Towards accurate calculations of Zn^{2+} binding free energies in zinc finger proteins

UNDERGRADUATE HONORS RESEARCH THESIS

Presented in Partial Fulfillment of the Requirements for the Bachelor of Science with Honors Research Distinction in the College of Engineering of the Ohio State University

By

Hok Hei Tam

Undergraduate Program in Chemical Engineering

The Ohio State University

2012

Thesis Committee:

Michael. E. Paulaitis, Advisor

Sherwin J. Singer

Abstract

Zinc fingers are Zn^{2+}-bound peptide motifs that bind DNA specifically and have great potential in gene therapy. However, the ion binding strength of the zinc finger is not well known, and computing this quantity will allow for the design of more stable zinc finger treatments. Ions in solution are a model system. Molecular dynamics (MD) simulations and the inverse potential distribution theorem were used to estimate the solvation free energies of zinc ions. The zinc coordination shells were stable and the initial coordination shell stayed throughout the 20 ns simulations. Quasi-chemical (QC) calculations are free energy calculations that partition the system into an inner shell, treated using quantum mechanics, and an outer shell, treated using continuum electrostatics. The theory was extended to multiple ligands in solution and used on Zn^{2+} in water/methanol mixtures, with the inner shell consisting of the six solvent molecules coordinated to the ion and the outer shell consisting of all other solvent. Increasing methanol coordinated to the zinc led to lower inner shell formation free energies but higher outer shell free solvation energies. A six-water coordination shell was found to be most stable. Using quasi-chemical theory with different concentrations in the outer shell did not yield major differences, but this could have been due to an insufficient treatment of the van der Waals forces. A quasi-chemical approximation using MD to treat the outer shell would fix such problems and will be useful in computing zinc finger ion binding free energies.

Table of Contents

List of Figures

List of Tables

1 Introduction

Metal ions are important biologically because of their structural and reactive role in many proteins. It is estimated that one-third of all proteins are metalloproteins (metal-containing proteins) [1]. They play important roles in reduction/oxidation reactions due to their electron transfer properties. Their highly charged nature also enhances the structural stability of many proteins. One class of such metalloproteins are zinc finger proteins.

Zinc finger proteins are DNA binding proteins with structural Zn^{2+} ions that have demonstrated a high degree of specificity for certain DNA strands. Because of this, they have found applications in gene therapy, as they can be attached to nucleases and used to excise specific DNA sequences. Some researchers have recently demonstrated success in using zinc-finger nucleases to treat viral diseases such as HIV [2] and Hepatitis B [3] by destroying their DNA or required cellular co-receptors. However, the role of the divalent ion in the stability of the protein is not well known. Understanding the effect of mutations on the thermodynamic stability of the Zn^{2+} ion in the zinc finger allows for more stable synthetic zinc fingers to be constructed, increasing the ease and efficacy of using zinc-finger nucleases and other similar therapeutics. Also, as highly specific zinc-binding proteins, these could be used with an embedded fluorescent group as a sensitive zinc detector [4]. Understanding the comparative thermodynamic stability of such a protein to other ions will allow for the development of better models of metalloprotein binding.

Modern experimental methods, such as X-ray diffraction (XRD) and nuclear magnetic resonance (NMR) allow for the characterization of crystal structures. However, XRD cannot be used to determine solution structures, as the protein must be crystallized. Computer simulations must be used to solvate such structures. NMR can be used to determine solution structures, but computer simulations are used to refine such structures to yield more accurate descriptions of the protein conformation. One of the major computer simulation methods used is molecular dynamics (MD). MD simulations are atomistic simulations that use classical mechanics to determine trajectories of the atoms from an initial configuration. These simulations allow for the refinement of experimental structures as well as determination of basic thermodynamic quantities such as solvation free energies and interaction energies. The atomistic nature of MD simulations also allows insight into the mechanisms of atomic processes. MD simulations can be used to generate trajectories that are several ns in length for systems with tens of thousands of atoms. However, for metal ions, classical molecular dynamics simulations are unable to capture the quantum mechanical effects such as charge transfer and polarization. Other computational models must be used to accurately simulate the behavior of the metal ion.

Another method of computational structure determination is based on electronic structure theory. Quantum mechanical (QM) calculations, also known as *ab initio* calculations, derive results from the wavefunctions of the electrons in the system by solving self-consistent field equations. Thermodynamic properties can then be calculated from those wavefunctions using the harmonic approximation [5]. In general, this method is more accurate for calculating structural and thermodynamic properties than molecular dynamics, but is much more computationally demanding. Because of this limitation, QM calculations are only feasible for small systems on the order of tens of atoms. Also, QM does not capture the dynamics of the system. Newer methods such as ab initio molecular dynamics (AIMD) are able to perform dynamics

on such systems, but can only be feasibly used with tens of atoms over periods of under 1 ns.

In order to simulate systems with hundreds of atoms accounting for quantum effects on nanosecond timescales, several methods have been devised to link quantum calculations with faster approximations. One is the hybrid quantum mechanics/-molecular mechanics (QM/MM) method [6], where a subset of the atoms are treated with quantum mechanics and the rest of the atoms are treated using molecular (classical) mechanics. The two types of simulations are linked together using link atoms that are treated both ways and are attached to atoms from both sets. However, this method is not yet feasible for generating trajectories of protein simulations (tens of thousands of atoms) of sufficient length (several ns) for determining thermodynamic properties.

Another method is the quasi-chemical approximation [7], which partitions molecules into an inner shell and an outer shell. Under this model, the process of solvation can be decomposed into two parts: the formation of the inner shell and the solvation of the inner shell in the outer shell. Quantum mechanics are used to calculate free energies of formation of the inner shell complex, and the free energy of solvation is determined using continuum electrostatics. This method has been used with success on ions in water [8]. Such an approximation is much faster than any quantum simulation, yet also allows for accurate determination of free energies of solvation. However, this approach has not been used for mixtures of water and other solvents, which is a plausible model for considering a protein as a "solvation shell" for an ion.

2 Methods

2.1 Theory

2.1.1 Molecular Dynamics Simulations

Molecular dynamics simulations represent each atom as a point charge and mass with a Lennard-Jones 12-6 potential for van der Waals type forces. Interaction energies are provided from force fields and numerical integration is used to predict movement of atoms, generating an ensemble of states. Free energies can be calculated from these simulations using the principles of statistical mechanics. From the inverse potential distribution theorem [7], the excess chemical potential of a molecule X is

$$\mu_X^{ex} = \frac{1}{\beta} \log \left[\int_{-\infty}^{\infty} e^{\beta\epsilon} P_X(\epsilon)\, d\epsilon \right] \tag{2.1}$$

where $\beta = \frac{1}{k_B T}$, ϵ is the interaction energies between X and all other species, and P_X is the probability density function of ϵ. Assuming a gaussian distribution for the energies ($P_X(\epsilon) \propto e^{\frac{(\epsilon - \langle\epsilon\rangle)^2}{2\sigma^2}}$) and substituting this into Equation 2.1 results in

$$\mu_X^{ex} = \langle\epsilon\rangle + \frac{\beta\sigma^2}{2} \tag{2.2}$$

2.1.2 Quantum Mechanical Calculations

In QM calculations, an initial geometry is specified, and the wavefunctions of the electrons are solved to determine the internal energy of the system. The second derivatives of the energy functions can then be analyzed to determine the electronic vibrational structure of the system. Such frequency calculations can then be used to determine free energies of formation based on the harmonic approximation [5].

The method of solving the wavefunctions used in this case is density functional theory (DFT), where functionals of the spatial density of electrons are derived from

experiment or higher-level ab initio calculations. Since the functionals contain information about different atoms, such an approach is faster than pure ab initio approaches, which do not consider such information.

2.1.3 Continuum electrostatics

In lieu of explicit solvent, solvent effects can be accounted for by treating them from an electrostatics perspective. This is due to the fact that at the distances considered as the outer shell (several $\mathring{A}$), the van der Waals forces are typically much smaller than the electrostatic interaction with the zinc ion. The generalized Born solvation model [9] was used for the determination of the free energy changes. In this model, the free energy of solvation is approximated as the free energy of taking a charged object of a certain shape from a medium with a dielectric constant of 1 to a medium with the dielectric constant of the desired solvent.

2.1.4 Quasi-Chemical Model

In the quasi-chemical approximation, we decompose the system into two parts: we define an inner shell and an outer shell, and combine the two to determine the solvation free energy. The contribution of the outer shell is defined as the solvation free energy of the complex, and can be calculated using continuum electrostatics. The free energy of the inner shell is defined chemically: given a species X and ligand L, the free energy of the formation of a cluster XL_n is the free energy change of the following reaction:

$$X + nL \rightleftharpoons XL_n$$

This is equivalent to the expression $-kT \ln K_n^{(0)}$, where $K_n^{(0)}$ is the ideal gas reaction rate constant at 1 atm and 298 K. This value can be calculated using QM methods. In the case of ions in solvent, the X is the ion species, and the L are the solvent molecules around the ion. A more thorough derivation of quasi-chemical theory can

be found in [7].

The free energies are calculated in an ideal gas state at 298 K and 1 atm, but we are interested in the liquid state. To correct for this state, the free energy change due to a pressure change is $nkT \ln\left(\frac{P_{liq}}{P_0}\right)$, where P_{liq} is $RT\rho_w$ and ρ_w is the molar density of the solvent. This yields the ideal gas pressure at liquid densities. P_0 is 1 atm. This term is also written as $-kT \ln \rho^n$, as it is a density factor, and can be combined with the previous term as $-kT \ln K_n^{(0)} \rho^n$. A further discussion on this topic can be found in Grabowski et al [10].

A particular species X may also have coordination shells with different n. In that case, a term $kT \ln p_X(n)$ is included, where $p_X(n)$ is the probability of X being coordinated by n L's. This term accounts for the fact that this particular arrangement only contributes partially to the free energy. A first-order approximation is to consider only the most probable coordination number $\tilde{n}$. A more thorough discussion of this term can be found in Asthagiri et al [11].

Combining these terms together results in the expression

$$\mu_X^{ex} = -kT \ln K_{\tilde{n}}^{(0)} \rho^{\tilde{n}} + kT \ln p_X(\tilde{n}) + \left(\mu_{XL_{\tilde{n}}}^{ex} - \tilde{n}\mu_L^{ex}\right) \tag{2.3}$$

To extend this method to multiple ligand species, we can consider two ligands, L_1 and L_2. The reaction is then

$$X + nL_1 + mL_2 \rightleftharpoons XL_{1n}L_{2m}$$

For a first order approximation of the density factor, one can calculate the density of the solution by assuming that this is an ideal solution and approximating the liquid molar density as $\rho_{L_1L_2} = x_{L_1}\rho_{L_1} + x_{L_2}\rho_{L_2}$. This leads to $P_{L_1L_2} = kT\rho_{L_1L_2}$. Although this may not be true, the effect of the term is not very significant in the calculations (see Section 3.2.4). A more accurate approximation could also be per-

formed considering excess molar volumes. The corresponding density factor is then $(n+m)\,kT\ln\left(\frac{P_{L_1L_2}}{P_0}\right)$. Like before, we may write $\frac{P_{L_1L_2}}{P_0}=\rho$. The probability term can be modified to account for changes in the coordination numbers of n and m. The free energy is then determined to be

$$\begin{aligned}\mu_X^{ex} = &- kT\ln K^0_{n,m}\rho^{(n+m)} + kT\ln p_X(n,m)\\ &+ \left(\mu^{ex}_{XL_{1n}L_{2m}} - n\mu^{ex}_{L\ 1} - m\mu^{ex}_{L\ 2}\right)\end{aligned} \tag{2.4}$$

2.2 Calculation details

All calculations were performed on the GLENN cluster of the Ohio Supercomputer Center on dual socket nodes with two quad-core 2.5 GHz AMD Opteron processors. Calculation files can be found in Appendix A. More details on the specific calculations performed can be found in the Results section (Section 3).

2.2.1 Ions in solution

Simulation files were prepared using VMD 1.8.7 and the solvate plugin. A PDB file containing one ion was created by hand and solvated with a 32 $\mathring{A}$ TIP3P water box.

Molecular dynamics simulations were performed using NAMD 2.6b software [12] and the CHARMM 27 force field. The force field file can be found in Appendix A.2. NAMD input files can be found in Appendix A.3. Long-range electrostatics were treated using particle mesh Ewald (PME) and periodic boundary conditions were used. The nonbonded cutoff was set at 12 $\mathring{A}$. Langevin bath and a Nose-Hoover barostat at 298 K and 1 atm were used to generate NPT ensembles. A 2 fs timestep was used. Waters and methanols were constrained with the SHAKE algorithm. The system was minimized at 0K for 7500 steps, velocities were reinitialized to 298K, and the system was equilibrated for 200 ps. Production runs of 20 ns were performed. Frames were sampled every 250 fs. Pair interaction energies of the different frames

in the trajectory were calculated using NAMD. The free energy was then calculated according to Equation 2.2 using MATLAB R2011b.

A QM calculation of the Zn^{2+} ion was performed. DFT calculations were performed using Gaussian 09 Revision A.01 [13]. The Becke three-parameter Lee-Yang-Parr (B3LYP) hybrid exchange-correlation functional and the 6-311++G(2d,p) basis set was used. Initial structures were obtained from MD simulations. Structures were first optimized on a 6-31+ basis set, then on a 6-311+G(d,p) basis set, and finally on the 6-311++G(2d,p) at the VeryTight optimization level. Free energies were obtained using frequency analysis on optimized structures. No imaginary frequencies were found, indicating stable structures. Population analysis according to the CHELPG [14] procedure was used to determine charges. Radii for the different atoms were taken from Stefanovich et al [15]. The radius for Zn^{2+} was set at 2.2 $\mathring{A}$, as the value did not change the calculation much. Free energies of solvation were then calculated according to the generalized Born solvation model using the Adaptive Poisson-Boltzman Solver (APBS) 1.3 [16].

2.2.2 Zn^{2+} in methanol/water solutions

Coordination shells with specific numbers of methanols and waters were created by hand-modifying a PDB file generated from the QM optimization of the Zn^{2+} ion above. These were energy-minimized at 0 K in NAMD using the OPLS-ua force field to produce optimized structures. The force field file can be found in A.2. The optimized structures were put in a 2500-atom methanol/water solvent box with 0%, 5%,10%, 15% or 100% MeOH mol% using PACKMOL [17].

Molecular dynamics simulations were performed using NAMD 2.6b software [12] and the OPLS-ua force field. The force field file can be found in Appendix A.2. NAMD input files can be found in Appendix A.3. Long-range electrostatics were treated using particle mesh Ewald (PME) and periodic boundary conditions were

used. The nonbonded cutoff was set at 12 $\AA$. Langevin bath and a Nose-Hoover barostat at 298 K and 1 atm were used to generate NPT ensembles. A 2 fs timestep was used. Waters and methanols were constrained with the SHAKE algorithm. The system was minimized at 0K for 7500 steps, velocities were reinitialized to 298K, and the system was equilibrated for 200 ps. Production runs of 20 ns were performed. Frames were sampled every 250 fs. Pair interaction energies of the different frames in the trajectory were calculated using NAMD. The free energy was then calculated according to Equation 2.2 using MATLAB R2011b.

DFT calculations were performed using Gaussian 09 Revision A.01 [13]. The Becke three-parameter Lee-Yang-Parr (B3LYP) hybrid exchange-correlation functional and the 6-311++G(2d,p) basis set was used. Initial structures were derived from MD simulations. Structures were first optimized on a 6-31+ basis set, then on a 6-311+G(d,p) basis set, and finally on the 6-311++G(2d,p) at the VeryTight optimization level. Free energies were obtained using frequency analysis on optimized structures. No imaginary frequencies were found, indicating stable structures. Population analysis according to the CHELPG [14] procedure was used to determine charges. Radii for the different atoms were taken from Stefanovich et al [15]. The radius for Zn^{2+} was set at 2.2 $\AA$, as the value did not change the calculation much. Free energies of solvation were then calculated according to the generalized Born solvation model using the Adaptive Poisson-Boltzman Solver (APBS) 1.3 [16].

