Charge Transport in amorphous organic semiconductors

Dissertation

zur Erlangung des Grades

"Doktor der Naturwissenschaften"

am Fachbereich Physik
der Johannes Gutenberg-Universität Mainz

Alexander Lukyanov
geb. in Glasow, Russland

Max-Planck-Institut für Polymerforschung
Mainz, März 2011

Abstract

Organic semiconductors with the unique combination of electronic and mechanical properties may offer cost-effective ways of realizing many electronic applications, e. g. large-area flexible displays, printed integrated circuits and plastic solar cells. In order to facilitate the rational compound design of organic semiconductors, it is essential to understand relevant physical properties e. g. charge transport. This, however, is not straightforward, since physical models operating on different time and length scales need to be combined. First, the material morphology has to be known at an atomistic scale. For this atomistic molecular dynamics simulations can be employed, provided that an atomistic force field is available. Otherwise it has to be developed based on the existing force fields and first principle calculations. However, atomistic simulations are typically limited to the nanometer length- and nanosecond time-scales. To overcome these limitations, systematic coarse-graining techniques can be used.

In the first part of this thesis, it is demonstrated how a force field can be parameterized for a typical organic molecule. Then different coarse-graining approaches are introduced together with the analysis of their advantages and problems. When atomistic morphology is available, charge transport can be studied by combining the high-temperature Marcus theory with kinetic Monte Carlo simulations.

The approach is applied to the hole transport in amorphous films of tris(8-hydroxyquinoline)aluminium (Alq_3). First the influence of the force field parameters and the corresponding morphological changes on charge transport is studied. It is shown that the energetic disorder plays an important role for amorphous Alq_3, defining charge carrier dynamics. Its spatial correlations govern the Poole-Frenkel behavior of the charge carrier mobility. It is found that hole transport is dispersive for system sizes accessible to simulations, meaning that calculated mobilities depend strongly on the system size. A method for extrapolating calculated mobilities to the infinite system size is proposed, allowing direct comparison of simulation results and time-of-flight experiments. The extracted value of the nondispersive hole mobility and its electric field dependence for amorphous Alq_3 agree well with the experimental results.

Zusammenfassung

Organische Halbleiter weisen eine Kombination von elektronischen und mechanischen Eigenschaften auf, die eine kostengünstige Realisierung vieler elektronischen Anwendungen, z. B. großflächige flexible Displays, gedruckte integrierte Schaltungen oder Kunststoff-Solarzellen ermöglichen können. Im Zuge des Entwurfs neuer organischer Halbleiter ist es wichtig, relevante physikalische Eigenschaften zu verstehen, z. B. Ladungstransport. Dies ist jedoch mit Schwierigkeiten verbunden, da mehrere Modelle, die Effekte auf unterschiedlichen Zeit- und Längenskalen beschreiben, kombiniert werden müssen. Zunächst muss die Materialmorphologie mit atomistischer Auflösung bekannt sein. Diese kann durch Molekulardynamik-Simulationen generiert werden, unter der Voraussetzung, dass ein atomistisches Kraftfeld zur Verfügung steht, welches auf der Basis bestehender Kraftfelder und *ab-initio* Rechnungen entwickelt werden kann. Allerdings sind atomistische Simulationen auf Längen- und Zeitskalen beschränkt, die in der Größenordnung einiger Nanometer bzw. Nanosekunden liegen. Um größere Skalen zu erschließen können systematische Vergröberungstechniken (Coarse-Graining) verwendet werden.

Im ersten Teil dieser Arbeit wird gezeigt, wie ein Kraftfeld für ein typisches organisches Molekül parametrisiert werden kann. Dann werden verschiedene Vergröberungsansätze eingeführt und deren Vorteile und Probleme diskutiert. Sobald eine atomistische Morphologie zur Verfügung steht, kann der Ladungstransport durch eine Kombination der Hochtemperatur-Marcus-Theorie mit der kinetischen Monte-Carlo-Methode simuliert werden.

Dieser Ansatz wird verwendet, um den Löchertransport in amorphen Schichten von Tris-(8-Hydroxychinolin)-Aluminium zu simulieren. Zunächst wird der Einfluss der Kraftfeldparameter und der entsprechenden morphologischen Veränderungen auf den Ladungstransport untersucht. Es wird gezeigt, dass die energetische Unordnung im System eine wichtige Rolle spielt. Diese beeinflusst die Ladungsträgerdynamik erheblich, entsprechende räumliche Korrelationen beeinflussen das Poole-Frenkel-Verhalten der Ladungsträgermobilität. Es ist zu beobachten, dass Löchertransport dispersiven Charakter besitzt, d. h. berechnete Beweglichkeiten hängen stark von der Systemgröße ab. Es wird ein Verfahren zur Extrapolation der berechneten Beweglichkeiten zur unendlichen Systemgröße vorgeschlagen, das einen direkten Vergleich der Simulationsergebnisse zu Time-Of-Flight-Experimenten ermöglicht. Die extrahierten Werte der nichtdispersiven Löchermobilität und ihre Abhängigkeit vom elektrischen Feld stimmen für amorphes Alq_3 gut mit den experimentellen Ergebnissen überein.

Contents

Related Publications

[1] **A. Lukyanov**, C. Lennartz, and D. Andrienko. Amorphous films of tris(8-hydroxyquinolinato)aluminium: Force-field, morphology, and charge transport. *Phys. Stat. Sol.(a)*, 206:2737–2742, 2009.

[2] V. Rühle, C. Junghans, **A. Lukyanov**, K. Kremer, and D. Andrienko. Versatile object-oriented toolkit for coarse-graining applications. *J. Chem. Theory Comput.*, 5(12):3211–3223, 2009.

[3] **A. Lukyanov**, A. Malafeev, V. Ivanov, H.-L. Chen, K. Kremer, and D. Andrienko. Solvated poly-(phenylene vinylene) derivatives: conformational structure and aggregation behavior. *J. Mater. Chem.*, 20:10475–10485, 2010.

[4] **A. Lukyanov** and D. Andrienko. Extracting nondispersive charge carrier mobilities of organic semiconductors from simulations of small systems. *Phys. Rev. B.*, 82(19):193202, 2010.

Chapter 1

Organic electronics

When the first transistor was invented in the middle of 20th century [5], inorganic semiconductors such as silicon and germanium began to play a dominating role in electronics, where metals were prevailing before. Today science and industry are working on a new class of materials, known as organic semiconductors, which promise to revolutionize electronics again. Organic semiconductors have electronic properties approaching those of inorganic counterparts combined with mechanical properties of plastic materials, offering low-cost processing and a possibility to realize new applications, such as large-area flexible displays, low-cost printed integrated circuits and plastic solar cells [6]. Moreover, by varying the chemistry slightly, it is possible to modify material properties of interest (e.g. band gap), to achieve the desired device performance.

There are two major classes of organic semiconductors: low molecular weight materials (usually processed in vacuum) and polymers (usually processed by wet chemical techniques). Typical examples are shown in Fig. 1.1. Both have in common a conjugated π-electron system, which is responsible for their semiconducting properties. Thin films of organic semiconductors are typically used in three types of devices: (1) Organic light emitting diodes (OLED), (2) Organic field-effect transistors (OFET), (3) organic photovoltaic cells (OPVC). Typical structures of these devices are depicted in Fig. 1.2 and Fig. 1.3. An organic light emitting diode (OLED) emits light in response to an electric current [7]. Solar cell does the opposite: it converts the energy of sunlight directly into electricity [8]. Organic field-effect transistor controls the conductivity of a channel, made of organic semiconductor, by applying electric field [9].

In the following section we will discuss operation principles of the organic electronics devices in detail.

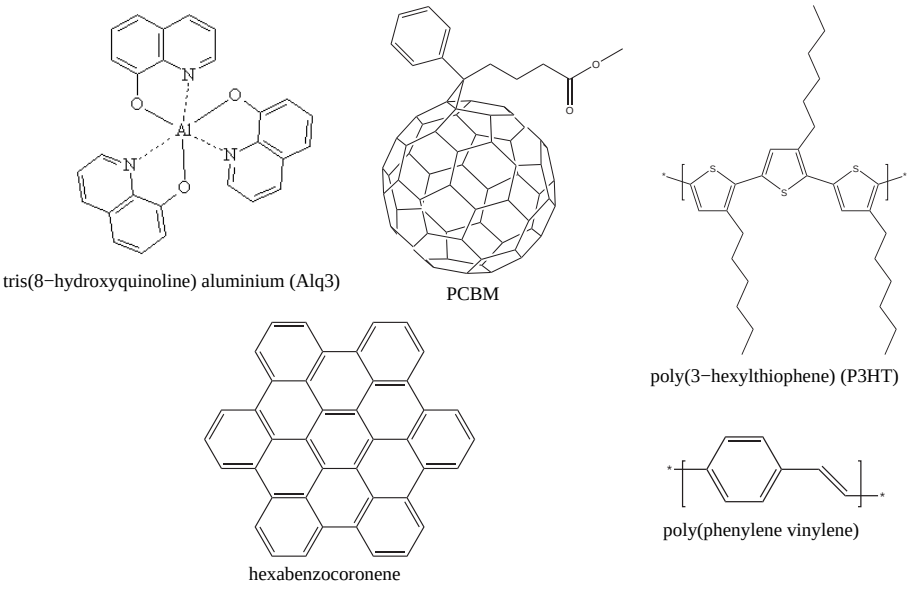

Figure 1.1: **Typical compounds used in organic electronics applications.**

1.1 Devices

1.1.1 Organic light emitting diodes

A typical OLED is composed of several layers of organic materials sandwiched between two electrodes. Charge carriers of both signs are injected from the electrodes and recombine, forming a neutral exciton, a bound state of an electron-hole pair. The decay of an exciton results in the emission of radiation whose frequency is in the visible range. The frequency of this radiation depends on the optical band gap of the material.[1] For the operation of the device, charge carriers must be efficiently injected to the organic film from the electrodes. This requires low energetic barriers at the metal-organic interfaces for both contacts in order to inject equally high amounts of electrons and holes and to provide a balanced charge carrier flow. In single layer devices hole and electron concentrations might be imbalanced, for example due to different electron and hole mobilities, which leads to low efficiency. Better efficiency is obtained if additional layers are added to the device. The hole (electron) transporting layer is used to provide intermediate energy states to allow holes (electrons) to cascade through smaller gaps

[1]Optical band gap is related to the exciton binding energy and the energy difference between the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO).

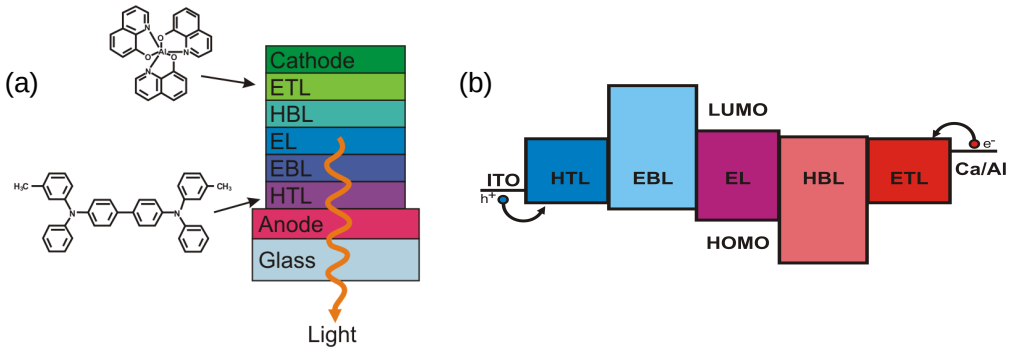

Figure 1.2: **Organic light emitting diode.** (a) Typical device consists of the following layers: electron transporting layer (ETL), hole blocking layer (HBL), emission layer (EL), electron blocking layer (EBL), hole transporting layer (HTL). Adapted from [6]; (b) Energy levels in OLED layers.

in case there is a barrier between the metal work function and the HOMO (LUMO) of the emission layer. Hole (electron) blocking layers have very deep HOMO (or high LUMO) levels which help to prevent holes (electrons) from passing through the device to the opposite electrode without forming an exciton. Typical OLED structure is shown in Fig. 1.2.

The first high-performance OLED was reported in 1987 [10]. These first devices used tris(8-hydroxyquinoline)aluminium (Alq_3) (see Fig. 1.1) as an electron transporting layer. After more than two decades of intensive research and development of OLEDs, Alq_3 continues to be the workhorse in low-molecular weight materials for these devices. It is used as an electron-transporting layer, as well as an emission layer where green light emission is generated by the electron-hole recombination in Alq_3. It also serves as a host material for various dyes, helping to tune the emission color from green to red [11]. Many studies have been focused on the optimization of OLED efficiency and long-term stability, by means of the understanding of charge transport properties of amorphous thin films, see e.g. Refs. [12, 13, 14, 15, 16, 17].

Understanding microscopic mechanisms of charge transport in amorphous Alq_3 is one of the topics of this thesis. It is covered in detail in Chapter 8.

1.1.2 Organic field effect transistors

Organic field effect transistors (OFETs) are the basic building blocks for flexible integrated circuits and displays. An OFET can be viewed as a resistor, that can be adjusted by applying external voltage. A typical OFET structure is shown in Fig. 1.3a. The current between the source and the drain

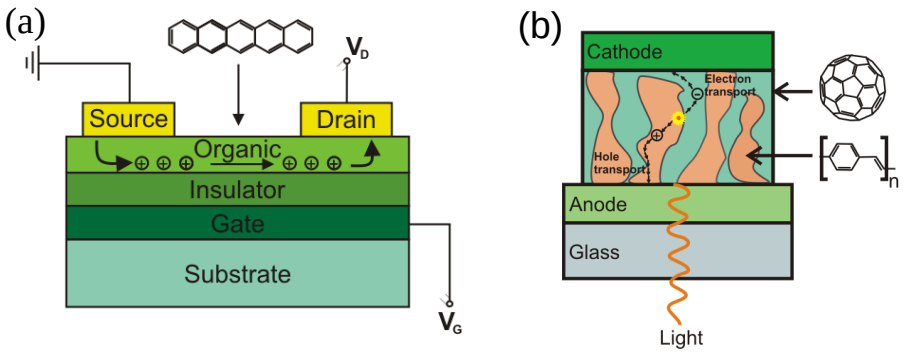

Figure 1.3: **Organic electronics devices** (a) field effect transistor; (b) organic solar cell. Adapted from [6].

electrodes is modulated by applying voltage to the gate electrode, which is separated from the transporting layer by a dielectric. Application of voltage to the gate electrode changes the charge carrier density in the organic semiconductor, thus modifying the resistance. In an ideal device there is no conductance without the application of the gate voltage ("off" state). When the gate voltage is applied, current between source and drain appears ("on" state). The on-off current ratio characterizes the ability of the device to "switch off". Currently values up to 10^8 can be achieved [18]. Device performance also critically depends on charge carrier mobility, which must be high enough in order to obtain source-drain currents, which can be modulated by a reasonable gate voltage. Therefore, designing materials with high charge carrier mobilities is a key to high-performance OFETs.

1.1.3 Organic solar cells

Organic solar cells convert sunlight into electric current. This process can be schematically described by the following steps: (1) absorption of a photon leading to the formation of a bound electron-hole pair (exciton); (2) exciton diffusion to a region where exciton dissociation (charge separation) occurs; and (3) charge transport within the organic semiconductor to the respective electrodes [19].

The large optical band gap in organic materials (normally higher than 2 eV) limits light harvesting to 30 % [19], reducing the device efficiency. To overcome this problem, low band gap polymers must be designed [20, 21]. Another problem is that primary photoexcitations do not lead directly to free charge carriers, but to coulomb-bound electron-hole pairs (excitons). Since exciton binding in organic semiconductors is of the order of $10 - 50\,k_BT$, thermal energy can not drive charge separation, as it would happen in inor-

ganic materials. Instead, an additional mechanism is required for the charge separation to occur [22]. One way to achieve this is to use a mixture of donor and acceptor compounds (e.g. conjugated polymers with fullerenes). By tuning the difference between the lowest unoccupied molecular orbitals (LUMO) of a donor and an acceptor it is possible to compensate for the exciton binding energy. In this case, electron transfers to the acceptor, leaving a hole in the donor. After the separation, free carriers diffuse to the electrodes and generate electric current in the external circuit.

Since the typical exciton diffusion length in organic materials is of the order of $10 - 20\,\mathrm{nm}$, only excitons created within this distance from the interface can reach it. This leads to the loss of absorbed photons further away from the interface and results in low efficiencies [23]. One way to solve this problem is to use bulk heterojunctions - blends of the donor and acceptor components in a bulk volume, which exhibit phase separation in a $10-20\,\mathrm{nm}$ length scale (see Fig. 1.3b). In such a nanoscale interpenetrating network, each interface is within a distance less than the exciton diffusion length from the absorbing site. By using the bulk heterojunction concept it is possible to increase the interfacial area between the donor and acceptor phases by orders of magnitude, significantly improving the solar cell efficiency [19].

1.2 Mobility measurements

The devices mentioned above share one common feature: their performance critically depends on the efficiency with which charge carriers move within the π-conjugated materials. The physical quantity which describes this efficiency is charge carrier mobility. Mobility is one of the key parameters of interest - both towards realizing improved device performance, as well as understanding the underlying semiconductor physics in these materials. It is in particular important for the efficiency of transistors (how fast they can switch) and solar cells (how fast the separated charges can "run away" from each other). Mobility strongly depends on the processing, chemical structure and purity of a material. Organic semiconductors have lower mobilities compared to their inorganic counterparts[2]: values of $0.1 - 1.0\,\mathrm{cm^2/Vs}$ are considered to be good for organic semiconductors. Most materials have mobilities which are orders of magnitude smaller. For example, hole mobilities in amorphous Alq_3 are of the order $10^{-9} - 10^{-8}\,\mathrm{cm^2/Vs}$ [26].

A number of techniques was designed in order to measure charge carrier mobility experimentally. An extensive review is given in Ref. [27]. Here we

[2]Typical electron mobility for crystalline silicon at room temperature (300 K) is $1400\,\mathrm{cm^2/Vs}$ and the hole mobility is around $450\,\mathrm{cm^2/Vs}$ [24, 25].

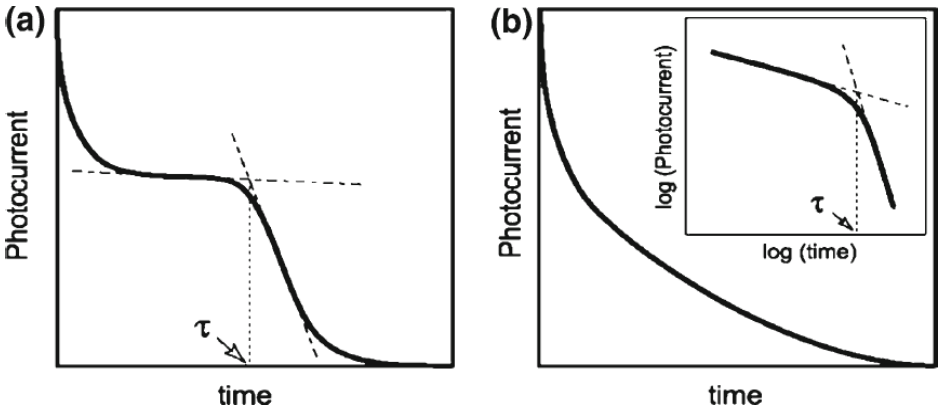

Figure 1.4: **Typical transient photocurrents** (a) non-dispersive; (b) dispersive. *Inset*: double logarithmic plot. Taken from [27].

will briefly describe the time-of-flight, diode and transistor measurements.

1.2.1 Time-of-flight measurements

The time-of-flight (TOF) method is based on the measurement of the carrier transit time τ, namely, the time required for a packet of carriers to drift across the organic semiconducting layer. In the TOF setup the material of interest is sandwiched between two electrodes, one of which is transparent. Charges are generated by photo-excitation of the film through irradiation with a short laser pulse. Subsequently, carriers propagate along the electric field and generate a displacement current, which flows until charge carriers arrive at the other electrode. Typical photocurrents are shown in Fig. 1.4. From the cusp of the *non-dispersive* photocurrent (Fig. 1.4a) the transit time τ can be determined. Then the charge carrier mobility is calculated as

$$\mu = \frac{v}{E} = \frac{d}{E\tau} = \frac{d^2}{V\tau} \tag{1.1}$$

where V is the applied voltage, $E = V/d$ is the electric field, and d is the sample thickness.

Polymers and amorphous glasses often exhibit *dispersive* photocurrents, without any definite cusp, as shown in Fig. 1.4b. In this case, τ can still be determined from the double logarithmic plot. Although mobilities determined from dispersive transients are thickness-dependent. This means that the dispersive mobility is not a material constant, but depends on the system size. First microscopic statistical model of dispersive transport was proposed in Ref. [28]. As opposed to the Gaussian packet (nondispersive transport), where the peak and the mean are located at the same position and move

with the same velocity, the mean carrier of a "dispersive" packet propagates with a velocity which decreases with time as it separates from the peak which remains nearly fixed at the point of origin of the carriers.

Due to the requirement that the absorption depth of the laser excitation must be much smaller than the sample thickness, TOF method requires thick samples (usually in the range from 4 to 20 μm). Thus, the measured bulk mobilities are sensitive to the positional and orientational disorder and defects in the sample [29]. Note that a typical TOF setup has very low charge carrier densities, meaning that single carrier mobilities are measured. An advantage of the time-of-flight method is that hole and electron mobilities can be studied separately.

1.2.2 Diode and transistor measurements

An alternative approach to measure the mobility of an organic material is to embed it as a functional layer into a device and extract the mobility from the device characteristics [30]. In an OLED, the material of interest is sandwiched between two electrodes, which are chosen in such a way that only holes or electrons are injected at low voltage. In the absence of traps and at low electric fields, current-voltage characteristics may be expressed as [27]:

$$I = \frac{9}{8}\epsilon_0\epsilon_r\mu\frac{U^2}{L^3} \tag{1.2}$$

Here ϵ_0 (ϵ_r) is dielectric permittivity of free space (of the semiconducting layer) and L is the device thickness. μ is charge carrier mobility, which is assumed to be constant throughout the sample. The prefactor 9/8 comes from the assumption that the diode has a rectangular geometry. Dependence (1.2) is characteristic of a space-charge limited current (SCLC). Space-charge limitation of the current means that the number of charge carriers in the material is not limited by the injection but by the amount of the carriers already present in the sample. Their electrostatic potential prevents injection of the additional charges [31]. In the SCLC scheme, the charge density is not uniform across the material and is the largest close to the injecting electrode. In the presence of traps and at high fields, Eq. 1.2 must be modified [27].

Similarly, carrier mobilities can be extracted from the electrical characteristics measured in a field-effect transistor (FET) configuration. The $I-V$ (current-voltage) expression in a linear regime is given by

$$I_{SD} = \frac{W}{L}\mu\left(V_G - V_T\right)V_{SD} \tag{1.3}$$

and in a saturated regime by

$$I_{SD} = \frac{W}{2L}\mu C \left(V_G - V_T\right)^2 \tag{1.4}$$

Here, I_{SD} and V_{SD} are the current and the voltage bias between source and drain, respectively, V_G denotes the gate voltage, V_T is the threshold voltage at which the current starts to rise, C is the capacitance of the gate dielectric, and W and L are the width and the length of the conducting channel.

In a FET setup mobilities are measured at high charge carrier densities. Since charge transport occurs in a narrow channel, it is affected by structural defects at the interface, surface topology, or polarity of the dielectric. This influence might be irrelevant for amorphous materials, but becomes important for crystalline or liquid-crystalline materials.

1.2.3 Summary: mobility measurements

Different techniques discussed above operate at different conditions: single carrier bulk mobilities are measured by the time-of-flight technique, whereas interface mobilities at high charge carrier densities are extracted from OFET characteristics. Other techniques exist, such as Pulse Radiolysis - Time Resolved Microwave Conductivity (PR-TRMC) [32], which measure local mobilities of small well-ordered domains. As a result, mobility values obtained by different methods for the same material can differ by orders of magnitude [27]. This should be kept in mind when comparing simulations with experimental data. Single carrier kinetic Monte Carlo simulations described in the next Chapter mimic time-of-flight experiments and thus must be compared to them. Nevertheless care must be taken when the simulation box contains only a few hopping sites as explained in detail in Chapter 8.

1.3 Importance of theory and simulation

Organic electronic devices presented in this chapter are promising candidates to replace (at least for some applications) their silicon-based analogues. This would simplify the production process, since cost-efficient techniques such as spin coating and ink-jet printing can be employed. Combination of mechanical and semiconducting properties of conjugated polymers allows design of flexible electronics, such as bendable solar cells, rollable light sources and displays.

In order to be competitive on the market, organic electronic devices must possess high enough efficiency and stability. At the moment only OLEDs are

routinely used in commercial applications: Samsung, Sony and HTC fabricate TV and smartphone devices with OLED displays [33]. However, solar cells reported so far show very low power conversion efficiency (about 5% [20]) and are not stable long enough under ambient conditions. To systematically improve the device performance and stability, fundamental understanding of the underlying physical and chemical processes is required. It is important to understand the relationships between the chemical structure of the compound, its morphology and finally the properties of the resulting thin film or device. In order to achieve this understanding, models and methods, applicable on different length- and time scales should be combined. In the past decades methods were developed to describe charge transport in organic materials on microscopic, mesoscopic and macroscopic levels. On a microscopic level Marcus theory is normally used (see Chapter 2), which provides rates for the charge hopping between neighboring molecules based on first principles calculations. On the mesoscopic level charge carrier mobilities can be calculated by solving a master equation for a system of hopping sites, given some assumptions about the charge hopping rates. This approach is usually called the Gaussian disorder model and it is described in detail in Chapter 2. On the macroscopic level device characteristics can be calculated using drift-diffusion equations, given charge carrier mobility is known as a function of temperature, electric field and charge carrier density [34].

In this thesis microscopic and mesoscopic levels of description are combined to calculate charge carrier mobilities in amorphous films of tris(8-hydroxyquinoline)aluminium (Alq_3). To model amorphous films classical molecular dynamics simulations are used. The approach works as follows: first, an atomistic force-field for Alq_3 is developed and validated, based on existing force-fields, first principle calculations and experimental data. Second, realistic material morphologies are obtained using molecular dynamics simulations. Semi-classical Marcus theory is then used to calculate charge hopping rates between all neighboring molecules in the amorphous morphology. Finally, kinetic Monte Carlo simulations are used to calculate charge carrier mobilities.

The above mentioned techniques are used to study how non-bonded force-field parameters affect the morphology of amorphous Alq_3 and how sensitive charge transport properties are to the corresponding morphological changes. It is shown that in the particular case of Alq_3, the energetic disorder plays an important role, defining charge carrier dynamics, and its spatial correlations govern the Poole-Frenkel behavior of the charge carrier mobility. It is found that hole transport in amorphous Alq_3 is dispersive for the system sizes accessible to simulations, meaning that calculated mobilities depend strongly on the system size. A method for extrapolating calculated mobil-

ities to the infinite system size is proposed, allowing direct comparison of simulation results and time-of-flight experiments. The extracted value of the non-dispersive hole mobility and its electric field dependence for amorphous Alq_3 agree well with the experimental results.

Chapter 2

Theoretical description of charge transport

Charge transport mechanisms and their descriptions in organic semiconductors can vary significantly depending on the degree of structural order. In the extreme case of highly purified molecular crystals at low temperatures band transport is observed [35]. This means that charge carriers are delocalized and their mobility is determined from their effective mass and the mean relaxation time of the band states [30]. However, electronic delocalization is weak compared to inorganic semiconductors (typically only a few k_BT at room temperature). Therefore room temperature mobilities in molecular crystals only reach values in the range from 1 to 10 cm^2/Vs [36]. A power law temperature dependence of charge mobility is a characteristic feature of band transport:

$$\mu \propto T^{-n} \quad \text{with} \quad n = 1 \ldots 3 \tag{2.1}$$

and mobility decreases with increasing the temperature.

The other extreme case is an amorphous solid, where charge carriers are strongly localized. In this case charge transport can be described by hopping of charge carriers between localized states, which can be entire molecules or conjugated segments in case of polymers. Charge localization results in much lower mobility values (around $10^{-3}\,cm^2/Vs$, and lower). In the hopping regime the temperature dependence shows an activated behavior and depends on the applied electric field E:

$$\mu(E,T) \propto \exp\left[(-\alpha/k_BT)^2\right] \cdot \exp(\beta\sqrt{E}) \tag{2.2}$$

where α and β are numerical constants, k_BT is a thermal energy. Typical value of α is of the order of 0.1 eV. For β values of the order of $10^{-3}\,(cm/V)^{0.5}$ are observed [26]. Electric field and temperature dependence of the charge carrier mobility in the hopping regime are discussed later in this chapter.

An intermediate between the band and the hopping transport regimes corresponds to the situation when a charge carrier is spread over several neighboring molecules. In this case semi-classical dynamics (for details see [37, 38]) is a suitable method of description.

In this thesis the focus is on amorphous highly disordered materials (Alq_3), where charge transport occurs by thermally activated hopping. In order to better understand the microscopic picture of charge transport, the derivation of the Marcus rate equation in the high-temperature limit is outlined in the next section.

2.1 High temperature semi-classical Marcus theory

Let us consider a single electron hop from a molecule D (donor) to a molecule A (acceptor), which can be viewed as an electron transfer reaction:

$$D^- A \rightarrow D A^- \tag{2.3}$$

where D and A denote donor and acceptor states respectively. D^- denotes the reactant state with an excess electron localized on the donor. After the electron has moved to the acceptor the product state is formed.

Taking into account that in organic semiconductors the intermolecular interactions are weak, the donor $|D\rangle$ and acceptor $|A\rangle$ states can be approximated by the non-interacting molecular orbitals, or diabatic states, and the electronic Hamiltonian takes its *tight-binding* form [39]:

$$H^{el} = E_D |D\rangle \langle D| + E_A |A\rangle \langle A| + J (|D\rangle \langle A| + |A\rangle \langle D|) \tag{2.4}$$

where E_D, E_A are the energies of the individual states (sometimes referred to as *site energies*), and J is the electronic coupling (*transfer integral*) for the two states.

In order to describe charge transfer reactions which are coupled to the nuclear motion, it is helpful to introduce a *reaction coordinate* q, related to the actual positions of the nuclei, which connects the donor and acceptor states (see Fig. 2.1).

To simplify the derivation, all nuclear motions are treated classically. We also assume that the potential energy surfaces for the initial and final states are harmonic with identical curvature.[1] The Hamiltonian of Eq. 2.4 can then

[1] A more general treatment, where different curvatures of the initial and final states are allowed is also possible [41]

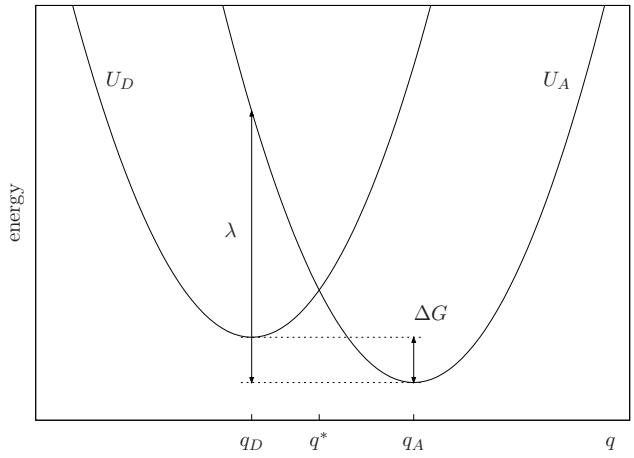

Figure 2.1: **Potential energy surfaces for a DA complex in a harmonic approximation.** The driving force (free energy difference) ΔG and the reorganization energy λ are indicated. Adapted from [40].

be rewritten as:

$$\begin{aligned} H^{el} = & |D\rangle \langle D| \left\{ E_D + \frac{1}{2}\omega_q^2 \left(q - q_D\right)^2 \right\} \\ & + |A\rangle \langle A| \left\{ E_A + \frac{1}{2}\omega_q^2 \left(q - q_A\right)^2 \right\} \\ & + J \left(|D\rangle \langle A| + |A\rangle \langle D| \right) \end{aligned} \tag{2.5}$$

where ω_q is the vibrational frequency of the mode promoting the charge transfer. Taking into account that the coupling J is small, it is possible to describe the electron transfer reaction within the framework of the perturbation expansion with respect to J where diabatic states (non-interacting donor and acceptor molecules) represent the zeroth-order Hamiltonian [42]. The first-order correction is then given by the Fermi's Golden Rule [43]:

$$k_{DA} = \frac{2\pi}{\hbar} \int dq\, p(q) |J^2| \delta \left(U_D(q) - U_A(q) \right) \tag{2.6}$$

where

$$\begin{aligned} U_D(q) &= E_D + \frac{1}{2}\omega_q^2 \left(q - q_D\right)^2 \\ U_A(q) &= E_A + \frac{1}{2}\omega_q^2 \left(q - q_A\right)^2 \end{aligned} \tag{2.7}$$

and the averaging is weighted by the canonical distribution of the nuclear positions:

$$p(q) \propto \exp\left[-U_D(q)/k_B T\right] \tag{2.8}$$

In the approximation of parabolic potential energy surfaces (PES), the integration can be carried out analytically, providing the expression for the charge transfer rate [44, 45, 46]:

$$k_{DA} = \frac{2\pi}{\hbar}|J|^2\sqrt{\frac{1}{4\pi k_B T\lambda}}\exp[-(\Delta G+\lambda)^2/4\lambda k_B T] \tag{2.9}$$

where λ is the so-called reorganization energy and it is related to the curvature of the parabolic PES. ΔG is the reaction driving force and is given by the difference of the PES minima (see Fig. 2.1). In a more general case, when the number of vibrational degrees of freedom is macroscopic, ΔG has to be understood as a free energy difference during the reaction. Eq. 2.9 is often referred to as a Marcus rate in the high-temperature limit.

From Eq. 2.9 it is clear that in order to calculate the charge hopping rate, one has to know several parameters: (1) electronic coupling element or transfer integral J, (2) reorganization energy λ and (3) $\Delta G = E_D - E_A$ - free energy difference between reactant and product states. E_D and E_A are normally referred to as site energies. Below we discuss how to extract these parameters from quantum chemical/classical calculations.

2.2 Reorganization energy

The reorganization energy is one of the key quantities that controls the rates for charge transfer. From Eq. 2.9 one can see, that in the normal regime ($|\Delta G| < \lambda$) the rate decreases exponentially with the increase of λ. Therefore, if high mobility is required for a particular application, compounds with low reorganization energies should be used.

Usually the reorganization energy is divided into two parts: inner and outer contributions. The inner (intramolecular) contribution arises from the change in the equilibrium geometry of the donor and acceptor molecules in the charge transfer reaction. The outer reorganization energy is due to the electronic and nuclear polarization/relaxation of the surrounding medium. In many cases these contributions are of the same order of magnitude [47]. Methods to estimate the outer reorganization energy were mainly developed to describe charge transfer in solutions [40], so it is desirable to extend these standard models to be able to calculate outer reorganization energy for a wide range of organic materials. In our calculations presented in Chapters 7 and 8 we neglected the outer reorganization energy.

