



Simple and systematic optimization of a polarizable water model

Lee-Ping Wang¹, Teresa Head-Gordon², Jay Ponder³, Pengyu Ren⁴, John Chodera⁵, Todd J. Martinez¹ and Vijay S. Pande¹

¹Dept of Chemistry, Stanford University, Stanford, CA, USA.

²Dept of Chemistry, U.C. Berkeley, Berkeley, CA, USA.

³Dept of Chemistry, Washington University, St. Louis, MO, USA.

⁴Dept of Bioengineering, University of Texas, Austin, TX USA.

⁵California Institute for Quantitative Biosciences (QB3), Berkeley, CA, USA.



Introduction

Our understanding of the physical interactions between atoms is often modeled using empirical *force fields*. Simple yet accurate force fields are broadly useful for physical chemistry but very challenging to develop.

We developed a program called **ForceBalance** to construct accurate molecular models from *theoretical data combined with experimental data* using systematic optimization methods and strict regularization schemes.

Here we apply ForceBalance to optimize an inexpensive polarizable water model largely based on the AMOEBA model but using a simplified model of electronic polarization. Our new model surpasses AMOEBA in accuracy for several properties of water.

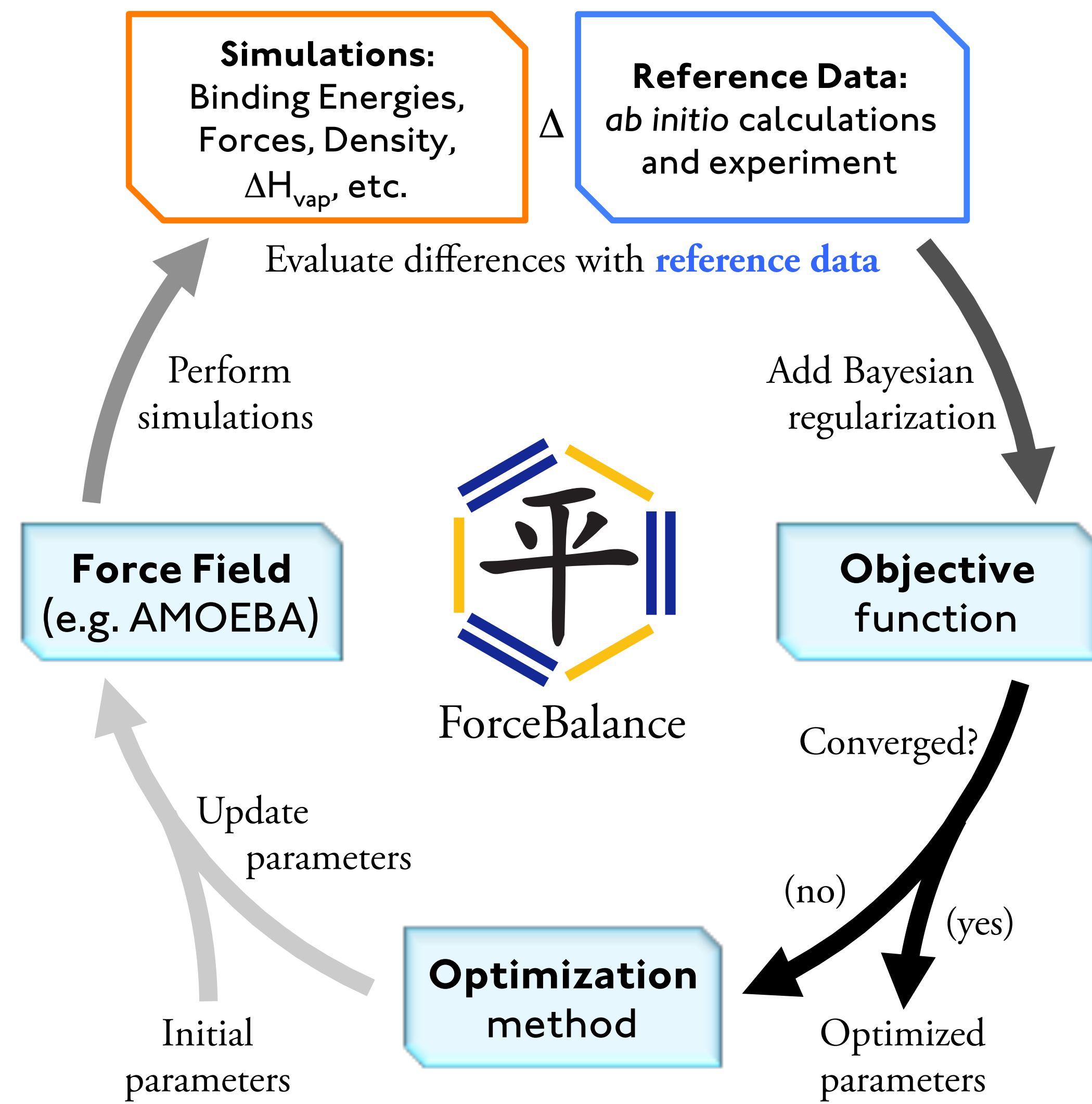


Figure 1: The structure of ForceBalance affords a high degree of flexibility with respect to the force field's functional form, the experimental and theoretical reference data, and the optimization method.