2.2.3 Zn^{2+} in zinc finger protein

The zinc finger motif used consisted of residues 42 to 71 of chain A of the TFIIIA zinc finger (PDB: 1TF6). The zinc finger motif was extracted from the NMR structure, and then hydrogen atoms were added using the PSFGEN module of VMD 1.9.1 [18]. The histidine residues coordinated to the zinc were changed to the neutral (HID) form and the cysteine residues were deprotonated. The protein was then solvated in

a $48\mathring{A} \times 48\mathring{A} \times 54\mathring{A}$ TIP3P water box.

Molecular dynamics simulations were performed using NAMD 2.6b software [12] and two force fields. The CHARMM 27 force field modified to include a deprotonated cysteine, which was obtained from parameters for methylthiolate. The AMBER FF09 force field was not modified. The force field files can be found in Appendix A.2. NAMD input files can be found in Appendix A.3. Long-range electrostatics were treated using particle mesh Ewald (PME) and periodic boundary conditions were used. The nonbonded cutoff was set at 12 $\mathring{A}$. Langevin bath and a Nose-Hoover barostat at 298 K and 1 atm were used to generate NPT ensembles. A 2 fs timestep was used. Waters were constrained with the SHAKE algorithm. The system was minimized at 0K for 10000 steps and then heated from 0 K to 298 K over 50 ps and equilibrated for 200 ps. Production runs of 10 ns were performed. Frames were sampled every 250 fs. Pair interaction energies of the different frames in the trajectory were calculated using NAMD. The free energy was then calculated according to Equation 2.2 using MATLAB R2011b.

DFT calculations were performed using Gaussian 09 Revision A.01 [13]. The Becke three-parameter Lee-Yang-Parr (B3LYP) hybrid exchange-correlation functional and the 6-311++G(2d,p) basis set was used. Initial structures were obtained from the NMR structure. Structures were first optimized on a 6-31+ basis set, then on a 6-311+G(d,p) basis set, and finally on the 6-311++G(2d,p) at the VeryTight optimization level. Free energies were obtained using frequency analysis on optimized structures. No imaginary frequencies were found, indicating stable structures. Population analysis according to the CHELPG [14] procedure was used to determine charges. Radii for the different atoms were taken from Stefanovich et al [15]. The radius for Zn^{2+} was set at 2.2 $\mathring{A}$, as the value did not change the calculation much. Free energies of solvation were then calculated according to the generalized Born solvation model using the Adaptive Poisson-Boltzman Solver (APBS) 1.3 [16].

QM/MM calculations were performed using NWCHEM 6.1 [19]. The α-carbons of the coordinating protein residues were considered to be the boundary of the QM region. The rest of the protein and a water shell were considered using molecular mechanics. The Amber FF09 force field was used for the MM portion. DFT using the LANL08DZ basis set, which is optimized for transition metals, was used with the B3LYP hybrid functional. Hydrogen link atoms were used. Optimization was performed using the conjugate gradient algorithm. The QM/MM boundary cutoff was set at 9 $\AA$. Electrostatic potential fitting of the QM region was used for parameterization of the QM atoms during the MM optimization. Ten cycles of optimization were performed, with a maximum of 50000 solvent iterations, 10000 protein iterations, and 500 QM core iterations. The resulting optimized structure was then further refined using the DFT calculation methodology above except with the alpha carbons fixed at the optimized positions.

3 Results

3.1 Ions in solution

3.1.1 MD simulations of Na^+, K^+ and Zn^{2+} in water

MD simulations of ions in water were performed to verify that results obtained from NAMD were accurate. Interaction energies of the ion with the waters were calculated and fitted to a normal distribution (see Figure 1 and Figure 2).

From these plots, it was determined that the energies were indeed normally distributed, so the free energies were calculated using Eqn 2.2. Results for the different ions (Na^+, K^+, and Zn^{2+}) are tabulated in Table 1.

The experimental and quasi-chemical values were from Asthagiri et al [8]. From these calculations, it becomes evident that although the free energies calculated using

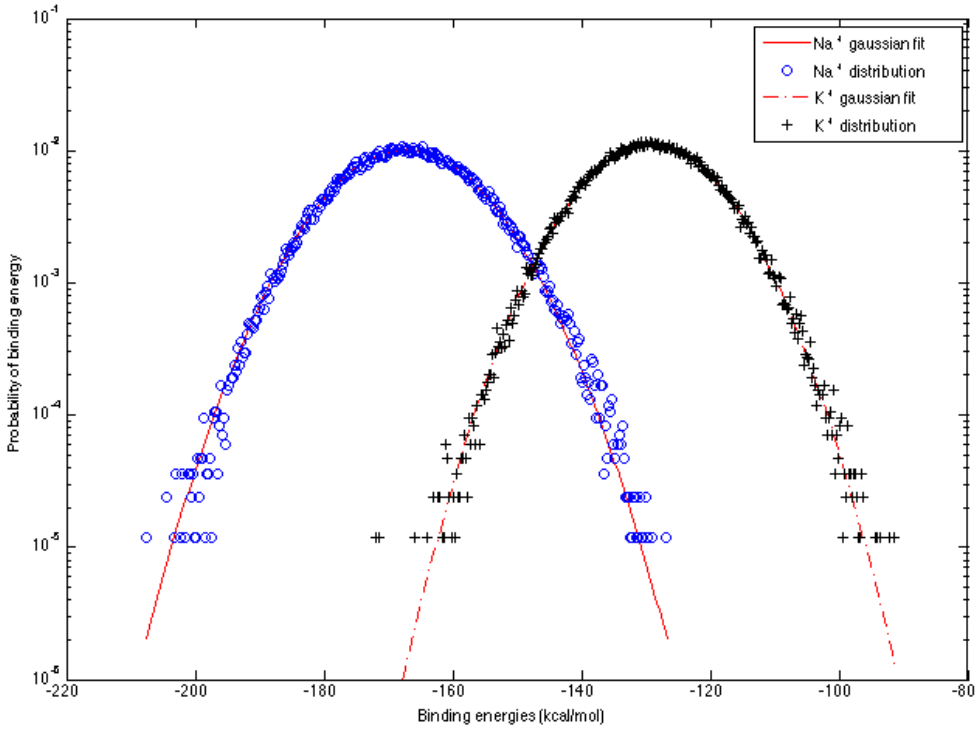


Figure 1: Na^+ and K^+ interaction energy distributions and gaussian fits from MD simulations.

this method are reasonable, the error increases with increasing atomic number and may be larger than desired. Because of this, a more accurate method will be used to determine free energies.

3.1.2 Quasi-chemical calculation of Zn^{2+} hydration free energy

Quasi chemical theory was applied to the zinc ion in particular to determine more accurate binding free energies. The inner shell radius was determined from the first minimum of the Zn^{2+}-O radial distribution function (see Figure 3). The radial distribution function was calculated from the MD simulations of Zn^{2+} in water.

It was found to contain 6 water molecules. Trajectory analysis determined that the waters coordinating the ligands did not exchange, thus there was a sharply-defined inner shell. QM calculations of the core were then performed according to

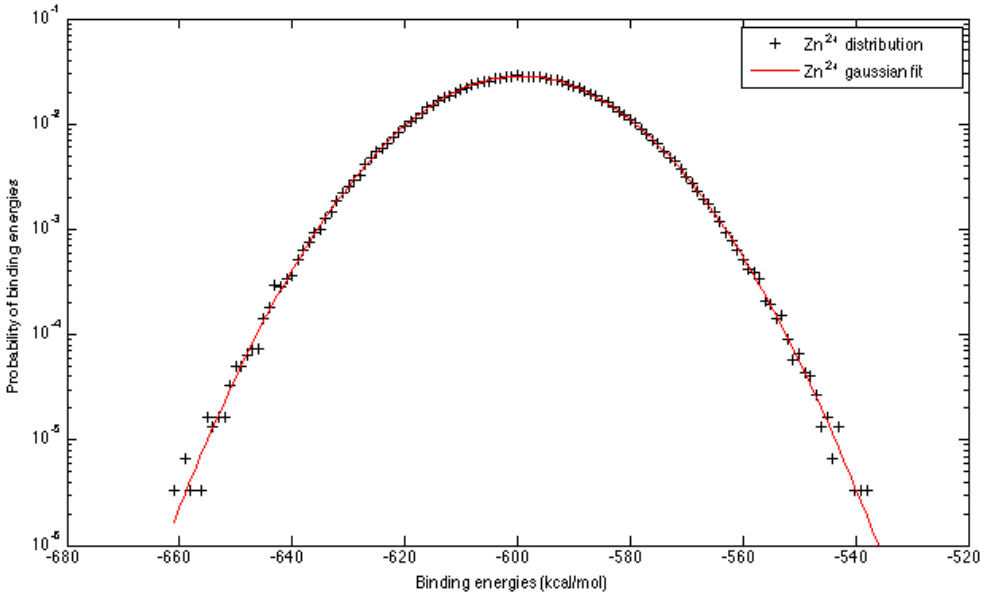


Figure 2: Zn^{2+} interaction energy distribution and gaussian fit from MD simulation.

Table 1: Solvation free energies for ions in water using MD compared to experimental values and quasi-chemical values

Units in $\left(\frac{kcal}{mol}\right)$	$\langle\epsilon\rangle$	$\beta\frac{\sigma^2}{2}$	μ_X^{ex}	QC [8]	Exp. [8]
Na^+	-166.9 ± 0.3	81.8 ± 1.4	-85.1 ± 1.5	-96.1	-91.5
K^+	-129.3 ± 0.2	67.2 ± 0.6	-61.9 ± 0.5	-75.2	-74.6
Zn^{2+}	-598.9 ± 0.2	165.2 ± 1.5	-433.8 ± 1.5	-458.1	-471.1

Standard deviations of values from dividing the simulation into 4 time blocks follow $\pm$.

the methodology in 2.2.3. The zinc solvation free energy and different components of the quasi-chemical free energy of solvation expression (Equation 2.3) can be found in Table 2.

The results agreed rather well with the results from Asthagiri et al [8].

3.2 Zn^{2+} in methanol / water mixtures

In order to extend quasi-chemical theory to mixtures, the Zn^{2+} ion was solvated in mixtures of methanol and water. Parameters for these simulations were derived from

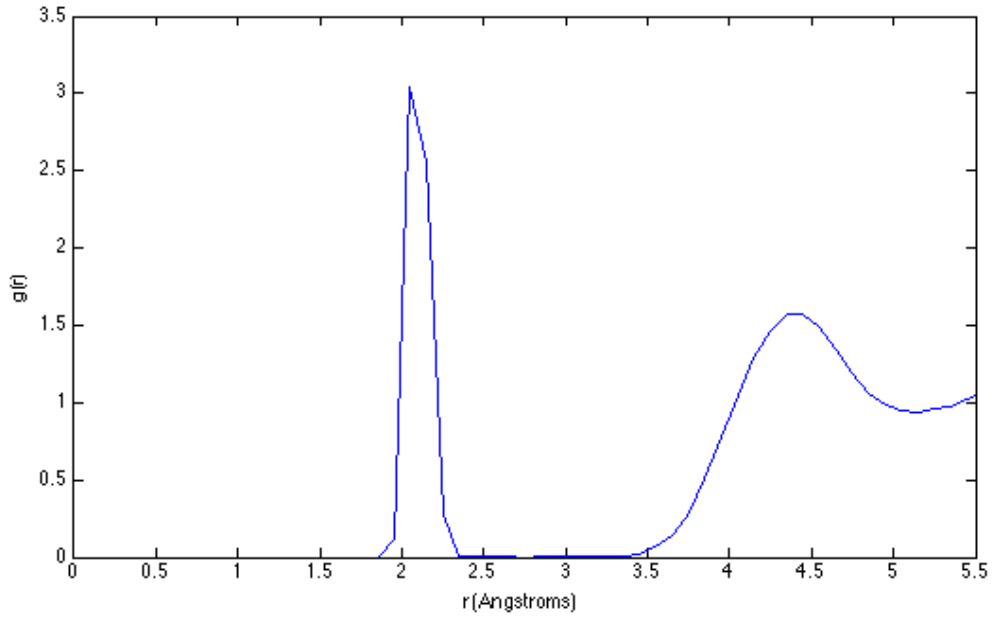


Figure 3: Zn^{2+}-OH_2 radial distribution function

Table 2: Quasi-chemical calculation of Zn^{2+} hydration free energy

$\left(\frac{kcal}{mol}\right)$	$-kT \ln K^{(0)}_{\tilde{n}}$	$\mu^{ex}_{XL_n}$	μ^{ex}_{L}	μ^{ex}_{X}	Experimental [8]
Asthagiri et al [8]	-279.2	-199.3	-7.7	-458.1	-471.1
This work	-282.8	-202.6	-8.1	-462.4	-471.1

the Optimized Potentials for Liquid Simulations United-Atom (OPLS ua) force field. In this force field, the methyl group of the methanol was approximated as a united atom. Simulations of 2500 solvent atoms and a zinc ion were run at methanol mol% of 0%, 5%, 10%, 15%, and 100%. It was found that the Zn^{2+} ion bound the six atoms around it from the initial configuration very tightly: no atoms coordinated to the zinc ion after minimization exchanged with other atoms in any of the 20 ns NPT production runs at any methanol mole fraction. Based on these results, MD simulations of the Zn^{2+} ion in water were performed with different coordination shells containing i waters and $6 - i$ methanols for i ranging from 0 to 4.

3.2.1 Free energies of solvation

Interaction energies between the zinc ion and the all outer shell (non-coordinated solvent) were calculated. The energies were normally distributed. A sample energy histogram can be found in Figure 4.

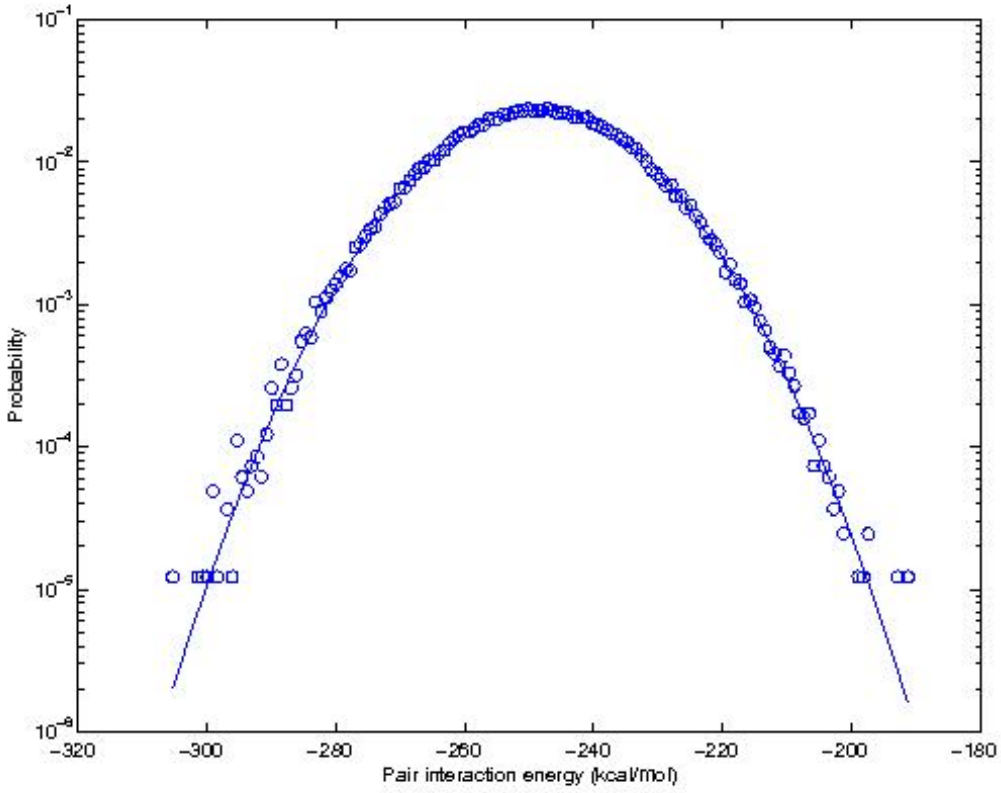


Figure 4: Interaction energy distribution of Zn^{2+} with non-coordinated solvent molecules (outer shell radius: 3.5 $\AA$) in 15% MeOH and water

Histograms for the other systems can be found in Appendix C.2. The gaussian approximations of the interaction energy distributions were deemed valid, and the solvation free energies were calculated with Equation 2.2. The free energies of solvation are tabulated in Table 3.

