Atomistic Modeling of the Solubility of Cellulose in Ionic Liquids

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II
Chemie, Physik und Mathematik

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt

von Frau Eliane Roos

ii

Erstgutachter: Prof. Dr. D. Sebastiani
    (Martin-Luther Universität Halle-Wittenberg)

Zweitgutachter: Prof. Dr. M. Brehm
    (Universität Paderborn)

Drittgutachter: Prof. Dr. M. Sulpizi
    (Johannes Gutenberg-Universität Mainz)

Tag der öffentlichen Verteidigung: 28.11.2024

Contents

iv

List of Abbreviations

ADF	**A**ngular **D**istribution Function
AGU	**A**nhydro**g**lucopyranose Units
AIMD	*Ab initio* **M**olecular **D**ynamics Simulation
BILFF	**B**io-Polymers in **I**onic **L**iquids **F**orce **F**ield
CDF	**C**ombined **D**istribution **F**unction
DACF	**D**imer **A**uto**c**orrelation **F**unction
DFT	**D**ensity **F**unctional **T**heory
GGA	**G**eneralized **G**radient **A**pproximation
GTH	**G**oedecker-**T**eter-**H**utter (pseudopotential)
IL	**I**onic **L**iquid
MD	**M**olecular **D**ynamics (Simulation)
MSD	**M**ean **S**quare **D**isplacement
NHC	**N**-**h**eterocyclic **C**arbene
NpT	Isothermic-isobaric Ensemble
NVE	Microcanonical Ensemble
NVT	Canonical Ensemble
RDF	**R**adial **D**istribution **F**unction
SDF	**S**patial **D**istribution **F**unction

1 Introduction

Lignocellulosic biomass is widely regarded as an almost inexhaustible source of raw material to meet the growing demand for environmentally friendly and biocompatible products.[1–4] This remarkable potential is primarily attributed to cellulose, a key component of lignocellulosic biomass. Cellulose has been used by humans for thousands of years in many applications, including construction, textile production, as packaging material, and writing surface.[5] However, the technical synthesis of celluloid in 1870 marked a significant turning point as the first synthesized thermoplastic polymer material. This breakthrough ushered in a new era of chemically modifying cellulose to produce innovative materials that could be manufactured on an industrial scale.[3, 6] Since this discovery, the range of applications for cellulose has continued to expand. In addition to the previously mentioned applications, cellulose and its derivatives are now being employed in the production of various other products, including bioplastics,[7] biofuels,[8] anti fungal membranes,[9] laminates,[10] high-performance electronics,[11] phosphorescent materials,[12] and numerous other innovative uses.

Cellulose and biomass in general play an important role in facilitating the development of more sustainable technologies and (smart) materials, as cellulose is one of the most abundant biopolymers and accounts for approximately 1.5×10^{12} tons of annual biomass production.[3] In contrast, the rising plastic production poses a significant environmental challenge. From 2008 to 2019, the annual plastic production increased from 245×10^6 tons to 368×10^6 tons.[2] This upward trend in plastic production, coupled with the finite nature of petroleum resources and the hazardous impact of plastic pollution on ecosystems,[13, 14] underscores the urgent need for a transition to non-fossil fuel-based alternatives. One solution is the development of biodegradable plastics, whose production capacity is projected to triple from approximately 2.2×10^6 tons in 2023 to around 7.4×10^6 tons by 2028.[15] Unlike conventional plastics usually derived from polyolefin, biodegradable bioplastics can be synthesized from renewable biopolymers such as polylactic acid, polyhydroxyalkanoates, and polybutylene succinate, as well as materials like starch, proteins, lipids, and, as previously mentioned, cellulose.[2, 16] The significant potential of cellulose fibers to replace fossil-fuel-based materials is highlighted by their ability to serve as packaging material in the food industry.[17] Notably, approximately 40 % of all plastic produced is dedicated to food packaging applications.[18] Moreover, cellulose fibers can find utility in various other fields, replacing nylon as suture material in surgical applications,[19] as a basis for hydrogels[20] and as 3D printable material,[21] among numerous other possibilities. The preference of cellulose fibers over other natural fibers, such as flax, is a result of their uniform fiber length, uniform morphology, and well-defined biochemical composition.[4] In addition, cellulose fibers can be generated from a wide range of sources, including food waste and agricultural byproducts,[22] as well as dead leaves.[7] Nevertheless, wood pulp remains the primary resource for cellulose processing.

Cellulose fibers are industrially produced through the Viscose or Lyocell procedures. These processes require dissolving cellulose, but the poor solubility of cellulose in conventional solvents presents challenges to both methods. The Viscose process involves the use of corrosive or toxic chemicals such as sodium hydroxide (NaOH) and carbon disulfide (CS_2) and the release of harmful gases, which poses risks to

both workers and the environment. On the other hand, the Lyocell process, which relies on the solution in *N*-methylmorpholine-*N*-oxide (NMMO), encounters several limitations. These include issues like poor solvent stability, degradation of both the solvent and cellulose, a substantial requirement for stabilizers and lower productivity levels.[4, 5]

A direct dissolution of raw cellulose not only simplifies processing, but also prevents the formation of byproducts. Therefore, there is a strong interest in finding environmentally friendly and non-toxic solvents that dissolve cellulose without derivatization. The use of solvents with negligible vapor pressure offers the advantage of minimizing the risk of atmospheric contamination and associated health concerns. In addition, non-flammability as well as thermal and chemical stability are desired for ease of handling. At the same time, a high cellulose solubility has to be ensured. As an alternative to conventional industrial solvents, ionic liquids (ILs) offer a new approach to the use of solvents while meeting the requirements described:[23–27] ILs are characterized by their organic salt nature with a melting point below 100 °C. The combination of an asymmetric organic cation with either organic or inorganic anions perturbs the crystal lattice structure and favors the existence of the system in the liquid state.[4] With an impressive range of over more than a trillion different ILs, they can be precisely tailored to specific reactions and physical properties by changing the choice of the anion or cation, allowing optimization of yield and product purity.[28] Because of their versatility, they are referred to as "designer solvents".[29] The adaptability of ILs in meeting chemical and physical criteria emphasizes the significance of systematically investigating the various factors that influence their capacity to act as solvents. This is essential to fully harness the potential of ILs in cellulose processing. The adaptability opens up opportunities not only in cellulose fiber production, but also in the in situ chemical modification of cellulose during the spinning process.[30] Since the pioneering discovery by Swatloski et al. that certain ILs can dissolve cellulose without derivatization,[31] researchers have been wondering why cellulose is soluble in ILs but not in other polar solvents like water. Even after more than two decades, some questions remain unanswered.

Multiple investigations have emphasized the substantial impact of various factors on the cellulose solubility in ILs, including process parameters (dissolution time, temperature and heating method),[31] cellulose-related factors (degree of polymerization and crystallinity of cellulose)[32–34] and IL-related aspects (purity, water content and the specific molecular structure of the ILs).[4, 31, 32, 35–39] It is discussed, that a driving force for the cellulose solubility is the formation of hydrogen bonds between the anion of the IL and the cellulose strands. The intra- and intermolecular hydrogen bonds of the hydroxyl groups of cellulose are broken by the formation of stronger interactions to the solvent. It was observed that certain anions such as dicyanamide with a low basicity and non-coordinating anions like hexafluorophosphate (PF_6) exhibit a low cellulose solubility. In contrast, acetate, formate, and chloride anions are regarded as more effective solvents for cellulose, a phenomenon discussed as being attributed to their superior capacity for forming hydrogen bonds.[31, 36, 40] However, it is controversial whether hydrogen bonding alone is the central cellulose-solvent interaction. For instance, there is ongoing discussion whether van der Waals forces also influence the solubility and how large their contribution is.[41]

Until recently, the influence of the chirality of the anion in the ILs was not known. We have demonstrated a significant difference in the solubility of cellulose in 1-ethyl-3-methyl-imidazolium-L-lactate ([EMIm][OLa-L]) and 1-ethyl-3-methyl-imidazolium-D-lactate ([EMIm][OLa-D]). Moreover, our experimental results showed that the solubility of cellulose in racemic ILs is lower compared to the two enantiopure forms,

possibly due to the formation of aggregates containing both L- and D-lactate anions.[37] The influence of the cation on the solubility has also long been debated. However, studies show that imidazolium, pyridinium, and ammonium cations facilitate a considerable degree of cellulose solubility.[40] One hypothesis is that aromatic-based ILs present lower relative interaction strengths between the cations and anions because of the reduction of the electrostatic strength due to the charge delocalization in the aromatic rings. As a result, more free anions are present for cellulose solubility.[42]

In addition, a decrease in solubility occurs when the chain length of the side chains of (imidazolium) cations is increased. It is suggested that this reduces the effective anion concentration and the hydrophobic interactions between the cations, which may reduce their ability to screen the anion.[31]

To increase the solubility of cellulose in ILs, co-solvents such as dimethyl sulfoxide (DMSO), *N,N*-dimethylformamide (DMF) and 1,3-dimethyl-2-imidazolidinone (DMI) can be used.[43, 44] These co-solvents reduce the viscosity of the mixture, accelerate the dissolution rate, and minimize the time required for cellulose dissolution.[5] Furthermore, it is discussed that the interaction of DMSO with the IL weakens the association between the cations and anions, facilitating a stronger interaction between the anion and the cellulose strand.[44–46]

To enhance the cellulose solubility in ILs and identify the optimal solvent mixture for application, a comprehensive understanding of the intricate process of the cellulose dissolution is essential. This involves considering the numerous interactions within the system. A comprehensive understanding of the interplay between the solvent and cellulose can be achieved through the use of molecular dynamics (MD) simulations.

The present work provides a tool for MD simulations of these large systems over long time scales: a force field for simulating the solubility of cellulose in four different ILs and DMSO in the presence and absence of water, called *BILFF* (*B*io-*P*olymers in *I*onic *L*iquids *F*orce *F*ield). The force field comprises parameters for cellulose, the ILs 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]), 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]), 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]), and 1-ethyl-3-methylimidazolium benzoate ([EMIm][OBz]), along with DMSO.

The IL [EMIm][OAc] is one of the most commonly used cellulose solvents. It has a lower melting point, lower toxicity and higher cellulose solubility than other ILs such as the well-known 1-butyl-3-methylimidazolium chloride ([BMIm][Cl]).[47] In the processing of cellulose with [EMIm][OAc], the IL can serve not only as a solvent but also as an active reagent, forming the industrially important derivative, cellulose acetate.[48] A disadvantage of [EMIm]$^+$ is its ability to abstract a proton of the imidazolium ring, which can lead to the formation of *N*-heterocyclic carbenes (NHCs). NHCs are extremely reactive and form dimers or unwanted byproducts. Thus, they can also cause the undesirable degradation of cellulose.[49] However, by replacing the carbon atom with a third nitrogen atom and forming [EMTr][OAc], proton abstraction can be avoided. At the same time, [EMTr][OAc] exhibits only a slightly reduced cellulose solubility. The potential use of triazolium-containing ILs as a non-derivatizing cellulose solvent is almost unexplored and has been described for the first time by Brehm et al. five years ago.[35] Therefore, there are almost no in-depth results on the effect of these ILs with cellulose.

The influence of the aromaticity of the anion on the interactions in the cellulose/IL system is also poorly understood, although it is known that the aromatic anion [OBz]$^-$ has a negative effect on cellulose solubility.[35, 36] The cause of this effect is

not yet clear. Therefore, in the present work, a force field was optimized for the ions [EMTr]$^+$ and [OBz]$^-$ for the first time in the literature.

Intentionally swapping these four different ions within the force field MD simulation, such as exchanging [EMTr]$^+$ with [EMIm]$^+$ and [OBz]$^-$ with [OAc]$^-$, enables insightful observations of their impact on cellulose solubility. Furthermore, this comprehensive dataset provides detailed insights into the influence of co-solvents and water on the simulated systems.

Force field MD simulations have previously offered initial insights into the hydrogen bonding network among cellulose, [EMIm][OAc], and DMSO.[36, 50–53] In these simulations, general force field parameters were employed, on the basis of AMBER[54–56] and OPLS-AA[57–65] as well as others.[66] However, these force field parameters were not specifically adjusted to accurately reproduce the hydrogen bond network and local structure of the cellulose/IL/DMSO/water system, which may lead to inaccurate results. In addition, there is ongoing research to create polarizable force fields for ionic liquids,[67] but none are yet available for the aforementioned ILs.

In this thesis, the force field BILFF is presented, which has the following characteristics:

- *Fast MD simulations of cellulose/IL/DMSO/water mixtures:*
  By using BILFF, force field MD simulations of the quinary systems can be performed within a few days with an accuracy approaching quantum chemical precision. This includes not only structural parameters such as ion pairing, solvation shell structure, and hydrogen bonding but also dynamic properties such as diffusion coefficients, and hydrogen bond lifetimes. By adjusting the force field parameters both in the presence and absence of water, the scope of BILFF is extended to force field simulations of not only the solubility of cellulose in ILs but also precipitation of cellulose by water.

- *Focus on accurate and precise modeling of interactions:*
  BILFF is the first force field to provide optimized force field parameters for the ILs [EMIm][OAc], [EMIm][OBz], [EMTr][OAc], [EMTr][OBz], as well as DMSO, specifically tailored to investigate the hydrogen bond network and solubility of cellulose in these systems. For the first time, the bridging acetal oxygen atom and the ring oxygen atom of cellulose were distinguished atomistically and treated as different atom types.

- *Reproduction of experimental data:*
  The calculated results from force field simulations with BILFF are in agreement not only with other quantum chemical simulations, but also with experimental data, including the thermal volume expansion coefficient, compressibility, self-diffusion coefficient, density, enthalpy of vaporization (of pure DMSO) and, in the case of cellulose, X-ray diffraction data of the crystal structure (XRD data).

- *Transferability:*
  Despite the fact that the force field parameters are not specifically optimized for [EMTr][OBz], but rather are designed for [EMIm][OBz] and [EMTr][OAc], it can be shown that quantum chemical simulations can still be accurately reproduced.

- *Applicability over a wide temperature range:*
  It is demonstrated that force field MD simulations with BILFF provide accurate

results over a wide temperature range of 350–550 K in agreement with quantum chemical simulations and experimental data, providing opportunities for further process optimization.

This cumulative thesis is structured as follows:
In the second chapter, the state of the art of the cellulose solubility is presented. The molecular structure of cellulose is explained as it is responsible for its physico-chemical properties, such as its insolubility in conventional solvents. The chapter discusses different solvents and presents the current research regarding the mechanism of dissolving cellulose in ILs. Finally, methods for performing MD simulations using force fields and quantum chemical approaches are explained, and the applied methods for analyzing the microstructure and dynamics of the simulated systems are described.

Chapter 3 presents the development of BILFF and its application in studying the solubility of cellulose in the different ILs, DMSO, and water. The content consists of four published articles and an additional section covering unpublished further work. The articles include the description of the optimization process and validation of BILFF, alongside its application in MD simulations. At the same time, the hydrogen bond network in the respective solvents is investigated and conclusions for the cellulose solubility are drawn.

Chapter 4 summarizes the results from the published articles and the continuing work. The main findings are highlighted and possible further applications for the use of BILFF are discussed. BILFF's force field parameters, along with additional analysis results, can be found in the appendix.

2   Theoretical Framework

2.1   Interplay of Cellulose and Ionic Liquids: Structure, Interactions and Solubility

In this section, the structure of the bio-polymer cellulose and the resulting intra- and intermolecular interactions are explained in more detail. In this context, the solubility of cellulose in ionic liquids is discussed. A comprehensive description of cellulose structure and its solubility is presented in Ref. [5] and [40], while Ref. [41] and [68] briefly discuss the characteristics of cellulose interactions, as summarized below.

2.1.1   Molecular Structure of Cellulose

Cellulose is a linear polymer of D-glucose units, which are linked by β-(1→4)-glycosidic bonds. Due to the constraints of the β-linkage and intramolecular interactions, every second monomer is rotated by 180° compared to its neighbors. The glucose monomers in the cellulose polymer are often referred to as *anhydroglucopyranose units* (AGU) (cf. Fig. 2.1). The number of AGU determines the degree of polymerization (*DP*) of cellulose, which varies depending on the origin of the bio-polymer. For example the degree of polymerization varies between 9 000–10 000 in native wood, 300–1700 in wood pulp[a] and 250–500 in regenerated cellulose fibers. In cellulose, one end of the polymer contains the anomeric carbon, which is the carbon atom involved in the formation of a glycosidic bond. This particular end is known as the non-reducing end. At the opposite end of the polymer, the anomeric carbon remains unbound and can form an aldehyde structure. This side is referred to as the reducing end of the polymer, where the glycosidic linkage and the aldehyde structure coexist in equilibrium.
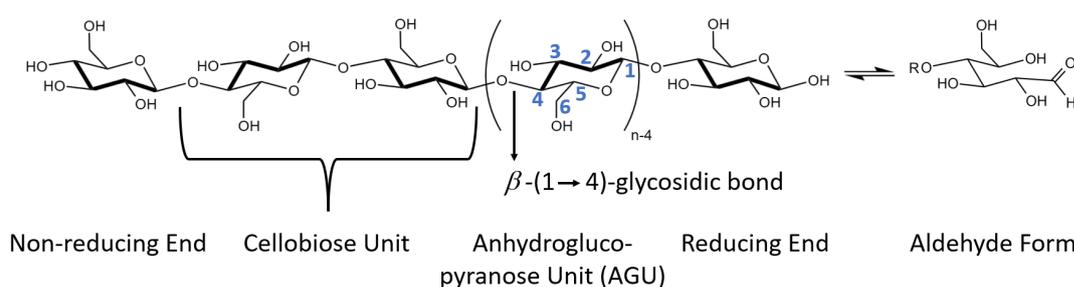
FIGURE 2.1: Molecular structure of cellulose (n = degree of polymerization). The reducing end of the polymer coexists in equilibrium with an aldehyde form.

2.1.2   Hydrogen Bonds and Van der Waals Forces

The equatorial positioning of its hydroxyl groups and the presence of axial CH groups enables both hydrophilic and hydrophobic properties of the cellulose polymer. Due to intramolecular hydrogen bonds between the hydroxide groups in one

---

[a]Pulp: lignocellulosic fibrous material that is obtained by the chemical or mechanical separation of cellulose fibers from wood, fiber crops, or waste paper.[40]

cellulose strand as well as intermolecular hydrogen bonds and van der Waals interactions between the single sheet structures, the polymer strands exhibit a semi-ordered structure (cf. Section 2.1.3). Because of their importance to the structure of cellulose and its solubility, these two types of interactions are discussed briefly below.

Hydrogen bonds occur even between two isolated molecules and involve two bonding partners referred to as "donor" (D) and "acceptor" (A). In a hydrogen bond, D donates a proton to A, resulting in the formation of a bond known as D-H$\cdots$A. The bond strength varies depending mainly on the involved atoms and ranges from low energies of approximate $1\ k_B T$ in C-H$\cdots$C bonds over the range of 5–15 $k_B T$ in OH$\cdots$O bonds to very high energies of over 50 $k_B T$ in HF.[41]
The exact nature of hydrogen bonding is still under debate. It is discussed that it is a combination of electrostatic, covalent, and dispersion forces.[69–71] The presence of a positive charge density on the hydrogen atom and a negative charge density on A contributes to an electrostatic nature of the bond. However, for example, Compton scattering experiments of ice, where the intermolecular bonding consists primarily of hydrogen bonds, support the idea of a charge transfer or even covalence of the bond.[69]

The induction and dispersion effects between atoms can be summarized as so-called *van der Waals forces*. These forces occur even between the hydrophobic groups of cellulose and play a crucial role in its overall thermodynamic stability. Van der Waals forces consist of attractive and repulsive components. The attractive component is given by dispersion forces (also known as London forces) and arise from the formation of temporary dipoles during fluctuations in the electron clouds of surrounding atoms. When an atom develops an instantaneous dipole, it can induce a dipole in neighboring atoms, leading to an attractive force between them. When the atoms are at a considerable distance from each other, the attractive potential is relatively small. However, as the two atoms approach each other, the potential energy becomes increasingly negative, intensifying the attraction between them. This phenomenon continues as the atoms move closer, causing their electron densities to overlap, resulting in a phenomenon known as interpenetration. This results in a strong repulsion, referred to as Pauli repulsion, which causes the energy to rapidly increase. Pauli repulsion arises from the prohibition of any two electrons in a system from having the same set of quantum numbers.[72] The attractive and repulsive components of the van der Waals forces can be summarized in the Lennard-Jones (LJ) potential, which is discussed in Section 2.2.2.

2.1.3   Crystal Structure of Cellulose

Three different polymorphs of cellulose have been identified: cellulose I, II and III, which differ in their hydrogen bonding. (In some literature also a fourth additional crystal form, cellulose IV, is distinguished.[5, 74, 75]) The most common form is cellulose I and can be categorized into I_α and I_β,[76, 77] whereas I_α is dominating in bacteria and algae and I_β in higher plants.[5] Cellulose I_α exhibits a triclinic unit cell with a single chain. Cellulose I_β is monoclinic and characterized by two distinct layers arranged in alternating sheets (cf. Fig. 2.2). In both polymorphs of cellulose I, strong hydrogen bonds are only present within the individual layers. Between the layers, van der Waals interactions dominate,[1, 78] although it is discussed whether weak
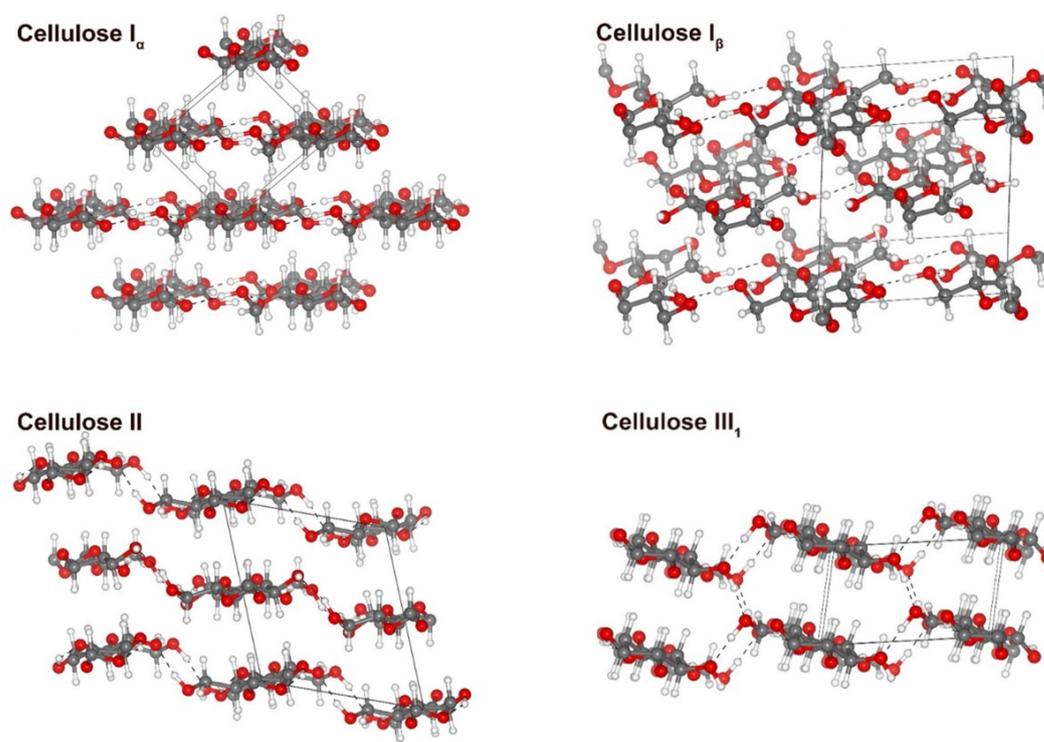
FIGURE 2.2: Crystal structure of the cellulose allomorphs (atom color code: red – O; grey – C; white – H). The hydrogen bonds between the sheets and across the layers are shown with dashed lines. The Figure was taken from Ref. [73].

C-H$\cdots$O hydrogen bonds also play a role.[79] Through processes such as regeneration or alkali and ionic liquid treatment, cellulose I undergoes a facile and irreversible transformation into the monoclinic cellulose II by rearranging the cellulose strands into antiparallel sheets with a three-dimensional network of intra- and inter-layer hydrogen bonds. Cellulose III can be obtained from cellulose I and cellulose II by treatment with anhydrous ammonia. This form has a parallel strand orientation like cellulose I, but with intra- and interlayer hydrogen bonding. The layers are staggered, while in cellulose I they are "flat".[1, 5, 34] (Some literature also distinguishes between cellulose III$_I$ and III$_{II}$, indicating that it originated from cellulose I and II, respectively.[1])

The varying stacking configurations and hydrogen bonding within and between the layers in the different cellulose morphologies impact the decrystallization and depolymerization. For example, studies have shown that the decrystallization work required by the enzyme cellulase is higher for strands of cellulose I$_\alpha$ and I$_\beta$ than for celluloses II and III.[33] Furthermore, the crystal allomorph and relative crystallinity of cellulose have been shown to influence the range of primary products produced by fast pyrolysis. In addition the different cellulose allomorphs have been observed to have an effect on the phase change and viscoelastic properties of cellulose during the rapid heating.[34]

In this context, it is important to note that cellulose does not exist as a single crystal in its natural state. The crystalline regions in cellulose are interrupted by regions of amorphous structures, resulting in a heterogeneous composition, i.e. a semi-crystalline state. However, it is still not understood exactly how the amorphous regions are formed.[1]

2.1.4   Cellulose Solvents

To enable a homogeneous chemistry of cellulose and facilitate its characterization and shaping, it is often desirable to place the polymer in solution. However, preparing a cellulose solution remains challenging: While molecules such as glucose, cellobiose, and cellulose oligomers with a *DP* of less than 10 are soluble in simple solvents such as water, cellulose itself is not. To understand the difference in solubility of glucose and cellulose, the *Gibbs free energy* ΔG can be examined

$$\Delta G = \Delta H - T\Delta S, \tag{2.1}$$

where T represents the temperature, and ΔS denotes the entropy change. In an ideal solution, the enthalpy of mixing (ΔH) is zero. Consequently, there is no energetic advantage or disadvantage for the components to interact homomolecular or heteromolecular. In the case of polymers in solution, usually non-ideal mixing occurs. Considering the equation for Gibbs free energy (eq. 2.1), the solubility of a substance depends on the balance between the enthalpy change and the temperature-dependent entropy change ($T\Delta S$). If the dissolved state has a lower Gibbs free energy ($\Delta G < 0$), the dissolution is energetically favorable and the substance is soluble. The enthalpy change accounts for the energy involved in breaking and forming of intermolecular interactions, while the entropy change represents the increase in disorder or randomness associated with the dissolution process. When comparing the interactions between a glucose molecule and its solvent and the interactions between cellulose and its solvent, it is expected that similar types of interactions will occur. The entropy gain of glucose monomers is significantly higher than that of cellulose. This discrepancy is a contributing factor why solvents like water easily dissolve glucose molecules due to their greater entropy gain, whereas cellulose does not exhibit the same level of solubility.[5, 80]

However, water can cause cellulose to swell. Swelling is a process in which the solvent penetrates the polymer structure to some extent, interact with it and cause significant changes in volume and physical properties. Despite these changes, the cellulose remains in a solid or semi-solid state. In contrast, dissolution completely destroys the supramolecular structure of cellulose due to a stronger interaction between the solvent and the polymer.

The structure of cellulose is largely based on hydrogen bonds. To disrupt these hydrogen bonds within cellulose and facilitate its dissolution, solvents with a substantial capacity for hydrogen bonding and/or robust Coulomb interactions are essential. Diverse solvent options are available:

Based on the potential interactions between cellulose and solvents Turbak classified cellulose solvents into four main categories:[49, 81]

1. Cellulose as a base with an acidic solvent, e.g. H_2SO_4

2. Cellulose as an acid with a basic solvent, e.g. NaOH

3. Cellulose as a ligand with a complexing agent as the solvent, e.g. Cuoxam

4. Cellulose as a reactive compound, leading to the formation of soluble transient derivatives or intermediates, e.g. xanthate.

Another classification of cellulose solvents is the distinction between non-derivatizing and derivatizing solvents, and aqueous and non-aqueous solvents. Non-derivatizing solvents dissolve cellulose only through intermolecular interactions.
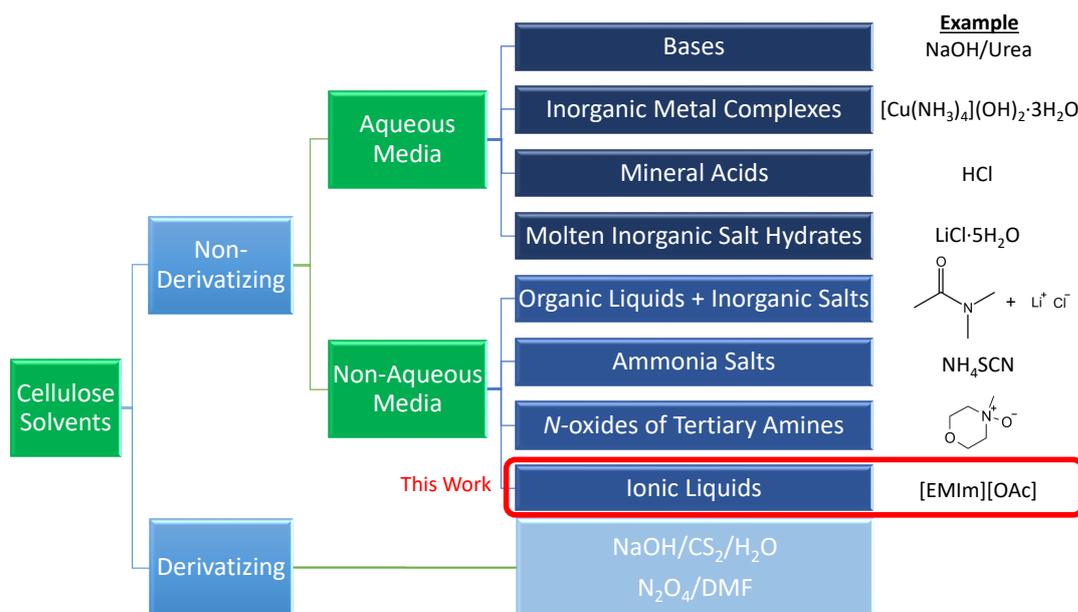
FIGURE 2.3: Classification of cellulose solvents with some examples.

Derivatizing solvents, however, dissolve cellulose while forming ether, ester, or acetal derivatives. A possible combination of these two classifications for cellulose solvents is shown in Fig. 2.3. The selection of a solvent depends on the intended purpose of keeping cellulose in solution. An ideal solvent for cellulose should possess specific characteristics. Firstly, it should be efficient in dissolving cellulose while being inert towards the cellulose strand itself. Additionally, the solvent should be recyclable and non-hazardous to the environment and human health. Furthermore, the final degradation of the solvent-cellulose system should yield non-hazardous and biodegradable products.

From the variety of solvents, selected solvent classes are explained in more detail below in accordance with Fig. 2.3:[b]

Cellulose is partially soluble in a 10 wt.% aqueous solution of sodium hydroxide, depending on its degree of polymerization.[5, 82] However, the amphiphilic properties of cellulose result in the formation of aggregates rather than complete dissolution.[83] To improve the dissolution process, NaOH solutions can be modified with additives such as poly(ethylene glycol), urea and/or thiourea.[84–86] These additives are environmentally friendly and have a low toxicity. However, the amount of additives required in these systems makes recovering and reusing processes an economic necessity. Various analytical techniques, including NMR spectroscopy, small-angle neutron scattering, and wide-angle X-ray scattering, suggest that the dissolution mechanism in NaOH/urea involves the formation of hydrates of NaOH, which are able to form hydrogen bonds with cellulose at low temperatures. The urea molecules surround the cellulose/NaOH/H_2O inclusion complex, preventing cellulose aggregation.[5, 87, 88] Between urea and cellulose, no direct interactions have been found yet.

The solutions of cellulose in aqueous alkali with urea and/or thiourea are extensively studied for shaping purposes, such as regenerating cellulose into fibers and

---

[b]A comparison of the specific solubility of all the solvents mentioned is currently not feasible due to insufficient data using consistent measurement parameters.

membranes.[89]

An interesting alternative as cellulose solvents are inorganic metal complexes. These complexes consist of transition metal ions and nitrous ligands, offering unique properties for cellulose dissolution. According to Saalwächter et. al., Cuoxam ($[Cu(NH_3)_4](OH)_2 \cdot 3\,H_2O$), Ni-tren ($[Ni(NH_2-CH_2-CH_2)_3N](OH)_2$), and Cd-tren ($[Cd(NH_2-CH_2-CH_2)_3N](OH)_2$) have shown a high efficiency of this type of cellulose solvent. Between the cation and cellulose a coordinative binding takes place, where the cation binds to two deprotonated OH groups at the C2 and C3 positions of the AGU, acting as a chelating ligand. On the other hand, in Cu-en ($[Cu(H_2N-CH_2-CH_2-NH_2)_2](OH)_2$), there is no coordinative binding but rather pure Coulomb interactions with cellulose. However, there are some problems associated with these solvents. Most of them are strongly colored, which can pose issues in analytical applications that rely on light scattering or refractive index measurements as the color of these solvents may interfere with the accuracy and reliability of the analytical results.[5, 90] Additionally, some representatives of inorganic metal complex-based solvents can cause a cellulose depolymerization even in the presence of trace amounts of oxygen.[91] This is a significant concern when working with such solvents, as uncontrolled depolymerization of cellulose should be avoided.

Other solvents for cellulose in aqueous media are molten inorganic salts such as $MgCl_2 \cdot 6H_2O$ and $LiCl \cdot 5H_2O$. While these solvents can be very effective, they drastically reduce the chain length of cellulose.[92] When using the solvent $Ca(SCN)_2$, cellulose is likely to form a complex and an addition compound with the thiocyanate anion. The crystallinity of the regenerated cellulose product seems to be influenced by the precipitant employed. Specifically, using water as the precipitant results in the recovery of only cellulose II. In contrast, when methanol is utilized, the predominant product is mainly cellulose I.[5, 93]

Another solvent mixture for cellulose is certain organic liquids in combination with inorganic salts. For example, for homogeneous modifications and analytical characterization, a mixture of dimethylacetamide (DMAc) and lithium chloride ($C_4H_9NO + LiCl$) is often used as a solvent system (cf. Fig. 2.3 for molecular structure). The lithium atom interacts with the carboxyl carbon of DMAc, forming a macrocation. This macrocation then interacts with the hydroxyl groups of cellulose through a hydrogen bond bridged by chloride ions.[5, 94] A major advantage of this solvent mixture is that the solutions are colorless and compatible with gel permeation chromatography columns. However, it is necessary to prevent the presence of water, otherwise the interactions between the salt and the cellulose cannot take place. To ensure a water-free system, the DMAc/LiCl/cellulose system must be distilled and the solvent itself dried.[94, 95] These processes require additional effort and resources, and therefore are not widely applied.

An industrially used non-derivatizing cellulose solvent is *N*-methylmorpholine-*N*-oxide (NMMO), a *N*-oxide of a tertiary amine (cf. Fig. 2.3 for molecular structure). NMMO has the ability to directly dissolve high concentrations of cellulose without altering the chemical properties of the cellulose strand through hydrogen bonding. Additionally, it allows for the recycling of over 99 % of the solvent after usage. The typical dissolving process in NMMO, the Lyocell process, involves mixing a suspension of cellulose in the solvent with a large excess of water (10–15% pulp, 50–60% NMMO, 20–30% water) in the first step. The excess water facilitates low

viscosity and superior mixing. Subsequently, the water is removed at temperatures between 100 and 120 °C, under reduced pressure, until complete cellulose dissolution is achieved.[5, 96]

A significant challenge associated with NMMO is its thermal instability due to its energy-rich N-O bond. NMMO is an oxidant and is sensitive to any form of catalytic impurities in the spinning dope, which can lead to side reactions. To address this issue, stabilizers must be added at the beginning of the dissolution process. Furthermore, NMMO degradation can occur in the presence of transition metal ions such as iron and copper. Therefore, strict exclusion of these impurities is necessary in the process.[3, 5, 96] In addition, it is important to note that there are two stable hydrates of NMMO: the monohydrate $NMMO \cdot H_2O$ and $NMMO \cdot 2.5 H_2O$. The latter hydrate can pose challenges, as cellulose precipitation occurs when the water content surpasses $1.2 H_2O$.[5, 97]

The viscose method is another widely employed industrial process and uses a $NaOH/CS_2$ solution as a solvent. In the initial stage, cellulose undergoes swelling by forming hydrates in the presence of NaOH and H_2O: The solvent permeates the amorphous regions of cellulose, initiating its expansion. Subsequently, alkali cations diffuse into the crystalline regions, resulting in the formation of an intermediate known as "alkali cellulose". This intermediate undergoes irreversible swelling, a phenomenon recognized as mercerization. In the subsequent xanthation process, the previously mercerized cellulose reacts with gaseous CS_2 to produce sodium cellulose xanthate, which is then dissolved in diluted alkali to yield a uniform viscose dope for spinning. The resulting cellulose xanthate is subsequently precipitated as cellulose fibers in water.[5, 98] However, it's crucial to note that this process requires a significant amount of environmentally unfriendly CS_2. This leads to the generation of toxic waste, contributing to environmental pollution and posing risks to human health. Additionally, only a small fraction of the CS_2 is recovered, while the majority is released into the environment.[98]

Solubility of Cellulose in Ionic Liquids

Another non-derivative solvent for cellulose that is non-toxic, recyclable and environmentally friendly, depending on the choice of molecules, is ionic liquids (ILs). ILs are organic salts with a melting point below 100 °C. They are characterized by a low vapor pressure, low volatility, non-flammability, excellent solvation potential, and high thermal and chemical stability.[28, 99] The immense variety of possible combinations of cations and anions allows for tuning the chemical and physical properties of the resulting salts, as the number of possible ILs is estimated to be between 10^{12} and 10^{18}.[28]

Since the synthesis of ethylammonium nitrate in 1914 as the first IL at room temperature[24], the field has experienced significant growth. As early as 1934, Graenacher discovered that liquefied quaternary ammonium salts can dissolve cellulose.[100] In 2002, Swatloski demonstrated the ability of the neat imidazolium salt 1-butyl-3-methylimidazolium chloride ([BMIm][Cl]) to dissolve cellulose.[31]

Nowadays, ILs find applications in numerous areas of research and industry. They have proven to be highly versatile and valuable in various fields such in materials science,[101] as pharmaceuticals,[102] electrolytes in batteries,[103] solvents for synthesis and catalysis,[23, 104] as well as e.g. lubricant additives,[105] among many others.

Alkylammonium   *N*-alkylpyridinium   Dialkylimidazolium
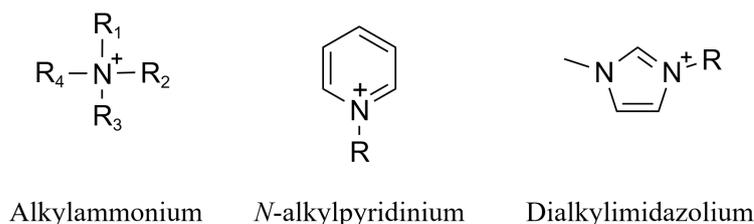
FIGURE 2.4: Cations of ionic liquids for cellulose processing.

ILs have also proven to be effective solvents in cellulose technology, including dissolution for regeneration,[38] homogeneous derivatization,[106] and biomass processing, such as wood component separation.[107]

There are several classes of ionic liquids that can dissolve cellulose, including alkylammonium-, *N*-alkylpyridinium-, and dialkylimidazolium-based ILs (cf. Fig. 2.4). The ability to dissolve cellulose increases from the first to the last one.[40]

It is discussed that the anion of the IL infiltrates the cellulose structure and disrupts the natural cellulose structure by competing with its hydrogen bonds. ^{13}C-NMR relaxation measurements have confirmed that in 1-butyl-3-methylimidazolium chloride ([BMIm][Cl]), the chloride ions interact approximately stoichiometrically with the hydroxyl groups of cellulose. This leads to the formation of an electron donor – electron acceptor complex, which disrupts the interactions between the polysaccharide chains and subsequently enables the polymer dissolution.[5, 108]

The solubility of cellulose depends on the type and structure of the anion. Previously unknown was the fact that the chirality of the IL anion plays a role. We were able to show that the solubility of cellulose is significantly higher in 1-ethyl-3-methylimidazolium-L-lactate ([EMIm][OLa-L]) than in 1-ethyl-3-methylimidazolium-D-lactate ([EMIm][OLa-D]) (cf. Tab. 2.1). Furthermore, the solubility of cellulose in the racemic IL is lower compared to the two enantiopure forms. One hypothesis is an intriguing interplay between lactate anions of opposite chirality, potentially leading to the formation of aggregates containing both L- and D-lactate anions in racemates. This might lead to a reduction in the number of free anions accessible for cellulose solvation.[37] Furthermore, the influence of aromatic or aliphatic anions is still under investigation.[36] An example is the comparison of solubility in 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) and benzoate ([EMIm][OBz]) in Tab. 2.1. (See also Chapter 3 of this thesis.)

However, the role of the cation in the dissolution process is still a subject of debate. It has been observed that while the butylimidazolium salt is capable of dissolving cellulose, the corresponding salts with higher alkyl chains are unable to do so. This phenomenon can be attributed to the relationship between the ability to dissolve cellulose and the concentration of anions in the immediate vicinity of the polymer. Smaller cations enable a greater number of anions to surround the cellulose chain, thereby facilitating effective dissolution.[5, 31]

Furthermore, the solubility of cellulose in imidazolium salts undergoes changes when the acidic proton at C2 is substituted with a methyl group,[32, 38] as shown in Tab. 2.1 (cf. entries [BMIm][Cl] and [BdMIm][Cl]). Accordingly, not only the length of the alkyl chain is important, but also its position on the ring, which influences other important properties such as glass transition temperature, viscosity, and thus the cellulose solubility.[31, 39, 109, 110]

In the context of cellulose solubility, the Kamlet–Taft parameters and their shift to a point where cellulose becomes soluble in ILs are also discussed.[43] The Kamlet–Taft

parameters consist of the dipole polarizability (π^*), the acidity of the hydrogen bond donor (α) and the basicity of the hydrogen bond acceptor (β).[111–113] Of particular importance is the β-value, which is primarily influenced by the anionic species of the IL.[114] It has been observed that there is a linear correlation between this value and the solubility of cellulose in pure ILs:[115] ILs with a β-value ranging from 0.80 to 1.20 are considered to be effective cellulose solvents.[116] For instance, the excellent cellulose solvent [EMIm][OAc] possesses a basicity of $\beta = 1.11$,[117] whereas the less effective solvent [BMIm][Cl] has a value of $\beta = 0.87$[118] (cf. cellulose solubilities in Tab. 2.1). On the other hand, a non-solvent of cellulose, [BMIm][PF$_6$], exhibits a β-value of 0.21.[114]

In addition, studies have shown that the ion mobility also affects cellulose solubility. Tomimatsu et al. found that the solubility of cellulose in an IL/DMSO mixture gradually increased from a DMSO mole fraction of 0 to about 0.8. (The solubility of cellulose was determined by systematically increasing its concentration, stirring for 6 hours after each addition, and subsequently verifying complete cellulose dissolution.) Interestingly, the β value remained constant within this concentration range. They suggests that the increase in cellulose solubility was not only influenced by the β-value, but was also due to an increase in the ion mobility.[44] In this process, DMSO acts as a co-solvent. Aprotic and highly polar co-solvents such as DMSO, *N,N*-dimethylformamide (DMF) and 1,3-dimethyl-2-imidazolidinone (DMI) are able to significantly increase the cellulose solubility (cf. Tab. 2.1).[43, 44] The absence of acidic protons ensures that the co-solvent does not disrupt the hydrogen bonds between the anions and the cellulose molecules, which are crucial for an effective dissolution.[119] These co-solvents serve to reduce the viscosity of the solution, leading to a faster dissolution rate and ultimately minimizing the time required for the cellulose dissolution. Moreover, the incorporation of a co-solvent enhances the overall manageability and ease of handling of the solution.[5]

Aside from reducing viscosity, the process by which co-solvents enhance cellulose solubility is still a topic of ongoing discussion. The proposed mechanism for the DMSO/IL/cellulose dissolution suggests that it comprises two steps, as illustrated in Fig. 2.5. At first, in the pure IL, there is a strong association between the ions leading to the formation of tightly bound ion pairs. When co-solvents such as DMSO are introduced into the system, this strong ion association is partially disrupted. As a result, larger amounts of free cations and anions are available. In the second step of the mechanism, the dissociated free anions interact with cellulose, facilitating its dissolution.[44–46] (See also Section 3.4 of this thesis.)

However, although the mechanism of cellulose dissolution appears to be mainly based on the disruption of hydrogen bonding of cellulose by the solvent, there are studies suggesting that other interactions and effects are also involved.[41] Thus, a debate has arisen to understand cellulose more as an amphiphilic molecule due to its hydrophilic and hydrophobic groups.[41, 120, 121] It is not yet clear how large the influence of e.g. van der Waals interactions as well as the above-mentioned factors (aromaticity and chirality of the anion, choice of cation, process parameters, ...) is. A deeper insight into the interaction between ILs and cellulose can be obtained by using MD simulations, whereas this work provides a tool for fast and precise force field MD simulations specifically designed for this investigative purpose.

After the dissolution of cellulose, a typical pretreatment process involves the recovery and desiccation of the polymer. In order to precipitate cellulose, protonic antisolvents such as water, ethanol, methanol, and others, can be used to disrupt the IL-cellulose interactions.[31, 119, 125] The macroscopic morphology of the regenerated
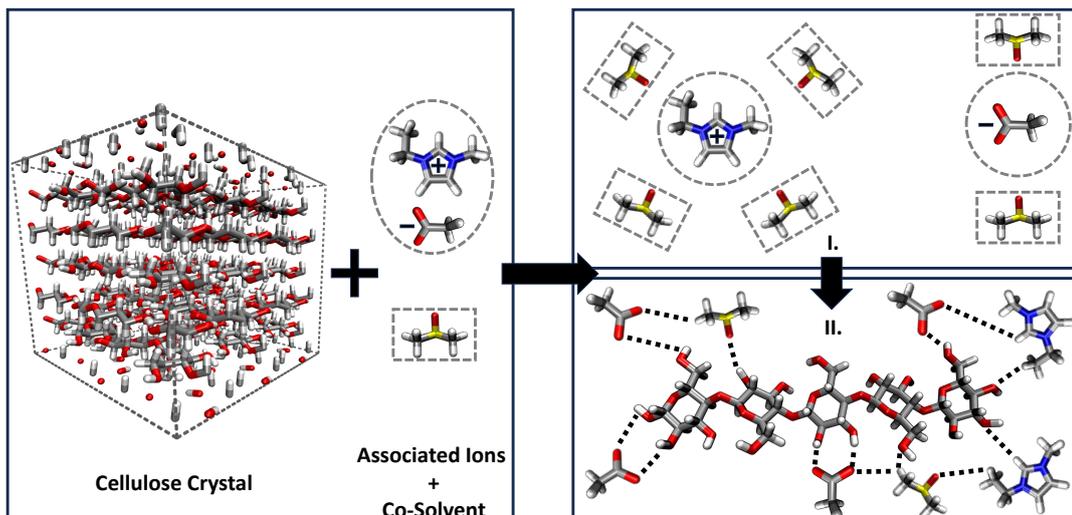
FIGURE 2.5: Illustration of the proposed reaction mechanism for the cellulose solvation in an ionic liquid/co-solvent system with two key steps: I.) The co-solvent (here DMSO) separates the ions of the IL (here [EMIm][OAc]) from each other. II.) The IL/co-solvent mixture interacts with cellulose (atom color code: yellow – *S*; red – *O*; blue – *N*; grey – *C*; white – *H*). The molecular structures and the unit cell were created with Avogadro (Version 1.2.0),[122] VMD[123] and Tachyon[124].

cellulose depends on how the cellulose/IL solution comes into contact with the regeneration solvent: If the cellulose/IL solution is rapidly mixed with an aqueous stream, cellulose precipitates as a powdery floc.[31] Alternatively, extrusion of the cellulose/IL solution into an anti-solvent can produce thin fibers and rods.[31, 126] Casting the cellulose/IL solution onto a glass plate results in the formation of thin cellulose films.[127] Furthermore, washing regenerated gel with liquid carbon dioxide before supercritical drying can yield nanofibrillar cellulose aerogels.[128]

Studies have shown that after the process of dissolution and regeneration, there is a change in the crystalline structure from cellulose I to cellulose II (see Fig. 2.2 for cellulose crystal structures).[107] This reduced crystallinity of the regenerated cellulose makes it much more susceptible to enzymatic hydrolysis, which greatly enhances the conversion of cellulose to glucose.[129]

After the cellulose solubilization process, the IL can be recycled, which is of great economic and environmental importance. Some methods are currently being investigated for IL recycling, such as the use of supercritical CO_2[130] as well as direct distillation[131] and microfiltration.[132]

TABLE 2.1: Solubility of cellulose in various imidazolium and triazolium based ionic liquids. The ILs considered in this thesis and related publications are marked in bold. The cellulose source used in the measurements and its degree of polymerization are given in parentheses in column 3.

IL Name	Structure	Cellulose Solubility / wt.%	Temp./ °C
[EMIm][OAc]		21.1 (MCC, *DP*=n.m.)[35]	80
[EMIm][OAc] + DMSO		28.0 (Avicel, *DP*=n.m.)[44]	25
[BMIm][OAc]		15.5 (MCC, *DP*=229)[133]	70
[BMIm][Cl]		18.0 (Avicel, *DP*=286)[32]	80
		12.0 (Avicel, *DP*=307)[32]	80
[BdMIm][Cl]		9.0 (Avicel, *DP*=377)[32]	80
[EMTr$_{123}$][OAc]		19.9 (MCC, *DP*=n.m.)[35]	80
[EMTr$_{124}$][OAc]		17.5 (MCC, *DP*=n.m.)[35]	80
[EMIm][OBz]		7.4 (MCC, *DP*=n.m.)[36]	80
[EMTr$_{123}$][OBz]		8.5 (MCC, *DP*=n.m.)[36]	80
[EMIm][OLa$_L$]		18.5 (MCC, *DP*=n.m.)[37]	80
[EMIm][OLa$_D$]		7.0 (MCC, *DP*=n.m.)[37]	80

[EMIm]: 1-ethyl-3-methylimidazolium; [BMIm]: 1-butyl-3-methylimidazolium; [BdMIm]: 1-butyl-2,3-dimethylimidazolium; [EMTr$_{123}$]: 1-ethyl-3-methyl-1,2,3-triazolium; [EMTr$_{124}$]: 4-ethyl-1-methyl-1,2,4-triazolium; [OAc]: acetate; [OBz]: benzoate; [OLa-L]: L-lactate; [OLa-D]: D-lactate; MCC: microcrystalline cellulose; Avicel: microcrystalline cellulose with a defined particle size; *DP*: degree of polymerization; n.m.: not mentioned.

2.2   Force Fields

To optimize the cellulose solvation process in ionic liquids, a comprehensive understanding of the intricate interplay between the solvent molecules and cellulose is essential. This can be achieved through molecular dynamics simulations. Despite well-known intermolecular interactions, a complex array of quantum mechanical forces operates between numerous electrons orbiting atoms. Because of the complex nature of the interactions, conducting quantum chemical calculations is too resource-intensive for enabling dynamic simulations over extended periods. This limitation restricts their ability to model macroscopic system properties and phenomena occurring over prolonged time scales. A more feasible approach is to use empirical potential energy functions in *force fields* describing the interactions of all particles in force field MD simulations. The force acting between the particles is calculated using the potential energy U of the system, which can be written as the sum of the bonded and nonbonded interactions:

$$U(\vec{r}) = \sum U_{\text{bonded}}(\vec{r}) + \sum U_{\text{nonbonded}}(\vec{r}). \tag{2.2}$$

Here, $\vec{r}_i(t)$ are the spatial coordinates of the particle. A detailed explanation of the empirical functions used to determine the interactions in force fields can be found in Ref. [68, 72] and [134], and is summarized below. The force $\vec{F}_i$ acting on particle i at time t is then calculated using

$$\vec{F}_i(t) = -\nabla_i U(\vec{r}_1(t), ..., \vec{r}_N(t)). \tag{2.3}$$

With knowledge of the forces acting on each particle, it becomes possible to calculate its change in position over time by solving Newton's equation of motion (see Section 2.4).

2.2.1   Bonded Interactions in Force Fields

To describe the bonded interactions in a force field, four energy contributions are considered: the potential energy of bonds, bond angles, dihedral angles, and the out-of-plane angle (as shown in Figure 2.6).

In a molecule consisting of two atoms A and B, a chemical bond results from the transfer of one or more electrons from A to B or from the presence of one or more electrons in shared orbitals. The latter case results in a lower energy than if the electron belonged to a single atom. This bond energy is particularly strong when it is mediated by two electrons. The requirement for this so-called *covalent bond* is that both atoms have available orbitals in the appropriate shell. The potential energy of a typical bond follows the pattern shown in Fig. 2.7. An effective approximation for this is the *Morse potential*. However, it is rarely utilized in force fields due to its high computational complexity. Furthermore, the incorporation of the Morse potential necessitates the introduction of additional parameters, which increases the complexity of the force field. Additionally, calculating the exponential function in the Morse potential is more computationally demanding than evaluating a polynomial function.[68, 72] As a result, *Hooke's Law* is often preferred for describing bonds: In a simplified representation, the covalent bond is treated as a spring that exerts a force between atoms A and B. The force required to stretch or compress the "spring"
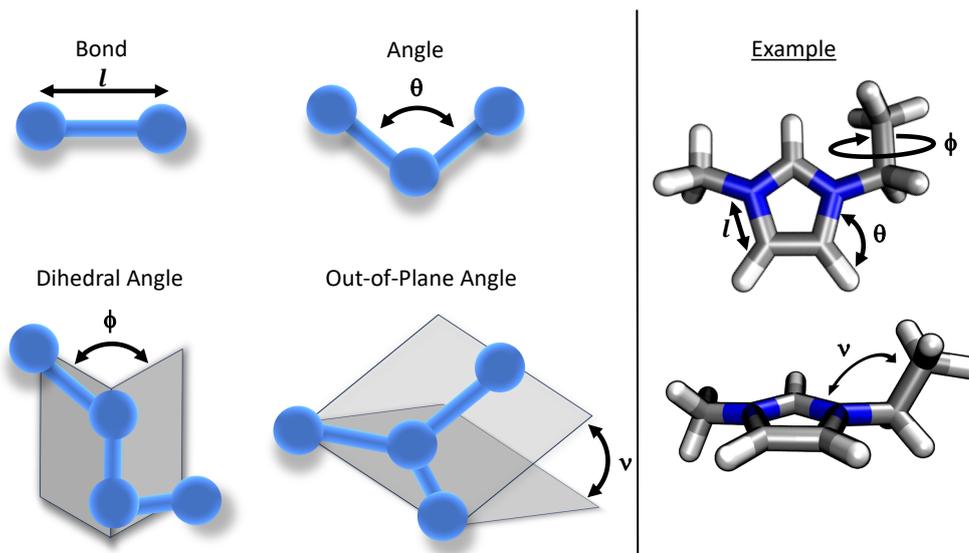
FIGURE 2.6: Illustration of the different types of bonded interactions in force fields. As an example, the 1-ethyl-3-methyl-imidazolium cation is shown.

is directly proportional to the displacement (extension or compression) and is influenced by the force constant k_l. Following this notion, the potential energy of a covalent bond is

$$U_{\text{Hooke}}(l) = \frac{k_1}{2}(l - l_0)^2, \tag{2.4}$$

where l_0 is the reference bond length.

By comparing the harmonic potential with the Morse potential (cf. Fig. 2.7), it is shown that Hooke's law, as an approximation for describing the potential of molecular bonds, provides valid results for distances corresponding to bonding in the ground state of molecules. The harmonic potential is also applied in BILFF.

To calculate the *bond angle* θ between two atoms, the harmonic potential is also commonly used and is utilized in BILFF:

$$U_{\text{angle}}(\theta) = \frac{k_\theta}{2}(\theta - \theta_0)^2, \tag{2.5}$$

where k_θ represents the angle force constant.

The *dihedral angle* ϕ (also called torsion angle) is another important aspect in force fields and is illustrated in Fig. 2.6. Dihedral angles can be crucial in determining the conformations of molecules. Since the torsion itself is periodic, the torsional potential energy must also exhibit periodicity. The potential energy of a torsion angle $U_{\text{dihedral}}(\phi)$ can be described by a Fourier series:

$$U_{\text{dihedral}}(\phi) = \sum_{n=1}^{4} \left[\frac{V_n}{2} [1 + cos(n\phi_n - \gamma_n] \right]. \tag{2.6}$$

The torsion potential is determined primarily by minimizing steric hindrance and electronic repulsion between atoms. The coefficient V_n provides a qualitative
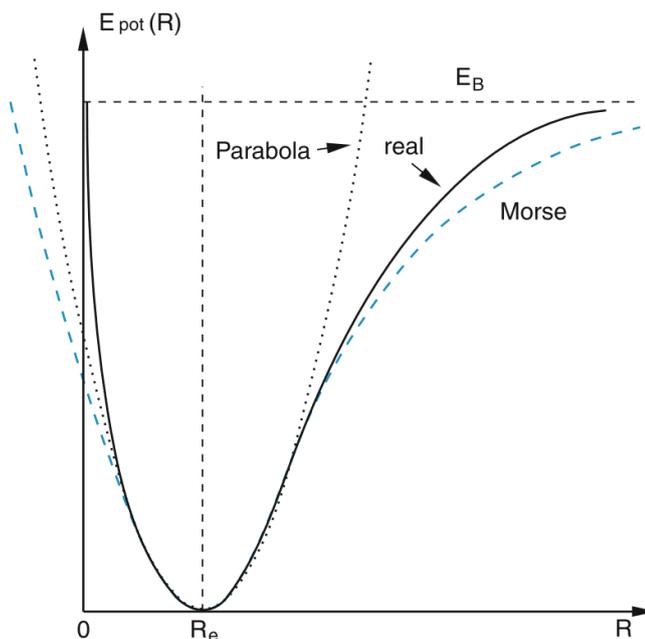
FIGURE 2.7: Comparison of the harmonic potential and Morse potential with the experimental (real) potential from Ref. [135]. The variable E_B represents the binding energy.

indication of the relative barriers to the rotation. The parameter n in the equation represents the multiplicity of the torsion potential, determined by the number of minimum points in the function as the bond rotates through an angle ϕ from 0 to 360°. The parameter γ_n, which is also called the phase factor, determines the position of the torsion angle at which the potential passes through its minimum value.

The number of the cosine terms depends on the force field and the torsion angles being described. For example, the MM2 force field[136–139] employs three terms. Directly interpreting these terms can be challenging, given the various factors influencing the torsion potential. An attempt to provide a physical interpretation was made for simple fluorinated hydrocarbons through the analysis of *ab initio* molecular dynamic simulations.[68, 72]

In *n*-butane (cf. Fig. 2.8), the one- and threefold term are discussed to represent the steric stresses in the molecule. Whereas the threefold fit shows a very good agreement with the full rotational profile, the second cosine term is negligible here.[68]

In the very common OPLS force field,[57–59] a fourth cosine term is sometimes added to describe torsional vibrations that are not adequately captured by the first three terms and is also used in BILFF. These torsional oscillations can occur due to complex molecular geometries or specific interactions between atoms.

For some molecules, an additional expression in the force field is necessary to accurately describe their structure. This term is known as the *out-of-plane bending term*. It is particularly important for molecules such as benzoate, imidazolium, and triazolium, where the ring protons energetically prefer to be in the plane of the ring as well as to maintain the planarity of the ring itself. There are various ways to implement the out-of-plane bending term in the force field. One approach, as also employed in BILFF, is to treat them as improper torsion angles and assign a torsion potential to them.
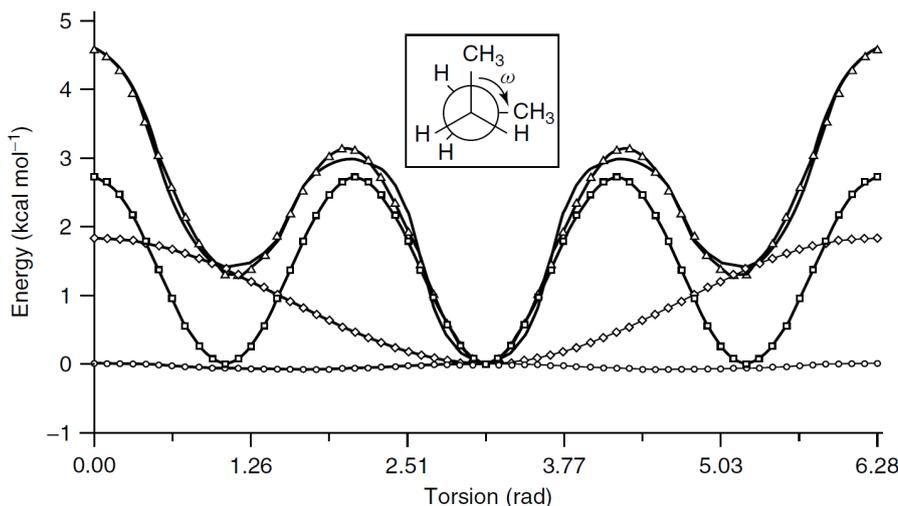
FIGURE 2.8: The torsional energy for a rotation about the C–C bond of *n*-butane (bold black curve). The figure was taken from Ref. [68]. The Fourier sum ($\triangle$) is in a good agreement with it. The symbolized curves show the individual terms: onefold ($\Diamond$), twofold ($\circ$) and threefold ($\Box$). (The inscribed variable ω represents the dihedral angle and corresponds to the parameter ϕ in eq. 2.6.)

2.2.2   Nonbonded Interactions in Force Fields

Nonbonded interactions play an essential role in accurately describing and understanding the properties of both cellulose and ionic liquids. They are the sum of pairwise electrostatic and van der Waals forces:

$$U_{\text{nonbonded}} = \sum_{pairs} \left[U_{ij}^{\text{vdW}} + U_{ij}^{\text{Coul}} \right]. \tag{2.7}$$

Van der Waals forces are a collective term for several types of interactions (cf. Section 2.1.2). The attractive and repulsive components of the van der Waals forces can be simplified and summarized in the *Lennard-Jones potential* (cf. Fig. 2.9). This empirical function is commonly used and also applied in BILFF. Here, the attractive part between two atoms *i* and *j* follows $-r^{-m}$ and the repulsive part follows r^{-n}:

$$U_{\text{LJ}}(r_{ij}) = a_{\text{LJ}}\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^n - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^m \right];$$
$$a_{\text{LJ}} = \frac{n}{n-m} \left(\frac{n}{m} \right)^{m/(n-m)} \tag{2.8}$$

In most cases, the exponents *n* and *m* are specified as $n = 12$ and $m = 6$, so that the prefactor a_{LJ} results as $a_{\text{LJ}} = 4$. While the distance dependence of r^{-6} in the attractive term represents the actual physical behavior of dispersion interactions, the choice of a twelfth power dependence for the repulsive term has no theoretical justification. Rather, it was chosen to reduce the computational cost compared to, for example, using an exponential function, and it can be calculated using the square of r^{-6}.[68, 140] In some force fields, alternative exponents such as 9 or 10 are also used for the repulsive part of the potential, resulting in a less steep energy curve. This can allow for a better adaptation to the specific characteristics of the system under
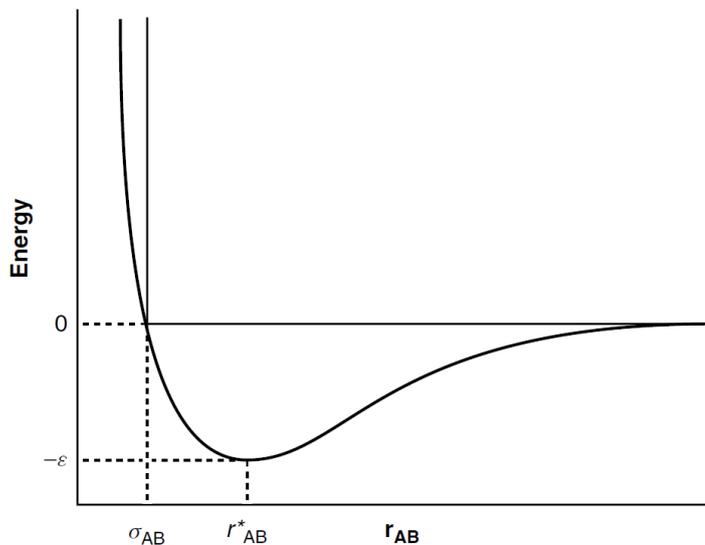
FIGURE 2.9: Lennard-Jones potential (curve) and hard-sphere potential (straight lines). In
the hard-sphere potential, atoms are defined as non-penetrable spheres with a
radius of r_{vdw}, which cannot overlap. The potential energy becomes infinitely
high for $r < 2\,r_{vdW}$.[68] (The inscribed parameters ϵ and σ_{AB} correspond to ϵ_{ij}
and σ_{ij} in equation 2.8.)

consideration, e.g. for systems in the glass state.[72, 141, 142] BILFF focuses on accu-
rately reproducing the hydrogen bonding of cellulose and ILs, utilizing the well-
established (12,6) potential commonly used for ILs.[57–63]

The Lennard-Jones parameters ϵ_{ij} and σ_{ij} define the depth of the potential well
and the particle-particle distance at which the potential energy is zero (cf. Fig. 2.9).
The calculation of the Lennard-Jones parameters between different atoms can be per-
formed by using *mixing rules*. There are various mixing rules, such as the Lorentz–
Berthelot mixing rule. The rule states that the parameter σ_{ij} for the interaction be-
tween species i and j (here: atoms) is the arithmetic mean of the values for the two
pure species, while ϵ is calculated as a geometric mean. The mixing rule is particu-
larly successful when applied to similar species, as it provides good agreement with
experimental findings.[72] To ensure the most accurate reproduction of the experi-
mental findings, the OPLS force field and BILFF uses the geometric description for
both parameters:

$$\epsilon_{ij} = \sqrt{\epsilon_i\epsilon_j}$$
$$\sigma_{ij} = \sqrt{\sigma_i\sigma_j}. \tag{2.9}$$

As a rule of thumb, it can be seen that the value for ϵ_i increases with the atomic
number and thus with the polarizability of the atom type. The parameter σ_i increases
with each period in the periodic table, but decreases from left to right across a pe-
riod with the increasing charge of the nucleus.[143] This trend is also noticeable in the
force field parameters of BILFF, where carbon, for instance, exhibits a larger σ_i value
compared to nitrogen and oxygen in the ILs.

Coulomb interactions result from the interaction of charged particles or multipoles.
The Coulomb potential between two partial charges q_i and q_j can be calculated by

$$U_{ij}^{\text{Coul}} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{2.10}$$

with the vacuum permittivity ϵ_0 and the distance between the particles r_{ij}.
In the simplest case, the partial charge is permanent and is assigned consistently to all atoms of the same type, regardless of conditions or environment. In this case, the atomic charges are fixed and serve as constant parameters in all simulations. An alternative approach is to consider the electronegativity of the atom in question, along with the electronegativities of the atoms to which it is chemically bonded. By doing so, the atomic electronegativity is introduced as a parameter within the calculation process, which may result in varying atomic charges while maintaining a high degree of flexibility. However, it is important to recognize that while charge polarization approaches offer improved accuracy in certain scenarios, they come at a computational cost. Calculating the electrostatic interaction energy within a so-called *polarizable force field* demands approximately an order of magnitude more computational resources than the same simulation using a static force field. This highlights the trade-off between accuracy and computational efficiency in molecular simulations.[68] BILFF therefore uses static charges optimized to reproduce experimental and calculated observations of the molecules.

Since both the van der Waals and the Coulomb interactions are pairwise interactions, they scale quadratically with the number of particles, which leads to an enormous computational effort as the size of the system increases. To address this, a common strategy is to use a finite spherical *cutoff* on the interactions by considering only the interactions within a certain distance and neglecting those beyond the cutoff. However, long-range interactions such as electrostatic interactions require special consideration. One method is the *Ewald summation*,[144] which splits the interactions into a short-ranged and a long-ranged part. In the short-range environment of the charged particles, the function is calculated in real space. In the long-range environment, the calculation takes place in reciprocal space using a Gaussian charge distribution. This reduces the scaling from N^2 to $N^{3/2}$.[145] Related to this method is the Particle-Particle-Particle-Mesh (PPPM) method,[146] which was applied in this work. For further insights into the Ewald summation, the PPPM algorithm, and their technical implementation, refer to Ref. [144] and [146].

2.2.3 Force Field Equation and Parameterization

As mentioned at the beginning of this section, force fields calculate the potential energy of the system using a summation of the bonded and nonbonded interactions (cf. eq. 2.2). The respective energy contributions of these two interaction types result in the general force field equation, which is also the potential energy function of BILFF:

$$
\begin{aligned}
U(r^N) = &\sum_{i \in \text{bonds}} \frac{k_{l,i}}{2}(l_i - l_{i,0})^2 + \sum_{i \in \text{angles}} \frac{k_{\theta,i}}{2}(\theta_i - \theta_{i,0})^2 \\
&+ \sum_{i \in \text{dihedrals}} \left[\frac{V_{1,i}}{2}[1 + \cos(\phi_i)] + \frac{V_{2,i}}{2}[1 - \cos(2\phi_i)] \right. \\
&\qquad\qquad \left. + \frac{V_{3,i}}{2}[1 + \cos(3\phi_i)] + \frac{V_{4,i}}{2}[1 - \cos(4\phi_i)] \right] \\
&+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right] + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}} \right] f_{ij},
\end{aligned}
\tag{2.11}
$$

where f_{ij} represents a scaling factor. (According to OPLS–AA,[57–59] the non-bonded Lennard-Jones and Coulomb interactions are neglected for 1–2 and 1–3 neighbors. 1–4 interactions are scaled down by a factor of f_{ij}=0.5, otherwise f_{ij}=1.0) All of the force field parameters ($k_{l,i}, l_{i,0}, k_{\theta,i}, \theta_{i,0}, V_{n,i}, \epsilon_{ij}, \sigma_{ij}, q_i, q_j$) in eq. 2.11 are determined individually. Therefore, a force field is typically tailored to predict a specific set of properties, and its parameters are adjusted accordingly.

However, a challenge arises because force field parameters are statistically non-independent, leading to the potential distortion of optimized values through coupling with other parameters. Thus, the choice of the bond force constant not only influences the reference bond length but also impacts the dipole moment of the bond simultaneously. In contrast, the angle term remains unaffected.

One approach to optimize the force field parameters to reproduce experimental data, for example, is the use of a *penalty function Z*, which quantifies the magnitude of the discrepancy between the calculated values $calc_{ij}$ and the experimental values exp_{ij}

$$
Z = \left[\sum_{i}^{\text{Observables}} \sum_{j}^{\text{Occurences}} \frac{(calc_{ij} - exp_{ij})^2}{\omega_i^2} \right]^{1/2}
\tag{2.12}
$$

with the weighting factor ω_i. This factor contains units to make Z dimensionless, and takes into account the different number of data points for the different observables, as well as the degree of tolerance that the penalty function will have for them. The aim is to minimize Z using statistical or quasi-statistical techniques.[68] However, an automated parameter adjustment procedure can be very difficult due to a ragged high-dimensional potential energy hypersurface of the system which possesses a large number of local minima.

Chemical knowledge can facilitate the selection of a single parameter set. For example, a parameter set that assigns a positive partial charge to oxygen atoms would seem chemically unreasonable. Similarly, many force constants and equilibrium values can be identified with unusually large or small values. Based on this concept, another approach follows a step-by-step optimization process. Initially, several parameters are optimized to reproduce a smaller set of desired properties. Subsequently, these parameters are fixed while other parameters are optimized over a larger set of properties. This process continues until all parameters have been adjusted.[68] This procedure was used to optimize the force field parameter of BILFF in order to obtain the most accurate reproduction of experimental as well as calculated values from quantum chemical simulations for the complex systems.

Finally, the computed potential energy serves as a foundation for examining dynamic processes within the studied system. This is realized by molecular dynamics simulations, which are explained in more detail in Section 2.4.

2.3   Electronic Structure Methods

While utilizing a force field to compute the potential energy of large molecular systems is an efficient way to reduce computational time, specific problems such as chemical reactions require the time-consuming study of the precise electronic configuration of the system. Electronic structure methods are used to calculate many-body properties and provide an accurate characterization of intermolecular interactions. These methods are considered a first-principles approach.

The description of a many-body particle system resides in its many-body wave function. In the stationary case, the *Schrödinger equation* is given by

$$\hat{H}\Psi(\vec{r}, \vec{R}) = E\Psi(\vec{r}, \vec{R}). \tag{2.13}$$

Here, E is the eigenvalue of the molecular wave function Ψ, while $\vec{R}$ and $\vec{r}$ represent the positions of the nuclei and electrons, respectively. The spin degrees of freedom have been omitted for simplicity.

The time-independent *Hamilton operator* $\hat{H}$ is defined as

$$\hat{H} = \hat{T}_\text{e} + \hat{T}_\text{N} + \hat{V}_\text{ee} + \hat{V}_\text{NN} + \hat{V}_\text{eN}. \tag{2.14}$$

The operator $\hat{T}_\text{e}$ accounts for the kinetic energies of the electrons, while $\hat{T}_\text{N}$ represents the kinetic energy associated with the nuclei. The interactions among the electrons are described by the operator $\hat{V}_\text{ee}$, and, similarly, the interactions between nuclei are captured by $\hat{V}_\text{NN}$. Lastly, the operator $\hat{V}_\text{eN}$ represents the interactions between the electrons e and the nuclei N. In this context, $\hat{H}$ can be expressed as

$$
\begin{aligned}
\hat{H} = & \sum_{i=1}^{e} -\frac{1}{2}\nabla_i^2 + \sum_{A=1}^{N} -\frac{1}{2M_A}\nabla_A^2 + \sum_{i=1}^{e}\sum_{j>i}^{e}\frac{1}{|\vec{r}_i - \vec{r}_j|} \\
& + \sum_{A=1}^{N}\sum_{B>A}^{N}\frac{Z_A Z_B}{|\vec{R}_A - \vec{R}_B|} + \sum_{i=1}^{e}\sum_{A=1}^{N} -\frac{Z_A}{|\vec{r}_i - \vec{R}_A|}.
\end{aligned}
\tag{2.15}
$$

The indices i and j denote electrons, while A and B indicate nuclei. M_A symbolizes the masses of the nuclei. Furthermore, Z_A and Z_B corresponds to the atomic number of nuclei.[c]

In the following, an effective simplification of the complex wave function, the Born–Oppenheimer approximation as well as the Kohn–Sham density functional theory are briefly explained. Additionally, a concise introduction to pseudopotentials is presented. For a more in-depth explanation of these topics, refer to Ref. [68, 72] and [147].

---

[c]The equation is formulated in atomic units, where the elementary charge e, the reduced Planck constant $\hbar$, the mass m and $4\pi\epsilon_0$ are normalized to 1 for simplification.

2.3.1 Born–Oppenheimer Approximation

The exact quantum mechanical treatment of a many-body system is only feasible for small system sizes due to the computational complexity. Therefore, approximations are necessary. The *Born–Oppenheimer approximation* is an effective simplification of the complex wave function and separates the total many-body wavefunction $\Psi(\vec{r}, \vec{R})$ into the multiplication of two different components: the electronic $\Psi_e(\vec{r}, \vec{R})$ and the nuclear wave function $\Psi_N(\vec{R})$:

$$\Psi(\vec{r}, \vec{R}) = \Psi_e(\vec{r}, \{\vec{R}\}) \, \Psi_N(\vec{R}) \qquad (2.16)$$

This approximation is based on the assumption that the nuclei of the molecular framework move much slower than the electrons, which is justified by the three orders of magnitude larger mass of the nucleus. This makes it possible to calculate the electronic part of the wave function, which depends only parametrically on the nuclear coordinates. By excluding the kinetic energy contributions of the nuclei, the resulting electronic Schrödinger equation with the corresponding eigenvalue E_e takes the form of

$$\hat{H}_e \Psi_e(\vec{r}, \{\vec{R}\}) = E_e(\{\vec{R}\}) \Psi_e(\vec{r}, \{\vec{R}\}). \qquad (2.17)$$

While the electron is quantum mechanically described using the time-independent Schrödinger equation, the atomic nucleus can be approximated as a classical particle. The interaction between nuclei and electrons $\hat{V}_{eN}$ can be replaced using a generalized external potential $v_{ext}(\vec{r})$. For the Hamiltonian of e electrons follows:

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{eN}$$
$$\hat{H}_e = \sum_i^e -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^e \sum_{j>i}^e \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{i=1}^e v_{ext}(\vec{r}_i). \qquad (2.18)$$

The total electronic energy E_{el} is then

$$E_{el} = \varepsilon_{tot}[\Psi] = \left\langle \Psi \left| \sum_{i=1}^e -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^e \sum_{j>i}^e \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{i=1}^e v_{ext}(\vec{r}_i) \right| \Psi \right\rangle. \qquad (2.19)$$

However, even in the case of extremely small systems, the electronic Schrödinger equation still does not have an analytical solution. Various techniques have been developed to find approximate solutions to equation 2.17: a wave-function-based approach (Hartree–Fock) and a density-based approach (density functional theory). In the following, the basic principle of the density functional theory is briefly explained.

2.3.2 Kohn–Sham Density Functional Theory

Even using the Born–Oppenheimer approximation, obtaining an exact quantum mechanical description of a system is a significant challenge. The computational requirements for such precision are prohibitively high, making such approaches infeasible for larger systems. One approach for handling the intricacies of the high-dimensional many-body problem is through the application of the *density functional theory* (DFT).

The *first Hohenberg–Kohn theorem*,[148] a fundamental principle within DFT, states that the ground state wavefunction of a system consisting of interacting electrons is determined by a unique functional $E_{v_0}[n]$ of the electron density $n(\vec{r})$. In other words, all ground state electronic quantities can be extracted from the electron density alone, without knowing the exact form of the wave function. Thus, the difficult task of finding the exact wave function can be transformed into the more computable task of determining the electron density within a specified external potential v_0:

$$E_{v_0}[n] = \langle \Psi[n], \hat{H}\, \Psi[n] \rangle. \tag{2.20}$$

The *second Hohenberg–Kohn theorem* states that the energy functional $E_{v_0}[n]$ reaches a minimum if and only if the electron density $n(\vec{r})$ corresponds to the actual ground-state electron density. This implies that the wavefunction minimizing the energy functional $E_{v_0}[n]$ is indeed the solution of the ground-state Schrödinger equation. For all v-representable densities other than the ground state, a distinct original wave function can be associated, which differs from the variational wave function. Consequently, the energy functional is not minimized for these densities. Unfortunately, a simple minimization of the energy functional could lead to a non-physical density. A less stringent condition needs to be applied.[149] The Kohn–Sham approach[150] provides a remedy for the problem of finding the correct electronic ground state density in practice.