Method

The force field is parameterized by minimizing an objective function in the parameter space. The objective function contains differences between the model predictions and the theoretical and experimental reference data.

ForceBalance computes parametric derivatives of the objective function for nonlinear optimization. When fitting simulated ensemble-average properties to experiment, the parametric derivative is given as follows:

$$\nabla_{\kappa} \langle A \rangle \equiv \nabla_{\kappa} \left(\frac{\int d\mathbf{r} e^{-\beta E(\mathbf{r}, \kappa)} A(\mathbf{r})}{Z} \right) = -\beta (\langle \nabla_{\kappa} E \cdot A \rangle - \langle \nabla_{\kappa} E \rangle \langle A \rangle),$$

where $\langle A \rangle$ is the Boltzmann average of the observable A , κ represents the force field parameters, and Z is the canonical partition function. The RHS is a two-point correlation function between A and the energy derivative $\nabla_{\kappa} E$.

We apply ideas from Bayesian probability theory to prevent overfitting of parameters. All of the force field parameters in this work represent a deviation of $< 25\%$ from their original starting values.

Figure 2: Molecular Simulation Methods

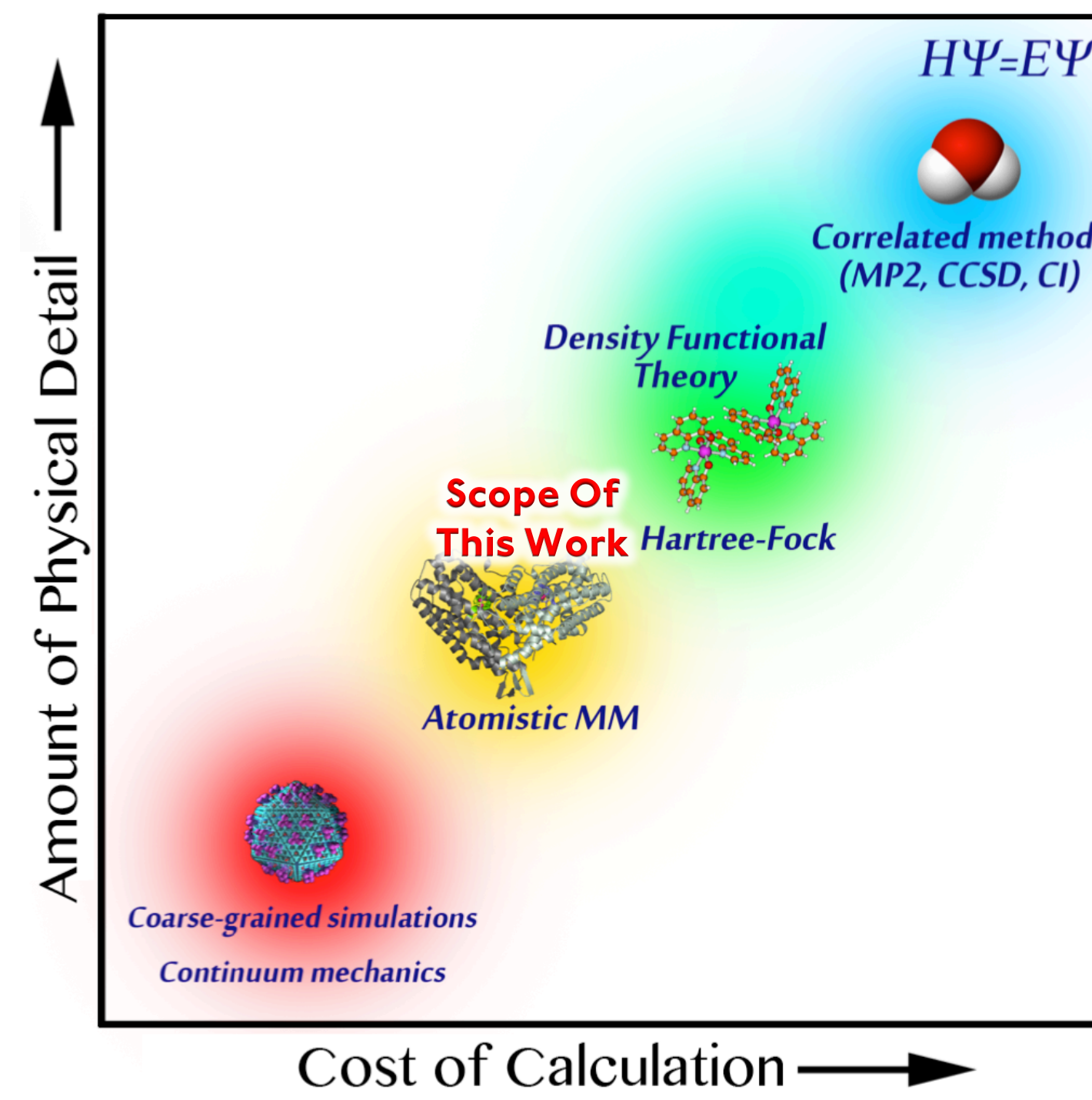


Figure 4: Fit to Experimental Data

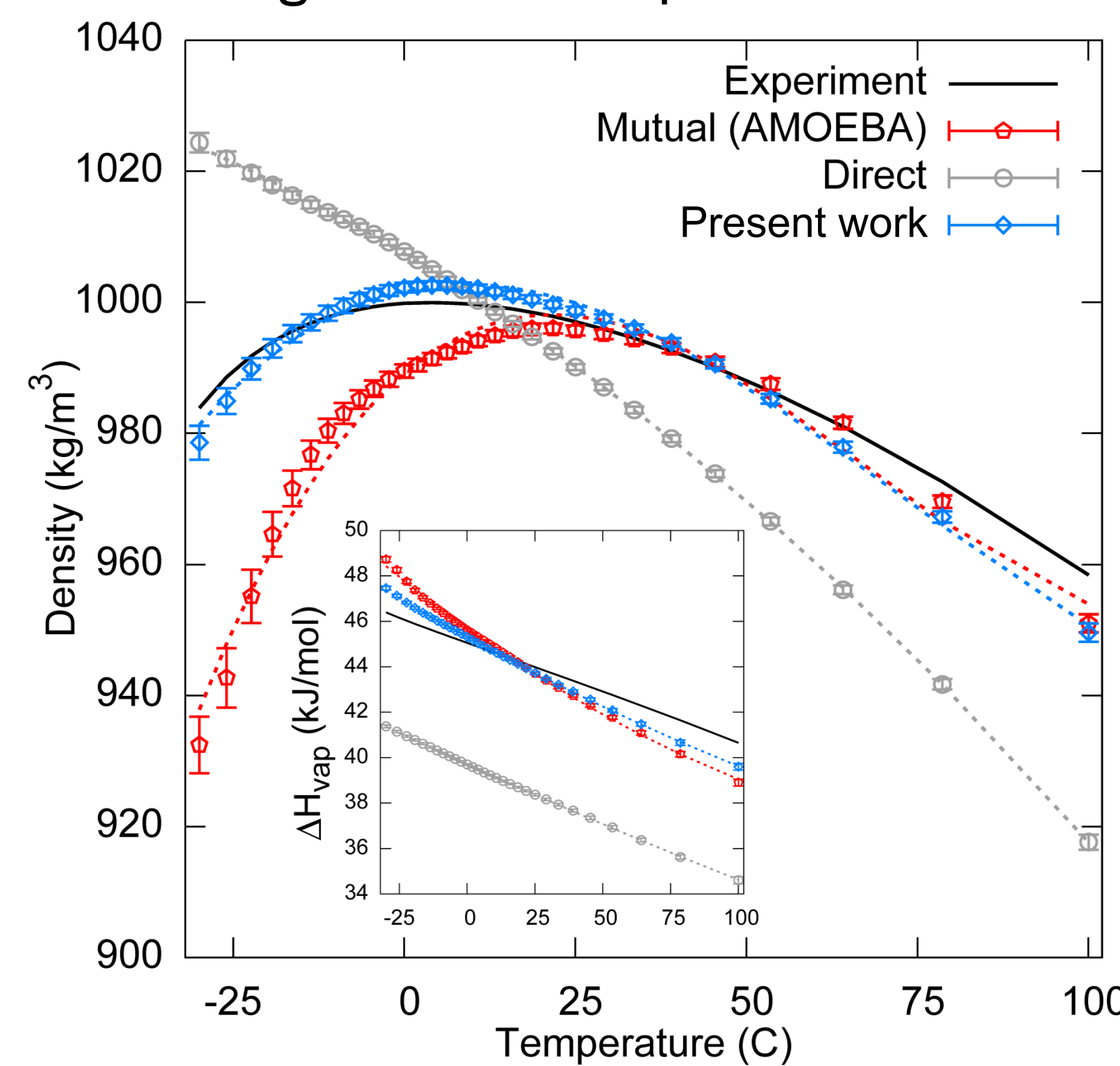


Figure 2 (Upper left): In this work, we develop a force field for highly detailed atomistic MM simulation.