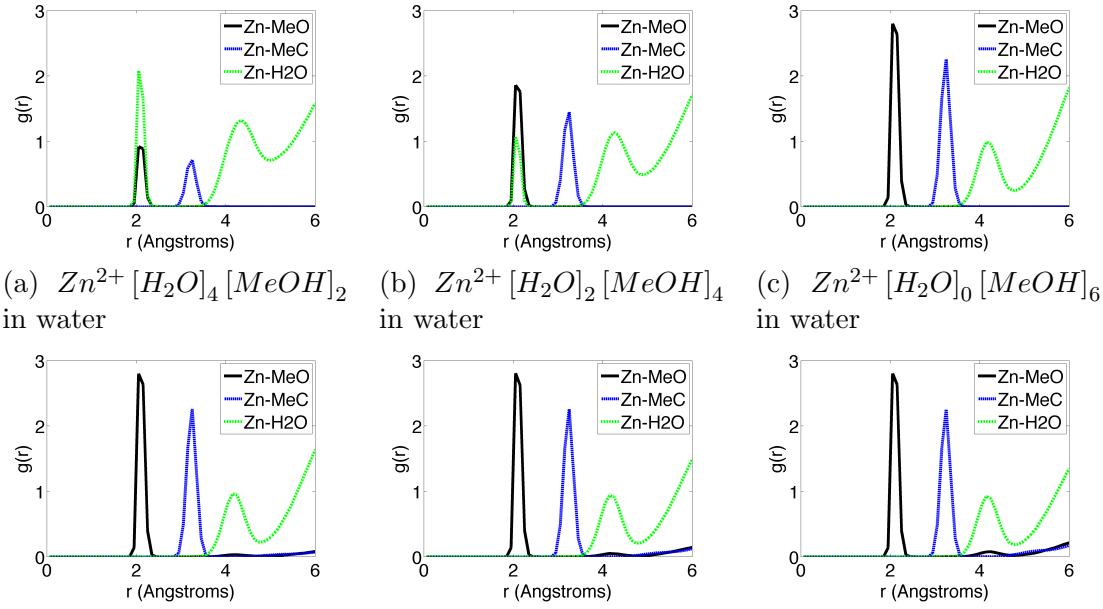
Table 3: Calculation of Zn^{2+} ion solvation free energies from MD simulations for various solvents

Inner shell methanols	Methanol mol%	Zinc ion binding free energy
2	0	-443
3	0	-438
4	0	-433
5	0	-432
6	0	-428
6	5	–429
6	10	-432
6	15	-434

3.2.2 Radial distribution functions

Radial distribution functions between Zn^{2+} and water oxygen and methanol oxygen and methyl groups were calculated using VMD 1.9.1. [18]. Some are shown below in Figure 5. The other radial distribution functions can be found in Appendix C.1.

From the radial distribution functions, we see that the locations of the peaks do not move with changing solvent methanol composition or coordination shell methanol composition, but the heights of the peaks change. This shows that effect of the methanol on the general structure of the inner shell is minimal, but it does push the outer shell further away from the zinc. This is what most likely explains the trend observed in the solvation free energies above: adding more methanol to the inner shell decreases the stability of the system by virtue of its size. The plots also show that a radius of 3.5 $\mathring{A}$ for the outer shell is a reasonable choice.



(a) $Zn^{2+}\,[H_2O]_4\,[MeOH]_2$ in water (b) $Zn^{2+}\,[H_2O]_2\,[MeOH]_4$ in water (c) $Zn^{2+}\,[H_2O]_0\,[MeOH]_6$ in water

(d) $Zn^{2+}\,[H_2O]_0\,[MeOH]_6$ in 5% MeOH (e) $Zn^{2+}\,[H_2O]_0\,[MeOH]_6$ in 10% MeOH (f) $Zn^{2+}\,[H_2O]_0\,[MeOH]_6$ in 15% MeOH

Figure 5: Radial distribution functions

3.2.3 Occupancy probabilities

Occupancy probabilities were determined using a VMD Tcl script (See Appendix B.2). For the quasi-chemical approximation, a specific number of each kind of atom in the inner shell is assumed. To verify that this assumption is reasonable, the probability that the area around the ion is occupied by a certain number of waters and methanols was calculated. Table 4 shows the probabilities that an *incorrect* number of atoms each non-hydrogen type were observed. Correctness was defined as the following: for simulations with n MeOH in the coordinated to the ion, there should be n methanol oxygens, n methanol methyl groups, and $6 - n$ water oxygens within the 3.5 $\mathring{A}$ boundary.

From the small probabilities in Table 4, we see that the assumption that the shell consists of n methanols and $6 - n$ waters is reasonable.

Table 4: Probabilities that incorrect numbers of each atom type are within 3.5 $\mathring{A}$ of Zn^{2+}

$n = MeOH$	MeOH mol%	$P\,(Me\mathbf{O}H \neq n)$	$P\,(\mathbf{Me}OH \neq n)$	$P\,(\mathbf{O}H_2 \neq 6-n)$
2	0	0	9.1×10^{-3}	2.2×10^{-2}
3	0	0	1.4×10^{-2}	1.6×10^{-2}
4	0	0	1.9×10^{-2}	1.2×10^{-2}
5	0	0	2.6×10^{-2}	9.9×10^{-3}
6	0	0	3.2×10^{-2}	7.4×10^{-3}
6	5	6.7×10^{-5}	3.2×10^{-2}	6.8×10^{-3}
6	10	2.3×10^{-4}	3.3×10^{-2}	7.0×10^{-3}
6	15	2.4×10^{-3}	3.2×10^{-2}	6.9×10^{-3}

These are the probabilities that the specified number of the centers of the atom type in **bold** is within 3.5$\mathring{A}$ of the ion.

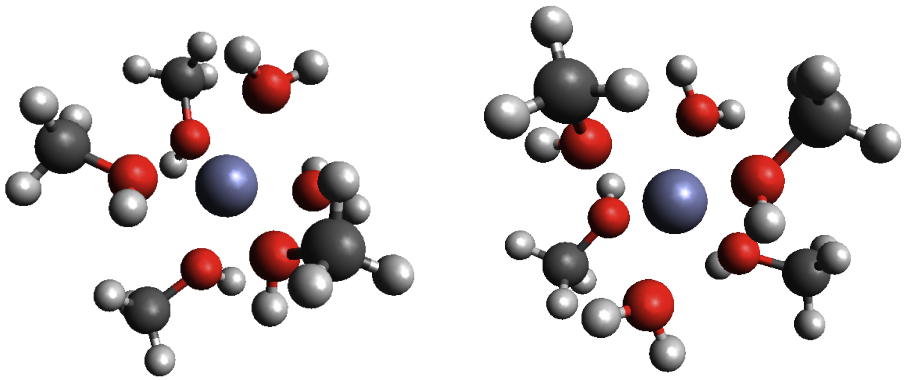
3.2.4 Quasi-chemical calculations of free energies

Based on the observation that the zinc ion bound the nearest six molecules tightly, and that it always exhibited a hexacoordinated geometry, QM calculations of the Zn^{2+} ion with different coordination shells containing i waters and $6 - i$ methanols for i ranging from 0 to 6 were performed, and Equation 2.4 was used to determine the free energies for all the coordination states. It was determined that there were several isomers of some solvation shells. An example of this is shown in Figure 6

QM optimizations of the geometries from Figure 6 revealed that the energy differences were negligible ($< 0.1\ \frac{kcal}{mol}$), hence only one isomer was used.

A table (Table 5) of the solvation free energies in water of each coordination state (neglecting the probability term in Equation 2.4) and their contributions is shown below.

Also, we assume ideal solutions to calculate the standard state correction ($-kT \ln \rho^6$) (see Section 2.1.4). However, since the term is not very large, such an approximation



(a) Waters adjacent (b) Waters nonadjacent

Figure 6: Two isomers of a Zn^{2+} ion coordinated with 4 MeOH and 2 H_2O.

Table 5: Quasi-chemical calculation of Zn^{2+} solvation free energies with different coordination numbers

Units in $\left(\frac{kcal}{mol}\right)$	$kT \ln K^{(0)}_{n,m}$	$-kT \ln \rho^6$	$\mu^{ex}_{XL_{1n}L_{2m}}$	$n\mu^{ex}_{L_1} + m\mu^{ex}_{L_2}$	μ^{ex}_X
$Zn^{2+}\left[H_2O\right]_6\left[MeOH\right]_0$	-283	-25.6	-203	-49	-462
$Zn^{2+}\left[H_2O\right]_5\left[MeOH\right]_1$	-285	-25.3	-195	-46	-459
$Zn^{2+}\left[H_2O\right]_4\left[MeOH\right]_2$	-287	-24.9	-187	-43	-455
$Zn^{2+}\left[H_2O\right]_3\left[MeOH\right]_3$	-289	-24.5	-180	-41	-453
$Zn^{2+}\left[H_2O\right]_2\left[MeOH\right]_4$	-291	-24.0	-174	-39	-451
$Zn^{2+}\left[H_2O\right]_1\left[MeOH\right]_5$	-292	-23.5	-168	-36	-447
$Zn^{2+}\left[H_2O\right]_0\left[MeOH\right]_6$	-294	-22.8	-162	-34	-445

will not affect the calculation much.

To account for the different solvents, the dielectric constant was adjusted using the methodology in [20]. Table 6 shows the solvation free energies in various mixtures.

Table 6: Quasi-chemical calculation of Zn^{2+} solvation free energies in water/methanol mixtures

MeOH mol%	0%		15%		100%	
Dielectric Constant	78.4		65		33	
Units in $\left(\frac{kcal}{mol}\right)$	$\mu^{ex}_{XL_{1n}L_{2m}}$	$\mu^{ex}_{L_{1n}L_{2n}}$	$\mu^{ex}_{XL_{1n}L_{2m}}$	$\mu^{ex}_{L_{1n}L_{2n}}$	$\mu^{ex}_{XL_{1n}L_{2m}}$	$\mu^{ex}_{L_{1n}L_{2n}}$
$Zn^{2+}\left[H_2O\right]_6\left[MeOH\right]_0$	-203	-49	-202	-48	-199	-47
$Zn^{2+}\left[H_2O\right]_5\left[MeOH\right]_1$	-195	-46	-194	-46	-191	-45
$Zn^{2+}\left[H_2O\right]_4\left[MeOH\right]_2$	-187	-44	-186	-44	-183	-42
$Zn^{2+}\left[H_2O\right]_3\left[MeOH\right]_3$	-180	-41	-179	-41	-176	-40
$Zn^{2+}\left[H_2O\right]_2\left[MeOH\right]_4$	-174	-39	-174	-39	-171	-38
$Zn^{2+}\left[H_2O\right]_1\left[MeOH\right]_5$	-168	-36	-167	-36	-165	-35
$Zn^{2+}\left[H_2O\right]_0\left[MeOH\right]_6$	-162	-34	-162	-34	-159	-33

We see that there are no large differences in free energies from this method along the range of MeOH mol% studied. The main differences in free energy arise from the $\mu^{ex}_{XL_{1n}L_{2m}}$ term. From this calculation, it appears that the water-coordinated state is most favorable even in pure methanol.

Since no exchanges were observed between atoms in the inner shell and atoms outside the inner shell in the simulations, it was not possible to determine the relative probabilities of each coordination state from the trajectories. However, based on the free energy differences of each coordination state, it is possible to determine the relative probabilities: for each coordination state, μ^{ex}_X will be the same, with the differences between the free energies accounted for with the $-kT\ln p_X\left(0\right)$ term. Solving for the relative probability in this manner, we obtain the probabilities of each configuration shown in Table 7.

Table 7: Probabilities of observing different coordination states of Zn^{2+}

Units in $\left(\frac{kcal}{mol}\right)$	μ_X^{ex}	$p_X\,(n,m)$
$Zn^{2+}\,[H_2O]_6\,[MeOH]_0$	-462	0.997
$Zn^{2+}\,[H_2O]_5\,[MeOH]_1$	-459	2.33×10^{-3}
$Zn^{2+}\,[H_2O]_4\,[MeOH]_2$	-455	7.02×10^{-6}
$Zn^{2+}\,[H_2O]_3\,[MeOH]_3$	-453	7.14×10^{-8}
$Zn^{2+}\,[H_2O]_2\,[MeOH]_4$	-451	3.09×10^{-9}
$Zn^{2+}\,[H_2O]_1\,[MeOH]_5$	-447	1.03×10^{-11}
$Zn^{2+}\,[H_2O]_0\,[MeOH]_6$	-445	1.35×10^{-13}

We see that at equilibrium, Zn^{2+} would be coordinated to six waters. Since the free energies in differing solvent do not change by much (see Table 6), we conclude that a six water coordinated Zn^{2+} ion will be the most stable in any water/methanol solvent.

3.2.5 Charge transfer on Zn^{2+} ions in QM calculations

The charges on the Zn^{2+} ion from the different QM calculations were computed using electrostatic fitting (see Table 8).

From these results, we can see that charge transfer plays a significant role in terms of quantum effects in this system. This verifies the need to treat the core with an electronic structure method.

3.3 TFIIIA zinc finger motif

3.3.1 AMBER FF09 simulations

MD simulations were also performed for the zinc finger motif present in TFIIIA. The zinc finger motif was taken from chain A of the TFIIIA, residues 42-71. These residues

Table 8: Charge transfer off of Zn^{2+} ion

Number of Methanols	Charge on $Zn^{2+}(e)$
0	2.09
1	1.92
2	1.75
3	1.60
4	1.34
5	1.09
6	0.80

P-F-P-***C***-K-E-E-G-***C***-E-K-G-F-T-S-L-H-H-L-T-R-***H***-S-L-T-***H***-T-G-E-K

Figure 7: Sequence of peptide used. Zinc ligands are in bold.

were chosen as this peptide's properties had been characterized by early studies [4]. The sequence of the peptide is shown in Figure 7.

The NMR structure of the protein with the zinc binding site is shown in Figure 8. Note that the zinc ion appears to be tetrahedrally coordinated by the four amino acids.

The motif was solvated in a $48\mathring{A} \times 48\mathring{A} \times 54\mathring{A}$ water box, minimized for 10000 steps, and heated to 298 K over 50 ps. Ten nanosecond NPT production runs were performed and were used for trajectory analysis. The CHARMM27 force field was unable to produce a stable structure with the zinc ion within the binding site. However, the AMBER FF09 force field was able to do so. Interaction energies between the Zn^{2+} ion and all other atoms were obtained from the trajectory. The interaction energies were normally distributed (see Figure 9). The solvation free energy was calculated using Equation 2.2, and the components are tabulated in Table 9.

In the equilibrated structure, the Zn^{2+} ion was coordinated to two waters in addition to the four protein ligands as seen in the NMR structure (see Figure 10). In

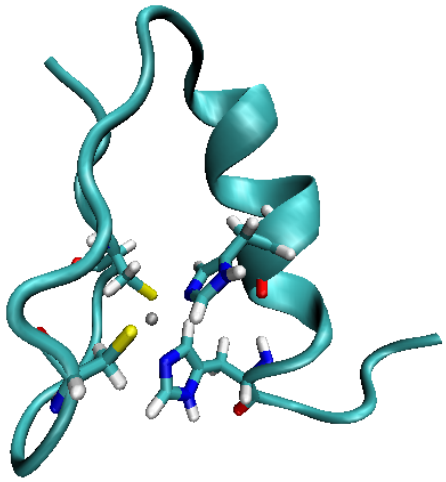


Figure 8: TFIIIA Zinc finger motif NMR structure with binding site. Grey ball is zinc, yellow is sulfur. The cysteine side chains are on the right, the histidines are on the left.

Table 9: Components of zinc finger solvation free energy

$\langle\epsilon\rangle$ $\left(\frac{kcal}{mol}\right)$	$\frac{\beta\sigma^2}{2}$ $\left(\frac{kcal}{mol}\right)$	μ_X^{ex} $\left(\frac{kcal}{mol}\right)$
-652	130	-522

order to reproduce the NMR structure, other methods were used to reparameterize the inner core. From the literature, it was determined that the cause of the extra waters in the binding site was inadequate consideration of quantum effects such as charge transfer and polarization [21].