Below, the intramolecular reorganization energy is defined in terms of vibronic modes. In order to understand separate contributions to λ, it is convenient to switch to a "monomer" picture. The PES of the donor and

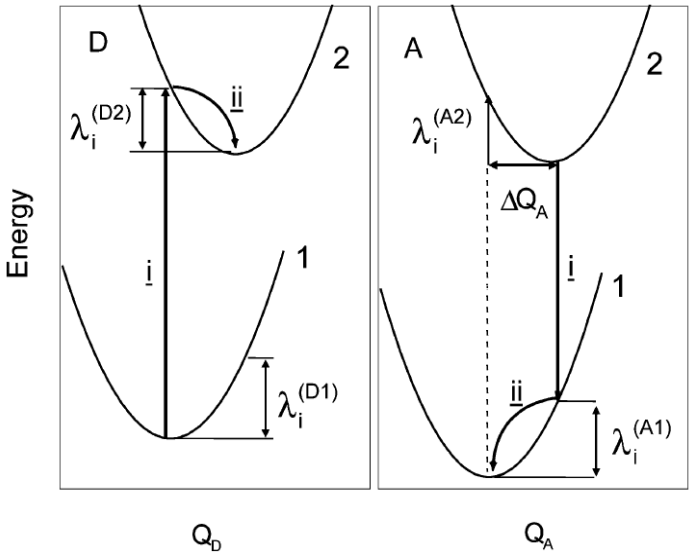

Figure 2.2: **Potential energy surfaces of a donor and an acceptor molecules related to charge transfer.** See text for details. Taken from [47].

acceptor molecules, involved in a transfer reaction $D + A^+ \rightarrow D^+ + A$ (hole transport) are shown separately in Fig. 2.2. Electronic states A1(D1) and A2(D2) correspond to the neutral and cationic states of the acceptor (donor), respectively. Charge transfer process can be formally divided into two steps: (1) simultaneous reduction of A^+ and oxidation of D at frozen reactant geometries, corresponding to the vertical transition from the minimum of $D1$ surface to $D2$ and a similar $A2$ to $A1$ transition; (2) relaxation of the product nuclear geometries.

Thus, the intramolecular reorganization energy consists of two terms [47] (see Fig. 2.2):

$$\lambda_i = \lambda_i^{(A1)} + \lambda_i^{(D2)} \tag{2.10}$$

with $\lambda_i^{(A1)} = E^{(A1)}(A^+) - E^{(A1)}(A)$ and $\lambda_i^{(D2)} = E^{(D2)}(D) - E^{(D2)}(D^+)$. Here $E^{(A1)}(A^+)$ and $E^{(A1)}(A)$ are the energies of the neutral acceptor A at the cation geometry and optimal ground-state geometry, respectively. $E^{(D2)}(D)$ and $E^{(D2)}(D^+)$ are the energies of the cation D^+ at the neutral geometry and optimal cation geometry.

Formula 2.10 allows to compute the reorganization energy of a cation (hole transport) based on four quantum chemical calculations. Similar procedure can be used to calculate the reorganization energy for an electron transport.

2.3 Transfer integrals

A number of computational techniques has been developed for calculation of electronic couplings, or transfer integrals, J_{ij} [48]. A widely used approach is to use Koopmans' theorem and to estimate the transfer integrals for holes (electrons) as half the splitting of the HOMO (LUMO) levels in a system made of two molecules in the neutral state [47]. In this approach a quantum chemical calculation has to be performed for each pair of neighboring molecules, making it computationally demanding for big systems. Also care must be taken, when splitting approach is used for asymmetric dimers. In such a situation, a part of the electronic splitting can simply arise from the different local environments experienced by the two interacting molecules, so additional correction terms are required [49].

Another approach, often called a projective method, relies on the projection of molecular orbitals of monomers onto the manifold of the molecular orbitals of the dimer within a Counterpoise basis set [50, 51]. Recently the relation between calculation parameters, (e.g. basis set, model chemistry) and associated computational costs for the projective method were systematically evaluated, finding that systems up to several thousand molecules can be treated on a DFT level [52].

An alternative approach to evaluate transfer integrals was reported by J. Kirkpatrick [53]. It is based on Zerner's Independent Neglect of Differential Overlap (ZINDO) Hamiltonian [54], which requires only a single self-consistent field calculation on an isolated molecule to be performed in order to determine the transfer integral for all pairs of molecules. The advantage of this method is that the density matrix for a pair of molecules is not calculated explicitly, but constructed based on the relative geometry of the two molecules and the transporting orbitals of the isolated molecules. Therefore, only a single ZINDO calculation for each type of molecule is needed for all pairs. Additionally, overlap integrals for atomic orbitals can be precalculated and stored, which further improves performance of the method. In this work the ZINDO method was used to calculate transfer integrals for systems containing up to 14.000 molecules, which is beyond the capabilities of the other methods.[2]

[2]It is practically impossible to estimate the accurracy of the ZINDO method *a priori*. In practice one can compare the results obtained by different levels of theory for a specific system to judge, whether the most efficient method still provides a reasonable accuracy.

2.4 Free energy difference

Hopping from a molecule i to another molecule j is driven by the free energy difference ΔG. One can see from Eq. 2.9, that the charge transfer rate depends exponentially on ΔG_{ij}, which makes the free energy term very important, especially in the case of materials with large energetic disorder, where the width of the distribution of site energies E_i can be of the order of 0.1 eV.

In general ΔG_{ij} can have several contributions. If charge transport is studied under the externally applied electric field $\boldsymbol{E}$, the corresponding contribution reads $\Delta G_{ij}^{ext} = e\boldsymbol{E}\boldsymbol{r}_{ij}$ where e is elementary charge and $\boldsymbol{r}_{ij}$ is the distance between molecules i and j.

If molecules have large dipole moments, the electrostatic contribution to ΔG might be significant [55]. The corresponding ΔG_{ij}^{el} arises from the interaction of the charge carrier with the surrounding dipoles. It can be calculated classically using partial charges for charged and neutral molecules in the ground state obtained from density functional theory calculations as described in [56].

Finally, if molecules are highly polarizable, electronic polarization must be taken into account in addition to the simple electrostatic picture. Electrostatic interactions between molecules induce dipole moments on them, which in turn produce an electric field, so the problem must be treated self-consistently [57].

When polymers are studied, charge carriers are localized not on the whole polymer chains, but on the so-called conjugated segments, which can have different lengths [58]. If the hopping occurs between conjugated segments of different lengths, the difference in HOMO (LUMO) levels must be also taken into account. Another term accounts for the energetic difference when the hopping occurs between chemically identical molecules with different conformations at finite temperatures, if the conformational disorder is significant. This contribution is ignored in this thesis.

2.5 Gaussian disorder model

Using Eq. 2.9 together with the techniques to calculate charge transport parameters (e.g. transfer integrals) mentioned in the previous sections, it is possible to study charge transport in realistic morphologies (based on molecular dynamics simulations) without any fitting parameters [59]. Before describing this type of simulations in detail it is useful to discuss the so-called Gaussian disorder model (GDM), in which charge transport parameters are not calcu-

lated, but obtained by fitting experimental data, or simply prescribed [60]. The Gaussian disorder model is a generic model for hopping transport. GDM (or its minor modifications) can explain experimentally observed transport characteristics, such as field and temperature dependence of charge carrier mobility, dispersive/non-dispersive transport. It captures all the relevant physics of hopping transport, that is why it is essential to understand GDM before performing simulations based on realistic morphologies.

In GDM charge carriers are assumed to be localized on the sites of a cubic lattice. Instead of Marcus rates (Eq. 2.9) the Gaussian disorder model uses Miller-Abrahams rates, originally used for inorganic semiconductors [61]:

$$k_{ij} = \nu_0 \exp(-2\gamma R_{ij}) \begin{cases} \exp\left(-\dfrac{\epsilon_j - \epsilon_i}{k_B T}\right) & \text{for } \epsilon_j > \epsilon_i \\ 1 & \text{for } \epsilon_j < \epsilon_i \end{cases} \tag{2.11}$$

Here ν_0 is a material-specific prefactor, R_{ij} is the separation between sites i and j, γ is the overlap factor, and ϵ_i and ϵ_j are the site energies. The first exponential term describes the decrease in electronic coupling with molecular separation, thus modelling the decay of the overlap of the wave functions of neighbors. The obvious simplification here is that it does not depend on the relative orientations of the molecules. The last term is the Boltzmann factor for an upward jump and is equal to 1 for a jump downward in energy. Energetic disorder is simulated by assigning site energies taken from the Gaussian distribution with variance σ:

$$p(\epsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right) \tag{2.12}$$

The "degree" of energetic disorder in the system is characterized by the dimensionless parameter $\hat{\sigma} = \sigma / k_B T$. Positional disorder is modelled by allowing the wave function overlap parameter, $\Gamma_{ij} = 2\gamma R_{ij}$, to fluctuate in a random manner. It is done by considering $\Gamma_{ij} = \Gamma_i + \Gamma_j$, each varying randomly according to the Gaussian probability density of standard deviation $\delta\Gamma$. The variance of Γ_{ij} is $\Sigma = \sqrt{2}\delta\Gamma$. It characterizes the "amount" of positional disorder in the system.

GDM was extensively studied by Bässler and coworkers using time-of-flight type (see Sec. 1.2.1) kinetic Monte Carlo simulations [60, 62, 63]. In the following, their main findings are briefly summarized.

2.5.1 Temperature dependence

At low electric fields, charge carrier mobilities were found to depend on temperature in a non-Arrhenius fashion:[3]

$$\mu(T) = \mu_0 \exp\left[-\left(\frac{2\sigma}{3k_BT}\right)^2\right] = \mu_0 \exp\left[-\left(\frac{T_0}{T}\right)^2\right] \tag{2.13}$$

with σ representing the standard deviation of the site energy distribution from Eq. 2.12. This expression provides a way to extract the energetic disorder parameter σ from an experimentally measured temperature dependence. Note that positional disorder Σ does not affect temperature dependence of charge mobility. It is solely defined by the energetic disorder parameter σ.

2.5.2 Field dependence

When energetic disorder is not present or very small ($\hat{\sigma} \lesssim 2$), charge mobility decreases with the increase of electric field. For ($\hat{\sigma} \gtrsim 2$) the opposite behavior is observed. In all cases, the field dependence of the mobility approaches a $\log \mu \propto \beta E^{1/2}$ law at moderately high fields ($\gtrsim 7 \times 10^5\,\mathrm{V/cm}$), known as Poole-Frenkel behavior [65]. Poole-Frenkel dependence is routinely observed in time-of-flight measurements for amorphous organic compounds. However, experimentally it is also observed at lower electric fields. To obtain Poole-Frenkel behavior at low fields, site energy *spatial correlations* must be taken into account [55, 66]. Fundamental question about the origin of Poole-Frenkel behavior in general is still open [67, 34, 68].

2.5.3 The nondispersive to dispersive transition

As already discussed in Sec. 1.2.1, charge transport becomes dispersive when the time required for a packet of carriers to reach the dynamic equilibrium becomes comparable to the transient time. Under these conditions current transients do not show a plateau region and extracted carrier mobilities depend on the system size. A dispersive - nondispersive transition was observed in GDM simulations, in agreement with the experimental data [69]. An empirical relationship between the critical temperature T_c (equivalent to the critical disorder parameter $\hat{\sigma}_c$) and the sample thickness was also established:

$$\hat{\sigma}_c^2 = 44.8 + 6.7 \log L \tag{2.14}$$

[3] This formula was obtained by fitting the experimental data. Analytical results for one-dimensional system show that the real dependence is slightly more complicated, although the leading term $\exp\left[-\left(\frac{T_0}{T}\right)^2\right]$ is correct [64].

with L being the sample length (in dimensionless cm). For a given sample thickness L, transport is dispersive for $\hat{\sigma} > \hat{\sigma}_c$ and nondispersive otherwise.

2.5.4 Expression for the mobility

From the results of Monte Carlo simulations, the general behavior of the mobility as a function of both temperature and electric field in the presence of both positional and energetic disorder is given by:

$$\mu(\hat{\sigma}, \Sigma, E) = \mu_0 \exp\left[-\left(\frac{2}{3}\hat{\sigma}\right)^2\right] \times \begin{cases} \exp[C(\hat{\sigma}^2 - \Sigma^2)E^{1/2}] & \Sigma \geq 1.5 \\ \exp[C(\hat{\sigma}^2 - 2.25)E^{1/2}] & \Sigma < 1.5 \end{cases} \tag{2.15}$$

where C is an empirical constant. This formula is obtained by analyzing the simulation data and is widely used to analyze experimental results.

2.6 Role of site energy spatial correlations

As already mentioned in Sec. 2.5.2, GDM fails to reproduce Poole-Frenkel behavior of charge carrier mobility at low fields ($\lesssim 7 \times 10^5$ V/cm), routinely observed in experiments [70]. The reason is that site energies in GDM are fully random and uncorrelated, which appears to be too severe an approximation for many systems, especially in the case when molecules have permanent dipole moments. The physical reason for spatial correlations is the long-range nature of the dipole's electrostatic potential. It was shown by S. Novikov and V. Vannikov, that the electrostatic potential in a system of randomly oriented dipoles is strongly correlated [71] , see Fig. 2.3(a). This holds true even when the dipoles are orientationally uncorrelated. As a consequence, site energies for this system are also correlated and the correlation function decays slowly with the intersite separation r [55]:

$$C(r) = \langle \epsilon(0)\epsilon(r) \rangle \sim \sigma_d^2 a/r \tag{2.16}$$

where a is a minimal charge-dipole separation and σ_d is the rms width of the dipolar energetic disorder. An empirical relation was obtained for the field dependence of the non-dispersive mobility in correlated (e.g. dipolar) media [66]:

$$\mu = \mu_0 \exp\left[-\left(\frac{3\hat{\sigma}_d}{5}\right)^2 + C_0\left(\hat{\sigma}_d^{3/2} - \Gamma\right)\sqrt{\frac{eaE}{\sigma_d}}\right] \tag{2.17}$$

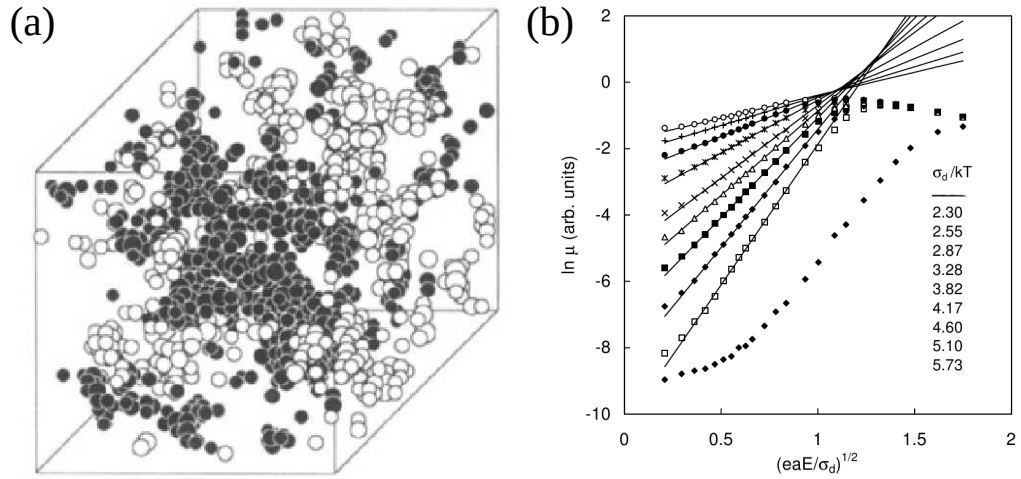

Figure 2.3: **Role of site energy correlations.** (a) Distribution of electrostatic potential in a finite sample ($31 \times 31 \times 31$) of point dipoles. Black and white spheres represent sites with positive and negative values of electrostatic potential ϕ respectively. Taken from [71]; (b) Results of the correlated disorder model (CDM) simulations for different $\hat{\sigma}_d$ (from top curve downward). The lowest curve is the mobility for the GDM for $\hat{\sigma} = 5.10$. Taken from [66].

where $\hat{\sigma}_d = \sigma_d / k_B T$, $C_0 = 0.78$, and $\Gamma = 2$. Analogous to the Gaussian disorder model, this model is usually referred to as *correlated* disorder model (CDM). The difference in field dependence of charge mobility between GDM and CDM is illustrated in Fig. 2.3(b).

One can understand without complicated calculations why spatial correlations of site energies enhance Poole-Frenkel behavior at low fields. Physically, a strong field dependence should occur when the potential drop $\delta U = eEl$ across a relevant length of the system is comparable to $k_B T$. With uncorrelated energies the only length scale in the problem is the average distance between the hopping sites. Correlations introduce a new length scale, namely, the correlation length associated with energetic disorder, thereby decreasing the critical field [55].

2.7 Charge transport in realistic morphologies

In spite of the success of the GDM (and similar models) in describing hopping transport, this model does not have predictive power. In order to study a specific material, parameters for the model (e.g. energetic disorder parameter σ) must be obtained by fitting experimental data. In that sense, they

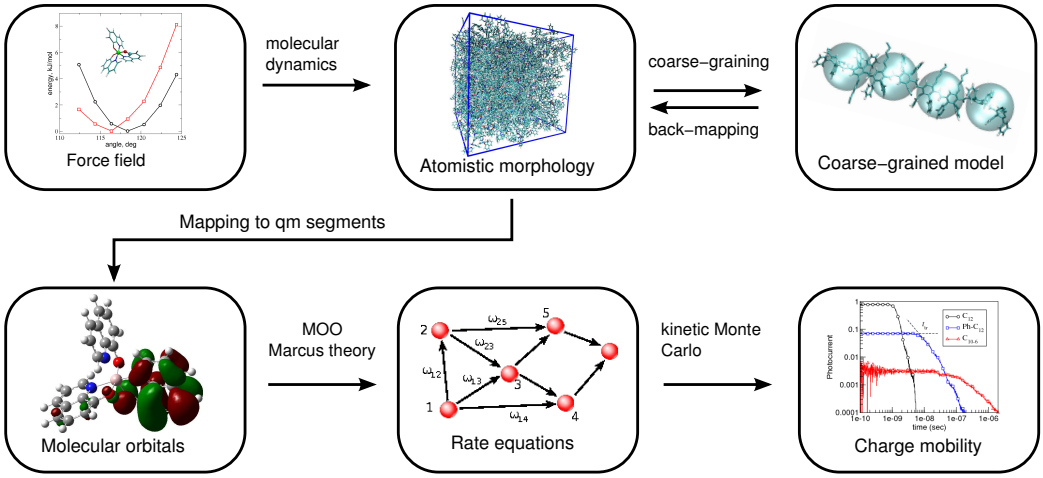

Figure 2.4: **Outline of the multiscale methodology.**

are merely adjustable parameters, without any microscopic meaning. Miller-Abrahams rates used in the GDM only depend on the distance between two molecules, but not on their mutual orientation, making it impossible to study how the morphology of material affects charge transport. In order to understand the effect of morphology and to relate charge transport properties to the underlying chemical structure, a multiscale approach should be used, where charge hopping rates are calculated based on the realistic morphologies. In this thesis the methodology is used, which was initially developed to study charge transport in discotic liquid crystals and avoids using fitting parameters and regular grids [59, 72, 56, 73, 74].

Realistic morphologies are obtained by means of atomistic molecular dynamics simulations [75]. Since for organic molecules force-fields are not readily available, force-field parameterization should be done starting from the force-field parameters for similar compounds and quantum chemical calculations [1, 76, 77]. The developed force-field must be validated, for example, by comparing structural and thermodynamical properties extracted from the simulations to experimental data [3]. This is illustrated in Chapter 3. In some cases, length- and time-scales of atomistic simulations are not sufficient to equilibrate the system. For example, this is the case for polymer melts, where equilibration time scales with the third power of the chain length $\tau \propto N^3$ [78]. Coarse-graining techniques help to overcome length- and time-scale limitations of atomistic simulations by reducing the number of degrees of freedom representing a system of interest [79]. Coarse-graining methods and their limitations are covered in detail in Chapters 4 and 5, where we also compare the performance of different coarse-graining techniques for the model

systems.

When a realistic morphology is available, parameters required to obtain Marcus hopping rates (transfer integrals, site energies) can be calculated as described in Sec. 2.3 and Sec. 2.4. Together with reorganization energies from DFT calculations (Sec. 2.2), these parameters are used to obtain charge hopping rates for each pair of neighboring molecules, using semiclassical Marcus theory (Eq. 2.9). Once the hopping rates are known, one can treat each molecule as a structureless hopping site, located at the molecule's center of mass.

Finally, a kinetic Monte Carlo algorithm is used to simulate charge dynamics and calculate charge carrier mobilities. An outline of the methodology is given in Fig. 2.4.

Chapter 3

Morphology simulations

3.1 Molecular dynamics simulations

Molecular dynamics (MD) simulations describe evolution of molecular systems in a classical limit, meaning that atomic nuclei are treated as point particles, while electronic degrees of freedom are not taken into account explicitly. Instead interactions between nuclei are described by a potential energy function (usually referred to as a force field), which implicitly represents forces arising from electronic degrees of freedom.

The basic functional form of a force field encapsulates both bonded terms relating to atoms that are linked by covalent bonds, and nonbonded (also called "noncovalent") terms describing the long-range electrostatic and van der Waals forces. All atomistic simulations in this work are based on all-atom OPLS force field [80], which has the following form:

$$\begin{aligned} U\left(\{\vec{r}_i\}\right) = & \sum_{bonds} \frac{1}{2} k_b \left(r - r_0\right)^2 + \frac{1}{2} \sum_{angles} k_\theta \left(\theta - \theta_0\right)^2 \\ & + \sum_{torsions} \left\{ \frac{V_1}{2}[1 + \cos(\varphi)] + \frac{V_2}{2}[1 - \cos(2\varphi)] \right. \\ & \left. + \frac{V_3}{2}[1 + \cos(3\varphi)] \right\} + \sum_{impropers} k_\gamma \left(\gamma - \gamma_0\right)^2 \\ & + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right\}, \end{aligned} \tag{3.1}$$

First two sums in this expression account for the bond stretching (2-particle interactions) and angle bending (3-particles). The third sum represents torsion angles (4-particles), which are modeled by the first four terms of Fourier

series. Torsions are periodic functions of the angle and are used to describe rotation around the chemical bonds. The fourth sum represents improper dihedral angles (4-particles), which are introduced to keep atoms in a plane (e. g. aromatic rings), or to prevent molecules from flipping over to their mirror images. The last sum over all pairs represents non-bonded interactions, which are modeled by Lennard-Jones (steric repulsion and London dispersion) and Coulomb (electrostatics) terms.

To propagate the system in the phase-space, Newton's equations of motions are integrated numerically. From the point of view of statistical mechanics, this corresponds to the microcanonical ensemble, since total energy of the system is conserved. It is possible, however, to simulate other ensembles of statistical mechanics, for example, NVT or NPT [81, 82, 83, 84, 85].

For more information about molecular dynamics simulations the reader is referred to the textbooks [86, 75, 87].

In this thesis molecular dynamics simulations are used to obtain realistic morphologies of organic materials. In the following section we will discuss how such force field can be obtained for a typical organic molecule.

3.2 Force field development

Since force fields for novel organic compounds are generally not available, they must be developed for every compound of interest. Force field development is an area of research in itself, and many efforts have been undertaken to automate the parameterization procedure [80, 88, 89]. In this section a simplified procedure is described, which was used in the course of this thesis to parameterize force fields for conjugated polymers poly(2,3-diphenylphenylene vinylene)(DP-PPV) [3], poly[2,6-(4,4-bis-(2-ethylhexyl)-4H -cyclopenta[2,1-b;3,4-b]-dithiophene)-alt -4,7-(2,1,3-benzothiadiazole)]
(PCPDTBT) as well as for tris(8-hydroxyquinoline)aluminium(Alq_3) [1].

3.2.1 DP-PPV force field as an example

In this section a typical parameterization procedure is described using DP-PPV as an example. DP-PPVs have been considered as a family of green-emitting materials for polymer LED applications due to their good mechanical and optical properties [90, 91, 92]. Chemical structure of DP-PPV is shown in Fig. 3.1a.

As a starting point, we use the OPLS all-atom force-field [80]. Parameters for bonds, angles as well as van der Waals parameters for non-bonded interactions are taken from this force-field. Partial charges and missing bonded

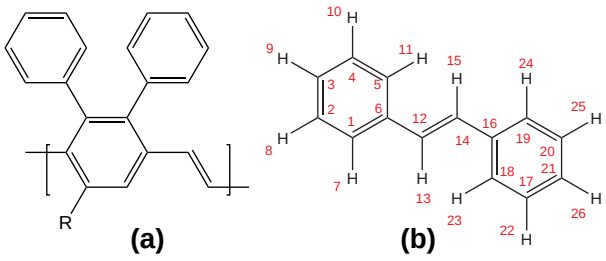

Figure 3.1: (a) Chemical structure of DP-PPV derivatives. R = C_6H_{13} corresponds to DP6-PPV and R = $C_{10}H_{21}$ to DP10-PPV. (b) Trans-stilbene - a monomer of poly-(phenylene vinylene) used to validate the re-parametrized atomistic force-field.

Level of theory	atom No. 6	atom No. 12
B3LYP 6-31G(d,p)	0.227	-0.187
B3LYP 6-311G	0.278	-0.241
B3LYP 6-311G(d)	0.295	-0.248
B3LYP 6-311G(d,p)	0.283	-0.237
B3LYP 6-311+G(d,p)	0.294	-0.243
B3LYP 6-311++G(d,p)	0.301	-0.241
B3LYP 6-311++G(2d,2p)	0.264	-0.216
B3LYP 6-311G(2df,2pd)	0.254	-0.210
MP2 6-31G(d,p)	0.259	-0.222

Table 3.1: Partial charges of atoms 6 and 12 as a function of the basis set size. Atom labeling is shown in Fig. 3.1b.

interactions are determined using first principles calculations [1, 77, 72]. The force-field parametrization is then verified by simulating several thermodynamic properties of trans-stilbene, whose chemical structure is shown in Fig. 3.1b.

Partial atomic charges were calculated using the CHELPG procedure [93]. For geometry optimization we used hybrid DFT functional B3LYP [94] as well as Møller-Plesset second order perturbation theory (MP2). To illustrate the basis set convergence, the charges of the atoms 6 and 12 (see Fig. 3.1b) are listed in Table 3.1 as a function of the basis set. One can see that for small basis sets, the variation is about 20%. Saturation is achieved for a rather large basis set, 6-311G++(2d,2p). The DFT values agree well with MP2 calculations, especially for large basis sets. To assess the values of partial charges in a polymer, we have also performed calculations for tetra- and octamers. No significant variations were found.

dihedral	ϕ_0, deg	k_ϕ, kJ/mol
1-6-12-14	0	7
6-12-14-16	0	30
6-1-5-12	0	270

Table 3.2: Dihedral parameters. See Fig. 3.1b and Fig. 3.2 for notations.

3.2.1.1 Parametrization of backbone dihedrals

To refine the force-field parameters for dihedral angles, we first considered the backbone without the side chains as shown in Fig. 3.1b. Three dihedral potentials, which are not present in the OPLS force-field, determine the rigidity and conformation of the backbone. To obtain parameters for these potentials, the angle of interest was scanned by optimizing the molecular geometry for a fixed value of the dihedral. The scan provides a set of optimized molecular structures and total energies for each angle value. Subsequently, the energy of each optimized conformation was evaluated with the help of the force-field, where the dihedral of interest was switched off. To do this, the molecular geometry was again optimized for each value of the constrained dihedral angle and the difference between the two energies was fitted, providing the desired dihedral parameters [95].

For the dihedrals (1-6-12-14) and (6-12-14-16), which describe rotation around the bonds (see Fig. 3.1b), the functional form given by Eq. (3.2) was used, while for the improper dihedral (6-1-5-12), which keeps the atoms in plane, the functional form given by Eq. (3.3) was used

$$V = k_\phi \left[1 + \cos(2\phi - \phi_0)\right], \tag{3.2}$$

$$V = \frac{1}{2} k_\phi \left(\phi - \phi_0\right)^2, \tag{3.3}$$

where ϕ_0 is the equilibrium angle and k_ϕ is the fitted force constant.

The results of fitting are shown in Fig. 3.2. For the first dihedral, (1-6-12-14), different levels of theory provide different equilibrium values of the dihedral angle. MP2 calculations suggest that the ground state of trans-stilbene is nonplanar contrary to the DFT calculations. In fact, the discrepancy between these methods is a known issue. A more detailed study of trans-stilbene shows that it is planar and the value of the torsional barrier is 14.3 kJ/mol [96]. This value was used for fitting. The results for all three dihedrals are summarized in Table 3.2.

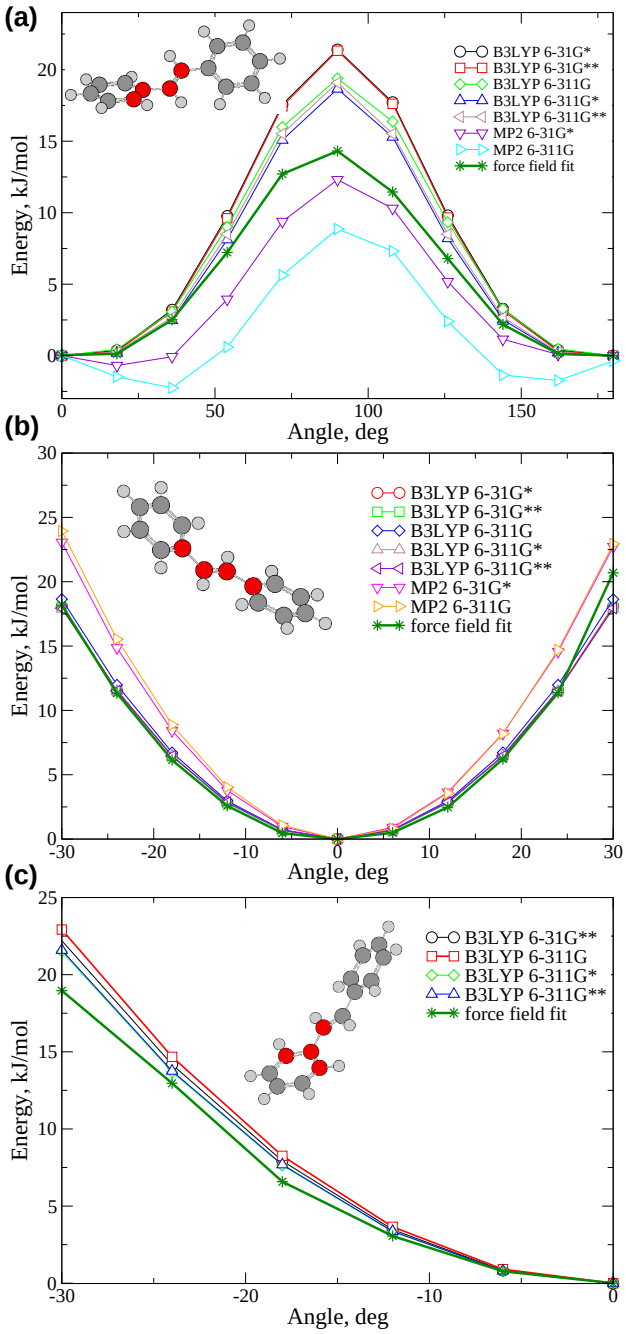

Figure 3.2: Energies calculated using first principle methods as well as fitted force-field potentials for the dihedrals: (a) 1-6-12-14 (b) 6-12-14-16 (c) 6-1-5-12. The scanned dihedrals are depicted in the insets. Different methods and basis sets are shown.

	Experiment [97]	MD simulations
a	12.287 ± 0.003	12.09 ± 0.06
b	5.660 ± 0.003	5.38 ± 0.06
c	15.478 ± 0.005	16.9 ± 0.2
β, deg	112.03 ± 0.1	110.0 ± 0.2

Table 3.3: Monoclinic unit cell parameters of trans-stilbene. All distances are given in Å.

	C_0	C_1	C_2	C_3	C_4	C_5
(i)	8.60	0.0	-30.81	0.0	21.66	0.0
(ii)	-4.22	-0.027	9.44	0.48	-5.15	-0.47

Table 3.4: Ryckaert-Belleman parameters for (i) the dihedral linking two phenyl rings and (ii) the dihedral connecting the backbone phenyl ring with the alkyl side chain. All constants are in kJ/mol.

3.2.1.2 Force-field validation

To validate the force-field, we compared the dimensions of the simulated and experimentally measured unit cell of trans-stilbene crystal and its melting temperature.

The monoclinic unit cell of trans-stilbene [97] was multiplied as $2a \times 4b \times 2c$ to be able to use 0.9Å cutoff distance for van der Waals interactions. After energy minimization with the conjugate gradient method [98], a 200 ps molecular dynamics run in the NPT ensemble (anisotropic Berendsen thermostat [81], $P = 1\,\text{bar}, T = -160°\text{C}$) was performed. After equilibration, an NPT production run of 600 ps was performed. The simulated density of trans-stilbene was $1161\,\text{kg/m}^3$, which is in a good agreement with the experimental value of $1200\,\text{kg/m}^3$ as well as the crystallographic parameters given in Table 3.3.

To simulate the crystal melting, we performed a simulated annealing run, increasing the temperature from $-160°$C to 180°C during 1200 ps (heating rate 0.283°C/ps). While monitoring the mean squared displacement and density of the compound, as shown in Fig. 3.3. We concluded the melting point to be $127 \pm 25°$C. To reduce the error bars, a set of 200 ps NPT simulations at 102°C, 112°C, 122°C, and 132°C were performed. Up to 122°C, the system remains in the crystalline state, melting completely at 132°C. Our predicted melting point of $127 \pm 5°$C agrees well with the experimental value of $124 \pm 1°$C. To ensure that the system size and the heating rate do not affect the results, we also annealed a $4a \times 6b \times 4c$ cell. The results are shown in Fig. 3.3, indicating that there are no significant finite size effects. Within

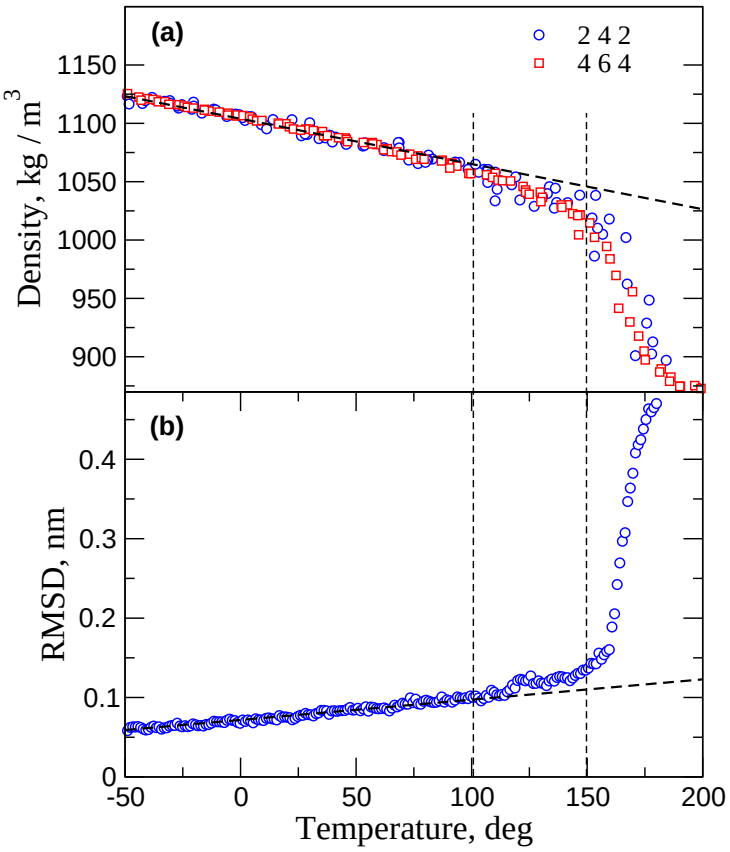

Figure 3.3: (a) Density as a function of temperature during a 1200 ps simulated annealing run. Two system sizes are shown: $2a \times 4b \times 2c$ and $4a \times 6b \times 4c$. The results suggest that the melting point of trans-stilbene is 127 ± 25°C. (b) Root mean square deviation from the equilibrium crystalline structure as a function of temperature. 1200 ps simulated annealing. Melting occurs at 1020 ± 80 ps, which corresponds to 127 ± 25°C.

the range of $0.07 - 0.3\,\mathrm{deg/ps}$, we did not observe any dependence on the heating rate.