Figure 3 (Upper right): The water force field in this work uses atomic multipoles and direct polarization.

Figure 4 (Lower left): Plot of water density and ΔH_{vap} (inset) showing accuracy of the model in this work.

Figure 5 (Lower right): Comparison of forces and energies from force field and *ab initio* calculations.

Figure 3: Polarizable Atomic Multipoles

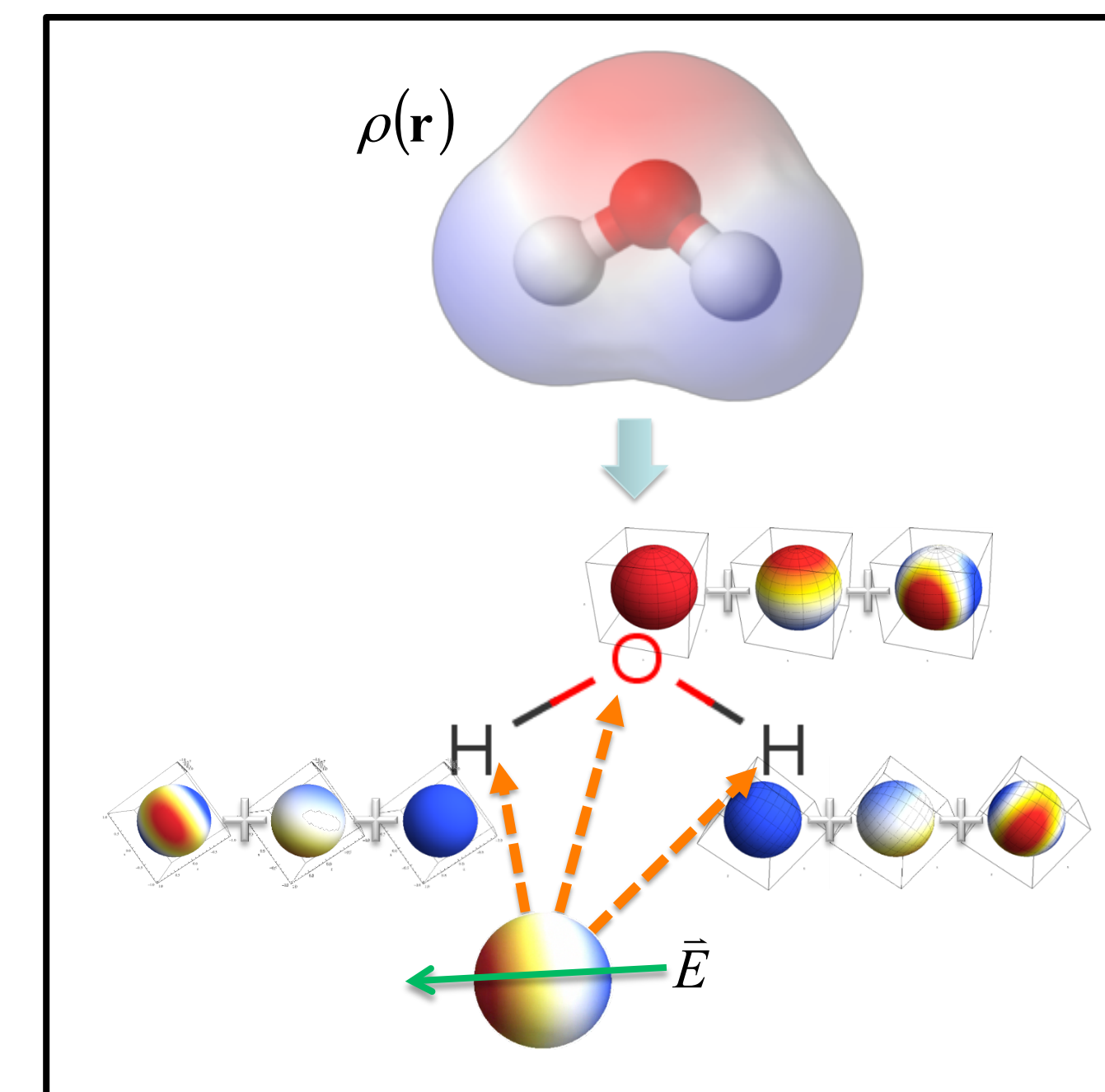


Figure 5: Fit to Theoretical Calculations

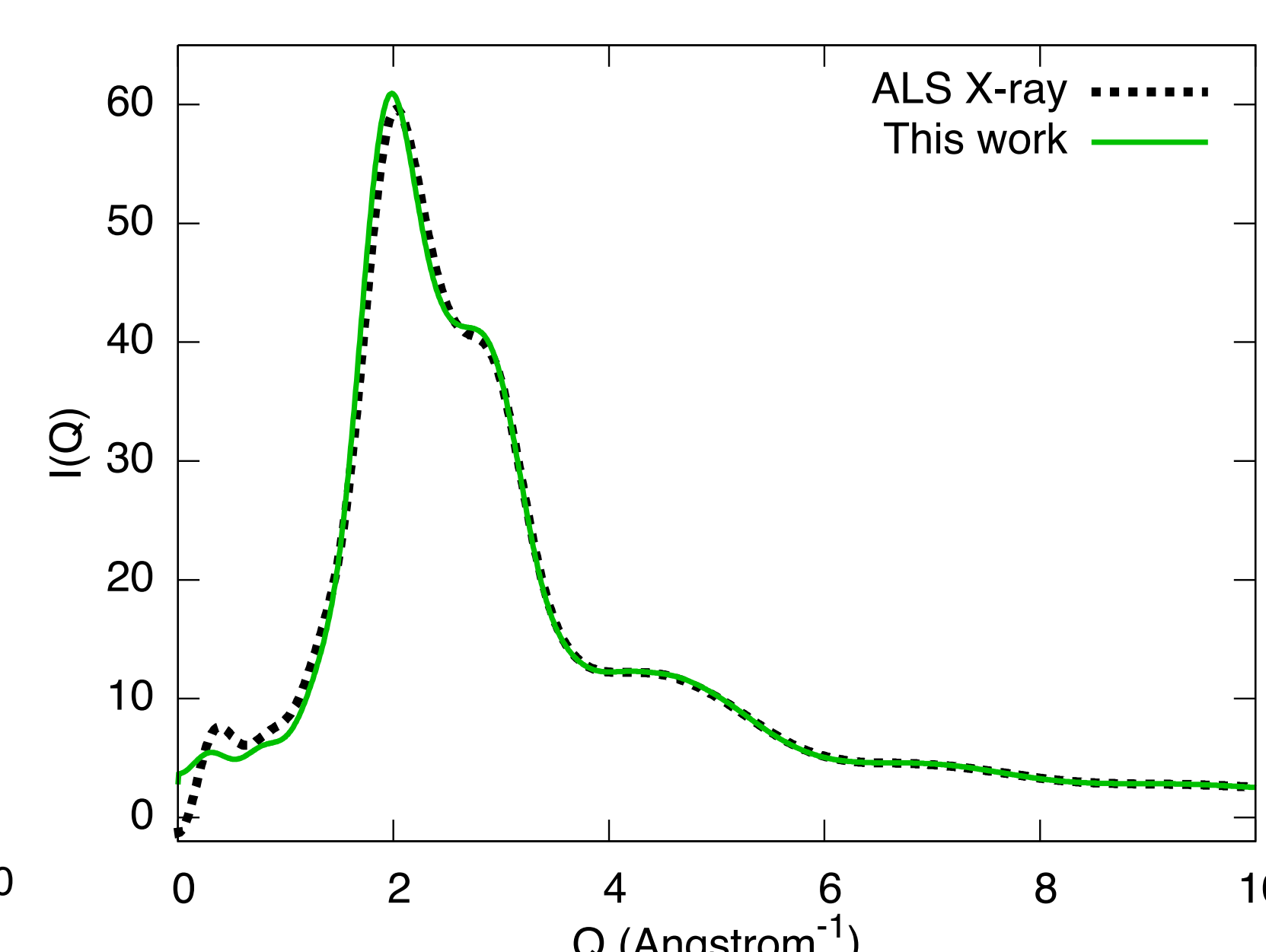
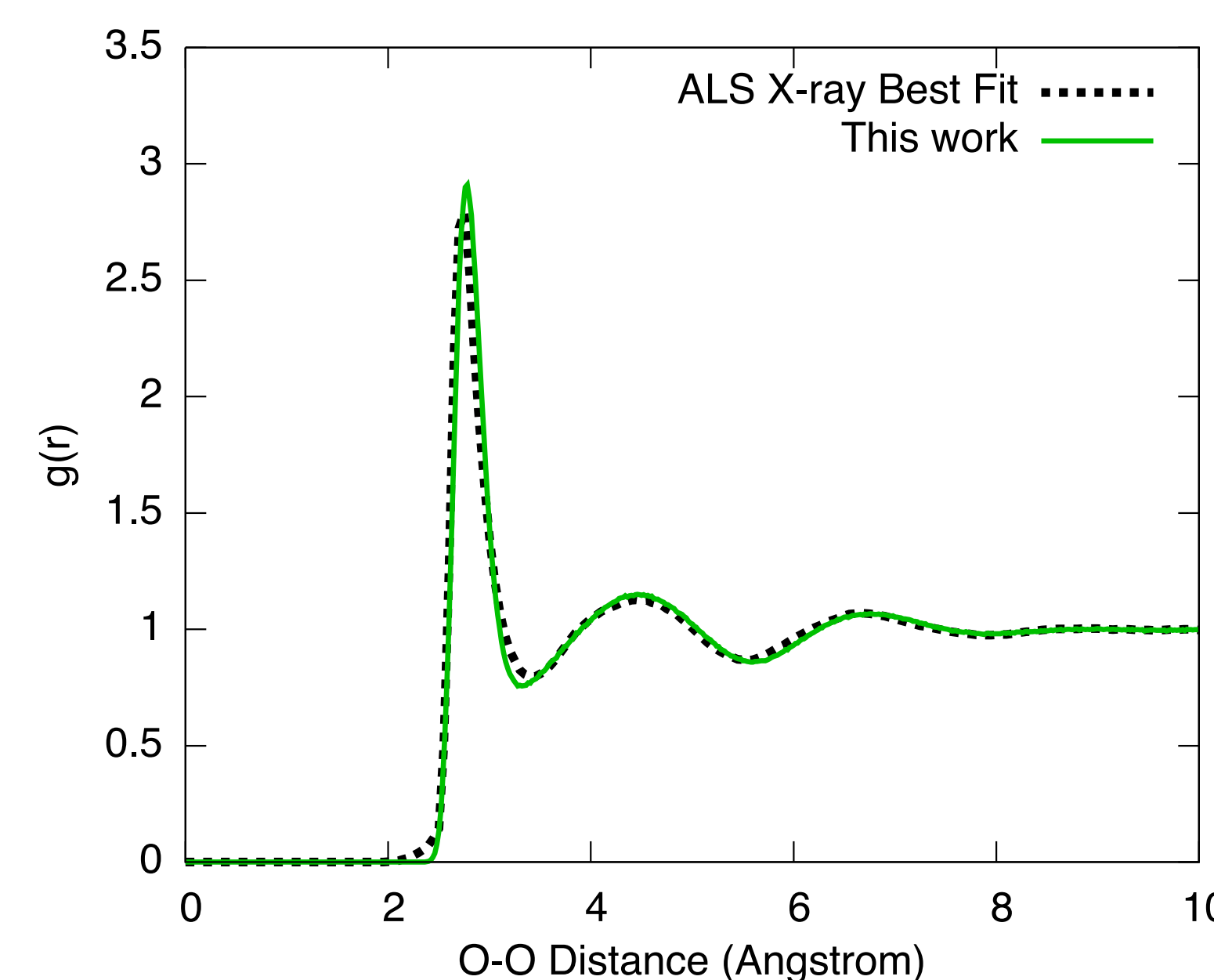
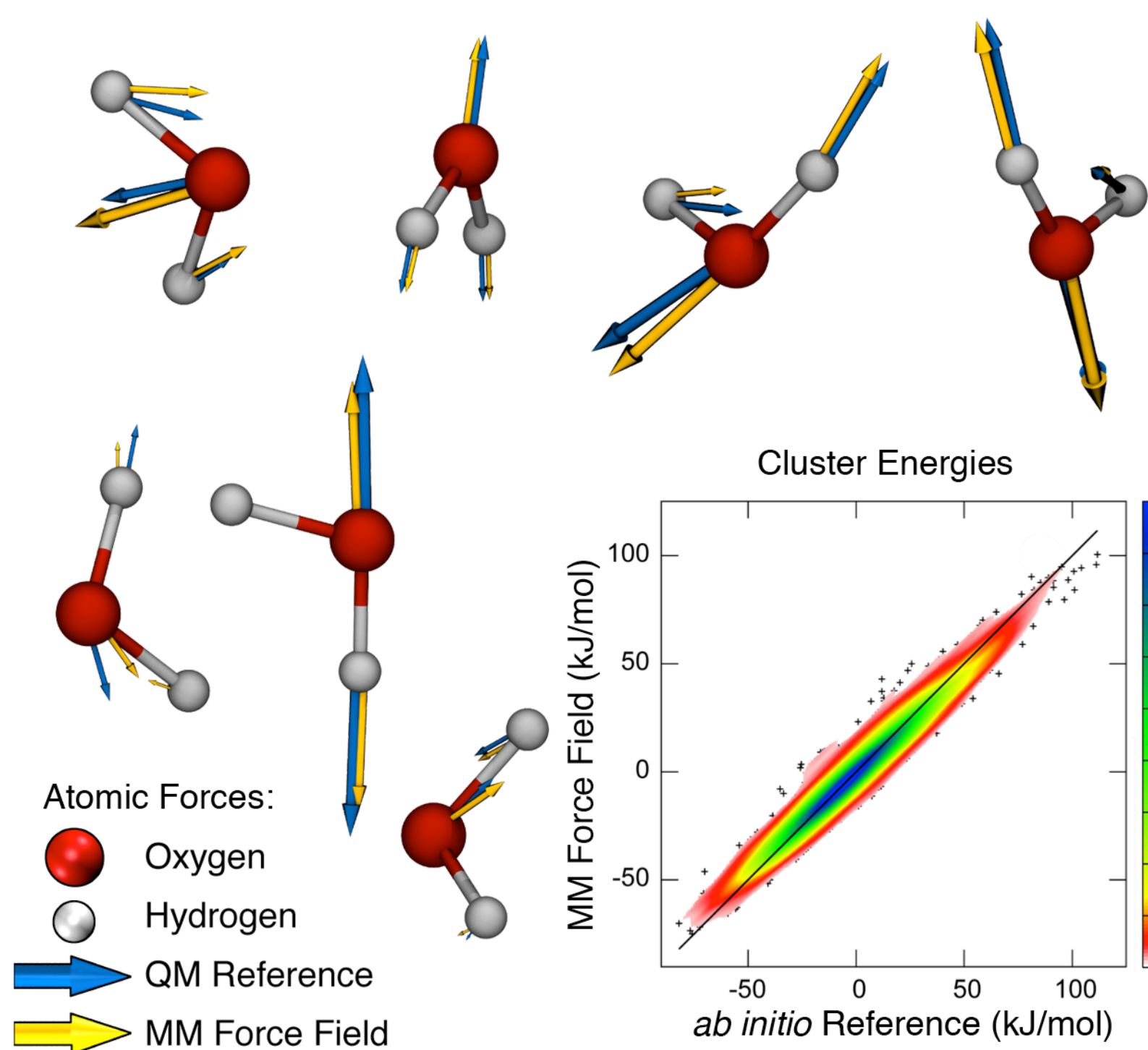


