

ADVANCES IN METAL ION MODELING

By

Pengfei Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry—Doctor of Philosophy

2016

ABSTRACT

ADVANCES IN METAL ION MODELING

By

Pengfei Li

Metal ions play fundamental roles in geochemistry, biochemistry and materials science. With the tremendous increasing power of the computational resources and largely inventions of the computational tools, computational chemistry became a more and more important tool to study various chemical processes. Force field modeling strategy, which is built on physical background, offered a fast way to study chemical systems at atomic level. It could offer considerable accuracy when combined with the Monte Carlo or Molecular Dynamics simulation protocol. However, there are various metal ions and it is still challenging to model them using available force field models. Generally there are several models available for modeling metal ions using the force field approach such as the nonbonded model, the bonded model, the cationic dummy atom model, the combined model, and the polarizable models. Our work concentrated on the nonbonded and bonded models, which are widely used nowadays. Firstly, we focused on filling in the blanks of this field. We proposed a noble gas curve, which was used to describe the relationship between the van der Waals radius and well depth parameters in the 12-6 Lennard-Jones potential. By using the noble gas curve and multiple target values (the hydration free energy, ion-oxygen distance, coordination number values), we have consistently parameterized the 12-6 Lennard-Jones nonbonded model for 63 different ions (including 11 monovalent cations, 4 monovalent anions, 24 divalent cations, 18 trivalent cations, and 6 tetravalent cations) combined with three widely used water models (TIP3P, SPC/E,

and TIP4P_{EW}). Secondly, we found there is limited accuracy of the 12-6 model, which makes it hard to simulate different properties simultaneously for ions with formal charge equal or larger than +2. By considering the physical origins of the 12-6 model, we proposed a new nonbonded model, named the 12-6-4 LJ-type nonbonded model. We have systematically parameterized the 12-6-4 model for 55 different ions (including 11 monovalent cations, 4 monovalent anions, 16 divalent cations, 18 trivalent cations, and 6 tetravalent cations) in the three water models. It was shown that the 12-6-4 model could reproduce several properties at the same time, showing remarkable improvement over the 12-6 model. Meanwhile, through the usage of a proposed combining rule, the 12-6-4 model showed excellent transferability to mixed systems. Thirdly, we have developed the MCPB.py program to facilitate building of the bonded model for metal ion containing systems, which can largely reduce human efforts. Finally, an application case of a metallochaperone - CusF was shown, and based on the simulations we hypothesized an ion transfer mechanism.

*This dissertation is dedicated to my parents, who always love, trust,
and support me.*

ACKNOWLEDGEMENTS

I thank God for leading my life, loving me, and being my Lord. I thank my family especially my parents for their generous love. My parents have raised me and spared no effort to offer me a good life and excellent education. They have set excellent examples for me about hard work and good behavior. They created a loving home, always love me and are my powerful backing. I would like to thank the brothers and sisters in the Lansing Chinese Christian Church, who have shared the Gospel with me and helped me to know God. I learned a lot about Love and how to practice love in the Church, which really benefits my life. I also thank the Chinese community in the Greater Lansing area, which has enriched my life at Michigan State University (MSU).

I acknowledge my advisor Professor Kenneth M. Merz Jr. for his patient love and kind support of my research. I thank him for being a model theoretical chemist. His positive attitude, broad thoughts, and deep insights educated me greatly. I thank my guidance committee members: Professor Robert I. Cukier, Professor Benjamin G. Levine, and Professor James E. Jackson, and my guidance committee when I was in the University of Florida (UF): Professor Adrian E. Roitberg, Professor Samuel B. Trickey, Professor Benjamin W. Smith, and Professor Clifford R. Bowers, as well as my research advisor during my undergraduate, Professor Zexing Cao in Xiamen University, for their kind help and support during my research career. I thank the members of the Merz research group for their company. Being a member of the Merz research group is a precious memory for me. Among them, I specifically thank Ben for handing me the project of designing van

der Waals parameters for the zinc ion, which is the beginning of this research story. I would like to thank Dhruva for his guidance in the early stage of my PhD research, and for his kind help and suggestions during my PhD career. I thank Lin for his hard work concerning parameterization of the nonbonded model for monovalent and highly charged ions. I would like to thank Yipu for his help with computer programming, which helped me greatly in my start-up PhD research. I would also like to thank Mike and David for their help with the language of my manuscripts.

Besides the Merz group members, I thank Shuai Wang (UF) for his help of writing the analysis code for determining the ion-oxygen distance. I would like to thank the AMBER developers' team. It is an honor to be a part of the team and I really enjoyed the meetings, which broadened my horizons and offered me precious opportunities to talk with the experts in related fields. Among them I especially thank Professor David A. Case, Dr. Jason M. Swails, and Dr. Hai Nguyen for their generous help with computer programming in the AMBER software package. I thank the financial support from the United States National Institutes of Health (RO1's GM044974 and GM066859), and the computing support from the high performance computing centers at UF and MSU. I thank the kind support from the secretary of my advisor, Ms. Carey Byerrum and one of the graduate secretaries in the Department of Chemistry at MSU, Ms. Heidi Wardin, who are always kind and efficient. Finally I thank other teachers, colleagues, administrators, secretaries, and friends who have enriched my life and supported my research.

TABLE OF CONTENTS

LIST OF TABLES.....	ix
LIST OF FIGURES	xii
KEY TO ABBREVIATIONS	xvi
CHAPTER 1: INTRODUCTION.....	1
1.1 General Introduction of Models for Metal Ions.....	1
1.2 The 12-6 Nonbonded Model.....	3
1.3 The 12-6-4 Nonbonded Model	9
1.4 The Bonded Model and MCPB.py Program	11
1.5 The CusF Metallochaperone System.....	14
CHAPTER 2: PARAMETERIZATION OF THE 12-6 NONBONDED MODEL.....	15
2.1 Target Values.....	15
2.2 Methods	19
2.3 Validation Tests of Several Available LJ Parameters for M(II) Ions.....	31
2.4 Parameter Space Scanning.....	33
2.5 Parameter Determination.....	41
2.6 Parameter Assessment	47
2.7 Error Analysis.....	53
2.8 Conclusions	58
CHAPTER 3: PARAMETERIZATION OF THE 12-6-4 NONBONDED MODEL	62
3.1 Model Origin	62
3.2 Methods	64
3.3 Parameter Space Scanning and Parameter Determination	68
3.4 Discussion.....	72
3.5 Parameter Assessment.....	85
3.6 Design Your Own Parameters	91
3.7 Conclusions	93
CHAPTER 4: DEVELOPMENT OF THE MCPB.PY PROGRAM	96
4.1 Code Structure	96
4.2 Workflow	97
4.3 Usage	100
4.4 Examples	103
4.5 Conclusions	111
CHAPTER 5: APPLICATION CASE OF THE CUSF METALLOCHAPERONE SYSTEM	112

APPENDICES	116
APPENDIX A: TABLES	117
APPENDIX B: FIGURES	164
APPENDIX C: COPYRIGHT NOTICE	208
BIBLIOGRAPHY	209

LIST OF TABLES

Table 1. Target values of the HFE, IOD and the CN of the first solvation shell for monovalent ions.....	117
Table 2. Target values of the HFE, IOD and the CN of the first solvation shell for divalent metal ions.....	118
Table 3. Target values of the HFE, IOD and the CN of the first solvation shell for trivalent and tetravalent metal ions.....	119
Table 4. Parameters of four different water models.	120
Table 5. Absolute HFEs, IODs and CNs of M(II) ions employing parameters available in the AMBER package.	121
Table 6. Absolute HFEs, IODs and CNs of Zn(II) employing parameters developed by Stote and Karplus.....	122
Table 7. Absolute HFEs, relative HFEs (relative to the Cd^{2+} ion), IODs and CNs of M(II) ions with the parameters taken from Babu and Lim.....	123
Table 8. Standard deviations of the IOD values for the LJ grids of divalent metal ions.....	124
Table 9. The experimental LJ parameters for noble gas atoms. ^a	125
Table 10. Comparison of results of different scaling methods in the VDW-disappearing and VDW-appearing steps of the HFE calculations for the Be^{2+} and Ba^{2+} ions. ^a	126
Table 11. Final optimized HFE parameter set of divalent metal ions for different water models.....	127
Table 12. The simulated HFE, IOD and CN values of the HFE parameter set for divalent metal ions. ^a	128
Table 13. Final optimized IOD parameter set of divalent metal ions. ^a	130
Table 14. The simulated HFE, IOD and CN values of the IOD parameter set for divalent metal ions. ^a	131
Table 15. Final optimized CM parameter set of divalent metal ions for different water models.....	133
Table 16. The simulated HFE, IOD and CN values of the CM parameter set for divalent metal ions. ^a	134

Table 17. Final optimized HFE parameter set of monovalent ions for different water models.....	136
Table 18. The simulated HFE, IOD and CN values of the HFE parameter set for monovalent metal ions. ^a	137
Table 19. Final optimized IOD parameter set of monovalent ions for all the three water models.....	139
Table 20. The simulated HFE, IOD and CN values of the IOD parameter set for monovalent metal ions. ^a	140
Table 21. Estimated HFE parameter set of highly charged metal ions for different water models.....	142
Table 22. Estimated IOD parameter set of highly charged metal ions for different water models.....	143
Table 23. Comparison between the $R_{\min}/2$ parameters from the HFE parameter set for the TIP3P water model and the VDW radii calculated from QMSP method.....	144
Table 24. The activity derivatives for six different ionic solutions in the SPC/E water model (under ~0.3 M condition).....	145
Table 25. The average IOD errors for the HFE parameter set of the mono-, di-, tri- and tetravalent cations.	146
Table 26. The average HFE errors for the IOD parameter sets of the mono-, di-, tri- and tetravalent cations.	147
Table 27. Final optimized 12-6-4 parameter set of divalent metal ions for different water models.....	148
Table 28. Computed C_4 values for the 12-6-4 parameter set of divalent metal ions for different water models. ^a	149
Table 29. The simulated HFE, IOD and CN values of the 12-6-4 parameter set for divalent metal ions. ^a	150
Table 30. Final optimized 12-6-4 parameter set of monovalent ions for different water models.....	151
Table 31. The simulated HFE, IOD and CN values of the 12-6-4 parameter set for monovalent ions. ^a	153
Table 32. Final optimized 12-6-4 parameter set of highly charged metal ions for different water models.....	154

Table 33. The simulated HFE, IOD and CN values of the 12-6-4 parameter set for highly charged metal ions. ^a	156
Table 34. Polarizabilities used in the 12-6-4 model for different atom types in the AMBER FF.....	158
Table 35. The IODs of the Mg ²⁺ and Cl ⁻ ions at different concentrations. ^a	159
Table 36. Comparison of the structural properties predicted by the three different Mg ²⁺ parameters used in the simulations of a Mg ²⁺ -nucleic acid system.....	160
Table 37. Polarizability values of the AMBER atom types used in the 12-6-4 nonbonded model simulation of a metalloprotein system. ^a	161
Table 38. Small model capping schemes for coordinated amino acid residues.....	162
Table 39. The RMSD values between the CSD, QM optimized and MM minimized structures. ^a	163

LIST OF FIGURES

Figure 1. Atomic ions that were parameterized for the 12-6 model in current work. These ions are shown with blue background. The monovalent anions, monovalent cations, divalent cations, trivalent cations, and tetravalent cations are shown in light green, yellow, orange, red, dark red, respectively. The ions that have multiple oxidation states are shown in white, which are $\text{Cr}^{2+}/\text{Cr}^{3+}$, $\text{Fe}^{2+}/\text{Fe}^{3+}$, $\text{Cu}^{+}/\text{Cu}^{2+}$, $\text{Ag}^{+}/\text{Ag}^{2+}$, $\text{Ce}^{3+}/\text{Ce}^{4+}$, $\text{Sm}^{2+}/\text{Sm}^{3+}$, $\text{Eu}^{2+}/\text{Eu}^{3+}$, and $\text{Tl}^{+}/\text{Tl}^{3+}$. Besides these atomic ions, the H_3O^{+} and NH_4^{+} ions were also parameterized in current work.	164
Figure 2. Thermodynamics cycle used to determine the HFE values.	165
Figure 3. HFE fitting curves for 24 M(II) metal ions for the TIP3P water model.	166
Figure 4. HFE fitting curves for Zn^{2+} in four different water models.	167
Figure 5. IOD fitting curves for Zn^{2+} in four different water models.	168
Figure 6. CN switching during a MD simulation when $R_{\min}/2 = 2.2 \text{ \AA}$ and $\epsilon = 0.1 \text{ kcal/mol}$ for a M(II) ion in the TIP3P water model.	169
Figure 7. HFE and IOD fitting curves of six representative M(II) metal ions in the TIP3P water model.	170
Figure 8. Determination of the three parameter sets for Zn^{2+} in the TIP3P water model.	171
Figure 9. Radial distribution functions of the three parameter sets for Zn^{2+} in the TIP3P water model.	172
Figure 10. Fitting curves between the HFE and IOD values for the positive (upper) and negative (lower) monovalent ions in the three water models together with the target values of the ions investigated in the present work.	173
Figure 11. Fitting curves between the HFE and $R_{\min}/2$ values for the positive (upper) and negative (lower) monovalent ions in the three water models.	174
Figure 12. Fitting curves between the IOD and $R_{\min}/2$ values for the positive (upper) and negative (lower) monovalent ions in the three water models.	175
Figure 13. Unsigned average errors of the VDW radii of different parameter sets using QMSP calculated values as reference. The parameter sets developed herein are shown in the seven green columns on the right side of the figure.	176
Figure 14. PDB entry 4BV1. Waters are not shown in the figure, the ferric ions are shown as silver spheres. This picture was created using VMD. ¹	177

Figure 15. Chain C in PDB entry 4BV1 (upper) and a close-up figure of the metal site in Chain C (lower). The ferric ion is represented as a silver sphere and it is coordinated by one Cys, four His and one water molecule. The figures were made using VMD. ¹	178
Figure 16. RMSDs of the backbone CA, C, N atoms (upper) and the metal site heavy atoms (lower) for the simulations with the IOD and 12-6-4 parameter sets using the initial structure (experimental structure) as reference.	179
Figure 17. RMSFs of the heavy atoms of the protein residues in the simulations using the IOD and 12-6-4 parameter sets.....	180
Figure 18. HFE percent errors for different parameter sets of monovalent ions in the TIP3P water model. HZ ⁺ and HE ⁺ represent the H ⁺ Zundel and Eigen ions respectively. Since we did not design HFE set of parameters for the HZ ⁺ and HE ⁺ ions, they are not shown in the figure.	181
Figure 19. IOD percent errors for different parameter sets of monovalent ions in the TIP3P water model. HZ ⁺ and HE ⁺ represent the H ⁺ Zundel and Eigen ions respectively. Since we did not design IOD set of parameters for the HZ ⁺ and HE ⁺ ions, they are not shown in the figure.	182
Figure 20. The average errors (upper) and average percentage errors (lower) of the IODs for the HFE parameter set for the mono-, di-, tri- and tetravalent cations in the TIP3P water model.	183
Figure 21. The average errors (upper) and average percentage errors (lower) of the HFEs for the IOD parameter set for the mono-, di-, tri- and tetravalent cations in the TIP3P water model.	184
Figure 22. Fitting curves between the HFE and R _{min} /2 values (upper) and fitting curves between the IOD and R _{min} /2 values (lower) for the positive and negative monovalent ions in different water models.....	185
Figure 23. Scheme representing intermolecular interactions: the green double-headed arrow and red double-headed arrow represent the interactions that are included and not included in the 12-6 nonbonded model, respectively.	186
Figure 24. Atomic ions that were parameterized for the 12-6-4 model in current work. These ions are shown with blue background. The monovalent anions, monovalent cations, divalent cations, trivalent cations, and tetravalent cations are shown in light green, yellow, orange, red, dark red, respectively. The ions that have multiple oxidation states are shown in white, which are Cr ²⁺ /Cr ³⁺ , Fe ²⁺ /Fe ³⁺ , Cu ⁺ /Cu ²⁺ , Ce ³⁺ /Ce ⁴⁺ , and Tl ⁺ /Tl ³⁺ . Besides these atomic ions, the H ₃ O ⁺ and NH ₄ ⁺ ions were also parameterized in current work.....	187
Figure 25. HFE and IOD fitting curves for the Cu ²⁺ , Ni ²⁺ , Zn ²⁺ , Co ²⁺ , and Cr ²⁺ ions in the TIP3P water model.	188

Figure 26. HFE and IOD fitting curves for the Fe^{2+} , Mg^{2+} , V^{2+} , Mn^{2+} , and Hg^{2+} ions in the TIP3P water model.	189
Figure 27. HFE and IOD fitting curves for the Cd^{2+} , Ca^{2+} , Sn^{2+} , Sr^{2+} , and Ba^{2+} ions in the TIP3P water.	190
Figure 28. Workflow of the MCPB.py program.....	191
Figure 29. Examples of capping different metal bound amino acid residues in the small model. Structures on the left are the coordinated amino acids in the protein, and the structures on the right are the capped residues in the small models built by MCPB.py. Panels A and B are for PDB entry 1E67, panels C, D, and E are for PDB entries 1AK0, 1Y79, and 2BZS, respectively. Zinc coordination is shown as dotted lines with the bond lengths given in Å.....	192
Figure 30. Metal ions currently supported by the MCPB.py program. The metals with a blue background use the VDW parameters from section 2 while the metals with green background use the VDW parameters adapted from UFF. ²	193
Figure 31. Structure of chain A from PDB entry 1E67 with the zinc ion represented by a VDW sphere and the metal site residues indicated by sticks. This figure was made using VMD. ¹	194
Figure 32. Content of the MCPB.py input file for the construction of the metal site found in chain A of PDB entry 1E67.....	195
Figure 33. Commands used for the construction of the metal site found in chain A of PDB entry 1E67 using MCPB.py. The second to the fifth lines are for the QM calculations using Gaussian03, while the last line is to build the topology and coordinate files using tleap.....	196
Figure 34. Bond parameters (above) and angle parameters (below) for the metal site in chain A from the 1E67 PDB structure. Here M1 is the atom type of the zinc ion, while Y1, Y2, Y3, Y4 are the atom types for the four atoms bound to the zinc ion, which are the GLY45 backbone oxygen atom, the d nitrogen atom in HIE46, the sidechain sulfur atom in CYM112 and the d nitrogen atom in HIE117, respectively. Here HIE and CYM are the “AMBER style” residue names: HIE means a His residue which has the e nitrogen protonated, CYM indicates a Cys residue which is negatively charged (<i>i.e.</i> with a CH_2S^- sidechain group).	197
Figure 35. RMSD values of the protein backbone N, CA and C atoms (upper) and the heavy atoms in the metal site (lower) over 20 ns of simulation. The RMSD values were calculated using CPPTRAJ. ³	198
Figure 36. Structure of residue 1 from CSD entry FAJYAR01. This figure was made using VMD. ¹	199

Figure 37. Content of the MCPB.py input file for the construction of the metal site found in residue 1 of CSD entry FAJYAR01.	200
Figure 38. Commands used for the construction of the metal site found in residue 1 of CSD entry FAJYAR01 using MCPB.py. The second to the fifth lines are for the QM calculations using Gaussian03, while the last line is to build the topology and coordinate files using tleap.	201
Figure 39. Bond parameters (above) and angle parameters (below) for the metal site of residue 1 from CSD entry FAJYAR01. Here M1 is the atom type of the Os^{2+} ion, while Y1 to Y6 are the atom types for the six metal bound nitrogen atoms.	202
Figure 40. Comparison of the QM and MM calculated normal modes.	203
Figure 41. Cartoon representation of Cu(I)•CusF (PDB entry 2VB2). Cu^+ is shown as a pink sphere, and the metal binding residues are represented as sticks.	204
Figure 42. Cu^+ –Trp44 distance and Glu46–Met47 Φ dihedral coordinates (black lines) for our two-dimensional PMF MD simulations shown in (a) a closed structure and (b) a wide-open structure of Cu(I)•CusF.	205
Figure 43. Scheme employed to generate the starting conformations for the two-dimensional PMF calculation. (Left) Initial conformations were generated by sampling along the diagonal from the closed to the wide-open conformations. Arrows indicate the direction of sampling along the reaction coordinate. Equilibrated conformation from a window was used as the starting coordinate of the adjacent window as we moved along the diagonal. (Right) Scheme employed to sample conformations starting from windows along the diagonal (solid circles) to the hollow circles.	206
Figure 44. (Upper) A two-dimensional free energy landscape calculated for the Cu–Trp44 distance <i>versus</i> the Glu46–Met47 Φ dihedral from our classical MD simulations of Cu(I)•CusF. Cu(I)•CusF is unable to make the transition from the closed to the open conformations. The closed conformations are represented by smaller Cu^+ –Trp distances, while the open conformations are represented by larger Cu^+ –Trp44 distances. (Lower) PMF profile based on the umbrella sampling simulations in which the Cu^+ –Trp44 distance and Glu46–Met47 Φ dihedral were gradually modified to force the transition from the closed conformation to the wide-open conformation of Cu(I)•CusF. Error bars were calculated using a bootstrapping algorithm implemented in Alan Grossfield’s WHAM code (http://membrane.urmc.rochester.edu/content/wham). Error bars were calculated along each axis of the 2D PMF profile. The error bars along the dihedral dimension are between 0.002 and 0.040 kcal/mol with an average of 0.022 kcal/mol. The error bars along the distance dimension are in the range of 0.001 to 0.128 kcal/mol with an average of 0.012 kcal/mol.	207

KEY TO ABBREVIATIONS

AMBER	Assisted Model Building with Energy Refinement
CHA	Charge hydration asymmetry
CHARMM	Chemistry at Harvard macromolecular mechanics
CM	Compromise
CN	Coordination number
CPMD	Car-Parrinello molecular dynamics
FF	Force field
GAFF	General AMBER force field
GROMACS	Groningen machine for chemical simulations
GROMOS	A force field in the GROMACS software package
HFE	Hydration free energy
IOD	Ion-oxygen distance
KB	Kirkwood-Buff
LAMMPS	Large-scale atomic/molecular massively parallel simulator
LJ	Lennard-Jones
MC	Monte Carlo
MCPB	Metal center parameter builder
MD	Molecular dynamics
MM	Molecular mechanics
NAMD	Nanoscale molecular dynamics
NGC	Noble gas curve

OPLS-AA	Optimized potentials for liquid simulations - all atom
PBC	Periodic boundary condition
PDB	Protein databank
PME	Particle mesh Ewald
QM	Quantum mechanics
QMCf	quantum mechanical charge field
QMSP	Quantum scaling principle
RDF	Radial distribution function
RESP	Restrained electrostatic potential
TI	Thermodynamic integration
UAE	Unsigned average error
VDW	van der Waals
ZAFF	Zinc AMBER force field
pyMSMT	Python metal site modeling toolbox

CHAPTER 1: INTRODUCTION

1.1 General Introduction of Models for Metal Ions

Metal ions play significant roles in chemical disciplines such as geochemistry, biochemistry and materials science. There are 87 metals among the first 112 elements in the periodic table. The elements Al, Fe, Ca, Na, K, Mg and Ti occupy about a quarter of the earth's crust. There are about one third of the proteins in the protein databank (PDB) that contain metal ions.⁴ Metal ions such as calcium, zinc, iron, copper, manganese, nickel, and magnesium ions form complexes with surrounding amino acid residues, and serve significant functional roles including structural, electron transfer and catalytic functions.⁵⁻¹⁸ There are more than 80% large scale industry processes rely on solid catalysis, which are usually related to the transition metal chemistry.¹⁹

There are different ways to simulate these ions using theoretical approaches, for example, the quantum mechanics (QM) method,²⁰⁻²² the molecular mechanics (MM) method^{4, 23, 24} and the combined QM/MM method.²⁵⁻²⁸ Classical force fields (FFs), which use analytical functions to represent the relationship between the energy and configuration of a system, have significant speed advantages over the QM based methods. It is a state-of-the-art tool to study systems at the atomic level when combined with Molecular Dynamics (MD)^{29, 30} or Monte Carlo (MC) method.^{31, 32} There are a number of strategies that have been employed in the classical FFs: the bonded model,^{4, 33-36} the nonbonded model,³⁷⁻⁴⁰ the hybrid model,⁴¹ the cationic dummy model,^{23, 42, 43} the constrained nonbonded model,⁴⁴

the points-on-a-surface model⁴⁵ and the polarizable model⁴⁶⁻⁴⁸ have been described and parameterized to study a broad range of metal containing complexes.

1.2 The 12-6 Nonbonded Model

$$\begin{aligned}
 U_{nonbonded}(r_{ij}) &= \frac{e^2 Q_i Q_j}{r_{ij}} + \frac{C_{12}^{ij}}{r_{ij}^{12}} - \frac{C_6^{ij}}{r_{ij}^6} = \frac{e^2 Q_i Q_j}{r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right] = \\
 &\frac{e^2 Q_i Q_j}{r_{ij}} + 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]
 \end{aligned} \tag{1}$$

The nonbonded model treats the metal ion as a point with an integer charge while the interactions are represented by Columbic and Lennard-Jones (LJ) terms (see equation 1). The first term is a classical Coulomb potential while the second term is a 12-6 LJ potential. In which r_{ij} is the distance between two particles i and j , and the Q_i and Q_j are the point charges of the two particles while e is the proton charge. For the metal ion, its point charge is usually treated as an integer number according to their oxidation state, hence the parameters (C_{12}^{ij}, C_6^{ij}) , $(R_{min,ij}, \varepsilon_{ij})$ or $(\sigma_{ij}, \varepsilon_{ij})$ are the only two parameters that need to be determined. In the nonbonded model, the coordination of the metal ion is flexible, which allows coordination number (CN) switching and ligand exchange at the metal center. However, this model oversimplifies the interaction between the ions and their surrounding residues. In addition to Columbic and van der Waals (VDW) interactions, charge transfer, polarization and even covalent interactions could exist between a metal ion and its surrounding ligands.⁴⁹⁻⁵¹ Furthermore, a single point poorly represents the charge distribution of most ions. The electronic cloud is usually non-symmetrically distributed around the metal ion, which could also further change and redistribute in response to changes in the surrounding environment. Nevertheless, due to the simple form, computational efficiency and excellent transferability characteristics of

the nonbonded model,^{37, 38, 40, 52} it is still extensively used for metal ions in MD simulations even though more sophisticated potential forms exist.