The *Kohn–Sham approach* presents an artificial auxiliary system of non-interacting particles with the same ground state density as the interacting system. The individual one-electron Kohn–Sham orbitals ϕ_i are introduced and related to the ground state density n_o:

$$n_o(\vec{r}) = \sum_i |\phi_i(\vec{r})|^2 \tag{2.21}$$

The kinetic energy of the non-interacting system is given by

$$\tau[\phi_i] = \sum_i \langle \phi_i, -\frac{1}{2}\nabla_i^2 \phi_i \rangle. \tag{2.22}$$

The Kohn–Sham concept involves defining a new total energy functional by subtracting several terms computed from the wave function of a non-interacting electron system having the same density as that which would result from the exact solution of interacting particles. For the individual energy functionals, the following terms result: The functional of the interaction energy of the electron gas is given by

$$\varepsilon_{\mathrm{H}}[n] = \frac{1}{2} \iint \frac{n(\vec{r})n(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}\, d^3r\, d^3r' \tag{2.23}$$

and is referred to as the Hartree energy of the system. The functional of the interaction with the external potential is

$$\varepsilon_{\mathrm{ext}}[n] = \int v_{\mathrm{ext}}(\vec{r})n(\vec{r})\, d^3r. \tag{2.24}$$

The electron-electron interaction functional ε_{ee} is a decomposition of the Hartree energy functional and the exchange correlation functional $\varepsilon_{xc}[n]$:

$$\varepsilon_{ee}[n] = \varepsilon_H[n] + \varepsilon_{xc}[n]. \tag{2.25}$$

The exchange-correlation functional $\varepsilon_{xc}[n]$ encompasses the discrepancy between the energy functional of the interacting real system ε_{tot} (cf. eq. 2.19) and that of the non-interacting system ε_{KS}. Thereby ε_{KS} is the sum of the kinetic energy of the non-interacting system (eq. 2.22), the Hartree energy (eq. 2.23) and the functional of the external potential (eq. 2.24):

$$
\begin{aligned}
\varepsilon_{xc}[n] &= \varepsilon_{tot}[n] - \varepsilon_{KS}[n] \\
&= \varepsilon_{tot}[n] - \left(\tau[\phi_{[n]}] + \varepsilon_H[n] + \varepsilon_{ext}[n] \right).
\end{aligned}
\tag{2.26}
$$

This discrepancy occurs even if identical electronic densities are assumed. It includes all the many-body effects as well as the difference in kinetic energy between the fictive non-interacting system and the real one, and cannot be exactly addressed. Therefore, the exchange-correlation functional can only be approximated and remains the subject of research. There are several approaches. The simplest method is the *Local Density Approximation*, which models the functional's dependence solely on the local electron density. Another method is the *Generalized Gradient Approximation* (GGA). This method extends the previous approach of approximating the unknown functional by an integral over a function ϵ_{xc}^{GGA} that depends on the electron density and its gradient at a particular point in space:[147]

$$\varepsilon_{xc}^{GGA}[n] = \int n(\vec{r}) \epsilon_{xc}^{GGA} \left[n(\vec{r}); \nabla n(\vec{r}) \right] d^3r. \tag{2.27}$$

The BLYP functional,[151, 152] which is used in this work, belongs to this category and is very common because of its low computational time. In particular, hydrogen bonding, which plays a crucial role in the systems investigated, is well described.[153] In addition to these approaches, there are hybrid methods that combine the exact exchange from Hartree–Fock theory with other approximations. This combination aims to strike a balance between accuracy and computational efficiency.

Deriving the energy functionals (eq. 2.23, 2.24, 2.26) with respect to the electron density yields the corresponding potentials, so that for the Kohn–Sham Hamiltonian follows:

$$
\begin{aligned}
\hat{H}_{KS} &= -\frac{1}{2}\nabla^2 + v_{ee}(\vec{r}) + v_{eN}(\vec{r}) \\
\hat{H}_{KS} &= -\frac{1}{2}\nabla^2 + v_H(\vec{r}) + v_{xc}(\vec{r}) + v_{ext}(\vec{r}).
\end{aligned}
\tag{2.28}
$$

The potentials are summarized to the Kohn–Sham potential:

$$v_{eff}(\vec{r}) = v_H(\vec{r}) + v_{xc}(\vec{r}) + v_{ext}(\vec{r}). \tag{2.29}$$

In the end, the resulting Kohn–Sham equation is

$$\left[-\frac{1}{2}\nabla_i^2 + v_{eff}(\vec{r}) \right] \phi_i(\vec{r}) = \varepsilon_i \, \phi_i(\vec{r}) \tag{2.30}$$
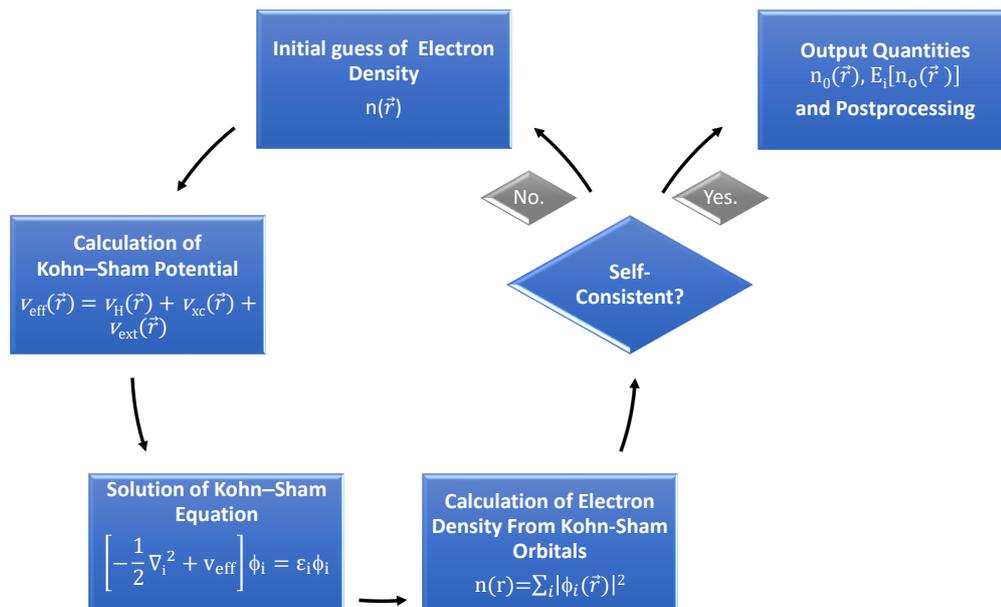
FIGURE 2.10: Self-consistent loop for the iterative solution of the Kohn–Sham equations.

To solve the Kohn–Sham equation, a self-consistent loop is applied (cf. Fig. 2.10). Starting with an initial guess for the electron density, the electronic Hamiltonian is constructed using this density, and the Kohn–Sham equation is solved to obtain a new set of electronic wavefunctions. The electronic density is then recalculated from these new wavefunctions. If the loop does not converge, it indicates that the estimated electron density is too far away from the true solution. In such cases, a more sophisticated initial guess may be needed, and the process is repeated.

For the Kohn–Sham orbitals ϕ_i, a set of different well-defined functions, called *basis functions*, can be used. A possible basis function are *plane waves*. Plane waves are favored as a basis set for characterizing periodic systems because of their correspondence to Fourier series and their intrinsic periodic behavior. Each orbital wave function can be represented as a linear combination of plane waves. However, a significant challenge arises when attempting to incorporate them directly into the Kohn–Sham equations. Due to the rapid oscillations of the orbitals as a result of Coulomb repulsion effects, as well as the Pauli exclusion principle, an accurate description of these oscillations would require a large number of basis sets, thus incurring high computational costs. For simplification, the focus is set on the valence electrons, which are responsible for chemical bonding. The inner electrons are considered inert and excluded from the calculations. A common approach is to replace the "true" potential in the core regions with a significantly weaker potential, called a *pseudopotential*. The basic concept is to ensure that the resulting pseudo wave function remains smooth near the nuclei. However, this approach implies that properties heavily dependent on the wave function near the nuclei cannot be straightforwardly derived from such calculations. Multiple functional forms are available for pseudopotentials and differ in terms of the number of plane waves required for their representation and the extent to which they can be extrapolated to diverse atomic environments. For practical utility, a pseudopotential designed for an atom should be unaffected by its chemical surroundings. This requirement of transferability is fulfilled by norm-conserving pseudopotentials. In this case, apart

from reproducing the all-electron valence wave function beyond a certain core radius, the charge of the pseudo-wave function within the core region must precisely match the corresponding charge in an all-electron calculation. Such a pseudopotential, the *Goedecker–Teter–Hutter (GTH) pseudopotential*,[154, 155] was used in this work.

The potential resulting from the solution of the electronic structure problem can finally be used to calculate time-dependent processes of the system via so-called *ab initio* molecular dynamics simulations. A detailed explanation of molecular dynamic simulations is given in the following section.

2.4   Molecular Dynamics Simulations

Molecular dynamics (MD) simulations are a widely used technique to study structural, dynamic and thermodynamic properties at the atomic level. They provide insights into material properties that may not be observable from experiments alone. With the current state of technology, MD simulations are computable over an observation period of a few picoseconds to microseconds. The molecules thereby obey Newton´s laws, which state:

1. "A body remains in a state of rest or uniform linear motion, unless it is forced to change its state by acting forces."

2. "The change of motion of a body is proportional to the action of the moving force and occurs according to the direction of the straight line along which the force acts."

3. "For every action, there is an equal and opposite reaction. If a body *A* exerts a force on another body *B* (actio), an equal but opposite force acts from body *B* on body *A* (reactio)."

Newton's second law can be expressed as $\vec{F} = m\vec{a}$, where $\vec{a}$ represents acceleration, and $\vec{a} = \ddot{\vec{r}}_i(t)$. In the context of a system of interacting particles, this relationship is further described by the equation

$$\vec{F}_i(t) = m_i\ddot{\vec{r}}_i(t) = -\nabla_i U(\vec{r}_1(t), ..., \vec{r}_N(t))$$
$$= -\nabla_i min_{\Psi_0}\{\langle\Psi_0, \hat{H}_e\Psi_0\rangle\}, \tag{2.31}$$

where $\vec{F}_i(t)$ is the force acting on particle i, m_i its mass, and $\vec{r}_i(t)$ the spatial coordinates of the particle at time t. The variable Ψ_0 stands for the ground state wave function. To determine the forces $\vec{F}_i(t)$, the gradient of the potential energy U can be used. The calculation of the potential energy can be done by different methods, such as electronic structure methods (see Section 2.3)[d] as well as by using force fields (see Section 2.2).

Once the forces acting on the particles in the system are calculated, *Newton's equation of motion* (eq. 2.31) can be solved to observe the temporal evolution of the molecular system. Various algorithms are available for the numerical solution of the equation. An example of such an algorithm is the Verlet algorithm,[156, 157] which is detailed in Ref. [68, 72, 143] and introduced in the subsequent section.

2.4.1   Numerical Solution of the Equations of Motion

Algorithms for solving Newton's equation of motion (eq. 2.31) rely on the calculation of the position $\vec{r}_i$ of the i-th particle based on its position at time t, the position at the previous time step $t - \delta t$ as well as the new position at $t + \delta t$. A method for the calculation is provided by the commonly used *Verlet Algorithm*,[156, 157] which is based on a Taylor expansion.

$$\vec{r}_i(t - \delta t) = \vec{r}_i(t) - \dot{\vec{r}}_i(t)\delta t + \frac{1}{2}\ddot{\vec{r}}_i(t)\delta t^2 - \frac{1}{6}\dddot{\vec{r}}_i(t)\delta t^3 + ... \tag{2.32}$$

$$\vec{r}_i(t + \delta t) = \vec{r}_i(t) + \dot{\vec{r}}_i(t)\delta t + \frac{1}{2}\ddot{\vec{r}}_i(t)\delta t^2 + \frac{1}{6}\dddot{\vec{r}}_i(t)\delta t^3 + ... \tag{2.33}$$

---

[d]The term AIMD has been used in this thesis to refer to Born–Oppenheimer molecular dynamics simulations.

Combining both equations and rearranging them to express $\vec{r}_i(t + \delta t)$ leads to the following approximation:

$$\vec{r}_i(t + \delta t) = 2\vec{r}_i(t) - \vec{r}(t - \delta t) + \ddot{\vec{r}}_i(t)\delta t^2. \tag{2.34}$$

The Verlet integrator achieves a higher level of accuracy compared to relying solely on the simple Taylor expansion, as it becomes evident from the elimination of the first and third order terms in the Taylor expansion.
The application of Newton's second law (eq. 2.31) to the expression for $\ddot{\vec{r}}_i(t)$ results in the well-known Verlet algorithm

$$\vec{r}_i(t + \delta t) = 2\vec{r}_i(t) - \vec{r}(t - \delta t) + \frac{\vec{F}_i(t)}{m_i}\delta t^2. \tag{2.35}$$

During the time step δt, the forces acting on the particle are assumed to be constant. However, a disadvantage of the algorithm is that the velocity of the particle is not directly accessible. The velocity at time t is not known until the subsequent time step $t + \delta t$ is calculated:

$$\vec{v}_i(t) = \frac{\vec{r}_i(t + \delta t) - \vec{r}_i(t - \delta t)}{2\delta t}. \tag{2.36}$$

A drawback is that positions and velocities are not synchronized, preventing the simultaneous calculation of the kinetic energy contribution to the total energy along with the positions. The *Velocity-Verlet algorithm*[158] remedies this deficiency. The procedure includes two steps: first the position at time $t + \delta t$ is calculated according to eq. 2.33. Then the new acting forces are calculated at the new position, resulting in the new acceleration of the particle. Applying Newton's second law, results in:

$$\vec{v}_i(t + \delta t) = \vec{v}_i(t) + \frac{\vec{F}_i(t) + \vec{F}_i(t + \delta t)}{2m_i}\delta t. \tag{2.37}$$

Thus, the kinetic energy at the time $t + \delta t$ is accessible.
To determine the initial velocity, one common approach is to use random numbers that follow the Maxwell–Boltzmann distribution.[72]

2.4.2 Periodic Boundary Condition

For the simulation of a realistic system, hundreds of molecules are necessary. *Periodic boundary conditions* can be applied to describe an (infinite) macrosystem using a microsystem called a unit cell, avoiding surface effects. Molecules at the center of the system would otherwise experience different forces than those at the edge of the simulation cell due to the reduced number of neighboring particles.
The molecules are placed in an (often) cubic box as an initial configuration. This unit cell is surrounded by simulation cells with identical configurations, referred to as images. When a molecule moves within the original unit cell, its replica in the neighboring boxes moves simultaneously. If a particle exits the simulation cell on one side, it re-enters on the opposite side.
In order to simulate large molecules such as the bio-polymer cellulose surrounded by solvent molecules, it is necessary to ensure that the original simulation cell is at least the size of the largest molecule plus twice the largest cutoff distance. Otherwise, interatomic interactions could be counted at least twice. In practice, one often goes far beyond this minimum requirement.[68]
Long-range forces such as Coulomb forces, which extend beyond the simulation cell,

pose a problem for this method, and as a result would lead to an infinite summation of interactions. This issue is addressed in more detail in Section 2.2.2.

The periodic boundary conditions should not be used for liquids near the liquid-gas transition, where macroscopic fluctuations can occur. However, they have little influence on the equilibrium thermodynamic properties of liquids that are far from phase transitions,[143] as is also the case in the simulations presented in the current work.

2.4.3 Temperature and Pressure Control

To model the realistic behavior of particles and derive macroscopic properties of the system, microscopic systems are often subjected to certain macroscopic conditions such as pressure, and temperature. Therefore, statistical *ensembles* are commonly employed in practical applications. An ensemble is a collection of microscopic states that share a common set of macroscopic properties.[159]

In the *microcanonical ensemble*, also known as the NVE ensemble, the number of particles N, the volume V, and the total energy of the system E are kept constant. Consequently, changes occur in the pressure p and temperature T. This ensemble is similar to an isolated system due to the non-exchange of energy with the environment. In many cases, however, simulations of different ensembles are required. For studying the behavior of a system in contact with a heat source, the *canonical ensemble*, also known as the NVT ensemble, can be used with a constant number of particles N, volume V, and temperature T. As a result, the energy of the system can fluctuate.

A constant pressure p and temperature T, as present in most experimental measurements, is achieved in the *isothermal-isobaric NpT ensemble*.

Among numerous approaches to temperature and pressure control, this section briefly reviews a selection of thermostats and barostats utilized in this thesis. For more detailed explanations, refer to Ref. [72, 159] and [145].

Under the consideration that the temperature is defined by the mean kinetic energy of all particles, a constant temperature can be achieved by scaling the velocities of all particles by a factor or by coupling the system to a heat bath that supplies or dissipates energy to the system. One method of scaling the velocity of particles was developed by Andersen in 1980.[160] At the beginning, a time interval is set during which the *Andersen thermostat* changes the velocity of a randomly selected particle in the system. When the interval ends, another particle is selected at random, and its velocity is recalculated according to a given Maxwell–Boltzmann distribution. This velocity change corresponds to a collision of the particle with an imaginary heat bath that maintains the system at a specified temperature. The system then proceeds along a surface of constant energy until the velocity of the selected particle undergoes another change. Changing the velocity of the particle causes the system to jump to a different energy surface. By repeating this process at consistent intervals, the system gains the ability to thoroughly explore all significant regions within its configuration space. This results in dynamic behavior without irreversible states. One can distinguish between minor collisions, wherein only one (or a few) particles are impacted, and massive collisions, which result in alterations to the velocities of all particles.

The Andersen thermostat is most suitable for systems with weak interactions between particles, as it assumes that the interaction between particles and the heat source is weak enough to be considered as "scattering interaction". In systems with

strong interactions, the thermostat can lead to inaccurate results.

A very common alternative to the Andersen thermostat is the *Berendsen thermostat*.[161] The basic concept is to connect the system to an external heat bath maintained at the desired temperature that exchanges thermal energy with the system. This gradual coupling ensures that the temperature is slowly adjusted to the desired value. The velocity of the particles is continuously adjusted according to scaling parameter λ:

$$\lambda^2 = 1 + \frac{\delta t}{\tau} \left(\frac{T_{\text{bath}}}{T(t)} - 1 \right) \tag{2.38}$$

where $T(t)$ represents the temperature of the system at time t.[72, 161] The system is thus coupled with a heat bath of temperature T_{bath} in each time step δt with the coupling parameter τ, which determines how quickly the system temperature is rescaled to the desired temperature. A higher value of τ leads to a slower adjustment of the system temperature.
Therefore, the Berendsen thermostat can effectively compensate for significant temperature deviations and equilibrate a system at a new temperature with minimal computational effort. However, it comes with certain drawbacks. The scaling of velocity can artificially extend temperature differences among system components, even if the overall system temperature is at the desired value. This may lead to an uneven energy distribution among components and motion modes. As a result, the canonical ensemble cannot be generated correctly. Furthermore, due to the lack of statistical equilibrium conditions, the Berendsen thermostat may struggle to accurately describe dynamic properties.

A different method of temperature control is achieved through the use of the *Langevin thermostat*, which employs the principles of stochastic dynamics for temperature regulation. The Langevin equation, utilized in the algorithms presented in Ref. [162] and [163], describes the motion of particles in a solvent and is composed of three main forces. One force is the particle-particle interaction $\vec{F}_{\text{intra},i}$. The second one is a force resulting from the motion of the molecule through the solvent, similar to a friction force. This force is proportional to the particle's velocity and the collision frequency γ_i. Here, γ_i is defined as $\gamma_i = \xi_i / m_i$, with ξ_i representing the friction coefficient and the mass m_i of the i-th particle. The third force $\vec{F}_{\text{random},i}$ arises from random fluctuations caused by interactions with solvent molecules. This force averages out to zero and follows a Gaussian distribution. The friction term and the random force are added to the intramolecular forces $\vec{F}_{\text{intra},i}$ resulting in the Langevin equation of motion

$$m_i \ddot{\vec{r}}_i = -\xi_i \dot{\vec{r}}_i + \vec{F}_{\text{intra},i} + \vec{F}_{\text{random},i}. \tag{2.39}$$

The random force can be considered as a temperature-like quantity that adds energy to the system, while the friction term removes energy. However, due to the calculation of the additional acting forces, the computational effort is substantially increased compared to the Berendsen thermostat. In addition, the dynamics of the system are significantly affected.

One of the most commonly used thermostats is the *Nosé–Hoover thermostat*. The Nosé–Hoover thermostat extends the Hamiltonian of the system to

$$\hat{H}_{\text{Nosé}} = U(\vec{r}_1(t), ..., \vec{r}_N(t)) + \sum_{i=1}^{N} \frac{p_i^2}{2m_i s^2} + \frac{p_s^2}{2Q} + g k_B T \ln s. \qquad (2.40)$$

To the potential $U(\vec{r}_1(t), ...\vec{r}_N(t))$, which describes the interactions between the particles of position $\vec{r}_i$, a term for the kinetic energy of the system and the kinetic energy of the thermostat is added. The kinetic energy is calculated by summing over the squared momenta p_i of all particles divided by their respective masses m_i. Simultaneously, a term is added for the thermostat with a fictitious degree of freedom s and its momentum p_s. The variable Q determines the time scale on which the thermostat operates. In the last term $g k_B T \ln s$, g corresponds to the total number of spatial degrees of freedom within the system, which is calculated as $3N + 1$ (one additional degree of freedom for s). The variables k_B and T represent the Boltzmann constant and the temperature, respectively.

The preservation of the Hamiltonian $\hat{H}_{\text{Nosé}}$ ensures the maintenance of a microcanonical ensemble (NVE). As shown in Ref. [159], the fluctuations within the microcanonical ensemble of the Nosé-Hamiltonian are equivalent to those observed in the canonical ensemble (NVT) of the original system. In this way, canonical sampling can be performed in a deterministic manner.

For the Nosé–Hoover dynamics, the following set of equations is defined:

$$\dot{\vec{r}}_i = \frac{\vec{p}_i}{m_i}$$

$$\dot{\vec{p}}_i = \vec{F}_i - \vec{p}_i \frac{p_\eta}{Q}$$

$$\dot{\eta} = \frac{p_\eta}{Q} \qquad (2.41)$$

$$\dot{p}_\eta = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{m_i} - d N k_B T.$$

An additional term has been introduced into the momentum equation of $\dot{\vec{p}}_i$ that can be associated with a friction term. This term is used to counteract temperature differences and excessive deviations from the desired temperature and can be either positive or negative. By controlling fluctuations in kinetic energy, it mimics the behavior of a heat bath.

However, the Nosé–Hoover thermostat encounters difficulties when the system has multiple conservation laws, since each conservation law introduces restrictions on the phase space. This may lead to a nonergodic behavior. One possible attempt for overcoming this problem is to couple the p_η variable with a Nosé–Hoover-like thermostat. The new variables $\tilde{\eta}$ and $p_{\tilde{\eta}}$ are introduced. By doing so, $p_{\tilde{\eta}}$ is also again coupled to a Nosé–Hoover thermostat. This can lead to an endless chain of thermostats, but must terminate at some point. This results in the so-called *Nosé–Hoover chain thermostat* of length N. Each variable in the chain is contributing to the temperature control. The coupling between the variables causes the thermal energy to be uniformly distributed in the system, resulting in a stable temperature control even under multiple conservation laws.

As discussed in this section, there are various approaches available for temperature control in molecular dynamics simulations. Each thermostat has its own set of advantages, disadvantages, and specific use cases. In the MD simulations performed

in this work, the previously discussed thermostats were employed: the Berendsen thermostat,[161] the Langevin thermostat,[162, 163] and the Nosé–Hoover thermostat, including its extended form known as the Nosé–Hoover chain thermostat.[164–166]

Nevertheless, not only a temperature control is of relevance. Certain thermodynamic properties require a constant pressure. Moreover, numerous experimental measurements are conducted under conditions of both constant pressure and temperature.
A constant pressure can be achieved by changing the volume of the simulation cell. The magnitude of the volume change is related to the compressibility:

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \tag{2.42}$$

The compressibility is directly included as a proportionality factor in the *Berendsen barostat*.[161] Similar to the Berendsen thermostat, the barostat[161] scales the volume of the simulation cell by a factor λ, which is equivalent to a scaling of the atomic coordinates by $\lambda^{1/3}$:

$$\lambda = 1 - \kappa\frac{\delta t}{\tau_p}\left(p - p_{\text{bath}}\right)$$
$$\vec{r}_i' = \lambda^{1/3}\vec{r}_i. \tag{2.43}$$

An alternative is the *Nosé–Hoover barostat*, which follows the same basic idea as its thermostat counterpart and represents an extension of it. Additional variables, p_S and s, are added to the Hamiltonian of the sytem, which are coupled to the volume and external pressure of the system.[165]

2.5   Trajectory Analysis

In molecular dynamics simulations, Newton's equations of motion (eq. 2.31) are solved to determine the velocity and position of each particle at each time step, stored in a *trajectory*. A trajectory can be viewed as a path through a $3N$-dimensional space, where N represents the number of particles. For systems with hundreds of particles, this poses a high-dimensional problem. Various software packages are available for analyzing such large datasets. In this thesis, various structural and dynamic analyses were performed using the software package TRAVIS[167, 168] developed by M. Brehm et al. These analyses are explained in more detail in the following section. The shown examples of the analyses are based on the results of this thesis.[169, 170] For more information about the analyses and TRAVIS see Ref. [159, 167, 168, 171].

2.5.1   Analysis of the Microstructure

Radial and Angular Distribution Function

Characterizing the equilibrium properties of real systems can be challenging due to the diverse behavior resulting from particle interactions. One method for the calculation of the microstructure is the use of a distribution function of averaged distances. For this purpose, a distance histogram between a reference particle and the observed particle in the system can be calculated. The calculation of the so-called *radial distribution function* (RDF) is employed, which gives the probability of an observed particle being found at a certain distance from a reference particle, relative to the uniform density of the observed particles throughout the simulation cell (cf. Fig. 2.11). The uniform density is the density that would be found if all observed particles were uniformly distributed in the simulation cell. Maxima of the RDF greater than 1 indicate that particles at these specific distances are more likely than average to be found at this distance, and are indicative of the average structural motifs present in the system. For large distances, the value of the RDF should tend toward 1, reflecting a more uniform distribution of particles in the system. RDFs are defined as

$$g_{ab}(r) := \frac{V}{\frac{4}{3}\pi r^2 N_a N_b T} \sum_{t=1}^{T} \sum_{i=1}^{N_a} \sum_{j=1}^{N_b} \delta\left(r - \left|\vec{r}_i(t) - \vec{r}_j(t)\right|\right) \qquad (2.44)$$

between particles of type a and b at a distance r. The variables N_a and N_b represent the number of atoms of the reference particle a and the observed particle b, respectively. The position vectors $\vec{r}_i(t)$ and $\vec{r}_j(t)$ correspond to the i-th and j-th atoms at time t. The term δ represents the Dirac delta distribution function, T the total number of trajectory steps and V the volume of the simulation box.
The prefactor serves as a correction factor to prevent the histograms from increasing in the order of r^2 for larger values of r, as more and more neighboring particles can be found with increasing distance.[171]
RDFs not only provide information about microstructures within a system but also play an important role in understanding the complex interplay of attractive and repulsive interactions between particles under investigation. These interactions are directly influenced by adjustments in the force field parameters, for example the partial charge q and the Lennard-Jones parameters σ and ϵ (cf. eq. 2.11). Any modifications to these parameters result in observable changes in the curve behavior of the

RDF. The partial charge of the atoms influences the attractive force between them, determining the number of reference particles present at a specific distance from the observed particle and, consequently, mainly affecting the height of the maxima in the RDF. The Lennard-Jones parameters influence the distance at which particles start to repel each other, thus mainly affecting the position of these maxima. As a result, RDFs serve as a valuable tool for optimizing force field parameters (e.g. in BILFF) with respect to microstructure, allowing for a fine-tuning of these parameters to the desired properties.

Similar to RDFs, the histogram of the angles between the reference and the observed particle can be calculated giving information about the overall orientation of the particles to each other (cf. Fig. 2.11). The definition of a (cone-corrected) *angular distribution function* (ADF) is:

$$f_{abc}(\alpha) := \frac{1}{\sin(\alpha) N_a N_b N_c T} \sum_{t=1}^{T} \sum_{i=1}^{N_a} \sum_{j=1}^{N_b} \sum_{k=1}^{N_c} \delta\Big(\alpha - \sphericalangle \big(\vec{r}_i(t), \vec{r}_j(t), \vec{r}_k(t) \big) \Big). \quad (2.45)$$

The cone correction factor $1/\sin(\alpha)$ is introduced to account for possible distortions of the uniform angular distribution (see Ref. [168, 171]).

In order to investigate correlations between different one-dimensional histograms, such as RDFs and ADFs, a *combined distribution function* (CDF) can be used. For each pair of molecules under consideration, the one-dimensional functions are combined into an n-tuple that contributes to an n-dimensional histogram, forming the CDF. The z-axis represents the probability of particle presence based on the selected criteria using a color scale. For example, when angle and distance histograms are combined, the contour plot reveals distinct preferences for certain angles at certain distances, providing valuable insight into the spatial arrangements and correlations of the system. In Fig. 2.11, a distance-angle CDF of the hydrogen bond of [EMIm]$^+$ and the oxygen atoms of [OAc]$^-$ in aqueous [EMIm][OAc] is shown as an example.

Spatial Distribution Function

Offering valuable insights into the overall structural arrangement, spatial distribution functions (SDFs) provide a comprehensive perspective on the relative orientations of molecules within a complex system. These functions illustrate the probability density of locating a specific atom at a particular point in Cartesian space relative to a fixed reference molecule. The location of surrounding molecules is represented by *isosurfaces* of the corresponding probability density. Isosurfaces are three-dimensional representations of a given data set where all points on the surface have the same value. They are expressed in units of particle density, such as nm^{-3}, which specify the particle density along the isosurface. In other words, the isosurface indicates how many particles are in the vicinity of this surface (cf. Fig. 2.12 as an example).
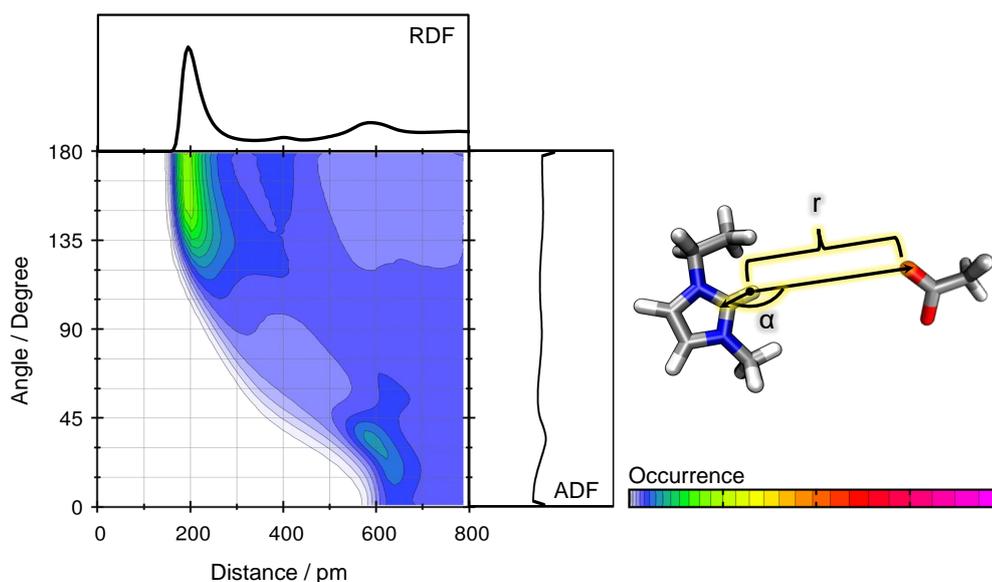
FIGURE 2.11: Distance–angle combined distribution function (CDF) between two particles (here a ring proton of [EMIm]$^+$ and the oxygen atoms of [OAc]$^-$ in aqueous [EMIm][OAc]) based on a radial distribution function (RDF) and an angular distribution function (ADF). Color code in arbitrary units. Concept based on Ref.[171].

2.5.2   Analysis of the Dynamics

Dimer Autocorrelation Function

To evaluate the lifetime of an aggregated species under a given geometric condition, a *dimer existence autocorrelation function* (DACF) can be calculated. A pair of particles is considered a dimer if certain geometric criteria, such as distances and angles between particles i and j, are fulfilled. They can be defined e.g. by the maxima occurring in a distance-angle CDF (cf. Section 2.5.1). The function $\beta_{ij}(t)$ is used to determine whether the criteria for molecules i and j are met at time t:

$$\beta_{ij}(t) = \begin{cases} 1, & \text{if criteria are fulfilled between } i \text{ and } j \text{ in step } t \\ 0, & \text{else.} \end{cases} \tag{2.46}$$

By considering all possible i and j pairs and averaging, the DACF is calculated:

$$DACF(\tau) = \frac{1}{N^2(T-\tau)} \sum_{t=0}^{T-\tau} \sum_{i=1}^{N} \sum_{j=1}^{N} \beta_{ij}(t)\beta_{ij}(t+\tau). \tag{2.47}$$

Here, N represents the number of particles, and T is the total number of simulation steps. At $\tau = 0$, the autocorrelation function exhibits an initial value of 1. As time progresses and the system undergoes dynamic changes, dimers may dissociate, and the criteria may no longer be fulfilled, causing the DACF to tend toward 0. However, there are two ways to define the DACF depending on the way the dimer is treated if the criteria for the dimer are violated in between the simulation:
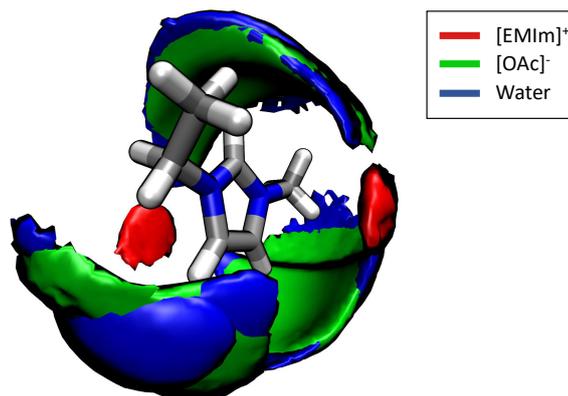
FIGURE 2.12: Spatial distribution function of the arrangement of molecules around an example reference molecule (here [EMIm]$^+$ in aqueous [EMIm][OAc]) with water (blue, $26\,\text{nm}^{-3}$), the oxygen atom of the anion (here [OAc]$^-$, green, $7\,\text{nm}^{-3}$) as well as the ring center of the cation (red, $7\,\text{nm}^{-3}$).

1. the continuous DACF, which does not allow the reformation of dimers (If the dimer is reformed, it would be considered a different dimer, even though it was formed from the same two molecules.)

2. the intermittent DACF, where reformations of dimers are allowed (as long as the duration of the dissociation does not exceed the time criterion).

By calculating the integral of the DACF, it is possible to determine the lifetime of a dimer (see eq. 2.48). However, a common problem with many autocorrelation functions is that they tend to approach zero too slowly. As a result, the function may still have a finite value at the maximum correlation depth, which is limited by the length of the trajectory. To accurately determine the integral up to $t = \infty$, the function must be extrapolated to larger correlation depths using a multi-exponential fit

$$\int_{\tau=0}^{\infty} c_1 \exp(-e_1\tau) + c_2 \exp(-e_2\tau) + ... \, d\tau = \frac{c_1}{e_1} + \frac{c_2}{e_2} + ... \qquad (2.48)$$

with the fitting variables c_i and e_i. TRAVIS[167, 168] performs four fits with one to four mono-exponential terms. The fit with the highest determination coefficient R^2 is ultimately used for evaluation purposes.

Mean Square Displacement

In order to investigate the mobility of the molecules in a system, the diffusion coefficient can be calculated. The diffusion coefficient is derived from the *mean square displacement* (MSD).The MSD is determined by observing the deviation of the position of a particle with respect to a reference position in a given time interval τ. Using the number T of simulation steps, the total number of particles N, and the position vector $\vec{r}_i(t)$ of particle i at time t, the MSD is defined as

$$MSD(\tau) = \frac{1}{N(T-\tau)} \sum_{t=0}^{T-\tau} \sum_{i=1}^{N} \left| \vec{r}_i(t) - \vec{r}_i(t+\tau) \right|^2$$

$$= \left\langle \left| \vec{r}_i(t_0) - \vec{r}_i(t_0+\tau) \right|^2 \right\rangle_{i,t}. \tag{2.49}$$

The self-diffusion coefficient D is obtained by applying the Einstein relation

$$\frac{\delta \langle r^2(t) \rangle}{\delta t} = 2dD \tag{2.50}$$

considering the MSD in an infinite time τ obtained with the dimensionality of the system d ($d = 3$ for an isotropic system):

$$2dD = \lim_{\tau \to \infty} \frac{\left\langle \left| r_i(t_0) - r_i(t_0+\tau) \right|^2 \right\rangle_{i,t}}{\tau}. \tag{2.51}$$

According to this relationship, the diffusion coefficient can be calculated from the slope of a linear regression of a correlation depth–MSD plot (cf. Fig. 2.13). TRAVIS[167, 168] automatically performs a linear fit to the second half of the obtained function to ensure a sufficient diffusion of the system.
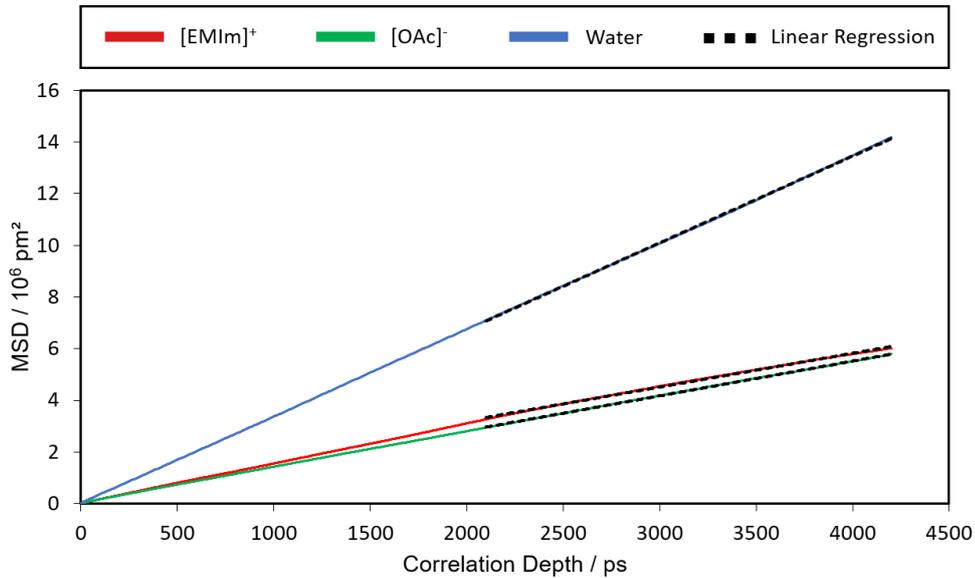
FIGURE 2.13: Mean square displacement (MSD) of the center of mass of three different molecules (here [EMIm]$^+$, [OAc]$^-$ and water). The self-diffusion coefficient of the molecules can be obtained by calculating the slope of a linear regression of the curves.

3 Results: Development and Application of a Force Field for Bio-Polymers in Ionic Liquids (BILFF)

3.1 Overview

To understand the complex process of dissolving cellulose in ionic liquids, it is crucial to consider the many interactions involved. A comprehensive understanding of the interplay between cellulose and the solvent molecules can be achieved by using MD simulations. The requirements for a realistic simulation include employing precise methods to describe molecules and interactions, appropriate system dimensions, and sufficient simulation time. However, the intricate nature of these systems and the long timescales involved in solubility processes often result in a lack of computational resources required for quantum chemical simulations. This thesis introduces a force field, called *BILFF* (*B*io-Polymers in *I*onic *L*iquids *F*orce *F*ield), designed for conducting MD simulations of large cellulose/IL systems over extended time scales with near quantum chemical accuracy. BILFF is specifically tailored to accurately reproduce the microstructure and dynamics of the hydrogen bonding network in ILs and is presented in four articles in this thesis. Force field parameters for cellulose, the ILs [EMIm][OAc], [EMTr][OAc], [EMIm][OBz] and [EMTr][OBz], as well as the cellulose co-solvent DMSO are presented. An overview of the molecules included in BILFF and which sections of this chapter contain information about them is given in Fig. 3.1. Sections 3.2 and 3.3 introduce the optimization process of the force field parameters of BILFF, while Section 3.4–3.6 extend BILFF to additional solvents and applications.

In **Section 3.2**, the parametrization of [EMIm][OAc] under both anhydrous and aqueous conditions at $350\,\mathrm{K}$ is presented. In the case of the aqueous system, the water content was chosen to be $\chi_{\mathrm{H_2O}} = 0.75$, where χ represents the molar fraction. At this concentration of water, cellulose is already precipitated again.[172] A comprehensive comparison between the RDFs and distance–angle CDFs characterizing the hydrogen bond network in the system was performed, along with a statistical analysis of bond lengths, angles, and dihedral angles by comparing them to the results of reference AIMD simulations. The partial charges and the Lennard-Jones radii have been adjusted iteratively by a trial-and-error method in order to achieve an accurate reproduction of the hydrogen bonds as well as of the experimental system densities. The bond lengths, angles, and torsional parameters were adjusted by changing the equilibrium values and force constants or by changing the height of the potential barriers. We found that reducing the total ion charge to ± 0.82 provided the best reproduction of the experimental diffusion coefficients. Reduced ion charges have been frequently used in the literature to account for charge screening effects.[173–176] By comparing the microstructure of the force field simulations using BILFF with results from ab initio MD simulations, we demonstrate the outstanding reproduction of the hydrogen bond network. These results are compared with those of the literature force field, highlighting the need for an optimized force field to accurately
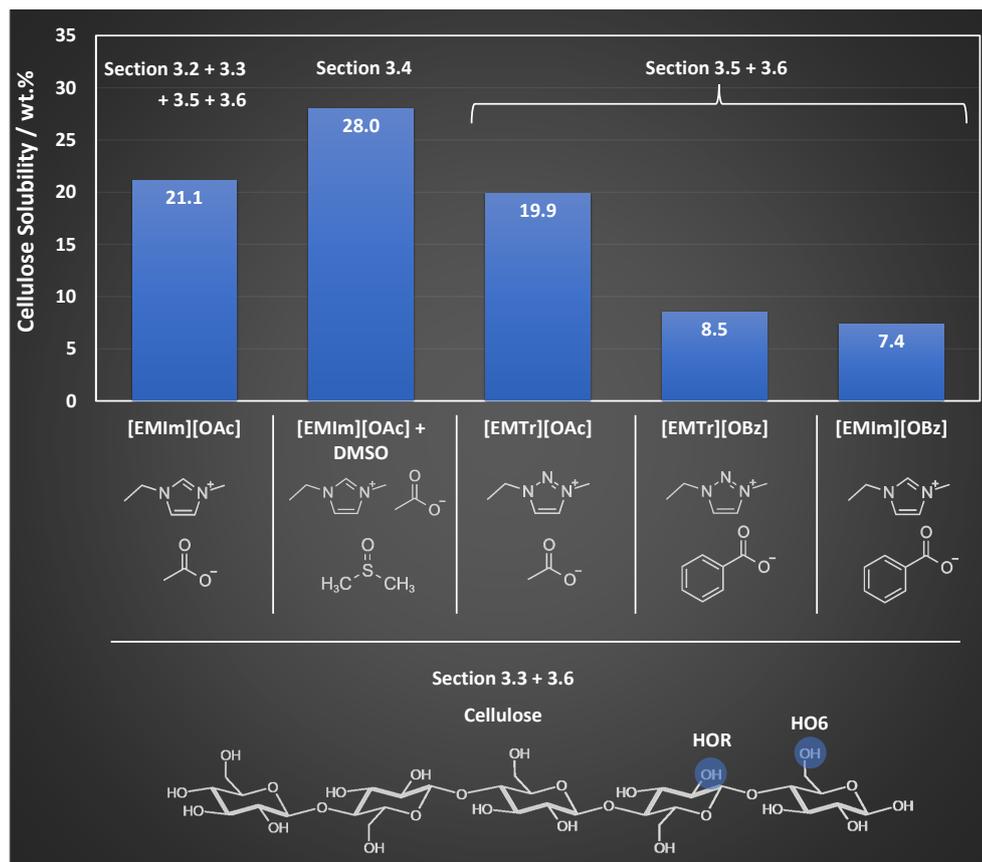
FIGURE 3.1: Overview of the molecules for which force field parameters are available in BILFF, along with experimental data on cellulose solubility at 80 °C (pure ILs) and 25 °C ([EMIm][OAc] + DMSO)[35, 36, 44] as well as references to the relevant sections in this thesis.

represent the microstructure of the system. Non-trivial interactions, such as π-π stacking of the cations, solvation shells, as well as dynamic parameters like hydrogen bond lifetimes and molecule diffusion, can also be accurately reproduced, as demonstrated by comparison to theoretical and experimental results. Finally, we examined the application of the force field over a wide temperature range from 350 K to 550 K and compared it with the AIMD results, observing a remarkable agreement. The utilization of BILFF in the force field MD simulations has facilitated a more thorough characterization of the hydrogen bond formation in anhydrous and aqueous [EMIm][OAc] and water: In the presence of water, the strength[a] and lifetime of the cation–anion hydrogen bond decrease as a result of the competition between [EMIm]$^+$ and water for the oxygen atom of acetate. The acetate oxygen atoms are partially occupied and shielded from [EMIm]$^+$ by the hydrogen atoms of the water molecules. Notably, this anion–water hydrogen bond represents the most stable one in the system. The outcomes of the calculated interaction energies are in line with these observations.

Section 3.3 continues the introductory work of BILFF and shows the adjustment of the force field parameters for cellulose in (aqueous) [EMIm][OAc] using a similar approach as in Section 3.2 with the aim of reproducing the microstructure from

---

[a]The strength of the hydrogen bonds is inferred based on the height of the first peak in the RDF, which indicates an increased probability of finding the observed atom near the reference atom.

AIMD simulations. The reference AIMD simulation for cellulose utilized a strand containing three β-D-glucose monomers. In contrast, the force field MD simulations employed a cellulose strand with five monomers (cf. Fig. 3.1) to benefit from the reduced computational cost of force field MD simulations and to achieve a more accurate representation of the statistical occurrence of the hydrogen bonding. To extend the physical simulation time and enhance sampling, 50 individual AIMD simulations were conducted, each initialized from different starting configurations, for cellulose in anhydrous [EMIm][OAc]. Additionally, another set of 50 AIMD simulations was performed for cellulose in aqueous [EMIm][OAc], all at a temperature of 350 K. The RDFs and CDFs were calculated from the total trajectory.

The hydroxyl groups directly on the carbon ring (on carbon atoms C2 and C3), called HOR, and on the methylene group (on carbon atom C6), called HO6, were distinguished in the analysis (cf. Fig. 3.1). This marks the first occurrence in the literature where such a distinction has been made in the calculation of RDFs. Moreover, BILFF stands out as the first force field in the literature to differentiate in the parameterization of the bridging acetal oxygen atom and the ring oxygen atom of cellulose. This enhancement allows for a more precise simulation outcome, as shown by comparing the microstructure with the results of simulations of the literature force field and the AIMD. It was demonstrated that not only the hydrogen bond network but also the experimental density, even at 550 K, and X-ray diffraction (XRD) data of the crystal structure of cellulose can be excellently reproduced. Furthermore, BILFF was employed for force field MD simulations to investigate the competition of hydrogen bonds among cellulose, [EMIm][OAc] and water. The investigation utilized various analyses, such as Sankey diagrams, a contact matrix, and SDFs. It was found that in anhydrous [EMIm][OAc], both hydroxyl groups, HOR and HO6, form similarly strong cellulose–anion hydrogen bonds. In the aqueous system, however, they are shielded from the anion by the water to a different degree. This difference is primarily due to the freely rotatable nature of the HO6 hydroxyl group, which makes it more accessible to water, along with a higher partial charge of its protons compared to the HOR group. The spatial arrangement of the IL around cellulose also reveals interactions not only between the anions and the cellulose strand but also involving the cations of [EMIm][OAc]. Weak hydrogen bonds are formed between the ring protons of [EMIm]$^+$ and the oxygen atoms of the hydroxyl group. However, the side chains of [EMIm]$^+$ do not form cellulose–cation hydrogen bonds. Furthermore, intramolecular hydrogen bonding occurs in cellulose, confirming findings in the literature,[36, 50–53] primarily facilitated by HO6.

Section 3.4 takes up the force field optimization methodology of BILFF to extend the force field to the cellulose co-solvent DMSO. The force field parameters thereby provide accurate results not only for pure and aqueous DMSO, but also for DMSO in anhydrous and aqueous [EMIm][OAc] at 350 K. The applicability was also investigated for higher temperatures of 450 K. The results of the force field MD simulations with BILFF were validated with experimental data on compressibility, thermal expansion coefficient, diffusion coefficient and density. The force field MD simulations with BILFF of DMSO in [EMIm][OAc] show that hydrogen bonds are mainly formed between the ions, but DMSO competes strongly as a hydrogen bond acceptor, confirming results from the literature.[53] With respect to the stability of the hydrogen bonds, the presence of DMSO reduced the lifetime of the cation–anion hydrogen bonding. This is consistent with the common theory in the literature that DMSO shields the ions from each other and thus increases the solubility of cellulose due to the higher number of free anions.[44–46] However, the opposite effect occurs in

DMSO/[EMIm][OAc]/water. The presence of DMSO leads to a significant increase in the lifetime of the hydrogen bonds among the ions and with water, by more than twice as compared to DMSO-absent IL/water blends. The impact of DMSO, coupled with the dual role of water as both hydrogen bond donor and acceptor in the system, not only affects the stability and strength of hydrogen bonds but also influences system density and molecular mobility.

In **Section 3.5**, the parameterization of BILFF was extended to include the ions [EMTr]$^+$ and [OBz]$^-$ for conducting precise force field MD simulations of the ILs [EMTr][OAc], [EMIm][OBz] and [EMTr][OBz] under anhydrous and aqueous conditions. This is the first time that optimized force field parameters for these ILs, specifically tailored to investigate the hydrogen bond network and solubility of cellulose in these systems, have been published in the literature.
By utilizing the already optimized parameter set for [EMIm]$^+$ and [OAc]$^-$, the force field parameters of [EMTr]$^+$ in [EMTr][OAc] and [OBz]$^-$ in [EMIm][OBz] were adjusted at a temperature of 350 K. The transferability of these force field parameters in the combination of both ions in [EMTr][OBz] was validated by comparison with AIMD simulations at temperatures ranging from 350 to 550 K. The transferability to the other IL [EMTr][OBz] confirms the accuracy of the underlying force field parameters of the ions.
The MD simulation results utilizing BILFF were employed to thoroughly investigate the hydrogen bonding networks within the examined IL systems. It was found that the cation–anion hydrogen bonding of the benzoate anion is slightly stronger than that of acetate, although the oxygen atoms of benzoate have a lower partial charge ($-0.55\,e$ vs. $-0.66\,e$). However, this trend reverses in the presence of water. One possible factor for this increased hydrogen bonding is the restricted rotational/translational freedom due to the T-shaped arrangement of the benzoate anions, resulting in a more stable spatial arrangement and longer hydrogen bond lifetime. With respect to the cations, the cation–anion and cation–water hydrogen bonding was found to be stronger for triazolium compared to imidazolium. Since the only molecular distinction is the CH/N substitution, the increased strength is most likely due to the higher partial charges of the CH ring protons ($+0.19\,e$ vs. $+0.15\,e$). This difference in the hydrogen bonding strength of the cations and anions plays an ambivalent role with respect to the solubility of cellulose in the various ILs. The hydrogen bonding ability, mainly of the anions, is crucial to dissolve cellulose. However, the simultaneously increased cation–anion hydrogen bond strength is counteracting this process, since the anions are not freely available to interact with cellulose. Comparing simulation results with the solubility of cellulose reveals an apparent discrepancy: While the solubility of cellulose increases in the following order [EMIm][OBz] < [EMTr][OBz] < [EMTr][OAc] < [EMIm][OAc],[35, 36] the strength of the cation–anion hydrogen bonds (in the anhydrous IL) does not behave simultaneously: [EMTr][OBz] > [EMTr][OAc] > [EMIm][OBz] > [EMIm][OAc]. The best solvent, [EMIm][OAc], exhibits the weakest interaction between cation and anion, while the IL with the strongest ion-pair interaction, [EMTr][OBz], is the second poorest solvent for cellulose. This illustrates the complexity of the underlying solubility mechanism. Further investigations are necessary to comprehensively understand the interactions between ILs and cellulose.

This question is further addressed in **Section 3.6** by analyzing the hydrogen bonding between cellulose and the ILs. BILFF has been used to perform force field

MD simulations of cellulose in these four ILs with and without water, allowing simulations with a high number of particles and extended simulation times. The results have not yet been published. However, they provide an insight into some of the possible applications of BILFF:

Section 3.6.1 investigates the impact of the cations [EMIm]$^+$ and [EMTr]$^+$ as well as the influence of a varying water amount (0 to 20 wt.% water) on the system cellulose/IL/water at 350 K. To focus on the influence of the cation, [OAc]$^-$ was chosen as fixed anion. It was confirmed that [EMTr]$^+$ forms slightly stronger cation–anion hydrogen bonds than [EMIm]$^+$. The anion therefore interacts less with the cellulose strand in anhydrous [EMTr][OAc] than in [EMIm][OAc], which is consistent with the slightly reduced solubility of cellulose in [EMTr][OAc].[35] However, in the presence of water, this effect is compromised by strong competitive hydrogen bonding with water. At the same time, the interaction between the anion and cellulose is influenced by water. The protons of the freely rotating hydroxyl group on the methylene group (HO6) are less affected by the presence of water than the hydroxyl protons directly on the monomer (HOR) and form stronger anion–cellulose hydrogen bonds.

In **Section 3.6.2** the results of force field MD simulations using BILFF of cellulose in the different ILs [EMTr][OAc], [EMIm][OAc], [EMIm][OBz], and [EMTr][OBz] are presented both in the presence and absence of water. In the anhydrous system, the anion in [EMIm][OAc] and [EMTr][OAc] tended to interact weaker with cellulose than in [EMIm][OBz] and [EMTr][OBz]. Given the common assumption that cellulose dissolves in an IL primarily due to pronounced hydrogen bonding with the anion, this result is surprising since [EMIm][OAc] is the most effective of the investigated solvents. This apparent discrepancy has not yet been clearly clarified and serves as the starting point for further investigations using force field MD simulations with BILFF. It appears that the solubility of cellulose depends not only on the strength of the hydrogen bonds, but possibly also on other factors such as the aromaticity of the ions and the possible formation of an anion–cellulose chelate complex. When water is introduced, the strength of the hydrogen bond network between the anion and cellulose is significantly reduced, depending on the specific IL. However, the spatial arrangement of the hydrogen bonds between the cellulose and the anion is nearly unchanged and independent of the anion.

3.2 Article I: A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 1: [EMIm][OAc]/Water Mixtures

<u>E. Roos</u>, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 1: [EMIm][OAc]/Water Mixtures", *Phys. Chem. Chem. Phys.*, 23, 1242–1253, 2021

Author contributions: In this article, I performed all calculations, was responsible for data processing, data analysis, visualization of the results, and the writing of the original article draft. M. Brehm was responsible for supervising the work and the conceptualization of the project. Both authors contributed to subsequent review iterations of the article.

The following article [E. Roos, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 1: [EMIm][OAc]/Water Mixtures", *Phys. Chem. Chem. Phys.*, 23, 1242–1253, 2021] is reproduced from Ref. [169] with permission from the PCCP Owner Societies. No changes were made.

ROYAL SOCIETY
OF CHEMISTRY

Check for updates

A force field for bio-polymers in ionic liquids (BILFF) – part 1: [EMIm][OAc]/water mixtures†

Eliane Roos [ID] and Martin Brehm [ID]*

We present BILFF, a novel force field for bio-polymers in ionic liquids. In the first part of our study, we introduce optimized force field parameters for mixtures of the ionic liquid (IL) 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) with water. This imidazolium-based IL is of particular practical importance as it can dissolve significant amounts of cellulose even at room temperature. An understanding of this dissolution process *via* molecular dynamics simulations requires a quantitative description of the microscopic structure and the strong hydrogen bonds with a method able of simulating at least several dozen nanoseconds, which is the main aim of our novel force field. To reach this goal, we optimize the force field parameters to reproduce radial, spatial, and combined distribution functions, hydrogen bond lifetimes, diffusion coefficients, and several other quantities from reference *ab initio* molecular dynamics (AIMD) simulations. Non-trivial effects such as dispersion interactions between the side chains and π–π stacking of the cations are reproduced very well. We further validate the force field by comparison to experimental data such as thermal expansion coefficients, bulk modulus, and density at different temperatures, which yields good agreement and correct trends. No other force field with optimized parameters for mixtures of [EMIm][OAc] and water has been presented in the literature yet. Optimized force field parameters for cellulose and other ILs will be published in upcoming articles.

1 Introduction

Ionic liquids (ILs) are an extremely interesting class of substances with some remarkable properties: since the synthesis of ethylammonium nitrate in 1914 as the first ionic liquid at room temperature,[1] the interest in this substance class for industrial and research applications has grown steadily. Ionic liquids are commonly defined as salts with a melting point below 100 °C and have a very low vapor pressure. Due to their low volatility, non-flammability, outstanding solvation potential and thermal and chemical stability,[2,3] they find application as solvents for synthesis and catalysis,[4–9] as lubricant additives,[10] in electrochemical sensors,[11] batteries,[12–15] solar cells,[16–18] for telescope mirrors,[19] and even to transport medicines through the skin.[20] Due to the great variability of cations and anions, the number of possible ionic liquids is estimated to be between 10^{12} and 10^{18},[21,22] which suggests many other previously unknown applications.

Ionic liquids are divided into several classes, including alkylammonium-, phosphonium-, *N*-alkylpyridinium-, and dialkylimidazolium-based ionic liquids. The latter moved into the focus of research because of the stability against oxidation and reduction of the cation and the relatively low viscosity.[2] One of these ILs is 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]). The chemical structure is shown in the lower panel of Fig. 1.

An important and thoroughly investigated feature of imidazolium-based ILs is the hydrogen bond donor ability of the cation.[23–26] [EMIm][OAc] forms particularly strong hydrogen bonds[27–29] and is able to dissolve some poorly soluble substances such as cellulose even at room temperature and without derivatization,[30–32] which has numerous applications in various industrial sectors, *e.g.* the paper and packaging industry, for pharmaceuticals, cosmetics, in the electrical industry, and in 3D printing technology.[33,34] Due to these numerous applications and the problem that cellulose is not soluble in water or other conventional solvents, there is a great general interest in understanding why and how [EMIm][OAc] is able to dissolve cellulose and which role intermolecular hydrogen bonds play in this process. Therefore, various investigations of the dissolution of cellulose in [EMIm][OAc] as a good solvent have already been performed, using different methods such as NMR spectroscopy,[35] small angle X-ray scattering,[36] and molecular dynamics simulations.[37–42]

The optimization of the solvation process requires a fundamental understanding of the solvation mechanism on a microscopic scale. This requires a consideration on the molecular level.

Institut für Chemie – Theoretische Chemie, Martin-Luther-Universität Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany. E-mail: Martin_Brehm@gmx.de
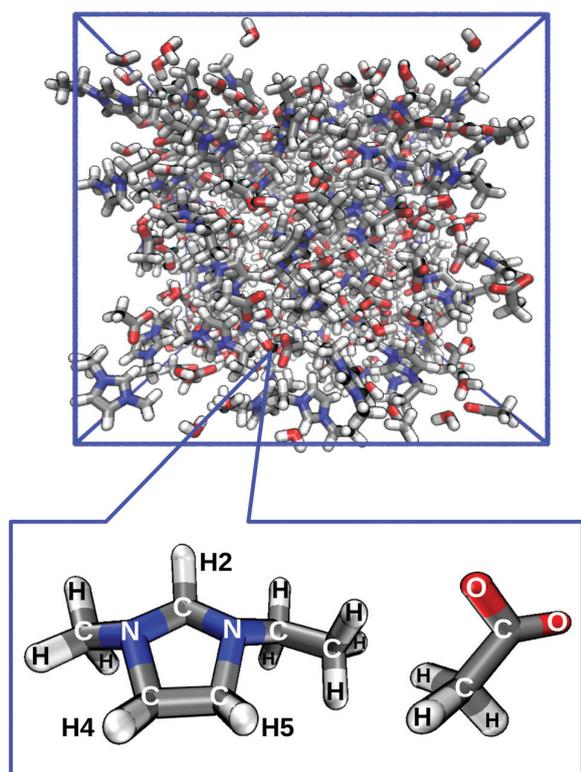
Fig. 1 Snapshot of the aqueous [EMIm][OAc] simulation cell (upper panel) and atom nomenclature used in this article (lower panel).

A very accurate description of the intermolecular interactions can be achieved by using *ab initio* molecular dynamics (AIMD) simulations; however, the computational effort is very high, and the time scales are too short to observe processes relevant for dissolution. A more suitable approach is the application of force field molecular dynamics (FFMD), which is far less time consuming and can simulate much larger time intervals. Many FFMD simulations of ionic liquids have already been published, and many promising force fields for different kinds of ILs have been presented in the literature within the last two decades.[43–56]

First results on hydrogen bonds between cellulose and [EMIm][OAc] from force field MD simulations are already known,[30,57–59] but by using a force field without specially adapted Lennard-Jones radii, bond lengths, bond angles, and torsion parameters for this system. Existing force fields for ionic liquids on the basis of AMBER[45,49] and OPLS-AA[46,47,50,54,56] have often been developed without a special emphasis on hydrogen bonding and local structure, so that the description of ILs with very strong hydrogen bonding (such as [EMIm][OAc]) is not optimal. There already exist a few force fields for ILs with a focus on hydrogen bonding which have been derived from AIMD simulations, such as the OPLS-VSIL force field[60] which utilizes virtual sites. However, those do not contain parameters optimized for [EMIm][OAc].

To enable a more precise modeling of the solvation of bio-polymers such as cellulose in ILs, we started to develop a force

field which is optimized for an accurate description of the hydrogen bonding between bio-polymers and ILs, with a special emphasis on microstructure and dynamics. We call this force field the "Bio-Polymers in Ionic Liquids Force Field" (BILFF). In this first article, we present optimized force field parameters for [EMIm][OAc]/water mixtures. Compatible force field parameters for cellulose and other ionic liquids (and possibly also other bio-polymers) will be presented in upcoming articles. Our force field is based on the functional form of OPLS-AA[61–63] and can therefore be used in a variety of simulation program packages. In aqueous systems, water molecules are described *via* the widely used TIP4P-EW model without modifications.[64]

By iterative adaptation to reference AIMD (*ab initio* molecular dynamics) simulations, the force field parameters (partial charges, Lennard-Jones parameters, bond lengths, angles, and torsion parameters) are optimized. The radial distribution functions (RDFs) and combined distribution functions (CDFs) of the hydrogen bonds as well as the spatial distribution functions (SDFs) and the diffusion coefficients are compared with the AIMD. The force field adjustment was performed on pure [EMIm][OAc] and [EMIm][OAc] in water in a molar ratio of 1 : 3 at a temperature of 350 K, which is a commonly used temperature for experimental studies of the solubility of cellulose in ionic liquids.[31,32,65] Our force field is further validated by comparing the system density and diffusion coefficients to experimental data at different temperatures. We will show that even at higher temperatures, the corresponding AIMD simulations can be reproduced well by a force field MD using BILFF.

2 Computational details

To develop the force field BILFF, AIMD simulations of pure and aqueous [EMIm][OAc] were performed in a first step. To do so, trajectories of pure [EMIm][OAc] as well as a [EMIm][OAc]/water mixture, which were investigated in literature before,[27,66,67] were used as pre-equilibrated starting configurations. The simulations were performed with the program package CP2k[68–70] using the Quickstep method[71] and orbital transformation (OT).[72] The composition and cell size of the simulated systems can be found in Table 1. The electronic structure was described *via* density functional theory (DFT),[73,74] using the BLYP functional[75,76] with the dispersion correction D3(BJ) from Grimme[77,78] using revised damping parameters from Smith *et al.*[79]

Table 1 Overview of the simulation parameters of the *ab initio* molecular dynamics (AIMD) and force field molecular dynamics (FFMD) simulations of pure and aqueous [EMIm][OAc]

Method	System	Composition	Sim. time/ps	Box size/ pm	Density/ g cm^{-3}
AIMD	Pure	36 ion pairs	235	2121.24	1.066
	Aqueous	27 ion pairs 81 water	167	2158.40	0.999
FFMD	Pure	128 ion pairs	10 000	3227.58	1.078
	Aqueous	81 ion pairs 243 water	10 000	3112.85	1.070

A plane-wave energy cutoff of 350 Ry was applied. Basis sets of the kind DZVP-MOLOPT-SR-GTH[80] and GTH pseudopotentials[81,82] were used. The physical simulation time of the AIMD was 235 ps for the system of pure [EMIm][OAc] and 167 ps for the system of aqueous [EMIm][OAc] with an integrator time step of 0.5 fs. The temperature has been set to 350 K by a Nosé–Hoover chain thermostat[83–85] with a coupling constant of 100 fs (*i.e.*, *NVT* ensemble).

In an iterative process, the force field parameters were adapted to the AIMD. For this purpose, force field molecular dynamics simulations with the LAMMPS program package[86] were performed as follows: 128 ion pairs of [EMIm][OAc] in a simulation box with an edge length of $a = b = c = 3227.58$ pm for the pure IL system and 81 ion pairs of [EMIm][OAc] with 243 water molecules in a simulation box with an edge length of $a = b = c = 3112.85$ pm for the aqueous IL system were distributed using Packmol[87] (see Table 1) and pre-equilibrated. A first equilibration was performed for 2.5 ps in an *NVE* ensemble at a temperature of 500 K using a Berendsen thermostat[88] with a time constant of 1.0 fs. Afterwards, the temperature was ramped down to 350 K during a time interval of 50 ps. Later, a simulation of 100 ps in an *NpT* ensemble using a Nosé–Hoover thermostat[83–85] with a coupling constant of 100 fs and a Nosé–Hoover barostat with a coupling constant of 2000 fs was conducted. The shock waves resulting from the size change of the simulation box were damped in a subsequent run using a Langevin thermostat[89,90] with a coupling constant of 100 fs for a duration of 50 ps. A further simulation step without the Langevin thermostat for a duration of 2.5 ns followed. In an another *NpT* simulation lasting 7.5 ns using the Nosé–Hoover thermostat with a coupling constant of 100 fs, the average value of the volume of the simulation box was computed and the resulting density and box size were noted. Subsequently, the volume of the simulation cell was changed to the previously obtained average value during a 100 ps run. The resulting shock waves were again damped in another simulation using the Langevin thermostat with a coupling constant of 100 fs and a simulation time of 250 ps. To equilibrate the system, a simulation in a *NVT* ensemble lasting 2.5 ns was performed by using a Nosé–Hoover thermostat with a coupling constant of 100 fs. The final production run was conducted in *NVT* ensemble using a Nosé–Hoover thermostat with a coupling constant of 100 fs and a simulation time of 10 ns. The positions of the particles were written to trajectory every 100 fs and evaluated in further analyses. The integrator time step was set to 0.5 fs in all force field simulations. The bonds and angles in TIP4P-EW water were kept rigid *via* the RATTLE algorithm.[91,92] No other bonds or angles were constrained. We used a Coulomb and Lennard-Jones cutoff radius of 800 pm and a PPPM long-range Coulomb solver (as implemented in LAMMPS).[86]

All trajectory analyses have been performed with the TRAVIS program package.[93,94] The molecular structures as well as the snapshot of the simulation cell was created with VMD[95] and Tachyon.[96] The plots were prepared with xmgrace,[97] while the contour plots have been plotted with Wolfram Mathematica.[98]

3 Optimization procedure

The starting point for the development of our new force field BILFF was the CL&P force field of J. N. Canongia Lopes and A. A. H. Pádua,[46,47,50] which is an extension of the OPLS-AA force field.[61–63] The functional form of OPLS-AA is expressed as

$$U(r^N) = \sum_{\text{bonds}}^{i} \frac{k_l}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}}^{i} \frac{k_\theta}{2}(\theta_i - \theta_{i,0})^2$$
$$+ \sum_{\text{dihedrals}}^{i} \left[\frac{V_1^i}{2}[1 + \cos(\phi_i)] + \frac{V_2^i}{2}[1 - \cos(2\phi_i)] \right.$$
$$\left. + \frac{V_3^i}{2}[1 + \cos(3\phi_i)] \right]$$
$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] + \frac{q_i q_j e^2}{4\pi\varepsilon_0 r_{ij}} \right] f_{ij}.$$
$$(1)$$

The first 3 terms correspond to the bonded interactions. The bond lengths, angles, and torsion parameters were adapted to the values from the AIMD by modifying the equilibrium values, force constants, and the height of the potential barriers (final parameters see Tables S1–S6 in the ESI†). The non-bonded interactions in the last term include the Lennard-Jones potential and Coulomb interactions. The parameter f_{ij} is a scaling factor which scales the Lennard-Jones and Coulomb interactions of 1-2 and 1-3 neighbors to zero, and those of 1-4 neighbors by 0.5 (in accordance to OPLS-AA).

To adjust the partial charges and Lennard-Jones parameters, a comparison of the radial distribution functions (RDFs) of the intermolecular hydrogen bonds of the ring hydrogen atoms H2, H4 and H5 of [EMIm]$^+$ (see Fig. 1) with the oxygen atoms of acetate and water as well as the mean squared displacement (MSD) from the AIMD and the FFMD was performed after each adjustment. These protic ring hydrogen atoms form the most important hydrogen bonds in the system.[27,28] The optimal total charge of the ions was found to be ±0.82 by comparison of the diffusion coefficients. Reduced ion charges for ionic liquids have often been used in the literature to account for charge screening effects as an alternative to explicitly polarizable force fields.[99–102] The optimization procedure was repeated as an iterative process until a good agreement of the results with the reference AIMD simulation was achieved. Because of the small expected influence of the methyl and ethyl side chains on the charge distribution, it was assumed that the hydrogen atoms H4 and H5 of [EMIm]$^+$ are equivalent, so that identical force field parameters were used for them. For water molecules, the TIP4P-EW force field[64] was applied without adjustments. As noted above, the optimized force field parameters from this work can be found in Tables S1–S6 in the ESI.†

3.1 Microstructure

The development of BILFF is based on the parameters of the CL&P force field,[46,47,50] which were further optimized by comparing the results of the FFMD with the AIMD. In the following, we briefly

1244 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 1242–1253

This journal is © the Owner Societies 2021

analyze the RDFs of the AIMD before comparing them with the results of FFMD simulations using the CL&P force field,[46,47,50] the force field of Z. Liu, S. Huang and W. Wang[45] and our novel force field BILFF.

The RDFs of an AIMD of the pure ionic liquid show a large residence probability of the acetate oxygen atoms close to the hydrogen atoms H2, H4 and H5 of $[EMIm]^+$ (see Fig. 2). The distance of the first maximum between the acetate oxygen and the hydrogen atoms of $[EMIm]^+$ is 190 pm in the case of H2 and 205 pm in the case of H4 and H5. The fact that the oxygen atoms of $[OAc]^-$ have their highest residence probability in the vicinity of H2 of $[EMIm]^+$ (black curve) suggests that this interaction is the strongest hydrogen bond in the system.
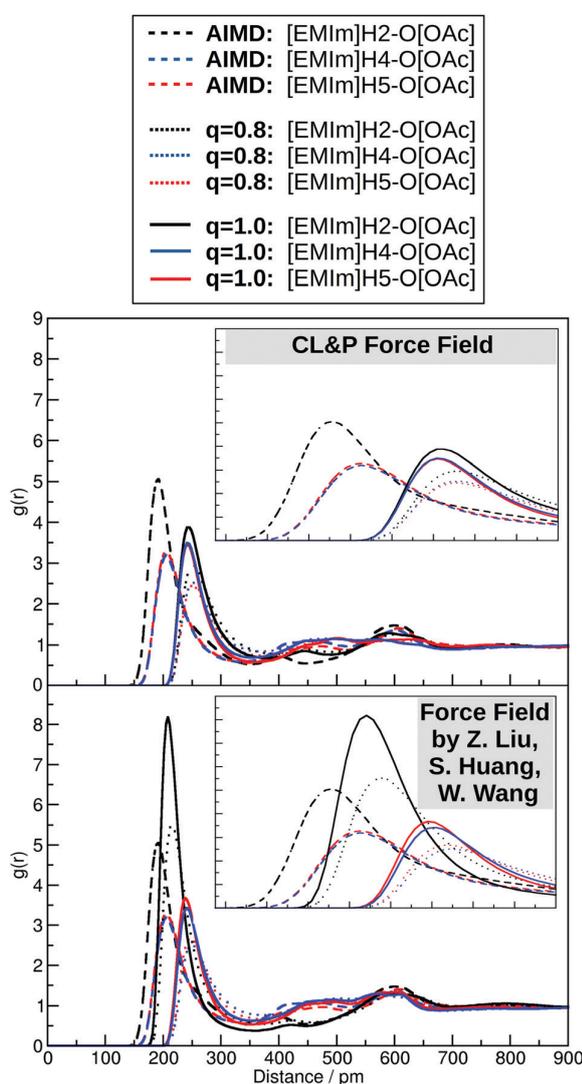
Fig. 2 Comparison of the radial distribution functions of the AIMD and FFMD simulations using the existing CL&P force field[46,47,50] (top panel) and the existing force field developed by Z. Liu *et al.*[45] (bottom panel) between the hydrogen atoms H2, H4 and H5 of $[EMIm]^+$ and the oxygen atoms of acetate in the pure ionic liquid with different total charges of the ions.

As expected, the mostly equivalent hydrogen atoms H4 and H5 show similar behavior and form significantly weaker hydrogen bonds to acetate. This can be explained by the slightly higher partial charge of H2 due to the closeness to both nitrogen atoms. To compare the radial distribution functions of the hydrogen bonds in the AIMD with the results of the FFMD simulations using the CL&P force field[46,47,50] and the force field by Z. Liu, S. Huang and W. Wang,[45] simulations with a total ion charge of ± 1.0 and ± 0.8 were performed. Both force fields were originally developed for an ionic charge of ± 1.0, while a reduced ionic charge of ± 0.8 has been widely used in the literature.[99–102] Compared to the AIMD, the RDF maxima of the FFMD simulations using both force fields with total ion charges of ± 1.0 and ± 0.8 show different peak heights with simultaneously larger particle distances (see Fig. 2). This fact shows the necessity of adaptation of the partial charges, which influence the strength of the interaction, and the adaption of the Lennard-Jones radii which influence the particle distances (please note that both force fields do not include optimized parameters for acetate anions). In contrast, the results of the novel force field BILFF fits well to the AIMD for these hydrogen bonds in both the pure and the aqueous ionic liquid (see Fig. 3).

In the aqueous ionic liquid, the first maximum of the RDF of the AIMD shows that the acetate oxygen atoms form hydrogen bonds with H2, H4, and H5 at a distance of about 200 pm and 220 pm. Due to the presence of water, the hydrogen bonds are weakened in total. This effect can be explained by the fact that the oxygen atoms of the acetate ions are partly occupied by the hydrogen atoms of the water molecules (see Fig. 6). H2 forms a stronger hydrogen bond than H4 and H5 to both the oxygen atom in acetate and the oxygen atom in water (see Fig. 3 and Fig. S2 in the ESI†). The strongest hydrogen bonds are formed between the oxygen atoms of acetate ions and the hydrogen atoms of water at a particle distance of 175 pm (see Fig. 6).

By comparing the FFMD using BILFF with the AIMD, our novel force field shows also a good agreement for the aqueous system – not only in consideration of the hydrogen bonds between $[EMIm]^+$ and $[OAc]^-$ (see Fig. 3), but also with regard to the hydrogen bonds between $[EMIm]^+$ and water (see Fig. S2 in the ESI†). The RDFs of these hydrogen bonds in the pure and aqueous $[EMIm][OAc]$ are also in accordance with several investigations in literature that have been published before[27,103–105] and confirm the accuracy of BILFF.

So far, only the ring hydrogen atoms of $[EMIm]^+$ have been considered. Fig. 4 extends this consideration to the side chains of $[EMIm]^+$. The radial distribution function between the carbon atoms of the ethyl side chain of two cations can also be reproduced well by BILFF. The maximum residence probability of two carbon atoms is about 390 pm.

To consider the angular distribution of the atoms of the hydrogen bonds as well, combined distribution functions (CDFs) were investigated. Fig. 5 shows the combined distance/angle distribution function between the hydrogen atom H2 of the cation and the acetate oxygen atoms in the pure ionic liquid

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 1242–1253 | **1245**
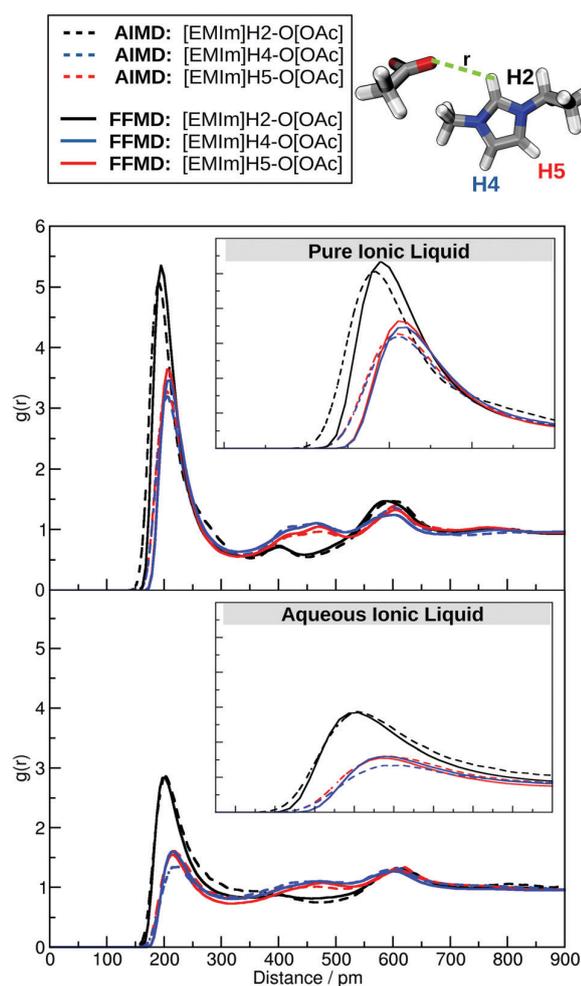
Fig. 3 Comparison of the radial distribution function of the AIMD and the FFMD simulations using BILFF between the hydrogen atoms H2, H4 and H5 of [EMIm]$^+$ and the oxygen atoms of acetate in the pure (top panel) and the aqueous ionic liquid (bottom panel).

Fig. 4 Comparison of the radial distribution functions from the AIMD and the force field MD using BILFF between the terminal carbon atoms of the ethyl side chains of [EMIm]$^+$ ions in the pure and the aqueous ionic liquid.

from the trajectories of the AIMD and the FFMD using BILFF. The angle spanned between the connection vector from H2 to the neighboring ring carbon and the connection vector from H2 to the acetate oxygen atom was considered.

The combined distribution function in Fig. 5 shows a good agreement of the FFMD using BILFF and the AIMD. Slight deviations can occur due to simulation-related noise caused by local density fluctuations. The maximum residence probability for the interaction between H2 of [EMIm]$^+$ and the acetate oxygen atoms ranges from 110° to 180° in the AIMD and from around 105° to 180° in the FFMD, while the distances are found in a range of about 140 pm to 305 pm and 150 pm to 295 pm, respectively. Since hydrogen bonds are characterized by a bond length of about 200 pm and an angle close to 180° between the hydrogen atom and the oxygen atom, the existence of a hydrogen bond is indicated. The black rectangles in the figure represent structural conditions for the existence of a hydrogen bond

between the atoms studied. These conditions will be used later in the work to calculate the lifetime of the hydrogen bonds (see Section 3.2). The second peak at around 390 pm and 180° shows the presence of the second oxygen atom of the acetate ion in the presence of the observed hydrogen bond. The AIMD and the FFMD are also in good agreement on this point, thus confirming the correctness of BILFF.

Fig. 6 depicts the residence probabilities of the hydrogen atoms of water in the presence of the acetate oxygen atoms. For this purpose, the enclosed angle between the connection vector of the hydrogen and oxygen atoms of water and the connection vector of the water hydrogen atom and the acetate oxygen atom are analyzed. The large intensity of the peak in the range of about 140° to 180° and 130 pm to 240 pm indicates a strong hydrogen bond. In a distance of 390 pm to the hydrogen atom of water, the second acetate oxygen is located. The CDF of the FFMD simulation using BILFF agrees very well with the AIMD. Furthermore, these CDFs are in agreement with already known results published in literature,[66,106] and therefore validate the force field. Additional CDFs of the simulations can be found in Fig. S4–S6 in the ESI,† such as a CDF between the [EMIM]$^+$ cations which shows that our force field nicely reproduces the π–π stacking of the cations, which is well-known in the literature[66,67,107,108] and also seen in the reference AIMD simulations (see Fig. S5, ESI†).