In summary, we can conclude that the performance of our force-field for trans-stilbene is adequate.

3.2.1.3 DP-PPV force-field

To derive the force-field for DP6-PPV and DP10-PPV, we followed the same strategy. We first calculated the partial charges of a DP-PPV monomer unit and then parametrized two additional dihedral potentials. The first one, linking phenyl rings to the backbone and the second one, connecting the backbone phenyl ring and the alkyl side chain. The Ryckaert-Belleman functional form [99] was used to parametrize these two dihedrals

$$V_{rb}(\phi) = \sum_{n=0}^{5} C_n \cos^n \phi, \tag{3.4}$$

where $\phi = 0$ corresponds to the trans-conformation. The obtained constants, C_n, are given in Table 3.4. For the alkyl side chains, we used the OPLS united atom force-field [80].[1]

Derived force field was used to study conformational properties and aggregation behavior of DP-PPV in organic solutions [3].

3.3 Limitations of MD simulations

All-atom molecular dynamics simulations provide key insight into the structure and dynamics of soft matter systems by providing a model of molecular motion with angstrom level detail and femtosecond resolution. However, all-atom simulations are limited to the nanometer length- and nanosecond time-scales, even for the most powerful modern hard- and software [100]. This limitation may be crucial, for example if one wants to simulate an equilibrium morphology of a polymer melt. Equilibration time for a melt scales with $\tau \propto N^3$, where N is the backbone length [78]. To access larger simulated time scales and system sizes, a further simplification of the molecular model is needed. Systematic coarse-graining methods, which are discussed in the next chapter, help to overcome these limitations. Combined with efficient backmapping procedures [101, 102], which allow to reintroduce back

[1] Since the side chains are practically neutral, this approximation should not affect the conformational structure of the backbone significantly. If charge transport is studied, side chains can also be neglected, because the charge is normally delocalized over the conjugated backbone [59].

the atomistic details, coarse-graining methods can be used to simulate large scale morphologies [101, 103, 104].

Chapter 4

Coarse-graining techniques

4.1 Introduction

Computational materials science deals with phenomena covering a wide range of length- and time-scales, from Ångstrøms (typical bond lengths) and femtoseconds (bond vibrations) to micrometers (crack propagation) and milliseconds (a single polymer chain relaxation). Depending on the characteristic time- and length-scales involved, the system description can vary from first principles and atomistic force-fields to coarse-grained models and continuum mechanics. The role of bottom-up coarse-graining, in a broad sense, is to provide a systematic link between these levels of description.

This chapter is focused on coarse-graining techniques that link two *particle-based* descriptions with a different number of degrees of freedom. The system with the larger number of degrees of freedom is denoted as the *reference* system. The system with the reduced number of the degrees of freedom is referred to as the *coarse-grained* system. An example is an all-atom (reference) and a united atom (coarse-grained) molecular representations, where the number of the degrees of freedom is reduced by embedding hydrogens into heavier atoms. Another example, which is treated in detail here, is an all-atom (three sites) and a single site model of water. Other examples can be readily found in the literature [103, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115].

It is assumed that the following prerequisites are satisfied:

- Both the reference and the coarse-grained descriptions are represented by a set of point sites, $\boldsymbol{r} = \{\boldsymbol{r}_i\}$, $i = 1, 2, \ldots, n$ in case of the reference system, and $\boldsymbol{R} = \{\boldsymbol{R}_j\}$, $j = 1, 2, \ldots, N$ in case of the coarse-grained system.

- A mapping scheme, i.e. a relation between $\boldsymbol{r}$ and $\boldsymbol{R}$, can be expressed as $\boldsymbol{R} = \hat{M}\boldsymbol{r}$, where $\hat{M}$ is an $n \times N$ matrix.

- For the reference system, trajectory that samples a canonical ensemble is available.

Then the prime task of systematic coarse-graining is to derive a potential energy function of the coarse-grained system, $U(\boldsymbol{R})$.

To do this, one can use several coarse-graining approaches. From the point of view of implementation, these approaches can be divided in *iterative* and *non-iterative* methods. *Boltzmann inversion* is a typical example of a non-iterative method [103]. In this method, which is exact for *independent* degrees of freedom, coarse-grained interaction potentials are calculated by inverting the distribution functions of the coarse-grained system. Another example of a non-iterative method is *force matching*, where the coarse-grained potential is chosen in such a way that it reproduces the forces on the coarse-grained beads [116, 108]. Configurational sampling [117], which matches the potential of mean force, also belongs to this category. Boltzmann inversion and force matching only require a trajectory for a reference system. Once that is known, coarse-grained potentials can be calculated for any mapping matrix $\hat{M}$. Note that this is often a "special" trajectory which is designed to decouple the degrees of freedom of interest, e. g. a single polymer chain in vacuum with appropriate exclusions [103].

Iterative methods refine the coarse-grained potential $U(\boldsymbol{R})$ by re-iterating coarse-grained simulations and calculating corrections to the potential on the basis of the reference and coarse-grained observables (e. g. radial distribution function or pressure). The simplest example is the iterative Boltzmann inversion method [118], which is an iterative analogue of the Boltzmann inversion method. More sophisticated (in terms of the update function) is the inverse Monte Carlo approach [119].

One can also classify systematic coarse-graining approaches by micro- and macroscopic observables used to derive the coarse-grained potential, such as structure-based [119, 120, 103], force-based [116, 121, 108], and potential-based approaches [122], where the name identifies the quantities used for coarse-graining. Note that hybrids of these methods are also possible [106, 115].

With a rich zoo of methods plus their combinations available at hand, it is natural to ask about an optimal method for a specific class of systems. On a more fundamental level one might question whether the different methods provide the same coarse-grained potential and whether it is possible to formulate a set of (even empirical) rules favoring one method with respect

to another. It is obvious this is a difficult task to be treated analytically, especially for realistic systems. To assess the quality of a particular coarse-graining technique one needs to apply all available methods to a certain number of systems and to compare and quantify the degree of discrepancy between the coarse-grained and reference descriptions.

This is, however, cumbersome due to the absence of a single package where all these methods are implemented with the same accuracy and same level of technical detail. That is why we started to implement such a package, the Versatile Object-oriented Toolkit for Coarse-graining Applications (VOTCA) [2] in collaboration with V. Rühle and C. Junghans. The author of this thesis was focused on implementation of the force matching method.

In the following sections the basic ideas behind each method are described, paying special attention to the technical issues one has to overcome when implementing them. Then the methods are illustrated by coarse-graining systems of different complexity: a three-site SPC/E water, methanol, propane, and hexane.

For more information the reader is referred to a (far from complete) list of reviews which cover various aspects of generating coarse-grained potentials [123, 124, 125, 126, 79, 127, 102].

4.2 Boltzmann inversion

Boltzmann inversion is the simplest method one can use to obtain coarse-grained potentials [103]. It can only be used for *bonded* potentials, such as bonds, angles, and torsions. Boltzmann inversion is structure-based and only requires positions of atoms.

The idea of Boltzmann inversion stems from the fact that in a canonical ensemble *independent* degrees of freedom q obey the Boltzmann distribution, i. e.

$$P(q) = Z^{-1} \exp\left[-\beta U(q)\right] \ , \tag{4.1}$$

where $Z = \int \exp\left[-\beta U(q)\right] \mathrm{d}q$ is a partition function, $\beta = 1/k_{\mathrm{B}}T$. Once $P(q)$ is known one can invert eq. 4.1 and obtain the coarse-grained potential, which in this case is a potential of mean force:

$$U(q) = -k_{\mathrm{B}}T \ln P(q) \tag{4.2}$$

Note that the normalization factor Z is not important since it would only enter the coarse-grained potential $U(q)$ as an irrelevant additive constant.

In practice, $P(q)$ is computed from the trajectory of the reference system which is sampled either by Monte Carlo, molecular dynamics, stochastic

dynamics, or any other integrator that ensures a canonical distribution of states.

Boltzmann inversion is simple to implement, however one has to be careful with the rescaling of the probability P due to orientational entropy as well as computational issues. The probability rescaling can be explained on a particular example of coarse-graining of a single polymer chain by beads with bond, angle and torsion potentials. In this case the coarse-grained potential U depends on three variables, bond length $\boldsymbol{r}$, angle θ and torsion angle φ.

Assuming, as before, a canonical distribution and independence of the coarse-grained degrees of freedom, one can write

$$\begin{aligned} P(r,\theta,\varphi) &= \exp\left[-\beta U(r,\theta,\varphi)\right] , & (4.3)\\ P(r,\theta,\varphi) &= P_r(r)P_\theta(\theta)P_\varphi(\varphi) . & (4.4) \end{aligned}$$

If now the histograms for the bonds $H_r(r)$, angle $H_\theta(\theta)$, and torsion angle $H_\varphi(\varphi)$ are computed, they must be rescaled in order to obtain the volume normalized distribution functions:

$$P_r(r) = \frac{H_r(r)}{4\pi r^2} , P_\theta(\theta) = \frac{H_\theta(\theta)}{\sin\theta} , P_\varphi(\varphi) = H_\varphi(\varphi) . \quad (4.5)$$

The coarse-grained potential can then be calculated by Boltzmann inversion of the distribution functions:

$$\begin{aligned} U(r,\theta,\varphi) &= U_r(r) + U_\theta(\theta) + U_\varphi(\varphi) , & (4.6)\\ U_q(q) &= -k_\mathrm{B}T \ln P_q(q),\ q = r,\theta,\varphi . \end{aligned}$$

On the technical side, the implementation of the Boltzmann inversion method requires *smoothing* of $U(q)$ to provide a continuous force. Splines can be used for this purpose. Poorly and unsampled regions, that is regions with high $U(q)$, shall be *extrapolated*. Since the contribution of these regions to the canonical density of states is small the exact shape of the extrapolation is less important.

Another crucial issue is the cross-correlation of the coarse-grained degrees of freedom. Independence of the coarse-grained degrees of freedom is the main assumption that allows factorization of the probability distribution, eq. 4.4, and the potential, eq. 4.6, hence, one has to carefully check whether this assumption holds in practice. This can be done by performing coarse-grained simulations and comparing cross-correlations for all pairs of degrees of freedom in atomistic and coarse-grained resolution, e. g. using a two-dimensional histogram, analogous to a Ramachandran plot. Note, that checking the linear correlation coefficient does not guarantee statistical independence of variables, for example $c(x, x^2) = 0$ if x has a symmetric

probability density $P(x) = P(-x)$. This case is often encountered in systems used for coarse-graining [128, 129]. The concept is illustrated in sections 5.3 and 5.4 for liquid propane and a single molecule of hexane.

4.3 Iterative Boltzmann inversion

Iterative Boltzmann inversion (IBI) is a natural extension of the Boltzmann inversion method. Since the goal of the coarse-grained model is to reproduce the distribution functions of the reference system as accurately as possible, one can also iteratively refine the coarse-grained potentials using some numerical scheme. Depending on the update function, this can be done either by using iterative Boltzmann inversion [118] or inverse Monte Carlo [119, 120] method.

In the iterative Boltzmann inversion, the coarse-grained potential is refined according to the following scheme:

$$\begin{aligned} U^{(n+1)} &= U^{(n)} + \Delta U^{(n)} \ , \\ \Delta U^{(n)} &= k_{\mathrm{B}} T \ln \frac{P^{(n)}}{P_{\mathrm{ref}}} = U^{\mathrm{ref}}_{\mathrm{PMF}} - U^{(n)}_{\mathrm{PMF}} \ . \end{aligned} \tag{4.7}$$

One can easily see that convergence is reached as soon as the distribution function $P^{(n)}$ matches the reference distribution function P_{ref}, or, in other words, the potential of mean force, $U^{(n)}_{\mathrm{PMF}}$ converges to the reference potential of mean force.

Note that eq. 4.7 is nothing else but a numerical scheme that allows one to match the coarse-grained and the reference distribution functions. It can be seen as a first-order correction to the interaction potential with respect to a gas of non-interacting particles. Indeed, in an ideal gas, the probability of finding two particles at a distance r is $P^{(0)} \propto 4\pi r^2$, which is equivalent to the statement that the radial distribution function of an ideal gas is 1. Substituting $P^{(0)}$ into eq. 4.7 one obtains the first iteration $U^{(1)} = -k_{\mathrm{B}} T \ln(P_{\mathrm{ref}}/4\pi r^2)$, which is the potential of mean force, eq. 4.2.

IBI can be used to refine both bonded and non-bonded potentials. It is primarily used for simple fluids with the aim of reproducing the radial distribution function of the reference system in order to obtain non-bonded interactions [118]. It can have convergence problems for multicomponent systems, since it does not account for cross-correlation correction terms, that is the updates for P_{AA}, P_{AB}, and P_{BB} are not coupled (the subscript enumerates a single component in a multicomponent system). For such systems, the

inverse Monte Carlo method works better. The scheme can be stabilized by multiplying the update function, $\Delta U^{(n)}$, by a factor $\eta \in [0..1]$.

On the implementation side, IBI has the same issues as the inverse Boltzmann method, i. e. smoothing and extrapolation of the potential must be implemented.

One should also mention that, according to the Henderson theorem [130, 127], which is a classical analogue of the Hohenberg-Kohn theorem, the pairwise coarse-grained potential $U(r)$ is unique up to an additive constant and exists [131, 132], which, in principle, states that all structure-based iterative methods must converge to the same coarse-grained potential, provided that their aim is to exactly reproduce pair correlation functions of the reference system. However, this is often not the case in practice (see Sec. 5.2), since small changes in the radial distribution function often lead to big changes in the pair potential, i. e. it is difficult to control systematic errors during the calculation of the potential update.

Another issue of coarse-graining is that coarse-grained models cannot reproduce all the statistical or thermodynamic properties of the reference system. Pressure, compressibility, or viscosity [133] are often very different from those of the reference system. In some cases, however, one can correct for some of these. For example, the viscosity can be adjusted by tuning the parameters of the thermostat [85] and the pressure can be corrected iteratively by adding a linear term to the non-bonded potential:

$$\Delta U^{\text{pressure}} = A \left(1 - \frac{r}{r_{\text{cut}}} \right), \tag{4.8}$$

where A is either a constant, e. g. $-0.1\, k_{\text{B}}T$ [118], or can be estimated from the virial expansion [134]. Compressibility and pressure, however, cannot be corrected simultaneously.

4.4 Inverse Monte Carlo

Inverse Monte Carlo (IMC) is another iterative procedure that refines the coarse-grained potentials until the coarse-grained model reproduces a set of reference distribution functions. It is very similar to IBI except that the update of the potential, ΔU, is calculated using rigorous thermodynamic arguments.

The name "inverse Monte Carlo" is somehow confusing and is due to the fact that the original algorithm was combined with Monte Carlo sampling of the phase space [119]. However, practically any sampling method can be

used (e.g. molecular dynamics or stochastic dynamics) as long as it provides a canonical sampling of the phase space.

A detailed derivation of the IMC method can be found in Ref. [119]. Here a brief recapitulation of the more compact version for non-bonded interactions is given (Ref. [127]), with the emphasis on technical problems encountered during implementation and application of the method.

The idea of IMC is to express the potential update ΔU in a thermodynamically consistent way in terms of measurable statistical properties, e. g. radial distribution function $g(r)$. Considering a single-component system as an example one can write the Hamiltonian of the system as:

$$H = \sum_{i,j} U(r_{ij}) \ , \tag{4.9}$$

where $U(r_{ij})$ is the pair potential and we assume that all interactions depend only on the distance, r_{ij}, between particles i and j. It is further assumed that this potential is short-ranged, i.e. $U(r_{ij}) = 0$ if $r_{ij} \geq r_{\mathrm{cut}}$.

The next step is to tabulate the potential $U(r)$ on a grid of M points, $r_\alpha = \alpha \Delta r$, where $\alpha = 0, 1, \ldots, M$, and $\Delta r = r_{\mathrm{cut}}/M$ is the grid spacing. Then the Hamiltonian, eq. 4.9, can be rewritten as:

$$H = \sum_{\alpha} U_\alpha S_\alpha \ , \tag{4.10}$$

where S_α is the number of particle pairs with interparticle distances $r_{ij} = r_\alpha$ which correspond to the tabulated value of the potential U_α.

On one hand, the average value of S_α is related to the radial distribution function $g(r)$:

$$\langle S_\alpha \rangle = \frac{N(N-1)}{2} \frac{4\pi r_\alpha^2 \Delta r}{V} g(r_\alpha) \ , \tag{4.11}$$

where N is the number of atoms in the system ($\frac{1}{2}N(N-1)$ is then the number of all pairs) and Δr is the grid spacing, r_{cut}/M, V is the total volume of the system.

On the other hand, $\langle S_\alpha \rangle$ is a function of the potential U_α and hence can be expanded in a Taylor series with respect to small perturbations of U_α, ΔU_α:

$$\Delta \langle S_\alpha \rangle = \sum_{\gamma} \frac{\partial \langle S_\alpha \rangle}{\partial U_\gamma} \Delta U_\gamma + \mathcal{O}(\Delta U^2) \ . \tag{4.12}$$

The derivatives $\partial \langle S_\alpha \rangle / \partial U_\gamma$ can be obtained by using the chain rule:

$$\begin{aligned} A_{\alpha\gamma} &= \frac{\partial \langle S_\alpha \rangle}{\partial U_\gamma} \\ &= \frac{\partial}{\partial U_\gamma} \frac{\int dq S_\alpha(q) \exp\left[-\beta \sum_\lambda U_\lambda S_\lambda(q)\right]}{\int dq \exp\left[-\beta \sum_\lambda U_\lambda S_\lambda(q)\right]} \\ &= \beta \left(\langle S_\alpha \rangle \langle S_\gamma \rangle - \langle S_\alpha S_\gamma \rangle \right) \ . \end{aligned} \tag{4.13}$$

Using equations 4.11, 4.12, and 4.13 one can calculate the correction for the potential by solving a set of linear equations:

$$\langle S_\alpha \rangle - S_\alpha^{\mathrm{ref}} = A_{\alpha\gamma} \Delta U_\gamma \ , \tag{4.14}$$

where S_α^{ref} is given by the target radial distribution function. The procedure is then repeated until convergence is reached.

A clear advantage of the IMC compared to the IBI method is that the update of the potential is rigorously derived using statistical mechanics and hence the iterative procedure shall converge faster with the IMC update than with the empirical IBI update. Another advantage is that, in the case of multicomponent mixtures, IMC takes into account correlations of observables, that is updates for U_{AA}, U_{AB}, and U_{BB} are interdependent (A and B denote different particle types). In the IBI method these updates are independent which may lead to convergence problems for multicomponent systems.

The advantages come, of course, at a computational cost. As it is clear from eq. 4.13, one has to calculate cross-correlations of S_α. This requires much longer runs to get statistics that is good enough to calculate the potential update to a similar accuracy as IBI. The accuracies of the update functions of IMC and IBI methods are compared in section 5.1 for the case of a coarse-grained model of water.

Another issue of the IMC method is the stability of the scheme. Several factors can influence it: the first, and rather technical, point is that $g^{\mathrm{ref}}(r_\alpha)$ has to be calculated using exactly the same convention for the grid as S_α (e.g. the function value should be assigned to the middle of the interval), otherwise the scheme becomes unstable. Second, inversion of $A_{\alpha\gamma}$ requires that it shall be well defined. This means that one has to remove the regions which are not sampled, such as those at the beginning of the radial distribution function. The convergence can be significantly improved if a smoothing of the potential update ΔU is used. Note that it is better to do smoothing of the update function, not the potential itself, since the latter has more features which can be lost due to too aggressive smoothing. The convergence can also be improved by introducing a multiplicative prefactor for the update function or using a regularization procedure by adding thermodynamic constraints [135].

Finally, it was also noticed that the systematic error in $\langle S_\alpha S_\gamma \rangle$ is always higher in the vicinity of the cutoff, which leads to a shift in the tail of the interaction potential and, as a result, to a large offset of pressure. The cross-correlation term $\langle S_\alpha S_\gamma \rangle$ is also very sensitive to the box size, and special care must be taken in order to converge the results with respect to system size. Finite size effects are discussed in detail in section 5.2, where liquid methanol is coarse-grained.

4.5 Force matching

Force matching (FM) is another approach to evaluate corse-grained potentials [116, 108, 136]. In contrast to the structure-based approaches, its aim is not to reproduce various distribution functions, but instead try to match *forces* on coarse-grained beads as closely as possible. FM is a non-iterative method and hence is less computationally demanding.

The method works as follows. The first assumption is that the coarse-grained force-field (and hence the forces) depends on M parameters $g_1, ..., g_M$. These parameters can be prefactors of analytical functions, tabulated values of the interaction potentials, or coefficients of splines used to describe these potentials.

In order to determine these parameters, the reference forces on coarse-grained beads are calculated by properly re-weighting the forces on the atoms:

$$\vec{f}_i^{\,\mathrm{ref}} = M_i \sum_\alpha \frac{w_\alpha \vec{f}_\alpha}{m_\alpha} \ , \tag{4.15}$$

where $M_i = \left(\sum_\alpha w_\alpha^2/m_\alpha\right)^{-1}$ is the mass of the bead i, index α numbers all atoms belonging to this bead, $\vec{f}_\alpha$ is the force on the atom α, m_α is its mass, w_α are mapping coefficients used to obtain the position of the coarse-grained bead, $\vec{R}_i = \sum_\alpha w_\alpha r_\alpha$. If the center of mass is used in the mapping, eq. 4.15 simplifies to the sum of the forces.

By calculating the reference forces for L snapshots one can write down $N \times L$ equations:

$$\vec{f}_{il}^{\,\mathrm{cg}}(g_1, ..., g_M) = \vec{f}_{il}^{\,\mathrm{ref}},\ i = 1, \ldots, N, l = 1, \ldots, L \ . \tag{4.16}$$

Here $\vec{f}_{il}^{\,\mathrm{ref}}$ is the force on the bead i, $\vec{f}_{il}^{\,\mathrm{cg}}$ is the coarse-grained representation of this force. Index l enumerates snapshots picked for coarse-graining. By running the simulations long enough one can always ensure that $M < N \times L$. In this case the set of equations 4.16 is overdetermined and can be solved in a least-squares sense.

Though the underlying idea of FM is very simple, implementation wise it is the most complicated method. Here we briefly outline the problems, which are then discussed in more detail in Appendix A.

Going back to the set of equations 4.16 one can see that $\boldsymbol{f}_{il}^{\mathrm{cg}}$ is, in principle, a non-linear function of its parameters $\{g_i\}$. It is, therefore, useful to represent the coarse-grained force-field in such a way that equations (4.16) become linear functions of $\{g_i\}$. This can be done using splines to describe the functional form of the forces [108].

An adequate sampling of the system requires a large number of snapshots L. Hence, the applicability of the method is often constrained by the amount of available memory. To remedy the situation, one can split the trajectory into blocks, find the coarse-grained potential for each block and then perform averaging over the blocks. More details on the technical implementation of force matching using cubic splines is given in Appendix A.

4.5.1 FM method: formal statistical mechanical derivation

In the classification of CG methods given in Sec. 4.1, FM method was defined as force-based as opposed to structure-based methods, such as iterative Boltzmann inversion and inverse Monte Carlo. In this section it is shown that FM equations follow from the requirement that the ensemble of low resolution structures observed with a CG model is a low resolution representation of the ensemble that would be observed using an atomistically detailed system. This fact shows, that the classification given above is purely nomenclatural and that the force matching may be attributed to the structure-based methods as well. In order to show this mathematically, one has to introduce some notations. Description in this section follows the derivation given in [137].

It is assumed that the instantaneous state of the atomistic system is described by specifying the values of the Cartesian coordinates and momenta:

$$\boldsymbol{r}^n = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}, \tag{4.17}$$

$$\boldsymbol{p}^n = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n\}. \tag{4.18}$$

of the n atoms in the system. The atomistic hamiltonian is:

$$h(\boldsymbol{r}^n, \boldsymbol{p}^n) = \sum_{i=1}^{n} \frac{1}{2m_i} \boldsymbol{p}_i^2 + u(\boldsymbol{r}^n) \tag{4.19}$$

In the canonical ensemble, positions and momenta are distributed according to:

$$p_{rp}(\boldsymbol{r}^n, \boldsymbol{p}^n) = p_r(\boldsymbol{r}^n) p_p(\boldsymbol{p}^n) \tag{4.20}$$

where

$$p_r(\boldsymbol{r}^n) \propto \exp(-u(\boldsymbol{r}^n)/k_BT) \tag{4.21}$$

$$p_p(\boldsymbol{p}^n) \propto \exp(-\sum_{i=1}^{n} \boldsymbol{p}_i^2/2m_ik_BT) \tag{4.22}$$

Similarly, on a coarse-grained level, the coordinates and momenta are specified by the positions and momenta of CG sites:

$$\boldsymbol{R}^N = \{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}, \tag{4.23}$$

$$\boldsymbol{P}^N = \{\boldsymbol{P}_1, \ldots, \boldsymbol{P}_N\}. \tag{4.24}$$

Hamiltonian and canonical distributions of CG system can be obtained by formally substituting lower case letters with the capital ones in Eq. 4.19-4.22.

The physical meaning of the positions of the CG sites is specified by a linear mapping operator $\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n) = \{\boldsymbol{M}_{\boldsymbol{R}1}^N(\boldsymbol{r}^n), ..., \boldsymbol{M}_{\boldsymbol{R}N}^N(\boldsymbol{r}^n)\}$ of the form

$$\boldsymbol{M}_{\boldsymbol{R}I}^N(\boldsymbol{r}^n) = \sum_{i=1}^{n} c_{Ii}\boldsymbol{r}_i \qquad \text{for I = 1,. . . ,N} \tag{4.25}$$

Thus $\boldsymbol{M}_{\boldsymbol{R}I}^N(\boldsymbol{r}^n)$ is the physical meaning of $\boldsymbol{R}_I$ in terms of an atomistic model. Similarly, physical meaning of the CG momentum $\boldsymbol{P}_I$ is

$$\boldsymbol{M}_{\boldsymbol{P}I}^N(\boldsymbol{p}^n) = M_I \sum_{i=1}^{n} c_{Ii}\dot{\boldsymbol{r}}_i = M_I \sum_{i=1}^{n} \frac{c_{Ii}}{m_i}\boldsymbol{p}_i \qquad \text{for I = 1,. . . ,N} \tag{4.26}$$

Coefficients c_{Ii} are usually called a *mapping scheme*. For any reasonable mapping scheme translational invariance should be satisfied. That is, if an atomistic system is translated by a constant vector, the corresponding coarse-grained system is also translated by the same vector. This implies that, for all I:

$$\sum_{i=1}^{n} c_{Ii} = 1. \tag{4.27}$$

Atomistic equilibrium probability distributions in Eq. 4.20 together with the mapping operators in eqns. 4.25 and 4.26 imply the following equilibrium distribution for the CG positions and momenta:

$$p_R(\boldsymbol{R}^N) = \int d\boldsymbol{r}^n p_r(\boldsymbol{r}^n)\delta(\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N) \tag{4.28}$$

$$p_P(\boldsymbol{P}^N) = \int d\boldsymbol{p}^n p_p(\boldsymbol{p}^n)\delta(\boldsymbol{M}_{\boldsymbol{P}}^N(\boldsymbol{p}^n) - \boldsymbol{P}^N) \tag{4.29}$$

The requirement $p_R(\boldsymbol{R}^N) = P_R(\boldsymbol{R}^N)$, $p_P(\boldsymbol{P}^N) = P_P(\boldsymbol{P}^N)$ is equivalent to the following equations:

$$\exp(-U(\boldsymbol{R}^N)/k_BT) \propto \int d\boldsymbol{r}^n \exp(-u(\boldsymbol{r}^n)/k_BT) \times\delta(\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n) - \boldsymbol{R}^N) \tag{4.30}$$

$$\exp\left(-\sum_{I=1}^{N} \boldsymbol{P}_I^2/2M_Ik_BT\right) \propto \int d\boldsymbol{p}^n \exp\left(-\sum_{i=1}^{n} \boldsymbol{p}_i^2/2m_ik_BT\right) \times\delta(\boldsymbol{M}_{\boldsymbol{P}}^N(\boldsymbol{p}^n) - \boldsymbol{P}^N) \tag{4.31}$$

Equations 4.30 and 4.31 are the consistency conditions for the CG model. If they are satisfied, the CG model will generate the same statistical ensemble as the low resolution representation of the underlying atomistic model. Below such CG models are referred to as *consistent* CG models.

Equation 4.30 implies, that the coarse-grained force-field for a consistent CG model $U(\boldsymbol{R}^N)$ is completely determined by the atomistic force-field $u(\boldsymbol{r}^n)$ and the mapping scheme $\boldsymbol{M}_{\boldsymbol{R}}^N(\boldsymbol{r}^n)$. Consistent CG force-field is a many-body potential of mean force (PMF), which is a conditional free energy surface in the coordinate space of the CG variables. In general $U(\boldsymbol{R}^N)$ does not factorize, and contains 2-body, 3-body, ..., interactions.

It can be shown, that equation 4.30 is equivalent to the force-matching equations 4.16 for the broad range of mapping operators [137]. Thus, FM equations can be obtained by requiring that structural properties of the atomistic model are reproduced. In practice, however, FM method rarely reproduces structural properties, such as radial distribution functions (see chapter 5), because the basis set used to represent the CG force-field is not complete. Due to efficiency reasons, most modern MD codes use only pairwise spherical-symmetric nonbonded potentials. Thus, CG force field is an effective pair potential, that represents an approximate decomposition of the many-body interaction obtained from a formal integration over uninteresting degrees of freedom. For many cases this is not enough to represent many-body potential of mean force $U(\boldsymbol{R}^N)$ (see Sec. 5.1).

4.5.2 FM method: connections to the liquid state theory

Structure-based coarse-grained modeling is similar to the "inverse problem" of liquid-state theory [138]. Both attempt to determine an interaction potential reproducing an observed structure. The theory of the Yvon-Born-Green

(YBG) equations provides a direct solution to this inverse problem [138], given that such an interaction potential exists. The YBG equation provides an exact relation between a given two-body interaction potential and the n- and $(n+1)$-particle distribution functions obtained from equilibrium simulations employing this potential.[1] Therefore, a CG pair potential may be determined by inverting the YBG equation for the observed two- and three-particle CG distribution functions. This relationship shows a role of high-order correlations in deducing a pair potential that will reproduce the observed CG structure.

It was recently shown by W. Noid *et al* [136] that for homogeneous isotropic systems force matching equations 4.16 are equivalent to the generalized Yvon-Born-Green equations [138]. For such systems, FM procedure explicitly considers the two- and three-particle correlations between CG sites within an atomistic MD simulation, assumes, that these distributions were generated by a pairwise decomposable force field, and then inverts the resulting YBG equation to determine this force field.

4.6 Relationship between CG methods

In this section a comparison of force matching and RDF-based methods (IBI, IMC) is given for one-component homogeneous isotropic systems.

Iterative Boltzmann inversion and Inverse Monte Carlo methods determine CG potentials, which reproduce a given radial distribution functions (RDF). However, these potentials are not guaranteed to reproduce higher-order correlation functions. For example, it is known that IBI and IMC potentials for water do not reproduce three-particle correlations, although the RDF is quantitatively accurate [134, 139]. In contrast, force matching method is able to reproduce both RDF and 3-body correlations, if 3-body interactions are included in the coarse-grained force field [140].

As discussed in Sec. 4.5.2, force matching method implicitly measures two- and three-body correlation functions describing CG beads within an atomistic MD simulation and then directly inverts the YBG equation to determine a 2-body central CG potential that would generate these distribution functions, if such a potential exists. In general such a potential may not exist, and consequently, simulations with the resulting FM potential will not reproduce either the two- or the three-particle distributions exactly. Nevertheless, FM method clearly takes three-particle correlations into account. Since FM does not necessarily reproduce the pair correlation functions, comparison of

[1] YBG equations for a homogeneous liquid are presented in Appendix B.

atomistic and CG RDFs is a useful measure of the validity of the CG model (see Chapter 5).

In contrast, IBI and IMC methods only consider the two-particle correlations and attempt to solve the YBG equation for the pair interaction that reproduces the target RDF while allowing the three-particle correlation functions to vary as necessary [136]. Succesive Molecular Dynamics or Monte Carlo simulations used for iteratively updating the pair potential may be considered as a nonlinear regression algorithm that solves the YBG equation for a pair potential reproducing a fixed RDF. If three-body correlations were not significant in the YBG equation, then the pair potential would be simply the two-body PMF, which is often implemented as an initial guess in IBI and IMC. During successive iterations the simulated RDF converges to the target RDF, measured from atomistic simulations. IBI and IMC implicitly incorporate information regarding three-body correlations by updating the force field to improve agreement between the measured and the target RDF, but in the iterative simulations the three-body correlations may change. The YBG equation that is implicitly solved through the IBI or IMC method incorporates the target RDF and is guaranteed to reproduce this RDF. However, the three-particle correlations in the final YBG equation may be different than those in the original atomistic representation of the system.

Chapter 5

Comparison of coarse-graining techniques

After the discussion of coarse-graining techniques in detail in Chapter 4 it is useful to consider practical examples. In this chapter iterative Boltzmann inversion, inverse Monte Carlo and force matching are applied to model systems of different complexity: a three-site SPC/E water, methanol, propane, and a single chain of hexane. Strengths and drawbacks of different methods are discussed for every example. This study was done together with V. Rühle and C. Junghans within the VOTCA project [2].