There are numerous systematic studies that have been reported in recent decades regarding the parameterization of the LJ nonbonded model for atomic ions. For example, Åqvist pioneered the development of LJ parameters for the alkali and alkaline-earth metal cations.³⁷ Dang and co-workers developed a series of LJ parameters for alkali metal and halide ions from 1992 to 2012 for either non-polarized or polarized water models.⁵³⁻⁶⁰ Peng and Hagler parameterized the 9-6 potential for alkali metal cations and halide anions.⁶¹ Jensen and Jorgensen have parameterized the LJ potential for the halide ions, alkali metal ions and the ammonium ion using the TIP4P water model.⁶² Roux and co-workers have parameterized the nonbonded model for the alkali metal and halide ions using the SWM4-DP polarized water model.^{47, 63} Babu and Lim re-optimized LJ parameters for biologically relevant +2 metal (M(II)) cations based on experimental relative hydration free energy (HFE) values while the nonbonded interactions were truncated by an atom-based force switching function.⁶⁴ Joung and Cheatham have developed the LJ parameters for alkali metal cations and halide anions for three commonly used water models (TIP3P,⁶⁵ SPC/E,⁶⁶ and TIP4P_{EW}⁶⁷) for use in particle mesh Ewald (PME) simulation.³⁹ Netz and co-workers have designed different parameter sets for alkali metal ions in SPC/E water by treating single-ion and ion-pair properties as targets.⁶⁸ Hasse and co-workers have developed different parameter sets for alkali metal ions and halide ions in the SPC/E water in order to reproduce different experimental endpoints.^{69, 70} Reif and Hunenberger have created LJ nonbonded model parameters for

alkali metal and halide ions using the SPC or SPC/E water model.⁷¹ These parameters have been developed for different combining rules and simulation protocols. Various experimental/theoretical target values were used in the previous parameterization work to simulate ions in various environments (*e.g.*, the gas phase, liquid phase and solid phase, interfacial phase, *etc.*). For example, QM calculated gas phase ion-water interaction energies for the monohydrates, experimental HFEs, enthalpies and entropies, the ion-oxygen distance (IOD) and CN of the first solvation shell, diffusion coefficients, mean residence times, radial distribution functions (RDFs), electric conductivity, dynamic hydration numbers, ion-pair properties, osmotic coefficients, lattice constants, lattice energies, *etc.* Overall, it is hard to reproduce all of the available experimental/theoretical results at the same time due to accuracy limitations of the classical nonbonded model.

LJ parameters always have limited transferability between different water models, mixing rules and simulation conditions.⁷²⁻⁷⁵ The PME method is now the *de facto* standard method used to calculate the long-range electrostatic energy in periodic boundary simulation cells.⁷⁶⁻⁷⁹ It calculates the short-range interactions in real space while the long-range interactions are developed in Fourier space. Importantly, it decreases the time complexity of MD simulations from $O(N^2)$ to $O(N\log N)$ with reliable precision.⁷⁶⁻⁷⁹ However, there was no systematic parameterization work on various ions, which range from the monovalent to tetravalent, performed specifically for MD simulations employing the PME method. By employing the thermodynamic integration (TI) method⁸⁰⁻⁸⁷ and MD approach, we designed LJ parameters for more than 60 ions (range from monovalent to tetravalent) for three widely used water models (TIP3P, SPC/E and

TIP4P_{EW}) respectively. These parameters are specifically designed for the PME method⁷⁷⁻⁷⁹ and different physical properties (*e.g.* HFE, IOD and CN).

Based on our simulation results, we found that the transferability of existing LJ parameters for divalent ions is limited. Meanwhile, we also found that the simulated HFE, IOD and CN values are highly correlated and there appears to be a one-to-one correspondence between them. Results showed that different water models have different properties and should be treated separately when designing the parameters. Moreover, due to the simplicity of the 12-6 nonbonded model, we could not reproduce all HFE, IOD and CN values simultaneously for divalent metal ions. And the stronger of the coordination interaction between the divalent metal ion and water molecules, in general, the larger of the errors in the nonbonded model. Furthermore, we found that generally the TIP3P, SPC/E, TIP4P_{EW} and TIP4P water models experience a successive increase of error.

Besides the error of the 12-6 model described above, which brings about no ideal LJ parameters for divalent metal ions, there are innumerable combinations of the VDW radius and well depth parameters ($R_{\min}/2$ and ϵ) that can generate the same HFE or IOD values,³⁹ which offers a challenge for determining the final parameters. There are different strategies to solve this problem. In some parameter sets, the parameters were designed using fixed ϵ values for all the negative or positive ions.^{60, 62, 69, 70} Jensen and Jorgensen designed “big and small” ϵ parameters for negative and positive ions, respectively, in their parameterization work.⁶² Joung and Cheatham have determined the

parameters empirically based on a compromise between different experimental properties, including the HFE, lattice energy and lattice constant.³⁹ For the alkali metal and halide ion series there is an apparent trend in this parameter set where as $R_{\min}/2$ increases for the bigger ions the ϵ parameter also increases.³⁹ In present work, via a consideration of the physical meaning of the VDW interaction, we fitted a curve (herein termed as the “noble gas curve” – NGC) to describe the relationship between the $R_{\min}/2$ and ϵ values based on the experimental results for noble gas atoms. According to NGC we determined the LJ parameters for various ions.

Towards different potential applications, we designed three parameter sets for more than 60 ions (range from monovalent to tetravalent) in TIP3P, SPC/E and TIP4P_{EW} water models based on the NGC: one for reproduction of experimental HFEs, one for experimental IODs, and a final set representing a compromise between the former two properties (we term these the HFE, IOD, and CM parameter sets, respectively). Overall, the HFE parameter set achieved an error range of $\sim\pm 1.0$ kcal/mol for the absolute HFEs and $\sim\pm 2.0$ kcal/mol for the relative HFEs, the CM set could well reproduce the CN values and achieved an error range of $\sim\pm 2.0$ kcal/mol for the relative HFEs, and the IOD set achieved an error range of $\sim\pm 0.01$ Å for the IODs.

We found that the derived $R_{\min}/2$ parameters coincide well with the VDW radii calculated using the quantum scaling principle (QMSP) method.⁸⁸ The unsigned average errors (UAEs) of the present parameter sets are smaller than those of other parameter sets with respect to the QMSP calculated VDW radii. We also carried out test simulations on six

different ionic solutions. Using the experimental activity derivatives as a basis for comparison, in general, our parameter sets showed better performance than the parameter sets of Joung and Cheatham³⁹ and Horinek *et al.*⁶⁸ This supports our assertion that the current parameters have better transferability due to a better balance between the $R_{\min}/2$ and ϵ parameters from NGC. Meanwhile, in test simulations on a Fe(III) containing protein system, stable trajectories were obtained with the metal binding site being well conserved, which further supports the excellent transferability of these parameters. This work marks a systematic investigation and determination of LJ parameters for various ions that can be employed in PME based MD simulations and these parameters are compatible with FFs such as AMBER,⁸⁹ CHARMM,⁹⁰ OPLS-AA,⁹¹ and GROMOS⁹² when used with the PME model.

Furthermore, error analysis was carried out and results showed that the 12-6 model could well simulate some monovalent ions with strong ionic characteristics (such as Na^+ and K^+ ions). However, it yields non-negligible error, making it could not reproduce several experimental properties simultaneously for monovalent ions that have strong polarization interaction with surroundings (such as Tl^+ and Ag^+) and ions with formal charge equal or larger than +2. Results also showed that the 12-6 model has a much larger underestimation of the ion-water interaction with the increasing of ion's formal charge. Hence, we developed a 12-6-4 LJ-type nonbonded model for metal ions and introduced it below.

1.3 The 12-6-4 Nonbonded Model

The 12-6 LJ nonbonded model (see equation 1) explicitly includes the Pauli repulsion and induced dipole-induced dipole interactions via the LJ potential, while the charge-charge, charge-dipole, and dipole-dipole interactions are represented by the Coulomb potential. However, it doesn't take into account the charge-induced dipole interaction and the dipole-induced dipole interaction explicitly. The former interaction is the dominant one of these two types of interactions in the case of metal ions and its inclusion, in principle, would greatly improve a nonbonded model representing ions. The standard 12-6 LJ nonbonded model is reasonable for neutral systems, while for highly charged systems, the charge-induced dipole interaction, which is proportional to r^{-4} , becomes very significant and needs to be considered. With these considerations, we decided to add an r^{-4} term into the standard 12-6 LJ nonbonded model in order to include the ion-induced dipole interaction.

We have parameterized the 12-6-4 model for more than 50 ions. The parameterization was also performed separately for TIP3P, SPC/E and TIP4P_{EW} these three widely used water models. Unlike the 12-6 LJ nonbonded model, the new 12-6-4 LJ-type model can reproduce different target values (HFE, IOD and CN) with good accuracy at the same time after selecting appropriate parameters. Afterwards, independent research from Panteva *et al.* proved the outperformance of the 12-6-4 models over the 12-6 models for simulating the Mg^{2+} -aqueous system.⁹³

Meanwhile, we found that the magnitude of final C_4 parameters were consistent with the analytical formulation of the charge-induced dipole potential. Furthermore, results also showed that the error of the 12-6 model for cations is approximately proportional to the square of the cation's formal charge. This trend is also consistent with the original equation of the charge-induced dipole interaction, which kind of interaction is omitted in the 12-6 nonbonded model. Based on the charge-induced dipole interaction, we proposed a specific combining rule for the C_4 terms. By employing the new combining rule it was shown that the 12-6-4 model was readily transferable to mixed systems such as salt solutions, nucleic acid and protein systems. In the salt solution simulations, the 12-6-4 parameter set showed the best performance among the investigated parameter sets, which further illustrated it is a superior model relative to the 12-6 LJ nonbonded model.

Furthermore, the 12-6-4 nonbonded model is compatible with FFs like AMBER,⁸⁹ CHARMM,⁹⁰ OPLS-AA,⁹¹ and GROMOS⁹² with almost no additional computing cost when comparing to the 12-6 model. Even so, attention should be paid when the new nonbonded model is employed in systems with strong charge-transfer effects since it may be unphysical to treat the central metal atom as an actual “ion” in such cases.

1.4 The Bonded Model and MCPB.py Program

The bonded model is widely used in contemporary FFs like AMBER,⁸⁹ CHARMM,⁹⁰ GROMOS⁹² and OPLS-AA,⁹¹ and the generic functional form is shown in equations 1-3 (CHARMM has an additional 1-3 Urey-Bradley nonbonded term, which is not shown in these equations). The total energy of a system is represented by its bonded part and nonbonded part (see equation 2). The bonded energy consists of the bond, angle and torsion terms (see equation 3). The bond and angle terms are described using harmonic equations while the torsion potentials is represented by a Fourier expansion. In equation 3, k_r , r_{ij} , $r_{ij,eq}$ are the bond force constant, bond length and equilibrium bond length; K_θ , θ_{ij} , and $\theta_{ij,eq}$ are the angle force constant, angle amplitude and equilibrium angle values; V_n , n , ω , γ are the torsion barrier, periodicity, torsion angle, and phase, respectively. The nonbonded term is described by equation 1 and was introduced in section 1.2.

$$U_{total} = U_{bonded} + U_{nonbonded} \quad (2)$$

$$U_{bonded}(r_{ij}) =$$

$$\sum_{All\ bonds} k_r (r_{ij} - r_{ij,eq})^2 + \sum_{All\ angles} k_\theta (\theta_{ij} - \theta_{ij,eq})^2 + \sum_{All\ torsions} \sum_n \frac{1}{2} V_n [1 + \cos(n\omega - \gamma)] \quad (3)$$

There are several ways to obtain force constants parameters: through empirical methods,⁹⁴ through experiments (*e.g.*, X-ray, NMR, normal mode analysis of spectra *etc.*),^{95, 96} or based on theoretically calculated Hessian matrices.^{97, 98} In general, the empirical method could be applied more broadly but usually offers limited accuracy. Deriving parameters from experimental information is very challenging and time-consuming, thereby, restricting it to specific systems. Parameter determination based on

quantum calculations offers considerable accuracy and is applicable to a wide range of molecules.

Seminario proposed a method which uses the Cartesian Hessian matrix to calculate the force constant (referred as the Seminario method) and validated it through a series of small organic molecules.⁹⁸ Nilsson *et al.* developed the Hess2FF software to calculate the force constants through the Seminario method. They applied it to 5 different systems with some containing either a Fe or Zn ion.⁹⁹ Lin and Wang applied the Seminario method and the restrained electrostatic potential (RESP) fitting scheme to zinc complexes with the general AMBER force field (GAFF).¹⁰⁰ They showed that the bonded model with RESP fitted charges showed the best performance among the models investigated.¹⁰⁰ Peters *et al.* developed the metal center parameter builder (MCPB) software based on the MTK++ software package using the C++ language.⁴ It is a semi-automatic tool for the parameterization of metal ion containing molecules. The zinc AMBER force field (ZAFF) has been developed for four-coordinated zinc complexes in protein system using the Seminario/ChgModB combination (shown as the best combination) using MCPB.⁴ After the availability of MCPB a broad range of metal ion containing systems have been parameterized using this program.^{34, 36, 101, 102} However, even with the considerable time saving afforded by MCPB, the process is still overly complicated for the non-expert.

Herein we introduce the MCPB.py software, a python based metal center parameter builder, which streamlines much of the functionality of MCPB into a much easier to use program. It was built on the python metal site modeling toolbox (pyMSMT) in

AmberTools15.¹⁰³ It uses a much more optimized workflow, offering a more user-friendly experience with far fewer steps requiring user input. It supports Z-matrix, Seminario and empirical methods for the parameterization. In the current version it supports a variety of AMBER FFs, more than 80 ions, and two widely used QM software packages (Gaussian¹⁰⁴ and GAMESS-US¹⁰⁵). Together with ParmEd¹⁰³ or ACEYPE¹⁰⁶ the topology and coordinate files of AMBER can be converted to the formats used by CHARMM¹⁰⁷ or GROMACS.¹⁰⁸ We expect the MCPB.py application to further expand and expedite the modeling of ions in metalloproteins and organometallic compounds using a range of packages and FFs.

1.5 The CusF Metallochaperone System

CusF is an Ag^+/Cu^+ bound metallochaperone in the bacterial pathogen *Escherichia coli*. It is a member of the CusC(F)BA proteins, which belongs to the resistance-nodulation-cell division superfamily that provide Ag^+ and Cu^+ resistance for the bacterial.^{109, 110} It has a specific cation- π coordination in which the metal ion coordinated to the 6-member aromatic ring in a tryptophan residue. In a previous study, the $\text{Cu}^+/\text{Ag}^+-\pi$ interaction in CusF was evaluated to be on the order of ~ 10 kcal/mol.¹¹¹ In a prevailing hypothesis of metal ion efflux, CusF metallochaperone transports Cu^+/Ag^+ ions from the periplasm to the CusCBA metal ion transporter protein-complex, which expels the metal ion from the cell.^{112, 113} The mechanism of metal ion capture by CusF, and its subsequent transfer to CusB, however, remains poorly understood. Recent experiments showed that CusF interact transiently with the N-terminal disordered region of CusB when either one of them is in the metal bound form, which allows metal ion transfer reversibly.¹¹⁴ Furthermore, metal ion binding quenches conformational dynamics in both proteins. These results indicate that the conformational dynamics of CusF is important for the process of CusF-CusB complex formation and metal ion transferring. Using a bonded/nonbonded hybrid model together with the umbrella sampling technique, we studied the potential of mean force (PMF) changes of the Cu^+ bounded CusF from its closed to wide-open conformations. Results showed that the transformation barrier is ~ 8 kcal/mol, which supports the hypothesis that the open conformation is important for metal ion releasing from the CusF metallochaperone.

CHAPTER 2: PARAMETERIZATION OF THE 12-6 NONBONDED MODEL

2.1 Target Values

In this work we designed parameters based on three different target values: HFE, IOD and the CN of the first solvation shell. HFE values represent the thermodynamic properties while the other two represent structural properties. Target HFEs, IODs and CNs for investigated ions are given in Tables 1-3. Overall, we have parameterized the 12-6 LJ nonbonded model for 63 different ions (see Figure 1, including 4 monovalent anions, 11 monovalent cations, 24 divalent cations, 18 trivalent cations, and 6 tetravalent cations) in three widely used water models (TIP3P, SPC/E and TIP4P_{EW}) respectively.

There are different sets of experimental HFE values for ions that exist in the literature.^{115,}

¹¹⁶ Usually these values were obtained based on the proton's HFE value but final agreement on this value has not been reached. The general range of it is from -250 to -265 kcal/mol.^{71, 115, 117-120} For example, in Reif and Hünenberger's work on design of the monovalent ion parameters,⁷¹ they supported a value of -1100 kJ/mol (~-263 kcal/mol) for the HFE of proton. Herein we used the experimental HFEs of ions from Marcus¹¹⁵ and Schmid *et al.*¹¹⁶ Marcus obtained the HFEs of cations based on $\Delta_{\text{hyd}}G^\circ(\text{H}^+) = -1056$ kJ/mol (~ -252 kcal/mol) from $\Delta_{\text{hyd}}H^\circ(\text{H}^+) = -1094$ kJ/mol (~ -261 kcal/mol) and $\Delta_{\text{hyd}}S^\circ(\text{H}^+) = -131$ J/(K•mol) or $S_\infty[\text{H}^+(\text{aq})] = -22.2$ J/(K•mol). It is one of the most complete databases regarding the thermodynamic properties of ions. Schmid *et al.* refitted the HFE values of the atomic ions based on $\Delta_{\text{hyd}}H^\circ(\text{H}^+) = -1078$ kJ/mol (~ -258

kcal/mol).¹¹⁶ For monovalent cations, there is a trivial difference (~ 0.7 kcal/mol on average) between the data sets of Marcus and Schmid *et al.* For alkaline earth metal ions, the average difference is ~ 3.9 kcal/mol between the two sets. We have used Marcus' HFE values as target values since this set is more complete. The HFE value of the H_3O^+ ion is from Palascak and Shields¹²¹ due to its absence in the data sets of Marcus¹¹⁵ and Schmid *et al.*¹¹⁶ For the halide ions, Schmid *et al.*'s values are ~ 8 kcal/mol lower than Marcus' values on average. Herein, we used Schmid *et al.*'s data set as our target since it is consistent with a broader range of experimental values.^{39, 122}

Most of the IOD values of the monovalent ions come from Marcus' review¹²³ except those of the Tl^+ , Cu^+ , NH_4^+ and H^+ ions (see Table 1). The Tl^+ ion is similar to the Sn^{2+} ion which all have a lone pair electron in the outmost electron shell, resulting in the observation of two IODs in the first solvation shell in the aqueous phase. This effect is hard to reproduce using classical MD simulations due to the assumption of isotropic behavior. Typically, the reported CN value of the Tl^+ ion is between 6-8 while Persson *et al.* determined the CN of Tl^+ as 4 for aqueous systems.¹²⁴ In their work they found a 4-coordinated Tl^+ ion with two different IOD values representing two water molecules at 2.73 Å and two at 3.18 Å. The IOD value of Tl^+ (2.96 Å) used in the present work is the average of these two values. Shannon calculated the effective ionic radii of 6- and 8-coordinated Tl^+ ion as 1.50 Å and 1.59 Å respectively.¹²⁵ By adding the effective ionic radius of O^{2-} (1.40 Å) from Pauling, we can estimate that the IOD values for Tl^+ are 2.90 Å and 2.99 Å for 6 and 8-coordinated structures, respectively, which is consistent with the IOD value we are using (2.96 Å). Vchirawongkwin *et al.* performed QM/MM MD

simulations on the aqueous Ti^+ system and observed two different IODs (with 2.79 and 3.16 Å) and determined the average CN as 5.9.¹²⁶ Their theoretical work supports the “structure-breaking” character of Ti^+ in bulk system.¹²⁶ It is hard to obtain IOD value for the Cu^+ ion since it is easily oxidized by water. The IOD value for Cu^+ used in present work comes from quantum calculations of Burda *et al.*¹²⁷ There is no reliable experimental IOD value for NH_4^+ . We estimated the IOD value (2.85 Å) based on the ionic radius of the NH_4^+ ion (1.45 Å) from Detellier and Laszlo¹²⁸ and the ionic radius of O^{2-} (1.40 Å) from Pauling.¹²⁹ This value is consistent with classical and Car-Parrinello molecular dynamics (CPMD) simulations.^{130, 131} H^+ in aqueous solution is thought to exist either as the Zundel (H_5O_2^+), Eigen (H_9O_4^+) or hydronium ion (H_3O^+). The IOD value of H^+ in the Zundel form was obtained from the quantum calculations of Meraj and Chaudhari.¹³² While the IOD value of H^+ in the Eigen ion was taken from the calculations of Sobolewski and Domcke done at the MP2/6-31+G** level of theory.¹³³ Blauth *et al.* investigated aqueous Ag^+ by using the quantum mechanical charge field (QMCF) MD method, in which they found the CN of the first solvation shell to be 6. The Ag^+ ion also acts as a “structure-breaking” factor in the bulk system.¹³⁴ Blumberger *et al.* investigated the Cu^+ , Cu^{2+} , Ag^+ and Ag^{2+} ions in the bulk systems using the CPMD simulation method, in which they found the CN of these ions to be 2, 5-6, 4 and 5 respectively.¹³⁵

Most of the IOD and CN values for M(II) ions were referred from Marcus’ review,¹²³ while the rest were from some other references (see Table 2). The IOD and CN values for M(III) ions were taken from Marcus’ review,¹²³ while the IOD and CN values for the M(IV) ions were taken from a number of sources (see Table 3).¹³⁶⁻¹³⁹ The experimental

effective ionic radii of M(III) and M(IV) ions were obtained from Shannon.¹²⁵ Based on the IOD values and effective ionic radii, we estimated the effective radii of the coordinated water and display the data in Table 3. Some highly charged ions (with charge larger than +2) which readily hydrolyze water such as As^{3+} , Sn^{4+} and Pb^{4+} ions²⁶ were not considered in the present work.

2.2 Methods

For $R_{min,ij}$ (or σ_{ij}) in equation 1, there are two combining rules - the Lorentz combining rule (see equation 4) and the Good-Hope combining rule (see equation 5) that are widely used in MM simulations.

$$R_{min,ij} = R_{min,i} + R_{min,j} \text{ or } \sigma_{ij} = \sigma_i + \sigma_j \quad (4)$$

$$R_{min,ij} = \sqrt{R_{min,i} \times R_{min,j}} \text{ or } \sigma_{ij} = \sqrt{\sigma_i \times \sigma_j} \quad (5)$$

For the ε_{ij} (potential well depth) parameter in equation 1, the Berthelot combining rule is commonly used (see equation 6).

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \times \varepsilon_j} \quad (6)$$

The Lorentz-Berthelot combining rules (equations 4 and 6) is used in the AMBER⁸⁹ and CHARMM⁹⁰ FFs while the geometric-mean combining rules (equations 5 and 6) is used in the OPLS-AA⁹¹ FF. Parameters using one combining rule usually need to be modified when moving to another combining rule. Herein all the simulations use the Lorentz-Berthelot combining rules, but if one wants to use the geometric combining rules, they can adapt the parameters by using equations 4-6. The LJ parameters for the water models (TIP3P,⁶⁵ SPC/E,⁶⁶ TIP4P⁶⁵ and TIP4P_{EW}⁶⁷) employed in this work are shown in Table 4.