To further validate BILFF, spatial distribution functions (SDFs) of the FFMD and the reference AIMD were calculated and compared. Fig. 7 shows the comparison of the SDFs from the AIMD and the FFMD using BILFF, considering the acetate oxygen and the ring centers of neighboring [EMIm]$^+$ ions around a fixed cation in the pure IL (see Fig. 7 upper and lower panel) and the water oxygen atoms in the aqueous system (see Fig. 7, mid panel). Again, BILFF shows a good agreement with the AIMD. As expected, the oxygen atoms of acetate and water accumulate around the ring protons H2, H4 and H5 due to the hydrogen bond. The ring centers (RC) of the [EMIm]$^+$ cations arrange themselves on top of each other due to the π–π stacking which was discussed above.

Paper
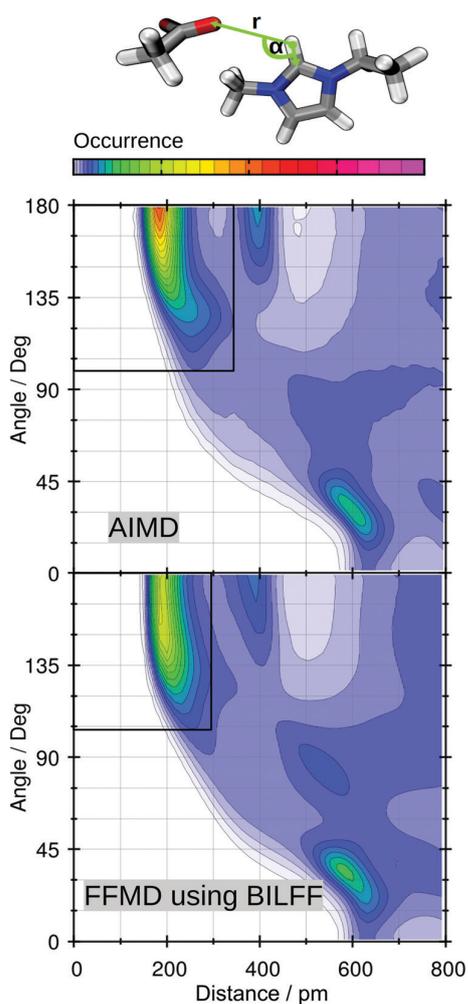
Fig. 5 Combined distribution functions from the AIMD (top panel) and the force field MD using BILFF (bottom panel) between the hydrogen atom H2 of [EMIm]$^+$ and the acetate oxygen in the pure ionic liquid with marked geometric criterion for the lifetime calculation of this hydrogen bond (black rectangles).
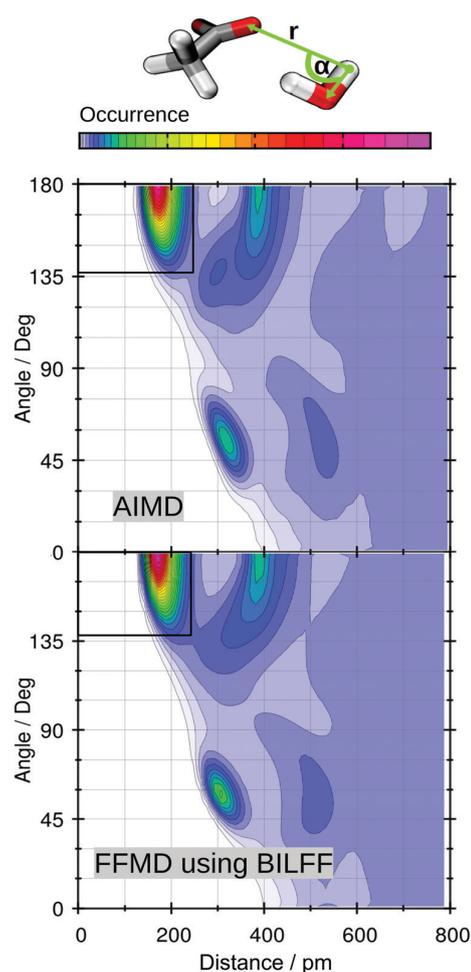
Fig. 6 Combined distribution functions (CDFs) from the AIMD (top panel) and the force field MD using BILFF (bottom panel) between the hydrogen atom of water and the acetate oxygen in the aqueous ionic liquid with marked geometric criterion for the lifetime calculation of this hydrogen bond (black rectangles).

3.2 Hydrogen bond lifetimes

The applicability of a force field for a specific task depends not only on how accurately structural properties can be reproduced, but also on the reproduction of dynamical properties. In the following, BILFF is validated by the comparison of hydrogen bond lifetimes, diffusion coefficients, and system densities with respect to the AIMD.

Table 2 shows a comparison of the lifetime of the hydrogen bonds from the trajectories of the AIMD and the FFMD using BILFF at a temperature of 350 K. The structural criteria for the formation of all hydrogen bonds have been determined according to the location of the maxima of the residence probability in the angle-distance CDFs (see black rectangles in Fig. 5 and 6). The analysis of the hydrogen bond lifetimes is based on the widely used autocorrelation formalism[109] as implemented in TRAVIS.[94] While the intermittent lifetimes allow for the reformation of the hydrogen bond after a short interruption, the

continuous lifetimes do not allow for interruptions.[110] From the results of the intermittent and continuous hydrogen bond lifetimes, the following conclusions can be drawn for the hydrogen bonds in [EMIm][OAc]:

• Measured in terms of lifetime, the hydrogen atom H2 of [EMIm]$^+$ forms more stable hydrogen bonds than H4 and H5 in both pure and aqueous ionic liquid.

• In the aqueous ionic liquid, the hydrogen atoms H4 and H5 tend to form similarly stable hydrogen bonds to the oxygen atom of acetate and water.

• The presence of water reduces the lifetime and thus the stability of each hydrogen bond between the hydrogen atoms of [EMIm]$^+$ and the acetate oxygen atoms.

• The hydrogen bonds between the hydrogen atoms of [EMIm]$^+$ and the oxygen atom of water have a shorter lifetime and thus a lower stability than the hydrogen bonds between the hydrogen atoms of [EMIm]$^+$ and the acetate oxygen atoms.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 1242–1253 | **1247**
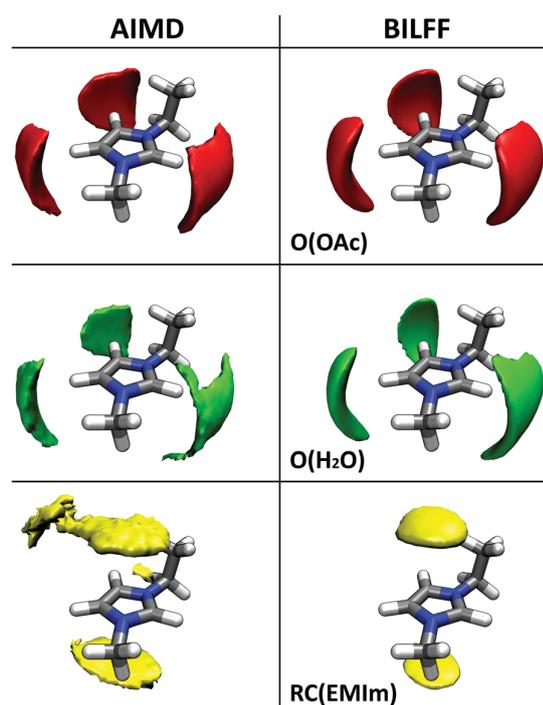
Fig. 7 Spatial distribution functions (SDFs) around the [EMIm]$^+$ cation from our reference AIMD simulations (left-hand side) and simulations with our new force field BILFF (right-hand side). Upper panel: Acetate oxygen atoms in the pure IL (red, iso-value 25 nm^{-3}), mid panel: water oxygen atoms in the aqueous system (green, iso-value 40 nm^{-3}), lower panel: ring centers (RC) of other [EMIm]$^+$ cations in the pure IL (yellow, iso-value 10 nm^{-3}).

Table 2 Overview of the lifetime τ of hydrogen bonds in the AIMD and the force field MD using BILFF at 350 K. All numbers in ps

	Intermittent lifetime		Continuous lifetime	
Hydrogen bond	τ(AIMD)	τ(FFMD)	τ(AIMD)	τ(FFMD)
Pure [EMIm][OAc]				
[EMIm]H2–O[OAc]	736.25	506.77	10.46	14.12
[EMIm]H4–O[OAc]	312.18	385.93	7.28	6.71
[EMIm]H5–O[OAc]	306.14	353.49	6.39	7.60
Aqueous [EMIm][OAc]				
[EMIm]H2–O[OAc]	104.81	126.11	4.28	4.69
[EMIm]H4–O[OAc]	52.50	74.14	2.48	2.74
[EMIm]H5–O[OAc]	39.10	67.34	1.81	2.50
[EMIm]H2–O(H$_2$O)	52.08	41.04	1.41	1.54
[EMIm]H4–O(H$_2$O)	41.11	29.62	1.73	1.27
[EMIm]H5–O(H$_2$O)	46.33	30.74	2.35	1.29
(H$_2$O)H–O[OAc]	192.11	199.31	0.76	1.94

• Considering the intermittent lifetimes in the aqueous [EMIm][OAc], the most stable hydrogen bond in the presence of water is found between the hydrogen atoms of water and the oxygen atoms of acetate.

The continuous hydrogen bond lifetime shows that the first bond breakage of a hydrogen bond occurs after only a few

picoseconds. It should be noted that the lifetime of some hydrogen bonds has a greater value than the simulation time of the AIMD (235.5 ps for pure or 167.5 ps for the aqueous system), so that the lifetime is also the subject of a large error interval in this case. This may explain deviations in the results of the AIMD and the FFMD with BILFF. The lifetime of the hydrogen bonds from the AIMD can be approximately reproduced with the simulations with BILFF. A better agreement of the lifetime of the hydrogen bonds of the FFMD with the AIMD is possible by variation of the Lennard-Jones radii and partial charges of the atoms in the force field. However, the change of these parameters would negatively influence the agreement of the RDFs between the FFMD and the AIMD in terms of the intensity of the hydrogen bonds.

3.3 Directed interaction energies

To get some insight into the directed interactions in our simulations, we computed the specific pairwise interaction energies between the ions and molecules. As only pairwise interactions are considered in standard FFMD (without explicit polarizability and three-body terms) as performed here, this can be easily achieved by re-computing the potential energy of snapshots from the original trajectory with certain Lennard-Jones and Coulomb interactions switched off. It can be easily understood that in ILs, the Coulomb interaction between the ions is dominant, and all other (possibly more interesting) interactions are more or less shadowed by this huge contribution. Therefore, we compute so-called directed interaction energies. Consider the example of the interaction between anion and cation in the pure IL, *e.g.*, *via* a hydrogen bond. The directed interaction energy between anion (A) and cation (C) is computed as $E_{\text{A–C}}^{\text{dir}} = E_{\text{A–C}} - E_{\text{A}} - E_{\text{C}}$, where $E_{\text{A–C}}$ contains all non-bonded Lennard-Jones and Coulomb interactions present in the system, E_{A} contains only non-bonded interactions among anions and has the total charge of the missing cations uniformly distributed in the simulation cell as a background charge, and E_{C} contains only non-bonded interactions among cations and has the total charge of the missing anions uniformly distributed in the simulation cell as a background charge. Bonded interactions are completely neglected. This approach ensures that all snapshot calculations are electrically neutral in total, which is required to compute the long range Coulomb interactions. A positive value of the directed interaction energy therefore does not mean that the two ions repel each other – it just means that the interaction is less favorable than the hypothetical case in which both ions see the total charge of each other uniformly distributed in the cell as a background charge. A negative value of E^{dir} indicates the presence of a directed interaction such as a hydrogen bond, while a value close to zero or even positive indicates the absence of such directed interactions. We have already demonstrated this approach several times in the literature.[31,32]

The computed directed interaction energies E^{dir} between the ions and molecules in the FFMD simulations of pure and aqueous [EMIm][OAc] using BILFF are presented in Table 3; all values are normalized to one ion pair. In the pure IL, we find

Table 3 Directed interaction energies E^{dir} between anion, cation, and water in pure and aqueous FFMD simulations using BILFF; all values normalized to one ion pair

Interaction energy/kJ mol^{-1} per IP	Pure [EMIm][OAc]	Aqueous [EMIm][OAc]
Anion–cation	−25.97	3.26
Anion–water	—	−170.63
Cation–water	—	−26.85
Water–water	—	−37.04

that the directed interaction energy between anion and cation has a value of around −26 kJ mol^{-1} per ion pair, which indicates the presence of a directed interaction between anion and cation – most likely the hydrogen bond which was discussed above. Note again that this value is not the absolute hydrogen bond energy, but rather the energy difference by which the hydrogen bond is more favorable than the hypothetical case in which each ion sees the total charge of its hydrogen bond partner as uniform background charge. In the aqueous system, on the other hand, the directed interaction energy between anion and cation is close to zero and even slightly positive, which allows the conclusion that the hydrogen bond between anion and cation becomes significantly weaker if water is added, which was already found when discussing the RDFs above. This effect is certainly caused by the extremely strong hydrogen bond between acetate and water, characterized by a directed interaction energy of more than −150 kJ mol^{-1} per acetate ion. The hydrogen bond between [EMIm]$^+$ cations and water has a much lower directed interaction energy of around 25 kJ mol^{-1}.

3.4 Temperature dependence

BILFF was developed at a temperature of 350 K. To validate the force field at different temperatures, FFMD simulations of aqueous [EMIm][OAc] with the same simulation parameters (see Table 1) were performed. In a first step, the radial distribution functions of the hydrogen bond between H2 and H5 of [EMIm]$^+$ and the oxygen atoms of acetate were investigated (see Fig. 8). As expected, the residence probability $g(r)$ of the ring hydrogen atoms in the vicinity of the acetate oxygen decreases with increasing temperature. The average distance between the particles is not affected. Due to the fact that the relative position of the first peak of the RDF of the AIMD and the FFMD using BILFF is nearly unchanged, it is shown, that the FFMD can reproduce the results even at higher temperatures.

To support this statement and validate our new force field, the system densities and mean squared displacements (MSDs) at different temperatures were also investigated. A comparison of the system density of both the pure and the aqueous ionic liquid from the FFMD using BILFF shows a good reproducibility of extrapolated experimental values at different temperatures taken from ref. 111 (see Table 4). At 450 K and 550 K, the simulated densities are slightly lower than the experimental values. As expected, the density decreases with increasing temperature.

The self-diffusion coefficients were calculated from the slope of the mean squared displacements (MSDs) using the Einstein
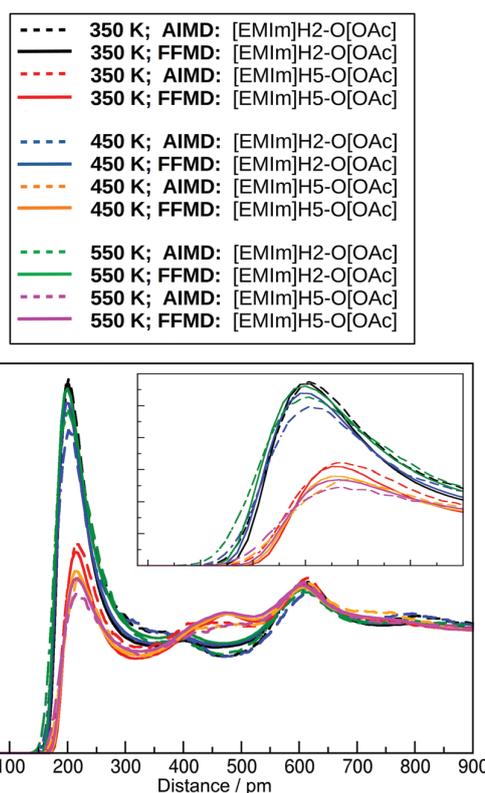
Fig. 8 Radial distribution functions (RDFs) of the hydrogen bonds between the ring hydrogen atoms H2 and H5 of [EMIm]$^+$ and the oxygen atoms of acetate in the aqueous ionic liquid at different temperatures, simulated by AIMD and force field MD using BILFF.

relation and are presented in Table 5. Experimentally determined diffusion coefficients for these molecules can be found in the literature for pure [EMIm][OAc].[112,113] The values of the temperature dependent measurements from ref. 113 were extrapolated to a temperature of 350 K by assuming an Arrhenius-type temperature dependence, as suggested in the original article.[113] Diffusion coefficients for an aqueous [EMIm][OAc] solution from ref. 112 were measured at 313.15 K with a molar fraction of $x_{IL} = 0.3$. Since the measurements were performed at a

Table 4 Comparison of the system density from the force field MD using BILFF of pure and aqueous [EMIm][OAc] ($x_{IL} = 0.25$) with extrapolated experimental values at different temperatures

Temperature	ρ(FFMD)/g cm^{-3}	ρ(Lit.)^a/g cm^{-3}
Pure [EMIm][OAc]		
350 K	1.078	1.066
450 K	0.996	1.006
550 K	0.913	0.945
Aqueous [EMIm][OAc]		
350 K	1.070	1.069
450 K	0.973	1.003
550 K	0.856	0.937

^{a} Extrapolated values of temperature dependent measurements ($x_{IL} = 0.252$ in the aqueous system) from ref. 111.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 1242–1253 | **1249**

Table 5 Computed self-diffusion coefficients D from FFMD using BILFF of pure and aqueous [EMIm][OAc] (x_{IL} = 0.25) at different temperatures together with experimental results (last column)

Ion	Temperature/K	D(FFMD)/ 10^{-11} m^2 s^{-1}	D(Lit.)^{ab}/ 10^{-11} m^2 s^{-1}
Pure [EMIm][OAc]			
[EMIm]$^+$	350	9.26	14.1/0.98
	450	81.76	—/—
	550	207.59	—/—
[OAc]$^-$	350	6.72	12.8/0.84
	450	66.83	—/—
	550	170.77	—/—
Aqueous [EMIm][OAc]			
[EMIm]$^+$	350	21.77	—/4.17
	450	171.09	—/—
	550	436.00	—/—
[OAc]$^-$	350	22.42	—/4.61
	450	155.76	—/—
	550	463.61	—/—
H$_2$O	350	55.97	—/10.47
	450	378.47	—/—
	550	1160.70	—/—

^{a} Extrapolated to 350 K via Arrhenius plot from temperature-dependent experimental measurements at T = 283–333 K, see ref. 113. ^{b} Measurement at T = 313.15 K with a molar fraction of [EMIm][OAc] of x_{IL} = 0.3 in the aqueous system.[112]

lower temperature and with a different substance amount fraction, the differences between experiment and simulation (x_{IL} = 0.25) are expected and quite acceptable. The larger cation diffuses faster than the smaller anion in pure ionic liquid. All diffusion coefficients are highly increased by the presence of water. In the aqueous ionic liquid, the increase in the diffusion coefficient of the anion is greater than for the cation. Thus, in contrast to the pure ionic liquid, the anion diffuses faster than the cation in the aqueous system. This fact can be confirmed by experimental data taken from ref. 112. However, this effect does not occur at 450 K. Water as the smallest molecule has the highest self-diffusion coefficient.

There is evidence in literature that the self-diffusion coefficient of ionic liquids obeys an Arrhenius-type behavior with respect to temperature.[114] To check whether this is also the case in our simulations, Arrhenius plots of the diffusion coefficients at different temperatures is presented in Fig. S7 and S8 as well as Table S7 in the ESI;† it turns out that both the pure and the aqueous system are close to Arrhenius-type behavior, as seen by the almost linear shape in the logarithmic-reciprocal plots.

Apart from that, it has been reported in literature that the dynamics in ionic liquids can be sub-diffusive in an intermediate correlation time range,[115] resembling the behavior of supercooled liquids before the glass transition.[116] To check this, the parameter can be computed according to the equation

$$\beta(t) = \frac{d \log_{10}\langle \Delta r(t)^2 \rangle}{d \log_{10} t} \quad (2)$$

from ref. 117. We present the corresponding plots of our pure and aqueous simulations at three different simulation temperatures in
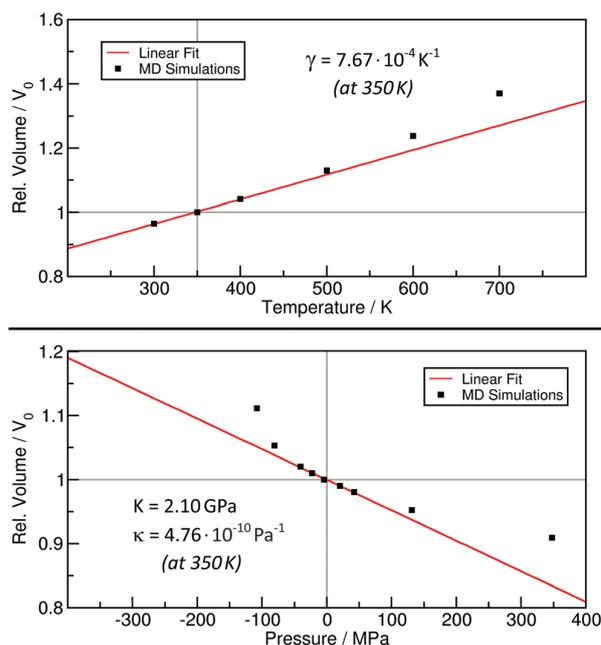
Fig. 9 Upper panel: Computing the thermal volume expansion coefficient γ from multiple FFMD simulations using BILFF at constant pressure (1 bar) and different temperatures; lower panel: computing the compressibility κ and the bulk modulus K from multiple FFMD simulations using BILFF at constant temperature (350 K) and different pressures.

Fig. S9 in the ESI.† Due to the fact that approaches the value 1 with increasing correlation depth, it can be assumed that the dynamics of our simulations is in the diffusive regime.

3.5 Thermal expansion and bulk modulus

Another interesting category of physicochemical properties for the characterization of liquids are linear response coefficients such as the thermal volume expansion coefficient and the bulk modulus. We used BILFF to perform force field molecular dynamics simulations in the NpT ensemble at different temperatures and pressures and monitored the volume change. Based on these data, we estimated the thermal volume expansion coefficient γ and the bulk modulus K at T = 350 K and p = 101.3 kPa by linear regression to the results close to that point, see Fig. 9. In Table 6, we compare our computed values (first row) to experimental results from literature (second row). The experimental thermal volume expansion coefficient was measured at 348 K,[118] while the bulk modulus was determined at 353 K.[119] When compared to the experimental results, the

Table 6 Computed thermal volume expansion coefficient γ and bulk modulus K from FFMD of pure [EMIm][OAc] at 350 K using BILFF (first row, see Fig. 9) together with experimental results from the literature[118,119] (second row)

	Volume expansion coeff./10^{-4} K^{-1}	Bulk modulus/GPa
FFMD with BILFF	7.67	2.10
Experiment	5.43 [ref. 118]	2.88 [ref. 119]

thermal volume expansion coefficient and bulk modulus from BILFF show deviations of 41% and 37%, respectively, which is still in the acceptable range.

4 Conclusion

In this work, the new force field BILFF (Bio-Polymers in Ionic Liquids Force Field) is introduced. In this first part of the study, we have presented optimized force field parameters for mixtures of the ionic liquid [EMIm][OAc] with water. BILFF is a force field specifically designed to accurately describe the microstructure and dynamics of directed interactions such as hydrogen bonds in this particular system, and enables simulations with high precision combined with small simulation times. To achieve this, the force field parameters were iteratively optimized by comparing the radial and combined distribution functions of the intermolecular hydrogen bonds and the diffusion coefficients of the molecules to those from reference AIMD simulations of a pure and an aqueous system at a temperature of 350 K. The total ion charges, the partial atomic charges, and the Lennard-Jones radii were iteratively adjusted. The bond lengths, angles and torsion parameters were adapted by changing the equilibrium values and force constants or by changing the height of the potential barriers. A good agreement between FFMD simulation using BILFF and the reference AIMD was achieved in terms of combined distance/angle distribution functions between the hydrogen bonding sites as well as between the ring centers of the cation, correctly reproducing well-known effects such as the π–π stacking of the cations,[66,67,107,108] which is also found in the AIMD simulations. The radial distribution function between the ethyl side chains of two cations was also reproduced well, indicating a correct description of dispersion interactions. In a comparison of the radial distribution functions of the hydrogen bonds from simulations with BILFF and the frequently used force field of J. N. A. Canongia Lopes and A. A. H. Pádua,[46,47,50] BILFF even shows a better agreement with the AIMD (please note that the CL&P force field does not contain optimized parameters for the acetate anion). Furthermore, the spatial distribution functions (SDFs) for several atom types around the cation from the FFMD with BILFF agree well with the AIMD, and fully reproduce, e.g., the π–π stacking of the cations.

Dynamical properties such as the average lifetimes of the hydrogen bonds as well as the self-diffusion coefficients are reproduced well with a force field MD using BILFF for both the pure and the aqueous system as revealed by the comparison with AIMD simulations. The force field parameters were validated at elevated temperatures by additional simulations at 450 K and 550 K: the radial distribution functions of the hydrogen bonds between the ring hydrogen atoms of [EMIm]$^+$ and the oxygen atoms of acetate, the system density and the diffusion coefficients were analyzed. The comparison of the simulated diffusion coefficients, the system density, the thermal volume expansion coefficient, and the bulk modulus with experimental data shows a good agreement even at higher temperatures.

The validation of the force field demonstrates that BILFF offers a very cost-effective and accurate new way of simulating mixtures of [EMIm][OAc] with water, in particular if the microstructure and the hydrogen bonding needs to be described correctly. An attractive possible application is the simulation of solvation processes of bio-polymers such as cellulose in [EMIm][OAc]. Therefore, BILFF will be extended to cellulose (and possibly also other bio-polymers) as well as further ionic liquids based on cations such as 1,2,3-triazolium and anions such as benzoate, salicylate and lactate – for which we already have measured the cellulose solubility[31,32] – in upcoming articles.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 P. Walden, *Bull. Acad. Imp. Sci. St.-Petersbourg*, 1914, **1800**, 405–422.
2 K. Ghandi, *Green Sustainable Chem.*, 2014, **04**, 44–53.
3 F. Malberg, M. Brehm, O. Hollóczki, A. S. Pensado and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18424–18436.
4 Z. M. Judeh, C. B. Ching, J. Bu and A. McCluskey, *Tetrahedron Lett.*, 2002, **43**, 5089–5091.
5 L. A. Blanchard, D. Hancu, E. J. Beckman and J. F. Brennecke, *Nature*, 1999, **399**, 28–29.
6 T. Welton, *Chem. Rev.*, 1999, **99**, 2071–2084.
7 R. Sheldon, *Chem. Commun.*, 2001, 2399–2407.
8 D. Zhao, M. Wu, Y. Kou and E. Min, *Catal. Today*, 2002, **74**, 157–189.
9 B. Ullah, J. Chen, Z. Zhang, H. Xing, Q. Yang, Z. Bao and Q. Ren, *Sci. Rep.*, 2017, **7**, 42699.
10 Y. Zhou and J. Qu, *ACS Appl. Mater. Interfaces*, 2017, **9**, 3209–3222.
11 M. J. A. Shiddiky and A. A. J. Torriero, *Biosens. Bioelectron.*, 2011, **26**, 1775–1787.
12 H. Sakaebe, H. Matsumoto and K. Tatsumi, *Electrochim. Acta*, 2007, **53**, 1048–1054.
13 M. Armand, F. Endres, D. R. MacFarlane, H. Ohno and B. Scrosati, *Nat. Mater.*, 2009, **8**, 621–629.
14 M. A. Navarra, *MRS Bull.*, 2013, **38**, 548–553.
15 G. A. Elia, U. Ulissi, S. Jeong, S. Passerini and J. Hassoun, *Energy Environ. Sci.*, 2016, **9**, 3210–3220.
16 Y. Bai, Y. Cao, J. Zhang, M. Wang, R. Li, P. Wang, S. M. Zakeeruddin and M. Grätzel, *Nat. Mater.*, 2008, **7**, 626–630.
17 G. P. S. Lau, J.-D. Décoppet, T. Moehl, S. M. Zakeeruddin, M. Grätzel and P. J. Dyson, *Sci. Rep.*, 2015, **5**, 18158.
18 S. Denizalti, A. K. Ali, Ç. Ela, M. Ekmekci and S. Erten-Ela, *Chem. Phys. Lett.*, 2018, **691**, 373–378.

19 R. D. Rogers, *Nature*, 2007, **447**, 917–918.

20 E. E. L. Tanner, A. M. Curreri, J. P. R. Balkaran, N. C. Selig-Wober, A. B. Yang, C. Kendig, M. P. Fluhr, N. Kim and S. Mitragotri, *Adv. Mater.*, 2019, **31**, e1901103.

21 F. Endres and S. Zein El Abedin, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2101–2116.

22 N. V. Plechkova and K. R. Seddon, *Chem. Soc. Rev.*, 2008, **37**, 123–150.

23 K. Fumino, A. Wulf and R. Ludwig, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8790–8794.

24 C. Roth, T. Peppel, K. Fumino, M. Köckerling and R. Ludwig, *Angew. Chem., Int. Ed.*, 2010, **49**, 10221–10224.

25 K. Fumino, T. Peppel, M. Geppert-Rybczyńska, D. H. Zaitsau, J. K. Lehmann, S. P. Verevkin, M. Köckerling and R. Ludwig, *Phys. Chem. Chem. Phys.*, 2011, **13**, 14064–14075.

26 T. Peppel, C. Roth, K. Fumino, D. Paschek, M. Köckerling and R. Ludwig, *Angew. Chem., Int. Ed.*, 2011, **50**, 6661–6665.

27 M. Brehm, H. Weber, A. S. Pensado, A. Stark and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2012, **14**, 5030–5044.

28 Y. Chen, Y. Cao, X. Sun and T. Mu, *J. Mol. Liq.*, 2014, **190**, 151–158.

29 R. Giernoth, A. Bröhl, M. Brehm and Y. Lingscheid, *J. Mol. Liq.*, 2014, **192**, 55–58.

30 Y. Li, X. Liu, S. Zhang, Y. Yao, X. Yao, J. Xu and X. Lu, *Phys. Chem. Chem. Phys.*, 2015, **17**, 17894–17905.

31 M. Brehm, M. Pulst, J. Kressler and D. Sebastiani, *J. Phys. Chem. B*, 2019, **123**, 3994–4003.

32 M. Brehm, J. Radicke, M. Pulst, F. Shaabani, D. Sebastiani and J. Kressler, *Molecules*, 2020, **25**, 3539.

33 T. Watermann and D. Sebastiani, *Z. Phys. Chem.*, 2018, **232**, 989–1002.

34 A. Sharma, M. Thakur, M. Bhattacharya, T. Mandal and S. Goswami, *Biotechnol. Rep.*, 2019, **21**, e00316.

35 J. Zhang, H. Zhang, J. Wu, J. Zhang, J. He and J. Xiang, *Phys. Chem. Chem. Phys.*, 2010, **12**, 1941–1947.

36 T. Endo, S. Hosomi, S. Fujii, K. Ninomiya and K. Takahashi, *Molecules*, 2017, **22**, 178.

37 Z.-D. Ding, Z. Chi, W.-X. Gu, S.-M. Gu, J.-H. Liu and H.-J. Wang, *Carbohydr. Polym.*, 2012, **89**, 7–16.

38 Y. Li, X. Liu, S. Zhang, Y. Yao, X. Yao, J. Xu and X. Lu, *Phys. Chem. Chem. Phys.*, 2015, **17**, 17894–17905.

39 Y.-R. Liu, K. Thomsen, Y. Nie, S.-J. Zhang and A. S. Meyer, *Green Chem.*, 2016, **18**, 6246–6254.

40 B. Manna and A. Ghosh, *J. Biomol. Struct. Dyn.*, 2019, **37**, 3987–4005.

41 T. Ishida, *J. Phys. Chem. B*, 2020, **124**, 3090–3102.

42 P. Griffin, S. Ramer, M. Winfough and J. Kostal, *Green Chem.*, 2020, **22**, 3626–3637.

43 R. Faller, H. Schmitz, O. Biermann and F. Müller-Plathe, *J. Comput. Chem.*, 1999, **20**, 1009–1017.

44 D. Reith, H. Meyer and F. Müller-Plathe, *Comput. Phys. Commun.*, 2002, **148**, 299–313.

45 Z. Liu, S. Huang and W. Wang, *J. Phys. Chem. B*, 2004, **108**, 12978–12989.

46 J. N. Canongia Lopes, J. Deschamps and A. A. H. Pádua, *J. Phys. Chem. B*, 2004, **108**, 2038–2047.

47 J. N. Canongia Lopes and A. A. H. Pádua, *J. Phys. Chem. B*, 2006, **110**, 19586–19592.

48 T. Köddermann, K. Fumino, R. Ludwig, J. N. Canongia Lopes and A. A. H. Pádua, *ChemPhysChem*, 2009, **10**, 1181–1186.

49 X. Zhong, Z. Liu and D. Cao, *J. Phys. Chem. B*, 2011, **115**, 10027–10040.

50 J. N. Canongia Lopes and A. A. H. Pádua, *Theor. Chem. Acc.*, 2012, **131**, 3330.

51 T. Köddermann, D. Reith and R. Ludwig, *ChemPhysChem*, 2013, **14**, 3368–3374.

52 F. Dommert, K. Wendler, B. Qiao, L. Delle Site and C. Holm, *J. Mol. Liq.*, 2014, **192**, 32–37.

53 J. Michalowsky, L. V. Schäfer, C. Holm and J. Smiatek, *J. Chem. Phys.*, 2017, **146**, 054501.

54 B. Doherty, X. Zhong, S. Gathiaka, B. Li and O. Acevedo, *J. Chem. Theory Comput.*, 2017, **13**, 6131–6145.

55 F. Uhlig, J. Zeman, J. Smiatek and C. Holm, *J. Chem. Theory Comput.*, 2018, **14**, 1471–1486.

56 K. Goloviznina, J. N. Canongia Lopes, M. Costa Gomes and A. A. H. Pádua, *J. Chem. Theory Comput.*, 2019, **15**, 5858–5871.

57 H. Liu, K. L. Sale, B. M. Holmes, B. A. Simmons and S. Singh, *J. Phys. Chem. B*, 2010, **114**, 4293–4301.

58 S. Gehrke, K. Schmitz and O. Holóczki, *J. Phys. Chem. B*, 2017, **121**, 4521–4529.

59 B. Manna and A. Ghosh, *J. Biomol. Struct. Dyn.*, 2019, **37**, 3987–4005.

60 B. Doherty, X. Zhong and O. Acevedo, *J. Phys. Chem. B*, 2018, **122**, 2962–2974.

61 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

62 J. W. Ponder and D. A. Case, *Adv. Protein Chem.*, 2003, **66**, 27–85.

63 S. V. Sambasivarao and O. Acevedo, *J. Chem. Theory Comput.*, 2009, **5**, 1038–1050.

64 H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, **120**, 9665–9678.

65 R. P. Swatloski, S. K. Spear, J. D. Holbrey and R. D. Rogers, *J. Am. Chem. Soc.*, 2002, **124**, 4974–4975.

66 M. Brehm, H. Weber, A. S. Pensado, A. Stark and B. Kirchner, *Z. Phys. Chem.*, 2013, **227**, 177–204.

67 M. Brehm and D. Sebastiani, *J. Chem. Phys.*, 2018, **148**, 193802.

68 The CP2K developers group, 2020. CP2K is freely available from http://www.cp2k.org/, accessed August 20, 2020.

69 J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 15–25.

70 T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß,

1252 | *Phys. Chem. Chem. Phys.*, 2021, **23**, 1242–1253

This journal is © the Owner Societies 2021

M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack and J. Hutter, *J. Chem. Phys.*, 2020, **152**, 194103.

71 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.*, 2005, **167**, 103–128.

72 J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2003, **118**, 4365–4369.

73 P. Hohenberg and W. Kohn, *Phys. Rev. B: Solid State*, 1964, **136**, 864.

74 W. Kohn and L. Sham, *Phys. Rev.*, 1965, **140**, 1133.

75 A. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.

76 C. Lee, W. Yang and R. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.

77 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

78 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

79 D. G. A. Smith, L. A. Burns, K. Patkowski and C. D. Sherrill, *J. Phys. Chem. Lett.*, 2016, **7**, 2197–2203.

80 J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2007, **127**, 114105.

81 S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 1703–1710.

82 C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 3641–3662.

83 S. Nose, *J. Chem. Phys.*, 1984, **81**, 511–519.

84 S. Nose, *Mol. Phys.*, 1984, **52**, 255–268.

85 G. Martyna, M. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.

86 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.

87 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.

88 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

89 T. Schneider and E. Stoll, *Phys. Rev. B: Solid State*, 1978, **17**, 1302–1322.

90 B. Dünweg and W. Paul, *Int. J. Mod. Phys. C*, 1991, **02**, 817–827.

91 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

92 H. C. Andersen, *J. Comput. Phys.*, 1983, **52**, 24–34.

93 M. Brehm and B. Kirchner, *J. Chem. Inf. Model.*, 2011, **51**(8), 2007–2023.

94 M. Brehm, M. Thomas, S. Gehrke and B. Kirchner, *J. Chem. Phys.*, 2020, **152**, 164105.

95 W. Humphrey, A. Dalke and K. Schulten, *et al.*, *J. Mol. Graphics*, 1996, **14**, 33–38.

96 J. Stone, MSc thesis, Computer Science Department, University of Missouri-Rolla, 1998.

97 Grace, (c) 1996–2008 Grace Development Team, see http://plasma-gate.weizmann.ac.il/Grace.

98 Wolfram Research, Inc., Mathematica, Version 8.0, Champaign, IL, 2010.

99 X. Wu, Z. Liu, S. Huang and W. Wang, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2771–2779.

100 B. L. Bhargava and S. Balasubramanian, *J. Chem. Phys.*, 2007, **127**, 114510.

101 H. Liu and E. Maginn, *J. Chem. Phys.*, 2011, **135**, 124507.

102 P. Ray, R. Elfgen and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2019, **21**, 4472–4486.

103 A. Mele, C. D. Tran and S. H. de Paoli Lacerda, *Angew. Chem., Int. Ed.*, 2003, **42**, 4364–4366.

104 M. Moreno, F. Castiglione, A. Mele, C. Pasqui and G. Raos, *J. Phys. Chem. B*, 2008, **112**, 7826–7836.

105 Q.-G. Zhang, N.-N. Wang and Z.-W. Yu, *J. Phys. Chem. B*, 2010, **114**, 4747–4754.

106 J. Thar, M. Brehm, A. P. Seitsonen and B. Kirchner, *J. Phys. Chem. B*, 2009, **113**, 15129–15132.

107 M. O. Sinnokrot, E. F. Valeev and C. D. Sherrill, *J. Am. Chem. Soc.*, 2002, **124**, 10887–10893.

108 R. P. Matthews, T. Welton and P. A. Hunt, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14437–14453.

109 D. C. Rapaport, *Mol. Phys.*, 1983, **50**, 1151–1162.

110 S. Gehrke, M. von Domaros, R. Clark, O. Hollóczki, M. Brehm, T. Welton, A. Luzar and B. Kirchner, *Faraday Discuss.*, 2018, **206**, 219–245.

111 E. Quijada-Maldonado, S. van der Boogaart, J. H. Lijbers, G. W. Meindersma and A. B. de Haan, *J. Chem. Thermodyn.*, 2012, **51**, 51–58.

112 C. A. Hall, K. A. Le, C. Rudaz, A. Radhi, C. S. Lovell, R. A. Damion, T. Budtova and M. E. Ries, *J. Phys. Chem. B*, 2012, **116**, 12810–12818.

113 S. M. Green, M. E. Ries, J. Moffat and T. Budtova, *Sci. Rep.*, 2017, **7**, 8968.

114 H. Tokuda, K. Hayamizu, K. Ishii, M. A. B. H. Susan and M. Watanabe, *J. Phys. Chem. B*, 2004, **108**, 16593–16600.

115 M. G. Del Pópolo and G. A. Voth, *J. Phys. Chem. B*, 2004, **108**, 1744–1752.

116 J. Qian, R. Hentschke and A. Heuer, *J. Chem. Phys.*, 1999, **110**, 4514–4522.

117 S. L. Perkins, P. Painter and C. M. Colina, *J. Phys. Chem. B*, 2013, **117**, 10250–10260.

118 F. Yang, X. Wang, Q. Chen and H. Tan, *J. Mol. Liq.*, 2019, **284**, 82–91.

119 S. Stevanovic, A. Podgoršek, A. A. H. Pádua and M. F. Costa Gomes, *J. Phys. Chem. B*, 2012, **116**, 14416–14425.

This journal is © the Owner Societies 2021

Phys. Chem. Chem. Phys., 2021, **23**, 1242–1253 | 1253

3.3 Article II: A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 2: Cellulose in [EMIm][OAc]/Water Mixtures

Author contributions: In this article, I performed all calculations, was responsible for data processing, data analysis, visualization of the results, and the writing of the original article draft. M. Brehm was responsible for the choice of computational methods and for supervising the progress of the data processing. D. Sebastiani was responsible for supervising the work and the conceptualization of the project. All authors contributed to further review iterations of the article.

The following article [E. Roos, D. Sebastiani, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 2: Cellulose in [EMIm][OAc]/Water Mixtures", *Phys. Chem. Chem. Phys.*, 25, 8755–8766, 2023] is reproduced from Ref. [177] with permission from the PCCP Owner Societies. No changes were made.

A force field for bio-polymers in ionic liquids (BILFF) – part 2: cellulose in [EMIm][OAc]/water mixtures†

Eliane Roos, [ID] Daniel Sebastiani [ID] and Martin Brehm [ID] *

We present the extension of our force field BILFF (Bio-Polymers in Ionic Liquids Force Field) to the bio-polymer cellulose. We already published BILFF parameters for mixtures of ionic liquid 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) with water. Our all-atom force field focuses on a quantitative reproduction of the hydrogen bonds in the complex mixture of cellulose, [EMIm]$^+$, [OAc]$^-$ and water when compared to reference *ab initio* molecular dynamics (AIMD) simulations. To enhance the sampling, 50 individual AIMD simulations starting from different initial configurations were performed for cellulose in solvent instead of one long simulation, and the resulting averages were used for force field optimization. All cellulose force field parameters were iteratively adjusted starting from the literature force field of W. Damm *et al.* We were able to obtain a very good agreement with respect to both the microstructure of the reference AIMD simulations and experimental results such as the system density (even at higher temperatures) and the crystal structure. Our new force field allows performing very long simulations of large systems containing cellulose solvated in (aqueous) [EMIm][OAc] with almost *ab initio* accuracy.

1. Introduction

Cellulose is one of the most abundant bio-materials on Earth due to its occurrence as a structural material in plants and in some fungi, algae, tunicates and extracellular products of some bacteria.[1–3] Per year, approximately 1.1×10^{11} to 1.5×10^{12} tons of lignocellulosic biomass is formed,[2,4] whose main components are cellulose (40–50 wt%), hemicellulose (25 wt%), and lignin (25 wt%).[5] For thousands of years, cellulose has been used by humanity for building shelters, producing textiles and as fuel, packaging material, and writing material.[2] In the course of time, the range of applications has steadily expanded. Apart from the uses mentioned above, cellulose is now also applied for the production of biofuels,[6,7] pharmaceuticals,[3] cement,[8] photonic hydrogels,[9] and many more.[10]

The applicability of cellulose is often limited by its poor solubility in most conventional solvents as well as water, which increases the cost of production. Substances dissolving cellulose can be classified into derivatizing and non-derivatizing (direct) solvents. A direct dissolution of cellulose not only

Institut für Chemie – Theoretische Chemie, Martin-Luther-Universität Halle–Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany.
E-mail: Martin_Brehm@gmx.de

simplifies processing, but also the solvent can be recycled more easily since no by-products are generated.[2,11,12]

Ionic liquids (ILs) are a promising class of solvents for a sustainable chemical product cycle, and many of them possess features such as low volatility, non-flammability, outstanding solvation potential as well as thermal and chemical stability.[13–18] In particular, imidazolium-based ILs moved into the focus of research due to their (electro)chemical stability as well as their relatively low viscosity.[18] A prominent example of imidazolium-based ILs is 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]), which has been discussed a lot in the literature and is used for many different applications now.[19–23]

In 2002, Swatloski *et al.* discovered that certain ILs are able to dissolve cellulose without derivatization.[24] Since then, various other imidazolium- and triazolium-based ILs which possess that ability were found, [EMIm][OAc] was one among them.[25–28] However, many aspects of cellulose solubility in ionic liquids are not yet fully understood, such as the role of the cation and the various inter- and intramolecular interactions.[29–31] The dissolution process of cellulose in [EMIm][OAc] has already been investigated in many studies *via* NMR spectroscopy,[32,33] small angle X-ray scattering,[34] and UV/vis spectroscopy.[35] In various molecular dynamics simulations, it was observed that the anions of the IL form strong hydrogen bonds with the hydroxyl groups of cellulose, which might be a possible cause for cellulose solubility.[26–28,36–44]

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 8755–8766 | 8755

To reach sufficiently long simulation times, it is advantageous to apply the approach of force field molecular dynamics (FFMD) simulations which require less computational resources than ab initio MD (AIMD) simulations. However, in many such studies, force field parameters were used that were not specifically developed to describe the dissolution of cellulose in [EMIm][OAc] as well as [EMIm][OAc]–water mixtures. Existing force fields for ILs on the basis of AMBER[45,46] and OPLS–AA[47–51] have often been developed without a special emphasis on hydrogen bonding and local structure, so that the description of ILs with very strong hydrogen bonding (such as [EMIm][OAc]) is not optimal. There already exist a few force fields for ILs with a focus on hydrogen bonding which have been derived from AIMD simulations, such as the OPLS–VSIL force field[52] which utilizes virtual sites. Furthermore, there is ongoing work on developing polarizable force fields for ionic liquids.[53] However, those do not contain parameters optimized for cellulose in [EMIm][OAc]–water mixtures. This can lead to an inaccurate description of the hydrogen bonds between cellulose and the ionic liquid and thus to an inaccurate representation of the microstructure of the complex system, which has an important impact on understanding cellulose solubility.

In a previous work, we have presented a new force field for mixtures of the IL [EMIm][OAc] and water with a special emphasis on the hydrogen bonds, called BILFF (Bio-Polymers in Ionic Liquids Force Field).[54] Now, we present the extension of this force field to cellulose in pure as well as in aqueous [EMIm][OAc]. By comparing the radial and combined distance–angle distribution functions (RDFs and CDFs, respectively) and the statistical occurrence of preferred bond lengths, angles, and dihedral angles resulting from our force field with reference to AIMD simulations, all force field parameters were iteratively adjusted focusing on the intra- and intermolecular hydrogen bonds of cellulose.

As the experimental dissolution process of cellulose can take several hours,[55,56] it is essential to go beyond the simulation times and system sizes accessible via AIMD in order to understand such processes based on MD simulations. This is enabled by our new force field BILFF without sacrificing the accurate description of the hydrogen bond network. In this article, the development of the extension of BILFF to cellulose is presented and discussed.

2. Computational details

The force field parameters for cellulose were iteratively adjusted to reference AIMD simulations of a pure and aqueous cellulose–[EMIm][OAc] system with the aim of enabling very long force field MD simulations of large systems with almost the accuracy of an AIMD simulation (see Section 3). In order to reduce the computational cost of AIMD simulations and still achieve a large simulation time and a good statistical sampling of the system, 50 AIMD simulations starting from different initial configurations were performed for cellulose in pure and aqueous [EMIm][OAc], respectively, instead of just one long simulation.

Table 1 Simulation parameters of the final equilibrated ab initio (AIMD) and force field (FFMD) molecular dynamics simulations of cellulose in pure and aqueous [EMIm][OAc]

System	Composition	Sim. time/ ps	Box size/ pm	Density/ g cm^{-3}
AIMD				
Pure	1 Cellulose trimer 42 [EMIm][OAc]	729	2285	1.07
Aqueous	1 Cellulose trimer 29 [EMIm][OAc] 87 Water	1011	2266	1.00
FFMD				
Pure	1 Cellulose pentamer 192 [EMIm][OAc]	50 000	3715	1.09
Aqueous	1 Cellulose pentamer 150 [EMIm][OAc] 450 Water	50 000	3760	1.08

Different initial configurations were chosen at equidistant intervals from a 200 ns long force field simulation of a cellulose strand consisting of three β-D-glucose monomers in 42 IL ion pairs with a cell size of $a = b = c = 2285$ pm as well as 29 IL ion pairs and 87 water molecules in a cell size of $a = b = c = 2266$ pm, respectively (see Table 1). By simultaneously adjusting the force field parameters for cellulose in pure [EMIm][OAc] as well as in the additional presence of water, we ensured that the force field can be used both for the dissolution process of cellulose and its subsequent precipitation by adding water. Here, a ratio of IL ion pairs to water of 1 : 3 (ω_{H_2O} = 23.5 wt%) was chosen for the aqueous system, at which cellulose is already precipitated again.[35] These force field simulations were performed with the force field parameters of W. Damm[57] for cellulose, our previously published force field BILFF[54] for [EMIm][OAc], and TIP4P–EW[58] for water.

It should be noted here that the use of a non-polarizable force field in the presence of polar groups or even ionic species in the system can lead to certain issues, such as diffusion coefficients deviating by several orders of magnitude.[53] In the particular case studied here, such issues do not seem to arise, as it is suggested by the good reproduction of the experimental diffusion coefficient, thermal expansion coefficient, and bulk modulus for pure [EMIm][OAc] in our previously published article[54] as well as the agreement of the crystal structure of cellulose with XRD data in this article (see Section 3.3).

The force field simulations for the initial configurations were performed as follows. For a more detailed description of the simulation protocol, please see our previous article.[54]

The molecules listed above were distributed in the simulation cell using Packmol.[59] The simulation was then equilibrated with LAMMPS[60] for 25 ps at 350 K in the NVT ensemble with a time step of 0.5 fs, using a Berendsen thermostat[61] with a coupling constant of 1.0 fs (velocity rescaling). Subsequently, a simulation interval of 3.5 ns using a Nosé–Hoover chain thermostat[62–64] with a time constant of 100 fs followed. After the first 50 ps of this second simulation interval, a Langevin thermostat[65,66] with a coupling constant of 100 fs was switched on for 1.0 ns in order to dampen acoustic shock waves in the cell. Finally, the production run was performed with the same thermostat settings for a

8756 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 8755–8766

This journal is © the Owner Societies 2023

simulation time of 200 ns. Every 2 ps, the position of each particle was stored to the trajectory. As in our previous work,[54] the Lennard-Jones cutoff radius was set to 800 pm, which is relatively short when compared to some other simulation studies, but was implicitly accounted for while fitting the parameters as indicated by the very good agreement to the reference AIMD simulation. Long-range electrostatic interactions were computed using a PPPM solver as implemented in LAMMPS.[60]

Based on these force field simulation trajectories, 50 pre-equilibrated initial configurations for the subsequent AIMD simulations were picked both for the pure and aqueous systems. The AIMD simulations were performed with the CP2k program package[67–69] using the Quickstep method[70] and orbital transformation (OT).[71] To describe the electronic structure, density functional theory (DFT) was applied,[72,73] using the BLYP functional[74,75] with the dispersion correction D3(BJ) of Grimme[76,77] together with the revised damping parameters of Smith *et al.*[78] The simulations were performed with a plane-wave energy cutoff of 350 Ry, atom-centered basis sets of type DZVP–MOLOPT–SR–GTH,[79] and GTH pseudopotentials.[80,81] Each of the individual simulations was running for approx. 20 ps for the pure system and approx. 24.5 ps for the aqueous system. The time step was set to 0.5 fs. The first 4 ps of the simulations of the pure system and 3 ps of the simulations of the aqueous system were discarded as equilibration. This resulted in a total simulation time of 729 ps for the pure cellulose–IL system and 1011 ps for the aqueous system. The simulations were performed at a temperature of 350 K, which is a frequently used temperature for experimental studies of the solubility of cellulose in ILs.[24,27,28] The temperature was kept at the target value using a Nosé–Hoover chain thermostat[62–64] with a coupling constant of 100 fs.

The AIMD simulations were finally used as a reference for optimizing the force field parameters. The force field simulations of each adjustment steps were performed as follows: one cellulose pentamer and 192 IL ion pairs for the pure system as well as one cellulose pentamer, 150 IL ion pairs and 450 water molecules for the aqueous system were placed in cubic simulation boxes of cell size 3715 pm and 3760 pm, respectively, at 350 K. Fig. 1 shows a snapshot of the simulation cell for cellulose in aqueous [EMIm][OAc]. By using a cellulose pentamer in the force field simulations instead of a cellulose trimer such as in the reference AIMD simulations, the statistical occurrence of the investigated hydrogen bonds could be improved with only a slight increase in computational cost. The monomer units were weighted according to their occurrence in the molecule to ensure a meaningful comparison of the cellulose trimer and pentamer for the comparison of the RDFs. Further information about this is given in Section 3. The force field simulations for the adjustment of the force field parameters were carried out analogously to the force field simulations in our already published article.[54] The simulation parameters are summarized in Table 1. However, for the force field development in this work, the length of the first pre-equilibration interval was increased to a physical simulation time of 25 ps. The production run amounted to 50 ns for both the pure and aqueous cellulose–IL systems.
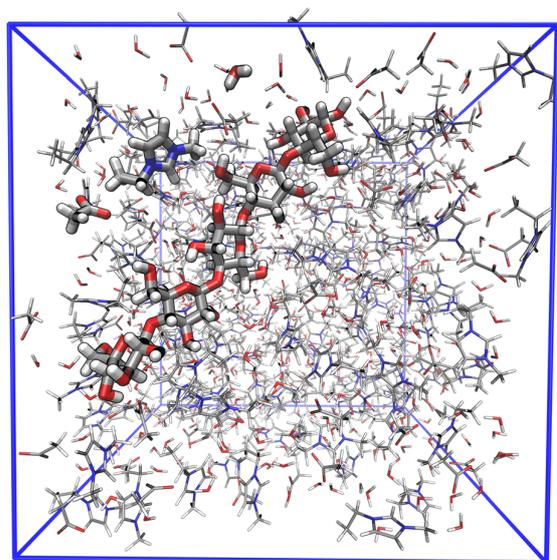
Fig. 1 Snapshot of the simulation cell of the cellulose pentamer in aqueous [EMIm][OAc] used for force field development.

The trajectories were analyzed using the TRAVIS program package.[82,83] For visualization, xmgrace[84] and Wolfram Mathematica[85] were used. The chemical structures and snapshots were generated using VMD[86] and the Tachyon renderer.[87]

3. Optimization procedure

The starting point for the development of the cellulose parameter set in BILFF was the cellulose force field of W. Damm.[57] Our previously published force field BILFF[54] was used for [EMIm][OAc] and TIP4P–EW[58] (with constrained bonds and angles) for water without any modifications. These force fields are based on the functional form of OPLS–AA:[88–90]

$$
\begin{aligned}
U(r^N) =\ & \sum_{i \in \text{bonds}} k_{l,i}(l_i - l_{i,0})^2 + \sum_{i \in \text{angles}} k_{\theta,i}(\theta_i - \theta_{i,0})^2 \\
& + \sum_{i \in \text{dihedrals}} \left[\frac{V_{i,1}}{2}[1 + \cos(\phi_i)] + \frac{V_{i,2}}{2}[1 - \cos(2\phi_i)] \right. \\
& \left. + \frac{V_{i,3}}{2}[1 + \cos(3\phi_i)] + \frac{V_{i,4}}{2}[1 - \cos(4\phi_i)] \right] \\
& + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[4\varepsilon_{ij}\left[\left(\frac{ij}{r_{ij}}\right)^{12} - \left(\frac{ij}{r_{ij}}\right)^6 \right] + \frac{q_i q_j e^2}{4\pi\varepsilon_0 r_{ij}} \right] f_{ij}.
\end{aligned}
$$

$$(1)$$

As described in section 2, a cellulose trimer was considered in the AIMD simulations and a cellulose pentamer in FFMD simulations. The labeling of the cellulose atom types is visualized in Fig. 2. The low computational cost of FFMD was utilized to further increase the statistical significance of the hydrogen bonds of cellulose in the systems to fit the force field parameters to the AIMD simulation: the single RDFs of the
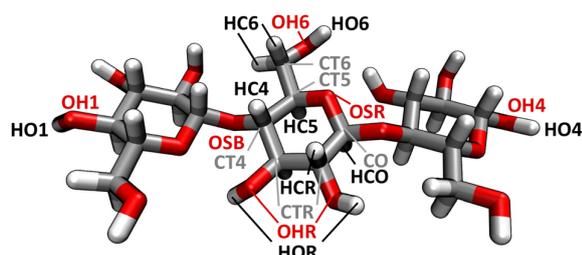
Fig. 2 Nomenclature of cellulose atom types in our force field BILFF.

hydrogen bonds of different glucose units were averaged and weighted according to their position in the molecule, thus creating one averaged RDF per atom type. As a first step, the RDFs of the two HOR protons (see Fig. 2) per monomer were averaged. In the second step, the resulting RDFs of different monomers were weighted relatively to each other. Each of the two terminal monomers was weighted by 12.5% and the middle monomer(s) by 75%, so that the results are comparable with the case of the pentamer with a weighting factor of 25% for each middle monomer unit. This ensured that the hydrogen bonds of the terminal glucose units in the strand play a small role in order to account for the behavior of a longer cellulose strand and thus a larger number of non-terminal glucose units in the central part. The resulting RDFs of the FFMD and AIMD simulations were compared to each other, forming a basis for the adjustment of the atomic partial charges and the Lennard-Jones Sigma parameters of the hydrogen bond donor and acceptor atoms. In order to optimize the parameters for the specific systems while preserving the properties of the initially used literature force field (as far as possible), the partial charges and Sigma values were iteratively adjusted until an acceptable reproduction of the reference AIMD simulations could be obtained. It should be noted that the use of an automated parameter adjustment procedure—such as, *e.g.*, force matching—is very difficult in the specific case due to the ragged high-dimensional penalty functional of the system which possesses a large number of local minima.

The magnitude of the parameter adjustments is shown in Section 3.1.1. When modifying the charges, the charge neutrality of the glucose monomers was maintained. The force field parameters for bonded interactions like bond lengths, angles, and dihedral angles were adapted to the average values from the AIMD simulations by changing the equilibrium values and force constants or the height of the potential barrier, respectively. All optimized force field parameters are presented in the ESI† (see Tables S1–S5).

As a consequence of the chemically identical structure of the individual monomers, the same atom types were assigned to all monomers. Within one monomer, some atoms were considered chemically equivalent caused by a nearly identical chemical environment and assigned to the same atom type (see Fig. 2). Due to the simultaneous force field development for the pure and aqueous systems, no further equivalences were assumed to ensure the best possible reproducibility of reference AIMD results. The two terminal hydroxyl groups

OH1–HO1 and OH4–HO4 were assigned to a separate atom type due to their unique location in the strand. A comparison of the dihedral angles between the FFMD and the AIMD simulation also showed that a splitting of the oxygen atom type OS as found in the literature force field[57] into the atom classes OSB for the bridging acetal oxygen atom and OSR for the ring oxygen atom is required to describe the intermolecular interactions more accurately (see Fig. S2 in the ESI†). This resulted in significant modifications of the bonded interactions of these atom classes compared to the literature force field.[57] Thus, new bond angles, lengths, and dihedral angles were defined for OSB and OSR and the corresponding values of the literature force field[57] for OS were adjusted: the force constants of the bonds of these affected atom types were varied by up to 41% and the force constants of the angles up to 30% in order to reproduce the AIMD results more accurately.

Within the scope of this work, only the force field parameters for cellulose were optimized. The parameters for (aqueous) [EMIm][OAc] have already been optimized and published in our first article on BILFF.[54] There, the total ion charge for [EMIm][OAc] was determined to be ±0.82. As shown there, charge screening effects can be taken into account by scaling down the total ionic charge of the ionic liquid. The optimized value of ionic charge was shown to be the most appropriate based on a comparison of diffusion coefficients between experimental data and the FFMD simulations using BILFF.[54] This downscaling of the ionic charge is a commonly applied procedure in the literature. Although there are also critical dissenting statements concerning a more accurate simulation of the structure of ionic liquids using polarizable force fields,[91] the charge reduction is sometimes considered an alternative to them.[50,92–95]

3.1. Microstructure of the system

3.1.1. Radial distribution functions. In the following, the reproducibility of the microstructure of both systems as seen in the reference AIMD simulation *via* FFMD simulations using the force field of W. Damm *et al.*[57] and BILFF is discussed. First, the hydrogen bond networks from the AIMD simulations are analyzed in more detail and are then compared with the results from a FFMD simulation with the literature force field[57] and the newly developed BILFF parameters.

The radial distribution functions (RDFs) of the individual hydrogen bonds were weighted and averaged according to Section 3, so that one RDF per atom type is obtained. Fig. 3 and 4 show a comparison of the RDFs between the cellulose protons HO6 and HOR to the acetate oxygen atoms of the reference AIMD simulation to an FFMD simulation using the force field of W. Damm *et al.*[57] and our new force field parameters. The RDFs from the AIMD simulations indicate strong hydrogen bonds of HO6 and HOR in the pure cellulose–[EMIm][OAc] system at a particle distance of about 160 pm, which are significantly weakened in the presence of water. The cause is the competition of the acetate oxygen atoms with the oxygen atom of water for the cellulose protons, which will be discussed in more detail below. Significant differences in the
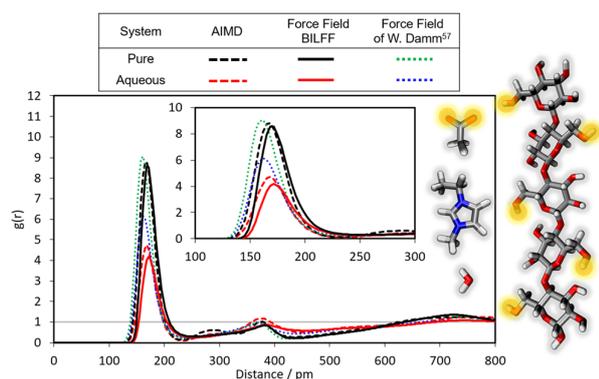
Fig. 3 Comparison of the RDF between the cellulose protons HO6 and the acetate oxygen atoms for a reference AIMD simulation and a force field MD simulation with the force field of W. Damm *et al.*[57] and our new force field BILFF.
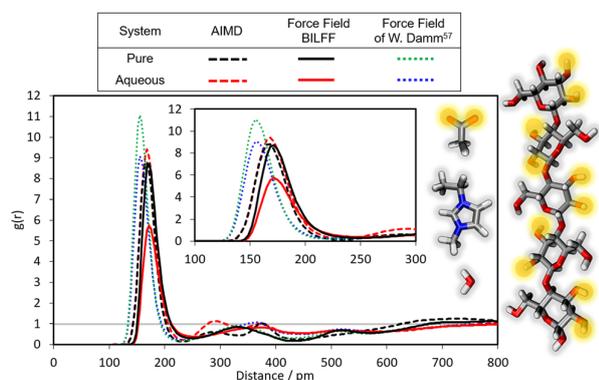
Fig. 4 Comparison of the RDF between the cellulose protons HOR and the acetate oxygen atoms for a reference AIMD simulation and a force field MD simulation with the force field of W. Damm *et al.*[57] and our new force field BILFF.

height and position of the first maxima in the RDFs between the AIMD and FFMD simulations in both systems become apparent for the literature force field,[57] indicating non-optimal partial charges and Sigma values of the cellulose protons. A reduction of the partial charges of the atom types HO6 by 2% from 0.4160 to 0.4067 (see Fig. 3) and HOR by 13% from 0.4280 to 0.3790 (see Fig. 4), as well as an increase of the partial charges of the neighboring oxygen atoms OH6 by 3% from -0.680 to -0.660 (see Fig. 3) and OHR by 2% from -0.6970 to -0.6850 (see Fig. 4) and the resulting shift of the charge distribution within the hydroxyl groups resulted in a significant improvement of the cellulose hydrogen bonds compared to the literature force field.[57] The overall charge neutrality of the single glucose units was maintained.

The too short hydrogen bond lengths in the literature force field[57] were compensated by increasing the repulsion between donor and acceptor *via* setting the Sigma values of HO6 to 0.97 Å and HOR to 1.00 Å, resulting in an excellent agreement with the RDFs of AIMD simulations in both pure and aqueous

cellulose–[EMIm][OAc] systems using BILFF (see Fig. 3 and 4). It should be noted here that Lennard-Jones parameters of zero were defined for the atom types HO6 and HOR in the literature force field of W. Damm *et al.*[57] and thus no Lennard-Jones interactions can be formed in that model, which are, however, essential for a correct description of the system.[43] The new Sigma values were iteratively developed by varying the parameters until an acceptable agreement with the RDFs of the reference AIMD was achieved. The values were validated by comparing the crystal structure of cellulose with experimental data. For more information, see Section 3.3.

After these adjustments, our new force field for cellulose shows an excellent agreement with the AIMD simulations in both pure and aqueous systems (see Fig. 3 and 4). Only the hydrogen bonds between HOR and the acetate oxygen atoms are shown to be slightly too weak in the aqueous system (see Fig. 4). Here, the second maximum of the RDF at larger particle distances indicates the interaction of the cellulose proton with the second acetate oxygen atom. Since both acetate oxygen atoms are assigned to the same atom types due to their chemical equivalence, a further optimization of this interaction is not possible here.

If we now compare the hydrogen bonds of HO6 and HOR in the reference AIMD simulations of the pure and aqueous cellulose–IL system with each other (see Fig. 3 and 4), it becomes apparent that in the pure system both protons form hydrogen bonds with the anion of the IL of approximately equal strength. Water weakens the hydrogen bond of HO6. HOR, on the other hand, is not affected, which is explained by the easier accessibility of HO6 by water.

In Fig. 5 and 6, RDFs are presented, which suggest that the influence of water on the cellulose–anion hydrogen bond is indeed caused by competing hydrogen bonds with water. The hydrogen bonds of the cellulose protons to the oxygen atom of water occur at a HO6–water distance of about 168 pm as well as a HOR–water distance of 171 pm in the reference AIMD simulation. The residence probability $g(r)$ and thus the strength of the hydrogen bond involving HO6 (see Fig. 5) is slightly
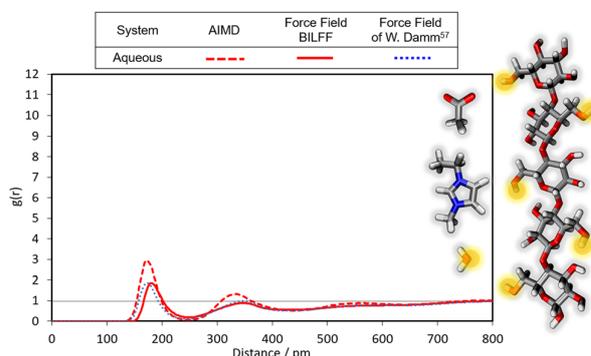
Fig. 5 Comparison of the RDF between the cellulose protons HO6 and the oxygen atom of water for a reference AIMD simulation and a force field MD simulation with the force field of W. Damm *et al.*[57] and our new force field BILFF.
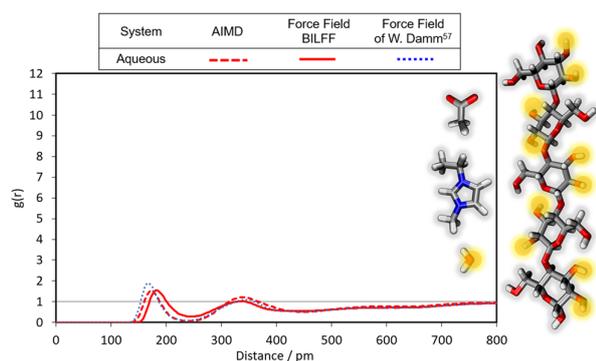
Fig. 6 Comparison of the RDF between the cellulose protons HOR and the oxygen atom of water for a reference AIMD simulation and a force field MD simulation with force field W. Damm *et al.*[57] and our new force field BILFF.

larger than the residence probability of HOR (see Fig. 6). This might be caused by a better accessibility of HO6 for hydrogen bond acceptors and its higher partial charge compared to HOR.

The "hydro-shielding" of HO6 also explains the reduced cellulose–IL hydrogen bond strength in the aqueous system compared to HOR. A comparison of the results of the cellulose–water hydrogen bonds from the reference AIMD simulation with the literature force field[57] and our new force field parameters in BILFF shows that BILFF can describe the HOR–water interaction better than the literature force field.[57] The HO6–water hydrogen bond, however, is slightly too weak in both force fields (see Fig. 5). Further adjusting this interaction would lead to a significant deterioration in agreement with the HO6–[OAc]⁻ interaction between AIMD and FFMD simulations, which is essential for cellulose solubility, so we decided to accept this compromise. The RDFs between the hydroxyl oxygen atoms of cellulose and the ring protons of the $[EMIm]^+$ cation as well as water are not discussed in detail here, since no significant hydrogen bonds could be found for that case.

3.1.2. Combined distance–angle distribution functions. For a more insightful description of the hydrogen bond geometry in a system, not only the hydrogen bond distance distribution is important, but also the angles at which the bond is formed. For this purpose, combined distance–angle distribution functions (CDFs) are considered, which establish a correlation between these two quantities. Fig. 7 and 8 show the distance–angle distribution functions between a proton of atom type HO6 and the oxygen atoms of acetate (see Fig. 7) and water (see Fig. 8) in the aqueous cellulose–IL system. To define the angle, connection vectors were defined between the cellulose hydroxyl proton of interest and the neighboring oxygen atom as well as between this cellulose proton and the oxygen atom(s) of acetate or water, so that a linear hydrogen bond would correspond to 180°. At an angle of about 137–180° and a particle distance of about 120–230 pm, the residence probability of the considered atoms HO6 and the acetate oxygen atoms possess a maximum in the AIMD simulation. The hydrogen bond between HO6 and the oxygen atom of water is recognizable in the AIMD simulations by a maximum between 135–180° and 130–240 pm (see Fig. 7). The
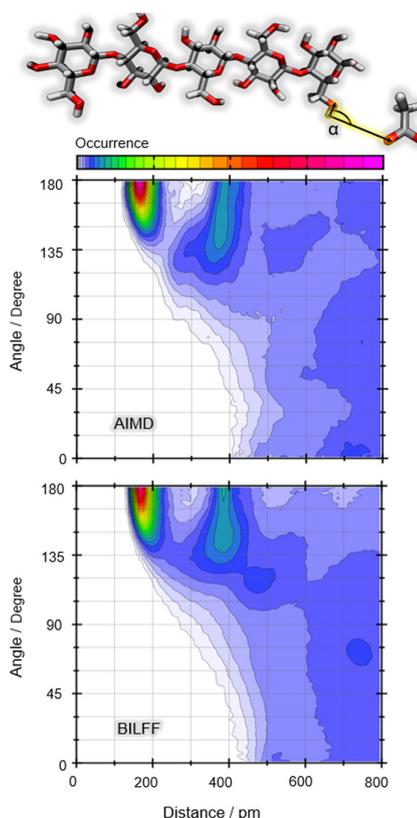
Fig. 7 Comparison of the distance–angle CDF between a proton of type HO6 of cellulose and the acetate oxygen atom in the aqueous system for a reference AIMD simulation and a force field MD simulation using the new force field parameters.

average hydrogen bond lengths are thus in agreement with the literature results.[26–28] The second maximum at a larger particle distance represents the residence probability of the second acetate oxygen atom or the water protons.

The figures also show a comparison of the results from the reference AIMD with the FFMD simulations using our new parameters. The FFMD can reproduce the results of the AIMD simulations quantitatively, thus validating the correctness of BILFF for these hydrogen bond geometries. It should be noted that no particular force field parameter is responsible for the shown distance–angle correlation according to eqn (1), and thus no direct optimization is possible.

3.1.3. Spatial distribution functions. Due to the formation of hydrogen bonds towards cellulose, a preferential spatial arrangement of the acceptor molecules around the protons of cellulose can be observed. In Fig. 9 and 10, spatial distribution functions (SDFs) based on the FFMD simulations with the new force field parameters are shown; averaged over all glucose monomers with a radius of the distribution function of 700 pm. The shown iso-values are indicated in the captions of the figures. Acetate (colored in green) accumulates close to the cellulose protons HO6 and HOR. In the pure cellulose–IL system, an arrangement of the $[EMIm]^+$ cation (colored in red) close to
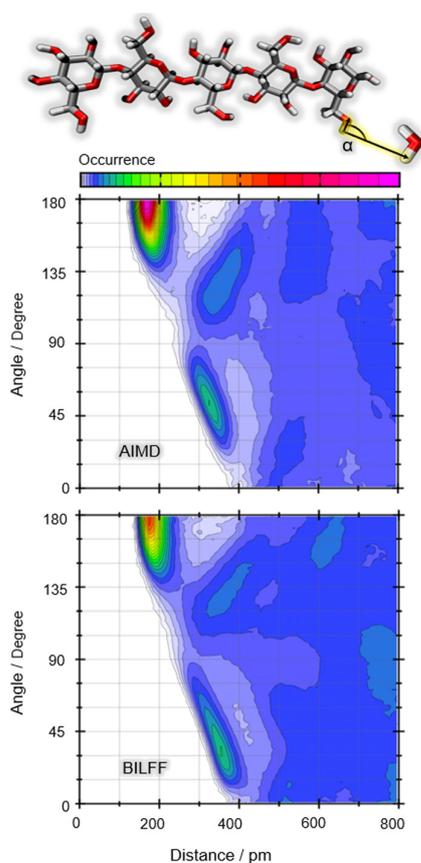
Fig. 8 Comparison of the distance–angle CDF between a proton of type HO6 of cellulose and the water oxygen atom in the aqueous system for a reference AIMD simulation and a force field MD simulation using the new force field parameters.

Fig. 10 Spatial distribution function (SDF) of the ring protons of [EMIm]$^+$ (red, 22 nm^{-3}), acetate (green, 12 nm^{-3}) and water (blue, 25 nm^{-3}) averaged over all glucose units in the cellulose–IL–water system calculated with BILFF.

spatial competition of the hydrogen acceptors as well as donors becomes obvious. Another interesting aspect is the ring-shaped arrangement of the cation and water around one HOR hydroxyl group, which occurs due to the free rotatability of the group. The results of the SDFs based on the AIMD simulations can be found in the ESI† (see Fig. S3 and S4).

3.2. Competing hydrogen bonds

In the previous considerations, mainly the hydrogen bonds between the protons of cellulose and the oxygen atoms of acetate and water were discussed. However, due to the large number of possible hydrogen bond donors and acceptors, various other competing hydrogen bonds can occur. One possible visualization for such a situation are Sankey diagrams, which can be computed by the TRAVIS package.[83] A Sankey diagram depicts the average hydrogen bond count—and thus also the hydrogen bond strength—for all pairs of hydrogen bond donors and acceptors in the mixture. The width of the

the ring oxygen atom and both hydroxyl oxygen atom types can also be observed. This indicates that the cellulose oxygen atoms also form hydrogen bonds to a smaller extent, but the surrounding acetate restricts the accessibility of the cellulose oxygen atoms for the [EMIm]$^+$ cation. Similar observations can be obtained in the aqueous system (see Fig. 10): water (colored in blue) arranges itself—as expected—in front of the cellulose hydroxyl protons and shields them against acetate and [EMIm]$^+$. Here, the direct

Fig. 9 Spatial distribution function (SDF) of the ring protons of [EMIm]$^+$ (red, 13 nm^{-3}) and the acetate oxygen atom (green, 14 nm^{-3}) averaged over all glucose units in the cellulose–IL system calculated with BILFF.

Fig. 11 Sankey diagram[83] illustrating hydrogen bonding in the pure cellulose–IL system calculated from the reference AIMD and a force field MD simulation using BILFF. Numbers indicate the average hydrogen bond count.

outgoing and incoming bars is proportional to the average number of hydrogen bonds formed by an individual donor or acceptor of that type (averaged over all molecules and the full simulation length). Fig. 11 shows the Sankey diagrams of the pure cellulose–IL system from the reference AIMD and the FFMD simulations using the new force field parameters. The hydrogen bond donors are shown on the left and the acceptors on the right hand side. The inserted numbers correspond to the average hydrogen bond count of the respective donor and acceptor atoms, which are determined from the number integral of the underlying RDFs up to the first minimum (*i.e.*, coordination numbers). Here, the coordination numbers calculated from the reference AIMD were compared with the results from the force field MD using our new force field parameters. It can be observed that there is a competition between the protons of cellulose and [EMIm]$^+$ cation for the oxygen atoms of acetate, which is coordinated by about four to five protons. However, cellulose is also known to form non-negligible intramolecular hydrogen bonds,[96] which are also formed in our simulations. A comparison of the coordination numbers shows that the interactions between the molecules as well as the coordination numbers from the AIMD simulations can be reproduced very well by our force field parameters.

A more detailed investigation of the intramolecular hydrogen bonds in cellulose is enabled in a contact matrix as computed by the TRAVIS package.[83] For this purpose, the height of the first maximum of the underlying radial distribution functions of the observed atoms $g(r)$ is color-coded in an acceptor–donor matrix.[83] It should be noted that the peak heights in $g(r)$ do not directly correspond to a pairwise probability distribution—which can rather be obtained as the product of $g(r)$ and the density of the respective observed atoms. Since the underlying RDFs are all normalized to the uniform density of the corresponding atom types, this was implicitly taken into account.

Fig. 12 shows all possible hydrogen bonds in the pure cellulose–[EMIm][OAc] system. The numbering of the cellulose atoms in the contact matrices is shown in Fig. S8 in the ESI.†

The protons are plotted against the oxygen atoms of the observed molecules. The strongest hydrogen bonds in the system, colored in red, are mainly formed between cellulose hydroxyl protons and the acetate oxygen atoms as well as between the ring protons H1,4,5 of [EMIm]$^+$ and the cellulose hydroxyl oxygen atoms. The ring and bridging oxygen atoms play a subordinate role, as expected, due to their shielding from surrounding protons and electron-donating CH$_2$ groups. The protons of the methyl and ethyl group of [EMIm]$^+$ don't form hydrogen bonds with cellulose. In some cases, intramolecular hydrogen bonds between the hydroxyl groups of cellulose occur. The strongest ones are formed by the protons of the atom type HO6, which have a higher freedom of rotation due to their position on an additional carbon. The same tendencies are also seen in the aqueous system, where the protons of the atom type HOR also form strong intramolecular hydrogen bonds (see Fig. 13). Meanwhile, the oxygen atoms of acetate and water compete for the protons HO6 and HOR, confirming the previous results. Also, the protons of water are in competition with the ring protons of [EMIm]$^+$ for the hydroxyl oxygen
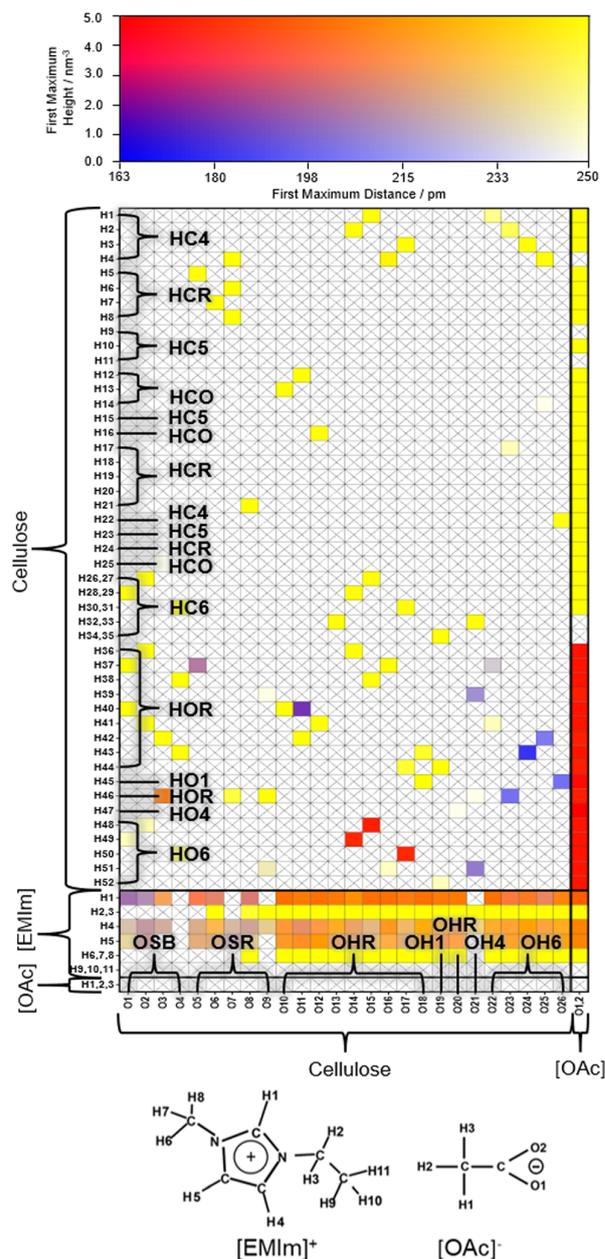
Fig. 12 Contact matrix[83] illustrating hydrogen bonding in the pure cellulose–IL system from the force field simulation with BILFF. Hydrogen bond donors on vertical axis, acceptors on the horizontal axis. The atom types of cellulose are noted in the matrix and the atom numbers of the relevant atoms of the ionic liquid are shown below.

atoms of cellulose. The non-identical behavior of chemically equivalent atoms can be attributed to insufficient statistical sampling.

