Development of the Drude Polarizable Force Field for H_2S and QM/MM Study of Aqueous Mg^{2+} and Zn^{2+}

by

©Saleh Riahi

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Masters of Chemistry

Department of Chemistry

Memorial University of Newfoundland

May 2014

St. John's   Newfoundland

Abstract

In this work, the effect of induced polarization in molecular dynamics simulations of liquids and solutions is explored. We developed a Drude polarizable model for liquid hydrogen sulfide (H_2S). The calculated thermodynamic properties for this model are in good agreement with experiment over the temperature range 212–298 K. This model is also accurate for aqueous solutions of H_2S, including the hydration free energy and diffusion coefficient. The interfacial properties of water under $H_2S(g)$, including free energy profile and surface tension, indicate that H_2S is a powerful surfactant. To explore more sophisticated computational models, the performance of Drude model in the prediction of hydration properties of Mg^{2+} and Zn^{2+} ions was also studied. Our results show that the Drude model qualitatively describes the solvation free energy of these two ions, but QM/MM simulations are able to achieve quantitative accuracy.

Acknowledgments

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Portions of this chapter are reproduced from articles published by the author.[1;2;3]

1.1 Industrial and Biological Importance of H_2S

Hydrogen sulfide (H_2S) is of major importance in industrial chemistry, as it is a corrosive and toxic component of natural gas and crude oil.[4] The presence of H_2S complicates the extraction, processing, and transportation of these commodities and creates challenges for engineering safe, reliable, and environmentally responsible processes. Desulfurization of natural gas streams via the Claus process is an expensive and complex step in gas processing.[5] Depending on the oil and gas reservoirs, H_2S concentrations varies between 80–150 ppm and field concentrations significantly increase at higher temperatures. The Claus process can reduce the H_2S content to roughly 15 ppm, however the energetic cost of this process is significant and acheiving lower sulfur levels using this method is difficult.[6]

The solvation of hydrogen sulfide in water is of particular industrial importance. Produced water generated from oil and gas extraction often contains elevated concentrations of

H$_2$S, requiring treatment.[7] H$_2$S solvation is also important in reservoir souring, where H$_2$S is introduced to a low sulfur oil reservoir due to the bacterial reduction of sulfate anions from injected seawater into H$_2$S.[8] At low temperatures and high pressures, H$_2$S will form stable hydrates.[9;10] The interaction of water and H$_2$S at the water/vapor interface is also used industrially in the production of heavy water through the Girdler sulfide process.[11] H$_2$S has been shown to have complex solvation and interfacial properties, acting as a powerful surfactant at the water/vapor interface.[12;13]

The presence of even low levels of H$_2$S in crude oil presents serious health and safety risks to workers.[14] These safety issues stem from the acute and long-term neurotoxicity of H$_2$S. Green et al. proposed that the respiratory toxicity results from H$_2$S acting as a surfactant at the air/liquid interface.[15] Interestingly, H$_2$S has also been discovered to have an intrinsic biological role as a neurotransmitter in very low concentrations.[16] H$_2$S can readily permeate cell membranes without a facilitator.[17] Cuevasabta et al. determined that this high permeability is related to the low solubility of H$_2$S in water in comparison to aliphatic solvents, an environment that is similar to the membrane interior.[18] Therefore, H$_2$O–H$_2$S interactions are vital to understanding the biological roles of H$_2$S.

Molecular dynamics (MD) simulations are a valuable tool in developing new chemical processes to manage H$_2$S, providing valuable data that are difficult to obtain experimentally.[19;20;21;22;23;24] Moreover, molecular simulation of the physiochemical properties of H$_2$S in aqueous and hydrocarbon solutions holds promise to aid in the development of these processes and manage the toxicity of H$_2$S.

1.2   Solvation of Mg^{2+} and Zn^{2+} Ions

Molecular simulation of ion–water systems have had important contributions in understanding the thermodynamics of these solutions. The solvation of ions is an enduring fascination

2

of physical, inorganic, and analytical chemists. The interaction between an ion and the solvent is strongly dependent on the properties of the ion, such as its size, valency, polarizability, and Lewis acidity/basicity.[25;26] These interactions ultimately affect the macroscopic physical properties of the solution in the form of colligative properties, electrokinetic phenomena, and electrical conductivity.

Ions in aqueous solutions are surrounded by water molecules and interact with them through intermolecular forces. The interaction between ion–water is mainly electrostatic and it is often stronger than other intermolecular interactions occuring in the solution, such as water–water hydrogen bonding. The strength of this interaction increases with net charge of the ion. On the other hand, ions with larger ionic radii have larger ion–water distances and therefore weaker interactions. Depending on the size, charge, and Lewis acidity of an ion, the number of water molecules that directly interact with the ions changes, but many divalent ions like Mg^{2+} and Zn^{2+} have a set of 6 tightly-bound water molecules in an octahedral arrangement around the ion. These water molecules are referred to as the first coordination sphere.

The hydration of the divalent ions Mg^{2+} and Zn^{2+} have particularly interesting parallels. Although Mg^{2+} is an alkaline earth metal and Zn^{2+} is a d^{10} transition metal, they have very similar ionic radii (0.72 Å and 0.74 Å, respectively).[27] The Born model of ion solvation predicts that two ions of the same valency and radii should have the same solvation free energies, however the reported free energy of solvation of Zn^{2+} is approximately 30 kcal mol^{-1} more negative than that of Mg^{2+}.[28] This suggests that difference between these ions stems from more subtle effects. Molecular simulation techniques can reveal these effects and differences at the microscopic level.

1.3 Molecular Dynamics

Intermolecular and intramolecular interactions have a strong influence on the properties of condensed phase systems, such as solids, liquids, and solutions. Intramolecular interactions include all the covalent interactions, such as bond, bond angle, and dihedral angle interactions. Intermolecular interactions consist of electrostatic, London dispersion, and Pauli repulsion interactions.

Molecular mechanics (MM) is a popular simulation method that is used to study condensed phase systems, polymers, and biomolecules. In this method, intramolecular bond and angular potentials are approximated as harmonic oscillators, and torsional interactions are approximated as a sum of periodic functions. Electrostatic interactions are approximated as pairwise Coulombic interactions between atom-centered point charges. London dispersion and Pauli repulsive interactions are calculated by the Lennard-Jones 6–12 potential.

$$U = \sum_{\text{bonds}} K_b(r - r_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedral}} K_\chi[1 + cos(n\chi - \sigma)]$$
$$+ \sum_{\substack{\text{nbonded} \\ \text{pairs}}} \left(E_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\epsilon r_{ij}} \right). \tag{1.1}$$

In the Eqn. 1.1, r_0, θ_0, K_b, and K_θ are the equilibrium bond length and bond angles and their force constants, respectively. n, σ, and K_χ are the periodicity, offset, and barrier height of the dihedral potentials, respectively. The electrostatic parameters, q_i and q_j, are the constant point charges located at the center of atom i and j at the distance r_{ij}. The $R_{\text{min},ij}$ and $E_{\text{min},ij}$ are the position and depth of the minimum in the Lennard-Jones potential, as is illustrated in Figure 1.1.

The various parameters in Eqn. 1.1 are referred to as force field parameters. One way to
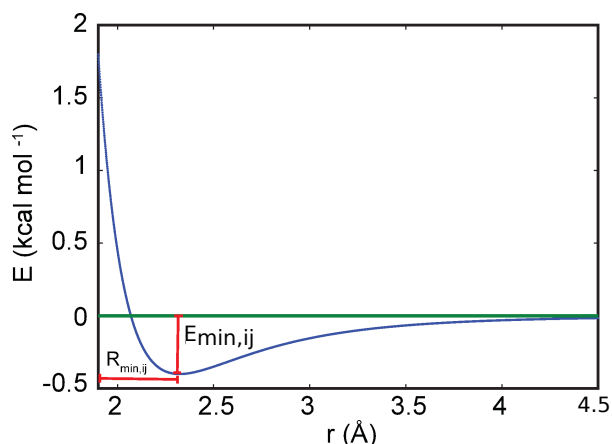
Figure 1.1: A representation of the Lennard-Jones function. The $R_{\min,ij}$ is the position of the minimum and $E_{\min,ij}$ is the well depth of the Lennard-Jones potential.

produce these parameters, developed by William Jorgensen,[29] is through adjusting parameters so that the simulations using these potentials recreate the experimental thermodynamic properties of the liquid. Thus, the performance of a model in predicting accurate thermodynamic properties depends on the properties that have been used in the parameterization process.[30]

One of the drawbacks of using a simplified potential like Eqn. 1.1 is that some aspects of intermolecular interactions are neglected. For instance, the partial atomic charges in Eqn. 1.1 are fixed. Therefore, induced polarization, which arises from changes in the electronic structure of molecules influenced by the external electric field created by the neighboring molecules, is totally ignored (for this reason Eqn. 1.1 is called the non-polarizable model).[31] Induced polarization is a significant issue in the systems that were studied in this thesis. Hydrogen sulfide is a highly polarizable molecule with a gas phase molecular polarizability of 3.63 Å^3,[32] more than twice that of water. In the case of Mg^{2+} and Zn^{2+}, the polarization of water molecules near the divalent charge of these ions is a significant term for these ion–water interactions. Accurate computational models of these systems should include this effect.

In order to overcome the problem described above, new methods have been developed

to introduce induced polarization in the calculation of properties in the condensed phase. The induced dipole model,[33] the Drude polarizable model,[34] and the fluctuating charge model[35] are different approaches that include polarizability. In this thesis, we will focus on the Drude model.

1.4 Computational Aspects of Molecular Simulation

1.4.1 Time Averages

The measurement of a given mechanical property is in fact the average of a function, G, which is position and momentum dependent, over a time interval, τ:

$$G_{\text{obs}} = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} G(p_t, q_t) dt. \tag{1.2}$$

In this equation, G_{obs} is the value that is measured experimentally. τ is the time length that G has been measured over. $G(p, q)$ implies that the physical property G depends upon the position and momentum of all the particles that comprise the system under measurement.

According to the principles of classical mechanics, having the position and momentum of particles in a given time, their next and previous positions and momentum can be obtained at any required time. p and q in the Hamiltonian form are defined for a particle, i, by the equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$
$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$
$$H = T + V,\ T = \frac{p_i^2}{2m},\ V = V(q) \tag{1.3}$$

The full set of possible values for the variables p_i and q_i is called phase space.

6

For an isolated system which is under constant energy conditions (the microcanonical ensemble), the momenta and coordinates of particles evolve according to Hamilton's equation of motion (Eqn. 1.3). As the system is under the constant energy regime, the trajectory calculated from the Eqn. 1.3 will also move on the constant energy surface. According to the ergodic hypothesis, given enough time, a finite number of particles will ultimately be able to sample all the possible phase space configurations on the constant energy surface. Therefore, the time averages of G on the trajectory generated by the appropriate number of particles corresponds to the experimentally observable property of that system.[36;37]

$$\langle G \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau G(q_t, p_t) dt. \tag{1.4}$$

In molecular dynamics simulations, Eqn. 1.3 is solved numerically by replacing the partial derivatives with the discrete approximation. One of the advantages of molecular dynamic methods over other molecular simulation methods, like Monte Carlo, is that molecular dynamics simulations provide information for both equilibrium and dynamic properties of the system.

1.4.2 Integrator

The easiest way to integrate Hamilton's equations of motion is through a Taylor series expansion of the positions of the particles (r) as a function of time (t),

$$r_i(t + \Delta t) \approx r_i(t) + v_i(t)\Delta t + \frac{F_i(t)}{2m_i}\Delta t^2 + O(\Delta t^3). \tag{1.5}$$

All the terms higher than second order in Δt are ignored. This equation is inappropriate for molecular dynamics simulation due to its large error, $O(\Delta t^3)$. A more efficient and accurate integration method is Verlet integrator.[38] This integrator is derived from two Taylor

7

expansions, one forward and one backward in time,

$$r_i(t + \Delta t) \approx 2r_i(t) - r_i(t - \Delta t) + \Delta t^2 \frac{F_i(t)}{m_i} + O(\Delta t^4). \tag{1.6}$$

There are other integrators developed for different applications, which can be found in Ref. 30.

The time interval Δt is always selected to be smaller (approximately 1/10) than the fastest oscillation in the system. A reasonable estimate of Δt for typical liquids is 1–2 femtosecond (1 fs=10^{-15} s).

1.4.3   Starting Point

In order to perform an MD simulation, the initial coordinates and velocities must be provided. In some systems, such as biological molecules or molecular crystals, the coordinates can be obtained from the X-ray crystal structure. For a simple liquid, any appropriate random distribution of coordinate with correct density can be used.

The atomic velocities are assigned according to the Maxwell–Boltzmann distribution with the desired temperature, T:[30]

$$f(v) = \left(\frac{m}{2\pi RT}\right)^{1/2} e^{-mv^2/2RT}. \tag{1.7}$$

1.4.4   Periodic Boundary Conditions

Depending on the power of the computational facilities, our simulations are only capable of describing systems containing 10000–100000 atoms. This number of particles is trivial even compared to a droplet of liquid, which contains on the order of 10^{23} atoms. In other words, the boundary conditions significantly affects the result of the MD simulation of such a small system. In order to simulate a bulk liquid using small number of atoms, the periodic

boundary condition (PBC) strategy is used. In the PBC method, the unit cell of the system is repeated infinitely in three dimensions. If a particle exits the unit cell, its image enters from the opposite side of the unit cell. Particles interact with particles in the periodic images of the unit through intermolecular interactions. This allows bulk liquids to be represented using a small number of particles. [39]

1.4.5 Truncation of Short Range Interactions

The dispersion interaction in the Lennard-Jones function is short range (decreases by the factor of $\frac{1}{r^6}$, where r is the interatomic distance), therefore particles far from each other do not have a significant contribution to the potential energy of the system due to this interaction. In order to improve the efficiency of simulation, the Lennard-Jones interaction is only calculated for atomic pairs at distances less than $10 - 16$ Å from each other. This distance is called the cutoff distance, r_c.

$$U_{lj}^{trunc} = \begin{cases} U_{lj} & r \leq r_c \\ 0 & r > r_c \end{cases} \tag{1.8}$$

In order to avoid a discontinuity in U_{lj} at r_c, a switching function is introduced to smoothly decrease the potential function to zero in a short interval. A form of switching function that can be used is, [40]

$$S(r) = \begin{cases} 1 & r \leq r_c \\ \frac{(r_{off}^2 - r^2)^2 (r_{off}^2 + 2r^2 - 3r_{on}^2)}{(r_{off}^2 - r_{on}^2)^3} & r_{on} < r \leq r_{off} \end{cases} \tag{1.9}$$

where $[r_{on}, r_{off}]$ is the interval that the switching function is applied over.

1.4.6 Electrostatic Interactions

Since electrostatic interactions have $\frac{1}{r}$ dependency, they cannot be truncated in the same manner as Lennard-Jones interactions are. The Ewald sum method was developed to calculate the electrostatic interaction in a computationally efficient and accurate way compatible with the PBC. Under PBC conditions, the Coulombic potential of particle i, located at position r_i, with charge q_i is:

$$\phi(r) = \sum_{n} \sum_{i=1}^{N} \frac{q_i}{|r_i + nL - r|}, \tag{1.10}$$

where L is the length of the unit cell and $n = (n_1, n_2, n_3)$, that corresponds to each PBC unit cell. In the Ewald sum method, Eqn. 1.10 is divided into short range and long range components. In the Ewald approximation, it is considered that each charge is surrounded by continuous positive and negative charge densities. Gaussian form densities are assigned to each charge. The sign of these functions are opposite for the positive and negative charges. Using Poisson's equation, the electrostatic potential applied from the counterions (particles with the opposite charge) on the sample charge A is,[30]

$$\phi_{\text{counterions},A} = \sum_{i} q_{\text{counterions},i} \operatorname{erfc}(\beta|r_i - r|)/|r_i - r| , \tag{1.11}$$

where $\operatorname{erfc}()$ is the complementary error function resulting from the Gaussian charge density with the width β. The $\phi_{counter,A}$ potential is a short range potential.

Considering the periodic form for the co–ions distribution (which are distributed in each PBC image as well), the potential resulting from the counterion charge distribution, $\phi_{\text{counterions},A}$, can be calculated using Fourier series. The final form of the $\phi_{\text{counterions},A}$ is:[30]

$$\phi_{counterions,A} = \frac{1}{\pi L} \sum_{m \neq 0} \frac{\exp[-\pi^2 m^2/\beta^2 L^2]}{m^2} \sum_{i}^{N} q_i \exp\left[2\pi \frac{m \cdot (r - r_i)}{L}\right], \tag{1.12}$$

where m is an integer number representing each unit cell under PBC condition.

The total potential on the particle **A** can be obtained by removing the potential $q_A/|r_A - r|$ from $\phi_{\text{counterions},A} + \phi_{counterions,A}$ and then taking the limit as $r \to r_A$. The result is:[30]

$$\phi_{Ew}(r_A) = \sum_m \sum_j q_j \frac{\text{erfc}(\beta|r_j - r_A + n|)}{|r_j - r_A + n|} - q_A \frac{2\beta}{\sqrt{\pi}}$$
$$+ \frac{1}{\pi L} \sum_{m \neq 0} \frac{\exp[-\pi^2 m^2/\beta^2 L^2]}{m^2} \sum_i^N q_i \exp\left[2\pi \frac{m \cdot (r - r_i)}{L}\right]. \tag{1.13}$$

The procedure used to assign these charge densities is by designing a uniform 3 dimensional grid in the unit cell.

1.4.7 Non-Hamiltonian Molecular Dynamics

Systems of practical interest are typically under isothermal–isochoric or isothermal–isobaric conditions, rather than the constant energy conditions described by a straightforward Verlet algorithm MD simulation. Properties like the enthalpy of evaporation, Gibbs energy of solvation, and diffusion coefficient are in this group. Molecular dynamics methods have been developed to sample phase space according to the desired thermodynamics condition.

Nosé developed a method to run MD simulation under constant temperature.[41] In this method an additional coordinate, s, is introduced that functions like a thermal reservoir. This coordinate allows dynamic energy flow to the system and back according to the following equation of motion:

$$H_{Nos\acute{e}} = \sum_{i=1}^N \frac{p_i^2}{2m_i s^2} + V(q_i) + \frac{p_s^2}{2Q} + \frac{g}{\beta} \ln s, \tag{1.14}$$

where β is $1/k_B T$, Q is an effective mass associated with s, and g is a parameter related to the degrees of freedom in the system.

One popular constant pressure technique is the Andersen method.[42] It controls pressure

based on coupling the system to a variable with the L^3 dimension to mimic the action of a piston in the real system. This piston has a mass equals to M_v, and it is introduced in the equations of motion as follow:

$$\dot{r}_i = \frac{p_i}{m_i} + \frac{\dot{V}}{3V} r_i$$
$$\dot{p}_i = F_i - \frac{\dot{V}}{3V} r_i$$
$$\ddot{V} = \frac{1}{M_v}[P(t) - P_0], \tag{1.15}$$

where V is the volume, $P(t)$ is the instantaneous pressure, P_0 is the desired pressure.

1.5   Other Molecular Simulation Techniques

In this section, the theoretical background for the other molecular simulation methods that are used in this work are provided.

1.5.1   The Drude Model

Among current polarizable force fields, the Drude model is attractive, as it is simple and computationally efficient. In the Drude model, polarizability is introduced by Drude particles. Drude particles are charged particles attached to non-hydrogen, "parent", atoms through a harmonic spring, interacting with other particles only through electrostatic interactions. In this model, polarizability is expressed as $\alpha = \frac{q_D^2}{k_D}$,[31] where α is the atomic polarizability, k_D is the restraining constant of the Drude oscillator (typically with value of 1000 N/m), and q_D is the charge of the Drude particle. In this model, $q_D + q_A$ represents the total charge of the atom. Therefore, given the total charge and polarizability of the parent atom, the charge of the associated Drude particle can be specified. The inclusion of these Drude particles results in extra expressions in the non-bonded energy ($U_{\text{non-bond}}$) function,

which accounts for the induced dipole interaction,[34]

$$
\begin{aligned}
U_{\text{non-bond}} = {} & \sum_A^N \sum_{B>A}^N \frac{q_c(A) \cdot q_c(B)}{|r(A) - r(B)|} + \sum_A^N \sum_{B>A}^{N_D} \frac{q_D(A) \cdot q_c(B)}{|r_D(A) - r(B)|} + \\
& \sum_A^{N_D} \sum_{B>A}^{N_D} \frac{q_D(A) \cdot q_D(B)}{|r_D(A) - r_D(B)|} + \frac{1}{2} \sum_A^{N_D} k_D(|r_D(A) - r(A)|)^2 + \\
& \sum_A^N \sum_{B>A}^N E_{\text{min},AB} \left[\left(\frac{R_{\text{min},AB}}{|r(A) - r(B)|} \right)^{12} - 2 \left(\frac{R_{\text{min},AB}}{|r(A) - r(B)|} \right)^6 \right].
\end{aligned}
\tag{1.16}
$$

In Eqn. (1.16), the first term is the simple charge-charge electrostatic interaction be-
tween the parent atoms, while the second, third, and fourth expressions result from Drude–
parent, Drude–Drude, and the potential energy of the Drude harmonic restraint, respec-
tively. The last term is the pair-wise Lennard-Jones interactions for repulsive and dispersive
interactions.

Although the Drude model is able to approximate the effect of induced polarization, in
some cases, such as ion solvation (Chapter 4), it may not be sufficiently accurate. In these
cases, QM/MM models and ab initio MD, described in the next sections, are more realistic
models that can be employed.

1.5.2   QM/MM Molecular Dynamics

QM/MM (quantum mechanical/molecular mechanical) is a popular hybrid ab initio/molec-
ular mechanics method. In this method, the system is divided into two parts. The part of
the system that is chemically or physically important is treated using a quantum mechanical
method. This part of the system is called inner region. The rest of the system, designated as
the outer region, is represented by a molecular mechanical model. For each new configura-
tion of the the system, corresponding to each MD step, the energy and forces in the inner
region is updated by a calculation using a quantum mechanical method.