Figure 6: O-O radial distribution function (left) and X-ray scattering spectrum (right) of liquid water at 298.15 K, 1 atm computed using the model in this work (green) compared to experimental measurements (black).

Application: Direct polarization water model

We sought to create an inexpensive water model using polarizable atomic multipoles (Figure 3). The 2003 AMOEBA model contains mutually polarizing dipoles that incur a large computational cost. Using *direct polarization* greatly speeds up the calculation (up to 5x) but alters the physical interactions, so the force field must be reparameterized.

Our reference data includes elements from theory and experiment. We calculated single-point energies and forces using *ab initio* quantum chemistry (MP2/aug-cc-pVTZ) for $>25,000$ cluster geometries extracted from the liquid. We considered the density and heat of vaporization ranging from -30°C to 100°C at atmospheric pressure. Gas-phase properties were also considered including the binding energies of clusters.

The parameterization was performed using ForceBalance and the GPU-accelerated OpenMM simulation software. Convergence was achieved in ~ 10 Levenberg-Marquardt iterations.

The quality of fit is shown in Figures 4 and 5. The density is accurate to within 0.5% over the temperature range, and the temperature of maximum density is accurate to within 5°C . ΔH_{vap} is reproduced to within 1 kJ/mol. The *ab initio* energies are reproduced to within 6 kJ/mol (20%).

To validate our model we calculated a number of liquid properties that were not fitted (Table 1), including the radial distribution function (Figure 6) and freezing point (Figure 7). Equilibrium and kinetic properties of the liquid are accurately reproduced.