3.3. Validation with respect to the experiment

To validate our new force field for cellulose, we compare the calculated system density from simulations using our force field BILFF with extrapolated experimental data at 350 K and

8762 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 8755–8766

This journal is © the Owner Societies 2023
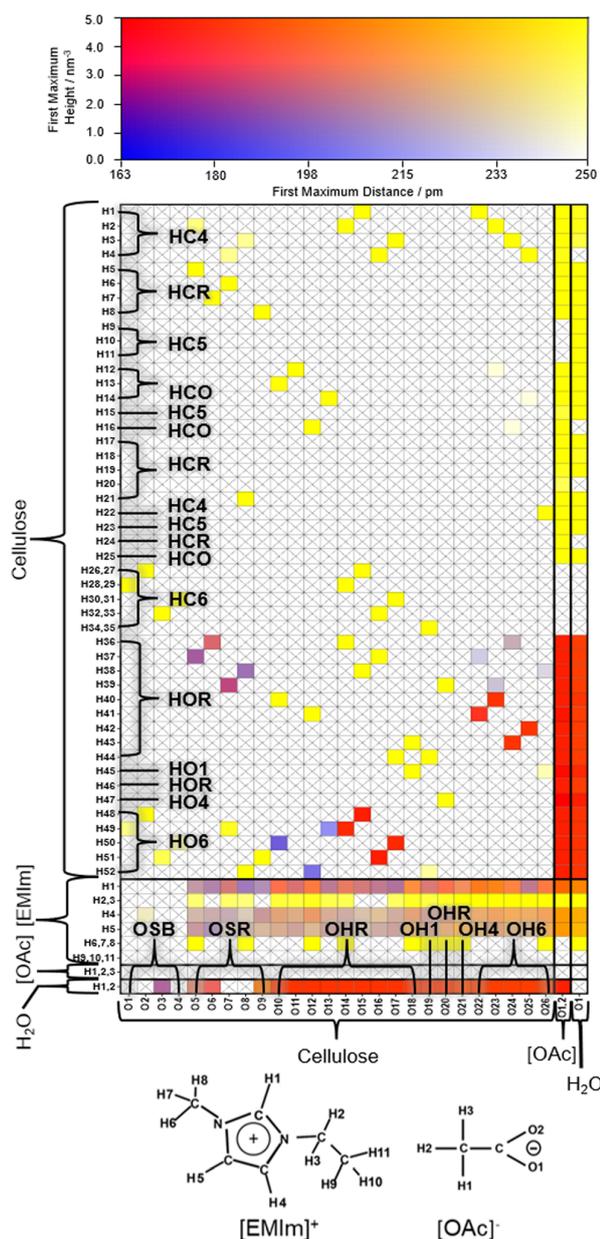
Fig. 13 Contact matrix[83] illustrating hydrogen bonding in the aqueous cellulose–IL system from the force field simulation with BILFF. The atom types of cellulose are noted in the matrix and the atom numbers of the relevant atoms of the ionic liquid are shown below.

Table 2 Comparison of the system density from the force field MD using BILFF of cellulose in pure and aqueous [EMIm][OAc] with extrapolated experimental values[97] at different temperatures

Temp.		Calculation		Experiment[97]
		BILFF	Damm[57]	
Pure solvent				
350 K	Cell size/pm	3715	3714	
	ρ/g cm^{-3}	1.085	1.086	1.072
550 K	Cell size/pm	3926	3924	
	ρ/g cm^{-3}	0.920	0.921	0.972
Aqueous solvent				
350 K	Cell size/pm	3760	3759	
	ρ/g cm^{-3}	1.077	1.077	—
550 K	Cell size/pm	4 047	4045	
	ρ/g cm^{-3}	0.863	0.865	—

be valid in a range of 25–120 °C and that deviations are possible at a temperature of 550 K, so that the nevertheless good agreement with the calculated density is remarkable.

For further validation, we calculated the crystal structure of pure cellulose (see Fig. 14) and compared it with experimental XRD data here. To do so, an experimentally determined crystal structure[98] was used as a starting point for a cell optimization calculation with our new force field using LAMMPS.[60] The protons were placed at the most frequent crystal positions. The unit cell was replicated along the three cell vectors *A*, *B*, and *C* by factors 3, 3, and 2, respectively. The crystal parameters are shown in Table 3. The calculated lattice constants are in outstanding agreement with the experimental literature data. The cell density shows a deviation of only 0.2%. The calculated lattice constant *A* is marginally too large, while the lattice constant *C* has a slightly too small value and compensates the small deviations.

Both the density of the pure and aqueous cellulose–[EMIm][OAc] system (see Tables 2 and 3) and the calculated cellulose crystal parameters are very similar in the force field
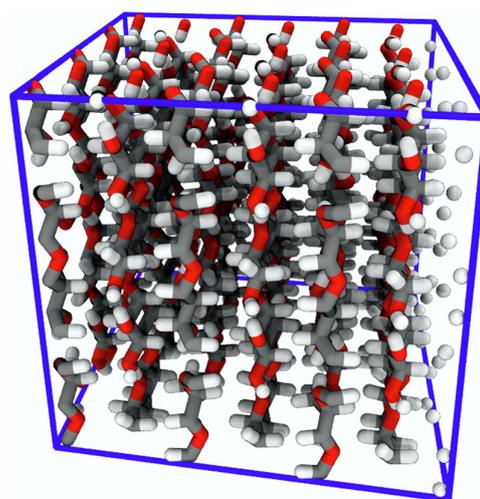
Fig. 14 Snapshot of the optimized cellulose crystal structure using force field BILFF.

550 K, whereas BILFF has been developed at a temperature of 350 K (see Table 2). The extrapolation of density is based on the experimental observations of J. Schuermann *et al.* who found a linear density/temperature correlation for cellulose dissolved in [EMIm][OAc] and characterized the correlation as $\rho_{[\text{EMIm}][\text{OAc}]} = -0.0005T + 1.1104$, where temperature *T* is given in units of °C.[97] The simulated system densities differ from the extrapolated literature data by only 1.2% at 350 K and by 5.7% at 550 K. It is worth mentioning that the empirical equation was found to

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 8755–8766 | **8763**

Table 3 Cellulose crystal parameters computed from force field cell optimizations using BILFF and the force field of Damm[57] as well as experimental XRD data.[98] The unit cell was replicated by factors 3, 3, and 2 along the vectors *A*, *B*, and *C*, respectively

	Calculation		
	Damm[57]	BILFF	Experiment[98]
Lattice constant A/pm	2235.4	2261.8	2335.(2)
Lattice constant B/pm	2443.4	2482.2	2460.(3)
Lattice constant C/pm	2088.7	2123.0	2076.(0)
Cell angle α/°	90.0	90.0	90.0
Cell angle β/°	90.0	90.0	90.0
Cell angle γ/°	96.9	97.0	96.5
Cell volume/nm^3	11.326	11.830	11.849
Cell density/g cm^{-3}	1.744	1.639	1.635

simulations using the literature force field[57] and BILFF. Our new force field thus allows the reproduction of an AIMD simulation of cellulose in pure and aqueous [EMIm][OAc] with a more accurate representation of the microstructure (see Section 3.1.1) while at the same time calculating the system density and crystal parameters as accurately as simulations with the force field of W. Damm.[57]

4. Conclusions

In this manuscript, we present the extension of our all-atom force field BILFF (Bio-Polymers in Ionic Liquids Force Field), already published for mixtures of the ionic liquid [EMIm][OAc] with water,[54] to the bio-molecule cellulose. The force field focuses on a quantitative reproduction of the hydrogen bonds as observed in reference AIMD simulations for the complex mixture of cellulose, [EMIm]$^+$, [OAc]$^-$ and water. For this purpose, the radial and the combined distance–angle distribution functions (RDFs and CDFs, respectively) of various hydrogen bonds of cellulose as both donor and acceptor as well as the statistical occurrence of bond lengths, angles, and dihedral angles were compared with the results of the reference AIMD simulation. To make use of the lower computational costs of the force field molecular dynamics simulations and to better reflect the statistical occurrence of hydrogen bonds, a cellulose strand consisting of five β-D-glucose monomers was used in the force field simulations and a strand with three β-D-glucose monomers in the more time-consuming reference AIMD simulation. To increase the physical simulation time and enhance the sampling, 50 single AIMD simulations starting from different initial configurations were performed for cellulose in pure and aqueous [EMIm][OAc] at a temperature of 350 K, respectively, and the RDFs and CDFs were calculated from the total trajectory. Based on the comparison of AIMD with the force field MD simulations, all force field parameters were iteratively adjusted starting from the force field of W. Damm,[57] and a very good agreement was obtained both for the microstructure of the reference AIMD simulations in pure and aqueous [EMIm][OAc] and experimental results such as the system density even at higher temperatures and XRD data of the cellulose crystal structure.

BILFF thus enables an accurate reproduction of the data from AIMD simulations with a focus on the complex hydrogen bond network found in the quaternary mixture and combines high accuracy with low simulation costs. This allows for long simulations of both the dissolution and precipitation processes of cellulose in [EMIm][OAc] and [EMIm][OAc] with water, which can take hours in experimental studies, and can provide insights into the amphiphilic character of cellulose. The development of BILFF is not complete, and further solvents, cosolvents, and bio-polymers will be parametrized and published during our ongoing work.

Author contributions

E. Roos was responsible for data curation, data analysis, visualization of the results, and the writing of the original manuscript draft. M. Brehm was responsible for the choice of computational methods and for supervising the progress of the data processing. D. Sebastiani was responsible for supervising the work and the conceptualization of the project. All authors contributed to the analysis and interpretation of the results, as well as further review iterations of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 M. Iguchi, S. Yamanaka and A. Budhiono, *J. Mater. Sci.*, 2000, **35**, 261–270.
2 C. Olsson and G. Westman, *Cellul.: Fundam. Aspects*, 2013, 143–177.
3 M. M. Abeer, M. C. I. Mohd Amin and C. Martin, *J. Pharm. Pharmacol.*, 2014, **66**, 1047–1061.
4 K. C. Badgujar and B. M. Bhanage, *Bioresour. Technol.*, 2015, **178**, 2–18.
5 L. Petrus and M. A. Noordermeer, *Green Chem.*, 2006, **8**, 861–867.
6 Y. Sun and J. Cheng, *Bioresour. Technol.*, 2002, **83**, 1–11.
7 M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady and T. D. Foust, *Science*, 2007, **315**, 804–807.
8 X. Sun, Q. Wu, S. Lee, Y. Qing and Y. Wu, *Sci. Rep.*, 2016, **6**, 31654.
9 J. A. Kelly, A. M. Shukaliak, C. C. Y. Cheung, K. E. Shopsowitz, W. Y. Hamad and M. J. MacLachlan, *Angew. Chem., Int. Ed.*, 2013, **52**, 8912–8916.

8764 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 8755–8766

This journal is © the Owner Societies 2023

10 B. Thomas, C. R. Midhun, K. B. Athira, M. H. Rubiyah, J. Jithin, M. Audrey, L. D. Glenna and C. Sanchez, *Chem. Rev.*, 2018, **118**, 11575–11625.

11 T. S. Anokhina, T. S. Pleshivtseva, V. Y. Ignatenko, S. V. Antonov and A. V. Volkov, *Pet. Chem.*, 2017, **57**, 477–482.

12 W. Ge, J. Shuai, Y. Wang, Y. Zhou and X. Wang, *Polym. Chem.*, 2022, **13**, 359–372.

13 Z. M. Judeh, C. B. Ching, J. Bu and A. McCluskey, *Tetrahedron Lett.*, 2002, **43**, 5089–5091.

14 L. A. Blanchard, D. Hancu, E. J. Beckman and J. F. Brennecke, *Nature*, 1999, **399**, 28–29.

15 T. Welton, *Chem. Rev.*, 1999, **99**, 2071–2084.

16 R. Sheldon, *Chem. Commun.*, 2001, 2399–2407.

17 F. Malberg, M. Brehm, O. Hollóczki, A. S. Pensado and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18424–18436.

18 K. Ghandi, *Green Sustainable Chem.*, 2014, **04**, 44–53.

19 M. Brehm, H. Weber, A. S. Pensado, A. Stark and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2012, **14**, 5030–5044.

20 M. Brehm, H. Weber, A. S. Pensado, A. Stark and B. Kirchner, *Z. Phys. Chem.*, 2013, **227**, 177–204.

21 Y. Chen, Y. Cao, X. Sun and T. Mu, *J. Mol. Liq.*, 2014, **190**, 151–158.

22 M. Brehm and D. Sebastiani, *J. Chem. Phys.*, 2018, **148**, 193802.

23 A. Sharma, M. Thakur, M. Bhattacharya, T. Mandal and S. Goswami, *Biotechnol. Rep.*, 2019, **21**, e00316.

24 R. P. Swatloski, S. K. Spear, J. D. Holbrey and R. D. Rogers, *J. Am. Chem. Soc.*, 2002, **124**, 4974–4975.

25 Y. Tomimatsu, Y. Yoshimura and A. Shimizu, *Aust. J. Chem.*, 2019, **72**, 669–673.

26 B. Manna and A. Ghosh, *J. Biomol. Struct. Dyn.*, 2019, **37**, 3987–4005.

27 M. Brehm, M. Pulst, J. Kressler and D. Sebastiani, *J. Phys. Chem. B*, 2019, **123**, 3994–4003.

28 M. Brehm, J. Radicke, M. Pulst, F. Shaabani, D. Sebastiani and J. Kressler, *Molecules*, 2020, **25**, 3539.

29 B. Lu, A. Xu and J. Wang, *Green Chem.*, 2014, **16**, 1326–1335.

30 B. Lindman, B. Medronho, L. Alves, C. Costa, H. Edlund and M. Norgren, *Phys. Chem. Chem. Phys.*, 2017, **19**, 23704–23718.

31 A. Xu, Y. Zhang, Y. Zhao and J. Wang, *Carbohydr. Polym.*, 2013, **92**, 540–544.

32 J. Zhang, H. Zhang, J. Wu, J. Zhang, J. He and J. Xiang, *Phys. Chem. Chem. Phys.*, 2010, **12**, 1941–1947.

33 T. G. A. Youngs, J. D. Holbrey, C. L. Mullan, S. E. Norman, M. C. Lagunas, C. D'Agostino, M. D. Mantle, L. F. Gladden, D. T. Bowron and C. Hardacre, *Chem. Sci.*, 2011, **2**, 1594–1605.

34 T. Endo, S. Hosomi, S. Fujii, K. Ninomiya and K. Takahashi, *Molecules*, 2017, **22**, 178.

35 C. Olsson, A. Idström, L. Nordstierna and G. Westman, *Carbohydr. Polym.*, 2014, **99**, 438–446.

36 H. Liu, K. L. Sale, B. M. Holmes, B. A. Simmons and S. Singh, *J. Phys. Chem. B*, 2010, **114**, 4293–4301.

37 Y. Li, X. Liu, S. Zhang, Y. Yao, X. Yao, J. Xu and X. Lu, *Phys. Chem. Chem. Phys.*, 2015, **17**, 17894–17905.

38 Z.-D. Ding, Z. Chi, W.-X. Gu, S.-M. Gu, J.-H. Liu and H.-J. Wang, *Carbohydr. Polym.*, 2012, **89**, 7–16.

39 Y.-R. Liu, K. Thomsen, Y. Nie, S.-J. Zhang and A. S. Meyer, *Green Chem.*, 2016, **18**, 6246–6254.

40 S. Gehrke, K. Schmitz and O. Hollóczki, *J. Phys. Chem. B*, 2017, **121**, 4521–4529.

41 T. Watermann and D. Sebastiani, *Z. Phys. Chem.*, 2018, **232**, 989–1002.

42 T. Ishida, *J. Phys. Chem. B*, 2020, **124**, 3090–3102.

43 P. Griffin, S. Ramer, M. Winfough and J. Kostal, *Green Chem.*, 2020, **22**, 3626–3637.

44 J. Blasius, R. Elfgen, O. Hollóczki and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2020, **22**, 10726–10737.

45 Z. Liu, S. Huang and W. Wang, *J. Phys. Chem. B*, 2004, **108**, 12978–12989.

46 X. Zhong, Z. Liu and D. Cao, *J. Phys. Chem. B*, 2011, **115**, 10027–10040.

47 J. N. Canongia Lopes, J. Deschamps and A. A. H. Pádua, *J. Phys. Chem. B*, 2004, **108**, 2038–2047.

48 J. N. Canongia Lopes and A. A. H. Pádua, *J. Phys. Chem. B*, 2006, **110**, 19586–19592.

49 J. N. Canongia Lopes and A. A. H. Pádua, *Theor. Chem. Acc.*, 2012, **131**, 3330.

50 B. Doherty, X. Zhong, S. Gathiaka, B. Li and O. Acevedo, *J. Chem. Theory Comput.*, 2017, **13**, 6131–6145.

51 K. Goloviznina, J. N. Canongia Lopes, M. Costa Gomes and A. A. H. Pádua, *J. Chem. Theory Comput.*, 2019, **15**, 5858–5871.

52 B. Doherty, X. Zhong and O. Acevedo, *J. Phys. Chem. B*, 2018, **122**, 2962–2974.

53 D. Bedrov, J.-P. Piquemal, O. Borodin, A. D. MacKerell, B. Roux and C. Schröder, *Chem. Rev.*, 2019, **119**, 7940–7995.

54 E. Roos and M. Brehm, *Phys. Chem. Chem. Phys.*, 2021, **23**, 1242–1253.

55 N. Sun, H. Rodríguez, M. Rahman and R. D. Rogers, *Chem. Commun.*, 2011, **47**, 1405–1421.

56 H. Wang, G. Gurau and R. D. Rogers, *Chem. Soc. Rev.*, 2012, **41**, 1519–1537.

57 W. Damm, A. Frontera, J. Tirado-Rives and W. L. Jorgensen, *J. Comput. Chem.*, 1997, **18**, 1955–1970.

58 H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, **120**, 9665–9678.

59 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.

60 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.

61 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

62 S. Nose, *J. Chem. Phys.*, 1984, **81**, 511–519.

63 S. Nose, *Mol. Phys.*, 1984, **52**, 255–268.

64 G. Martyna, M. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2535–2643.

65 T. Schneider and E. Stoll, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1978, **17**, 1302–1322.

66 B. Dünweg and W. Paul, *Int. J. Mod. Phys. C*, 1991, **02**, 817–827.

67 The CP2K developers group, 2017. CP2K is freely available from **http://www.cp2k.org/**.

68 J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *WIREs Comput. Mol. Sci.*, 2014, **4**, 15–25.

69 T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack and J. Hutter, *J. Chem. Phys.*, 2020, **152**, 194103.

70 J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Commun.*, 2005, **167**, 103–128.

71 J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2003, **118**, 4365–4369.

72 P. Hohenberg and W. Kohn, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1964, **136**, 864.

73 W. Kohn and L. Sham, *Phys. Rev.*, 1965, **140**, 1133.

74 A. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098–3100.

75 C. Lee, W. Yang and R. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.

76 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

77 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

78 D. G. A. Smith, L. A. Burns, K. Patkowski and C. D. Sherrill, *J. Phys. Chem. Lett.*, 2016, **7**, 2197–2203.

79 J. VandeVondele and J. Hutter, *J. Chem. Phys.*, 2007, **127**, 114105.

80 S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 1703–1710.

81 C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 3641–3662.

82 M. Brehm and B. Kirchner, *J. Chem. Inf. Model.*, 2011, **51**, 2007–2023.

83 M. Brehm, M. Thomas, S. Gehrke and B. Kirchner, *J. Chem. Phys.*, 2020, **152**, 164105.

84 Grace, (c) 1996-2008 Grace Development Team, see **http://plasma-gate.weizmann.ac.il/Grace**.

85 *Mathematica, Version 8.0*, Wolfram Research, Inc., Champaign, IL, 2010.

86 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.

87 J. Stone, MSc thesis, Computer Science Department, University of Missouri-Rolla, 1998.

88 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

89 J. W. Ponder and D. A. Case, *Adv. Protein Chem.*, 2003, **66**, 27–85.

90 S. V. Sambasivarao and O. Acevedo, *J. Chem. Theory Comput.*, 2009, **5**, 1038–1050.

91 J. G. McDaniel and A. Yethiraj, *J. Phys. Chem. Lett.*, 2018, **9**, 4765–4770.

92 X. Wu, Z. Liu, S. Huang and W. Wang, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2771–2779.

93 B. L. Bhargava and S. Balasubramanian, *J. Chem. Phys.*, 2007, **127**, 114510.

94 H. Liu and E. Maginn, *J. Chem. Phys.*, 2011, **135**, 124507.

95 P. Ray, R. Elfgen and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2019, **21**, 4472–4486.

96 M. Wohlert, T. Benselfelt, L. Wagberg, I. Furo, L. A. Berglund and J. Wohlert, *Cellulose*, 2022, **29**, 1–23.

97 J. Schuermann, T. Huber, D. LeCorre, G. Mortha, M. Sellier, B. Duchemin and M. P. Staiger, *Cellulose*, 2016, **23**, 1043–1050.

98 Y. Nishiyama, P. Langan and H. Chanzy, *J. Am. Chem. Soc.*, 2002, **124**, 9074–9082.

8766 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 8755–8766

This journal is © the Owner Societies 2023

3.4   Article III: A Force Field for the Solubility of Cellulose in DMSO/Ionic Liquids

Author contributions: In this article, I conducted the simulations as well as the calculations in Sections 3.1, 3.2, and 3.3.3. I was responsible for data processing, analysis, and visualization of the results, as well as drafting the original article. C. Gradaus performed the calculations of the hydrogen bond lifetime, diffusion coefficients and RDFs in Section 3.3 as well as writing text passages on the density and diffusion coefficient in Section 3.3.3. M. Brehm was responsible for the choice of computational methods and for supervising the progress of the data processing. D. Sebastiani was responsible for supervising the work and the conceptualization of the project. All authors contributed to further review iterations of the article.

ORIGINAL RESEARCH

A force field for the solubility of cellulose in DMSO/Ionic liquids

**Eliane Roos · Cora Gradaus ·
Daniel Sebastiani · Martin Brehm**

Abstract We present the extension of our all-atom force field BILFF (*B*io-polymers in *I*onic *L*iquids *F*orce *F*ield) to the co-solvent dimethyl sulfoxide (DMSO). BILFF already includes force field parameters for several imidazolium- and triazolium-based ionic liquids, water, and the bio-molecule cellulose. DMSO is known to increase the cellulose solubility of [EMIm][OAc] when applied as an additive. Our focus is on a correct reproduction of the hydrogen bonds in the ternary system of $[EMIm]^+$, $[OAc]^-$, water and DMSO compared to *ab initio* molecular dynamics simulations. For this purpose, the results from force field MD simulations such as radial and distance–angle distribution functions are compared with the reference AIMD simulation. Based on this, the force field parameters (starting from OPLS–AA) are iteratively adjusted. Four systems, pure and aqueous DMSO as well as DMSO in pure and aqueous [EMIm][OAc], are considered and additionally compared to DMSO-free [EMIm][OAc] systems. A very good agreement with respect to the microstruc-

ture of the quantum chemical reference simulations as well as to experimental data such as density, diffusion coefficients, enthalpy of vaporization, compressibility and thermal expansion coefficients can be observed over a wide temperature range. BILFF thus enables accurate simulations of larger systems of solvated cellulose in (aqueous) [EMIm][OAc] and the co-solvent DMSO.

Keywords DMSO · Force field molecular dynamics simulation · Cellulose Co-solvent · Ionic liquid · Solvent structure · Hydrogen bond lifetime · Calculated volume expansion coefficient · Calculated compressibility

Introduction

Cellulose is recognized as a nearly limitless resource to fulfill the growing need for environmentally friendly and biocompatible products (Azimi et al. 2022, Klemm et al. 2005, Nanda et al. 2022, Payne et al. 2015). It has diverse applications, spanning from packaging materials and writing surfaces to biofuels (Himmel et al. 2007), bioplastics (Fang et al. 2023), anti fungal membranes (Wanichapichart et al. 2012), laminates (Yousefi Shivyari et al. 2016), and phosphorescent materials (Zhang et al. 2022), among other innovative uses (Abeer et al. 2014, Sun and Cheng 2002, Thomas et al. 2018). Some processes require the dissolution of cellulose. However, the solubility of cel-

E. Roos · C. Gradaus · D. Sebastiani · M. Brehm (✉)
Institut für Chemie – Theoretische Chemie, Martin-Luther-Universität Halle–Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany
Martin_Brehm@gmx.de

M. Brehm
Department Chemie, Universität Paderborn, Warburger Straße 100, 33098 Paderborn, Germany

lulose in conventional solvents is a challenge due to its extensive inter- and intramolecular hydrogen bonding. Industrially used derivatizing solvents such as carbon disulfide (CS$_2$) or N-methylmorpholine N-oxide (NMMO) have several disadvantages such as toxicity, insufficient solvent stability, degradation of both cellulose and solvent and significant stabilizer requirements (Azimi et al. 2022, Olsson and Westman 2013). A direct dissolution of raw cellulose not only simplifies the processing but also circumvents byproduct generation. Therefore, there is a strong interest in discovering environmentally friendly and non-toxic solvents capable of dissolving cellulose without requiring derivatization. In 2002, Swatloski discovered that such a cellulose solvent are ionic liquids (ILs) (Swatloski et al. 2002). A technically relevant cellulose solubility is achieved, for example, in the IL 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]), which has become the focus of research due to its additional high chemical stability (Froschauer et al. 2013, Le et al. 2012). However, while IL-based methods for processing biomass show promise for large-scale applications, certain ILs can be expensive and have a high viscosity, creating challenges for scaling up (Li et al. 2013). A combination of an IL and polar organic co-solvents, such as dimethyl sulfoxide (DMSO) exhibits a lower viscosity than the pure IL and improves the solubility of cellulose (Anokhina et al. 2017, Mohan et al. 2016, Tomimatsu et al. 2019, Xu et al. 2013, Yang et al. 2019b). However, it is still debated how precisely DMSO enhances the cellulose solubility, while water, also a polar molecule, reduces it: To dissolve cellulose, it is necessary to replace the inter- and intramolecular cellulose hydrogen bonds with solvent–cellulose hydrogen bonds formed primarily by the anion of the IL. However, these interactions are hindered by a strong association between the cation and the anion (Brehm et al. 2019, 2020a). A common theory is that DMSO separates the ions from each other and, by additionally reducing the viscosity and thereby increasing the mass transfer rate (Le et al. 2014), promotes the formation of anion–cellulose hydrogen bonds (Anokhina et al. 2017, Brehm et al. 2019, 2020a, Manna and Ghosh 2019, Mohan et al. 2016, Paiva et al. 2022, Tomimatsu et al. 2019).

Although several studies on cellulose in [EMIm][OAc] and DMSO have been published, the influence of water on the underlying hydrogen bond network has hardly been investigated. However, in view

of the fact that the IL is hygroscopic and that cellulose reprecipitates even in the presence of small amounts of water (Froschauer et al. 2013, Le et al. 2012), the importance of water should not be ignored. To gain a thorough understanding of the intricate interactions present within the cellulose/IL/DMSO/water system at the atomic level, molecular dynamic (MD) simulations can be utilized. However, despite its accuracy, quantum chemical techniques are impractical for studying these large systems due to their high computational demand and time-consuming nature. To overcome these limitations, force fields can be utilized as an empirical alternative for representing the interactions. A force field describes the interactions between molecules in a simplified way by representing the total potential energy as the sum of the contributions from bonds, bond angles, torsion angles, and electrostatic and van der Waals energies. These energy components are parameterized using empirical functions, allowing a force field to focus on specific application fields.

There are already several force field parameter sets available for the molecule DMSO, which have been optimized to accurately describe the thermodynamic properties of pure DMSO and aqueous DMSO (Bordat et al. 2003, Fox and Kollman 1998, Geerke et al. 2004, Strader and Feller 2002). However, there is currently no force field specifically designed to accurately describe the microstructure between an ionic liquid, particularly [EMIm][OAc], water, and DMSO, as well as the hydrogen bonds between them. Inaccurate modeling of these directed interactions may lead to an imprecise description of the delicate hydrogen bond network. Therefore, we have developed optimized force field parameters for DMSO in this complex mixture to achieve the level of accuracy comparable to quantum chemical MD simulations (*ab initio* MD (AIMD) simulations).

For the investigation of the bio-molecule solution process with force field simulations, we have already published optimized force field parameters for the ionic liquids 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) (Roos and Brehm 2021), 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]), 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]), 1-ethyl-3-methylimidazolium benzoate ([EMIm][OBz]) (Roos et al. 2023a), water, and cellulose. This article presents the extension of the force field BILFF (Roos and Brehm 2021, Roos et al. 2023b, a) (*B*io-Polymers in *I*onic *L*iquids *F*orce *F*ield) to the co-solvent DMSO. For this purpose, we compared vari-

ous analyses regarding the microstructure and dynamics of the system from force field MD simulations with the results of reference quantum chemical MD simulations and minimized the deviations by varying the force field parameters. This allowed us to develop an optimal parameter set with a focus on the hydrogen bonds of the system simultaneously in anhydrous and aqueous [EMIm][OAc]. For validation, further simulations of pure and aqueous DMSO in the absence of the ionic liquid were calculated and the system densities, self-diffusion coefficients, thermal volume expansion coefficient, and the compressibility were compared with experimental data. In addition, the reproducibility of the quantum chemical results at elevated simulation temperature was investigated. Furthermore, the lifetime of the investigated hydrogen bonds, densities and self-diffusion coefficients were compared with the results from force field simulations of DMSO-free [EMIm][OAc] and thus the influence of DMSO on the system was discussed.

Computational details

The force field parameters for DMSO in mixtures with [EMIm][OAc] were developed simultaneously in both the presence and absence of water at 350 K. The force field parameters were iteratively adjusted by comparing various analyses of the force field MD simulation with reference AIMD simulations (see "Optimization procedure"). The initial configurations for these AIMD simulations were generated from the final configuration of a 20 ns NpT force field simulation with the force field parameters of OPLS–AA (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009) for DMSO, BILFF (Roos and Brehm 2021, Roos et al. 2023b, a) for [EMIm][OAc], and TIP4P–EW (Horn et al. 2004) (with constrained bonds and angles using the RATTLE algorithm (Andersen 1983, Ryckaert et al. 1977) for water. The simulation parameters such as the number of molecules, the size of the simulation cell and the final system density are listed in Table 1. The AIMD simulations were performed—analogously to those in our previously published article on the force field parameters for [EMIm][OAc] (Roos and Brehm 2021)—with CP2k (Hutter et al. 2014, Kühne et al. 2020, The CP2K developers group 2017). The Quickstep method (VandeVondele et al. 2005) and orbital transformation (OT) (VandeVondele and Hutter 2003)

were used. For the description of the electronic structure, density functional theory (Hohenberg and Kohn 1964, Kohn and Sham 1965) was applied using the BLYP functional (Becke 1988, Lee et al. 1988) and the dispersion correction D3(BJ) of Grimme et al. (2010, 2011) together with the revised damping parameters of Smith et al. (2016). The DZVP–MOLOPT–SR–GTH (VandeVondele and Hutter 2007) were applied as basis sets, and GTH pseudopotentials (Goedecker et al. 1996, Hartwigsen et al. 1998) were utilized. The plane-wave energy cutoff was set to 350 Ry.

The first 15 ps of the AIMD simulations of the DMSO/IL/(H_2O) systems and the first 20 ps of the AIMD simulations of pure DMSO and DMSO/H_2O were discarded as equilibration. The resulting physical simulation times are listed in Table 1.

The force field MD simulations for the force field optimization were performed analogously to our previous article on BILFF (Roos and Brehm 2021). Using Packmol (Martínez et al. 2009), the molecules were statistically distributed in the simulation box. Pre-equilibrations were performed at first at 500 K in the NVE ensemble using a Berendsen thermostat (Berendsen et al. 1984) with a coupling constant of 1.0 fs and a physical simulation time of 25 ps, followed by 100 ps of equilibration at 350 K in the NpT ensemble using a Nosé–Hoover thermostat (Martyna et al. 1992, Nose 1984a, b) (coupling constant of 100 fs) and a Nosé–Hoover barostat (coupling constant of 2000 fs). The resulting acoustic shock waves were subsequently damped in another simulation interval with the help of a Langevin thermostat (Dünweg and Paul 1991, Schneider and Stoll 1978). After calculating the final volume of the simulation box as average over the NpT simulation, the simulation cell size was adjusted in a 10 ps simulation interval. In a next step, the shock waves were damped again and the system was equilibrated. Finally, the production run was performed in the NVT ensemble using a Nosé–Hoover thermostat (Martyna et al. 1992, Nose 1984a, b) with a physical simulation time of 10 ns and an integration time step of 0.5 fs. As in our previous studies, the Lennard-Jones cutoff radius was set to 800 pm. For the electrostatic interactions, the PPPM long-range solver from LAMMPS (Plimpton 1995) was utilized.

In addition, to validate the force field, an AIMD simulation and a force field MD simulation of pure and aqueous DMSO at 350 K were performed. To investigate the temperature dependence of the force field,

Table 1 Simulation parameters of the final equilibrated *ab initio* and force field MD simulations of DMSO in pure and aqueous [EMIm][OAc] for development and validation of the force field

System	Number of Molecules	Sim. time / ps	Box size / pm	Density / g cm^{-3}
AIMD				
DMSO	64 DMSO	90	1962	1.100
DMSO/H$_2$O	24 DMSO	31	1687	0.947
	48 Water			
DMSO/IL	12 DMSO	135	2052	1.064
	27 IL pairs			
DMSO/IL/H$_2$O	12 DMSO	120	2055	1.067
	26 IL pairs			
	12 Water			
FFMD				
DMSO	250 DMSO	10 000	3143	1.045
DMSO/H$_2$O	150 DMSO	10 000	3031	1.022
	300 Water			
DMSO/IL	60 DMSO	10 000	3485	1.086
	135 IL pairs			
DMSO/IL/H$_2$O	36 DMSO	10 000	2945	1.088
	78 IL pairs			
	36 Water			

Fig. 1 Snapshot of the force field MD simulation of a) pure DMSO, b) aqueous DMSO as well as c) DMSO/IL, and d) DMSO/IL/H$_2$O (atom color code: yellow – S; red – O; blue – N; gray – C; white – H)
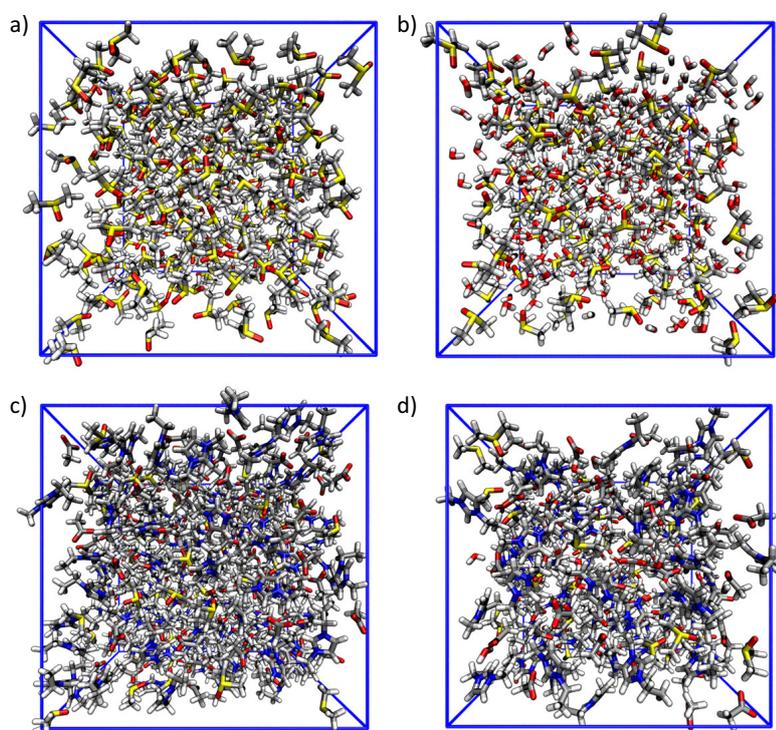
AIMD simulations of DMSO in [EMIm][OAc] and [EMIm][OAc]/water, as well as force field MD simulations of all four systems at an elevated temperature of 450 K, were carried out. After discarding the first 15 ps of the AIMD simulation of the DMSO/IL/(H$_2$O) system for equilibration, the resulting physical simulation lengths of the AIMD simulations at 450 K are 91 ps (DMSO/IL) and 90 ps (DMSO/IL/H$_2$O).

Snapshots of the force field MD simulations of the four different systems are shown in Fig. 1.

The analyses shown in this article were performed with the trajectory analysis program package TRAVIS (Brehm and Kirchner 2011, Brehm et al. 2020b) and visualized using xmgrace (Grace Development Team 1996), Wolfram Mathematica (Wolfram Research, Inc. 2010), and VMD (Humphrey et al. 1996) with the Tachyon renderer (J. Stone 1998). To calculate the lifetime of the hydrogen bonds, the autocorrelation formalism (Rapaport 1983) was applied.

Optimization procedure

In developing the force field parameters for DMSO in [EMIm][OAc], special attention was paid to the hydrogen bonds in the system, in particular the hydrogen bond between the oxygen atom of DMSO and the ring protons of the imidazolium cation. In the literature, the interaction between the cation and DMSO is considered to play an important role for the increased cellulose solubility (Mohan et al. 2016). The DMSO–DMSO interaction was also considered for the force field development. However, since this hydrogen bond is increasingly weakened by the ionic liquid above a molar fraction beyond $\chi_{IL} = 0.46$, its optimization was subordinated to the DMSO–cation hydrogen bond (Paiva et al. 2022).

The force field adjustment was performed simultaneously for both pure and aqueous [EMIm][OAc] at 350 K to also accurately reproduce the interactions of the system for a consideration of the subsequent precipitation process of cellulose with water. A molar fraction of DMSO of $\chi_{DMSO} = 0.31$ was chosen in pure [EMIm][OAc], since at this point the cellulose solubility is sufficiently high (Tomimatsu et al. 2019, Xu et al. 2013).

For DMSO in aqueous [EMIm][OAc], a molar fraction of DMSO of $\chi_{DMSO} = \chi_{water} = 0.24$ (equivalent to 3.9 wt.% water) was determined, since cellulose becomes insoluble in [EMIm][OAc] above 15 wt.% water content even at higher temperature of $323 - 353$ K (Froschauer et al. 2013, Le et al. 2012).

To optimize the force field parameters, the radial and distance–angle distribution functions of the cation–DMSO hydrogen bond calculated from force field MD simulations were compared with reference AIMD simulations.

The force field parameters q and σ for all atoms of DMSO were iteratively adjusted via a trial-and-error method until the deviations of the radial distribution function were minimized.

Furthermore, a comparison of the system density with experimental data was used for the optimization of σ (cf. Table 3). In the ESI†, also a comparison of the partial charges of DMSO with common other literature force fields (Bordat et al. 2003, Fox and Kollman 1998, Geerke et al. 2004, Strader and Feller 2002) based on AMBER (Cornell et al. 1995) and GROMOS (Liu et al. 1995) can be found.

The bonded interactions were adjusted by comparing the statistical occurrence of the bond lengths, angles, and dihedral angles appearing in the reference AIMD simulations with the force field simulations and modifying the corresponding force constants and height of the potential barriers. This resulted in, for example, a modification of the C–S bond length from 179.0 pm to 184.3 pm to reflect the statistically most frequent equilibrium bond length. Furthermore, it should be noted that dihedral angles were not specified in the reference force field OPLS–AA (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009) and were introduced in our force field. The nomenclature used for atom types can be found in Fig. 2. All force field parameters are shown in the ESI†.
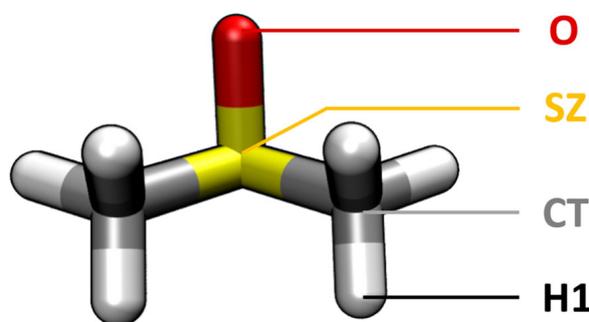
Fig. 2 Nomenclature of atom types in our force field BILFF

Microstructure of the system

Radial distribution functions

To analyze the microstructure, a distribution function of the averaged distances between a reference and an observed particle is calculated. The resulting radial pair distribution function (RDF) gives the probability of finding the observed particle at a given distance from a reference particle, relative to the uniform density of observed particles throughout the simulation cell. In the following, the hydrogen bonding of the different DMSO systems in the presence and absence of water is investigated via RDFs. The reproducibility of the microstructure of the AIMD simulations using BILFF for DMSO is discussed in comparison to the literature force field OPLS–AA (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009). Figure 3 shows the RDF between the oxygen atom of DMSO and the ring protons of [EMIm]$^+$ in both the pure and aqueous DMSO/IL systems. The RDFs of the individual hydrogen bonds were averaged over all three ring protons. The oxygen atom of DMSO form a hydrogen bond to the cation [EMIm]$^+$ with a particle distance of about 208 pm and a $g(r) = 1.5 - 1.6$, as can be seen from the first maximum of the RDF. The agreement of the RDF with the AIMD-based result is very good. To achieve this, the partial charge of the DMSO oxygen atom was increased by about 13% from -0.420 to -0.474 compared to the literature force field (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009). A comparison between the results of the literature force field and the AIMD simulation shows clear deviations and emphasizes the need for optimization of the force field parameters in order to accurately reproduce the hydrogen bonds. Comparing the DMSO–cation RDF in the presence and absence of water reveals similar residence probabilities, indicating that water does not influence the strength of the hydrogen bonding.

Considering the hydrogen bond between DMSO and water (cf. Fig. 4), a maximum in the RDF calculated from the reference AIMD simulation at a particle distance of about 180 pm and a height of $g(r) = 4.2$ can be observed. While the averaged hydrogen bonding distance between these atoms in the AIMD simulation is correctly reproduced by the force field MD simulation with BILFF and agrees very well with quantum chemical distances calculated in the literature (Kirchner and

Reiher 2002), the residence probability is slightly too low. Nevertheless, the agreement between the results of the reference AIMD simulation and a force field simulation with BILFF is much better than the result calculated with the literature force field. Finding a balance between the accuracy of the description of the complicated interactions between the protons of the cation vs. water and the DMSO oxygen atom is not trivial. However, our main goal is to accurately reproduce the cellulose/[EMIm][OAc]/DMSO system, so we have accepted the deviation in the replication of the DMSO-water hydrogen bond. As a result, the force field parameters of DMSO have not been further modified. A comparison between the DMSO–cation and the DMSO–water hydrogen bond shows a significantly higher $g(r)$ in the second case, which allows conclusions to be drawn about a stronger interaction with water.

In order to verify the transferability of the force field to different concentration ratios of DMSO and water, additional simulations were performed with a mixing ratio of DMSO to water of 1:1. An RDF of the hydrogen bond between DMSO and water can be found in Fig. S–3 in the ESI†. Again, the particle distance from the AIMD simulation can be reproduced correctly, but the height of the $g(r)$ is a little too low.

The analysis of the DMSO–DMSO interactions shows that the $g(r)$ between the DMSO carbon atoms and the protons (see Fig. 5) in the DMSO/IL and DMSO/IL/H$_2$O systems in the AIMD simulation differ slightly from each other. The g(r) of the force field simulation with BILFF lies within the average of these two RDFs. However, the RDF of the hydrogen bond between two DMSO molecules (cf. Fig. S-2) in the force field simulations shows a slight deviation from the AIMD-based results regarding the location and intensity of the initial peak. This difference is observed in both force field simulations utilizing the literature force field and BILFF. It underscores the challenge of accurately capturing the intricate interactions within this complex system.

Combined distance–angle distribution functions

Considering the combined distance–angle distribution functions of two molecules, information is provided whether certain molecule arrangements occur preferentially. Such a combined distribution function (CDF) between an example ring proton of [EMIm]$^+$ and the

Fig. 3 Comparison of the RDFs between the reference AIMD simulation and a force field MD simulation with the OPLS–AA force field (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009) and BILFF between the DMSO oxygen atom and the ring protons of [EMIm]$^+$. The RDFs are averaged over all three ring protons
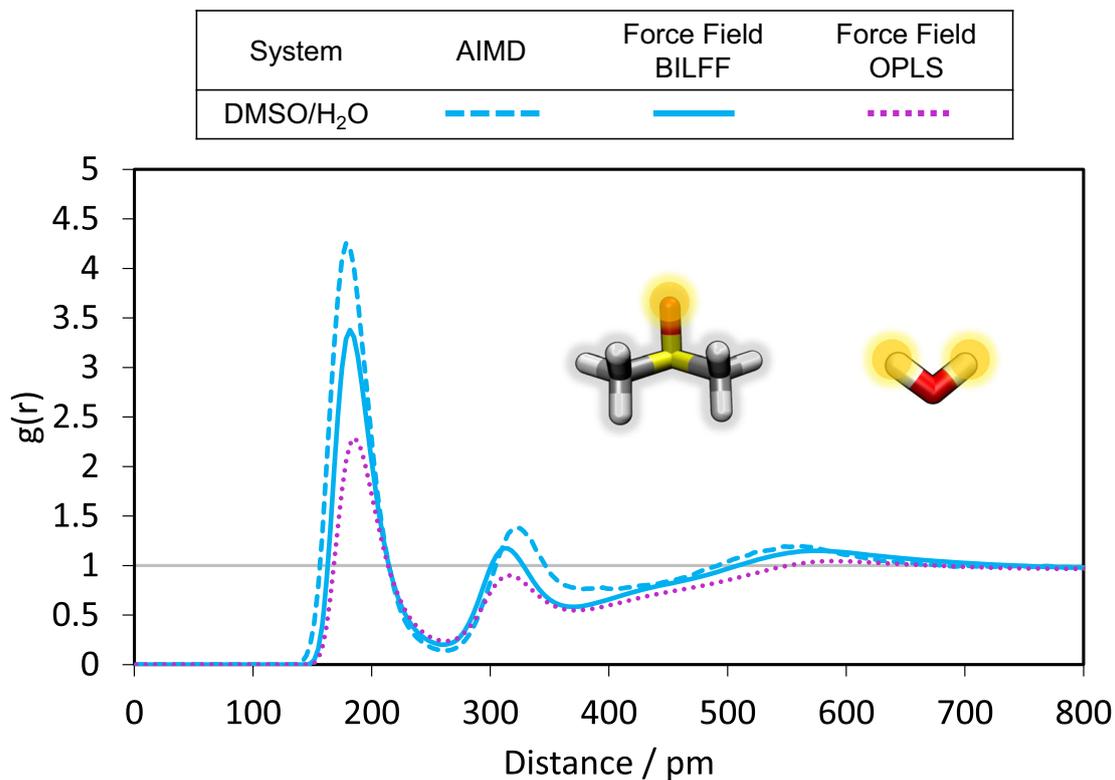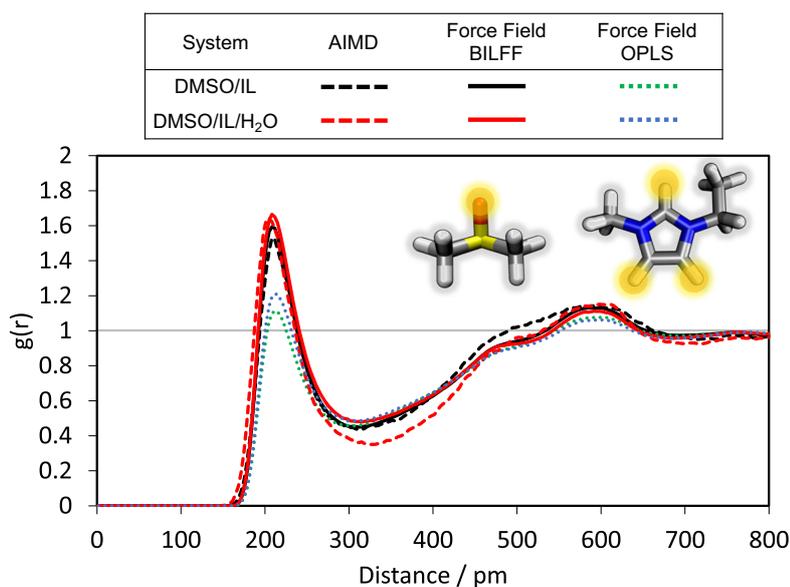
System	AIMD	Force Field BILFF	Force Field OPLS
DMSO/IL	– – –	——	······
DMSO/IL/H$_2$O	– – –	——	······

System	AIMD	Force Field BILFF	Force Field OPLS
DMSO/H$_2$O	– – –	——	······

Fig. 4 Comparison of the RDFs between the reference AIMD simulation and a force field MD simulation with the OPLS–AA force field (Jorgensen et al. 1996, Ponder and Case 2003, Sam-basivarao and Acevedo 2009) and BILFF between the DMSO oxygen atom and the water protons in the DMSO/H$_2$O system
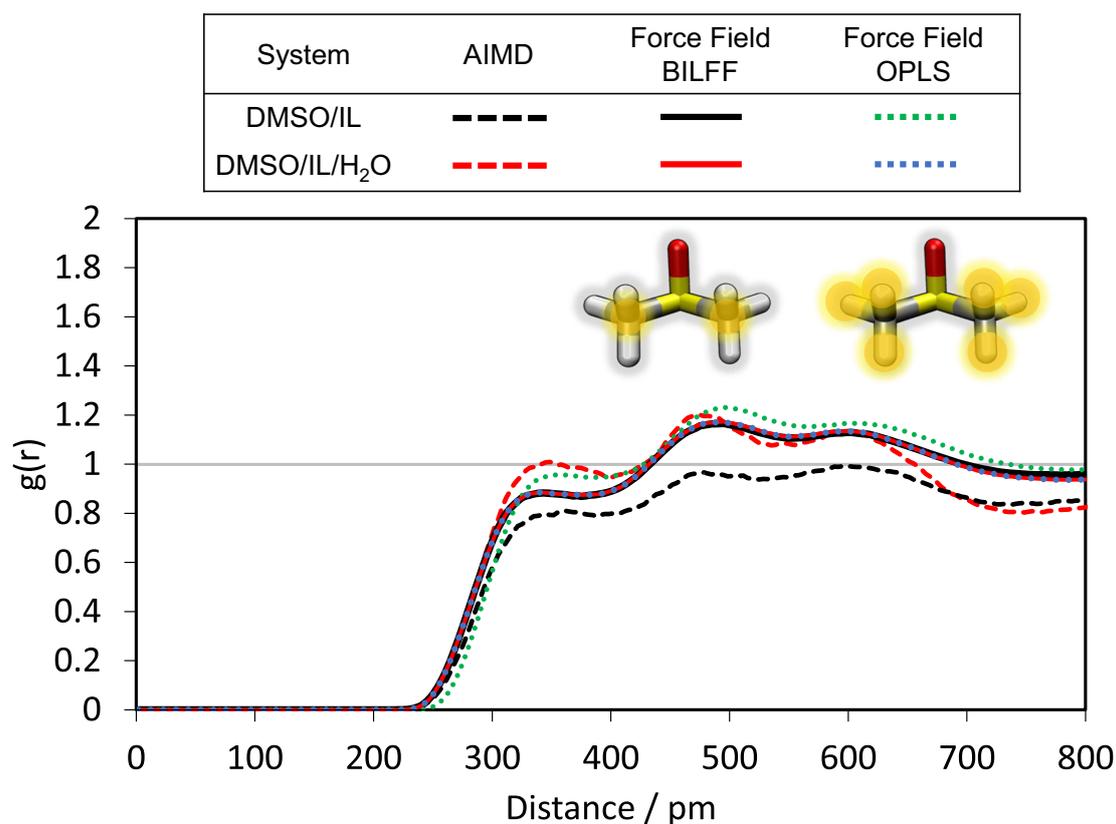
Fig. 5 Comparison of the RDFs between the reference AIMD simulation and a force field MD simulation with the OPLS–AA force field (Jorgensen et al. 1996, Ponder and Case 2003, Sam-basivarao and Acevedo 2009) and BILFF between the DMSO carbon atom and the DMSO protons

DMSO oxygen atom is presented in Fig. 6 showing the distance $O_{DMSO} \cdots H_{EMIm}$ and the corresponding angle $\measuredangle(O_{DMSO}, CH_{EMIm}, H_{EMIm})$ as depicted in the illustration. The CDF of the AIMD simulation (top) and the FFMD simulation with BILFF (bottom) exhibit a strong overall agreement. The peak at 80–280 pm, 117–180° represents the hydrogen bond between the two molecules. Here, the force field MD simulation depicts a slightly wider angle range, plausibly due to a better statistical sampling than the AIMD simulation. The second peak (395–520 pm, 50–80°) corresponds to the hydrogen bonding involving the isolated CH group of $[EMIm]^+$, resulting in an increased presence of DMSO on the "opposite" side of $[EMIm]^+$.

Considering the CDF of DMSO and water (cf. Fig. 7), a similar good agreement of the overall pattern can be obtained between the results of the AIMD and FFMD simulations. The first peak (30–240 pm, 145–180°) shows the hydrogen bond between DMSO and water. The second maximum at 295–350 pm and 40–70° arises from the "second" water proton. Again, the AIMD simulation infers a marginally stronger localization than the force field.

Spatial distribution functions

The microstructure of the system can be further investigated using spatial distribution functions (SDFs). Figure 8 shows the arrangement of the different compounds in the DMSO/IL/H₂O system around a DMSO molecule, where preferential orientations can be identified. Around the oxygen atom of a DMSO molecule, protons from neighboring DMSO molecules (shown in gray), $[EMIm]^+$ cations (shown in red), and water (shown in blue) are arranged in a circular formation in layers, resulting from the directional dependence of hydrogen bonding and the free rotation of the DMSO oxygen atom. Here, the competition of the
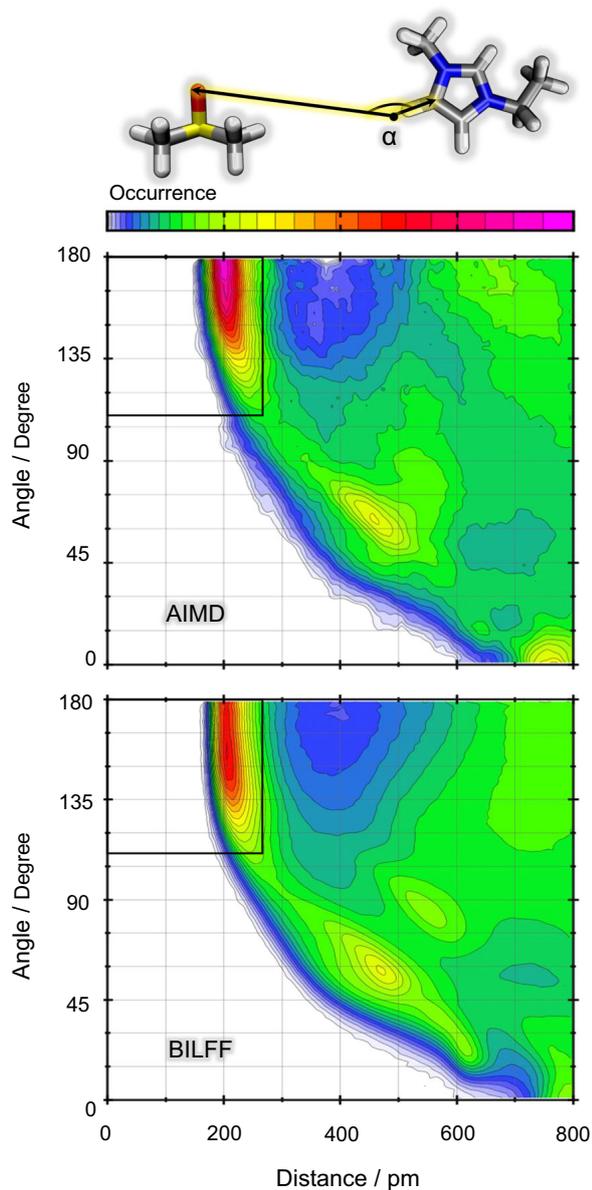
Fig. 6 Distance–angle distribution function between an example ring proton of [EMIm]$^+$ and the DMSO oxygen atom in the DMSO/IL system comparative as a result of a reference AIMD simulation and a force field MD simulation with BILFF. The black rectangle shows the geometric criterion for calculating the lifetime of the hydrogen bonds. The depicted angle is for illustrative purposes only

Fig. 7 Distance–angle distribution function between the protons of water and the DMSO oxygen atom in the DMSO/H$_2$O system comparative as a result of a reference AIMD simulation and a force field MD simulation with BILFF. The black rectangle shows the geometric criterion for calculating the lifetime of the hydrogen bonds. The depicted angle is for illustrative purposes only

hydrogen bond donors for the acceptor oxygen atom of DMSO becomes apparent, which is also discussed in the Sankey diagrams in Figs. 10 and 11. Within the molecule layers around the DMSO oxygen atom, a water bi-layer is formed around [EMIm]$^+$, resulting in being isolated from DMSO.

The protons of a DMSO molecule are surrounded by acetate (shown in green), water (shown in blue), and other DMSO molecules (shown in gray) in a competitive manner. The directed interactions of the hydrogen bonding partners, combined with the repulsion associated with the dihedral angles and the rotational free-
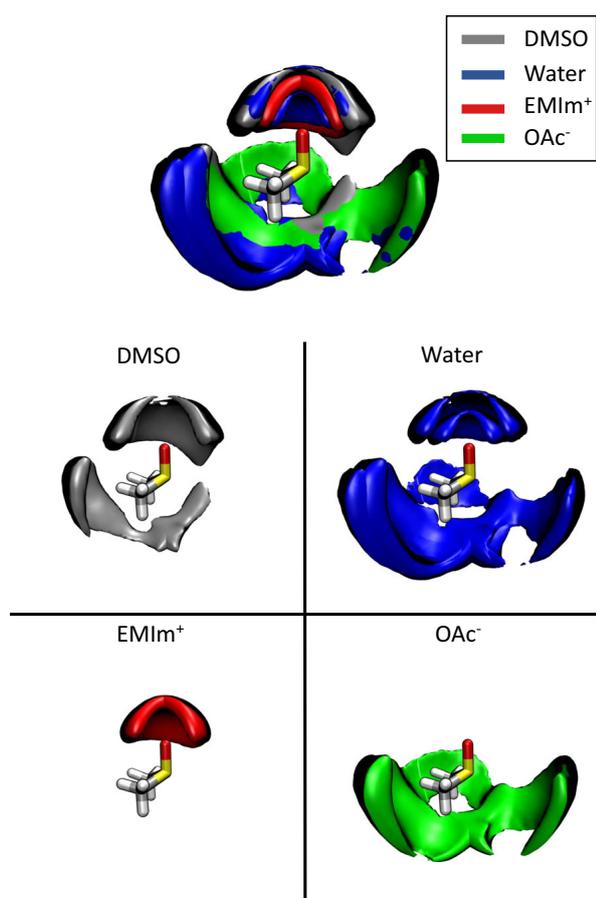
Fig. 8 Spatial distribution function (SDF) of the arrangement of molecules around DMSO with protons and oxygen atoms of DMSO (gray, oxygen: $4\,\mathrm{nm}^{-3}$; proton: $2\,\mathrm{nm}^{-3}$) and water (blue, $2.5\,\mathrm{nm}^{-3}$), the ring protons of [EMIm]$^+$ (red, $5\,\mathrm{nm}^{-3}$) and the oxygen atom of acetate (green, $5\,\mathrm{nm}^{-3}$) in the DMSO/IL/H$_2$O system calculated with BILFF

Fig. 9 Spatial distribution function (SDF) of the arrangement of molecules around [EMIm]$^+$ with the protons and oxygen atoms of DMSO (gray, $3\,\mathrm{nm}^{-3}$) and water (blue, $6\,\mathrm{nm}^{-3}$), the ring center of [EMIm]+ (red, $7\,\mathrm{nm}^{-3}$) and the oxygen atom of acetate (green, $10\,\mathrm{nm}^{-3}$) in the DMSO/IL/H$_2$O system calculated with BILFF

dom of the methyl groups in DMSO, result in a circular alignment of these molecules in three distinct regions.

A consideration of the spatial distribution functions of the neighboring molecules around an [EMIm]$^+$ cation (cf. Fig. 9) shows that DMSO arranges nearly uniformly around the cation without a preferred orientation. This is in contrast to the orientation of the oxygen atoms of water and acetate around the ring protons of [EMIm]$^+$ as well as an orientation of the ring center of the next cation in top of one another due to $\pi-\pi$-stacking. These results are in agreement with the observations for the [EMIm][OAc]–water system in the absence of DMSO, which were discussed in our previously published article on BILFF (Roos and Brehm 2021). Thus, the presence of DMSO in the aqueous

ionic liquid does not significantly change the arrangement of the anion and water around the cation, though it does compete with the hydrogen bonds of the anion and water for the cation, which will be discussed in the following section.

The SDFs calculated from the reference AIMD simulations can be found in the ESI†(cf. Figs. S-4 and S-5).

Competing hydrogen bonds

The multiple hydrogen bond donors and acceptors in the quaternary system of DMSO, [EMIm][OAc] and water engage in a competition with one another. In order to analyze the intricate network of the hydrogen bonds, Sankey diagrams are used (cf. Figs. 10 and 11).

Fig. 10 Sankey diagram of DMSO in anhydrous [EMIm][OAc] calculated from a force field MD simulation using BILFF with a comparison to the coordination numbers from a reference AIMD simulation
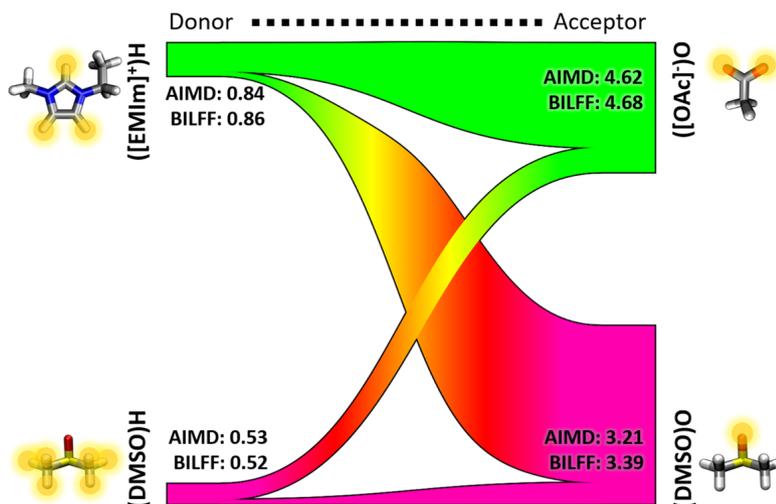
Fig. 11 Sankey diagram of DMSO in aqueous [EMIm][OAc] calculated from a force field MD simulation using BILFF with a comparison to the coordination numbers from a reference AIMD simulation
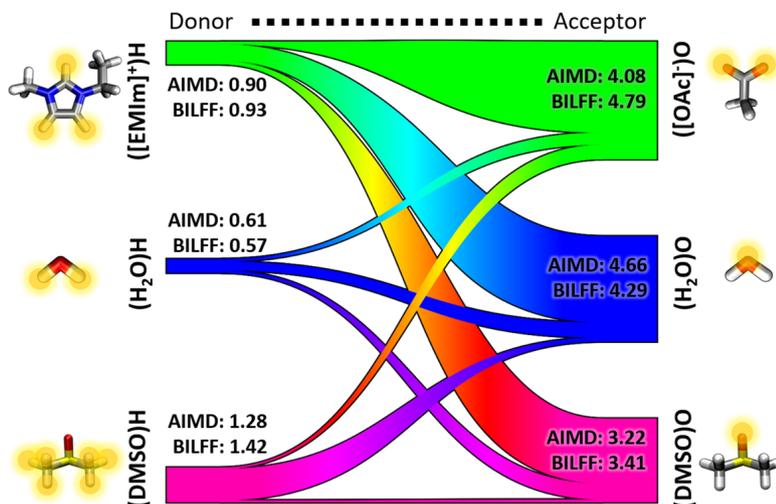
Fig. 12 Radial distribution function of the hydrogen atoms of [EMIm]$^+$ averaged over all ring protons, with the oxygen atom of DMSO in the anhydrous DMSO/IL system at 350 and 450 K simulated by AIMD and FFMD using BILFF
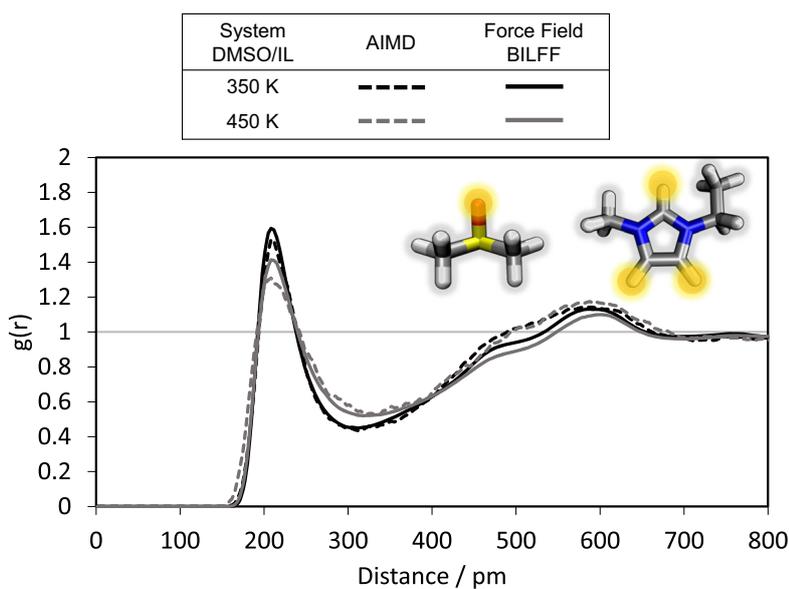
The left side of the diagram shows the hydrogen bond donors, while the right side corresponds to the hydrogen bond acceptors. The width of the bars is proportional to the number of hydrogen bonds that are formed per donor/acceptor site on average. The numbers in the bars represent the coordination numbers and result from the number integral of the corresponding RDFs up to the first minimum. In the anhydrous DMSO/IL system (cf. Fig. 10), it is obvious that hydrogen bonds are formed mainly between [EMIm]$^+$ and [OAc]$^-$ but DMSO competes strongly as a hydrogen bond acceptor. Thus, the oxygen atom of [OAc]$^-$ is surrounded by about 5 protons of mostly [EMIm]$^+$, but also of other DMSO molecules. The DMSO oxygen atom, on the other hand, is surrounded by 3 protons of other molecules. This finding is in agreement with literature simulations (Paiva et al. 2022). Similar results are found in the presence of water (cf. Fig. 11). The coordination number of the DMSO protons calculated from the force field MD simulation increases from 0.52 in the anhydrous system to 1.42 in the aqueous system. This is due to the hydrate shell around DMSO. The competition for the protons of [EMIm]$^+$, formerly between [OAc]$^-$ and DMSO, is now dominated by water.

The coordination numbers from the AIMD simulations, and thus the microstructure of the systems, can be well reproduced by force field simulations using BILFF in both the presence and absence of water.

Temperature dependence

Radial distribution functions at higher temperatures

BILFF was developed at a temperature of 350 K for DMSO in anhydrous and aqueous [EMIm][OAc]. For validation at different temperatures, force field MD simulations with the same simulation parameters (cf. Table 1) were performed at 450 K. In addition, force field MD simulations of anhydrous and aqueous DMSO were performed and included in the validation of our force field.

The radial distribution functions of the ring protons of [EMIm]$^+$ and the oxygen atom of DMSO in the anhydrous and aqueous DMSO/IL system are compared with RDFs calculated from the AIMD simulation (cf. Figs. 12 and 13). As expected, the residence probability $g(r)$ of the observed atoms is decreasing with increasing temperature in both systems. The average distance between the molecules, which is in both systems at around 210 pm, is not affected by temperature.

The force field simulation with our optimized parameters can also reproduce the results of the AIMD simulation at 450 K.

Hydrogen bond lifetime

In order to describe the microstructure of the systems and to validate BILFF, several structural quan-

Fig. 13 Radial distribution function of the hydrogen atoms of [EMIm]$^+$, averaged over all ring protons, with the oxygen atom of DMSO in the aqueous DMSO/IL system at 350 and 450 K simulated by AIMD and FFMD using BILFF
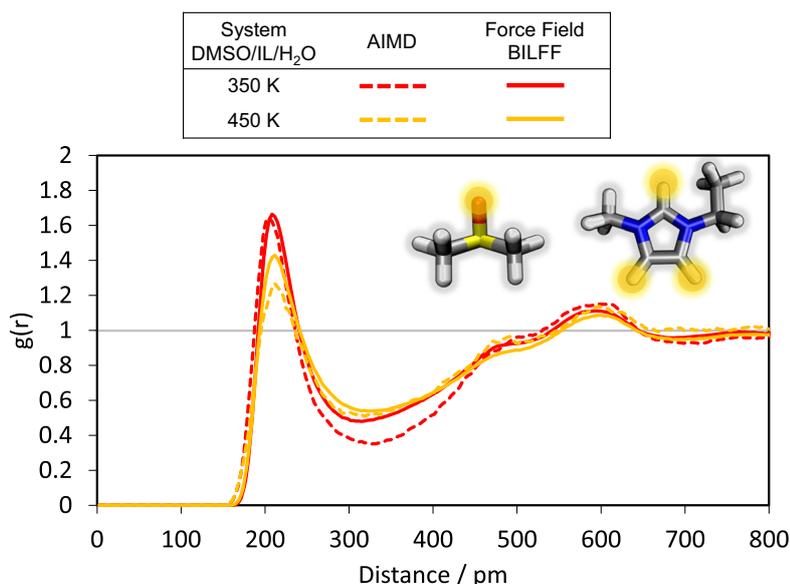
tities have been investigated so far. To further analyze the dynamics, we performed hydrogen bond lifetime calculations at two different temperatures and compared the results for six different systems, including anhydrous DMSO, aqueous DMSO, DMSO–[EMIm][OAc], DMSO–[EMIm][OAc]–water, and anhydrous and aqueous [EMIm][OAc]. Additionally, we compared the results of force field MD simulations with BILFF to the AIMD simulation.

Table 2 shows the lifetime of the most relevant hydrogen bonds of the molecules in the investigated systems at 350 K and 450 K resulting from AIMD simulations as well as force field simulations using BILFF. A distinction was made between the calculation of a continuous and intermittent hydrogen bond lifetime. While the intermittent lifetime allows breaking and reformation of hydrogen bonds, the continuous lifetime only considers hydrogen bonds without interruption in the autocorrelation. The distance and angle criteria between the hydrogen bond partners of interest were selected according to the maximum probability of residence in the combined distance–angle distribution function and are tabulated in the ESI†(cf. Table S-7). An example of the determination of these criteria is given in Figs. 6 and 7 as shown by the black rectangle in the CDFs.

The results of the lifetime calculation for the hydrogen bonds of the individual ring protons of $[EMIm]^+$ are also listed in the ESI†(cf. Table S-8). The following general conclusions can be drawn from the lifetime of the hydrogen bonds:

- As expected, a significant decrease in hydrogen bond lifetime can be observed with increasing temperature and thus a decrease in their stability.
- In any case, the hydrogen bond lifetime until the first interruption (the continuous lifetime) lasts only a few picoseconds, which is a typical time interval for hydrogen bonds (Brehm and Sebastiani 2018, Gehrke et al. 2018, Roos and Brehm 2021).
- The water–anion and the cation–anion hydrogen bonds generally have the longest observed lifetimes and thus the largest stability in all four systems studied.
- In the presence of DMSO, the intermittent lifetimes of the cation–anion, water–anion, and cation–water hydrogen bonds in the aqueous systems are significantly increased by a factor of up to more than two compared to the systems without DMSO in our previously published article (Roos and Brehm 2021). However, in the anhydrous system, the presence of DMSO reduces the lifetime of the cation–anion hydrogen bond.

Table 2 Overview of the lifetime τ of hydrogen bonds in the AIMD simulation and the force field MD simulation using BILFF at different temperatures. Lifetimes in ps

Temp.	Intermittent			Continuous	
	τ(AIMD)	τ(FFMD)		τ(AIMD)	τ(FFMD)
DMSO					
(DMSO)H-O(DMSO)					
350 K	22.79	22.33		1.82	1.38
450 K	–	7.04		–	0.95
DMSO / H_2O					
(DMSO)H-O(DMSO)					
350 K	12.17	23.01		1.27	1.07
450 K	–	5.24		–	0.76
$(H_2O)H-O(DMSO)$					
350 K	45.54	32.92		3.86	5.48
450 K	–	6.93		–	1.81
DMSO / [EMIm][OAc]					
(DMSO)H-O(DMSO)					
350 K	52.55	121.08		2.10	1.54
450 K	26.74	20.20		1.18	1.00

Table 2 continued

Temp.	Intermittent			Continuous	
	τ(AIMD)	τ(FFMD)		τ(AIMD)	τ(FFMD)
([EMIm])H−O(DMSO)					
350 K	64.06	95.21		2.02	2.40
450 K	(103.69)	17.88		0.92	1.28
([EMIm])H−O([OAc])					
350 K	424.66	547.30		4.12	4.78
450 K	48.88	86.59		1.69	2.11
DMSO / [EMIm][OAc] / H_2O					
(DMSO)H-O(DMSO)					
350 K	114.34	118.41		2.22	1.52
450 K	19.07	18.05		0.81	0.98
([EMIm])H−O(DMSO)					
350 K	78.17	93.75		2.13	2.31
450 K	19.95	15.99		0.91	1.24
([EMIm])H−O([OAc])					
350 K	164.51	436.43		2.92	3.97
450 K	57.04	67.32		1.40	1.89
([EMIm])H−O(H_2O)					
350 K	34.89	99.74		0.73	0.93
450 K	14.63	13.99		0.43	0.59
(H_2O)H−O(DMSO)					
350 K	(12.98)	141.73		0.38	3.71
450 K	79.99	19.31		0.71	1.48
(H_2O)H−O([OAc])					
350 K	443.35	455.55		0.02	0.20
450 K	34.49	27.22		0.02	0.15
[EMIm][OAc] [a]					
([EMIm])H−O([OAc])					
350 K	471.98	779.74		2.95	4.45
450 K	87.72	95.83		1.61	1.95
[EMIm][OAc] / H_2O [a]					
([EMIm])H−O([OAc])					
350 K	73.15	146.00		1.26	1.83
450 K	37.75	23.96		2.33	1.12
([EMIm])H−O(H_2O)					
350 K	31.18	40.36		0.57	0.75
450 K	9.81	5.70		0.40	0.50
(H_2O)H−O([OAc])					
350 K	153.15	165.65		0.16	0.78
450 K	44.89	19.83		0.10	0.45

[a] *Calculated from the MD simulations of our already published article* (Roos and Brehm 2021) *with the same angle and distance criteria*

– A consideration of the hydrogen bonds of the individual ring protons of [EMIm]$^+$ shows that the ring proton H2 (between both nitrogen atoms) forms longer lasting hydrogen bonds to water and [OAc]$^-$ than the other two ring protons H4 and H5 (cf. Table S-8 in the ESI†). This is in agreement with the results of a DMSO-free IL/(H$_2$O) system (Roos and Brehm 2021). However, the [EMIm][OAc]–DMSO hydrogen bonds are not affected by this phenomenon.

– The DMSO–H$_2$O hydrogen bonds are longer-lived than the DMSO–DMSO hydrogen bonds.

– The presence of water decreases the lifetime and thus the stability of the cation–anion hydrogen bond. This is also the case in the DMSO-free system (Roos and Brehm 2021). However, the DMSO hydrogen bonds are not significantly affected.

The AIMD-based results can be reproduced well in the case of the continuous hydrogen bond lifetime. Larger deviations occur for the intermittent hydrogen bond lifetime. On the one hand, dynamical properties—and in particular hydrogen bond lifetimes—are very sensitive to both the experimental conditions and the potential energy surface, which is defined by the force field parameters. Due to this high sensitivity, values can easily be off by an order of magnitude, and the agreement found here is still within acceptable limits (Gehrke et al. 2018, Kohagen et al. 2011, Malberg et al. 2013). On the other hand, it is noteworthy that certain lifetimes exceed the total simulation time of the AIMD simulations ($\approx$ 100 ps). While this is in principle possible via Rapaport's autocorrelation formalism, these AIMD-based values bear a large uncertainty and should be considered rough estimates. The force field simulations, on the other hand, are certainly long enough so that we can report the force field-based hydrogen bond lifetimes with greater confidence.

Validation with respect to experiment

To further validate our optimized force field, the densities, self-diffusion coefficients, thermal volume expansion coefficient, and bulk modulus were investigated and compared to experimental data. In addition, the effect of DMSO on the self-diffusion coefficients is investigated by comparison with DMSO-free [EMIm][OAc] systems (Roos and Brehm 2021).

A consideration of the densities of the different systems shows that the density of DMSO is significantly influenced by the addition of water and the ionic liquid (cf. Table 3). Comparing the density with literature results (Ciocirlan and Iulian 2012, Paknejad et al. 2019) shows an excellent reproduction of the density of pure DMSO at 350 K (cf. Table 3). The deviation is less than 1%. At 450 K, the calculated system density from the force field simulation is slightly lower than experimental data (Paknejad et al. 2019), but is still within an acceptable range with a deviation of 4.7%.

In aqueous DMSO, the calculated system density can also reproduce the experimental data with a deviation of less than 3.8%, given that it was measured at a lower temperature of 338 K with a slightly different molar ratio than used in ref. Cowie and Toporowski (1961) and ref. Lü et al. (2016). Although the force field was not optimized for pure and aqueous DMSO, the nevertheless good agreement of the system density with experimental data further validates the force field.

The density of the DMSO/IL system is in excellent agreement with experimental data (Nisa et al. 2019) and is replicated without any deviation.

As shown in our previous publication (Roos and Brehm 2021), the density of [EMIm][OAc] in water can be well reproduced using BILFF. To our knowledge, no literature values of the density for the complex DMSO/IL/H$_2$O system were published in a comparable temperature and concentration range. But given that the density of DMSO, DMSO/H$_2$O as well as DMSO/IL and IL/H$_2$O show a good agreement with experimental data, it can be assumed that the calculated density of the aqueous DMSO/IL system is also in a realistic range.

In all four systems, simulations utilizing the literature force field OPLS–AA (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009) exhibit a slightly too low system density when compared to experimental data.

Furthermore, self-diffusion coefficients at a temperature of 350 K and 450 K were calculated based on the mean squared displacement (MSD) using the Einstein equation and are shown in Table 4. The resulting diffusion coefficients allow the following conclusions:

– In all systems, the self-diffusion coefficient increases with increasing temperature as expected.

– In general, the diffusion coefficients in DMSO/IL/(H$_2$O) are increased compared to DMSO-free IL/(H$_2$O), which were considered in our previously published article (Roos and Brehm 2021).