3.3.2 Non-QM methods for accounting for quantum effects

To account for charge transfer, the charges on the Zn^{2+} ion and the surrounding ligands were modified by hand using the charges from Li et al [21], who used the electron densities from a QM optimization calculation to determine the charge transfer. This method did not, however, exclude the waters from entering the zinc finger. Using the CHARMM 27 force field with this parameterization resulted in a stable protein, but

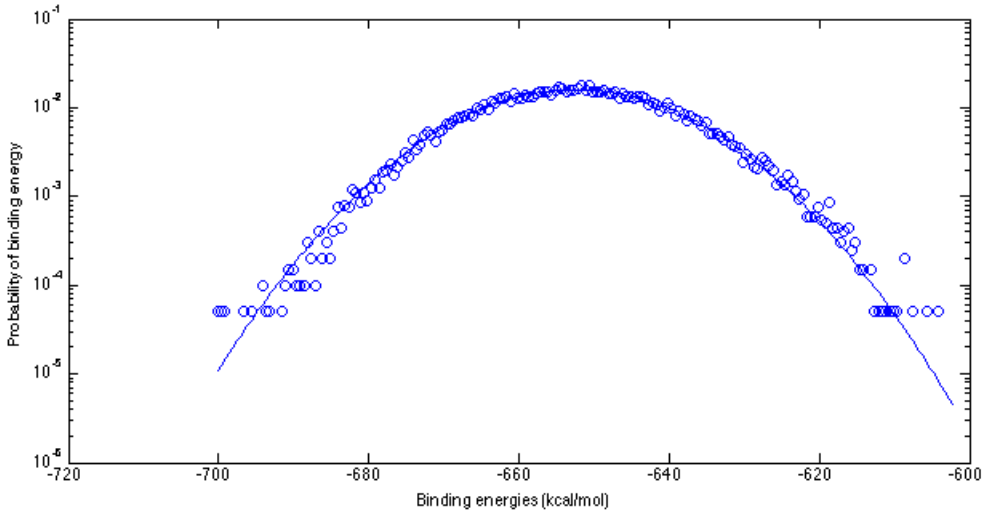


Figure 9: Zinc finger Zn^{2+} interaction energy distribution and gaussian fit.

with the waters in the binding site as shown in Figure 10.

Simulations were also performed using the cationic dummy atom model [22], which uses four charged "dummy atoms" attached to the zinc ion by springs in order to approximate polarization effects. This model was able to successfully predict the structure, but since it constrained the ligands, the vibrational modes would be affected. It was decided to explore chemical and quasi-chemical approaches for determining the free energies.

3.3.3 QM/MM calculations

QM/MM calculations were performed on the zinc finger to determine an optimized QM geometry for the core as well as an optimized MM geometry for the rest of the protein. The side chains of the protein and the zinc ion were treated with DFT using the LANL08DZ basis set, which is optimized for transition metals. The B3LYP hybrid functional was used. The rest of the atoms were treated using the AMBER FF09 force field. QM/MM dynamics at the time lengths required for determination of the free energies were not feasible. QM/MM methods were used to optimize geometries for the zinc finger core in order to have accurate geometries for usage with the quasi-chemical

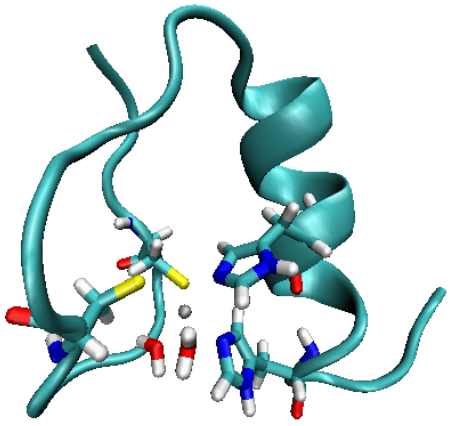


Figure 10: Zinc finger equilibrated structure. Note the two water molecules in the binding site.

method.

4 Discussion

4.1 Ions in water

4.1.1 MD simulations

From the results in Table 1, we conclude that these simulations do indeed contain good measures of the binding free energy of the Zn^{2+} ion to water. While there are errors on the order of tens of $\frac{kcal}{mol}$, the trends are the same. However, since we are extending these results to measuring stability changes in mutations of proteins, a more accurate method may be necessary.

4.1.2 Quasi-chemical calculation of Zn^{2+} hydration free energy

The hydration free energies calculated using the quasi-chemical method agreed very well with those of Asthigiri et al [8], which demonstrates that the procedure used

was accurate enough. The differences are mainly due to the usage of a larger basis set and a different method of solving the Poisson-Boltzman equation.

4.2 Zinc ion in water and methanol

4.2.1 Inner-shell stability

The zinc ion binds its coordinating ligands very tightly. Because of this, any exchanges were not able to be modeled, as the initial configuration of the zinc coordination shell did not change throughout the simulation. Increasing temperature to increase the fluctuations of the atoms was unsuccessful at dislodging any of the coordinating ligands. This effect can also be seen in the relatively sharp peaks of the radial distribution functions (see C.1). Because of this, it was impossible to use simulations to determine the most stable coordination state, so it was done using solvation free energies.

4.2.2 Effect of coordination shell composition

From the quasi-chemical calculations of the free energy of solvation with different coordination states (Table 5, the largest quantity in the calculation of the solvation free energy is the equilibrium constant term $-kT \ln K^{(}0)_{n,m}$. It decreases with increasing methanol around the zinc finger. Such an effect could be due to the fact that the larger system (more atoms) allows the electrons to delocalize more than with water. The outer shell contribution increases greatly with increasing methanols, and it dominates the differences in free energy. This is most likely due to the increasing size of the complex, which can be seen from the radial distribution functions C.1. The decreasing second-shell water peak and increasing third-shell water peaks are due to more of the water being pushed away from the ion. Since it takes more work to take a larger charged sphere into an electric field, it is reasonable that increasing the size would destabilize the system. We also see this pattern in the MD simulations (see

Table 3), as the cluster solvation free energies also increase by a similar amount (2-4 $\frac{kcal}{mol}$). Overall, the results suggest that the aqueous coordinated state is the most stable.

There is no experimental data regarding the solvation free energy of the zinc ions in water/methanol mixtures, so it is not possible to compare the simulation data to actual data. However, as the MD simulations and QC calculations agree in terms of the relative differences in free energy between the coordination states, these may still be good estimates of the actual values.

4.2.3 Effect of solvent composition

One concern with the simulations is the different trends with regards to changing the solvent. In the MD simulations (Table 3), it appears that increasing methanol in the solvent actually tends to stabilize the cluster by a significant amount. However, the opposite appears to be true concerning the QC calculations (Table 6). Additional methanol in the solvent (lowering the dielectric constant) is unfavorable, but the effect is not very large. These discrepancies could be due to the nature of the approximations made. In the QC calculations, the outer shell is treated as a purely electrostatic entity. However, along the edges of the cluster, the van der Waals forces could be playing an important role. Because of this, a more advanced quasi-chemical approximation with a higher-level representation of the outer shell may be necessary.

4.3 Zinc fingers

While this work did not truly explore the stability of zinc fingers, some insight was still gained into methods of modeling zinc fingers. Firstly, classical methods are unable to reproduce the tetracoordinated geometry of the zinc finger. This may be due to the nature of the parameterization of the zinc ion in force fields, as in solution, the zinc ion is indeed hexacoordinated.

4.3.1 Application of zinc ion in water/methanol mixtures to zinc finger thermodynamics

Two methods of applying the quasi-chemical method to zinc fingers arise. Firstly, one can consider the entire protein as the ligand in the inner shell. Secondly, one can consider the sidechains of the protein ligands as the inner shell, and the rest of the protein and the water as the outer shell. Both methods have challenges. In the first method, modeling the entire protein using QM would be unfeasible. Also, in order to use the quasi-chemical method, we need to know the solvation free energy of the protein without the zinc ion (see Equation 2.4). Since we do not know the conformation of the protein without the ion, it may be very difficult to estimate accurately. However, if we consider relative binding free energies of different ions instead, we can approximate the relative reaction free energy changes in the protein-zinc system by free energy changes in the binding site. The term with the protein without the ion cancels out. Thus, this method can be used to determine relative binding free energies.

For the second method, since the quasi-chemical model treats the outer shell electrostatically using the dielectric constant, a dielectric constant for the zinc finger protein is required. This can be calculated from MD trajectories. However, if we mutate only one residue, it may be difficult to have a significant effect on the dielectric constant. As can be seen from Table 6, small dielectric constant changes do not affect the solvation free energy much in that approximation. Also, in a protein, van der Waals and bonded interactions of the sidechains in the inner shell with the outer shell are more important.

To overcome these limitations, the outer shell component of the solvation free energy can also be approximated using the inverse potential distribution theorem and MD with the method described in Section 2.2.3. Several factors need to be addressed before such a calculation can take place. Firstly, the bonded interactions need to be

considered. One of the principle difficulties in QM/MM calculations, another method seeking to bridge quantum and classical simulations, is in the coupling of the two forms of simulation. Methodologies from QM/MM may also be applicable to QC calculations. Secondly, a method of simulating the protein with MD that gives good estimates of the outer-shell free energies is required. Several methods exist, but their effects on the calculation of free energies has yet to be considered. Thirdly, an accurate estimate of the geometry of the inner shell is necessary. This was performed using QM/MM geometry optimizations to determine an accurate inner shell geometry that also considered the overall conformation of the protein.

5 Conclusions

Several methods can be used to estimate the free energy of hydration of ions in solution. Two of these are quasi-chemical theory and the inverse potential distribution theorem. Using the IPDT with MD simulation data yields results accurate to within tens of $\frac{kcal}{mol}$. Quasi-chemical theory uses electronic structures to determine the formation and solvation free energies of a certain complex. The formulation of the IPDT lends itself naturally to mixtures in solutions.

QC theory was also expanded to include different ligands coordinating the species of interest. A derivation can be found in section 2.1.4. The solvation free energy of different coordination states of the Zn^{2+} ion was found to follow two opposite trends. Firstly, the free energy of formation of the inner shell was found to decrease with increasing methanols in the coordination shell. This may be due to the fact that a larger system allows for better electron delocalization, stabilizing the chemical equilibrium. Secondly, the free energy contribution of solvation of the inner shell cluster was found to increase with increasing methanols in the coordination shell. This may be due to the fact that a larger system also requires more work to solvate.

This was supported by MD simulations. The second trend was more dominant in the zinc-methanol-water system, and the most favorable hexacoordinated state was the zinc ion with six water molecules.

Extending QC theory to mixtures of solvents reduces to changing the dielectric constant of the solvent and using that to solvate the inner shell cluster. However, this yielded discrepancies with the MD simulations. These discrepancies require further investigation. It was observed, however, that the dielectric constant made little impact on the free energies in the case of water and methanol solutions. This may not be true for mixtures of fluids with more disparate dielectric constants.

From these observations, in order to adapt these methods for determining the binding free energies of the Zn^{2+} ions in zinc fingers, it will be beneficial to use one of two methods. First, the free energy change for different ions in the protein can be measured instead using a quasi-chemical model where the entire protein is in the inner shell as the ion ligand. Secondly, a combined method could be used. In this quasi-chemical framework, the "inner shell" would be considered using QM calculations as before. However, the "outer shell" would be treated using MD simulations and the IPDT to determine solvation free energies of the cluster. Some theoretical issues in computing the binding free energies include computing the core structure accurately, simulating the protein accurately using MD, and the treatment of bonds and van der Waals forces. Several approaches exist for the first and second challenges. The third problem has similarities with problems in calculating energies using QM/MM methods and similar approaches can be used. The ability to accurately compute binding free energies of such proteins will allow for the better design of stable synthetic zinc finger therapies for many genetic and viral disorders.

References

[1] K. J. Waldron and N. J. Robinson, "How do bacterial cells ensure that metalloproteins get the correct metal?," *Nat Rev Micro*, vol. 7, no. 1, pp. 25–35, 2009.

[2] N. Holt, J. Wang, K. Kim, G. Friedman, X. Wang, V. Taupin, G. Crooks, D. Kohn, P. Gregory, M. Holmes, *et al.*, "Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to ccr5 control hiv-1 in vivo," *Nature biotechnology*, vol. 28, no. 8, pp. 839–847, 2010.

[3] T. Cradick, K. Keck, S. Bradshaw, A. Jamieson, and A. McCaffrey, "Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis b virus dnas," *Molecular Therapy*, vol. 18, no. 5, pp. 947–954, 2010.

[4] J. M. Berg and H. A. Godwin, "Lessons from zinc-binding peptides.," *Annu Rev Biophys Biomol Struct*, vol. 26, pp. 357–371, 1997.

[5] J. W. Octerski, "Thermochemistry in gaussian." `http://www.gaussian.com/g_whitepap/thermo/thermo.pdf`, June 2000.

[6] H. Lin and D. Truhlar, "Qm/mm: what have we learned, where are we, and where do we go from here?," *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, vol. 117, no. 2, pp. 185–199, 2007.

[7] T. Beck, M. Paulaitis, and L. Pratt, *The potential distribution theorem and models of molecular solutions*. Cambridge Univ Pr, 2006.

[8] D. Asthagiri, L. R. Pratt, M. E. Paulaitis, and S. B. Rempe, "Hydration structure and free energy of biomolecularly specific aqueous dications, including zn2+ and first transition row metals.," *J Am Chem Soc*, vol. 126, no. 4, pp. 1285–1289, 2004 Feb 4.

[9] W. Still, A. Tempczyk, R. Hawley, and T. Hendrickson, "Semianalytical treatment of solvation for molecular mechanics and dynamics," *Journal of the American Chemical Society*, vol. 112, no. 16, pp. 6127–6129, 1990.

[10] P. Grabowski, D. Riccardi, M. Gomez, D. Asthagiri, and L. Pratt, "Quasichemical theory and the standard free energy of h+ (aq)," *The Journal of Physical Chemistry A*, vol. 106, no. 40, pp. 9145–9148, 2002.

[11] D. Asthagiri, P. Dixit, S. Merchant, M. Paulaitis, L. Pratt, S. Rempe, and S. Varma, "Ion selectivity from local configurations of ligands in solutions and ion channels," *Chemical Physics Letters*, vol. 485, no. 1-3, pp. 1–7, 2010.

[12] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–1802, 2005.

[13] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Or-

tiz, J. Cioslowski, and D. J. Fox, "Gaussian 09 Revision A.1." Gaussian Inc. Wallingford CT 2009.

[14] C. BRENEMAN and K. WIBERG, "Determining atom-centered monopoles from molecular electrostatic potentials - the need for high sampling density in formamide conformational-analysis," *Journal of Computational Chemistry*, vol. 11, pp. 361–373, APR 1990.

[15] E. Stefanovich and T. Truong, "Optimized atomic radii for quantum dielectric continuum solvation models," *Chemical physics letters*, vol. 244, no. 1-2, pp. 65–74, 1995.

[16] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, "Electrostatics of nanosystems: application to microtubules and the ribosome.," *Proc Natl Acad Sci U S A*, vol. 98, no. 18, pp. 10037–10041, 2001 Aug 28.

[17] L. Martínez, R. Andrade, E. Birgin, and J. Martínez, "Packmol: A package for building initial configurations for molecular dynamics simulations," *Journal of computational chemistry*, vol. 30, no. 13, pp. 2157–2164, 2009.

[18] W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *Journal of Molecular Graphics*, vol. 14, pp. 33–38, 1996.

[19] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, *et al.*, "Nwchem: a comprehensive and scalable open-source solution for large scale molecular simulations," *Computer Physics Communications*, vol. 181, no. 9, pp. 1477–1489, 2010.

[20] P. Wang and A. Anderko, "Computation of dielectric constants of solvent mixtures and electrolyte solutions," *Fluid phase equilibria*, vol. 186, no. 1, pp. 103–122, 2001.

[21] W. Li, J. Zhang, J. Wang, and W. Wang, "Metal-coupled folding of cys2his2 zinc-finger," *Journal of the American Chemical Society*, vol. 130, no. 3, pp. 892–900, 2008.

[22] Y. Pang, "Successful molecular dynamics simulation of two zinc complexes bridged by a hydroxide in phosphotriesterase using the cationic dummy atom method," *Proteins: Structure, Function, and Bioinformatics*, vol. 45, no. 3, pp. 183–189, 2001.