TI method^{80, 82-87} is a powerful tool to study the free energy difference between two different states of one system. It was used to simulate the HFEs in the present work. It employed a mixed potential of the initial and final states (V_0 and V_I respectively in equation 7) during the simulation process. Herein k is an integer number which equals 1

when linear mixing is employed, λ is a number between 0 and 1, representing the mixing extent of the two states ($V(\lambda)$ is equal to V_0 when $\lambda=0$ while $\lambda=1$ results in $V(\lambda)=V_1$).

$$V(\lambda) = (1 - \lambda)^k V_0 + [1 - (1 - \lambda)^k] V_1 \quad (7)$$

To simulate the VDW disappearance or appearance process, the soft-core scaling method with linear mixing (with $k=1$ in equation 7) was utilized.⁸¹ It employs a λ dependent modified LJ equation (see equation 8). Herein r_{ij} is the distance between the vanishing atom and another atom, α is a constant set to 0.5 and σ equals $R_{min,ij}/(2^{1/6})$. When $\lambda=0$ it is identical to a normal 12-6 LJ equation while when λ approaches 1 it displays a smooth interaction between the “soft-core” atom and a surrounding particle, allowing them to approach each other closely with a finite energy penalty. It prevents the “end-point catastrophe” when we “turn on” the particle, thereby eliminating related artifacts in an effective way.

$$V_{soft-core VDW} = 4 \varepsilon (1 - \lambda) \left[\frac{1}{\left[\alpha \lambda + \left(\frac{r_{ij}}{\sigma} \right)^6 \right]^2} - \frac{1}{\alpha \lambda + \left(\frac{r_{ij}}{\sigma} \right)^6} \right] \quad (8)$$

Herein we run a mixture of NVT and NPT simulations, so we make the approximation:

$$\Delta A = \Delta G \quad (9)$$

The free energy difference of two states is obtained by integration of $\partial V / \partial \lambda$ values along the λ coordinate (see equation 10) in the NVT and NPT ensembles. The results could be fit to a cubic spline or quadratic curve, while in present work we employed Gaussian quadrature to calculate the integration (see equation 11, where w_i are the weights for the different i values).

$$\Delta A = A(\lambda = 1) - A(\lambda = 0) = \int_0^1 \langle \partial V / \partial \lambda \rangle_\lambda d\lambda \quad (10)$$

$$\Delta A = \sum w_i \langle \partial V / \partial \lambda \rangle_i \quad (11)$$

Based on the consideration of balance of accuracy and speed, we used three methods (*i.e.* Methods 1-3) to simulate the HFE, IOD and CN values of ions. For Methods 1 and 2 all the simulations were carried out using the AMBER 11 suite of programs¹⁴⁰ while the modeling and data analyses were performed using the AmberTools suite of programs.¹⁴⁰ For Method 3 the AMBER 12 suite of programs¹⁴⁰ was used to perform the simulations while the AmberTools suite of programs¹⁴⁰ was utilized to carry out the data analysis. The PME⁷⁷⁻⁷⁹ method and periodic boundary condition (PBC) were used during the MD and TI simulations. The “tin-foil” boundary condition was employed for the systems that are not neutral. The time-step was set as 1 fs and the cut-off was set to 10 Å unless specified. The Langevin algorithm was used to control the temperature with a collision frequency of 5 ps⁻¹. For simulations performed in the NPT ensemble, Berendsen’s barostat with isotropic position scaling was utilized for pressure control with the relaxation time set as 10 ps and 1 ps for the TI and normal MD simulations, respectively. The SHAKE algorithm¹⁴¹ was used to constraint the distances between hydrogen atoms and their attached heavy atoms with a tolerance of 1.0×10^{-5} Å while the “three-site” algorithm was used for the water molecules.¹⁴² In present work, λ values were set to 0.1127, 0.5 and 0.88729 in a 3-window TI simulation. λ values were treated as 0.1127, 0.5, 0.88729 and 0.98 in a 4-window TI simulation. λ values were chosen as 0, 0.1127, 0.5, 0.88729 and 1 for a 5-window TI simulation. λ values were set to 0, 0.04691, 0.23076, 0.5, 0.76923, 0.95308 and 1 for a 7-window TI simulation. While for a 9-window TI simulation λ values were treated as 0, 0.2544, 0.12923, 0.29707, 0.5, 0.70292, 0.87076, 0.97455, and 1. Each window began from the final snapshot of the previous window, the windows of $\lambda=0$, 0.98 (only used in the 4-window TI simulations), and 1

served to equilibrate the system and were not considered in the final free energy calculation using equation 11. Herein the thermodynamic cycle shown in Figure 2 was used to simulate HFEs of ions. Generally, in present work, we estimated the uncertainties of the simulated HFE values using the 12-6 model as $\sim\pm 1$ kcal/mol for the monovalent¹⁴³ and divalent ions,¹⁴⁴ and $\sim\pm 2$ kcal/mol for the highly charged ions.¹⁴⁵

First, we created a $\sim(46\text{\AA}\times 46\text{\AA}\times 46\text{\AA})$ cubic water box surrounding a dummy atom with the closest water molecule at least $1.5\text{\AA}$ away from the dummy atom. In total, there were 2439 water molecules in the system for the TIP3P and SPC/E water models while for the TIP4P and TIP4P_{EW} water models this number was 2389. We performed 1000 steps of minimization using the steepest descent algorithm followed by 1000 steps of minimization using the conjugate gradient algorithm. Afterwards a 1 ns heating process was simulated in the NVT ensemble to heat the system from 0 to 300 K. And then a second 1 ns simulation was carried out at 300K in the NVT ensemble to equilibrate the system. To correct the system density, a 1 ns NPT simulation was performed under 1 atmosphere (atm) and 300 K with the final structure was treated as the starting structure for TI simulations using Method 2 (details are shown below). Finally another 1 ns NVT simulation was conducted to prepare the initial structure for TI simulations using Method 1 (details are shown below).

Method 1 was used for two-dimensional scanning of the LJ parameter space of M(II) ions. In this protocol, we performed LJ parameter space scanning for the M(II) ion with a fixed mass of 65.4 g/mol (referred as a Zn^{2+} ion, the choice of mass has a limited influence on

the simulated energetic and structural properties). The range of $R_{\min}/2$ was chosen as 0.3-2.5 Å with a 0.1 Å interval and ϵ was evenly distributed in the range of 10^{-6} -1 kcal/mol in the logarithmic scale with a interval of 1 for $-\log(\epsilon)$ (in total there are 23 different $R_{\min}/2$ values, 7 different ϵ values, forming $23 \times 7 = 161$ different combinations of LJ parameters for the M(II) ions). All combinations of the LJ parameters were investigated for each water model in the present work. To balance speed and accuracy, we used a one-step method (turn on the VDW and electrostatic interactions of the metal ion in one step) to obtain ΔG_{Total} and $-\Delta G_{\text{Total}}$ (see Figure 2). For each LJ parameter combination, we performed the simulation as described below. First, the ion hydration process was simulated in the NVT ensemble using a 9-window linear TI simulation where each window had a 200 ps simulation time with the $dV/d\lambda$ values in the last 150 ps were collected and averaged. Then the ΔG_{Total} value was obtained via Gaussian quadrature using equation 11. Subsequently, we performed a 1 ns MD simulation with snapshots collected every 1000 steps over the last 500 ps simulation. From these snapshots the ion-oxygen RDF was obtained with a resolution of 0.01 Å based on the average volume of the trajectory. The IOD value was obtained from two times of quadratic fittings based on the RDF: the first quadratic fitting was done for the data within ± 0.1 Å of the peak of the first solvation shell (in total 21 points with a 0.01 Å interval along the RDF were considered). The second quadratic fitting was performed over the RDF data within ± 0.1 Å of the point that was closest to the apex of the first fitting. The IOD value was obtained from this final fit and kept with two decimal places. The CN value was determined via integrating the ion-oxygen RDF from the origin to its first minimum. Finally, we carried out the backwards TI simulation to obtain the $-\Delta G_{\text{Total}}$ value using the same method as

determining the ΔG_{Total} value. Finally the $-\Delta G_{\text{Total}}$ and ΔG_{Total} values were used to determine the HFE value.

For our final determination of the LJ parameters for M(II) ions, we employed Method 2, which is a more consistent method to obtain HFE, IOD and CN values. Using TI simulations, we determined the ΔG_{VDW} , ΔG_{Ele} , $-\Delta G_{\text{Ele}}$ and $-\Delta G_{\text{VDW}}$ values (which correspond to the free energy changes for the VDW-appearing, charge-appearing, charge-disappearing and VDW-disappearing steps, respectively – see Figure 2) in the NPT ensemble consecutively. For ΔG_{VDW} and $-\Delta G_{\text{VDW}}$, we employed a 3-window soft-core linear scaling method due to the better performance of soft-core method over both the linear and nonlinear scaling methods without using a soft-core potential.⁸¹ For the VDW scaling simulations, each window was equilibrated for 100 ps followed by 200 ps of production while for the electrostatic scaling simulations, each window was equilibrated for 50 ps followed by 150 ps of production. Herein the HFE value are computed using $\text{HFE} = \frac{1}{2} \times (\Delta G_{\text{VDW}} + \Delta G_{\text{Ele}} - (-\Delta G_{\text{Ele}} - \Delta G_{\text{VDW}}))$. Finally we modeled the ion-aqueous system again and performed a 2000 step minimization (including 1000 steps of minimization using the steepest descent algorithm followed by 1000 steps of minimization using the conjugate gradient algorithm), a 500 ps NVT heating, a 500 ps NPT equilibration and then a 2 ns NPT production run. Snapshots were stored every 1000 steps during the production run (in total 2000 snapshots were collected). Afterwards the IOD and CN values were determined based on these snapshots and the analysis method described in Method 1.

Method 3 was used to simulate the HFE, IOD and CN values of monovalent, trivalent and tetravalent ions. Same as Method 2, we obtained the HFE value based on the free energy changes associated with four processes: ΔG_{VDW} , ΔG_{Ele} , $-\Delta G_{Ele}$ and $-\Delta G_{VDW}$. At first, a dummy atom was solvated in a 13 Å thick water box with the closest water molecule at least 1.5 Å away from it (which offers a cubic box with a length of 29 Å, and a total VDW size of $\sim(32\text{Å} \times 32\text{Å} \times 32\text{Å})$). There are 721, 721 and 732 water molecules in the TIP3P, SPC/E and TIP4P_{EW} water boxes, respectively. Then 1000 steps of steepest descent minimization followed by 1000 steps of conjugate gradient minimization were performed to minimize the initial structure. Afterwards a 500 ps simulation in the NVT ensemble was carried out to heat the system from 0 to 300 K. And then a 500 ps simulation in the NPT ensemble was performed under 300 K and 1 atm to correct the system density and further equilibrate the system. The final snapshot was used as the initial structure for the TI simulations in the NPT ensemble. Afterwards, the HFE value was obtained in the same way as described in Method 2 except that we used a 4-window linear soft-core simulations to obtain the ΔG_{VDW} value (in which the final window was not considered in the free energy calculation but was just used to further equilibrate the system). For determination of the IOD and CN values, we used the same method as described in Method 1, except that the snapshots were stored every 500 fs in the final production run, yielding 4000 snapshots for analysis.

Fitting procedure of the HFE quadratic fitting curves:

- (1) There are 7 different ϵ values for each row in Table SI.1 (SI means supporting information), we performed the quadratic fitting and got the equation for HFE *versus* $-\log(\epsilon)$ for each row with a certain $R_{\min}/2$ value.
- (2) Solve the equation to determine the $-\log(\epsilon)$ value to reproduce the target HFE of a metal ion with the fixed $R_{\min}/2$ for each row. Then for each row we obtained a point $(R_{\min}/2, -\log(\epsilon))$ which can reproduce the target HFE. In total there are 23 rows so we can get 23 points, each of which can reproduce the target HFE.
- (3) Discarded the $(R_{\min}/2, -\log(\epsilon))$ points which were not in the scanned LJ space (with $R_{\min}/2$ in 0.3-2.5 Å and $-\log(\epsilon)$ in 0-6) and performed another quadratic fitting of $-\log(\epsilon)$ *versus* $R_{\min}/2$ for the remaining points to get the final fitting curve for the target HFE.

Fitting procedure of the IOD quadratic fitting curves:

- (1) There are 7 different ϵ values for each row in Table SI.2, we performed the quadratic fitting and got the equation for IOD *versus* $-\log(\epsilon)$ for each row with a certain $R_{\min}/2$ value.
- (2) Solve the equation to determine the $-\log(\epsilon)$ value to reproduce the target IOD of a metal ion with the fixed $R_{\min}/2$ for each row. Then for each row we obtained a point $(R_{\min}/2, -\log(\epsilon))$ which can reproduce the target IOD. In total there are 23 rows so we can get 23 points, each of which can reproduce the target IOD.
- (3) Discarded the $(R_{\min}/2, -\log(\epsilon))$ points which were not in the scanned LJ space (with $R_{\min}/2$ in 0.3-2.5 Å and $-\log(\epsilon)$ in 0-6) and performed another quadratic fitting of $-\log(\epsilon)$ *versus* $R_{\min}/2$ for the remaining points to get the final fitting curve for the target IOD.

$\log(\epsilon)$ versus $R_{\min}/2$ for the remaining points to get the final fitting curve for the target IOD.

OpenMM^{146, 147} (in version 6.1) was used for the ionic solution simulations. 13 Na⁺ ion and 13 Cl⁻ ions were solvated in a water box with a size of $\sim(46\text{\AA}\times46\text{\AA}\times46\text{\AA})$, in which there were 2414 SPC/E water molecules (giving a molarity of ~0.3 M). Firstly, 2000 steps of minimization employing the L-BFGS algorithm was performed to optimize the structure, then 1 ns simulation in the NVT ensemble was performed to heat the system with temperature increasing gradually from 0 to 300 K. Afterwards another 1 ns of simulation was performed in the NVT ensemble to equilibrate the system. And then 2 ns of simulation in the NPT ensemble was carried out to equilibrate the system. Finally 10 ns of production run was performed in the NPT ensemble with structures stored for each 0.5 ps, yielding 20000 snapshots for the final analysis. 10 Å cut-off was used for the nonbonded interaction. The PME method and PBC were used in these MD simulations and the Langevin algorithm with a 1.0 ps⁻¹ friction coefficient was employed for the temperature control. A MC barostat was utilized for the pressure control with the volume change attempt frequency set as 25 fs⁻¹. The time step was set to 1.0 fs with a “three-point” SHAKE algorithm¹⁴² was used to constrain the water geometry. The Kirkwood-Buff integrals were calculated based on equation 12, then the activity derivatives a_{cc} were obtained from equation 13.

$$G_{ij} = \int_0^\infty 4\pi r^2 [g_{ij}(r) - 1] dr \quad (12)$$

$$a_{cc} = \frac{1}{1 + \rho_c(G_{cc} - G_{cw})} = \frac{1}{1 + N_{cc} - \frac{\rho_c \times N_{cw}}{\rho_w}} \quad (13)$$

Herein $g_{ij}(r)$ is the RDF between species i and j . The cut-off of the RDF was set to 12 Å and dr was treated as 0.01 Å for the RDF calculation. ρ_c and ρ_w are the number densities of the ion (here the positive and negative ions are treated as indistinguishable) and water respectively, while N_{cc} and N_{cw} are the excess CNs between the ion-ion and ion-water pairs.

For the simulations of a Fe(III) containing protein, the AMBER 12¹⁴⁰ and AmberTools¹⁴⁰ suites of programs were used for the system building, structure minimizations, MD simulations and data analysis. The structure of Chain C of PDB entry 4BV1 was used for our modeling. The H++ web server¹⁴⁸ was utilized to add hydrogen atoms to the protein system with setting the pH, salinity, internal dielectric constant, and external dielectric constant as 7.2, 0.15, 4 and 80, respectively. Different names of the His residues were assigned according to their protonation states. Afterwards the Cys residue that binds to the iron ion was renamed to “CYM” with its sulfur linked hydrogen atom was deleted. ACE and NME groups were used to cap the protein system. A TIP3P⁶⁵ water box with thickness of 10 Å was employed to solvate the protein system. Afterwards the minimizations and MD simulations were performed as follows:

- (1) 2000 steps of steepest descent minimization, plus 3000 steps of conjugate gradient minimization were performed with the protein system (excluding the capped residues) and metal ion being held by a force restraint of 500 kcal/mol•Å⁻².
- (2) 2000 steps of steepest descent minimization followed by 3000 steps of conjugate gradient minimization were carried out for the system with a 500 kcal/mol•Å⁻² force

constant on the heavy atoms of the protein (including the capped residues) and metal ion.

- (3) 10000 steps of steepest descent minimization and afterwards 10000 steps of conjugate gradient minimization were performed for the system using a $200 \text{ kcal/mol} \cdot \text{\AA}^{-2}$ restraint on the backbone C, CA and N atoms of the protein (including the capped residues).
- (4) 5000 steps of steepest descent minimization and then 40000 steps of conjugate gradient minimization were carried out for the entire system.
- (5) 400 ps of simulation using the NVT ensemble was performed to heat the system from 0 to 300 K with the protein system (including the capped residues) and metal ion having a force restraint of $10 \text{ kcal/mol} \cdot \text{\AA}^{-2}$.
- (6) 200 ps of simulation using the NVT ensemble was carried out to equilibrate the system at 300 K.
- (7) A 2 ns simulation using the NPT ensemble was performed under 300 K and 1 atm to correct the density and further equilibrate the system.
- (8) Finally, 10 ns of production simulation was performed using the NPT ensemble at 300 K and 1 atm with snapshots being stored every 2 ps. In total there were 5000 frames collected for the data analysis.

The Langevin algorithm was used to control the temperature with a collision frequency set at 5.0 ps^{-1} . An isotropic position scaling algorithm was employed to control the pressure with a relaxation time of 2.0 ps. The cut-off value of the nonbonded interaction was set to 10 Å. PME was used to handle the long-range electrostatic interactions.⁷⁷⁻⁷⁹ SHAKE¹⁴¹ was employed during the simulation to constrain the bonds involving

hydrogen atoms, while for the water molecules a “three-point” algorithm¹⁴² was employed.

2.3 Validation Tests of Several Available LJ Parameters for M(II) Ions

The general philosophy of parameter design is to make the best compromise estimate for different physical properties at the same time. In the first part of our work, we tested the transferability of LJ parameters, using the PME approach and Method 2, for some previously developed parameters for M(II) ions. The data from these simulations are shown in Tables 5-7: Table 5 is for the LJ parameters of Mg^{2+} , Ca^{2+} and Zn^{2+} found in the AMBER ff99 which is provided in the AMBER Package.¹⁴⁰ In this FF the LJ parameters of Mg^{2+} and Ca^{2+} ions were adopted from Åqvist³⁷ by utilizing the Lorentz-Berthelot combining rules while the LJ parameters of Zn^{2+} ion were obtained from Merz.⁴⁰ Table 6 shows Zn^{2+} LJ parameters designed by Stote and Karplus³⁸ while Table 7 contains the LJ parameters designed by Babu and Lim.⁶⁴ These results suggest that the differences between treating the long-range electrostatics with PME method *versus* other methods cannot be simply overlooked. For example, as shown in Table 5, for the Ca^{2+} LJ parameters in AMBER ff99, there are large differences of the HFE and IOD values of Ca^{2+} ion obtained herein, which are based on the PME method, from the HFE and IOD values determined in Åqvist's earlier work,³⁷ which are based on a different method. The Zn^{2+} LJ parameters in Table 6 don't have big differences (0.02 Å) in the simulated IODs but have significant differences in the simulated HFEs (~70 kcal/mol). The values in Table 7 indicate that Babu and Lim's parameters are shifted by ~45 kcal/mol from the experimental absolute HFEs, while they show good agreement with the PME simulations with respect to reproducing the relative HFEs. However, some metal ions still have notable differences (for example, Be^{2+} has a 10 kcal/mol difference in the relative HFEs

and 0.12 Å difference for the IODs between the two different methods). Based on the results shown above, we decided that it was necessary to design a set of parameters for the M(II) metal ions using the state-of-the-art PME based MD simulations.

It is extremely important to note that we are not condemning the earlier efforts all of which were excellent.^{37, 38, 40} It simply reflects the fact that simulation protocols have evolved to the point where PME is the accepted standard for the treatment of long-range interactions and it is possible to carry out very long MD simulations using this model with, for example, GPU technology.¹⁴⁹⁻¹⁵⁶

2.4 Parameter Space Scanning

We cannot design a satisfactory unified ion parameter set for all popular water models since the force field parameters of respective water models are different (see Table 4). Hence, we performed simulations for different combinations of the $R_{\min}/2$ and ϵ parameters for the TIP3P, SPC/E, TIP4P and TIP4P_{EW} water models, respectively. Based on the HFE values for TIP3P LJ grids (see Table SI.1a), quadratic fitting (the fitting procedure was shown in section 2.2.) was done for each of the 24 M(II) metal ions in the TIP3P water model and the fitting curves are depicted in Figure 3.

Figure 3 indicates similar trends exist with respect to the fitting curves for monovalent ions from the previous work of Joung and Cheatham.³⁹ HFE increases with a decrease of $R_{\min}/2$, and smaller $R_{\min}/2$ value with a large ϵ parameter can yield similar HFE values as a larger $R_{\min}/2$ parameter coupled with a smaller ϵ value. This can be explained by the form of the LJ potential function (see equation 1): since $C_{12} = \epsilon \times R_{\min}^{12}$ and $C_6 = 2\epsilon \times R_{\min}^6$, smaller $R_{\min}/2$ parameter with bigger ϵ value and smaller ϵ parameter with larger $R_{\min}/2$ value yield similar C_{12} and C_6 values.¹⁵⁷ Furthermore, ϵ is directly proportional to C_{12} and C_6 while $R_{\min}/2$ is raised to the twelfth and sixth power in the expression for C_{12} and C_6 , respectively. This is the reason why the HFE is quasi-linearly dependent on $R_{\min}/2$ while its dependency on ϵ is logarithmic.

Figure 3 shows that generally all of the fitting curves have a similar shape but different Y-intercepts for the TIP3P water model. The HFE values of LJ grids for different water

models (see Table SI.1) show that the HFE differences amongst the same LJ parameters within different water models could not be neglected. To clarify the differences of the HFE fitting curves for different water models, we treated the Zn^{2+} ion as an example and illustrated its HFE fitting curves within four different water models in Figure 4. It can be seen that the two 3-site models (TIP3P and SPC/E) show very similar results and are distinctly different from the two 4-site models (TIP4P and TIP4P_{EW}). Meanwhile, the two 4-site water models also showed a remarkable difference from each other. Therefore, we concluded that it is necessary for us to design different parameters for the same metal ions for use with different water models.

Meanwhile, the IOD and CN values of LJ grids for different water models (see Table SI.2) show that the four water models generated very similar IOD and CN values when using the same LJ parameters for the M(II) metal ion. To elucidate the difference in IOD values for the four studied water models, we carried out a standard deviation analysis of the IOD values between each pair of the four water models and display the data in Table 8. Our results indicate the TIP3P and SPC/E water models have nearly the same IOD values with a 0.00 Å systematic difference and 0.01 Å standard deviation with each other when using the same LJ parameters for the metal ions while a similar situation exists between the TIP4P and TIP4P_{EW} water models. Generally, there is a 0.02 Å systematic difference between the 3-site and 4-site water models for the IOD values when the same LJ parameters are utilized for the metal ions. Meanwhile, IOD fitting curves were obtained from Table SI.2 by following the similar procedure as for the HFE fitting curves (see section 2.2). Again, we treated the Zn^{2+} ion as an example and depicted the IOD

fitting curves in Figure 5. It can be seen from the figure that the two 3-site water models share one curve and the two 4-site water models share the other curve (although in the latter instance there is a slight difference), which is in agreement with the standard deviation analysis (see Table 8).

As discussed in the introduction, the nonbonded model can simulate CN switching processes. There are several non-integer CN values in Table SI.2, which suggests there is CN switching occurring during these simulations. As an example, we show CN switching in the MD simulation of a M(II) metal ion with LJ parameters of $R_{\min}/2 = 2.2 \text{ \AA}$ and $\epsilon = 0.1 \text{ kcal/mol}$ in a TIP3P water box in Figure 6. We observe the CN switches between 8, 9, 10, and 11 during the simulation.

A detailed examination of the HFE, CN and IOD values of LJ grids for different water models (see Tables SI.1 and SI.2) indicates that, for each water model, for the parameter combinations corresponding to the same HFE, they have almost the same IOD and CN. These results suggest there is likely a relationship between the HFE, IOD and CN values for the metal ion-water systems, hinting at the strong correlation between various solvation properties, which is consistent with earlier work.¹⁵⁷ Figure 7 shows the HFE and IOD fitting curves for six representative metal ions with different sizes in the TIP3P water model. From Figure 7 we find that the IOD and HFE fitting curves for each metal ion are almost parallel with each other and do not have any intersection points in the investigated range, implying it is hard to find a parameter to reproduce the experimental HFE and IOD values at the same time for these metal ions. At the same time, the figure

also shows that the distance between the HFE and IOD fitting curves of metal ion begins to decrease along with increase of the ion size, which may be due to the simplicity of the nonbonded model. The electrostatic plus LJ potential approximation underestimates the interaction energy of the metal ion and ligating residues at short range, especially when there is a strong charge transfer, polarization or even covalent interaction between them. In this situation, if one wants to reproduce the experimental HFEs, one should have smaller IOD values than the experimental values. Meanwhile, the 12-6 nonbonded model is more appropriate for the monovalent metal ions since polarization and charge transfer effects are likely to decrease (and these ions tend to be mostly ionic in nature), allowing the parameters to be designed to reproduce both the experimental HFE and IOD values simultaneously.³⁹

Therefore, there appears to be no single “perfect” LJ parameter set for the M(II) ions since none are able to reproduce the experimental HFE and IOD values simultaneously in a simulation. Hence, we concluded that it was necessary to design several sets of parameters for these M(II) metal ions to meet different demands. Since our intention is to design LJ parameters for the M(II) metal ions specifically for PME based MD simulations, we only designed parameters for the TIP3P, SPC/E and TIP4P_{EW} water models. The TIP4P water model was modified to produce the TIP4P_{EW} model, which is designed specifically for PME based simulations.⁶⁷ First, by treating the experimental HFE values as the target property we designed the HFE parameter set for each of the three water models. Next, we designed the IOD parameter set to reproduce the experimental IOD values (due to the limited experimental data set for IOD values, only

16 of 24 divalent ions have IOD parameter sets). In the case of the IOD parameter sets, they ended up being the same for the three water models. In the end, we designed the CM (short for compromise) parameter set for the three water models respectively, which is a compromise between the HFE and IOD parameter sets using the experimental relative HFE and the CN values as targets.

Since numerous points exist on the fitting curves capable of reproducing almost the same HFE, IOD and CN values, it is problematic to pick a single point among them to determine the final LJ parameters. Initially, we wanted to do simulations on the solid-state salts of these M(II) metal ions together with anions for the different combinations of LJ parameters, as employed in the protocol of Joung and Cheatham.³⁹ However, it is difficult to find valid and consistent experimental data for the salts containing the M(II) metal ions we are dealing with. Our second approach was to pick the point that is also capable of reproducing the QM calculated interaction energy between the metal ion and one or several surrounding water molecules. Unfortunately, although for some metal ions (such as Ca^{2+}) we could get reasonable results, we could not obtain valid results for most of the ions, especially for the metal ions capable of strong covalent interactions with the surrounding waters. This likely reflects the simplicity of the nonbonded model. Furthermore, it is also difficult to find a standard QM method protocol, which is largely due to the various possible electronic states for some of the metal ions.

