universität wien

MASTERARBEIT

Titel der Masterarbeit

„Coarse-Graining in Molecular Dynamics Simulations“

verfasst von

Daniel Braun, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 066 863
Studienrichtung lt. Studienblatt:	Masterstudium Biologische Chemie
Betreut von:	Univ.-Prof. Dr. Othmar Steinhauser

Danksagung

Zuerst möchte ich Allen am Institut für computergestützte biologische Chemie danken. Insbesondere gilt dieser Dank Prof. Dr. Othmar Steinhauser, der mich im Verlauf dieser Arbeit mit grösster Fürsorge und Kompetenz, weit über naturwissenschaftliche Belangen hinaus, angeleitet hat. Gleichermaßen danke ich Prof. Dr. Stefan Boresch, durch dessen Kooperationsbereitschaft, Einsatz und Expertise diese Arbeit in der Form erst ermöglicht wurde. Weiters danke ich meinen Kollegen Dr. Christian Schröder, Michael Schmollngruber, Sonja Maurer, und Dr. Marcello Sega, die mich in professioneller, wie auch in persönlicher Hinsicht beraten und mit viel Humor und Herz begleitet haben.

Natürlich gebührt mein größter Dank meiner Familie- meinen Eltern, die mir die Wunder dieser Welt gezeigt haben und mich immer in allen Vorhaben unterstützt haben, und meinen Brüdern Sebastian und Benjamin, für ihre Liebe und Anteilnahme.

Ich danke allen Menschen, die mich auf meinem bisherigen Lebensweg begleitet, unterstützt, und inspiriert haben.

Zusammenfassung

Diese Arbeit befasst sich mit der Simulation chemischer Systeme im Hinblick auf biologische Fragestellungen. Da in der Biologie sowohl räumliche als auch zeitliche Dimensionen mehrere Größenordnungen überstreichen, sind Überlegungen zur Auflösung des Models entscheidend. Während exate Darstellungen oft zu aufwendig sind, um sie in einer vernünftigen Zeit am Computer simulieren zu können, bieten abstraktere Modelle oft nicht genug Einblick in die physikochemischen Eigenschaften, die verantwortlich sind für das Auftreten emergenter Phänomene in der Biologie.

Auf der einen Seite besteht der Wunsch komplexe Systeme zu simulieren, auf der anderen Seite möchte man entscheidende Teile desselben in genauerer Auflösung betrachten können. Daraus entwickelte sich die Idee eines Multi-Scale Ansatzes. In der Molekulardynamik ist man meist an den Eigenschaften des biologischen Makromoleküls interessiert. Trotzdem verbraucht die Berechnung des Lösungsmittels den größten Teil der Rechenleistung. Deshalb erscheint es sinnvoll, zuerst nur das Lösungsmittel in einer geringern Auflösung zu simulieren. In biologischen Systemen in das Lösungsmittel Wasser. Demzufolge ist es für die Anwendung der Molekulardynamik in Biologie und Medizin von enormer Wichtigkeit ein effizientes, aber dennoch physikalisch repräsentatives Wassermodell zu finden.

Chapter 1 gibt eine kurze Einführung in grundlegende Prinzipien der Molekulardynamik. Chapter 2 beinhaltet die Untersuchung dreier Wassermodelle in unterschiedlichen Auflösungen: BMW, ein coarse-grained Wassermodell, und SPC/E und TIP3P, die beide atomistische Wassermodelle sind. Im Fokus sind die Effizienz und vor allem die Genauigkeit der physikalischen Eigenschaften der drei Modelle, und deren Vergleich. Das Modell geringster Auflösung, BMW, zeigt ein überraschend gutes Verhalten nicht nur bezüglich statischer sondern auch dynamischer Eigenschaften. Außerdem wurde im Zuge dieser Analyse festgestellt, dass die gesamte Dynamik systematisch vom Zeitschritt, mit dem das System simuliert wurde, abhängt. Das Manuskript in Chapter II wurde in *The Journal of Chemical Physics* zur Publikation eingereicht.

Abstract

This work is concerned with the simulation of chemical systems in a biological context. Since biological systems are complex and cover a huge range of spatial and temporal scales, the development of meaningful models certainly involves considerations on resolution. Exact descriptions are often computationally too costly to simulate systems of biological interest in reasonable amounts of time, while more abstract models sometimes offer not enough insight into the physicochemical basis of emergent biological phenomena.

From the desire to simulate large and complex systems, and at the same time being able to investigate critical parts in great detail, there arises the idea of multiscale approaches. In molecular dynamics, while one is usually interested in the behavior of the biomolecular solute, calculation of the solvent consumes most of computational resources. Thus, it makes sense to describe the solvent in lower resolution. In biological systems water is the most prominent solvent. For molecular dynamics simulation and its application in biology and medicine, it is of prime importance to find an efficient, yet physically accurate water model.

While Chapter 1 introduces into some basic principles of molecular dynamics, Chapter 2 comprises the investigation of three different water models in different resolution: BMW, a coarse-grained water model, and SPC/E and TIP3P, two all-atomistic water models. Their performance regarding efficiency and physical accuracy is critically assessed. The model in lower resolution, BMW, performs surprisingly well not only concerning static but also dynamical properties. Aside from that, it is shown that system dynamics depends on the time step in a systematic way.

The manuscript provided in Chapter 2 was submitted to *The Journal of Chemical Physics* for publication.

Contents

1. Introduction

1.1. Molecular Dynamics: a powerful method

Terrestrial life evolved in the liquid phase and biological systems occur in solution [1]. Liquid character is even attributed to biological macromolecules, e.g., proteins, themselves [2]. This means that the structure of a biomolecule is not rigid but highly flexible, and therefore its average is a crude simplification. In fact, a more complete and better description of a biomolecular system and its functionality is given by its dynamics, its behavior over time [2]. Molecular dynamics (MD) is a computational method by which large chemical systems, especially fluids can be well characterized, and consequently is an optimal choice to treat biological questions, which require chemical resolution.

The beginnings of computational MD dates back to 1957. Alder and Wainwright [3] simulated so-called hard spheres. While this has little to do with real chemistry, seven years later, Rahman [4] managed to simulate systems of liquid argon interacting through a Lennard-Jones potential. He analyzed the trajectories with respect to various static and dynamical quantities and compared them to experiment. In 1971 the first aqueous system was simulated [5]. The water model (BNS) was critically assessed with respect to experimental data by the authors themselves. This work served as a well-founded starting point for the development of further water models

for MD simulations- an ongoing task in computational (bio)chemistry, and also the main interest of this work.

The simulation of a physical system on the computer is confronted with technical limitations. While a physical system is effectively infinite and evolves smoothly in time, of course, a model is finite in size and precision. The evolution of a modeled system in time is an especially crucial topic and will be discussed in the next section.

1.2. Time evolution

In molecular dynamics the motion of the system is calculated by solving the Newtonian equations of motion for all particles simultaneously [3]. These equations relate the force $\vec{F}$ and the change of motion:

$$\frac{d\vec{v}}{dt} \cdot m = \vec{F} \tag{1.1}$$

with $\vec{v}$ being the velocity and m being the mass of the particle. The force on a particle $\vec{F}(t)$ itself is a highly complex function of time, or more precisly a function of the time evolution of the potential exerted from all other particles (cf. next section). In other words, we have a large set of coupled differential equations, which cannot be solved analytically, but only numerically as a set of difference equations [4]. The time evolution of coordinates $\vec{r}(t)$ and velocities $\vec{v}(t)$ of all particles is calculated at discrete time points, i.e. in small increments of time Δt. After n steps the system has arrived at time $(t_0 + n \cdot \Delta t)$ or (t_n). The quality of this approximation mainly relies on the size of the time step Δt and on the specific finite difference algorithm used. While the first issue will be addressed in great detail later (cf. Chapter 2), a brief overview of finite difference algorithms of motion is presented in this section.

If we approximate $\vec{r}(t_n + \Delta t)$ – the coordinates in the next time step – as a Taylor series expansion about the instantaneous time t_n, we get

$$\vec{r}(t_n+\Delta t) = \vec{r}(t_n)+\dot{\vec{r}}(t_n)\cdot\Delta t+\ddot{\vec{r}}(t_n)\cdot\frac{\Delta t^2}{2}+\dddot{\vec{r}}(t_n)\cdot\frac{\Delta t^3}{6}+...+\mathcal{O}(\Delta t^k) \quad (1.2)$$

Because Δt is supposed to be very small, the higher order terms have little contribution and under certain conditions can be neglected.

To consider only the zeroth and first term results in the formulation of the Euler scheme for classical motion:

$$\vec{r}(t_n + \Delta t) = \vec{r}(t_n) + \vec{v}(t_n) \cdot \Delta t + \mathcal{O}(\Delta t^2) \quad (1.3)$$

However, the Euler scheme is a first order integration scheme, and the error grows linearly with time. This introduces a systematic energy drift, contradicting the conservation laws of physics and involving a biased sampling of phase space.

An integration scheme which meets most requirements of the molecular world is the Störmer-Verlet scheme [6, 7]. It can be deduced with an additional Taylor series expansion

$$\vec{r}(t_n-\Delta t) = \vec{r}(t_n)-\dot{\vec{r}}(t_n)\cdot\Delta t+\ddot{\vec{r}}(t_n)\cdot\frac{\Delta t^2}{2}-\dddot{\vec{r}}(t_n)\cdot\frac{\Delta t^3}{6}\pm...\pm\mathcal{O}(\Delta t^k) \quad (1.4)$$

Adding Eq. 1.2 and Eq. 1.4 yields

$$\vec{r}(t_n + \Delta t) + \vec{r}(t_n - \Delta t) = 2 \cdot \vec{r}(t_n) + \ddot{\vec{r}}(t_n) \cdot \Delta t^2 + \mathcal{O}(\Delta t^4) \quad (1.5)$$

which can be also written as

$$\vec{r}(t_n + \Delta t) = 2 \cdot \vec{r}(t_n) - \vec{r}(t_n - \Delta t) + \frac{\vec{F}(t_n)}{m} \cdot \Delta t^2, \quad (1.6)$$

the coordinate-Verlet scheme. From Eq. 1.5 one sees that the Verlet scheme is of third order in $\vec{r}$ (the third term cancles out), and therefore, with small time steps, is reasonably accurate. It also preserves the microreversibility of time, because

$$\vec{r}(t_n - \Delta t) = 2 \cdot \vec{r}(t_n) - \vec{r}(t_n + \Delta t) + \frac{\vec{F}(t_n)}{m} \cdot \Delta t^2. \tag{1.7}$$

Eq. 1.6 and 1.7 can alternatively be converted by time reversal ($\Delta t \to -\Delta t$). The drawback of the coordinate-Verlet scheme is that the velocities are not included explicitly, which are needed to compute e.g. the kinetic energy. They can be calculated approximately from the coordinates by

$$\vec{v}(t_n) = \frac{\vec{r}(t_n + \Delta t) - \vec{r}(t_n - \Delta t)}{2\Delta t}. \tag{1.8}$$

The leapfrog scheme [8], another variant of the Störmer-Verlet scheme, introduces velocities explicitly by defining the simple relation

$$\vec{v}(t_n + \frac{\Delta t}{2}) = \frac{\vec{r}(t_n + \Delta t) - \vec{r}(t_n)}{\Delta t}, \tag{1.9}$$

which is equivalent to

$$\vec{v}(t_n - \frac{\Delta t}{2}) = \frac{\vec{r}(t_n) - \vec{r}(t_n - \Delta t)}{\Delta t}. \tag{1.10}$$

Eq. 1.9 can be directly written as the first part of the leapfrog scheme

$$\vec{r}(t_n + \Delta t) = \vec{r}(t_n) + \vec{v}(t_n + \frac{\Delta t}{2}) \cdot \Delta t. \tag{1.11}$$

Solving Eq. 1.10 for $\vec{r}(t_n - \Delta t)$ and substituting the result in Eq. 1.6, together with Eq. 1.9 directly yields the formula for the velocities as the second part of the leapfrog scheme

$$\vec{v}(t_n + \frac{\Delta t}{2}) = \vec{v}(t_n - \frac{\Delta t}{2}) + \frac{\vec{F}(t_n)}{m} \cdot \Delta t. \tag{1.12}$$

This results in a trajectory with exact coordinates at full times $\vec{r}(t)$ and velocities at half times $\vec{v}(t + \Delta t)$. Approximate velocities at full times can be obtained through a forumla

$$\vec{v}(t_n) = \frac{1}{2} \cdot \left(\vec{v}(t_n + \frac{\Delta t}{2}) + \vec{v}(t_n - \frac{\Delta t}{2}) \right) \tag{1.13}$$

Because of this approximation, concerns were raised about a correct calculation of velocities within the leapfrog scheme [9, 10].