The challenging part of QM/MM models is the formulation of the interactions between outer region and inner region:

$$E_{QM/MM}^{tot} = E_{MM}(out) + E_{QM}(in) + E_{QM-MM}(in, out). \tag{1.17}$$

The interaction between outer and inner regions, $E_{QM-MM}(in, out)$, consists of bonding, electrostatic, and attractive and repulsive VDW interactions:

$$E_{QM-MM}(in, out) = E_{QM-MM}^{bond} + E_{QM-MM}^{VDW} + E_{QM-MM}^{elec}. \tag{1.18}$$

Different methodologies have been developed to calculate each of the components in Eqn. 1.18 which can be found in the Ref. 43. In the case of ion solvation, there is no bonding interaction between outer and inner regions; only Lennard-Jones and electrostatic interactions are included in $E_{QM-MM}^{in,out}$. To calculate the electrostatic interactions between the outer region and inner region, the interaction between the atomic point charges of water molecules in the outer region and electrons and protons of the inner region are estimated using:

$$F_{QM-MM,\mu\nu}^{elec} = F_{\mu\nu}^{QM} - \sum_{n}^{MM} \langle \mu | \frac{q_n}{r_n} | \nu \rangle. \tag{1.19}$$

In this equation, $F_{\mu\nu}$ is the element of the Fock matrix corresponding to the basis functions μ and ν obtained from QM calculation of inner region and q_n is the charge of nth particle in the outer region.[44]

1.5.3 Ab Initio Molecular Dynamics

Ab initio molecular dynamics technique (AIMD) is another quantum-mechanics based molecular simulation method. This technique relies on the calculation of time-independent Schrödinger equation:

$$E\Psi = \hat{H}\Psi \tag{1.20}$$

$$\hat{H} = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + V_{n-e}\left(\{r_i\}, \{R_I\}\right). \tag{1.21}$$

In this equation I and i refer to the nuclei and electrons, respectively. ∇^2 is the kinetic energy operator and V_{n-e} is an operator that includes all the electron–electron, electron–nuclei, and nuclei–nuclei potential interactions. Since the beginning of quantum mechanics, different methodologies, such as Ehrenfest and Born–Oppenheimer molecular dynamics, were explored to calculate Eqn. 1.21. The Car–Parrinello technique, first published in 1985, was the first method that made it possible to perform ab initio molecular dynamics simulations on relatively large systems.

In some instances, Car–Parrinello MD (CPMD) is more computationally efficient than other methods because the electronic degrees of freedom are propagated simultaneously to the nuclear degrees of freedom, avoiding a costly electronic SCF procedure at each time step.[45] CPMD achieves this by assigning a fictitious mass to each electronic orbital and coupling these orbitals to a low temperature thermostat and solving the equations of motion using Lagrangian mechanics.

The hybrid Gaussian and plane waves (GPW) method is another technique that allows AIMD simulations to be performed efficiently. In the GPW method, the Gaussian type basis sets are assigned to the electrons to describe the wave function. The plane-wave functions are employed as an auxiliary basis set that represents the density on a regular grid. The plane waves are combined so that the result has the same value as the density matrix calculated from the Gaussian basis sets. Fast Fourier transforms (FFT) makes this combination computationally efficient, so 100 ps long simulations can be performed routinely on systems containing up to 1000 atoms.[46]

The efficiency of the GPW method is partly results from the pseudo-potential method. As the core electrons are not significantly affected by the chemical environment of the atom, they are removed from the system. The interactions of the core electrons on the valence electrons are replaced by pseudo-potentials to mimic the effect of these electrons. These pseudo–potentials are then used in the Kohn–Sham equation along with the other terms, such as exchange–correlation potential, to calculate the energy of the system:[46;47]

$$E[n] = \int V(r)n(r) \, dr + \frac{1}{2} \int \int \frac{n(r)n(r')}{|r - r'|} \, dr \, dr' + A[n]$$

$$E[n] = E^T[n] + E^V[n] + E^H[n] + E^{XC}[n] + E^{II}$$

$$= \sum_{\mu\nu} P^{\mu\nu} \langle \phi_\mu(r)| - \frac{1}{2}\nabla^2 |\phi_\nu(r)\rangle$$

$$+ \sum_{\mu\nu} P^{\mu\nu} \langle \phi_\mu(r)|V_{loc}^{PP}(r)|\phi_\nu(r)\rangle$$

$$+ \sum_{\mu\nu} P^{\mu\nu} \langle \phi_\mu(r)|V_{nl}^{PP}(r,r')|\phi_\nu(r')\rangle$$

$$+ 2\pi\Omega \sum_G \frac{\tilde{n}^*(G)\tilde{n}(G)}{G^2} + \int e^{XC}(r) \, dr$$

$$+ \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|R_I - R_J|}. \tag{1.22}$$

where $V(r)$ is a static potential and $A(r)$ is a universal functional of the density. $n(r)$ and $\tilde{n}(r)$ are the densities calculated from the Gaussian and plane wave functions, respectively. $E^T[n]$ is the electronic kinetic energy, $E^V[n]$ is the $e - n$ interaction which is described by local and non-local pseudo-potential (V_{loc}^{PP} and V_{nl}^{PP}). $E^H[n]$ is the electronic Hartree energy, which is calculated using plane wave functions and $E^{XC}[n]$ is the exchange-correlation energy. E^{II} is the $n - n$ interaction.

16

1.6 Calculating Thermodynamic Properties From MD Simulations

The time averages of calculated from MD simulations can be related to thermodynamic properties of bulk material. In this section, methods that are used to calculate thermodynamic properties calculated in this work are explained.

1.6.1 Enthalpy of Vaporization

By assuming the vapor phase of a material behaves as an ideal gas, with effectively no intermolecular interactions, the enthalpy of vaporization can be calculated from following equation:

$$\Delta_{vap}H = RT - \langle U \rangle_{liq}, \qquad (1.23)$$

for the rigid models. An additional degree of freedom in the gas phase is introduced by the flexible models, that is the harmonic vibrational energy term. This term can be calculated analytically using the equipartition theorem. By including this term for the gas phase potential energy, the enthalpy of vaporization of the flexible models is calculated from:

$$\Delta_{vap}H = \frac{3}{2}RT - \langle U \rangle_{liq}. \qquad (1.24)$$

The gas phase enthalpy of the Drude polarizable model includes an additional term due to the oscillation of Drude particles, although this term is small because these degrees of freedom are regulated to a very low temperature ($T^* = 1$ K).

$$\Delta_{vap}H = RT - (\langle U \rangle_{liq} - \frac{3}{2}RT^*). \qquad (1.25)$$

17

1.6.2 Gibbs Energy of Hydration

The Gibbs energy of solvation ($\Delta G_{\text{hydr}}^{\circ}$) is calculated using a free energy perturbation (FEP) decoupling of the solute from the solvent.[48] This procedure divides $\Delta G_{\text{hydr}}^{\circ}$ into electrostatic, dispersive and repulsive components,

$$\Delta G_{\text{hydr}} = \Delta G_{\text{elec}} + \Delta G_{\text{disp}} + \Delta G_{\text{repul}} \tag{1.26}$$

These energies are calculated from the relative free energies of two states using a coupling parameter, λ,

$$U(\lambda) = (1 - \lambda)U_i + \lambda U_j \tag{1.27}$$

In this equation, U_i and U_j, correspond to the potential energy of the i and j states, respectively. Following the theory of thermodynamic integration[37] each component in Eqn. 1.26 is calculated from:

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \langle U_j - U_i \rangle_\lambda \, \mathrm{d}\lambda. \tag{1.28}$$

1.6.3 Dielectric Constant

The dielectric constant is calculated using,[49]

$$\epsilon = \epsilon_\infty + \frac{4\pi(\langle M^2 \rangle - \langle M \rangle^2)}{3\langle V \rangle k_B T}, \tag{1.29}$$

where M is the dipole moment of the unit cell, $\langle V \rangle$ is the total volume of the system and ϵ_∞ is infinite frequency dielectric constant. ϵ_∞ is calculated from the Clausius–Mossotti equation,

$$\frac{\epsilon_\infty - 1}{\epsilon_\infty + 2} = \frac{4\pi\alpha}{3\nu} \tag{1.30}$$

where α is the molecular polarizability and ν is the molecular volume obtained from simulation.

1.6.4 Diffusion and Shear Viscosity Coefficient

The diffusion coefficient can be evaluated using the Einstein relation,[39]

$$D_{PBC} = \frac{1}{6t}\langle |r_i(t) - r_i(0)|^2 \rangle, \tag{1.31}$$

the PBC indicates that correction from the system size effect has not been applied. The correction for finite size effects of Yeh and Hummer[50] is applied by,

$$D = D_{PBC} + 2.837297\frac{k_B T}{6\pi\eta L}, \tag{1.32}$$

where η is the shear viscosity coefficient calculated from the Green–Kubo relation,[39]

$$\eta = \frac{V}{k_B T}\int_0^\infty \mathrm{d}t \langle p_{xy}(0) \cdot p_{xy}(t) \rangle_0, \tag{1.33}$$

where p_{xy} is the xy component of the pressure tensor.

1.6.5 Surface Tension

The surface tension is calculated from an isothermomal-isochoric simulation of a periodic slab of the liquid with gas layers adjacent to the liquid phase. This theory is described in Refs. 51 and 39,

$$\gamma = \frac{1}{2} L_z [\langle P_{zz} \rangle - \frac{1}{2} (\langle P_{xx} \rangle + \langle P_{yy} \rangle)] \tag{1.34}$$

where P_{ii}'s are the diagonal elements of the pressure tensor and L_z is the length of the unit cell along the z-axis.

1.7 Outline

This thesis is divided into three chapters. The Drude force field parameterization procedure for H_2S is described in Chapter 2. The thermodynamic properties of H_2S are obtained and compared to the other existing H_2S models. In Chapter 3, the thermodynamic properties, such as Gibbs hydration energy, diffusion coefficient, dipole moment, and surface tension, of the H_2S–H_2O system have been explored. In Chapter 4, the solvation of Mg^{2+} and Zn^{2+} is simulated. Three different molecular simulation methods, including non-polarizable force field, Drude polarizable force field, and QM/MM simulation are employed. The performance of each model is discussed. Finally, in Chapter 5, the conclusions and future work directions are discussed.

Bibliography

[1] Riahi, S.; Rowley, C. N. *J. Phys. Chem. B* **2013**, *117*, 5222–5229.

[2] Riahi, S.; Rowley, C. N. *Submitted to J. Phys. Chem. B–jp-2013-096198*

[3] Riahi, S.; Roux, B.; Rowley, C. N. *Can J. Chem.* **2013**, *91*, 552–558.

[4] Reiffenstein, R. J.; Hulbert, W. C.; Roth, S. H. *Annu. Rev. Pharmacol. Toxicol.* **1992**, *32*, 109–134.

[5] Fischer, H. *Chem. Ing. Tech.* **1967**, *39*, 515–520.

[6] Wolf Jr., T., McLean *Removing Hydrogen Sulfide From Sour Crude in the Field*; 1979.

[7] Lee, K., Neff, J., Eds. *Produced Water: Environmental Risks and Advances in Mitigation Technologies*; Springer, 2011.

[8] Maxwell S., T., Spark *Souring of Reservoirs by Bacterial Activity During Seawater Waterflooding*; 2005.

[9] Liang, S.; Kusalik, P. G. *J. Phys. Chem. B.* **2010**, *114*, 9563–9571.

[10] Liang, S.; Kusalik, P. G. *Chem. Sci.* **2011**, *2*, 1286–1292.

[11] Miller, F.; Vandome, A.; John, M. *Girdler Sulfide Process*; VDM Publishing, 2010.

[12] Herrick, C. S.; Gaines, G. L. *J. Phys. Chem.* **1973**, *77*, 2703–2707.

[13] Strathdee, G. G.; Given, R. M. *J. Phys. Chem.* **1976**, *80*, 1714–1719.

[14] O'Rourke, D.; Connolly, S. *Annu. Rev. Env. Resour.* **2003**, *28*, 587–617.

[15] Green, F. H.; Schurch, S.; De Sanctis, G. T.; Wallace, J. A.; Cheng, S.; Prior, M. *J. Appl. Physiol.* **1991**, *70*, 1943–1949.

[16] Li, L.; Rose, P.; Moore, P. K. *Annu. Rev. Pharmacol. Toxicol.* **2011**, *51*, 169–187.

[17] Mathai, J. C.; Missner, A.; Kügler, P.; Saparov, S. M.; Zeidel, M. L.; Lee, J. K.; Pohl, P. *PNAS* **2009**, *106*, 16633–16638.

[18] Cuevasanta, E.; Denicola, A.; Alvarez, B.; Möller, M. N. *PLoS ONE* **2012**, *7*, e34562.

[19] Delhommelle, J.; Boutin, A.; Fuchs, A. H. *Mol. Simulat.* **1999**, *22*, 351–368.

[20] Nath, S. K. *J. Phys. Chem. B* **2003**, *107*, 9498–9504.

[21] Vorholz, J.; Rumpf, B.; Maurer, G. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4449–4457.

[22] Nieto-Draghi, C.; Mackie, A. D.; Avalos, J. B. *J. Chem. Phys.* **2005**, *123*, 014505.

[23] Galliero, G.; Nieto-Draghi, C.; Boned, C.; Avalos, J. B.; Mackie, A. D.; Baylaucq, A.; Montel, F. *Ind. Eng. Chem. Res.* **2007**, *46*, 5238–5244.

[24] Ungerer, P.; Wender, A.; Demoulin, G.; Bourasseau, E.; Mougin, P. *Mol. Simulat.* **2004**, *30*, 631–648.

[25] Ohtaki, H.; Radnai, T. *Chem. Rev.* **1993**, *93*, 1157–1204.

[26] Roux, B.; Bernèche, S.; Egwolf, B.; Lev, B.; Noskov, S. Y.; Rowley, C. N.; Yu, H. *J. Gen. Physiol.* **2011**, *137*, 415–426.

[27] Shannon, R. D. *Acta Cryst.* **1976**, *A32*, 751–767.

[28] Marcus, Y. *Biophys. Chem.* **1994**, *51*, 111–127.

[29] Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638–6646.

[30] Becker, O.; MacKerell, A.; Roux, B.; Watanabe, M. *Computational Biochemistry and Biophysics*; Marcel Dekker Incorporated, 2001.

[31] Lopes, P.; Roux, B.; MacKerell, A. *Theor. Chem. Acc.* **2009**, *124*, 11–28.

[32] Olney, T. N.; Cann, N.; Cooper, G.; Brion, C. *Chem. Phys.* **1997**, *223*, 59–98.

[33] Ma, B.; Lii, J.-H.; Allinger, N. L. *J. Comp. Chem.* **2000**, *21*, 813–825.

[34] Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

[35] Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

[36] Hill, T. *Statistical Mechanics: Principles and Selected Applications*; Dover Books on Physics Series; McGraw-Hill, 1956.

[37] Tuckerman, M. *Statistical Mechanics and Molecular Simulations*; Oxford Graduate Texts; Oxford University Press, UK, 2008.

[38] Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.

[39] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 1989.

[40] Steinbach, P. J.; Brooks, B. R. *J. Comp. Chem.* **1994**, *15*, 667–683.

[41] Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511–519.

[42] Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384–2393.

[43] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

[44] Rowley, C. N.; Roux, B. *J. Chem. Theory Comput.* **2012**, *8*, 3526–3535.

[45] Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods; Cambridge University Press, 2009.

[46] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005**, *167*, 103–128.

[47] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[48] Deng, Y.; Roux, B. *J. Phys. Chem. B* **2004**, *108*, 16567–16576.

[49] Neumann, M.; Steinhauser, O. *Chem. Phys. Lett.* **1984**, *106*, 563–569.

[50] Yeh, I.-C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.

[51] Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1949**, *17*, 338–343.

Chapter 2

A Drude Polarizable Model for Liquid Hydrogen Sulfide

Portions of this chapter have been published as: Riahi, S., Rowley, C.N. A Drude Polarizable Force Field for Liquid Hydrogen Sulfide. *J. Phys. Chem. B* 117 (17), 5222–5229, 2013. The parameterization process and the calculation of properties were performed by the author. The manuscript was prepared by the author, with assistance by C. Rowley.

2.1 Introduction

To study the thermodynamics of liquid H_2S, a robust model for H_2S that is able to accurately describe the mechanical, thermodynamic, and transport properties of liquid hydrogen sulfide is required. Several H_2S models have been developed for this purpose. Jorgensen proposed one of the earliest models for H_2S — a three point model parameterized to the enthalpy of vaporization and density of pure H_2S.[1] The lack of satisfactory structural prop-

erties of this model led Forester et al.[2] to propose a four point model with a lone pair on the bisector of the ∠HSH. Kristóf and Liszi[3] reparameterized the four site Forester model to match the experimental enthalpy of vaporization and density along coexistence curve over the temperature range 212–352 K. Most recently, Potoff et al.[4] developed a new model based on SPC/E water model to study binary mixtures of H_2S and *n*-pentane.

While these models are capable of reproducing some of the properties of H_2S, they all neglect induced polarization, using fixed atomic charges to describe electrostatic interactions. Recently, the Drude polarizable force field[5] has been developed for a range of molecular liquids, including water,[6;7] alkanes,[8] amides,[9] and benzene,[10] alcohols,[11;12] amines,[13] and lipids.[14]

The Drude model extends the established form of a force field by adding charged "Drude" particles that move in response to electric fields, mimicking the effect of induced electron polarization. This model is now available in the popular molecular mechanical codes CHARMM[15] and NAMD.[16;17] This model can be parameterized to reproduce the mechanical and thermodynamic properties of these liquids without overestimating the dielectric constant and molecular dipole moment, a major problem for non-polarizable force fields. MacKerell and Zhu developed a Drude model force field for series of sulfur-containing compounds, although this study did not include a model for H_2S.[18]

In this study, a Drude polarizable model for H_2S is developed by adjusting the parameters to correctly predict the experimental mechanical, thermodynamic, and electric properties of liquid H_2S at 212 K (the boiling point of H_2S at 1 atm). Parameters were determined for this model through a systematic search of parameter space, similar to that used to develop the SWM4-NDP[7] model. We recalculated the properties for the existing non-polarizable models using a consistent size of system and methods for calculating non-bonded interactions and compare them to the results of our Drude model. The S–S radial distribution functions (rdf) of these H_2S models are also compared to ab initio molecular

dynamics and neutron diffraction data.[19]

2.2 Theory and Methods

2.3 The Drude Model

Rigorously, the position of the Drude particles should be energy-minimized in each config-
uration, so as to adjust in response to the electric field and remain on the self consistent field
(SCF) surface. As minimizations are computationally expensive, an extended Lagrangian
method was developed to determine the position of the Drude particles by assigning them
a fictitious mass and propagating their positions dynamically.[20] A low temperature thermo-
stat ($T^* \approx 1$ K) is applied to these degrees of freedom to ensure the simulation remains
close to SCF energy surface.

2.3.1 Molecular Mechanical Simulations

All the simulations were performed using CHARMM, version c36b2.[15] All H_2S molecules
were made rigid using the SHAKE algorithm,[21] with exception to the flexible Potoff model.
Isobaric–isothermal molecular dynamics simulations were used to calculate all the mechan-
ical and thermodynamics properties, using an extended Lagrangian method implemented
within the velocity–Verlet propagation algorithm. A dual Nosé–Hoover thermostat[22] was
applied; the first was coupled to the atomic sites and the second was coupled to the Drude
particles. The relaxation time of the atomic system was 0.1 ps and while the relaxation
time of the Drude system was 0.005 ps. The Andersen–Hoover barostat[23] with a relaxation
time of 0.2 ps was used to regulate the pressure. A cubic cell of 560 H_2S molecules under
periodic boundary conditions was used to calculate the bulk properties of each model. A
smooth real-space cutoff of 12 Å was applied to non-bonded interactions. For electrostatic

interactions, the particle mesh Ewald (PME) method with the coupling parameter $\kappa =0.33$ and sixth order B-spline interpolation was applied. Where the error on a calculated property is reported, 10 different 1 ns MD simulations with a 1 fs step size were performed to obtain the standard deviation.

2.3.2 Ab Initio Molecular Dynamics Simulations

An ab initio molecular dynamics (AIMD) simulation was also performed for a 16.3^3Å^3 cubic cell containing of 64 H_2S molecules. A 50 ps of AIMD simulation was performed using the QUICKSTEP module of the CP2K program.[24] The time step of this simulation was 1 fs and the PBE functional[25] with the TZV2P-MOLOPT-GTH basis set, GTH-PBE-q6 pseudopotentials, and a 280 Ry cutoff density grid were used. The S–H bond lengths were constrained to 1.34 Å using SHAKE algorithm.[21] Langevin MD with a frictional coefficient (γ) of 1 ps^{-1} was employed to sample the canonical ensemble of this system. The DFT-D correction for dispersion developed by Grimme et. al. was applied.[26]

2.4 Results and Discussion

2.4.1 Parameterization

For the parameterization of the polarizable H_2S model, the same strategy as SWM4-NDP water model[6] were applied. In our model, the H_2S molecules are constrained to their experimental gas phase geometry, where the S–H bond lengths are 1.34 Å and the ∠HSH is $92°$. An additional point charge, referred to as the lone pair, is located at the bisector of ∠HSH. Its distance from the S atom chosen such that the model has the correct gas phase dipole moment. The standard approximation of only having Lennard-Jones interactions between sulfur atoms and not H atoms was made. Based on this description, this model has five

free parameters: the polarizability α, the charge and position of the lone pair q_m and l_m, and Lennard-Jones parameters E_{min} and R_{min}. The same constraint which is used in the SWM4-NDP model is imposed, where the net charge of the sulfur is zero ($q_S + q_D = 0$), which simplifies the parameterization.

Initially, a very broad scan over parameters of α, l_m, q_m, E_{min}, and R_{min} was performed, where the dimerization energy (E_{dim}) and S–S distance (R_{SS}) of the H_2S–H_2S dimer was calculated for each set of parameters. E_{dim} and R_{SS} were compared to the counterpoise-corrected CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pVTZ H_2S–H_2S dimer structure and energy. Of these parameters, models with reasonable R_{SS} distances and dimerization energies greater than -4.61 kJ mol^{-1} all yielded very dense liquids with high dielectric constants and enthalpy of vaporizations. This led us to adjust the parameter space that was searched to include parameters that lead to dimerization energies in the interval of -3.77 to -4.61 kJ mol^{-1}. Parameter combinations within this range of dimer properties were chosen for the second stage of parameterization. A similar tolerance in the dimerization energy was also necessary in the development of the SWM4-DP and SWM4-NDP water models.[6;7]

Using these bounds on the parameters, the optimal parameters were then determined using central composite design (CCD) using the R project package.[27] In the CCD designing process, two values of α were chosen: 2.5 and 3.0 Å^3. The other parameters were assigned four evenly spaced values in the following intervals: l_M 0.17–0.20 Å, E_{min} between 1.67–2.72 kJ mol^{-1}, and R_{min} 4.06–4.20 Å. MD simulations were performed with these sets of parameters under isobaric–isothermal condition at 212 K. The density, enthalpy of vaporization, and dielectric constant for the parameter combinations generated from CCD were calculated at 212 K. These data were fitted to the second order (quadratic) polynomial response function with 15 coefficients,

$$\mathcal{F}_{\{\alpha, l_m, q_m, \epsilon, R_{min}\}} = a_1 \alpha^2 + ... + a_5 R^2_{min} + b_1 \alpha l_m + ... + b_{10} \epsilon R_{min}, \qquad (2.1)$$

where $\mathcal{F}$ can be replaced by any of density, enthalpy of vaporization or dielectric constant. The coefficient a_i and b_i are determined by the fitting procedure. The minimum of $\{\mathcal{F}_{expt.} - \mathcal{F}_{\{\alpha, l_m, q_m, \epsilon, R_{min}\}}\}^2$ were obtained for each of three properties. A narrow grid search was performed in the vicinity of the minimum to find the optimal parameters.