5.1 Coarse-graining of water

Water is one of the most studied liquids, both from the point of view of all-atom representations and coarse-grained models [141, 142]. In this section coarse-graining of the all-atom SPC/E water model [143, 144] is presented. The corresponding parameters of this 3-site model are given in the caption to Fig. 5.1. Note that this is a rigid model, i. e. the distances between two hydrogens as well as oxygen and hydrogens are constrained during the molecular dynamics runs. For coarse-graining one-site representation with a pair potential $U(R_{ij})$ is used, where R_{ij} connects the centers of mass of water molecules i and j.

The all-atom system consisting of 2180 water molecules was first equilibrated in the NPT ensemble at 300 K and 1 bar for 100 ns using the Berendsen thermostat and barostat [81]. The last 80 ns were used to determine the equilibrium box size of 4.031 nm, which was then fixed during the 45 ns production run in the NVT ensemble using a stochastic dynamics algorithm [145]. For all further analysis, only the last 40 ns were used. The radial distribution

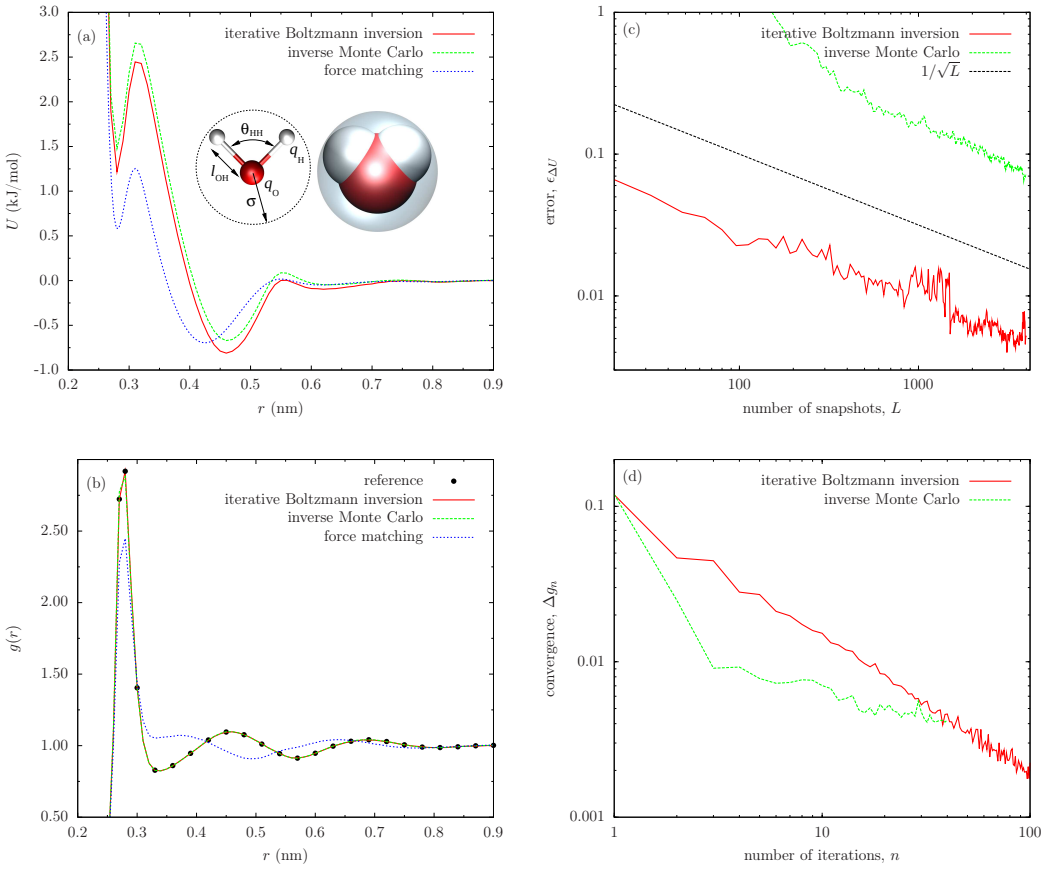

Figure 5.1: **Water** (a) Coarse-grained potentials for SPC/E water obtained using different coarse-graining techniques. (b) Corresponding radial distribution functions. (c) Average error of the potential update function versus number of snapshots used for calculating the update function. (d) Root mean square deviation of reference and current radial distribution function versus iteration step. One can see that IMC converges faster than IBI. Inset of (a) shows Van der Waals excluded volume and coarse-grained representations of a single water molecule, as well as parameters used: $\sigma = 3.166\,\text{Å}$, $\epsilon = 0.650\,\text{kJ mol}^{-1}$, $l_{\text{OH}} = 1.0000\,\text{Å}$, $q_{\text{H}} = +0.4238\,e$, $q_{\text{O}} = -0.8476\,e$, $\theta_{\text{HH}} = 109.47°$.

function was calculated using a 0.01 nm grid spacing. The snapshots were output every 0.4 ps.

Force matching potentials were calculated using blocks of 6 snapshots each. Spline grid spacing of 0.02 nm was used in the interval from 0.24 to 1 nm. For the iterative procedures, the potential of mean force was taken as an initial guess for the interaction potential. The coarse-grained box had the same system size as in the atomistic simulations. Simulations of the coarse-grained liquid were done using a stochastic dynamics algorithm. [145] 300 iterations of 100 ps each were performed when using IBI. For IMC we used 10 iterations of 500 ps each. Additionally, two iterations of triangular smoothing were applied to the IMC potential update, ΔU. The cut-off was chosen at 0.9 nm with a grid spacing of 0.01 nm.

The reference radial distribution function, $g^{\mathrm{ref}}(r)$, coarse-grained potentials and corresponding radial distribution functions are shown in fig. 5.1a,b. IBI and IMC give practically the same interaction potential. Although the force-matched potential has a very similar structure with two minima, the corresponding radial distribution function is very different from the target one. Reasons for these discrepancies are discussed in Refs. [136, 126, 127] and in Chapter 4 and stem from the fact that FM aims to reproduce the many-body potential of mean force, which does not necessarily guarantee perfect pairwise distribution functions, considering the fact that the basis sets in the coarse-grained force field may be limited. It is worth mentioning, that inclusion of 3-body interactions into the coarse-grained force field remedy the situation for the FM method: in this case both radial distribution function and 3-particle correlations are reproduced [140].

Note that all three methods lead to a different pressure of the coarse-grained system: 8000 bar (IBI), 9300 bar (IMC), and 6500 bar (FM). Different pressures for the iterative methods is due to different accuracy of the potential update. Indeed, small changes of pressure can significantly affect the potential, especially its long tail [118, 146, 134]. However, they hardly change the radial distribution function due to small compressibility of water. One can improve the agreement between the iterative methods on the cost of the compressibility by using pressure correction terms for the update [134].

The performance of the iterative methods depends on two factors: (i) the average (over all bins) error of the potential update $\epsilon_{\Delta U}$ and (ii) the number of iterations required for convergence. We define the average error as:

$$\epsilon_{\Delta U} = \frac{1}{N} \sum_{i=0}^{N} \epsilon(\Delta U(r_i)) \ , \tag{5.1}$$

where N is the number of bins and $\epsilon(\Delta U(r_i))$ is the error of the update

function at a separation r_i. $\epsilon(\Delta U(r_i))$ was calculated using a Jackknife analysis [147].

The average error of the potential update is shown in Fig. 5.1c as a function of the run length. One can see that, for both methods, the error decreases as $1/\sqrt{L}$, where L is the number of snapshots used for averaging. However the prefactor for the IBI update error, which is based on the radial distribution function, is at least ten times smaller than for the IMC update error, which makes use of cross-correlations of S_α. This observation implies that, in order to have the same accuracy of the update function, IMC needs significantly longer sampling.

This disadvantage is of course compensated by the efficiency of the update function, which is assessed by computing the root mean square deviation, Δg_n, of the current and target radial distribution functions:

$$\Delta g_n^2 = \int \left[g^{\mathrm{ref}}(r) - g^{(n)}(r) \right]^2 \mathrm{d}r \ . \tag{5.2}$$

Δg_n is plotted as a function of the number of iterations, n, in Fig. 5.1d. It is clear that IMC converges much faster than IBI, though the root mean square deviation saturates after some number of iterations.

5.2 Coarse-graining of methanol

Liquid methanol (see the inset in Fig. 5.2) is the second example of coarse-graining of *non-bonded* interactions that is presented here. In fact, FM has already been used to coarse-grain this system [146] and, contrary to water, the liquid structure (radial distribution function) is well reproduced by the FM coarse-grained potential. In addition, the excluded volume of methanol is larger than that of water and the undulations of the radial distribution function extend up to 1.5 nm. As will be shown, this leads to pronounced finite size effects for IMC, since it has a non-local potential update. FM and IBI do not have this problem, since the IBI potential energy update is local and FM is based on pair forces. The range of the latter is much shorter than the correlation length of structural properties (such as undulations of the radial distribution function), which may propagate over the boundaries for small boxes.

Simulation parameters were taken from ref. [146] and OPLS [148, 149] all-atom force field was used. Atomistic simulations were performed with 1000 methanol molecules in a cubic box (4.09 nm box size) at 300 K using the Nose-Hoover thermostat [82, 83]. The system was equilibrated for 2 ns followed

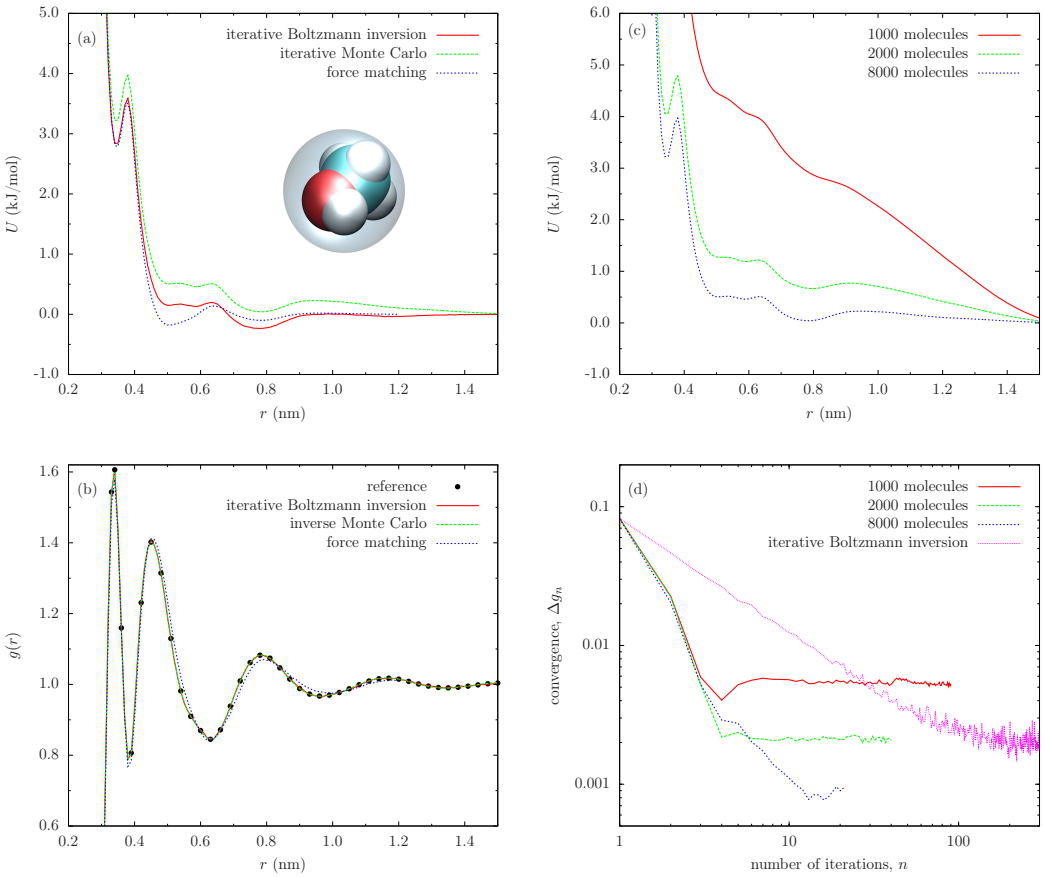

Figure 5.2: **Methanol** (a) Coarse-grained potentials. (b) Corresponding radial distribution functions. (c) coarse-grained potentials using 10 IMC iterations for simulation boxes with 1000, 2000 and 8000 methanol molecules (box size 4.09 nm, 5.15308 nm, and 8.18 nm) equilibrated at the same density. (d) Root mean square deviation of reference and the current radial distribution function versus number of iterations. Similar to liquid water, IMC converges faster than IBI. The convergence saturates and the saturation error strongly depends on the system size. The inset of (a) shows the Van der Waals excluded volume and coarse-grained representations of a methanol molecule.

by a production run of 18 ns. The reference radial distribution function was calculated using snapshots every 0.5 ps and is shown in Fig. 5.2b.

The FM potential was calculated using blocks of 6 frames each and a spline grid of 0.02 nm. With this potential, coarse-grained simulations were performed using a stochastic dynamics integrator, 1000 beads and with the same box size and the same temperature as in the atomistic simulations. The system was equilibrated for 40 ps followed by a production run of 160 ps. Snapshots were stored every 5 ps and used to calculate the radial distribution function.

For the iterative procedures, the potential of mean force was taken as an initial guess. The cutoff was chosen at 1.54 nm with a grid spacing of 0.01 nm. For IBI, 300 iterations were performed using stochastic dynamics with the same parameters used in the FM-based procedure. The IMC iterations were performed with 8000 molecules and a box size of 8.18 nm. The total length of the run was 1 ns and snapshots were stored every 0.2 ps. Two smoothing steps were used at each iteration for the potential update, ΔU.

The coarse-grained potentials for all methods are shown in Fig. 5.2a. In spite of small differences between the coarse-grained potentials, the agreement between the reference and coarse-grained radial distribution functions is excellent, as can be seen from Fig. 5.2b.

It is important to mention that the IMC method, which has a non-local update, is prone to systematic errors due to finite size effects and hence requires much larger simulation boxes in order to calculate the potential update. This is due to artificial cross-correlations of S_α at large distances, which lead to a small difference of tails between the coarse-grained and the reference radial distribution functions, and, as a consequence, to a much higher pressure of the coarse-grained system and a significantly different coarse-grained potential. In contrast, IBI and FM work well with system sizes of the order of two radial distribution function cutoff lengths.

To illustrate this point, simulation boxes of three different sizes were prepared, with 1000, 2000 and 8000 methanol molecules (box size of 4.09 nm, 5.15308 nm and 8.18 nm, simulation times of 3 ns, 2 ns and 1 ns respectively). The IMC iterative procedure was repeated until the potentials converged, and these are shown in Fig. 5.2c. One can see that the potentials significantly differ from each other. These differences lead to small deviations in the tail of the radial distribution function, which, however, vanish, in a systematic way, for bigger boxes, as illustrated in Fig. 5.2d, where the integral of the difference of the reference and current distribution functions is plotted. More detailed analysis have shown that, for small boxes, an additional linear term in the potential update at large separations appear. To determine the origin of this term, ΔU was calculated using the full matrix $A_{\alpha\beta}$ as well as only

its diagonal elements. The potential after 50 IBI iterations was taken as an initial guess. Without the off-diagonal elements ΔU was small once the reference and coarse-grained radial distribution functions were matching each other. Inclusion of the off-diagonals elements always resulted in an additional, practically linear, term in the potential update which became smaller for large boxes. Based on this observation is was concluded that the off-diagonal elements of the matrix $A_{\alpha\beta}$ systematically change with the box size.

To summarize, IMC should be used with care for small systems. The potential update (or the coarse-grained potential) must be converged with respect to the simulation box size. In the case of methanol coarse-graining, a box of size three times the radial distribution function cut-off was not enough to achieve the converged potential for IMC, even though this is sufficient for IBI and FM methods.

5.3 Liquid propane: from an all- to an united-atom description

So far coarse-graining of *non-bonded* degrees of freedom was illustrated using liquid water and methanol as examples. In this section it is shown how *bonded* interactions can be coarse-grained by deriving a united atom model (i.e. hydrogens embedded into heavier atoms) from an all-atom model of liquid propane.[1] The mapping scheme, as well as the bonded coarse-grained variables (two bonds, b, and one angle, θ) are shown in the inset of Fig. 5.3. Note that this coarse-graining scheme has two different bead types: an inner bead, of type B, with two hydrogens, and two outer beads, of type A, with three hydrogens. As a result, three types of non-bonded interactions, U_{AA}, U_{BB}, and U_{AB} must be determined.

As before, atomistic simulations were performed using the OPLS all atom force field [148, 149]. A box of liquid propane was first equilibrated at 200 K and 1 bar in the NPT ensemble for 10 ns using the Berendsen thermostat and barostat [81]. The equilibrated box of the size $4.96337 \times 5.13917 \times 4.52386\,\mathrm{nm}^3$ was then simulated for 10 ns in the NVT ensemble at 200 K using velocity rescaling with a stochastic term [84]. No bond constraints were used during

[1]The united atom model we use here shall not be confused with the united atom models commonly used in the atomistic force-field community, for example OPLS-UA force-field [148, 149]. The latter models map the potentials, which are analytical functions of bonds, angles, and dihedral angles, onto thermodynamic properties of the corresponding substances. In our case coarse-grained potentials are tabulated functions of coarse-grained variables and only the mapping (hydrogens embedded into heavier atoms) is similar to that of the united atom force-fields.

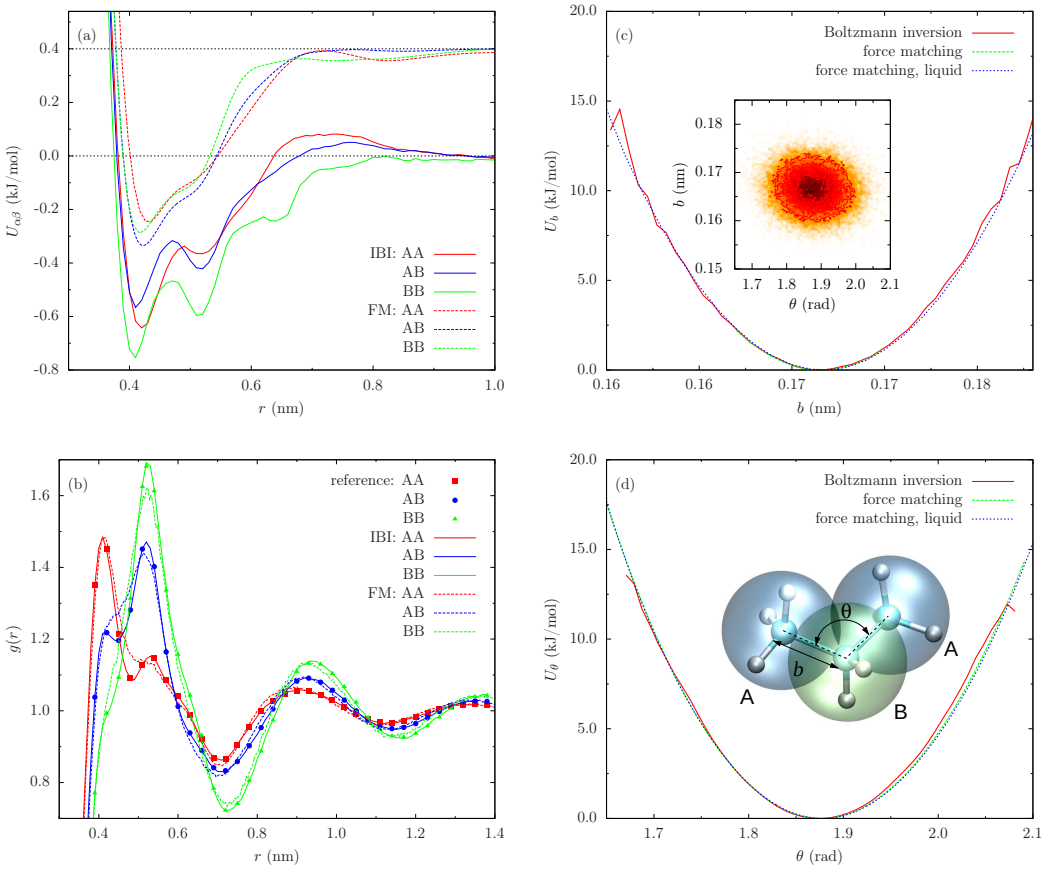

Figure 5.3: **Propane** (a) Non-bonded interaction potentials U_{AA}, U_{BB}, and U_{AB} obtained with IBI and FM methods. For clarity, FM potentials are offset along the y axis. (b) Corresponding radial distribution functions, plotted together with the atomistic radial distribution function. (c) Bond potential obtained for a single molecule in vacuum by Boltzmann-inverting the corresponding distribution function, using FM for a single propane molecule in vacuum and force matching for liquid propane. (d) Angular coarse-grained potentials. FM-based distributions for a single molecule and the liquid are on top of each other. The inset of (c) shows the correlations of b and θ. The inset of (d) shows all-atom and coarse-grained representations of a propane molecule, bead types, and coarse-grained bonded degrees of freedom (bond b and angle θ).

the simulations and hence the integration timestep was 1 fs. Snapshots were written every 1 ps.

In the case of iterative methods, the bonded potentials (bond and angle) were calculated by Boltzmann-inverting the corresponding distribution functions of a single molecule in vacuum, according to eq. 4.5. The propane molecule in vacuum was simulated in an stochastic dynamics run [145] for 100 ns with snapshots stored every 2 ps. Non-bonded potentials were iteratively refined by using IBI with a grid spacing of 0.01 nm and a cutoff of 1.36 nm (1.38 nm) for A-A, A-B (B-B) interaction types, respectively. The run length for each iteration was 50 ps with snapshots written every 0.5 ps. At every iteration step only one interaction type was corrected. When using the FM method, both bonded and non-bonded potentials were obtained at the same time, since FM does not require the explicit separation of bonded and non-bonded interactions.

The obtained potentials are shown in Fig. 5.3a,c,d. FM and Boltzmann inversion-derived bond and angle potentials (Fig. 5.3c,d) perfectly agree with each other. The non-bonded potentials, shown in Fig. 5.3a, are not completely identical, but have similar shapes and barrier heights. This of course results in a good reproducibility of the propane liquid structure by the FM-based coarse-grained potentials, as can bee seen from the radial distribution functions shown in figure Fig. 5.3b. Again, as expected, IBI reproduces the reference radial distribution functions exactly.

To summarize, the united atom model of liquid propane is an ideal example of coarse-graining where the structure- and force-based methods result in similar bonded and non-bonded interaction potentials. As will be shown later, this is due to (i) the completeness of the basis set used to construct the coarse-grained force-field; and (ii) independence of bond and angular degrees of freedom. The latter can be understood with the help of a histogram showing the correlation of b and θ, depicted in the inset of Fig. 5.3c.

In the next section coarse-graining of a single molecule of hexane is considered, for which this is not the case.

5.4 Angular potential of a hexane molecule

The last example discussed here is the angular potential of a hexane coarse-grained into a three-bead chain, with two carbon atoms per bead (see the inset in Fig. 5.4a). Atomistic simulations of a single hexane molecule in vacuum were performed using the all-atom OPLS force field and a stochastic dynamics integrator [145]. The run length was 1000 ns and snapshots were stored every 2 ps.

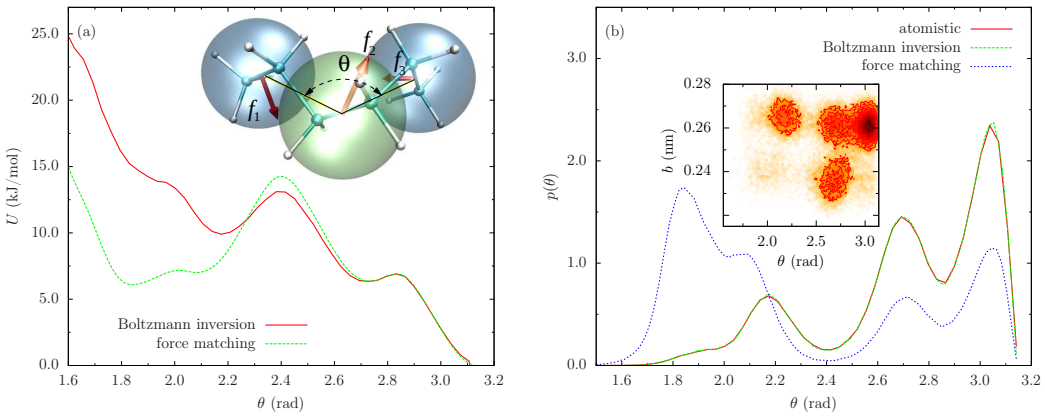

Figure 5.4: **Hexane** (a) Coarse-grained angular potentials obtained using Boltzmann inversion (no iterations) and FM for a single hexane molecule in vacuum. The inset of (a) shows hexane molecule and its coarse-grained representation. Arrows indicate the directions of the forces on three beads for a specific snapshot. (b) Probability density (probability distribution normalized by the interval) obtained from the atomistic run as well as from the runs using coarse-grained angular potentials. The inset of (b) shows the correlation of b and θ).

The coarse-grained angular potential was again obtained by Boltzmann-inverting the angular distribution function or by using the FM method (blocks of 50000 frames each, a spline grid of 0.05 nm, sampling in the $\theta \in [1.6, 3.14]$ interval). Both coarse-grained potentials are shown in Fig. 5.4a. The corresponding distribution functions, together with the reference function obtained from the atomistic simulations, are shown in Fig. 5.4b.

It is obvious that the distribution which corresponds to simple Boltzmann inversion is practically identical to the reference distribution, while the FM-based distribution samples small angles much more often, which is a direct consequence of a very deep local minimum in the angular potential at these angles. It is easy to understand why FM fails to predict the relative height of this minimum. On a coarse-grained level the change of the angle from large to small values corresponds to the reorientation of the dihedral angles at the atomistic level. This reorientation results in instantaneous forces, $\boldsymbol{f}_1$, $\boldsymbol{f}_2$, $\boldsymbol{f}_3$, on the beads which have an *out of plane* component, where the plane is defined by the centers of the beads (see also the inset of Fig. 5.4a). The coarse-grained potential, however, has only an angular term, U_θ, and hence can only capture forces which lie *in the plane* in which the angle θ is defined. Hence, only the *projections* of the forces on this plane are used in FM, and this clearly leads to underestimation of the position of the second minimum, since the work conducted by the out-of-plane forces is completely ignored.

Additionally, this mapping scheme does not have independent variables,

e. g. bond and angle degrees of freedom are coupled, as can be seen from the Ramachandran plot shown in the inset of Fig. 5.4(b). This means that, even though Boltzmann inversion reproduces correct distributions, sampling of the configurational space is incorrect because of the lack of cross-correlation terms in the coarse-grained potential.

This example clearly shows that coarse-graining shall be used with understanding and caution, the methods should be cross-checked with respect to each other, as well as with respect to the reference system. In the next Chapter a more complicated example is considered, namely a solvent-free coarse-grained model of poly(2,3-diphenylphenylene vinylene) with hexyl (DP6-PPV) and decyl side chains (DP10-PPV) in chloroform.

Chapter 6

Conformational structure of PPV derivatives

After introducing different coarse-graining techniques in Chapter 4 and discussing their strengths and weaknesses in Chapter 5 we will consider a more complicated example, namely a coarse-grained model for the dilute solutions of conjugated polymer poly(2,3-diphenyl phenylene vinylene).

Conjugated polymers have attracted much interest due to their unique optical and semiconducting properties, making them the materials of choice for opto-electronic applications, e.g. polymer light emitting diodes and plastic conductive layers [150, 151, 152].

Even in the early stages of the design of polymeric optoelectronic devices, it already became clear that the electronic properties of thin films of conjugated polymers are extremely sensitive to the global as well as the local arrangement of chains. By choosing different processing techniques, e.g. spin-coating or drop-casting, and processing conditions, such as solvent and temperature, it is possible to obtain different morphologies and hence control the electronic properties of the film.

The generic molecular architecture of conjugated polymers is comprised of a semi-rigid conjugated backbone, responsible for charge conductance, and flexible side chains that insure solubility and facilitate processing. When drop-casted or spin-coated from solution, the resulting film morphology depends on the chain conformation in dilute solution. Hence, understanding of the conformational structure of conjugated polymers in solutions is necessary in order to control the morphology of a thin organic semiconducting layer.

Experimentally, techniques such as light or neutron scattering are used to study structural properties of polymer solutions. These properties are extracted by fitting scattering profiles to predefined analytical models [153]. However, even an excellent fit alone can not guarantee that the underlying

analytical model is valid. In this situation, computer simulations may be employed to validate the conclusions of experimental studies and help to link morphology and electronic structure to charge or exciton mobility (see Chapter 2). In order to do this, one first has to be able to generate large-scale material morphologies at an atomistic scale resolution. This involves (a) parametrization of atomistic force-fields, since these are not readily available for most organic compounds (Chapter 3) and (b) development of solvent-free coarse-grained models, capable of back-mapping (Chapters 4 and 5). The latter is essential for extending time- and length-scales accessible to those of classical molecular dynamics simulations.

In this chapter, we illustrate how the latter task may be tackled for dilute solutions of conjugated polymers.[1] Specifically, we study two derivatives of poly(2,3-diphenylphenylene vinylene) (DP-PPV) with hexyl (DP6-PPV) and decyl (DP10-PPV) side chains, which were presented in Chapter 3 (see Fig. 3.1).

DP-PPVs have been considered as a family of green-emitting materials for LED applications due to their high glass transition temperature, high fluorescence efficiency, and ease of monomer and polymer synthesis [90, 91, 92]. Conformational structure and aggregation behavior of DP6-PPV have been studied by means of small angle neutron scattering (SANS) and dynamic light scattering (DLS) [154]. It could be shown that DP6-PPV tends to aggregate in chloroform and toluene, yielding network aggregates whose internal structure can be characterized by a certain fractal dimension. Two types of segmental association with distinct stability were identified for the toluene solution. The highly stable segmental association was attributed to the $\pi - \pi$ complex already present in the DP6-PPV powder, while the labile segmental association was ascribed to the poor affinity of the aliphatic side chains of DP6-PPV to toluene. An analogous study for DP10-PPV showed that only a minor fraction of the polymer undergoes segmental association in chloroform, whereas in toluene disk-like clusters are formed [155]. The difference in aggregation behavior between DP6- and DP10-PPV was attributed to the more pronounced steric repulsion of the longer side-chains in the latter polymer.

This chapter is organized as follows. First the results of the atomistic simulations of dilute DP-PPV solutions are presented. Special attention is paid to side-chain stretching, backbone orientational correlations, chain tacticity, and the influence of side groups on the backbone planarity. Subsequently, atomistic trajectories of a single chain in a solvent in combination with the potential of mean force calculations are used to parametrize a coarse-grained

[1] The former task is considered in Chapter 3.

model with two repeat units per bead and coarse-grained simulations are performed. These then serve to study large-scale systems and to determine how static structure factors and persistence length change as a function of polymer concentration. The work is summarized by comparing simulation results to experimental data and by commenting on the limitations of our multiscale approach.

6.1 Atomistic molecular dynamics

In this section, we present results of our atomistic molecular dynamics simulations such as side chain stretching, planarity of the backbone, and chain persistence length in two solvents. Throughout the text, we use the following abbreviations for different polymer derivatives, solvents, and backbone lengths: each system name begins either with D6 or D10, corresponding to DP6-PPV or DP10-PPV respectively. An additional letter denotes the solvent: "c" is for chloroform and "t" for toluene. Finally, a number for the backbone chain length, in repeat units, is added. For example, D6c20 is a 20 monomer units long DP6-PPV derivative in chloroform. Unless otherwise stated, all simulations were performed at 300 K.

6.1.1 Alkyl side chain stretching in solvated DP-PPV

Experimentally, chloroform is considered to be a relatively good solvent for both DP6- and DP10-PPV derivatives, whereas toluene is a relatively poor one. Additional segmental aggregation in toluene relative to chloroform is often ascribed to the poor affinity of the aliphatic side chains of DP-PPV for toluene [154, 155]. To check this, we analyzed alkyl side chain stretching in both solvents.

A 10-mer of DP6- or DP10-PPV was simulated in chloroform and toluene for 40 ns and the corresponding distributions of the side chain end-to-end distances are shown in Fig. 6.1. As one can see, the distributions are identical for both solvents, i.e. the difference in solvent quality does not affect the side chain conformations, at least in our atomistic model. The same conclusion could be made when simulating an all-atom hexane chain (which corresponds to a side chain of DP6-PPV) in chloroform and toluene, where again, no effect of solvent quality on chain conformations was detected.

Experimentally, solvent quality can be characterized via the second virial coefficient A_2, which can be determined from static light scattering experiments [155]. $A_2 = 6.7 \times 10^{-6}\mathrm{mol\,dm^3\,g^{-2}}$ was reported for chloroform and $3.7 \times 10^{-6}\,\mathrm{mol\,dm^3\,g^{-2}}$ for toluene. This difference is rather small and

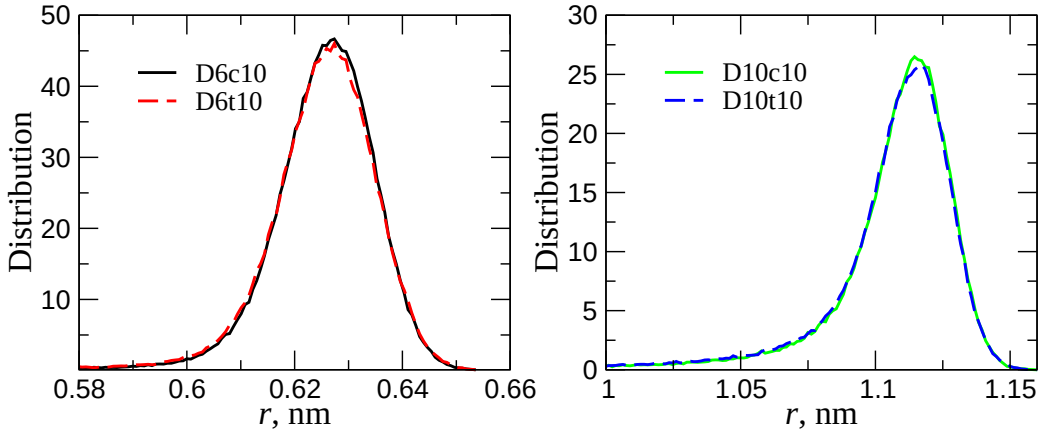

Figure 6.1: Alkyl side chain end-to-end distance distributions for DP6- and DP10-PPV chains of 10 repeat units. Chain stretching does not depend on the solvent. Name abbreviations are explained in the beginning of Sec. 6.1

poses the question of whether toluene is a poorer solvent than chloroform only because of the additional chain aggregation observed in toluene solutions [154, 155]. It is also interesting that toluene is often reported to be a good solvent for another PPV derivative, MEH-PPV (Ref. [156]) as well as for polyfluorenes with longer alkyl side chains [157, 158]. This further suggests that both toluene and chloroform might be relatively good solvents for DP-PPVs and that aggregation is an artifact of an initial non-equilibrium state of polymer chains in a powder.