A method to include velocities at full times explicitly into the integration algorithm and hence into the propagation of the simulated system is the velocity-Verlet scheme [11]. It can be obtained by solving Eq. 1.13, multiplied by 2, for $\vec{v}\ (t_n\ -\ \Delta t/2)$ which combined with Eq. 1.12 yields

$$\vec{v}(t_n + \frac{\Delta t}{2}) = \vec{v}(t_n) + \frac{\vec{F}(t_n)}{2m} \cdot \Delta t \tag{1.14}$$

This together with Eq. 1.11 yields the formula for the coordinates in the velocity-Verlet scheme

$$\vec{r}(t_n + \Delta t) = \vec{r}(t_n) + \vec{v}(t_n) \cdot \Delta t + \frac{\vec{F}(t_n)}{2m} \cdot \Delta t^2. \tag{1.15}$$

Since Eq. 1.12 is equivalent to

$$\vec{v}(t_n + \frac{3 \cdot \Delta t}{2}) = \vec{v}(t_n + \frac{\Delta t}{2}) + \frac{\vec{F}(t_n + \Delta t)}{m} \cdot \Delta t \tag{1.16}$$

and Eq. 1.13 is equivalent to

$$\vec{v}(t_n + \Delta t) = \frac{1}{2} \cdot \left(\vec{v}(t_n + \frac{3 \cdot \Delta t}{2}) + \vec{v}(t_n + \frac{\Delta t}{2}) \right) \tag{1.17}$$

we can write

$$\vec{v}(t_n + \Delta t) = \vec{v}(t_n + \frac{\Delta t}{2}) + \frac{\vec{F}(t_n + \Delta t)}{2m} \cdot \Delta t. \tag{1.18}$$

Now the formula for the velocities in the velocity-Verlet scheme can be deduced from Eq. 1.14 and Eq. 1.18 as

$$\vec{v}(t_n + \Delta t) = \vec{v}(t_n) + \frac{\vec{F}(t_n + \Delta t) + \vec{F}(t_n)}{2m} \cdot \Delta t. \tag{1.19}$$

With any of the algorithms presented a whole trajectory can be calculated recursively from initial conditions $\vec{r}(0)$ and $\vec{v}(0)$. Generally speaking, the new coordinates and velocities for all particles in the system are calculated from the already known ones and the forces at the current time. This is done until the desired length of the trajectory is reached [12].

Probably the greatest challenge in constructing an integration algorithm for MD is to yield the "correct" trajectories even if rather large time steps Δt are employed. This is so crucial because it is the calculation of forces $\vec{F}(t)$ in every time step, which consumes most of computational resources [1]. Algorithms which allow for rather large time steps without perturbing the trajectory too much, thus, are very efficient in terms of computer time.

Not only the integration algorithm but also the calculation of forces themselves has to be optimized. Some basic considerations and requirements and mainly the idea of optimizing the calculation of interactions through coarse-graining will be presented in the next section.

1.3. Calculation of interactions

As shortly mentioned in the previous section the force $\vec{F}(t)$ acting on a particle is a function of time and consequently must be calculated anew in every time step for every particle in the system. During a simulation this is done reams of times. Hence, it is a critical process demanding a number of considerations and simplifications to be feasible at all.

In a system of N particles the force on particle i is calculated from the potential energy of its surrounding consisting of all particles j as

$$\vec{F_i} = -\vec{\nabla} \sum_{j \neq i}^{N} U_{ij}(r_{ij}) \tag{1.20}$$

where $\vec{\nabla}$ is the Nabla operator and stands for the gradient. For the sake of efficiency, in classical MD the potential U_{ij} does not explicitly include quantum-mechanical potentials. Rather, the idea of the Born-Oppenheimer approximation is practically applied by averaging over electronic degrees of freedom [13]. Both, nuclear and electronic contributions are usually employed at the sites of the nuclei – if one particle in the simulation corresponds to exactly one atom, this is called an all-atom (AA) description. Furthermore, the potential U_{ij} is a two-body interaction, which is an approximation to the many-body potential [1, 4]. In other words, the total potential acting on i is a superposition of effective pair potentials between particle i and each one of all other particles j. It is still a tremendous effort to obtain all the distance vectors $\vec{r}_{ij}$ between the particles $(1/2\ (N-1)N)$, and calculate the respective forces [1]. Neighbour lists and cut-off scheme methods are implemented to reduce the computational effort by considering only the contribution of the nearest neighbours in the force calculation. Moreover, long-range interactions (electrostatics) and the spatial finiteness of the model require special treatment, such as the implementation of peri-

odic boundary conditions. Those are highly critical topics in MD simulation but they fall out of the scope of this work.

Systems of biological interest are often complex and exhibit spatial and temporal scales which are currently inaccessible to AA simulations. Although computational resources and algorithms for MD simulations are evolving tremendously, it does not suffice to wait until these developments meet the requirements for treating certain scientific questions. This is easily exemplified: simulating a protein of twice the diameter of a protein that we are currently able to simulate requires a box which is at least twice as large and thus includes $2^3 = 8$ times more particles. Considering Moore's law we would have to wait three years until this is reasonably feasible. Now, in biology it is often the interaction of two or more macromolecules which is of interest. This, of course, amounts to much greater spatial scales. But it does not end there: ambitions (and not least the medical need) exist not only to simulate systems consisting of a few proteins, but whole parts of a cell or even whole cells [14]. On top of all that, the required temporal scales for the accurate sampling of physical quantities grows very rapidly with size and complexity [15]; an example just mentioned is the time scale on which intramolecular motions occur vs. the time scale of intermolecular dynamics of macromolecules. It becomes clear that there exists an urge to describe biological systems at lower resolution and, in the course of that, expand the range of problems to be able to investigate.

Aside from that, it remains the desire to deduce biological phenomena from basic physicochemical principles [13]. Therefore, it is necessary to find a solution which not only allows for simulating larger systems, but simulating them correctly. Since in biomolecular simulations the solvent "can easily comprise more than 80 % of the particles and computational effort in an atomistic MD simulation" [13], and it is usually the solute which is of interest, it seems reasonable to consider possibilities of simplifying the

solvent first. Usually solvent molecules move much faster than the biomolecular solutes. This is a physical motivation to eliminate all fluctuations of the solvent and include its average effect on the solute implicitly. Implicit solvation is a frequently used method, because it entails a huge gain in computational efficiency. However, although this is tempting (and also justified in some applications), many important features, especially in biological applications, are lost. For example, macromolecules that are not in direct contact with one another still transfer linear and angular momentum via intervening water molecules [14]. Furthermore, water-induced effects contribute to hydrophobic forces that have considerable effects in protein folding [13]. These are just a few of the reasons to stick with a particle-based approach. A suitable compromise is coarse-graining (CG), a still particle-based approach, but in lower resolution. While in AA models each particle represents one atom, in CG models the atom is not considered a strict upper boundary and further degrees of freedom are eliminated. Thus, CG particles present groups of atoms or even molecules, ideally still holding the essential and desired features of the underlying chemical system.

The reduced number of particles in CG systems entails a reduced number of interactions and so a reduced number of costly force calculations in each time step of a MD simulation (see, e.g., Eq. 1.6). Also, usually a larger time step Δt can be applied (for a detailed discussion see Chapter 2) for a further gain in computational efficiency. Taking it all together, CG models provide a suitable basis for physically accurate and sufficiently detailed, but also computationally much more efficient MD simulations [16].

Actually, coarse-graining in MD is a relatively common approach. For example, in one of the first works on MD simulations of proteins studying protein dynamics – Karplus et al. (1977) [2] – all hydrogens were united with their respective heavier atoms (today known as united-atom approach).

Nonetheless, the interest of the scientific community in CG has dramatically and steadily been increasing especially in the last 15 years [16].

There are a number of different approaches for developing a CG model, depending on the application [17]. A general distinction can be made between "top-down" and "bottom-up" approaches [13]. While the former relies on and tries to emergently reproduce experimental observations, the latter employs information from more detailed descriptions of the system, mostly from AA models or polarizable [18] AA models. Also, the latter allows to deduce emergent (biological) phenomena from "chemical properties", because the particles often correspond to chemical entities such as functional groups, or at least have some other characteristics stemming from the microscopic description, as we will see shorlty. Of course, the distinction between the two approaches becomes difficult, as more and more models mix them. For example, the CG water model BMW [19] includes information about the dipole moment of water clusters of four water molecules from various AA simulations, and uses the remaining degrees of freedom to correctly reproduce experimental quantities such as the density and the static dielectric constant (cf. Chapter 2).

Care must be taken, though, because the effective potentials in CG force fields, like in AA force fields, are parametrized under certain assumptions, e.g., for a certain thermodynamic range [17] and perform well only in some particular applications and under some particular physical conditions [16].

While there exist a number of CG models in different resolutions for biological macromolecules themselves (e.g., for proteins, nucleic acids, membranes, etc.), first, it is important to find an adequate low resolution model for water as the most prominent solvent in biological systems. Since a water molecule is relatively small consisting of only three atoms, the contraction of one water molecule to a single site offers little efficiency [20], but the tetrahedral structure and characteristic electrostatic properties are already

lost [13]. Consequently, several multi-water CG models, e.g., the aforementioned BMW model, have been developed. For an overview please refer to Chaper II, or in greater detail to Ref [20].

CG water models can be applied in multiscale simulations [21], with different parts of the system modeled in different resolution. Of course, the parts of particular interest, e.g., the biomolecule or only its reactive center, and small solutes such as ligands are modeled in high resolution, while the surrounding (e.g., the solvent), having only an indirect, but still very important effect, can be modeled in lower resolution. Such multiscale approaches are highly promising in terms of expanding the spatiotemporal range of modeled systems, and at the same time they enable to investigate critical parts of the system in high resolution. With this in mind, it is apparent that the (bio)molecular level and the systems level come closer together and begin to connect. Hopes are, that in the near future this is going to advance biology and medicine to the next level [14].

1.4. Connection with Experiment

Models and their respective potentials are developed either on the basis of quantum-mechanical calculations, on the basis of experimental data, or both. Pair-potentials between atoms are either of bonded or non-bonded nature, and always require a large set of different experiments to be formulated and parametrized.

Once this cumbersome task is completed and the system can be simulated, the respective model can be validated by comparison to another set of experiments [4]; usually they are of a more collective nature. Examples are hydrodynamics with the viscosity as its central quantity, testing for the fluidity of the system, or dielectric spectroscopy, which probes the charge anisotropy of the system. Its frequency-dependence reveals diverse

processes of collective nature in time and is a highly critical measure for a simple model. Moreover, a comparison of simulation to experimental data in terms of collective observables is independent of certain details of the model, and takes into account mostly emergent, collective phenomena. Consequently, this way CG models can simultaneously be compared to experiment and to AA models in a meaningful way.

"Eventually, if the model is a good one, the simulator hopes to offer insights to the experimentalist, and assist in the interpretation of new results." [1] This is important, because the experimental observables are often a superposition of different components and are difficult or impossible to interpret. While the experimentalist is stuck here, the simulator has the trajectories at hand, which can be used to analyze the system in great depth and from different perspectives. For example, the translational and rotational contributions in a dielectric spectrum [22–24] or in solvation dynamics [25] can be identified. The bottom line is "computer simulation provides a direct route from the microscopic details of a system [...] to macroscopic properties of experimental interest." [1]

Chapter 2 analyzes three different water models of different resolution: BMW, a CG water model, and SPC/E and TIP3P, which are both AA water models. Transport and dielectric properties of the water models are assessed by comparing them to experimental measurements and among themselves. In the course of that, a number of theoretical and methodical insights are delivered. For example, it is shown that the viscosity can consistently and accurately be calculated from single particle dynamics only. Furthermore, system dynamics is shown to depend on the size of the time step Δt in a systematic way.

The manuscript provided in Chapter 2 was submitted to *The Journal of Chemical Physics* for publication.

Bibliography

[1] M. P. Allen and D. J. Tildesley, *Computer simulations of liquids*, Oxford Press, New York, 1989.

[2] J. A. McCammon, B. R. Gelin, and M. Karplus, *Nature* **267**, 585 (1977).

[3] B. J. Alder and T. W. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957).

[4] A. Rahman, *Phys. Rev.* **136**, A405 (1964).