Finally, we selected an alternate way to design the 12-6 LJ parameters. The LJ potential, which was first proposed by Sir John Edward Lennard-Jones to represent the interaction

between noble gas atoms in 1924,¹⁵⁸ is remarkably accurate for the noble gases and a very good approximation for neutral atoms and molecules. In the 12-6 LJ function the r^{-12} term represents the interaction caused by Pauli repulsion due to the overlap of the molecular orbitals at close distance. The r^{-6} term describes the long-range attraction between molecules due to the dispersion force. Generally, the more dispersive electronic cloud one particle has, the bigger $R_{\min}/2$ and ϵ value it should have, as shown in the experimental data.¹⁵⁹ Using the experimental data¹⁵⁹ and the Lorentz-Berthelot combining rules, we obtained the $R_{\min}/2$ and ϵ parameters for the He, Ne, Ar, Kr, and Xe atoms (see Table 9). For all of the metal ions treated here, they should have smaller $R_{\min}/2$ and ϵ values than those of Xe since the biggest metal ion herein, Ba^{2+} has a smaller $R_{\min}/2$ than Xe because of the same electronic structure but a larger nuclear charge. Furthermore, if one metal ion has a $R_{\min}/2$ value between the $R_{\min}/2$ values of the Kr and Xe atoms, it should have a ϵ value between the ϵ values of the Kr and Xe atoms as well. Furthermore, if we could get a curve to represent the relationship between $R_{\min}/2$ and ϵ , together with the HFE and IOD fitting curves we obtained in the former part, we could determine the LJ parameters for the M(II) metal ions.

To be consistent with the HFE and IOD fitting curves, we produced the curve fitting between the $-\log(\epsilon)$ and $R_{\min}/2$ values shown in Table 9. By treating $R_{\min}/2$ as x and $-\log(\epsilon)$ as $f(x)$, we attempted several fits with different functions. Finally, we found that the Slater function $f(x) = C_1 \times e^{-C_2 x}$ with $C_1=57.36$ and $C_2=2.471$ had a better R^2 value (0.98265) than the quadratic fitting ($R^2=0.94509$). We named this curve as the noble gas curve (NGC) and determined the final LJ parameters from the points on this curve.

It was found that different combinations of parameter pairs could reproduce the same HFE value in both Joung and Cheatham's work³⁹ and our work described above. In light of this and to balance the accuracy and speed, we constrained one of the two parameters ($R_{\min}/2$ and ϵ) in a physically reasonable manner and then vary the other one during the parameter space scanings for the monovalent, trivalent and tetravalent ions. To accomplish this we used NGC described above to select an appropriate ϵ value followed by a one-dimensional scan over a series of $R_{\min}/2$ values.

For the monovalent cations, we performed the parameter scanning using the Na^+ ion (which has a mass of 22.99 g/mol) as our reference. We only carried out the scanning at one mass since the atomic mass has only a small effect on the computed thermodynamic and structural properties. We scanned $R_{\min}/2$ from 0.8 to 2.3 Å with an interval of 0.1 Å. For the monovalent anions, we did the parameter scanning using the Cl^- ion (which has a mass of 35.45 g/mol) as our reference. The scanning range of $R_{\min}/2$ is from 1.7 to 3.2 Å with an interval of 0.1 Å. A range of larger $R_{\min}/2$ values was chosen for the anions because that, in general, they have more dispersed electronic clouds than the cations. The parameters, and the simulated HFE, IOD and CN values of these parameter scanings are shown in Tables SI.3 and SI.4. For trivalent and tetravalent metal ions, we performed the parameter space scanning similar to the monovalent ions. We used the Fe^{3+} (mass of 55.85 g/mol) and Th^{4+} (mass of 232.04 g/mol) ions as our references for the parameter scanings of the trivalent and tetravalent metal ions, respectively. We scanned the $R_{\min}/2$ from 0.9 to 2.3 Å with an interval of 0.1 Å while the ϵ value is obtained for each $R_{\min}/2$

value based on NGC. Parameter points, and the simulated HFE, IOD and CN values are collected in Tables SI.5 and SI.6.

2.5 Parameter Determination

The HFE, CM and IOD fitting curves for the Zn^{2+} ion and the NGC are shown in Figure 8. It can be seen that the CM fitting curve for the Zn^{2+} ion is almost in the middle of the HFE and IOD fitting curves. The original LJ parameters can be obtained as the intersection points between the HFE, CM and IOD fitting curves with the NGC. After slightly tuning the parameters, the final LJ parameters can be determined. We employed Method 2 in this part, which is a more accurate way to obtain the HFE, IOD and CN values. In the VDW-disappearing and VDW-appearing steps we employed the soft-core scaling method instead of the linear or nonlinear normal scaling methods due to its better performance over the latter two.⁸¹ In the present work, we also conducted tests among the different scaling methods and the data is given in Table 10. The L, K4, K6 and SC in Table 10 represent the linear scaling, nonlinear scaling with $k=4$, nonlinear scaling with $k=6$ and soft-core scaling methods, respectively, while all the windows involved a 300 ps simulation with the last 200 ps used to determine the free energy changes. It can be seen that the soft-core scaling method gives better-converged and consistent results (*i.e.*, the free energies of the VDW-appearing and VDW-disappearing processes have the opposite sign) than the other methods. The linear and nonlinear scaling methods could give consistent results for the small Be^{2+} ion, but had more difficulty with larger ion like Ba^{2+} ion.

The HFE parameter set of divalent metal ions is shown in Table 11, and its simulated HFE, IOD and CN values are shown in Table 12. The HFE parameter set achieved a ± 1

kcal/mol accuracy of reproducing experimental HFEs. The IOD parameter set, and its simulated HFE, IOD and CN values are shown in Tables 13-14, respectively. Generally, the IOD parameter set reproduced the experimental IOD values with an accuracy of ± 0.01 Å. The CM parameter set, and its simulated HFE, IOD and CN values are listed in Tables 15-16, respectively. During our parameterization efforts we found that it was impossible to simulate all the CN values while simultaneously reproducing the relative HFE values for the CM parameter set so we compromised on the reproduction of the CNs for the Be^{2+} and Sn^{2+} ions and tried to best reproduce their relative HFE values. This led to the CM parameter set having an average error of ~ 25 kcal/mol in the absolute HFE (while reproducing the relative HFE) for the TIP3P and SPC/E water models, while for the TIP4P_{EW} water model this value increased to ~ 40 kcal/mol. Generally the CM parameter set reached a ± 2 kcal/mol accuracy of the relative HFEs while keeping the CNs of most M(II) metal ions. Here we treated the Zn^{2+} ion as an example again and showed the RDFs for the different parameter sets in Figure 9. It could be seen that the CM parameter set yields a first solvation peak between those of the other two parameter sets. Meanwhile, the IOD and CN values are 1.67 Å with 4.1, 1.93 Å with 6.0, and 2.08 Å with 6.0 for the HFE, CM, and IOD parameter sets respectively.

We performed curve fits for HFE *versus* IOD along the scanning range for $R_{\text{min}}/2$ for both the monovalent cations and anions (see Figure 10). The target values are also shown in the figure as blue stars with error bars. There are different phenomena that have been observed for the positive and negative ions. For the positive ions, the error of the 12-6 LJ nonbonded model is relatively small. For example, from Figure 10, we can see that it is

possible to reproduce the target HFE and IOD values simultaneously for each of the Na^+ , K^+ , Rb^+ , Cs^+ and NH_4^+ ions. Ramaniah *et al.* analyzed the K^+ -aqueous system using *ab initio* MD and also demonstrated that classical potentials can yield good predictions for this system.¹⁶⁰ While for the Li^+ , Tl^+ , Cu^+ , Ag^+ and H_3O^+ ions, the error of the 12-6 model is relatively large because of their stronger charge-induced dipole interactions with the surrounding water molecules. The proton is not shown in the figure because it is out of range. From the discussion below, we note a 60-90 kcal/mol difference between the HFE value of the IOD parameter set and the experimental HFE value for the proton. We did not design a CM parameter set for the monovalent ions due to the relatively small error of the 12-6 model that for almost half of the positive ions it is possible to reproduce the IOD values using the HFE parameter set. For the monovalent anions, the TIP3P water model showed the most consistency with the experimental values when employing the 12-6 LJ nonbonded model. Meanwhile, it is intriguing that the 12-6 nonbonded model overestimated the HFEs of the halide ions. It is hard to reproduce both the experimental HFE and IOD values simultaneously and this may be due to the charge hydration asymmetry (CHA) effect, which is further discussed below.

The parameter sets developed herein attempt to reproduce certain target values through a trial and error process. Based on the results of our parameter scans with the 12-6 LJ potential, we have fit curves of HFE *versus* $R_{\text{min}}/2$ for the monovalent positive and negative ions for the three water models (see Figure 11). For the HFE fitting curves of the positive ions, the TIP3P and SPC/E water models are very close with each other with TIP4P_{EW} being a little further away. This is consistent with our work on divalent metal

ions. Intriguingly, for the negative ions, the SPC/E and TIP4P_{EW} water models are similar while the TIP3P water model gives more positive HFE values for the ions with the same LJ parameters. This may also come from the CHA effect (discussed below). Since there is a nontrivial difference between the three water models, we performed parameterizations of the monovalent ions for these three water models separately to reproduce the target HFE values. The HFE parameter set is shown in Table 17. Its simulated HFE, IOD and CN values are given in Table 18. These parameters reproduce HFEs within ± 1.0 kcal/mol of the target values.

We also analyzed the IOD values *versus* $R_{\min}/2$ values for the monovalent cations and anions with the three water models (see Figure 12). The three water models nearly share the same curve in the figure, for both the cations and anions. This is also consistent with our work on divalent metal ions. The only difference between the cations and anions is that TIP4P_{EW} predicts slightly larger IOD values for the cations but slightly smaller IOD values for the anions than TIP3P and SPC/E do when using the same LJ parameters. Therefore, we designed a united IOD parameter set for all three water models. This parameter set is shown in Table 19 while the simulated HFE, IOD and CN values are shown in Table 20. The final IOD parameter set reproduced the target IOD values within ± 0.01 Å for most of the investigated ions in all of the three water models.

In general, our parameter sets reproduce the target CN values with reasonable accuracy. The largest errors are found for the Ti^+ , Cu^+ , Ag^+ and NH_4^+ , H_3O^+ ions, which are ions that do not share electronic structures with the noble gas atoms. The electronic structure

of the outermost shell of Ti^+ , Cu^+ and Ag^+ are $6s^2$, $3d^{10}$ and $4d^{10}$ respectively. While for the NH_4^+ and H_3O^+ ions, it is hard to reproduce their hydrogen bond network because the current model doesn't parameterize the hydrogen atoms explicitly. For the Ti^+ ion, the two different IODs were not reproduced in the present parameter sets due to the isotropic nature of the nonbonded model. For the Cu^+ ion, we could not reproduce the target CN value (which is 2) with any of the solvent models. In all cases the CN of the Cu^+ ion is 4. A similar situation has found for the Ag^+ ion where the target CN value ranges from 2 to 4 while our parameter sets give CNs from 4.8 to 6.0. Blauth *et al.* predicted the CN value of Ag^+ as 6.0 in the aqueous system using the QMCF MD simulations¹³⁴ while Blumberger *et al.* simulated this value as 4 using the CPMD method.¹³⁵ It is hard to obtain the experimental CN value for the hydronium ion, but it has been proposed that the CNs of water and the hydronium ion are similar to each other.¹²³ The HFE, IOD, and 12-6-4 parameter sets (in which the 12-6-4 parameter set is introduced in section 3) predict the CN value of the H_3O^+ ion in the range of 4.7-5.2, 7.0-7.1, and 8.2-8.4, respectively. The former one is consistent with the experimental CN value of water (determined as 5.2 from Soper¹⁶¹) while the later two CN values appear somewhat overestimated. The CN value of the NH_4^+ ion is in the range of 4-11 from the review of Ohtaki and Radnai,¹⁶² but a CN value around 4-5 seems more reasonable based on the consideration of its hydrogen bond network. Our parameter sets predicted this CN in the range of 6.8-7.9 while the CPMD simulations done by Brugé *et al.* predicted the CN as 5.3.¹³¹

Based on the quadratic fittings of the data points from the parameter scanning for the M(III) and M(IV) ions, we estimated their HFE and IOD parameter sets. The two

parameter sets are shown in Tables 21-22. Due to the large error of the 12-6 LJ nonbonded model for these highly charged ions, we did not design the CM parameter set for them because it would yield huge errors for both the HFE and IOD properties.

2.6 Parameter Assessment

In the parameter space, it is possible to reproduce the same properties using different combinations of the $R_{\min}/2$ and ϵ values. This is because different combinations of $R_{\min}/2$ and ϵ can generate similar C_{12} and C_6 values for a certain pair of atom types. However, a poor balance between the $R_{\min}/2$ and ϵ values can cause transferability issues since all of the other C_{12} and C_6 values need to be calculated based on the combining rules for the mixed system. In this way, the balance of the two parameters plays a significant role in parameter transferability.

Our final parameter sets are consistent with each other since they were built in an analogous manner. In particular, they show a clear trend that when an ion has a bigger VDW radius, it has a deeper well depth. In the work of Peng *et al.*, they proposed the following relationships between the VDW parameters of the ions which share the same electronic structure with noble gas atoms:

- (1) The ion radii and well depth should increase for the ions inside a group;
- (2) The ion radius of the negative ion should be bigger than the neutral atom then the positive ion, which both have the same electronic configuration;
- (3) The dispersion term for the negative ion should be larger, comparing that of the neutral atom then that of the positive ion, which both have the same electronic configuration, due to the extent of the electron cloud.⁶¹

Our parameters follow these trends precisely. Also, our parameter set follows the trend that when two cations share the same electronic structure, the cation with a larger charge

has a smaller VDW radius. For example, the HFE parameter set for the TIP3P water model offers $R_{\min}/2$ parameters of the Na^+ , Mg^{2+} , and Al^{3+} ions as 1.475, 1.284, and 0.981 Å, respectively.

Besides follow the physical trends, our parameters are also quantitatively agreed with experimental results and theoretical calculations based on the QMSP method. For example, experimental results show that the F^- ion has smaller ionic radius than K^+ ion¹²⁵ and the IOD parameter set also shows this trend. In Table 23 we compared the $R_{\min}/2$ values of 15 ions, which share the same electronic structures with the noble gas atoms, in the HFE parameter set to the VDW radii calculated by Stokes using the QMSP method.⁸⁸ Generally they are in strong agreement with each other (with the smallest, biggest, and average absolute percentage errors as 0.0%, 9.1%, and 4.0%, respectively). For example, for the three M(III) ions (Al^{3+} , Y^{3+} and La^{3+}), the estimated $R_{\min}/2$ values in the estimated HFE parameter set for the TIP3P water model are 0.981 Å, 1.454 Å, and 1.628 Å, respectively. The calculated VDW radii are 1.046 Å, 1.481 Å and 1.642 Å respectively based on the QMSP method. There are only 6.2%, 1.8%, and 0.9% differences between the two value sets for the three trivalent ions respectively. This further validates the physical consistency of our parameterization work.

Meanwhile, the $R_{\min}/2$ values of our parameter sets are more consistent with the VDW radii calculated by the QMSP method than the parameter sets presented elsewhere. In Table SI.7 we compare our parameters of Na^+ , K^+ , Rb^+ , Cs^+ , F^- , Cl^- , Br^- , and I^- ions to the parameter sets developed by other groups (where we adapted their parameter sets for the

Lorentz-Berthelot combining rules). The average errors, standard deviations of the errors, and UAEs (using the QMSP VDW radii as reference values) of each parameter set were calculated and are given in Table SI.7. We show the UAEs of the different parameter sets in Figure 13. All our parameter sets (in which the 12-6-4 parameter set is introduced in section 3) show the smallest UAEs (~ 0.08 Å) relative to earlier parameter sets. The metal ion parameters available in AMBER, which were adapted from the pioneering work of Åqvist where he fit the ions using a cut-off procedure (this was done prior to the emergence of the PME method), show the largest UAE. In this parameter set there is an imbalance between C_{12} and C_6 values, in which the C_6 value is too small while the C_{12} value is closer to our values (based on the relative ratio). Because of this the transferability of the parameter set will be affected since $R_{\min,ij}$ is overestimated due to the too small C_6 value (recall $R_{\min,ij} = \sqrt[6]{\frac{2 \times C_{12,ij}}{C_{6,ij}}}$). For example, this causes the alkali metal ions to have an inverse trend amongst the ϵ parameters. The Jensen parameter set has a relatively large UAE due to the choice of the ϵ value for the positive ions (which is too small). The UAE of the VDW radii of the positive ions is ~ 1.08 Å for this parameter set while the corresponding UAE of the negative ions is only ~ 0.08 Å. This is because the ϵ value of the positive ions (which is 0.0005 kcal/mol) is too small while the ϵ value of the negative ions (which is 0.71 kcal/mol) is more reasonable. The large UAE value for the parameter set 5 of Netz and co-workers (2009_HMN-5) is due to the small ϵ value for the positive ions (~ 0.0006 kcal/mol).⁶⁸ The parameter set from Joung and Cheatham has too small ϵ values for the halide ions, for example, the ϵ value of Na^+ is always bigger than that of F^- (~ 30 to 100 times).³⁹ Dang's parameter set also shows small UAE value

because its ϵ values are in the range of 0.1-0.2 kcal/mol,⁵³⁻⁶⁰ which are reasonable for the monovalent ions.

Moreover, the $R_{\min}/2$ values in our 12-6 LJ parameter sets could be used as the VDW radii for the RESP charge fitting procedure. For example, in the work of Kuznetsov *et al.*,¹⁶³ they used 1.4 Å as the VDW radius for the RESP charge fitting for both the Fe^{2+} and Fe^{3+} ions. Herein, the $R_{\min}/2$ of Fe^{2+} was determined as 1.409 Å in the IOD parameter set, and the IOD parameter set for the Fe^{3+} ion estimated the $R_{\min}/2$ as 1.386, 1.386, and 1.375 for the TIP3P, SPC/E, and TIP4P_{EW} water models, respectively.

To further validate our parameter sets, we have carried out MD simulations of different ion pair solutions. The Kirkwood-Buff (KB) integrals were calculated to investigate the ionic solution systems. Finally the activity derivatives were computed to evaluate the parameter sets. In the work of Mouka *et al.*,¹⁶⁴ they found that the parameter set from Joung and Cheatham³⁹ (the JC parameter set herein) and the parameter set 5b from Horinek *et al.*⁶⁸ (the H2 parameter set herein, the parameters were adapted using the Lorentz-Berthelot combining rules and are shown in Table SI.8) are the best among 13 different parameter sets for simulating the NaCl solutions when using the SPC/E water model. In the work of Fyta and Netz, they re-optimized the mixing rules of the ion pairs to reproduce the activity derivatives of several different kinds of ionic solutions.⁷⁵ They found it was hard to reproduce the experimental activity derivatives of the NaF, KF, NaI, and CsI salt solutions. When compared to the values obtained using the un-optimized (standard) Lorentz-Berthelot combining rules, they found that a larger ϵ_{ij} value for NaF, a

larger $R_{\min,ij}$ value for KF, and a smaller ϵ_{ij} value for NaI and CsI were necessary to reproduce the experimental activity derivatives.

In the present work, we have carried out simulations on NaCl, KCl, NaF, KF, NaI, and CsI salt solutions at a concentration of ~ 0.30 M. The SPC/E water model was used in these simulations. The simulation procedure is detailed in the section 2.2. Comparison between our parameter sets (HFE, IOD and the 12-6-4 parameter sets for the SPC/E water model, where the 12-6-4 parameter set is introduced in section 3 below) with the JC and H2 parameter sets were made and are shown in Table 24. From Table 24 we can see that the average errors of the parameter sets follow the sequence $12-6-4 < HFE \sim IOD < JC < H2$ while the UAEs follows the trend that $12-6-4 < IOD < JC < HFE < H2$. In general our three parameter sets showed improved results over the JC and H2 parameter sets, indicating their superior transferability.

Meanwhile, we also performed simulations of a metalloprotein system. PDB entry 4BV1 was used to obtain the starting coordinates for this modeling exercise. It is a superoxide reductase (SOR) found in *Nanoarchaeum equitans*. It is a protein tetramer with each monomer has a metal site containing a Fe^{3+} ion. The structure has been determined by using X-ray crystallography to a resolution of $1.90 \text{ \AA}$. The tetramer structure is shown in Figure 14 while Chain C with its metal site is shown in Figure 15. The metal site contains 4 His groups, 1 Cys group, and 1 water molecule. By treating Chain C as the initial structure, we performed three simulations with different parameter sets (the HFE, IOD, and 12-6-4 parameter sets, in which the 12-6-4 parameter set is introduced in section 3).

The TIP3P water model was employed during the simulations. Details of the simulation procedure are given in section 2.2. Totally 10 ns production was performed during the simulation and snapshots were stored every 500 fs. The HFE parameter set prefers a smaller CN (of 4) and the metal ion moves out from the binding pocket, while stable metal complex structures were obtained for the simulations using the IOD and 12-6-4 parameter sets.

A RMSD analysis was performed over the protein backbone CA, C, N atoms and the metal site (metal ion plus the ligating residues and the binding water molecule) heavy atoms for the simulations by treating the initial structure (experimental structure) as reference. The results are depicted in Figure 16. The RMSD of the protein backbone CA, C, N atoms fluctuated around ~ 1.2 Å while the RMSD of the metal site heavy atoms is ~ 0.5 Å. These values illustrate that the metal binding site is stable during the simulations.

We have also performed an RMSF analysis of the heavy atoms for the protein residues together with the oxygen atom in the metal site binding water. The results are shown in Figure 17. From this figure it can be seen that the metal site residues: residue His10 (residue number 11), His 35 (residue number 36), His 41 (residue number 42), Cys 97 (residue number 98) and His 100 (residue number 101) all have relatively small RMSF values (~ 0.5 Å) for their heavy atoms. The oxygen atom of the metal site binding water (residue number 115, is not shown in the figure since the protein ends at residue number 112) has RMSFs of ~ 0.6 Å for the simulations using the IOD and 12-6-4 parameter set. These results further validated that the metal ion site is stable during the simulations.

2.7 Error Analysis

To further assess the final parameter sets, we performed error analysis. The resultant values of monovalent ions are shown in Table SI.9. Using the TIP3P water model as an example, we depict the error analysis results of monovalent ions in Figures 18-19. For the positive ions, the 12-6 LJ nonbonded model largely reproduces both the experimental HFE and IOD values for Na^+ , K^+ , Rb^+ , Cs^+ and NH_4^+ ions while for the Li^+ , Tl^+ , Ag^+ , H^+ and H_3O^+ ions, the 12-6 LJ nonbonded model yields larger errors. We also analyzed the percentage errors with respect to the experimental HFE and IOD values for the parameter sets of M(II) ions. These results are summarized in Table SI.10. From these data we can estimate the maximum error ranges. The values given in square brackets in Table SI.10a are for the IOD set, while the unbracketed values given in Table SI.10b are for the HFE parameter set, with the former indicating the error in HFEs (*e.g.* ~18% in TIP3P for Be^{2+}) if we get the IODs correct, while the latter is the error we see in the IOD values (*e.g.* ~30% in TIP3P for Be^{2+}) if we get HFEs correct. Hence, these values indicate the maximum error range associated with the modeling of M(II) cations using an unpolarized nonbonded model. Moreover we observe the trend that the TIP3P, SPC/E, TIP4P_{EW} and TIP4P water models have increasing errors successively. For all four water models, ions like Be^{2+} , Cu^{2+} and Zn^{2+} have larger errors presumably due to their strong coordination interaction with the surrounding waters. The alkaline-earth metal ions, except for Be^{2+} , have the smallest errors likely due to their preference to form ionic bonds. Meanwhile, the estimated absolute and percentage errors of different parameter sets for highly charged ions were shown in Table SI.11. We can see that there are significant errors

associated with the 12-6 LJ nonbonded model for highly charged ions, we did not carry out further refinement work on the estimated 12-6 LJ parameters since the resultant parameters would be of limited usefulness.

Furthermore, we have summarized the (estimated) average errors and (estimated) average percent errors for the HFE and IOD parameter sets for the mono-, di-, tri- and tetravalent cations in Tables 25-26. Again using the TIP3P water model as an example we graphically summarize our results in Figures 20-21. From these tables and figures we observe that, not surprisingly, the average error of the monovalent cations are significantly smaller than that of the cations with higher oxidation states. The TIP3P and SPC/E water models show very similar results while the TIP4P_{EW} water model shows the biggest deviation. For example, for the TIP3P water model, the average error of the IOD values for the HFE parameter set increases from -0.14 Å for M(I), to -0.27 Å for M(II) ions, -0.29 Å for M(III) ions and -0.58 Å for M(IV) cations. While the average error goes from ~20 kcal/mol for M(I), to ~50 kcal/mol for M(II) ions, ~80 kcal/mol for M(III), and ~240 kcal/mol for M(IV) ions for the IOD parameter set. These results, taken as a whole, shows that the underestimation the 12-6 LJ nonbonded model increases dramatically as the charge on the cation increases.

In present work we used the experimental HFEs of cations from Marcus' data set. These data were obtained based on the conventional HFEs of cations, and the HFE of proton, which was treated as -1056 kJ/mol (~ -252 kcal/mol) by Marcus, using the following equation (where Q is the ion's formal charge):

$$\Delta_{hyd}G^0 = \Delta_{hyd}G_{conv}^0 + Q \times \Delta_{hyd}G_{proton}^0 \quad (14)$$

Other values for the HFE of proton in the general range of -250 to -265 kcal/mol (see section 2.1) can also be used, which would be more likely to decrease the experimental HFEs used herein and then increase the underestimation extent of the 12-6 model for cations. This is because Marcus used a value close to the upper band (-250 kcal/mol).