6.1.2 Orientational correlations in solvated DP-PPV

Orientational correlations of the polymer backbone can be used to calculate the persistence length of a chain, which can be directly compared to experimental values.

A single chain of 20 monomer units was solvated in previously equilibrated solvent boxes, containing 23550 and 17728 molecules of chloroform and toluene respectively. After 10 ns of equilibration, a production run of 40 ns was performed. During the production run, the orientational correlations of repeat units were calculated as

$$\cos\theta_n = \langle e_i \cdot e_{i+n} \rangle \,, \tag{6.1}$$

where e_i is a unit vector giving the orientation of the i-th repeat unit along the backbone. This orientation is defined by a pair of carbon atoms for a backbone benzene ring, such as atoms 3 and 6 or 16 and 21 in Fig. 3.1b.

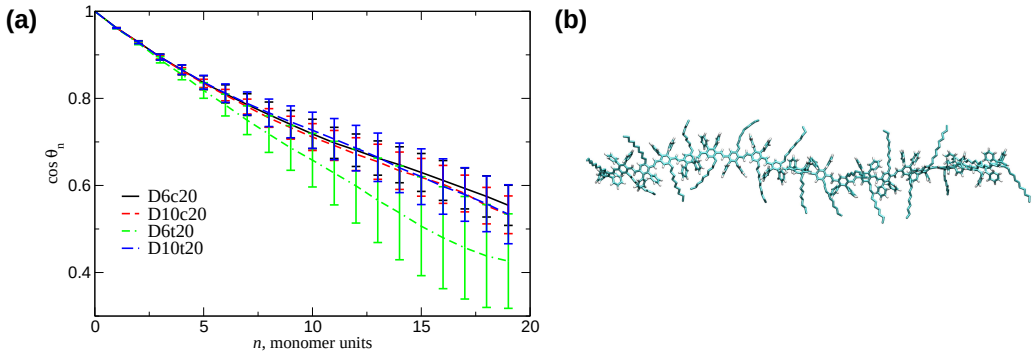

Figure 6.2: (a) Orientational correlations for 20 repeat unit chains of DP6- and DP10-PPV in chloroform and toluene. 40 ns run, average of 20000 configurations. (b) Typical conformation of a DP10-PPV chain in chloroform.

$\langle \cdots \rangle$ denotes the time average and the average over all pairs, i, $i+n$, along the chain. The persistence length, l_p, can be estimated as $\log(\cos\theta_n) \propto -nl/l_p$, where l is the length of the repeat unit. Note, that this is only possible if correlations decay exponentially, which might not necessarily be the case [159, 160, 161].

Orientational correlations for different solvents and different side-chain lengths are shown in Fig. 6.2a. It is clear that the backbone is quite rigid (see also Fig. 6.2b, where a typical conformation of a DP10-PPV chain in chloroform is shown). Moreover, within available accuracy, the decay of the correlation function does not depend on the side-chain length or solvent. A fit yields a rough estimate of the persistence length, $l_p \sim 17-25$ Å. Improved statistical averages are needed for more accurate estimates of l_p, which can be obtained by using solvent-free coarse-grained models, as described in Sec. 6.2.

6.1.3 Planarity and tacticity of solvated DP-PPV

To facilitate charge transport along a conjugated chain, planarity of the backbone is required [162]. In principle, conjugation already enforces a planar backbone conformation. However, non-bonded (Coulomb and van der Waals) contributions can favor twists in the backbone. Here, we study how the chemical structure of a repeat unit, in particular the side chain groups, affects the planarity and tacticity of the backbone.

A PPV repeat unit has two dihedral angles which control its planarity, as depicted in the inset of Fig. 6.3a. In Fig. 3.2a the dihedral potential of trans-stilbene, corresponding to these angles is shown. It has two minima separated by a barrier significantly higher than $k_{\mathrm{B}}T$. Since both minima are rather shallow, thermal fluctuations can easily lead to a 45 degree twist in

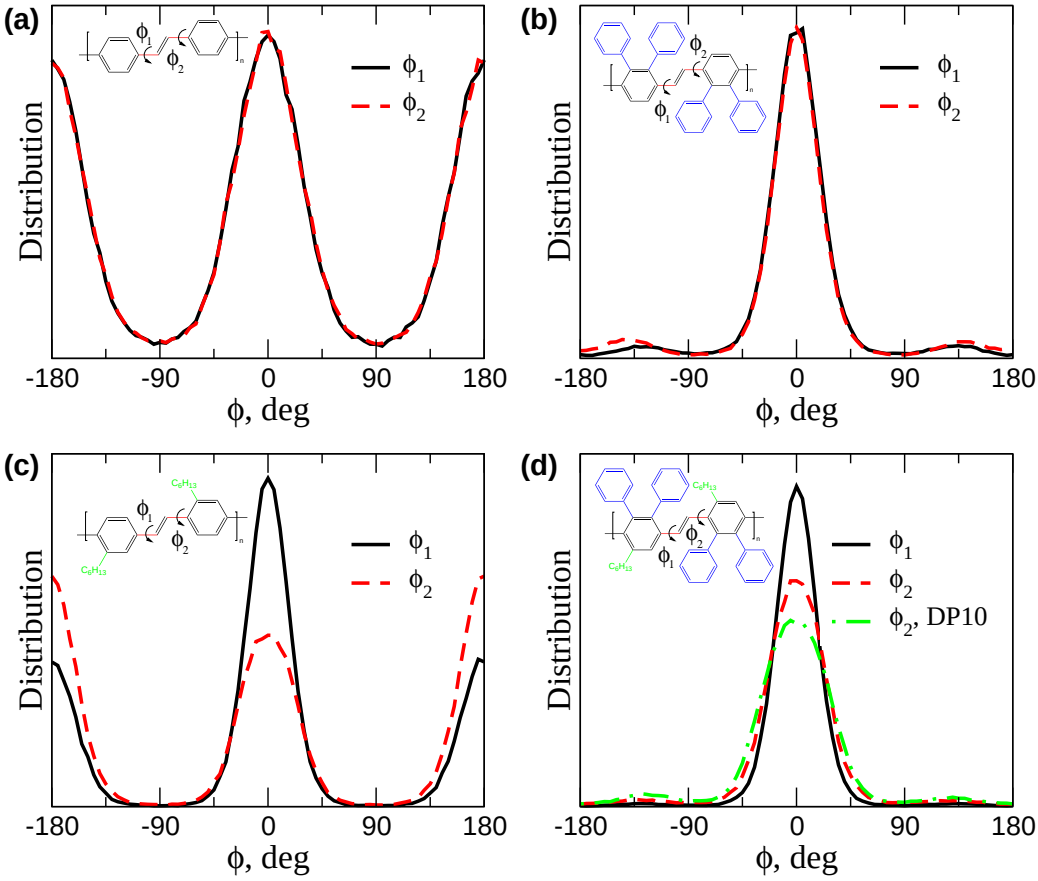

Figure 6.3: Distributions of dihedral angles controlling planarity of the PPV unit. (a) Bare PPV, without the side chains and phenyl rings. (b) Phenyl rings only. (c) Side chains only. (d) DP6-PPV. All results stem from a single chain in vacuum. Insets depict the chemical structure of the corresponding compounds.

the backbone, breaking the conjugation.

To study the effect of varying side chains on the dihedral distributions, we simulated three model systems, all based on a DP6-PPV trimer. The first one, referred to as a "bare backbone" system, had both alkyl side chains and phenyl rings (not belonging to the backbone) substituted with hydrogens. The second system did not have aliphatic side chains, while the last system had the non-bonded interactions of the phenyl rings not belonging to the backbone switched off.

The distributions of the dihedral angles for these three cases are shown in Fig. 6.3. The distributions for a backbone without any side groups, Fig. 6.3a, are rather broad with the maxima located at 0 and ± 180 deg. Due to the symmetry of the "bare backbone" PPV, both distributions are

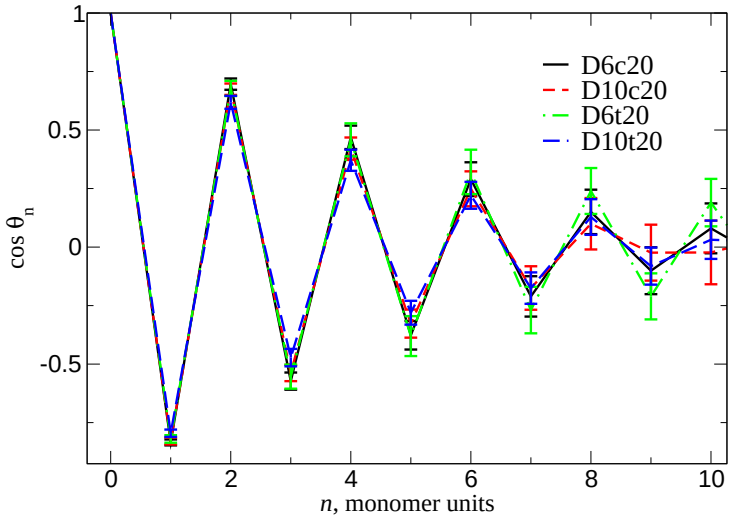

Figure 6.4: Orientational correlations of neighboring monomer units. Correlation functions show that DP-PPV polymers are syndiotactic.

identical and probabilities of finding the system with angles of 0 deg and ±180 deg are the same. The distributions became asymmetric for the systems with the alkyl side chains, as shown in Fig. 6.3c. Here, one of the dihedral angles samples conformations around 0 deg and the other those of around 180 deg. This reflects the fact that the alkyl side chains repel each other and are attached in an asymmetric way with respect to the two dihedrals. Hence, the symmetry of the distributions with respect to the angles is broken. If instead, only the phenyl rings are attached, the cis conformation becomes improbable, as can be seen in Fig. 6.3b. In this case, the $\phi = 0$ deg conformation is more preferable than the $\phi = \pm 180$ deg one. Both effects add up in the distributions of DP6-PPV, which are shown in Fig. 6.3d. Here, the heights of the distributions of the dihedral angles ϕ_1 and ϕ_2 are different at $\phi = 0$ deg, which is due to alkyl side chains. Additionally, conformations with $\phi = \pm 180$ deg are strongly suppressed due to the presence of the phenyl rings. For the DP10-PPV derivative, the situation is qualitatively similar. Namely, longer alkyl side chains lead to an even broader distribution of the dihedral angle ϕ_2.

The final issue we would like to address here is the tacticity of a DP-PPV polymer chain. It is important for choosing an appropriate mapping for a coarse-grained model, as discussed in Sec. 6.2. By analyzing the distributions of the backbone dihedral angles, we have concluded that the presence of the phenyl side groups leads to chain conformations with opposite orientations of alkyl side chains of neighboring repeat units, i.e. DP-PPV is a syndio-

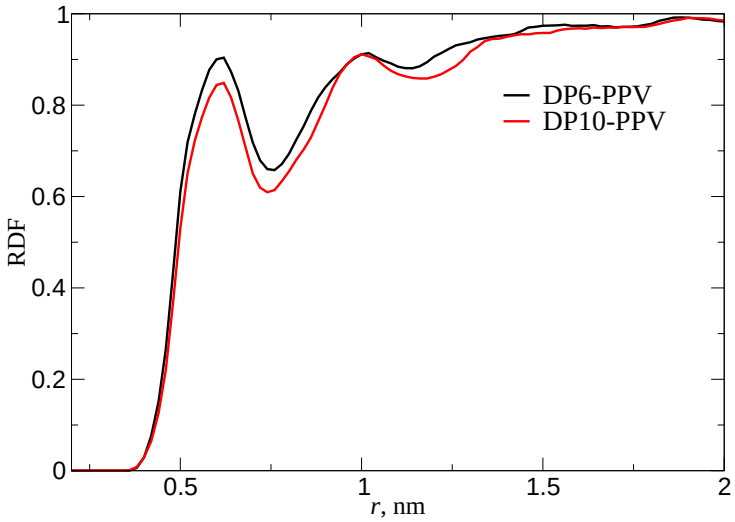

Figure 6.5: Intermolecular radial distribution function for phenyl rings in the melt of DP-PPV oligomers.

tactic polymer. In order to see the effect of solvent and side chain length on correlations of monomer orientations, we calculated the correlation function of a cross product of vectors connecting the atoms 2-4 and 18-19 (see Fig. 3.1b for atom designations). This cross-product, which is a measure of a chain deviation from planarity, is shown in Fig. 6.4. Again, within available accuracy, our model cannot differentiate between toluene and chloroform or hexyl and decyl side chains.

6.1.4 DP-PPV melt

An interesting observation made on the basis of WAXS experiments is that, in a powder, DP6-PPV forms a $\pi-\pi$ complex but DP10-PPV does not [154]. This conclusion was made by analyzing WAXS profiles, which have a sharp maximum at 3 Å in the case of DP6-PPV, and ascribed to the distance between the π-stacked phenyl rings. This maximum was absent in the DP10-PPV powder. The difference in ring packing was explained in terms of the bulkier side chains of DP10-PPV as compared to those of DP6-PPV.

The radial distribution function of the centers of mass of the phenyl rings is shown in Fig. 6.5. As one can see, it is practically identical for the melts of DP6- and DP10 oligomers. Hence, in a melt of oligomers, the average distance between phenyl rings is not sensitive to the length of side chains. Of course, the melt morphology might differ from the non-equilibrium semicrystalline morphology of a polymer powder. Nevertheless, some $\pi-\pi$ com-

plexes should also appear in the melt and their amount and the $\pi-\pi$ distance should depend on the length of the side chains. Since this is not the case, atomistic models hint that the absence of $\pi - \pi$ complexes in DP10-PPV powder is due to the non-equilibrium state of the polymer.

Another interesting experimental observation is the crystallization of polymer backbones upon annealing of a spin-coated film. Crystallization leads to a significantly more pronounced X-ray diffraction peak at a $\pi - \pi$ stacking distance, which is reported to be 3 Å [154]. Fig. 6.5 indeed shows that after annealing, the first maximum of the radial distribution function is located at a separation of about 6 Å. In fact, this separation corresponds to the minimum of the potential of mean force of two coarse-grained beads in vacuum as shown in Fig. 6.6c. It obviously overestimates the $\pi - \pi$ staking distance since the system is in a non-equilibrium state. However, the reported in experiments separation of 3 Å is too small and already in the range of the repulsion of two cofacially aligned repeat units, even without any side chains attached to them [163].

6.1.5 Atomistic molecular dynamics: summary

In summary, the all-atom molecular dynamics simulations give similar and often identical results for DP6-PPV and DP10-PPV in both solvents, i.e. we could not capture the difference in solvent quality. Longer side chains of DP10-PPV as compared to DP6-PPV do neither affect phenyl ring packing in a melt nor dimer-dimer interactions in vacuum and solution, as will be shown in Sec. 6.2.2. Hence, experimentally observed fractal aggregates of DP6-PPV in chloroform or disk-like aggregates of DP10-PPV in toluene, can not be rationalized by atomistic models. However, reasonable agreement with experiments can be expected for a dilute solution of DP10-PPV in chloroform, since in this case no aggregation occurs. In the following section, we will develop a solvent-free coarse-grained model of DP10-PPV in chloroform.

6.2 A coarse-grained model for DP10-PPV

As has already been pointed out in Sec. 6.1, atomistic simulations do not allow an accurate estimation of polymer persistence lengths. Moreover, for studying dilute solutions of DP-PPV in chloroform or toluene, systems with many chains have to be simulated at different concentrations, which is practically impossible at the atomistic level of detail.

In this section, we develop a solvent-free coarse-grained model for DP10-PPV in chloroform and use it to calculate persistence length and static struc-

ture factor as a function of polymer concentration. Both are then compared to the experiment.

To represent a DP-PPV chain on a coarse-grained level, we mapped two monomer units onto one spherical bead, as illustrated in Fig. 6.6d. Two (instead of one) monomer units per bead were used because PPV is syndiotactic (see Sec. 6.1.3). Hence, a coarse-grained representation with one bead per repeat unit is problematic, since repeat units are asymmetric and cannot be properly described by spherically-symmetric pair potentials.

6.2.1 Bonded interaction potentials

The next step of systematic coarse-graining is to determine bonded and non-bonded potentials. Here, only two types of bonded interactions, are included, namely a bond stretching potential between two successive beads and an angle potential between three successive beads. Dihedral potentials are not included as DP-PPVs have a rigid backbone which can be accounted for by the angular potential only. To parametrize the bonded potential, an approach was used in which bonded interactions are obtained from canonical sampling of a single chain in a solvent. The potentials are then obtain by simple Boltzmann inversion. Specifically, a single chain of 20 repeat units of DP10-PPV in chloroform was used. Resulting distributions for both atomistic and coarse-grained sampling are in perfect agreement with each other as shown in Fig. 6.6a,b.

6.2.2 Non-bonded interaction potentials

For non-bonded coarse-grained interaction potentials we used the potential of mean force (PMF) between dimers. In this case, PMF is a free energy of a dimer pair at a specific separation, averaged over all possible mutual orientations of the dimers, as well as positions and orientations of solvent molecules, if present.

Note that non-bonded potentials can not be obtained by IBI or IMC in this case, because it is practically impossible to obtain accurate RDFs for a dilute solution. Hence, to simulate dilute solutions, we used PMF in a solvent as a non-bonded potential.

PMFs for both DP6- and DP10-PPV were calculated in vacuum and solvents, using configurational sampling [117]. Results are shown in Fig. 6.6c. In vacuum, PMF predicts strong dimer-dimer attraction, which might lead to polymer aggregation. On the other hand, fully solvated dimers have a purely repulsive interaction. Calculated PMFs, and hence the coarse-grained non-bonded potentials, correspond to infinitely dilute systems. In practice,

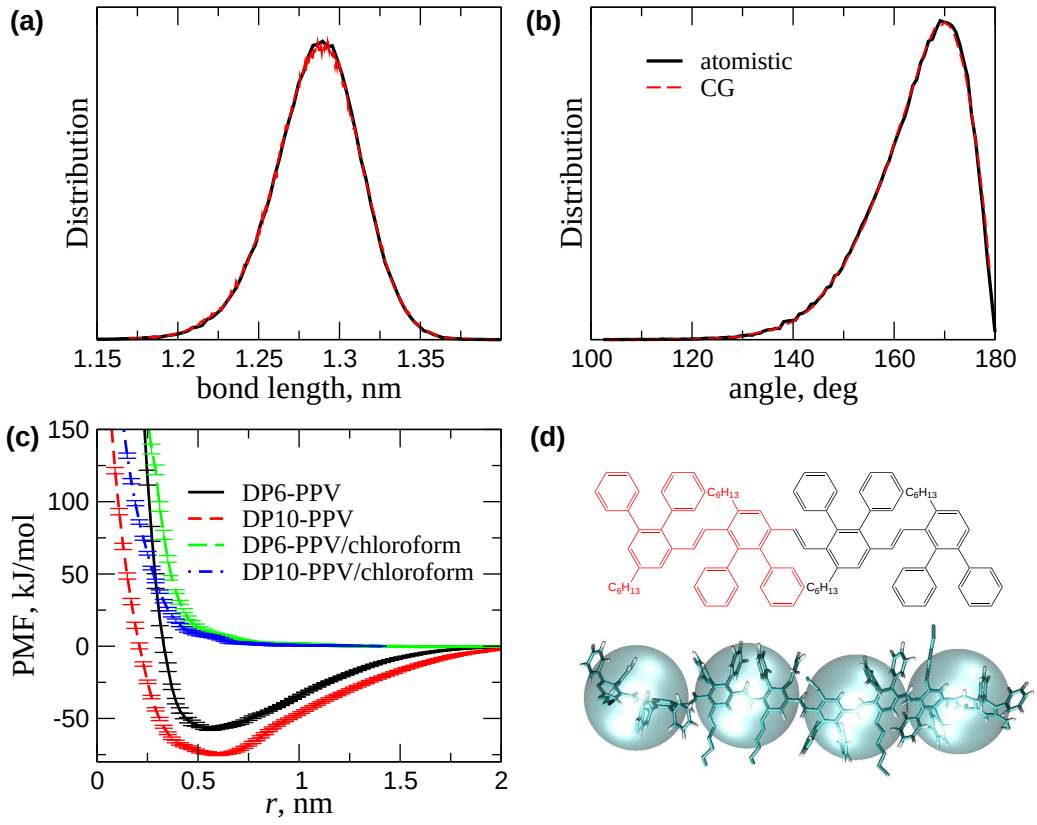

Figure 6.6: Coarse-grained bond (a) and angle (b) distribution functions as obtained from atomistic and coarse-grained MD simulations. A chain of 20 monomers of DP10-PPV in chloroform was used for sampling. Potentials are obtained by Boltzmann-inverting the distributions. (c) Dimer-dimer potential of mean force (PMF), calculated using configurational sampling. PMFs in toluene are identical to those in chloroform. (d) Mapping scheme for the coarse-grained model. When determining centers of coarse-grained beads, no alkyl side chains were included.

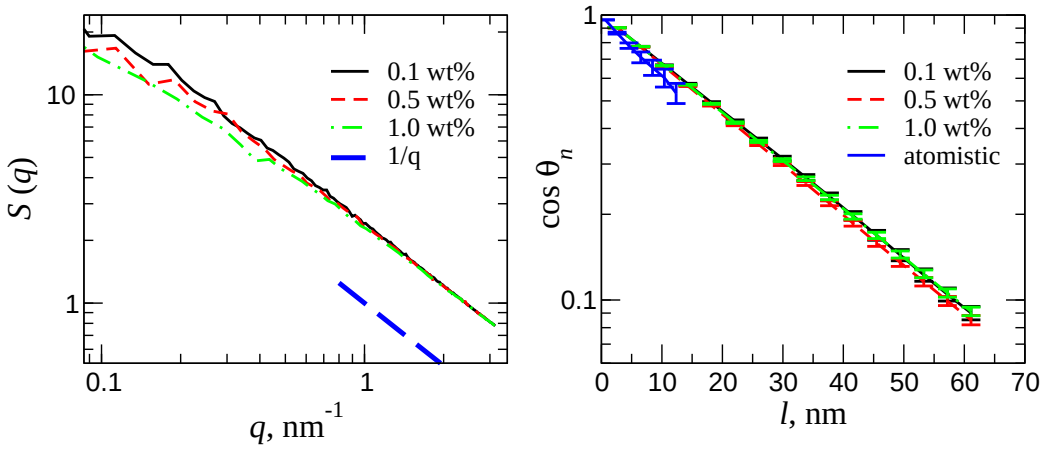

Figure 6.7: Backbone orientational correlations for DP10-PPV/chloroform solutions for different concentrations. Results from atomistic simulations are shown for comparison.

potentials might depend on polymer concentration. However, for systems of up to 1 wt% this dependence may be neglected. In addition, PMFs in both solvents are purely repulsive and only ensure that coarse-grained beads have a certain excluded volume. Hence, simulation results should not be sensitive to the actual functional form of the potential.

6.2.3 Coarse-grained simulations of DP10-PPV in chloroform

Experimentally, three different concentrations of DP10-PPV/chloroform solutions were studied: 0.1, 0.5, and 1.0 wt% [155]. We used the same concentrations as in experiment and compared simulated and measured polymer persistence lengths and static structure factors.

The systems consisted of 512 polymer chains, 50 beads each, which corresponds to 100 monomer units. NVT simulations were performed for 16 ns and 800 frames were used to compute static structure factors. All coarse-grained simulations were performed at 300 K.

The calculated structure factors, $S(q)$, for different polymer concentrations are shown in Fig. 6.7a. One can see that $S(q)$ scales as q^{-1} in the high-q region, irrespective of concentration. This is due to the rod-like nature of the segments constituting the DP10-PPV chains [164]. In the low-q region, intensity decreases with an increase of concentration. This dependence is determined by the dynamic network structure formed by the interchain overlap in the semidilute solution [165].

concentration, wt%	l_p, nm
0.1	6.4
0.5	13.3
1.0	23.0
MD	25.0 ± 0.5

Table 6.1: Estimations of persistence length from SANS experiments [155] and coarse-grained MD simulations. MD predicts the same value for all three concentrations.

Backbone orientational correlations are shown in Fig. 6.7b. These correlations decay exponentially on the accessible length scale, which allows estimations of polymer persistence lengths as discussed in Sec. 6.1.2. The calculated persistence length of 25 ± 0.5 nm is independent of concentration, contrary to experimental findings summarized in Table. 6.1. In experiments, persistence length was extracted by fitting SANS profiles to the wormlike chain model with excluded-volume interactions [155].

The increase with concentration can be attributed to chain aggregation. At low concentrations, chains do not interact and there is no aggregation. Upon increasing the concentration, aggregates start to form and inside those aggregates, chains become locally more extended which increases their persistence length. Our model can not capture this effect since we have repulsive effective potentials. Hence, no driving force is present, which might lead to polymer aggregation at higher concentrations.

6.3 Discussion

Our simulations confirm experimental evidence that both chloroform and toluene are good solvents for DP-PPVs. This is based on alkyl side chain stretching, backbone orientational correlations, and the potential of mean force (PMF) of DP-PPV dimers in the solvent. However, the experimentally reported quantitative results on solvent quality could not be reproduced.

Atomistic simulations show that the difference in side chain lengths of DP6- and DP10-PPV does not affect chain packing in a melt of oligomers, which is practically identical for both DP6- and DP10-PPV. Interdimer interactions in solutions are also not strongly affected by the difference in side chain lengths, namely the interdimer PMF is repulsive for both DP6- and DP10-PPV. This seems to contradict experimental observations which find that DP6-PPV forms aggregates in solution, whereas DP10-PPV does not. A possible reason for this inconsistency might be that in experiments, the system is in its non-equilibrium state, since powder, in which polymer chains

are semicrystalline, was used to prepare the solution. Furthermore, as can be found in Ref. [155], DP10-PPV does aggregate in toluene solutions. However, if the system is heated up to 85°C and then cooled down, these aggregates tend to disappear. Recently, an alternative explanation of anomalous aggregation in good polymer solutions has been proposed [166]. This mechanism is applicable to stereoregular polymers, such as DP-PPV, which indeed shown to be syndiotactic in our simulations.

Coarse-grained simulations show that, for very low polymer concentrations (0.1 – 1.0 wt%), polymer persistence length does not depend on polymer concentration. This again contradicts the experimental picture, where the increase of persistence length at 1.0 wt% is rationalized as chain aggregation. The overestimation of the polymer persistence length in simulations might be due to the presence of tetrahedral chemical defects in real samples [167]. In these defects conjugated carbon-carbon bonds are replaced by tetrahedral ones. They divide polymer chains into structurally identifiable quasi-straight segments and reduce orientational correlations, which leads to a decrease in persistence length.

Finally, we should mention that various approximations in our simulation models might also lead to a disagreement between experiments and simulations. For example, as experimental data for conjugated compounds is rather sparse, force-field validation becomes problematic. To this end, although we have reproduced density, melting temperature, and crystal structure of stilbene, this is no guaranty that thermodynamic properties such as solvation free energy are correctly reproduced. Other issue can be limited accessible length- and time-scales. Explicit solvents and rather stiff polymer backbones prohibit the study of global chain conformations and chain self-interactions via atomistic molecular dynamics. Solvent-free coarse-grained models are capable of simulating bigger boxes on longer timescales. These, however, lead to additional approximations, which cannot be easily controlled.

Chapter 7

Morphology and charge transport in amorphous Alq_3

7.1 Introduction

Tris(8-hydroxyquinoline)aluminium (Alq_3, Fig. 7.1a) is a commonly used organic semiconductor with a higher electron than hole mobility [168]. The first organic light emitting diodes were made with Alq_3 and since then it has played the role of a "guinea pig" compound in organic electronics [6]. The distinctive properties of Alq_3 are its green light emission and good electron mobility. It is also used as a host material for emissive dopants of lower emission energy.

In organic semiconductors both emission and charge transport properties are extremely sensitive to the material morphology. The latter can be controlled by an appropriate processing or compound derivatization. Understanding the link between the structure, morphology, and macroscopic properties of organic compounds is the first step towards their rational de-

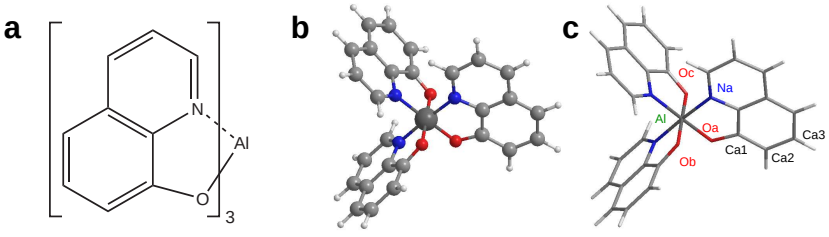

Figure 7.1: (a) Chemical structure of tris-8(hydroxyquinoline aluminium) (Alq_3). (b) Meridional isomer of Alq_3. (c) Atom labeling used to show the distributions of angles and dihedrals. Largest linear dimension of the molecule is ≈ 1 nm.

sign [74].

In spite of numerous experimental studies, there were only two theoretical/computational attempts to relate the morphology of amorphous Alq_3 to its charge carrier mobility [169, 68]. In Ref. [169] the rigid-body approximation for the intramolecular structure was used together with the non-bonded parameters taken from the Dreiding force field. Approximate force constants for bonded parameters and non-bonded parameters of the Williams 99 force field were used in Ref. [68]. In both cases neither the force field nor the generated amorphous morphology was checked against the experimental data, and its characterization was reduced to pair correlation functions and density at ambient conditions.

In this chapter, we first derive bonded parameters for the Alq_3 molecule using first-principles potential energy scans. Second, we compare OPLS and Williams 99 force fields to each other as well as to the available experimental data. We further assess how the morphological difference, resulting from using different force field parameters, affects charge transport. Finally, we demonstrate that the dipolar energetic disorder (see Sec. 2.4) must be taken into account in order to describe charge dynamics correctly. We analyze the influence of energetic disorder in terms of pathways and occupation probabilities and show that the spatial correlations of site energies must be taken into account in order to reproduce Poole-Frenkel behavior (see Sec. 2.6) of the charge carrier mobility.

7.2 Force-field parameters

As already explained in Chapter 3, force field refinement can be split into several steps: (i) determination of partial charges; (ii) calculation of the force field parameters for the bonded interactions; (iii) parameterization of the non-bonded interactions.

Partial charges were obtained from electrostatic potentials using a grid based method (CHELPG) [93] after the geometry optimization of a meridional Alq_3 isomer[1] in vacuum (see Fig. 7.1b) using the B3LYP functional and 6-311g(d) basis set. The corresponding molecular dipole moment was $d = 4.4\,\mathrm{D}$. The DFT results were checked against MP2 calculations with 6-311g(d) basis set, for which the dipole moment is 5.4 D. No substantial differences were noticed with either B3LYP or MP2 charges used. All results

[1] Another isomer which is commonly observed in the octahedral complexes of the type MN_3O_3, where M is a trivalent metal, namely the *facial* isomer, has not been experimentally identified for amorphous Alq_3 [170].

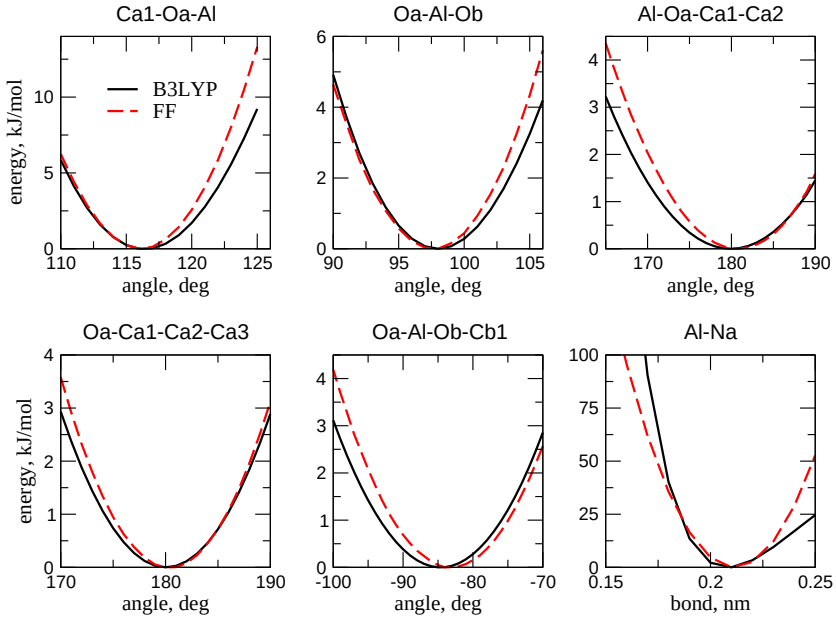

Figure 7.2: B3LYP/6-311+g(d,p) energy scans (full lines) together with fits using OPLS force field (dashed lines) for several representative degrees of freedom.

reported here were, therefore, performed with charges based on calculations using B3LYP functional.

Bonded interactions were parameterized by matching the first-principle potential energy scans with the corresponding force field based scans [77, 76, 72]. Note that the meridional isomer of Alq_3 is not symmetric, for example oxygen-aluminium-oxygen angle for Oa-Al-Ob is different from Ob-Al-Oc, as shown in Fig. 7.1c. The same holds for the dihedrals O-Al-O-C. In total 16 scans (see table 7.1) were performed in order to capture the potential energy surface of the molecule correctly. The potential energy scans were performed with GAUSSIAN 03 program [171] using B3LYP functional and 6-311g+(d,p) basis set. The results, together with the fits based on the atomistic force field for some of them are shown in Fig. 7.2. The resulting force field constants are summarized in table 7.1. These constants are used to calculate angle and dihedral potentials using the following expression: $V(r) = \frac{1}{2}k_\theta(\theta - \theta_0)^2$. The harmonic spring potential between Al and N was added to model the ionic bond of the Alq_3 molecule. For simplicity, the scan was performed only for the Al-Na bond (see Fig. 7.2), but the resulting potential was applied also to Al-Nb and Al-Nc pairs.

From Fig. 7.2 one can see that the parameterized dihedral potentials (Al-Oa-Ca1-Ca2, Oa-Ca1-Ca2-Ca3, Oa-Al-Ob-Cb1 and the corresponding

Degree of freedom	Constant k_θ, $kJmol^{-1}rad^{-2}$
Ca1-Oa-Al	600
Cb1-Ob-Al	500
Cc1-Oc-Al	400
Oa-Al-Ob	320
Ob-Al-Oc	200
Oc-Al-Oa	420
Al-Oa-Ca1-Ca2	55
Al-Ob-Cb1-Cb2	50
Al-Oc-Cc1-Cc2	50
Oa-Ca1-Ca2-Ca3	200
Ob-Cb1-Cb2-Cb3	200
Oc-Cc1-Cc2-Cc3	200
Oa-Al-Ob-Cb1	50
Oa-Al-Oc-Cc1	30
Oc-Al-Oa-Ca1	20
bond Al-Na	62459, $kJmol^{-1}nm^{-2}$

Table 7.1: Bonded interactions parameters, obtained using the fitting described in the text.