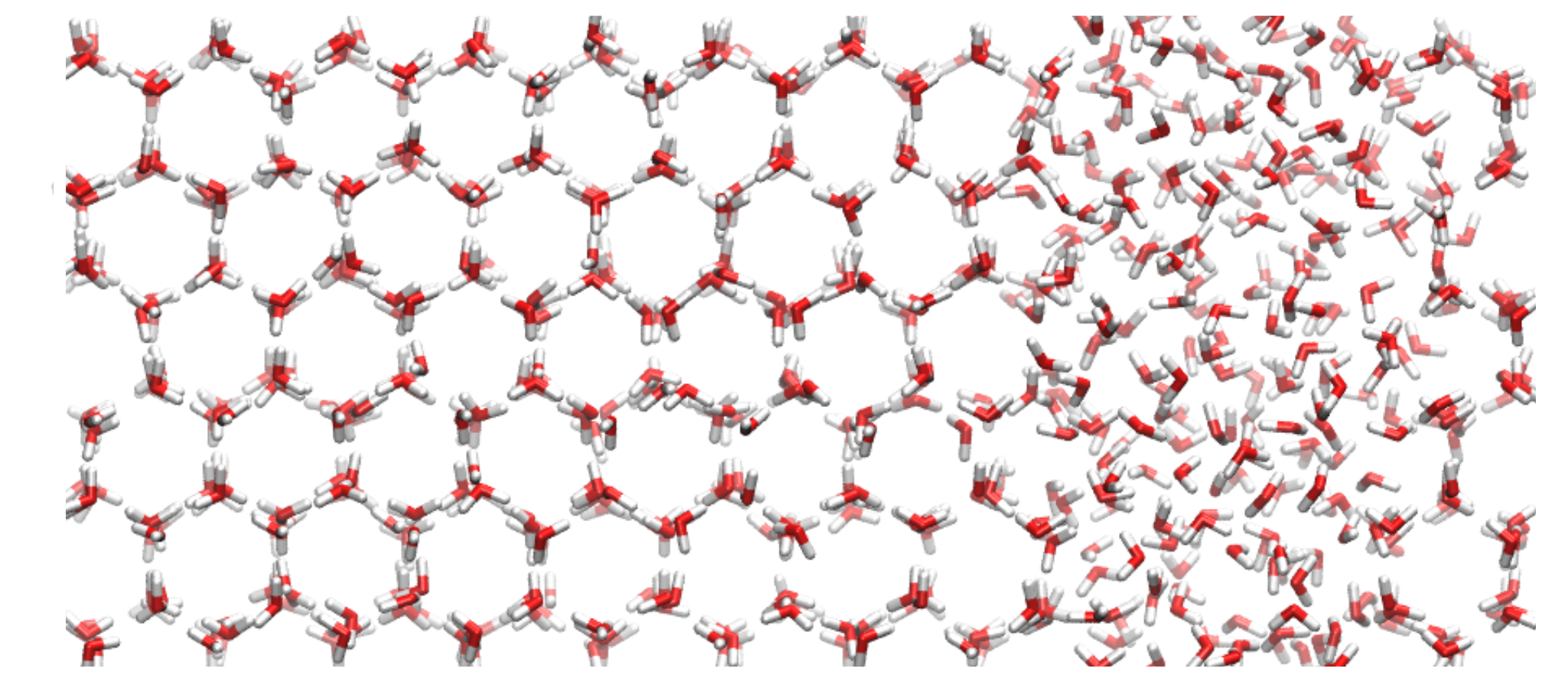


Figure 7: Simulating the freezing of water at a solid / liquid interface.

	Density ρ (g cm ⁻³)	Density maximum (°C)	Heat of vaporization ΔH_{vap} (kJ mol ⁻¹)	Self-diffusion constant D (10 ⁻⁵ cm ² s ⁻¹)	Dielectric constant ϵ
TIP3P	1.002(1)	-91 ± 5	10.41(1)	6.05(5)	92 ± 5
TIP4P-2005	0.998	+5 ± 3	10.89	2.59(6)	62 ± 4
SWM4-NDP	0.999	-121	10.51	2.78(4)	79 ± 3
AMOEBA	1.000(1)	+25	10.48(8)	2.0	81 ± 2
This work	0.9971(3)	4 ± 3	10.56(1)	2.61(2)	82 ± 3
Experiment	0.9970	+4	10.518	2.30	78.3
	Melting point (°C)	Shear viscosity η (10 ⁻³ Pa s)	Isothermal compressibility κ_T (10 ⁻⁶ atm ⁻¹)	Isobaric heat capacity c_p (cal mol ⁻¹ K ⁻¹)	Thermal expansion α_p (10 ⁻⁴ K ⁻¹)
TIP3P	-127 ± 2	0.308(10)	64 ± 5	20.0(6)	9.2(8)
TIP4P-2005	-23 ± 2	0.575(8)	46.5	18.9	2.8
SWM4-NDP	< -150	0.70			
AMOEBA			55	21.3 ± 2.0	
This work	-14 ± 2	0.86(2)	42.1	18.3	3.3
Experiment	0	0.89	45.8	18.0	2.55

Table 1: Bulk properties of water models vs. experimental measurements.

Conclusion

This project demonstrates the importance of developing force fields using combined theoretical and experimental data. We hope that the water model in this work will be inexpensive and accurate for broad applications.

ForceBalance and OpenMM are freely available at <https://simtk.org/>.