Table 2.1: Parameters of the non-polarizable and Drude H$_2$S models

	Jorgensen	Potoff et al.	Forester et al.	Kristóf and Liszi	Drude model
d_{SH} (Å)	1.340	1.340	1.340	1.340	1.340
θ_{HSH} (deg)	92	92	92	92	92
d_l (Å)			0.1933	0.1862	0.2020
q_S (e)	−0.470	−0.380	0.661	0.400	2.744
q_H (e)	0.235	0.190	0.278	0.250	0.137
q_l (e)			−1.217	−0.900	−0.274
q_D (e)					−2.744
α (Å^3)					−2.5
E_{min} (kJ mol^{-1})	1.047	1.930	2.238	2.0801	2.378
R_{min} (Å)	4.1532	4.1756	4.1418	4.1866	4.176

These final parameters of the Drude polarizable model, along with all the previous H$_2$S models, are presented in Table 2.1. This type of optimization is designed to yield the best possible parameters given the constraints imposed on the parameter and the target data, without ad hoc adjustments by the researcher.

2.4.2 Properties

2.4.2.1 Dimer Properties and Molecular Dipoles

The H$_2$S–H$_2$S dimerziation energy (E_{dim}) and sulfur–sulfur distance (R_{SS}) of Drude model and the non-polarizable models are presented in Table 2.2. E_{dim} of Drude model is −4.10 kJ mol^{-1}, which is 1.76 kJ mol^{-1} less than the CCSD(T)/aug-cc-pVTZ//MP2/aug-cc-pVTZ single point calculation. The R_{SS} of the dimer is 0.06 Å larger than the MP2/aug-cc-pVTZ optimized structure. This deviation seems to be inevitable for this model, as all

putative parameters generated during the Drude force field parameterization procedure that had dimerization energies closer to the reference resulted in liquids with high densities and enthalpies of vaporization.

Table 2.2: The structure and interaction energies of the H_2S dimer calculated with the non-polarizable, Drude, and high level ab initio models.

Model	E_{dim} (kJ mol^{-1})	R_{SS} (Å)
CCSD(T)/aug-cc-pVTZ[a]	−5.86	3.99
Jorgensen	−10.80	3.62
Potoff et al.	−7.37	3.89
Forester et al.	−7.24	3.83
Kristóf and Liszi	−6.53	3.90
Drude model	−4.10	4.06

[a] Single-point calculation at MP2/aug-cc-pVTZ optimized structure. NWChem[28] was used for ab initio calculations. The calculations were performed with the Boys–Bernardi correction for Basis Set Superposition Error (BSSE).[29]

Among the non-polarizable models, the Jorgensen model deviates most significantly, overestimating the dimerization energy by more than 4.19 kJ mol^{-1} and predicting a S–S distance that is 0.37 Å too short. The Potoff and Forester models have similar dimerization energies, roughly 1.51 kJ mol^{-1} higher than the reference value, and underestimate the S–S distance by 0.1 Å and 0.15 Å, respectively. The Kristóf model has a dimerization energy that is within 0.67 kJ mol^{-1} of the ab initio result, the most accurate dimerization energy predicted by any model, although the S–S distance is 0.09 Å too short.

These trends in dimerization energies and structures can be interpreted based on the static molecular dipole moments (μ_0) of these models (Table 2.3). The Drude model was parameterized subject to the constraint that the molecular dipole moment of the model is equal to the experimental gas phase dipole moment of 0.98 D. All the non-polarizable models have static dipole moments that are somewhat larger than the experimental value. The 3 point models have the largest dipole moments; 2.1 D and 1.7 D for the Jorgensen and Potoff

Table 2.3: Calculated physical properties of the non-polarizable and Drude H_2S models at 212 K and 1 atm.

Model	Density (kg m^{-3})	$\Delta_{vap}H$ (kJ mol^{-1})	ϵ	μ_0[a] (D)	η (10^{-4} Ps·s)	Diffusion (10^{-5} cm^2 s^{-1})
Jorgensen	959.3 ± 0.5	18.07 ± 0.00	48.60 ± 1.14	2.1	1.41 ± 0.13	9.25 ± 0.21
Potoff et al.	961.6 ± 1.7	19.74 ± 0.02	25.09 ± 0.75	1.7	3.04 ± 0.29	4.43 ± 0.15
Forester et al.	1016.5 ± 0.4	20.39 ± 0.01	9.73 ± 0.15	1.4	4.68 ± 0.46	2.88 ± 0.11
Kristóf and Liszi	958.7 ± 0.4	18.67 ± 0.00	11.59 ± 0.27	1.4	3.62 ± 0.26	3.73 ± 0.09
Drude model	951.4 ± 0.4	17.81 ± 0.00	8.18 ± 0.13	0.98	3.84 ± 0.29	3.27 ± 0.09
Expt.[c]	949	18.68	8.04	0.98	3.45	4.00

[a] Gas phase dipole moment.
[b] Density and enthalpy from Ref. 30, dielectric from Ref. 31, gas phase dipole from Ref. 32, viscosity from Ref. 33, diffusion from Ref. 34.

models, respectively. The 4 point Forester and Kristóf models both have dipole moments of 1.4 D, 43% larger than the experimental value. This overestimation of the dipole moment of the molecules results in larger electrostatic interactions between molecules and hence larger dimerization energies.

The polarizability of the final Drude model is 2.5 $Å^3$ which is about 69% of the experimental gas phase polarizability of 3.631 $Å^3$.[35] This is consistent with the trend in Drude polarizable models for other molecular liquids,[9;18] where molecular polarizabilities that are 60–70% of their gas phase values were found to be optimal. For example, the polarizability of SWM4-NDP water model is 73% of the experimental gas phase polarizability of water.[7]

2.4.2.2 Density

All the models that were evaluated have a density that is very close to experimental value of 949 kg m^{-3} at 212 K, with exception of the Forester model, which overestimates the density by 67 kg m^{-3}. The Drude model is in particularly good agreement, with a density of 951.4 kg m^{-3}. This agreement is to be expected, as these models are typically parameterized to reproduce this experimental target value.

To test whether these models are successful at higher temperatures, the density of H_2S over the temperature range 212–298 K is calculated (Figure 2.1). The Jorgensen model

Figure 2.1: The density of H_2S calculated with the non-polarizable and Drude models along the liquid–vapor coexistence curve over the temperature range 212–298 K. The experimental values are taken from Ref. 30.

diverges from the experimental density as the temperature increases. For instance, at 298 K it calculates the density to be about 100 kg m^{-3} less than experiment. The Forester model systematically overestimates the density over the entire temperature range. The Kristóf, Potoff, and Drude models are all in good agreement with experiment, deviating by less than 20 kg m^{-3} even at 298 K.

2.4.2.3 Enthalpy of Vaporization

The enthalpies of vaporization of the Drude and non-polarizable models calculated from Eqn.1.23 and 1.25. The calculated values at 212 K are collected in Table 2.3. All the models predict the enthalpy of vaporization within 2 kJ mol^{-1} of the experimental value, with exception to the Forrester and Potoff models, which overestimate the enthalpy of vaporization by 1.71 and 1.06 kJ mol^{-1}, respectively. The Drude model underestimates the experimental enthalpy of vaporization by 0.87 kJ mol^{-1} at 212 K. The Kristóf model per-

forms exceptionally well, yielding an enthalpy of vaporization that is in essentially perfect agreement with experiment at 212 K.
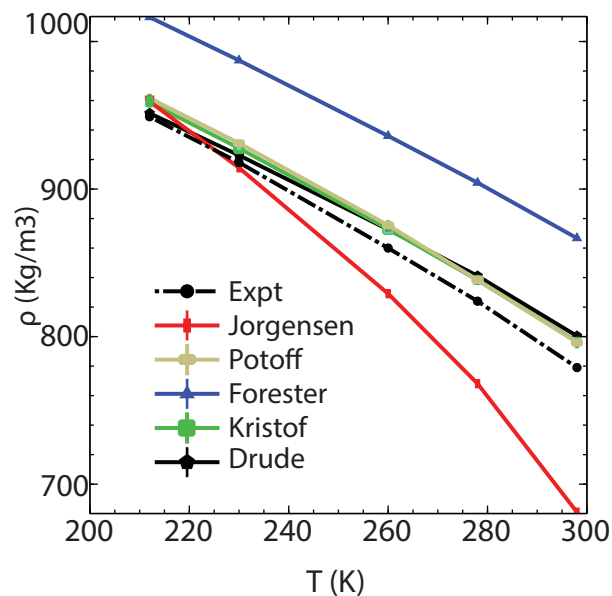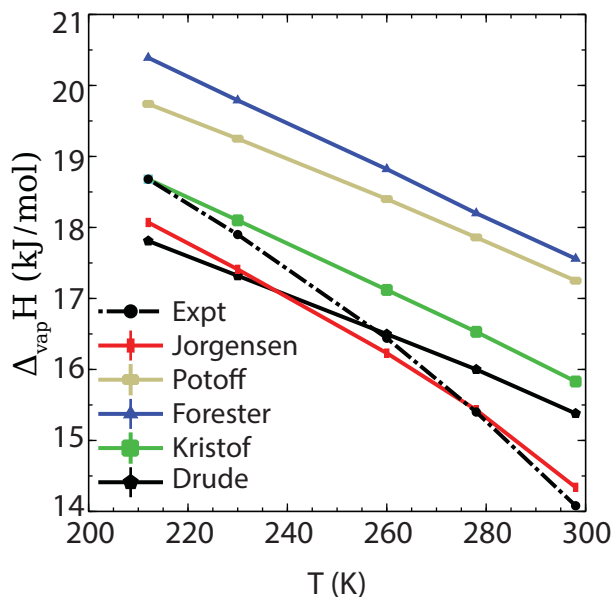

Figure 2.2: The enthalpy of vaporization of the non-polarizable and Drude models along the liquid–vapor coexistence curve for the temperature range 212–298 K. The experimental values are were from Ref. 30.

The temperature dependence of the enthalpy of vaporization is more complex. This property, calculated using the Drude and non-polarizable models, is plotted along with the experimental values over the range of 212–298 K in Figure 2.2. As it was inferred for the densities, the Forester model has the largest deviation from experiment, significantly overestimating the enthalpy of vaporization across the full range of temperatures. The Potoff model also shows a sizable deviation from experiment, a divergence that increases with temperature; the calculated enthalpy of vaporization is more than $3.00 \, \text{kJ mol}^{-1}$ higher than the experimental value at 298 K. Other models are in acceptable agreement with experimental value and their errors are about $1 \, \text{kJ mol}^{-1}$ across the temperature range. The enthalpy of vaporization calculated using the Kristóf model is in exact agreement with experiment at 212 K, but diverges as the temperature is increased and is $2 \, \text{kJ mol}^{-1}$ higher than the exper-

imental value at 298 K. The Jorgensen model underestimates the enthalpy of vaporization by roughly 1 kJ mol^{-1} at 212 K, but shows better agreement at higher temperatures.

In general, all these models correctly predict a decrease in the enthalpy of vaporization as the temperature increases, but underestimate the rate of decline. As a result, each model is very effective at one particular temperature, but deviates from experiment as the temperature changes. Overall, these changes are modest; the change in the experimental enthalpy of vaporization is roughly 4.5 kJ mol^{-1} over 80 K. Nevertheless, none of the models evaluated are able to describe this effect fully.

2.4.2.4 Dielectric Constant

The dielectric constants at 212 K calculated using the Drude and non-polarizable models are calculated from Eqn. 1.29 and are presented in Table 2.3. The non-polarizable models all have significantly larger dielectric constants than the experimental value of 8.04. This can be related to their large gas phase molecular dipole moments, which are all significantly larger than the experimental value of 0.98 D. The three point models developed by Jorgensen and Potoff have considerably larger dielectric constants than the four point models, suggesting that the lone pair site is important for simultaneously describing the mechanical and electric properties of the liquid. The Drude model shows a significant improvement over the non-polarizable models; its calculated dielectric constant of 8.18 ± 0.13, which matches the experimental value of 8.04 within statistical error. We attribute the improved accuracy of this model to its realistic gas phase molecular dipole moment, the incorporation of induced polarization effects through the Drude oscillator, and the direct parameterization of the model to reproduce the dielectric constant. The ability of the Drude model to predict accurately the dielectric constants of molecular liquids is one of the most significant successes of this model.[6;8;9]

2.4.2.5 Self-Diffusion and Shear Viscosity

The calculated and experimental self-diffusion coefficients and shear viscosities coefficients are calculated fron Eqn. 1.32 and 1.33. The calculated values are plotted over the temperature range 212–298 K in Figure 2.3 and 2.4, respectively. The Jorgensen model yields liquids that are very diffusive and hence less viscous.
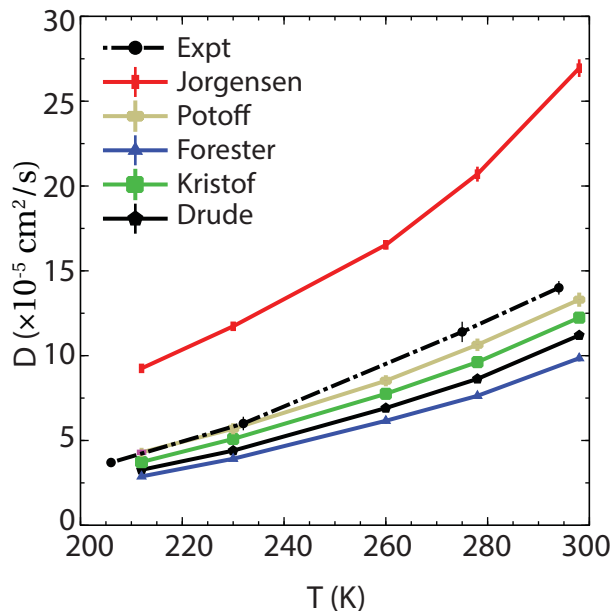
Figure 2.3: The self-diffusion coefficient of H_2S calculated using the non-polarizable and Drude models along the liquid–vapor coexistence curve for the temperature range of 212–298 K. The experimental values were taken from Ref. 34.
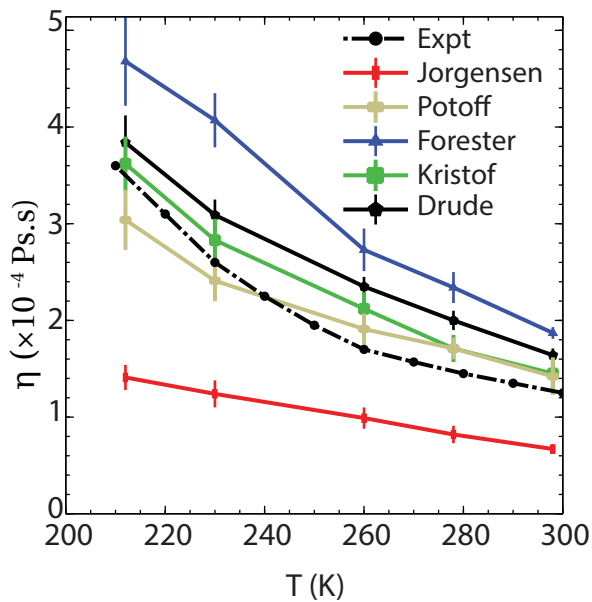
Figure 2.4: The viscosity coefficient of H_2S calculated using non-polarizable and Drude models along the liquid–vapor coexistence curve for the temperature range 212–298 K. The experimental values were taken from Ref. 33.

The Forester model, on the other hand, yields a less diffusive and more viscous liquid. The Potoff model predicts both these transport properties in excellent agreement with experiment. The Kristóf and Drude models show very similar trends for both the self-diffusion and shear viscosity and are both in good agreement with experiment.

2.4.2.6 Radial Distribution Function

The radial distribution function $(g_{SS}(r))$ of a liquid provides the correlation function between the bulk density of a liquid and its radially dependent density as a function of the distance between two atoms in the liquid (r). This provides a rigorous and transferable descriptor of the structure of a liquid. This function can be compared to data derived from experimental neutron and X-ray diffraction data and integrated to give the coordination number $(n_c = 4\pi \int_0^r g(r)r^2dr)$. The S–S radial distribution functions $(g_{SS}(r))$ of liquid H_2S
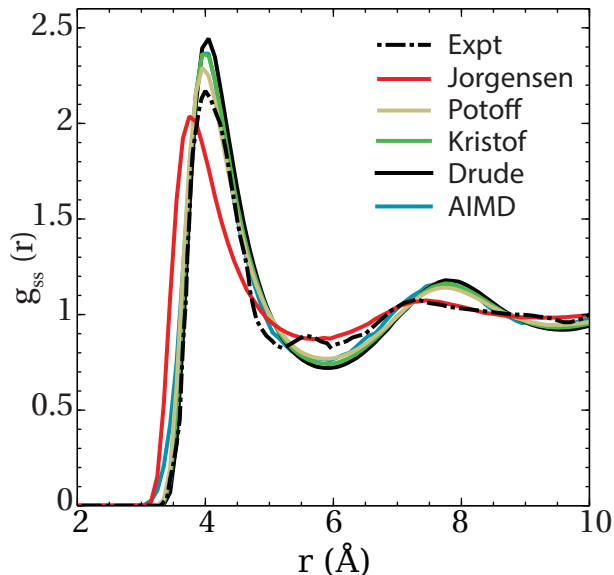
Figure 2.5: Sulfur–sulfur radial distribution function at 298 K and 19.8 atm. The experimental data was determined by neutron diffraction by Santoli et al.[36]

calculated using the Drude and non-polarizable models are plotted along with the experimental neutron scattering and AIMD data in Figure 2.5. (The radial distribution functions were calculated at at 298 K and 19.9 atm, which are the conditions the neutron diffraction data were collected for. An AIMD simulation was performed under NVT conditions at the equivalent density ($\rho_N = 0.0148$ Å^{-3}) to provide a comparison based on a first-principles model).

With exception of the Jorgensen model, all the models yield similar radial distribution functions that are in good agreement with the experimental and ab initio data, with the maximum of the first peak occurring at a S–S distances of 3.95–4.05 Å. The first peak of the rdf calculated using the Jorgensen model occurs at a distance of 3.75 Å, a significantly shorter inter-atomic distance than the other models predict. The height of first peak of the rdf's of the force field models are sharper than the height in this peak in the experimentally-derived rdf. This is a common problem in molecular mechanical force fields that use the Lennard-Jones function to describe interatomic Pauli repulsive interactions. The $E \propto r^{-12}$

term in the Lennard-Jones function tends to overestimate the steepness of the short range repulsive interaction, leading to sharper peaks in the rdf.[6]

Table 2.4: The position of the first and second peak of the rdf and coordination number of first shell calculated from MD simulations at 212 K and 1 atm for the non-polarizable and Drude H_2S models. For comparison, the values for Ar(l) and the SWM4-NDP water model are also included.

Model	R_1	R_2	$\frac{R_2}{R_1}$	n_1
Jorgensen	3.75	7.35	1.96	9.0
Potoff et al.	3.95	7.65	1.94	11.1
Forester et al.	3.95	7.55	1.91	11.9
Kristóf and Liszi	4.05	7.65	1.89	12.0
Drude model	4.05	7.65	1.89	12.1
AIMD	3.95	7.75	1.96	12.0
Ar	3.65	7.15	1.96	12.7
SWM4-NDP	2.85	4.45	1.56	4.8

Analysis of the rdf provides some insight into the structure of liquid hydrogen sulfide. The ratios of the positions of the second and first peaks are presented in Table 2.4. For all the H_2S models, this ratio is near 2, which is the characteristic of a simple fluid with a structure that is dominated by packing forces.[37] In contrast, this ratio is 1.56 for SWM4-NDP water. Similarly, the integration of the first peak of $g_{SS}(r)$ gives the average number of S atoms in the first coordination sphere, known as coordination number (n_c). With exception of the Jorgensen model, this value is between 11 to 12 for all the MM and AIMD models, reflecting a close-packing arrangement that is typical of simple Lennard-Jones fluids like Ar(l). In contrast, estimates of this ratio for liquid water are typically between 4 and 5, due to a restriction of the possible structures in order to form strong hydrogen bonds.

The S–H and H–H radial distribution functions were also calculated (Figure 2.6 and Figure 2.7, respectively). With exception to the Jorgensen model, all models are in good agreement in terms of the height and location of the peaks of these distributions. These distributions are in reasonable agreement with the distribution determined using neutron diffraction, given the limited resolution of the experimental data.[19;36] Broad peaks occur
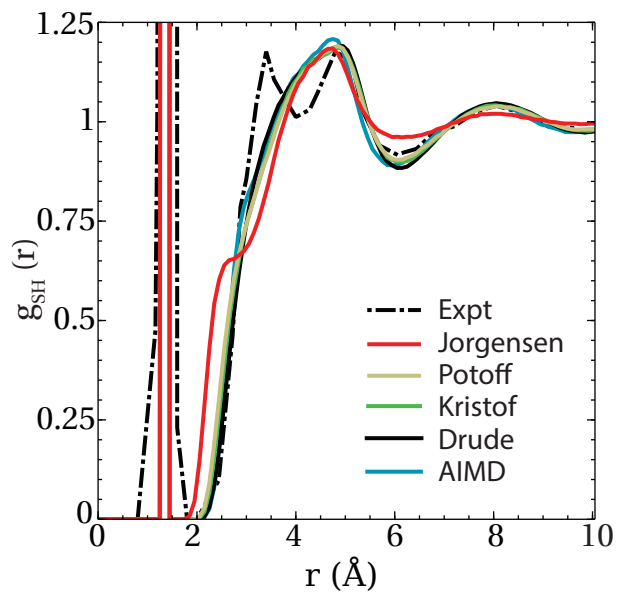
Figure 2.6: Sulfur–hydrogen radial distribution function at 298 K and 19.8 atm.
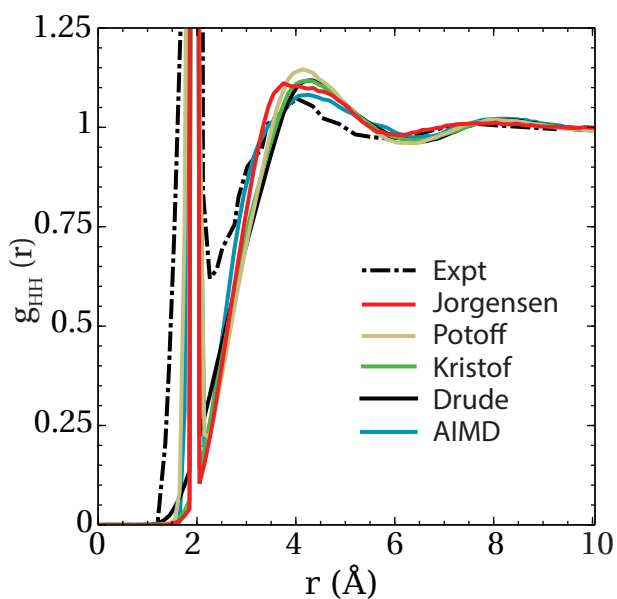
Figure 2.7: Hydrogen–hydrogen radial distribution function at 298 K and 19.8 atm.

around $r_{S-H} = 5.0$ Å and at $r_{H-H} = 4.2$ Å, respectively. The O–H rdf of water shows a strong correlation at $r = 1.85$ Å due to hydrogen bonds between water molecules at with O–O separations near $r_{O-O} = 2.8$ Å. The corresponding peak in the $g_{S-H}(r)$ of liquid

H$_2$S should occur near $r_{S-H} = 2.56$ Å if H$_2$S had a similar hydrogen bonding structure. The absence of this feature confirms that H$_2$S lacks a water-like hydrogen bonding network and instead behaves like a simple liquid, consistent with the results of simulations using the Kristóf model by Nieto-Draghi et al.[38]

2.4.2.7 Performance of Extended Lagrangian Dynamics

In order to confirm that use of extended Lagrangian dynamics to propagate the positions of the Drude particles is accurate for these highly polarizable molecules, a 500 ps MD simulation using the SCF method at 212 K was performed. During this simulation, the positions of the Drude particles were adjusted to minimize the potential energy at each step, which ensures that the simulation remains on the SCF surface. The properties derived from this simulation were identical to those calculated using extended Lagrangian dynamics within statistical uncertainty; for the SCF trajectory, the calculated density is 947 kg m^{-3}, the enthalpy of vaporization is 17.66 kJ mol^{-1}, and the dielectric constant is 8.1.