Table 3 Comparison of the system densities from force field MD simulations (FFMD) using the literature force field OPLS–AA (Jorgensen et al. 1996, Ponder and Case 2003, Sambasivarao and Acevedo 2009) and BILFF with experimental data at two different temperatures

Temperature	Box size / pm	ρ(OPLS) /g·cm^{-3}	ρ(BILFF) / g·cm^{-3}	ρ(Lit.) / g·cm^{-3}
DMSO				
350 K	3142.57	0.99	1.05	1.04 [a] / 1.04 [b]
450 K	3305.94	–	0.90	0.94 [c]
DMSO / H_2O				
350 K	3030.53	0.98	1.02	1.06 [d] / 1.06 [e]
450 K	3193.82	–	0.87	–
DMSO / [EMIm][OAc]				
350 K	3496.76	1.06	1.07	1.07 [f]
450 K	3599.64	–	0.99	–
DMSO / [EMIm][OAc] / H_2O				
350 K	2953,89	1.07	1.08	–
450 K	3041.72	–	0.99	–

The molar fractions are χ_{DMSO}=0.33 in the DMSO/H_2O system and χ_{DMSO}=0.31 in the DMSO/IL mixture

[a] *Measurements at 353.15 K from ref.* Ciocirlan and Iulian (2012).

[b] *Measurements at T=353.15 K from ref.* Paknejad et al. (2019).

[c] *Measurements at T=453.15 K from ref.* Paknejad et al. (2019).

[d] *Measurements at T=338.15 K with a molar fraction of DMSO of x_{DMSO}=0.349 from ref.* Cowie and Toporowski (1961).

[e] *Measurements at T=338.15 K with a molar fraction of DMSO of x_{DMSO}=0.349 from ref.* Lü et al. (2016).

[f] *Measurements at T=353.15 K with a molar fraction of DMSO of x_{DMSO}=0.294 from ref.* Nisa et al. (2019)

Table 4 Self-diffusion coefficients D from force field MD simulations using BILFF of pure and aqueous DMSO as well as DMSO in pure and aqueous [EMIm][OAc] at different temperatures

	Temperature	D(FFMD) / 10^{-11} m^2 s^{-1}	D(Lit.) / 10^{-11} m^2 s^{-1}
DMSO			
DMSO	350 K	171.8	126.4 [a]
DMSO / H_2O			
DMSO	350 K	114.5	65.0 [b]
	450 K	581.7	–
H_2O	350 K	200.9	110.0 [b]
	450 K	1038.7	–
DMSO / [EMIm][OAc]			
DMSO	350 K	27.5	31.4 [c]
	450 K	158.3	–
[EMIm]$^+$	350 K	16.2	20.6 [c]
	450 K	112.0	–
[OAc]$^-$	350 K	14.0	20.2 [c]
	450 K	86.7	–

Table 4 continued

	Temperature	D(FFMD) $/\ 10^{-11}\,\mathrm{m^2\,s^{-1}}$	D(Lit.) $/\ 10^{-11}\,\mathrm{m^2\,s^{-1}}$
DMSO / [EMIm][OAc] / H$_2$O			
DMSO	350 K	26.3	–
	450 K	196.0	–
[EMIm]$^+$	350 K	17.4	–
	450 K	102.7	–
[OAc]$^-$	350 K	11.7	–
	450 K	95.8	–
H$_2$O	350 K	19.7	–
	450 K	183.6	–
[EMIm][OAc] [f]			
[EMIm]$^+$	350 K	9.26	14[d] / 0.98[e]
	450 K	81.76	311[d]
[OAc]$^-$	350 K	6.72	13[d] /0.84[e]
	450 K	66.83	307[d]
[EMIm][OAc] / H$_2$O [f]			
[EMIm]$^+$	350 K	21.77	4.17[e]
	450 K	171.09	–
[OAc]$^-$	350 K	22.42	4.61[e]
	450 K	155.76	–
H$_2$O	350 K	55.97	10.47[e]
	450 K	378.47	–

The molar fractions in the simulations are χ_{DMSO}=0.33 in DMSO–H$_2$O, χ_{DMSO}=0.31 in DMSO/IL and χ_{DMSO}=0.24 in DMSO/IL/H$_2$O. For comparison, the diffusion coefficients of DMSO-free [EMIm][OAc] and [EMIm][OAc]/H$_2$O are also shown from our previously published article (Roos and Brehm 2021)

[a] *Measurements at 328.15 K from ref.* Holz et al. (2000)

[b] *Measurements at T=308.2 K and* χ_{DMSO}*=0.3 from ref.* Packer and Tomlinson (1971)

[c] *Extrapolated to 350 K via Arrhenius plot from temperature-dependent experimental measurements at T=283–333 K, see ref.* Radhi et al. (2015)

[d] *Linearly extrapolated values of temperature dependent measurements at* $283-333\,K$ *from ref.* Green et al. (2017)

[e] *Measurement at T=313.15 K with a molar fraction of [EMIm][OAc] of* χ_{IL}*=0.3 in the aqueous system* (Hall et al. 2012)

[f] *Calculated values of force field simulations with BILFF at 350 K from our previously published article* (Roos and Brehm 2021). χ_{IL}*=0.25 in the aqueous system*

– In the pure and aqueous DMSO/IL mixture, the self-diffusion coefficients are lower than in both DMSO systems without ionic liquid. Accordingly, the ionic liquid reduces the diffusion rate of DMSO.

– Considering the simulations at a temperature of 350 K, water also lowers the diffusion coefficient of DMSO. This is consistent with the increase of density in the presence of the ionic liquid and water (cf. Table 3) although the viscosity decreases (Le et al. 2014).

– The decrease in the diffusion rate caused by water in a DMSO/IL/H$_2$O mixture at 350 K thereby shows the opposite effect to the increase of diffusion in DMSO-free IL/H$_2$O (Roos and Brehm 2021), demonstrating the complexity of the underlying hydrogen bond network.

– In both anhydrous and aqueous DMSO/IL, the calculated self-diffusion coefficient of DMSO is slightly larger than the self-diffusion coefficients of [EMIm]$^+$ and [OAc]$^-$, which is in agreement with experimental values for anhydrous DMSO/IL (Radhi et al. 2015). The cation diffuses faster than the anion at both temperatures. Furthermore, water molecules diffuse faster than [EMIm][OAc] but

Fig. 14 Calculation of the thermal volume expansion coefficient γ of the DMSO/IL system at 350 K from multiple FFMD simulations using BILFF at constant pressure (1 bar) and different temperatures
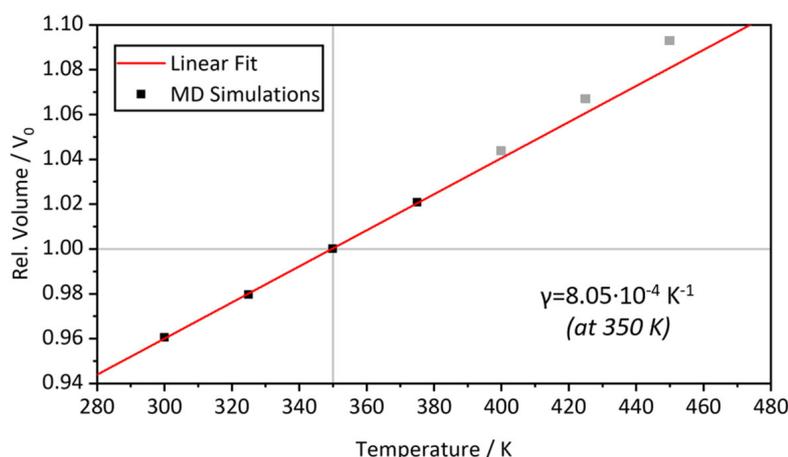
slower than DMSO in the DMSO/IL/H$_2$O mixture, while in the absence of the ionic liquid, water shows a significantly higher diffusion coefficient.

The results of the force field MD simulations with BILFF reflect the general patterns observed in the experimental data. Despite occasional numerical fluctuations in the individual diffusion values and slightly different temperatures and molar fractions of DMSO in the measurements, the overall agreement of this dynamic parameter with the experimental results remains acceptable. The temperature dependence is replicated across a range spanning more than one order of magnitude. It should be noted that especially the quaternary mixture DMSO/[EMIm][OAc]/water has been little studied so far, so no experimental data are available.

All validations discussed above were focused on the microscopic structure and dynamics of the liquid, which is the main aim of our force field. However, to

obtain a more complete picture, we also performed a few validations based on (macroscopic) thermodynamic properties, as it will be discussed in the following.

The bulk modulus of pure DMSO (cf. Fig. 14) and the thermal expansion coefficient of the DMSO/IL system (cf. Fig. 15) were calculated and compared with experimental data. For this purpose, force field simulations were performed in the *NpT* ensemble at different temperatures/pressures, and the relative volume change was represented as a graph. The two quantities were calculated using a linear fit to the results close to a temperature of 350 K and a pressure of 1 bar and were compared to experimental data in Table 5. The compressibility was measured at 353.15 K and 5 bar (Paknejad et al. 2019) as well as at a temperature of 323.15 K and a pressure of 1.01 bar (Egorov et al. 2010). The thermal expansion coefficient for the DMSO/IL mixture was measured at 353.15 K and a molar fraction

Fig. 15 Calculation of the compressibility κ and the bulk modulus K of pure DMSO at 1 bar from multiple FFMD simulations using BILFF at constant temperature (350 K) and different pressures
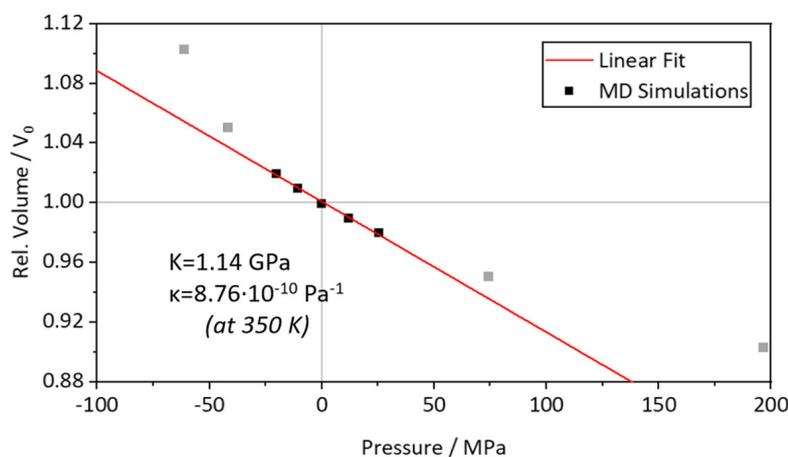
Table 5 Comparison of the calculated compressibility κ of a force field MD simulation using BILFF of pure DMSO and the calculated thermal expansion coefficient γ of DMSO in [EMIm][OAc] (χ_{IL}=0.69) at 350 K and 1 bar with experimental data (Egorov et al. 2010, Paknejad et al. 2019, Yang et al. 2019a)

	FFMD	Experiment
DMSO		
Compressibility κ / 10^{-10} Pa^{-1}	8.76	7.35[a] / 7.04[b]
DMSO/[EMIm][OAc]		
Thermal expansion coeff. γ / 10^{-4} K^{-1}	8.05	6.10[c]

[a] *Measurements at 353.15 K and 5 bar from ref.* Paknejad et al. (2019).
[b] *Measurements at T=323.15 K and 1.01 bar from ref.* Egorov et al. (2010).
[c] *Measurements at T=353.15 K and χ_{IL}=0.65 from ref.* Yang et al. (2019a)

of the ionic liquid of χ_{IL}=0.647 according to ref. Yang et al. (2019a).

A comparison of the calculated values for the compressibility of pure DMSO and the thermal expansion coefficient of DMSO in [EMIm][OAc] with experimental data shows a good agreement, although the thermal expansion was measured at a 30 K lower temperature and the force field was not optimized for pure DMSO. In addition to the density (cf. Table 3) and the self-diffusion coefficient (cf. Table 4), also these two quantities can be well reproduced by BILFF for both pure DMSO as well as a DMSO/IL mixture.

The final validation which we performed concerns the enthalpy of vaporization ΔH_{Vap} of pure DMSO. We have used the following computational protocol. During a liquid phase simulation of pure DMSO in NpT ensemble, we averaged the total potential energy of the system, which is equivalent to the internal energy U_{liquid}. In addition, we performed a set of simulations of one DMSO molecule in vacuum, starting from different initial configurations, and also averaged the potential energy, which corresponds to U_{gas}. Based on these results, we were able to obtain the internal energy of vaporization $\Delta U_{Vap} = U_{gas} - U_{liquid}$. The final step to the enthalpy is given via $\Delta H_{Vap} = \Delta U_{Vap} + p\Delta V$. From the experimental liquid and vapor density of DMSO, 1100 g L^{-1} and 3.43 g L^{-1}, respectively *(both at standard conditions)*, we obtained $\Delta V =$

Table 6 Experimental and computed enthalpy of vaporization ΔH_{Vap} of pure DMSO at 293 K and 1 bar, including the pV work term $p\Delta V \approx 2.3$ kJ mol^{-1}. Experimental data taken from ref. Douglas (1948)

	ΔH_{Vap} / kJ mol^{-1}
Experiment[a]	52.89 ± 0.42
Simulation (BILFF)	54.14
Simulation (OPLS–AA)	51.13

[a] *Measurement at 298.15 K from ref.* Douglas (1948)

22.71 L mol^{-1}, and therefore $p\Delta V = 2.3$ kJ mol^{-1}, which needs to be added to the internal energies of vaporization.

The results are presented in Table 6. We performed the computational protocol with both the original DMSO force field (OPLS–AA) and our optimized force field (BILFF) and compare the obtained values to the experimental reference value. While the resulting enthalpy of vaporization based on OPLS–AA is 3.4% too small, the value based on BILFF is 2.4% too large. We conclude that the resulting enthalpies of vaporization for pure DMSO based on both force fields are reasonable and very well within the expected range, and that the result based on BILFF is slightly better in comparison to experiment.

Conclusion

In this article, we present the extension of the all-atom force field for *B*io-Polymers in *I*onic *L*iquids (short: BILFF) to the co-solvent DMSO, which significantly increases the solubility of cellulose in ionic liquids such as [EMIm][OAc] (Anokhina et al. 2017, Mohan et al. 2016, Tomimatsu et al. 2019, Xu et al. 2013, Yang et al. 2019b). Previously, we published optimized force field parameters for the ionic liquids 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) (Roos and Brehm 2021), 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]), 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]), and 1-ethyl-3-methylimidazolium benzoate ([EMIm][OBz]) (Roos et al. 2023a) and cellulose (Roos et al. 2023b).

The force field was developed for an accurate reproduction of the directed interactions in DMSO/[EMIm][OAc] in the presence and absence of water with a special focus on the hydrogen bonds. For this purpose, the radial distribution functions, the

distance–angle distribution functions, and the density and statistical frequency of occurring bond lengths, angles, and dihedral angles were analyzed and compared with the results of reference AIMD simulations as well as experimental density measurements. Based on this comparison, the force field parameters were iteratively adjusted until the deviations between the results from the force field MD and AIMD simulation were minimized. The force field was validated by experimental data on density, self-diffusion coefficients, enthalpy of vaporization, thermal expansion coefficient and compressibility of the different systems, showing excellent or at least reasonable agreement with literature data. Furthermore, an applicability of our optimized force field parameters at higher temperatures was shown based on a comparison with reference AIMD simulations.

The simulation results show that DMSO forms a solvation shell around $[EMIm]^+$ and thus shields it from $[OAc]^-$. This is consistent with the prevailing theory in the literature that $[OAc]^-$ is more freely available for interaction with cellulose due to the DMSO shielding, thus increasing cellulose solubility (Mohan et al. 2016, Tomimatsu et al. 2019, Xu and Zhang 2015). In the aqueous system, the underlying hydrogen bond network is more complicated, as water is also a strong hydrogen bond donor for DMSO. The lifetime of the cation–anion hydrogen bond in DMSO/IL/water is even increased. Due to the complexity of the underlying interaction network, no trivial answer to this apparent contradiction can be given here.

The diffusion rate is increased by DMSO compared to the pure IL, which is also consistent with the literature mechanism (Le et al. 2014, Tomimatsu et al. 2019) and is considered to be another reason for the increased cellulose solubility in DMSO/IL.

Data Availability Statement The corresponding author [MB] can be contacted for available data supporting the findings of this study.

Declarations

Conflicts of interest There are no conflicts of interest to declare.

Ethics approval Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

Abeer MM, Amin M, Iqbal MC, Martin C (2014) A review of bacterial cellulose-based drug delivery systems: their biochemistry, current approaches and future prospects. J Pharm Pharmacol 66(8):1047–1061. https://doi.org/10.1111/jphp.12234

Andersen HC (1983) Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. J Comput Phys 52(1):24–34. https://doi.org/10.1016/0021-9991(83)90014-1

Anokhina TS, Pleshivtseva TS, Ignatenko VY, Antonov SV, Volkov AV (2017) Fabrication of composite nanofiltration membranes from cellulose solutions in an [emim]oac–dmso mixture. Pet Chem 57(6):477–482. https://doi.org/10.1134/S0965544117060020

Azimi B, Maleki H, Gigante V, Bagherzadeh R, Mezzetta A, Milazzo M, Guazzelli L, Cinelli P, Lazzeri A, Danti S (2022) Cellulose-based fiber spinning processes using ionic liquids. Cellulose 29(6):3079–3129. https://doi.org/10.1007/s10570-022-04473-1

Becke A (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A 38:3098–3100. https://doi.org/10.1103/PhysRevA.38.3098

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81(8):3684–3690. https://doi.org/10.1063/1.448118

Bordat P, Sacristan J, Reith D, Girard S, Glättli A, Müller-Plathe F (2003) An improved dimethyl sulfoxide force field for molecular dynamics simulations. Chem Phys Lett 374(3–4):201–205. https://doi.org/10.1016/S0009-2614(03)00550-5

Brehm M, Kirchner B (2011) Travis - a free analyzer and visualizer for monte carlo and molecular dynamics trajectories. J Chem Inf Model 51(8):2007–2023. https://doi.org/10.1021/ci200217w

Brehm M, Sebastiani D (2018) Simulating structure and dynamics in small droplets of 1-ethyl-3-methylimidazolium acetate. J Chem Phys 148(19):193802. https://doi.org/10.1063/1.5010342

Brehm M, Pulst M, Kressler J, Sebastiani D (2019) Triazolium-based ionic liquids: a novel class of cellulose solvents. J Phys Chem B 123(18):3994–4003. https://doi.org/10.1021/acs.jpcb.8b12082

Brehm M, Radicke J, Pulst M, Shaabani F, Sebastiani D, Kressler J (2020) Dissolving cellulose in 1,2,3-triazolium- and imidazolium-based ionic liquids with aromatic anions. Molecules 25(15):3539–3559. https://doi.org/10.3390/molecules25153539

Brehm M, Thomas M, Gehrke S, Kirchner B (2020) Travis-a free analyzer for trajectories from molecular simulation. J Chem Phys 152(16):164105. https://doi.org/10.1063/5.0005078

Ciocirlan O, Iulian O (2012) Properties of pure 1-butyl-2,3-dimethylimidazolium tetrafluoroborate ionic liquid and its binary mixtures with dimethyl sulfoxide and acetonitrile. J Chem Eng Data 57(11):3142–3148. https://doi.org/10.1021/je3007474

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 117(19):5179–5197. https://doi.org/10.1021/ja00124a002

Cowie JMG, Toporowski PM (1961) Association in the binary liquid system dimethyl sulfoxide – water. Can J Chem 39(11):2240–2243. https://doi.org/10.1139/v61-296

Douglas TB (1948) Vapor pressure of methyl sulfoxide from 20 to 50°. calculation of the heat of vaporization. J Am Chem Soc 70(6):2001–2002. https://doi.org/10.1021/ja01186a005

Dünweg B, Paul W (1991) Brownian dynamics simulations without gaussian random numbers. Int J Mod Phys C 02(03):817–827. https://doi.org/10.1142/S0129183191001037

Egorov GI, Makarov DM, Kolker AM (2010) Densities and volumetric properties of ethylene glycol + dimethylsulfoxide mixtures at temperatures of (278.15 to 323.15) k and pressures of (0.1 to 100) mpa. J Chem Eng Data 55(9):3481–3488. https://doi.org/10.1021/je100089s

Fang S, Lyu X, Tong T, Lim AI, Li T, Bao J, Hu YH (2023) Turning dead leaves into an active multifunctional material as evaporator, photocatalyst, and bioplastic. Nat Commun 14(1):1203. https://doi.org/10.1038/s41467-023-36783-8

Fox T, Kollman PA (1998) Application of the resp methodology in the parametrization of organic solvents. J Phys Chem B 102(41):8070–8079. https://doi.org/10.1021/jp9717655

Froschauer C, Hummel M, Iakovlev M, Roselli A, Schottenberger H, Sixta H (2013) Separation of hemicellulose and cellulose from wood pulp by means of ionic liquid/cosolvent systems. Biomacromol 14(6):1741–1750. https://doi.org/10.1021/bm400106h

Geerke DP, Oostenbrink C, van der Vegt NFA, van Gunsteren WF (2004) An effective force field for molecular dynamics simulations of dimethyl sulfoxide and dimethyl sulfoxide−water mixtures. J Phys Chem B 108(4):1436–1445. https://doi.org/10.1021/jp035034i

Gehrke S, von Domaros M, Clark R, Hollóczki O, Brehm M, Welton T, Luzar A, Kirchner B (2018) Structure and lifetimes in ionic liquids and their mixtures. Faraday Discuss 206:219–245. https://doi.org/10.1039/C7FD00166E

Goedecker S, Teter M, Hutter J (1996) Separable dual-space gaussian pseudopotentials. Phys Rev B 54:1703–1710. https://doi.org/10.1103/PhysRevB.54.1703

Grace Development Team (1996) see http://plasma-gate.weizmann.ac.il/grace

Green SM, Ries ME, Moffat J, Budtova T (2017) Nmr and rheological study of anion size influence on the properties of two imidazolium-based ionic liquids. Sci Rep 7(1):8968. https://doi.org/10.1038/s41598-017-09509-2

Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. J Chem Phys 132(15):154104. https://doi.org/10.1063/1.3382344

Grimme S, Ehrlich S, Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. J Comput Chem 32(7):1456–1465. https://doi.org/10.1002/jcc.21759

Hall CA, Le KA, Rudaz C, Radhi A, Lovell CS, Damion RA, Budtova T, Ries ME (2012) Macroscopic and microscopic study of 1-ethyl-3-methyl-imidazolium acetate-water mixtures. J Phys Chem B 116(42):12810–12818. https://doi.org/10.1021/jp306829c

Hartwigsen C, Goedecker S, Hutter J (1998) Relativistic separable dual-space gaussian pseudopotentials from h to rn. Phys Rev B 58:3641–3662. https://doi.org/10.1103/PhysRevB.58.3641

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science 315(5813):804–807. https://doi.org/10.1126/science.1137016

Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. Phys Rev B 136:864. https://doi.org/10.1103/PhysRev.136.B864

Holz M, Heil SR, Sacco A (2000) Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate 1h nmr pfg measurements. Phys Chem Chem Phys 2(20):4740–4742. https://doi.org/10.1039/B005319H

Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T (2004) Development of an improved four-site water model for biomolecular simulations: Tip4p-ew. J Chem Phys 120(20):9665–9678. https://doi.org/10.1063/1.1683075

Humphrey W, Dalke A, Schulten K (1996) Vmd: visual molecular dynamics. J Mol Graph 14(1):33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Hutter J, Iannuzzi M, Schiffmann F, VandeVondele J (2014) Cp2k: atomistic simulations of condensed matter systems. Comput Mol Sci 4(1):15–25. https://doi.org/10.1002/wcms.1159

J Stone (1998) Msc thesis. Computer Science Department, University of Missouri-Rolla

Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 118(45):11225–11236. https://doi.org/10.1021/ja9621760

Kirchner B, Reiher M (2002) The secret of dimethyl sulfoxide-water mixtures. a quantum chemical study of 1dmso-n water clusters. J Am Chem Soc 124(21):6206–6215. https://doi.org/10.1021/ja017703g

Klemm D, Heublein B, Fink HP, Bohn A (2005) Cellulose: fascinating biopolymer and sustainable raw material. Angew Chem 44(22):3358–3393. https://doi.org/10.1002/anie.200460587

Kohagen M, Brehm M, Lingscheid Y, Giernoth R, Sangoro J, Kremer F, Naumov S, Iacob C, Kärger J, Valiullin R, Kirchner B (2011) How hydrogen bonds influence the mobility of imidazolium-based ionic liquids. a combined theoretical and experimental study of 1-n-butyl-3-methylimidazolium bromide. J Phys Chem B 115(51):15280–15288. https://doi.org/10.1021/jp206974h

Kohn W, Sham L (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140:1133. https://doi.org/10.1103/PhysRev.140.A1133

Kühne TD, Iannuzzi M, Del Ben M, Rybkin VV, Seewald P, Stein F, Laino T, Khaliullin RZ, Schütt O, Schiffmann F, Golze D, Wilhelm J, Chulkov S, Bani-Hashemian MH, Weber V, Borštnik U, Taillefumier M, Jakobovits AS, Lazzaro A, Pabst H, Müller T, Schade R, Guidon M, Andermatt S, Holmberg N, Schenter GK, Hehn A, Bussy A, Belleflamme F, Tabacchi G, Glöß A, Lass M, Bethune I, Mundy CJ, Plessl C, Watkins M, VandeVondele J, Krack M, Hutter J (2020) Cp2k: An electronic structure and molecular dynamics software package - quickstep: efficient and accurate electronic structure calculations. J Chem Phys 152(19):194103. https://doi.org/10.1063/5.0007045

Le KA, Sescousse R, Budtova T (2012) Influence of water on cellulose-emimac solution properties: a viscometric study. Cellulose 19(1):45–54. https://doi.org/10.1007/s10570-011-9610-3

Le KA, Rudaz C, Budtova T (2014) Phase diagram, solubility limit and hydrodynamic properties of cellulose in binary solvents with ionic liquid. Carbohydr Polym 105:237–243. https://doi.org/10.1016/j.carbpol.2014.01.085

Lee C, Yang W, Parr R (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785–789. https://doi.org/10.1103/PhysRevB.37.785

Li C, Tanjore D, He W, Wong J, Gardner JL, Sale KL, Simmons BA, Singh S (2013) Scale-up and evaluation of high solid ionic liquid pretreatment and enzymatic hydrolysis of switchgrass. Biotechnol Biofuels 6(1):154. https://doi.org/10.1186/1754-6834-6-154

Liu H, Mueller-Plathe F, van Gunsteren WF (1995) A force field for liquid dimethyl sulfoxide and physical properties of liquid dimethyl sulfoxide calculated using molecular dynamics simulation. J Am Chem Soc 117(15):4363–4366. https://doi.org/10.1021/ja00120a018

Lü P, Zhao G, Zhang X, Yin J, Bao J (2016) Measurement and prediction on the surface properties of dimethyl sulfoxide/water mixtures. Chem Res Chin Univ 32(1):100–105. https://doi.org/10.1007/s40242-016-5297-1

Malberg F, Brehm M, Hollóczki O, Pensado AS, Kirchner B (2013) Understanding the evaporation of ionic liquids using the example of 1-ethyl-3-methylimidazolium ethylsulfate. Phys Chem Chem Phys 15(42):18424–18436. https://doi.org/10.1039/C3CP52966E

Manna B, Ghosh A (2019) Dissolution of cellulose in ionic liquid and water mixtures as revealed by molecular dynamics simulations. J Biomol Struct Dyn 37(15):3987–4005. https://doi.org/10.1080/07391102.2018.1533496

Martínez L, Andrade R, Birgin EG, Martínez JM (2009) Packmol: a package for building initial configurations for molecular dynamics simulations. J Comput Chem 30(13):2157–2164. https://doi.org/10.1002/jcc.21224

Martyna G, Klein M, Tuckerman M (1992) Nosé-Hoover chains: The canonical ensemble via continuous dynamics. J Chem Phys 97:2535–2643. https://doi.org/10.1063/1.463940

Mohan M, Banerjee T, Goud VV (2016) Effect of protic and aprotic solvents on the mechanism of cellulose dissolution in ionic liquids: a combined molecular dynamics and experimental insight. ChemistrySelect 1(15):4823–4832. https://doi.org/10.1002/slct.201601094

Nanda S, Patra BR, Patel R, Bakos J, Dalai AK (2022) Innovations in applications and prospects of bioplastics and biopolymers: a review. Environ Chem Lett 20(1):379–395. https://doi.org/10.1007/s10311-021-01334-4

Nisa LdP, Mugang M, Bermejo MD, Segovia Puras JJ, Pérez E (2019) Density and melting points for the binary mixtures dimethyl sulfoxide (dmso) + 1-ethyl-3-methylimidazolium acetate and dmso + choline acetate. J Chem Eng Data 64(7):2923–2928. https://doi.org/10.1021/acs.jced.8b01048

Nose S (1984a) A molecular dynamics method for simulations in the canonical ensemble. Mol Phys 52:255–268. https://doi.org/10.1080/00268978400101201

Nose S (1984b) A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 81:511–519. https://doi.org/10.1063/1.447334

Olsson C, Westman G (2013) Direct dissolution of cellulose: Background, means and applications. Cellulose-Fundamental Aspects, pp 143–178. https://doi.org/10.5772/52144

Packer KJ, Tomlinson DJ (1971) Nuclear spin relaxation and self-diffusion in the binary system, dimethylsulphoxide (dmso)+ water. Trans Faraday Soc 67:1302–1314. https://doi.org/10.1039/TF9716701302

Paiva TG, Zanatta M, Cabrita EJ, Bernardes CE, Corvo MC (2022) Dmso/il solvent systems for cellulose dissolution: binary or ternary mixtures? J Mol Liq 345:117810. https://doi.org/10.1016/j.molliq.2021.117810

Paknejad A, Mohammadkhani R, Zarei H (2019) Experimental high-temperature, high-pressure density measurement and perturbed-chain statistical associating fluid theory modeling of dimethyl sulfoxide, isoamyl acetate, and benzyl alcohol. J Chem Eng Data 64(12):5174–5184. https://doi.org/10.1021/acs.jced.9b00396

Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Ståhlberg J, Beckham GT (2015) Fungal cellulases. Chem Rev 115(3):1308–1448. https://doi.org/10.1021/cr500351c

Plimpton S (1995) Fast parallel algorithms for short-range molecular dynamics. J Comp Phys 117:1–19. https://doi.org/10.1006/jcph.1995.1039

Ponder JW, Case DA (2003) Force fields for protein simulations. Adv Protein Chem 66:27–85. https://doi.org/10.1016/S0065-3233(03)66002-X

Radhi A, Le KA, Ries ME, Budtova T (2015) Macroscopic and microscopic study of 1-ethyl-3-methyl-imidazolium acetate-dmso mixtures. J Phys Chem B 119(4):1633–1640. https://doi.org/10.1021/jp5112108

Rapaport DC (1983) Hydrogen bonds in water. Mol Phys 50(5):1151–1162. https://doi.org/10.1080/00268978300102931

Roos E, Brehm M (2021) A force field for bio-polymers in ionic liquids (bilff) - part 1: Emimoac/water mixtures. PCCP 23(2):1242–1253. https://doi.org/10.1039/d0cp04537c

Roos E, Sebastiani D, Brehm M (2023b) A force field for bio-polymers in ionic liquids (bilff) - part 2: cellulose in emimoac/water mixtures. PCCP. https://doi.org/10.1039/D2CP05636D

Roos E, Sebastiani D, Brehm M (2023a) Bilff: All-atom force field for modeling triazolium- and benzoate-based ionic liquids. Molecules 28(22). https://doi.org/10.3390/molecules28227592

Ryckaert JP, Ciccotti G, Berendsen HJ (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23(3):327–341. https://doi.org/10.1016/0021-9991(77)90098-5

Sambasivarao SV, Acevedo O (2009) Development of opls-aa force field parameters for 68 unique ionic liquids. J Chem Theory Comput 5(4):1038–1050. https://doi.org/10.1021/ct900009a

Schneider T, Stoll E (1978) Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. Phys Rev B 17(3):1302–1322. https://doi.org/10.1103/PhysRevB.17.1302

Smith DGA, Burns LA, Patkowski K, Sherrill CD (2016) Revised damping parameters for the d3 dispersion correction to density functional theory. J Phys Chem Lett 7(12):2197–2203. https://doi.org/10.1021/acs.jpclett.6b00780

Strader ML, Feller SE (2002) A flexible all-atom model of dimethyl sulfoxide for molecular dynamics simulations. J Phys Chem A 106(6):1074–1080. https://doi.org/10.1021/jp013658n

Sun Y, Cheng J (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour Technol 83(1):1–11. https://doi.org/10.1016/S0960-8524(01)00212-7

Swatloski RP, Spear SK, Holbrey JD, Rogers RD (2002) Dissolution of cellulose with ionic liquids. J Am Chem Soc 124(18):4974–4975. https://doi.org/10.1021/ja025790m

The CP2K developers group (2017) Cp2k is freely available from http://www.cp2k.org/

Thomas B, Raj MC, B AK, H RM, Joy J, Moores A, Drisko GL, Sanchez C (2018) Nanocellulose, a versatile green platform: From biosources to materials and their applications. Chem Rev 118(24):11575–11625. https://doi.org/10.1021/acs.chemrev.7b00627

Tomimatsu Y, Suetsugu H, Yoshimura Y, Shimizu A (2019) The solubility of cellulose in binary mixtures of ionic liquids and dimethyl sulfoxide: influence of the anion. J Mol Liq 279:120–126. https://doi.org/10.1016/j.molliq.2019.01.093

VandeVondele J, Hutter J (2003) An efficient orbital transformation method for electronic structure calculations. J Chem Phys 118(10):4365–4369. https://doi.org/10.1063/1.1543154

VandeVondele J, Hutter J (2007) Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. J Chem Phys 127:114105. https://doi.org/10.1063/1.2770708

VandeVondele J, Krack M, Mohamed F, Parrinello M, Chassaing T, Hutter J (2005) Quickstep: fast and accurate density functional calculations using a mixed gaussian and plane waves approach. Comput Phys Commun 167(2):103–128. https://doi.org/10.1016/j.cpc.2004.12.014

Wanichapichart P, Taweepreeda W, Nawae S, Choomgan P, Yasenchak D (2012) Chain scission and anti fungal effect of electron beam on cellulose membrane. Radiat Phys Chem 81(8):949–953. https://doi.org/10.1016/j.radphyschem.2012.02.035

Wolfram Research, Inc (2010) Mathematica, version 8.0, champaign, il

Xu A, Zhang Y (2015) Insight into dissolution mechanism of cellulose in [c4mim][ch3coo]/dmso solvent by 13c nmr spectra. J Mol Struct 1088:101–104. https://doi.org/10.1016/j.molstruc.2015.02.031

Xu A, Zhang Y, Zhao Y, Wang J (2013) Cellulose dissolution at ambient temperature: role of preferential solvation of cations of ionic liquids by a cosolvent. Carbohydr Polym 92(1):540–544. https://doi.org/10.1016/j.carbpol.2012.09.028

Yang F, Wang X, Chen Q, Tan H (2019) Improvement of the properties of 1-ethyl-3-methylimidazolium acetate using organic solvents for biofuel process. J Mol Liq 284:82–91. https://doi.org/10.1016/j.molliq.2019.03.144

Yang M, Zhao W, Wang S, Yu C, Singh S, Simmons B, Cheng G (2019) Dimethyl sulfoxide assisted dissolution of cellulose in 1-ethyl-3-methylimidazoium acetate: small angle neutron scattering and rheological studies. Cellulose 26(4):2243–2253. https://doi.org/10.1007/s10570-018-2218-0

Yousefi Shivyari N, Tajvidi M, Bousfield DW, Gardner DJ (2016) Production and characterization of laminates of paper and cellulose nanofibrils. ACS 8(38):25520–25528. https://doi.org/10.1021/acsami.6b07655

Zhang X, Cheng Y, You J, Zhang J, Yin C, Zhang J (2022) Ultralong phosphorescence cellulose with excellent anti-bacterial, water-resistant and ease-to-process performance. Nat Commun 13(1):1117. https://doi.org/10.1038/s41467-022-28759-x

3.5 Article IV: BILFF: All-Atom Force Field for Modeling Triazolium- and Benzoate-Based Ionic Liquids

<u>E. Roos</u>, D. Sebastiani, M. Brehm, "BILFF: All-Atom Force Field for Modeling Triazolium- and Benzoate-Based Ionic Liquids", *Molecules*, 28, 7592, 2023

Author contributions: In this article, I performed all calculations, was responsible for data processing, data analysis, visualization of the results, and the writing of the original article draft. M. Brehm was responsible for the conceptualization of the project and for supervising the progress of the data processing. D. Sebastiani supervised the project and provided me with valuable advice during the writing of the manuscript. All authors contributed to further review iterations of the article.

Article

BILFF: All-Atom Force Field for Modeling Triazolium- and Benzoate-Based Ionic Liquids

Eliane Roos [1] , **Daniel Sebastiani** [1] **and Martin Brehm** [1,2,*]

1   Institut für Chemie—Theoretische Chemie, Martin-Luther-Universität Halle–Wittenberg,
    Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany; eliane.roos@gmx.de (E.R.);
    daniel.sebastiani@chemie.uni-halle.de (D.S.)
2   Department Chemie, Universität Paderborn, Warburger Straße 100, 33098 Paderborn, Germany
*   Correspondence: martin_brehm@gmx.de

Abstract: We present an extension of our previously developed all-atom force field BILFF (Bio-polymers in Ionic Liquids Force Field) to three different ionic liquids: 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]), 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]), and 1-ethyl-3-methylimidazolium benzoate ([EMIm][OBz]). These ionic liquids are of practical importance as they have the ability to dissolve significant amounts of cellulose even at room temperature. Our force field is optimized to accurately reproduce the strong hydrogen bonding in the system with nearly quantum chemical accuracy. A very good agreement between the microstructure of the quantum chemical simulations over a wide temperature range and experimental density data with the results of BILFF were observed. Non-trivial effects, such as the solvation shell structure and π–π stacking of the cations, are also accurately reproduced. Our force field enables accurate simulations of larger systems, such as solvated cellulose in different (aqueous) ionic liquids, and is the first to present the optimized parameters for mixtures of these solvents and water.

Keywords: ionic liquid; cellulose solvent; force field molecular dynamics simulation; solubility; hydrogen bond; solvation shell; imidazolium; triazolium; benzoate; acetate

1. Introduction

Ionic liquids (ILs) have become an attractive alternative to conventional solvents in various industries due to their unique properties like their non-volatility, non-flammability, and their very good thermal and chemical stability [1,2]. The properties of ILs can be tailored to suit specific applications by changing the cation or anion components [3,4]. This flexibility has led to their use in a range of different applications, including in catalysis, electrochemistry, material science, energy, and biotechnology [5–11]. One of the promising usage of ILs is the dissolution and processing of cellulose [12–17], which enables the production of cellulose-based derivatives in a homogeneous manner [18,19] and the processing of biomass by the separation of wood components [20,21].

Cellulose is characterized by its highly structured composition of glucose monomers, forming long chains that can range from several hundred to tens of thousands of units in length. These chains are highly ordered and arranged in parallel. Intramolecular hydrogen bonding stiffens the polymer chains, while intermolecular hydrogen bonding facilitates the arrangement of these linear polymers into sheet-like structures. These sheets are densely packed together through hydrophobic interactions and form crystalline structures [22,23]. To dissolve cellulose in ILs, the intermolecular hydrogen bonds between the cellulose strands must be disrupted and replaced by interactions with the solvent. This can be achieved through the formation of new hydrogen bonds to the solvent molecules. Several studies have demonstrated that the effectiveness of ILs in dissolving cellulose depends on several factors such as the type of cation and anion, the basicity as well as the position and length of the alkyl chains of the cation [15–17,24,25].

Therefore, the choice of the anion plays a more decisive role than the choice of the cation; the IL 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]), for example, shows one of the highest cellulose solubilities with 36.0 g per mol IL, while 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]) has a barely reduced solubility of 34.0 g per mol IL. In contrast, the ILs 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]) and 1-ethyl-3-methylimidazolium benzoate ([EMIm][OBz]) have decreased cellulose solubilities of 21.7 g per mol IL and 18.6 g per mol IL, respectively [15,16].

The two cations, [EMTr]$^+$ and [EMIm]$^+$, are distinguished by the replacement of one ring carbon atom in [EMTr]$^+$ by a third nitrogen atom. As a result, the third ring proton in the 1,2,3-triazolium is missing. This structural change leads to a lower probability of proton abstraction and thus to a lower formation of *N*-heterocyclic carbenes (NHC). NHCs are highly reactive and form dimers or undesirable byproducts. Thus, they can also cause the unwanted decomposition of cellulose [26]. The potential use of triazolium-containing ILs as non-derivatizing cellulose solvents is nearly unexplored and was first described by Brehm et al. in 2019 [15]. By understanding the complex mechanisms involved in cellulose dissolution in different ILs, it is possible to develop more efficient and sustainable processes for the production of cellulose-based materials. This requires a thorough understanding of the interactions between ILs and cellulose, which can be achieved through molecular dynamics (MD) simulations. However, the complexity of the system, together with the problem that solubility processes take place on time scales that exhaust the resources of quantum chemical MD simulations, necessitates the use of force field MD (FFMD) simulations. To accurately simulate the interactions between ILs and cellulose, it is crucial to have well-optimized force field parameters for the individual ions. Until now, there has been no optimized force field available for the ions [EMTr]$^+$ and [OBz]$^-$, which limited the accuracy of simulations involving these ILs. In this article, we present the development of a force field for (aqueous) [EMTr][OAc], [EMIm][OBz] and their combination [EMTr][OBz], which is an extension of our previously published force field for [EMIm]$^+$ and [OAc]$^-$ called BILFF (Bio-Polymers in Ionic Liquids Force Field) [27,28]. The force field development is focused on the accurate reproduction of the microstructure of the ILs, especially the hydrogen bonds, based on results of ab initio MD (AIMD) simulations. To optimize the force field parameters, we compared the results of different analysis like radial and combined distribution functions of the hydrogen bonds along with statistical analyses of bond lengths, angles, and dihedral angles with the results from AIMD simulations for both the pure and aqueous ILs. Based on this comparison, the force field parameters were adjusted and optimized for a simulation of the ILs both in the absence and presence of water at 350 K. For the aqueous system, a ratio of IL ion pairs to water of 1:3 was chosen, at which cellulose is already precipitated again [29–31].

The results demonstrate that the new force field effectively reproduces the structural and dynamic properties of the ILs, including ion pairing, solvation shell structure, and hydrogen bonding with respect to their geometry and lifetime, over a wide temperature range. Furthermore, we contrasted and compared the results of the three ILs [EMTr][OAc], [EMIm][OBz] and [EMTr][OBz] with experimental density data, as well as with the previously published results for [EMIm][OAc].

The optimized force field parameters enable the simulation of larger cellulose systems in various ILs on larger time scales, which would otherwise be too computationally expensive when using AIMD simulations. This enables a deeper understanding of the underlying interactions for the solubility of cellulose in the different ILs and thus the identification of novel and improved cellulose solvents.

2. Optimization Procedure

In a previously published article [27], the optimized force field parameters for (aqueous) [EMIm][OAc] were presented. The aim of this article is the extension of the parameter set to the ions [EMTr]$^+$ and [OBz]$^-$ using the already optimized parameter set for

[OAc]$^-$ and [EMIm]$^+$. For this purpose, simulations of [EMTr][OAc] and [EMIm][OBz] were performed. Since it is known that many imidazolium- and triazolium-based ionic liquids are hygroscopic and that this might drastically affect their properties [32,33], the optimization of the force field parameters was carried out simultaneously in the presence and absence of water. For this purpose, simulations of pure [EMIm][OBz] and [EMTr][OAc] as well as the combination of both ions as [EMTr][OBz] in water, were performed (cf. Figure 1). The latter simulation is used to investigate both parameter sets of [EMIm][OBz] and [EMTr][OAc] in the presence of water. The optimization of the force field parameters was performed by comparing various analyses of force field MD simulations with reference AIMD simulations of these three systems, which are discussed in more detail in the following sections.

[EMTr][OAc]

[EMIm][OBz]

[EMTr][OBz] + H$_2$O
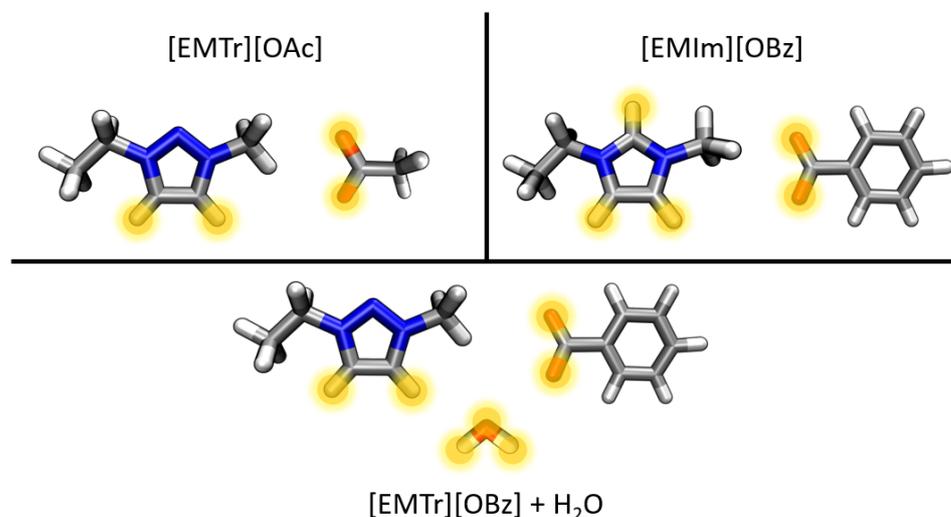
Figure 1. Overview of the investigated systems containing the ions 1-ethyl-3-methyl-1,2,3-triazolium ([EMTr]$^+$), 1-ethyl-3-methylimidazolium ([EMIm]$^+$), benzoate ([OBz]$^-$), and acetate ([OAc]$^-$) in this article with highlighted relevant hydrogen bonds donors and acceptors (atom color code: blue—N; grey—C; red—O; white—H).

The aim of optimizing the force field parameters for [EMTr]$^+$ and [OBz]$^-$ in the presence and absence of water is to enable force field MD simulations of these ions with nearly the accuracy of quantum chemical calculations to compute larger systems with less computational effort.

The focus was set on the correct reproduction of the hydrogen bonds between the investigated ions [EMTr]$^+$ as well as [OBz]$^-$ and their counterions [OAc]$^-$ and [EMIm]$^+$ calculated from quantum chemical simulations. [EMIm]$^+$ and [OAc]$^-$ were chosen as counterions for [OBz]$^-$ and [EMTr]$^+$, since optimized force field parameters are already available for these two ions (see our previously published articles on BILFF [27,28]).

The hydrogen bonds in the aqueous system were investigated using an AIMD simulation of aqueous [EMTr][OBz]. A molar fraction of $\chi_{IL} = 0.25$ was simulated in order to cover a wide range of applications as it is known that, at this water content, cellulose is already precipitated again in other ILs such as [EMIm][OAc] [29–31].

The force field is based on the functional form of the OPLS–AA force field [34–36]:

$$U(r^N) = \sum_{i \in \text{bonds}} k_{l,i}(l_i - l_{i,0})^2 + \sum_{i \in \text{angles}} k_{\theta,i}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{i \in \text{dihedrals}} \left[\frac{V_{i,1}}{2}[1 + \cos(\phi_i)] + \frac{V_{i,2}}{2}[1 - \cos(2\phi_i)] \right.$$

$$\left. + \frac{V_{i,3}}{2}[1 + \cos(3\phi_i)] + \frac{V_{i,4}}{2}[1 - \cos(4\phi_i)] \right]$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}} \right] f_{ij}$$

(1)

$$\epsilon_{i,j} = \sqrt{\epsilon_i \epsilon_j}, \qquad \sigma_{i,j} = \sqrt{\sigma_i \sigma_j}.$$

(2)

In order to optimize the force field parameters, radial distribution functions of the reference AIMD simulations were compared with the results of the force field MD simulations. To retain the properties of the underlying force fields as much as possible, the force field parameters q and σ were iteratively adjusted using a trial-and-error approach. The force field parameters were varied until the deviations of the RDFs were minimized. The results are discussed in Section 2.1.1. In addition, σ was optimized with respect to a good reproduction of the experimental density (see Table 7).

The bonded interactions were adjusted based on a comparison of the statistical occurrence of bond lengths, bond angles and dihedral angles, and by varying the force constants and height of the potential barrier. The nomenclature of the atom types can be found in Figure 2.
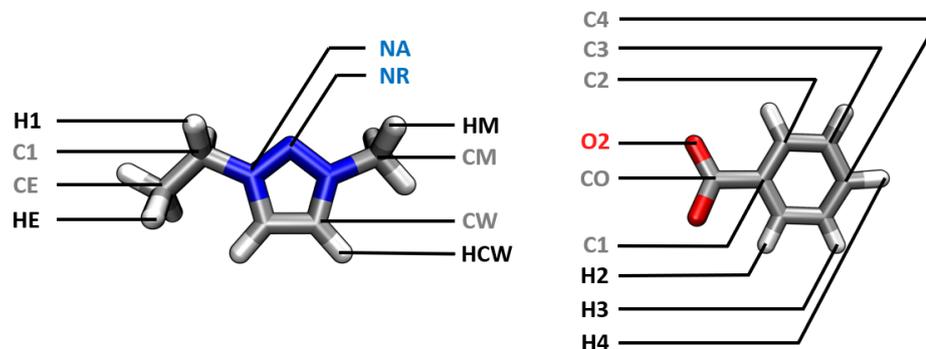
Figure 2. Nomenclature of the atom types of [EMTr]$^+$ and [OBz]$^-$ in our force field BILFF.

To correctly reproduce the reference AIMD simulations, the equilibrium bond length of the NA–NR bond in [EMTr]$^+$, for example, was modified by +0.3%. The NA–CW bond length was reduced by about 2%, while the bond force constant was increased by about 12% compared to the value in [EMIm]$^+$.

In [OBz]$^-$, the greatest adjustment in the bond length and force constant occurred in the CO–O2 bond, with changes of about +2% and +56%, respectively. The equilibrium bond angle, on the other hand, was not changed for both ions.

A flowchart of the force field optimization process can be found in Figure 3. The force field parameters and nomenclature for the atom types and classes can be found in Tables 1–5.
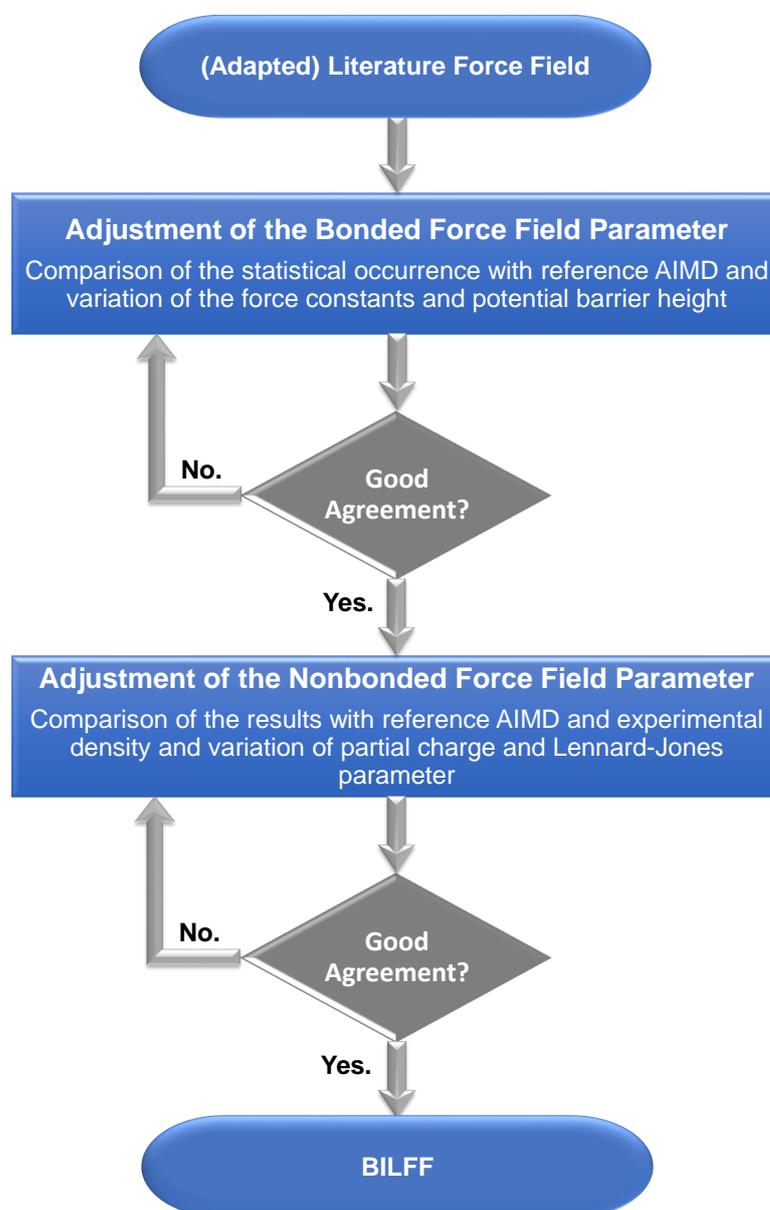
Figure 3. Flowchart of the optimization process of the force field parameters for BILFF. As a starting point for the optimization of the force field parameters, BILFF [27,28] for [EMTr]$^{+}$ was used. For [OBz]$^{-}$ the Lennard-Jones parameters and force constants were adapted from OPLS–AA [34–36]. The partial charges were calculated using the restrained electrostatic potential (RESP) methodology. The equilibrium bond length was calculated using a geometry optimization.

Table 1. Nomenclature of the atom types and atom classes of [EMTr]$^+$ and [OBz]$^-$ in BILFF. The atom types are applied for the non-bonded interactions (see Table 2), while the atom classes are used for the bonded interactions (see Tables 3–5).

Atom Type	Atom Class
[EMTr]$^+$	
C1	CT
CE	CT
CM	CT
CW	CW
HCW	HA
H1	HC
HE	HC
HM	HC
NR	NR
NA	NA
[OBz]$^-$	
C1	CA
C2	CA
C3	CA
C4	CA
CO	CO
H2	HA
H3	HA
H4	HA
O2	O2

Table 2. Optimized atomic partial charges q and Lennard-Jones parameters σ and ϵ of [EMTr]$^+$ and [OBz]$^-$ in BILFF.

Atom Type	q / e	σ / Å	ϵ / kJ mol^{-1}
[EMTr]$^+$			
C1	−0.187	3.34	0.2760
CE	−0.054	3.34	0.2760
CW	−0.144	3.38	0.2930
HCW	0.191	1.48	0.1260
HC	0.070	2.38	0.1260
H1	0.148	2.38	0.1260
NR	−0.204	3.10	0.7110
NA	0.204	3.10	0.7110
[OBz]$^-$			
C1	0.005	3.70	0.2929
C2	−0.118	3.70	0.2929
C3	−0.121	3.70	0.2929
C4	−0.299	3.70	0.2929
CO	0.398	3.90	0.4393
H2	0.070	2.42	0.1255
H3	0.157	2.42	0.1255
H4	0.200	2.42	0.1255
O2	−0.550	2.80	0.8786

Table 3. Optimized bond equilibrium lengths l_0 and force constants k_1 of [EMTr]$^+$ and [OBz]$^-$ in BILFF.

Bond	l_0 / Å	k_1 / kJ mol^{-1} Å^{-2}
	[EMTr]$^+$	
NA–NR	1.344	3199.2
CW–HA	1.088	2633.8
CW–NA	1.375	3108.7
CW–CW	1.386	3773.2
NA–CT	1.488	2046.3
HC–CT	1.099	2679.4
CT–CT	1.533	2125.5
	[OBz]$^-$	
CA–CA	1.387	3274.1
CA–HA	1.088	2707.4
CA–CO	1.504	1906.9
CO–O2	1.282	4273.1

Table 4. Optimized angle equilibrium values θ_0 and force constants k_θ of [EMTr]$^+$ and [OBz]$^-$ in BILFF.

Angle	θ_0 / Deg	k_θ / kJ mol^{-1} rad^{-2}
	[EMTr]$^+$	
CW–NA–NR	112.1	568.7
NR–NA–CT	118.6	396.5
NA–NR–NA	104.4	610.1
NA–CT–CT	110.9	361.2
NA–CW–CW	107.0	579.7
NA–CW–HA	120.8	200.9
CW–CW–HA	131.7	190.5
NA–CT–HC	107.2	375.9
CT–CT–HC	111.4	296.2
HC–CT–HC	109.2	226.5
CW–NA–CT	125.2	242.9
	[OBz]$^-$	
CA–CA–CA	120.0	446.0
CA–CA–HA	120.0	258.1
CA–CA–CO	120.0	397.6
CA–CO–O2	117.0	550.2
O2–CO–O2	126.0	735.9

2.1. Microstructure of the Systems

2.1.1. Radial Distribution Functions

In the following, the radial distribution functions (RDFs) in the different ILs are compared with the results of the corresponding AIMD simulations. Furthermore, conclusions are drawn about the differences in the formation of hydrogen bonds between the cations [EMTr]$^+$ and [EMIm]$^+$ and their respective bonding partners, as well as between the anions [OBz]$^-$ and [OAc]$^-$ and their respective hydrogen bond donors. First, the hydrogen bonds of the cation [EMTr]$^+$ are analyzed (see Figures 4–6). The ring protons of [EMTr]$^+$ form a strong hydrogen bond to the anion [OAc]$^-$ with a particle distance of 195 pm (cf. Figure 4) and a $g(r) = 5.4$, as can be seen from the first maximum of the RDF. The agreement of the RDF with the AIMD result is excellent. To achieve this, the partial charge of the ring protons

was increased by about 27%, from 0.150 to 0.191, compared to the force field parameters for [EMIm]$^+$ [27]. The value of σ was decreased from 1.62 Å to 1.48 Å by about 12%, reducing the atom repulsion.

Table 5. Optimized torsional coefficients V_n of [EMTr]$^+$ and [OBz]$^-$ in BILFF.

Torsion Angle	V_1 / kJ mol^{-1}	V_2 / kJ mol^{-1}	V_3 / kJ mol^{-1}	V_4 / kJ mol^{-1}
	[EMTr]$^+$			
CW–NA–NR–NA	0.0000	19.4600	0.0000	0.0000
CT–NA–NR–NA	0.0000	19.4600	0.0000	0.0000
NR–NA–CW–CW	0.0000	12.5500	0.0000	0.0000
NR–NA–CW–HA	0.0000	12.5500	0.0000	0.0000
NR–NA–CT–HC	0.0000	0.0000	0.0000	0.0000
NR–NA–CT–CT	0.1000	1.0000	0.1000	−0.3000
CT–NA–CW–CW	0.0000	12.5500	0.0000	0.0000
CT–NA–CW–HA	0.0000	12.5500	0.0000	0.0000
NA–CW–CW–NA	0.0000	65.0000	0.0000	0.0000
NA–CW–CW–HA	0.0000	44.9800	0.0000	0.0000
HA–CW–CW–HA	0.0000	30.0000	0.0000	0.0000
CW–NA–CT–HC	0.1000	0.2000	0.0000	0.0000
CW–NA–CT–CT	0.4000	1.0000	0.0000	0.2000
NA–CT–CT–HC	0.0000	0.0000	0.3670	0.0000
HC–CT–CT–HC	0.0000	0.0000	1.2552	0.0000
	[OBz]$^-$			
CA–CA–CA–CA	0.0000	30.334	0.0000	0.0000
HA–CA–CA–CA	0.0000	30.334	0.0000	0.0000
HA–CA–CA–CO	0.0000	30.334	0.0000	0.0000
HA–CA–CA–HA	0.0000	30.334	0.0000	0.0000
CA–CA–CA–CO	0.0000	30.334	0.0000	0.0000
CA–CA–CO–O2	0.0000	8.000	0.0000	0.0000

Figure 4. Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OAc]$^-$ and the ring protons of [EMTr]$^+$ calculated from a reference AIMD and force field MD simulations using adapted literature force field parameters [37–39] and BILFF. The RDFs are averaged over both ring protons.
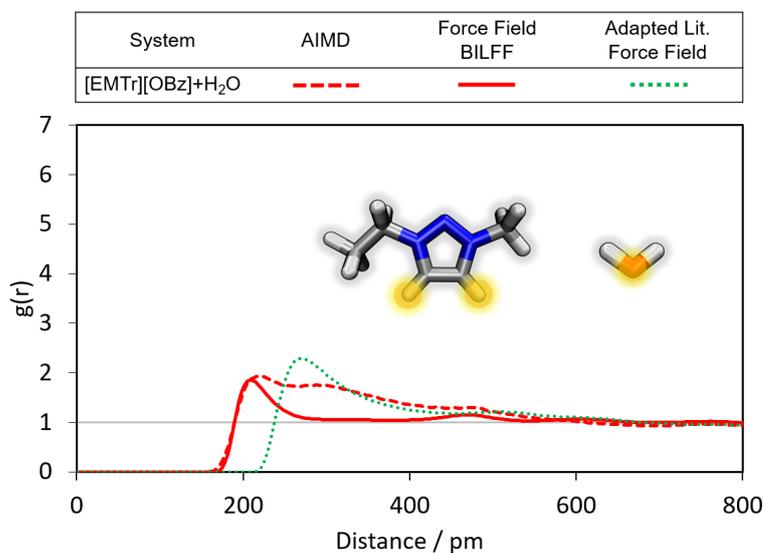
Figure 5. Comparison of the RDFs of the hydrogen bond between the marked oxygen atom of water and the ring protons of [EMTr]$^+$ calculated from a reference AIMD and force field MD simulations of aqueous [EMTr][OBz] using adapted literature force field parameters [34–39] and BILFF. The RDFs are averaged over both ring protons.
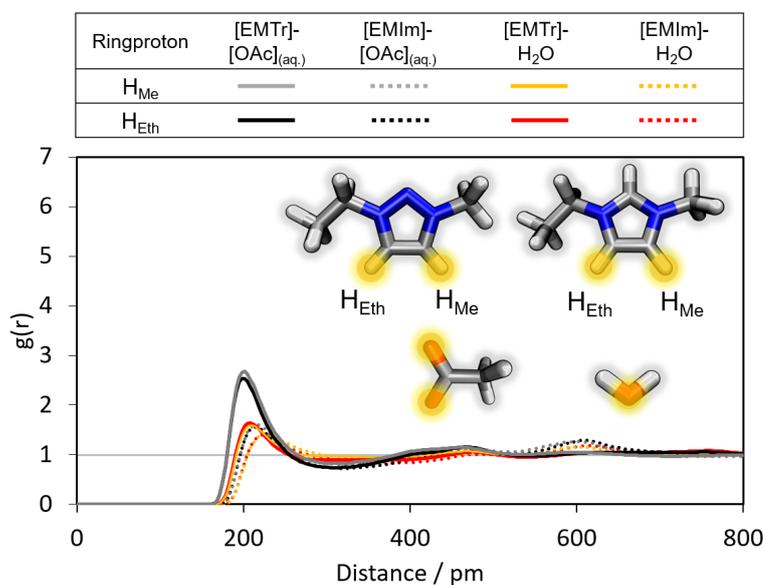
Figure 6. Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OAc]$^-$/water and the ring protons of [EMIm]$^+$/[EMTr]$^+$ calculated from force field MD simulations of aqueous [EMTr][OAc] and [EMIm][OAc] using BILFF.

According to the height of the RDFs (Figures 5 and 6), the hydrogen bond between [EMTr]$^+$ and the oxygen of water is comparatively weaker than the hydrogen bond between [EMTr]$^+$ and [OAc]$^-$.

A comparison of the hydrogen bond of [EMTr]$^+$ and H$_2$O in aqueous [EMTr][OBz] between an AIMD and force field MD simulation with BILFF is shown in Figure 5. It should be noted that the force field parameters were specifically adjusted for [EMTr]$^+$ in [EMTr][OAc] and [OBz]$^-$ in [EMIm][OBz], but not for the combination of both ions as [EMTr][OBz]. However, the results of the RDF calculated from AIMD simulations

can still be well reproduced for the $[EMTr]^+ \cdots H_2O$ hydrogen bond, indicating the good transferability of the force field parameters for both molecules.

When comparing our results with the adapted literature force fields for $[EMTr]^+$ [37–39] and $[OBz]^-$ [34–36] large deviations with respect to the results of the AIMD become visible, which underlines the importance of an optimization of the force field parameters for a correct reproduction of the hydrogen bonds.

Additionally, it is interesting to compare the strength of the cation–anion/cation–water hydrogen bonds of the two different aqueous ILs [EMTr][OAc] and [EMIm][OAc] shown in Figure 6. For $[EMTr]^+$, the first maximum of $g(r)$ is observed with a higher intensity at slightly smaller particle distances for both types of hydrogen bonds (cation–anion and cation–water). The ring protons of the cation $[EMTr]^+$ thus form stronger hydrogen bonds to both $[OAc]^-$ and water than the cation $[EMIm]^+$ even though the 1,2,3-triazolium cation has a considerably weaker acidity than imidazolium ($pK_A([EMTr][OAc]) = 24$, $pK_A([EMIm][OAc]) = 20$–23) [15]. The trend persists upon an exchange of the anion ($[OAc]^-/[OBz]^-$) (cf. Figures S1 and S2 in the Supplementary Materials). One reason is the 27% higher partial charge of the ring protons of the triazolium cation due to the additional nitrogen atom in the ring in comparison to imidazolium. The reduced number of ring protons in $[EMTr]^+$ also affects the competition for hydrogen bond donors, increasing the intensity of each individual interaction.

Furthermore, it was investigated whether the two aromatic ring protons in both cations differ in terms of the intensity of hydrogen bonding. The corresponding RDFs are shown in Figure 6 (black/grey curve and red/orange curve). An almost identical behavior of the curves can be observed, which occurs due to the equivalence of the ring protons in each of the two molecules.

In the following, the hydrogen bonds of the second considered ion $[OBz]^-$ are analyzed (cf. Figures 7–9). With a $g(r)$ of 3.7 and a particle distance of 205 pm in the first maximum of the RDF, the anion forms strong hydrogen bonds to $[EMIm]^+$ (cf. Figure 7). By increasing the partial charge of the oxygen atom of $[OBz]^-$ by about 5% from -0.524 to -0.550, and reducing the atom repulsion by modifying σ by about 5% from 2.96 Å to 2.80 Å compared to the adapted force field parameter of OPLS–AA [34–36] (cf. Table S1), an excellent agreement with the AIMD simulation was obtained. In contrast, the adapted literature force field [34–36] shows slight deviations in the position and height of the first maximum.
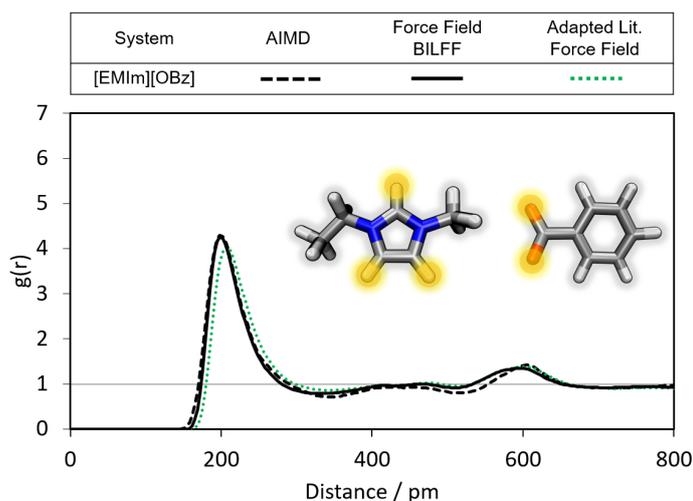
Figure 7. Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of $[OBz]^-$ and the ring protons of $[EMIm]^+$ calculated from a reference AIMD and force field MD simulations using adapted literature force field parameters [34–36] and BILFF. The RDFs are averaged over all three ring protons.
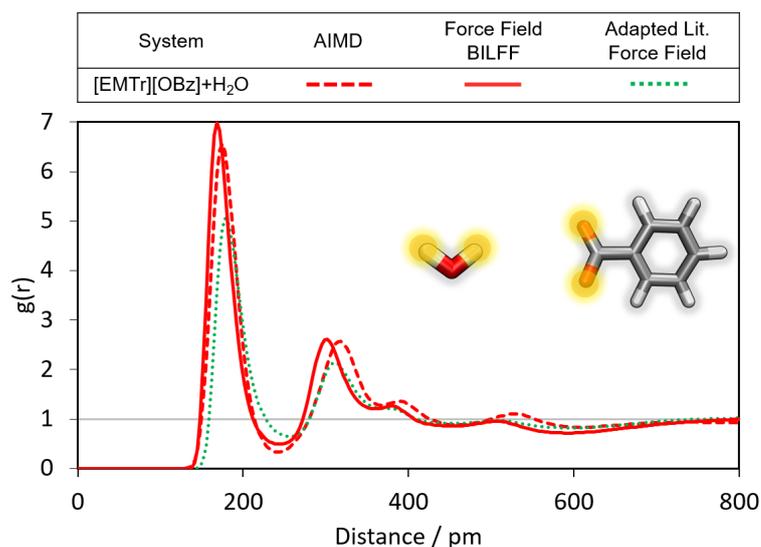
Figure 8. Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OBz]$^-$ and the protons of water calculated from a reference AIMD and force field MD simulations using adapted literature force field parameters [34–36] and BILFF.
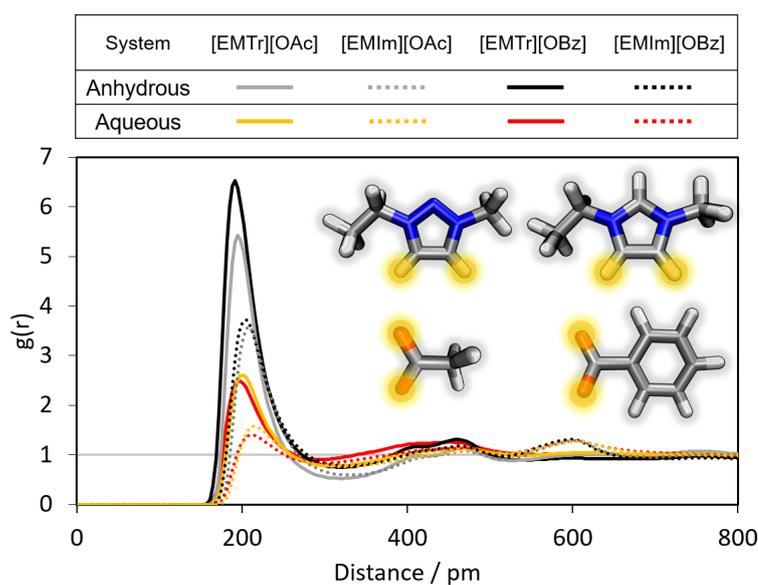
Figure 9. Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OBz]$^-$ as well as [OAc]$^-$ and the ring protons of [EMIm]$^+$ and [EMTr]$^+$ calculated from force field MD simulations using BILFF. The RDFs are averaged over the marked ring protons.

However, benzoate forms an even stronger hydrogen bond to the protons of water than to [EMIm]$^+$ (cf. Figures 7 and 8). Besides the obvious difference in the O–H vs. C–H dipole moment of water and [EMIm]$^+$, another reason can be steric effects. The oxygen atoms of benzoate are more easily accessible by the small water molecules than by the large imidazolium cation.

In comparison, the adapted literature force field gives rise to a significantly weaker peak intensity of the corresponding RDF (cf. Figures 7 and 8, green curve).

In addition, a comparison of the cation–anion hydrogen bonding intensities of all four ion pair combinations is analyzed in Figure 9. The first maximum in the RDF of $[OBz]^- \cdots [EMIm]^+$ is slightly increased compared to $[OAc]^- \cdots [EMIm]^+$ in the anhydrous system (black and grey curves). This indicates that $[OBz]^-$ forms slightly stronger hydrogen bonds to the cation than $[OAc]^-$, as already observed before [40]. We can therefore confirm the previous reports in the literature. This effect is even more pronounced in the triazolium-containing ILs with a difference in the first maximum of a $g(r)$ of 1.1. The same trend is observed when comparing the corresponding RDFs of the AIMD simulations (cf. Figure S3 in the Supplementary Materials).

In the presence of water, the respective cation$\cdots$anion hydrogen bonds of both anion species are weakened (cf. Figure 9, red and orange curves). This is caused by the additional formation of an anion$\cdots$water hydrogen bond, which is much stronger than the cation–anion hydrogen bond (cf. Figure S4 in the Supplementary Materials).

Due to the stronger hydrogen bond, the hydration shell around $[OBz]^-$ is more pronounced than that of $[OAc]^-$. As a result, there is a greater attenuation of the cation$\cdots [OBz]^-$ interaction relative to the cation$\cdots [OAc]^-$ hydrogen bond in the presence of water.

The hydration shell around the ions, as well as the coordination numbers, are discussed in Section 2.2.

In summary, the following trends are thus obtained from the comparison of the RDFs of the different ILs:

- Triazolium forms stronger hydrogen bonds with both the anion and water than imidazolium.
- Benzoate forms stronger hydrogen bonds with the cation than acetate in the anhydrous system.
- Benzoate forms stronger hydrogen bonds with water than acetate, resulting in a greater attenuation of the cation–anion interaction in the aqueous system.

2.1.2. Combined Distance–Angle Distribution Functions

While radial distribution functions provide an important initial assessment of the quality of the simulated hydrogen bonding structure, the directional aspect should not be neglected.

Figure 10 shows the combined distance–angle distribution functions (CDFs) of the distance $H_{EMIm} \cdots O_{OBz}$ and the corresponding angle $\angle(CH_{EMIm}, H_{EMIm}, O_{OBz})$ as illustrated in the sketch. The overall pattern of the CDFs of the AIMD simulation (top) and the FFMD simulation with BILFF (bottom) agree very well. The main hydrogen bonding peak at a particle distance of about 200 pm shows a broader angular distribution in the BILFF simulation, which may be the result of better statistical sampling compared to the AIMD simulation.

The second peak at about 400 pm originates from both the "second" oxygen of $[OBz]^-$ and the "second" hydrogen of the $[EMIm]^+$ molecule. Again, the AIMD simulation yields a slightly stronger localization than the force field.

The third peak, observed at approximately 600 pm, occurs at an angle of around 15° and corresponds to the hydrogen bonding involving the isolated CH group, leading to an increased presence of benzoate ions on the "upper" side of the molecule.

Considering the CDFs of the 1,2,3-triazolium cation and the anions acetate and benzoate, a similar good agreement of the CDFs of the AIMD and FFMD simulations can be obtained (cf. Figures S8 and S9 in the Supplementary Materials). Additional CDFs regarding the cation$\cdots$anion, cation$\cdots$cation, and anion$\cdots$water interactions in [EMIm][OAc] can be found in the previously published article [27].

However, for a correct reproduction of the microstructure of the AIMD simulation, not only the hydrogen bonds are important, but also the description of the π–π interactions between the aromatic rings of $[EMTr]^+$ and $[OBz]^-$.
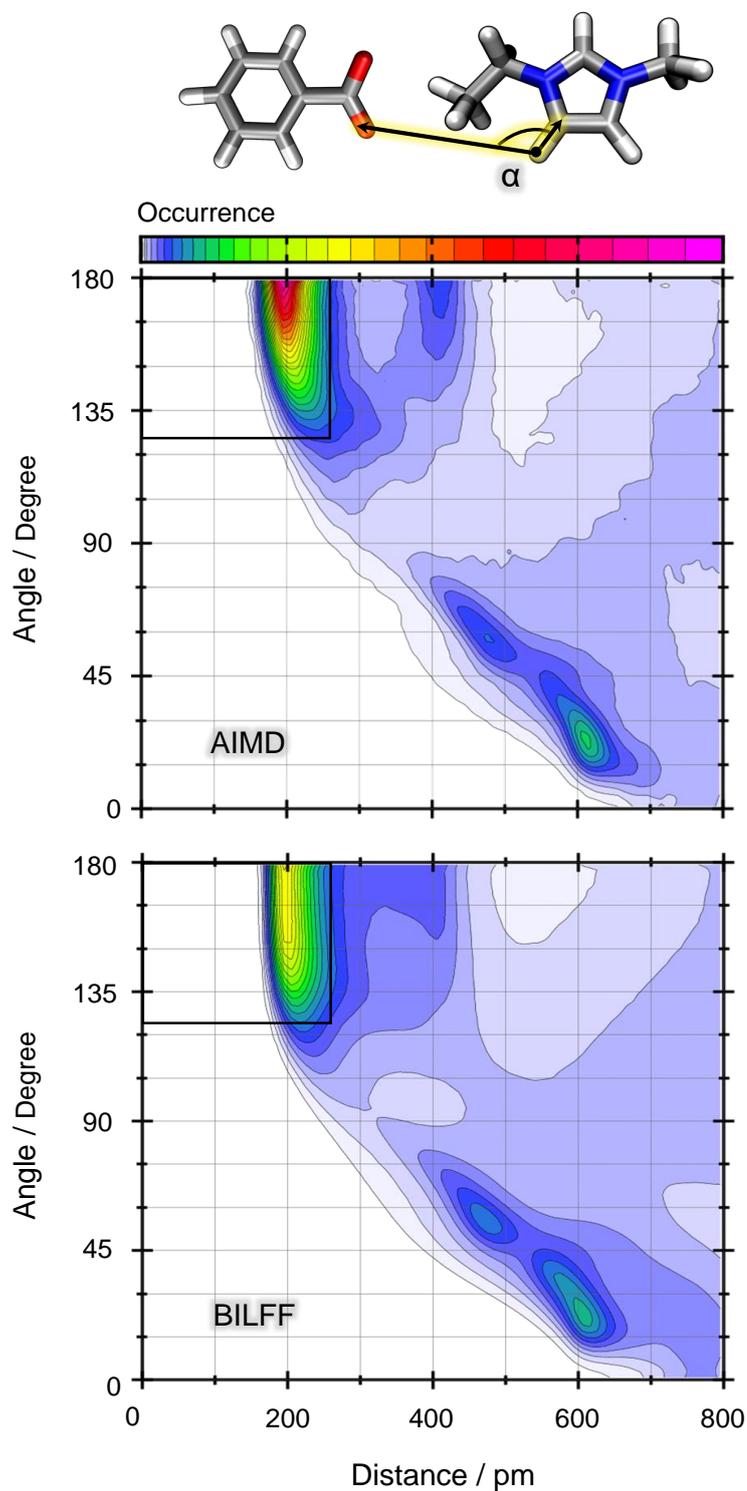
Figure 10. Distance–angle distribution function between an example ring proton of [EMIm]$^+$ and the oxygen atoms of [OBz]$^-$ in pure [EMIm][OBz] as a result of a reference AIMD simulation (**top**) and a force field MD simulation with BILFF (**bottom**). Color code in arbitrary units. The black rectangle demonstrates the geometric criterion for the existence of a hydrogen bond, as used in the lifetime calculation.

Considering the homomolecular distance–angle distribution functions of the aromatic cations, [EMTr]$^+$ and [EMIm]$^+$ (cf. Figure 11, left panel and Ref. [27]), symmetric peaks are observed at 400 pm, 0° and 400 pm, 180°, but no residence probability is found around 400 pm, 90°. This pattern indicates the occurrence of π–π stacking. The absence of a significant probability for the cations being at a 90° angle to each other implies that a T-shaped arrangement between the cations is unlikely.
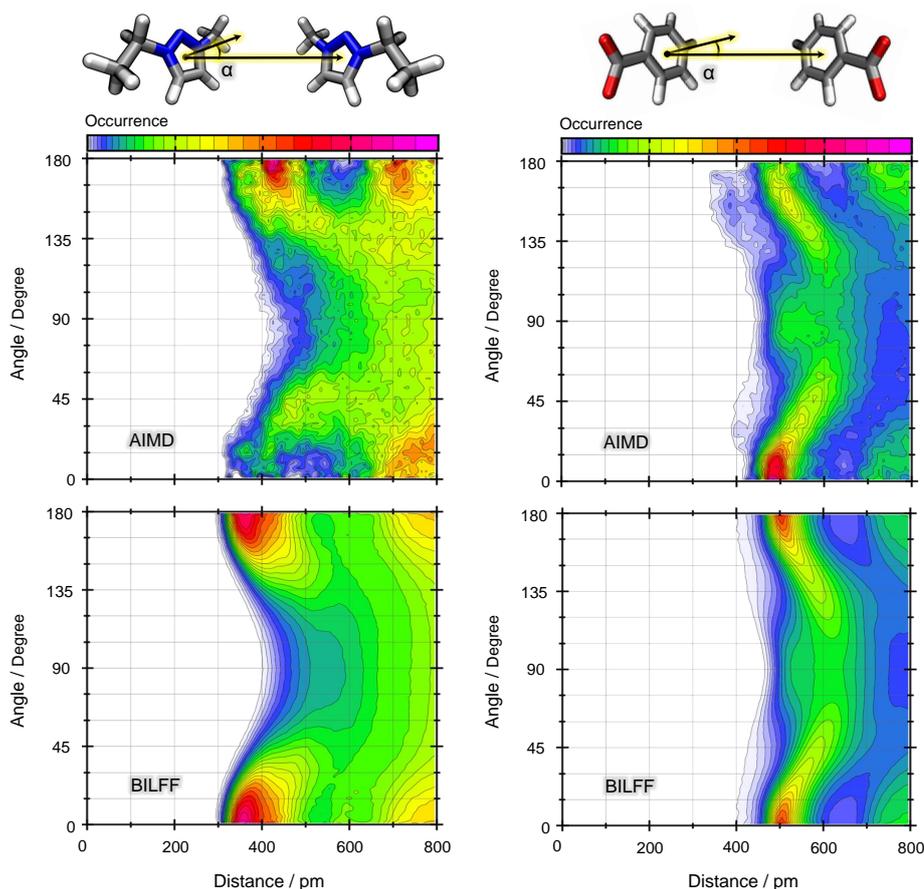
Figure 11. Distance–angle distribution function between two [EMTr]$^+$ (**left**) and two [OBz]$^-$ (**right**) ring centers in aqueous [EMTr][OBz] as a result of a reference AIMD simulation (**top**) and a force field MD simulation with BILFF (**bottom**). Color code in arbitrary units.