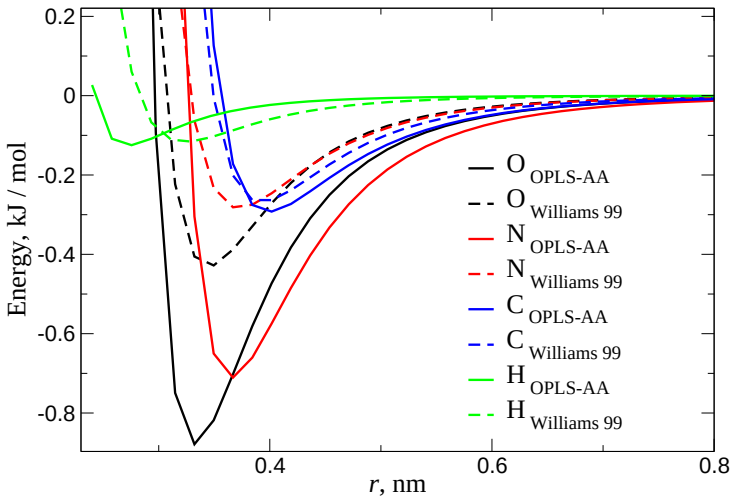

Figure 7.3: Non-bonded O-O, N-N, C-C and H-H interactions for the OPLS and Williams 99 force fields. OPLS force field has bigger binding energies for N and O. Note that interactions between different atom types, e. g. N and O are calculated using the OPLS mixing rules [80].

atom	OPLS		W99 fitted	
	σ, nm	ϵ, kJ/mol	σ, nm	ϵ, kJ/mol
Carbon	0.355	0.29288	0.348873	0.263802
Hydrogen	0.242	0.12552	0.277899	0.107032
Oxygen	0.296	0.87864	0.305762	0.428992
Nitrogen	0.325	0.71128	0.327197	0.282361

Table 7.2: Non-bonded parameters of OPLS and Williams 99 force fields used in our simulations. For Williams 99, effective Lennard-Jones parameters obtained by fitting Buckingham potential are given (W99 fitted).

potentials for the other ligands) are quite soft. This is consistent with the fact that the Al-N bonds can be characterized as ionic rather than covalent. Consequently the angle values we used for parametrization do not cover the whole region of the dihedral angles, which can be sampled at 300 K during the MD simulation. However, the extension of the scans beyond those values is problematic, because at large deviations from an equilibrium geometry van-der-Waals non-bonded interactions between ligands start to play a significant role, which can not be correctly described by the density functional theory. Instead of extending the region of the scans we verified that the potential of mean force (PMF) of the dihedral distributions for amorphous morphology (see below) reproduces the scans obtained from the DFT. Additionally we compared the distributions of the representative degrees of freedom obtained from the force field simulations with the results of *ab-initio* molecular dynamics (see Appendix C).

The Lennard-Jones parameters for all atoms were taken either from the OPLS [80] or Williams 99 [172] force field. Note that the OPLS force field uses the Lennard-Jones (12-6) potential while Williams 99 is based on the Buckingham potential, $V(r) = A\exp(-Br) - \frac{C}{r^6}$. The only difference between the two functional forms is the repulsive part, and one can refit the Buckingham potential with the Lennard-Jones one, which also results in a speed up of about 2.5. The non-bonded force field parameters are summarized in table 7.2.

The two force fields are compared in Fig. 7.3. One can see that the binding energy for nitrogen and oxygen is much bigger for the OPLS than Williams 99 force field and should result in a more dense molecular packing. Aromatic carbons have similar parameters.

To verify the derived force field, we studied amorphous morphologies of Alq_3. The systems were prepared by first arranging the molecules on a cubic lattice with the density adjusted in order to avoid molecular overlaps and

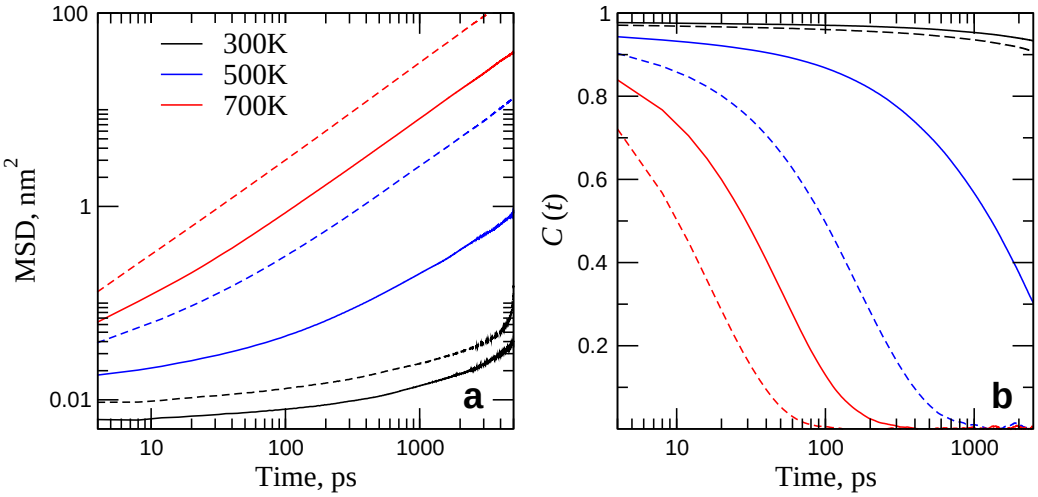

Figure 7.4: (a) Log-log plot of the mean squared displacement of the center of mass of Alq_3 molecules for a set of temperatures. (b) Correlation function of a Alq_3 molecular orientation, defined as a normal vector to the plane containing the three oxygens. Solid line: OPLS force field. Dashed line: Williams 99 force field.

then equilibrating at 700 K in an NPT ensemble, with a stochastic velocity rescaling thermostat [84] and the Berendsen barostat [81], and finally cooling down to room temperature. All simulations were performed using the GROMACS package [173]. Note that two alternative ways of system preparation were reported previously. The first one, based on the Monte Carlo algorithm mimicking the film deposition process [169], is limited to rigid molecules. The second, based on a compression of a simulation box of a randomly positioned molecules at low density [68] is sensitive to the compression rate and often results in voids.

To monitor the equilibration dynamics we calculated relaxation times from the correlation functions for the rotational molecular motion as well as mean squared displacement (MSD) of their translational motion. Both are shown in figs. 7.4a,b for a set of temperatures. Figs. 7.4a,b indicate that systems can be equilibrated at 700 K after about 1 ns run, which is the time needed for the center of mass of a molecule to travel a distance comparable to its size. In addition, after this period of time molecular orientations are fully uncorrelated. At lower temperatures significantly longer runs are required. Therefore, to prepare large-scale morphologies, we annealed our systems at 700 K and then cooled them down to the desired temperature with the cooling rate 0.2 ps/K.

To compare molecular packing in amorphous morphologies generated with

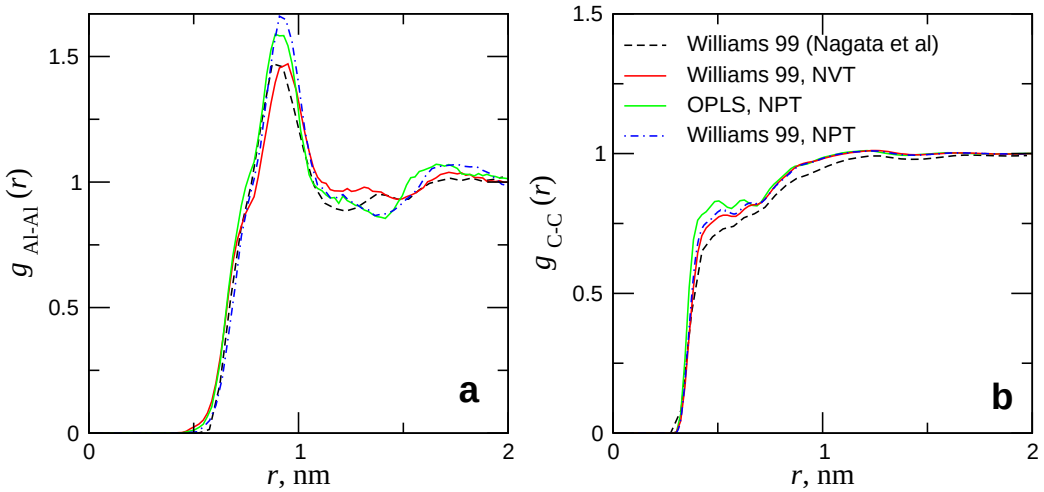

Figure 7.5: (a) Al-Al and (b) C-C radial distribution functions for OPLS and Williams 99 force fields obtained after equilibration in NPT ensemble at 300 K. For comparison, results of NVT simulations (note different density) of Ref. [68] are also shown.

the OPLS and Williams 99 force fields, we calculated two radial distribution functions, which are shown in figs. 7.5a,b. One can see that different *bonded* parameters used in this work and Ref. [68] do not significantly change the radial distribution function. Note, however, that these two curves are compared for a fixed density of Ref. [68]. On the other hand, if the systems are equilibrated in the NPT ensemble (which results in higher density), radial distribution functions have a more pronounced structure (better resolved peaks) for both OPLS- and Williams 99 based force fields. The difference between the two force fields is small, however. One can conclude that the radial distribution function is not an optimal measure of molecular packing, due to its limited sensitivity.

Other two macroscopic quantities are known from experiments: the density of amorphous Alq_3 at ambient conditions and its glass transition temperature [174, 176]. The density versus temperature plot is shown in Fig. 7.6a. One can see that the system density strongly depends on the Lennard-Jones parameters (and does not change if we substitute ESP charges derived from the B3LYP- with MP2-based electron density). One can see that OPLS predicts a higher density than Williams 99 and it is also closer to the experimental values.[2] Densities used in previous simulation studies by Y. Nagata et

[2]Note, that we intentionally did not try to reproduce the experimental density of the amorphous Alq_3 from Ref. [174] exactly. In Sec. 7.3 it is demonstrated, that the difference in density does not influence charge transport properties significantly.

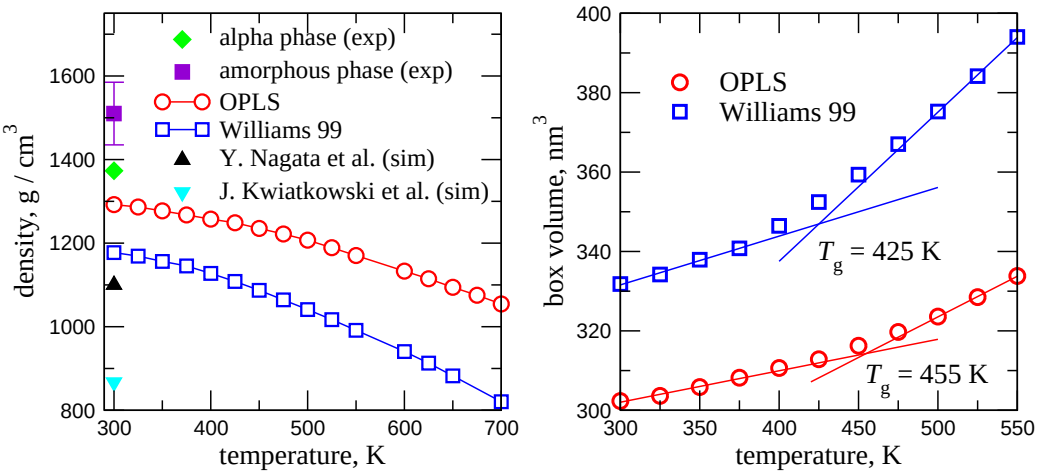

Figure 7.6: (a) Density versus temperature for OPLS and Williams 99-based force fields together with experimental results: amorphous density is taken from [174], alpha-phase density is taken from [175]. Density used in previous studies by Y. Nagata et al. [68] and J. Kwiatkowski et. al [169] are also shown. (b) Volume of the simulation box versus temperature for amorphous Alq_3. Intersection of two fitting lines is used to predict glass transition temperature, which is estimated to be 455 K for OPLS and 425 K for Williams 99 force field. Experimental value is 448 K [176].

al. [68] (density is mentioned in the paper) and J. Kwiatkowski et. al. [169] (density is calculated using number of molecules and the box size specified in the paper) are also shown for comparison.

To estimate the glass transition temperature we calculated the volume of the simulation box as a function of temperature. These dependencies are shown in Fig. 7.6b. By performing a linear fit of the low and high temperature parts of the curves one can find the intersection point which provides an estimate of the glass transition temperature [177]. The latter is estimated to be 455 K for OPLS and 425 K for Williams 99 force field. Experimental value is 448 K [176], which is again closer to the OPLS-predicted value.

Summarizing, we can conclude that the OPLS-based force field performs better for the amorphous Alq_3 films than the Williams 99 force field by predicting closer to the experimental values of density and glass transition temperature.

7.3 Relationship between morphology and charge transport

In this section we use the multiscale approach presented in Sec. 2.7 to simulate charge dynamics in two different morphologies described above and to understand how the change in morphology influences charge carrier mobility. For simplicity we ignore the energetic disorder in these considerations, but the more detailed analysis shows, that the conclusions drawn from the simplified picture stay valid also if electrostatic contribution to the energetic disorder is taken into account.

Within the Marcus theory under the assumption that energetic disorder is negligible, charge transport is governed only by two parameters: the reorganization energy and the transfer integral (see Chapter 2). The reorganization energy λ was computed as discussed in Sec. 2.2 using B3LYP functional with a triple zeta split basis set, 6-311g+(d,p), using the GAUSSIAN 03 [171] program. The computed values for the reorganization energies are $\lambda_h = 0.23\,\text{eV}$ and $\lambda_e = 0.28\,\text{eV}$. The transfer integrals J were computed using the intermediate neglect of differential overlap level of theory (ZINDO) as described in Sec. 2.3. Molecular orbitals are calculated using the GAUSSIAN 98 package [178]. Since the transfer integral is related to the molecular overlap, it is very sensitive to relative orientations and positions of the neighbors. The corresponding HOMO and LUMO orbitals are shown in Fig. 7.7a. Note that ZINDO predicts that the LUMO is localized on one of the arms, while DFT calculations of Ref. [169], as well as our MP2 calculations (see Fig. 7.7a) show that the LUMO is delocalized over two of them. Different localization may affect charge dynamics, especially the ratio between the hole and electron mobilities, since the delocalization leads to the less pronounced dependence of transfer integrals on the mutual orientation of neighbors.

However, here we do not aim at a quantitative description of charge dynamics, since we anyway ignore several essential ingredients required for its correct description, such as energetic disorder as well as electrostatic and polarization contributions to it. Our main goal is to understand how the change of morphology influences charge carrier mobility.[3] Hence, here we analyze only the hole transport, since all three first-principles methods predict similar localization pattern for HOMO.

For charge dynamics simulations amorphous morphologies containing 4096 molecules in a cubic box were prepared using both OPLS and Williams 99 force fields and the aforementioned equilibration procedure. The size of the

[3]Of course one has to keep in mind that the energetic disorder might also be sensitive to the underlying morphology.

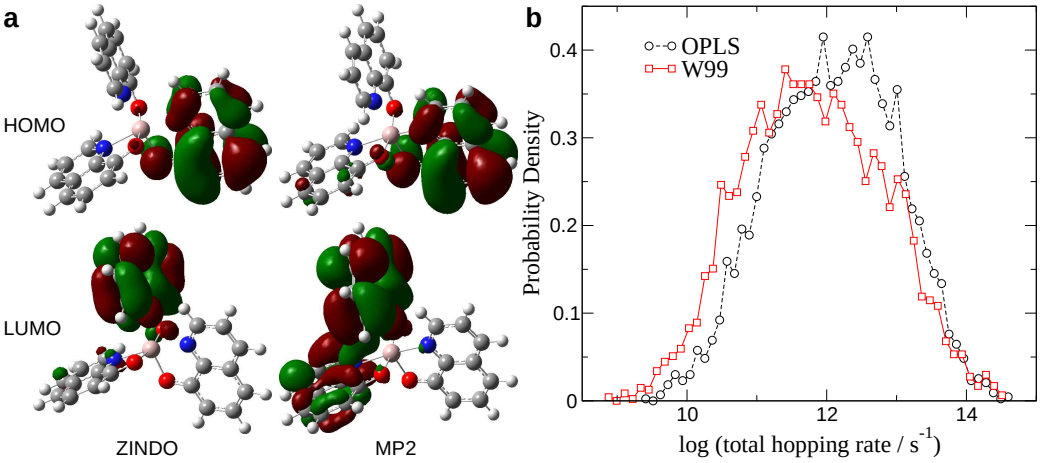

Figure 7.7: (a) HOMO and LUMO orbitals of a meridional isomer of Alq$_3$ calculated with ZINDO and Hartree-Fock (using MP2-optimized geometry) methods. (b) Probability distribution of the total hopping rate away from the molecule for OPLS and Williams 99 force fields. Only hole transport was considered. ZINDO orbitals were used for calculations.

cubic box was 13.4263 nm, the length of the KMC run was 10^{-5} sec, and the mobilities were averaged over 10 starting points. A cutoff of 15 Å was used to calculate the hopping rates between neighboring molecules.

We first have a look at the probability distributions of the total hopping rate away from a molecule:

$$\omega_i^{total} = \sum_j \omega_{ij} \tag{7.1}$$

where the sum is taken over all neighbors of the molecule i.

The result is shown in Fig. 7.7b. One can notice that the distribution that corresponds to the OPLS force field is shifted to higher values of outgoing hopping rates compared to the Williams 99 force field. This is expected, since the OPLS force field predicts more dense molecular packing.

Charge carrier mobilities were calculated by monitoring the projection of the charge velocity on the field direction. The simulations were done for a single charge diffusing in an external field through the configuration of one molecular dynamics snapshot with periodic boundary conditions applied in all directions. The mobility was calculated as $\mu = v/E$, where v is the averaged over a hundred different starting positions for every frame velocity of the charge carrier along the field E.

The calculated hole mobilities are $\mu^h_{\mathrm{OPLS}} = (3.7 \pm 0.1) \times 10^{-3}\,\mathrm{cm^2V^{-1}s^{-1}}$ and $\mu^h_{\mathrm{W99}} = (1.7 \pm 0.1) \times 10^{-3}\,\mathrm{cm^2V^{-1}s^{-1}}$. Both values are several orders of

magnitude higher than the experimentally measured ones: electron mobilities of amorphous Alq_3 are of the order of $10^{-5} - 10^{-6}\,cm^2V^{-1}s^{-1}$ [168, 179] while hole mobilities are of the order of $10^{-8}\,cm^2V^{-1}s^{-1}$ [168]. The discrepancy between the experimental and calculated values is due to the fact that we ignore the energetic disorder, which can significantly influence resulting mobilities [169, 68]. The effect of energetic disorder is discussed in the next section.

To conclude, the difference in density and packing induced by different force field parameters does not significantly affect the charge dynamics, leading to slightly higher mobilities in case of a more dense system (OPLS-based morphology).

7.4 Role of energetic disorder

Considerations of the Sec. 7.3 ignore energetic disorder. However, the dipole moment of an Alq_3 molecule is rather large ($d = 4.4\,D$, see Sec. 7.2). Hence, dipolar disorder cannot be neglected (see Chapter 2). In this section we study the influence of energetic disorder on charge dynamics in amorphous Alq_3 by comparing carrier mobilities calculated with and without the disorder. Visualization of hopping site occupation probabilities and current filaments further suggests the important role of energetic disorder.

7.4.1 Computational details

We assume that there are two contributions to the free energy difference ΔG_{ij} (see Eq. 2.9): one due to an externally applied electric field and the other one due to different electrostatic interactions of the excess charge with the surrounding neutral molecules: $\Delta G_{ij} = e\mathbf{Er} + \Delta G_{el}$. The electrostatic contribution, $\Delta G_{el} = E_i - E_j$, is calculated classically as it is described in Ref. [56] using CHELPG-fitted partial charges for a charged and neutral Alq_3 molecule in the ground state obtained from DFT calculations using the B3LYP hybrid functional and the 6-311g(d) basis set. We denote the partial charge of the atom a in the molecule i as q_i^a if the molecule is neutral and as q_i^{a*} if the molecule is charged. ΔG_{ij}^{el} is then given by [56]:

$$\Delta G_{ij}^{el} = \sum_c \sum_{k\neq i} \sum_n \frac{q_i^{c*} q_k^n}{4\pi\epsilon_0\epsilon r_{cn}} - \sum_c \sum_{k\neq j} \sum_n \frac{q_j^{c*} q_k^n}{4\pi\epsilon_0\epsilon r_{cn}}\,, \tag{7.2}$$

with the permittivity of free space ϵ_0 and a relative macroscopic dielectric constant ϵ. The distance between two atoms is given by r_{cn}. The first

sum runs over all partial charges q_i^{c*} of the initially charged molecule i that interact with all partial charges q_k^n in all other $k \neq i$ neutral molecules. It represents the energy of the system if the molecule i is charged. Second term is the energy of the system when the molecule j is charged. This procedure is exact in the limit of an infinite system. We use no spherical cutoff but apply the nearest image convention, and have checked that this converges already for a few hundered molecules in the simulation box.

Polarization effects are treated phenomenologically by using a distance-dependent dielectric constant [180]:

$$\epsilon = \epsilon(r_{cn}) = \epsilon_\infty - (\epsilon_\infty - 1)\mathrm{e}^{-sr_{cn}}(1 + sr_{cn} + 1/2s^2r_{cn}^2). \tag{7.3}$$

where $\epsilon_\infty = 3.0$ (experimental value) [181] and $s = 0.3\,\text{Å}^{-1}$ [68]. Effectively, atoms that are close in space feel the unscreened Coulomb interaction with $\epsilon = 1$ which is then screened to the bulk value ϵ_∞ at larger distances.

Charge dynamics is simulated using the kinetic Monte Carlo technique using the (extended) VOTCA package [2]. We use periodic boundary conditions and calculate charge carrier mobility as

$$\mu = \frac{\langle v_E \rangle}{E} \tag{7.4}$$

where $\langle v_E \rangle$ is the averaged projection of carrier velocity on the direction of the applied electric field E.

7.4.2 Poole-Frenkel behavior

Mobilities with and without disorder, for different values of an applied electric field are shown in Fig. 7.8. One can see that taking energetic disorder into account leads to significantly lower mobilities. More important, it results in the correct electric field dependence of the mobility: $\mu \propto \exp(\beta E^{1/2})$ which is known as a Poole-Frenkel behavior (see Chapter 2) and was observed experimentally for amorphous Alq_3 [168, 26, 179]. The distribution of site energies resulting from electrostatic interactions is shown on the inset of Fig. 7.8. It has a Gaussian shape with a standard deviation of $\sigma = 0.14\,\text{eV}$.

Another important observation is that, when energetic disorder is taken into account, charge carrier mobilities are system-size dependent. This indicates that charge transport becomes dispersive, as was discussed in Chapter 2. As a consequence, calculated dispersive mobilities disagree with the experimental reference data that is obtained in the nondispersive regime (see Fig. 7.8). A way of extracting the absolute value of nondispersive mobility from simulations of small systems is addressed in Chapter 8.

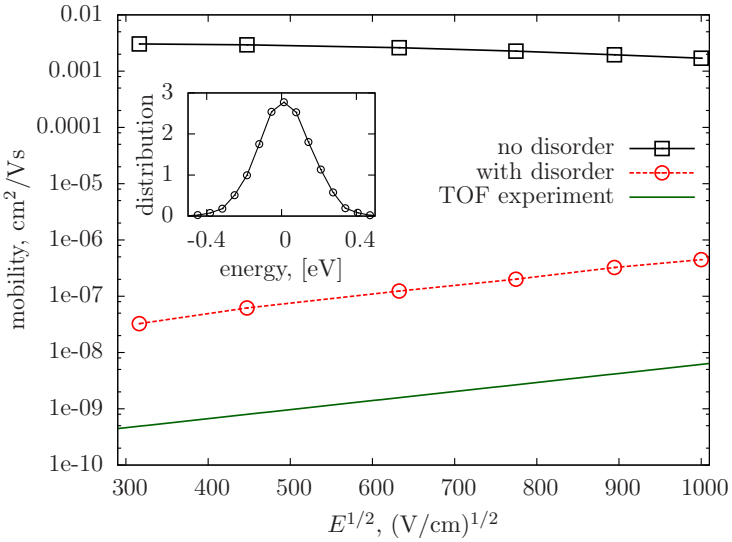

Figure 7.8: Hole mobility as a function of applied electric field of a system with 13824 Alq_3 molecules with and without energetic disorder. Results of the time-of-flight experiment are taken from Ref. [26]. Inset shows the site energy distribution. It is Gaussian with $\sigma = 0.14\,\text{eV}$.

7.4.3 Charge dynamics visualization

In the previous section it was shown that energetic disorder significantly affects charge dynamics. Taking energetic disorder into account lowers calculated mobilities and changes its electric field dependence. Here we visualize the difference between the systems with and without disorder using charge carrier pathways and site occupations.

Occupation probability of a hopping site is defined as a fraction of time a charge carrier spends on it. It is normalized in a sense that the sum of all the occupation probabilities over all hopping sites is 1. Here we present the results for an amorphous morphology of 512 Alq_3 molecules.

Fig. 7.9 compares the distributions of occupation probabilities calculated with and without energetic disorder. When energetic disorder is ignored, the distribution of occupation probabilities is narrow, since all probabilities are of the same order of magnitude. When energetic disorder is taken into account, the distribution broadens, spanning several orders of magnitude. This is illustrated in the inset of Fig. 7.9, where hopping sites are shown with colored spheres. Each color corresponds to a particular order of magnitude of the occupation probability. The radius of a sphere is proportional to the $\log_{10}$ of the occupation probability. When energetic disorder is ignored, all spheres have the same color, meaning that each site has the same order of

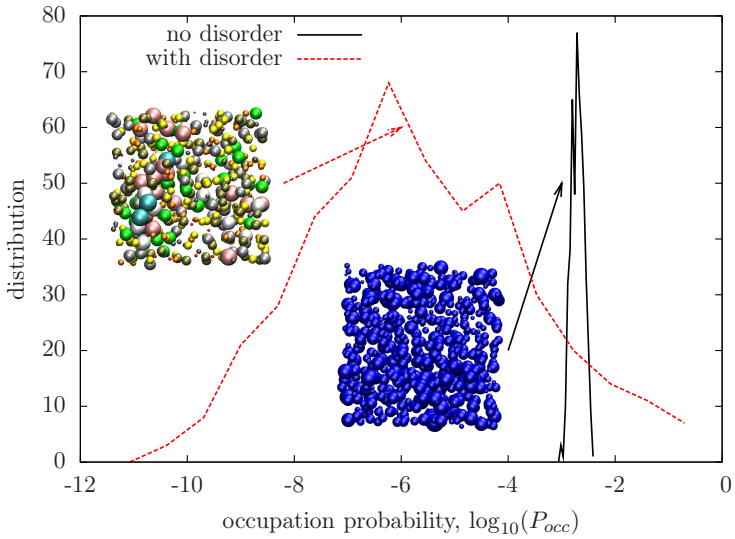

Figure 7.9: Distribution of occupation probabilities with and without energetic disorder. Insets show the occupation probabilities of the same order of magnitude with the same color. Occupation probabilities were calculated for one trajectory snapshot in a 10^{-3} sec KMC simulation

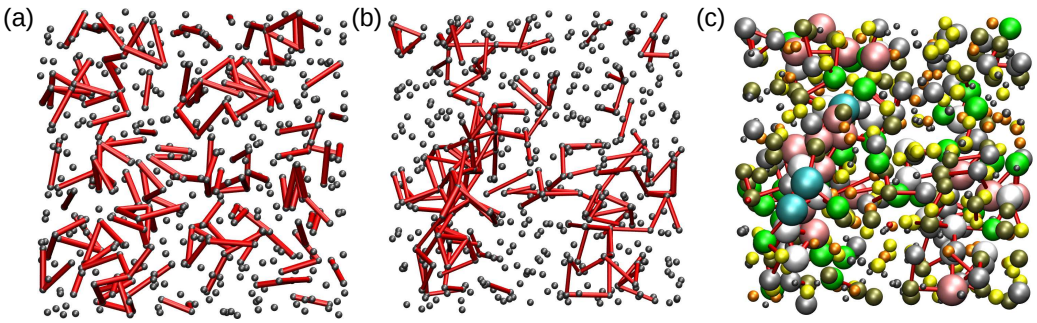

Figure 7.10: (a) Probability current graph. Without (a) and with (b) energetic disorder; (c) The same graph as in (b), but with occupation probabilities as in Fig. 7.9

magnitude for the occupation probability. When energetic disorder is taken into account, a colorful picture is obtained instead, which allows to identify hopping sites with the highest occupation probabilities. These sites, which also have the largest radius, are shown in cyan and pink. One can see that the sites with the largest occupation probability are not randomly distributed in space. They form percolating pathways, which span the simulation box from bottom to top, see Fig. 7.9. The use of percolation theory seems appealing for analyzing the finite-size effects in such a system [182]. If energetic disorder is small cluster analysis approach of Ref. [58] can be used to relate the "effective" transfer integral in the system with the charge carrier mobility. In our case energetic disorder is significant and transfer integral J_{ij} alone does not determine the charge carrier mobility. The use of the charge transfer rate ω_{ij} instead is also problematic, because of its asymmetry: $\omega_{ij} \neq \omega_{ji}$. Influence of the finite system size on the charge carrier dynamics is analyzed in Chapter 8.

To quantify the spatial distribution of occupation probabilities, we also calculated the probability current. It is defined by each pair of neighboring hopping sites i and j as follows:

$$J_{ij}^{v} = \left| \sum_{\text{hops } i \to j} \frac{d_{ij}}{\tau_i} - \sum_{\text{hops } j \to i} \frac{d_{ij}}{\tau_j} \right| . \tag{7.5}$$

Here the first sum goes over all hops from i to j observed in a kinetic Monte Carlo run. The second sum is over all backward hops. τ_i and τ_j are the corresponding waiting times, d_{ij} is the distance between the sites i and j. A large value of the probability current for a particular pair of sites means that this pair considerably contributes to the probability current. A collection of such pairs thus represents most favorable pathways.

In order to visualize the probability current we show the bonds between 200 pairs with the highest value of the current in Fig. 7.10. Without the energetic disorder, the most favorable pathways are uniformly distributed in space and a charge carrier is equally likely to take any of the pathways. When the energetic disorder is taken into account, a charge carrier moves on a very rough energy landscape, and some pathways are sampled more frequently than the others. The distribution of the most favorable pathways, Fig. 7.10(b), clearly correlates with the occupation probabilities, Fig. 7.9(c). To emphasize the correlation they are shown together in Fig. 7.10(c).

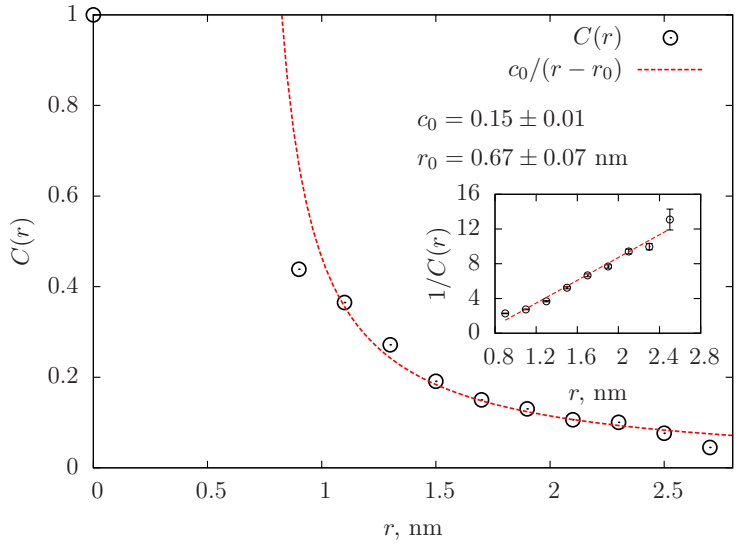

Figure 7.11: Reduced site energy correlation function calculated for the system containing 13824 Alq$_3$ molecules. Inset illustrates the linear fit.

7.4.4 Spatial correlations of energetic disorder

It is interesting to note that there is a direct relationship between the energy and occupation probability of a hopping site. Indeed, in the canonical ensemble occupation probabilities are Boltzmann-distributed:

$$p_i \propto \exp\left[-E_i/k_BT\right] \tag{7.6}$$

Thus, the sites with high occupation probabilities are those with the lowest energies. We have already mentioned that the sites with the highest occupation probabilities are not randomly distributed in space (Sec. 7.4.3) but form pathways. Keeping in mind that these sites have the lowest energies, one may conclude that site energies are correlated in space. In fact, this result might be anticipated, since the Alq$_3$ molecule has a large dipole moment (see Sec. 7.2), which is known to lead to the spatial correlations in the site energies and to often affect the field dependence of the mobility (see Sec. 2.6).

Site energy correlations can be characterized by the following function:

$$C(r) = \frac{\langle (E_i - \langle E\rangle)(E_j - \langle E\rangle)\rangle}{\langle (E_i - \langle E\rangle)^2\rangle}, \tag{7.7}$$

which shows the "degree" of energetic correlation of two sites, separated by the distance r. For the system of randomly oriented point dipoles, the

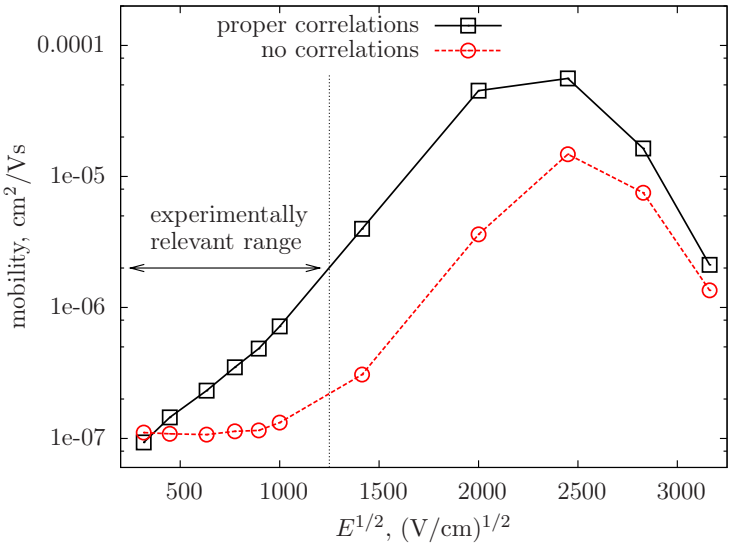

Figure 7.12: Poole-Frenkel plot for a system with 4096 molecules: (a) site energies are spatially correlated; (b) site energy correlation is removed by a random shuffling of the site energies.

correlation function decays as $1/r$ at large distances (see Sec. 2.6). Since Alq_3 molecules have a finite size (linear dimensions around 1 nm), the correlation function should be at least corrected for this:[4]

$$C(r) = \frac{c_0}{r - r_0}, \tag{7.8}$$

where r_0 represents the characteristic length-scale, which corresponds to the molecule size. The calculated correlation function for amorphous Alq_3 is presented in Fig. 7.11 together with the fitting to the functional form of Eq. 7.8. Fitting yields satisfactory results in the range of $1.0 - 2.5$ nm. Apparent deviation in the range from 0 to 1.0 nm can be explained by the fact that there are no molecule pairs in this region due to the excluded volume interactions. Parameters extracted by fitting are: $c_0 = 0.15 \pm 0.01$ and $r_0 = 0.67 \pm 0.07$ nm. The value of r_0 gives a reasonable estimation of the linear size of an Alq_3 molecule.[5]

[4] Of course, Alq_3 molecules are not point dipoles, so higher multipole moments also exist, which poses a question, whether the $1/r$ form is applicable at all.