[5] A. Rahman and F. H. Stillinger, *J. Chem. Phys.* **55**, 3336 (1971).

[6] C. Störmer, *Arch. Sci. Phys. Nat.* **24**, 221 (1907).

[7] L. Verlet, *Phys. Rev.* **159**, 98 (1967).

[8] R. W. Hockney, *Meth. Comput. Phys.* **9**, 136 (1970).

[9] M. A. Cuendet and W. F. van Gunsteren, *J. Chem. Phys.* **127**, 184102 (2007).

[10] S. J. Marrink, X. Periole, D. P. Tieleman, and A. H. de Vries, *Phys. Chem. Chem. Phys.* **12**, 2254 (2010).

[11] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Williams, *J. Chem. Phys.* **76**, 637 (1982).

[12] M. Griebel, S. Knapek, and G. Zumbusch, Springer, Berlin, Heidelberg, New York, 2007.

[13] W. G. Noid, *J. Chem. Phys.* **139**, 090901 (2013).

[14] M. Gruebele and D. Thirumalai, *J. Chem. Phys.* **139**, 121701 (2013).

[15] M. Deserno, *Macromol. Rapid. Commun.* **30**, 752 (2009).

[16] B. A. Merchant and J. D. Madura, *Annu. Rev. Comput. Chem.* **7**, 67 (2011).

[17] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein, *J. Phys. Condens. Matter* **16**, R481 (2004).

[18] T. A. Halgren and W. Damm, *Curr. Opin. Struct. Biol.* **11**, 236 (2001).

[19] Z. Wu, Q. Cui, and A. Yethiraj, *J. Phys. Chem. B* **114**, 10524 (2010).

[20] K. R. Hadley and C. McCabe, *Mol. Sim.* **38**, 671 (2012).

[21] Z. Wu, Q. Cui, and A. Yethiraj, *J. Phys. Chem. Lett.* **2**, 1794 (2011).

[22] C. Schröder, J. Hunger, A. Stoppa, R. Buchner, and O. Steinhauser, *J. Chem. Phys.* **129**, 184501 (2008).

[23] C. Schröder and O. Steinhauser, *J. Chem. Phys.* **132**, 244109 (2010).

[24] C. Schröder, *J. Chem. Phys.* **135**, 024502 (2011).

[25] M. Schmollngruber, C. Schröder, and O. Steinhauser, *J. Chem. Phys.* **138**, 204504 (2013).

Transport and Dielectric Properties of Water and the Influence of Coarse-Graining: Comparing BMW, SPC/E and TIP3P models

Daniel Braun, Stefan Boresch, and Othmar Steinhauser
University of Vienna, Department of Computational Biological Chemistry, Währingerstrasse 17, A-1090 Vienna, Austria

Submitted to *The Journal of Chemical Physics*

Long-term molecular dynamics simulations are used to compare the single particle dipole reorientation time, the diffusion constant, the viscosity and the frequency-dependent dielectric permittivity of the coarse-grained BMW water model to two common atomistic three-point water models, SPC/E and TIP3P. In particular, the agreement between the calculated viscosity of BMW and the experimental viscosity of water is satisfactory. We also discuss contradictory values for the static dielectric properties reported in the literature. Employing molecular hydrodynamics, we show that the viscosity can be computed from single particle dynamics, circumventing the slow convergence of the standard approaches. Furthermore, our data indicates that the Kivelson relation connecting single particle and collective reorientation time holds true for all systems investigated.

Since simulations with coarse-grained force fields often employ extremely large time steps, we also investigate the influence of time step on dynamical properties. We observe a systematic acceleration of system dynamics when increasing the time step. Carefully monitoring energy/temperature conservation is found to be a sufficient criterion for the reliable calculation of dynamical properties. By contrast, criteria based on the ratio of fluctuations of total vs. kinetic energy are not sensitive enough.

I. INTRODUCTION

For over half a century molecular dynamics (MD) simulations have provided numerous valuable insights into structure and dynamics of complex systems at molecular resolution. Thanks to the ever-increasing speed of computers, numerous algorithmic developments and the continued improvements of force fields it is nowadays not only possible but almost routine to apply simulations to challenging biological problems, despite the requirements in terms of spatial and temporal scales of such calculations. Biomolecular processes occur in aqueous solution, and solute–solvent interactions are extremely important and must not be disregarded. While the solvent can be modeled implicitly with good success in many cases, the averaging underlying implicit solvent models cannot correctly reproduce certain intra- and intermolecular processes in biomolecules.[1–5] Examples are packing, the hydrophobic effect, or electrostatic steering.[6–8] Use of explicit solvent, in which each water molecule is represented by three, four or more interaction sites, on the other hand, demands high computational resources, significantly larger than those required for just simulating the biomolecular solute. To overcome this dilemma, a number of so-called coarse-grained (CG) force fields and water models have been developed. While still a particle based description of interactions, the lower resolution of CG models provides a suitable compromise between physical detail and computational efficiency in many cases.

The most widely used force fields for biomolecular simulations employ an all-atomistic (AA) representation of interactions, i.e., each atom is an interaction site. In CG descriptions multiple atoms are merged into one interaction site. The number of atoms incorporated into such a site defines the

level of coarse-graining.[9] Depending on the model a single site can represent a functional group, or even a complete amino acid side chain. While the computational complexity $O(N \ln N)$ remains unchanged, the computational cost is lowered since in a CG representation N is the number of interaction sites as opposed to the number of atoms in the AA case. In both CG and AA representations the complex nature of physical interactions is replaced by a set of effective potentials, limiting their applicability to the physical conditions they have been designed for. "In this respect AA and CG models are on the same footing".[10]

In addition to the reduced number of interaction sites, computational efficiency in CG simulations is enhanced further since larger time steps can be used. In MD simulations the time evolution of coordinates $\vec{r}(t)$ and velocities $\vec{v}(t)$ is approximated by discrete integration algorithms based on a truncated Taylor series of the Newtonian equations of motion. Consider, e.g., the leapfrog algorithm:[11]

$$\vec{v}(t_n + \frac{\Delta t}{2}) = \vec{v}(t_n - \frac{\Delta t}{2}) + \frac{\Delta t}{m}\vec{F}(t_n) + \mathcal{O}(\Delta t^3) \tag{1}$$

$$\vec{r}(t_n + \Delta t) = \vec{r}(t_n) + \vec{v}(t_n + \frac{\Delta t}{2})\Delta t + \mathcal{O}(\Delta t^3) \tag{2}$$

Here t_n denotes the time after n integration steps, and $\vec{F}(t)$ is the force acting on the particle at time t. The numerical stability of the leapfrog integrator depends on the suitable choice of the time step Δt. However, the multiplicative factor of $\vec{F}(t)$ in Eq. 1 is not just Δt, but $\Delta t/m$. The mass m of CG sites is larger than that of atomic sites in AA force fields. Because of the proportionality to $1/m$, time steps in CG simulations can be correspondingly larger than in the AA case.

This paper is concerned with coarse-grained descriptions of water. Reducing a water molecule to one or two interaction sites saves only marginal simulation time.[12] Instead, a number of multi-water CG models were developed. Probably the simplest one is the MARTINI model,[13] which consists of a single Lennard-Jones (LJ) site and treats electrostatics as if the solute was immersed in a dielectric. The model was refined by the introduction of explicit charges to account for the electrostatic coupling to a solute.[14] Several other models with explicit electrostatics,[12,15,16] such as the "big-multipole-water" (BMW) model,[17] followed. BMW mimics a cluster of four water molecules and was parametrized with respect to the density, dielectric constant, isothermal compressibility, surface tension, and potential at the air-water interface. Analogous to the most widely used three-point AA water models SPC/E[18,19] and TIP3P[20,21], the three charged sites of the BMW model are located on a rigid isosceles triangle. An important difference concerns the functional form of the repulsive term of the van der Waals interactions, which is a superposition of a LJ potential and the softer Born-Mayer-Huggins potential. For BMW good static properties, such as density or dielectric constant, were reported,[17,22] which suggests its utility in biomolecular simulations, in particular multiscale approaches with the biomolecular solute in full atomistic detail.[23] However, if one is interested in dynamic properties of a solute, one has to know the dynamical properties of the solvent; ideally, they are in reasonable agreement with experiment. Therefore, a primary goal of this work is the calculation of transport properties (e.g., viscosity η) and the frequency-dependent dielectric permittivity $\varepsilon(\omega)$ of BMW. The results are not only compared to experimental data, but also to the respective values of SPC/E

and TIP3P. Particularly for TIP3P, widely differing values for the dielectric constant can be found in the literature;[17,24–26] in order to address these discrepancies, we recomputed the transport and dielectric properties also for the atomistic water models rather than relying on literature data.

There are different guiding principles for developing and assessing CG force fields, depending on the scope of application.[9] For CG water models mapping multiple molecules into a super-molecule, as is the case for BMW, the comparison of single particle properties, e.g., rotational dynamics, to other CG or AA models, as well as experimental data is difficult; obviously these quantities will depend on how many waters the super-molecule represents. Collective properties, such as the viscosity η or the frequency-dependent dielectric permittivity, on the other hand, should be practically invariant to the details of the model. They describe the system at a mesoscopic level suitable for the comparison to experiment, and are independent of the level of graining of the underlying model. Furthermore, the two theories complement each other well: hydrodynamics is sensitive to size, shape and interaction of molecules; the viscosity η as its central quantity is a measure for the complete dynamics of a system. Dielectric theory provides a means to describe charge distribution and polarity on a collective scale and also as a function of time. Unfortunately, the conventional methods for computing the viscosity η converge slowly and tend to be rather imprecise. However, we show that for the systems at hand η can be obtained consistently and with good precision by application of molecular hydrodynamics.[27]

While in CG simulations larger time steps are permissible compared to AA representations, many workers employ time steps that far exceed what is jus-

tified by the higher masses of CG beads and the ratio $\Delta t/m$ (see Eq. (1)).[28,29] Pushing the time step beyond what is considered safe is done in AA simulations as well,[30,31] though less frequently. Choosing a suitable time step is a balancing act. Too large a time step results in unstable simulations; too small time steps are computationally inefficient. In connection with CG simulations, concerns were raised recently that too large Δt may result in artifacts. The discussion focused on energy conservation and structural properties,[32,33] in line with most of the existing literature on this topic. E.g., Berendsen et al.[34] analyzed energy conservation and performance of different integration algorithms as a function of Δt. By contrast, we are not aware of studies investigating the influence of time step on dynamic properties, such as η or $\varepsilon(\omega)$, with the notable exception of an early study by Fincham,[35] who included the diffusion constant in his analysis. Given our focus on transport and frequency-dependent dielectric properties, we decided to study systematically the influence of the time step on system dynamics, in addition to standard criteria such as conservation of energy and of average temperature.

II. THEORY

We start by describing the theoretical basis of the connection between single particle and collective dynamics. First, we introduce relevant hydrodynamic relations (Sect. II A), which are then adapted for the application at the molecular level (Sect. II B), showing the connection between single particle dynamics and hydrodynamics. We end this section with an overview of computer adapted dielectric theory, outline how to compute the frequency-dependent dielectric permittivity, and show how the single particle dipolar reorientation time is related to the collective dipolar reorientation time (Sect. II C).

A. Hydrodynamic background

The motion of a molecule i without intramolecular interactions is described by that of a rigid body:

$$m_i \frac{d\vec{v}_i}{dt} = \vec{F}_i \tag{3}$$

$$I_i \frac{d\vec{\omega}_i}{dt} = \vec{N}_i \tag{4}$$

where m_i and I_i are the total mass and the moment of inertia of molecule i; $\vec{v}_i$ and $\vec{\omega}_i$ are its translational and angular velocities, and $\vec{F}_i$ and $\vec{N}_i$ are the net force and torque acting on the molecule. While an approximation for molecular liquids in general, Eqs. 3 and 4 apply exactly to the present case, since all three systems studied, TIP3P, SPC/E and BMW, are rigid water models.

