Phys. Chem. Chem. Phys.*, 15:20395–20414, 2013.

[356] J. Yoo and A. Aksimentiev. In situ structure and dynamics of dna origami determined through molecular dynamics simulations. *Proceedings of the National Academy of Sciences*, 110(50):20099–20104, 2013.

[357] R. Peierls. Model-making in physics. *Contemporary Physics*, 21(1):3–17, 1980.

[358] M. Suarez. Scientific representation. *Philosophy Compass*, 5(1):91–101, 2010.

[359] S. W. Yi. The nature of model-based understanding in condensed matter physics. *Mind & Society*, 3(1):81–91, 2002.

[360] A. Gelfert. Strategies of model-building in condensed matter physics: trade-offs as a demarcation criterion between physics and biology? *Synthese*, 190(2):253–272, 2013.

[361] T. Knuuttila. Modelling and representing: An artefactual approach to model-based representation. *Studies In History and Philosophy of Science Part A*, 42(2):262–271, 2011.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass die Dissertation von mir selbstständig angefertigt wurde und alle von mir genutzten Hilfsmittel angegeben wurden. Aus fremden Quellen entnommene Gedanken und Passagen sind als solche kenntlich gemacht.

Stefan Kesselheim