[5] The radius of gyration of a single Alq_3 molecule at 300 K is $R_g = 0.378 \pm 0.001$ nm. Another estimation of the molecular linear size can be obtained by dividing the volume of the equilibrated simulation box by the number of molecules, thus obtaining the effective volume of one molecule, which can be used to obtain molecular radius, assuming that the molecule is spherical. This "effective" radius is $R_{eff} = 0.53 \pm 0.01$ nm.

To study how site energy spatial correlations affect the Poole-Frenkel behavior of the charge mobility (see Sec. 2.5.2) we performed simulations, in which site energies are not correlated. Correlations were removed by random shuffling of the site energies, which does not change the site energy distribution. Results are shown in Fig. 7.12. One can see that in the range of fields which is typically used in time-of-flight experiments ($300 - 1300\,(\mathrm{V/cm})^{1/2}$) [168, 26, 179] only the model with proper correlations reproduces the correct Poole-Frenkel behavior. If site energies are uncorrelated, the mobility has a plateau at low fields. At higher fields, however, both models give PF behavior $\log \mu \propto \beta E^{1/2}$ with the same slope β. This is consistent with the more general study of the effects of spatially correlated disorder on charge carrier mobility (Sec. 2.6). At even higher fields, charge carrier mobility reaches its maximum and then decreases. This is known as an inverted region [40]. Our results also explain the fact that in the previous study field-independent mobilities were observed [169]. In this study the dipolar energetic disorder was ignored and therefore site energies were spatially uncorrelated.

To conclude, Poole-Frenkel behavior of the charge carrier mobility in amorphous Alq_3 is governed by the site energy spatial correlations in the range of fields which is relevant for the time-of-flight experiments.

Chapter 8

Extracting nondispersive mobilities

Bottom-up multiscale approach discussed in Sec. 2.7 allows, in principle, to study transport without fitting parameters. However, its bottleneck is the computationally demanding evaluation of hopping rates for each pair of neighboring molecules, in particular intermolecular transfer integrals [52].[1] If density functional theory is employed, systems of up to several thousand molecules can be treated. As discussed in Chapter 2, semi-empirical methods, such as ZINDO, are more efficient since they require only precomputed monomer orbitals [53]. These methods are, however, not applicable for a large class of, e. g. metal-coordinated, compounds. As a result, simulated systems are normally only several nanometers thick, and charge transport is dispersive at room temperature for many amorphous materials, where energetic disorder is significant [28, 62]. Hence, simulated mobility will be system-size dependent.

In simulations, the box is often duplicated in the direction of the field before charge dynamics is studied. This seemingly straightforward increase of the system size will indeed result in non-dispersive transients, but is still incorrect since all periodically duplicated boxes have exactly the same (small) number of site energies, defining the equilibrium energy of a charge. Hence, charge carriers would traverse the sample at a different (higher) temperature than in an infinitely large system. On the other hand, in time-of-flight (TOF) experiments, a typical sample thickness is in the μm range and transport is often non-dispersive. To link simulation and experiment, one needs to extract the non-dispersive mobility from the simulations of small systems,

[1] The number of molecule pairs for which transfer integrals have to be evaluated can be of the order of $10^5(-10^6)$ for a single trajectory snapshot, making these calculations very time-consuming, even if parallelization is used.

where charge transport is dispersive.

8.1 Mean carrier energy in a finite system

To address this problem, we first consider the Gaussian density of states (GDOS), where the site energies ϵ_i are distributed according to $p(\epsilon) = 1/\sqrt{2\pi\sigma^2}\exp(-\epsilon^2/2\sigma^2)$. For a finite number of hopping sites, the carrier mean energy in canonical ensemble reads:

$$E_N = \left\langle \frac{1}{Z_N} \sum_{n=1}^{N} \epsilon_n e^{-\beta\epsilon_n} \right\rangle, \tag{8.1}$$

where $Z_N = \sum_{n=1}^{N} e^{-\beta\epsilon_n}$, $\langle \cdots \rangle$ denotes the average over all realizations of N site energies, $\beta = 1/k_\mathrm{B}T$. In a system with an infinite number of sites, $N = \infty$, the mean energy is proportional to inverse temperature [183, 184]:

$$E_\infty = \frac{1}{Z} \int_{-\infty}^{\infty} \epsilon p(\epsilon) e^{-\beta\epsilon}\, d\epsilon = -\sigma^2/k_\mathrm{B}T, \tag{8.2}$$

where $Z = \int_{-\infty}^{\infty} p(\epsilon) e^{-\beta\epsilon}\, d\epsilon$.

The dependencies of E_N/σ versus the number of sites, N, and the inverse temperature, $\sigma/k_\mathrm{B}T$, are shown in Fig. 8.1. One can see that the carrier mean energies in finite size systems are systematically higher than E_∞. This implies that mobility simulations in a small box would be effectively performed at a higher temperature. Hence, one might expect that the mobility will decrease with the increase of system size.[2] A similar trend is observed for the temperature dependence of E_N/σ, where mean energies are higher than E_∞ for small temperatures and large values of energetic disorder. Analysis of these dependencies yields an empirical expression for the transition between the dispersive and non-dispersive transport for large N, similar to [62]:

$$(\sigma/k_\mathrm{B}T)^2 = -5.7 + 1.05 \log N \tag{8.3}$$

For $\sigma/k_\mathrm{B}T = 2.7$, the asymptotic value E_∞ is achieved for $N > 10^5$. For $\sigma/k_\mathrm{B}T = 5.4$, E_N is well above E_∞, even for $N > 10^5$, i.e. significantly

[2] In experiments, temporal relaxation is normally discussed [183, 184], where mean energy is a function of time and $E(t \to \infty) = E_\infty$. While this approach is suitable for the interpretation of experimental data and time-of-flight type simulations, where charges are injected on one and collected on the other side of the sample, in simulations with periodic boundary conditions it is more appropriate to consider mean carrier energy as a function of the total number of hopping sites N.

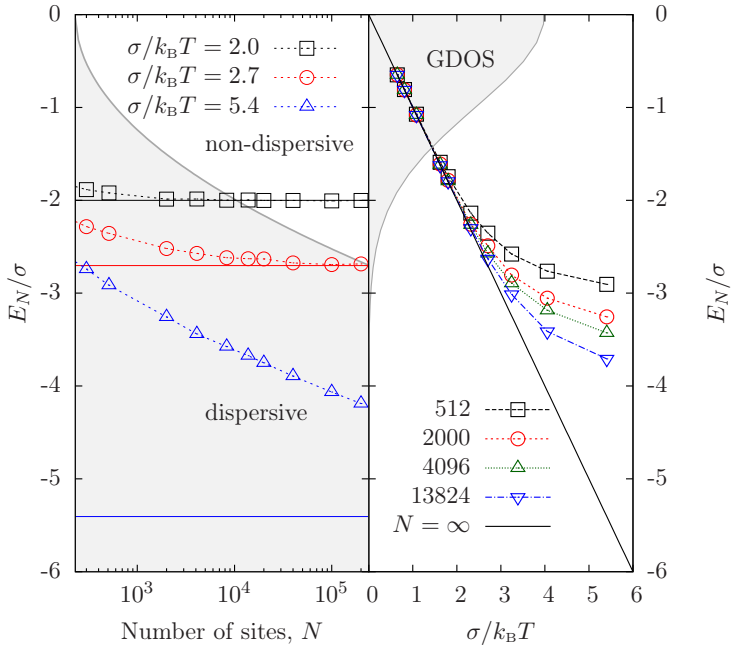

Figure 8.1: E_N (symbols) and E_∞ (solid lines) as a function of number of sites N (left) and inverse temperature (right). E_N is calculated by choosing N random numbers from the Gaussian distribution and evaluating the sum in Eq. (8.1). $\sigma/k_BT = 5.4$ corresponds to 300 K for Alq_3, which has an energetic disorder of $\sigma = 0.14$ eV. The grey area on the left plot (small N, large σ) defines the parameter space of the dispersive transport. GDOS is shown on the right plot.

bigger systems are required to equilibrate a charge carrier in the GDOS with $\sigma = 0.14$ eV at 300 K, which is a typical value of the energetic disorder of Alq_3. Note that much smaller systems are often used, $N = 294$ [68], $N = 1137$ [169], $N = 320$ [185] but the results are still compared directly to experiments.

There is a very simple mathematical argument, behind the fact that E_N is always larger than E_∞. Site energies ϵ_n contribute to the average E_N with the weight $e^{-\beta\epsilon_n}$. Roughly speaking, low-lying energies ϵ_n give the biggest contribution to the average E_N, whereas the highest energies can be neglected. However, for small N there is only a very small probability, that those low-lying ϵ_n are encountered. When increasing N, more and more low-lying energies are sampled, thus decreasing the E_N.

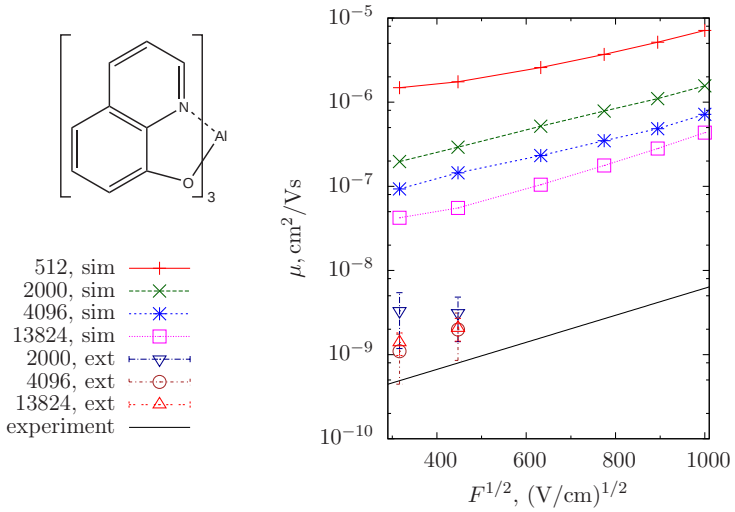

Figure 8.2: Chemical structure of Alq_3, simulated (sim) and extrapolated (ext) Alq_3 hole mobilities as a function of electric field for a set of box sizes (number of hopping sites, N). Strong variation of simulated mobilities with system size is typical for dispersive transport. Simulated values collapse on a single curve (in a zero field limit) after extrapolation. Experimental mobilities are taken from Ref. [26]. Poole-Frenkel behavior of the mobility [26, 16, 168], $\mu = \mu_0 \exp(\gamma\sqrt{F})$, is also reproduced.

8.2 Temperature extrapolation

Eq. (8.3) predicts that a brute-force increase of the number of sites N cannot resolve the problem for compounds with large energetic disorder, since N increases exponentially with σ^2. Eq. (8.3) also hints at a possible solution. Indeed, the relevant dimensionless parameter is the half width of GDOS divided by temperature, $\sigma/k_{\mathrm{B}}T$. Hence, the transition between dispersive and non-dispersive transport can be shifted to lower values of N by simply increasing the temperature. At some elevated temperature, transport will eventually become non-dispersive even for small box sizes. Provided the temperature dependence of the non-dispersive mobility is known, its value can then be extrapolated to experimentally relevant temperatures.

Two prerequisites are needed to perform such an extrapolation. First, the simulation temperatures should be high enough for transport to be non-dispersive for a given size of the simulation box. To do this, the transition temperature T_{ND}, can be estimated via Eq. (8.3) from the value of energetic disorder σ (which can be obtained from the site energy distribution even in small systems) and the number of hopping sites, N. Non-dispersive mobilities can then be calculated for a set of temperatures above T_{ND}. Second, an explicit temperature dependence of charge carrier mobility is needed. In one

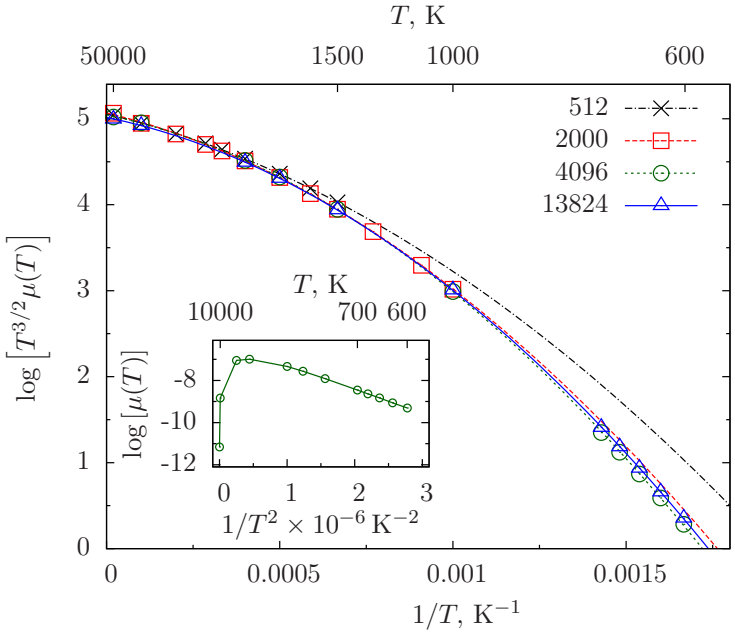

Figure 8.3: Temperature dependence of non-dispersive hole mobility. $F = 10^5\,\mathrm{V/cm}$. Points correspond to the simulation results, lines to fittings to Eq. (8.4). Inset shows that widely used GDM temperature dependence [60, 62], $\mu(T) = \mu_0 \exp[-(T_0/T)^2]$, can only be used in a limited temperature range, failing at high temperatures [64].

dimension, the exact analytical expression for mobility is known for an arbitrary set of rates [186, 64, 187]. In the case of Marcus rates, the temperature dependence of the non-dispersive mobility at zero field reads [187]:

$$\mu(T) = \frac{\mu_0}{T^{3/2}} \exp\left[-\left(\frac{a}{T}\right)^2 - \left(\frac{b}{T}\right)\right], \tag{8.4}$$

where a, b, and μ_0 are material constants. Strictly speaking, this temperature dependence is valid in one dimension only. As we will see, it can still be used in a three-dimensional case in a very broad temperature range. A fit to Eq. (8.4) allows for the parameters μ_0, a, and b to be extracted. Finally, the same expression is used to obtain the mobility at a desired temperature. Since the actual temperature dependence of mobility is not of interest, the morphology equilibrated at room temperature is kept fixed. To study temperature dependence, e.g. an effect of a glass transition [62], the procedure should be repeated for several, equilibrated at different temperatures, morphologies.

To illustrate and test the extraction of non-dispersive mobilities, we consider hole transport in amorphous films of tris(8-hydroxyquinoline) aluminum

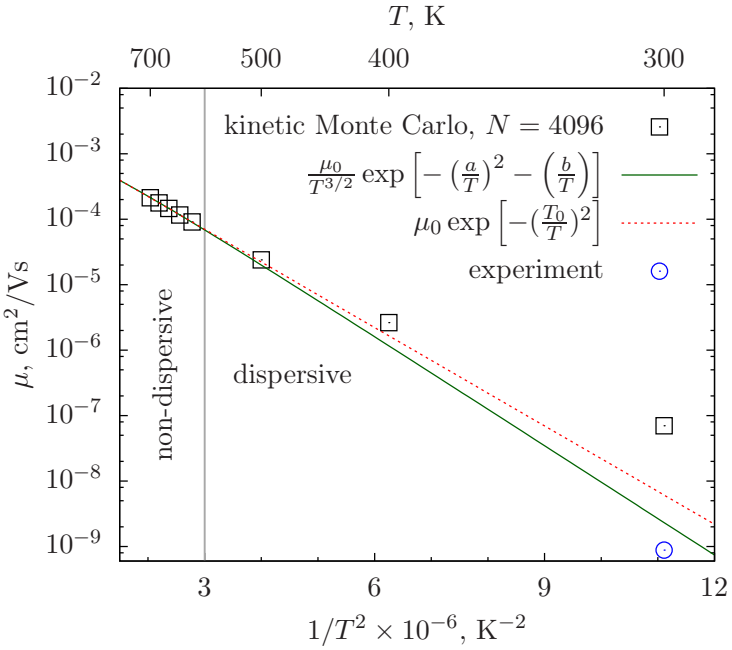

Figure 8.4: Alq_3 hole mobility as a function of temperature. $F = 10^5\,\mathrm{V/cm}$. In the dispersive regime, mobility deviates from Eq. (8.4) as well as from a GDM-based low field limit temperature dependence of charge carrier mobility. As long as high temperatures are not required (for fitting), the difference between the GDM-based and Eq. (8.4) functional forms is small.

(Alq_3). Hole mobilities in amorphous Alq_3 are known to be non-dispersive for a μm range thickness [26, 168, 16].

The Poole-Frenkel plot, namely, room temperature mobility as a function of the square-root of an applied electric field, is shown in Fig. 8.2 for simulation boxes of 512, 2000, 4096, and 13824 molecules. Details of the simulations are described in Sec. 7.4.1. The mobility is clearly system size dependent and is several orders of magnitude larger than the experimentally reported value. After estimating T_{ND} for each N (for example, $T_{\mathrm{ND}}(N = 4096) \approx 600\,\mathrm{K}$), the mobility was calculated for a broad temperature range from 700 K to 50000 K for $T > T_{\mathrm{ND}}(N)$. Simulation results together with the fit to Eq. (8.4) are shown in Fig. 8.3. The agreement is excellent and validates the ansatz used for the temperature dependence. A small deviation exists only for $N = 512$, which is due to the limited accuracy of the fit since there are only a few points available above T_{ND}. One should note that this is an additional limitation in the case of too small system sizes.

The results of the extrapolation are illustrated in Fig. 8.4, where both simulated and extrapolated mobilities are shown. At low temperatures, when

transport is dispersive, mobility is systematically higher than the value prescribed by Eq. (8.4). In this region, carrier mobility is system-size dependent. Above T_{ND}, mobility does not depend on the system size and its temperature dependence agrees well with Eq. (8.4).

Finally, we calculated the non-dispersive mobility at 290 K for two low values of the electric field, $F = 10^5\,\mathrm{V/cm}$ and $2 \times 10^5\,\mathrm{V/cm}$. The results are shown in Fig. 8.2 together with the TOF results obtained at 290 K [26]. One can see that both absolute values as well as the Poole-Frenkel behavior are very well reproduced. Given systematic errors of transfer integrals and site energies and the fact that TOF experiments provide slightly different mobilities depending on the coating rate for the amorphous film preparation [26, 16, 168], the agreement between simulated and experimental values is excellent.

To summarize, an approach was proposed which can be used to obtain non-dispersive charge carrier mobilities from simulations in small systems, which provides a way to bridge length-scales with different carrier dynamics.

Conclusion and Outlook

In this thesis a multiscale approach which combines molecular dynamics simulations for generating material morphologies, electronic-structure calculations for determining charge hopping rates and kinetic Monte Carlo for simulating charge dynamics was applied to study charge transport in amorphous films of tris(8-hydroxyquinoline) aluminium (Alq_3). First, an atomistic force field was parameterized by matching the first-principles potential energy scans with the corresponding force field based scans. The resulting force field was then validated against available experimental data, reproducing the density and glass transition temperature. It was also shown that the morphological changes caused by the slight variation of the nonbonded force field parameters do not significantly affect charge dynamics.

Amorphous morphologies of Alq_3 were then used to simulate the charge carrier dynamics. It was demonstrated that the large molecular dipole moment is responsible for the energetic disorder, which significantly affects the charge transport. Spatial correlations of the site energies are shown to govern the Poole-Frenkel behavior of the charge mobility. It was also found that the hole transport is dispersive for system sizes accessible to simulations, implying that simulated mobilities are system size dependent. A method for extrapolating simulated mobilities to the infinite system size was proposed, which allows to compare calculated mobilities directly to the time-of-flight experiments. The extracted value of the nondispersive hole mobility and its electric field dependence were found to be in agreement with the experimental results.

As a natural extension of this work, the techniques must be developed to calculate different contributions to the energetic disorder in organic materials, since those can significantly affect the carrier dynamics. For example, the explicit treatment of the atomic polarization might lead to a significant source of disorder, which was only phenomenologically taken into account in this work. Another possible contribution is related to the fact, that every molecule in the amorphous morphology has a different conformation, which is different from its ground state. Thus, apart from the corresponding ener-

getic difference, reorganization energies must be explicitly calculated for each molecule instead of using the value obtained for the ground state geometry.

Taking these additional sources of disorder into account will provide a deeper microscopic understanding of the charge transport in amorphous materials. Specifically, it might help to understand why electron mobilities in amorophous Alq_3 are two orders of magnitude higher than the hole mobilities and why electron transport is reported to be dispersive in most time-of-flight studies, while showing higher mobilities than the nondispersive hole transport. In the long run microscopic understanding of the structure-property relationships in organic semiconductors will help to design more efficient devices with improved stability.

For amorphous Alq_3 large scale morphologies can be generated solely using atomistic molecular dynamics. However, atomistic simulations can only cover nanometer length- and nanosecond time-scales, which might be not enough to equilibrate a system with long relaxation times (e. g. polymeric systems). In order to overcome this limitation, systematic coarse-graining methods can be used. Iterative Boltzmann inversion (IBI), inverse Monte Carlo (IMC) and force matching (FM) were applied to parameterize coarse-grained models for the SPC/E water, methanol, propane and the single molecule of hexane. It was demonstrated that, while IMC has a more efficient update than IBI, it sometimes suffers from finite-size effects. It was shown that force matching is a very efficient method, but in many cases it fails to reproduce the reference distributions, because the basis set used to represent the coarse-grained force field is not complete. However, if 3-body interaction terms are included in the coarse-grained force field, FM would reproduce both 2-body and 3-body correlation functions, which is not possible to achieve with IBI or IMC. As an extension of this work, the development of the "hybrid" coarse-graining techniques is needed. One can think of a scheme which combines the strength of both IBI and IMC. This scheme must have an efficient update, comparable to that of IMC, combined with the stability of IBI. Another possibility is a combination of IBI for bonded interactions (since in this case the iterative procedure is rarely required) and FM for the nonbonded interactions (since FM reproduces nonbonded reference distributions in many cases). Computational cost of this "hybrid" scheme will be negligible compared to the pure IBI, since iterative CG simulations are not required.

Appendix A

Force matching using cubic splines

Implementations of force matching using different basis functions (linear splines, cubic splines, step functions) and different methods for solving the least-squares problem (QR decomposition, singular value decomposition, iterative techniques, normal matrix approach) are discussed in detail in Ref. [146].

Here the outline is given for the implementation which uses cubic splines as basis functions, QR-decomposition for solving the least-squares problem, and block averaging to sample large trajectories.

In this case the force $\vec{f}_{\gamma_i}(\{r_k\})$ acting on bead i due to an interaction γ_i with the potential U_{γ_i} can be written as

$$\begin{aligned} \vec{f}_{\gamma_i}(\{r_k\}) &= -\vec{\nabla}_i U\left(\kappa\left(\{\vec{r}_k\}\right)\right) \\ &= -\frac{\partial U}{\partial \kappa}\vec{\nabla}_i \kappa\left(\{\vec{r}_k\}\right) \\ &= -f_{\gamma_i}\vec{\nabla}_i \kappa\left(\{\vec{r}_k\}\right) \ , \end{aligned} \tag{A.1}$$

where $\kappa = r, b, \theta, \varphi$ denotes the type of interaction, and $\vec{\nabla}_i$ the gradient with respect to the coordinates $\vec{r}_i$ of bead i. The variable κ can label non-bonded interactions, bonds, angles or dihedral angles, which are given by the distance between two beads, the bond length, the angle which depends on 3 beads or the dihedral angle defined using 4 beads, respectively. Now, the total force $\vec{f}_i^{\,\mathrm{cg}}$ acting on coarse-grained bead i can be expressed in terms of the coarse-grained interactions and equation 4.16 can be rewritten as

$$\sum_{\gamma_i} f_{\gamma_i}(\kappa)\, \vec{\nabla}_i \kappa\left(\{\vec{r}_{kl}\}\right) = \vec{f}_{il}^{\,\mathrm{ref}} \ , \tag{A.2}$$

where γ_i enumerates all interactions acting on bead i.

$f(\kappa)$ is interpolated using cubic splines connecting a set of points $\{\kappa_k\}$

$$\begin{aligned} S_n(\kappa, \{\kappa_k\}, \{f_k\}, \{f_k^{''}\}) &= A_n(\kappa) f_n \\ &+ B_n(\kappa) f_{n+1} \\ &+ C_n(\kappa) f_n^{''} \\ &+ D_n(\kappa) f_{n+1}^{''} \ . \end{aligned} \tag{A.3}$$

where $\{f_k\}, \{f_k^{''}\}$ are tabulations of $f(\kappa)$ and its second derivative on the grid $\{\kappa_k\}$, the parameters $\{f_k\}, \{f_k^{''}\}$ are obtained from the fit, $\kappa \in [\kappa_n, \kappa_{n+1}]$, and the coefficients A_n, B_n, C_n, and D_n have the following form

$$\begin{aligned} A_n(\kappa) &= 1 - \frac{\kappa - \kappa_n}{h_{n+1}} \ , \\ B_n(\kappa) &= \frac{\kappa - \kappa_n}{h_{n+1}} \ , \\ C_n(\kappa) &= \frac{1}{2}(\kappa - \kappa_n)^2 - \frac{1}{6}\frac{(\kappa - \kappa_n)^3}{h_{n+1}} - \frac{1}{3}h_{n+1}(\kappa - \kappa_n) \ , \\ D_n(\kappa) &= \frac{1}{6}\frac{(\kappa - \kappa_n)^3}{h_{n+1}} - \frac{1}{6}h_{n+1}(\kappa - \kappa_n) \ , \end{aligned} \tag{A.4}$$

where $h_n = \kappa_{n+1} - \kappa_n$.

An additional requirement on the spline coefficients is the continuity of the first derivatives

$$\begin{aligned} A_n(\kappa_{n+1})' f_n + B_n(\kappa_{n+1})' f_{n+1} + \\ C_n(\kappa_{n+1})' f_n'' + D_n(\kappa_{n+1})' f_{n+1}'' = \\ A_{n+1}(\kappa_{n+1})' f_{n+1} + B_{n+1}(\kappa_{n+1})' f_{n+2} + \\ C_{n+1}(\kappa_{n+1})' f_{n+1}'' + D_{n+1}(\kappa_{n+1})' f_{n+2}'' \ . \end{aligned} \tag{A.5}$$

If the total number of grid points is $N+1$ ($n = 0, 1, \ldots, N$) then these conditions are specified for the points $n = 0, 1, \ldots, N-1$. For non-periodic potentials the end points are treated using normal boundary conditions, i. e. $f_0'' = 0$, $f_N'' = 0$.

Due to the spline interpolation, equation A.2 simplifies to a set of linear equations with respect to the fitting parameters $f_n, f_n^{''}$. The complete set of equations to solve therefore consists of equations A.2 and constraints A.5. Strictly speaking, this set of equations cannot be solved in a least-squares sense using simple QR decomposition. The reason is that the constraints shall be satisfied *exactly* to ensure the continuity of the first derivative of the potential, which is not the case if they are treated in a least-squares sense. To solve the problem, one, in principle, has to use a constrained least-squares

solver [188]. From a practical point of view, however, it is simpler to treat the constraints in a least-squares sense for each block. This will only lead to a piecewise-smooth potential, but the smoothness can be "recovered" by averaging over the blocks.

APPENDIX A. FORCE MATCHING USING CUBIC SPLINES

Appendix B

Yvon-Born-Green equations for a homogeneous liquid

Consider a system, consisting of N identical beads. The hamiltonian is defined as:

$$H(\boldsymbol{R}_N, \boldsymbol{P}_N) = \sum_i^N \frac{1}{2M} \boldsymbol{P}_i^2 + V^N(\boldsymbol{R}_N) \tag{B.1}$$

Assuming that the potential function V^N is pairwise decomposable, the total force on each bead $\boldsymbol{F}_i = -\partial V^N / \partial \boldsymbol{R}_i$ arises from a sum of pair interactions:

$$\boldsymbol{F}_i = \sum_j \boldsymbol{F}_{ij}(\boldsymbol{R}_i, \boldsymbol{R}_j) \tag{B.2}$$

For a homogeneous system with $\rho(\boldsymbol{R}_i) = \rho = \text{const}$, the following Yvon-Born-Green (YBG) equation holds, which relates the equilibrium two- and three-particle distribution functions to the pairwise decomposable force field $\boldsymbol{F}_{ij}$ [138]:

$$\left(k_B T \frac{\partial}{\partial \boldsymbol{R}_i} - \boldsymbol{F}_{ij}(\boldsymbol{R}_i, \boldsymbol{R}_j) \right) g^{(2)}(\boldsymbol{R}_i, \boldsymbol{R}_j) = \rho \int d\boldsymbol{R}_k \, \boldsymbol{F}_{ik} \, g^{(3)}(\boldsymbol{R}_i, \boldsymbol{R}_j, \boldsymbol{R}_k) \tag{B.3}$$

APPENDIX B. YVON-BORN-GREEN EQUATIONS FOR A HOMOGENEOUS LIQUID

Appendix C

Validation of the Alq_3 force field

To assess the quality of the developed Alq_3 force field we compare the distribution of the selected degrees of freedom obtained in different simulations: (1) Car-Parrinello molecular dynamics simulations [189] with the dispersion-corrected BLYP functional [190]; (2) Classical MD for a single molecule; (3) Classical MD of the amorphous morphology consisting of 512 molecules. Force field parameterization described in Chapter 7 was done on a single molecule using B3LYP scans. That implies that nonbonded dispersion forces were not properly taken into account. To check whether this introduces a serious mistake we performed CPMD simulations with the dispersion-corrected functional. To quantify how the molecular packing in the amorphous morphology affects the molecule shape we also added MD simulations of the amorphous morphology for comparison. Resulting distributions are shown in Fig. C.1.

The results of both classical simulations are practically identical, showing that packing does not play an important role. On the other hand, CPMD results[1] show deviations for some of the degrees of freedom (angle Ca1-Oa-Al, dihedral Oa-Al-Ob-Cb1). These deviations might be due to the use of different functionals (BLYP for the CPMD run, B3LYP for the force field parameterization) or due to the fact that dispersion forces were not taken into account properly during the parameterization. Nevertheless, the quality of the resulting force field is sufficient for the type of calculations performed in this thesis.

[1]Since CPMD simulations are limited to the picosecond time scales even for a single molecule, only very poor statistics was obtained.

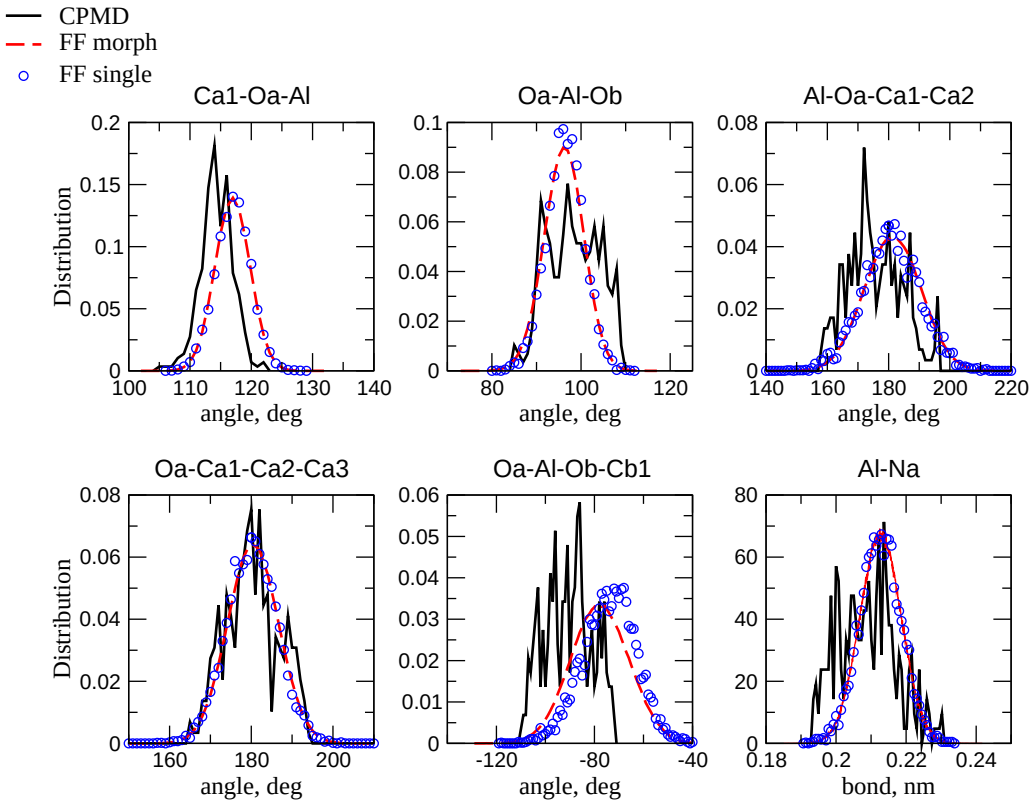

Figure C.1: Distributions of the several representative degrees of freedom of an Alq$_3$ molecule obtained using Car-Parrinello MD (dispersion-corrected BLYP) (full lines), classical MD with the developed force field for an amorphous morphology (dashed lines) and for a single molecule in vacuum (dots).

Bibliography

[1] **A. Lukyanov**, C. Lennartz, and D. Andrienko. Amorphous films of tris(8-hydroxyquinolinato)aluminium: Force-field, morphology, and charge transport. *Phys. Stat. Sol.(a)*, 206:2737–2742, 2009.

[2] V. Rühle, C. Junghans, **A. Lukyanov**, K. Kremer, and D. Andrienko. Versatile object-oriented toolkit for coarse-graining applications. *J. Chem. Theory Comput.*, 5(12):3211–3223, 2009.

[3] **A. Lukyanov**, A. Malafeev, V. Ivanov, H.-L. Chen, K. Kremer, and D. Andrienko. Solvated poly-(phenylene vinylene) derivatives: conformational structure and aggregation behavior. *J. Mater. Chem.*, 20:10475–10485, 2010.

[4] **A. Lukyanov** and D. Andrienko. Extracting nondispersive charge carrier mobilities of organic semiconductors from simulations of small systems. *Phys. Rev. B.*, 82(19):193202, 2010.

[5] J. Bardeen and W. H. Brattain. The transistor, a semi-conductor triode. *Phys. Rev.*, 74(2):230–231, 1948.

[6] W. Brütting. *Physics of Organic Semiconductors*. Wiley-VCH, Berlin, 2005.