In Brownian dynamics the trajectory of a molecule i is modeled as diffusion through a viscous medium of the surrounding molecules j. The friction tensor

$\overleftrightarrow{\xi}_{ij}$ represents the drag forces and torques, which act from the medium on $\vec{v}_i$ and $\vec{\omega}_i$. The resulting loss in energy is compensated by random forces $\vec{F}_i'$ and torques $\vec{N}_i'$. Newtonian equations of motion (Eq. (3) and (4)) are replaced by Langevin equations

$$m_i \frac{d\vec{v}_i}{dt} = -\sum_j \left(\overleftrightarrow{\xi}_{ij}^{VV} \cdot \vec{v}_j + \overleftrightarrow{\xi}_{ij}^{V\Omega} \cdot \vec{\omega}_j \right) + \vec{F}_i'(t) \tag{5}$$

$$I_i \frac{d\vec{\omega}_i}{dt} = -\sum_j \left(\overleftrightarrow{\xi}_{ij}^{\Omega V} \cdot \vec{v}_j + \overleftrightarrow{\xi}_{ij}^{\Omega\Omega} \cdot \vec{\omega}_j \right) + \vec{N}_i'(t) \tag{6}$$

where V and Ω refer to translational and rotational motion, respectively. As shown by Dahler et al.[36] the friction tensors $\overleftrightarrow{\xi}_{ij}$ are related to diffusion tensors $\overleftrightarrow{D}_{ij}$ by a generalized Einstein relation

$$\overleftrightarrow{D}_{ij}^{XY} = k_B T \cdot \left(\overleftrightarrow{\xi}_{ij}^{XY} \right)^{-1} \tag{7}$$

with T being the temperature and k_B the Boltzmann constant, and $X, Y \in \{V, \Omega\}$.

Expressing the drag forces and torques on the right hand side of Eqs. (5) and (6) by their hydrodynamic analogue involving the Oseen tensor and its derivatives, Wolynes and Deutch[37] have given explicit representations of Eq. (7):

$$\overleftrightarrow{D}_{ij}^{VV} = k_B T \cdot \overleftrightarrow{T}_{ij}^{VV} \tag{8}$$

$$\overleftrightarrow{D}_{ij}^{V\Omega} = k_B T \cdot (1 - \delta_{ij}) \, \overleftrightarrow{T}_{ij}^{V\Omega} = \overleftrightarrow{D}_{ij}^{\Omega V} \tag{9}$$

$$\overleftrightarrow{D}_{ij}^{\Omega\Omega} = k_B T \cdot \overleftrightarrow{T}_{ij}^{\Omega\Omega} \tag{10}$$

$\overleftrightarrow{T}_{ij}$ is a function of the mutual distance vector $\vec{r}_{ij} = \vec{r}_j - \vec{r}_i$ between two points in the medium, i.e., the index ij is a short hand notation for $(\vec{r}_{ij})$.

Consequently, these relations show the dependence of $\overset{\leftrightarrow}{D}_{ij}$ on $\vec{r}_{ij}$, mediated by the aforementioned Oseen tensor and its derivatives, describing hydrodynamic interactions. The strength of interaction is represented by one central quantity, i.e., the viscosity η:

$$\overset{\leftrightarrow}{T}_{ij}^{VV} = \frac{1}{8\pi\eta}\frac{1}{r_{ij}}\Big(\frac{\vec{r}_{ij}:\vec{r}_{ij}}{r_{ij}^2} + \overset{\leftrightarrow}{I}\Big) \tag{11}$$

$$\overset{\leftrightarrow}{T}_{ij}^{\Omega V} = \frac{1}{2}\vec{\nabla}\times \overset{\leftrightarrow}{T}_{ij}^{VV} = \overset{\leftrightarrow}{T}_{ij}^{V\Omega} \tag{12}$$

$$\overset{\leftrightarrow}{T}_{ij}^{\Omega\Omega} = \frac{1}{2}\vec{\nabla}\times \overset{\leftrightarrow}{T}_{ij}^{V\Omega} \tag{13}$$

The colon stands for the dyadic product, and $\vec{\nabla}$ is the Nabla operator. These tensors are Green functions of the linearized Navier-Stokes equation, i.e., the Stokes equation.

The mathematics of hydrodynamics and electrostatics are highly similar,[38] with the viscosity η being the analogue of the dielectric constant $\varepsilon(0)$. Further, $\overset{\leftrightarrow}{T}_{ij}^{\Omega\Omega}$ of Eq. 13 can be written

$$\overset{\leftrightarrow}{T}_{ij}^{\Omega\Omega} = \frac{1}{16\pi\eta}\,\overset{\leftrightarrow}{T}_{ij}^{\vec{\mu}\vec{\mu}}, \tag{14}$$

where $\overset{\leftrightarrow}{T}_{ij}^{\vec{\mu}\vec{\mu}}$ is the dipole-dipole tensor of electrostatics,

$$\overset{\leftrightarrow}{T}_{ij}^{\vec{\mu}\vec{\mu}} = \frac{1}{r_{ij}^3}\Big(\frac{3\vec{r}_{ij}:\vec{r}_{ij}}{r_{ij}^2} - \overset{\leftrightarrow}{I}\Big), \tag{15}$$

and $\vec{\mu}$ denotes the molecular dipole moment.

B. Molecular Hydrodynamics

Alder et al.[39] showed that the application of hydrodynamics at the molecular level is feasible; i.e., single particle dynamics of a molecule can be con-

sistently described by hydrodynamics.[27] Based on the summary just presented, we now derive a key relation used throughout the remainder of this manuscript.

A molecule is treated as spherical Brownian particle, and only self terms $i = j$ are considered. Formally, this corresponds to the limit $r_{ij} = 0$. Practically, the particle is spatially extended, and the interaction is calculated for a distance $r_{ij} = a$ (the hydrodynamic radius) from the rotationally averaged diffusion tensors:

$$\overleftrightarrow{D}_{ii}^{VV} = \frac{k_B T}{6\pi\eta a} \overleftrightarrow{I} \tag{16}$$

$$\overleftrightarrow{D}_{ii}^{\Omega\Omega} = \frac{k_B T}{8\pi\eta a^3} \overleftrightarrow{I} \tag{17}$$

With interaction uniform in every direction, $\overleftrightarrow{D}_{ii}^{VV}$ reduces to a scalar D, the translational self-diffusion coefficient.

$$D = \frac{k_B T}{6\pi a\eta}. \tag{18}$$

The rotational diffusion equation relates $\overleftrightarrow{D}_{ii}^{\Omega\Omega}$ to the single particle reorientation time τ_l

$$\frac{1}{\tau_l} = l(l+1)\frac{k_B T}{8\pi\eta a^3}. \tag{19}$$

For the reorientation of dipoles the angular momentum $l = 1$, and we obtain

$$\tau_s = \frac{4\pi a^3}{k_B T}\eta. \tag{20}$$

In MD simulations τ_s can be obtained straightforwardly from the time correlation function (TCF) of the molecular dipole moment; and D from the mean square displacement (MSD) of the molecular center of mass, cf. Methods.

The central observation is that the viscosity η cancels from the product $D \cdot \tau_s$,

$$D \cdot \tau_s = \frac{2}{3} a^2 \tag{21}$$

By substituting a from Eq. (21) either in Eq. (18) or (20), one gets an explicit expression for the viscosity η in terms of τ_s and D,

$$\eta = \eta_{\tau_s} = \frac{k_B T}{3\pi\sqrt{6}} \frac{1}{\sqrt{D^3 \tau_s}} \tag{22}$$

Eq. 22 permits the calculation of the viscosity from single particle dynamical properties only.

Throughout this paper we denote values for the viscosity computed in this manner as η_{τ_s}, in order to distinguish them from values obtained by the standard approaches based on the pressure tensor, cf. Methods.

C. Dielectric permittivity and collective reorientation

While hydrodynamics considers the velocity field in a medium, dielectric theory explicitly probes the anisotropy of charge distribution. The latter is characterized by the collective rotational dipole moment $\vec{M}_D = \sum_{j=1}^{N} \vec{\mu}_j$, where $\vec{\mu}_j$ are the molecular dipole moments. $\vec{M}_D$ is related to the frequency-dependent dielectric permittivity $\varepsilon(\omega)$ according to

$$\varepsilon(\omega) - 1 = \frac{4\pi}{3 V k_B T} \left(\left\langle \vec{M}_D^2 \right\rangle + i\omega \mathcal{L} \left[\langle \vec{M}_D(0)\, \vec{M}_D(t) \rangle \right] \right). \tag{23}$$

As one sees from Eq. 23, the spectrum of $\varepsilon(\omega)$ consists of a real part $\varepsilon'(\omega)$, the so-called dielectric dispersion, and an imaginary part $\varepsilon''(\omega)$, the dielectric absorption. In MD simulations both contributions can be obtained from the

TCF [40–42]

$$\Phi(t) = \langle \vec{M}_D(0) \cdot \vec{M}_D(t) \rangle. \tag{24}$$

In terms of the so-called Kirkwood correlation factor G_K,

$$G_K = \frac{\langle \vec{M}_D^2 \rangle}{N\mu^2} = 1 + \frac{1}{N\mu^2} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \vec{\mu}_i \cdot \vec{\mu}_j \right\rangle \tag{25}$$

the dielectric constant $\varepsilon(\omega = 0) = \varepsilon(0)$ is given by

$$\varepsilon(0) - 1 = \frac{4\pi}{3Vk_BT} N\mu^2 G_K \tag{26}$$

The TCF of the collective dipole moment $\Phi(t)$ can also be written as a superposition of single-molecule self- and cross-correlations functions

$$\Phi(t) = \sum_{i=1}^{N} \Big(\langle \vec{\mu}_i(0)\vec{\mu}_i(t) \rangle + \sum_{j \neq i}^{N} \langle \vec{\mu}_i(0)\vec{\mu}_j(t) \rangle \Big) \tag{27}$$

with initial value $\Phi(0) = \mu^2 G_K$. Since in the condensed phase reorientation of a single molecular dipole, characterized by the TCF $\varphi(t) = \langle \vec{\mu}_i(0)\vec{\mu}_i(t) \rangle$, strongly depends on interactions with neighboring dipoles, there must be a relationship between $\varphi(t)$ and $\Phi(t)$. For the mono-exponential case $\Phi(t) = \exp(-t/\tau_D)$, Kivelson et al.[43,44] derived a linear relationship between the collective reorientation time τ_D of $\Phi(t)$ and the single particle reorientation time τ_s of $\varphi(t)$,

$$\tau_D = \tau_s \cdot \frac{G_K}{J_K} \tag{28}$$

In Ref. 44 the so-called dynamical Kirkwood correlation factor J_K was left undetermined. Two years later Wolynes and Deutch[37] developed in a very thorough and elucidating paper that an approximate expression for J_K in static terms is given by

$$J_K \propto \int_0^\infty dr\ 4\pi\rho r^2 \left\langle \vec{\mu}_i \overleftrightarrow{T}_{ij}^{\vec{\mu}\vec{\mu}} \vec{\mu}_j \right\rangle, \tag{29}$$

where $\overleftrightarrow{T}_{ij}^{\vec{\mu}\vec{\mu}}$ denotes the dipole-dipole tensor already defined in Eq. (15). Eq. 29 represents the angle-averaged energy between pairs of molecular dipoles, separated by a distance r and integrated over spherical shells. For light scattering a rigorous expression for J_K was deduced in Ref. 45, which for dielectric relaxation takes the special form

$$J_K = 1 + \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \frac{\int_0^\infty dt \langle \dot{\vec{\mu}}_i(0) \dot{\vec{\mu}}_j(t) \rangle}{\int_0^\infty dt \langle \dot{\vec{\mu}}_i(0) \dot{\vec{\mu}}_i(t) \rangle} \tag{30}$$

III. METHODS

A. Simulations

All simulations were carried out with version c38b1 of CHARMM.[46] The code was extended locally to handle the superposition of a Born-Mayer-Huggins potential and a Lennard-Jones potential needed for BMW.[17] All simulation systems represented 2100 waters; i.e., the BMW systems consisted of 525 "super-molecules". Some details of the simulation systems are summarized in Table I. Each system was equilibrated by simulations under constant pressure and temperature conditions (CPT) for 4 ns for BMW and for 2 ns for the two AA models. No significant dependence of the average density on the time step was found during the CPT simulations; see also Results. Therefore, the average side length of the cubic simulation box found in the second half of the CPT equilibration was used in the constant volume, constant temperature (NVT) production simulations; they are listed in Table I. Electrostatic interactions were calculated using the particle mesh Ewald (PME) method[47] with the Ewald κ parameter set to 0.41 and a 48×48×48 grid for the reciprocal space interactions. The cut-off for the real space electrostatic interactions was 12 Å for TIP3P and SPC/E, and 14 Å for BMW. Van der Waals interactions were switched off smoothly between 10 and 12 Å for the AA models and between 10 and 14 Å for the CG model.