In contrast to this, the corresponding CDF of [OBz]$^-$ $\cdots$ [OBz]$^-$ (cf. Figure 11; right panel) shows two maxima at 500 pm for angles of 0° and 180°, along with a broad distribution in the center region (500 pm, 90°). The unusually large distance between the molecules combined with the higher probability of $\alpha = 90°$ shows that the coplanar [OBz]$^-$ $\cdots$ [OBz]$^-$ motif is no longer the dominant structure (as opposed to either cations). Instead, they are twisted relative to each other, adopting a T-shaped arrangement rather than a stacked configuration.

In both the anion$\cdots$anion and cation$\cdots$cation CDFs, the results obtained from FFMD simulations exhibit a good agreement with AIMD simulations. It is important to note here that the force field parameters were developed for [EMTr]$^+$ in [EMTr][OAc] and [OBz]$^-$ in [EMIm][OBz] but not for the combination [EMTr][OBz]. The fact that the microstructure of the AIMD simulation can nevertheless be reproduced well even in the case of the cross combination [EMTr][OBz] in water shows a good transferability of the force field parameters. This aspect will be discussed in more detail in Section 2.3.3.

2.1.3. Spatial Distribution Functions

Providing valuable insights into the overall structural arrangement, spatial distribution functions (SDF) provide a comprehensive picture of the relative orientations of molecules within the complex system involving the different cations, anions, and water. Figure 12 shows the results of the force field MD simulation using BILFF for all four systems. The oxygen atoms of water and the anion arrange themselves competitively at the respective ring protons of the cation. Comparing the imidazolium and triazolium patterns, the isolated aromatic ring proton of $[EMIm]^+$ acts as a strong attractor for the oxygen atoms, resulting in deflation around the aliphatic side chains, while the substituted nitrogen is avoided, resulting in a much more pronounced oxygen density around the methyl/ethyl side chains.
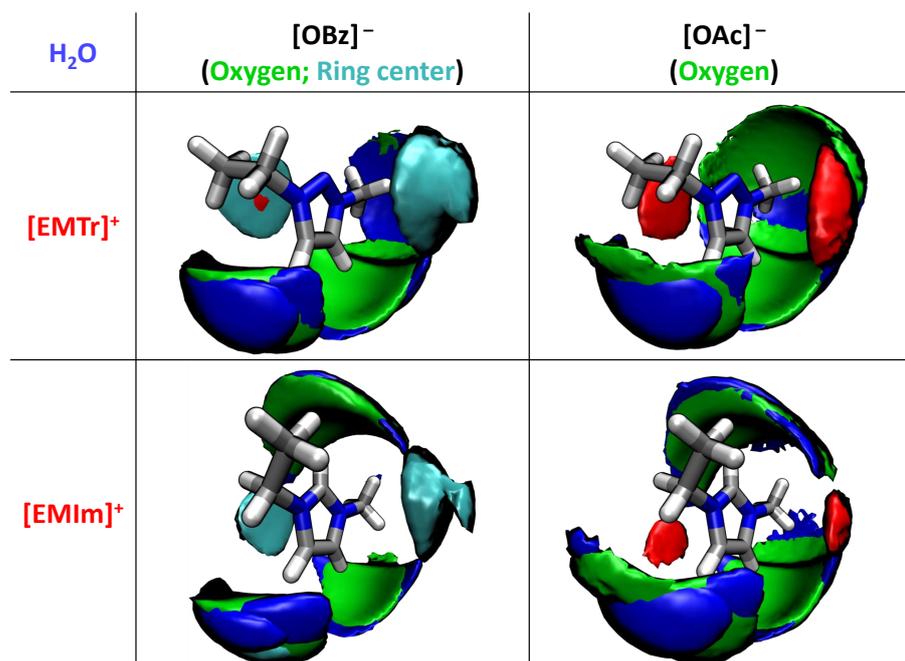
Figure 12. Spatial distribution function of the arrangement of molecules around the two different cations with the protons and oxygen atoms of water (blue, $26\,nm^{-3}$) and the oxygen atom of the anion (green, $7\,nm^{-3}$) as well as the ring center of the cation (red, $7\,nm^{-3}$) and anion (cyan, $10\,nm^{-3}$) in all four systems calculated with BILFF.

The spatial distribution function also shows distinct π–π stacking motifs above/below the aromatic cations. Interestingly, the stacking partner is determined by the anion type; for $[OAc]^-$ a homomolecular cation–cation stacking (red areas in Figure 12) is observed, while for $[OBz]^-$ the corresponding regions are preferentially occupied by the benzoate anions (cyan areas). The ring stacking of the benzoate is in agreement with results in the literature [16]. The results from the force field MD simulation of aqueous [EMTr][OBz] also agree well with the results obtained in the reference AIMD simulation (cf. Figure S10 in the Supplementary Materials).

2.2. Competing Hydrogen Bonds

In the ternary mixtures of ionic liquids in water, all direct interactions between the ions are brought into competition with hydrogen bonding interactions with water. The topology of the interaction pattern is analyzed in the form of a Sankey diagram (cf. Figure 13). The coordination numbers were calculated from the integral under the curve of the corresponding RDFs to the first minimum. These numbers indicate the number of hydrogen bonds per atom, which is proportional to the thickness of the bars in the Sankey diagram.
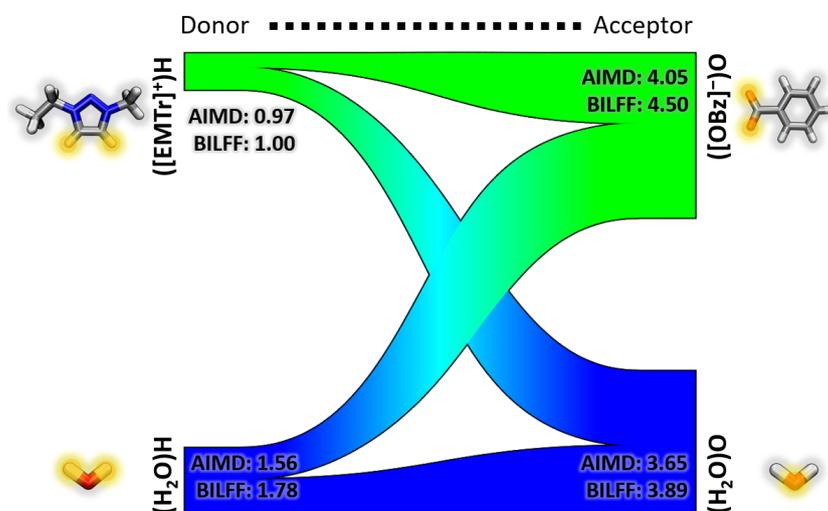
Figure 13. Visualization of the competing hydrogen bonds in aqueous [EMTr][OBz] as Sankey diagram calculated from a force field MD simulation with BILFF. The inserted numbers represent coordination numbers and are shown comparatively to the AIMD simulation.

The triazolium cation exhibits a coordination number of 1.0 in aqueous [EMTr][OBz] (cf. Figure 13), and 1.21 in aqueous [EMTr][OAc] (cf. Figure S11 in the Supplementary Materials). The imidazolium cation has similar coordination numbers of 0.95 in aqueous [EMIm][OBz] and 1.15 in aqueous [EMTr][OAc] (cf. Figures S12 and S13 in the Supplementary Materials).

Considering the anions, benzoate is surrounded by about 4.5 other molecules of the cation or water in the corresponding ILs (cf. Figures 12 and S12 in the Supplementary Materials). The smaller acetate, on the other hand, has a slightly higher coordination number of about 5 (cf. Figures 12 and S12 in the Supplementary Materials).

Compared to the results for [EMTr][OBz] from the AIMD simulation, the coordination numbers of BILFF can be well reproduced, especially for the cation. It should be noted, that the force field parameters were not explicitly optimized for this combination, but for [EMTr][OAc] and [EMIm][OBz].

2.3. Temperature Dependence

2.3.1. Radial Distribution Functions at Higher Temperatures

To verify the transferability of BILFF to [EMTr][OBz] force field MD simulations were performed and compared with reference AIMD simulations. Also, the application of the force field at elevated temperatures has been investigated and is discussed in this section. The radial distribution functions (RDFs) of the hydrogen bonds between the ring protons of the cation and the oxygen atoms of the anion are shown in Figure 14.

At 350 K, the first maximum of the RDF occurs at a particle distance of 205 pm and a $g(r) = 2.5$. As expected, with increasing temperature, the maximum decreases slightly to a $g(r) = 2.1$ at a temperature of 550 K. As the temperature rises, the mobility of the molecules increases, so that the hydrogen bonds become more short-lived, resulting in a smaller $g(r)$.

Similar to the AIMD simulation, the height of the maxima in the RDFs of the force field MD simulation does not change drastically with increasing temperature. Although the force field parameters were not directly optimized for aqueous [EMTr][OBz], the agreement with the reference AIMD simulation is very good, even at high temperatures, underlining the high accuracy of the optimized force field parameters of BILFF.
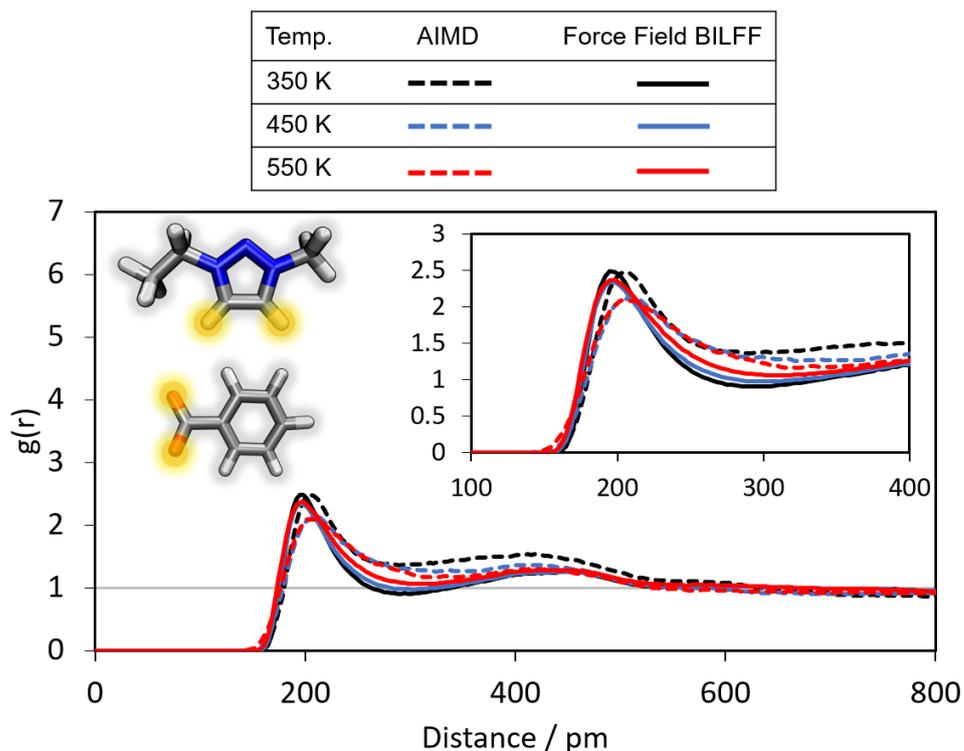
Figure 14. Radial distribution functions (RDFs) of the hydrogen bonds between the marked oxygen atoms of [OBz]$^-$ and the ring protons of [EMTr]$^+$ in aqueous [EMTr][OBz] at different temperatures, simulated by AIMD simulation and force field MD simulation using BILFF.

2.3.2. Hydrogen Bond Lifetime

Up to this point, a number of structural quantities have been computed, mainly expressed as distribution functions. Exemplary dynamical properties have also been calculated, in particular the hydrogen bond lifetime.

In Table 6, the hydrogen bond lifetime calculated from the AIMD and FFMD simulation using BILFF are compared. Furthermore, a comparative analysis of the lifetimes of the different hydrogen bonds in the investigated ILs [EMTr][OAc], [EMIm][OBz] and [EMTr][OBz] is shown. In addition, a comparison with the IL [EMIm][OAc] is made. The lifetime of the hydrogen bonds is studied in the presence and absence of water.

The calculation of the lifetime is differentiated in the intermittent and continuous calculation whereby, in the intermittent calculation, the breaking and reforming of the hydrogen bond is allowed. The continuous lifetime, on the other hand, represents the duration until the very first breakage of the hydrogen bond. To calculate the lifetime, geometric criteria were defined on the basis of the first maximum in the underlying distance–angle distribution functions. These criteria are listed in Table S5 in the Supplementary Materials. In Figure 10, the definition of the geometric criterion was exemplified as a black rectangle. The lifetime values presented in Table 6 are, respectively, averaged over all ring protons of the imidazolium (three ring protons) and 1,2,3-triazolium cation (two ring protons). The lifetime of the hydrogen bonds of the single protons H_{Me} and H_{Eth} can be found in Table S6 in the Supplementary Materials. (The lifetime of the hydrogen bonds [EMIm]$^+ \cdots$[OAc]$^-$ as well as [EMIm]$^+ \cdots$ water, with respect to all three individual ring protons can be found in the previous article on BILFF [27]).

Table 6. Overview of the lifetime τ of the hydrogen bonds in all four systems comparing the results of the reference AIMD simulation and the force field MD simulation using BILFF at the given temperatures (C^+ = Cation, A^- = Anion). (No AIMD simulations of anhydrous [EMTr][OBz] as well as [EMTr][OAc]/H_2O and [EMIm][OBz]/H_2O have been performed, so no data are available for these).

Temp.	Intermittent		Continuous	
	τ(AIMD)	τ(FFMD)	τ(AIMD)	τ(FFMD)
/ K	/ ps	/ ps	/ ps	/ ps
[EMTr][OAc]				
$(C^+)H\cdots O(A^-)$				
350	627.1	855.3	4.0	4.1
[EMTr][OAc]/H_2O				
$(C^+)H\cdots O(A^-)$				
350	–	117.6	–	1.9
$(C^+)H\cdots O(H_2O)$				
350	–	32.9	–	1.0
$(H_2O)H\cdots O(A^-)$				
350	–	58.3	–	0.2
[EMIm][OAc] [a]				
$(C^+)H\cdots O(A^-)$				
350	472.0	779.7	3.0	4.5
[EMIm][OAc]/H_2O [a]				
$(C^+)H\cdots O(A^-)$				
350	73.2	146.0	1.3	1.8
$(C^+)H\cdots O(H_2O)$				
350	31.2	40.4	0.6	0.8
$(H_2O)H\cdots O(A^-)$				
350	153.2	165.7	0.2	0.8
[EMIm][OBz]				
$(C^+)H\cdots O(A^-)$				
350	242.2	1841.6	1.0	2.0
[EMIm][OBz]/H_2O				
$(C^+)H\cdots O(A^-)$				
350	–	207.3	–	1.2
$(C^+)H\cdots O(H_2O)$				
350	–	236.0	–	1.7
$(H_2O)H\cdots O(A^-)$				
350	–	299.7	–	6.6
[EMTr][OBz]				
$(C^+)H\cdots O(A^-)$				
350	–	2821.1	–	6.3
[EMTr][OBz]/H_2O				
$(C^+)H\cdots O(A^-)$				
350	105.0	203.8	1.6	2.5
450	25.5	27.8	0.8	1.4
550	15.5	11.2	0.6	1.0
$(C^+)H\cdots O(H_2O)$				
350	38.1	48.3	0.7	1.0
450	5.7	6.4	0.4	0.6
550	–	2.3	0.3	0.5
$(H_2O)H\cdots O(A^-)$				
350	466.6	273.1	4.4	8.8
450	36.7	34.8	1.0	2.2
550	15.6	11.4	0.5	1.3

[a] Calculated from the MD simulations of our already published article [27].

The following conclusions can be drawn:

The longest hydrogen bond lifetime is observed in the anhydrous system, primarily due to strong Coulomb interactions, with $[EMTr]^+$ forming more durable hydrogen bonds with the anion than $[EMIm]^+$. However, this trend is reversed in the aqueous ILs.

The introduction of water results in the reduction of the hydrogen bond lifetime between the cations and anions in all ILs, due to a competition for oxygen atoms surrounding the cation, as well as a lower viscosity. This competition between water and the anion can also be observed in the radial and spatial distribution functions (see Sections 2.1.1 and 2.1.3) as well as in the Sankey diagrams (see Section 2.2). At the same time, the anion $[OBz]^-$ forms more long-lasting hydrogen bonds to water than $[OAc]^-$. This is also reflected in the frequency of the occurrence of the hydrogen bond of the ILs measured by the height of the first maximum of the RDF.

However, the lifetime of the hydrogen bonds can be affected not only by water, but also by an increased temperature. Due to the resulting higher mobility of the molecules, the lifetime of the hydrogen bonds decreases, as seen in aqueous [EMTr][OBz].

In all four ILs, the first breakup of the hydrogen bonds take place after only a few picoseconds (about 1–7 ps). Therefore, the two cation ring protons H_{Me} and H_{Eth} show similar values of lifetime (see Table S6 in the Supplementary Materials). In BILFF, they are regarded as equivalent.

It should be noted that dynamical properties, such as lifetimes, depend exponentially on the potential energy surface and are therefore subject to larger variations (with respect to the experiment, but also with respect to variations of the theoretical methods). From this perspective, the agreement between the force field MD simulation with BILFF and the AIMD simulation is still very good. It is also noteworthy that some of the lifetimes are in excess of the total simulation time of the AIMD simulation.

2.3.3. Validation of the Density and Diffusion

To test the reproducibility of experimental quantities, one approach is to compare the calculated system densities of the different ILs with corresponding experimental data (see Table 7). As expected, a decrease in density is observed in the presence of water. Pure [EMTr][OBz] has the highest density and (aqueous) [EMIm][OAc] the lowest. This difference can be attributed to the larger molecular volume occupied by $[EMTr]^+$ and $[OBz]^-$. The agreement of the density of pure [EMTr][OAc] as well as of the combined IL [EMTr][OBz] for which the force field parameters were not specifically optimized, is outstanding. The deviation is only 0.9% from the experimental measurement. The experimental density of [EMIm][OBz] is reproduced without any deviation at all.

Table 7. Comparison of the system densities from force field MD simulations using BILFF at 350 K. The molar fraction of the ionic liquids in the aqueous systems is $\chi_{IL} = 0.25$.

System	Temp. / K	Box Size / pm	ρ(FFMD) / g cm^{-3}	ρ(Lit.) / g cm^{-3}
[EMTr][OAc]	350	3198	1.11	1.12 [a]
[EMTr][OAc]/H_2O	350	3027	1.09	–
[EMIm][OAc]	350	3225	1.08 [b]	1.07 [c]
[EMIm][OAc]/H_2O	350	3043	1.07 [b]	1.07 [c]
[EMIm][OBz]	350	3550	1.10	1.10 [d]
[EMIm][OBz]/H_2O	350	3286	1.09	–
[EMTr][OBz]	350	3529	1.13	1.14 [e]
	350	3271	1.10	–
[EMTr][OBz]/H_2O	450	3369	1.01	–
	550	3502	0.90	–

[a] Measurements at 323.15 K from Ref. [15]. [b] Calculated from the MD simulations of our already published article [27]. [c] Extrapolated values of temperature dependent measurements from Ref. [41] (x_{IL} = 0.252 in the aqueous system). [d/e] Measurements at 358.15 K from Ref. [16].

The mobility of the molecules also plays a crucial role in accurately describe the behavior of the ILs. Below, the self-diffusion coefficients in all investigated ionic liquids calculated from the force field MD simulations are compared with the reference AIMD simulation (see Table 8). The following observations can be made:

Table 8. Self-diffusion coefficients D from force field MD simulations using BILFF. For comparison, the diffusion coefficients of [EMIm][OAc] are also shown from our previously published article [27].

Molecule	Temp. / K	D(AIMD) / 10^{-11} m^2 s^{-1}	D(FFMD) / 10^{-11} m^2 s^{-1}
		[EMTr][OAc]	
[EMTr]$^+$	350	4.02	5.94
[OAc]$^-$		4.79	5.35
		[EMTr][OAc]/H$_2$O	
[EMTr]$^+$	350	−	24.42
[OAc]$^-$		−	22.15
H$_2$O		−	54.88
		[EMIm][OAc] [a]	
[EMIm]$^+$	350	7.09	9.26
[OAc]$^-$		6.61	6.72
		[EMIm][OAc]/H$_2$O [a]	
[EMIm]$^+$	350	24.95	21.77
[OAc]$^-$		28.10	22.42
H$_2$O		46.88	55.97
		[EMIm][OBz]	
[EMIm]$^+$	350	10.33	3.72
[OBz]$^-$		8.92	1.89
		[EMIm][OBz]/H$_2$O	
[EMIm]$^+$	350	−	16.01
[OBz]$^-$		−	11.42
H$_2$O		−	42.83
		[EMTr][OBz]	
[EMTr]$^+$	350	−	1.77
[OBz]$^-$		−	1.13
		[EMTr][OBz]/H$_2$O	
[EMTr]$^+$	350	9.41	15.30
	450	96.73	123.40
	550	244.91	370.79
[OBz]$^-$	350	11.75	10.54
	450	97.40	102.42
	550	186.37	310.49
H$_2$O	350	25.56	48.66
	450	252.32	358.21
	550	562.49	102.76

[a] Calculated from the MD simulations of our already published article [27].

- According to the FFMD results, the bulky cation generally diffuse faster than the anion [42]. An exception is [EMIm][OAc] in water, where the anion diffuses faster. This phenomenon is already known and is in agreement with experimental data for [EMIm][OAc] [43].
- The diffusion coefficient is higher at elevated temperatures.
- In the presence of water, the diffusion coefficient of all ionic liquids increases significantly.
- The diffusion coefficient is influenced by both the cations and anions in the IL. When comparing the diffusion of the cations in the different ILs containing either [EMTr]$^+$ or [EMIm]$^+$, it is observed that the triazolium cations exhibit a slower

diffusion compared to imidazolium. Furthermore, when comparing the diffusion rates of the anions, it is found that benzoate diffuses slower than acetate in both pure and aqueous systems. This is due to the different occupied volumes and thus the bulkiness of the molecules as well as their more pronounced hydrogen bond formation.

- The lowest diffusion rate, and, at the same time, the highest density, can be found in pure [EMTr][OBz].

The results of the force field MD simulations reflect the trends of the reference AIMD simulations for the imidazolium-containing ILs. In the case of the triazolium-containing ILs, the AIMD simulation shows a slightly faster diffusion of the anion compared to the cation at 350 K and 450 K, and thus an opposite trend to the FFMD. Despite occasional numerical fluctuations in the specific diffusion values, the overall agreement of this dynamic quantity with the first-principles reference simulation is very good. The temperature dependence is reproduced over more than one order of magnitude. It should be noted that [EMTr][OBz] has been synthesized recently, so no experimental data for the diffusion coefficient are available yet. For [EMIm][OAc] the diffusion coefficient agrees very well with experimental data ($D_{\text{[EMIm], exp.}} = 14.1 \cdot 10^{-11} \, \text{m}^2 \, \text{s}^{-1}$; $D_{\text{[OAc], exp.}} = 12.8 \cdot 10^{-11} \, \text{m}^2 \, \text{s}^{-1}$) [44].

3. Computational Details

In this section, the computational details of the underlying reference AIMD simulations and force field MD simulations are presented. As a starting configuration for the AIMD simulations, the last time step of a force field MD simulation using the bonded and Lennard-Jones parameters of OPLS–AA [34–36] and the CL&P force field [37–39] and RESP partial charges (see Ref. [15]) as well as TIP4P–EW [45] (with constrained bonds and angles using the SHAKE algorithm [46,47]) for water was used. The physical simulation time was 5 ns for [EMIm][OBz], 0.5 ns for [EMTr][OAc] and 20 ns for aqueous [EMTr][OBz]. The final AIMD simulations were performed in analogy to the previous articles on BILFF [27,28] with CP2k [48–50] using the quickstep method [51] and orbital transformation (OT) [52]. Under application of the density functional theory [53,54] using the BLYP functional [55,56], dispersion correction D3(BJ) of Grimme [57,58] and the revised damping parameters of Smith et al. [59], the electron density was calculated. The DZVP-MOLOPT-SR-GTH [60] and GTH pseudopotentials [61,62] with a plane-wave energy cutoff of 350 Ry were used as basis sets. To equilibrate the AIMD simulations, the first 30 ps for [EMTr][OAc] and the first 50 ps for the systems with [OBz]$^-$ were discarded, respectively. The physical simulation time, size of the simulation box, and particle numbers can be found in Table 9.

The force field MD simulations for optimizing the force field parameter were also performed analogously to the previous articles on BILFF [27,28]. After constructing the simulation box using Packmol [63], an equilibration was performed at 500 K in the *NVT* ensemble using a Berendsen thermostat [64] (coupling constant of 1.0 fs) in a 25 ps simulation followed by a 100 ps simulation at a temperature of 350 K in the *NpT* ensemble using a Nosé–Hoover thermostat [65–67] (coupling constant of 100 fs) and a Nosé–Hoover barostat (coupling constant of 2000 fs). Using the Langevin thermostat [68,69], the acoustic shock waves were damped after the equilibration runs and adjustment of the final box volume. The production run was finally computed in the *NVT* ensemble using the Nosé–Hoover thermostat [65–67] (coupling constant of 100 fs) and a time integrator of 0.5 fs. The PPPM long-range solver of LAMMPS [70] was applied with a Lennard-Jones cutoff radius of 800 pm to calculate the electrostatic interactions.

As a starting point for the optimization of the force field parameters BILFF [27,28] for [EMTr]$^+$ was used. The reference bond length and angle of atom type NR were determined using a geometry optimization. For [OBz]$^-$, the Lennard-Jones parameters and the force constants were adapted from OPLS–AA [34–36]. The partial charges were calculated from quantum chemical calculations via the restrained electrostatic potential (RESP) methodology. The equilibrium bond length was calculated using a geometry optimization. The starting force field parameters are listed in the Supplementary Materials.

Table 9. Simulation parameters and physical simulation time (sim. time) of the final equilibrated ab initio and force field MD simulations of [EMIm][OBz] and aqueous [EMTr][OBz] for the development and validation of the force field.

System	Composition	Sim. Time / ps	Box Size / pm	Density / g cm^{-3}
		AIMD		
[EMTr][OAc]	36 [EMTr][OAc]	30 + 223	2121	1.072
[EMIm][OBz]	36 [EMIm][OBz]	50 + 46	2319	1.114
[EMTr][OBz] /H$_2$O	27 [EMTr][OBz] 81 Water	50 + 103	2319	1.033
		FFMD		
[EMTr][OAc]	128 [EMTr][OAc]	10,000	3198	1.113
[EMTr][OAc] /H$_2$O	81 [EMTr][OBz] 243 Water	10,000	3027	1.092
[EMIm][OBz]	128 [EMIm][OBz]	10,000	3550	1.103
[EMIm][OBz] /H$_2$O	81 [EMTr][OBz] 243 Water	10,000	3286	1.086
[EMTr][OBz]	128 [EMTr][OBz]	10,000	3529	1.128
[EMTr][OBz] /H$_2$O	81 [EMTr][OBz] 243 Water	10,000	3271	1.105

The ionic charge was set to ±0.82, as in the force field development of [EMIm][OAc] [27]. Reducing the total ion charge is discussed as an alternative to the usage of computationally more expensive polarizable force fields in the literature and is widely applied [71–75].

To validate the force field, additional force field MD simulations of pure and aqueous [EMTr][OBz] at 350 K as well as force field and reference AIMD simulations of aqueous [EMTr][OBz] at elevated temperatures of 450 K and 550 K were calculated. It should be noted that, for this IL, the force field parameters have not been optimized and the transferability of the parameters for [EMTr]$^+$ in [EMTr][OAc] and [OBz]$^-$ in [EMIm][OBz] to this IL will be tested. The same simulation parameters as described above were used. The first 30 ps of the AIMD simulations were again discarded for equilibration. The resulting physical simulation time is 88 ps in each case. An analysis of the temperature dependence of the simulations can be found in Section 2.3.

To analyze the trajectories, the program package TRAVIS [76,77] was applied. Xmgrace [78], Wolfram Mathematica [79], and VMD [80] with the Tachyon renderer [81] were used to create the figures.

4. Conclusions

In this article, we present the extension of our all-atom force field BILFF (Bio-Polymers in Ionic Liquids Force Field) to the ionic liquids [EMTr][OAc], [EMIm][OBz] and [EMTr][OBz] in absence and presence of water. BILFF was previously tuned for an optimal balance of the competing interactions in the ternary mixture of [EMIm][OAc] and water in view of its application to the challenge of modeling the solvation structure of cellulose (as solute) by [EMIm][OAc]/water [27,28]. The focus of this work is the generalization of the [EMIm][OAc] force field to additional ions ([EMTr]$^+$, [OBz]$^-$) including all four combinations.

This article thus presents the first force field specifically optimized for [EMTr][OAc] and [EMIm][OBz]. The objective was to ensure an accurate reproduction of the quantum chemical microstructure as well as dynamic properties such as hydrogen bond lifetimes and diffusion in these ILs. To achieve this, an iterative adjustment of the force field parameters to the results of reference AIMD simulations and experimental system densities was performed. In addition, to verify the transferability of the optimized force field parameters for [EMIm][OBz] and [EMTr][OAc], the results of (aqueous) [EMTr][OBz] were evaluated.

By comparing the simulation results of the four ILs [EMTr][OAc], [EMIm][OAc], [EMTr][OAc] and [EMTr][OBz], the hydrogen bonding strength of benzoate was found to

be slightly stronger than that of acetate, in spite of the smaller partial charges on the oxygen atoms ($-0.55\,e$ vs. $-0.66\,e$). This trend is inverted by water. This observation is surprising at first glance, and indeed we cannot provide a "single" answer to this phenomenon. One contributing factor to this enhancement of hydrogen bonding strength is the reduced rotational/translational freedom due to the T-formation of the anions, which leads to a more stable spatial arrangement and a longer lifetime of the hydrogen bonds. Regarding the cations, the hydrogen bond strength is stronger for triazolium compared to imidazolium. With the CH/N substitution as the only molecular difference, the enhancement is most likely due to the higher partial charges of the CH ring protons ($+0.19\,e$ vs. $+0.15\,e$). This different ionic hydrogen bonding strengths play an ambiguous role, in view of the solubility of cellulose in the different ILs; while a good hydrogen bonding ability (mainly of the anions) is essential to break the cellulose–cellulose bonds, this effect is counterbalanced by an equally increased hydrogen bond strength between the ion pairs. These two effects are in competition with each other, and again, no "single argument" is able to resolve this issue unambiguously.

Experimentally, the solubility of cellulose increases in the following order: [EMIm][OBz] < [EMTr][OBz] < [EMTr][OAc] < [EMIm][OAc] [15,16]. The results of the simulations show that the best solvent [EMIm][OAc] has the weakest cation–anion interaction in its pure form, whereas the IL with the strongest ion pair interactions, [EMTr][OBz], is the second poorest cellulose solvent. However, the even lower solubility in [EMIm][OBz] cannot be explained by this scheme, as its cation–anion interaction is somewhat weaker than that of [EMTr][OBz].

Clearly, further factors must be considered to establish a complete picture of cellulose solubility in such complex solvents. One important aspect is the mere spatial size of the anion, as it strongly influences its diffusivity and thus its ability to intercalate into the cellulose crystal/fibrils and initiate the disruption of the internal cellulose hydrogen bonding network. Another point to consider is the influence of the van der Waals effects, which also alter the balance of the competing interactions between the anion/cation/water/cellulose molecules.

Abbreviations

The following abbreviations are used in this article:

AIMD   Ab initio Molecular Dynamics
BILFF   Bio-polymers in Ionic Liquids Force Field
CDF   Combined Distribution Function
FFMD   Force Field Molecular Dynamics
IL   Ionic Liquid
MD   Molecular Dynamics
RDF   Radial Distribution Function
SDF   Spatial Distribution Function

References

1. Plechkova, N.V.; Seddon, K.R. Applications of ionic liquids in the chemical industry. *Chem. Soc. Rev.* **2008**, *37*, 123–150. [CrossRef] [PubMed]
2. Hallett, J.P.; Welton, T. Room-temperature ionic liquids: Solvents for synthesis and catalysis. 2. *Chem. Rev.* **2011**, *111*, 3508–3576. [CrossRef]
3. Zhang, S.; Sun, N.; He, X.; Lu, X.; Zhang, X. Physical properties of ionic liquids: Database and evaluation. *J. Phys. Chem. Ref. Data* **2006**, *35*, 1475–1517. [CrossRef]
4. MacFarlane, D.R.; Tachikawa, N.; Forsyth, M.; Pringle, J.M.; Howlett, P.C.; Elliott, G.D.; Davis, J.H.; Watanabe, M.; Simon, P.; Angell, C.A. Energy applications of ionic liquids. *Energy Environ. Sci.* **2014**, *7*, 232–250. [CrossRef]
5. Zhao, D.; Wu, M.; Kou, Y.; Min, E. Ionic liquids: Applications in catalysis. *Catal. Today* **2002**, *74*, 157–189. [CrossRef]
6. MacFarlane, D.R.; Forsyth, M.; Howlett, P.C.; Pringle, J.M.; Sun, J.; Annat, G.; Neil, W.; Izgorodina, E.I. Ionic liquids in electrochemical devices and processes: Managing interfacial electrochemistry. *Acc. Chem. Res.* **2007**, *40*, 1165–1173. [CrossRef]
7. Obadia, M.M.; Drockenmuller, E. Poly(1,2,3-triazolium)s: A new class of functional polymer electrolytes. *Chem. Commun.* **2016**, *52*, 2433–2450. [CrossRef]
8. Guo, N.; Li, M.; Sun, X.; Wang, F.; Yang, R. Enzymatic hydrolysis lignin derived hierarchical porous carbon for supercapacitors in ionic liquids with high power and energy densities. *Green Chem.* **2017**, *19*, 2595–2602. [CrossRef]
9. Elia, G.A.; Ulissi, U.; Mueller, F.; Reiter, J.; Tsiouvaras, N.; Sun, Y.K.; Scrosati, B.; Passerini, S.; Hassoun, J. A long-life lithium ion battery with enhanced electrode/electrolyte interface by using an ionic liquid solution. *Chemistry* **2016**, *22*, 6808–6814. [CrossRef]
10. Fletcher, J.T.; Sobczyk, J.M.; Gwazdacz, S.C.; Blanck, A.J. Antimicrobial 1,3,4-trisubstituted-1,2,3-triazolium salts. *Bioorg. Med. Chem.* **2018**, *28*, 3320–3323. [CrossRef]
11. Tan, W.; Li, Q.; Dong, F.; Zhang, J.; Luan, F.; Wei, L.; Chen, Y.; Guo, Z. Novel cationic chitosan derivative bearing 1,2,3-triazolium and pyridinium: Synthesis, characterization, and antifungal property. *Carbohydr. Polym.* **2018**, *182*, 180–187. [CrossRef]
12. Swatloski, R.P.; Spear, S.K.; Holbrey, J.D.; Rogers, R.D. Dissolution of cellulose correction of cellose with ionic liquids. *J. Am. Chem. Soc.* **2002**, *124*, 4974–4975. [CrossRef]
13. Brandt, A.; Ray, M.J.; To, T.Q.; Leak, D.J.; Murphy, R.J.; Welton, T. Ionic liquid pretreatment of lignocellulosic biomass with ionic liquid–water mixtures. *Green Chem.* **2011**, *13*, 2489. [CrossRef]
14. Gupta, K.M.; Jiang, J. Cellulose dissolution and regeneration in ionic liquids: A computational perspective. *Chem. Eng. Sci.* **2015**, *121*, 180–189. [CrossRef]
15. Brehm, M.; Pulst, M.; Kressler, J.; Sebastiani, D. Triazolium-based ionic liquids: A novel class of cellulose solvents. *J. Phys. Chem. B* **2019**, *123*, 3994–4003. [CrossRef]
16. Brehm, M.; Radicke, J.; Pulst, M.; Shaabani, F.; Sebastiani, D.; Kressler, J. Dissolving cellulose in 1,2,3-triazolium- and imidazolium-based ionic liquids with aromatic anions. *Molecules* **2020**, *25*, 3539. [CrossRef]
17. Radicke, J.; Roos, E.; Sebastiani, D.; Brehm, M.; Kressler, J. Lactate–based ionic liquids as chiral solvents for cellulose. *J. Polym. Sci.* **2023**, *61*, 372–384. [CrossRef]
18. Zhang, L.; Ruan, D.; Gao, S. Dissolution and regeneration of cellulose in NaOH/thiourea aqueous solution. *J. Polym. Sci. B Polym. Phys.* **2002**, *40*, 1521–1529. [CrossRef]
19. Heinze, T.; Schwikal, K.; Barthel, S. Ionic liquids as reaction medium in cellulose functionalization. *Macromol. Biosci.* **2005**, *5*, 520–525. [CrossRef] [PubMed]
20. Sun, N.; Rahman, M.; Qin, Y.; Maxim, M.L.; Rodríguez, H.; Rogers, R.D. Complete dissolution and partial delignification of wood in the ionic liquid 1-ethyl-3-methylimidazolium acetate. *Green Chem.* **2009**, *11*, 646. [CrossRef]
21. Fort, D.A.; Remsing, R.C.; Swatloski, R.P.; Moyna, P.; Moyna, G.; Rogers, R.D. Can ionic liquids dissolve wood? Processing and analysis of lignocellulosic materials with 1-n-butyl-3-methylimidazolium chloride. *Green Chem.* **2007**, *9*, 63–69. [CrossRef]
22. Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal structure and hydrogen-bonding system in cellulose Ibeta from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082. [CrossRef]
23. Olsson, C.; Westm, G. Direct dissolution of cellulose: Background, means and applications. *Cellul. Fundam. Asp.* **2013**. [CrossRef]
24. Ren, F.; Wang, J.; Yu, J.; Zhong, C.; Xie, F.; Wang, S. Dissolution of cellulose in ionic liquid-DMSO mixtures: Roles of DMSO/IL ratio and the cation alkyl chain length. *ACS Omega* **2021**, *6*, 27225–27232. [CrossRef] [PubMed]

25. Li, W.; Sun, N.; Stoner, B.; Jiang, X.; Lu, X.; Rogers, R.D. Rapid dissolution of lignocellulosic biomass in ionic liquids using temperatures above the glass transition of lignin. *Green Chem.* **2011**, *13*, 2038. [CrossRef]

26. Liebert, T.F.; Heinze, T.J.; Edgar, K.J. *Cellulose Solvents: For Analysis, Shaping and Chemical Modification*; American Chemical Society: Washington, DC, USA, 2010; Volume 1033. [CrossRef]

27. Roos, E.; Brehm, M. A force field for bio-polymers in ionic liquids (BILFF)—Part 1: EMImOAc/water mixtures. *Phys. Chem. Chem. Phys.* **2021**, *23*, 1242–1253. [CrossRef]

28. Roos, E.; Sebastiani, D.; Brehm, M. A force field for bio-polymers in ionic liquids (BILFF)—Part 2: Cellulose in EMImOAc/water mixtures. *Phys. Chem. Chem. Phys.* **2023**, *25*, 8755–8766. [CrossRef]

29. Olsson, C.; Idström, A.; Nordstierna, L.; Westman, G. Influence of water on swelling and dissolution of cellulose in 1-ethyl-3-methylimidazolium acetate. *Carbohydr. Polym.* **2014**, *99*, 438–446. [CrossRef]

30. Le, K.A.; Sescousse, R.; Budtova, T. Influence of water on cellulose-EMIMAc solution properties: A viscometric study. *Cellulose* **2012**, *19*, 45–54. [CrossRef]

31. Froschauer, C.; Hummel, M.; Iakovlev, M.; Roselli, A.; Schottenberger, H.; Sixta, H. Separation of hemicellulose and cellulose from wood pulp by means of ionic liquid/cosolvent systems. *Biomacromolecules* **2013**, *14*, 1741–1750. [CrossRef]

32. Bowron, D.T.; D'Agostino, C.; Gladden, L.F.; Hardacre, C.; Holbrey, J.D.; Lagunas, M.C.; McGregor, J.; Mantle, M.D.; Mullan, C.L.; Youngs, T.G.A. Structure and dynamics of 1-ethyl-3-methylimidazolium acetate via molecular dynamics and neutron diffraction. *J. Phys. Chem. B* **2010**, *114*, 7760–7768. [CrossRef] [PubMed]

33. Yaghini, N.; Pitawala, J.; Matic, A.; Martinelli, A. Effect of water on the local structure and phase behavior of imidazolium-based protic ionic liquids. *J. Phys. Chem. B* **2015**, *119*, 1611–1622. [CrossRef] [PubMed]

34. Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236. [CrossRef]

35. Ponder, J.W.; Case, D.A. Force fields for protein simulations. *Adv. Protein Chem.* **2003**, *66*, 27–85. [CrossRef] [PubMed]

36. Sambasivarao, S.V.; Acevedo, O. Development of OPLS-AA force field parameters for 68 unique ionic liquids. *J. Chem. Theory Comput.* **2009**, *5*, 1038–1050. [CrossRef] [PubMed]

37. Canongia Lopes, J.N.; Deschamps, J.; Pádua, A.A.H. Modeling ionic liquids using a systematic all-atom force field. *J. Phys. Chem. B* **2004**, *108*, 2038–2047. [CrossRef]

38. Canongia Lopes, J.N.; Pádua, A.A.H. Molecular force field for ionic liquids III: Imidazolium, pyridinium, and phosphonium cations; chloride, bromide, and dicyanamide anions. *J. Phys. Chem. B* **2006**, *110*, 19586–19592. [CrossRef]

39. Canongia Lopes, J.N.; Pádua, A.A.H. A generic and systematic force field for ionic liquids modeling. *Theor. Chem. Acc.* **2012**, *131*, 3330. [CrossRef]

40. Pike, S.J.; Hutchinson, J.J.; Hunter, C.A. H-Bond Acceptor Parameters for Anions. *J. Am. Chem. Soc.* **2017**, *139*, 6700–6706. [CrossRef]

41. Quijada-Maldonado, E.; van der Boogaart, S.; Lijbers, J.H.; Meindersma, G.W.; de Haan, A.B. Experimental densities, dynamic viscosities and surface tensions of the ionic liquids series 1-ethyl-3-methylimidazolium acetate and dicyanamide and their binary and ternary mixtures with water and ethanol at T = (298.15 to 343.15K). *J. Chem. Thermodyn.* **2012**, *51*, 51–58. [CrossRef]

42. Urahata, S.M.; Ribeiro, M.C.C. Single particle dynamics in ionic liquids of 1-alkyl-3-methylimidazolium cations. *Chem. Phys.* **2005**, *122*, 024511. [CrossRef]

43. Hall, C.A.; Le, K.A.; Rudaz, C.; Radhi, A.; Lovell, C.S.; Damion, R.A.; Budtova, T.; Ries, M.E. Macroscopic and microscopic study of 1-ethyl-3-methyl-imidazolium acetate-water mixtures. *J. Phys. Chem. B* **2012**, *116*, 12810–12818. [CrossRef] [PubMed]

44. Green, S.M.; Ries, M.E.; Moffat, J.; Budtova, T. NMR and rheological study of anion size influence on the properties of two imidazolium-based ionic liquids. *Sci. Rep.* **2017**, *7*, 8968. [CrossRef]

45. Horn, H.W.; Swope, W.C.; Pitera, J.W.; Madura, J.D.; Dick, T.J.; Hura, G.L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678. [CrossRef]

46. Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [CrossRef]

47. Andersen, H.C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34. [CrossRef]

48. The CP2K Developers Group. 2023. Available online: http://www.cp2k.org/ (accessed on 18 June 2023).

49. Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic simulations of condensed matter systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15–25. [CrossRef]

50. Kühne, T.D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V.V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R.Z.; Schütt, O.; Schiffmann, F.; et al. CP2K: An electronic structure and molecular dynamics software package—Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys* **2020**, *152*, 194103. [CrossRef] [PubMed]

51. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103–128. [CrossRef]

52. VandeVondele, J.; Hutter, J. An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.* **2003**, *118*, 4365–4369. [CrossRef]

53. Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev. B* **1964**, *136*, 864. [CrossRef]

54. Kohn, W.; Sham, L. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, 1133. [CrossRef]

55. Becke, A. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100. [CrossRef] [PubMed]

56. Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789. [CrossRef]

57. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys* **2010**, *132*, 154104. [CrossRef]

58. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [CrossRef] [PubMed]

59. Smith, D.G.A.; Burns, L.A.; Patkowski, K.; Sherrill, C.D. Revised damping parameters for the D3 dispersion correction to density functional theory. *J. Phys. Chem. Lett.* **2016**, *7*, 2197–2203. [CrossRef]

60. VandeVondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, 114105. [CrossRef]

61. Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710. [CrossRef]

62. Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641–3662. [CrossRef]

63. Martínez, L.; Andrade, R.; Birgin, E.G.; Martínez, J.M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164. [CrossRef] [PubMed]

64. Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys* **1984**, *81*, 3684–3690. [CrossRef]

65. Nose, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268. [CrossRef]

66. Nose, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511–519. [CrossRef]

67. Martyna, G.; Klein, M.; Tuckerman, M. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1192**, *97*, 2535–2643. [CrossRef]

68. Schneider, T.; Stoll, E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B* **1978**, *17*, 1302–1322. [CrossRef]

69. Dünweg, B.; Paul, W. Brownian dynamics simulations without Gaussian random numbers. *Int. J. Mod. Phys. C* **1991**, *2*, 817–827. [CrossRef]

70. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.* **1995**, *117*, 1–19. [CrossRef]

71. Bhargava, B.L.; Balasubramanian, S. Refined potential model for atomistic simulations of ionic liquid [bmim][PF6]. *J. Chem. Phys.* **2007**, *127*, 114510. [CrossRef]

72. Liu, H.; Maginn, E. A molecular dynamics investigation of the structural and dynamic properties of the ionic liquid 1-n-butyl-3-methylimidazolium bis(trifluoromethanesulfonyl)imide. *J. Chem. Phys.* **2011**, *135*, 124507. [CrossRef]

73. Doherty, B.; Zhong, X.; Gathiaka, S.; Li, B.; Acevedo, O. Revisiting OPLS force field parameters for ionic liquid simulations. *J. Chem. Theory Comput.* **2017**, *13*, 6131–6145. [CrossRef] [PubMed]

74. Doherty, B.; Acevedo, O. OPLS force field for choline chloride-based deep eutectic solvents. *J. Phys. Chem. B* **2018**, *122*, 9982–9993. [CrossRef] [PubMed]

75. Wang, Y.L. Competitive microstructures versus cooperative dynamics of hydrogen bonding and Pi-type stacking interactions in imidazolium bis(oxalato)borate ionic liquids. *J. Phys. Chem. B* **2018**, *122*, 6570–6585. [CrossRef] [PubMed]

76. Brehm, M.; Thomas, M.; Gehrke, S.; Kirchner, B. TRAVIS – A free analyzer for trajectories from molecular simulation. *J. Chem. Phys.* **2020**, *152*, 164105. [CrossRef] [PubMed]

77. Brehm, M.; Kirchner, B. TRAVIS—A free analyzer and visualizer for Monte Carlo and molecular dynamics trajectories. *J. Chem. Inf. Model.* **2011**, *51*, 2007–2023. [CrossRef]

78. Grace, 1996–2023 Grace Development Team. Available online: http://plasma-gate.weizmann.ac.il/Grace (accessed on 18 June 2023).

79. Wolfram Research, Inc. *Mathematica*, Version 8.0; Wolfram Research, Inc.: Champaign, IL, USA, 2010.

80. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]

81. Stone, J. An Efficient Library for Parallel Ray Tracing and Animation. Ph.D. Thesis, Computer Science Department, University of Missouri-Rolla, Rolla, MO, USA, 1998.

3.6 Ongoing Unpublished Work: Applications of BILFF

To provide a brief insight into the wide range of questions that can be investigated using force field molecular dynamics simulations with BILFF, this section is devoted to two previously unpublished application examples.

In Section 3.6.1, force field MD simulations with BILFF are used to address the controversial question of whether and how the cation of the IL influences the dissolution process of cellulose, as well as the impact of varying water amounts (0 to 20 wt.% water) on the cellulose/IL/water system. The two ILs, 1-ethyl-3-methyl-imidazolium acetate ([EMIm][OAc]) and 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]), both recognized as effective cellulose solvents, were investigated, and the intricate hydrogen bond network in the quaternary system was examined. Section 3.6.2 extends the above investigations to the influence of the anions on the cellulose/IL/(water) hydrogen bond network, comparing the effects of an aliphatic anion, acetate, with an aromatic anion, benzoate ([OBz]$^-$). For this purpose, the results of force field MD simulations using BILFF of the ILs [EMIm][OAc], [EMTr][OAc], [EMIm][OBz] and [EMTr][OBz] with cellulose in the presence and absence of water are considered.

Since triazolium-based ILs were discovered as cellulose solvents only five years ago,[35] detailed results on the dissolution mechanism in these ILs are not yet available. Both studies are therefore intended to provide a starting point for further investigations. Since the length of this thesis is limited, only a brief overview of the results shall be given here, with special emphasis on hydrogen bonding networks.

3.6.1 Is the hydrogen bond network of cellulose/IL/H$_2$O mixtures affected by the choice of the cation?

Although the two cations 1-ethyl-3-methylimidazolium and 1-ethyl-3-methyl-1,2,3-triazolium differ only in one single atom, experimental data show a different solubility behavior of cellulose in the two ILs 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) and 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]):[35] The only difference in the structure of these two cations is the replacement of one C-H unit in the ring of [EMIm]$^+$ by a nitrogen atom in [EMTr]$^+$. As a result, the latter molecule has the advantage of avoiding side reactions such as the formation of NHCs.[179] The seemingly minor structural change reduces the cellulose solubility in [EMTr][OAc]. The causes have not yet been conclusively clarified.

Additionally, the influence of water on the hydrogen bond network was studied. Since many imidazolium and triazolium based ionic liquids are hygroscopic,[180, 181] the influence of water on the interactions between the ILs and cellulose should not be underestimated, as water can cause the precipitation of cellulose.[93] To better understand these complex mechanisms, force field MD simulations were performed with BILFF, varying the water content in the two ILs from 0 wt.% up to 20 wt.%, where cellulose is completely precipitated.[172] These simulations were performed as part of the bachelor thesis of Luisa Voigt under my supervision. Luisa Voigt performed the MD simulations with water contents of 5, 10 and 20 wt.% water and analyzed the simulations. I performed and analyzed the simulations of the anhydrous systems. Daniel Sebastiani and I designed the research and supervised the project. The interpretation of the results was done by both Luisa Voigt and me.

Computational Details

Each simulated system consists of a cellulose strand of 5 β-D-glucose monomers, 130 IL pairs and different amounts of water (see Tab. 3.1). The simulation cell was generated using Packmol.[182] In all simulations, the force field BILFF was applied. For water the force field TIP4p-EW[183] (with constrained bonds and angles using the RATTLE algorithm[184, 185]) was used. All force field MD simulations followed the procedure outlined in Section 3.3 and were performed at 350 K.
The molecular structures presented were created using VMD[123] and Tachyon.[124] All analyses were performed using the TRAVIS software package.[36, 168]

TABLE 3.1: Simulation parameters and physical simulation time (sim. time) of the equilibrated force field MD simulations of [EMIm][OAc] and [EMTr][OAc] at different water concentrations. The mass fraction of water refers to the solvent mixture.

System	Composition	Sim. time /ns	Box size /pm	Density /g·cm^{-3}
	[EMTr][OAc]			
0 wt.% H$_2$O	1 cellulose pentamer 192 [EMTr][OAc]	50	3682	1.12
5 wt.% H$_2$O	1 cellulose pentamer 130 [EMTr][OAc] 70 H$_2$O	50	3393	1.12
10 wt.% H$_2$O	1 cellulose pentamer 130 [EMTr][OAc] 150 H$_2$O	50	3368	1.12
20 wt.% H$_2$O	1 cellulose pentamer 130 [EMTr][OAc] 320 H$_2$O	50	3512	1.11
	[EMIm][OAc]			
0 wt.% H$_2$O	1 cellulose pentamer 192 [EMIm][OAc]	50	3715	1.09
5 wt.% H$_2$O	1 cellulose pentamer 130 [EMIm][OAc] 70 H$_2$O	50	3328	1.09
10 wt.% H$_2$O	1 cellulose pentamer 130 [EMIm][OAc] 150 H$_2$O	50	3393	1.09
20 wt.% H$_2$O	1 cellulose pentamer 130 [EMIm][OAc] 320 H$_2$O	50	3532	1.08

Microstructure of the System

As seen in the previous sections, the interaction between the cation and anion is of great importance for the dissolution of cellulose in ILs. Therefore, the strength of the cation–anion hydrogen bonds in the ILs [EMIm][OAc] and [EMTr][OAc] was investigated at different water concentrations. Figure 3.2 shows the corresponding RDFs. The results of the two marked ring protons of the respective cations have been averaged, since they are equivalent to each other. When comparing the simulations of both ILs at the same water content, it is noticeable that for [EMTr]$^+$ the first maximum of $g(r)$ is observed with higher intensity at smaller particle distances about 195 pm. The ring protons of the cation [EMTr]$^+$ thus form stronger hydrogen bonds to [OAc]$^-$ than [EMIm]$^+$. One of the underlying reasons is the 27 % higher partial charge of the ring protons of the triazolium cation due to the additional nitrogen atom in the ring compared to imidazolium. The lower number of ring protons in [EMTr]$^+$ also affects the competition for hydrogen bond donors and increases the intensity of each interaction. These results are in agreement with the findings in Section 3.5 of this thesis. (Furthermore, in that section, it is shown that even when averaging the RDF over all three protons of [EMIm]$^+$, [EMTr]$^+$ exhibits stronger hydrogen bonding with its anion (cf. Fig. 7 and Fig. 9 in Section 3.5).) In both ILs, an increased water content leads to a weakening of the cation–anion hydrogen bonds due to the formation of a hydration shell. The shielding effect of water is evident, for example, in the RDFs of the cation–water interaction as well as the strong anion–water hydrogen bonds (cf. Fig. S2 and S5 in Section A.4 as well as Fig. 12 in Section 3.5 of this thesis).

Since [EMTr]$^+$ forms stronger cation–anion hydrogen bonds than [EMIm]$^+$, it is expected that this effect counteracts cellulose dissolution by reducing the availability of anions to interact with cellulose. To confirm this hypothesis, RDFs of the anion–cellulose hydrogen bonding were analyzed to investigate in which IL the hydrogen bonding is more pronounced (cf. Fig. 3.3). A distinction was made between hydroxyl groups directly bonded to the carbon ring (called HOR) and those attached to the methylene groups (called HO6). To generate an averaged RDF for the two atom types, the corresponding RDFs of the individual protons were averaged and weighted based on their molecular positions in cellulose as described in Section 3.3 of this thesis. The RDFs show that the cellulose-acetate interaction tends to be stronger in [EMIm][OAc] (dotted lines) than in [EMTr][OAc] (straight lines). This is consistent with the finding that in [EMIm][OAc] the anions are freer due to a weaker cation–anion interaction and therefore interact stronger with cellulose. This agrees with the experimental data that [EMIm][OAc] is a slightly better solvent for cellulose than [EMTr][OAc][35].

To investigate if the same trends also occur with increasing water content, the RDFs of the cellulose–anion hydrogen bonds in the aqueous ILs are examined. By comparing the RDFs of the two hydroxyl groups HO6 and HOR (cf. Fig. 3.3) with acetate at the same water concentration in the systems, it can be seen that the HOR–[OAc]$^-$ interactions seem to be more affected by water than the HO6–[OAc]$^-$ hydrogen bond, even at low water concentrations. A possible explanation for this could be the formation of an energetically favored complex of water with the two neighboring HOR hydroxyl groups, resulting in a high shielding effect against acetate. At the same time, HO6 can rotate freely and is thus more accessible to the larger acetate molecules without encountering steric hindrance. To investigate these non-trivial effects, more MD simulations and further analysis are required. For the necessary force field MD simulations of these quaternary systems with long simulation times, the force-field

parameters of BILFF can be used to obtain accurate results.[b]
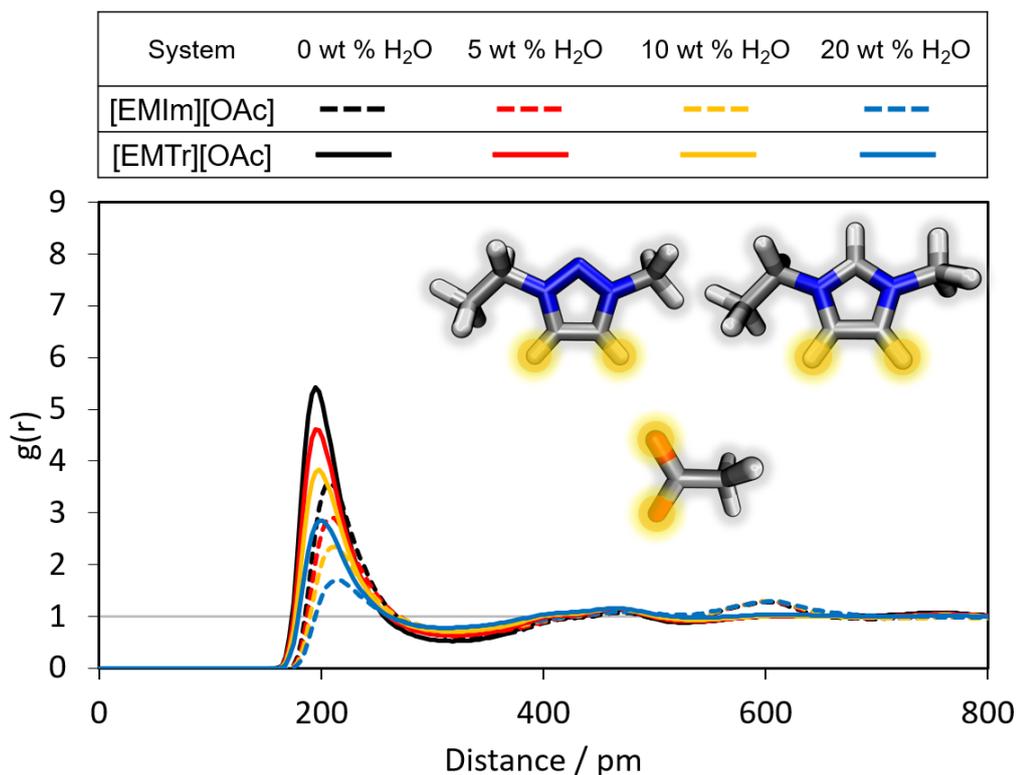
FIGURE 3.2: Radial distribution functions of the hydrogen bond between the marked ring protons of the cations 1-ethyl-3-methylimidazolium ([EMIm]$^+$) and 1-ethyl-3-methyl-1,2,3-triazolium ([EMTr]$^+$) and their respective anion acetate at different water concentrations.

Conclusion

This project demonstrates the applicability of BILFF for force field MD simulations to study the influence of [EMIm]$^+$ and [EMTr]$^+$ cations on the cellulose/IL/water system as a function of the water concentration (0–20 wt.% water, 350 K). Both structural and dynamic properties were investigated.

[EMTr]$^+$ forms stronger hydrogen bonds to [OAc]$^-$ than [EMIm]$^+$, both in the presence and absence of water. The anion therefore interacts less strongly with the cellulose strand in [EMTr]$^+$ compared to [EMIm]$^+$, which is consistent with the reduced solubility of cellulose in [EMTr][OAc].[35]

Additionally, the position of the hydroxyl groups in the cellulose strand plays a crucial role in the hydrogen bond formation. The protons of the freely rotating hydroxyl group on the methylene group (HO6) can form stronger hydrogen bonds to the anion and are less affected by the presence of water than the hydroxyl protons directly on the monomer (HOR). The exact reasons for these observations are not fully understood and require further investigation. One potential explanation involves the formation of an energetically favored water complex with the two neighboring HOR hydroxyl groups, leading to a significant shielding effect against acetate. Simultaneously, the free rotation of HO6 renders it more accessible to the larger acetate

---

[b]Additional investigations of the microstructure and dynamics of the systems, including a comparative analysis of the diffusion with existing literature data and the calculation of the lifetime of the hydrogen bonds, are presented in the bachelor thesis of Luisa Voigt.
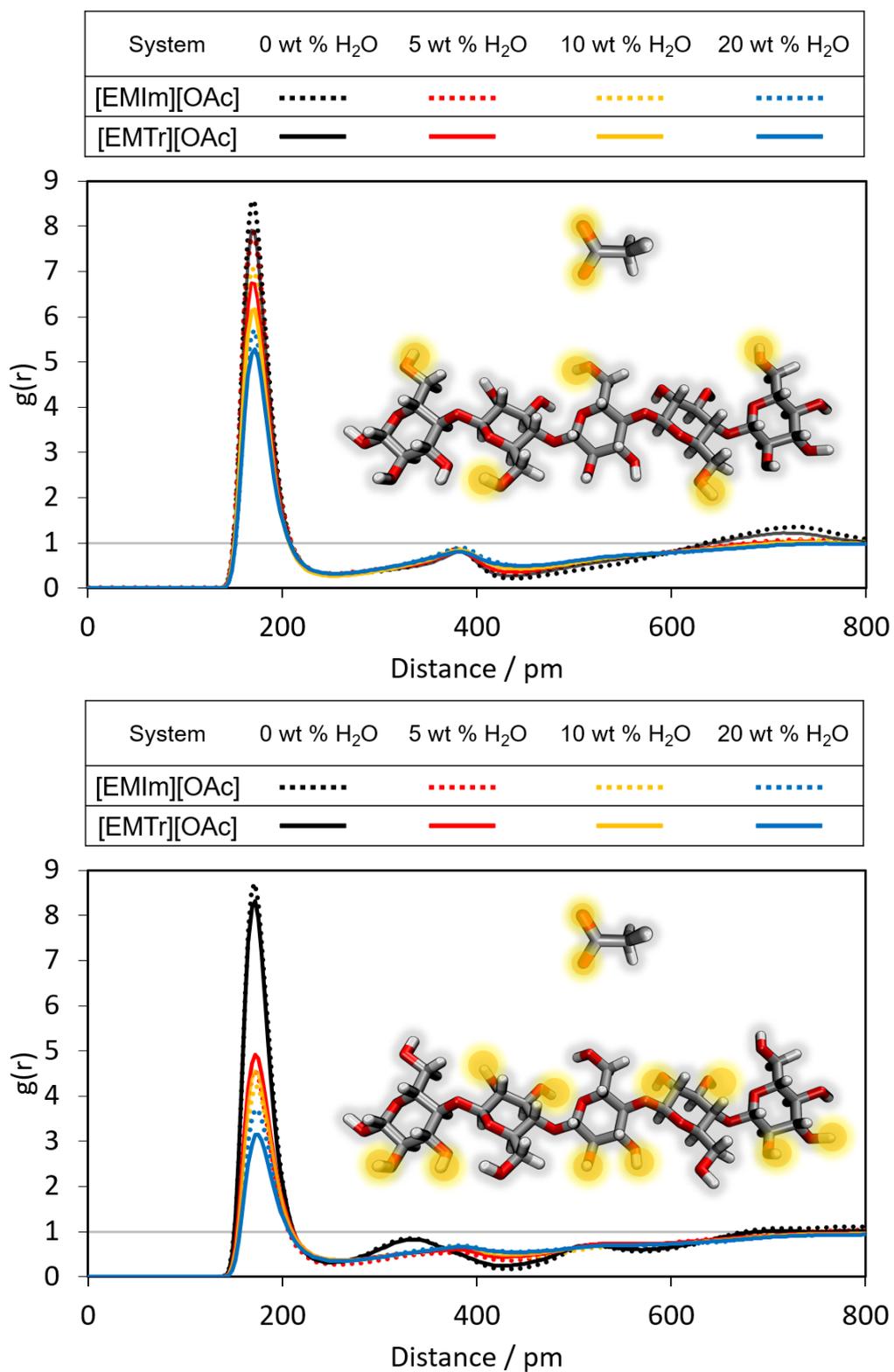
FIGURE 3.3: Radial distribution functions of the hydrogen bond between the marked protons (HO6 (**top**) and HOR (**bottom**)) of cellulose and acetate in [EMIm][OAc] and [EMTr][OAc] at different water concentrations.

molecules without encountering steric hindrance.

In summary, both cations affect the hydrogen bond network between cellulose and the anion of IL differently, but the effects are not very pronounced. This result is consistent with the discussion in the literature.[31, 32, 38, 109, 110] Thus, the choice between [EMIm]$^+$ and [EMTr]$^+$ as cations is not crucial for the solubility of cellulose. Triazolium can therefore be considered as an alternative cation to imidazolium to avoid side reactions such as the formation of NHCs.

3.6.2   How does the anion affect the hydrogen bond network?

The cation appears to exert only a minor influence on the solubility of cellulose, as demonstrated in Section 3.6.1, whereas the selection of the anion appears to be more significant: the aliphatic anion acetate exhibits a greater cellulose solubility than the aromatic anion benzoate. 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]) has one of the highest cellulose solubilities of 36 g per mol IL, while 1-ethyl-3-methyl-imidazolium benzoate ([EMIm][OBz]) has a lower cellulose solubility of 18.6 g per mol IL. The same tendency is observed for the cation [EMTr]$^+$: 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]) has a cellulose solubility of 34 g per mol IL and is therefore a better solvent than 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]) with a cellulose solubility of 21.7 g per mol IL.[35, 36] The reason for this drastic difference in cellulose solubility is still unclear.

The investigations in Section 3.5 show that the cation–anion hydrogen bonding of benzoate is slightly stronger than that of acetate, despite the smaller partial charges on the oxygen atoms of benzoate ($-0.55\,e$ vs. $-0.66\,e$). Notably, this trend is reversed when water is added to the system due to even stronger anion–water hydrogen bonds. This section extends the investigation to the anion–cellulose interaction and examines which of the two anions, [OBz]$^-$ or [OAc]$^-$, forms stronger hydrogen bonds with cellulose, thus providing initial insights into the difference in solubility. Since water is crucial for the processing and subsequent precipitation of cellulose, the simulations are conducted for both cases: when cellulose is dissolved in the pure IL and when it is precipitated again at a molar fraction of the IL of $\chi_{IL} = 0.25$. Regarding the cation, both [EMTr]$^+$ and [EMIm]$^+$ were selected to broaden the material basis and derive general results. The applicability of BILFF is thereby illustrated by a brief consideration of the microstructure of the system.

Computational Details

Force field MD simulations were performed with a cellulose strand of 5 β-D-glucose monomers in the anhydrous and aqueous ILs [EMTr][OAc], [EMIm][OAc], [EMIm][OBz] and [EMTr][OBz] (cf. Tab. 3.2). The simulation cell was generated using the program Packmol.[182] The force field BILFF was applied for the ILs and cellulose, while TIP4P-EW was used for water,[183], with constraints on the bonds and angles using the RATTLE algorithm.[184, 185] All simulations followed the procedure described in Section 3.3 of this thesis and were performed at a temperature of 350 K. The molecular structures were created using VMD[123] and Tachyon.[124] All analyses were performed using the TRAVIS software package.[36, 168]

Microstructure of the System

In this section RDFs were used to study the strength of the anion–cellulose hydrogen bond in the four different ILs. In Fig. 3.4, the RDFs between the hydroxyl

TABLE 3.2: Simulation parameters and physical simulation time (sim. time) of the final equilibrated force field MD simulations of the different ILs in absence and presence of water with a molar fraction of $\chi_{IL} = 0.25$. The mass fraction of water depicted in the table refers to the solvent mixture.

System	Composition	Sim. time /ns	Box size /pm	Density /g·cm^{-3}
	[EMTr][OAc]			
0 wt.% H$_2$O	1 cellulose pentamer 192 [EMTr][OAc]	50	3683	1.12
23 wt.% H$_2$O	1 cellulose pentamer 150 [EMTr][OAc] 450 H$_2$O	50	3740	1.10
	[EMIm][OAc]			
0 wt.% H$_2$O	1 cellulose pentamer 192 [EMIm][OAc]	50	3715	1.09
24 wt.% H$_2$O	1 cellulose pentamer 150 [EMIm][OAc] 450 H$_2$O	50	3760	1.08
	[EMTr][OBz]			
0 wt.% H$_2$O	1 cellulose pentamer 192 [EMTr][OBz]	50	4059	1.13
18 wt.% H$_2$O	1 cellulose pentamer 150 [EMTr][OBz] 450 H$_2$O	50	4036	1.11
	[EMIm][OBz]			
0 wt.% H$_2$O	1 cellulose pentamer 192 [EMIm][OBz]	50	4082	1.11
19 wt.% H$_2$O	1 cellulose pentamer 150 [EMIm][OBz] 450 H$_2$O	50 000	4054	1.09
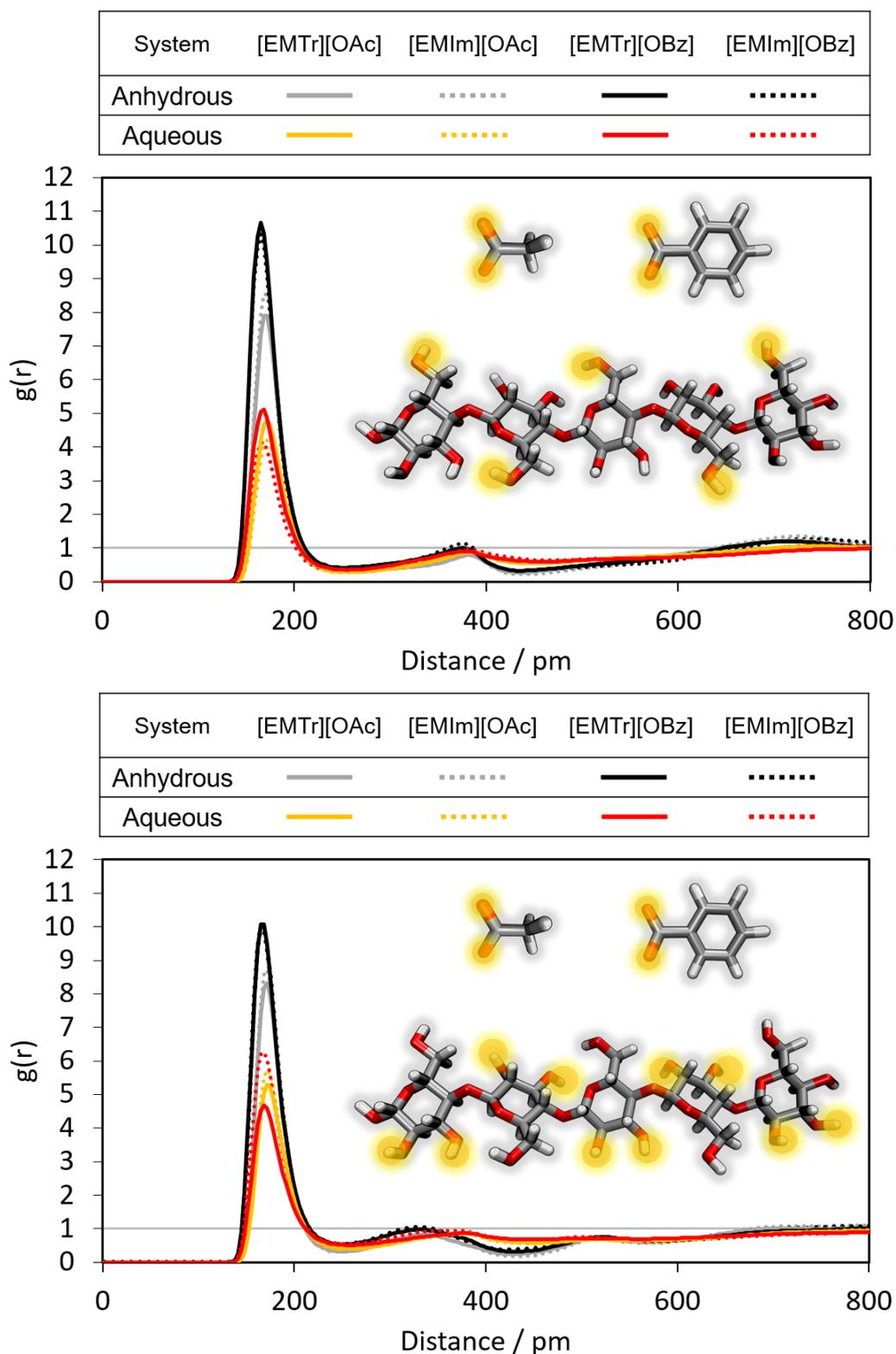
FIGURE 3.4: Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OBz]⁻ as well as [OAc]⁻ and the marked protons (HO6 (**top**) and HOR (**bottom**)) of cellulose in anhydrous and aqueous [EMTr][OAc], [EMIm][OAc], [EMTr][OBz] and [EMIm][OBz] calculated from force field MD simulations using BILFF. The RDFs are averaged over the marked protons, respectively.

groups HO6/HOR and the anions $[OBz]^-/[OAc]^-$ in the four ILs are compared. The impact of the cation on the cellulose–anion interaction does not exhibit a distinct trend, whereas the choice of anion significantly influences the attraction between the molecules: the particle distance of the hydrogen bond partners in the benzoate-containing ILs is smaller than in the acetate-containing ones (165 pm vs. 170 pm). In the anhydrous system, the anion in [EMIm][OAc] and [EMTr][OAc] forms relatively weaker interactions with cellulose compared to [EMIm][OBz] and [EMTr][OBz]. This effect is surprising in view of the higher cellulose solubility in [EMIm][OAc]. In fact, no trivial answer to the problem can be given here. It appears that the solubility of cellulose does not only depend on the hydrogen bond strength, but possibly also on other factors such as steric hindrance, aromaticity of the ions and other factors such as the possible formation of an anion–cellulose chelate complex. Further investigations are still needed, and BILFF provides the necessary tool for force field MD simulations of such large systems.

In the presence of water, the hydrogen bonds between the hydroxyl groups and the anions are weakened. This is caused by competing hydrogen bonds between cellulose and water and shielding of the anion by water, as shown in Section 3.3 (cf. Fig. 5 and 6) and A.4 (cf. Fig. S5) of the thesis. The influence of water on the cellulose–anion hydrogen bond varies depending on the IL in question, and the positioning of the hydroxyl groups also plays a role (also cf. 3.6.1). However, the intricate nature of the hydrogen bond network makes it difficult to discern clear trends in the relative strengths of these hydrogen bonds in the aqueous ILs.

Furthermore, while in an aqueous system the positioning of the hydroxyl groups appears to have an effect, in the anhydrous system the two protons of the hydroxyl groups exhibit a nearly identical hydrogen bonding behavior. This is in agreement with the results from Section 3.6.1 and 3.3 (cf. Fig. 3 and 4).

To investigate the influence of the anion on the spatial arrangement of the anion around cellulose, the distance–angle combined distribution function of the cellulose–anion hydrogen bond in an acetate-containing IL, [EMIm][OAc], is compared with a benzoate-containing one such as [EMIm][OBz] (cf. Fig. 3.5). The figures show the combined distance–angle distribution functions (CDFs) of the distance $H_{cellulose} \cdots O_{anion}$ and the corresponding angle $\angle(OH_{cellulose}, H_{cellulose}, O_{anion})$ as shown in the sketch, respectively. At an angle of about 135–180° and a particle distance of about 120–230 pm, the residence probability of the cellulose proton and the oxygen atoms of the anion has a maximum in both cases. Differences can be seen in the angular range of the second peak at about 400 pm, which is more delocalized and shifted to slightly smaller angles in the benzoate-containing system. This second peak originates from the second oxygen atom of the anions respectively.

Conclusion

In this section, as examples of applications of BILFF, the results of force field MD simulations of cellulose in four different ILs [EMTr][OAc], [EMIm][OAc], [EMIm][OBz], and [EMTr][OBz] were presented both in the presence and absence of water. The focus was on the comparative study of the strength of hydrogen bonds between cellulose and the anions in the different ILs, using RDFs and CDFs as the analysis tool. In the anhydrous system, the simulations showed that the anion in [EMIm][OAc] tended to interact less strongly with cellulose than in [EMTr][OBz] and [EMIm][OBz]. Given the common assumption that cellulose dissolves in ILs primarily due to pronounced hydrogen bonding with the anion, this result is surprising since
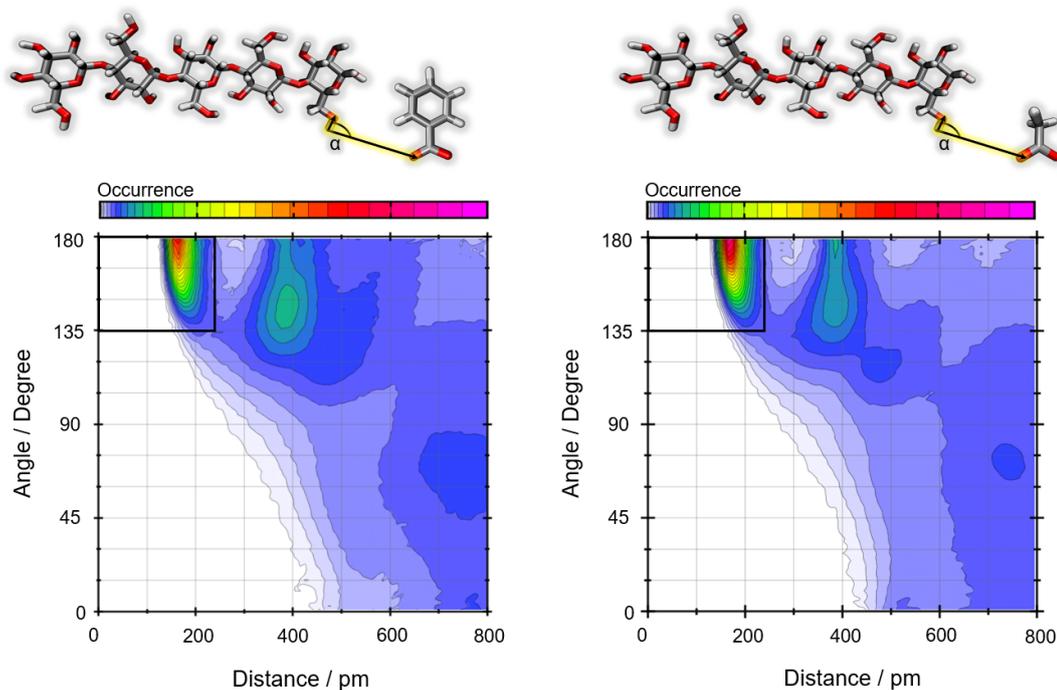
FIGURE 3.5: Distance–angle distribution function between an example proton of cellulose and the oxygen atoms of [OBz]$^-$ in aqueous [EMIm][OBz] (*left panel*) as well as [OAc]$^-$ in aqueous [EMIm][OAc] (*right panel*). Color code in arbitrary units. The black rectangle demonstrates the geometric criterion for the existence of a hydrogen bond, as used in the lifetime calculation.