2.5 Conclusion

In this study, a polarizable model for liquid hydrogen sulfide based on the Drude model, following the same methodology and constraints on the parameters as was used to develop the SWM4-NDP water model was developed. This model predicts the density, enthalpy of vaporization, self-diffusion coefficient, and shear viscosity coefficients of liquid hydrogen sulfide along the liquid-vapor coexistence curve between 212–298 K, in good agreement with experiment. The Drude polarizable model developed for H$_2$S also has the correct dipole moment for H$_2$S in the gas phase and has a dielectric constant that is in good agreement with the experimental value at 212 K. This model has been compared to existing non-polarizable models for H$_2$S. While each of these models is effective for modeling some

features of liquid hydrogen sulfide, each model has limitations, particularly for the dielectric constant, which all non-polarizable models significantly overestimate. This is likely due the anomalously high static dipole moment in these models, a consequence of the lack of induced polarization in fixed-charge models. Of the non-polarizable models evaluated, the four point model developed by Kristóf and Liszi model shows the best agreement with experiment for most of the properties evaluated, however it also overestimates gas phase dipole moment and the liquid dielectric constant. In these two respects, the Drude model presented here shows the best performance of all the models evaluated.

The radial distribution function calculated using Drude model H_2S is consistent with the distribution calculated from an AIMD simulation and neutron diffraction data. The analysis of this distribution indicates that although H_2S is a structural analog of H_2O, H_2S forms a simple liquid that does not have the persistent hydrogen bonding network that is present in liquid water.

Bibliography

[1] Jorgensen, W. L. *J. Chem. Phys.* **1986**, *90*, 6379–6388.

[2] Forester, T.; McDonald, I.; Klein, M. *Chem. Phys.* **1989**, *129*, 225–234.

[3] Kristóf, T.; Liśzi, J. *J. Phys. Chem. B* **1997**, *101*, 5480–5483.

[4] Kamath, G.; Lubna, N.; Potoff, J. J. *J. Chem. Phys.* **2005**, *123*, 124505.

[5] Lopes, P.; Roux, B.; MacKerell, A. *Theor. Chem. Acc.* **2009**, *124*, 11–28.

[6] Lamoureux, G.; Alexander D. MacKerell, J.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

[7] Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; Jr., A. D. M. *Chem. Phys. Lett.* **200**, *418*, 245–249.

[8] Vorobyov, I. V.; Anisimov, V. M.; MacKerell, A. D. *J. Phys. Chem. B* **2005**, *109*, 18988–18999.

[9] Harder, E.; Anisimov, V. M.; Whitfield, T.; MacKerell, A. D.; Roux, B. *J. Phys. Chem. B* **2008**, *112*, 3509–3521.

[10] Lopes, P. E. M.; Lamoureux, G.; Roux, B.; MacKerell, A. D. *J. Phys. Chem. B* **2007**, *111*, 2873–2885.

[11] Anisimov, V. M.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D. *J. Chem. Theory Comput.* **2007**, *3*, 1927–1946.

[12] He, X.; Lopes, P. E. M.; MacKerell, A. D. *Biopolymers* **2013**, *99*, 724–738.

[13] Orabi, E. A.; Lamoureux, G. *J. Chem. Theory Comput.* **2013**, *9*, 2035–2051.

[14] Chowdhary, J.; Harder, E.; Lopes, P. E. M.; Huang, L.; MacKerell, A. D.; Roux, B. *J. Phys. Chem. B* **2013**, *117*, 9142–9160.

[15] Brooks, B. R.; Brooks, I. C. L.; Mackerell, J. A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al., *J. Comp. Chem.* **2009**, *30*, 1545–1614.

[16] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comp. Chem.* **2005**, *26*, 1781–1802.

[17] Jiang, W.; Hardy, D. J.; Phillips, J. C.; MacKerell, A. D.; Schulten, K.; Roux, B. *J. Phys. Chem. Lett.* **2011**, *2*, 87–92.

[18] Zhu, X.; Mackerell, A. D. *J. Comp. Chem.* **2010**, *31*, 2330–2341.

[19] Andreani, C.; Merlo, V.; Ricci, M.; Soper, A. *Mol. Phys.* **1991**, *73*, 407–415.

[20] Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

[21] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.

[22] Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.

[23] Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

[24] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005**, *167*, 103–128.

[25] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[26] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys* **2010**, *132*, 154104.

[27] R Development Core Team, R: A Language and Environment for Statistical Comput-ing. R Foundation for Statistical Computing: Vienna, Austria, 2011; ISBN 3-900051-07-0.

[28] Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Dam, H. V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. *Comp. Phys. Comm.* **2010**, *181*, 1477–1489.

[29] Boys, S.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.

[30] Clarke, E. C. W.; Glew, D. N. *Can. J. Chem* **1970**, *48*, 764–775.

[31] Havriliak, S.; Swenson, R. W.; Cole, R. H. *J. Chem. Phys.* **1955**, *23*, 134–135.

[32] Lide, D. R. *CRC Handbook of Chemistry and Physics, 93th Edition (CRC Handbook of Chemistry & Physics)*, 93rd ed.; CRC Press, 2012; Chapter 9, pp 51–59.

[33] Galliero, G.; Boned, C. *Fluid Phase Equilib.* **2008**, *269*, 19–24.

[34] Dupré, F.; Piaggesi, D.; Ricci, F. *Phys. Lett. A* **1980**, *80*, 178–180.

[35] Olney, T. N.; Cann, N.; Cooper, G.; Brion, C. *Chem. Phys.* **1997**, *223*, 59–98.

[36] Santoli, G.; Bruni, F.; Ricci, F. P.; Ricci, M. A.; Soper, A. K. *Mol. Phys.* **1999**, *97*, 777–786.

[37] Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press, 1987.

[38] Nieto-Draghi, C.; Mackie, A. D.; Avalos, J. B. *J. Chem. Phys.* **2005**, *123*, 014505.

Chapter 3

The Solvation of Hydrogen Sulfide in Liquid Water and at the Water/Vapor Interface Using a Polarizable Force Field

Portions of this Chapter have been submitted for publication as: The Solvation of Hydrogen Sulfide in Liquid Water and at the Water/Vapor Interface Using a Polarizable Force Field. *J. Phys. Chem. B.* The parameterization and the calculation of the properties were performed by the author. The manuscript was prepared by the author, with assistance by C. Rowley.

3.1   Introduction

MD simulations are a powerful tool for studying the changes that occur in the solvation of a molecule as it moves across the water/vapor interface.[1,2] A few notable examples include studies of atmospheric gases,[3] ions,[4] alcohols,[5,6] benzene,[7] and CCl_4.[8] While many

of these studies used non-polarizable force fields, some concluded that induced polarization plays an important role in describing adsorption at the water/vapor interface.[9;10] For example, Vrbka et al. concluded that soft ions will have an excess concentration at the interface, in part due to their high polarizability.[4] As H_2S is a highly polarizable molecule, with a polarizability of 3.63 $\mathring{A}^3$,[11] we anticipate that induced polarization will affect its interfacial and solution properties.

The Drude SWM4-NDP water model accurately describes interfacial properties, such as surface tension,[12] and has been used to study the interfacial behavior of ions in water.[13;14] Here, we use the SWM4-NDP model and the Drude polarizable model of H_2S[15] (described in Chapter 2) to investigate the solution and interfacial properties of H_2S in liquid water.

3.2 Theory and Methods

3.2.1 MD Simulation Details

All the molecular dynamics simulations were performed with CHARMM, version c36b2.[16] The solvation of H_2S in bulk water was modeled using a 26 $\mathring{A}$ × 26 $\mathring{A}$ × 26 $\mathring{A}$ cubic unit cell, containing 560 H_2O molecules and 1 H_2S molecule. The Lennard-Jones interactions were calculated with cutoff distance of 12 $\mathring{A}$. Electrostatic interactions were calculated using the particle mesh Ewald method[17] with $\kappa = 0.33$, sixth-order B-spline interpolation and a grid spacing of 1 $\mathring{A}$ in the interfacial simulations and 0.8 $\mathring{A}$ for bulk liquid simulations. Both the H_2O and H_2S molecules were kept rigid using the SHAKE algorithm.[18] Isothermal-isobaric MD simulations were used to calculate the solution properties. A dual Nosé–Hoover thermostat with relaxation time of 0.1 ps and 0.005 ps was applied to control the temperature of heavy atoms and Drude particles, respectively. The Anderson–Hoover barostat with relaxation time of 0.2 ps was employed to maintain the pressure at 1 atm.

Simulations of the interfacial properties were performed under isothermal-isochoric

conditions using a Langevin thermostat,[19] with friction coefficients of 5.0 and 10.0 ps^{-1} on heavy atoms and Drude particles, respectively.

3.2.2   Ab Initio Molecular Dynamics

To validate the solvation structure of H_2S in water calculated using the Drude model, an ab initio molecular dynamics (AIMD) simulation was performed using the QUICKSTEP module of the CP2K package.[20] The simulation periodic unit cell was comprised of 64 H_2O and 1 H_2S molecules in a 12.7 Å $\times$ 12.7 Å $\times$ 12.7 Å cubic unit cell. The PBE functional[21] with the TZV2P-MOLOPT-GTH basis set, GTH-PBE-q6 (q1 for H atoms) pseudo-potentials, and a 280 Ry cutoff density grid were used in the ab initio calculations. The DFT-D method[22] was employed to incorporate dispersion effects. The S–H and O–H bonds were constrained to 1.340 Å and 0.957 Å, respectively, using the SHAKE algorithm. A 1 fs time step was used. The system was simulated using Langevin dynamics with a bath temperature of 298 K and a friction coefficient of $\gamma = 1$ ps^{-1}. A 50 ps simulation was performed to equilibriate the system prior to a 200 ps production trajectory.

3.2.3   Free Energy Perturbation

The Gibbs energy of hydration was calculated using a free energy perturbation (FEP) described in Eqn. 1.26 The electrostatic (ΔG_{elec}) and dispersive components (ΔG_{disp}) of the solvation free energy were calculated by integrating the path between an initial state where the given component of the solute–solvent interaction is at full strength ($\lambda = 0$) and a final state where the interaction is not present ($\lambda = 1$). Simulations were performed at 11 values of λ: $0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$, and 1.0. Due to the end point singularity in the repulsive component of solvation free energies,[23] ΔG_{repul} is calculated by introducing a staging parameter, s, and calculating the difference in Gibbs

energy using end point thermodynamic integration for the change between values of $s =$ $0.0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$, and 1.0. The full details of this method are described in Ref. 24.

For each window of the solvation energy calculations, 0.2 ns MD simulations were performed to equilibrate the system, followed by a 4 ns production simulation. The statistical uncertainty was estimated by calculating the solvation energy from 0.5 ns segments of this trajectory and calculating the standard deviation of these 8 estimates of the free energy. The Weighted Histogram Analysis Method (WHAM)[25] was used to calculate the dispersion and repulsion components of the Gibbs energy of solvation from the time series, while thermodynamic integration was used to calculate the electrostatic component.

3.2.4 Diffusion Coefficient

The diffusion coefficient of H_2S in water were calculated from the root mean square displacement (RMSD) of single H_2S molecule in periodic cell of 560 water molecules using the Einstein equation (Eqn. 1.31). The RMSD was calculated from a 5 ns isothermal–isochoric MD simulation.

3.2.5 Free Energy Profile

The free energy profile of the transition of one H_2O or H_2S molecule across the water/vapor interface was calculated using umbrella sampling.[26;27] These simulations were performed using a 19.56 Å $\times$ 19.56 Å $\times$ 58.68 Å tetragonal periodic cell containing a ~ 20 Å thick slab of 250 water molecules and the molecule undergoing adsorption (H_2O or H_2S). A representative configuration of this system is illustrated in Figure 3.1.

The free energy profile was calculated along the z-axis of the unit cell, perpendicular to the plane of the interface. The free energy profile was calculated over the interval of 0 Å
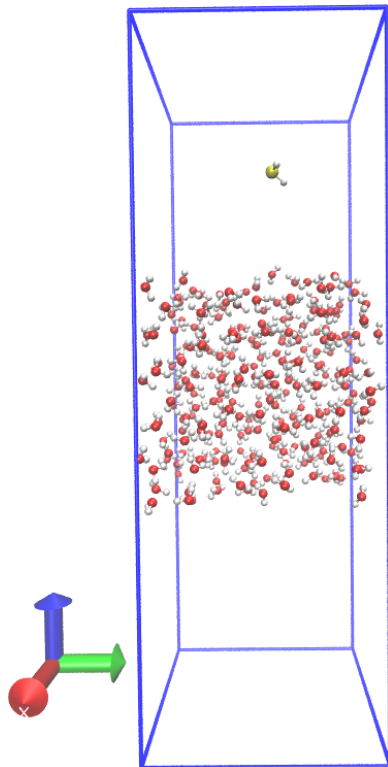
Figure 3.1: The periodic cell of the water slab that was used to calculate interfacial properties. The cell is tetragonal with dimensions 19.56 Å $\times$ 19.56 Å $\times$ 58.68 Å. The target H_2O or H_2S molecule was restrained to positions along the z-axis from the center of mass of the water slab, from $Z = 0$ Å, the center the slab, to $Z = 20$ Å, the vapor phase. These restrained simulations were used to calculate the free energy profile, average dipole, and orientation of H_2S in the transition across the water-vapor interface.

$\leq Z \leq 20$ Å, where Z is defined as the distance between the center of the mass of the water slab and the target molecule along the z-axis. In each window of the umbrella sampling simulation, the target molecule was restrained to a reference position along the Z coordinate at 0.5 Å spacings using a harmonic potential ($k = 5$ kcal mol^{-1} Å^{-2}). The reference positions of these simulations were placed at 0.5 Å separations along the coordinate, for a total of 41 windows. Each window was equilibriated for 1 ns before a 5 ns long trajectory was collected for analysis. The free energy profile was obtained from these trajectories using WHAM.[25]

3.2.6 Surface Tension

The surface tension was calculated from Eqn. 1.34.

3.3 Results and Discussion

3.3.1 Optimization of Lennard-Jones Parameters

The Drude polarizable force field describes non-bonded repulsive and dispersive interactions between two atoms, labeled A and B, through the Lennard-Jones potential,

$$U_{\text{LJ}} = \sum_A^N \sum_{B>A}^N E_{\text{min},AB} \left[\left(\frac{R_{\text{min},AB}}{|r(A) - r(B)|} \right)^{12} - 2 \left(\frac{R_{\text{min, }AB}}{|r(A) - r(B)|} \right)^{6} \right] \quad (3.1)$$

This function requires the definition of parameters E_{min} and R_{min} for each pair of interacting atoms. The parameters for homogeneous pairs of O and S atoms were determined by fitting to the bulk properties of water[12] and liquid hydrogen sulfide,[15] respectively. Reasonable estimates of the interaction parameters for heterogeneous pairs can be calculated from the parameters for homogeneous pairs using the Lorentz–Berthelot combining rule.[28]

$$E_{\text{min},AB} = \sqrt{E_{\text{min},AA} E_{\text{min},BB}}$$
$$R_{\text{min},AB} = \frac{1}{2}(R_{\text{min},AA} + R_{\text{min},BB}) \quad (3.2)$$

In some instances, simulations using parameters generated using the combining rule fail to give quantitatively accurate results.[29] Greater accuracy can be achieved by defining pair specific Lennard-Jones parameters, optimized to describe the heterogeneous system.[29;30] This is apparent in the hydration free energy of H_2S calculated when the Lorentz–Berthelot combining rule is used; the calculated $\Delta G_{\text{hydr}}^{\circ} = 0.24$ kcal mol^{-1}, which is 0.7 kcal mol^{-1} higher than the experimental hydration free energy. To improve the description of H_2S sol-

vation, H_2S–H_2O Lennard-Jones parameters were optimized to correctly predict the $\Delta G^\circ_{\text{hydr}}$ of H_2S through a 3 iteration grid search of the parameter space. This procedure is described in detail in Appendix B. S–O Lennard-Jones parameters of $R_{\text{min},S-O} = 3.808$ Å and $E_{\text{min},S-O} = 0.38618$ kcal mol^{-1} were found to be optimal. These parameters were used in all the simulations presented in this Chapter.

3.3.2 Solution Properties

3.3.2.1 Radial Distribution Function Analysis

The S–O and S–H radial distribution functions (rdf) of H_2S dissolved in water were calculated from a 4 ns trajectory to characterize the solvation structure. These functions are plotted in Figure 3.2.

At 298 K, there are two distinct maxima corresponding to the first and second coordination shells, centered at 3.6 Å and 6.5 Å, respectively. This is in good agreement with the rdf calculated from the AIMD trajectory, although the first peak of the AIMD rdf is lower and broader. The first maximum of the S–H rdf also occurs at 3.6 Å, although there is a shoulder on the right side of this peak, indicating that although waters in this shell are preferentially aligned tangentially to the solute, some extend radially to form hydrogen bonds with oxygen atoms in the second sphere. These distributions also indicate that there is only limited hydrogen bonding between the hydrogen sulfide solute and the solvent waters, as the shoulder on the right hand side of the S–H peak is minimal. This shoulder corresponds to the solute accepting hydrogen bonds from the solvent. The large molecular volume of H_2S and the limited polarity of S–H bonds make H_2S a relatively poor hydrogen bonding partner for water molecules.

The temperature dependence of the rdfs is also consistent with hydrophobic solvation. Guillot et al. noted that the second coordination sphere of hydrophobic solutes broaden and
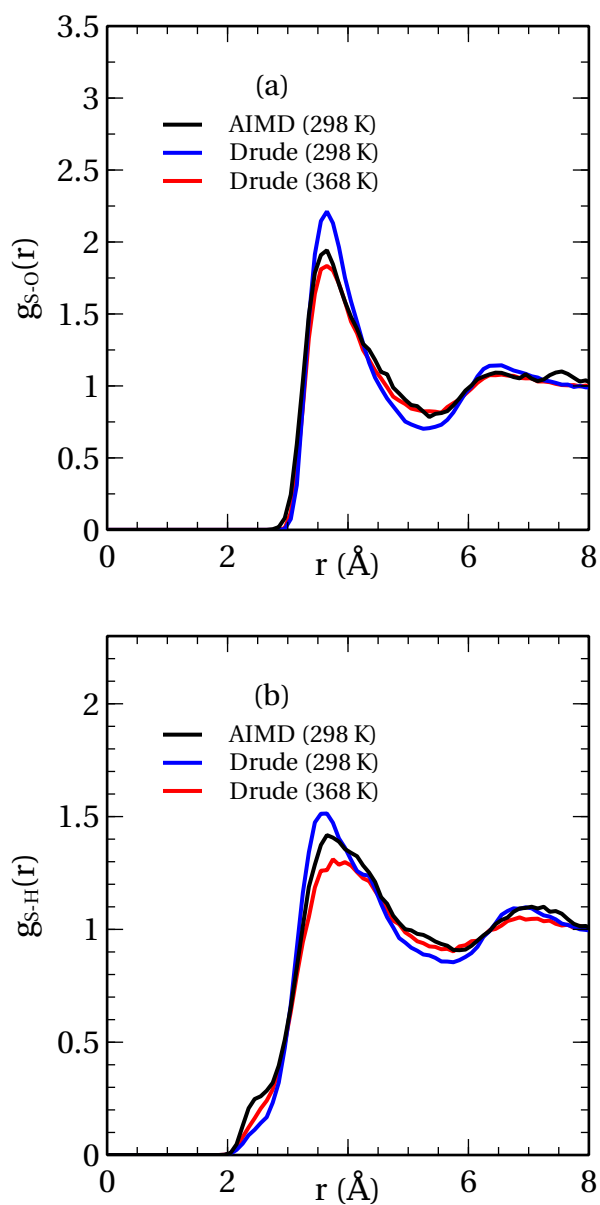
Figure 3.2: The S–O (a) and S–H (b) radial distribution functions calculated from the ab initio molecular dynamics and Drude model MD simulations of H_2S solvated in bulk water.

ultimately disappear at higher temperatures, as will the right shoulder of the X–H rdf.[31] This indicates that the hydrogen bonds formed between radially-oriented water molecules in the first coordination sphere disappear as the order of the hydrogen bonding network in the first sphere is lost at increased temperatures. This effect is apparent in the rdf of $H_2S(aq)$
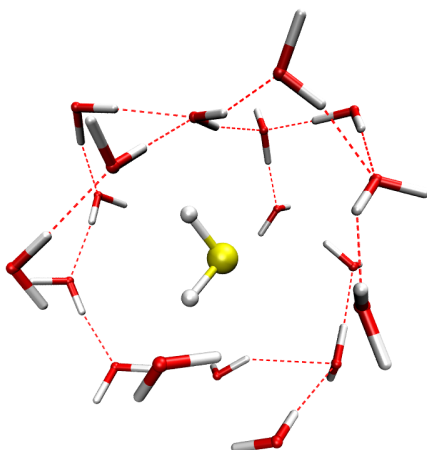
Figure 3.3: A representative configuration from a molecular dynamics simulation of H_2S (yellow) in liquid water. Hydrogen bonds are indicated by red dashed lines. The criteria for a hydrogen bonding interaction is for the O $\cdots$ H distance to be less than 3.0 Å and O–H $\cdots$ O angle to be greater than $150°$.

at 368 K in Figure 3.2.

This hydrophobic solvation is also evident in coordination number of H_2S in water, calculated by integrating the first peak of the S–O rdf. While molecules of bulk SWM4-NDP water have a coordination number of 4.8, H_2S molecules in water have a coordination number of 19.7. This is similar to those reported for hydrophobic solutes like argon and methane, which have coordination numbers of 16[32] and 20,[33;34] respectively. These coordination numbers reflect that the first coordination sphere of a hydrophobic solute like H_2S incorporates a large hydrogen-bonding network of water molecules.

This analysis is generally consistent with the view that H_2S is hydrated like a hydrophobic solute, with water molecules in the first coordination sphere forming an "iceberg"-like hydrogen bonding network around the solute.[35;36;37] A representative configuration of this solvation structure is shown in Figure 3.3.