Integration of the equation of motions was done via the leapfrog scheme,[11] which is the default in CHARMM. Since concerns regarding the correct calculation of velocities within the leapfrog scheme for large time steps were raised recently,[33,48] selected results were verified using a velocity-Verlet integrator[49]

TABLE I. Overview of selected simulation parameters

	BMW[17]	SPC/E[18,19]	TIP3P[20,21]
Type	CG	AA	AA
n_{mol}[a]	525	2100	2100
L/Å[b]	39.225	39.807	39.635
Δt/fs[c]	2, 5, 10, 20, 25	1, 2, 3, 4, 5	1, 2, 3, 4, 5

[a] Number of particles. Note that each BMW particle represents four water molecules.
[b] Side length of cubic simulation box used in NVT production simulations.
[c] Time Steps studied.

(VV2 integrator in CHARMM). In all simulations the molecules were held rigid using SHAKE.[50] The average temperature was held at 300 K by a Nosé-Hoover thermostat[51,52] with a coupling constant of 500 kcal mol^{-1} ps^2.

To investigate the dependence of various dynamical properties on the time step Δt, for each water model five simulations differing only in Δt were carried out (see Table I). The length of the production simulations ranged from 250 ns to 1 μs for BMW, and from 30 ns to 100 ns for the two AA models studied. To monitor energy conservation as a function of Δt, we also carried out additional microcanonical (NVE) simulations of at least 30 ns length.

B. Analysis

This work focused on the dynamical properties of the BMW, SPC/E and TIP3P water models. To characterize translational and rotational single particle dynamics, we studied the diffusion constant D and single particle dipolar

relaxation. Collective dynamical properties of the water models were compared by computing the viscosity η and the frequency dependent dielectric constant $\varepsilon(\omega)$.

The diffusion constant was derived from the mean square displacement (MSD) of the molecular center of mass using the Einstein relation[53]

$$6\,D\,t = \langle |\vec{r}_i(t) - \vec{r}_i(0)|^2 \rangle \tag{31}$$

where $\vec{r}_i$ denotes the center of mass position of water molecule i. The MSD raw data were fitted to a linear function in the usual manner. MSDs were computed for 1 ns; the data in the range 5–200 ps were used for the linear fit.

Single particle dipolar relaxation was studied by means of the normalized autocorrelation function

$$\varphi(t) = \frac{\langle \vec{\mu}(t) \cdot \vec{\mu}(0) \rangle}{\langle \vec{\mu}(0) \cdot \vec{\mu}(0) \rangle} \tag{32}$$

The resulting raw data were fitted to a Kohlrausch-Williams-Watts function, i.e., $\varphi(t) \propto \exp[-(t/\tau)^\beta]$. The value for β was approximately 0.8 in all cases. Stretched exponential fit functions of this type can be characterized by the mean reorientation time τ_s, defined by

$$\tau_s = \int_0^\infty dt\, \varphi(t) \;\approx\; \int_0^\infty dt\, e^{-(t/\tau)^\beta} = \frac{\tau}{\beta}\,\Gamma\left(\frac{1}{\beta}\right), \tag{33}$$

where $\Gamma(x)$ is the gamma function.

All direct approaches to compute η from simulation data are based on the off-diagonal elements of the pressure tensor $\overleftrightarrow{P}$, i.e.,

$$P_{\alpha\beta}(t) = \frac{1}{V} \cdot \sum_{j=1}^{N} \Big(m_j\, \vec{v}_j^{\,\alpha}(t) \cdot \vec{v}_j^{\,\beta}(t) - \vec{r}_j^{\,\alpha}(t) \cdot \vec{F}_j^{\,\beta}(t) \Big) \tag{34}$$

with $\alpha, \beta \in \{x, y, z\}$, $\alpha \neq \beta$. In the Green-Kubo approach[53] the slowly decaying statistical noise in the long time tails of the TCFs leads to inaccurate results. The Einstein-Helfand relation is, therefore, the preferred approach.[24,54,55] $\Delta G_{\alpha\beta}(t)$, the Helfand moment,[56] is the time integral of $P_{\alpha\beta}(t)$. Under periodic boundary conditions, η is the slope of the MSD of $\Delta G_{\alpha\beta}(t)$

$$\eta_P^{\leftrightarrow} \cdot t = \lim_{t\to\infty} \frac{V}{2\,k_B T} \left\langle \sum_{\alpha\beta} \Delta G_{\alpha\beta}(t)^2 \right\rangle \tag{35}$$

We write $\eta_P^{\leftrightarrow}$ to indicate the use of the Einstein-Helfand relation, and to distinguish it from viscosities calculated from molecular hydrodynamic theory, i.e., η_{τ_s} of Eq. (22). The MSDs $\Delta G_{\alpha\beta}(t)$ were always calculated based on all available simulation data, but only the first 0.1% of the raw MSDs were used for the linear fit.

As outlined in Sect. II C, all dynamical dielectric properties of interest rely on the TCF $\Phi(t)$ (Eq. (24)). In contrast to its single particle counterpart $\varphi(t)$, $\Phi(t)$ could always be fitted to a simple exponential function, i.e.,

$$\Phi(t) \propto A \exp[-(t/\tau_D)], \tag{36}$$

and the collective dipole reorientation time τ_D was obtained as a fit parameter, which corresponds to the time integral of the normalized $\Phi(t)$ from 0 to ∞.

IV. RESULTS

All results are summarized in Table II, which will be referenced throughout this section. It contains both data for all three water models studied, as well as the their dependence on the time step Δt used. The table starts with the average density ρ observed in the initial CPT simulations, which determined the size of the simulation box used in all later NVT simulations; cf. Methods and Table I. All lines similar to that for ρ, where a single value is given for all time steps, indicates quantities where no systematic dependence on Δt could be detected.

A. Single particle dynamics

Since all three water models studied are rigid, their single particle dynamics is described exactly by molecular translation and rotation (see Eqs. (3) and (4)). Hence, the translational diffusion coefficient D (Eq. (31)) and the single particle mean reorientation time τ_s (Eq. (33)) provide an excellent means to grasp the single particle dynamics of our systems. In Table II one sees immediately that both D and τ_s strongly depend on the time step used in the simulations, regardless of the particular water model studied. Both translational and rotational motion become faster as the time step increases. Before comparing the single particle properties of BMW, SPC/E and TIP3P, we, therefore, investigate this time step dependence in some detail.

TABLE II. Simulation results for BMW, SPC/E and TIP3P.

	BMW					SPC/E					TIP3P				
Δt/fs	2	5	10	20	25	1	2	3	4	5	1	2	3	4	5
ρ[a]			1.040					0.995					1.008		
D[b]	0.72	0.75	0.79	0.99	1.14	2.60	2.64	2.75	2.86	3.11	5.69	5.79	5.93	6.11	6.25
τ_s/ps[c]	31.7	30.7	29.5	23.3	19.7	4.76	4.71	4.42	4.19	3.88	2.08	2.05	2.02	1.94	1.86
$-\langle U^* \rangle$[d]	15.35	15.34	15.30	15.14	15.01	11.20	11.18	11.15	11.10	11.04	9.79	9.78	9.76	9.73	9.69
$\eta_{\overleftrightarrow{P}}$[e]	1.7	1.7	1.6	1.3	1.1	0.7	0.7	0.7	0.7	0.6	0.3	0.3	0.3	0.3	0.3
$(D\,\tau_s)$/Å^2	2.28	2.30	2.34	2.31	2.25	1.24	1.24	1.22	1.20	1.21	1.18	1.18	1.20	1.19	1.16
a[f]			1.86					1.35					1.33		
η_{τ_s}[g]	1.62	1.57	1.52	1.19	1.01	0.63	0.63	0.59	0.56	0.52	0.29	0.29	0.28	0.27	0.26
$\varepsilon(0)$[h]	63	63	62	63	65	73	74	72	72	69	103	104	103	103	102
τ_D/ps[i]	53.3	49.9	43.2	36.8	32.7	11.1	11.0	10.1	9.6	8.5	6.2	6.3	5.9	5.8	5.5
G_K[j]	2.76	2.78	2.72	2.77	2.89	3.83	3.90	3.78	3.81	3.65	5.44	5.53	5.36	5.42	5.39
J_K[k]			1.7					1.7					1.8		

[a] Average density in g cm^{-3} found during CPT equilibration

[b] Translational diffusion coefficient in 10^{-5}cm^{2}s^{-1}, cf. Eq. (31). The experimental value for water at 300 K is $\sim 2.4 \times 10^{-5}$cm^{2}s^{-1}.[57]

[c] Eq. (33). The statistical error is < 1 in the least significant digit shown. The experimental value is $2 - 7.5$ ps.[58]

[d] Mean potential energy per molecule in kcal/mol.

[e] Viscosity in Pa·s calculated from the pressure tensor (Eq. (35)). The experimental value for water at 300 K is ~ 0.9 Pa·s.[59]

[f] Hydrodynamic radius in Å, Eq. (21)

[g] Viscosity in Pa·s calculated via molecular hydrodynamics (Eq. (20)).

[h] Static dielectric constant. The statistical error is ± 2. The experimental value for water at 25°C is 78.36 (H. Weingärtner, private communication).

[i] Eq. (36); the experimental value for water at 25°C is 8.27 ps (H. Weingärtner, private communication).

[j] Eq. (25)

[k] Eq. (28)

1. Time step dependence

The results for D and τ_s tabulated in Table II both are based on fits. The raw data for τ_s, the single particle dipolar relaxation TCF $\varphi(t)$, can be obtained from the simulations with very high precision. For this reason, the following analysis of time step dependence focuses on $\varphi(t)$.

TCF $\varphi(t)$ (Eq. (32)) at three time resolutions is shown in Figs. 1 and 2.

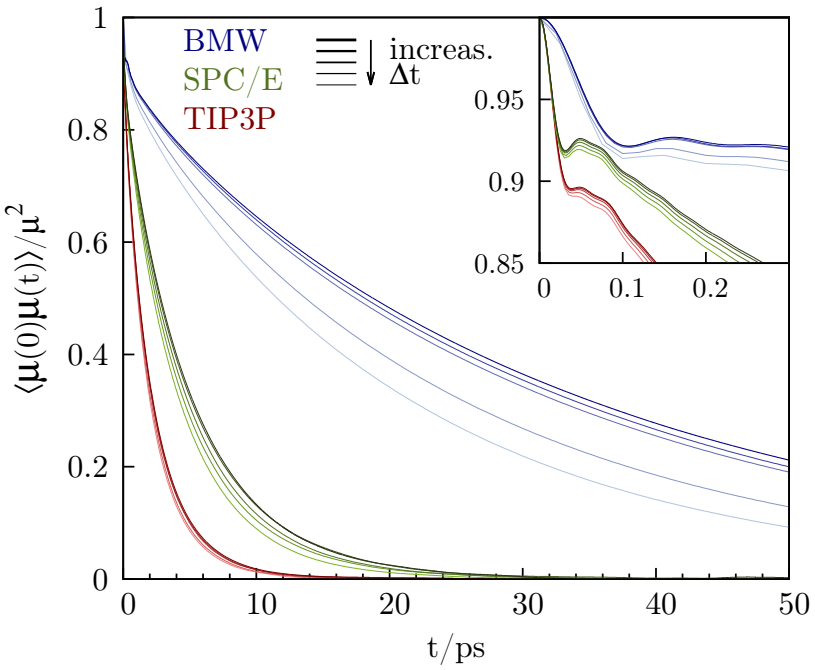

FIG. 1. Normalized reorientational autocorrelation function of the molecular dipole moment $\varphi(t)$ (Eq. (32)) for each water model and all time steps on the ps time scale. Lighter colors stand for larger time steps, cf. Table II. The insert shows the beginning of the curves, where librational motion occurs.