[EMIm][OAc] is the most effective of the investigated solvents. This apparent discrepancy has not yet been clearly clarified. It appears that the solubility of cellulose depends not only on the strength of the hydrogen bonds, but possibly also on other factors such as the aromaticity of the ions and the possible formation of an anion–cellulose chelate complex. In this context, it should be noted that the considerations made here have focused primarily on the hydrogen bonds and structural parameters, and that the solubilization process cannot be reduced to these factors alone.

In the presence of water, the hydrogen bond network between the anion and cellulose is significantly weakened depending on the choice of the IL. However, the spatial geometry of the hydrogen bonds between cellulose and the anion remains almost unchanged in the presence of the aliphatic anion [OAc]$^-$ and the aromatic anion [OBz]$^-$.

In order to gain a fundamental understanding of the solubility of cellulose in ILs at an atomic level and a deeper insight into the complex interaction networks of the quaternary system, further investigations such as calculating solvation enthalpies and employing spectroscopic methods are necessary. As shown in this section by a brief selection of structural analyses of force field MD simulations, BILFF is a useful tool for providing force field simulations for comprehensive investigations of the cellulose solubility in ILs.

4 Conclusion

Within the scope of this thesis, a force field for molecular dynamics simulations of the cellulose solubility in different ILs, called BILFF, is introduced. BILFF is the first force field to provide parameters for the ILs 1-ethyl-3-methylimidazolium acetate ([EMIm][OAc]), 1-ethyl-3-methylimidazolium benzoate ([EMIm][OBz]), 1-ethyl-3-methyl-1,2,3-triazolium acetate ([EMTr][OAc]), 1-ethyl-3-methyl-1,2,3-triazolium benzoate ([EMTr][OBz]), as well as the co-solvent dimethyl sulfoxide (DMSO), optimized for accurately describing the hydrogen bond networks and cellulose solubility in these systems. The force field parameters were simultaneously optimized in both the presence and absence of water. This approach was adopted due to the hygroscopic nature of many imidazolium- and triazolium-based ionic liquids, where water can exert a significant influence on their properties.[180, 181] Furthermore, water is utilized for precipitation,[5, 31, 93, 98] underscoring its importance in technical applications.

The optimization of the force field parameters was performed by comparing various analyses of the force field simulations with reference AIMD simulations at 350 K: The bonded interactions were adjusted based on a comparison of the statistical occurrence of bond lengths, bond angles, and dihedral angles with the results of the reference AIMD simulations and by varying the force constants and the height of the potential barrier. The nonbonded force field parameters underwent an iterative tuning process, employing a trial-and-error method. These parameters were systematically adjusted until the discrepancies in the radial distribution functions of the primary hydrogen bonds between the force field MD and AIMD simulations were minimized. Furthermore, the Lennard-Jones parameters were fine-tuned to ensure an accurate replication of the experimental density. The most accurate replication of the experimental diffusion coefficients was attained by assigning a total ionic charge of ± 0.82 to all four ILs.

BILFF is a toolbox for force field MD simulations that enables the simulation of cellulose/IL/DMSO/water systems within a few days. The accuracy of these simulations closely rivals that of quantum chemical calculations. This encompasses not only structural parameters such as ion pairing, solvation shell structure, $\pi - \pi$-stacking and hydrogen bonding but also dynamic properties such as diffusion coefficients and hydrogen bond lifetimes. It was shown that experimental data, such as the thermal volume expansion coefficient, compressibility, self-diffusion coefficient, density, enthalpy of vaporization (of pure DMSO), and, in the case of cellulose, X-ray diffraction data on the crystal structure, can be successfully reproduced with high accuracy. Furthermore, it was shown that the force field parameters for [EMTr]$^+$ in the IL [EMTr][OAc] and [OBz]$^-$ in [EMIm][OBz] can be transferred to [EMTr][OBz] without the necessity for adjustments, thereby substantiating the precision of the underlying force field optimization. Additionally, excellent agreement with quantum chemical simulations and experimental data across a wide temperature range of 350–550 K was shown, expanding the parameter space for optimizing the solution process.

The versatility of BILFF is demonstrated by force field molecular dynamics simulations of cellulose/IL/(DMSO)/(water). Both the microstructure and dynamics in the different systems were analyzed. A comparison of the anhydrous solvents [EMIm][OAc], [EMTr][OAc], [EMIm][OBz], and [EMTr][OBz] revealed that the

cation–anion hydrogen bond of the aromatic benzoate anion is slightly stronger than that of the aliphatic acetate, despite the lower partial charge of the oxygen atoms of benzoate ($-0.55\,e$ vs. $-0.66\,e$). However, this trend is reversed in the presence of water. One possible explanation for this heightened occurrence of hydrogen bonding of benzoate is linked to the reduced rotational and translational freedom caused by the T-shaped arrangement of benzoate, resulting in a more stable spatial arrangement and prolonged duration of hydrogen bonds. Moreover, it has been demonstrated that strong hydrogen bonds of the anion play an ambivalent role: The increased strength of the hydrogen bond reinforces the cation–anion interaction, counteracting cellulose dissolution by reducing the availability of anions to interact with cellulose. In terms of cations, triazolium demonstrated a stronger cation–anion and cation–water hydrogen bonding than imidazolium. Since the only molecular difference between the two cations is the CH/N substitution, this increased strength is likely due to the larger partial charges of the two remaining ring protons in $[\text{EMTr}]^+$ ($+0.19\,e$ vs. $+0.15\,e$). The disparity between $[\text{EMTr}]^+$ and $[\text{EMIm}]^+$ is also evident in $[\text{EMTr}]^+$ forming one less hydrogen bond and having lower cellulose solubility. This observation supports the hypothesis that hydrogen bonds in general play a critical role in cellulose solubility, a topic that remains under discussion.[31, 36, 40, 41]

An investigation of the four anhydrous ILs resulted in the following ranking of the strength of the cation–anion hydrogen bonds: [EMTr][OBz] > [EMTr][OAc] > [EMIm][OBz] > [EMIm][OAc]. Simultaneously, the strength of the anion–cellulose hydrogen bond follows the order [EMIm][OBz] = [EMTr][OBz] > [EMIm][OAc] > [EMTr][OAc]. The anion is not the only hydrogen bond partner for cellulose. Weak hydrogen bonds of cellulose with the cation and water as well as intramolecular cellulose hydrogen bonds also occur confirming the literature.[36, 50–53] Furthermore, a dependency of the hydrogen bond strength between cellulose and the anion on the position of the hydroxyl group in the cellulose chain was observed. It was demonstrated that a distinction exists between the hydroxyl groups on the carbon ring (at C2 and C3) and the freely rotatable, slightly more positively charged protons of the hydroxyl groups on the methylene group (at C6) in cellulose. Moreover, it was found that distinguishing between the bridging acetal oxygen atom and the ring oxygen atom is necessary to accurately describe the intermolecular interactions. Thus, BILFF is the first force field to differentiate between these two types of atoms in its parameterization.

The MD simulations also reveal that water plays an ambivalent role in cellulose solubility. On one hand, even minor amounts, such as 5 wt.%, are sufficient to hinder access for the anion by shielding the cellulose hydroxyl groups to varying degrees. On the other hand, water reduces the interaction between the cation and anion, thereby facilitating the anion–cellulose hydrogen bond. This dual effect underscores the ongoing debate regarding water's function not only as an antisolvent but also as a co-solvent, depending on its concentration.[186]

DMSO, one of the most common co-solvents, significantly increases the solubility of cellulose in ILs.[44–46] The simulation results prove that DMSO reduces the lifetime of the cation–anion hydrogen bond. This finding supports the hypothesis that DMSO shields the ions of the IL from each other and thus increases the solubility of cellulose by increasing the number of free anions.[44–46] However, a contrary effect occurs in the presence of water in the mixture DMSO/[EMIm][OAc]/water, highlighting the complexity of the underlying hydrogen bonding network. DMSO leads to a more than two-fold increase in the lifetime of hydrogen bonds within the IL, as well as between the IL and water, when compared to the DMSO-free IL/water mixture. The reason is most likely connected to the ambivalent influence of water in the

system, as well as the influence of DMSO not only on the hydrogen bond stability and strength, but also on system density and molecular mobility. Using BILFF, further insights into this phenomenon can be gained through concentration-dependent simulations, where both water and DMSO content are varied, along with exploring different IL compositions.

While BILFF was optimized to accurately describe the complex hydrogen bond network, it provides the basis for the investigation of various factors influencing cellulose solubility in the different ILs. One factor is the spatial size of the anion, which significantly affects its diffusivity and, consequently, its ability to penetrate the cellulose crystal or fibrils, initiating the disruption of the internal cellulose hydrogen bonding network. Additionally, the influence of van der Waals effects should be considered, as they impact the balance of interactions among the anion, cation, water, DMSO, and cellulose molecules. Furthermore, the chirality of the anion plays a role, as we have shown in a previous publication,[37] along with enthalpic effects.

As demonstrated through its applications, BILFF proves to be a fast and powerful tool for in-depth investigations into the solubility of cellulose in diverse solvent compositions via force field MD simulations. The remarkable accuracy of BILFF not only enables further in-depth studies of cellulose solubility in different solvents, as demonstrated in this work. The short simulation times of force field MD simulations with BILFF lead to new possibilities for optimizing the chemical processing of cellulose, including the mapping of phase diagrams even in complex systems such as IL/cellulose/water/co-solvent. Moreover, it allows for the identification of optimal solvent mixtures for cellulose.

A Appendix: Supporting Information of the Paper – Force Field Parameter and Additional Analysis

A.1 Supporting Information to Section 3.2 - Article I

A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 1: [EMIm][OAc] / Water Mixtures

Supporting Information

Eliane Roos and Martin Brehm*

Institut für Chemie - Theoretische Chemie, Martin-Luther-Universität Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany.

*E-mail: Martin_Brehm@gmx.de

*Website: https://brehm-research.de/

Force Field Parameters

In this section, we report our optimized force field parameters for the 1-ethyl-3-methylimidazolium cation and the acetate anion. The parameters are compatible to OPLS–AA.[1-3] In accordance to OPLS–AA, the non-bonded Lennard-Jones and Coulomb interactions for 1–2 and 1–3 neighbors are neglected, while those for 1–4 neighbors are scaled with a factor of 0.5 (see scaling factor f_{ij} in Equation 1 below). Geometric mixing rules are applied for the Lennard-Jones interactions. We used a Coulomb and Lennard-Jones cutoff radius of 800 pm and a PPPM long-range Coulomb solver *(as implemented in LAMMPS)*[4] in all our simulations.

When simulating mixtures with water, the TIP4P–EW force field[5] *(with fixed bonds and angles)* has been applied for the water molecules. These parameters have not been modified within this work, they are used as reported in the original publication.
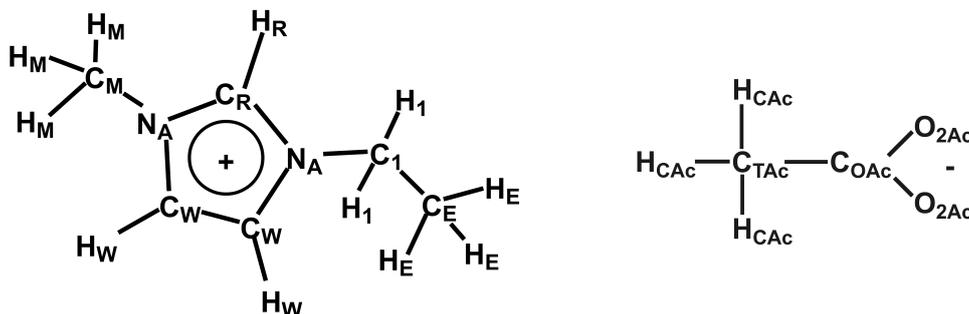
Figure S-1: Atom types in the 1-ethyl-3-methylimidazolium cation *(left)* and the acetate anion *(right)*. For the corresponding atom classes, see Table S-1.

Table S-1: Atom types *(see Figure S-1)* and corresponding atom classes for the 1-ethyl-3-methyl-imidazolium cation *(upper part)* and the acetate anion *(lower part)*. The atom types are used for the non-bonded interactions *(see Table S-2)*, while the atom classes are used for the bonded interactions *(see Tables S-3 to S-6)*.

Atom Type	Atom Class
CR	CR
C1	CT
CE	CT
CM	CT
CW	CW
HR	HA
HW	HA
H1	HC
HE	HC
HM	HC
NA	NA
COAc	CO
CTAc	CT
HCAc	HC
O2Ac	O2

The potential energy in OPLS–AA is calculated according to the following equation:

$$U(r^N) = \sum_{\text{bonds}}^{i} \frac{k_l}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}}^{i} \frac{k_\theta}{2}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{\text{dihedrals}}^{i} \left[\frac{V_1^i}{2}[1 + cos(\phi_i)] + \frac{V_2^i}{2}[1 - cos(2\phi_i)] + \frac{V_3^i}{2}[1 + cos(3\phi_i)]\right] \qquad (1)$$

$$+ \sum_{i=1}^{N}\sum_{j=i+1}^{N} \left[4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}}\right] f_{ij} .$$

Table S-2: Atomic partial charges q (total ion charge ±0.82) and Lennard-Jones parameters σ and ϵ for the 1-ethyl-3-methylimidazolium cation *(upper part)* and the acetate anion *(lower part)*. For the atom types, see Figure S-1.

Atom Type	q	σ / Å	ϵ / kJ mol^{-1}
C1	-0.147	3.34	0.276
CE	-0.042	3.34	0.276
CM	-0.147	3.34	0.276
CR	-0.096	3.38	0.293
CW	-0.113	3.38	0.293
H1	$+0.116$	2.38	0.126
HE	$+0.055$	2.38	0.126
HM	$+0.116$	2.38	0.126
HR	$+0.165$	1.43	0.126
HW	$+0.150$	1.62	0.126
NA	$+0.134$	3.10	0.711
COAc	$+0.575$	3.57	0.439
CTAc	-0.230	3.34	0.276
HCAc	$+0.049$	2.38	0.126
O2Ac	-0.656	2.96	0.879

Please note: The force constants k for the bonds and angles and the torsion parameters V do **not** include the factor $\frac{1}{2}$ from the potential energy equation *(see Equation 1)*.

Table S-3: Bond equilibrium lengths l_0 and force constants k_l for the 1-ethyl-3-methylimidazolium cation and the acetate anion. For the atom classes, see Table S-1.

Bond Type	l_0 / Å	k_l / kJ mol^{-1} Å^{-2}
CO–O2	1.284	4457
CR–HA	1.089	2570
CT–CO	1.554	1820
CT–CT	1.532	2097
CT–HC	1.099	3013
CW–CW	1.374	4019
CW–HA	1.085	2943
NA–CR	1.356	3636
NA–CT	1.485	2078
NA–CW	1.403	2775

Table S-4: Angle equilibrium values θ_0 and force constants k_θ for the 1-ethyl-3-methylimidazolium cation and the acetate anion. For the atom classes, see Table S-1.

Angle Type	θ_0 / Deg	k_θ / kJ mol^{-1} rad^{-2}
O2–CO–O2	127.2	790.6
CT–CO–O2	117.7	524.4
NA–CR–HA	126.2	244.9
NA–CR–NA	107.6	831.0
CO–CT–HC	110.3	303.6
CT–CT–HC	111.4	338.6
HC–CT–HC	109.2	308.0
NA–CT–CT	110.9	532.7
NA–CT–HC	107.2	411.7
CW–CW–HA	131.7	236.7
NA–CW–CW	107.0	805.7
NA–CW–HA	120.8	263.0
CR–NA–CW	110.0	842.8
CT–NA–CR	124.9	370.2
CT–NA–CW	125.2	367.8

Table S-5: Torsion parameters V_n for the 1-ethyl-3-methylimidazolium cation and the acetate anion. For the atom classes, see Table S-1.

Torsion Angle Type	V_1 / kJ mol^{-1}	V_2 / kJ mol^{-1}	V_3 / kJ mol^{-1}
CT–NA–CR–NA	0.000	19.460	0.000
CT–NA–CR–HA	0.000	19.460	0.000
CW–NA–CR–NA	0.000	39.000	0.000
CW–NA–CR–HA	0.000	19.460	0.000
CR–NA–CT–HC	0.000	0.000	0.000
CR–NA–CT–CT	10.540	0.000	0.000
CW–NA–CT–HC	0.000	0.000	0.519
CW–NA–CT–CT	14.300	−12.200	−1.590
CR–NA–CW–HA	0.000	12.550	0.000
CR–NA–CW–CW	0.000	25.000	0.000
CT–NA–CW–HA	0.000	12.550	0.000
CT–NA–CW–CW	0.000	12.550	0.000
HC–CT–CT–HC	0.000	0.000	1.255
NA–CT–CT–HC	0.000	0.000	0.367
HC–CT–CO–O2	0.000	0.000	0.000
HA–CW–CW–HA	0.000	30.000	0.000
NA–CW–CW–HA	0.000	44.980	0.000
NA–CW–CW–NA	0.000	65.000	0.000

Please note: Improper torsions are handled as torsions; there is no extra term for impropers in the potential energy equation.

Table S-6: Improper torsions for the 1-ethyl-3-methylimidazolium cation and the acetate anion. For the atom classes, see Table S-1.

Improper Torsion Type	V_1 / kJ mol^{-1}	V_2 / kJ mol^{-1}	V_3 / kJ mol^{-1}
NA–CW–CW–HA	0.000	8.200	0.000
CR–CW–NA–CT	0.000	8.370	0.000
CR–CW–NA–HA	0.000	8.370	0.000
NA–NA–CR–CT	0.000	9.200	0.000
NA–NA–CR–HA	0.000	9.900	0.000
CT–O2–CO–O2	0.000	90.000	0.000

Radial Distribution Functions
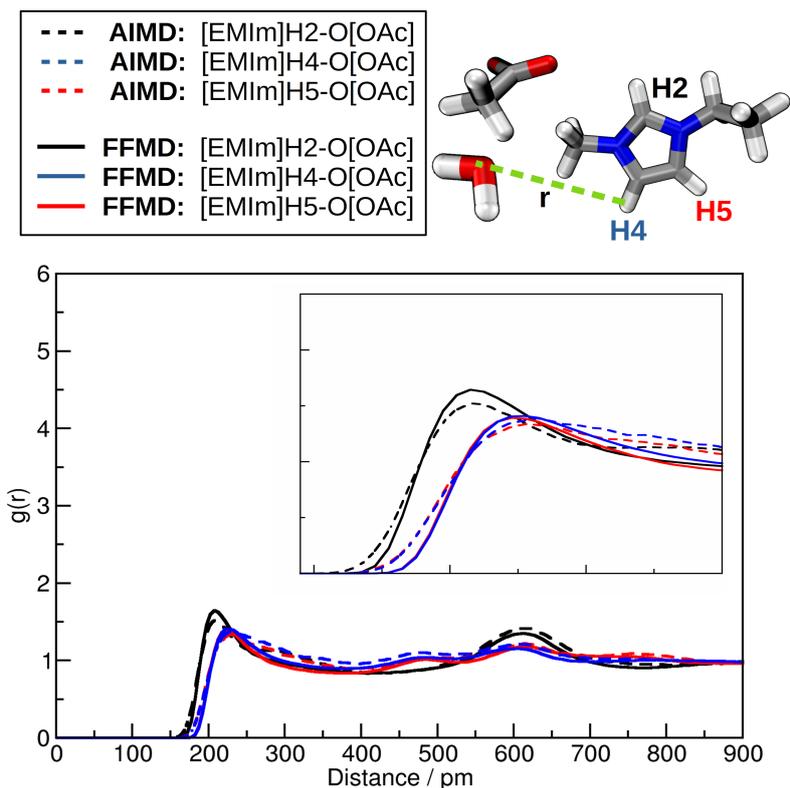
Figure S-2: Comparison of the radial distribution functions (RDFs) of the AIMD and the force field MD using BILFF between the hydrogen atoms H2, H4 and H5 of $[EMIm]^+$ and the oxygen atom of water in the aqueous ionic liquid.
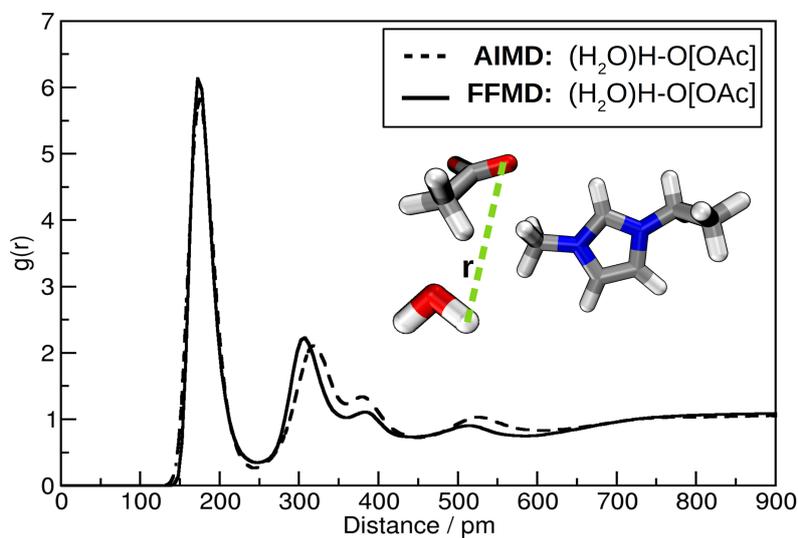
Figure S-3: Comparison of the radial distribution functions (RDFs) of the AIMD and the force field MD using BILFF between the hydrogen atoms of water and the oxygen atoms of acetate in the aqueous ionic liquid.

Combined Distribution Functions
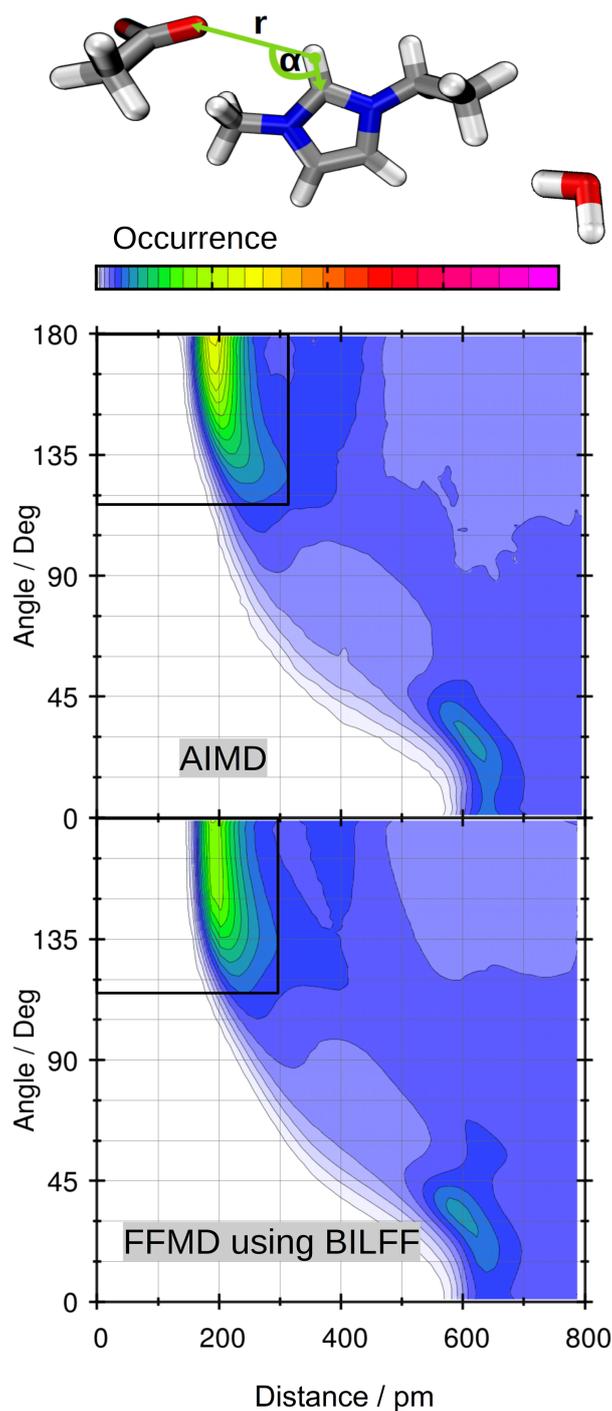
Figure S-4: Combined distribution functions (CDFs) from the AIMD *(upper panel)* and the force-field MD using BILFF *(lower panel)* between the hydrogen atom H2 of [EMIm]$^+$ and the acetate oxygen in the aqueous ionic liquid with marked geometric criterion for the dimer autocorrelation function of this hydrogen bond *(black rectangle)*.
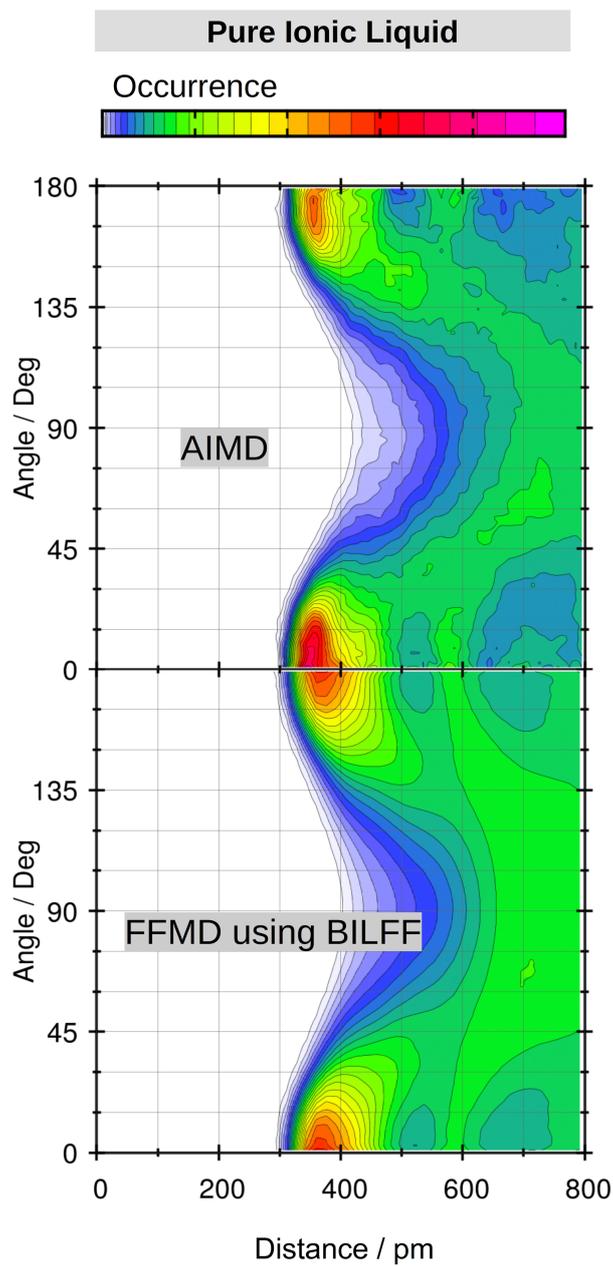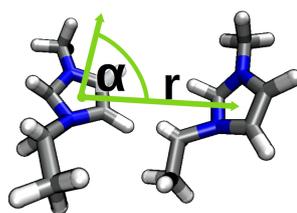
Figure S-5: Combined distribution functions (CDFs) from the AIMD *(upper panel)* and the force field MD using BILFF *(lower panel)* between the ring centers of two [EMIm]$^+$ cations in the pure ionic liquid.
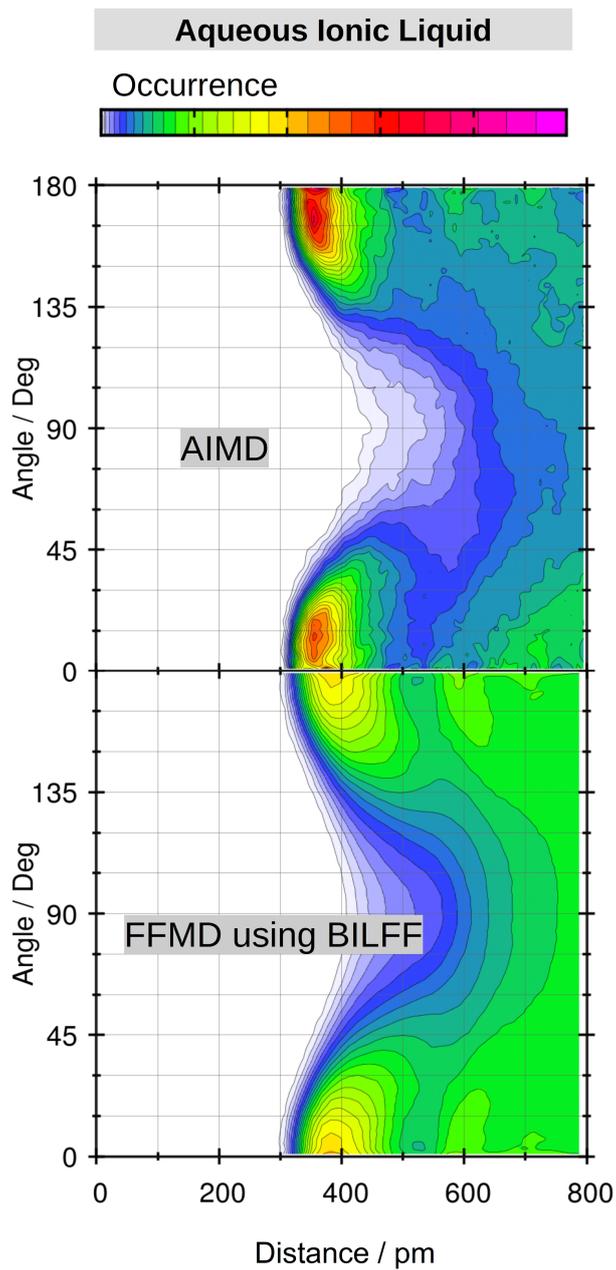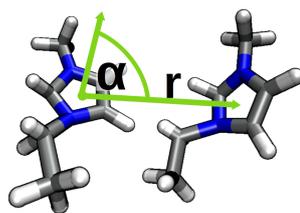
Figure S-6: Combined distribution functions (CDFs) from the AIMD *(upper panel)* and the force field MD using BILFF *(lower panel)* between the ring centers of two [EMIm]$^+$ cations in the aqueous ionic liquid.

Arrhenius Plot of the Diffusion Coefficients

Figures S-7 and Fig. S-8 show the Arrhenius plots of the calculated diffusion coefficients at different temperatures of pure and aqueous [EMIm][OAc]. The linear shape in the logarithmic–reciprocal plots indicates an Arrhenius behavior of the diffusion coefficients, which is frequently observed for ionic liquids.[6]
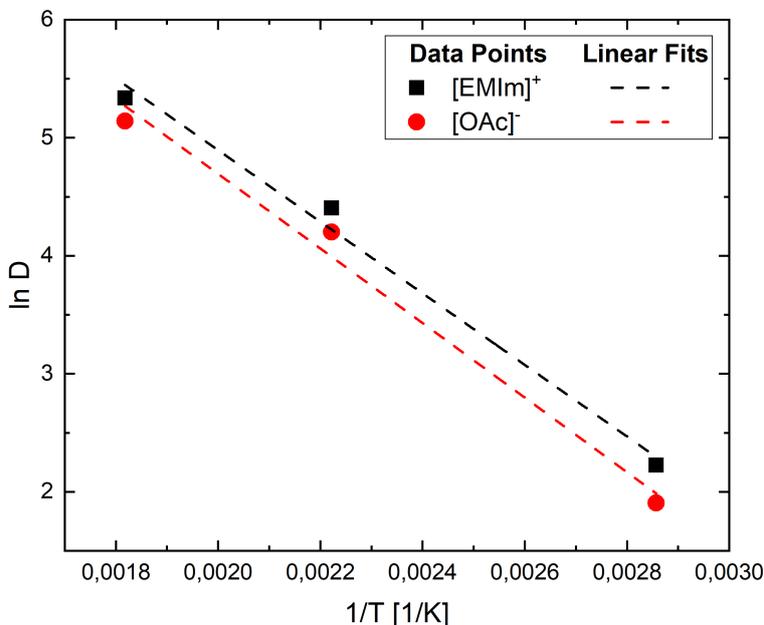
Figure S-7: Arrhenius plot of the calculated diffusion coefficients from our FFMD simulations of the pure system using BILFF.

Table S-7: Regression parameters of the linear regression $y = m \cdot x + n$ in the corresponding Arrhenius plots, Fig. S-7 and S-8.

System	Parameter	$[EMIm]^+$	$[OAc]^-$	H_2O
pure	n	10.96 ± 0.71	11.01 ± 0.82	–
	m	-3032 ± 304	-3159 ± 350	–
	R^2	0.980	0.976	–
aqueous	n	11.47 ± 0.59	11.50 ± 0.22	$12.40 \pm 0,15$
	m	-2917 ± 252	-2928 ± 95	-2926 ± 64
	R^2	0.985	0.998	0.999
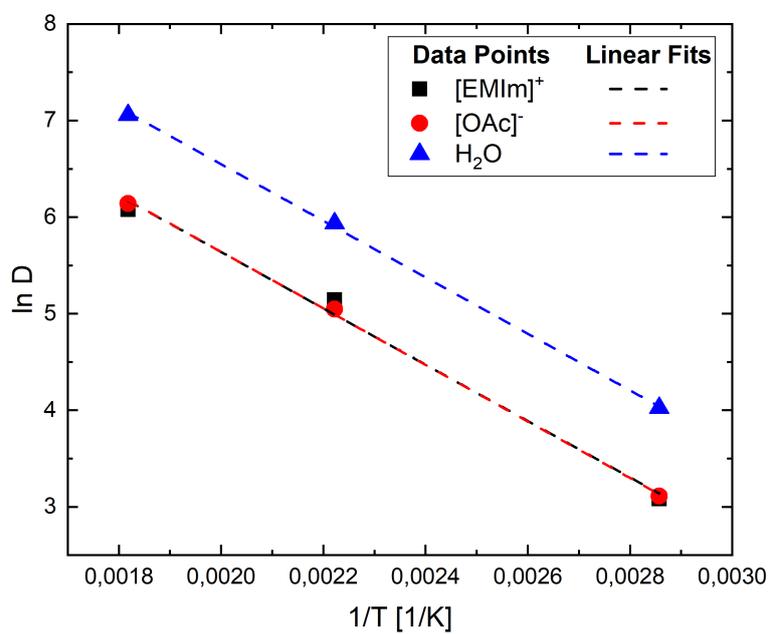
Figure S-8: Arrhenius plot of the calculated diffusion coefficients from our FFMD simulations of the aqueous system using BILFF.

Diffusive Regime

The double logarithmic plot of the mean squared displacement (MSD) over the correlation depth (see Fig. S-9) shows whether the dynamics of a system is diffusive or subdiffusive.[7,8] For both the pure and the aqueous system, a convergence towards a slope of 1 with increasing correlation depth is recognizable. At a temperature of $350\,\mathrm{K}$, a slope of 1 (and thus a diffusive dynamics) is only reached at large correlation depths of around $1\,\mathrm{ns}$ in the pure IL.
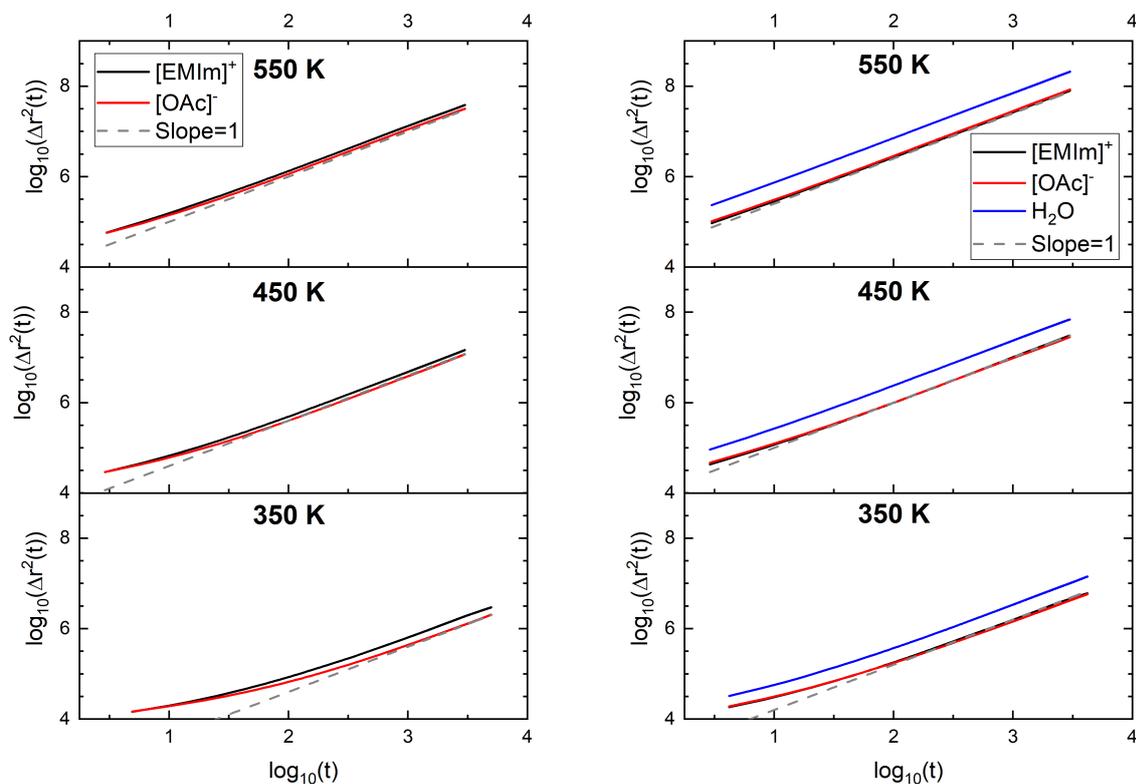
Figure S-9: The double logarithmic plot of the mean squared displacement (MSD) over the correlation depth of the pure (*left*) and aqueous (*right*) system at different temperatures. The dashed lines indicate a slope of 1, *i. e.*, diffusive behavior.

References

(1) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(2) Ponder, J. W.; Case, D. A. *Protein Simulations*; Advances in Protein Chemistry; Elsevier, 2003; Vol. 66; pp 27–85.

(3) Sambasivarao, S. V.; Acevedo, O. Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids. *J. Chem. Theory Comput.* **2009**, *5*, 1038–1050.

(4) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comp. Phys.* **1995**, *117*, 1–19.

(5) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(6) Tokuda, H.; Hayamizu, K.; Ishii, K.; Susan, M. A. B. H.; Watanabe, M. Physicochemical Properties and Structures of Room Temperature Ionic Liquids. 1. Variation of Anionic Species. *J. Phys. Chem. B* **2004**, *108*, 16593–16600.

(7) Del Pópolo, M. G.; Voth, G. A. On the Structure and Dynamics of Ionic Liquids. *J. Phys. Chem. B* **2004**, *108*, 1744–1752.

(8) Perkins, S. L.; Painter, P.; Colina, C. M. Molecular Dynamic Simulations and Vibrational Analysis of an Ionic Liquid Analogue. *J. Phys. Chem. B* **2013**, *117*, 10250–10260.

A.2 Supporting Information to Section 3.3 - Article II

english

A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 2: Cellulose in [EMIm][OAc] / Water Mixtures

Supporting Information

Eliane Roos, Daniel Sebastiani and Martin Brehm*

Institut für Chemie - Theoretische Chemie, Martin-Luther-Universität Halle-Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany.

*E-mail: Martin_Brehm@gmx.de

*Website: https://brehm-research.de/

1 Force Field Parameters

In this section, our optimized force field parameters for cellulose in pure and aqueous [EMIm][OAc] are presented. The nomenclature used for the atom types *(used for the non-bonded interactions)* can be found in Figure S–1 while the atom classes *(used for the bonded interactions)* are shown in Table S–1. Please note that the atom types OH1, HO1, OH4, and HO4 correspond to terminal hydroxyl groups and are found only once on each end of the cellulose strand. The force field is based on the potential energy equation of the OPLS–AA force field:[1–3]

$$
\begin{aligned}
U(r^N) = &\sum_{i \in \text{bonds}} k_{l,i}(l_i - l_{i,0})^2 + \sum_{i \in \text{angles}} k_{\theta,i}(\theta_i - \theta_{i,0})^2 \\
&+ \sum_{i \in \text{dihedrals}} \left[\frac{V_{i,1}}{2}[1 + \cos(\phi_i)] + \frac{V_{i,2}}{2}[1 - \cos(2\phi_i)] \right. \\
&\left. + \frac{V_{i,3}}{2}[1 + \cos(3\phi_i)] + \frac{V_{i,4}}{2}[1 - \cos(4\phi_i)] \right] \\
&+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j e^2}{4\pi \epsilon_0 r_{ij}} \right] f_{ij} \, .
\end{aligned}
\tag{S–1}
$$

$$
\epsilon_{i,j} = \sqrt{\epsilon_i \, \epsilon_j}, \qquad \sigma_{i,j} = \sqrt{\sigma_i \, \sigma_j}
\tag{S–2}
$$

The force constants k for the bonds and angles and the torsion parameters V in Tab. S–2, S–3, S–4, and S–5 do not include the factor $1/2$. In accordance with OPLS–AA, the interactions between the atomic neighbors 1–2 and 1–3 were not taken into account; the 1–4 interactions were scaled with $f_{ij} = 0.5$. For the Lennard-Jones cross terms, geometric mixing rules were applied *(see Equation S–2)*. The Coulomb and Lennard-Jones cutoff radius were set to 800 pm. For the Coulomb interactions, the PPPM long-range Coulomb solver (as implemented in LAMMPS)[4] was applied. For [EMIm][OAc], the force field BILFF[5] and for water TIP4P–EW[6] (with constrained bonds and angles) was used without modifications. The total charge of the ions amounts to ± 0.82 in BILFF.[5]
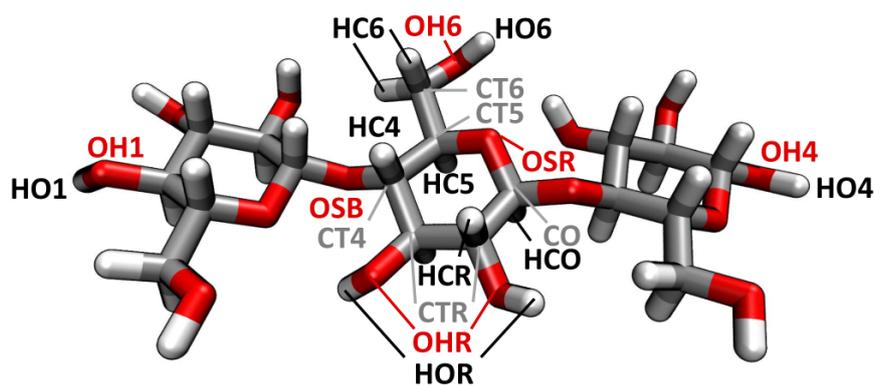
Figure S–1: Nomenclature of cellulose atom types in our force field BILFF. The corresponding atom classes are listed in Table S–1.

Table S–1: Nomenclature of atom types and atom classes in cellulose. The atom types are used for the non-bonded interactions *(see Table S–2)*, while the atom classes are used for the bonded interactions *(see Tables S–3*, S–4, and S–5).

Atom Type	Atom Class
CO	CO
CTR	CT
CT4	CT
CT5	CT
CT6	CT
OSB	OSB
OSR	OSR
OH1	OH
OH4	OH
OHR	OH
OH6	OH
HCO	HC
HCR	HC
HC4	HC
HC5	HC
HC6	HC
HOR	HO
HO1	HO
HO4	HO
HO6	HO

Table S–2: Optimized atomic partial charges q and Lennard-Jones parameters σ and ϵ for cellulose in comparison to the literature force field of W. Damm *et al.*[7]

Atom Type	BILFF			Literature[7]		
	q	σ	ϵ	q	σ	ϵ
		/ Å	/ kJ mol^{-1}		/ Å	/ kJ mol^{-1}
CO	0.3050	3.50	0.27614	0.3000	3.50	0.27614
CTR	0.2100	3.50	0.27614	0.2050	3.50	0.27614
CT4	0.2100	3.50	0.27614	0.2050	3.50	0.27614
CT5	0.1750	3.50	0.27614	0.1700	3.50	0.27614
CT6	0.1600	3.50	0.27614	0.1450	3.50	0.27614
OSB	−0.3890	2.90	0.58576	−0.3970	2.90	0.58576
OSR	−0.3887	2.90	0.58576	−0.4000	2.90	0.58576
OH1	−0.6920	3.07	0.71128	−0.7000	3.07	0.71128
OH4	−0.6180	3.07	0.71128	−0.7000	3.07	0.71128
OHR	−0.6850	3.07	0.71128	−0.7000	3.07	0.71128
OH6	−0.6600	3.12	0.71128	−0.6830	3.12	0.71128
HCO	0.1060	2.60	0.12552	0.1000	2.50	0.12552
HCR	0.0450	2.60	0.12552	0.0600	2.50	0.12552
HC4	0.0460	2.60	0.12552	0.0600	2.50	0.12552
HC5	0.0370	2.60	0.12552	0.0300	2.50	0.12552
HC6	0.0470	2.60	0.12552	0.0600	2.50	0.12552
HOR	0.3790	1.00	0.12552	0.4350	0.00	0.00000
HO1	0.4280	1.00	0.12552	0.4350	0.00	0.00000
HO4	0.4930	1.10	0.12552	0.4350	0.00	0.00000
HO6	0.4067	0.97	0.12552	0.4180	0.00	0.00000

Table S–3: Optimized bond equilibrium lengths l_0 and force constants k_l for cellulose in comparison to the literature force field of W. Damm *et al.*[7]

Bond	BILFF		Literature[7]	
	l_0	k_l	l_0	k_l
	/ Å	/ kJ mol^{-1} Å^{-2}	/ Å	/ kJ mol^{-1} Å^{-2}
CT–OSB	1.446	1815.7	1.410	1338.9
CT–OSR	1.442	1867.2	1.410	1338.9
CT–OH	1.440	2138.9	1.410	1338.9
CO–OSB	1.405	2001.7	1.380	1338.9
CO–OSR	1.441	1610.2	1.380	1338.9
CO–OH	1.441	2288.5	1.380	1338.9
CT–CT	1.530	1789.1	1.529	1121.3
CT–CO	1.527	1819.3	1.529	1121.3
CT–HC	1.101	2801.1	1.090	1422.6
CO–HC	1.101	2880.5	1.090	1422.6
OH–HO	0.965	3054.5	0.945	2313.8

Table S–4: Optimized angle equilibrium values θ_0 and force constants k_θ for cellulose in comparison to the literature force field of W. Damm *et al.*[7]

Angle	BILFF		Literature[7]	
	θ_0	k_θ	θ_0	k_θ
	/ Deg	/ kJ mol^{-1} rad^{-2}	/ Deg	/ kJ mol^{-1} rad^{-2}
CT–CT–CT	113.64	458.5	112.7	244.1
CT–CT–CO	112.38	407.4	112.7	244.1
CT–CT–HC	109.95	341.2	110.7	156.9
CO–CT–HC	109.02	323.9	110.7	156.9
CT–CO–HC	111.41	340.3	110.7	156.9
CT–CT–OH	110.87	480.2	109.5	209.2
CO–CT–OH	110.62	515.2	109.5	209.2
CT–CT–OSB	109.24	370.5	109.5	209.2
CT–CT–OSR	109.24	444.0	109.5	209.2
CT–CO–OSB	110.40	413.3	109.5	209.2
CT–CO–OSR	110.40	458.6	109.5	209.2
CT–CO–OH	114.22	506.3	109.5	209.2
CT–OSB–CO	108.12	320.9	109.5	251.0
CT–OSR–CO	108.10	307.7	109.5	251.0
CT–OH–HO	110.10	373.4	108.5	230.1
CO–OH–HO	113.67	364.8	108.5	230.1
HC–CT–HC	108.21	331.1	107.8	138.1
HC–CT–OSB	108.01	402.0	109.5	146.4
HC–CT–OSR	108.04	410.5	109.5	146.4
HC–CT–OH	109.84	365.0	109.5	146.4
HC–CO–OSB	108.60	312.9	109.5	146.4
HC–CO–OSR	108.63	378.8	109.5	146.4
HC–CO–OH	112.80	377.0	109.5	146.4
OSB–CO–OSR	108.80	589.5	111.6	387.4
OSR–CO–OH	106.75	731.5	111.6	387.4

Table S–5: Optimized torsional coefficients V_n for cellulose in comparison to the literature force field of W. Damm et al.[7]

Torsion Angle	BILFF				Literature[7]			
	V_1 kJ mol^{-1}	V_2 kJ mol^{-1}	V_3 kJ mol^{-1}	V_4 kJ mol^{-1}	V_1 kJ mol^{-1}	V_2 kJ mol^{-1}	V_3 kJ mol^{-1}	V_4 kJ mol^{-1}
HC–CT–CT–HC	0.0000	0.0000	1.3305	0.0000	0.0000	0.0000	1.3305	0.0000
HC–CO–CT–HC	0.0000	0.0000	1.3305	0.0000	0.0000	0.0000	1.3305	0.0000
HC–CT–CT–CT	0.0000	0.0000	1.5313	0.0000	0.0000	0.0000	1.5313	0.0000
HC–CO–CT–CT	0.0000	0.0000	1.5313	0.0000	0.0000	0.0000	1.5313	0.0000
HC–CT–CT–CO	0.0000	0.0000	1.5313	0.0000	0.0000	0.0000	1.5313	0.0000
CT–CT–CT–CT	7.2802	−0.6569	1.1673	0.0000	7.2802	−0.6569	1.1673	0.0000
CT–CT–CT–CO	7.2802	−0.6569	1.1673	0.0000	7.2802	−0.6569	1.1673	0.0000
HC–CT–CT–OH	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CT–CT–OSB	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CT–CT–OSR	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CT–CO–OSB	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CT–CO–OSR	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CO–CT–OH	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CT–CO–OH	0.0000	0.0000	1.9581	0.0000	0.0000	0.0000	1.9581	0.0000
HC–CT–OH–HO	0.0000	0.0000	1.8828	0.0000	0.0000	0.0000	1.8828	0.0000
HC–CO–OH–HO	0.0000	0.0000	1.8828	0.0000	0.0000	0.0000	1.8828	0.0000
HC–CT–OSB–CO	0.0000	0.0000	3.1798	0.0000	0.0000	0.0000	3.1798	0.0000
HC–CT–OSR–CO	0.0000	0.0000	3.1798	0.0000	0.0000	0.0000	3.1798	0.0000
HC–CO–OSB–CT	0.0000	0.0000	3.1798	0.0000	0.0000	0.0000	3.1798	0.0000
HC–CO–OSR–CT	0.0000	0.0000	3.1798	0.0000	0.0000	0.0000	3.1798	0.0000
CO–OSB–CT–CT	2.7196	−1.0460	2.8033	−4.1840	2.7196	−1.0460	2.8033	0.0000
CO–OSR–CT–CT	2.7196	−1.0460	2.8033	−4.1840	2.7196	−1.0460	2.8033	0.0000
CT–OSB–CO–CT	2.7196	−8.3680	8.3680	0.0000	2.7196	−1.0460	2.8033	0.0000
CT–OSR–CO–CT	2.7196	−8.3680	8.3680	0.0000	2.7196	−1.0460	2.8033	0.0000
CT–OSB–CO–OSR	−1.5690	−5.6819	4.1840	0.0000	−1.5690	−5.6819	0.0167	0.0000
CT–OSR–CO–OSB	−1.5690	−5.6819	4.1840	0.0000	−1.5690	−5.6819	0.0167	0.0000
CT–OSR–CO–OH	−1.5690	−5.6819	0.0167	0.0000	−1.5690	−5.6819	0.0167	0.0000
OSR–CO–OH–HO	−5.2593	−7.5563	0.0126	0.0000	−5.2593	−7.5563	0.0126	0.0000
CT–CT–CT–OH	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CT–CT–CT–OSB	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CT–CT–CT–OSR	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CT–CT–CO–OSB	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CT–CT–CO–OSR	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CT–CT–CO–OH	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CO–CT–CT–OH	−5.5898	0.0000	0.0000	0.0000	−5.5898	0.0000	0.0000	0.0000
CT–CT–OH–HO	11.1880	−12.0625	4.2928	0.0000	11.1880	−12.0625	4.2928	0.0000
CO–CT–OH–HO	11.1880	−12.0625	4.2928	0.0000	11.1880	−12.0625	4.2928	0.0000
CT–CO–OH–HO	11.1880	−12.0625	4.2928	0.0000	11.1880	−12.0625	4.2928	0.0000
OH–C –C –OS	18.0707	0.0000	0.0000	0.0000	11.1880	−12.0625	4.2928	0.0000
OH–CT–CT–OSB	18.0707	0.0000	0.0000	0.0000	18.0707	0.0000	0.0000	0.0000
OH–CT–CT–OSR	18.0707	0.0000	0.0000	0.0000	18.0707	0.0000	0.0000	0.0000
OSR–CT–CT–OSB	18.0707	0.0000	0.0000	0.0000	18.0707	0.0000	0.0000	0.0000
OH–CT–CO–OSB	18.0707	0.0000	0.0000	0.0000	18.0707	0.0000	0.0000	0.0000
OH–CT–CO–OSR	18.0707	0.0000	0.0000	0.0000	18.0707	0.0000	0.0000	0.0000
OH–CT–CT–OH	37.9321	0.0000	0.0000	0.0000	37.9321	0.0000	0.0000	0.0000
OH–CO–CT–OH	37.9321	0.0000	0.0000	0.0000	37.9321	0.0000	0.0000	0.0000

2 Distribution Functions
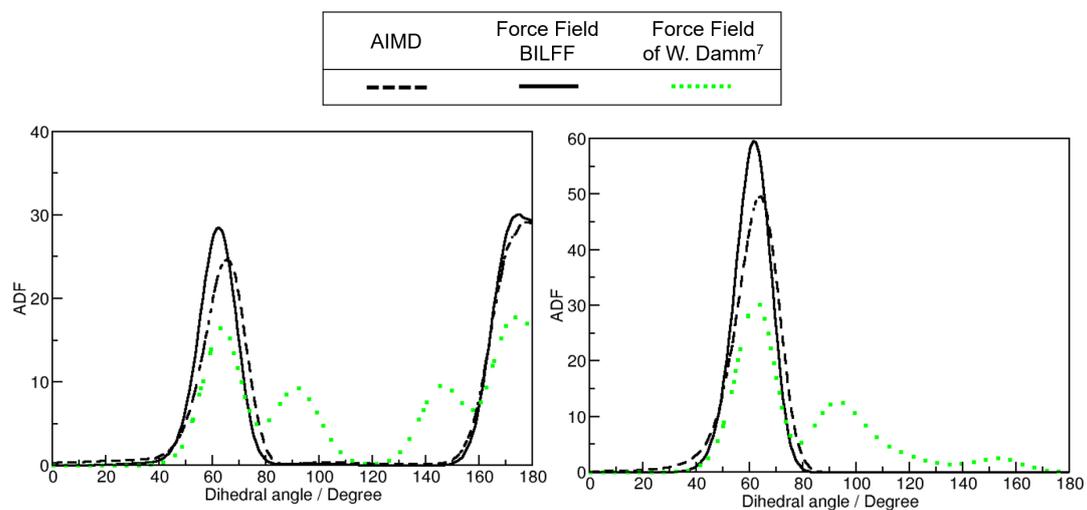
Figure S–2: Comparison of the angular distribution function of the dihedral angles CO–OSR–CT–CT *(left)* and CT–CO–OSR–CT *(right)* from the reference AIMD and the force field MD simulation using BILFF and the literature force field.[7]
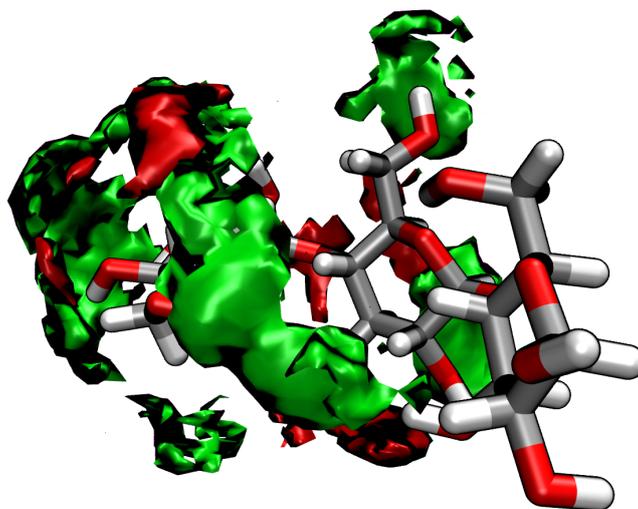
Figure S–3: Spatial distribution function of the ring protons of $[\mathrm{EMIm}]^+$ (red, $13\,\mathrm{nm}^{-3}$) and the acetate oxygen (green, $14\,\mathrm{nm}^{-3}$) averaged over all glucose units in the cellulose–IL system in the reference AIMD simulation.
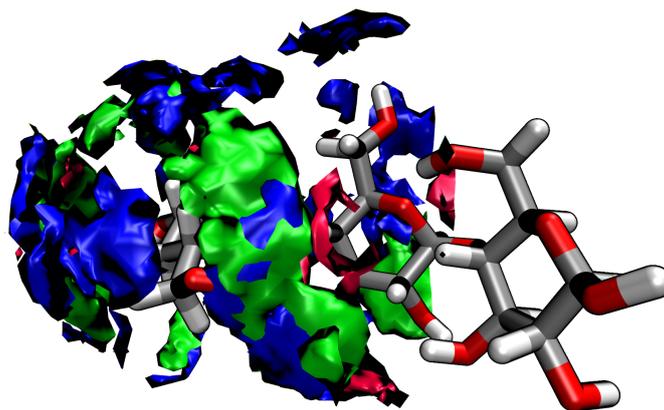
Figure S–4: Spatial distribution function of the ring protons of [EMIm]$^+$ (red, $9\,\mathrm{nm}^{-3}$), acetate (green, $19\,\mathrm{nm}^{-3}$) and water (blue, $25\,\mathrm{nm}^{-3}$) averaged over all glucose units in the cellulose–IL–water system in the reference AIMD simulation.

3 Investigation of the Statistical Variance of the Radial Distribution Function

In this section, the radial distribution functions of the hydrogen bonds between the hydroxyl protons of cellulose of type HO6 *(see Fig. S–5)* and HOR *(see Fig. S–6)* and the oxygen atoms of acetate are shown. As already described in the main text, the individual RDFs of the observed atoms were averaged and weighted according to their position in the molecule *(see Section 3 in the main text)*. The results were obtained from eleven force field MD simulations using our new parameters of the pure cellulose–[EMIm][OAc] system. Each of the simulations were performed with a different randomly generated velocity initialization but with the same force field parameters. All other simulation parameters are identical to each other. By differentiating the velocity initialization, a statistical variance of the results could be produced, which serve to evaluate our RDFs shown in the main text (here colored in black). This therefore represents a benchmark for the ergodicity of the simulations. As expected, a diversity of the height of the first maxima becomes observable. The position of those maxima is identical. This proves that the results of the force field simulations shown in the main text are within the variation range.
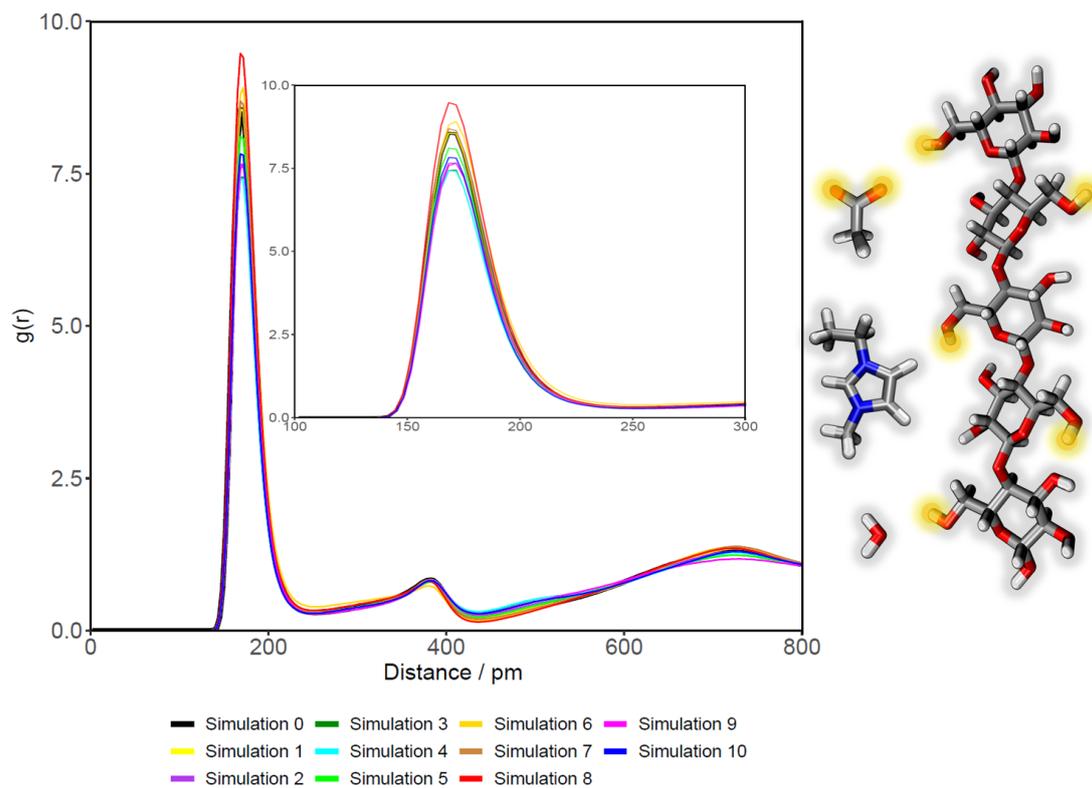
Figure S–5: Radial distribution functions of HO6-type hydroxyl protons from cellulose to acetate oxygen atoms in the pure cellulose–[EMIm][OAc] system. The force field simulations were performed with randomly generated different velocity initializations but with the same force field parameters of BILFF.
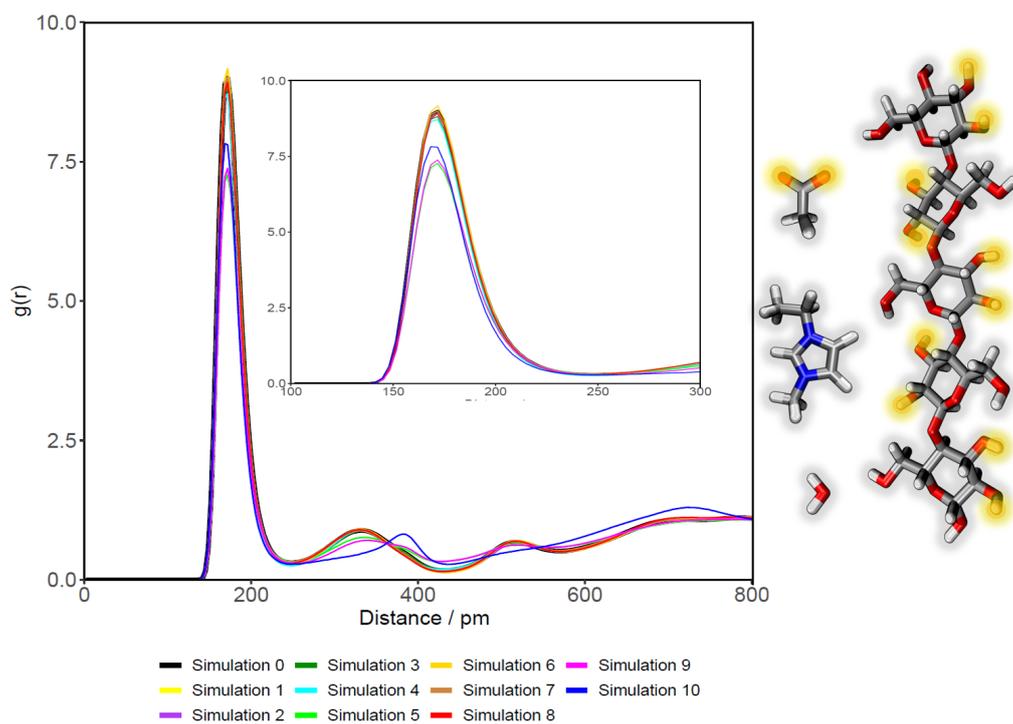
Figure S–6: Radial distribution functions of HOR-type hydroxyl protons from cellulose to acetate oxygen atoms in the pure cellulose–[EMIm][OAc] system. The force field simulations using BILFF were performed with randomly generated different velocity initializations.

4 Contact Matrices

In this section, the contact matrices from the AIMD simulations of the pure and aqueous cellulose–[EMIm][OAc] systems are shown. The numbering of the protons and oxygen atoms of cellulose used in the contact matrices calculated from the AIMDs is shown in Figure S–7. Figure S–8 presents the numbering of the cellulose atoms in the force field MD simulations.

In both cellulose–IL systems, hydrogen bonds between the hydroxyl protons of cellulose and the oxygen atoms of acetate and water are dominant, as well as hydrogen bonds between the protons of water and the oxygen atoms of cellulose. Hydrogen bonds are also formed between the ring protons of [EMIm]$^+$ and the oxygen atoms of cellulose. The AIMD simulations thus show the same trends of the hydrogen bond network as the force field simulations with BILFF.
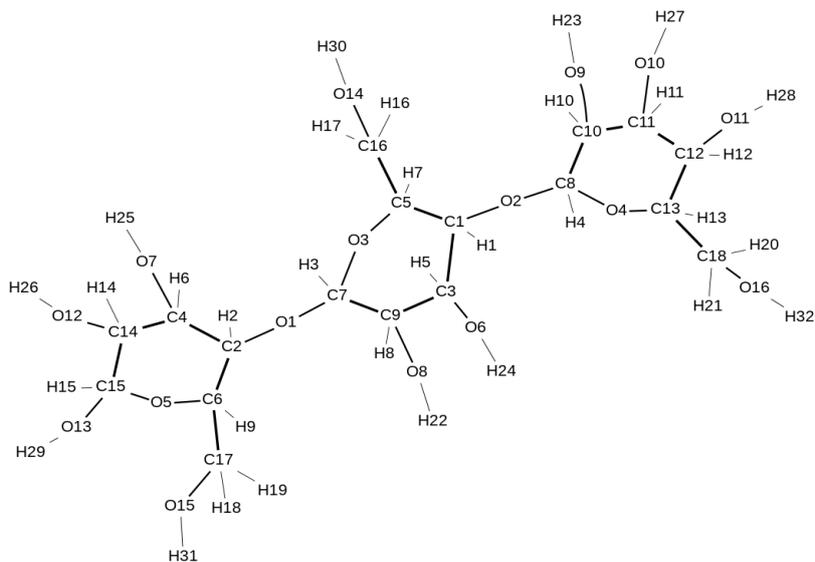
Figure S–7: Numbering of the protons and oxygen atoms of cellulose used in the contact matrices calculated from the AIMD simulations.
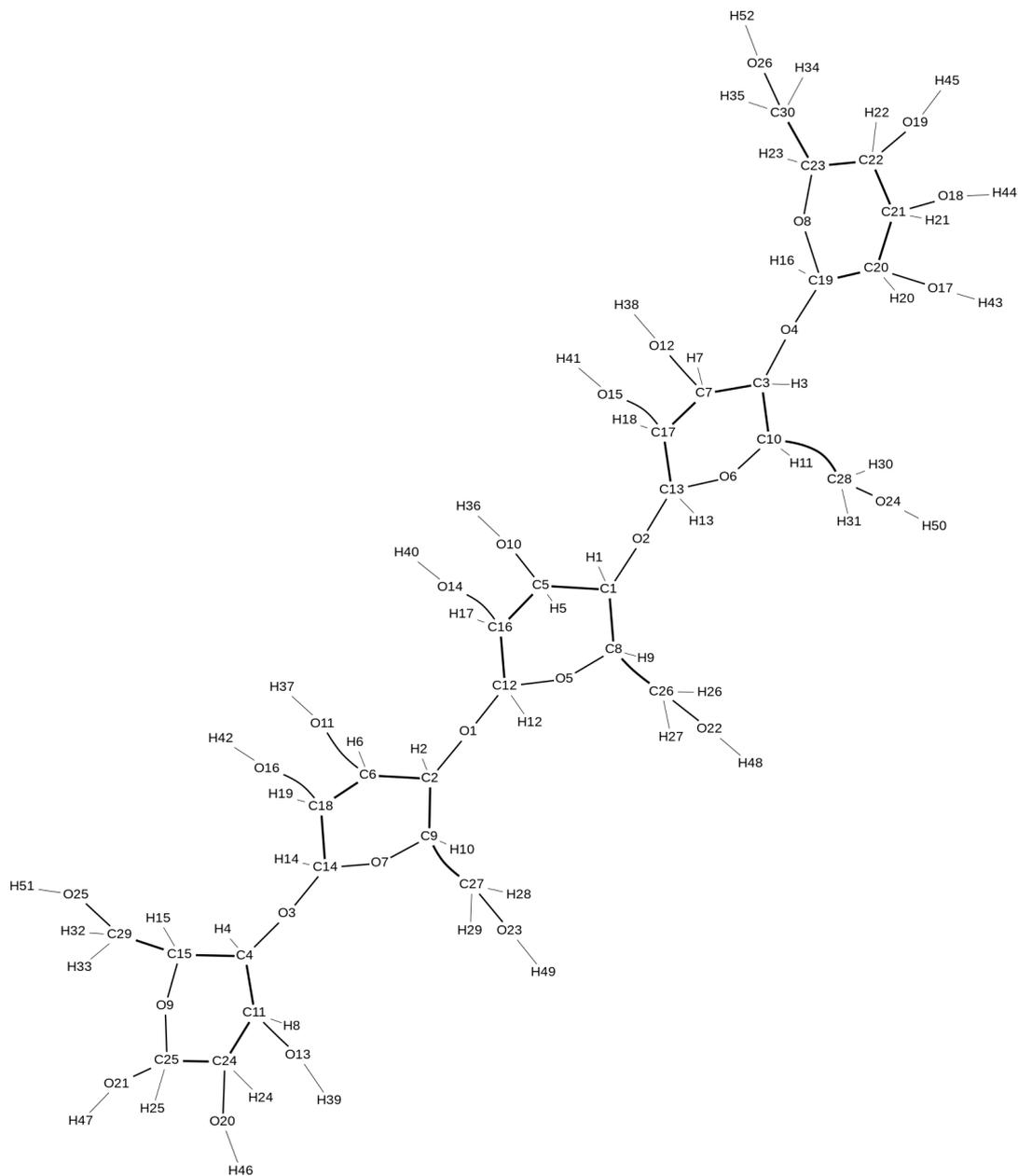
Figure S–8: Numbering of the protons and oxygen atoms of cellulose used in the contact matrices calculated from the force field MD simulations.
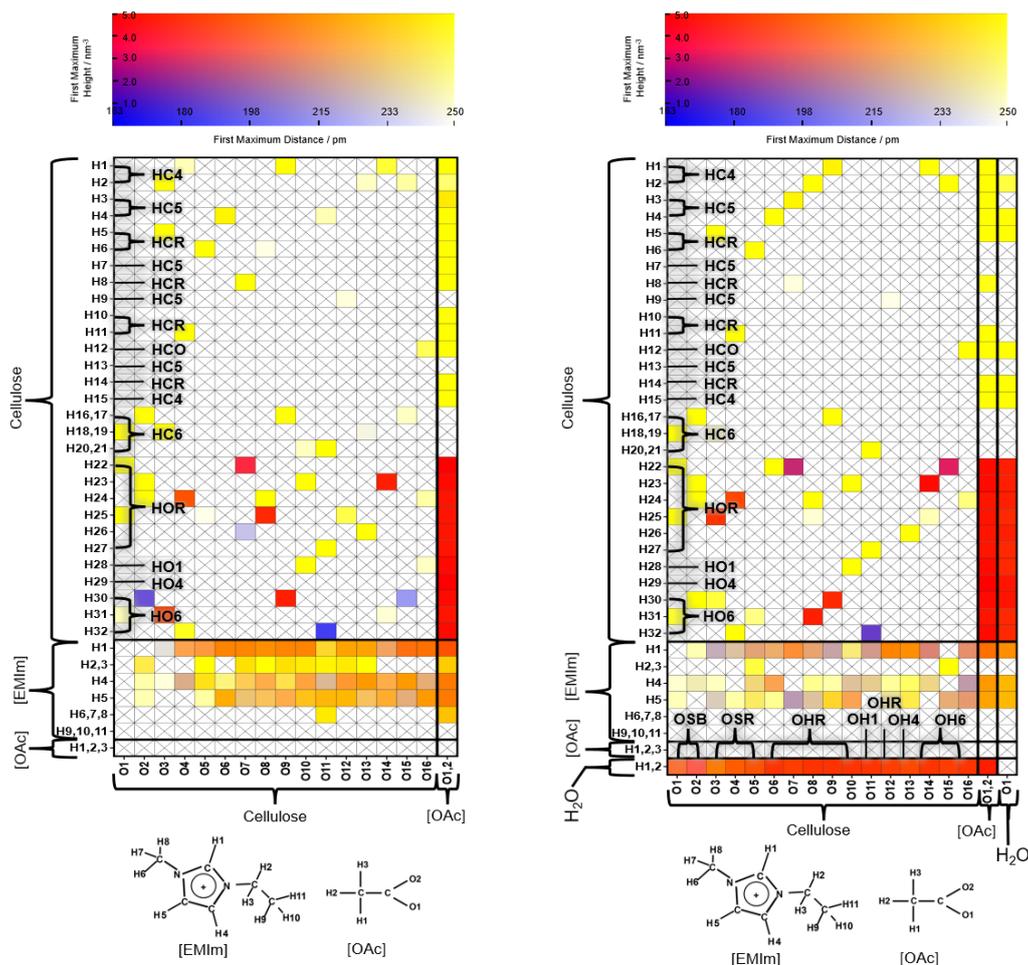
Figure S–9: Contact matrix illustrating hydrogen bonding in the cellulose–[EMIm][OAc] system in absence *(left)* and presence *(right)* of water from the reference AIMD simulation. For a better understanding, the atom types of cellulose are noted in the matrix and the atom numbers of the relevant atoms of the ionic liquid are shown below.

13

References

(1) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J. Am. Chem. Soc. **1996**, 118, 11225–11236.

(2) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. Adv. Protein Chem. **2003**, 66, 27–85.

(3) Sambasivarao, S. V.; Acevedo, O. Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids. J. Chem. Theory Comput. **2009**, 5, 1038–1050.

(4) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. J. Comp. Phys. **1995**, 117, 1–19.

(5) Roos, E.; Brehm, M. A force field for bio-polymers in ionic liquids (BILFF) - part 1: EMImOAc/water mixtures. Phys. Chem. Chem. Phys. **2021**, 23, 1242–1253.

(6) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J. Chem. Phys **2004**, 120, 9665–9678.

(7) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. OPLS all-atom force field for carbohydrates. J. Comput. Chem. **1997**, 18, 1955–1970.

A.3 Supporting Information to Section 3.4 - Article III

A Force Field for the Solubility of Cellulose in DMSO/Ionic Liquids

Supporting Information

Eliane Roos [1], Cora Gradaus [1], Daniel Sebastiani [1] and Martin Brehm [1,2,*]

[1] *Institut für Chemie – Theoretische Chemie, Martin-Luther-Universität Halle–Wittenberg, Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany.*

[2] *Department Chemie, Universität Paderborb, Warburger Straße 100, 33098 Paderborn, Germany.*

[*]E-mail: Martin_Brehm@gmx.de

[*]Website: https://brehm-research.de/

1 Force Field Parameter

In this section, the optimized force field parameters for DMSO in pure and aqueous [EMIm][OAc] are shown. The atom type nomenclature can be found in Figure S-1 and the atom classes in Table S-1. The force field is based on the potential equation of the OPLS–AA force field:[1–3]

$$
\begin{aligned}
U(r^N) = &\sum_{i\,\in\,\text{bonds}} k_{l,i}(l_i - l_{i,0})^2 + \sum_{i\,\in\,\text{angles}} k_{\theta,i}(\theta_i - \theta_{i,0})^2 \\
&+ \sum_{i\,\in\,\text{dihedrals}} \left[\frac{V_{i,1}}{2}[1 + \cos(\phi_i)] + \frac{V_{i,2}}{2}[1 - \cos(2\phi_i)] \right. \\
&\qquad\qquad \left. + \frac{V_{i,3}}{2}[1 + \cos(3\phi_i)] + \frac{V_{i,4}}{2}[1 - \cos(4\phi_i)] \right] \\
&+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}} \right] f_{ij}
\end{aligned}
\tag{1}
$$

$$
\epsilon_{i,j} = \sqrt{\epsilon_i\,\epsilon_j}, \qquad \sigma_{i,j} = \sqrt{\sigma_i\,\sigma_j}.
\tag{2}
$$

The non-bonded interactions are tabulated in Table S-2 and the bonded ones in Tables S-3, S-4, and S-5 in comparison to the literature force field OPLS–AA.[1–3] The force constants *(see Table S-3 and S-4)* and torsion parameters *(see Table S-5)* do not include the factor 1/2.

The non-bonded interactions between the atomic neighbours 1–2 and 1–3 were not considered. The 1–4 interactions were scaled with the factor $f_{ij} = 0.5$. The Lennard-Jones interactions were calculated with the geometric mixing rules *(see Eq. 2)*. For the Coulomb interactions, the PPPM long-range Coulomb solver (as implemented in LAMMPS)[4] was applied. The cutoff radius of the Coulomb and Lennard-Jones interactions was set to 800 pm.

For [EMIm][OAc], the force field BILFF[5] and for water TIP4P-EW[6] (with fixed bonds and angles) was applied without any modifications. The total charge of the ions was set to ±0.82.

Since no force field parameters for the protons of DMSO are defined in the literature force field OPLS–AA[1–3], the partial charges resulted from the charge neutrality of the molecule together with the defined charges for the other atoms O, S and C. For the Lennard-Jones parameters of the protons, the given values of a CH_2/CH_3 substitute in the proximity of electron-withdrawing atomic groups were adopted.
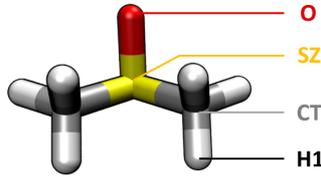
Figure S-1: Nomenclature of atom types in BILFF.

Table S-1: Nomenclature of atom types and atom classes of DMSO. The atom types are used for the non-bonded interactions *(see Table S-2)*. The atom classes are used for the bonded interactions *(see Tab. S-3, S-4, S-5).*

Atom Type	Atom Class
O	O
S	SZ
CT	CT
H1	H1

Table S-2: Optimized atomic partial charges q and Lennard-Jones parameters σ and ϵ of DMSO.

Atom Type	BILFF			Literature[1-3]		
	q	σ / Å	ϵ / kJ mol^{-1}	q	σ / Å	ϵ / kJ mol^{-1}
O	-0.474	2.93	1.172	-0.420	2.93	1.172
S	0.136	3.30	1.653	0.130	3.56	1.653
CT	-0.029	3.30	0.276	-0.035	3.50	0.276
H1	0.066	2.38	0.126	0.060	2.50	0.126

Table S-3: Optimized bond equilibrium lengths l_0 and force constants k_l of DMSO.

Bond	BILFF		Literature[1-3]	
	l_0 / Å	k_l / kJ mol^{-1} Å^{-2}	l_0 / Å	k_l / kJ mol^{-1} Å^{-2}
O–SZ	1.533	3664.3	1.530	2928.8
SZ–CT	1.843	1269.8	1.790	1422.6
CT–H1	1.100	2621.6	1.090	1422.6

Table S-4: Optimized angle equilibrium values θ_0 and force constants k_θ of DMSO.