3.3.2.2 Gibbs Energy of Hydration

The hydration free energies were calculated from Eqn. 1.26. The experimental solubility data reported by Clarke et al.[38] show that H_2S is weakly soluble in water, with hydration free energies between -0.85 kcal mol^{-1} and -0.35 kcal mol^{-1} over the temperature range $273 - 323$ K. (These data were converted to the standard state convention of 1 M gas and 1 M solution concentrations, as described in Abraham, M. H. *J. Chem. Soc., Faraday Trans. 1* **1984**, *80, 153–181*.) Although the H_2S–H_2O Lennard-Jones interaction term was optimized to $\Delta G^{\circ}_{\text{hydr}}$ at 298 K, our model is able to predict the Gibbs energy of hydration with good accuracy over the $273 - 323$ K range (Figure 3.4).
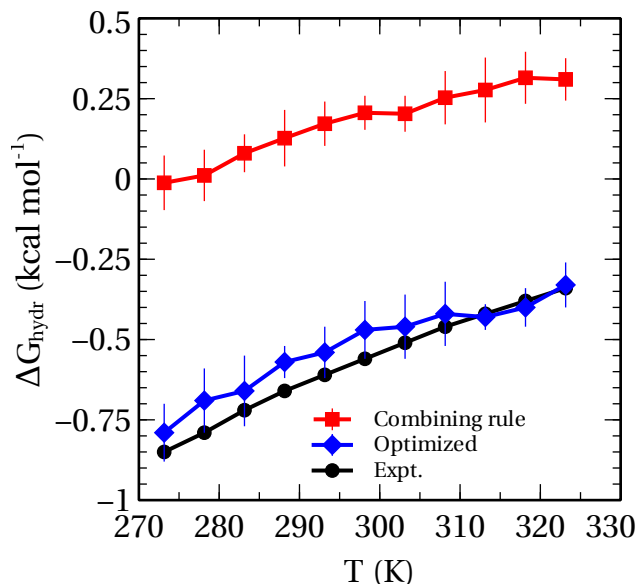
Figure 3.4: ΔG_{hydr} for H_2S over the $273 - 323$ K temperature range for Lennard-Jones parameters determined using the combining rule and parameters optimized for $\Delta G^{\circ}_{\text{hydr}}$.

The solvation energies calculated using the pairwise optimized S–O Lennard-Jones parameters are consistently more accurate than those calculated using Lennard-Jones parameters determined using the combining rule.

The components of $\Delta G^{\circ}_{\text{hydr}}$ provide additional insight into the nature of solvation (Ta-

Table 3.1: $\Delta G_{\mathrm{hydr}}^{\circ}$ of H_2O and H_2S calculated using FEP. All values are in kcal mol^{-1}.

$\Delta G_{\mathrm{hydr}}^{\circ}$	H_2O	H_2S
Electrostatic	-8.05	-0.95
Dispersive	-3.29	-7.08
Repulsive	5.62	7.54
Total	-5.72	-0.49

ble 3.1). We compare these components for the hydration of an H_2S molecule to the hydration of an H_2O molecule to illustrate the differences in their solvation. The electrostatic component of H_2S solvation is only -0.95 kcal mol^{-1} compared to -8.05 kcal mol^{-1} in H_2O solvation, where the electrostatic solvent–solute interaction is dominant. As H_2S has a deeper Lennard-Jones well, the dispersive interaction is more favorable than for a H_2O molecule: $\Delta G_{\mathrm{disp}} = -7.08$ kcal mol^{-1} for H_2S vs $\Delta G_{\mathrm{disp}} = -3.29$ kcal mol^{-1} for H_2O. The dispersive interaction for H_2S is largely canceled by a large repulsive component of 7.54 kcal mol^{-1}, reflecting the large molecular volume of H_2S. This shows that even though H_2S is highly polarizable, it has only limited electrostatic interactions with the water solvent due to its modest polarity and large size, resulting in primarily hydrophobic solvation.

3.3.2.3 Diffusion

The diffusion of H_2S in water is an important issue in process engineering and provides a test of how effective the Drude model is for describing H_2S mass transport in water. The diffusion coefficient values were calculated from Eqn. 1.31. The experimental diffusion coefficient of H_2S in water at 298 K is 1.75×10^{-5} cm^2 s^{-1}, which is somewhat smaller than the self-diffusion coefficient of water ($D = 2.3 \times 10^{-5}$ cm^2 s^{-1}),[39] but consistent with other hydrophobic solutes of similar molecular volume, like CH_4 and CO_2.[40] As Figure 3.5 shows, the diffusion coefficient calculated using the Drude model is in good agreement to the experimental data within the range of $293 - 333$ K.
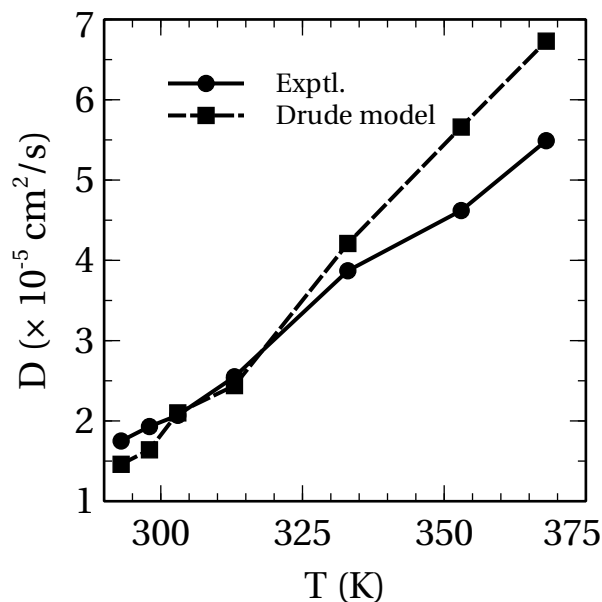
Figure 3.5: Diffusion coefficient of H_2S in water over the $293 - 368$ K temperature range.

3.3.3 Interfacial Properties

In order to investigate the interfacial properties of H_2S at the water/vapor interface, we calculated the free energy profile, the average dipole moment, and orientational distribution of a H_2S molecule as it moves from the vapor phase to the bulk water. For comparison, we also calculated these properties for a H_2O molecule. The system used in these calculations is described in the Theory and Methods section and is illustrated in Figure 3.1. These properties are calculated as function of Z, corresponding to the distance between the center of mass of water molecules and target molecule along the z-axis, normal to the interface. A position of $Z = 0$ Å corresponds to the target molecule being in the center of the slab — effectively a bulk solvent environment. We calculate these properties over the range 0 Å $< Z < 20$ Å, where $Z = 20$ Å corresponds to a point in the vapor phase roughly 10 Å from the interface.

3.3.3.1 Free Energy Profile

The free energy profile for the transition of H_2O and H_2S molecules through the interface is plotted in Figures 3.6a and 3.6b, respectively. These energies are referenced to the energy at $Z = 20$ Å, where the molecule in the vapor phase, so they correspond to the reversible work required to move the target molecule from the vapor phase to a given position along the Z coordinate. The end of the free energy profile, where $Z = 0$ Å, corresponds to the target molecule being at the center of the slab — effectively in the bulk solvent — so the profile at this point is approximately equal to $\Delta G^\circ_{\text{hydr}}$. The solvation energy of H_2S calculated from the free energy profile is -0.71 kcal mol^{-1}, in reasonable agreement with the $\Delta G^\circ_{\text{hydr}}$ of -0.47 ± 0.09 kcal mol^{-1} calculated using FEP.

The free energy profile reaches a minimum of ~ -1.9 kcal mol^{-1} in the range 7 Å$\leq Z \leq 13$ Å, which corresponds to H_2S being adsorbed at the water/vapor interface. The Gibbs energy of adsorption at the surface can be estimated from the difference in the free energy profile at this point and in the bulk solvent,[1] yielding $\Delta G^\circ_{\text{surf}} = 1.2$ kcal mol^{-1}. This surface adsorption energy is sizable for an uncharged solute and indicates that there is a significant surface excess for H_2S. This is consistent with experimental reports that determined that there is an adsorbed state for H_2S at the surface of water.[41;42] The barrier corresponding H_2S crossing from the bulk state to the surface-adsorbed state is only 0.3 kcal mol^{-1} and is not expected to significantly impede transfer of H_2S from the bulk to the surface. Past the interface, the H_2S water interaction diminishes rapidly, becoming effectively non-interacting ~ 8 Å from the interface. We see no barrier on the free energy profile to the adsorption of H_2S molecules to the interface from the vapor phase.

3.3.3.2 Dipole Moment

Polarizable force fields provide insight into the polarization of a solute as it moves from the gas phase into a solvent. The average dipole moment of H_2O and H_2S as a function of the
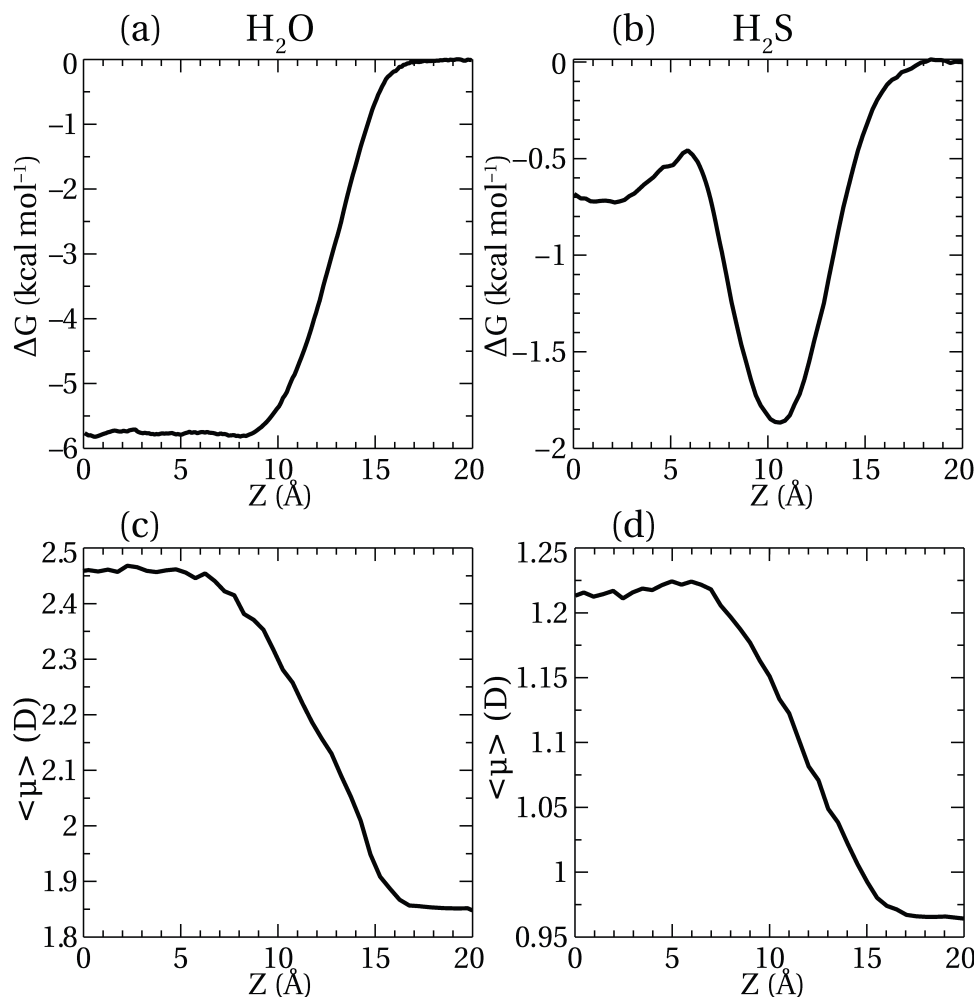
Figure 3.6: The free energy profile (top) and average dipole moments (bottom) of H_2O (a,c) and H_2S (b,d) molecules as they move across the water/vapor interface. Z is the distance along the z-axis (normal to the plane of the interface), with respect to the center of mass of the water slab. $Z = 0$ Å corresponds to the bulk water environment and $Z = 20$ Å corresponds to the vapor phase. The water/vapor interface occurs between 7 Å $< Z <$ 13 Å.

Z coordinate is plotted in the Figures 3.6c and 3.6d, respectively. The dipole moment of H_2S increases from its gas phase value of 0.98 D to the $1.1 - 1.2$ D range when adsorbed to the water surface and then increases to its maximum value of 1.23 D in solution. This polarization is somewhat smaller in magnitude than for H_2O; the average dipole of SWM4-NDP H_2O increases from 1.85 D in the gas phase to 2.46 D in the bulk liquid due to strongly polarizing hydrogen bonding interactions. The large molecular volume of H_2S and the
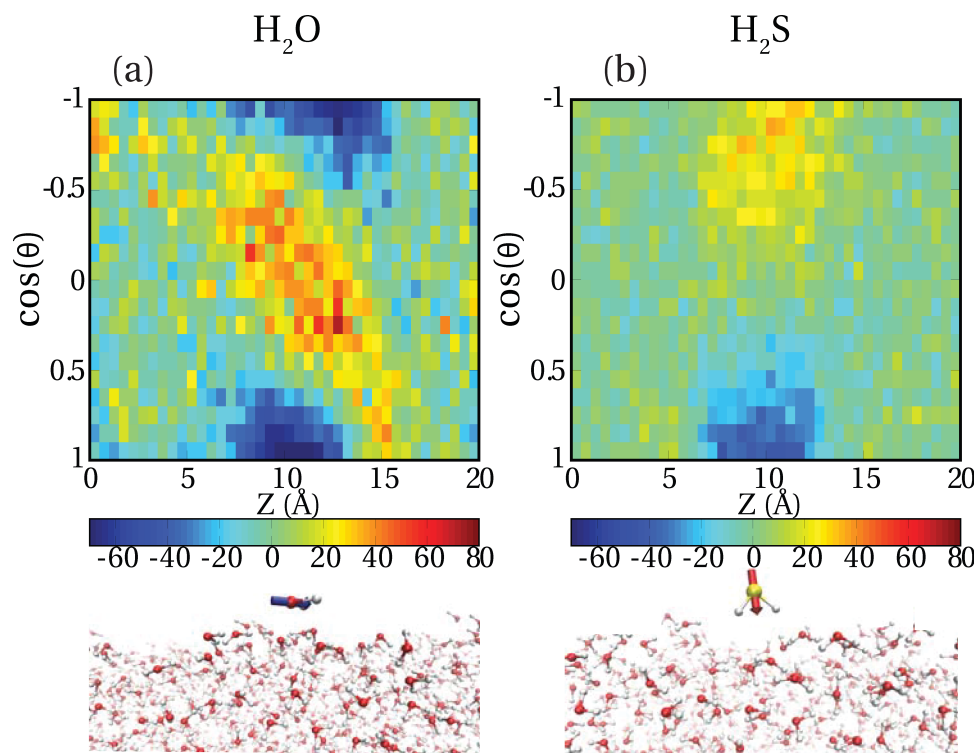
Figure 3.7: Orientational probability distributions of H_2O (left) and H_2S (right) molecules at the water/vapor interface. Z corresponds to the position of the target molecule, H_2O or H_2S, with respect to the center of mass of the water slab. $Z = 0$ Å is at the center of the slab, $Z = 20$ Å is in the vapor phase, and the water/vapor interface occurs between 7 Å $< Z <$ 13 Å. The y-axis corresponds to the cosine of the angle between the plane of the target molecule and the xy-plane.

limited hydrogen bonding between H_2S and the water solvent limits the degree of induced polarization.

3.3.3.3 Orientational Distribution

Another interesting feature of the liquid/vapor interface is the orientational distributions of molecules at the interface. Figures 3.7a and 3.7b show the probabilities of the molecular axis of having an angle with respect to the z-axis, relative to the uniform distribution, for H_2O and H_2S, respectively. There is no orientational preference for molecules in the gas or bulk solution phases; however, molecules at the interface (7 Å $\leq Z \leq$ 13 Å) show

60

non-uniform orientational distributions.

There is a $20 - 40\%$ increase in orientations where H_2S is aligned perpendicular to the surface with the S atom upwards and a corresponding decrease in perpendicular orientations where the S atom is closest to the surface. This inverse orientation has also been reported for chloromethane and bromomethane.[43,44] This is in contrast to molecules like ammonia and •OH that are preferentially oriented on the water/vapor interface such that they accept a hydrogen bond from a dangling surface O–H bond.[45,3] Given that the hydrogen bond accepting ability of hydrogen sulfide is poor, the orientation with the sulfur pointing away from the surface may be favored simply because it maximizes the interaction between the dipole of H_2S and water molecules at the interface while avoiding short range repulsive interactions between the sulfur atom and water molecules.

In comparison, surface water molecules have much higher orientational biases than surface-adsorbed H_2S molecules. There is a $40 - 80\%$ increase in the probabilities of H_2O molecules at the interface holding orientations where they are aligned parallel to the interface. This orientation allows for the hydrogen bonding interactions between surface water molecules to be maximized.[46,8] This repeats the trend we observed throughout this study, where hydrated H_2S molecules tend not to engage in the same type of strong hydrogen bonding interactions as water molecules.

3.3.3.4   Surface Tension

Molecules with high surface excesses act as surfactants, decreasing the surface tension of the interface they are adsorbed on.[47] Experiments by Broseta et al.[48] determined that the surface tension of water at 313 K decreases to 16 dyn cm^{-1} when exposed to gaseous H_2S with a density of 764 kg m^{-3}. The surface tension under these conditions was calculated by performing a simulation of a periodic cell containing a liquid slab of water in contact with a vapor phase of H_2S with a density of 764 kg m^{-3} from Eqn. 1.34. This corresponds to

250 H$_2$O molecules and 202 H$_2$S molecules, as illustrated in Figure 3.8.
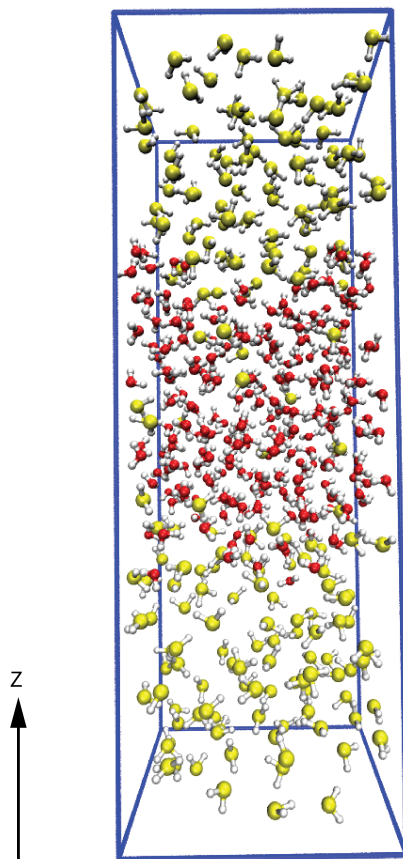
Figure 3.8: A representative configuration of the system used to calculate the surface tension of water under $\rho = 764$ kg m^{-3} H$_2$S(g).

The surface tension of pure water SWM4-NDP water is 63 dyn cm^{-1} at 313 K, but the calculated surface tension in the presence of the H$_2$S gas is decreased to 20 dyn cm^{-1}, in good agreement with the experimental result of 16 dyn cm^{-1}.

3.4 Conclusion

In this study, we determined optimal Lennard-Jones parameters for modeling a heterogeneous H$_2$S–H$_2$O system using the Drude polarizable force field. These parameters were used to predict the solution and interfacial behavior of H$_2$S in water. This model is able to

predict the ΔG_{hydr} of H_2S in good agreement with experimental gas solubility data from 273 to 323 K. The diffusion coefficient of H_2S in water is also predicted accurately between 298 and 368 K.

Analysis of the S–O and S–H radial distribution functions calculated using the Drude model and AIMD trajectories indicated that the hydrogen bonding between H_2S and solvent H_2O molecules is limited. Instead, the solvation structure of H_2S in aqueous solution is more similar to that of a hydrophobic solute, with the water molecules forming an "iceberg"-like hydrogen bonding network around the H_2S.

These simulations showed that H_2S is increasingly polarized by water as it crosses the water/vapor interface, with the average dipole moment of H_2S increasing from its gas phase value of 0.98 D to 1.25 D in bulk water. There is a significant orientational preference for H_2S molecules adsorbed at the surface, which tend to align perpendicular to the surface with the sulfur atom pointing away from the interface. The free energy profile along the coordinate corresponding to H_2S transitioning from the vapor to the liquid phase shows a large surface excess, with the free energy profile holding a minimum value of -1.9 kcal mol^{-1} at the interface but only -0.7 kcal mol^{-1} in the bulk solvent. This is in agreement with experimental and simulation data that indicates that H_2S is a powerful surfactant, decreasing the surface tension of bulk water by roughly 40 dyn cm^{-1} at high H_2S gas densities.

Bibliography

[1] Garrett, B. C.; Schenter, G. K.; Morita, A. *Chem. Rev.* **2006**, *106*, 1355–1374.

[2] Tobias, D. J.; Stern, A. C.; Baer, M. D.; Levin, Y.; Mundy, C. J. *Annu. Rev. Phys. Chem.* **2013**, *64*, 339–359.

[3] Vácha, R.; Slavíček, P.; Mucha, M.; Finlayson-Pitts, B. J.; Jungwirth, P. *J. Phys. Chem. A* **2004**, *108*, 11573–11579.

[4] Vrbka, L.; Mucha, M.; Minofar, B.; Jungwirth, P.; Brown, E. C.; Tobias, D. J. *Curr. Opin. Colloid Interface Sci.* **2004**, *9*, 67–73.

[5] Wilson, M. A.; Pohorille, A. *J. Phys. Chem. B* **1997**, *101*, 3130–3135.

[6] Morita, A. *Chem. Phys. Lett.* **2003**, *375*, 1–8.

[7] Dang, L. X.; Feller, D. *J. Phys. Chem. B* **2000**, *104*, 4403–4407.

[8] Chang, T.-M.; Dang, L. X. *J. Chem. Phys.* **1996**, *104*, 6772–6783.

[9] Shelley, J. C.; Sprik, M.; Klein, M. L. *Langmuir* **1993**, *9*, 916–926.

[10] Marrone, T. J.; Hartsough, D. S.; Merz, K. M. *J. Phys. Chem.* **1994**, *98*, 1341–1343.

[11] Olney, T. N.; Cann, N.; Cooper, G.; Brion, C. *Chem. Phys.* **1997**, *223*, 59–98.

[12] Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; Jr., A. D. M. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

[13] Neyt, J.-C.; Wender, A.; Lachet, V.; Ghoufi, A.; Malfreyt, P. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11679–11690.

[14] Archontis, G.; Leontidis, E.; Andreou, G. *J. Phys. Chem. B* **2005**, *109*, 17957–17966.

[15] Riahi, S.; Rowley, C. N. *J. Phys. Chem. B* **2013**, *117*, 5222–5229.

[16] Brooks, B. R. et al. *J. Comp. Chem.* **2009**, *30*, 1545–1614.