A Sample Input Files

A.1 Gaussian Input Files

A.1.1 Geometry Optimization

%chk=ZnMeOH_1MeOH.2.chk

%nproc=8

%mem=8GB

B3LYP/6−31+G(d,p) Opt=(MaxCycles=3000) SCF=Tight

ZnMeOH_1MeOH initial optimization

2 1

Zn

O	1	2.12893				
H	2	0.96755	1	126.17828		
H	2	0.96757	1	125.75120	3	170.40655
O	1	2.13616	2	87.12158	3	94.83202
H	5	0.96761	1	125.02553	2	359.41570
H	5	0.96745	1	127.44986	2	180.60039
O	3	4.91606	1	7.01409	2	166.75975
H	8	0.96777	1	125.97754	2	273.69230
H	8	0.96776	1	126.28971	2	84.80468
C	1	3.27698	2	77.67952	3	274.51543
O	11	1.46857	1	27.07727	2	179.43446
H	12	0.96628	11	109.55118	1	179.32941
H	11	1.08519	1	79.64376	2	0.77357
H	11	1.08816	1	120.94433	2	108.22552
H	11	1.08818	1	119.82954	2	253.25482
H	1	2.81424	2	88.72717	3	167.08164
O	17	0.96758	1	37.82510	2	92.74253
H	18	0.96749	17	107.47534	1	178.20228

```
O        1          2.13326      2         89.99628      3          5.86169
H       20          0.96751      1        126.00081      2        265.86735
H       20          0.96756      1        126.52056      2         85.96788

%chk=ZnMeOH_1MeOH.2.chk
%nproc=8
%mem=8GB
# B3LYP/6-311+G(d,p) Opt=(MaxCycles=3000) SCF=Tight Geom=check guess=read

ZnMeOH_1MeOH opt1 optimization

2 1

%chk=ZnMeOH_1MeOH.2.chk
%nproc=8
%mem=8GB
# B3LYP/6-311++G(2d,p) Opt=(VeryTight,MaxCycles=3000) SCF=Tight Geom=check guess=read

ZnMeOH_1MeOH opt2 optimization

2 1
```

A.1.2 Frequency and Population Analysis

```
%chk=ZnMeOH_1MeOH.2.chk
%nproc=8
%mem=8GB
# Freq Pop=(ChelpG, ReadRadii) b3lyp/6-311++g(2d,p) geom=checkpoint guess=read

ZnMeOH_1MeOH frequency and population analysis

2 1

Zn 2.2
C 2.096
O 1.576
H 1.172
```

A.2 NAMD Parameters

A.2.1 CHARMM27

Standard CHARMM27 parameters were used, with the CYM residue using the following parameterization:

```
RESI CYM          -1.00
! Thiolate form
! Foloppe, N., J. Sagemark, K. Nordstrand, K.D. Berndt, and L. Nilsson
! (2001). J. Mol. Biol. 310:449-470.
! Atom types and charges transfered from methythiolate
GROUP
ATOM N    NH1    -0.47  !      |
ATOM HN   H       0.31  !  HN-N
ATOM CA   CT1     0.07  !      |   HB1
ATOM HA   HB      0.09  !      |   |     -
GROUP                   !  HA-CA--CB--SG    (thiolate)
ATOM CB   CS      -0.38 !      |   |

ATOM HB1  HA      0.09  !      |   HB2
ATOM HB2  HA      0.09  !   O=C
ATOM SG   SS      -0.80 !      |
GROUP
ATOM C    C       0.51
ATOM O    O       -0.51
BOND CB CA   SG CB   N HN   N  CA
BOND O  C    C  CA   C +N   CA HA   CB HB1   CB HB2
IMPR N -C CA HN   C CA +N O
DONOR HN N
ACCEPTOR O C
IC -C   CA   *N   HN    1.3479 123.9300  180.0000 114.7700  0.9982
IC -C   N    CA   C     1.3479 123.9300  180.0000 105.8900  1.5202
IC N    CA   C    +N    1.4533 105.8900  180.0000 118.3000  1.3498
IC +N   CA   *C   O     1.3498 118.3000  180.0000 120.5900  1.2306
IC CA   C    +N   +CA   1.5202 118.3000  180.0000 124.5000  1.4548
IC N    C    *CA  CB    1.4533 105.8900  121.7900 111.9800  1.5584
IC N    C    *CA  HA    1.4533 105.8900 -116.3400 107.7100  1.0837
IC N    CA   CB   SG    1.4533 111.5600  180.0000 113.8700  1.8359
IC SG   CA   *CB  HB1   1.8359 113.8700  119.9100 107.2400  1.1134
IC SG   CA   *CB  HB2   1.8359 113.8700 -125.3200 109.8200  1.1124
```

A.2.2 AMBER FF09

Standard FF09 parameters were used.

A.2.3 OPLS-ua

The topology file

```
* DK's & HP's TOP
*
   22      1
MASS      1 HT      1.00800 H ! TIPS3P WATER HYDROGEN
MASS      2 OT     15.99940 O ! TIPS3P WATER OXYGEN
MASS      3 OH     15.99940 O ! Hydronium oxygen
MASS      4 HH      1.00800 H ! Hydronium hydrogen
MASS      5 HC      1.00800 H ! Hydrogen to carbon
MASS      6 HN      1.00800 H !
MASS      7 C      12.00000 C
MASS      8 N      14.00700 N
MASS      9 HQ      1.00800 H
MASS     10 OQ     16.00000 O
MASS     11 BE      9.00000 Be
MASS     12 HS      1.00800 H ! SPCE WATER HYDROGEN
MASS     13 OS     15.99940 O ! TIPS3P WATER OXYGEN
MASS     14 Kr     83.00000 Kr ! Check mass number
MASS     15 Ne     20.00000 Ne ! Check mass number
MASS     16 KG     83.00000 KG ! Krypton Guissani
MASS     17 NG     20.00000 NG ! Neon Guissani
MASS     18 HF      1.00800 H ! Formate hydrogen
MASS     19 CF     12.00000 C ! Formate Carbon
MASS     20 OF     15.99940 O ! Formate Oxygen
MASS     21 P      31.0000  P ! Phosphorus
MASS     22 CH4    16.0320  C ! Methane United atom check methane mass
MASS     23 CH3    15.0240  C ! Methyl
MASS     24 OA     15.99940 O ! Alcohol oxygen
MASS     25 HA     1.00800  H ! Alcohol hydrogen
MASS     26 CHE    15.0240  C ! Methyl in Ethane
MASS     27 CHK    15.0240  C ! Methyl KBFF for Methanol JPCB 109 2005 15080-15086
MASS     28 OAK    15.99940 O ! Alcohol oxygen KBFF
MASS     28 HAK    1.00800  H ! Alcohol hydrogen KBFF
MASS     29 NPC    12.0000  C ! Neopentane center carbon
```

```
MASS     30 NPX    15.99940 C ! Neopentane methyl groups
MASS    196 ZN    65.370000 ZN ! zinc (II) cation

AUTO ANGLES DIHE

RESI TIP3           0.000
GROUP
ATOM OH2   OT      -0.834
ATOM H1    HT       0.417
ATOM H2    HT       0.417
BOND OH2 H1 OH2 H2
ANGLE H1 OH2 H2
ACCEPTOR OH2
PATCHING FIRS NONE LAST NONE

RESI SPCE           0.000
GROUP
ATOM OH2   OS      -0.8476
ATOM H1    HS       0.4238
ATOM H2    HS       0.4238
BOND OH2 H1 OH2 H2 H1 H2
ANGLE H1 OH2 H2
ACCEPTOR OH2
PATCHING FIRS NONE LAST NONE

RESI  O2M          -1.000
GROUP
ATOM  OM1  OS      -0.5000
ATOM  OM2  OS      -0.5000
ACCEPTOR OM1
ACCEPTOR OM2
PATCHING FIRST NONE LAST NONE

RESI KRYP           0.000
GROUP
ATOM Kr     Kr      0.000
PATCHING FIRST NONE LAST NONE

RESI MET            0.000
GROUP
ATOM CH4   CH4    0.000
```

```
PATCHING FIRST NONE LAST NONE

RESI MEOH          0.000
GROUP
ATOM C     CH3     0.265
ATOM O     OA     -0.700
ATOM H     HA      0.435
BOND C O   O   H
ANGLE C O H
DONOR      H   O
ACCEPTOR   O   C
PATCHING FIRST NONE LAST NONE

RESI MOHK          0.000
GROUP
ATOM C     CHK     0.300
ATOM O     OAK    -0.820
ATOM H     HAK     0.52
BOND C O   O   H   C   H
ANGLE C O H
DONOR      H   O
ACCEPTOR   O   C
PATCHING FIRST NONE LAST NONE

RESI ETH           0.000
GROUP
ATOM C1    CHE     0.000
ATOM C2    CHE     0.000
BOND C1   C2
PATCHING FIRST NONE LAST NONE

RESI NEOP          0.000
GROUP
ATOM CC    NPC     0.000
ATOM C1    NPX     0.000
ATOM C2    NPX     0.000
ATOM C3    NPX     0.000
ATOM C4    NPX     0.000
BOND  CC C1   CC C2
BOND  CC C3   CC C4
ANGLE C1 CC C2   C2 CC C3
```

```
ANGLE C3 CC C4   C4 CC C1
ANGLE C1 CC C3   C2 CC C4
PATCHING FIRST NONE LAST NONE

RESI KRYG          0.000
GROUP
ATOM KG     KG     0.000
PATCHING FIRST NONE LAST NONE

RESI NEO           0.000
GROUP
ATOM Ne     Ne     0.000
PATCHING FIRST NONE LAST NONE

RESI NEG           0.000
GROUP
ATOM NG     NG     0.000
PATCHING FIRST NONE LAST NONE

RESI 3PIT          0.000
GROUP
ATOM OH2  OT       0.000
ATOM H1   HT       0.000
ATOM H2   HT       0.000
BOND OH2 H1 OH2 H2 H1 H2
ANGLE H1 OH2 H2
ACCEPTOR OH2
PATCHING FIRST NONE LAST NONE

RESI H3O           1.000
GROUP
ATOM OH2  OH      -0.521
ATOM H1   HH       0.507
ATOM H2   HH       0.507
ATOM H3   HH       0.507
! BOND OH2 H1  OH2  H2  OH2  H3
PATCHING FIRST NONE LAST NONE

RESI O3H           0.000
GROUP
ATOM OH2  OH       0.000
```

```
ATOM H1     HH        0.000
ATOM H2     HH        0.000
ATOM H3     HH        0.000
BOND OH2 H1   OH2   H2   OH2   H3
PATCHING FIRST NONE LAST NONE


RESI OH3              0.000
GROUP
ATOM O1        OS     0.000
ATOM O2        OS     0.000
ATOM O3        OS     0.000
ATOM O4        OS     0.000
ATOM H5        HS     0.000
ATOM H6        HS     0.000
ATOM H7        HS     0.000
ATOM H8        HS     0.000
ATOM H9        HS     0.000
ATOM H10       HS     0.000
ATOM H11       HS     0.000
PATCHING FIRST NONE LAST NONE

RESI H5O2             1.000
GROUP
ATOM H1     HQ        0.505
ATOM O2     OQ       -0.734
ATOM O3     OQ       -0.730
ATOM H4     HQ        0.474
ATOM H5     HQ        0.505
ATOM H6     HQ        0.475
ATOM H7     HQ        0.505
! BOND H1 O2   H1 O3
! BOND O2 H6   O2 H7
! BOND O3 H4   O3 H5
PATCHING FIRST NONE LAST NONE

RESI 2O5H             0.000
GROUP
ATOM H1     HQ        0.000
ATOM O2     OQ        0.000
ATOM O3     OQ        0.000
```

```
ATOM H4    HQ       0.000
ATOM H5    HQ       0.000
ATOM H6    HQ       0.000
ATOM H7    HQ       0.000
! BOND H1 O2   H1 O3
! BOND O2 H6   O2 H7
! BOND O3 H4   O3 H5
PATCHING FIRST NONE LAST NONE

RESI IMD          0.000
GROUP
ATOM N1   N      -0.090
ATOM C2   C       0.232
ATOM N3   N      -0.716
ATOM C4   C       0.217
ATOM C5   C      -0.375
ATOM H1   HN      0.318
ATOM H2   HC      0.102
ATOM H4   HC      0.082
ATOM H5   HC      0.230
PATCHING FIRST NONE LAST NONE

RESI DMI          0.000
GROUP
ATOM N1   N       0.000
ATOM C2   C       0.000
ATOM N3   N       0.000
ATOM C4   C       0.000
ATOM C5   C       0.000
ATOM H1   HN      0.000
ATOM H2   HC      0.000
ATOM H4   HC      0.000
ATOM H5   HC      0.000
PATCHING FIRST NONE LAST NONE

RESI IMP          1.000
GROUP
ATOM N1   N      -0.115
ATOM C2   C       0.011
ATOM N3   N      -0.123
ATOM C4   C      -0.140
```

```
ATOM C5   C      -0.122
ATOM H1   HN      0.399
ATOM H2   HC      0.230
ATOM H3   HN      0.403
ATOM H4   HC      0.232
ATOM H5   HC      0.225
PATCHING FIRST NONE LAST NONE


RESI PMI          0.000
GROUP
ATOM N1   N       0.000
ATOM C2   C       0.000
ATOM N3   N       0.000
ATOM C4   C       0.000
ATOM C5   C       0.000
ATOM H1   HN      0.000
ATOM H2   HC      0.000
ATOM H3   HN      0.000
ATOM H4   HC      0.000
ATOM H5   HC      0.000
PATCHING FIRST NONE LAST NONE

RESI BER          2.000
GROUP
ATOM BE1  BE      2.000
PATCHING FIRST NONE LAST NONE

RESI BE4          2.000
GROUP
ATOM BE1  BE      1.664
ATOM O2   OQ     -1.093
ATOM O3   OQ     -1.093
ATOM O4   OQ     -1.097
ATOM O5   OQ     -1.097
ATOM H6   HQ      0.589
ATOM H7   HQ      0.591
ATOM H8   HQ      0.590
ATOM H9   HQ      0.588
ATOM H10  HQ      0.589
ATOM H11  HQ      0.591
```

```
ATOM H12   HQ       0.590
ATOM H13   HQ       0.588
PATCHING FIRST NONE LAST NONE

RESI   FOR          -1.000
ATOM   H1   HF      -0.100
ATOM   C2   CF       0.620
ATOM   O3   OF      -0.760
ATOM   O4   OF      -0.760
PATCHING FIRST NONE LAST NONE

RESI BEO            0.000
GROUP
ATOM   O1   OF      -0.659097
ATOM   C1   CF       0.498766
ATOM   O2   OF      -0.393698
ATOM   C2   CF       0.484823
ATOM   O3   OF      -0.390139
ATOM   O4   OF      -0.655186
ATOM   BE1 BE        1.114531
PATCHING FIRST NONE LAST NONE


RESI BEP            1.00000
GROUP
ATOM   BE1   BE      1.70000
ATOM   P     P       1.90000
ATOM   O1    OF     -0.70700
ATOM   O2    OF     -1.30000
ATOM   O3    OF     -0.64400
ATOM   O4    OF     -0.70700
ATOM   H5    HF      0.37900
ATOM   H6    HF      0.37900
PATCHING FIRST NONE LAST NONE

RESI PHO            -1.00000
GROUP
ATOM   P     P       1.36617
ATOM   O1    OF     -0.78008
ATOM   O2    OF     -0.87714
ATOM   O3    OF     -0.70454
```

```
ATOM  O4   OF     -0.70712
ATOM  H5   HF      0.3526
ATOM  H6   HF      0.35011
PATCHING FIRST NONE LAST NONE

RESI ZN2        2.00 ! Zinc ion, Roland Stote
GROUP
ATOM ZN   ZN   2.00
PATCHING FIRST NONE LAST NONE

END
```