The error of the 12-6 model for the anions follows the opposite trend from the cations: the 12-6 model overestimated rather than underestimated the ion-water interaction for the halide anions. Which, we proposed, is due to the CHA effect. Rajamani *et al.* explained that the CHA effect arises from two effects: 1) a positive electric potential is induced by the surrounding water molecules on the surface of a neutral solute, so more energy will be released when changing the neutral solute into a negatively charged one; 2) water has different packing and orientations on the surface of positively and negatively charged solutes.¹⁶⁵ Onufriev and co-workers proposed that the CHA effect stems from the asymmetric electric multipole components of water molecules.¹⁶⁶ Later on their group developed a charge-asymmetric Born equation based on a statistical point of view, which incorporates the CHA effect efficiently.¹⁶⁷ For aqueous systems containing ions, different atoms of the water molecule are coordinated to the counter-ions due to their different electrical properties. It is the oxygen atoms, which have a partial negative charge, greater mass, and most of the electronic cloud coordinated to the positive ion. The opposite situation exists for the negative ions: it is the hydrogen atoms of the water molecules that orient towards the negative ions.

The CHA effect can be well reproduced by using the parameters developed herein. Using the K^+ and F^- ions as an example: they have relatively similar effective ionic radii (1.38 Å and 1.33 Å for the 6-coordinated K^+ and F^- ions, respectively),¹²⁵ while the HFE difference between them is ~41 or ~49 kcal/mol based on Marcus' or Schmid *et al.*'s data sets. In the IOD parameter set, they almost have the same $R_{min}/2$ parameters (1.745 and 1.739 Å for K^+ and F^- ions respectively), while the computed HFE differences between them are ~54, ~61 and ~65 kcal/mol for the TIP3P, SPC/E and TIP4Pew water models, respectively. This is consistent with previous work that shows the TIP3P, SPC/E, TIP4P water models have an increasing CHA effect.^{165, 166} In Table 20 we show the IOD values for the IOD parameter set for the three water models. As we can see, unlike the positive ions, the IOD value sequence is $TIP3P > SPC/E > TIP4P_{EW}$ for the negative ions. This is opposite to the sequence of the hydrogen charge on these water models, which are +0.417e, +0.4238e and +0.52422e for the TIP3P, SPC/E and TIP4P_{EW} water models, respectively. Hence, on one hand, for the positive ions, it seems the dipole moment of a water molecule plays a dominant role for the solvation properties of the ions: the TIP3P and SPC/E water models showed similar performance with each other and their dipole moments are all ~2.35 D, while the TIP4P_{EW} water model showed a bigger underestimation of the ion-water interaction and it has a dipole moment of ~2.32 D. On the other hand, for the negative ions, it appears that the charge of the hydrogen atoms is a dominant factor for the simulated properties of these ions.

To further clarify the CHA effect for the different water models employed herein, we have depicted the CHA effect using the data from our parameter scans and show the

results in Figure 22. In this figure we observe that the CHA effect of the negative ions has smaller IOD and more negative HFE values when compared to the positive ions employing the same LJ parameters. It is intriguing that this effect decreases considerably for the HFE values but not the IOD values as the ion's VDW radius increases.

2.8 Conclusions

First, we tested the transferability of LJ parameters determined in previous work and found that it was necessary to design new parameters for M(II) metal ions in PME simulations. Systematic studies were performed to determine the LJ parameters for divalent cations using different water models with the Lorentz-Berthelot combining rules in PME simulations. HFEs, IODs in the first solvation peaks as well as CNs were determined for various combinations of LJ parameters by employing TI and MD simulations and using the PME summation method to model long-range electrostatics. The results showed there is a correlated relationship for the simulated HFE, IOD and CN values. A series of curves were obtained using the quadratic fitting procedure towards the target values based on the parameter space scanning results. It was observed that different water models give different HFEs but highly similar structural properties when treating the same metal ions with identical parameters.

Generally, it is hard to reproduce all the target properties of the M(II) metal ions in aqueous solution using the 12-6 nonbonded model due to the model's simplicity, which agrees with the former work of Ponomarev *et al.*⁵⁰ Overall, the 12-6 nonbonded model underestimates the interaction energy between these metal ions and surrounding water molecules and the underestimation extent is water model dependent. More interaction terms other than the LJ and Coulomb potentials should be considered in the FF in order to perform more accurate modeling of the M(II) metal ions. Polarizable FFs^{47, 54, 56, 122, 168-171} and the short-long effective functions (SLEF),¹⁷² which consider short-range interactions

such as polarization and charge transfer effects in more accurate ways, could be promising methods to solve the dilemma of M(II) metal ion parameter design.

Through a consideration of the physical origins of the VDW interaction, we fit a curve from the experimental data of noble gas atoms to represent the relationship between the two parameters in the 12-6 LJ potential. Furthermore, we also investigated the monovalent, trivalent and tetravalent ions. We found that the 12-6 LJ nonbonded model gives rise to relative small errors for the monovalent ions: for almost half of the monovalent cations investigated, it could reproduce the target HFE and IOD values simultaneously. Meanwhile, we found that the 12-6 model yields huge errors for the highly charged ions.

Based on the parameter space scanning results and NGC, we have developed two 12-6 LJ parameter sets (HFE and IOD) for 15 monovalent ions (11 monovalent cations plus 4 monovalent anions) for the TIP3P, SPC/E, and TIP4P_{EW} water models, respectively. Meanwhile, we arrived at three parameter sets (HFE, IOD, and CM parameter sets) for 24 M(II) cations also with each of the TIP3P, SPC/E, and TIP4P_{EW} water models. Furthermore, we also estimated HFE and IOD parameter sets for 24 highly charged metal ions (18 trivalent cations plus 6 tetravalent cations) with the three water models. The HFE parameter set used experimental HFE values as the target; the target property for the IOD parameter set is the experimental IOD values; while the CM parameter set aimed at reproducing the experimental relative HFE and CN values. Generally these parameters accurately reproduced the target properties.

We found our final parameters follow the trends of VDW parameters within ion series. Meanwhile, these parameters are consistent with experimentally or theoretically determined results. Results showed that on average the $R_{\min}/2$ values in our HFE parameter set for the TIP3P water model are only $\sim 4.0\%$ deviated from the VDW radii determined using the QMSP method. We have also compared our final parameter sets with previous sets from a number of research groups. Using the VDW radii calculated based on the QMSP method as the reference values, we find that our parameter sets show the smallest UAEs among all the parameter sets investigated. This is due to a better balance between the $R_{\min}/2$ and ϵ values in our parameter sets through using NGC. Furthermore, we predicted the activity derivatives for six different ion pair solutions and showed that the parameter sets developed herein (HFE, IOD, and 12-6-4, in which the 12-6-4 parameter set is introduced in section 3) reproduced these quantities better than existing parameter sets, validating their improved transferability.

Finally we performed error analysis for the 12-6 model of ions. For some monovalent cations (*e.g.* Na^+ , K^+ , Rb^+ , Cs^+ and NH_4^+), it is possible to reproduce both the experimental HFE and IOD values simultaneously by employing the 12-6 LJ nonbonded model while for others like Ag^+ , Tl^+ the 12-6 model has non-negligible errors. We also investigated the underestimation extents of the HFEs by the nonbonded model for different M(II) metal ions and found the errors are larger for the metal ions that could form stronger coordination interactions (*i.e.*, covalent bonds) with surrounding waters. Moreover, we found that generally, with the increasing charge of the cation there is a notable decrease in the accuracy of the 12-6 LJ nonbonded model. Using TIP3P as an

example, the average underestimation of the HFE values increases from ~ 20 kcal/mol for M(I) cations, to ~ 50 kcal/mol for M(II) cations, ~ 80 kcal/mol for M(III) cations and ~ 240 kcal/mol for M(IV) cations when trying to reproduce the target IOD values. The average underestimation of the IOD values increases from $0.14 \text{ \AA}$, to $0.27 \text{ \AA}$, $0.29 \text{ \AA}$ and $0.58 \text{ \AA}$ for the M(I), M(II), M(III) and M(IV) cations respectively when trying to reproduce the experimental HFE values. The error trend of the 12-6 model for the halide anions was also discussed. The CHA phenomenon has been observed for the anions in the explicit water models we have investigated. Our results indicate that the CHA effect increases along the series of TIP3P, SPC/E, and TIP4P_{EW} with the TIP3P water model having the smallest error when compared to the experimental results.

CHAPTER 3: PARAMETERIZATION OF THE 12-6-4 NONBONDED MODEL

3.1 Model Origin

As discussed in chapter 2, we found that it is impossible to reproduce all the three experimental properties at the same time in simulations using the 12-6 nonbonded model with parameters spanning the typical LJ space when ion has a formal charge $\geq +2$. This is because the 12-6 nonbonded model underestimates the interaction energy between the ion and surrounding water molecules. This underestimation decreases with the increasing of the metal ion's size, and increases along with the ion charge.

FFs are designed to accurately describe the interactions in a complex system using potential functions that are relatively easy to compute and consist of terms representing the bonded and nonbonded interactions. The nonbonded model in typical FFs is composed of an electrostatic and a VDW term. Point charges are obtained in a number of ways, but in the AMBER FF RESP charges¹⁷³ are used and they approximate the charge-charge, charge-dipole and dipole-dipole interactions. The 12-6 LJ potential is used to represent the VDW interaction, which consists of the Pauli repulsion and induced dipole-induced dipole (so called dispersion) interactions. However, as shown in Figure 23, generally there is no term representing the charge-induced dipole and dipole-induced dipole interaction explicitly in the nonbonded model. Furthermore, among these two interactions, the charge-induced dipole interaction is the dominant one in the case of metal ions and it has a potential functional form proportional to r^{-4} . In light of this, we

added a new term into the 12-6 nonbonded model between the charged ion and the surrounding particles to represent the charge-induced dipole interaction. The new potential was given in equation 15 and the parameters that need to be determined are $R_{\min}/2$, ε and κ (or C_4). We determined the final parameters for 55 ions (range from monovalent to tetravalent, see Figure 24) using different water models (TIP3P, SPC/E, and TIP4P_{EW}) by utilizing the experimental HFE, IOD and CN values shown in Tables 1-3 (in terms of the divalent ions, we only designed parameters for the ions with all of the three target values available in Table 2). Details of the process used to design our parameter sets is described in the following section.

$$\begin{aligned}
 U_{ij}(r_{ij}) &= \frac{c_{12}^{ij}}{r_{ij}^{12}} - \frac{c_6^{ij}}{r_{ij}^6} - \frac{c_4^{ij}}{r_{ij}^4} + \frac{e^2 Q_i Q_j}{r_{ij}} = \varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min,ij}}{r_{ij}} \right)^6 \right] - \frac{c_4^{ij}}{r_{ij}^4} + \frac{e^2 Q_i Q_j}{r_{ij}} = \\
 &\varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min,ij}}{r_{ij}} \right)^6 \right] - \kappa \frac{c_6^{ij}}{r_{ij}^4} + \frac{e^2 Q_i Q_j}{r_{ij}} \quad (15)
 \end{aligned}$$

3.2 Methods

When simulating the HFE values using the 12-6-4 model, the $\Delta G_{\text{Ele+Pol}}$ and $-\Delta G_{\text{Ele+Pol}}$ instead of ΔG_{Ele} and $-\Delta G_{\text{Ele}}$ values were obtained (see Figure 2). For the divalent metal ions, a similar method to Method 3 was used to obtain the HFE values, except that a 3-window TI scaling procedure was used to obtain ΔG_{VDW} . We also used a similar method as Method 3 to obtain the IOD and CN values except that a bigger box with size of $\sim(36\text{\AA}\times 36\text{\AA}\times 36\text{\AA})$ was used for the TIP4P_{EW} water model, in which there were 1085 water molecules. For the monovalent and highly charged ions, Method 3 was used to simulate the HFE, IOD and CN values. The time-step was set to 1 fs for all of the MD simulations except for the simulations performed for the proton-aqueous system using the 12-6-4 model, for which a 0.5 fs time-step was used. We estimated the uncertainties of the simulated HFEs using the 12-6-4 model in present work are about ± 1 kcal/mol for the monovalent¹⁴³ and divalent ions,¹⁷⁴ while they are ± 2 kcal/mol for the highly charged ions.¹⁴⁵

Fitting procedure of the HFE quadratic fitting curves:

- (1) There are 7 different κ values for each row in Table SI.12. We performed the quadratic fitting and got the equation for HFE *versus* κ for each row with a fixed $R_{\text{min}}/2$ value.
- (2) We solved the equation to determine the κ value to reproduce the target HFE of a metal ion with the fixed $R_{\text{min}}/2$. Then for each row we obtained a point $(R_{\text{min}}/2, \kappa)$

which can reproduce the target HFE. In total there are 16 rows so we got 16 points, with each of which can reproduce the target HFE.

- (3) We discarded the points which were not in the parameter space (with $R_{\min}/2$ in 0.8-2.3 Å and κ in 0-6 Å⁻²) and performed another quadratic fitting of κ *versus* $R_{\min}/2$ for the remaining points to get the final fitting curve for the target HFE.

Fitting procedure of the IOD quadratic fitting curves:

- (1) There are 7 different κ values for each row in Table SI.13. We performed the quadratic fitting and got the equation for IOD *versus* κ for each row with a fixed $R_{\min}/2$ value.
- (2) We solved the equation to determine the κ value to reproduce the target IOD of a metal ion with the fixed $R_{\min}/2$. Then for each row we obtained a point ($R_{\min}/2, \kappa$) which can reproduce the target IOD. In total there are 16 rows so we got 16 points, with each of which can reproduce the target IOD.
- (3) We discarded the points which were not in range of $R_{\min}/2$ in 0.8-2.3 Å and κ in 0-10 Å⁻² and then performed another quadratic fitting of κ *versus* $R_{\min}/2$ for the remaining points to get the final fitting curve for the target IOD.

Mg²⁺ is one of the most commonly seen divalent cations in cells and play a significant role in the metabolism of organisms. In order to test the performance of our parameters in mixed systems, we carried out the simulations of MgCl₂ solutions at different concentrations. For simulating these MgCl₂ systems, the Cl⁻ LJ parameters were from the

work of Joung and Cheatham,³⁹ a TIP3P water box (including ~2000 water molecules) was used, and three different concentration levels (0.25 M, 0.5 M and 1.0 M) were modeled:

- (1) For the system with 0.25 M concentration: 10 Mg^{2+} and 20 Cl^- ions were solvated in a cubic water box with 2158 TIP3P water molecules;
- (2) For the system with 0.5 M concentration: 19 Mg^{2+} and 38 Cl^- ions were solvated in a cubic water box with 2184 TIP3P water molecules;
- (3) For the system with 1.0 M concentration: 38 Mg^{2+} and 76 Cl^- ions were solvated in a cubic water box with 2102 TIP3P water molecules.

The process for obtaining the IOD values was described in Method 3 in section 2.2.

Nucleic acids are the gene carriers. Mg^{2+} ions play very important roles in nucleic acid functions. We treated a nucleic acid system as an example to assess different parameter sets of the Mg^{2+} ion. The PDB structure of a DNA fragment (PDB ID: 1D23) was used in the simulation. The water molecules were removed from the PDB file and hydrogen atoms were added using the *reduce* program¹⁷⁵ in AmberTools. Afterwards, 7 Mg^{2+} ions were added to neutralize the system (plus the two Mg^{2+} ions already existing in the PDB structure, yielding 9 Mg^{2+} in the system in total). Then a cubic water box with 3547 TIP4P_{EW} waters was used to solvate the system.

The minimizations and MD simulations were performed in 7 stages as described follows:

- (1) Firstly, 2000 steps steepest descent minimization plus 3000 steps conjugate gradient minimization were performed to relax the water molecules and ions (with a 500.0 kcal/mol•Å⁻² force restraint on the other parts of the structure);

- (2) Afterwards, 2000 steps of steepest descent minimization plus 3000 steps of conjugate gradient minimization were performed to relax the hydrogen atoms, water molecules and ions (again, with a $500.0 \text{ kcal/mol} \cdot \text{\AA}^{-2}$ force restraint on the other parts of the structure);
- (3) Then 5000 steps of steepest descent minimization plus 40000 steps of conjugate gradient minimization were performed to relax the whole system.
- (4) 40 ps of simulation in the NVT ensemble was carried out to heat the system from 0 to 300 K.
- (5) Afterwards, 20 ps of simulation in the NVT ensemble at 300 K was performed to equilibrate the system.
- (6) Then 2 ns of simulation was performed using the NPT ensemble at 300 K and 1 atm to correct the system density and further equilibrate the system.
- (7) Finally 10 ns of production simulation in the NPT ensemble was performed at 300K and 1 atm with snapshots being stored every 2 ps for the final data analysis.

PBC and PME were employed during all the MD simulations. The time-step was set as 1 fs and the cut-off was set to 10 Å. Temperature was controlled by the Langevin dynamics with the collision frequency equals 1.0 ps^{-1} . The pressure relaxation time was set as 2.0 ps in the NPT ensemble. SHAKE^{141, 142} was used to constrain the bonds involving hydrogen atoms during the MD simulation. The distance between the Mg^{2+} and backbone phosphate was obtained in a similar way as the IOD value, using the procedure described in Method 3.

3.3 Parameter Space Scanning and Parameter Determination

First, we scanned the parameter space for divalent metal ions where $R_{\min}/2$ and κ represented the two axes. The range investigated for $R_{\min}/2$ was 0.8-2.3 Å with an interval of 0.1 Å interval, while κ ranged 0-6 Å⁻² with an interval of 1 Å⁻². The ϵ value was fixed for each $R_{\min}/2$ value according to NGC, where the $R_{\min}/2$ and ϵ values had a relationship of $-\log(\epsilon) = C_1 \times e^{C_2 \times R_{\min}/2}$ with $C_1 = 57.36$ and $C_2 = -2.471$. To be consistent with our work for the 12-6 model, the final parameters represent $R_{\min}/2$ with three decimal places, ϵ with eight, and κ with three. The HFE, IOD and CN values from our parameter space scanning for the TIP3P, SPC/E and TIP4P_{EW} water models are given in Tables SI.12 and SI.13. The HFE and IOD fitting curves were obtained based on these data. The curve fitting procedure is provided section 3.2. It is similar to the method used for the 12-6 model (see section 2.2) with the only difference being for the IOD fitting curve. The intention of this modification was to obtain enough meaningful points to fit the IOD curves.

Using the TIP3P water model as an example, we present the HFE and IOD fitting curves we obtained in Figures 25-27. In contrast to the 12-6 LJ nonbonded model of divalent metal ions, in which the HFE and IOD fitting curves for each metal ion were nearly parallel with each other (see Figure 7), the HFE and IOD fitting curves for each metal ion for the 12-6-4 model have an intersection-point. The fitting curves for the Be²⁺ ion is not shown in Figures 25-27 because the intersection-point of its HFE and IOD fitting curves is beyond the scanning range examined. Because of the presence of an intersection point

there exists a set of parameters that reproduce the experimental HFE and IOD values simultaneously. After obtaining the intersection-point between the two fitting curves and fine-tuning the resultant parameters, we obtained the final parameters for each divalent metal ion. These parameters are shown in Tables 27-28, the simulated HFE, IOD and CN values are given in Table 29. Herein, our final parameters reproduce the target HFEs within ± 1 kcal/mol and simultaneously reproduce most of the target IOD and CN values with good accuracy.

The 12-6 LJ nonbonded model still shows some unsatisfactory behavior in that it underestimates (for some positive ions) or overestimates (for the negative ions) ion-water interactions, which leads us to build a 12-6-4 LJ-type nonbonded model for the monovalent ions investigated. Meanwhile, the 12-6 model showed remarkable errors for highly charged ions, for which the 12-6-4 model was also parameterized.

Besides the 12-6 LJ scanning curves of monovalent and highly charged ions, we also scanned the corresponding parallel curves using the 12-6-4 model. These curves share the same $R_{\min}/2$ and ϵ values with the 12-6 scanning curves but have an additional fixed C_4 term. This C_4 term equals $100 \text{ kcal/mol} \cdot \text{\AA}^4$ for monovalent cations. While for the monovalent anions, this C_4 term equals $-100 \text{ kcal/mol} \cdot \text{\AA}^4$. And for the trivalent and tetravalent ions, this term equals $500 \text{ kcal/mol} \cdot \text{\AA}^4$. The parameters, and the simulated HFE, IOD and CN values of these parameter scans were shown in Tables SI.3 to SI.6. Via linear interpolations from the 12-6 to the 12-6-4 scanning curves we were able to obtain initial guesses for the final 12-6-4 parameter sets. After initial parameter selection

and subsequent fine-tuning, the final 12-6-4 parameters were determined to reproduce the target HFE and IOD values. The final 12-6-4 parameters of monovalent ions are shown in Table 30, while the corresponding computed HFE, IOD and CN values are shown in Table 31. These parameters simultaneously reproduce the target HFEs and IODs with considerable accuracy (within ± 1.0 kcal/mol of the target HFEs and within ± 0.01 Å of the target IODs). The final optimized 12-6-4 parameters for highly charged ions are given in Table 32 while the simulated HFE, IOD and CN values are shown in Table 33. These parameters reproduce the target HFE values within ± 1 kcal/mol and target IOD values within ± 0.01 Å for the M(III) ions while they reproduce the target HFE values within ± 2 kcal/mol and the target IOD values within ± 0.01 Å for the M(IV) ions. It can be seen that in the 12-6-4 parameter set for cations ranged from monovalent to tetravalent, the $R_{\min}/2$ terms are similar between the three water models while the C_4 term for TIP4P_{EW} water is generally larger than for the same metal ion in the other two water models. This may due to the smaller dipole of the TIP4P_{EW} water model (~ 2.32 D) relative to the TIP3P (~ 2.35 D) and SPC/E (~ 2.35 D) water models.

Comparing to the 12-6 model, we can see there is significant improvement of the accuracy using the 12-6-4 model, which is able to reproduce the experimental HFE and IOD values simultaneously. While for the 12-6 model, if you want to reproduce the experimental HFE value, the error in the simulated IOD value would increase along with the formal charge of the cation (see Table 25 and Figure 20). *Vice versa*, if you simulate the IOD value using the 12-6 model, the error of the calculated HFE would increase

markedly with an increase of the oxidation state of the cation (see Table 26 and Figure 21).

3.4 Discussion

The situations are different for the monovalent cations and anions with respect to the obtained parameters. For the positive ions, the C_4 terms are always positive, while the C_4 terms are negative values for the negative ions. This arises from the fact that the 12-6 LJ nonbonded model ignores the charge-induced dipole interaction for the positive ions, but overestimates the charge transfer effect for the negative ions (see discussion below). Meanwhile, it is interesting that the C_4 term decreases as the ionic radii increases for the halide ions when using the TIP3P water model. While for the SPC/E and TIP4P_{EW} water models the corresponding C_4 terms are very similar for different halide ions.

The 12-6 LJ nonbonded model represents the interaction between an ion and its environment via a summation of the electrostatic and VDW terms. It doesn't consider polarization and charge-transfer effects explicitly, which may be the reason for the negative C_4 terms for the halide ions. For the cations, it is the oxygen atom that acts as the coordinated atom while it is the hydrogen atom coordinated to the counter-ion in the anion-water system. Since the electronic cloud focuses on the oxygen atoms, the charge-induced dipole cannot be overlooked for the positive ion-water systems. However, the polarization effect in the halide ion-water interaction should be relatively weak given the nature of the electron cloud on the hydrogen atoms in a water molecule.

Meanwhile, the charge-transfer effect behaves differently in ion-water systems for the positive and negative ions. The charge transfer direction is from the water molecules to

the ions for the former systems but in the opposite direction for the later systems. Thompson and Hynes found that, for anion-water systems, there is charge transfer from the ion to the water, and the strength of this effect decreases as the halide ion's size increases.¹⁷⁶ This is counter-intuitive to what would be expected based solely on the electron affinities of the halogen atoms. They also found that when the distance between the halide ion and water is in the range of 0.7-1.0 Å, there is a strong polarization effect. However, as the distance continues to increase the charge transfer effect becomes the dominant part where the charge is mainly transferred between the halide ion and the oxygen atom but not the hydrogen atoms of the water molecule.¹⁷⁶ Moreover, the closer the halide ion is to a water molecule the stronger of the charge transfer effect. The red shift caused by the charge transfer effect is hard to simulate using the 12-6 model while it is demonstrated that a two-state valence bond model is capable of doing that.¹⁷⁶ This is consistent with our observation that as the halide ion size increases, the HFE error of the IOD parameter set decreases (true for the SPC/E and TIP4P_{EW} water models while it is almost a constant for the TIP3P water model). Taking the charge transfer effect into account, the charge of the halide ions are less than -1e in aqueous solution,¹⁷⁷ so the 12-6 LJ nonbonded model overestimates the halide ion-water interactions, making a negative C₄ term necessary in the 12-6-4 LJ-type nonbonded model.

For the alkaline earth metal ions the C₄ term decreases monotonically with an increase of the metal ion's radius. This may due to the shielding of the metal ion's ability to polarize the surrounding water molecules when the ion's radius increases. Another interesting observation is that although Be²⁺ is the smallest ion studied in the present work, it doesn't

have the largest C_4 value among all the metal ions. For example, in the TIP3P or SPC/E water model, the C_4 value of the Be^{2+} ion is around $190 \text{ \AA}^4 \text{ kcal/mol}$, which is $\sim 20\%$ less than the value obtained for Zn^{2+} . This may arise from a significant charge-transfer effect between Be^{2+} and its surrounding water molecules. Pavlov *et al.* showed that there is about $-1.28e$ transferred from the surrounding water molecules to the Be^{2+} ion in the $[\text{Be}(\text{H}_2\text{O})_4]^{2+}$ complex according to a Mulliken population analysis at the B3LYP/6-311+G(2d,2p) level of theory.¹⁷⁸

For the divalent metal ions in the first transition row, their C_4 terms are also consistent with the Irving-William series¹⁷⁹ where the magnitudes of the C_4 terms varies as $\text{Mn}^{2+} < \text{Fe}^{2+} < \text{Co}^{2+} < \text{Ni}^{2+} < \text{Cu}^{2+} > \text{Zn}^{2+}$, which arises due to the interplay between the covalent and ionic interactions.¹⁸⁰ Through density functional theory, it was found that the charge transfer effect follows the Irving-William series sequence.¹⁸⁰ There is also a larger bond order between the Cu^{2+} ion and its ligating ligands than for the other M(II) ions in the Irving-William series. Moreover, the ionic and covalent interactions can compensate one another. When there is a stronger charge transfer, there is usually a stronger covalent interaction but a weaker ionic interaction between the metal ion and its ligands.¹⁸⁰ Another thing to bear in mind is that the d^9 electronic structure of the $\text{Cu}(\text{H}_2\text{O})_6^{2+}$ complex results in the Jahn-Teller effect where the axial bond lengths are about $0.44 \text{ \AA}$ longer than the equatorial ones.¹²³ This effect could not be modeled using the present 12-6-4 nonbonded potential due to the isotropic approximation employed.