The main plot in Fig. 1 shows the full range of the TCF for the three water models and all time steps studied. Lighter colors indicate longer time steps; for all water models $\varphi(t)$ decays faster for longer time steps, in accord with the τ_s values reported in Table II. Dipolar relaxation of a single particle, as quantified by $\varphi(t)$, is strongly coupled to the dielectric behavior of its environment. However, for very short time scales (cf. insert of Fig. 1) $\varphi(t)$ is dominated by librational motion. In this time domain the relaxation of $\vec{\mu}(t)$ is mostly controlled by local potential energy barriers and can be considered as decoupled from the environment. Thus, it is interesting to see that even

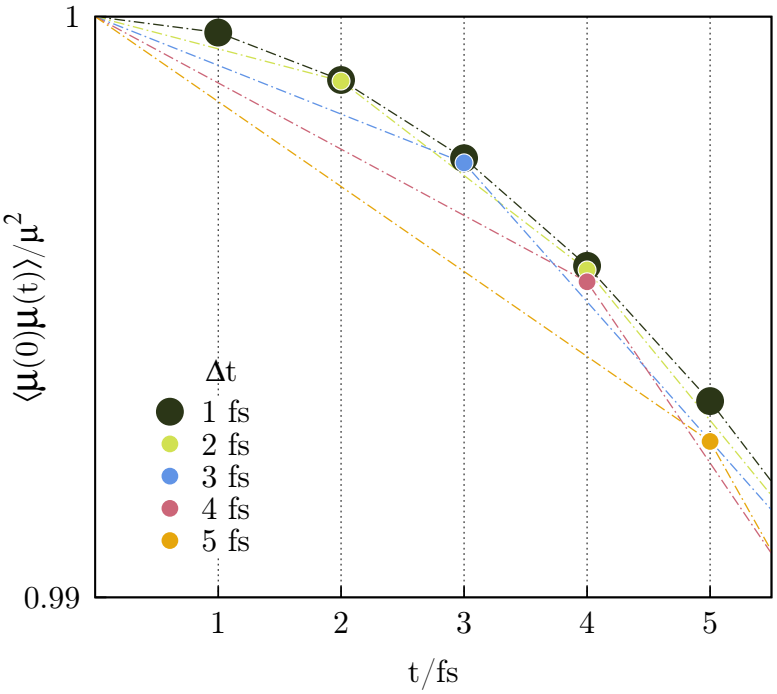

FIG. 2. Magnification of the very first time steps of $\varphi(t)$. The data shown are for SPC/E; analogous behavior was observed for TIP3P and BMW.

in the librational regime dynamics is always faster for longer time steps.

The time resolution is taken to the extreme in Fig. 2 which depicts $\varphi(t)$ for SPC/E as a function of time step Δt during the initial 5 fs. Completely analogous results were obtained for BMW and TIP3P, data not shown. Let us focus, e.g., on the values of $\varphi(t)$ for $t = 4\,\mathrm{fs}$ as obtained in four 1 fs steps (black dots), two 2 fs steps (light green dots) or a single 4 fs step (red dot). One sees that at $t = 4\,\mathrm{fs}$ the correlation for $\Delta t = 1$ fs is higher than for $\Delta t = 2$ fs, and both are higher than for $\Delta t = 4$ fs. In fact, we found that in all cases the condition

$$\langle \hat{\mu}(0)\hat{\mu}(t_n)\rangle_{\Delta t_1} > \langle \hat{\mu}(0)\hat{\mu}(t_n)\rangle_{\Delta t_2 = \Delta t_1 \cdot m} \tag{37}$$

holds for every time $t_n = \Delta t_1 \cdot n = \Delta t_2 \cdot n/m$, with $n,\ m > 1$ and $\hat{\mu} = \vec{\mu}/\,|\vec{\mu}|$. In other words, the average amount of correlation of a molecular dipole moment over a particular time interval decreases with increasing simulation time step. The scalar product $\hat{\mu}(0)\hat{\mu}(t) = \cos\alpha(t)$ describes an angular motion on the unit sphere. In terms of the angle α one can also state that

$$\langle \alpha(t_n)\rangle_{\Delta t_1} < \langle \alpha(t_n)\rangle_{\Delta t_2 = \Delta t_1 \cdot m} \tag{38}$$

with

$$\alpha(t_n) = \int_0^{t_n} dt\ \vec{\omega}\ \approx \sum_{l=1}^{n} \Delta t\ \vec{\omega}_l\,; \tag{39}$$

i.e., in a given time interval a water molecule rotates on average "further" for larger time steps $\Delta t_2 > \Delta t_1$.

The thermostat employed in NVT simulations maintains the average kinetic energy about a fixed target value; i.e., the average modulus of the angular velocity $\langle|\vec{\omega}|\rangle$ is the same regardless of the time step Δt used in the simulation. Consequently, the behavior described by Eq. (38) can only be the result

of changes of direction of $\vec{\omega}$. At larger time steps Δt the torques $\vec{N}$ (Eq. (4)) are calculated less frequently, and, therefore, the direction of $\vec{\omega}$ changes less frequently in a certain time interval. Let as before $\Delta t_2 = m\Delta t_1,\ m > 1$; then Eq. (38) and Eq. (39) can be combined to

$$\Big\langle \sum_{l=1}^{n} \sum_{k=1}^{m} \Delta t_1 \, \vec{\omega}_{l_k} \Big\rangle < \Big\langle \sum_{l=1}^{n} \Delta t_2 \, \vec{\omega}_l \Big\rangle \tag{40}$$

Eq. (40) is a triangle inequality for rotational motion and indicates that the discretized rotational motion is sub-additive. At smaller time steps, the direction of a particle's dipole moment changes more frequently as barriers are encountered. At larger time steps, some of these barriers are traversed. This suggests that at larger time steps the system will visit higher energy regions of phase space more frequently. Table II also lists the average potential energy per particle $\langle U^* \rangle$ for each water model and all time steps. Indeed, one sees that for larger time steps $\langle U^* \rangle$ becomes more positive (note that the table lists $-\langle U^* \rangle$!). Thus, simulations with larger time steps can be viewed as being more "direct", resulting in faster rotational dynamics.

So far, we have focused on $\varphi(t)$ as a function of Δt. Completely analogous dependence on the time step was observed for the center of mass MSD (Eq. (31)); data not shown. Similarly to what was just described for $\varphi(t)$, this acceleration is caused by fewer changes of translational direction per time interval. Thus, both translational and rotational single particle dynamics becomes faster for larger time steps.

Fig. 3 illustrates the dependence of τ_s on Δt for all three water models. Going from 1 fs to 2 fs for the AA models, and from 2 fs to 5 fs for the CG model, the change in τ_s is relatively small. For larger time steps, > 3 fs

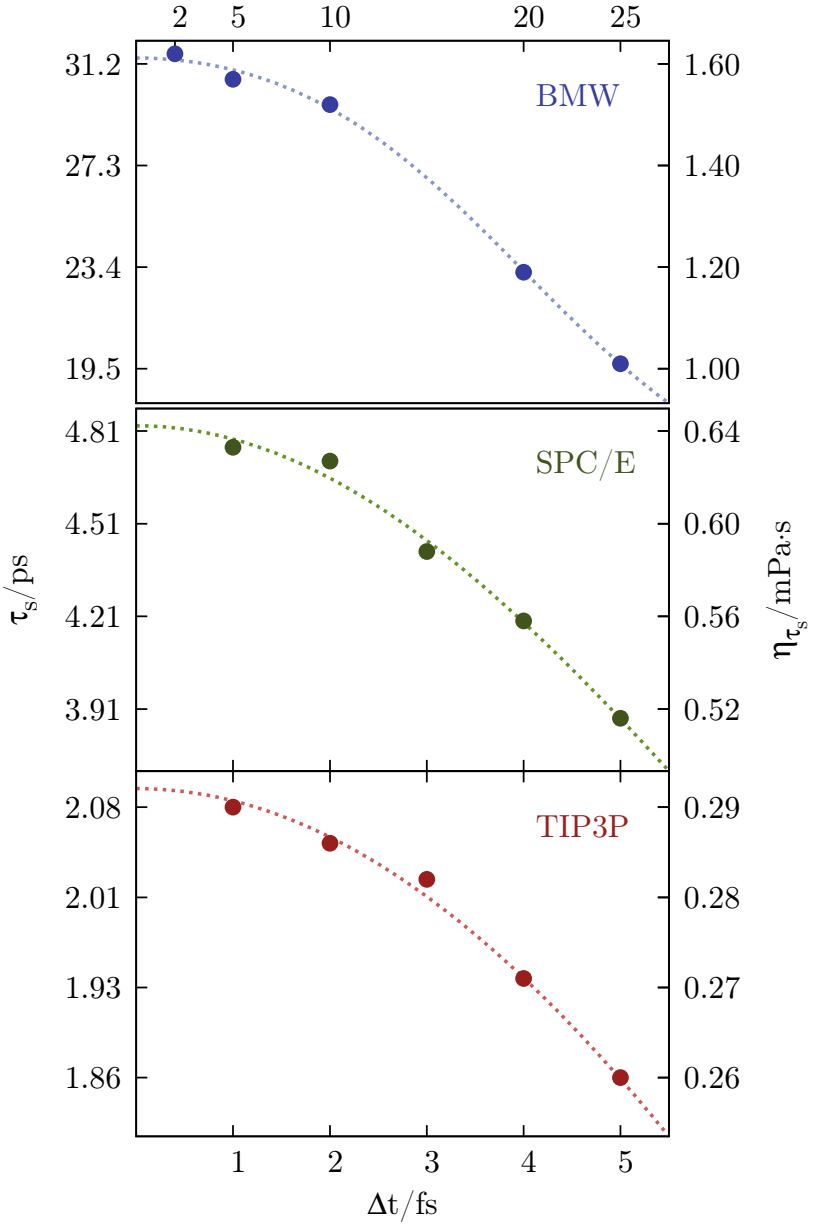

FIG. 3. Single particle reorientation time τ_s and viscosity η_{τ_s} (Eq. (20)) for each water model as a function of the time step. Note the different scaling of the abscissa for CG and AA models.

in the AA case and $>$ 10 fs in the CG case, the effect becomes more and more pronounced. Clearly, time steps of 20 fs and more should not be used in simulations with the BMW model if one is interested in dynamical properties. Since one BMW super-molecule represents four waters, its mass is also four times as high. Thus, a 10 fs time step for BMW corresponds to a 2.5 fs time step for SPC/E or TIP3P. Interestingly, 2.5 fs is commonly viewed as the largest safe time step for rigid atomistic water models; Fig. 3 supports this value with respect to dynamical properties and indicates that time steps for BMW should not exceed 10 fs if one is interested in reliable dynamical properties. The 10 fs threshold for BMW is in so far surprising as the Born-Mayer-Huggins potential of the BMW model is softer and smoother than the Lennard-Jones potentials of the AA models and could be expected to mitigate effects of larger time steps.[32] However, water–water interactions are dominated by electrostatics; in fact, van der Waals interactions account for only $\sim$ 15 % of the total energy for the AA models and $\sim$ 3 % for the CG model.

Traditionally, the choice of the largest possible time step is based on either conservation of total energy or the comparison of fluctuations of the total energy to fluctuations of the kinetic energy.[32,34,53] We, therefore, repeated a subset of simulations under NVE conditions. The fluctuations of the total energy in all these simulations are accurately proportional to $(\Delta t)^2$, indicating that no other sources of energy loss are present.[53] One rule of thumb postulates that the ratio of fluctuation of total energy to fluctuation of kinetic energy should not exceed 1/5.[32] This threshold is not reached in any of our simulations; according to this criterion the 5 fs TIP3P and 20 fs BMW simulations

are acceptable. In a different context, Elber and co-workers recently stressed the importance of monitoring energy conservation or, alternatively, conservation of temperature under NVE conditions.[60] For the systems studied here, TIP3P simulations with $\Delta t = 5$ fs and BMW simulations with $\Delta t = 20$ fs indeed showed a significant energy drift of $\sim 5\%/\mu s$ and $\sim 11\%/\mu s$, respectively. The loss of energy goes hand in hand with a rise in temperature, which was ~ 25 K/μs (TIP3P) and ~ 73 K/μs (BMW). We want to point out, that simulation lengths > 10 ns were required to detect the energy/temperature drifts; for simulation lengths of 500 ps or 1 ns everything looked well behaved. In simulations of 30 ns length with $\Delta t \leq 2$ fs (TIP3P) and $\Delta t \leq 10$ fs (BMW) neither energy, nor temperature drift could be detected.

2. Comparison of water models

The data in Table II show that D and τ_s of SPC/E are in good agreement with experiment. As documented elsewhere in the literature,[61] the single particle dynamics of TIP3P is too fast by about a factor of two. The D and τ_s values of BMW cannot be directly compared to those of the AA models as they describe translational and rotational motion of a "super-molecule" representing four waters. Nevertheless, knowing the self-diffusion constant of BMW may be useful in multiscale applications, in which, e.g., a solute and ions described by an atomistic force field is solvated by BMW, since the single particle dynamics of BMW will influence dynamical properties of the solute(s).

Once one knows D and τ_s, one can compute the hydrodynamic radius a according to Eq. (21); it is also tabulated in Table II. For SPC/E and TIP3P

one obtains very similar values of 1.35 and 1.33 Å, respectively. Using a to calculate a hydrodynamic volume $4\pi\, a^3/3$, one can compare the BMW and the two AA models at least qualitatively. The ratio of hydrodynamic volumes for, e.g., BMW and SPC/E is $(1.86/1.35)^3 = 2.62$. Given that one BMW molecule represents four waters, this value is plausible. In the neat liquid four waters will be tightly packed, so one would expect that the average volume occupied by these four waters be smaller than four times the volume of single water molecule.