The parameter file

```
* DK's and HP's param file
*

BONDS

HS   HS      0.000     1.6323  ! SPCE HS-HS distance
HT   HT      0.000     1.5139  ! Required for shake
OT   HT    450.000     0.9572  !
OS   HS    450.000     1.0000  ! SPCE OS-HS distance is different
OQ   HQ    450.000     0.9572  ! Zundel dummy values
OS   OS    450.000     1.3100  ! For superoxide ... largely dummy
OH   HH    450.000     0.9700
N    C     200.000     1.3600
N    HN    450.000     1.0000
C    HC    400.000     1.0800
C    C     200.000     1.3600
OA   HA    553.000     0.9600  ! JPCB 111 2007 4467-4476 &
CH3  OA    386.000     1.4250  ! JACS 106 1984 765-784
CH3  HA      0.000     1.9550  ! Required for shake
CHE  CHE     0.000     1.5300  ! JACS 106 1984 6638-6646
OAK  HAK   553.000     0.9450  ! JPCB 109 2005 15080-15086
CHK  OAK   386.000     1.4300  ! JACS 106 1984 765-784
CHK  HAK     0.000     1.9480  ! For shake
NPC  NPX   268.000     1.5300  ! Supplementary info JACS 118 Page 11225

ANGLES
HT   OT   HT     55.000   104.5200
HS   OS   HS     55.000   109.4000 ! SPCE uses tetrahedral
```

```
HH   OH   HH      55.000    114.0000
HQ   OQ   HQ       0.000      0.0000  ! Zundel dummy values
OQ   HQ   OQ       0.000      0.0000  ! Zundel dummy values
CH3  OA   HA      55.000    108.5000
CHK  OAK  HAK     55.000    108.5000  ! same as OPLS
NPX  NPC  NPX     58.35     109.4700  ! Fixing at tetrahedral angle

DIHEDRAL
OQ   HQ   OQ    HQ   0.00  0.000     ! Zundel dummy

NONBONDED nbxmod  5 atom cdiel shift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5

OT      0.000000  -0.152100     1.768200
HT      0.000000  -0.046000     0.224500
OS      0.000000  -0.155394     1.776600   ! SPCE Oxygen
HS      0.000000   0.000000     0.000000   ! SPCE hydrogen has no sigma
OQ      0.000000  -0.152100     1.768200
HQ      0.000000  -0.046000     0.224500
OH      0.000000  -0.152100     1.768200
HH      0.000000  -0.046000     0.224500
N       0.000000  -0.017000     1.824000
C       0.000000  -0.086000     1.908000
HN      0.000000  -0.015700     0.600000
HC      0.000000  -0.015000     1.359000
BE      0.000000  -0.018680     1.143900  ! Using Aqvist Li+ parameters
KG      0.000000  -0.335800     2.062500  ! Krypton from Guissani/Straatsma
Kr      0.000000  -1.500000     1.829600  ! Krypton fit to MP2 results
Ne      0.000000  -0.300000     1.433900  ! Neon fit to MP2 results
NG      0.000000  -0.036880     1.703300  ! Neon from Guissani/Straatsma
CF      0.000000  -0.070000     2.000000  ! Formate carbon
OF      0.000000  -0.120000     1.700000  ! Formate oxygen
HF      0.000000  -0.046000     0.224500  ! Formate hydrogen
P       0.000000  -0.585000     2.150000  ! Phosphorus atom
CH4     0.000000  -0.294000     2.093390  ! Jorgensen methane
CH3     0.000000  -0.207000     2.118650  ! Methonal (methyl) JPC 90 1986 1276-1284
OA      0.000000  -0.170000     1.722980  ! Alcohol oxygen JPC 90 1986 1276-1284
HA      0.000000   0.000000     0.000000  ! Alcohol hydrogen
CHE     0.000000  -0.207000     2.118650  ! Ethane (methyl) JACS 106 1984 6638-6646
CHK     0.000000  -0.207270     2.103490  ! Methonal (methyl) JPC 90 1986 1276-1284
OAK     0.000000  -0.155500     1.791450  ! Alcohol oxygen JPC 90 1986 1276-1284
```

```
HAK     0.000000   -0.021030      0.886750  ! Alcohol hydrogen
NPX     0.000000   -0.145000      2.222470  ! JACS 106 1984 6638-6646
NPC     0.000000   -0.050000      2.132670  ! JACS 106 1984 6638-6646

ZN      0.000000   -0.250000      1.090000 ! ALLOW  ION
                ! RHS March 18, 1990
END
```

A.3 NAMD Input files

A.3.1 Ion and Solvent MD

```
#############################################################
## JOB DESCRIPTION                                         ##
#############################################################

# Running simulation for ZnMeOH with 5 mol% MeOH

#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################

set simname         MeOH_15

structure           $simname.psf
coordinates         $simname.pdb
set temperature     298
set outputname      $simname

firsttimestep       0


#############################################################
## SIMULATION PARAMETERS                                   ##
#############################################################

# Input
paraTypeCharmm      on
parameters          par.inp
temperature         $temperature
```

```
# Force-Field Parameters
exclude             scaled1-4
1-4scaling          1.0
cutoff              12
switching           on
switchdist          10
pairlistdist        13.5


# Integrator Parameters
timestep            2.0  ;# 1fs/step
rigidBonds          none  ;# needed for 2fs steps
nonbondedFreq       1
fullElectFrequency  2
stepspercycle       10


# Constant Temperature Control
langevin            on    ;# do langevin dynamics
langevinDamping     5     ;# damping coefficient (gamma) of 5/ps
langevinTemp            $temperature
langevinHydrogen    off    ;# don't couple langevin bath to hydrogens

# Periodic Boundary Conditions
cellBasisVector1    35    0    0
cellBasisVector2     0   35    0
cellBasisVector3     0    0   35
cellOrigin          12   12   12

wrapAll             on

# Fixed Atoms
#fixedAtoms on
#fixedAtomsFile $simname.pdb
#fixedAtomsCol B
# PME (for full-system periodic electrostatics)
PME                 yes
PMEGridSizeX        64
PMEGridSizeY        64
PMEGridSizeZ        64
```

```
# Constant Pressure Control (variable volume)
useGroupPressure      yes ;# needed for rigidBonds
useFlexibleCell       no
useConstantArea           no

langevinPiston        on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  100
langevinPistonDecay    50
langevinPistonTemp    $temperature

# Output
outputName          $outputname
restartfreq         1000     ;# 1000steps = every 1ps
dcdfreq              125
xstFreq             1000
outputEnergies       500
outputPressure       500


#############################################################
## EXTRA PARAMETERS                                        ##
#############################################################


#############################################################
## EXECUTION SCRIPT                                        ##
#############################################################
minimize 7500
run 10200000
```

A.3.2 Protein MD

```
#############################################################
## JOB DESCRIPTION                                         ##
#############################################################

# Minimization, Equilibration, and measuring pair interactions between


#############################################################
## ADJUSTABLE PARAMETERS                                   ##
```

```
#############################################################

set simname                zinc_finger

amber                      yes
parmfile                   berg_dz.top
coordinates          berg_dz.pdb
bincoordinates       berg_dz.coor
binvelocities        berg_dz.vel

set temperature      298
set outputname       berg_dz

firsttimestep        0


#####################################################################
## SIMULATION PARAMETERS                                           ##
#####################################################################

# Force-Field Parameters
exclude              scaled1-4
1-4scaling           1.0
cutoff               12
switching            on
switchdist           10
pairlistdist         13.5


# Integrator Parameters
timestep             1.0  ;# 1fs/step
rigidBonds           all  ;# needed for 2fs steps
nonbondedFreq        1
fullElectFrequency   2
stepspercycle        10


# Constant Temperature Control
langevin             on    ;# do langevin dynamics
langevinDamping      5     ;# damping coefficient (gamma) of 5/ps
langevinTemp         $temperature
```

```
langevinHydrogen    off    ;# don't couple langevin bath to hydrogens


# Periodic Boundary Conditions
cellBasisVector1    47    0    0
cellBasisVector2     0   53    0
cellBasisVector3     0    0   47
cellOrigin           0    0    0

wrapAll              on


# PME (for full-system periodic electrostatics)
PME                  yes
PMEGridSizeX         64
PMEGridSizeY         64
PMEGridSizeZ         64


# Constant Pressure Control (variable volume)
useGroupPressure      yes ;# needed for rigidBonds
useFlexibleCell       no
useConstantArea           no

langevinPiston        on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  200
langevinPistonDecay   100
langevinPistonTemp    $temperature

# Output
outputName           $outputname

restartfreq          500     ;# 500steps = every 1ps
dcdfreq              500
xstFreq              500
outputEnergies        10
outputPressure        10

#############################################################
## EXTRA PARAMETERS                                        ##
```

```
#####################################################################


#####################################################################
## EXECUTION SCRIPT                                                 ##
#####################################################################
minimize 10000
for { set TEMP 50} { $TEMP < $temperature } {incr TEMP 50 } {
        reinitvels $TEMP
        langevinPistonTemp $TEMP
        langevinTemp $TEMP
        run 200
}
reinitvels $temperature
langevinPistonTemp $temperature
langevinTemp $temperature
run 5000000;
```

A.3.3 Trajectory Analysis

```
#####################################################################
## JOB DESCRIPTION                                                  ##
#####################################################################

# Calculating Interaction Energies between zinc and water


#####################################################################
## ADJUSTABLE PARAMETERS                                            ##
#####################################################################

set simname         MeOH_15

structure           $simname.psf
coordinates         $simname.pdb
extendedSystem      $simname.xsc

set temperature     298
set outputname      $simname.pair

firstTimestep       0
pairInteraction           on
```

```
pairInteractionGroup1  0
pairInteractionFile    $simname.meth.pdb
pairInteractionGroup2  1
pairInteractionCol         B
#####################################################################
## SIMULATION PARAMETERS                                           ##
#####################################################################

# Input
paraTypeCharmm       on
parameters           par.inp
temperature          $temperature


# Force-Field Parameters
exclude              scaled1-4
1-4scaling           1.0
cutoff               12
switching            on
switchdist           10
pairlistdist         13.5
COMmotion                        yes

# Periodic Boundary Conditions
cellBasisVector1     36    0    0
cellBasisVector2      0   36    0
cellBasisVector3      0    0   36
cellOrigin            0    0    0

wrapAll              on

# PME (for full-system periodic electrostatics)
PME                  yes
PMEGridSizeX         64
PMEGridSizeY         64
PMEGridSizeZ         64

# Output
outputName           $outputname

#####################################################################
```

```
## EXTRA PARAMETERS                                                ##
#####################################################################


#####################################################################
## EXECUTION SCRIPT                                                ##
#####################################################################

set ts 0
coorfile open dcd $simname.dcd
while { ![coorfile read] } {
   firstTimestep $ts
   run 0
   incr ts 1
}
coorfile close
```

A.4 NWCHEM Input files

A.4.1 Preparation

```
#Creates files for optimization of the berg zinc fingerG

start berg
prepare
 system berg
 source berg_orig.pdb
 new_top new_rst
 center
 orient
 solvate 5.0
 modify atom 202:ZN quantum
 modify atom 45:_CB quantum
 modify atom 45:3HB quantum
 modify atom 45:2HB quantum
 modify atom 45:_SG quantum
 modify atom 50:_CB quantum
 modify atom 50:3HB quantum
 modify atom 50:2HB quantum
 modify atom 50:_SG quantum
 modify atom 63:_NE2 quantum
```

```
 modify atom 63:_CG quantum
 modify atom 63:_ND1 quantum
 modify atom 63:_CE1 quantum
 modify atom 63:_CD2 quantum
 modify atom 63:_HD2 quantum
 modify atom 63:_HE1 quantum
 modify atom 63:_HD1 quantum
 modify atom 67:_NE2 quantum
 modify atom 67:_CG quantum
 modify atom 67:_ND1 quantum
 modify atom 67:_CE1 quantum
 modify atom 67:_CD2 quantum
 modify atom 67:_HD2 quantum
 modify atom 67:_HE1 quantum
 modify atom 67:_HD1 quantum
 update lists
 ignore
 write berg.rst
 write berg.pdb
end

task prepare
```

A.4.2 Optimization

```
memory total 8192 mb
scratch_dir scratch
start berg

#  LANL2DZ ECP  EMSL  Basis Set Exchange Library  2/8/12 1:45 PM
# Elements                             References
# --------                             ----------
# H  - Ne: T. H. Dunning Jr. and P. J. Hay, in Methods of Electronic Structure
# Theory, Vol. 2, H. F. Schaefer III, ed., PLENUM PRESS (1977)
# Na - Hg: P. J. Hay and W. R. Wadt, J. Chem. Phys. 82, 270 (1985).
# P. J. Hay and W. R. Wadt, J. Chem. Phys. 82, 284 (1985).
# P. J. Hay and W. R. Wadt, J. Chem. Phys. 82, 299 (1985).
#

BASIS "ao basis" PRINT
#BASIS SET: (4s) -> [2s]
H    S
```

```
      19.2384000                0.0328280
       2.8987000                0.2312040
       0.6535000                0.8172260
H     S
       0.1776000                1.0000000
#BASIS SET: (10s,5p) -> [3s,2p]
C     S
    4233.0000000                0.0012200
     634.9000000                0.0093420
     146.1000000                0.0454520
      42.5000000                0.1546570
      14.1900000                0.3588660
       5.1480000                0.4386320
       1.9670000                0.1459180
C     S
       5.1480000               -0.1683670
       0.4962000                1.0600910
C     S
       0.1533000                1.0000000
C     P
      18.1600000                0.0185390
       3.9860000                0.1154360
       1.1430000                0.3861880
       0.3594000                0.6401140
C     P
       0.1146000                1.0000000
#BASIS SET: (10s,5p) -> [3s,2p]
N     S
    5909.0000000                0.0011900
     887.5000000                0.0090990
     204.7000000                0.0441450
      59.8400000                0.1504640
      20.0000000                0.3567410
       7.1930000                0.4465330
       2.6860000                0.1456030
N     S
       7.1930000               -0.1604050
       0.7000000                1.0582150
N     S
       0.2133000                1.0000000
N     P
```

```
      26.7900000                0.0182540
       5.9560000                0.1164610
       1.7070000                0.3901780
       0.5314000                0.6371020
N     P
       0.1654000                1.0000000
#BASIS SET: (3s,3p) -> [2s,2p]
S     S
       1.8500000               -0.5324335
       0.4035000                1.2763801
S     S
       0.1438000                1.0000000
S     P
       4.9450000               -0.0608116
       0.4870000                1.0132686
S     P
       0.1379000                1.0000000
#BASIS SET: (3s,2p,5d) -> [2s,2p,2d]
Zn    S
       0.7997000               -0.6486112
       0.1752000                1.3138291
Zn    S
       0.0556000                1.0000000
Zn    P
       0.1202000                1.0000000
Zn    P
       0.0351000                1.0000000
Zn    D
      68.8500000                0.0258532
      18.3200000                0.1651195
       5.9220000                0.4468212
       1.9270000                0.5831080
Zn    D
       0.5528000                1.0000000
END
# Elements                              References
# --------                              ----------
# Na - Hg: P. J. Hay and W. R. Wadt, J. Chem. Phys. 82, 270 (1985).
#          P. J. Hay and W. R. Wadt, J. Chem. Phys. 82, 284 (1985).
#          P. J. Hay and W. R. Wadt, J. Chem. Phys. 82, 299 (1985).
#
```

```
ECP
S nelec 10
S ul
1     532.6685222            -10.0000000
2     108.1342248            -85.3593846
2      24.5697664            -30.4513290
2       7.3702438            -10.3745886
2       2.3712569             -0.9899295
S S
0     106.3176781              3.0000000
1     100.8245833             10.6284036
2      53.5858472            223.6360469
2      15.3706332             93.6460845
2       3.1778402             28.7609065
S P
0     101.9709185              5.0000000
1      93.2808973              6.0969842
2      65.1431772            285.4425500
2      24.6347440            147.1448413
2       7.8120535             53.6569778
2       2.3112730              8.9249559
Zn nelec 18
Zn ul
1     386.7379660            -18.0000000
2      72.8587359           -124.3527403
2      15.9066170            -30.6601822
2       4.3502340            -10.6358989
2       1.2842199             -0.7683623
Zn S
0      19.0867858              3.0000000
1       5.0231080             22.5234225
2       1.2701744             48.4465942
2       1.0671287            -44.5560119
2       0.9264190             12.9983958
Zn P
0      43.4927750              5.0000000
1      20.8692669             20.7435589
2      21.7118378             90.3027158
2       6.3616915             74.6610316
```

```
2       1.2291195              9.8894424
Zn D
2      13.5851800             -4.8490359
2       9.8373050              3.6913379
2       0.8373113             -0.5037319
END

dft
        xc b3lyp
end

md
        system berg
        noshake
end

qmmm
        bqzone 9
        region  qm      mm      solvent
        maxiter 500     10000   50000
        ncycles 10
        density static
        xyz berg.optim
end

task qmmm dft optimize
```