Cd^{2+} has a smaller C_4 value than that of Zn^{2+} , which is due to the longer distance between the Cd^{2+} ion and the surrounding water molecules. However, it is interesting that Hg^{2+} , the ion that has the highest atomic number among these +2 metal ions, has a C_4 value bigger than that of Zn^{2+} . Tai and Lim found that Hg^{2+} is a much better charge acceptor than Zn^{2+} according to their calculations,¹⁸¹ which they ascribed to the relativistic effect. Hg^{2+} can accept more negative charge from its surrounding ligands than the Zn^{2+} ion. This phenomenon not only happened when the ligating atom is the “soft” sulfur atom but also when the ligating atom is a “harder” atom like nitrogen or oxygen.¹⁸¹

The Sn^{2+} cation also has an unusual C_4 value among these +2 metal ions. This may be because its two outermost electrons occupy a 5s orbital and Sn also has a +4 oxidation state that is slightly more stable. Experiment has also revealed two different bond lengths with an asymmetric structure in the first hydration shell of Sn^{2+} , which can not be well reproduced with our model.¹⁶²

The main group and transition trivalent metal ions have much stronger ion-water interactions than the Ln^{3+} ions. They form a stable octahedral structure with water molecules in the first solvation shell. Data in Table 3 indicates that the average effective radius of the first solvation shell water is ~ 1.35 Å for the first several metal ions, which is consistent with strong interactions between the coordinated water molecules and these metal ions. Meanwhile, these values are close to previously proposed coordinated water radius (~ 1.34 Å) based on experimental data of Cesium alums.¹⁸² The corresponding average value is ~ 1.49 Å and ~ 1.46 Å for the Ln^{3+} and the M(IV) metal ions respectively,

which implies a smaller electronic cloud overlap between the metal ion center and each of the coordinated water molecules.

Al^{3+} , In^{3+} , Tl^{3+} are group IV ions. For the C_4 parameters derived herein we obtained a sequence of $\text{Tl}^{3+} > \text{Al}^{3+} > \text{In}^{3+}$. The Al^{3+} ion is the smallest M(III) ion, resulting in a relatively large C_4 term due to its strong covalent interaction with the coordinated water molecules. Tl has two oxidation states: +1 and +3 and the HFE values of the Tl^+ and K^+ ions are almost the same in Marcus' HFE set.¹¹⁵ Tl^{3+} could have very strong covalent interactions with the surrounding residues. The reduced electric potentials (electric potentials of reaction $\text{M(III)} + 3\text{e}^- = \text{M}$) are -1.67 eV, -0.3382 eV and +0.72 eV for Al^{3+} , In^{3+} and Tl^{3+} , respectively.¹⁸³ The positive reduction potential of Tl^{3+} makes it a very reactive species. It readily obtains electrons from its surroundings, which may be the reason for a strong charge-transfer effect between the Tl^{3+} ion and the surrounding water molecules. The 12-6-4 parameters of In^{3+} and Tl^{3+} ions gave an excellent prediction for the HFE and IOD values but overestimated the CNs (8 instead of 6). This may be due to the lack of a correction term for the water-water interactions in the first solvation shell during the simulations. The water-water interactions were parameterized to reproduce the pure liquid water properties in the original parameter designs. However, the first solvation shell waters of the highly charged metal ions should more strongly repel one another due to their bigger charge separations. This effect is smaller for the M(I) and M(II) metal ions, but it dramatically increases for the highly charged ions. Meanwhile, this kind of effect may decrease in protein systems due to the pre-organizations of the metal ion binding sites.

Fe^{3+} has a larger C_4 term than Al^{3+} , the smallest M(III) ion, which suggests that Fe^{3+} has a stronger interaction with its surrounding water molecules. This is consistent with QMCF MD simulations, which showed the force constant between the ion and the oxygen of the first solvation shell waters ($k_{\text{ion-O}}$) is 198 N/m for Fe^{3+} , compared to 185 N/m for Al^{3+} .²⁶ This is a consequence of both the electrostatic and covalent interactions. The Fe^{3+} ion has an average charge of +1.85e (from a Mulliken analysis) in the QMCF simulation¹⁸⁴ while the Al^{3+} ion has a corresponding value of +2.5e,¹⁸⁵ which implies there is a stronger charge-transfer effect between the Fe^{3+} ion and the coordinated water molecules than for the Al^{3+} ion. There is a slight overestimation of the CN for the Fe^{3+} ion. As discussed above, this may be due to the underestimation of the interactions between the first solvation water molecules. While this effect is operative in aqueous solution, it could be less of an issue in protein systems (see section 3.5).

The +3 oxidation state is the typical oxidation state of the Ln elements, with the exception that Eu^{2+} and Ce^{4+} could also be observed. This is because Eu^{2+} has a half-filled 4f orbital while Ce^{4+} has the same electronic configuration as Xe. The interaction of the Ln^{3+} ions with surrounding water molecules would be expected to have more ionic character than the M(III) ions discussed above. For example, the C_4 terms between the Ln^{3+} ions and water molecules are in the range of 152-282 kcal/mol $\cdot\text{\AA}^4$, which is smaller than the range of 258-456 kcal/mol $\cdot\text{\AA}^4$ for the other +3 metal ions discussed above. Previous simulations found that the $k_{\text{ion-O}}$ values are much smaller for the Ln^{3+} ions. For example, La^{3+} , Ce^{3+} , Lu^{3+} and Er^{3+} have $k_{\text{ion-O}}$ values ~ 110 N/m, while the values for the Al^{3+} and Fe^{3+} ions are 185 and 203 N/m, respectively.¹⁸⁶ The Ln^{3+} ions have effective

ionic radii in the range of 0.86-1.03 Å and IOD values in the range of 2.34-2.55Å. These values are similar to those of the Ca^{2+} ion (whose effective ionic radius and IOD value are 1.00 Å¹²⁵ and 2.46 Å,¹⁸⁷ respectively). Therefore, they have been used as probes to investigate the roles of the Ca^{2+} ions in biological systems.¹⁸⁸

From Table 32 we observe that the La^{3+} and Gd^{3+} ions have the smallest C_4 terms among the Ln^{3+} ions. This may be because they have either totally empty or half-filled 4f orbitals, making them more likely to form isolated ions, which reduces the covalent character of their bonds with the coordinated water molecules. It is easy to see the “lanthanide contraction” effect from Table 3. The effective ion radius decreases monotonically with an increase of the metal ions’ atomic number due to the poor shielding of the 4f electrons towards the 5s and 5p orbitals.¹⁸⁹ A similar tendency can also be seen for the HFE and IOD value along the series. Our final $R_{\text{min}}/2$ parameters are consistent with this pattern as well. Meanwhile, the CN also decreases along the Ln^{3+} ion series. Previous work reached the conclusion that the lighter Ln^{3+} ions (La^{3+} to Nd^{3+}) prefer a CN of ~9 and the heavier ions (Gd^{3+} to Tb^{3+}) prefer a CN of ~8 while the middle ions such as Sm^{3+} and Eu^{3+} , have CNs between these two values.¹⁹⁰⁻¹⁹⁵ It was proposed that the former Ln^{3+} ions form tricapped trigonal prism structures with their first solvation shells, while the structures shift to distorted bicapped trigonal prism structures for the heavier elements as one of the two capping water leaves the first solvation shell.¹⁹⁶ Generally speaking, the 12-6-4 parameter set could well reproduce the target HFE, IOD and CN values with the exception that some of the CN values were slightly overestimated. The final parameters gave CNs in the range of 9-10 for the Ln^{3+} ions, rather than the range of 8-9 reported in

the literature. As discussed above, this may be due to the fact that there is no water-water interaction correction term in the present parameterization process. Moreover, the CN values of the Ln^{3+} ions given by Marcus¹¹⁵ (as shown in Table 3) likely also vary under different experimental conditions (counter-ions used, solute concentration, *etc.*). Among all the trivalent metal ions investigated herein, Al^{3+} , Y^{3+} , and La^{3+} have the same electron configurations as the noble gas atoms Ne, Kr, and Xe respectively. Using the TIP3P water model as an example, we can see their C_4 values decrease monotonically from 399, to 216, and 152 kcal/mol $\cdot\text{\AA}^4$.

There are only a few M(IV) ions that exist in aqueous solution while the others are readily hydrolyzed into polynuclear complexes in water.¹³⁶⁻¹³⁹ Table 3 shows the M(IV) ions examined herein. These ions exist at least in highly acidic solutions. The CN values of these metal ions are greater than 8 with some of them being ~ 10 according to experiment.¹³⁶⁻¹³⁹ Earlier research found that Pu(IV), Th(IV) and U(IV) could strongly bind to transferrin, a iron-transport protein.¹⁹⁷ Hence, the parameters developed herein might facilitate theoretical research on the bio-toxicity of these M(IV) ions.

Zr^{4+} and Hf^{4+} are in the IVB group. Even though Hf^{4+} has a larger atomic number than Zr^{4+} , due to the “lanthanide contraction” effect, it has a smaller effective radius, a smaller IOD, a more negative HFE, and a bigger C_4 term than Zr^{4+} . These observations reflect its stronger interaction with the surrounding water molecules. In contrast, Ce^{4+} and Th^{4+} are in the same group where the one with larger atomic number (Th^{4+}) has a bigger ionic radius, a less negative HFE, and a smaller C_4 term. This may be because that Ce^{4+} and

Th^{4+} share the same electronic structures of Xe and Rn, respectively. Besides these ions, Zr^{4+} is another M(IV) ion which shares the same electronic structure with a noble gas atom (Kr). There is also a single trend in the C_4 terms of the Zr^{4+} , Ce^{4+} , Th^{4+} ions for each specific water model. For instance, their C_4 terms in the TIP3P water model are 761, 706, and 512 kcal/mol $\cdot\text{\AA}^4$, respectively.

Th, U and Pu elements are in the An series and are the largest elements investigated in the present work. Their tetravalent metal ions only exist in highly acidic solutions. Canaval *et al.* investigated Th^{4+} in aqueous solution using the QMCF MD simulation method. They found a stable 9-coordinate complex, and even third solvation shell water molecules had a bigger mean residence time than that of pure water, implying they are stabilized by the highly charged Th^{4+} ion.¹⁹⁸ U^{4+} fluoresces due to the electron transition between the $5f^16d^1$ and $5f^2$ electronic configurations.¹⁹⁹ The U^{4+} ion has the largest C_4 term in all the M(IV) metal ions investigated. Frick *et al.* investigated the U^{4+} ion in aqueous solution using the QMCF MD method and the CN value was characterized as 9 while the average charge of U^{4+} was predicted to be +2.68e from the Mulliken population analysis.²⁰⁰ Odoh *et al.* simulated the Pu^{3+} , Pu^{4+} , PuO_2^+ and PuO_2^{2+} ions in waters using the CPMD method.²⁰¹ They predicted the pK_a values of the first hydrolysis step for the Pu^{3+} , Pu^{4+} , PuO_2^+ and PuO_2^{2+} ions are 6.65, 0.17, 9.51, and 5.70 respectively, showing a general tendency that the larger of the charge of the metal center, the lower the pK_a value of the first hydrolysis reaction.

Hf⁴⁺, Zr⁴⁺ and Pu⁴⁺ have relatively smaller IOD values among the tetravalent ions where they all have experimental CNs of ~8.¹³⁹ Ce⁴⁺ was determined to have an experimental CN of ~9¹³⁶ while U⁴⁺ and Th⁴⁺ have CNs between 9-11.¹³⁹ Soderholm *et al.* proposed that counter-ions play a key role in the first solvation shell structure while the 9, 10 or 11-coordinated Th⁴⁺ have very small energy differences and are in a dynamic equilibration.²⁰² The simulated HFE and IOD values of the 12-6-4 parameter set are in excellent agreement with experiment. The simulated CN values of most of the M(IV) ions are ~10 with Hf⁴⁺ having a CN ~8 for the TIP4P_{EW} and SPC/E water models. Herein, the TIP3P model always predicted a larger CN value than the other two water models, which may be because it has a smaller C₁₂ term ($\sim 582.0 \times 10^5$ kcal·Å¹²/mol) than those of the SPC/E ($\sim 629.4 \times 10^5$ kcal·Å¹²/mol) and TIP4P_{EW} ($\sim 656.1 \times 10^5$ kcal·Å¹²/mol) models.

There are several redox pairs available in our 12-6-4 parameter set. We analyzed them below to explore the consistency of the C₄ parameters with respect to the behavior of these pairs in aqueous solution. We also calculated the relative HFEs inside some redox pairs for the TIP3P water model. A 9-window TI simulation (50 ps of equilibration and 150 ps of production for each window) was performed forward and backwards to obtain the final results. The simulated relative HFEs of Cr²⁺/Cr³⁺, Fe²⁺/Fe³⁺ and Ce³⁺/Ce⁴⁺ ion pairs were 516.9, 580.2, and 698.2 kcal/mol while the experimental values are 516.2, 579.6, and 697.6 kcal/mol respectively.¹¹⁵ These results further validated the method employed in the present work.

The Cr^{2+} and Cr^{3+} ions have the $[\text{Ar}]3d^4$ and $[\text{Ar}]3d^5$ electronic structures, respectively, where the $\text{Cr}(\text{H}_2\text{O})_6^{2+}$ complex has a strong Jahn-Teller effect while $\text{Cr}(\text{H}_2\text{O})_6^{3+}$ molecule has a standard octahedral configuration. The IOD values decreases from 2.08 to 1.96 Å for the Cr^{2+} and Cr^{3+} ions. Using the TIP3P water model as a representative example we observe that the $R_{\text{min}}/2$ parameter decreases from 1.431 to 1.415 Å while the C_4 term increases from ~ 137 to ~ 258 kcal/mol $\cdot\text{\AA}^4$ for the Cr^{2+} and Cr^{3+} ions, which trends well follow the intrinsic physics.

$\text{Fe}^{2+}/\text{Fe}^{3+}$ redox pair exists broadly in biologically related systems such as the Fe-S proteins and heme structures,^{203, 204} and it plays fundamental roles in many electron transfer processes. The experimental IOD value shrinks about ~ 0.08 Å from 2.11 Å of Fe^{2+} to 2.03 Å of Fe^{3+} ion. Moin *et al.* investigated the ferrous and ferric ions in water using the QMCF MD simulation method. They obtained the $k_{\text{ion-O}}$ force constant of 193 N/m for Fe^{3+} , which is almost twice as strong as that of Fe^{2+} (93 N/m). While the effective charges (from a Mulliken population analysis) are in the range of +1.25e to +1.45e (with an average of +1.36e) for the Fe^{2+} ion during the simulation and for the Fe^{3+} ion the effective charges are in the range of +1.70e to +1.95e (with an average of +1.85e).¹⁸⁴ By comparing the 12-6-4 parameters for the Fe^{2+} and Fe^{3+} ions in the TIP3P water model, we find that the $R_{\text{min}}/2$ decreases slightly as the outer shell electron number decreases while the C_4 term increases by about 2.5 times. This is consistent with the ratio between the C_4 terms of a trivalent and a divalent ion ($[3/2]^2 = 2.25$) that is derived from the original ion-induced dipole equation (see section 3.5).

The ratio between the experimental HFEs of the Cu^+ and Cu^{2+} ions is ~ 3.8 (using $-480.4/-125.5$), which is close to 4 and consistent with the Born equation. By treating the TIP3P water model as an example, we can see that the Cu^{2+} ion has a much bigger C_4 term than that of the Cu^+ ion (290.9 *versus* $7 \text{ \AA}^4 \text{ kcal/mol}$). However, the $R_{\min}/2$ parameter of the Cu^+ ion is less than that of the Cu^{2+} ion (1.217 *versus* $1.476 \text{ \AA}$), because it has a smaller IOD value ($1.87 \text{ \AA}$ of Cu^+ *versus* $2.11 \text{ \AA}$ of Cu^{2+}). Meanwhile, the 12-6-4 model could not reproduce the CN of the Cu^+ ion (giving it as 4 instead of 2) or the Jahn-Teller effect of the Cu^{2+} ion. These effects are related to the atomic orbitals and cannot be simulated using the 12-6-4 model which treats the ion as a symmetry particle.

The ratio between the experimental HFEs of the Tl^{3+} and Tl^+ ions is ~ 13.2 (using $-948.9/-71.7$), which is deviated from 9, the ratio between the squares of their oxidation states. Again, using the TIP3P water model as an example, we can see that the $R_{\min}/2$ parameter decreases by $0.322 \text{ \AA}$ when the Tl^+ ion loses two electrons. Meanwhile, the ratio between the C_4 terms of the Tl^{3+} and Tl^+ ions is ~ 9.1 (using $456/20$), which strongly agreed with the original ion-induced dipole equation (see section 3.5). These trends are consistent with the intrinsic physics and further validate the 12-6-4 model.

Ce has both the +3 and +4 oxidation states. Ce^{4+} is the most stable state because it shares the same electronic configuration with Xe. Just like the $\text{Cr}^{2+}/\text{Cr}^{3+}$, $\text{Fe}^{2+}/\text{Fe}^{3+}$, and $\text{Tl}^+/\text{Tl}^{3+}$ redox pairs discussed above we find that the $R_{\min}/2$ value decreases while the C_4 term increases significantly with the increasing of the ion' formal charge. For example, for the 12-6-4 parameters determined for the TIP3P water model, the $R_{\min}/2$ parameter decreases

by ~ 0.03 Å while the C_4 term increases by ~ 480 kcal/mol $\cdot$ Å⁴ when Ce³⁺ loses one electron.

3.5 Parameter Assessment

$$C_4 \approx -\frac{1}{2}\alpha_0 \left(\frac{q}{4\pi\epsilon_0\epsilon_r}\right)^2 \cos\theta \quad (16)$$

$$U_{q-\alpha_0} \approx -\frac{1}{2}\alpha_0 \left(\frac{q}{4\pi\epsilon_0\epsilon_r r^2}\right)^2 \cos \theta \quad (17)$$

To further assess our final parameters we employed equation 16, which is based on the equation for the charge-induced dipole interaction (equation 17), to approximate the C_4 term between the M(II) ion and the surrounding water molecules. In these two equations, q is the charge of the metal ion, α_0 is the polarizability of the particle interacting with the metal ion. θ is the angle between the induced dipole and electronic field generated by the metal ion. To calculate the C_4 term between an M(II) ion and a water molecule, we assumed $\alpha_0=1.444 \text{ \AA}^3$ (obtained from the book of Eisenberg and Kauzmann²⁰⁵), $q=+2e$, $\theta=0^\circ$ and $\epsilon_r=1$ (as in the AMBER FF⁸⁹). The calculation yields a value of $\sim 960.0 \text{ \AA}^4 \text{ kcal/mol}$ for the C_4 term, which is on the same order of magnitude (except for some large ions in the TIP3P or SPC/E water model) as but bigger than our final C_4 parameters for divalent metal ions. This may arise because the fixed-charge water models are over-polarized in their original designs. For example, the TIP3P and SPC/E water molecules have dipoles of $\sim 2.35 \text{ D}$ while the TIP4P_{EW} water molecule has a dipole of $\sim 2.32 \text{ D}$. These values are greater than the experimentally determined permanent dipole (1.855 D)²⁰⁶ for the gas-phase water molecule. These fixed-charge water models included the polarization effect to some extent by over-fitting the permanent dipole moment while omitting the induced dipole. This approximation may also give some insights into to the nonbonded model's tendency to underestimate the HFEs for +2 metal ions: in particular,

the three water models (TIP3P, SPC/E, and TIP4P_{EW}) all have smaller dipole moments than the dipole moment of a liquid phase water molecule determined from experiment (2.95 ± 0.2 D)²⁰⁷ or *ab initio* MD simulation (~ 3.0 D)²⁰⁸. In the work of Wu *et al.*, they found that the charge transfer effect between Zn²⁺ ion and water molecule could be incorporated into the polarization energy term when representing the charge and polarizability of Zn²⁺ as +2e and 0.260 Å³ respectively.²⁰⁹ This may come from the fact that the AMOEBA water model has a total dipole of 2.54 D,²¹⁰ which is bigger than the dipole moments of the three non-polarizable water models discussed here.

Meanwhile, as shown in Table 28, the C₄ term generally decreases when the metal ion's radius increases. This may arise from the $\cos\theta$ term in equation 16 because the nearby water molecules would be more readily polarized than the remote ones. When there is a greater distance between the metal ion and its surrounding water molecules, the latter would be more randomly oriented due to a reduction in the influence from the metal ion. Table 28 also shows that the TIP3P and SPC/E water models give very similar C₄ values while the TIP4P_{EW} model gives slightly larger values. This may arise from the fact that the negative charge of the TIP4P_{EW} water model is on the dummy atom, which is placed on the bisector of the hydrogen-oxygen-hydrogen angle, yielding a slightly smaller dipole, which is then responsible for the observed larger underestimation of the ion-water interaction. This representation results in a greater distance between the positive charge on the metal ion and the negative charge on water than those of the two 3-site water models and, thereby, increases the C₄ values.

Furthermore, from Table 26 and Figure 21 we observe that the underestimation extent of the absolute HFEs for the IOD parameter set increases roughly proportional to the charge square of the cation. This is consistent with the equation for the charge-induced dipole interaction (see equation 17), which interaction was not considered in the 12-6 LJ equation (see section 3.1). For example, the ratio of the underestimations of the absolute HFEs for the IOD parameter set in the TIP3P water model is roughly 1.0 : 3.0 : 4.8 : 14.2 between the mono-, di-, tri-, and tetravalent cations (calculated from 17.2 : 51.1 : 82.7 : 244.3 kcal/mol). This trend is close to the ratio: $1^2 : 2^2 : 3^2 : 4^2$, which is predicted from the original equation of charge-induced dipole interaction (see equation 16). This further validates the proposed C_4 term in the 12-6-4 model developed herein.

Herein we showed some test cases regarding the 12-6-4 model. These simulations proved the outperformance of the 12-6-4 model over the 12-6 model, and also the excellent transferability of the 12-6-4 model.

During the simulations of mixed systems using the 12-6-4 model, the C_4 parameter between the metal ion and each atom type (except the oxygen atom in water because the C_4 term between the metal ion and it has been parameterized) is evaluated using the following equation:

$$C_4(atom\ type) = \frac{C_4(H_2O)}{\alpha_0(H_2O)} \times \alpha_0(atom\ type) \quad (18)$$

We set the polarizability of a water molecule to $1.444 \text{ \AA}^3$, which was taken from Eisenberg and Kauzmann.²⁰⁵ Since the electron cloud almost centers on the oxygen atom

in water, here we treated the oxygen atom as having all of the polarizability of the water molecule while the hydrogen atoms have a polarizability equals zero.

Panteva *et al.* have benchmarked 17 different nonbonded models for simulating the Mg^{2+} -aqueous system in predicting the thermodynamic, structural, kinetic, and mass transport properties.⁹³ Results showed that none of the 12-6 models could reproduce all the properties at the same time. They also found that the 12-6-4 models offered improvement over the 12-6 models, and the 12-6-4 model combined with the SPC/E water model could accurately reproduce all the experimental properties simultaneously.

To further test the 12-6-4 nonbonded model proposed here, we performed simulations on aqueous MgCl_2 systems at different concentrations and on a Mg^{2+} -nucleic acid system. In total three parameter sets were tested in the simulations, which included the original AMBER FF parameters for Mg^{2+} , the CM and 12-6-4 parameter sets for Mg^{2+} developed herein. For the 12-6-4 parameter set, the $R_{\text{min}}/2$ and ϵ parameters of Mg^{2+} were from Table 27 based on the water model used and the C_4 parameter between Mg^{2+} and water oxygen atom is from Table 28 based on the water model used. Polarizabilities of different atom types, which were used to obtain the C_4 terms between the metal ion and different atom types in the 12-6-4 model, were shown in Table 34.

The IOD values of the Mg^{2+} and Cl^- ions in the MgCl_2 systems are shown in Table 35. It can be seen that both the 12-6-4 parameter set (which gives the IOD of ~ 2.11 Å for Mg^{2+}) and the CM parameter set (which gives the IOD of ~ 2.08 Å for Mg^{2+}) outperformed the

original AMBER FF parameters (which gives the IOD of ~ 2.00 Å for Mg^{2+}) with respect to the experimental IOD value of 2.09 ± 0.04 Å.¹²³ For the Mg^{2+} -nucleic acid system, Table 36 shows the IOD values and the Mg^{2+} -backbone phosphate distances. We can see that both the 12-6-4 and CM parameter sets offered remarkable improvements over the original Mg^{2+} parameter set in the AMBER FF while the 12-6-4 model gives the best results.