[17] Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

[18] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327–341.

[19] Tuckerman, M. E.; Berne, B. J. *J. Chem. Phys.* **1991**, *95*, 8362–8364.

[20] VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comp. Phys. Comm.* **2005**, *167*, 103–128.

[21] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[22] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys* **2010**, *132*, 154104.

[23] Shirts, M. R.; Mobley, D. L.; Chodera, J. D. In *Free Energy Calculations*; Spellmeyer, D., Wheeler, R., Eds.; Annual Reports in Computational Chemistry; Elsevier, 2007; Vol. 3; pp 41–59.

[24] Deng, Y.; Roux, B. *J. Phys. Chem. B* **2004**, *108*, 16567–16576.

[25] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

[26] Torrie, G.; Valleau, J. *J. Comp. Phys.* **1977**, *23*, 187–199.

[27] Frenkel, D.; Smit, B. *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science)*, 2nd ed.; Academic Press, 2001.

[28] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 1989.

[29] Song, W.; Rossky, P. J.; Maroncelli, M. *J. Chem. Phys.* **2003**, *119*, 9145–9162.

[30] Baker, C. M.; Lopes, P. E. M.; Zhu, X.; Roux, B.; MacKerell, A. D. *J. Chem. Theory Comput.* **2010**, *6*, 1181–1198.

[31] Guillot, B.; Guissani, Y. *J. Chem. Phys.* **1993**, *99*, 8075–8094.

[32] Broadbent, R. D.; Neilson, G. W. *J. Chem. Phys.* **1994**, *100*, 7543–7547.

[33] Jorgensen, W. L.; Gao, J.; Ravimohan, C. *J. Phys. Chem.* **1985**, *89*, 3470–3473.

[34] Sarma, R.; Paul, S. *J. Phys. Chem. B* **2012**, *116*, 2831–2841.

[35] Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945**, *13*, 507–532.

[36] Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521–533.

[37] Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683–3704.

[38] Clarke, E. C. W.; Glew, D. N. *Can. J. Chem.* **1971**, *49*, 691–698.

[39] Krynicki, K.; Green, C. D.; Sawyer, D. W. *Faraday Discuss. Chem. Soc.* **1978**, *66*, 199–208.

[40] Jähne, B.; Heinz, G.; Dietrich, W. *J. Geophys. Res.* **1987**, *92*, 10767–10776.

[41] Herrick, C. S.; Gaines, G. L. *J. Phys. Chem.* **1973**, *77*, 2703–2707.

[42] Strathdee, G. G.; Given, R. M. *J. Phys. Chem.* **1976**, *80*, 1714–1719.

[43] Harper, K.; Minofar, B.; Sierra-Hernandez, M. R.; Casillas-Ituarte, N. N.; Roe-selova, M.; Allen, H. C. *J. Phys. Chem. A* **2009**, *113*, 2015–2024.

[44] Habartová, A.; Valsaraj, K. T.; Roeselová, M. *J. Phys. Chem. A* **2013**, *117*, 9205–9215.

[45] Fu, C.-F.; Tian, S. X. *J. Phys. Chem. C* **2013**, *117*, 13011–13020.

[46] Wilson, M. A.; Pohorille, A.; Pratt, L. R. *J. Chem. Phys.* **1988**, *88*, 3281–3285.

[47] Israelachvili, J. N. *Intermolecular and Surface Forces*, 3rd ed.; Academic Press, 2010.

[48] Shah, V.; Broseta, D.; Mouronval, G.; Montel, F. *Int. J. Greenhouse Gas Control* **2008**, *2*, 594–604.

Chapter 4

QM/MM Molecular Dynamics Simulations of the Hydration of Mg(II) and Zn(II) Ions

Portions of this chapter have been published as: Riahi, S., Roux, B., Rowley, C.N. QM/MM Molecular Dynamics Simulations of the Hydration of Mg(II) and Zn(II) Ions. *Can. J. Chem.* 91(7), 552–558, 2013**.** The computation of the properties were performed by the author. The preparation and editing of the manuscript were perforemd by the colaboration of the author, C. Rowley, and B. Roux.

4.1   Introduction

Aqueous solutions of different ions, especially transition metal ions, show significant differences that are visible in their biological functions. Properties like the strength and stability of ion–water bond are significantly different in transition metal ions. These differences can

only be rationalized by exploring the physics of these systems at the atomic and molecular level. Molecular dynamics simulations are frequently used to examine the structure and dynamics of solvated ions. These simulations typically employ computationally efficient molecular mechanical (MM) models so that the full complexity of the ion and its environment can be represented. Small, divalent ions like Mg^{2+} and Zn^{2+} pose serious challenges for conventional molecular mechanical force fields because the ion–ligand distances are exceptionally short (e.g. $r_{M-O} \approx 2$ Å in water). At these distances, the divalent charge of these ions will strongly polarize the coordinating water molecules, so a conventional non-polarizable MM force field will not accurately describe this interaction. To address this issue, parameters for the interaction of Mg^{2+} and Zn^{2+} with water have been determined for both the AMOEBA[1;2] and Drude force fields,[3] which include terms to allow for the effects of induced electron polarization. These models show improved ion–water interaction energies and free energies of hydration.

Although polarizable force fields are capable of approximating the effects of induced polarization, several additional parameters must be defined based on limited experimental or quantum mechanical target data. As a result, it is not clear how accurately these models describe the structural features of the ionic solvation structure. Further, the absolute ion solvation free energies used to parameterize and validate these models are difficult to ascertain experimentally, introducing a degree of error. Experimental techniques for characterizing ion-water structure, such as neutron diffraction or X-ray scattering, provide only coarse structural features, so these models cannot be thoroughly validated by experimental data alone.

Ab initio molecular dynamics (AIMD) is an alternative strategy to examine the structure and thermodynamics of solvated ions. By performing a molecular dynamics simulation using a quantum mechanical (QM) representation of the ion and solvent, these methods can sample the configurational space of a solvated ion without the definition of force field pa-

rameters. Electron polarization and charge transfer effects are inherently included in these models. The downside of these QM models is that they are much more computationally expensive than MM models, so the number of solvent molecules that can be included is more limited and simulation time scales are much shorter. Further, these QM models are also inexact due to approximations in the density functional theory (DFT) exchange-correlation functionals, pseudopotential representations, and truncation of the basis set. These limitations can affect the calculated ion solvation structure, so previous reports based on AIMD simulations of Mg^{2+} and Zn^{2+} have a range of conclusions.[4;5;6;7;8;9;10;11] Hybrid QM/MM simulations are an obvious means to perform ab initio simulations of ion solvation at a reduced computational cost. There is a natural division of QM and MM regions, as the ion and the solvent molecules nearest to it can be represented using the QM methods while the rest of the solvent can be represented using the MM model. We employed this strategy in a recent study where high level QM/MM molecular dynamics simulations were used to model the solvation of Na^+ and K^+ in liquid water.[12] We attempted to limit the approximations in this model by using a large all-electron triple-ζ basis set and embedding the ion and nearest water molecules in a 14 Å sphere of polarizable MM water molecules.

The hydration of Mg^{2+} and Zn^{2+} are ideal subjects for this kind of study, as complex electronic effects are important in these ions and they have a small inner coordination sphere. In particular, an accurate first-principles description of their characteristic structural distribution functions, with a consistent computational methodology would establish how similar the solvation structure of these ions is. Additionally, calculating the relative solvation free energy of these ions using thermodynamic integration would test the surprising experimental result that Zn^{2+} has a more negative solvation free energy than Mg^{2+}. In this chapter, we report extended QM/MM molecular dynamics simulations on Mg^{2+}(aq) and Zn^{2+}(aq). The QM system was described using DFT and a large basis set, embedded in an MM sphere of Drude polarizable water molecules. These models are compared to the

results from the CHARMM non-polarizable and Drude polarizable force fields.

4.2   Computational Methods

All MM calculations were performed using CHARMM[13] version c37b2. QM/MM calculations were performed using CHARMM c37b2 interfaced with TURBOMOLE 6.3[14] through the QTURBO module. The equations of motion were propagated using the velocity Verlet algorithm,[15] modified by a Langevin thermostat to sample the canonical ensemble of configuration at a temperature of 298.15 K. For systems containing Drude particles, a dual Langevin thermostat was applied, where the first thermostat maintained the temperature of the atomic centers at 298.15 K and the second thermostat maintained the temperature of the Drude particles at 1 K. The ion was restrained at the center of a sphere of 451 water molecules with a radius of 14 Å. The water molecules were confined to this sphere by a half-cubic restraining function. All water molecules were constrained to their equilibrium liquid-phase bond lengths and angles using the SHAKE algorithm.[16] A 1 fs time step was used in all simulations.

The non-polarizable models used the TIP3P model for water[17] and standard CHARMM force field parameters for Mg^{2+} and Zn^{2+}. The Lennard-Jones parameters for these ions were $E_{min} = 0.015$ kcal mol^{-1} $R_{min} = 2.37$ Å for Mg^{2+} and $E_{min} = 0.25$ kcal mol^{-1} $R_{min} = 2.18$ Å for Zn^{2+}. The Drude polarizable models used the SWM4-NDP model for water[18] and the ionic parameters of Yu et al.[3] The non-polarizable and Drude MM models were equilibrated for 1 ns followed by a production run of 4 ns.

The QM/MM calculations employed the same sphere of water molecules as the MM models. The QM region included the ion and the six nearest water molecules. All the remaining water molecules were represented using the SWM4-NDP Drude polarizable water model.[18] SWM4-NDP is an accurate model for the mechanical, transport, and thermody-

namic properties of bulk water. In particular, the static dielectric constant of SWM4-NDP water is consistent with the experimental value of $\epsilon = 78$, whereas many non-polarizable models significantly overestimate the dielectric constant.

A QM region of six waters was chosen because previous reports found both Mg^{2+} and Zn^{2+} to have inner coordination spheres that are consistently hexacoordinate. An earlier QM/MM study of Zn^{2+}(aq) by Cauet et al.[9] found that there was no advantage to using a QM region larger than six water molecules. A report by Callahan et al. showed that $MgCl_2$ solutions do not form contact ion pairs even at very high concentrations, so we have focused on the infinitely dilute case of a single solvated dication.[19] This small QM region allows us to use a larger basis set and perform longer simulations. The QM region was represented using DFT with the PBE exchange-correlation functional.[20] This functional was chosen as it is non-empirical and has performed well across a wide range of chemical systems. As this is a pure functional and no exchange integrals were needed, it was possible to employ the highly efficient resolution of identity (RI) approximation.[21] The triple-ζ def2-TZVPP basis set was assigned for all atoms to limit basis set truncation and basis set superposition error.[22] An SCF convergence criteria of 10^{-7} was imposed to ensure accurate energies and gradients and the m5 grid was used for integration in the exchange-correlation term. The effect of the MM point charges to the QM region was calculated by including the one-electron integrals between the MM point charges and the QM basis functions in the Fock matrix. Lennard-Jones interactions between the QM and MM waters were adjusted to recreate the QM dimerization energies, as described in Ref. 12. Although we have developed an effective and rigorous method to prevent exchange of QM and MM water molecules from the solvent,[12] it was not necessary to use this method in this case because no such exchanges occurred during the time scale of our QM/MM simulations. These QM/MM models were equilibrated for 2 ps followed by a 98 ps production run. Although shorter simulations have been used in previous AIMD simulations, we performed these

longer simulations to ensure the distributions were sampled sufficiently.

4.2.1 Thermodynamic Integration

The difference in solvation free energies of Mg^{2+}(aq) and Zn^{2+}(aq) was calculated using alchemical thermodynamic integration, using a linear interpolation between the Mg^{2+}(aq) and Zn^{2+}(aq) states. 11 simulations were performed at values of λ between 0.0 and 1.0 separated by increments of 0.1. For the MM models, each value of λ was equilibrated for 400 ps followed by a 1 ns production period.

$$\Delta G_{Mg^{2+}(aq) \rightarrow Zn^{2+}(aq)} = \int_{\lambda=0}^{\lambda=1} \langle U_{Zn} - U_{Mg} \rangle_{\lambda} \, \mathrm{d}\lambda \tag{4.1}$$

The QM/MM relative solvation free energies were calculated using an analogous thermodynamic integration procedure, where the path between the QM/MM Mg^{2+} system ($\lambda = 0$) and QM/MM Zn^{2+} ($\lambda = 1$) system was calculated by interpolating between the two QM/MM Hamiltonians. The difference in electronic energy of the bare Mg^{2+} and Zn^{2+} ions was calculated separately and subtracted from the energy differences. This method is described in detail in an excellent review by Salahub and coworkers.[23] For each value of λ, the system was equilibrated for 2 ps before a 15 ps production period.

The free energy differences of each system were calculated from Eqn. 1.26 on the time series of our free energy perturbation (FEP) simulations data using the WHAM.[24] The intrinsic solvation free energies were corrected for the phase potential of the water models using the water model interfacial potentials determined in Ref. 25.

4.2.2 Absolute Solvation Free Energies

The absolute solvation free energies for the non-polarizable and Drude MM models were calculated using FEP/MD simulations by decoupling a single ion from a solvent of 450

(24 Å × 24 Å × 24 Å) water molecules under periodic boundary conditions. For electrostatic interactions, particle mesh Ewald (PME) scheme with $\kappa = 0.33$ and sixth order spline function was applied. Lennard-Jones and the real space electrostatic non-bonded interactions were calculated using a switching function between distances of 10 Å and 12 Å. For each window, a 400 ps equilibriation simulation was performed, followed by a 1 ns sampling period, with a time step of 1 fs. These simulations were performed using isothermal–isobaric molecular dynamics, with a Nosé–Hoover thermostat.[26] In the case of the Drude polarizable model, a dual thermostat was used, where the first thermostat at $T = 298.15$ K was coupled to the heavy atoms with a relaxation time of 0.1 ps and the second thermostat at $T^* = 1$ K with was coupled to the Drude particles with a relaxation time of 0.005 ps for the Drude particles. Fictious masses of 0.4 amu were assigned to the Drude particles. The Andersen–Hoover barostat[27] with a relaxation time of 0.2 ps was used to maintain a pressure of a 1 atm.

In order to calculate absolute energy of hydration of Mg^{2+} and Zn^{2+}, we used the decoupling methodology of Deng and Roux.[28] In this technique, the absolute hydration free energy the ion $\Delta_{hydr}G$ is calculated as the sum of three contributions corresponding to the electrostatic, dispersive, and repulsive interactions between the ion and solvent,

$$\Delta_{hydr}G = \Delta G_{elec} + \Delta G_{disp} + \Delta G_{rep} \qquad (4.2)$$

The electrostatic component (ΔG_{elec}) was calculated using thermodynamic integration of a path where the charge of the ion was reduced to zero through the coupling parameters λ. This path was calculated over 11 values of λ in increments of 0.1 between 0 and 1. An analogous FEP path as used to calculate the dispersion component of the Lennard-Jones interaction (ΔG_{disp}) using the Weeks–Chandler–Andersen decomposition. For the repulsive component (ΔG_{rep}), the staging parameter, s, was used to calculate the free energy

Table 4.1: Ion–O distances for Mg^{2+}(aq) and Zn^{2+}(aq). Values from computer simulations are taken as radius where the maximum of the M–O rdf occurs.

Method	Mg–O (Å)	Zn–O (Å)
Non-polarizable	1.98	2.09
Drude	2.05	2.03
QM/MM	2.11	2.13
AMOEBA	2.07[1]	1.98[2]
AMOEBA-VB	-	2.06[29]
Previous AIMD	2.10[4]	2.18[9]
	2.13[5]	2.18[10]
	2.08[6]	2.09[11]
	2.08[7]	
	2.13[8]	
exptl.	2.10[30]	2.08[31]

change at values of 0.0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. For each value of s, simulations were performed at $\lambda = 0$ and $\lambda = 1$. Using the time series of the FEP data from these simulations, ΔG_{elec} was calculated using trapezoidal rule, while ΔG_{disp} and ΔG_{rep} were calculated using WHAM.[24]

4.3 Results and Discussion

4.3.1 Ion Solvation Structure

4.3.1.1 Radial Distribution Functions and Coordination Numbers

The ion–O radial distribution function (rdf, $g(r)$) is a correlation function that relates the density of the solvent at a distance, r, from the ion, to the bulk density. Values less than one indicate a depletion of solvent density, while values greater than one indicate an increase. The location of the first coordination sphere of an ion is easily identifiable from the rdf, as there is a sharp peak at this distance. The rdfs calculated from our molecular dynamics
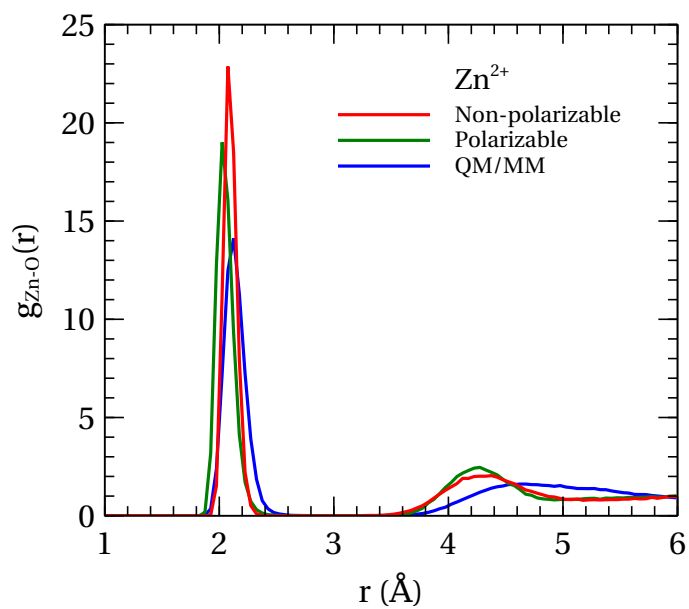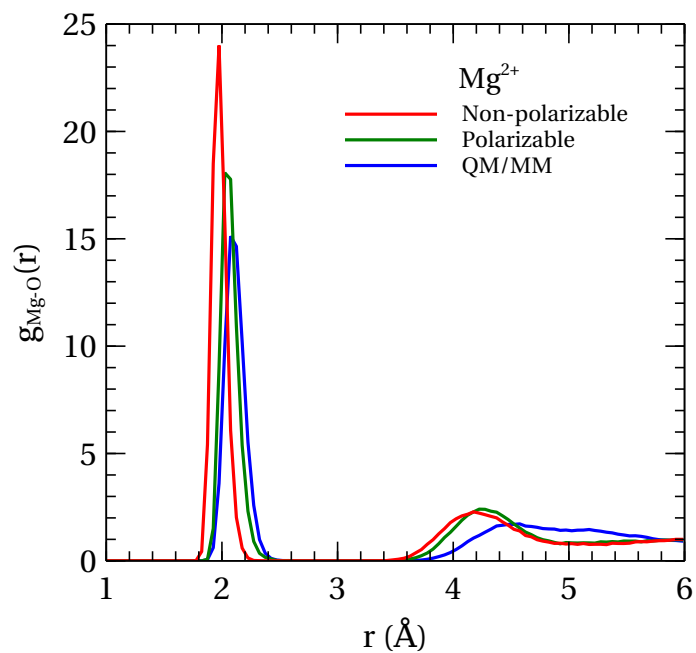
Figure 4.1: M–O radial distribution functions calculated from MD simulations with the CHARMM non-polarizable force field, Drude polarizable force field, and combined QM/MM approach.

simulations are presented in Figure 4.1.

The rdfs of Mg^{2+}(aq) are in generally good agreement, with a sharp peak for the first

coordination sphere near 2.1 Å. This distance is consistent with neutron diffraction data that gives a Mg–O distance of 2.10 Å[30] and with most previous AIMD simulations (Table 4.1).[1]

Table 4.2: Binding energy and M–O distance of $M(OH_2)_6^{2+}$.

Model	ΔE_{Mg} (kcal mol^{-1})	Mg–O (Å)	ΔE_{Zn} (kcal mol^{-1})	Zn–O (Å)
Non-polarizable	–340.1	1.98	–310.8	2.10
Drude	–323.1	2.08	–334.6	2.06
PBE/def2-TZVPP	–325.3	2.11	–356.5	2.13

These distances are close to those calculated from the optimized structure of $Mg^{2+}(OH_2)_6$ (Table 4.2), which shows that the inner coordination sphere remains near its gas phase potential energy minimum. The integral of this peak over spherical coordinations gives a coordination number of 6.0, indicating that the ion maintains a constant coordination number of 6 throughout the simulation. A representative configuration from the QM/MM simulation of Mg^{2+}(aq) that illustrates this inner sphere is presented in Figure 4.2.

The non-polarizable force field predicts the Mg^{2+}–O distance to be slightly shorter than the other methods, with a maximum near 2.0 Å. In contrast to alkali cations, there is a sharp separation between the first coordination sphere and the rest of the solvent; the rdf has a value of zero until the second coordination sphere that occurs near $r = 4.0$ Å. The rdfs calculated from the QM/MM and Drude simulations are in very good agreement, although the first peak of the Drude rdf is higher and more narrow the the QM/MM model. This reflects that the MM models use a Lennard-Jones potential to represent repulsive interactions, which is a harder repulsive force than Pauli repulsion forces are in reality.

The rdfs of Zn^{2+}(aq) show quite similar trends. In this case, all models predict a sharp peak for the first coordination sphere with centers that range between 2.09–2.13 Å. This distance is consistent with the experimental EXAFS and XANES spectra that indicate an

---

[1]The neutron diffraction data was collected from 0.2 M MgCl$_2$. Callahan et al.[19] determined that ion pairing is not significant at these concentrations.
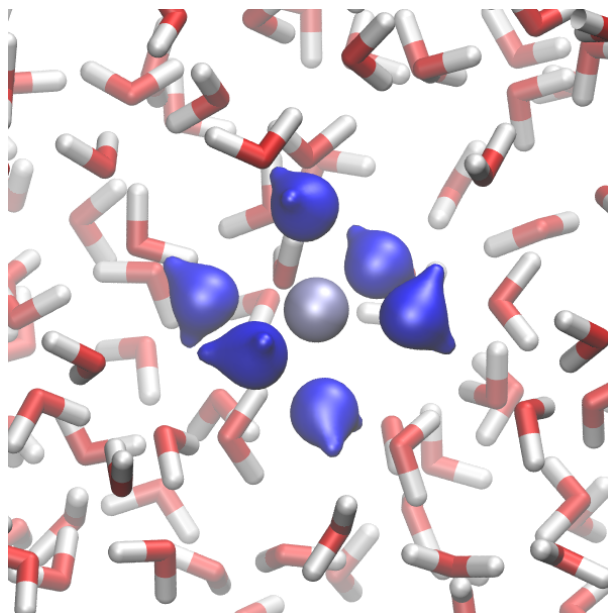
Figure 4.2: Representative configuration of QM/MM model of Mg^{2+}(aq). The electron density distribution for the QM region is shown in blue. The Mg^{2+} ion is shown in light blue.

average Zn^{2+}–O distance of 2.08 Å.[31;32] [2] Our simulations are in better agreement with the experimental value than some other AIMD simulations, which predict larger Zn^{2+}–O distances (Table 4.1). The Zn^{2+}–O distance appears to be sensitive to the QM method and basis set. Our use of an all electron triple-ζ basis set may explain the improved accuracy of our simulations. As for the Mg^{2+}(aq) rdf, the non-polarizable MM model predicts a sharper first peak than the other methods and the QM/MM has the broadest peak, reflecting a broader range of accessible configurations. The spherical integral of the Zn^{2+}–O radial distribution function is exactly 6.0 for all these methods, indicating that Zn^{2+} is also consistently predicted to be six coordinate.