B Calculation Scripts

B.1 Radial Distribution Function

VMD code:

```
mol new PPP
mol addfile DDD type dcd first FFF waitfor all
set ion [atomselect top "resname ZN2"]
$ion num
set meth_o [atomselect top "name O and pbwithin 15 of resname ZN2"]
$meth_o num
set meth_c [atomselect top "name C and pbwithin 15 of resname ZN2"]
$meth_c num
```

```
set wat_o [atomselect top "name OH2 and pbwithin 15 of resname ZN2"]
$wat_o num
set frames [molinfo 0 get numframes]
set frames [expr "$frames - 1"]
puts $frames
set res_o [measure gofr $ion $meth_o delta 0.1 rmax 10 first 0 last $frames usepbc true selupdate t
set res_c [measure gofr $ion $meth_c delta 0.1 rmax 10 first 0 last $frames usepbc true selupdate t
set wat_o [measure gofr $ion $wat_o delta 0.1 rmax 10 first 0 last $frames usepbc true selupdate tru
set out [open "OOO" w]
puts $out $res_o
puts $out $res_c
puts $out $wat_o
close $out
```

MATLAB Code

```
%Reads a gofr file generated from VMD and
%plots the radial distribution function
function plot_gofr(filename)
A=importdata(sprintf('%s.gofr.fmt',filename), ' ');
plot(A(1,:), A(4,:), '-k', A(5,:), A(8,:), '--b', A(9,:), A(12,:), '-.g', 'LineWidth',5)
lh=legend('Zn-MeO', 'Zn-MeC', 'Zn-H2O')
set(lh, 'FontSize',30);
xlabel('r (Angstroms)','FontSize',30)
ylabel('g(r)','FontSize',30);
axis([0 6 0 3]);
set(gca,'FontSize',30);
print(sprintf('thesis_plots/%s_gofr.png', filename), '-dpng', '-r300');
end
```

B.2 Occupancy Probability

```
mol new PPP
mol addfile DDD type dcd first BBB waitfor all
set wat_o [atomselect top "name OH2 and (pbwithin OOO of resname ZN2) and (not pbwithin III of resna
set meth_o [atomselect top "name O and (pbwithin OOO of resname ZN2) and (not pbwithin III of resnam
$meth_o num
set meth_c [atomselect top "name C and (pbwithin OOO of resname ZN2) and (not pbwithin III of resnam
$meth_c num
for {set counter 0} {$counter < 30} {incr counter} {
        lappend wat_o_ctr 0
```

```
        lappend meth_o_ctr 0
        lappend meth_c_ctr 0
        lappend num $counter
}
set frames [molinfo 0 get numframes]
for {set counter 0} {$counter < $frames } {incr counter} {
        $wat_o frame $counter
        $meth_o frame $counter
        $meth_c frame $counter
        $wat_o update
        $meth_o update
        $meth_c update
        set wat_o_ind [$wat_o num]
        if {$wat_o_ind >= [llength $wat_o_ctr]} {set $wat_o_ind [expr [llength $wat_o_ctr] - 1]}
        set meth_o_ind [$meth_o num]
        if {$meth_o_ind >= [llength $meth_o_ctr]} {set $meth_o_ind [expr [llength $meth_o_ctr] - 1]
        set meth_c_ind [$meth_c num]
        if {$meth_c_ind >= [llength $meth_c_ctr]} {set $meth_c_ind [expr [llength $meth_c_ctr] - 1]



        set wat_o_ctr [lreplace $wat_o_ctr $wat_o_ind $wat_o_ind [expr [lindex $wat_o_ctr $wat_o_ind
        set meth_o_ctr [lreplace $meth_o_ctr $meth_o_ind $meth_o_ind [expr [lindex $meth_o_ctr $meth
        set meth_c_ctr [lreplace $meth_c_ctr [$meth_c num] [$meth_c num] [expr [lindex $meth_c_ctr
}
set out [open "WWW" w]
puts $out $num
puts $out $wat_o_ctr
puts $out $meth_o_ctr
puts $out $meth_c_ctr
close $out
```

B.3 MD Free energy

After using the pair interaction energy script (Appendix A.3.3), the nrg file is further processed

```
#!/usr/bin/perl -w
my($input, $title, $output)=@ARGV;
print "
A=importdata('$input');
allnrg=A(50:17500,:); \%interaction energy
```

```
intnrg=allnrg(:,6);
elec=allnrg(:,3);
vdw=allnrg(:,4);
mean(intnrg)
std(intnrg)
beta=1/0.593; \%in kcal/mol
G=mean(intnrg)+var(intnrg)/2*beta;
B=exp(intnrg * beta);
C=sum(B) ./ length(B); \%This is the sum over the probabilities
D=log(C)  / beta \%solving e^(b*mu) for mu
[E,F]=hist(intnrg, 150);
E=E/sum(E);
[thefit,thegof]=fit(F', E', 'gauss1')
semilogy(F,E,'ob')
hold on
semilogy(F, feval(thefit, F'));
xlabel('Pair interaction energy (kcal/mol)')
ylabel('Probability')
\%title(sprintf('Zn^{2+}/non-first-shell interaction energies, $title.  \\\\mu^{ex}_{calc}=%6.4f kc
axis([min(F)-10, max(F)+10, min(E),max(E) ]);
print('-djpeg', '$output.hist.jpg');
disp('\\n')
disp(sprintf('Approx: %6.4f Gaussian: %6.4f %6.4f %6.4f %6.4f %6.4f',D,G,mean(elec), var(elec), mea
```