We also performed test simulations for several ionic solutions. The method was the same as the method used in the simulations of ionic solutions described in section 2.2 except a 12-6-4 model was used. We used equation 18 (for the C_4 terms between the cation and cation, between the anion and anion, and between the ion and the hydrogen atoms in water) and equation 19 (for the C_4 term between the cation and anion) to estimate the C_4 terms between the particle pairs (except for the C_4 term between cation and water oxygen atom, and the C_4 term between anion and water oxygen atom, both of which have been parameterized and were shown in Table 30). The polarizability values of various monovalent ions were taken from Sangster and Atwood.²¹¹

$$C_4 = \left(\frac{c_4^X(H_2O)}{\alpha_0(H_2O)} \times \alpha_0(M) + \frac{c_4^M(H_2O)}{\alpha_0(H_2O)} \times \alpha_0(X) \right) \quad (19)$$

As described in section 2.6, we compared the 12-6-4, HFE and IOD parameter sets to two of the best parameter sets developed previously regarding their abilities to reproduce the activity derivatives of different ionic solutions. The results are shown in Table 24. We can see that, indeed, the 12-6-4 parameter set showed the best performance among these parameter sets. Besides its better performance in simulating the single-ion properties, it

also showed improved results in reproducing the ion pair properties, further validating its advantage over the 12-6 model in simulating ions.

Moreover, a metalloprotein system was also tested. The method was the same as the method of simulating a Fe(III) containing protein described in section 2.2 except a 12-6-4 model was used. The C_4 term between the Fe^{3+} ion and water oxygen atom is from Table 32 and equation 18 was used to obtain the C_4 terms between the Fe^{3+} ion and other atom types. Polarizabilities of different atom types were shown in Table 37. As described in section 2.6, we examined the performance of the HFE, IOD and 12-6-4 parameter sets for simulating the Fe(III) containing protein system (see Figures 14-15). Results showed that the IOD and 12-6-4 parameter sets yielded stable metal complex structures (see Figures 16-17), while the HFE parameter set failed to do that.

3.6 Design Your Own Parameters

It is important to note that the HFEs of ions may vary among different experimental analyses and the subsequent theoretical treatments of the experimental data. The experimental HFEs of ions can be determined in different ways. In some works, the HFE values of ions were obtained based on the National Bureau of Standards (NBS) compilation of the conventional HFE values. In this situation, the HFEs of ions would change if the HFE of a proton changes. The present work relied on the approach of Marcus,¹¹⁵ in which he obtained the experimental HFEs of cations as described above (based on the proton's HFE and the conventional HFE values obtained from the NBS compilation,²¹² see equation 14). Marcus treated $\Delta_{\text{hyd}}G^0[\text{H}^+]=-1056 \text{ kJ/mol} \pm 6 \text{ kJ/mol}$ in his literature,¹¹⁵ which comes from $\Delta_{\text{hyd}}H^0[\text{H}^+]=-1094 \text{ kJ/mol}$, $\Delta_{\text{hyd}}S^0[\text{H}^+]=-131 \text{ J/(K}\cdot\text{mol)}$ or $S_{\infty}[\text{H}^+(\text{aq})]=-22.2 \text{ J/(K}\cdot\text{mol)}$. However, there are different HFE values that have been determined for the proton.^{117, 118, 120, 213, 214} For example, Tissandier *et al.* estimated the $\Delta_{\text{hyd}}G^0[\text{H}^+]$ value by employing the cluster-pair-approximation,¹¹⁸ while several other works concerning the computational estimation of the HFE of a proton have been published.^{117, 213, 214} Meanwhile, there were also experimental and theoretical efforts that predicted the HFE values for some of the investigated cations which were different from the target values used in this work.^{116, 215} For example, Schmid *et al.* proposed a new set of HFE values for several monovalent and divalent ions¹¹⁶ while Asthagiri *et al.* calculated the HFEs for the first transition row metal ions using a quasi-chemical theory of solutions.²¹⁵ Regardless of the choices made, the data in SI from our parameter space

scans for different water models would facilitate the parameter design targeting data sets other than we used herein.¹¹⁵

3.7 Conclusions

We showed that the electrostatic plus 12-6 LJ potential nonbonded model underestimates the interactions between metal ions and the surrounding water molecules in PME simulations (see section 2). Via a consideration of the physical origins, we hypothesized that the charge-induced dipole interaction is responsible for the majority of this underestimation. In light of this we proposed and parameterized a 12-6-4 LJ-type nonbonded model in order to take into account the charge-induced dipole interaction.

And we found that unlike the 12-6 LJ nonbonded model, the new 12-6-4 model could reproduce the experimental HFE, IOD and CN values simultaneously after systematic parameterization. It improves the accuracy of the 12-6 model remarkably with just a slight increase of the computational cost. In the present work we have parameterized the 12-6-4 LJ-type nonbonded model for 55 ions (including 11 monovalent cations, 4 monovalent anions, 16 divalent cations, 18 trivalent cations, and 6 tetravalent cations) for three extensively used water models (the TIP3P, SPC/E and TIP4P_{EW} water models). The 12-6-4 parameters have been parameterized to reproduce both the experimental HFE and IOD values simultaneously. Generally they reproduce the target HFEs within ± 1 kcal/mol for the ions that have formal charges equal or less than +3 and within ± 2 kcal/mol for the ions that have formal charges of +4, with reproducing the target IOD values within ± 0.01 Å.

Moreover, excellent quantitative and qualitative agreement with previous experimental and computational works supports the validity of the 12-6-4 LJ-type nonbonded model. For example, both the three parameters ($R_{\min}/2$, ϵ , and C_4) are in physical range. An independent study from Panteva *et al.* showed that the proposed 12-6-4 model offered improved performance over the 12-6 model for simulating the Mg^{2+} -aqueous system.⁹³ A new combining rule was proposed for the C_4 term, and tests in mixed systems further validated the new model: the 12-6-4 parameter set showed the best performance over all the parameter sets in the ionic solution tests, further supporting its advantage over the 12-6 LJ model, while simulations of aqueous $MgCl_2$, Mg^{2+} --nucleic acid, and metalloprotein systems consistently revealed good performance and transferability of the parameters determined herein.

In summary, the 12-6-4 LJ-type nonbonded model provides a significant improvement over the former 12-6 LJ nonbonded model. It reproduces several different kinds of experimental data at the same time, which eliminates the need to develop compromise parameters as was done for the former nonbonded model in PME simulations. It is easy to incorporate the present model into typical biomolecular FFs with minimal additional computational cost. We believe that the parameter sets developed herein will improve our ability to model ions in biological systems. However, further parameterization efforts may be needed to increase the performance of the current model. One caveat is that the new model doesn't consider the charge-transfer effect explicitly in the potential form, which may influence its ability to simulate systems with strong charge-transfer effects. Hence, care should be taken for systems like Be^{2+} and Hg^{2+} in aqueous solution due to the

existence of strong charge-transfer effects.^{178, 181} Finally, experimental values for the same metal ions may be variable due to different assumptions and standards employed. However, our data presented in SI allows for the straightforward design of LJ parameters for these ions, which could be helpful for the researchers who want to target different experimental HFE, IOD and CN values than used in our work.

CHAPTER 4: DEVELOPMENT OF THE MCPB.PY PROGRAM

4.1 Code Structure

MCPB.py is built using the molecule and atom classes in the pyMSMT package. There are different parsers (*e.g.* PDB parser, mol2 parser) available that read and write different format files. Users can build their own programs based on these data structures as well. The installed main program of MCPB.py is found in the \$AMBERHOME/bin directory. The source code for the MCPB.py program is freely available in the AmberTools suite of programs for those interested in delving into the details. The original version of the code is available as a part of the AmberTools15 package¹⁰³ and updated versions can be obtained from GitHub (<https://github.com/Amber-MD/pymsmt>). We suggest users that begin with the two examples in the SI to familiarize themselves with the program and then check the source code for particulars.

4.2 Workflow

Figure 28 illustrates the workflow of the MCPB.py program. It facilitates metal site modeling in classical FFs with parameters derived from QM calculations. It is designed as a bridge between several QM software packages (including Gaussian¹⁰⁴ and GAMESS-US¹⁰⁵) and MD software employing FFs (including AMBER,¹⁴⁰ CHARMM,¹⁰⁷ GROMACS,¹⁰⁸ NAMD,²¹⁶ OpenMM,¹⁴⁶ LAMMPS,²¹⁷ *etc.*) for the modeling of metal-containing systems. The version 1.0 of MCPB.py supports the following AMBER FFs: ff94, ff99, ff99SB, ff03, ff03.r1, ff10, ff12SB, ff14SB and GAFF. The default is ff14SB for the protein while GAFF is used for the small organic ligands. The program should be used with AmberTools software package because it uses code, data files and the resp program from it.

Similar to MCPB,⁴ MCPB.py uses two models to do the parameterization in order to strike a balance between accuracy and speed (as shown in Figure 28): a smaller one to obtain the metal associated bond and angle parameters and a larger one to parameterize the partial charges. Besides supporting the parameterization of metal centers which only have amino acid sidechains and ligands coordinating to the metal ion, MCPB.py further supports modeling for complexes, which have backbone, terminal oxygen or nitrogen atoms coordinated to the metal ions, or a system with a mixed sidechain/backbone/terminal/ligand binding mode. To differentiate it from the earlier MCPB program, we use “small model” instead of “sidechain model” because MCPB.py is more versatile in dealing with metal centers of different binding modes.

MCPB.py builds the small model based on the schemes shown in Table 38. The non-amino-acid residues are kept during the modeling process. Generally, MCPB.py uses three approaches to build the small model: (1) uses the CH_3R (where R represents a sidechain group) groups to mimic residues which have sidechain atoms bound to the metal ion; (2) uses a ACE (CO-CH_3) or NME (NH-CH_3) residue (together with a capped ACE or NME group, if necessary) to mimic the backbone when there is a backbone atom bound to the metal ion; or (3) builds a $\text{CH}_2\text{R-CO}$ or $\text{NH-CH}_2\text{R}$ group (with a capped ACE or NME group, if necessary) if both backbone and sidechain atoms of one residue are bound to the central metal ion. These approaches will mimic the chemical environment at an affordable computational cost with several examples shown in Figure 29. Similar to MCPB, MCPB.py supports modeling for those metal sites with multiple metal centers as well.

For the large model, MCPB.py uses an approach adapted from MCPB, in which the metal site amino acid residues are capped by ACE and NME groups, and if there are two residues that are both bound to the metal ion with less than 5 residues between them, the intermediate residues will be retained and simplified to GLY residues. There is a variable (`large_opt`) in MCPB.py to control whether a full geometry optimization, hydrogen-only optimization or no optimization is specified in the Gaussian input file of the large model. It is not suggested to carry out full geometry optimization for the large model due to the computational expense and the chance that the structure might undergo large structural changes in the absence of the protein environment constraints. In general, we recommend a hydrogen-only geometry optimization if some structural relaxation is necessary, but a

full optimization can be done if the computational resources allow and if it is confirmed that the structure doesn't dramatically change upon optimization.

In order to conserve the total charge of the metal site, RESP charge fits¹⁷³ for the large model are performed with the capped ACE, NME and intermediate GLY residues in their neutral forms. Similar to antechamber,²¹⁸ there are two stages of charge fitting performed in MCPB.py: the first stage is carried out with small restraints on the heavy atoms (0.0005 for each heavy atom) and no restraints on the hydrogen atoms. The second stage refits the charges on the CH₂ and CH₃ group (while the charges on the other groups are kept at the values from the first stage), including symmetrizing the charges on the hydrogen atoms bound to the same carbon atom, with a 0.001 restraint on each of the carbon atoms and no restraints on the hydrogen atoms.

Figure 30 shows the 81 metal ions presently supported by the MCPB.py program. Most of the VDW parameters (VDW radii for the RESP fits and LJ parameters) of the related metal ions are from section 2, while the remaining are adapted from UFF.²

4.3 Usage

The MCPB.py code was designed with an optimized structure and offers a more user-friendly experience than the MCPB program. The command line input for MCPB.py is:

MCPB.py -i input file -s step number

[--fchk Gaussian fchk file] [--logf Gaussian/GAMESS-US log file] (4)

Note: The --fchk and --logf options are not necessary, in the case where the fchk file or log file do not use the default name (users can consult the manual for further details).

There are a series of input variables available to meet different users' demands. In the examples below we kept the input files simple. For more advanced cases users can consult the software manual for full details. The modeling process of MCPB.py involves 4 steps (as shown in Figure 28) which are indicated as steps 1, 2, 3 and 4, respectively. The first step is preparing the input files for the small and large models (input files for the standard model will also be generated for the subsequent steps). The second step is generating the frmod file for the system, which contains the bond, angle, torsion and VDW FF parameters. The bond and angle parameters can be generated based on the Seminario method (default, also using step 2s), Z-matrix method (using step 2z), and an empirical method (using step 2e, only supports zinc currently). As in MCPB,⁴ MCPB.py assigns zero torsion barriers for dihedral angles.

The third step involves the RESP charge fitting steps and building the mol2 files for each metal site residue based on the fitted RESP charges. The program will rename these

residues automatically to differentiate them from the amino acid residues in the standard AMBER FF library. There are several options (with steps labeled as 3a, 3b, 3c(default) and 3d) available for performing the RESP charge fits. They correspond to charge fitting approaches involving: (1) all charges treated flexibly (ChgModA); (2) charges of the backbone heavy atoms (CA, N, C and O) are fixed (ChgModB); (3) charges of all the backbone atoms (N, H, C, O, CA, HA) are fixed (ChgModC); and (4) charges of all the backbone atoms and the CB atom are fixed (ChgModD). These fixed charges are assigned values from the AMBER FF employed. If a residue has a backbone atom bound to the metal ion, all atomic charges in it will be flexible no matter which algorithm is chosen. This is because metal binding may have a strong influence on the charge distribution of the atoms in the backbone.

The fourth step is generating the new PDB file with the renamed metal site residues and creating the leap input file to build the AMBER topology and coordinate files. If there are any missing parameters needed by leap, users can use the CartHess2FC.py program (available in <https://github.com/Amber-MD/pysmt>) to calculate all the bond and angle parameters of the small model and manually add the related parameters into the frcmod file. If there are any metal related dihedral parameters missing, users can give them a zero barrier in the frcmod file.

After completion of these steps users can transfer the AMBER topology and coordinate files to another format if they want to use alternate software to run the minimization and MD simulation. There are a few programs available to do this conversion: for example,

ParmEd¹⁰³ converts the AMBER format into CHARMM's format; ACEYPE¹⁰⁶ converts the AMBER format into that of GROMACS; finally, amber2lmp converts the AMBER format into LAMMPS's.

4.4 Examples

In what follows we delineate two sample examples: one is a metalloprotein and the other is an organometallic compound. The related modeling files could be found in the SI of ref 219.

Our first example is from PDB entry: 1E67, which is a zinc containing azurin protein from *Pseudomonas aeruginosa* and the structure has a resolution of 2.14 Å.²²⁰ The protein is a tetramer in which each monomer contains 128 amino acid residues and has a metal site with four residues bound to the central zinc ion (two HIS residues and one CYS residue which have sidechain atoms coordinated to the zinc ion, along with a GLY residue which has its backbone oxygen bound to the zinc ion). Here we only treat chain A (see Figure 31) as an example – the generated parameters are also applicable to the other three chains. First we prepare the system by using the H++ web server¹⁴⁸ to add hydrogen atoms (note: users need to manually add the zinc ion into the final PDB file and modify the residues which bind to the zinc ion because the H++ web server will delete the metal ions and water molecules, and ignore them when adding the hydrogen atoms). Followed this we use the MCPB.py input file to generate the necessary files for the small, standard, and large models.

In order to keep the example straightforward there are only a few variables in the MCPB.py input file (see Figure 32). For more complicated cases users can consult the manual for the details concerning all of the available options. Here we specify the PDB

file name we are using (after the original_pdb variable name), the group name of the system (the default group name is “MOL” while here it is 1E67), the bond cut-off value between the metal ion and ligands (here it is 2.7 Å and the default is 2.8 - sometimes adjustment of this value is needed to specify which residue is bound to the metal ion), the atom IDs of the central ions in the metal site of the PDB file (here we only have one number because it is a single center metal site), and the mol2 file of the zinc ion (in which we specify its atom type as “ZN” and its charge as +2, and to keep consistent with the zinc ion in the PDB file, we treat the zinc ion with a residue name of “ZN” and an atom name of “ZN” in the mol2 file as well).

Figure 33 lists the commands used during the MCPB.py modeling procedure. First we use the command: “*MCPB.py -i 1E67.in -s 1*”, which generates the small, standard and large models of the metal ion coordination sphere. In the fingerprint file for the standard model (see the files supplied in SI of ref ²¹⁹), there is atom type specification for each atom of the residues in the metal site (*i.e.*, the central metal ion plus its ligating residues), and the linkage information between the metal ion and its ligating atoms. The third and fifth columns of atom information show the original and new atom types, respectively. The “-s 1” flag tells the program to assign the atom types for the central metal ion and the metal bound atoms automatically (each one is assigned differently). Users can change the atom types and/or linkage information manually based on their own preferences (*e.g.* by treating some of the bound atoms identically).

With all the files in hand we then use the small model to calculate the Hessian matrix and the large model to perform the Merz-Kollman population analysis²²¹ by using Gaussian03,¹⁰⁴ Gaussian09²²² or GAMESS-US.¹⁰⁵ By default MCPB.py treats the output files as being from Gaussian03. If Gaussian09 or GAMESS-US is being used, users need to add “software_version g09” or “software_version gms” as a line in the input file before performing the second and third steps. Users can also modify the QM input files based on their own needs, *i.e.* changing the number of CPUs, memory usage, level of theory, *etc.* Here we used the B3LYP/6-31G* level of theory to perform the Gaussian calculations of example 1.

For the force constant determination there are two substeps: the first substep is QM geometry optimization (line 2 in Figure 33) and the second substep is the calculation of the Cartesian Hessian matrix (line 3 in Figure 33). The Cartesian Hessian matrix, which will be used by the Seminario method to obtain force constants, was stored in the Gaussian binary chk file for Gaussian03 and Gaussian 09 (the chk file should be converted to a fchk file using the *formchk* program). Alternatively, this matrix could be in a GAMESS-US log file (if using GAMESS-US) after the second QM calculation substep. The internal force constants (based on the Z-matrix method) are stored in the Gaussian log file, while Z-matrix method is not supported using the GAMESS-US software in the current version of MCPB.py. For the Merz-Kollman population analysis of the large model, there is only one calculation performed (line 5 in Figure 33), in which the VDW radius for the zinc ion was taken from the CM parameter set of the Zn^{2+} ion for the TIP3P water model.

After performing the QM calculations we can use “*MCPB.py -i 1E67.in -s 2*” to generate the final force constant parameters using the Seminario method. The resultant metal ion related bond and angle parameters have the “Created by Seminario method using MCPB.py” comment specified in the end of such lines (see Figure 34). We observe that the generated parameters are in a physically meaningful range: the bond and angle force constants are all lower than 100 (AMBER *fcmod* files use the unit of kcal/mol•Å⁻² for the bond force constants and kcal/mol•Rad⁻² for the angle force constants). Following this step we use the “*MCPB.py -i 1E67.in -s 3*” command to perform a RESP charge fit for the large model and generate the mol2 files for the metal site residues. From the RESP fit we obtain the zinc ion has a partial charge of ~+0.5e, which represents a strong electron transfer effect from the ligating residues to the zinc ion. This is consistent with a previous work from Merz *et al.*,²²³ which showed that refitting the charges of the metal site, other than simply assigning a +2 charge to the zinc ion, is needed to accurately mimic the metal site charge distribution. Finally, we use the “*MCPB.py -i 1E67.in -s 4*” command to generate the PDB file with the renamed metal site residues and the leap source file for the system. Afterwards, the topology and coordinate files can be created using the “*tleap -s -f 1E67_tleap.in > 1E67_tleap.out*” command. With all files in hand we can perform minimization and MD simulations using AMBER, or transfer the topology and coordinate files to another format if desired (see Figure 28). In this example, we used pmemd.MPI from AMBER 14²²⁴ to perform the energy minimization, MD heating simulation and subsequent MD equilibration simulation. Afterwards, pmemd.cuda^{225, 226} from AMBER 14²²⁴ was used for the production MD simulation.

Except the parameters obtained from MCPB.py, we used the AMBER ff14SB FF to model the protein system. The protein structure was solvated in a rectangular TIP3P⁶⁵ water box (with size of $\sim(67\text{\AA}\times 57\text{\AA}\times 65\text{\AA})$ and 5549 water molecules) with water molecules at least 1.5 $\text{\AA}$ away from the protein surface. No counter-ions were added because the system was neutral. Afterwards four minimization steps were performed. 1000 steps of steepest descent minimization and 1000 steps of conjugated gradient minimization were performed for each of the first three stages. The first stage involved a $200\text{ kcal/mol}\cdot\text{\AA}^{-2}$ restraint on the protein, and the second stage used a $200\text{ kcal/mol}\cdot\text{\AA}^{-2}$ restraint on the heavy atoms in the protein, while the third stage used a $200\text{ kcal/mol}\cdot\text{\AA}^{-2}$ restraint on the backbone N, CA, C atoms of the protein. In the fourth stage of minimization, 2000 steps of steepest descent minimization followed by 3000 steps of conjugated gradient minimization were carried out without any restraints. Afterwards 1 ns of simulation was performed in the NVT ensemble to heat the system from 0 to 298.15 K, followed by another 1 ns of simulation in the NVT ensemble to equilibrate the system at 298.15 K. Then a 1 ns of simulation was carried at 298.15 K and 1 atm in the NPT ensemble to correct the system density. Finally, 20 ns of production simulation was performed at 298.15 K in NVT ensemble with snapshots stored every 10 ps. In total, 2000 frames were collected for the final analysis. The Langevin algorithm was used to control the temperature with a collision frequency of 1.0 ps^{-1} while the Berendsen barostat²²⁷ was employed to control the pressure with a relaxation time of 1.0 ps. SHAKE¹⁴¹ was used to constrain the bonds between the hydrogen atoms and their connected heavy atoms while a specific “three-point” algorithm¹⁴² was used for the water molecules.

Figure 35 shows the RMSD values of the protein backbone N, CA, C atoms and the metal site heavy atoms over 2000 snapshots across 20 ns of sampling. The RMSD values of the metal site heavy atoms are fluctuating around 0.3 Å while the RMSD values of the whole protein backbone N, CA, C atoms are ~1.0 Å. These results indicate that the metal site structure was well conserved during the simulation, thereby, further validating the parameterization accomplished by MCPB.py.

Os²⁺ complexes are attracting interests due to their special electronic properties.²²⁸ Here we treat the Os[(phen)₃]²⁺ complex as the second example to prove the ability of MCPB.py of parameterizing the organometallic compounds. In the present work we use the Os[(phen)₃]²⁺ structure of residue 1 of the Cambridge Structural Database (CSD) entry: FAJYAR01 (see Figure 36) from the work of Demadis *et al.*²²⁸ The content of the MCPB.py input file is shown in Figure 37. The commands used for processing the MCPB.py construction are shown in Figure 38.

We employed the B3LYP/SDD level of theory to do the QM calculations. The VDW radius of Os²⁺ during the Merz-Kollman population analysis was treated as 1.56 Å, which was adapted from the VDW parameters of the “Os6+6” atom type in UFF.² The finally obtained bond and angle parameters are shown in Figure 39.

The MM minimization and normal mode analysis were carried out using the nucleic acid builder (NAB) module in AmberTools15.¹⁰³ The minimization was performed with two steps in gas-phase and a cut-off value of 100 Å. The first step used the conjugate gradient

minimization algorithm with the convergence criterion of the energy gradient RMS set to 5×10^{-5} kcal/mol over a maximum of 20000 steps. The second step used the Newton-Raphson algorithm with the convergence criterion of energy gradient RMS set as 2×10^{-12} kcal/mol over a maximum of 200 steps. Afterwards the normal mode analysis was performed based on the minimized structure.

We have calculated the RMSD values between each pair of the CSD, QM optimized and MM minimized structures (see Table 39). It is noted that the CSD structure is from a crystal, with two PF_6^- anions and 1/2 water molecule for each $\text{Os}[(\text{phen})_3]^{2+}$ complex while the QM optimization and MM minimization were all performed in the gas phase for the independent Os complex. We can see that in general the RMSD values are relative small (less than 0.4 Å).

We performed a linear fit of the QM and MM calculated normal modes. The R^2 value of the fitting is ~ 0.99 , which is reasonable. Meanwhile, we also compared the QM and MM calculated normal modes in Figure 40, with the normal mode numbers along the X axis, while the frequencies are on the Y axis. Generally the agreement is good with the exception of some modes occupy the medium (around 1500 cm^{-1}) and high (around 3200 cm^{-1}) frequencies. This may due the underestimation of the force constants of the hydrogen involving bonds in GAFF, as shown in the work regarding parameterization of zinc complexes by Lin and Wang.¹⁰⁰ However, even improvement could be obtained after tuning these parameters, Lin and Wang noted that hydrogen containing bonds are usually

constrained using SHAKE.¹⁴¹ Hence there would only be a small impact whether these parameters were fully optimized or not.

4.5 Conclusions

Metal ion modeling in FFs remains a challenging issue due to the range of coordination modes available to transition metal ions. MCPB.py is an efficient tool that facilitates the construction of reliable FF parameters for metal ion containing systems utilizing the bonded model. It supports two widely used QM software packages - Gaussian and GAMESS-US for the parameterization process. It has an optimized code structure and has more options than the original C++ based MCPB program. Far fewer steps overall and fewer steps requiring user interventions are capable of affording a reliable metal ion FF. The ability of MCPB.py to handle FF parameterizations for different metal centers has been shown by two examples outlined. It has a GNU_GPL_v3 license and is free to download and distribute. It lowers the barrier of the molecular modeling for metal ion containing systems and offers a clean interface for non-expert users interested in doing related simulations.

CHAPTER 5: APPLICATION CASE OF THE CUSF METALLOCHAPERONE SYSTEM

The CusF metallochaperone system has a special cation- π interaction in its metal site (see Figure 41), which attracts people's interests.²²⁹ In a former research theoretical calculations showed this cation- π interaction is bigger than 10 kcal/mol, which stabilizes the metal site structure considerably.¹⁰² Based on MD simulations we observed that the *apo*-CusF system has an open-close cycle at the microsecond time scale.²³⁰ In a putative mechanism of metal ion transfer between $\text{Cu}^+\cdot\text{CusF}$ and *apo*-CusB, it is likely that the interactions between the proteins drive $\text{Cu}^+\cdot\text{CusF}$ to sample the open conformation. Hence the conformational change of CusF may play an important role in the metal ion transfer process.