---

[2]EXAFS measurements in Ref. 31 were made of 0.2 M solutions of $ZnNO_3$, where ion pairing effects are not likely to be signficant.

4.3.1.2 Angular Distribution Functions

The angular distribution function (adf) is the probability distribution of the angles formed between the oxygen atoms of inner-sphere water molecules and the ion. High probabilities of finding water molecules trans to each other ($\theta = 180°$, $\cos\theta = -1$) and cis to each other ($\theta = 90°$, $\cos\theta = 0$) are indicative of an octahedral coordination mode. The adfs calculated from our molecular dynamics simulations are presented in Figure 4.3.

Figure 4.3: Water–ion–water angle distribution functions for waters in the first coordination sphere ($r_{M-O} < 2.5$ Å).

For Mg^{2+}, the adf of all methods show narrow, discrete distributions centered at $\cos\theta = -1$ and $\cos\theta = 0$. This indicates that at 298.15 K, Mg^{2+} makes only moderate oscillations around an octahedral coordination structu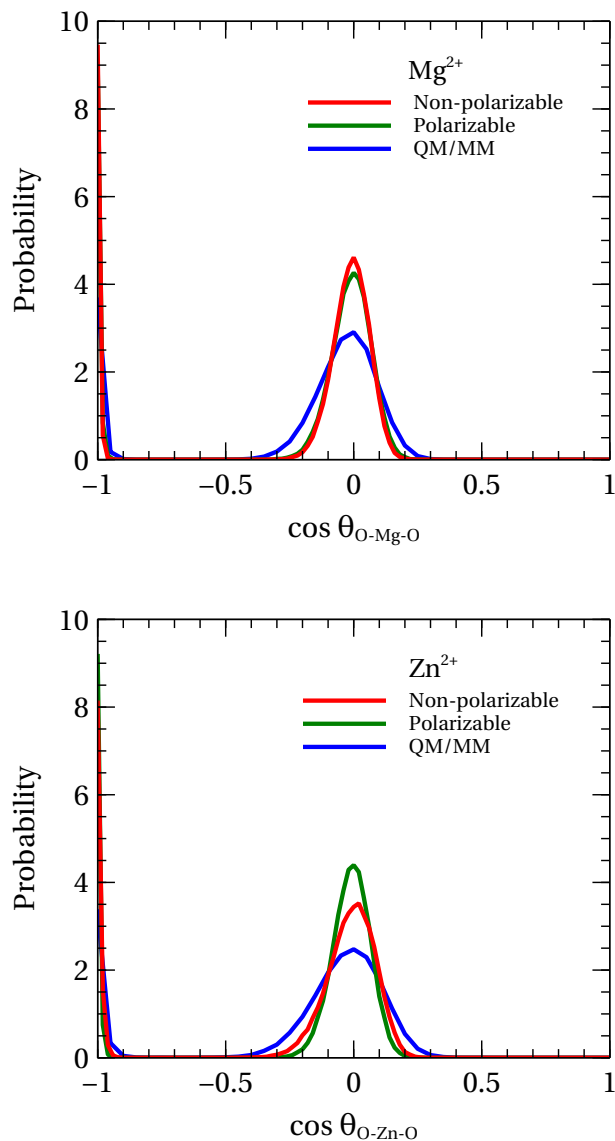re. The CHARMM and Drude MM models are in close agreement, with very sharp peaks in the distribution functions, although the QM/MM distribution is broader. These broader distributions can be explained by the rigorous treatment of electron polarization in the QM/MM model, which can stabilize configurations with more acute O–M–O angles. The approximate treatment of polarizability in the Drude model is not able to reproduce this effect fully.

The adfs of Zn^{2+} show greater variety. The QM/MM distribution very similar to that of QM/MM simulation for Mg^{2+}(aq), holding a fairly narrow six-coordinate octahedral orientation. The MM models yield more narrow distributions, although in contrast to the distributions for Mg^{2+}(aq), the simulation with the Drude model produces a more narrow distribution than the non-polarizable model. The use of a large Thole screening coefficient to correct for overpolarization in the Zn^{2+}–OH_2 interaction may account for this difference.

4.3.1.3 Tilt Distribution Function

The tilt distribution function shows the distribution of the angle (φ) formed between the ion, the oxygen atom of the inner sphere water molecules, and the bisector of the O–H bonds. When $\cos\varphi$ is equal to -1, the axis of the water molecule is aligned with the ion. This distribution is a measure of the strength of the ion-water charge-dipole interaction, which is strongest when the dipole moment of the water is aligned with ionic charge. The tilt distribution functions calculated from our molecular dynamics simulations are presented in Figure 4.4.

The tilt distributions of the Drude and QM/MM simulations of Mg^{2+}(aq) are in reasonably good agreement, with distributions that are peaked at $\cos\varphi = -1$ ($\varphi = 180°$) and decay to a low probability near $\cos\varphi = -0.5$ ($\varphi = 135°$), although the QM/MM distri-
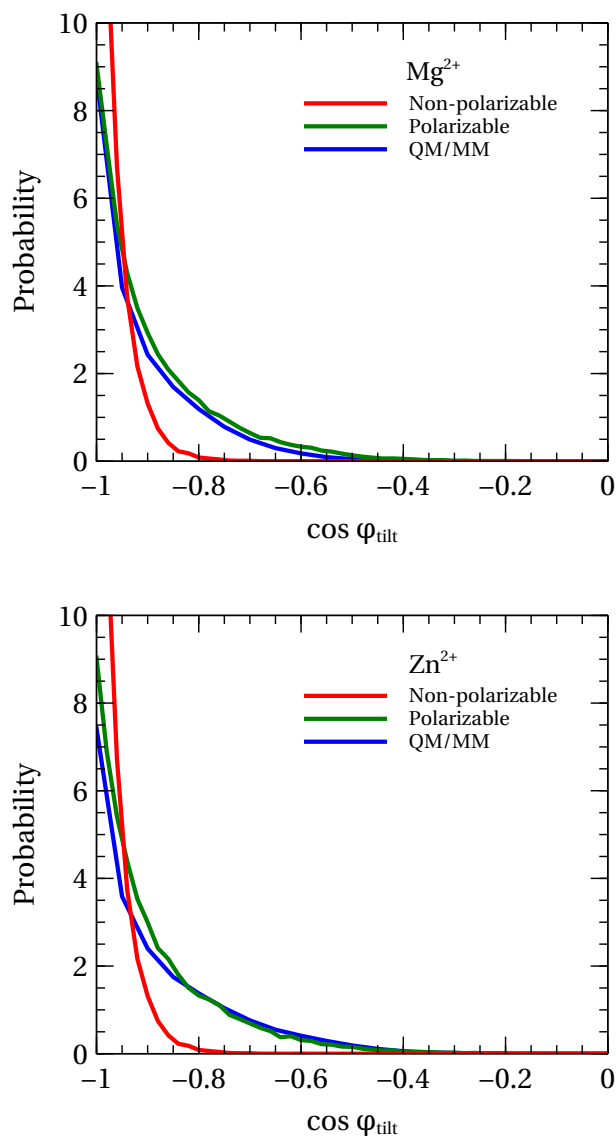
Figure 4.4: Tilt angle distributions (ion–O–H bisector) for waters in the first coordination sphere ($r_{M-O} < 2.5$ Å).

bution is slightly broader. The non-polarizable CHARMM MM distribution is sharpest, with the distribution decaying to zero near $\cos\varphi = -0.25$ ($\varphi = 104°$). This reflects the high permanent dipole of the TIP3P water model ($\mu_\circ = 2.347$ D) and the lack of induced electron polarization in this model.

The tilt distributions of Zn^{2+}(aq) show very similar trends and are generally very similar to the tilt distributions for Mg^{2+}(aq). This suggests that the tilt distributions of these ions

is primarily dependent on the size and valency of the ions and the description of electron polarization in the coordinating water molecules.

4.3.2 Solvation Free Energies

Absolute ion solvation free energies are difficult to calculate because the large change in free energy necessitates very long simulation times to achieve sampling to convergence. Relative solvation free energies are more straightforward, as they can be determined using thermodynamic integration of an alchemical path between two similar states. These relative solvation free energies can also be compared readily to the experimental relative solvation energies of two neutral salts with the same anion (e.g. $MgCl_2$ and $ZnCl_2$), which can be determined accurately without extrathermodynamic assumptions. We used the thermodynamic integration technique to calculate the relative hydration free energies of Mg^{2+} and Zn^{2+} for the non-polarizable, Drude, and QM/MM models (Table 4.3).

Table 4.3: Relative free energies of hydration of $Mg^{2+} \rightarrow Zn^{2+}$ ($\Delta\Delta G_{hydr}$).

Method	$\Delta\Delta G_{hydr}$ (kcal mol^{-1})
CHARMM	31.1
Drude	−13.2
AMOEBA	−27.8
QM/MM	−34.8
exptl.[3]	−30.2

The $\Delta\Delta G_{hydr}$ calculated using the QM/MM approach is in good agreement with experiment. This provides support for the experimental result that the relative solvation free energy of the two ions is roughly 30 kcal mol^{-1}. This difference is close to the relative potential energy of coordination for $Mg^{2+}(OH_2)_6$ and $Zn^{2+}(OH_2)_6$ (Table 4.2), suggesting that the difference in solvation free energies of the two ions is primarily due to stronger intermolecular interactions between the inner sphere water molecules and the ion, despite the ion–water distances being very similar. DFT calculations by Wu et al.[2] indicated that

there is a significant ligand to metal charge transfer between coordinating water molecules and a Zn^{2+} ion, which could account for this difference.

The non-polarizable CHARMM models perform least effectively, resulting in a relative solvation free energy of +30 kcal mol^{-1}, an error of 60 kcal mol^{-1} and incorrectly predicting Mg^{2+} to have the more negative ion hydration energy of the two ions. This difference in energy can be correlated to the smaller size of the Mg^{2+} ion that is apparent in the radial distribution function, which leads to stronger ion–water interactions. It is unlikely that any conventional non-polarizable model could accurately predict both the relative size and hydration energies of these two ions, as the more negative solvation free energy of Zn^{2+} occurs despite Zn^{2+} being essentially the same size as Mg^{2+} is counter to the expectation of a simple electrostatic description.

The Drude polarizable model performs better than the non-polarizable model in the calculation of relative solvation free energies, but falls short of quantitative accuracy. The solvation free energy is calculated to be –13.8 kcal mol^{-1}, in error by –16.2 kcal mol^{-1} from the experimental result. This discrepancy could be explained by the neglect of ion-water charge transfer and limitations of the Drude model for polarization. It is possible that the parameters of the Drude model could be revised to predict the free energy difference more accurately, as the original fitting was performed to reproduce a limited set of experimental data.

In principle, the Lennard-Jones parameters of the non-polarizable model could be adjusted to provide better agreement with experiment, although it is worth noting that classical arguments based on the Born model,

$$\Delta G_{\text{hydr}} = \frac{1}{2} \frac{q_{ion}^2}{R_{ion}} \left(\frac{1}{\epsilon} - 1 \right) \tag{4.3}$$

would require a difference of about 0.25 Å in the radius of Mg^{2+} and Zn^{2+} to match the

free energy difference of 30 kcal mol^{-1} that is observed experimentally. The effective Born radii R_{ion} entering Eqn. 4.3 are indirectly related to the position of the first maximum in the radial distribution function of these ions.[33] According to the radial distribution of Mg^{2+} and Zn^{2+} the radii should differ by at most 0.03 Å, pointing to a glaring failure of a purely classical treatment in this case. This supports the notion that QM effects underlie the large free energy difference.

Table 4.4: Absolute solvation free energy of Mg^{2+} and Zn^{2+} calculated molecular mechanical models.

Model	$\Delta_{hydr}G$ (kcal mol^{-1})	
	Mg^{2+}	Zn^{2+}
Non-polarizable	–407.7	–376.6
Drude	–450.2	–463.1
exptl. (est)	–435.4[1]	–467.7[2]

[1] Ref. 34
[2] Ref. 35

It is possible to calculate the absolute solvation free energies of molecular mechanical models using a decoupling method, which we report in Table 4.4. The limitations of the non-polarizable model are more apparent when we compute the absolute solvation free energies of these ions, which is straightfoward for molecular mechanical models. Although it is impossible to determine absolute solvation free energies experimentally without invoking extrathermdynamic assumptions, we can use the reported estimates values as rough guide to magnitude of the absolute solvation free energies. The Drude polarizable force field was parameterized to reproduce the solvation free energies of neutral salts and is in reasonable agreement with experimental estimates of the absolute solvation free energies. The non-polarizable models predict absolute solvation free energies that are lower than the experimental estimates by 27.7 kcal mol^{-1} and 91.1 kcal mol^{-1} for Mg^{2+} and Zn^{2+}, respectively. The non-polarizable force field performs relatively well for Mg^{2+}(aq) only because it underestimates the Mg–O distances, leading to stronger ion-water interactions. This under-

estimation of the absolute solvation free energies reflects the lack of induced polarization and charge transfer effects in the non-polarizable force field, an effect that is sizable in the solvation of these ions. The only way to produce more accurate solvation free energies with this model would be to decrease the radius of the ions to the point that the ion–water distances would be unrealistically small.

We can also compare to the solvation energies reported for the Mg^{2+} and Zn^{2+} ions parameterized for the AMOEBA polarizable force. Based on the difference of the reported absolute free energies for Mg^{2+} and Zn^{2+}, the relative solvation free energy of these two ions is –27.8 kcal mol^{-1}, which is close to the experimentally measured value of –30.0 kcal mol^{-1}. It should be noted that the AMOEBA model predicts the Zn–O distances to be roughly 0.1 Å smaller than the Mg–O distances. Neither the QM/MM or experimental scattering data support such a large difference in radii and instead predict Zn^{2+} to be slightly larger than Mg^{2+}. The Born model predicts that this 0.1 Å difference in radii would account for a –14.4 kcal mol^{-1} difference in the hydration free energy, so part of the success of the AMOEBA model may simply be because it underestimates the average Zn–O distance. The AMOEBA force field with Valence Bond terms (AMOEBA-VB) predicts more realistic Zn–O distances, although the hydration free energy for this model has not yet been reported.[29]

4.4   Conclusions

In this chapter, we have calculated the radial, tilt, and angular distribution functions for the solvation of the Mg^{2+} and Zn^{2+} ions in liquid water using molecular dynamics simulations. We compare the non-polarizable CHARMM force field, the Drude polarizable force field, and a QM/MM model. Our simulations confirm the established view that both these ions have a hexacoordinate inner coordination sphere with an octahedral structure. The inner coordination sphere is highly ordered and is separated from the bulk solvent. The non-

polarizable force field provides reasonable radial structures and coordination numbers of the two ions, although they do not correctly describe the relative differences of the two ions and overestimate the ion water coordination energies. The Drude MM model shows good agreement with the QM/MM model for ion–O distances and coordination energies in comparison to the non-polarizable model.

The relative solvation free energies of these ions were calculated using thermodynamic integration for the non-polarizable, Drude, and QM/MM methods. The non-polarizable MM method is in error by a significant amount and incorrectly predicts Mg^{2+} to have the more negative solvation free energy. The Drude MM force field underestimates the relative solvation free energy by –17 kcal mol^{-1}. The QM/MM method is most accurate, predicting the relative free energy within 5 kcal mol^{-1} of the experimental value. The QM/MM relative solvation free energy confirms the surprising inference from experimental data that Zn^{2+} has a significantly more negative hydration free energy in comparison to Mg^{2+} despite having the same net charge and a generally similar solvent structure. The implication is that the large free energy difference arises from some non-classical QM component that is not present in MM models, even in a polarizable model.

The QM/MM model used in this study is very effective for modeling the structures and thermodynamics of Mg^{2+} and Zn^{2+} ions in liquid water, as it is in good agreement with both the structural and thermodynamic experimental data. This method has potential to be a valuable tool for understanding the Mg^{2+} vs Zn^{2+} selectivity of binding sites in biomolecules such as proteins, nucleic acids, and phospholipids.

Bibliography

[1] Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 054511.

[2] Wu, J. C.; Piquemal, J.-P.; Chaudret, R.; Reinhardt, P.; Ren, P. *J. Chem. Theory Comput.* **2010**, *6*, 2059–2070.

[3] Yu, H.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D.; Roux, B. *J. Chem. Theory Comput.* **2010**, *6*, 774–786.

[4] Kulik, H. J.; Marzari, N.; Correa, A. A.; Prendergast, D.; Schwegler, E.; Galli, G. *J. Phys. Chem. B* **2010**, *114*, 9594–9601.

[5] Lightstone, F. C.; Schwegler, E.; Hood, R. Q.; Gygi, F.; Galli, G. *Chem. Phys. Lett.* **2001**, *343*, 549 – 555.

[6] Di Tommaso, D.; de Leeuw, N. H. *Crystal Growth & Design* **2010**, *10*, 4292–4302.

[7] Bhattacharjee, A.; Pribil, A. B.; Randolf, B. R.; Rode, B. M.; Hofer, T. S. *Chem. Phys. Lett.* **2012**, *536*, 39–44.

[8] Ikeda, T.; Boero, M.; Terakura, K. *J. Chem. Phys.* **2007**, *127*, 074503.

[9] Cauet, E.; Bogatko, S.; Weare, J. H.; Fulton, J. L.; Schenter, G. K.; Bylaska, E. J. *J. Chem. Phys.* **2010**, *132*, 194502.

[10] Fatmi, M. Q.; Hofer, T. S.; Randolf, B. R.; Rode, B. M. *J. Chem. Phys.* **2005**, *123*, 054514.

[11] Rega, N.; Brancato, G.; Petrone, A.; Caruso, P.; Barone, V. *J. Chem. Phys.* **2011**, *134*, 074504.

[12] Rowley, C. N.; Roux, B. *J. Chem. Theory Comput.* **2012**, *8*, 3526–3535.

[13] Brooks, B. R. et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

[14] TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from
`http://www.turbomole.com`.

[15] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637–649.

[16] van Gunsteren, W.; Berendsen, H. *Mol. Phys.* **1977**, *34*, 1311–1327.

[17] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

[18] Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; Jr., A. D. M. *Chem. Phys. Lett.* **2006**, *418*, 245–249.

[19] Callahan, K. M.; Casillas-Ituarte, N. N.; Roeselová, M.; Allen, H. C.; Tobias, D. J. *J. Phys. Chem. A* **2010**, *114*, 5141–5148.

[20] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[21] Von Arnim, M.; Ahlrichs, R. *J. Comput. Chem.* **1998**, *19*, 1746–1757.

[22] Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

[23] Zhang, R.; Lev, B.; Cuervo, J. E.; Noskov, S. Y.; Salahub, D. R. In *Combining Quantum Mechanics and Molecular Mechanics. Some Recent Progresses in QM/MM Methods*; Sabin, J. R., Ed.; Advances in Quantum Chemistry; Academic Press, 2010; Vol. 59; pp 353–400.

[24] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

[25] Lamoureux, G.; Roux, B. *J. Phys. Chem. B* **2006**, *110*, 3308–3322.

[26] Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.

[27] Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

[28] Deng, Y.; Roux, B. *J. Phys. Chem. B* **2004**, *108*, 16567–16576.

[29] Xiang, J. Y.; Ponder, J. W. *J. Comput. Chem.* **2012**, *34*, 739–749.

[30] Bruni, F.; Imberti, S.; Mancinelli, R.; Ricci, M. A. *J. Chem. Phys.* **2012**, *136*, 064520.

[31] D'Angelo, P.; Barone, V.; Chillemi, G.; Sanna, N.; Meyer-Klaucke, W.; Pavel, N. V. *J. Am. Chem. Soc.* **2002**, *124*, 1958–1967.

[32] D'Angelo, P.; Benfatto, M.; Della Longa, S.; Pavel, N. V. *Phys. Rev. B* **2002**, *66*, 064209.

[33] Roux, B.; Yu, H.; Karplus, M. *J. Phys. Chem.* **1990**, *94*, 4683–4–688.

[34] Schmid, R.; Miah, A. M.; Sapunov, V. N. *Phys. Chem. Chem. Phys.* **2000**, *2*, 97–102.

[35] Marcus, Y. *Biophys. Chem.* **1994**, *51*, 111–127.

Chapter 5

Conclusion and Future Directions

Induced polarization plays an important role in the physiochemical properties of condensed phase systems. This effect is neglected in the classical non-polarizable force fields. In this work, the importance of the induced polarization in two cases were explored. First, a Drude polarizable model was developed for liquid H_2S. This model is able to consistently predict the liquid properties of H_2S, such as enthalpy of vaporization, diffusion and viscosity coefficient, in close agreement with experiment within the 212–298 K temperature range, along the liquid–vapor coexistence curve. A prominent advantage of this model is in its ability to predict the dielectric constant of liquid H_2S, which is overestimated by all the previous models. This Drude model for H_2S also has the same gas phase dipole moment as the experimentally determined value, while previous models had a dipole moment that was significantly too high.

The thermodynamic properties of aqueous solutions of H_2S using Drude model is also in excellent agreement with the experimental values. The calculated Gibbs energy of solvation and the diffusion coefficient are in close agreement with the experimental values within a wide temperature range. The free energy profile of the solvation of a H_2S molecule in water shows the surfactant behavior of H_2S, which had been proposed based on previous

experimental studies. This behavior is further confirmed by the calculation of the surface tension of water slab exposed to H_2S vapor. H_2S significantly decreases the surface tension of water. Although our simulations show that H_2S is hydrated like a hydrophobic solute, the dipole moment of H_2S significantly increases during the solvation process. This effect is totally neglected by the non-polarizable models that were used in previous studies.

As the reliability of the Drude model of H_2S has been proven by recreating the thermodynamic properties of the liquid and aqueous H_2S systems, it is expected that this model is able to provide reliable data in the biological and industrial models. For instance, the mechanism of the biological effects of H_2S is not still fully understood.[1:2] Our H_2S model along with the other Drude models for the lipids, proteins, and water can be used to simulate H_2S in the biological systems and therefore help to understand the biophysics of H_2S. For example, it would be interesting to investigate the diffusion of H_2S in membrane and calculate the partition coefficient.

In Chapter 4, the ability of the Drude model in predicting the hydration properties of Mg^{2+} and Zn^{2+} ions was studied. Three different molecular simulation models were used, including non-polarizable MD, Drude polarizable MD, and a QM/MM method. Although the Drude models of Mg^{2+} and Zn^{2+} yield significantly more realistic properties of these two ions compared to the non-polarizable models, the relative Gibbs energy of solvation predicted by the Drude model is underestimated by 17 kcal mol^{-1}. QM/MM simulation shows the best agreement with the experiment. We attribute this to intermolecular interactions such as ligand-to-metal charge transfer that are not present in the non-polarizable and polarizable molecular mechanical models, which are limited to fixed-charge and induced-dipole electrostatic interactions, respectively.