C Results

C.1 Radial distribution functions

C.2 MD Interaction Energy Histograms

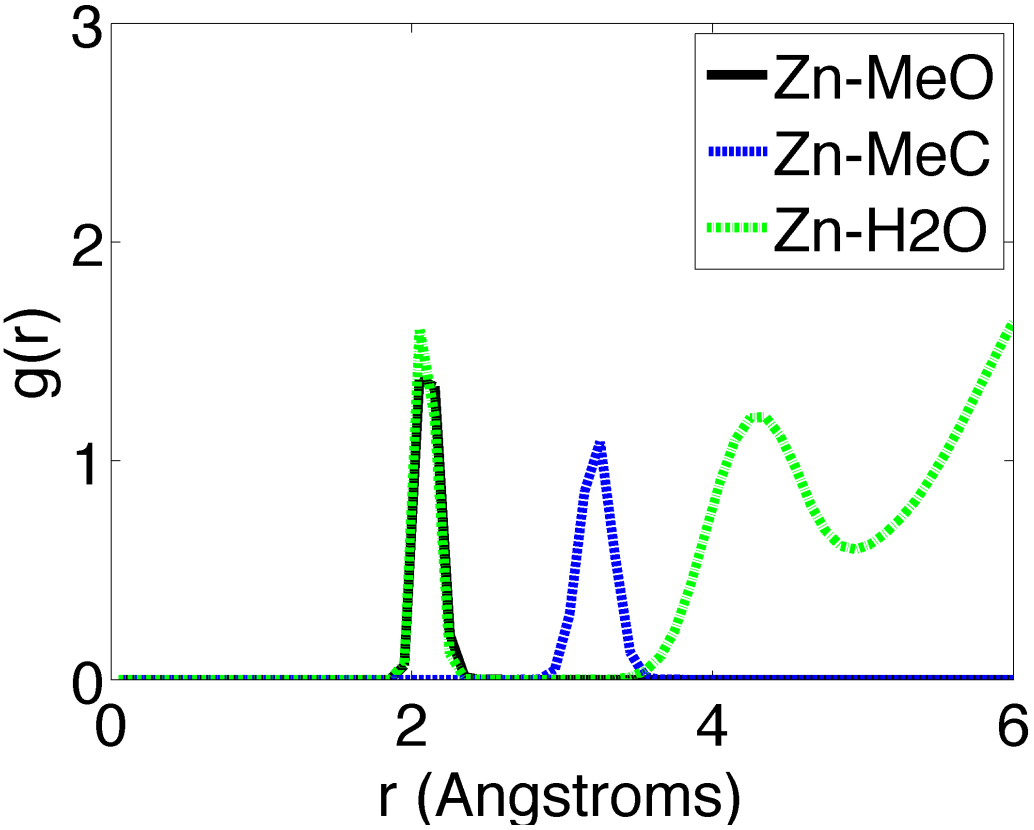


Figure 11: $Zn^{2+}\left[H_2O\right]_3\left[MeOH\right]_3$ in water

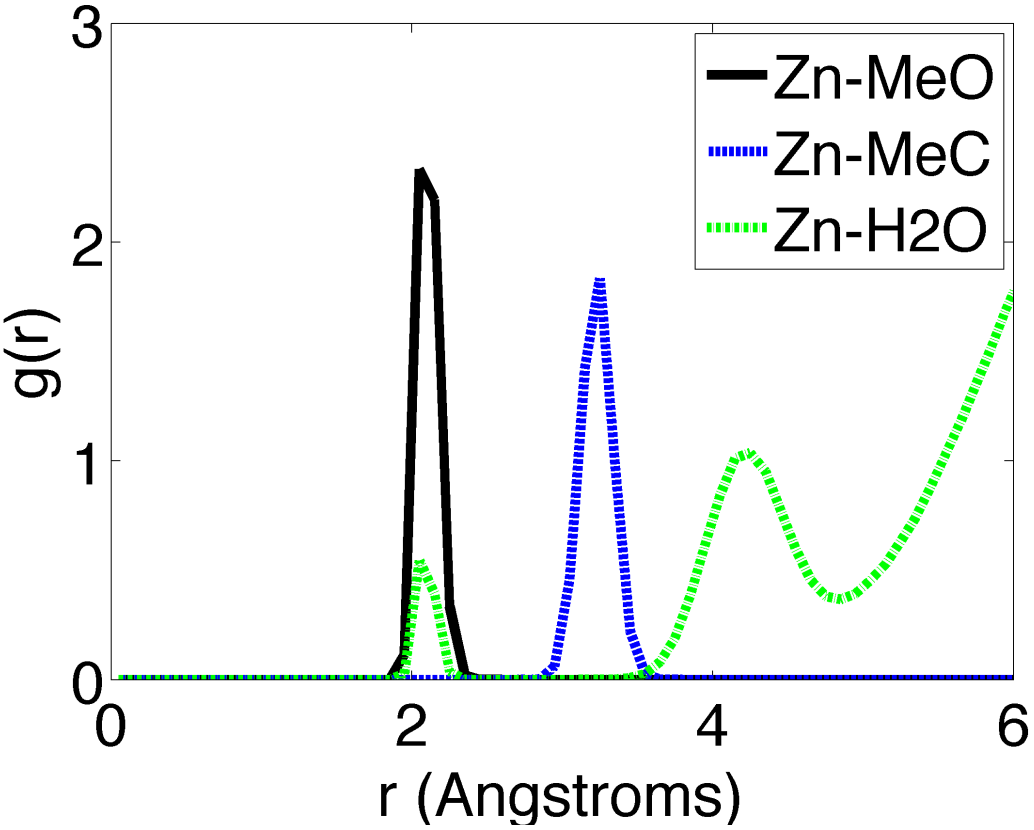


Figure 12: $Zn^{2+}\left[H_2O\right]_1\left[MeOH\right]_5$ in water

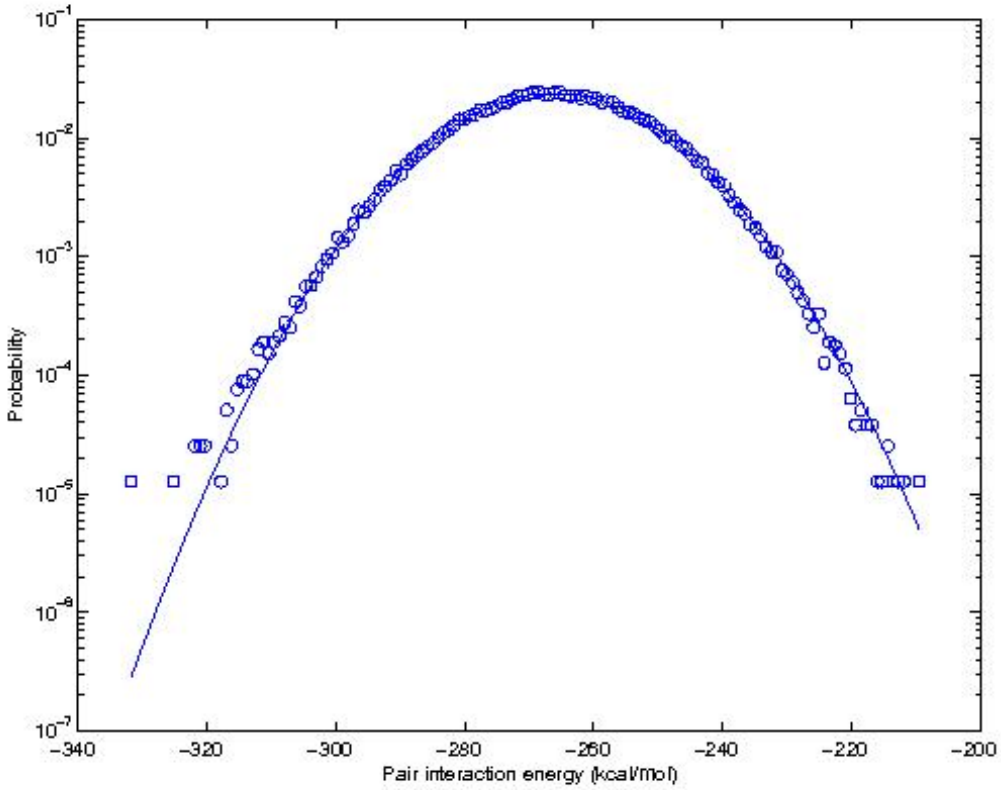


Figure 13: $Zn^{2+}\left[H_2O\right]_4\left[MeOH\right]_2$ in water

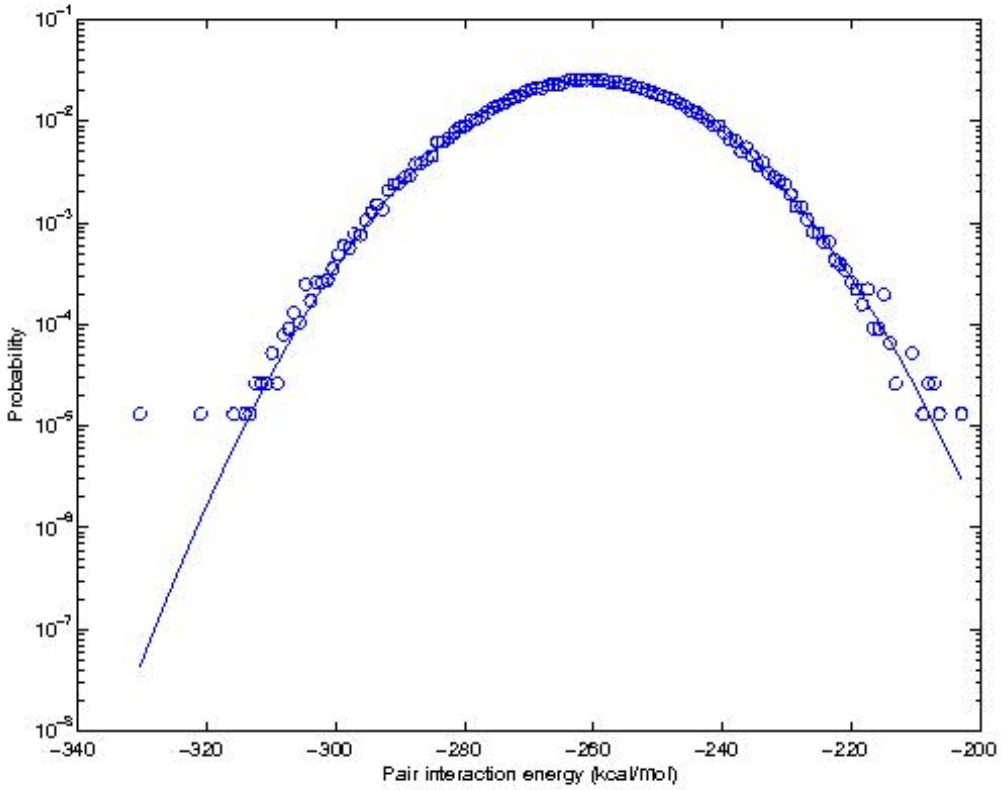


Figure 14: $Zn^{2+}\left[H_2O\right]_3\left[MeOH\right]_3$ in water

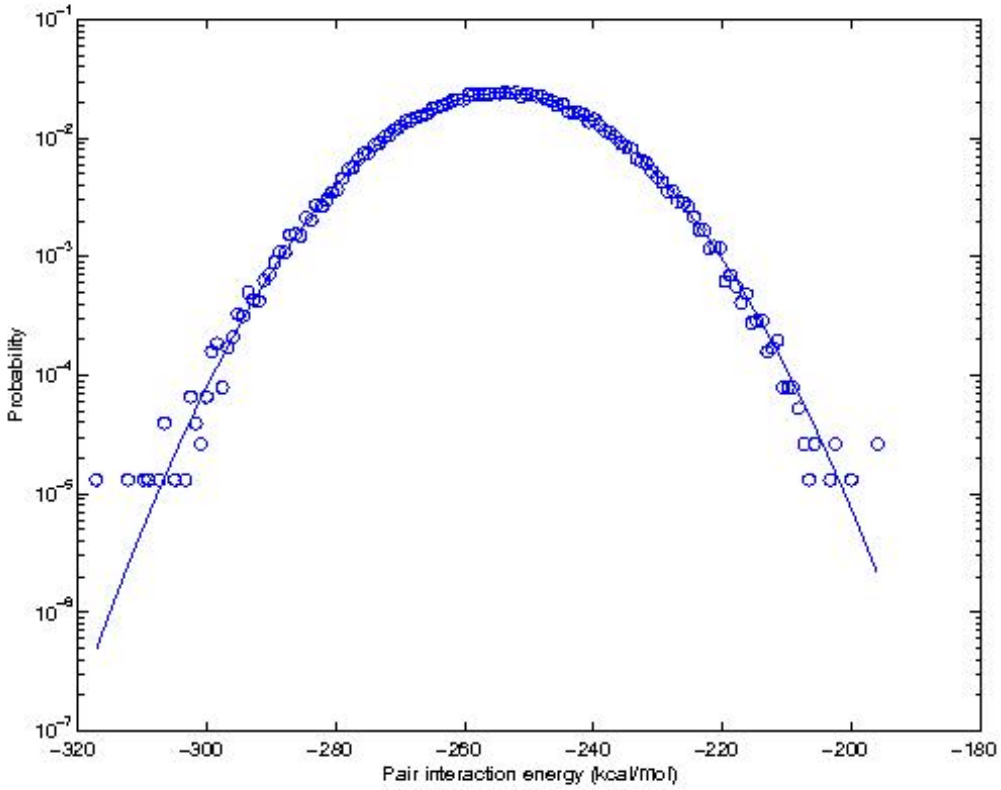


Figure 15: $Zn^{2+}\left[H_2O\right]_2\left[MeOH\right]_4$ in water

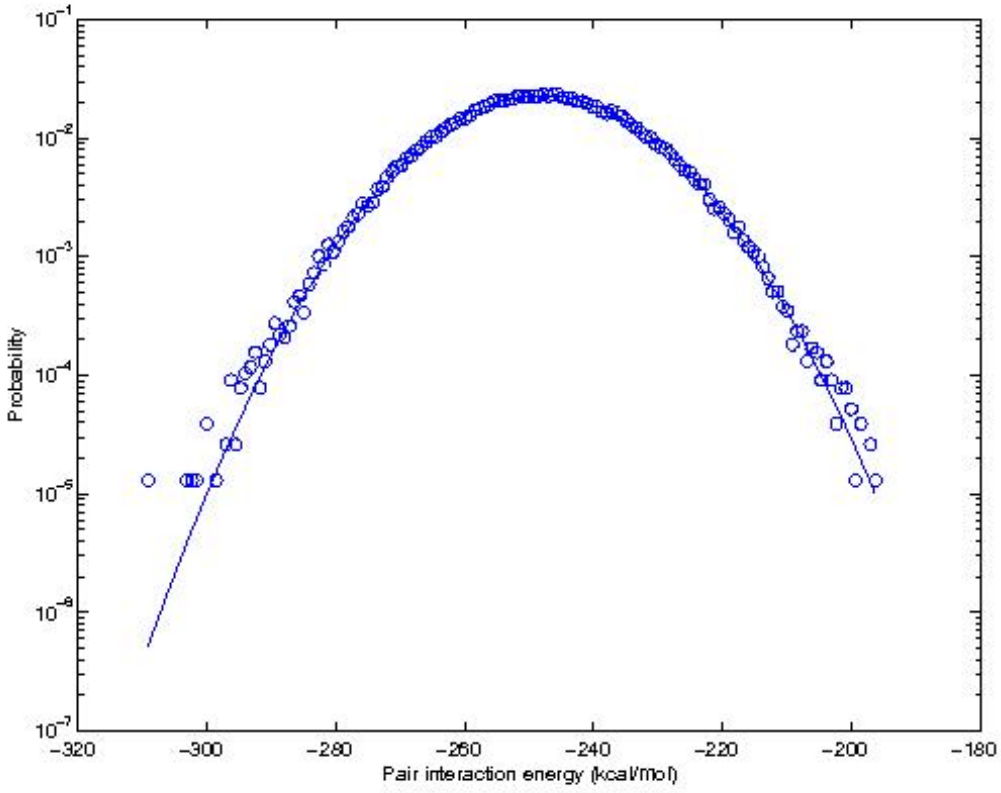


Figure 16: $Zn^{2+}\left[H_2O\right]_1\left[MeOH\right]_5$ in water

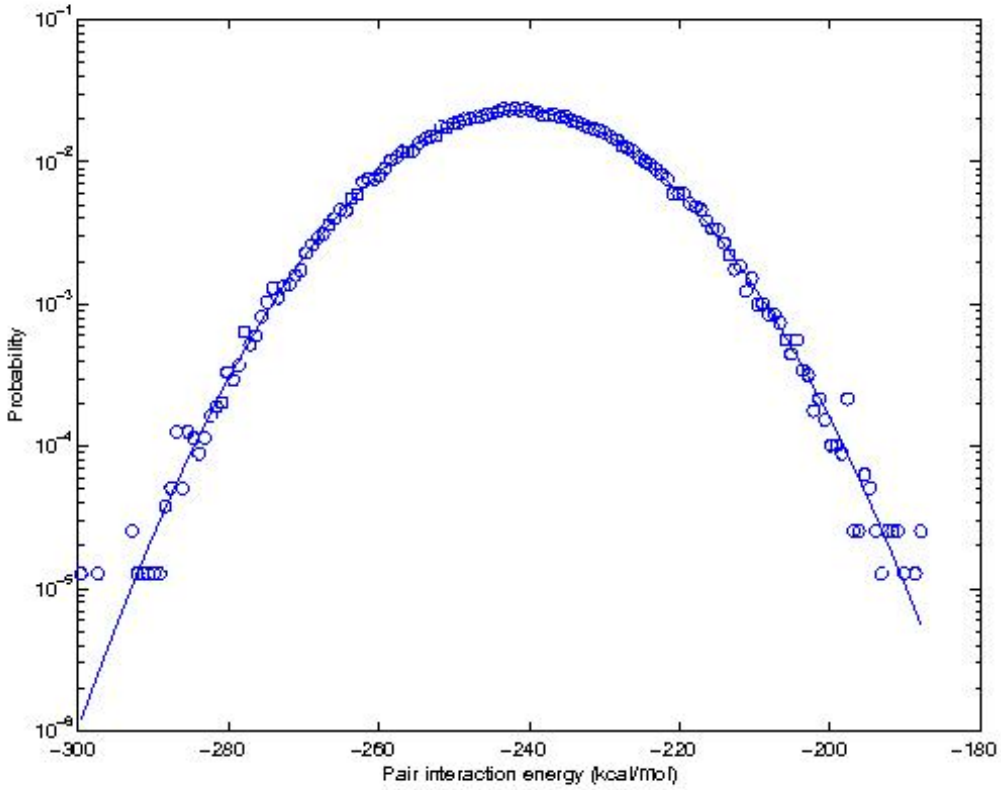


Figure 17: $Zn^{2+}\left[H_2O\right]_0\left[MeOH\right]_6$ in water

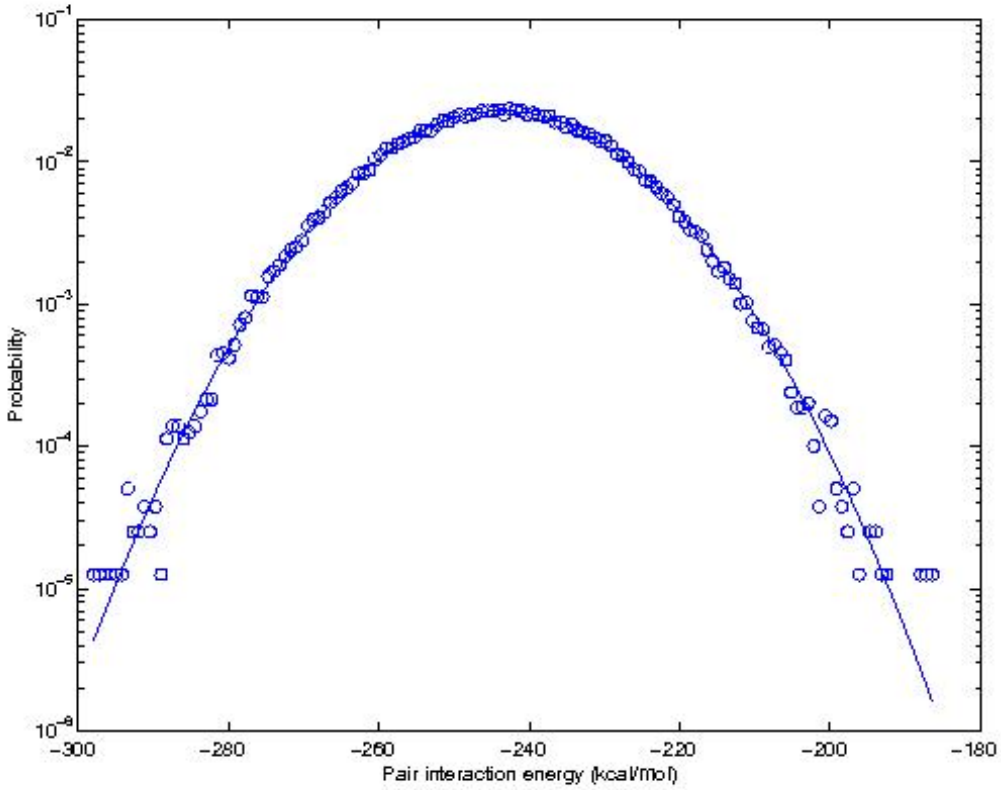


Figure 18: $Zn^{2+}\left[H_2O\right]_0\left[MeOH\right]_6$ in 5% MeOH and water

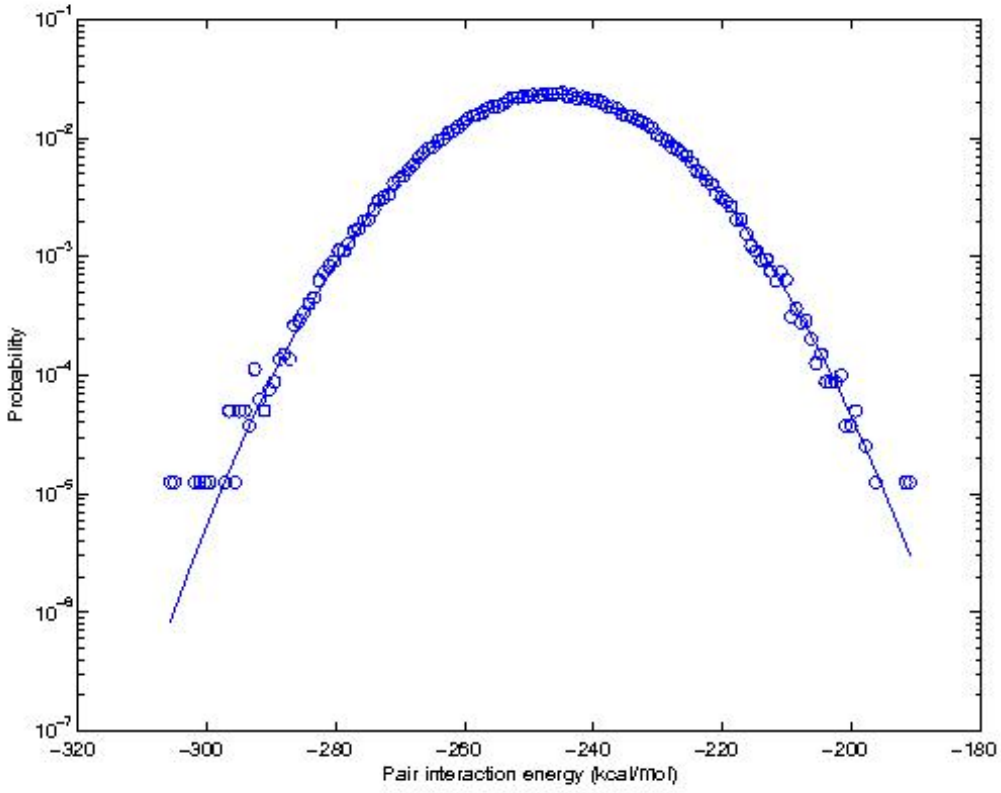


Figure 19: $Zn^{2+}\left[H_2O\right]_0\left[MeOH\right]_6$ in 10% MeOH and water

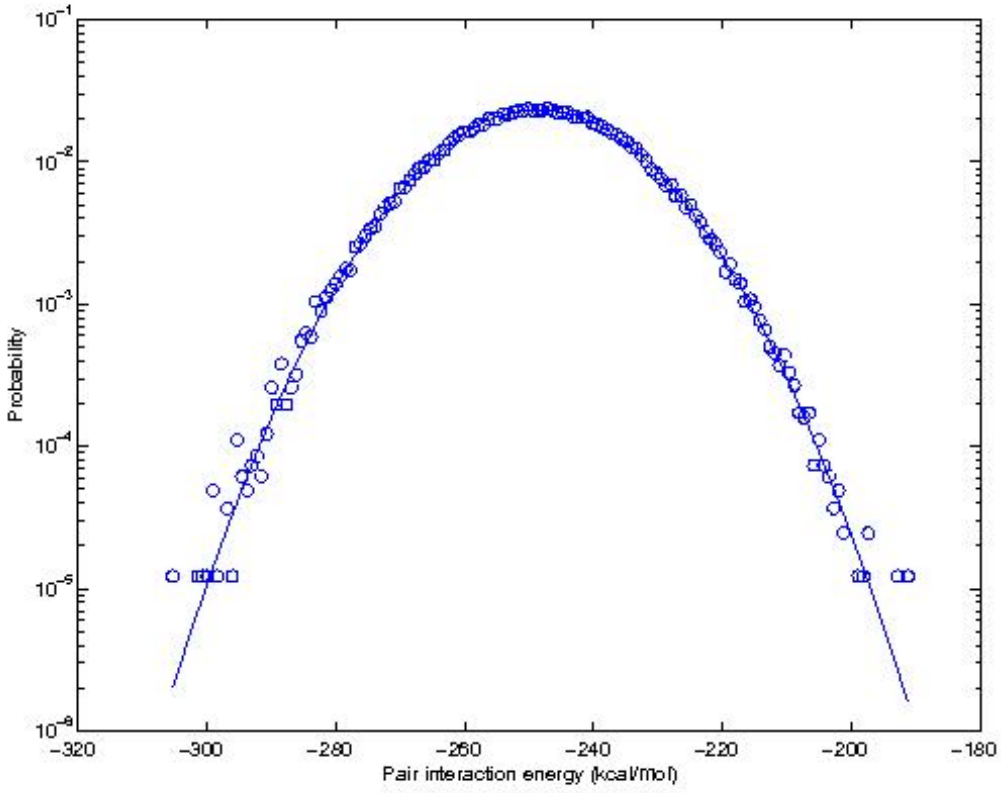


Figure 20: $Zn^{2+}\left[H_2O\right]_0\left[MeOH\right]_6$ in 15% MeOH and water