[7] J. Kalinowski. *Organic Light-Emitting Diodes: Principles, Characteristics & Processes*. Marcel Dekker Inc, collectors ed/. edition, 2004.

[8] C. Brabec, U. Scherf, and V. Dyakonov. *Organic Photovoltaics: Materials, Device Physics, and Manufacturing Technologies*. Wiley-VCH Verlag GmbH & Co. KGaA, 1. auflage edition, 2008.

[9] I. Kymissis. *Organic Field Effect Transistors: Theory, Fabrication and Characterization*. Springer-Verlag Gmbh, 1 edition, 2009.

[10] C. Tang and S. VanSylke. Organic electroluminescent diodes. *Appl. Phys. Lett.*, 51:913, 1987.

[11] C. Tang and S. VanSlyke. Electroluminescence of doped organic thin films. *J. Appl. Phys.*, 65:3610, 1989.

[12] J. Shi and C. Tang. Doped organic electroluminescent devices with improved stability. *Appl. Phys. Lett.*, 70:1665, 1997.

[13] L. S. Hung, C. W. Tang, and M. G. Mason. Enhanced electron injection in organic electroluminescence devices using an al/lif electrode. *Appl. Phys. Lett.*, 70:152–154, 1997.

[14] H. Aziz, Z. D. Popovic, N.-X. Hu, A.-M. Hor, and G. Xu. Degradation Mechanism of Small Molecule-Based Organic Light-Emitting Devices. *Science*, 283:1900–1902, 1999.

[15] M. Stössel, J. Staudigel, F. Steuber, J. Blassing, J. Simmerer, and A. Winnacker. Space-charge-limited electron currents in 8-hydroxyquinoline aluminum. *Appl. Phys. Lett.*, 76:115–117, 2000.

[16] W. Brütting, S. Berleb, and A. G. Mückl. Device physics of organic light-emitting diodes based on molecular materials. *Organic Electronics*, 2:1–36, 2001.

[17] S. Berleb and W. Brütting. Dispersive electron transport in tris(8-hydroxyquinoline)aluminum(alq3) probed by impedance spectroscopy. *Phys. Rev. Lett.*, 89:286601, 2002.

[18] T. P. Saragi, T. Fuhrmann-Lieker, and J. Salbeck. High on/off ratio and stability of amorphous organic field-effect transistors based on spiro-linked compounds. *Synthetic Metals*, 148(3):267 – 270, 2005.

[19] S. Günes, H. Neugebauer, and N. S. Sariciftci. Conjugated polymer-based organic solar cells. *Chem. Rev.*, 107(4):1324–1338, 2007.

[20] J. Peet, J. Kim, N. Coates, W. Ma, D. Moses, A. Heeger, and G. Bazan. Efficiency enhancement in low-bandgap polymer solar cells by processing with alkane dithiols. *Nature Materials*, 6:497, 2007.

[21] M. Schraber, M. Koppe, J. Gao, F. Cordella, M. Loi, P. Denk, M. Morana, H.-J. Egelhaaf, K. Forberich, G. Dennler, R. Gaudiana, D. Waller, Z. Zhu, X. Shi, and C. Brabec. Influence of the bridging atom on the performance of a low-bandgap bulk heterojunction solar cell. *Adv. Mater.*, 22:367–370, 2010.

[22] X. Yang and J. Loos. Toward High-Performance polymer solar cells: The importance of morphology control. *Macromolecules*, 40(5):1353–1362, 2007.

[23] C. Winder and N. Sariciftci. Low bandgap polymers for photon harvesting in bulk heterojunction solar cells. *J. Mater. Chem.*, 14:1077–1086, 2004.

[24] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta. Electron drift velocity in silicon. *Phys. Rev. B*, 12(6):2265–2284, 1975.

[25] G. Ottaviani, L. Reggiani, C. Canali, F. Nava, and A. Alberigi-Quaranta. Hole drift velocity in silicon. *Phys. Rev. B*, 12(8):3318–3329, 1975.

[26] H. Fong and S. So. Hole transporting properties of tris(8-hydroxyquinoline) aluminum (alq3). *J. App. Phys.*, 100:094502, 2006.

[27] S. Tiwari and N. Greenham. Charge mobility measurement techniques in organic semiconductors. *Optical and Quantum Electronics*, 41:69–89, 2009.

[28] H. Scher and E. W. Montroll. Anomalous transit-time dispersion in amorphous solids. *Phys. Rev. B*, 12:2455, 1975.

[29] F. Laquai, G. Wegner, C. Im, H. Bässler, and S. Heun. Comparative study of hole transport in polyspirobifluorene polymers measured by the charge-generation layer time-of-flight technique. *J. Appl. Phys.*, 99:023712, 2006.

[30] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Bredas. Charge transport in organic semiconductors. *Chem. Rev.*, 107(4):926–952, 2007.

[31] P. W. M. Blom, M. J. M. de Jong, and J. J. M. Vleggaar. Electron and hole transport in poly(p-phenylene vinylene) devices. *Appl. Phys. Lett.*, 68:3308–3310, 1996.

[32] P. P. Infelta, M. P. de Haas, and J. M. Warman. The study of the transient conductivity of pulse irradiated dielectric liquids on a nanosecond timescale using microwaves. *Radiation Physics and Chemistry*, 10(5-6):353 – 365, 1977.

[33] http://en.wikipedia.org/wiki/Organic_light-emitting_diode.

[34] W. F. Pasveer, J. Cottaar, C. Tanase, R. Coehoorn, P. A. Bobbert, P. W. M. Blom, D. M. de Leeuw, and M. A. J. Michels. Unified description of charge-carrier mobilities in disordered semiconducting polymers. *Phys. Rev. Lett.*, 94(20):206601, 2005.

[35] N. Karl. Charge carrier transport in organic semiconductors. *Synthetic Metals*, 133-134:649 – 657, 2003. Proceedings of the Yamada Conference LVI. The Fourth International.

[36] N. Karl and J. Marktanner. Electron and hole mobilities in high purity anthracene single crystals. *Mol. Cryst. Liq. Cryst.*, 355:149, 2001.

[37] A. Troisi and G. Orlandi. Charge-transport regime of crystalline organic semiconductors: Diffusion limited by thermal off-diagonal electronic disorder. *Phys. Rev. Lett.*, 96(8):086601, 2006.

[38] A. Troisi, D. L. Cheung, and D. Andrienko. Charge transport in semiconductors with multiscale conformational dynamics. *Phys. Rev. Lett.*, 102:116602, 2009.

[39] J. C. Slater and G. F. Koster. Simplified lcao method for the periodic potential problem. *Phys. Rev.*, 94(6):1498–1524, 1954.

[40] V. May and O. Kühn. *Charge and Energy Transfer Dynamics in Molecular Systems. Second, revised and enlarged edition.* Wiley-Vch, 2004.

[41] T. Kakitani and N. Mataga. Comprehensive study on the role of coordinated solvent mode played in electron-transfer reactions in polar solutions. *J. Phys. Chem.*, 91(24):6277–6285, 1987.

[42] V. Balzani. *Electron Transfer in Chemistry.* Wiley-VCH, Weinheim, 2001.

[43] L. Landau and L. Lifshitz. *Quantum Mechanics: Non-relativistic theory.* Butterworth-Heinemann, 3 edition, 1981.

[44] K. F. Freed and J. Jortner. Multiphonon processes in the nonradiative decay of large molecules. *J. Chem. Phys.*, 52(12):6272–6291, 1970.

[45] R. A. Marcus. On the theory of oxidation-reduction reactions involving electron transfer. i. *J. Chem. Phys.*, 24(5):966–978, 1956.

[46] R. A. Marcus. Electron transfer reactions in chemistry: Theory and experiment. *Rev. Mod. Phys.*, 65(3):599–610, 1993.

[47] J.-L. Bredas, D. Beljonne, V. Coropceanu, and J. Cornil. Charge-transfer and energy-transfer processes in pi-conjugated oligomers and polymers: A molecular picture. *Chem. Rev.*, 104(11):4971–5004, 2004.

[48] M. D. Newton. Quantum chemical probes of electron-transfer kinetics: the nature of donor-acceptor interactions. *Chem. Rev.*, 91(5):767–792, 1991.

[49] E. F. Valeev, V. Coropceanu, D. A. da Silva Filho, S. Salman, and J. Bredas. Effect of electronic polarization on Charge-Transport parameters in molecular organic semiconductors. *J. Am. Chem. Soc.*, 128(30):9882–9886, 2006.

[50] S. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies: Some procedures with reduced errors. *Mol. Phys.*, 19, 1970.

[51] S. Simon, M. Duran, and J. J. Dannenberg. How does basis set superposition error change the potential surfaces for hydrogen-bonded dimers? *J. Chem. Phys.*, 105, 1996.

[52] B. Baumeier, J. Kirkpatrick, and D. Andrienko. Density-functional based determination of intermolecular charge transfer properties for large-scale morphologies. *Phys. Chem. Chem. Phys.*, 12:11103–11113, 2010.

[53] J. Kirkpatrick. An approximate method for calculating transfer integrals based on the zindo hamiltonian. *Int. J. Quant. Chem.*, 108:51, 2008.

[54] J. Ridley and M. Zerner. An intermediate neglect of differential overlap technique for spectroscopy: Pyrrole and the azines. *Theoretica Chimica Acta*, 32(2):111–134, 1973.

[55] D. H. Dunlap, P. E. Parris, and V. M. Kenkre. Charge-dipole model for the universal field dependence of mobilities in molecularly doped polymers. *Phys. Rev. Lett.*, 77(3):542–545, 1996.

[56] J. Kirkpatrick, V. Marcon, K. Kremer, J. Nelson, and D. Andrienko. Columnar mesophases of hexabenzocoronene derivatives: Ii. charge carrier mobility. *J. Chem. Phys.*, 129:094506, 2008.

[57] S. Verlaak and P. Heremans. Molecular microelectrostatic view on electronic states near pentacene grain boundaries. *Phys. Rev. B*, 75(11):115127, 2007.

[58] V. Rühle, J. Kirkpatrick, and D. Andrienko. A multiscale description of charge transport in conjugated oligomers. *J. Chem. Phys.*, 132(13):134103, 2010.

[59] J. Kirkpatrick, V. Marcon, J. Nelson, K. Kremer, and D. Andrienko. Charge mobility of discotic mesophases: A multiscale quantum and classical study. *Phys. Rev. Lett.*, 98(22):227402, 2007.

[60] P. M. Borsenberger, L. Pautmeier, and H. Bässler. Charge transport in disordered molecular solids. *J. Chem. Phys.*, 94:5447, 1991.

[61] A. Miller and E. Abrahams. Impurity conduction at low concentrations. *Phys. Rev.*, 120(3):745–755, 1960.

[62] P. Borsenberger, E. Magin, M. van der Auweraer, and F. de Schryver. The role of disorder on charge transport in molecularly doped polymers and related materials. *Phys. Stat. Sol.(a)*, 140:9, 1993.

[63] H. Bässler. Charge transport in disordered organic photoconductors - a monte carlo simulation study. *Phys. Stat. Sol. B*, 175(1):15–56, 1993.

[64] H. Cordes, S. D. Baranovskii, K. Kohary, P. Thomas, S. Yamasaki, F. Hensel, and J. Wendorff. One-dimensional hopping transport in disordered organic solids. i. analytic calculations. *Phys. Rev. B*, 63(9):094201, 2001.

[65] W. D. Gill. Drift mobilities in amorphous charge-transfer complexes of trinitrofluorenone and poly-n-vinylcarbazole. *J. Appl. Phys.*, 43(12):5033–5040, 1972.

[66] S. Novikov, D. Dunlap, V. Kenkre, P. Parris, and A. Vannikov. Essential role of correlations in governing charge transport in disordered organic materials. *Phys. Rev. Lett.*, 81:4472, 1998.

[67] A. Ohno and J. Hanna. Simulated carrier transport in smectic mesophase and its comparison with experimental result. *Appl. Phys. Lett.*, 82(5):751–753, 2003.

[68] Y. Nagata and C. Lennartz. Atomistic simulation on charge mobility of amorphous tris(8-hydroxyquinoline) aluminum (alq3): Origin of poole-frenkel-type behavior. *J. Chem. Phys.*, 129:034709, 2008.

[69] P. M. Borsenberger, L. T. Pautmeier, and H. Bässler. Nondispersive-to-dispersive charge-transport transition in disordered molecular solids. *Phys. Rev. B*, 46(19):12145–12153, 1992.

[70] Y. Gartstein and E. Conwell. High-field hopping mobility in molecular systems with spatially correlated energetic disorder. *Chem. Phys. Lett.*, 245:351–358, 1995.

[71] S. Novikov and A. Vannikov. Cluster structure in the distribution of the electrostatic potential in a lattice of randomly oriented dipoles. *J. Phys. Chem.*, 99:14573, 1995.

[72] V. Marcon, V. T., J. Kirkpatrick, C. Jeong, D. Y. Yoon, K. Kremer, and D. Andrienko. Columnar mesophases of hexabenzocoronene derivatives. 1. phase transitions. *J. Chem. Phys.*, 129:094505, 2008.

[73] V. Marcon, W. Pisula, J. Dahl, D. W. Breiby, J. Kirkpatrick, S. Patwardhan, F. Grozema, and D. Andrienko. Understanding structure-mobility relations for perylene tetracarboxdiimide derivatives. *J. Am. Chem. Soc.*, 131:11426–11432, 2009.

[74] X. Feng, V. Marcon, W. Pisula, M. Hansen, J. Kirkpatrick, F. Grozema, D. Andrienko, K. Kremer, and K. Müllen. Towards high charge-carrier mobilities by rational design of the shape and periphery of discotics. *Nature Materials*, 8(5):421–426, 2009.

[75] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2 edition, 2004.

[76] T. Vehoff, J. Kirkpatrick, K. Kremer, and D. Andrienko. Atomistic force field and electronic properties of carbazole: from monomer to macrocycle. *Phys. Status Solidi B*, 245:839, 2008.

[77] V. Rühle, J. Kirkpatrick, K. Kremer, and D. Andrienko. Coarse-grained modelling of polypyrrole morphologies. *Phys. Status Solidi B*, 245:844, 2008.

[78] A. Grosberg and A. Khokhlov. *Statistical Physics of Macromolecules*. American Institute of Physics, 2002.

[79] G. A. Voth, editor. *Coarse-Graining of Condensed Phase and Biomolecular Systems*. Taylor and Francis, 2008.

[80] W. Jorgensen and J. Tirado-Rives. The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110:1657, 1988.

[81] H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, and J. R. Haak. Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.*, 81:3684–3690, 1984.

[82] S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52:255–268, 1984.

[83] W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695, 1985.

[84] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126:014101, 2007.

[85] 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.

[86] M. Allen and D. Tildesley. *Computer simulations of liquids*. Oxford science publications, 1989.

[87] D. Frenkel and B. Smit. *Understanding molecular simulation: From algorithms to applications*. Academic Press, 2 edition, 2002.

[88] R. Faller, H. Schmitz, O. Biermann, and F. Müller-Plathe. Automatic parameterization of force fields for liquids by simplex optimization. *J. Comp. Chem.*, 20(10):1009–1017, 1999.

[89] M. Hülsmann, T. Köddermann, J. Vrabec, and D. Reith. Grow: a gradient-based optimization workflow for the automated development of molecular models. *Comp. Phys. Commun.*, 181:499–513, 2010.

[90] B. Hsieh, H. Antoniadis, D. Bland, and W. Feld. Chlorine precursor route (cpr) chemistry to poly(p-phenylene vinylene)-based light emitting diodes. *Adv. Mater.*, 7:36, 1995.

[91] B. Hsieh, Y. Yu, E. Forsythe, G. Schaaf, and W. Feld. A new family of highly emissive soluble poly(p-phenylene vinylene) derivatives. a step toward fully conjugated blue-emitting poly(p-phenylene vinylenes). *J. Am. Chem. Soc.*, 120:231, 1998.

[92] W. Wan, H. Antoniadis, V. Choong, H. Razafitrimo, Y. Gao, W. Feld, and B. Hsieh. Halogen precursor route to poly[(2,3-diphenyl-p-phenylene)vinylene] (dp-ppv): Synthesis, photoluminescence, electroluminescence, and photoconductivity. *Macromolecules*, 30:6567, 1997.

[93] C. Breneman and K. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *J. Comp. Chem.*, 11:361, 1990.

[94] U. von Barth. Basic density functional theory - an overview. *Physica Scripta*, T109:9–39, 2004.

[95] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118:11225, 1996.

[96] S. Kwasniewski, L. Claes, J.-P. Francois, and M. Deleuze. High level theoretical study of the structure and rotational barriers of trans-stilbene. *J. Chem. Phys.*, 118:7823, 2003.

[97] A. Hoekstra, P. Meertens, and A. Vos. Refinement of the crystal structure of trans-stilbene (tsb). the molecular structure in the crystalline and gaseous phases. *Acta Cryst. B*, 31:2813, 1975.

[98] K. Zimmerman. Oral: All purpose molecular mechanics simulator and energy minimizer. *J. Comp. Chem.*, 12:310–319, 1991.

[99] J. P. Ryckaert and A. Bellemans. Molecular dynamics of liquid alkanes. *Far. Disc. Chem. Soc.*, 66:95–106, 1978.

[100] B. J. Peter, H. M. Kent, I. G. Mills, Y. Vallis, P. J. G. Butler, P. R. Evans, and H. T. McMahon. BAR Domains as Sensors of Membrane Curvature: The Amphiphysin BAR Structure. *Science*, 303(5657):495–499, 2004.

[101] W. Tschöp, K. Kremer, O. Hahn, J. Batoulis, and T. Burger. Simulation of polymer melts. ii. from coarse-grained models back to atomistic description. *Acta Polymerica*, 49:75–79, 1998.

[102] C. Peter and K. Kremer. Multiscale simulation of soft matter systems – from the atomistic to the coarse grained level and back. *Soft Matter*, 5:4357–4366, 2009.

[103] W. Tschöp, K. Kremer, J. Batoulis, T. Burger, and O. Hahn. Simulation of polymer melts. i. coarse-graining procedure for polycarbonates. *Acta Polymerica*, 49:61–74, 1998.

[104] D. Fritz, V. A. Harmandaris, K. Kremer, and N. F. A. van der Vegt. Coarse-grained polymer melts based on isolated atomistic chains: Simulation of polystyrene of different tacticities. *Macromolecules*, 42(19):7579–7588, 2009.

[105] J. Shelley, M. Shelley, R. Reeder, S. Bandyopadhyay, and M. Klein. A coarse grain model for phospholipid simulations. *J. Phys. Chem. B*, 105(19):4464–4470, 2001.

[106] C. Abrams and K. Kremer. Combined coarse-grained and atomistic simulation of liquid bisphenol a-polycarbonate: Liquid packing and intramolecular structure. *Macromolecules*, 36(1):260–267, 2003.

[107] T. Murtola, E. Falck, M. Patra, M. Karttunen, and I. Vattulainen. Coarse-grained model for phospholipid/cholesterol bilayer. *J. Chem. Phys.*, 121(18):9156–9165, 2004.

[108] S. Izvekov and G. Voth. Multiscale coarse graining of liquid-state systems. *J. Chem. Phys.*, 123(13):134105, 2005.

[109] Q. Sun and R. Faller. Systematic coarse-graining of a polymer blend: Polyisoprene and polystyrene. *J.Chem. Theo. Comp.*, 2(3):607–615, 2006.

[110] V. Harmandaris, N. Adhikari, N. van der Vegt, and K. Kremer. Hierarchical modeling of polystyrene: From atomistic to coarse-grained simulations. *Macromolecules*, 39(19):6708–6719, 2006.

[111] L. Yelash, M. Müller, P. Wolfgang, and K. Binder. How well can coarse-grained models of real polymers describe their structure? the case of polybutadiene. *J. Chem. Theor. Comp.*, 2(3):588–597, 2006.

[112] A. Shih, A. Arkhipov, P. Freddolino, and K. Schulten. Coarse grained protein-lipid model with application to lipoprotein particles. *J. Phys. Chem. B*, 110(8):3674–3684, 2006.

[113] A. Lyubartsev. Multiscale modeling of lipids and lipid bilayers. *Eur. Biophys. J.*, 35(1):53–61, 2005.

[114] J. Zhou, I. Thorpe, S. Izvekov, and G. Voth. Coarse-grained peptide modeling using a systematic multiscale approach. *Biophys. J.*, 92(12):4289–4303, 2007.

[115] A. Villa, N. van der Vegt, and C. Peter. Self-assembling dipeptides: including solvent degrees of freedom in a coarse-grained model. *Phys. Chem. Chem. Phys.*, 11(12):2068–2076, 2009.

[116] F. Ercolessi and J. B. Adams. Interatomic potentials from 1st-principles calculations - the force-matching method. *Europhys. Lett.*, 26(8):583–588, 1994.

[117] B. Hess, C. Holm, and N. van der Vegt. Modeling multibody effects in ionic solutions with a concentration dependent dielectric permittivity. *Phys. Rev. Lett.*, 96:147801, 2006.

[118] D. Reith, M. Pütz, and F. Müller-Plathe. Deriving effective mesoscale potentials from atomistic simulations. *J. Comp. Chem.*, 24(13):1624–1636, 2003.

[119] A. Lyubartsev and A. Laaksonen. Calculation of effective interaction potentials from radial-distribution functions - a reverse monte-carlo approach. *Phys. Rev. E*, 52(4):3730–3737, 1995.

[120] A. Soper. Empirical potential monte carlo simulation of fluid structure. *Chem. Phys.*, 202:295–306, 1996.

[121] S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth. Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching. *J. Chem. Phys.*, 120(23):10896–10913, 2004.

[122] G. Toth. Effective potentials from complex simulations: a potential-matching algorithm and remarks on coarse-grained potentials. *J. Phys. Cond. Mat.*, 19(33):335222, 2007.

[123] J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Müller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, and V. Tries. *Advances in Polymer Science: Viscoelasticity, Atomistic Models, Statistical Chemistry.* Springer Verlag, Heidelberg, 2000.

[124] K. Kremer. In M. E. Cates and M. R. Evans, editors, *Soft and fragile matter, nonequilibrium dynamics, metastability and flow*, Proceedings

of NATO ASI Workshop, pages 423–445, Bristol, UK, 2000. Institute of Physics.

[125] F. Müller-Plathe. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *Chem. Phys. Phys. Chem.*, 3(9):754–769, 2002.

[126] M. Johnson, T. Head-Gordon, and A. Louis. Representability problems for coarse-grained water potentials. *J. Chem. Phys.*, 126(14):144509, 2007.

[127] T. Murtola, A. Bunker, I. Vattulainen, M. Deserno, and M. Karttunen. Multiscale modeling of emergent materials: biological and soft matter. *Phys. Chem. Chem. Phys.*, 11(12):1869–1892, 2009.

[128] V. A. Harmandaris, D. Reith, N. F. A. Van der Vegt, and K. Kremer. Comparison between coarse-graining models for polymer systems: Two mapping schemes for polystyrene. *Macromol. Chem. Phys.*, 208(19-20):2109–2120, 2007.

[129] A. Villa, C. Peter, and N. F. A. van der Vegt. Self-assembling dipeptides: conformational sampling in solvent-free coarse-grained simulation. *Phys. Chem. Chem. Phys.*, 11(12):2077–2086, 2009.

[130] R. Henderson. Uniqueness theorem for fluid pair correlation-functions. *Phys. Lett. A*, A 49(3):197–198, 1974.

[131] J. Chayes and L. Chayes. On the validity of the inverse conjecture in classical density functional theory. *J. Stat. Phys.*, 36:471–488, 1984.

[132] J. Chayes, L. Chayes, and E. Lieb. The inverse problem in classical statistical-mechanics. *Comm. Math. Phys.*, 93(1):57–121, 1984.

[133] S. Leon, N. van der Vegt, L. Delle Site, and K. Kremer. Bisphenol a polycarbonate: Entanglement analysis from coarse-grained md simulations. *Macromolecules*, 38(19):8078–8092, 2005.

[134] H. Wang, C. Junghans, and K. Kremer. Comparative atomistic and coarse-grained study of water: What do we lose by coarse-graining? *Eur. Phys. J. E*, 28(2):221–229, 2009.

[135] T. Murtola, E. Falck, M. Karttunen, and I. Vattulainen. Coarse-grained model for phospholipid/cholesterol bilayer employing inverse monte carlo with thermodynamic constraints. *J. Chem. Phys.*, 126(7):075101, 2007.

[136] W. G. Noid, J. W. Chu, G. S. Ayton, and G. A. Voth. Multiscale coarse-graining and structural correlations: Connections to liquid-state theory. *J. Phys. Chem. B*, 111(16):4116–4127, 2007.

[137] W. G. Noid, J. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen. The multiscale coarse graining method. 1. a rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.*, 128:244114, 2008.

[138] J. Hansen and I. McDonald. *Theory of simple liquids*. Academic Press, San Diego, CA, 1990.

[139] P. Jedlovszky, I. Bako, G. Palinkas, T. Radnai, and A. K. Soper. Investigation of the uniqueness of the reverse monte carlo method: Studies on liquid water. *J. Chem. Phys.*, 105(1):245–254, 1996.

[140] L. Larini, L. Lu, and G. A. Voth. The multiscale coarse-graining method. 6. implementation of three-body coarse-grained potentials. *J. Chem. Phys.*, 132(16):164107, 2010.

[141] I. Nezbeda and J. Slovak. A family of primitive models of water: Three-, four and five-site models. *Mol. Phys.*, 90(3):353–372, 1997.

[142] A. Wallqvist and R. D. Mountain. Molecular models of water: Derivation and description. *Rev. Comp. Chem.*, 13:183–247, 2007.

[143] P. G. Kusalik and I. M. Svishchev. The spatial structure in liquid water. *Science*, 65:1219–1221, 1994.

[144] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91:6269–6271, 1987.

[145] W. F. V. Gunsteren and H. J. C. Berendsen. A leap-frog algorithm for stochastic dynamics. *Mol. Sim.*, 1(3):173, 1988.

[146] W. G. Noid, P. Liu, Y. Wang, J. Chu, G. Ayton, S. Izvekov, H. C. Andersen, and G. Voth. The multiscale coarse-graining method. 2. numerical implementation for coarse-grained molecular models. *J. Chem. Phys.*, 128:244115, 2008.

[147] W. Janke. *Statistical Analysis of Simulations: Data Correlations and Error Estimation*, volume 10 of *NIC Series, From Theory to Algorithms, Lecture Notes*. 2002.

[148] W. L. Jorgensen and J. Tirado-Rives. Development of the opls-aa force field for organic and biomolecular systems. *J. Chem. Soc., Abstr.*, 216:U696–U696, 23 1998.

[149] W. L. Jorgensen and J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci.*, 102(19):6665–6670, 2005.

[150] J. H. Burroughes, D. D. C. Bradly, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, and A. B. Holmes. Light-emitting diodes based on conjugated polymers. *Nature*, 347:539, 1990.

[151] R. H. Friend, R. W. Gymer, A. B. Holmes, J. H. Burroughes, R. N. Marks, C. Taliani, D. D. C. Bradley, D. A. dos Santos, J. L. Bredas, M. Löglund, and W. R. Salaneck. Electroluminiscence in conjugated polymers. *Nature*, 397:121, 1999.

[152] H. Spreitzer, H. Becker, E. Kluge, W. Kreuder, H. Schenk, R. Demandt, and H. Schoo. Soluble phenyl-substituted ppvs - new efficient materials for highly efficient polymer leds. *Adv. Mater.*, 10:1340, 1998.

[153] J. Pedersen and P. Schurtenberger. Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules*, 29:7602–7612, 1996.

[154] Y. C. Li, K. B. Chen, H. L. Chen, C. S. Hsu, C. S. Tsao, J. H. Chen, and S. A. Chen. Fractal aggregates of conjugated polymer in solution state. *Langmuir*, 22:11009, 2006.

[155] Y.-C. Li, C.-Y. Chen, Y.-X. Chang, P.-Y. Chuang, C. J-H, H.-L. Chen, C.-S. Hsu, V. Ivanov, P. Khalatur, and S.-A. Chen. Scattering study of the conformational structure and aggregation behavior of a conjugated polymer solution. *Langmuir*, 25:4668–4677, 2009.

[156] C. Collison, L. Rothberg, V. Treemaneekarn, and L. Y. Conformational effects on the photophysics of conjugated polymers: A two species model for meh-ppv spectroscopy and dynamics. *Macromolecules*, 34:2346–2352, 2001.

[157] M. Knaapila, V. Garamus, L. Dias, F.B. Almasy, F. Galbrecht, A. Charas, J. Morgado, H. Burrows, U. Scherf, and A. Monkman. Influence of solvent quality on the self-organization of archetypical hairy rods-branched and linear side chain polyfluorenes: Rodlike chains versus beta-sheets in solution. *Macromolecules*, 39:6505–6512, 2006.

[158] M. Grell, D. Bradley, G. Ungar, J. Hill, and K. Whitehead. Interplay of physical structure and photophysics for a liquid crystalline polyfluorene. *Macromolecules*, 32:5810–5817, 1999.

[159] H.-P. Hsu, W. Paul, and K. Binder. Standard definitions of persistence length do not describe the local "intrinsic" stiffness of real polymer chains. *Macromolecules*, 43:3094, 2010.

[160] L. Schäfer, A. Ostendorf, and J. Hager. Scaling of the correlations among segment directions of a self-repelling polymer chain. *J. Phys. A: Math. Gen.*, 32:7875, 1999.

[161] L. Schäfer and K. Elsner. Calculation of the persistence length of a flexible polymer chain with short-range self-repulsion. *Eur. Phys. J E*, 13:225, 2004.

[162] L. Siebbeles, F. Grozema, M. de Haas, and J. Warman. Effect of backbone structure on charge transport along isolated conjugated polymer chains. *Rad. Phys. Chem.*, 72:85–91, 2005.

[163] O. A. von Lilienfeld and D. Andrienko. Coarse-grained interaction potentials for polyaromatic hydrocarbons. *J. Chem. Phys.*, 124(5):054307, 2006.

[164] R. Roe. *Methods of X-ray and Neutron Scattering in Polymer Science*. Oxford: New York, 2000.

[165] M. Rubinstein and R. Colby. *Polymer Physics*. Oxford: New York, 2003.

[166] A. Semenov. Theory of cluster formation in homopolymer melts. *Macromolecules*, 42:6761–6776, 2009.

[167] D. Hu, J. Yu, K. Wong, B. Bagchi, P. Rossky, and P. Barbara. Collapse of stiff conjugated polymers with chemical defects into ordered, cylindrical conformations. *Nature*, 405:1030, 2000.

[168] S. Naka, H. Okada, Y. Onnagawa, T. Yamaguchi, and T. Tsutsui. Carrier transport properties of organic materials for el device operation. *Synth. Met.*, 111-112:331–333, 2000.

[169] J. J. Kwiatkowski, J. Nelson, H. Li, J.-L. Bredas, W. Wenzel, and C. Lennartz. Simulating charge transport in tris(8-hydroxyquinoline) aluminium (alq3). *PCCP*, 10:1852, 2008.

[170] M. Cölle and W. Brütting. Thermal, structural and photophysical properties of the organic semiconductor alq3. *Phys. Stat. Sol. (a)*, 201:1095, 2004.

[171] M. Frisch et al. Gaussian 03, Revision C.02, 2004. Gaussian, Inc., Wallingford, CT, 2004.

[172] D. Williams. Improved intermolecular force field for molecules containing h, c, n, and o atoms, with application to nucleoside and peptide crystals. *J. Comput. Chem.*, 22:1154, 2001.

[173] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. Gromacs 4: Algorithms for highly-efficient, load-balanced and scalable molecular simulation. *J. Chem. Theory Comput.*, 4:435, 2008.

[174] C. Maree, R. Weller, L. Feldman, K. Pakbaz, and H. Lee. Accurate thickness/density measurements of organic light-emitting diodes. *J. Appl. Phys.*, 84:4013, 1998.

[175] M. Brinkmann, G. Gadret, M. Muccini, C. Taliani, N. Masciocchi, and A. Sironi. Correlation between molecular packing and optical properties in different crystalline polymorphs and amorphous thin films of mer-tris(8-hydroxyquinoline)aluminum(iii). *J. Am. Chem. Soc.*, 122:5147, 2000.

[176] K. Naito and A. Miura. Molecular design for nonpolymeric organic dye glasses with thermal stability: ralations between thermodynamic parameters and amorphous properties. *J. Phys. Chem.*, 97:6240, 1993.

[177] S. W. Watt, J. Chisholm, W. Jones, and S. Motherwell. A molecular dynamics simulation of the melting points and glass transition temperatures of myo-and neo-inositol. *J. Chem. Phys.*, 121:9565, 2004.

[178] M. Frisch et al. Gaussian 98, 1998. Gaussian, Inc., Pittsburgh, PA, 1998.

[179] H. Murata, G. Malliaras, M. Uchida, Y. Shen, and Z. Kafafi. Non-dispersive and air-stable electron transport in an amorphous organic semiconductor. *Chem. Phys. Lett.*, 339:161–166, 2001.

[180] V. Daggett, P. A. Kollman, and I. D. Kuntz. Molecular dynamics simulations of small peptides: Dependence on dielectric model and ph. *Biopolymers*, 31(3):285–304, 1991.

[181] R. Martin, J. Kress, I. Campbell, and D. Smith. Molecular and solid-state properties of tris-(8-hydroxyquinolate)-aluminum. *Phys. Rev. B*, 61:15804, 2000.

[182] D. Stauffer. Scaling theory of percolation clusters. *Phys. Rep.*, 54:1–74, 1979.

[183] M. Grünewald, B. Pohlmann, B. Movaghar, and D. Würtz. Theory of non-equilibrium diffusive transport in disordered materials. *Philos. Mag. B.*, 49:341–356, 1984.

[184] B. Movaghar, M. Grünewald, B. Ries, H. Bässler, and D. Würtz. Diffusion and relaxation of energy in disordered organic and inorganic materials. *Phys. Rev. B.*, 33:5545, 1986.

[185] N. G. Martinelli, M. Savini, L. Muccioli, Y. Olivier, F. Castet, C. Zannoni, D. Beljonne, and J. Cornil. Modeling polymer dielectric/pentacene interfaces: On the role of electrostatic energy disorder on charge carrier mobility. *Adv. Funct. Mat.*, 19(20):3254–3261, 2009.

[186] B. Derrida. Velocity and diffusion constant of a periodic one-dimensional hopping model. *J. Stat. Phys.*, 31(3):433–450, 1983.

[187] K. Seki and M. Tachiya. Electric field dependence of charge mobility in energetically disordered materials: Polaron aspects. *Phys. Rev. B*, 65(1):014305, 2001.

[188] G. H. Golub and C. F. Van Loan. *Matrix Computations, 3rd Edition*. Johns Hopkins University Press, Baltimore, 1996.

[189] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471, 1985.

[190] O. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani. Optimization of effective atom centered potentials for london dispersion forces in density functional theory. *Phys. Rev. Lett.*, 93:153004, 2004.

BIBLIOGRAPHY