Angle	BILFF		Literature[1-3]	
	θ_0 / Deg	k_θ / kJ mol^{-1} rad^{-2}	θ_0 / Deg	k_θ / kJ mol^{-1} rad^{-2}
SZ–CT–H1	109.5	333.3	109.5	146.4
H1–CT–H1	107.8	507.9	107.8	138.1
CT–SZ–CT	96.0	510.4	96.0	259.4
CT–SZ–O	107.0	562.9	107.0	309.6

Table S-5: Optimized torsional coefficients V_n of DMSO.

| Torsion Angle | BILFF | | | | Literature[1–3] | | | |
	V_1 $\mathrm{kJ\,mol^{-1}}$	V_2 $\mathrm{kJ\,mol^{-1}}$	V_3 $\mathrm{kJ\,mol^{-1}}$	V_4 $\mathrm{kJ\,mol^{-1}}$	V_1 $\mathrm{kJ\,mol^{-1}}$	V_2 $\mathrm{kJ\,mol^{-1}}$	V_3 $\mathrm{kJ\,mol^{-1}}$	V_4 $\mathrm{kJ\,mol^{-1}}$
CT–SZ–CT–H1	6.4434	−6.1924	5.7739	−3.1798	0.0000	0.0000	0.0000	0.0000
H1–CT–SZ–O	6.4434	−6.1924	5.7739	−3.1798	0.0000	0.0000	0.0000	0.0000

2 Comparison of the Partial Charges of DMSO With Literature Data

In Tab. S-6, a comparison of the atomic charges of DMSO in BILFF and various literature force fields is presented. These force fields employ different approaches to parameter optimization and utilize different calculation methods for potential energy.

BILFF is based on the OPLS–AA force field.[1–3] The atomic charges are represented in columns 2 and 3, respectively. The force fields from Ref. 7 and 8 are compatible with the functional form of AMBER.[9] On the other hand, the force fields from Ref. 10 and Ref. 11 are based on GROMOS.[12] These two force fields are classified as united-atom model force fields, while the other ones are all-atom force fields. Therefore, a row for the methyl group has been added to the atom types in the table. Since the charges in BILFF are derived from the OPLS values[1–3] (columns 2 and 3), the values are quite similar.

When comparing the charges of oxygen and sulfur, the values in BILFF exhibit also a similarity to the charges in the united atom force fields of Geerke et. al.[10] and Bordat et. al.[11] (columns 6 and 7), while the force fields based on AMBER[9] show higher partial charges (column 4 and 5).

It is important to note that these force fields were developed with a focus on different properties. The force fields by Geerke et. al.,[10] Bordat et. al.,[11] and Fox et. al.,[7] for example, aimed to accurately describe thermodynamic properties. The first one was optimized specifically for aqueous DMSO, while the other two were optimized for pure DMSO. On the other hand, the force field of Strader et. al.[8] aimed to reproduce both the structural and thermodynamic properties of aqueous DMSO. In the case of BILFF, the focus was on reproducing the hydrogen bonding behavior of DMSO in the ionic liquid [EMIm][OAc] in the presence and absence of water. Therefore, it is not surprising that the atomic charges in BILFF differ from those in the other all-atom force fields, which are specialized in capturing thermodynamic properties.

Table S-6: Comparison of the partial charges of DMSO in BILFF and different literature force fields.

Atom Type	BILFF	OPLS–AA[1–3]	Fox et. al.[7]	Strader et. al.[8]	Geerke et. al.[10]	Bordat et. al.[11]
O	−0.474 00	−0.420 00	−0.520 50	−0.556 00	−0.447 53	−0.436 74
S	0.136 00	0.130 00	0.315 50	0.312 00	0.127 53	0.116 74
CT	−0.029 00	−0.035 00	−0.324 40	−0.148 00	–	–
H1	0.066 00	0.060 00	0.142 30	0.090 00	–	–
CH$_3$	–	–	–	–	0.160 00	0.160 00

3 Additional Distribution Functions
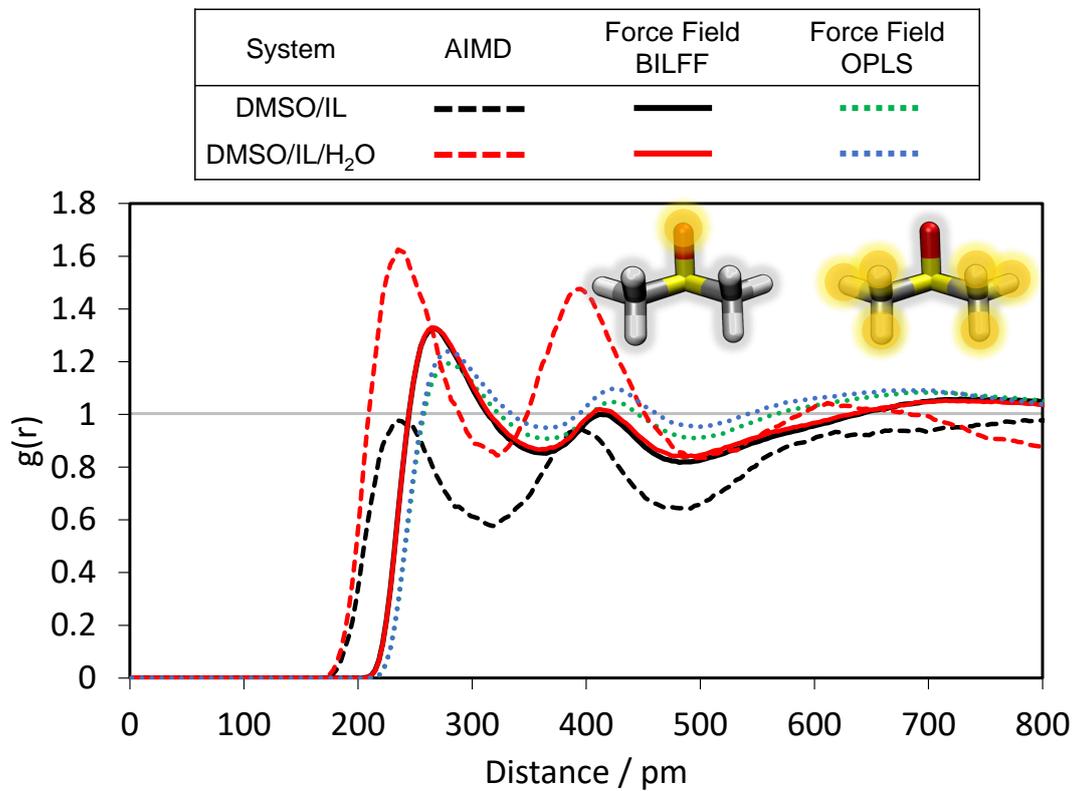
Figure S-2: Comparison of the RDFs between the reference AIMD and a force field MD simulation with the OPLS–AA force field [1–3] and BILFF between the oxygen atom and protons of DMSO.
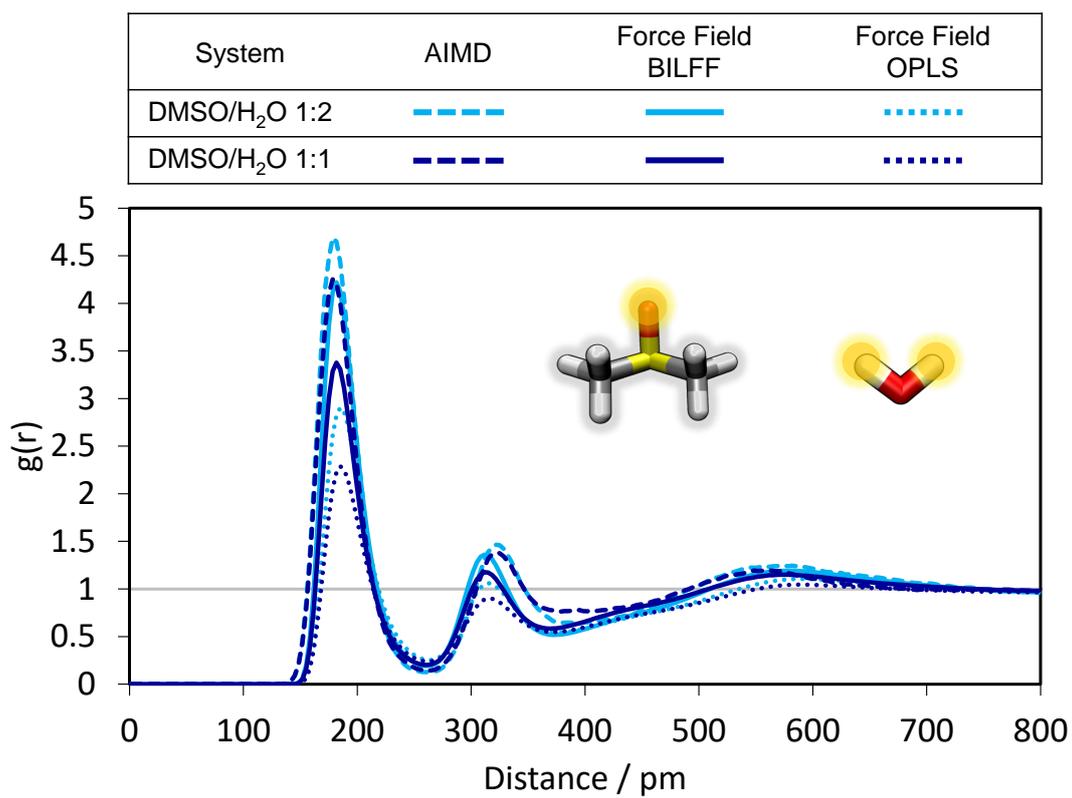
Figure S-3: Comparison of the RDFs between the DMSO oxygen and the water protons at two different DMSO/water mixing ratios (particle number ratio 1:2 and 1:1) calculated from the reference AIMD and force field MD simulations with the OPLS–AA force field [1–3] and BILFF.
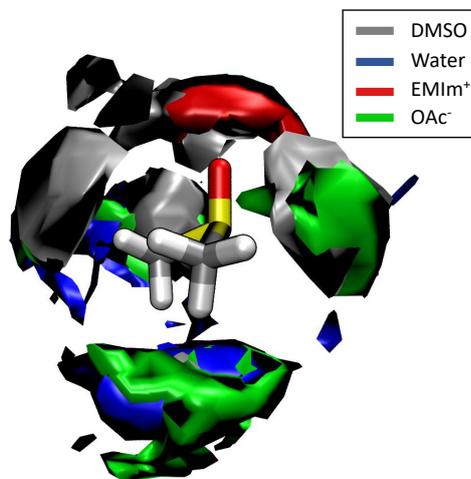
Figure S-4: Spatial distribution function (SDF) of the arrangement of molecules around DMSO with protons and oxygen atoms of DMSO (gray, oxygen: $6\,\mathrm{nm}^{-3}$; proton: $4\,\mathrm{nm}^{-3}$) and water (blue, $5\,\mathrm{nm}^{-3}$), the ring protons of $[\mathrm{EMIm}]^+$ (red, $7\,\mathrm{nm}^{-3}$) and the oxygen atom of acetate (green, $7\,\mathrm{nm}^{-3}$) in the DMSO–IL–$\mathrm{H_2O}$ system calculated from the reference AIMD simulation.
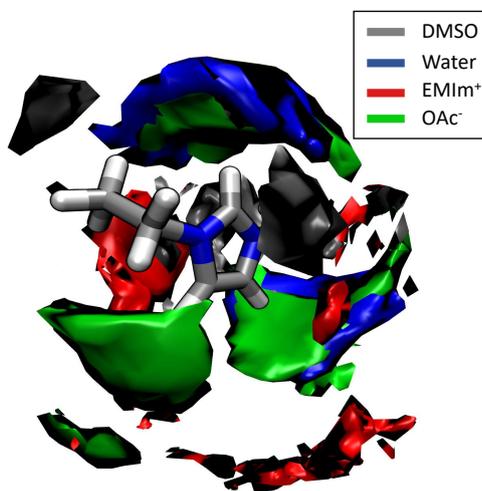
Figure S-5: Spatial distribution function (SDF) of the arrangement of molecules around $[\mathrm{EMIm}]^+$ with the protons and oxygen atoms of DMSO (gray, oxygen: $5\,\mathrm{nm}^{-3}$; proton: $3\,\mathrm{nm}^{-3}$) and water (blue, $6\,\mathrm{nm}^{-3}$), the ring center of $[\mathrm{EMIm}]^+$ (red, $9\,\mathrm{nm}^{-3}$) and the oxygen atom of acetate (green, $10\,\mathrm{nm}^{-3}$) in the DMSO–IL–$\mathrm{H_2O}$ system calculated from the reference AIMD simulation.

4 Lifetime of Additional Hydrogen Bonds

The calculation of the hydrogen bond lifetimes from the main text was based on specified angle and distance criteria. These were determined on the basis of occurring maxima in underlying combined distance–angle distribution functions and are listed in Tab. S-8. The same calculation criteria were used for all simulations.

Table S-7: Angle and distance criteria of the different hydrogen bonds for the calculation of the hydrogen bond lifetime.

Hydrogen Bond	Atom Distance	Angle
(DMSO)H–O(DMSO)	0–310 pm	90°–180°
([EMIm])H–O(DMSO)	0–280 pm	110°–180°
([EMIm])H–O([OAc])	0–260 pm	112°–180°
([EMIm])H–O(H_2O)	0–260 pm	120°–180°
(H_2O)H–O(DMSO)	0–250 pm	145°–180°
(H_2O)H–O([OAc])	0–180 pm	165°–180°
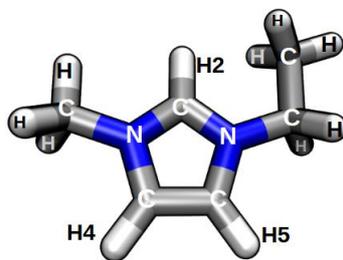
Figure S-6: Nomenclature of [EMIm]$^+$ in BILFF[5].

Table S-8: Overview of the lifetime τ of hydrogen bonds in the AIMD and the force field MD using BILFF at different temperatures. Lifetimes in ps. Please note that large lifetimes (> 100 ps) for the AIMD simulation are not reliable.

Temp.	Intermittent		Continuous	
	τ(AIMD)	τ(FFMD)	τ(AIMD)	τ(FFMD)
DMSO/[EMIm][OAc]				
([EMIm])H2–O([OAc])				
350 K	772.3	380.17	7.79	8.24
450 K	79.44	63.89	2.57	3.19
([EMIm])H4–O([OAc])				
350 K	230.11	–	1.80	2.59
450 K	34.16	–	0.98	1.36
([EMIm])H5–O([OAc])				
350 K	173.34	–	1.64	2.38
450 K	41.01	–	1.14	1.28
([EMIm])H2–O(DMSO)				
350 K	33.93	47.06	1.14	3.11
450 K	10.50	10.34	1.08	1.57
([EMIm])H4–O(DMSO)				
350 K	42.38	54.44	2.05	2.07
450 K	–	9.80	0.96	1.13
([EMIm])H5–O(DMSO)				
350 K	71.92	62.16	2.30	2.30
450 K	456.52	10.81	0.79	1.21
DMSO/[EMIm][OAc]/H_2O				
([EMIm])H2–O([OAc])				
350 K	185.05	307.861	5.03	6.56
450 K	59.38	52.59	2.03	2.79
([EMIm])H4–O([OAc])				
350 K	212.65	–	1.89	2.23
450 K	–	25.68	1.125	1.22
([EMIm])H5–O([OAc])				
350 K	114.44	–	1.47	2.07
450 K	–	24.71	0.78	1.16
([EMIm])H2–O(DMSO)				
350 K	44.64	48.69	2.67	3.05
450 K	4.77	9.41	0.59	1.52
([EMIm])H4–O(DMSO)				
350 K	46.80	53.04	1.78	1.97
450 K	–	8.90	0.80	1.10
([EMIm])H5–O(DMSO)				
350 K	635.37	57.24	2.01	2.20
450 K	18.10	10.11	1.14	1.18
([EMIm])H2–O(H_2O)				
350 K	67.34	80.96	0.95	1.14
450 K	21.30	10.69	0.53	0.71
([EMIm])H4–O(H_2O)				
350 K	26.75	54.77	0.55	0.77
450 K	16.43	6.50	0.35	0.50
([EMIm])H5–O(H_2O)				
350 K	11.89	54.37	0.57	0.84
450 K	3.81	6.52	0.38	0.52

References

[1] W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, J. Am. Chem. Soc., 1996, **118**, 11225–11236.

[2] J. W. Ponder and D. A. Case, Adv. Protein Chem., 2003, **66**, 27–85.

[3] S. V. Sambasivarao and O. Acevedo, J. Chem. Theory Comput., 2009, **5**, 1038–1050.

[4] S. Plimpton, J. Comp. Phys., 1995, **117**, 1–19.

[5] E. Roos and M. Brehm, PCCP, 2021, **23**, 1242–1253.

[6] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, J. Chem. Phys., 2004, **120**, 9665–9678.

[7] T. Fox and P. A. Kollman, J. Phys. Chem. B, 1998, **102**, 8070–8079.

[8] M. L. Strader and S. E. Feller, J. Phys. Chem. A, 2002, **106**, 1074–1080.

[9] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, J. Am. Chem. Soc., 1995, **117**, 5179–5197.

[10] D. P. Geerke, C. Oostenbrink, N. F. A. van der Vegt and W. F. van Gunsteren, J. Phys. Chem. B, 2004, **108**, 1436–1445.

[11] P. Bordat, J. Sacristan, D. Reith, S. Girard, A. Glättli and F. Müller-Plathe, Chem. Phys. Lett., 2003, **374**, 201–205.

[12] H. Liu, F. Mueller-Plathe and W. F. van Gunsteren, J. Am. Chem. Soc., 1995, **117**, 4363–4366.

A.4   Supporting Information to Section 3.5 - Article IV

BILFF: All-Atom Force Field for Modeling Triazolium- and Benzoate-Based Ionic Liquids

Supporting Information

Eliane Roos [1], Daniel Sebastiani [1] and Martin Brehm [1,2], *

1 *Institut für Chemie—Theoretische Chemie, Martin-Luther-Universität Halle–Wittenberg,*

Von-Danckelmann-Platz 4, 06120 Halle (Saale), Germany; eliane.roos@gmx.de (E.R.);

daniel.sebastiani@chemie.uni-halle.de (D.S.)

2 *Department Chemie, Universität Paderborn, Warburger Straße 100, 33098*

Paderborn, Germany

* Correspondence: martin_brehm@gmx.de

1 Force Field Parameters

This section presents the optimized force field parameters for [EMTr]$^+$ and [OBz]$^-$ in anhydrous and aqueous [EMTr][OAc], [EMIm][OBz], and [EMTr][OBz], comparing them to the adapted literature force field.[1–3]

Table S1 presents the non-bonded interactions while the bonded interactions are shown in Table S2-S4. The force constants *(see Tab. S2 and S3)* and torsion parameters *(see Tab. S4)* do not include the factor 1/2. The 1–2 and 1–3 interactions were not taken into account while the 1–4 interactions were scaled with $f_{ij} = 0.5$. The Lennard–Jones interactions were calculated using geometric mixing rules. The PPPM long-range Coulomb solver implemented in LAMMPS[4] was used for Coulomb interactions. The coulomb and Lennard-Jones interactions were truncated at a cutoff radius of 800 pm. The force field BILFF[5] was applied for [EMIm]$^+$ and [OAc]$^-$ while water, TIP4P–EW[6] (with fixed bonds and angles) was used without any modifications. The ions were assigned a total charge of ± 0.82.

To optimize the force field parameters for [EMTr]$^+$, we utilized BILFF[5,7] as a starting point. The Lennard-Jones parameters and force constants for [OBz]$^-$ were adapted from OPLS–AA,[1–3] while the partial charges were obtained using the restrained electrostatic potential (RESP) methodology based on quantum chemical calculations. Furthermore, the reference bond length was determined by geometry optimization. For [EMTr]$^+$ the Lennard-Jones parameter were adapted of [EMIm]$^+$ from BILFF,[5] while the reference bond length and angle of atom type NR was determined by a geometry optimization. The used starting force field parameters are listed in Tab. S–2-5 in comparison to the new optimized parameters.

Table S1: Comparison of the atomic partial charges q and Lennard-Jones parameter σ and ϵ of [EMTr]$^+$ and [OBz]$^-$ in BILFF and the adapted literature force field.[1–3,5,8–10]

Atom Type	BILFF			Literature (adapted)[1–3,5,8–10]		
	q	σ	ϵ	q	σ	ϵ
	$/\,e$	$/\,\text{Å}$	$/\,\text{kJ mol}^{-1}$	$/\,e$	$/\,\text{Å}$	$/\,\text{kJ mol}^{-1}$
			[EMTr]$^+$			
C1	-0.187	3.34	0.276	-0.147	3.34	0.276
CE	-0.054	3.34	0.276	-0.042	3.34	0.276
CW	-0.144	3.38	0.293	-0.133	3.38	0.293
HCW	0.191	1.48	0.126	0.150	1.54	0.126
HC	0.070	2.38	0.126	0.055	2.38	0.126
H1	0.148	2.38	0.126	0.116	2.38	0.126
NR	-0.204	3.10	0.711	-0.160	3.10	0.711
NA	0.204	3.10	0.711	0.160	3.10	0.711
			[OBz]$^-$			
C1	0.005	3.70	0.2929	0.006	3.55	0.2929
C2	-0.118	3.70	0.2929	-0.142	3.55	0.2929
C3	-0.121	3.70	0.2929	-0.025	3.55	0.2929
C4	-0.299	3.70	0.2929	-0.241	3.55	0.2929
CO	0.398	3.90	0.4393	0.371	3.75	0.4393
H2	0.070	2.42	0.1255	0.084	2.42	0.1255
H3	0.157	2.42	0.1255	0.069	2.42	0.1255
H4	0.200	2.42	0.1255	0.120	2.42	0.1255
O2	-0.550	2.80	0.8786	-0.524	2.96	0.8786

Table S2: Comparison of the bond equilibrium lengths l_0 and force constants k_1 of $[EMTr]^+$ and $[OBz]^-$ in BILFF and the adapted literature force field.[1–3,5,8–10]

Bond	BILFF		Literature (adapted)[1–3,5,8–10]	
	l_0	k_1	l_0	k_1
	/ Å	/ kJ mol^{-1} Å^{-2}	/ Å	/ kJ mol^{-1} Å^{-2}
		$[EMTr]^+$		
NA–NR	1.344	3199.2	1.340	3992.0
CW–HA	1.088	2633.8	1.085	2943.0
CW–NA	1.375	3108.7	1.403	2775.0
CW–CW	1.386	3773.2	1.374	4019.0
NA–CT	1.488	2046.3	1.485	2078.0
HC–CT	1.099	2679.4	1.099	3013.0
CT–CT	1.533	2125.5	1.532	2097.0
		$[OBz]^-$		
CA–CA	1.387	3274.1	1.400	1962.3
CA–HA	1.088	2707.4	1.084	1535.5
CA–CO	1.504	1906.9	1.529	1673.6
CO–O2	1.282	4273.1	1.252	2744.7

Table S3: Comparison of the angle equilibrium values θ_0 and force constants k_θ of $[EMTr]^+$ and $[OBz]^-$ in BILFF and the adapted literature force field.[1–3,5,8–10]

Angle	BILFF		Literature (adapted)[1–3,5,8–10]	
	θ_0	k_θ	θ_0	k_θ
	/ Deg	/ kJ mol^{-1} rad^{-2}	/ Deg	/ kJ mol^{-1} rad^{-2}
		$[EMTr]^+$		
CW–NA–NR	112.1	568.7	112.1	585.8
NR–NA–CT	118.6	396.5	118.6	585.8
NA–NR–NA	104.4	610.1	104.4	585.8
NA–CT–CT	110.9	361.2	110.9	532.7
NA–CW–CW	107.0	579.7	107.0	805.7
NA–CW–HA	120.8	200.9	120.8	263.0
CW–CW–HA	131.7	190.5	131.7	236.7
NA–CT–HC	107.2	375.9	107.2	411.7
CT–CT–HC	111.4	296.2	111.4	338.6
HC–CT–HC	109.2	226.5	109.2	308.0
CW–NA–CT	125.2	242.9	125.2	367.8
		$[OBz]^-$		
CA–CA–CA	120.0	446.0	120.0	263.6
CA–CA–HA	120.0	258.1	120.0	146.4
CA–CA–CO	120.0	397.6	120.0	355.6
CA–CO–O2	117.0	550.2	117.0	292.9
O2–CO–O2	126.0	735.9	126.0	334.7

Table S4: Comparison of the torsional coefficients V_n of [EMTr]$^+$ and [OBz]$^-$ in BILFF and the adapted literature force field. [1–3,5,8–10]

Torsion Angle	BILFF				Literature (adapted) [1–3,5,8–10]			
	V_1 / kJ mol^{-1}	V_2 / kJ mol^{-1}	V_3 / kJ mol^{-1}	V_4 / kJ mol^{-1}	V_1 / kJ mol^{-1}	V_2 / kJ mol^{-1}	V_3 / kJ mol^{-1}	V_4 / kJ mol^{-1}
				[EMTr]$^+$				
CW–NA–NR–NA	0.0000	19.4600	0.0000	0.0000	0.0000	19.4600	0.0000	0.0000
CT–NA–NR–NA	0.0000	19.4600	0.0000	0.0000	0.0000	19.4600	0.0000	0.0000
NR–NA–CW–CW	0.0000	12.5500	0.0000	0.0000	0.0000	12.5500	0.0000	0.0000
NR–NA–CW–HA	0.0000	12.5500	0.0000	0.0000	0.0000	12.5500	0.0000	0.0000
NR–NA–CT–HC	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
NR–NA–CT–CT	0.1000	1.0000	0.1000	−0.3000	−5.2691	0.0000	0.0000	0.0000
CT–NA–CW–CW	0.0000	12.5500	0.0000	0.0000	0.0000	12.5500	0.0000	0.0000
CT–NA–CW–HA	0.0000	12.5500	0.0000	0.0000	0.0000	12.5500	0.0000	0.0000
NA–CW–CW–NA	0.0000	65.0000	0.0000	0.0000	0.0000	65.0000	0.0000	0.0000
NA–CW–CW–HA	0.0000	44.9800	0.0000	0.0000	0.0000	44.9800	0.0000	0.0000
HA–CW–CW–HA	0.0000	30.0000	0.0000	0.0000	0.0000	30.0000	0.0000	0.0000
CW–NA–CT–HC	0.1000	0.2000	0.0000	0.0000	0.0000	0.0000	0.5190	0.0000
CW–NA–CT–CT	0.4000	1.0000	0.0000	0.2000	14.3000	−12.2000	−1.5900	0.0000
NA–CT–CT–HC	0.0000	0.0000	0.3670	0.0000	0.0000	0.0000	0.3670	0.0000
HC–CT–CT–HC	0.0000	0.0000	1.2552	0.0000	0.0000	0.0000	1.2552	0.0000
				[OBz]$^-$				
CA–CA–CA–CA	0.0000	30.334	0.0000	0.0000	0.0000	30.334	0.0000	0.0000
HA–CA–CA–CA	0.0000	30.334	0.0000	0.0000	0.0000	30.334	0.0000	0.0000
HA–CA–CA–CO	0.0000	30.334	0.0000	0.0000	0.0000	30.334	0.0000	0.0000
HA–CA–CA–HA	0.0000	30.334	0.0000	0.0000	0.0000	30.334	0.0000	0.0000
CA–CA–CA–CO	0.0000	30.334	0.0000	0.0000	0.0000	30.334	0.0000	0.0000
CA–CA–CO–O2	0.0000	8.000	0.0000	0.0000	0.0000	8.000	0.0000	0.0000

2 Additional Radial Distribution Functions

To further characterize the force field and describe the underlying molecular systems, additional radial distribution functions (RDFs) are shown below.
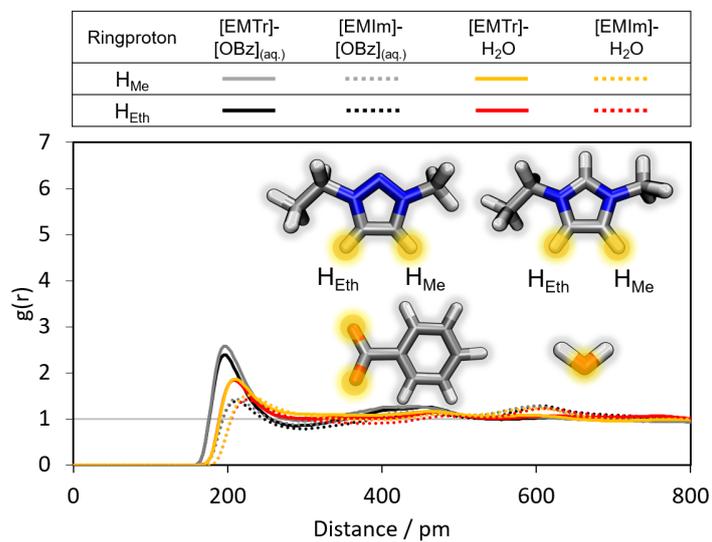
Figure S1: Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of $[OBz]^-$/water and the ring protons of $[EMIm]^+$/$[EMTr]^+$ calculated from force field MD simulations of [EMTr][OBz] and [EMIm][OBz] using BILFF.
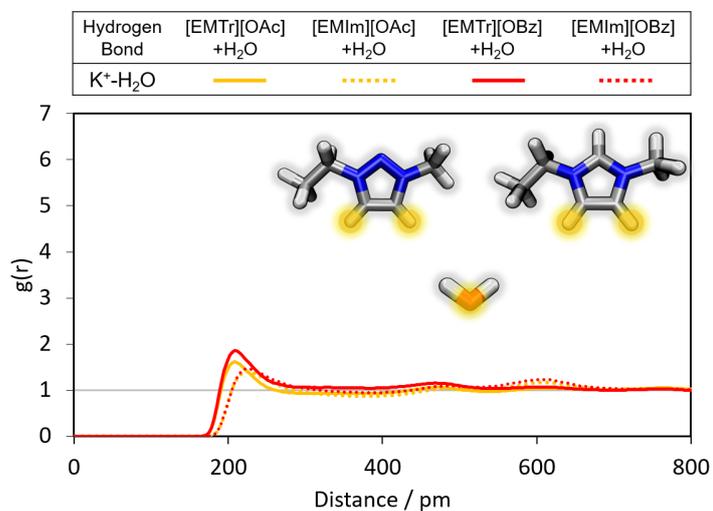
Figure S2: Comparison of the RDFs of the hydrogen bond between the marked oxygen atom of water and the ring protons of [EMTr]$^+$ and [EMIm]$^+$ calculated from force field MD simulations using BILFF in all four ILs.
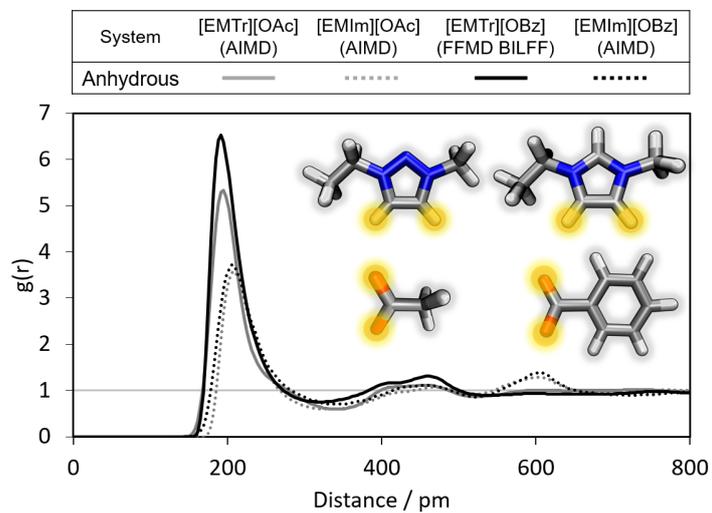
Figure S3: Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OBz]$^-$ as well as [OAc]$^-$ and the ring protons of [EMIm]$^+$ and [EMTr]$^+$ calulated from the reference AIMDs. For a comparison of the results with anhydrous [EMTr][OBz] the results from the force field MD simulation (FFMD) using BILFF are shown. The RDFs are averaged over the marked ring protons.
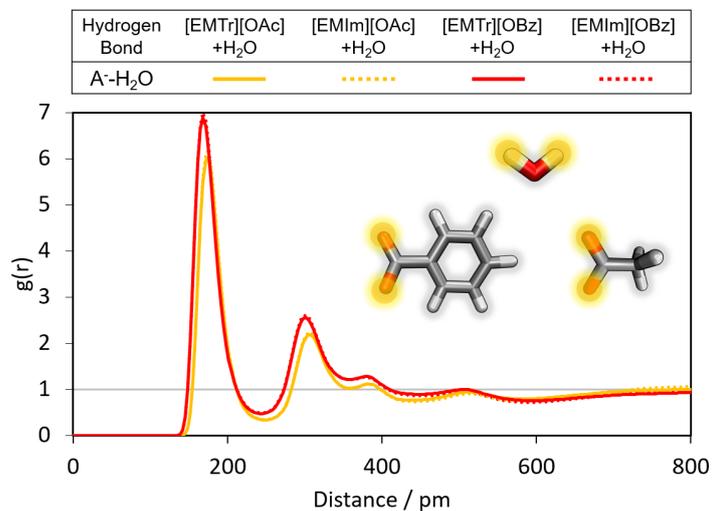
Figure S4: Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of $[OBz]^-/[OAc]^-$ and the protons of water calculated from force field MD simulations using BILFF in all four ILs.
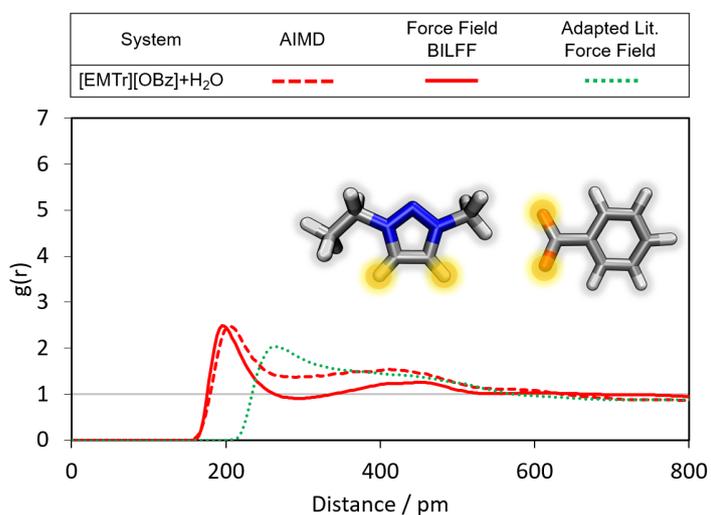
Figure S5: Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of $[OBz]^-$ and the ring protons of $[EMTr]^+$ calculated from a reference AIMD and force field MD simulations using adapted literature force field parameter [8–10] and BILFF. The RDFs are averaged over both ring protons.
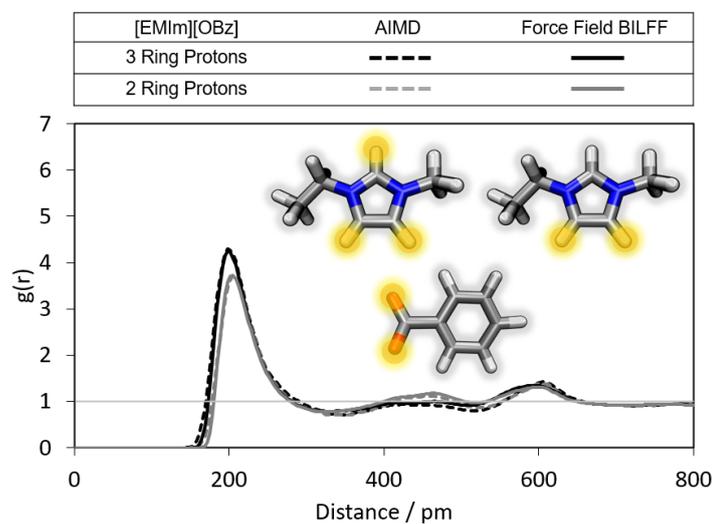
Figure S6: Comparison of the RDFs of the hydrogen bond between the marked oxygen atoms of [OBz]$^-$ and the two/three ring protons of [EMIm]$^+$ calculated from a reference AIMD and force field MD simulations using BILFF. The RDFs are averaged over the marked ring protons.

3 Additional Distance–Angle Combined Distribution Functions
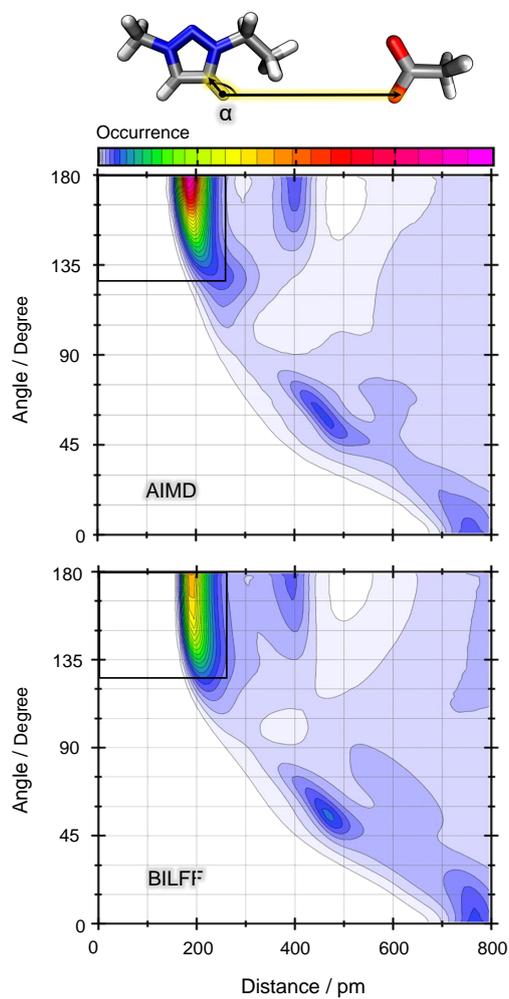
Figure S7: Distance–angle distribution function between an example ring proton of [EMTr]$^+$ and the oxygen atoms of [OAc]$^-$ in pure [EMTr][OAc] as a result of a reference AIMD (top) and a force field MD simulations with our new force field (bottom). The black rectangle demonstrates the geometric criterion for calculating the lifetime of the hydrogen bonds.
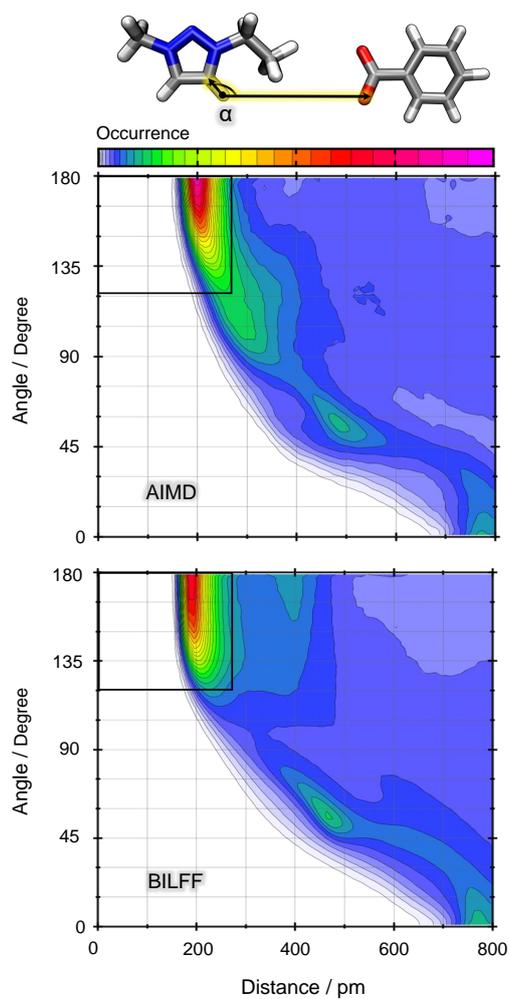
Figure S8: Distance–angle distribution function between an example ring proton of [EMTr]$^+$ and the oxygen atoms of [OBz]$^-$ in aqueous [EMTr][OBz] as a result of a reference AIMD (top) and a force field MD simulations with our new force field (bottom). The black rectangle demonstrates the geometric criterion for calculating the lifetime of the hydrogen bonds.

4 Additional Spatial Distribution Function
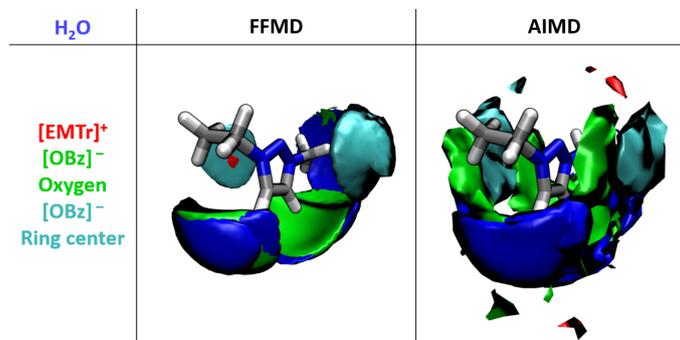
Figure S9: Spatial distribution function of the arrangement of molecules in aqueous [EMTr][OBz] around the cation with the protons and oxygen atoms of water (blue, $26\,\mathrm{nm}^{-3}$) and the oxygen atom of the anion (green, $7\,\mathrm{nm}^{-3}$) as well as the ring center of the cation (red, $7\,\mathrm{nm}^{-3}$) and anion (cyan, $10\,\mathrm{nm}^{-3}$) resulted from a force field MD simulation using BILFF and the reference AIMD simulation.

5 Sankey Diagrams
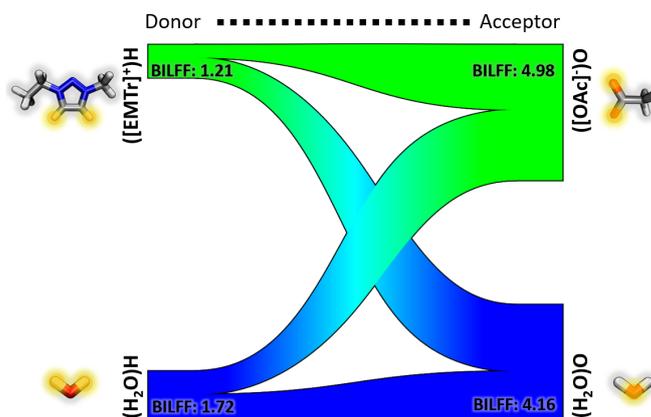
Figure S10: Sankey diagram of aqueous [EMTr][OAc] calculated from a force field MD simulation using BILFF.
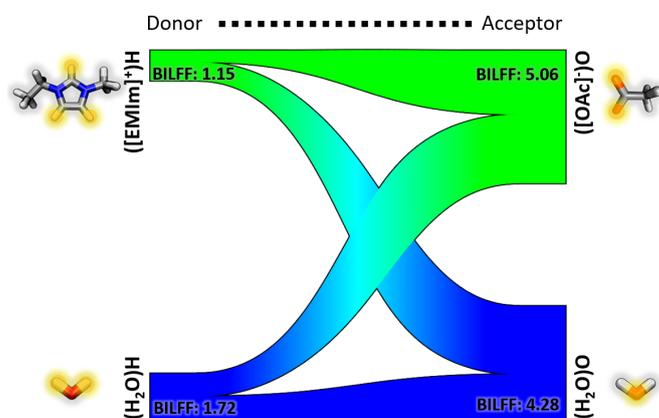
Figure S11: Sankey diagram of aqueous [EMIm][OAc] calculated from a force field MD simulation using BILFF.
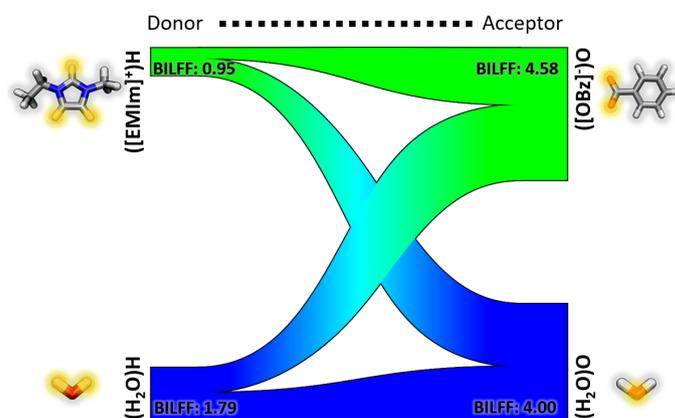
Figure S12: Sankey diagram of aqueous [EMIm][OBz] calculated from a force field MD simulation using BILFF.

6 Lifetime of Additional Hydrogen Bonds

Table S5: Angle and distance criteria of the different hydrogen bonds for the calculation of the hydrogen bond lifetime.

Hydrogen Bond	Atom Distance / pm	Angle / °
([EMTr])H–O([OAc])	0–260	125–180
([EMIm])H–O([OBz])	0–250	125–180
([EMTr])H–O([OBz])	0–270	122–180
([EMTr])H–O(H_2O)	0–260	118–180
(H_2O)H–O([OBz])	0–230	137–180
([EMIm])H–O([OAc])	0–260	112–180
([EMIm])H–O(H_2O)	0–260	120–180
(H_2O)H–O([OAc])	0–180	165–180

Table S6: Overview of the lifetime τ of the hydrogen bonds in all four systems comparing the results of the reference AIMD simulation and the force field MD simulation using BILFF at the given temperatures (C^+=Cation, A^-=Anion). (No AIMD simulations of anhydrous [EMTr][OBz] as well as [EMTr][OAc]/H_2O and [EMIm][OBz]/H_2O have been calculated, so no data are available for these.)

Temp.	Intermittent		Continuous	
	τ(AIMD)	τ(FFMD)	τ(AIMD)	τ(FFMD)
/ K	/ ps	/ ps	/ ps	/ ps
	[EMTr][OAc]			
$(C^+)H_{Eth}\cdots O(A^-)$				
350 K	513.8	556.6	4.1	4.3
$(C^+)H_{Me}\cdots O(A^-)$				
350 K	581.6	394.5	3.8	3.9
	[EMTr][OAc]/H_2O			
$(C^+)H_{Eth}\cdots O(A^-)$				
350 K	–	74.9	–	1.9
$(C^+)H_{Me}\cdots O(A^-)$				
350 K	–	81.9	–	1.8
$(C^+)H_{Eth}\cdots O(H_2O)$				
350 K	–	24.9	–	1.1
$(C^+)H_{Me}\cdots O(H_2O)$				
350 K	–	23.9	–	1.0
	[EMIm][OAc] [a]			
$(C^+)H_{Eth}\cdots O(A^-)$				
350 K	212.6	250.3	10.5	10.0
$(C^+)H_{Me}\cdots O(A^-)$				
350 K	282.4	260.7	11.9	11.5
	[EMIm][OBz]			
$(C^+)H_{Eth}\cdots O(A^-)$				
350 K	309.6	750.1	1.1	1.5
$(C^+)H_{Me}\cdots O(A^-)$				
350 K	161.7	657.3	1.0	1.4

13

[EMIm][OBz]/H$_2$O				
(C$^+$)H$_{Eth}$···O(A$^-$)				
350 K	–	71.8	–	0.9
(C$^+$)H$_{Me}$···O(A$^-$)				
350 K	–	69.1	–	0.8
(C$^+$)H$_{Eth}$···O(H$_2$O)				
350 K	–	25.2	–	0.6
(C$^+$)H$_{Me}$···O(H$_2$O)				
350 K	–	24.8	–	0.6
[EMTr][OBz]				
(C$^+$)H$_{Eth}$···O(A$^-$)				
350 K	–	1712.7	–	6.7
(C$^+$)H$_{Me}$···O(A$^-$)				
350 K	–	2014.9	–	5.8
[EMTr][OBz]/H$_2$O				
(C$^+$)H$_{Eth}$···O(A$^-$)				
350 K	80.7	128.2	1.8	2.5
450 K	25.8	18.3	0.9	1.4
550 K	10.9	7.6	0.6	1.0
(C$^+$)H$_{Me}$···O(A$^-$)				
350 K	83.8	134.5	1.4	2.4
450 K	17.9	9.2	0.7	1.3
550 K	10.8	7.8	0.6	1.0
(C$^+$)H$_{Eth}$···O(H$_2$O)				
350 K	38.9	33.0	0.7	1.1
450 K	4.4	4.6	0.4	0.7
550 K	–	1.7	0.3	0.5
(C$^+$)H$_{Me}$···O(H$_2$O)				
350 K	22.7	33.0	0.6	1.0
450 K	4.7	4.6	0.3	0.6
550 K	–	1.8	0.3	0.5

[a] *Calculated from the MD simulations of our already published article.* [5]

References

[1] W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, J. Am. Chem. Soc., 1996, **118**, 11225–11236.

[2] J. W. Ponder and D. A. Case, Adv. Protein Chem., 2003, **66**, 27–85.

[3] S. V. Sambasivarao and O. Acevedo, J. Chem. Theory Comput., 2009, **5**, 1038–1050.

[4] S. Plimpton, J. Comp. Phys., 1995, **117**, 1–19.

[5] E. Roos and M. Brehm, PCCP, 2021, **23**, 1242–1253.

[6] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, J. Chem. Phys., 2004, **120**, 9665–9678.

[7] E. Roos, D. Sebastiani and M. Brehm, PCCP, 2023, **25**, 8755–8766.

[8] J. N. Canongia Lopes, J. Deschamps and A. A. H. Pádua, J. Phys. Chem. B, 2004, **108**, 2038–2047.

[9] J. N. Canongia Lopes and A. A. H. Pádua, J. Phys. Chem. B, 2006, **110**, 19586–19592.

[10] J. N. Canongia Lopes and A. A. H. Pádua, Theor. Chem. Acc., 2012, **131**, 3330.

Bibliography

[1]   C. M. Payne, B. C. Knott, H. B. Mayes, H. Hansson, M. E. Himmel, M. Sand-
        gren, J. Ståhlberg and G. T. Beckham, *Chem. Rev.*, 2015, **115**, 1308–1448.

[2]   S. Nanda, B. R. Patra, R. Patel, J. Bakos and A. K. Dalai, *Environ. Chem. Lett.*,
        2022, **20**, 379–395.

[3]   D. Klemm, B. Heublein, H.-P. Fink and A. Bohn, *Angew. Chem.*, 2005, **44**, 3358–
        3393.

[4]   B. Azimi, H. Maleki, V. Gigante, R. Bagherzadeh, A. Mezzetta, M. Milazzo,
        L. Guazzelli, P. Cinelli, A. Lazzeri and S. Danti, *Cellulose*, 2022, **29**, 3079–3129.

[5]   C. Olsson and G. Westman, in *Cellulose- Fundamental Aspects*, pp. 143–178.

[6]   K. Balser, L. Hoppe, T. Eichler, M. Wendel and A.-J. Astheimer, *Ullmann´s
        Encyclopedia of Industrial Chemistry*, VCH, Vol. A5, 1986.

[7]   S. Fang, X. Lyu, T. Tong, A. I. Lim, T. Li, J. Bao and Y. H. Hu, *Nat. Commun.*,
        2023, **14**, 1203.

[8]   M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W.
        Brady and T. D. Foust, *Science*, 2007, **315**, 804–807.

[9]   P. Wanichapichart, W. Taweepreeda, S. Nawae, P. Choomgan and D. Yasen-
        chak, *Radiat. Phys. Chem.*, 2012, **81**, 949–953.

[10]   N. Yousefi Shivyari, M. Tajvidi, D. W. Bousfield and D. J. Gardner, *ACS*, 2016,
        8, 25520–25528.

[11]   Y. H. Jung, T.-H. Chang, H. Zhang, C. Yao, Q. Zheng, V. W. Yang, H. Mi, M.
        Kim, S. J. Cho, D.-W. Park, H. Jiang, J. Lee, Y. Qiu, W. Zhou, Z. Cai, S. Gong
        and Z. Ma, *Nat. Commun.*, 2015, **6**, 7170.

[12]   X. Zhang, Y. Cheng, J. You, J. Zhang, C. Yin and J. Zhang, *Nat. Commun.*, 2022,
        13, 1117.

[13]   M. E. Seeley, B. Song, R. Passie and R. C. Hale, *Nat. Commun.*, 2020, **11**, 2372.

[14]   Q. Wang, C. A. Adams, F. Wang, Y. Sun and S. Zhang, *Crit. Rev. Environ. Sci.
        Technol.*, 2022, **52**, 3211–3243.

[15]   European Bioplastics, *Bioplastics market data: www.european-bioplastics.org/market/*,
        13/12/2023.

[16]   F. Muneer, H. Nadeem, A. Arif and W. Zaheer, *Polym. Sci. Ser. C*, 2021, **63**,
        47–63.

[17]   S. Saedi, C. V. Garcia, J. T. Kim and G. H. Shin, *Cellulose*, 2021, **28**, 8877–8897.

[18]   F. Vilarinho, A. Sanches Silva, M. F. Vaz and J. P. Farinha, *Crit. Rev. Food Sci.
        Nutr.*, 2018, **58**, 1526–1537.

[19]   M. P. R. Guambo, L. Spencer, N. S. Vispo, K. Vizuete, A. Debut, D. C. White-
        head, R. Santos-Oliveira and F. Alexis, *Polymers*, 2020, **12**, 3042.

[20]   J. P. de Oliveira, G. P. Bruni, K. O. Lima, S. L. M. E. Halal, G. S. Da Rosa,
        A. R. G. Dias and E. R. Da Zavareze, *Food Chem.*, 2017, **221**, 153–160.

[21]   K. Jain, Z. Wang, L. D. Garma, E. Engel, G. C. Ciftci, C. Fager, P. A. Larsson
        and L. Wågberg, *Appl. Mater. Today*, 2023, **30**, 101703.

[22] Y.-R. Seo, J.-W. Kim, S. Hoon, J. Kim, J. H. Chung and K.-T. Lim, *Biosyst. Eng.*, 2018, **43**, 59–71.

[23] L. A. Blanchard, D. Hancu, E. J. Beckman and J. F. Brennecke, *Nature*, 1999, **399**, 28–29.

[24] P. Walden, *Bull. Acad. Imp. Sci. St.-Petersbourg*, 1914, 405–422.

[25] S. Sugden and H. Wilkins, *J. Chem. Soc., Chem.*, 1929, 1291–1298.

[26] R. Sheldon, *ChemComm.*, 2001, 2399–2407.

[27] F. Malberg, M. Brehm, O. Hollóczki, A. S. Pensado and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2013, **15**, 18424–18436.

[28] N. V. Plechkova and K. R. Seddon, *Chem. Soc. Rev.*, 2008, **37**, 123–150.

[29] M. Freemantle, *Chem. Eng. News Archive*, 1998, **76**, 32–37.

[30] F. Hermanutz, M. P. Vocht, N. Panzier and M. R. Buchmeiser, *Macromol. Mater. Eng.*, 2019, **304**, 1800450.

[31] R. P. Swatloski, S. K. Spear, J. D. Holbrey and R. D. Rogers, *J. Am. Chem. Soc.*, 2002, **124**, 4974–4975.

[32] S. Barthel and T. Heinze, *Green Chem.*, 2006, **8**, 301–306.

[33] G. T. Beckham, J. F. Matthews, B. Peters, Y. J. Bomble, M. E. Himmel and M. F. Crowley, *J. Phys. Chem. B .*, 2011, **115**, 4118–4127.

[34] C. Mukarakate, A. Mittal, P. N. Ciesielski, S. Budhi, L. Thompson, K. Iisa, M. R. Nimlos and B. S. Donohoe, *ACS Sustain. Chem. Eng.*, 2016, **4**, 4662–4674.

[35] M. Brehm, M. Pulst, J. Kressler and D. Sebastiani, *J. Phys. Chem. B .*, 2019, **123**, 3994–4003.

[36] M. Brehm, J. Radicke, M. Pulst, F. Shaabani, D. Sebastiani and J. Kressler, *Molecules*, 2020, **25**, 3539.

[37] J. Radicke, E. Roos, D. Sebastiani, M. Brehm and J. Kressler, *J. Polym. Sci.*, 2023, **61**, 372–384.

[38] B. Kosan, C. Michels and F. Meister, *Cellulose*, 2008, **15**, 59–66.

[39] J. Vitz, T. Erdmenger, C. Haensch and U. S. Schubert, *Green Chem.*, 2009, **11**, 417.

[40] H. Wang, G. Gurau and R. D. Rogers, *Chem. Soc. Rev.*, 2012, **41**, 1519–1537.

[41] M. Wohlert, T. Benselfelt, L. Wågberg, I. Furó, L. A. Berglund and J. Wohlert, *Cellulose*, 2022, **29**, 1–23.

[42] A. M. Fernandes, M. A. A. Rocha, M. G. Freire, I. M. Marrucho, J. A. P. Coutinho and L. M. N. B. F. Santos, *J. Phys. Chem. B*, 2011, **115**, 4033–4041.

[43] R. Rinaldi, *ChemComm*, 2011, **47**, 511–513.

[44] Y. Tomimatsu, H. Suetsugu, Y. Yoshimura and A. Shimizu, *J. Mol. Liq.*, 2019, **279**, 120–126.

[45] M. Mohan, T. Banerjee and V. V. Goud, *ChemistrySelect*, 2016, **1**, 4823–4832.

[46] A. Xu and Y. Zhang, *J. Mol. Struct.*, 2015, **1088**, 101–104.

[47] Y. Li, J. Wang, X. Liu and S. Zhang, *Chem. Sci.*, 2018, **9**, 4027–4043.

[48] S. Köhler, T. Liebert, M. Schöbitz, J. Schaller, F. Meister, W. Günther and T. Heinze, *Macromol. Rapid Commun.*, 2007, **28**, 2311–2317.

[49] T. F. Liebert, T. J. Heinze and K. J. Edgar, *Cellulose Solvents: For Analysis, Shaping and Chemical Modification*, J. Am. Chem. Soc., Washington, DC, 2010.

[50] Y. Li, X. Liu, S. Zhang, Y. Yao, X. Yao, J. Xu and X. Lu, *Physical Chem. Chem. Phys.*, 2015, **17**, 17894–17905.

[51] S. Gehrke, K. Schmitz and O. Hollóczki, *J. Phys. Chem. B*, 2017, **121**, 4521–4529.

[52] B. Manna and A. Ghosh, *J. Biomol. Struct. Dyn.*, 2019, **37**, 3987–4005.

[53] T. G. Paiva, M. Zanatta, E. J. Cabrita, C. E. Bernardes and M. C. Corvo, *J. Mol. Liq.*, 2022, **345**, 117810.

[54] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.

[55] Z. Liu, S. Huang and W. Wang, *J. Phys. Chem. B*, 2004, **108**, 12978–12989.

[56] X. Zhong, Z. Liu and D. Cao, *J. Phys. Chem. B*, 2011, **115**, 10027–10040.

[57] W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

[58] J. W. Ponder and D. A. Case, *Adv. Protein Chem.*, 2003, **66**, 27–85.

[59] S. V. Sambasivarao and O. Acevedo, *J. Chem. Theory Comput.*, 2009, **5**, 1038–1050.

[60] J. N. Canongia Lopes, J. Deschamps and A. A. H. Pádua, *J. Phys. Chem. B*, 2004, **108**, 2038–2047.

[61] J. N. Canongia Lopes and A. A. H. Padua, *J. Phys. Chem. B*, 2006, **110**, 19586–19592.

[62] J. N. Canongia Lopes and A. A. H. Pádua, *Theor. Chem. Acc.*, 2012, **131**, 1129.

[63] B. Doherty, X. Zhong, S. Gathiaka, B. Li and O. Acevedo, *J. Chem. Theory Comput.*, 2017, **13**, 6131–6145.

[64] K. Goloviznina, J. N. Canongia Lopes, M. Costa Gomes and A. A. H. Pádua, *J. Chem. Theory Comput.*, 2019, **15**, 5858–5871.

[65] W. Damm, A. Frontera, J. Tirado-Rives and W. L. Jorgensen, *J. Comput. Chem.*, 1997, **18**, 1955–1970.

[66] B. Doherty, X. Zhong and O. Acevedo, *J. Phys. Chem. B*, 2018, **122**, 2962–2974.

[67] D. Bedrov, J.-P. Piquemal, O. Borodin, A. D. MacKerell, B. Roux and C. Schröder, *Chem. Rev.*, 2019, **119**, 7940–7995.

[68] C. J. Cramer, *Essentials of computational chemistry: Theories and models*, John Wiley & Sons Ltd, Chichester, England, 2. ed., 2004.

[69] E. D. Isaacs, A. Shukla, P. M. Platzman, D. R. Hamann, B. Barbiellini and C. A. Tulk, *Phys. Rev. Lett.*, 1999, **82**, 600–603.

[70] S. J. Grabowski, *Chem. Rev.*, 2011, **111**, 2597–2625.

[71] A. Shahi and E. Arunan, *Phys. Chem. Chem. Phys.*, 2014, **16**, 22935–22952.

[72] A. R. Leach, *Molecular modelling: Principles and applications*, Pearson/Prentice Hall, Harlow, 2. ed., 12., 2001.

[73] D. Srivastava, M. S. Kuklin, J. Ahopelto and A. J. Karttunen, *Carbohydr. Polym.*, 2020, **243**, 116440.

[74] R. H. Atalla, J. D. Ellis and L. R. Schroeder, *J. Wood Chem. Technol.*, 1984, **4**, 465–482.

[75] P. Zugenmaier, *Prog. Polym. Sci.*, 2001, **26**, 1341–1417.

[76] R. H. Atalla and D. L. Vanderhart, *Science*, 1984, **223**, 283–285.

[77] D. L. VanderHart and R. H. Atalla, *Macromolecules*, 1984, **17**, 1465–1472.

[78] A. S. Gross and J.-W. Chu, *J. Phys. Chem. B .*, 2010, **114**, 13333–13341.

[79] Y. Nishiyama, J. Sugiyama, H. Chanzy and P. Langan, *J. Am. Chem. Soc.*, 2003, **125**, 14300–14306.

[80] B. Lindman, G. Karlström and L. Stigsson, *J. Mol. Liq.*, 2010, **156**, 76–81.

[81] A. F. Turbak, R. B. Hammer, R. F. Davies and H. L. Hegert, *Chemtech.*, 1980, **10**.

[82] K. Kamida, K. Okajima, T. Matsui and K. Kowsaka, *Polym. J.*, 1984, **16**, 857–866.

[83] C. Roy, T. Budtova and P. Navard, *Biomacromolecules*, 2003, **4**, 259–264.

[84] L. Yan and Z. Gao, *Cellulose*, 2008, **15**, 789–796.

[85] L. Zhang, D. Ruan and J. Zhou, *Ind. Eng. Chem. Res.*, 2001, **40**, 5923–5928.

[86] H. Jin, C. Zha and L. Gu, *Carbohydr. Res.*, 2007, **342**, 851–858.

[87] M. Egal, T. Budtova and P. Navard, *Cellulose*, 2008, **15**, 361–370.

[88] J. Cai, L. Zhang, S. Liu, Y. Liu, X. Xu, X. Chen, B. Chu, X. Guo, J. Xu, H. Cheng, C. C. Han and S. Kuga, *Macromolecules*, 2008, **41**, 9345–9351.

[89] X. Chen, C. Burger, F. Wan, J. Zhang, L. Rong, B. S. Hsiao, B. Chu, J. Cai and L. Zhang, *Biomacromolecules*, 2007, **8**, 1918–1926.

[90] K. Saalwächter, W. Burchard, P. Klüfers, G. Kettenbach, P. Mayer, D. Klemm and S. Dugarmaa, *Macromolecules*, 2000, **33**, 4094–4107.

[91] E. Sjöström, *Wood chemistry: Fundamentals and applications*, Academic Press, San Diego, 2. ed., 1993.

[92] H. Leipner, S. Fischer, E. Brendler and W. Voigt, *Macromol. Chem. Phys.*, 2000, **201**, 2041–2049.

[93] M. Hattori, T. Koga, Y. Shimaya and M. Saito, *Polym. J.*, 1998, **30**, 43–48.

[94] C. L. McCormick, P. A. Callais and B. H. Hutchinson, *Macromolecules*, 1985, **18**, 2394–2401.

[95] A. Potthast, T. Rosenau, R. Buchner, T. Röder, G. Ebner, H. Bruglachner, H. Sixta and P. Kosma, *Cellulose*, 2002, **9**, 41–53.

[96] T. Rosenau, A. Potthast, H. Sixta and P. Kosma, *Prog. Polym. Sci.*, 2001, **26**, 1763–1837.

[97] E. Maia, A. Peguy and S. Pérez, *Acta Cryst. B*, 1981, **37**, 1858–1862.

[98] H. Shen, T. Sun and J. Zhou, *Macromol. Mater. Eng.*, 2023, **308**, 2300089.

[99] J. P. Hallett and T. Welton, *Chem. Rev.*, 2011, **111**, 3508–3576.

[100] C. Graenacher, *pat.*, 1,943,176.

[101] Y. Bai, Y. Cao, J. Zhang, M. Wang, R. Li, P. Wang, S. M. Zakeeruddin and M. Grätzel, *Nat. Mater.*, 2008, **7**, 626–630.

[102] E. E. L. Tanner, A. M. Curreri, J. P. R. Balkaran, N. C. Selig-Wober, A. B. Yang, C. Kendig, M. P. Fluhr, N. Kim and S. Mitragotri, *Adv. Mater.*, 2019, **31**, e1901103.

[103] H. Sakaebe, H. Matsumoto and K. Tatsumi, *Electrochim. Acta*, 2007, **53**, 1048–1054.

[104] Z. M. Judeh, C. B. Ching, J. Bu and A. McCluskey, *Tetrahedron Lett.*, 2002, **43**, 5089–5091.

[105] Y. Zhou and J. Qu, *ACS Appl. Mater. Interfaces*, 2017, **9**, 3209–3222.

[106] T. Heinze, K. Schwikal and S. Barthel, *Macromol. Biosci.*, 2005, **5**, 520–525.

[107] N. Sun, M. Rahman, Y. Qin, M. L. Maxim, H. Rodríguez and R. D. Rogers, *Green Chem.*, 2009, **11**, 646.

[108] J. S. Moulthrop, R. P. Swatloski, G. Moyna and R. D. Rogers, *ChemComm.*, 2005, 1557–1559.

[109] W. Zheng, A. Mohammed, L. G. Hines, D. Xiao, O. J. Martinez, R. A. Bartsch, S. L. Simon, O. Russina, A. Triolo and E. L. Quitevis, *J. Phys. Chem. B .*, 2011, **115**, 6572–6584.

[110] F. Ren, J. Wang, J. Yu, C. Zhong, F. Xie and S. Wang, *ACS omega*, 2021, **6**, 27225–27232.

[111] M. J. Kamlet and R. W. Taft, *J. Am. Chem. Soc.*, 1976, **98**, 377–383.

[112] R. W. Taft and Mortimer J. Kamlet, *J. Am. Chem. Soc.*, 1976, **98**, 2886–2894.

[113] C. Reichardt, *Chem. Rev.*, 1994, **94**, 2319–2358.

[114] M. A. Ab Rani, A. Brant, L. Crowhurst, A. Dolan, M. Lui, N. H. Hassan, J. P. Hallett, P. A. Hunt, H. Niedermeyer, J. M. Perez-Arlandis, M. Schrems, T. Welton and R. Wilding, *PCCP*, 2011, **13**, 16831–16840.

[115] A. Brandt, J. Gräsvik, J. P. Hallett and T. Welton, *Green Chem.*, 2013, **15**, 550.

[116] L. K. J. Hauru, M. Hummel, A. W. T. King, I. Kilpeläinen and H. Sixta, *Biomacro-molecules*, 2012, **13**, 2896–2905.

[117] C. Froschauer, M. Hummel, M. Iakovlev, A. Roselli, H. Schottenberger and H. Sixta, *Biomacromolecules*, 2013, **14**, 1741–1750.

[118] H. Ohno and Y. Fukaya, *Chem. Lett.*, 2009, **38**, 2–7.

[119] D. L. Minnick, R. A. Flores, M. R. DeStefano and A. M. Scurto, *J. Phys. Chem. B*, 2016, **120**, 7906–7919.

[120] B. Lindman, B. Medronho, L. Alves, M. Norgren and L. Nordenskiöld, *Q. Rev. Biophys.*, 2021, **54**, e3.

[121] B. Medronho, A. Romano, M. G. Miguel, L. Stigsson and B. Lindman, *Cellulose*, 2012, **19**, 581–587.

[122] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminform.*, 2012, **4**, 17.

[123] W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

[124] J. Stone, *An efficient library for parallel ray tracing and animation*, MSc thesis, University of Missouri-Rolla, 1998.

[125] J.-P. Mikkola, A. Kirilin, J.-C. Tuuf, A. Pranovich, B. Holmbom, L. M. Kustov, D. Y. Murzin and T. Salmi, *Green Chem.*, 2007, **9**, 1229–1237.

[126] S.-L. Quan, S.-G. Kang and I.-J. Chin, *Cellulose*, 2010, **17**, 223–230.

[127] M. Bagheri, H. Rodríguez, R. P. Swatloski, S. K. Spear, D. T. Daly and R. D. Rogers, *Biomacromolecules*, 2008, **9**, 381–387.

[128] L.-Y. Long, Y.-X. Weng and Y.-Z. Wang, *Polymers*, 2018, **10**, 623.

[129] S. H. Lee, T. V. Doherty, R. J. Linhardt and J. S. Dordick, *Biotechnol. Bioeng.*, 2009, **102**, 1368–1376.

[130] A. M. Scurto, S. N. V. K. Aki and J. F. Brennecke, *ChemComm.*, 2003, 572–573.

[131] M. J. Earle, J. M. S. S. Esperança, M. A. Gilea, J. N. C. Lopes, L. P. N. Rebelo, J. W. Magee, K. R. Seddon and J. A. Widegren, *Nature*, 2006, **439**, 831–834.

[132] Q. Gan, M. Xue and D. Rooney, *Separation and Purification Technology*, 2006, **51**, 185–192.

[133] A. Xu, J. Wang and H. Wang, *Green Chem.*, 2010, **12**, 268–275.

[134] C. Gerthsen, *Physik*, Springer, Berlin, 23. ed., 2006.

[135] W. Demtröder, *Atoms, molecules and photons: An introduction to atomic-, molecular- and quantum physics*, Springer, Berlin and Heidelberg, 3. ed., 2018.

[136] N. L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127–8134.

[137] N. L. Allinger, Y. H. Yuh and J. H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551–8566.

[138] J. H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, 1989, **111**, 8566–8575.

[139] J. H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, 1989, **111**, 8576–8582.

[140] W. Demtröder, *Atome, Moleküle und Festkörper*, Springer, Berlin and New York, 3. ed., 2005, vol. 3.

[141] P. Bordat, F. Affouard, M. Descamps and K. L. Ngai, *Phys. Rev. Lett.*, 2004, **93**, 105502.

[142] P. Bordat, F. Affouard, M. Descamps and K. L. Ngai, *J. Non-Cryst. Solids*, 2006, **352**, 4630–4634.

[143] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, Clarendon Press, Oxford, 1987.

[144] P. P. Ewald, *Ann. Phys.*, 1921, **369**, 253–287.

[145] F. Jensen, *Introduction to computational chemistry*, Wiley, Chichester, Reprint, 2004.

[146] S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.

[147] D. Marx, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, Leiden, 2009.

[148] P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.

[149] R. G. Parr, in *Horizons of Quantum Chemistry*, ed. K. Fukui and B. Pullman, Springer Netherlands, Dordrecht, 1980, pp. 5–15.

[150] W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.

[151] A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.

[152] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

[153] M. Sprik, J. Hutter and M. Parrinello, *J. Chem. Phys.*, 1996, **105**, 1142–1152.

[154] S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B*, 1996, **54**, 1703–1710.

[155] C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B*, 1998, **58**, 3641–3662.

[156] L. Verlet, *Physical Review*, 1967, **159**, 98–103.

[157] L. Verlet, *Physical Review*, 1968, **165**, 201–214.

[158] W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *Chem. Phys.*, 1982, **76**, 637–649.

[159] M. E. Tuckerman, *Statistical mechanics: Theory and molecular simulation*, Oxford University Press, Oxford, 2010.

[160] H. C. Andersen, *Chem. Phys.*, 1980, **72**, 2384–2393.

[161] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys*, 1984, **81**, 3684–3690.

[162] T. Schneider and E. Stoll, *Phys. Rev. B*, 1978, **17**, 1302–1322.

[163] B. Dünweg and W. Paul, *Int. J. Mod. Phys. C*, 1991, **02**, 817–827.

[164] S. Nose, *J. Chem. Phys.*, 1984, **81**, 511–519.

[165] S. Nose, *Mol. Phys.*, 1984, **52**, 255–268.

[166] G. Martyna, M. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2535–2643.

[167] M. Brehm, M. Thomas, S. Gehrke and B. Kirchner, *Chem. Phys.*, 2020, **152**, 164105.

[168] M. Brehm and B. Kirchner, *J. Chem. Inf. Model*, 2011, **51**, 2007–2023.

[169] E. Roos and M. Brehm, *Phys. Chem. Chem. Phys.*, 2021, **23**, 1242–1253.

[170] E. Roos, D. Sebastiani and M. Brehm, *Molecules*, 2023, **28**, 7592.

[171] M. Brehm, *Analyzing Trajectories From Molecular Simulation*, PhD thesis, Leipzig, 2014.

[172] C. Olsson, A. Idström, L. Nordstierna and G. Westman, *Carbohydr. Polym.*, 2014, **99**, 438–446.

[173] X. Wu, Z. Liu, S. Huang and W. Wang, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2771–2779.

[174] B. L. Bhargava and S. Balasubramanian, *J. Chem. Phys.*, 2007, **127**, 114510.

[175] H. Liu and E. Maginn, *J. Chem. Phys.*, 2011, **135**, 124507.

[176] P. Ray, R. Elfgen and B. Kirchner, *Phys. Chem. Chem. Phys.*, 2019, **21**, 4472–4486.

[177] E. Roos, D. Sebastiani and M. Brehm, *Phys. Chem. Chem. Phys.*, 2023, **25**, 8755–8766.

[178] E. Roos, C. Gradaus, D. Sebastiani and M. Brehm, *Cellulose*, 2024.

[179] T. Liebert, in *Cellulose Solvents: For Analysis, Shaping and Chemical Modification*, ed. T. F. Liebert, T. J. Heinze and K. J. Edgar, J. Am. Chem. Soc, Washington, DC, 2010, vol. 1033, pp. 3–54.

[180] D. T. Bowron, C. D'Agostino, L. F. Gladden, C. Hardacre, J. D. Holbrey, M. C. Lagunas, J. McGregor, M. D. Mantle, C. L. Mullan and T. G. A. Youngs, *J. Phys. Chem. B*, 2010, **114**, 7760–7768.

[181] N. Yaghini, J. Pitawala, A. Matic and A. Martinelli, *J. Phys. Chem. B*, 2015, **119**, 1611–1622.

[182] L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.

[183] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, **120**, 9665–9678.

[184] J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

[185] H. C. Andersen, *J. Comput. Phys.*, 1983, **52**, 24–34.

[186] J. Shi, K. Balamurugan, R. Parthasarathi, N. Sathitsuksanoh, S. Zhang, V. Stavila, V. Subramanian, B. A. Simmons and S. Singh, *Green Chem.*, 2014, **16**, 3830–3840.

Curriculum Vitae

Personal Data

Name: Eliane Roos
Nationality: German

Education/Professional Career

since 11/2023	Project manager EWN – Entsorgungswerk für Nuklearanlagen GmbH Nuclear power plant Rheinsberg
since 01/2020	Doctoral studies Martin Luther University Halle-Wittenberg Institute of Chemistry (Theoretical Chemistry) Supervisor: Prof. Dr. Daniel Sebastiani
10/2017-10/2019	Graduate studies in Chemistry (*Master of Science*) Martin Luther University Halle-Wittenberg
10/2014 -11/2017	Undergraduate studies in Chemistry (*Bachelor of Science*) Martin Luther University Halle-Wittenberg
07/2014	University-entrance diploma Burg-Gymnasium Wettin, Wettin

Scholarships

10/2017-09/2018	Deutschlandstipendium (Graduate studies)
10/2014-03/2015	Deutschlandstipendium (Undergraduate studies)

List of Publications

[5]* <u>E. Roos</u>, C. Gradaus, D. Sebastiani, M. Brehm, "A Force Field for the Solubility of Cellulose in DMSO/Ionic Liquids", *Cellulose*, 2024, *accepted*

[4]* <u>E. Roos</u>, D. Sebastiani, M. Brehm, "BILFF: All-Atom Force Field for Modeling Triazolium- and Benzoate-Based Ionic Liquids", *Molecules*, 28, 7592, 2023

[3]* <u>E. Roos</u>, D. Sebastiani, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 2: Cellulose in [EMIm][OAc]/Water Mixtures", *Phys. Chem. Chem. Phys.*, 25, 8755-8766, 2023.

[2] J. Radicke, <u>E. Roos</u>, D. Sebastiani, M. Brehm, J. Kressler, "Lactate-based Ionic Liquids as Chiral Solvents for Cellulose", *J. Polym. Sci.*, 61, 5, 372-384, 2023.

[1]* <u>E. Roos</u>, M. Brehm, "A Force Field for Bio-Polymers in Ionic Liquids (BILFF) – Part 1: [EMIm][OAc]/Water Mixtures", *Phys. Chem. Chem. Phys.*, 23, 1242-1253, 2021.

*Publications marked with an asterisk are part of this thesis.

Scientific Talks and Posters

Listed below are conferences and meetings where I have given a scientific talk or presented a poster based on the results of this thesis. Several posters and talks presented at internal meetings of the graduate school GRK 2670-BEAM are not listed separately.

[5] 1st BEAM Symposium, September 6-8, 2023, Halle (Saale), Germany (*talk*)

[4] 58th Symposium on Theoretical Chemistry (STC 2022), September 18-22, 2022, Heidelberg, Germany (*poster*)

[3] 57th Symposium on Theoretical Chemistry (STC 2021), September 20-24, 2021, Würzburg (online), Germany (*poster*)

[2] Bunsen-Tagung 2021, May 10-12, 2021, Regensburg (online), Germany (*talk*)

[1] 55th Symposium on Theoretical Chemistry (STC 2019), September 22-26, 2019, Rostock, Germany (*poster*)

Statutory Declaration – Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht.
Ich erkläre, keine weiteren Promotionsversuche unternommen und die Dissertation in dieser oder einer anderen Fassung keiner anderen wissenschaftlichen Einrichtung vorgelegt zu haben.

Halle (Saale), May 14, 2024	
Ort, Datum	Unterschrift

Acknowledgements

I would like to take this opportunity to express my sincere gratitude and appreciation to the many people who have supported and inspired me along my academic journey.

First of all, I would like to thank *Prof. Dr. Daniel Sebastiani* for the opportunity to work on this very interesting project and for the advice, support and insightful discussions. Additionally, I am thankful for the enriching experiences gained from attending all the insightful conferences. Furthermore, I am grateful to have had the opportunity to organize the laboratory course in physical chemistry.

A great debt of gratitude is owed to *Prof. Dr. Martin Brehm* for his guidance, his support, and his introduction to the MD simulations. Our fruitful and motivating discussions helped me a lot both scientifically and personally.

My thanks also go to all members of the research group for the pleasant working atmosphere and for the enjoyable moments we shared together. In particular, I would like to thank the BILFF users in the group, especially *Cora Gradaus* for her work during her Bachelor's thesis and Master's thesis using BILFF to study the influence of DMSO on cellulose solubility. Likewise, my thanks go to *Luisa Voigt* for her work during her Bachelor's thesis investigating the influence of the cation of ionic liquids using BILFF.

I would also like to express my gratitude to *Prof. Dr. Jörg Kressler* and *Julian Radicke* for the synthesis of the ionic liquids, as well as for the measurements of their physical properties, and for the measurements of the cellulose solubility. I would also like to thank Julian for the scientific exchange, which productively contributed to this thesis.

The graduate school *BEAM – RTG 2670* provided me with valuable scientific exchanges and workshops that helped me improve skills in different areas. I am very grateful for the funding of my own research project under the incubator program, which focused on the effect of DMSO on cellulose solubility. The opportunity to be the spokesperson was an honor and a meaningful experience.

A very special thanks goes to my partner *Matthias*, who has always accompanied me since my studies and supported me with wisdom, humor and unwavering understanding.

I want to sincerely express my appreciation to my family, with a special mention to my *mother*, for her constant support throughout my entire educational journey.

A big thank you also goes to my two friends *Norman* and *Moritz*, on whom I have always been able to rely.

A huge thank you to everyone!

***