B. Viscosity

The viscosity of water is $\sim$ 0.9 mPa·s.[59] Looking at the entries for $\eta_P^{\leftrightarrow}$ in Table II, one sees that the calculated value for SPC/E of $\sim$ 0.7 mPa·s agrees best with experiment, and it is in reasonable agreement with earlier simulation results.[54,62] As is well documented in the literature, the viscosity of TIP3 is too low by a factor of three;[24,25,61,63] BMW, on the other hand, is almost twice as viscous compared to real water. Given that BMW is a CG model, however, the agreement with experiment is quite acceptable.

For TIP3P and SPC/E $\eta_P^{\leftrightarrow}$ seems not to be influenced by the time step used in the simulations. For BMW, however, there is a clear dependence on time step; the value obtained with a 25 fs time step is only 65% that of the value obtained with a 2 fs time step. Despite the use of the Helfand-Einstein relation, the statistical uncertainty of the computed viscosities $\eta_P^{\leftrightarrow}$ is high; this prompted us to explore the alternative route via molecular hydrodynamics described in Sect. II B.

The check for the applicability of molecular hydrodynamics is whether the

product $D \cdot \tau_s$ is independent of time step. We just saw that with increasing time step D increases, whereas τ_s decreases. The product of the two quantities is listed in Table II explicitly for each time step studied, and for each water model the values for $D \cdot \tau_s$ lie within a narrow range. The hydrodynamic radius calculated from Eq. (21) is for all practical purposes independent of Δt; therefore, we only list a single value for each of the water models. Thus, it seems justified to use Eq. (22) to compute viscosities; as already noted we refer to values calculated in this manner as η_{τ_s}. Comparison of η_{τ_s} with the respective $\eta_P^{\leftrightarrow}$ values shown in Table II shows good agreement. Deviations are in most cases well below ~ 0.1 mPa·s which is likely less than the statistical uncertainty of $\eta_P^{\leftrightarrow}$. Given the higher precision of η_{τ_s}, a systematic time step dependence of the viscosity becomes apparent for all three water models. Because of the manner how η_{τ_s} is calculated, this dependence is identical to that for τ_s. To illustrate this behavior, Fig. 3 also includes η_{τ_s} as a function of Δt.

C. Dielectric Properties

While we so far focused exclusively on dynamical properties, we start our discussion of dielectric properties by a detailed look at the static dielectric constant $\varepsilon(0)$. Looking at Table II, one sees that for each of the water models slightly varying results are obtained for the various time steps. However, as expected for static quantities, there is no systematic trend, and the differences are well within statistical error bars. The static dielectric properties are of interest for two reasons. Particularly for TIP3P wildly differing values, ranging from 80–110, were reported.[24–26,64] In the original BMW paper[17] a

rather unusual method to determine $\varepsilon(0)$ was used.

Wu et al.[17] scaled electrostatic interactions between BMW waters by 1/1.3, as if the system were immersed in a medium with dielectric constant 1.3. During the calculation of $\varepsilon(0)$, this scaling factor was not applied. In our opinion this approach is inconsistent. We, therefore, reproduced the BMW potential by scaling the point charges of BMW by $1/\sqrt{1.3}$, *but* these reduced charges were also used for the calculation the dielectric properties. Yethiraj and co-workers computed the dielectric constant of BMW and TIP3P by a method proposed by Hess et al.[65], which entails the calculation of potentials of mean force (PMF) between ion pairs in the solvent of interest. Using this PMF based approach they found a dielectric constant of 82 for TIP3P. Based on the theoretical approach used in this work, Berendsen and co-workers[25] reported 82–94 for $\varepsilon(0)$ of TIP3P. Their results clearly depended on details of the simulation setup, such as system size, cutoff radius, and settings for the reaction field. In an early study using Ewald summation, Feller et al.[24] reported a value of $\varepsilon(0) = 110$. Our group has regularly studied dielectric properties of simple water models, such as TIP3P.[64,66] Using Ewald summation and simulations of at least 2 ns length, we have consistently obtained values of $\sim 100 \pm 5$ for the dielectric constant of TIP3P, in agreement with the results reported in Table II. Within statistical error bars, identical results were obtained for a variety of system sizes and geometries, ranging from a truncated octahedron containing 432 waters[64] to > 2000 water molecules in a cubic box (this work).

Wu et al. reported a static dielectric constant of 74 for BMW. Here, based on Eqs. (23), (25) and (26) and using the scaled charges as just described,

we find $\varepsilon(0) = 63$. While this value is in slightly worse agreement with experiment, it is still quite acceptable and, e.g., comparable to the static dielectric constant of the SPC water model.[26] Of the three water models studied here, SPC/E with $\varepsilon(0) = 73$ is in best agreement with experiment. The differences in static dielectric constants are of course also reflected in the respective Kirkwood G-Factor (cf. Eq. 25), which is also listed in Table II. Just as for $\varepsilon(0)$, the small differences in results obtained for different time steps are not systematic and just reflect the statistical noise present in the data. G_K can be further resolved into contributions from concentric spherical shells of radius r:

$$g_K(r) = \frac{1}{\mu^2} \sum_{j|r_{ij} \leq r} \langle \vec{\mu}_i \cdot \vec{\mu}_j \rangle \tag{41}$$

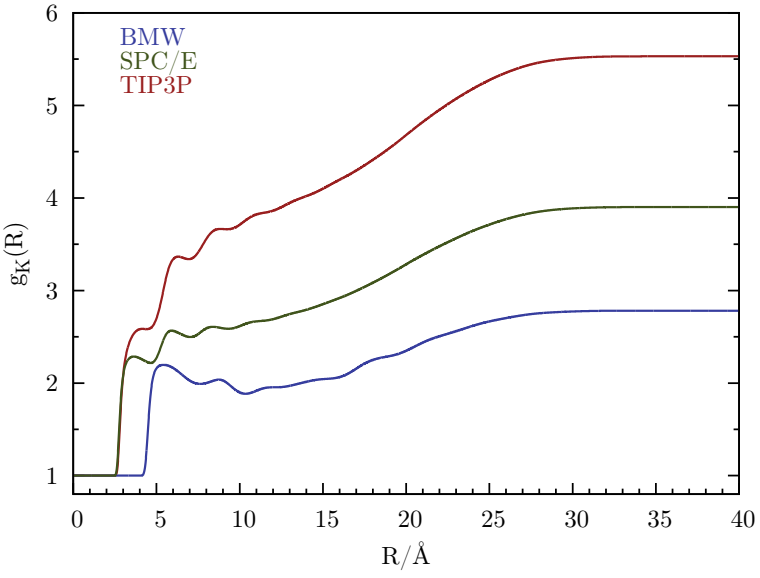

FIG. 4. $g_K(r)$ for each water model

$g_K(r)$ for BMW, SPC/E and TIP3P is shown in Fig. 4. Being a collective structural property, $g_K(r)$ is highly sensitive to type and strength of interactions, boundary conditions and other methodical details,[67] but it is invariant to the time step.

It was shown that for rigid three-point water models $g_K(r)$ depends on the geometry of the molecule, especially on the bond angle ϑ.[26] For the tetrahedral angle $\vartheta = 109.5°$ the compensation of dipolar orientations is at a maximum, and therefore, $G_K = g_K(r \to \infty)$ becomes minimal. SPC/E ($\vartheta = 109.5°$) indeed has a much lower G_K than TIP3P ($\vartheta = 104.5°$). In this respect, it is surprising to find that BMW has a lower G_K than SPC/E, although its $\vartheta = 120°$ differs significantly from the tetrahedral angle. However, because of its size, BMW, representing four waters, cannot be directly compared with AA models. In particular, one has to keep in mind that the average distance between BMW "super-molecules" is much larger. Therefore, despite higher point charges, dipolar orientation is somewhat more randomized.

In contrast to the static dielectric constant, the frequency dependent dielectric constant $\varepsilon(\omega)$ does depend on time step. This can be seen by the systematic dependence of τ_D (Eq. (36)) on Δt (see Table II), as well as in Fig. IV C which shows the dielectric permittivity, the imaginary part of $\varepsilon(\omega)$, for all three water models and time steps studied, and compares them to the experimental result (H. Weingärtner, personal communication). In all cases, larger time steps shift the curves to higher frequencies, i.e., shorter τ_D. Agreement with experiment is best for SPC/E, both in height and position of the peak.

The Kivelson-Madden relation (Eq. (28)) postulates a simple proportion-

ality between the single particle reorientation time τ_s and the reorientation time of the collective dipole moment τ_D. We find essentially the same proportionality factor J_K of 1.7–1.8 for all three models studied, cf. Table II; furthermore, J_K does not depend on the time step used in the simulations. Since τ_s is proportional to η (Eq. (20)), the Kivelson-Madden relation in turn connects η and τ_D. Thus, all dynamical properties including $\varepsilon(\omega)$ are interrelated,[69] which is reflected by the respective data in Table II.

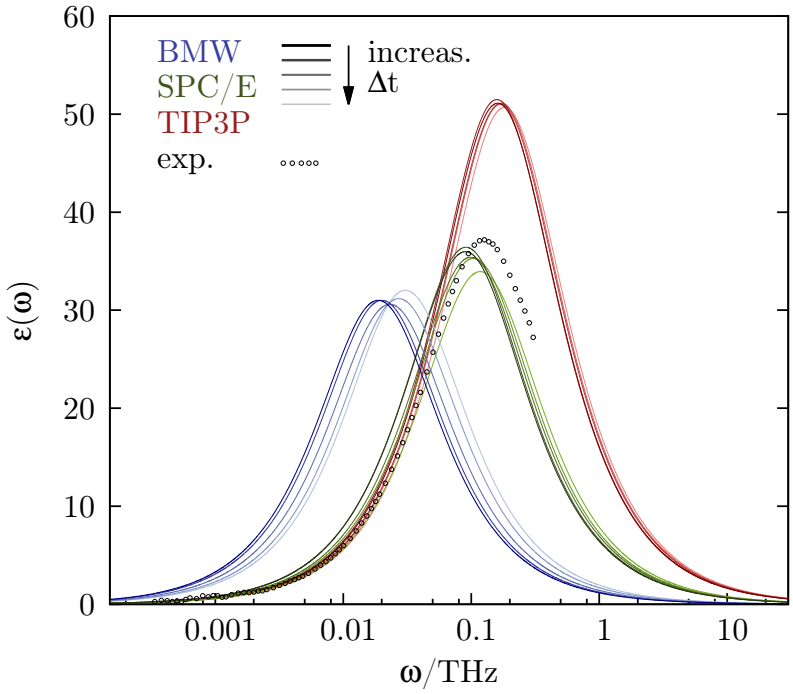

FIG. 5. Frequency-dependent dielectric permittivity $\varepsilon(\omega)$ for each water model and all time steps. The experimental data was provided by Prof. H. Weingärtner, using the method of Ref. 68

V. CONCLUDING DISCUSSION

In a liquid the translational and rotational motion of a single molecule is determined by the properties of this reference molecule and the interaction with its surrounding environment. In molecular hydrodynamics the molecular properties are reduced to a single quantity, the hydrodynamic radius a, and the interaction with the environment is described by the viscosity η. In this work we showed for AA and CG water models that translational diffusion and dipolar reorientation, as two examples of single particle dynamics, can be quantitatively characterized by a and η; see Eqs. (18) and (20). While the hydrodynamic radius was invariant to details of the simulation setup, this was not the case for the viscosity. In particular, the calculated viscosity depends on the time step used to integrate the equations of motion.

The dielectric spectrum of a liquid is determined by the relaxation of the total dipole moment $\vec{M}_D$ of the sample. Although the reorientation of individual, molecular dipoles $\vec{\mu}$ follows a Kohlrausch-William-Watts model, i.e, although the single particle correlation functions $\varphi(t)$ are stretched exponentials, the correlation function $\Phi(t)$ of their sum $\vec{M}_= \sum_i \vec{\mu}_i$ can be described perfectly by a single exponential; see Eq. (36). Moreover, the relaxation time τ_D of $\Phi(t)$ is proportional to the product of the single particle reorientation time τ_s and the static G-factor G_K. The latter is a measure of the collective order in the sample of molecular dipoles. The proportionality constant J_K is found to be more or less identical for all three water models studied. If one further takes into account that as just discussed τ_s is a function of a and η, one arrives at the observation that collective dynamics — with dielectric relaxation as a prominent example — can be quantitatively characterized by the

hydrodynamic radius, the viscosity and a static collective order parameter.