Mg^{2+} and Zn^{2+} have very different biological functions. For instance, Mg^{2+} mainly binds to the phosphates of nucleic acids. Zn^{2+}, on the other hand, typically binds to the cysteine and histidine side chains of proteins, forming catalytic sites or structural motifs

such as zinc fingers. QM/MM simulations of the biological systems, like the ones reported in this study, can be used to uncover the different functionality of these ions in biological systems.

Bibliography

[1] Mathai, J. C.; Missner, A.; Kügler, P.; Saparov, S. M.; Zeidel, M. L.; Lee, J. K.; Pohl, P. *PNAS* **2009**, *106*, 16633–16638.

[2] Motohiro, N. et al. *Nature* **2012**, *8*, 714–724.

Appendices

Appendix A

CHARMM Parameter File for Different H$_2$S Models

```
* H2S force field
* January 2013
! Change the the atom IDs if using this parameters with other
! parameter files to be compatible with other parameters
read rtf card

MASS 1      SD     32.06000 S        ! DRUDE H2S SULFUR
MASS 2      HD     1.00800 H         ! DRUDE H2S HYDROGEN
MASS 3      DSH2   0.00000 H         ! SULFUR DRUDE
MASS 4      SK     32.06000 S        ! KRISTOF   H2S SULFUR
MASS 5      HK     1.00800 H         ! KRISTOF H2S HYDROGEN
MASS 6      SF     32.06000 S        ! FORESTER H2S SULFUR
MASS 7      HF     1.00800 S         ! FORESTER H2S SULFUR
MASS 8      SJ     32.06000 S        ! JORGENSEN H2S SULFUR
MASS 9      HJ     1.00800 H         ! JORGENSEN H2S HYDROGEN
MASS 10     SP     32.06000 S        ! POTOFF H2S SULFUR
MASS 11     HP     1.00800 H         ! POTOFF H2S HYDROGEN
MASS 113    LP     0.00000 H         ! LONE PAIR

DEFA FIRS NONE LAST NONE

AUTOGENERATE ANGLES DIHEDRALS DRUDE          !note use of DRUDE

!POLARIZABLE 4 POINT H2S
RESI H2SD        0.000
GROUP
ATOM SH2  SD     0.00000 TYPE DSH2 ALPHA -2.5
ATOM SM   LP     -0.27400
```

```
ATOM H1   HD     0.13700
ATOM H2   HD     0.13700
BOND SH2 H1
BOND SH2 H2
BOND SH2 SM
BOND H1  H2                                    ! for SHAKE
ANGLE H1 SH2 H2
LONEPAIR bisector SM SH2 H1 H2 distance 0.202 angle 0.0 dihe 0.0
IC  H1   SH2  H2   H1   1.3400  92.00    0.00    44.00   1.9278
IC  H1   SM  *SH2  H2   1.3400  46.00  180.00   46.00   1.3400
IC  H2   H1   SH2  SM   0.0000   0.00    0.00   46.00   0.202

PATCH FIRST NONE LAST NONE


!KRISTOF 4 POINT H2S
RESI H2SK          0.000
GROUP
ATOM SH2  SK      0.40000
ATOM SM   LP     -0.90000
ATOM H1   HK      0.25000
ATOM H2   HK      0.25000
BOND SH2 H1
BOND SH2 H2
BOND SH2 SM
BOND H1  H2                                    ! for SHAKE
ANGLE H1 SH2 H2
IC  H1   SH2  H2   H1   1.3400  92.00   0.00  44.00  1.9278
IC  H1   SM  *SH2  H2   1.3400  46.00 180.00  46.00  1.3400
IC  H2   H1   SH2  SM   0.0000   0.00   0.00  46.00  0.1862
PATCH FIRST NONE LAST NONE

!FORESTER 4 POINT H2S
RESI H2SF          0.000
GROUP
ATOM SH2  SF      0.66100
ATOM SM   LP     -1.21700
ATOM H1   HF      0.27800
ATOM H2   HF      0.27800
BOND SH2 H1
BOND SH2 H2
BOND SH2 SM
BOND H1  H2                                    ! for SHAKE
```

```
ANGLE H1 SH2 H2
IC   H1  SH2  H2   H1   1.3400  92.00   0.00  44.00  1.9278
IC   H1  SM  *SH2  H2   1.3400  46.00 180.00  46.00  1.3400
IC   H2  H1   SH2  SM   0.0000   0.00   0.00  46.00  0.1933
PATCH FIRST NONE LAST NONE


!JORGENSEN 3 POINT H2S
RESI H2SJ         0.000
GROUP
ATOM SH2  SJ    -0.47
ATOM H1   HJ     0.235
ATOM H2   HJ     0.235
BOND SH2 H1
BOND SH2 H2
BOND H1  H2                                  ! for SHAKE
THET H1 SH2 H2
IC   H1  SH2  H2   H1   1.3400  92.00   0.00  44.00  1.9278
PATCH FIRST NONE LAST NONE


!POTOFF 3 POINT H2S
RESI H2SP         0.000
GROUP
ATOM SH2  SP    -0.380
ATOM H1   HP     0.190
ATOM H2   HP     0.190
BOND SH2 H1
BOND SH2 H2
THET H1 SH2 H2
IC   H1  SH2  H2   H1   1.3400  92.00   0.00  44.00  1.9278
PATCH FIRST NONE LAST NONE


end

read para card                              !append
* FF parameters
*
BONDS
!atom type    Kb        b0
!==========================================
SD      LP      0.00      0.202        ! DRUDE S-LP
SD      HD      398.00    1.3400       ! DRUDE S-H
SD      DSH2    500.00    0.000        ! DRUDE S-DRUDE
HD      HD      0.00      1.92438      ! DRUDE H-H
```

97

```
SK         LP        0.00        0.1862        ! KRISTOF S-LP
SK         HK        398.00      1.3400        ! KRISTOF S-H
HK         HK        0.00        1.92438       ! KRISTOF H-H
SF         LP        0.00        0.1933        ! FORESTER S-LP
SF         HF        398.00      1.3400        ! FORESTER S-H
HF         HF        0.00        1.92438       ! FORESTER H-H
SJ         HJ        398.00      1.3400        ! JORGENSE S-H
HJ         HJ        0.00        1.92438       ! JORGENSEN H-H
SP         HP        398.00      1.3400        ! POTOFF S-H


ANGLES
!atom types            Ktheta    Theta0
!=================================================
HD     SD     HD        39.6      92.0      ! DRUDE H-S-H
HD     SD     LP        0.0       46.065    ! DRUDE H-S-LP
HK     SK     HK        39.6      92.0      ! KRISTOF H-S-H
HK     SK     LP        0.0       46.065    ! KRISTOF H-S-LP
HF     SF     HF        39.6      92.0      ! FORESTER H-S-H
HF     SF     LP        0.0       46.065    ! FORESTER H-S-LP
HJ     SJ     HJ        39.6      92.0      ! JORGENSEN H-S-H
HP     SP     HP        65.1      92.0      ! POTOFF H-S-H


NONBONDED nbxmod  5 atom vatom cdiel vdistance switch vswitch -
cutnb 16.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5
SD     0.0000    -0.5680      2.0880    ! DRUDE S
HD     0.0000    -0.0000      0.0000    ! DRUDE H
DSH2   0.0000    -0.0000      0.0000    ! DRUDE DRUDE
SK     0.0000    -0.4968      2.0934    ! KRISTOF S
HK     0.0000    -0.0000      0.0000    ! KRISTOF H
SF     0.0000    -0.5345      2.0709    ! FORESTER S
HF     0.0000    -0.0000      0.0000    ! FORESTER H
SJ     0.0000    -0.2500      2.0766    ! JORGENSEN S
HJ     0.0000    -0.0000      0.0000    ! JORGENSEN H
SP     0.0000    -0.4610      2.0878    ! POTOFF S
HP     0.0000    -0.0000      0.0000    ! POTOFF H
LP     0.0000    -0.0000      0.0000    ! LP


END
```

Appendix B

Optimization procedure of Lennard-Jones Parameters for H$_2$S–H$_2$O system

To improve the description of H$_2$S solvation, H$_2$S–H$_2$O Lennard-Jones parameters were optimized to correctly predict the $\Delta G^{\circ}_{\text{hydr}}$ of H$_2$S through a 3 iteration grid search of the parameter space. The parameterization was performed using the $E_{\text{min},S-S}$ and $R_{\text{min},S-S}$ parameter space using the combining rule and the final parameters were converted to pairwise $E_{\text{min},S-O}$ and $R_{\text{min},S-O}$ terms. This procedure is depicted in Fig. B.1. In the first step, we searched a very broad parameter range, 0.45 kcal mol^{-1} $< E_{\text{min}} < 0.80$ kcal mol^{-1} with a spacing of 0.05 kcal mol^{-1} and 2.00 Å $< R_{\text{min}} < 2.10$ Å with a spacing of 0.02 Å. The area that predicted Gibbs hydration energies in the range of -1.03 kcal mol^{-1} to -0.36 kcal mol^{-1} was selected for a the second iteration of the parameterization, where finer grid search was performed. In the next step, the hydration energy was calculated for parameters between 0.65 kcal mol^{-1} $< E_{\text{min}} < 0.75$ kcal mol^{-1} with a spacing of 0.01 kcal mol^{-1} and 2.015 Å $< R_{\text{min}} < 2.025$ Å with a 0.005 Å spacing. The parameter space with Gibbs

energies of hydration between -0.75 to -0.40 kcal mol^{-1} were selected for the last stage of parameter search. The last step of grid search were performed on parameters with $2.020 < R_{\text{min}} < 2.025$ with a spacing of of 0.001 Å and 0.690 kcal mol$^{-1} < E_{\text{min}} < 0.710$ kcal mol^{-1} with a spacing of 0.005 kcal mol^{-1}. Parameters with R_{min} values in the $2.021 - 2.022$ Å range and E_{min} values in the $0.705 - 0.710$ kcal mol^{-1} range had the target hydration free energy. The middle of the two intervals, R_{min}=2.0215 Å and E_{min}= 0.707 kcal mol^{-1}, was selected as the final H$_2$O–H$_2$S parameters.
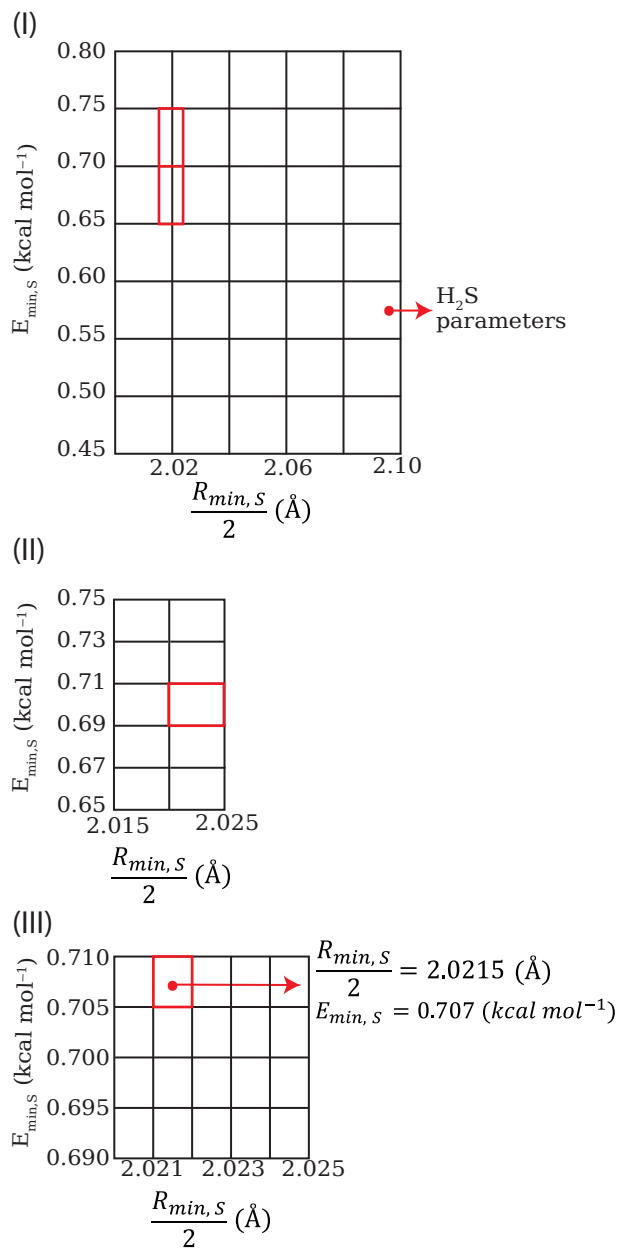
Figure B.1: Schematic of parameterization grid search method employed for optimization of S–O Lennard-Jones parameters for the H_2S–H_2O system.

Appendix C

CHARMM Parameter File For H$_2$S–H$_2$O system

```
* H2S force field
* January 2013
! Change the the atom IDs if using this parameters with other
! parameter files to be compatible with other parameters
read rtf card
MASS      1 SD        32.06000 S  ! DRUDE H2S SULFUR
MASS      2 HD         1.00800 H  ! DRUDE H2S HYDROGEN
MASS      3 DSH2       0.00000 H  ! SULFUR DRUDE
MASS    113 LP         0.00000 H  ! LONE PAIR
MASS    151 ODW       15.99940 O  ! WATER OXYGEN
MASS    152 HDW        1.00800 H  ! WATER HYDROGEN
MASS    153 DOH2       0.00000 H  ! WATER DRUDE


DEFA FIRS NONE LAST NONE


AUTOGENERATE ANGLES DIHEDRALS DRUDE !note use of DRUDE



!POLARIZABLE 4 POINT H2S
RESI H2SD            0.000
GROUP
ATOM SH2 SD          0.00000 TYPE DSH2 ALPHA -2.5
ATOM SM  LP         -0.27400
ATOM H1  HD          0.13700
ATOM H2  HD          0.13700
BOND SH2 H1
BOND SH2 H2
BOND SH2 SM
```

```
BOND H1 H2            ! for SHAKE
ANGLE H1 SH2 H2
LONEPAIR bisector SM SH2 H1 H2 distance 0.202 angle 0.0 dihe 0.0
IC H1 SH2 H2 H1 1.3400 92.00 0.00 44.00 1.9278
IC H1 SM *SH2 H2 1.3400 46.00 180.00 46.00 1.3400
IC H2 H1 SH2 SM 0.0000 0.00 0.00 46.00 0.202

PATCH FIRST NONE LAST NONE


! SWM4-NDP water
RESI SWM4          0.000
GROUP
ATOM OH2  ODW      0.00000  TYPE DOH2    ALPHA -0.97825258
ATOM OM   LP      -1.11466
ATOM H1   HDW      0.55733
ATOM H2   HDW      0.55733
BOND OH2 H1
BOND OH2 H2
BOND OH2 OM
BOND H1  H2      ! for SHAKE
ANGLE H1 OH2 H2
ACCEPTOR OH2
LONEPAIR bisector OM OH2 H1 H2 distance 0.24034492 angle 0.0 dihe 0.0
IC H1 OH2 H2  H1 0.9572 104.52 0.00 37.74 1.5139
IC H1 OM *OH2 H2 0.9572 52.26 180.00 52.26 0.9572
IC H2 H1  OH2 OM 1.5139 37.74 0.01 52.26 0.24034492
PATCH FIRST NONE LAST NONE

end

read para card !append
* FF parameters
*

BONDS
!atom type       Kb        b0
!=================================================
!
ODW     HDW    450.00    0.9572     ! SWM4-NDP O-H
ODW     LP     0.00      0.24034492 ! SWM4-NDP O-LP
ODW     DOH2   500.00    0.000      ! SWM4-NDP O-DRUDE
HDW     HDW    0.00      1.5139     ! SWM4-NDP H-H
```

```
SD      HD      398.00      1.3400          ! H2S DRUDE S-H
SD      LP      0.00        0.202           ! H2S DRUDE S-LP
SD      DSH2    500.00      0.000           ! H2S DRUDE S-DRUDE
HD      HD      0.00        1.92438         ! H2S DRUDE H-H


ANGLES
!atom types      Ktheta   Theta0
!=================================================
!
HDW   ODW   HDW      55.000  104.52    ! SWM4-NDP
HD    SD    HD       39.6    92.0       ! H2S DRUDE H-S-H


NONBONDED nbxmod 5 atom vatom cdiel vdistance switch vswitch -
cutnb 16.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5
!atom type ignored   epsilon        Rmin/2
!=================================================
!
HDW    0.0000   -0.0000         0.0000      ! SWM4-NDP  H
ODW    0.0000   -0.21094325    1.78692899  ! SWM4-NDP  O
SD     0.0000   -0.568          2.0880      ! H2S Drude S
HD     0.0000   -0.0000         0.0000      ! H2S DRUDE H
LP     0.0000   -0.0000         0.0000      ! LP
DOH2   0.0000   -0.0000         0.0000      ! Drude H2O
DSH2   0.0000   -0.0000         0.0000      ! Drude H2S


NBFIX
!atom1  atom2 (epsilonA*epsilonB)^0.5 (RminA+RminB)/2
!=================================================
!
SD      ODW     -0.38618                    3.80842899


END
```

Appendix D

CHARMM Input File to for the Simulation of Liquid H₂S

```
* CHARMM INPUT FILE FOR H2S
*

IOFORMAT EXTENDED
SET TEMP = 212
SET PRESSURE = 1.0
STREAM prm.prm
READ SEQUENCE H2SD 560
GENERATE MOL SETUP FIRST NONE LAST NONE NOANGLE NODIHEDRAL -
             DRUDE DMASS 0.4 HYPE HORD 4 KHYP  40000 RHYP 0.2

READ COOR FROM PDB NAME h2s.pdb

SET BOXLENGTH = 32.23

COOR COPY COMP
SHAKE PARAM
COOR SHAKE
COOR SDRUDE

SHAKE BOND
SHAKE BONH PARAM TOL 1.0e-12 SELECT .NOT. TYPE D* END
COOR STAT
COOR TRANS XDIR -?XAVE YDIR -?YAVE ZDIR -?ZAVE

CRYSTAL DEFINE CUBIC @boxlength @boxlength @boxlength 90.0 90.0 90.0
CRYSTAL BUILD NOPERATIONS 0 CUTOFF 16.0
IMAGE BYRESIDUES SELECT RESNAME H2SD END
```

```
NBOND ATOM VATOM VSWITCH LRC -
        CTOFNB 12.0 CUTNB 16.0 WMIN 1.5 -
        EWALD SPLINE KAPPA 0.33333 -
        PMEWALD ORDER 6 FFTX 32 FFTY 32 FFTZ 32 -
        INBFRQ -1
FASTER OFF

UPDATE
TPCONTROL NTHER 2 NSTEP 50 -
     THER 1 TAU 0.1 TREF @TEMP SELE .NOT. TYPE D* END -
     THER 2 TAU 0.002 TREF 1 SELECT TYPE D* END  -
     BARO BTAU 0.2 PREF @PRESSURE


OPEN UNIT 35 FORM WRITE NAME h2s.res
OPEN UNIT 64 FILE WRITE NAME  h2s.dcd

DYNA VV2 START NSTEP 1000000 TIMESTEP 0.001 -
     ISEED 8538 1397 3534 6939  IPRFREQ 1000 -
     IHTFRQ 500 IEQFRQ 500 NTRFRQ 1000  -
     IUNREA -1 IUNWRI 35 IUNCRD 64 IUNVEL -1 KUNIT -1 -
     NPRINT 1 NSAVC 500 NSAVV -1 IHBFRQ -1 INBFRQ -1  -
     FIRSTT @temp FINALT @temp  TSTRUCT @temp -
     IASORS 0 IASVEL 1 ISCVEL 0 ICHECW 0 TWINDH 0.0 TWINDL 0.0 -
     IMGFR -1
STOP
```

Appendix E

CP2K Input File for the Simulation of Liquid H₂S

```
&GLOBAL
  PROJECT h2s_64
  RUN_TYPE MD
  PRINT_LEVEL LOW
&END GLOBAL
&MOTION
  &CONSTRAINT
    &HBONDS
      ATOM_TYPE S
      MOLNAME  H2S
      TARGETS [angstrom] 1.34
    &END HBONDS
  &END CONSTRAINT

  &PRINT
    &TRAJECTORY
      LOG_PRINT_KEY T
       &EACH
         MD 1
       &END EACH
       FILENAME h2s.dcd
       FORMAT DCD
    &END TRAJECTORY
  &END  PRINT
  &MD
    ENSEMBLE LANGEVIN
    STEPS 2000
    TIMESTEP 1.0
```

```
        TEMPERATURE  212.0
        STEP_START_VAL         2500
        TIME_START_VAL         2500
        &LANGEVIN
          GAMMA 0.001
        &END LANGEVIN
    &END MD
&END MOTION
&FORCE_EVAL
    &PRINT
      &FORCES
      &END FORCES
    &END PRINT
    METHOD Quickstep
    &DFT
        BASIS_SET_FILE_NAME BASIS_SET
          POTENTIAL_FILE_NAME POTENTIAL
          WFN_RESTART_FILE_NAME wfn.wfn
          &SCF
            SCF_GUESS ATOMIC
          &END SCF
          &QS
            EPS_DEFAULT 1.0E-12
            MAP_CONSISTENT TRUE
            EXTRAPOLATION ASPC
            EXTRAPOLATION_ORDER 3
          &END QS
          &MGRID
            CUTOFF 280
          &END MGRID
          &XC
            DENSITY_CUTOFF     1.0000000000000000E-10
            GRADIENT_CUTOFF    1.0000000000000000E-10
            TAU_CUTOFF         1.0000000000000000E-10
            &XC_FUNCTIONAL  NO_SHORTCUT
              &PBE  T
              &END PBE
            &END XC_FUNCTIONAL
            &VDW_POTENTIAL
              POTENTIAL_TYPE  PAIR_POTENTIAL
              &PAIR_POTENTIAL
                TYPE  DFTD2
                REFERENCE_FUNCTIONAL PBE
```

```
            &END PAIR_POTENTIAL
          &END VDW_POTENTIAL
        &END XC
    &END DFT
    &SUBSYS
      &CELL
        ABC [angstrom] 16.3 16.3 16.3
        PERIODIC  XYZ
      &END CELL
      &COORD
        @INCLUDE  64.coor
      &END COORD
      &KIND H
        BASIS_SET TZV2P-MOLOPT-GTH
        POTENTIAL GTH-PBE-q1
      &END KIND
      &KIND S
        BASIS_SET TZV2P-MOLOPT-GTH
        POTENTIAL GTH-PBE-q6
      &END KIND
      &TOPOLOGY
        NUMBER_OF_ATOMS         192
        MULTIPLE_UNIT_CELL 1  1  1
      &END TOPOLOGY
    &END SUBSYS
&END FORCE_EVAL
```