The connection between single particle and collective properties just discussed also provides an alternative route for the calculation of viscosities. Eq. (22) is strictly valid only for rigid molecules, as is the case for the three water models studied here. However, we expect that our approach will also work for flexible solutes whose average geometry fluctuates within a narrow range. Further, molecular hydrodynamics can be applied to predict and/or rationalize how viscosities change in more complicated systems, cf. Ref. 70.

This work complements the detailed analysis of properties of the BMW model of Ref. 17. While we disagree with the originally reported value for the static dielectric constant, the value $\varepsilon(0) = 63$ for BMW is quite acceptable and arguably in better agreement with experiment than the value > 100 found for TIP3P. Similarly, while BMW's viscosity is too high by almost a factor of two, one should keep in mind that the value for TIP3P is too low by almost a factor of three. Because of its rather high τ_s value, the Kivelsen relation leads to a too high value for τ_D, i.e., the frequency-dependent dielectric constant is shifted noticeably to lower frequencies. This, however, in our opinion is an acceptable price for an at least fourfold increase in computational efficiency (cf. below).

The properties of the BMW model found in Ref. 17 and this work suggest its utility in multiscale simulations, where, e.g., a protein represented by an AA force field is solvated by BMW. In this respect, the detailed knowledge of the dynamical properties of BMW is particularly important. Recently, Venable et al. showed how dynamical properties of proteins can be, or, depending on one's point of view, have to be corrected for the use of the TIP3P water

model;[61] in the context of dielectric properties of protein solutions see, e.g., Refs. 71 and 72.

As discussed in the Introduction, CG models are considered computationally more efficient for two reasons, because of (i) the reduced number of interaction sites, and because of (ii) the possibility to use larger time steps. While there can be no dispute about argument (i), the findings of this study caution against undue optimism concerning (ii). First, our results show that there is a significant gap between the maximal time steps resulting in stable simulations, i.e., simulations that do not crash, and the maximal time step resulting in reliable dynamical properties. Not too surprisingly, the latter seems to be approximately four times the maximum time step than can be used in simulations with AA models (10 vs. 2.5 fs). However, since the viscosity of BMW is almost twice that of real water, conformational sampling of solutes in BMW will be slower by about that factor,[61] cancelling already by 50% the effect of a larger, "safe" time step. Further, if the solute, e.g. a protein, is described by an AA force-field, the practical combination of AA and CG force field will necessitate the use of a multiple time step (MTS) integrator in order to benefit from larger time steps possible for the solvent. Nevertheless, the inevitable overhead of a MTS integrator will reduce computational efficiency somewhat further. With respect to future multiscale applications, we, therefore, consider the primary performance enhancements of the BMW model resulting from the number of interaction sites.

Finally, aside from one early study by Fincham,[35] we are not aware of other studies looking specifically at the time step dependence of dynamical quantities. Our finding of faster dynamics for larger time steps is commen-

surate with an analysis given recently by Rao and Spichty, who showed that simulations with large(r) time steps effectively are equivalent to simulations at slightly higher temperatures.[31] While the results per se may not be too surprising, they stringently test some widely used criteria concerning the largest "safe" time step. While recommendations concerning energy fluctuations proved to be only a necessary criterion, checking for actual energy conservation, or, alternatively, temperature conservation in NVE simulation, seems to be a sufficient criterion as well. This agrees with a point of view promoted recently by Elber and co-workers in a slightly different context.[60] The practical downside of the latter type of checks is that several nanoseconds of simulation time seem to be necesssary to detect systematic energy (temperature) drifts. However, given that routine microsecond simulations are within reach, in particular when using CG models, exploratory simulations of $10 - 50$ ns seem an acceptable price to pay.

REFERENCES

[1]M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon, J. Phys. Chem. **97**, 3591 (1993).

[2]B. Honig and A. Nicholls, Science **268**, 1144 (1995).

[3]B. Roux and T. Simonson, Biophys. Chem. **78**, 1 (1999).

[4]M. Feig, W. Im, and C. L. Brooks III, J. Chem. Phys. **120**, 903 (2004).

[5]H. S. Muddana, N. V. Sapra, A. T. Fenley, and M. K. Gilson, J. Chem. Phys. **138**, 224504 (2013).

[6]G. Pellicane, G. Smith, and L. Sarkisov, Phys. Rev. Lett. **101**, 248102 (2008).

[7]M. A. Cuendet and O. Michielin, Biophys. J. **95**, 3575 (2009).

[8]B. A. Persson and M. Lund, Phys. Chem. Chem. Phys. **11**, 8879 (2009).

[9]B. A. Merchant and J. D. Madura, Annu. Rev. Comput. Chem. **7**, 67 (2011).

[10]S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein, J. Phys. Condens. Matter **16**, R481 (2004).

[11]R. W. Hockney, Meth. Comput. Phys. **9**, 136 (1970).

[12]K. R. Hadley and C. McCabe, Mol. Sim. **38**, 671 (2012).

[13]S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, J. Phys. Chem. **111**, 7812 (2007).

[14]S. O. Yesylevskyy, L. V. Schäfer, D. Segupta, and S. J. Marrink, PLoS Comput. Biol. **6** (2010).

[15]S. Riniker and W. F. van Gunsteren, J. Chem. Phys. **134**, 084110 (2011).

[16]L. Darré, M. R. Machado, P. D. Dans, F. E. Herrera, and S. Pantano, J. Chem. Theory Comput. **6**, 3793 (2010).

[17]Z. Wu, Q. Cui, and A. Yethiraj, J. Phys. Chem. B **114**, 10524 (2010).

[18]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981) p. 331.

[19]H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[20]W. L. Jorgensen, H. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[21]E. Neria, S. Fischer, and M. Karplus, J. Chem. Phys. **105**, 1902 (1996).

[22]Z. Wu, Q. Cui, and A. Yethiraj, J. Phys. Chem. Lett. **2**, 1794 (2011).

[23]T. A. Wassenaar, H. I. Ingolfsson, M. Priess, S. J. Marrink, and L. V. Schäfer, J. Phys. Chem. B **117**, 3516 (2013).

[24]S. E. Feller, R. W. Pastor, A. Rojnuckarin, S. Bogusz, and B. R. Brooks, J. Phys. Chem. **100**, 17011 (1996).

[25]D. van der Spoel, P. J. van Maaren, and H. J. C. Berendsen, J. Chem. Phys. **108**, 10220 (1998).

[26]P. Höchtl, S. Boresch, W. Bitomsky, and O. Steinhauser, J. Chem. Phys. **109**, 4927 (1998).

[27]J. P. Boon and S. Yip, *Molecular Hydrodynamics* (Dover Publications, Inc., New York, 1991).

[28]S. J. Marrink, A. H. de Vries, and A. E. Mark, J. Phys. Chem. B **108**, 750 (2004).

[29]L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, J. Chem. Theory Comput. **4**, 819 (2008).

[30]H. Oberhofer, C. Dellago, and S. Boresch, Phys. Rev. E **75**, 061106 (2007).

[31]F. Rao and M. Spichty, J. Comput. Chem. **33**, 475 (2012).

[32]M. Winger, D. Trzesniak, R. Baron, and W. F. van Gunsteren, Phys. Chem. Chem. Phys. **11**, 1934 (2009).

[33]S. J. Marrink, X. Periole, D. P. Tieleman, and A. H. de Vries, Phys. Chem. Chem. Phys. **12**, 2254 (2010).

[34]W. F. van Gunsteren and H. J. C. Berendsen, Mol. Phys. **34**, 1311 (1977).

[35]D. Fincham, Comput. Phys. Commun. **40**, 263 (1986).

[36]D. W. Condiff and J. S. Dahler, J. Chem. Phys. **44**, 3988 (1966).

[37]P. G. Wolynes and J. M. Deutch, J. Chem. Phys. **67**, 733 (1977).

[38]S. Kim and S. J. Karrila, "Microhydrodynamics. Principles and Selected Applications," (Dover, Mineola, New York, 2005) Chap. 2.

[39]B. J. Alder and W. E. Alley, Physics Today **37**, 56 (1984).

[40]O. Steinhauser, Ber. Bunsenges. Phys. Chem. **87**, 128 (1983).

[41]M. Neumann and O. Steinhauser, Chem. Phys. Lett. **95**, 417 (1983).

[42]M. Neumann, Mol. Phys. **50**, 841 (1983).

[43]T. Keyes and D. Kivelson, J. Chem. Phys. **56**, 1057 (1972).

[44]D. Kivelson and P. Madden, Mol. Phys. **30**, 1749 (1975).

[45]T. D. Gierke, J. Chem. Phys. **65**, 3873 (1976).

[46]B. R. Brooks, C. L. Brooks III, A. D. MacKerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, J. Comput. Chem. **30**, 1545 (2009).

[47]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G.

Pedersen, J. Chem. Phys. **102**, 8577 (1995).

[48]M. A. Cuendet and W. F. van Gunsteren, J. Chem. Phys. **127**, 184102 (2007).

[49]W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Williams, J. Chem. Phys. **76**, 637 (1982).

[50]J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).

[51]S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[52]W. G. Hoover, Phys. Rev. **A31**, 1695 (1985).

[53]M. P. Allen and D. J. Tildesley, *Computer simulations of liquids* (Oxford Press, New York, 1989).

[54]P. E. Smith and W. F. van Gunsteren, Chem. Phys. Lett. **215**, 315 (1993).

[55]C. Schröder, C. Wakai, H. Weingärtner, and O. Steinhauser, J. Chem. Phys. **126**, 084511 (2007).

[56]E. Helfand, Phys. Rev. **119**, 1 (1960).

[57]R. Mills, J. Phys. Chem. **77**, 685 (1973).

[58]D. Laage and J. Hynes, Science **311**, 832 (2006).

[59]J. Kestin, M. Sokolov, and W. A. Wakeham, J. Phys. Chem. Ref. Data **7**, 941 (1978).

[60]A. P. Ruymgaart, A. E. Cardenas, and R. Elber, J. Chem. Theory Comput. **7**, 3072 (2011).

[61]R. M. Venable, E. Hatcher, O. Guvench, A. D. MacKerell, Jr., and R. W. Pastor, J. Phys. Chem. B **114**, 12501 (2010).

[62]M. A. González and J. L. F. Abascal, J. Chem. Phys. **132**, 096101 (2010).

[63]W. Yu, P. E. M. Lopes, B. Roux, and A. D. MacKerell Jr., J. Chem. Phys.

138, 34508 (2013).

[64]S. Boresch, M. Willensdorfer, and O. Steinhauser, J. Chem. Phys. **120**, 3333 (1994).

[65]B. Hess, C. Holm, and N. van der Vegt, J. Chem. Phys. **124**, 164509 (2006).

[66]P. Höchtl, S. Boresch, W. Bitomsky, and O. Steinhauser, J. Chem. Phys. **109**, 4927 (1998).

[67]S. Boresch and O. Steinhauser, Ber. Bunsenges. Phys. Chem. **101**, 1019 (1997).

[68]R. Buchner, J. Barthel, and J. Stauber, Chem. Phys. Lett. **306**, 57 (1999).

[69]C. Schröder and O. Steinhauser, J. Chem. Phys. **128**, 224503 (2008).

[70]M. Haberler and O. Steinhauser, Phys. Chem. Chem. Phys. **13**, 17994 (2011).

[71]H. Weingärtner, A. Knocks, S. Boresch, P. Höchtl, and O. Steinhauser, J. Chem. Phys. **115**, 1463 (2001).

[72]C. Schröder, T. Rudas, and O. Steinhauser, J. Chem. Phys. **125**, 244506 (2006).

A. Appendix

A.1. Curriculum Vitae

Education

since 2012	Master thesis (Prof. Othmar Steinhauser)
	"Coarse-Graining in Molecular Dynamics Simulations"
since 2011	Studies of Biological Chemistry (University of Vienna)
2007 - 2011	Studies of Molecular Biology (University of Vienna)
(2006 - 2007	Civil service)
during 2003	High school exchange program
	(St.John's preparatory school, Minnesota, USA)
1998 - 2006	Secondary school (Stiftsgymnasium Melk)

Publications

"Transport and Dielectric properties of Water and the Influence of Coarse-Graining: Comparing BMW, SPC/E and TIP3P models"
D. Braun, S. Boresch, O. Steinhauser
submitted to *The Journal of Chemical Physics*