In an attempt to estimate the free energy difference between the closed and open conformations of $\text{Cu}^+\cdot\text{CusF}$, we performed a two-dimensional PMF calculation. Here the hybrid model from a former research was used to describe the metal binding site: the chemical bonds between the Cu^+ ion and the three ligating atoms in the His36, Met47, and Met49 residues were represented by the bonded model, while the cation- π interaction between the Cu^+ ion and the Trp44 residue was represented by a 12-6 nonbonded model. The 12-6 LJ parameters of the CE3 and CZ3 atoms in the Trp44 aromatic ring were adjusted to reproduce the QM calculated interaction strength between the Cu^+ ion and Trp44.¹⁰² Our simulations of the *apo* and metal-bound CusF systems suggest that the distance between the Cu^+ ion and the CZ3 atom in Trp44 and the protein backbone Φ dihedral angle between Glu36 and Met47 provided a useful metric to distinguish between

the closed and open states of this metallochaperone (see Figure 42). In a closed conformation, the Cu^+ -CZ3(Trp44) distance is 2.33 Å and the dihedral angle is -147.53° (see Figure 42). While in a wide-open conformation observed in our simulation, the corresponding Cu^+ -CZ3(Trp44) distance increases to 17.42 Å and the dihedral angle changes to 47.61° (see Figure 42).

In order to efficiently generate conformations for a detailed sampling study, PMF simulations were first performed to transition from state 1 (closed state) to state 2 (wide-open state) along the diagonal described in Figure 43. The total reaction pathway was divided into a number of windows such that values were incremented in each successive window. In this process, the final snapshot of the starting window was treated as the initial structure of the subsequent window. Conformations were equilibrated for 1 ns of MD simulation for each window. Conformations from the diagonal scan were used as starting structures for a detailed scan along both coordinates while maintaining the same overlap between adjacent windows. For the initial scan, the total reaction coordinate pathway was divided into 39 windows with an increment of 0.5 Å and 6.48° in each successive window along the distance and dihedral coordinates respectively (see the left panel in Figure 43). The starting and final conformations from these calculations along the diagonal have coordinates of (2.33 Å, -147.52°) and (21.33 Å, 98.72°) respectively, along the Cu^+ -Trp44 distance and the Glu46-Met47 dihedral angle. After the initial scan, the scan directions went in two directions (see the right panel in Figure 43). Similar to the initial scan, each window was equilibrated for 1 ns of MD simulation, and the equilibrated structure from a window was used as a starting conformation for the

neighboring window(s). For the final scan, we sampled an area covering 39 windows along the Cu^+ -Trp44 distance, and 44 windows across the Glu46-Met47 dihedral angle, adding up to 1716 windows. 5 ns of production PMF MD simulation was performed after the 1 ns of equilibration for each of the 1716 windows. Step size was set to 1 fs and data points were stored after every 10 fs of sampling time for all these simulations. A harmonic potential with a force constant of $10 \text{ kcal/mol} \cdot \text{\AA}^{-2}$ was applied to the Cu^+ -Trp44(CZ3) distance while a harmonic potential using a force constant of $500 \text{ kcal/mol} \cdot \text{Rad}^{-2}$ was applied to the dihedral angle in these calculations. In order to sample adequately around minima, an additional 125 ns of biased MD was sampled for each of the four windows with coordinates of (15.83 $\text{\AA}$, -56.80°), (16.83 $\text{\AA}$, 46.88°), (16.33 $\text{\AA}$, -56.80°), and (17.33 $\text{\AA}$, 46.88°) which characterize two minima, yielding over 9 μs of PMF-MD sampling in the end. The data from these biased simulations was collected and unbiased using the weighted histogram analysis method (WHAM) to calculate a two-dimensional free energy profile.

While these PMF calculations are limited in terms of the conformational changes owing to the choice of coordinates, they provide us a lower estimate of the energetic penalty ($\sim 8 \text{ kcal/mol}$) for the protein to undergo this transition (see Figure 44). Afterwards, a nonbonded model of Cu^+ , employing the HFE parameter set of Cu^+ for the TIP4P_{EW} water model, was used to study the Cu^+ •CusF system and it was found that the Cu^+ •CusF rapidly transitions to an open or askance conformation in the process of metal ion release and then returns to its closed conformation.²³⁰ While these simulations are qualitative and cannot be relied upon at this time to give time-scale information, they do provide support

for the general hypothesis that the open conformation plays an important role in metal ion release in CusF function.

APPENDICES

APPENDIX A: TABLES

Table 1. Target values of the HFE, IOD and the CN of the first solvation shell for monovalent ions.

Ions	Mass (g/mol)	Electronic structures	HFE (kcal/mol)	IOD (Å) ^d	CN ^d
Li ⁺	6.94	[He]	-113.5 ^a	2.08±0.06	4-6
Na ⁺	22.99	[Ne]	-87.2 ^a	2.35±0.06	4-8
K ⁺	39.10	[Ar]	-70.5 ^a	2.79±0.08	6-8
Rb ⁺	85.47	[Kr]	-65.7 ^a	2.89±0.10	---
Cs ⁺	132.91	[Xe]	-59.8 ^a	3.13±0.07	7-8
Tl ⁺	204.38	[Xe]4f ¹⁴ 5d ¹⁰ 6s ²	-71.7 ^a	2.96 ^e	4 ^e
Cu ⁺	63.55	[Ar]3d ¹⁰	-125.5 ^a	1.87 ^f	2 ^f
Ag ⁺	107.87	[Kr]4d ¹⁰	-102.8 ^a	2.41±0.02	2-4
NH ₄ ⁺	18.04	--	-68.1 ^a	2.85 ^g	4-11 ^j
H ⁺ (Zundel cation)	1.007	--	-251.0 ^a	1.24 ^h	2 ^h
H ⁺ (Eigen cation)	1.007	--	-251.0 ^a	1.01 ⁱ	1 ⁱ
H ₃ O ⁺	19.02	--	-103.4 ^b	2.75±0.01	4
F ⁻	19.00	[Ne]	-119.7 ^c	2.63±0.02	4.1-6.8
Cl ⁻	35.45	[Ar]	-89.1 ^c	3.18±0.06	6-8.5
Br ⁻	79.90	[Kr]	-82.7 ^c	3.37±0.05	6
I ⁻	126.9	[Xe]	-74.3 ^c	3.64±0.03	6-8.7

a. From Marcus.¹¹⁵ b. From Palascak and Shields.¹²¹ c. From Schmid *et al.*¹¹⁶ d. From Marcus unless specified otherwise. e. Weighted average value of four bonds (two at 2.73 Å and two at 3.18 Å) from Persson *et al.*¹²⁴ f. From Burda *et al.*¹²⁷ g. Obtained by addition of the ionic radius of NH₄⁺ from Detellier and Laszlo¹²⁸ and the effective ionic radius of O²⁻ from Pauling.¹²⁹ h. From quantum calculations done by Meraj and Chaudhari.¹³² i. From Sobolewski and Domcke at the MP2/6-31+G** level of theory.¹³³ j. From Ohtaki and Radnai.¹⁶²

Table 2. Target values of the HFE, IOD and the CN of the first solvation shell for divalent metal ions.

Ions	Electron configuration	HFE (kcal/mol) ^a	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	CN	IOD (Å)
Be ²⁺	[He]	-572.4	-152.9	4 ^b	1.67 ^b
Cu ²⁺	[Ar]3d ⁹	-480.4	-60.9	6 ^b	Eq: 1.96±0.04 Ax: 2.40±0.10 ^b Weighted mean distance: 2.11 ^f
Ni ²⁺	[Ar]3d ⁸	-473.2	-53.7	6 ^b	2.06±0.01 ^b
Pt ²⁺	[Xe]4f ¹⁴ 5d ⁸	-468.5	-49.0	n	n
Zn ²⁺	[Ar]3d ¹⁰	-467.3	-47.8	6 ^b	2.09±0.06 ^b
Co ²⁺	[Ar]3d ⁷	-457.7	-38.2	6 ^b	2.10±0.02 ^b
Pd ²⁺	[Kr]4d ⁸	-456.5	-37.0	n	n
Ag ²⁺	[Kr]4d ⁹	-445.7	-26.2	n	n
Cr ²⁺	[Ar]3d ⁴	-442.2	-22.7	6 ^d	Eq: 2.08 ^d
Fe ²⁺	[Ar]3d ⁶	-439.8	-20.3	6 ^b	2.11±0.01 ^b
Mg ²⁺	[Ne]	-437.4	-17.9	6 ^b	2.09±0.04 ^b
V ²⁺	[Ar]3d ³	-436.2	-16.7	6 ^c	2.21 ^c
Mn ²⁺	[Ar]3d ⁵	-420.7	-1.2	6 ^b	2.19±0.01 ^b
Hg ²⁺	[Xe]4f ¹⁴ 5d ¹⁰	-420.7	-1.2	6 ^b	2.41 ^b
Cd ²⁺	[Kr]4d ¹⁰	-419.5	0.0	6 ^b	2.30±0.02 ^b
Yb ²⁺	[Xe]4f ¹⁴	-360.9	58.6	n	n
Ca ²⁺	[Ar]	-359.7	59.8	8 ^e	2.46 ^e
Sn ²⁺	[Kr]4d ¹⁰ 5s ²	-356.1	63.4	6 ^d	Eq: 2.33-2.34; ^d Ax: 2.38-2.90 ^d Weighted mean distance: 2.62 ^b
Pb ²⁺	[Xe]4f ¹⁴ 5d ¹⁰ 6s ²	-340.6	78.9	n	n
Eu ²⁺	[Xe]4f ⁷	-331	88.5	n	n
Sr ²⁺	[Kr]	-329.8	89.7	8-15 ^d	2.64±0.04 ^b
Sm ²⁺	[Xe]4f ⁶	-328.6	90.9	n	n
Ba ²⁺	[Xe]	-298.8	120.7	9 ^g	2.83 ^g
Ra ²⁺	[Rn]	-298.8	120.7	n	n

a. From Marcus.¹¹⁵ b. From Marcus.¹²³ c. From Miyanaga *et al.*²³¹ d. From Ohtaki and Radnai.¹⁶² e. From Jalilehvand *et al.*¹⁸⁷ f. Calculated by the authors from the experimental data. g. From Smirnov and Trostin.²³² n. Either no experimental data were available or the data were deemed unreliable by Ohtaki and Radnai.¹⁶²

Table 3. Target values of the HFE, IOD and the CN of the first solvation shell for trivalent and tetravalent metal ions.

Metal ion	Electron configuration	HFE (kcal/mol) ^a	IOD (Å) ^b	CN ^b	Effective Ion Radii (Å)	First Shell Water Radii (Å)
Al ³⁺	[Ne]	-1081.5	1.88	6	0.54	1.34
Fe ³⁺	[Ar]3d ⁵	-1019.4	2.03	6	0.65	1.38
Cr ³⁺	[Ar]3d ³	-958.4	1.96	6	0.62	1.34
In ³⁺	[Kr]4d ¹⁰	-951.2	2.15	6	0.80	1.35
Tl ³⁺	[Xe]4f ¹⁴ 5d ¹⁰	-948.9	2.23	4-6	0.89	1.34
Y ³⁺	[Kr]	-824.6	2.36	8	0.90	1.46
La ³⁺	[Xe]	-751.7	2.52	8.0-9.1	1.03	1.49
Ce ³⁺	[Xe]4f ¹	-764.8	2.55	7.5	1.01	1.54
Pr ³⁺	[Xe]4f ²	-775.6	2.54	9.2	0.99	1.55
Nd ³⁺	[Xe]4f ³	-783.9	2.47	8.0-8.9	0.98	1.49
Sm ³⁺	[Xe]4f ⁵	-794.7	2.44	8.0-9.9	0.96	1.48
Eu ³⁺	[Xe]4f ⁶	-803.1	2.45	8.3	0.95	1.50
Gd ³⁺	[Xe]4f ⁷	-806.6	2.39	8.0-9.9	0.94	1.45
Tb ³⁺	[Xe]4f ⁸	-812.6	2.40	8.0-8.2	0.92	1.48
Dy ³⁺	[Xe]4f ⁹	-818.6	2.37	7.4-7.9	0.91	1.46
Er ³⁺	[Xe]4f ¹¹	-835.3	2.36	6.3-8.2	0.89	1.47
Tm ³⁺	[Xe]4f ¹²	-840.1	2.36	8.1	0.88	1.48
Lu ³⁺	[Xe]4f ¹⁴	-840.1	2.34	8	0.86	1.48
Hf ⁴⁺	[Xe]4f ¹⁴	-1664.7	2.16 ^c	8 ^c	0.85	1.31
Zr ⁴⁺	[Kr]	-1622.8	2.19 ^c	8 ^c	0.86	1.33
Ce ⁴⁺	[Xe]	-1462.7	2.42 ^d	9 ^d	0.87	1.55
U ⁴⁺	[Rn]5f ¹ 6d ¹	-1567.9	2.42 ^e	9-11 ^e	0.89	1.53
Pu ⁴⁺	[Rn]5f ⁴	-1520.1	2.39 ^f	8 ^f	0.86	1.53
Th ⁴⁺	[Rn]	-1389.8	2.45 ^e	9-11 ^e	0.94	1.51

a. From Marcus unless specified.¹¹⁵ b. Referenced or calculated from data from Marcus.¹²³ c. From Hagfeldt *et al.*¹³⁹ d. From Sham.¹³⁶ e. From Moll *et al.*¹³⁸ f. From Ankudinov *et al.*¹³⁷

Table 4. Parameters of four different water models.

Water model	Q(O)	Q(H) or Q(M) ^a	r(O-H) (Å)	H-O-H angle (degree)	r(O-M) ^b (Å)	R _{min} /2 for O atom (Å)	ε for O atom (kcal/mol)
TIP3P	-0.834	+0.417	0.9572	104.52	-----	1.7683	0.1520
SPC/E	-0.8476	+0.4238	1.0	109.47	-----	1.7767	0.1553
TIP4P	-1.04	+0.52	0.9572	104.52	0.15	1.7699	0.1550
TIP4P _{EW}	-1.04844	+0.5242	0.9572	104.52	0.125	1.77593	0.16275

a. Q(M) is for TIP4P and TIP4P_{EW} which are 4-site water models while Q(O) is for TIP3P and SPC/E which are 3-site water models. M represents the dummy atom in a 4-site water model. b. Distance between the dummy atom and oxygen atom in the water model, only valid for 4-site water models.

Table 5. Absolute HFEs, IODs and CNs of M(II) ions employing parameters available in the AMBER package.

	LJ parameters		Results from Method 2			Results from Åqvist ³⁷	
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)
Zn ²⁺	1.10	0.0125	-443.8	1.93	6.0	-----	-----
Mg ²⁺	0.7926	0.8947	-432.6	1.99	6.0	-455.9±2.6	2.00
Ca ²⁺	1.7131	0.459789	-307.0	2.70	8.9	-380.6±1.3	2.40

Table 6. Absolute HFEs, IODs and CNs of Zn(II) employing parameters developed by Stote and Karplus.

	LJ parameters		Results from Method 2			Results from Stote and Karplus ³⁸		
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN
Zn^{2+}	1.094	0.250	-399.9	2.10	6.0	-472.7	2.12	6.0

Table 7. Absolute HFEs, relative HFEs (relative to the Cd^{2+} ion), IODs and CNs of M(II) ions with the parameters taken from Babu and Lim.

	LJ parameters		Results from Method 2				Results from Babu and Lim ⁶⁴		
Ions	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	HFE (kcal/mol)	Relative HFE (kcal/mol)	IOD (Å)	CN	Relative HFE (kcal/mol)	IOD (Å)	CN
Be^{2+}	0.5637	0.0032	-521.4	-145.7	1.45	3.3	-156.3	1.57	4
Cu^{2+}	1.033	0.0427	-436.5	-60.8	1.96	6.0	-59.9	1.94	6
Ni^{2+}	1.0941	0.0366	-430.3	-54.6	1.98	6.0	-53.2	1.97	6
Pt^{2+}	1.1376	0.0332	-425.6	-49.9	1.99	6.0	-48.3	1.97	6
Zn^{2+}	1.1489	0.0325	-423.7	-48.0	2.00	6.0	-47.5	2.00	6
Co^{2+}	1.2267	0.0286	-414.6	-38.9	2.03	6.0	-38.0	2.02	6
Pd^{2+}	1.236	0.0282	-413.3	-37.6	2.03	6.0	-37.3	2.02	6
Ag^{2+}	1.3107	0.0266	-403.1	-27.4	2.07	6.0	-26.6	2.06	6
Cr^{2+}	1.3344	0.0264	-399.2	-23.5	2.08	6.0	-22.4	2.07	6
Fe^{2+}	1.3488	0.0264	-397.0	-21.3	2.09	6.0	-20.5	2.08	6
Mg^{2+}	1.3636	0.0266	-394.4	-18.7	2.10	6.0	-17.5	2.08	6
V^{2+}	1.3706	0.0266	-393.5	-17.8	2.10	6.0	-16.2	2.11	6
Mn^{2+}	1.4544	0.03	-377.2	-1.5	2.17	6.1	-1.5	2.16	6
Cd^{2+}	1.46	0.0304	-375.7	0.0	2.18	6.1	0.0	2.17	6
Yb^{2+}	1.9298	0.0309	-317.2	58.5	2.57	8.3	57.7	2.47	8
Ca^{2+}	1.9364	0.0318	-316.7	59.0	2.58	8.3	58.1	2.58	8
Sn^{2+}	1.954	0.0346	-313.2	62.5	2.61	8.5	62.6	2.58	8
Pb^{2+}	2.0195	0.0557	-298.2	77.5	2.71	8.9	78.2	2.68	8.5
Eu^{2+}	2.0846	0.0647	-288.7	87.0	2.78	9.0	88.8	2.74	9
Sr^{2+}	2.0923	0.0664	-287.8	87.9	2.79	9.0	89.3	2.75	9
Sm^{2+}	2.0997	0.068	-286.5	89.2	2.79	9.0	90.1	2.75	9
Ba^{2+}	2.2451	0.1993	-257.8	117.9	3.04	9.8	120.4	3.01	9.5

Table 8. Standard deviations of the IOD values for the LJ grids of divalent metal ions.

	TIP3P	SPC/E	TIP4P	TIP4P _{EW}
TIP3P	0.00Å	0.00Å±0.01Å	0.02Å±0.01Å	0.02Å±0.01Å
SPC/E	-----	0.00Å	0.02Å±0.02Å	0.02Å±0.01Å
TIP4P	-----	-----	0.00Å	0.00Å±0.01Å
TIP4P _{EW}	-----	-----	-----	0.00Å

Table 9. The experimental LJ parameters for noble gas atoms.^a

	$R_{\min}$ (Å)	ϵ (meV)	$R_{\min}/2$ (Å)	ϵ (kcal/mol) ^b	$-\log(\epsilon)$
He	2.97	0.92	1.485	0.02121603	1.67333588
Ne	3.10	3.6	1.55	0.08301924	1.08082125
Ar	3.76	12.2	1.88	0.28134298	0.55076392
Ke	4.00	17.2	2.00	0.39664748	0.40159530
Xe	4.40	24	2.20	0.55346160	0.25691251

a. Adapted from page 408 of the book of Radētisig *et al.*¹⁵⁹ according to the Lorentz-Berthelot combining rules. b. Using the conversion factor 1 eV=23.0609 kcal/mol.

Table 10. Comparison of results of different scaling methods in the VDW-disappearing and VDW-appearing steps of the HFE calculations for the Be^{2+} and Ba^{2+} ions.^a

		VDW-appearing				VDW-disappearing			
		L	K4	K6	SC	L	K4	K6	SC
Be^{2+}	3 Windows	-----	-----	-----	0.44	-----	-----	-----	-0.45
	5 Windows	0.25	0.19	0.17	-----	-0.26	-0.38	-0.38	-----
	7 Windows	0.28	0.24	0.23	-----	-0.29	-0.38	-0.39	-----
	9 Windows	0.30	0.27	0.25	-----	-0.29	-0.39	-0.37	-----
Ba^{2+}	3 Windows	-----	-----	-----	1.56	-----	-----	-----	-1.41
	5 Windows	-0.40	-1.34	-1.74	-----	0.54	-2.08	-1.76	-----
	7 Windows	0.12	-0.74	-1.10	-----	0.05	-1.83	-1.90	-----
	9 Windows	0.31	-0.45	-0.59	-----	-0.18	-1.74	-1.95	-----

a. In kcal/mol.

Table 11. Final optimized HFE parameter set of divalent metal ions for different water models.

	TIP3P		SPC/E		TIP4P _{EW}	
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)
Be ²⁺	0.907	0.00000080	0.915	0.00000105	0.815	0.0000000221
Cu ²⁺	1.144	0.00040203	1.149	0.00044254	1.078	0.00010063
Ni ²⁺	1.162	0.00056491	1.166	0.00060803	1.101	0.00016733
Pt ²⁺	1.173	0.00069036	1.176	0.00072849	1.114	0.00022027
Zn ²⁺	1.175	0.00071558	1.178	0.00075490	1.115	0.00022490
Co ²⁺	1.211	0.00132548	1.217	0.00146124	1.141	0.00037931
Pd ²⁺	1.215	0.00141473	1.217	0.00146124	1.145	0.00040986
Ag ²⁺	1.263	0.00294683	1.265	0.00303271	1.171	0.00066591
Cr ²⁺	1.273	0.00339720	1.276	0.00354287	1.181	0.00079606
Fe ²⁺	1.277	0.00359255	1.284	0.00395662	1.194	0.00099751
Mg ²⁺	1.284	0.00395662	1.288	0.00417787	1.208	0.00126172
V ²⁺	1.290	0.00429223	1.293	0.00446856	1.210	0.00130393
Mn ²⁺ /Hg ²⁺	1.339	0.00799176	1.338	0.00789684	1.276	0.00354287
Cd ²⁺	1.339	0.00799176	1.344	0.00848000	1.279	0.00369364
Yb ²⁺	1.526	0.04772212	1.518	0.04490976	1.464	0.02883819
Ca ²⁺	1.528	0.04844326	1.520	0.04560206	1.467	0.02960343
Sn ²⁺	1.543	0.05408454	1.532	0.04990735	1.479	0.03280986
Pb ²⁺	1.620	0.08965674	1.609	0.08389240	1.551	0.05726270
Eu ²⁺	1.666	0.11617738	1.656	0.11008622	1.596	0.07737276
Sr ²⁺	1.672	0.11991675	1.659	0.11189491	1.606	0.08235966
Sm ²⁺	1.680	0.12499993	1.667	0.11679623	1.606	0.08235966
Ba ²⁺ /Ra ²⁺	1.842	0.24821230	1.825	0.23380842	1.768	0.18767274

Table 12. The simulated HFE, IOD and CN values of the HFE parameter set for divalent metal ions.^a

	TIP3P				SPC/E			
	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN
Be ²⁺	-572.3	-152.5	1.14	2.0	-571.4	-152.4	1.15	2.0
Cu ²⁺	-481.2	-61.4	1.63	4.0	-481.2	-62.2	1.64	4.0
Ni ²⁺	-473.0	-53.2	1.65	4.0	-473.2	-54.2	1.66	4.0
Pt ²⁺	-467.8	-48.0	1.67	4.1	-468.4	-49.4	1.67	4.1
Zn ²⁺	-467.4	-47.6	1.67	4.1	-467.3	-48.3	1.68	4.3
Co ²⁺	-457.0	-37.2	1.87	6.0	-456.9	-37.9	1.89	6.0
Pd ²⁺	-457.0	-37.2	1.88	6.0	-456.8	-37.8	1.89	6.0
Ag ²⁺	-445.0	-25.2	1.93	6.0	-446.0	-27.0	1.94	6.0
Cr ²⁺	-441.6	-21.8	1.94	6.0	-441.9	-22.9	1.95	6.0
Fe ²⁺	-439.5	-19.7	1.94	6.0	-439.4	-20.4	1.96	6.0
Mg ²⁺	-437.7	-17.9	1.95	6.0	-437.6	-18.6	1.96	6.0
V ²⁺	-435.9	-16.1	1.95	6.0	-435.8	-16.8	1.96	6.0
Mn ²⁺	-419.9	-0.1	2.01	6.0	-420.6	-1.6	2.01	6.0
Hg ²⁺	-419.9	-0.1	2.01	6.0	-420.6	-1.6	2.01	6.0
Cd ²⁺	-419.8	0.0	2.01	6.0	-419.0	0.0	2.02	6.0
Yb ²⁺	-360.2	59.6	2.33	7.4	-360.2	58.8	2.30	7.0
Ca ²⁺	-360.2	59.6	2.33	7.4	-360.6	58.4	2.31	7.0
Sn ²⁺	-356.5	63.3	2.36	7.7	-356.4	62.6	2.33	7.2
Pb ²⁺	-340.9	78.9	2.46	8.0	-340.4	78.6	2.45	7.9
Eu ²⁺	-331.2	88.6	2.51	8.0	-330.9	88.1	2.51	8.0
Sr ²⁺	-329.7	90.1	2.52	8.1	-330.6	88.4	2.51	8.0
Sm ²⁺	-328.4	91.4	2.53	8.1	-328.6	90.4	2.52	8.0
Ba ²⁺	-299.4	120.4	2.74	9.0	-299.2	119.8	2.72	8.8
Ra ²⁺	-299.4	120.4	2.74	9.0	-299.2	119.8	2.72	8.8
Average Error	0.1	0.4	-0.27	-0.4	0.0	-0.5	-0.26	-0.5
Standard Deviation	0.5	0.5	0.14	1.0	0.5	0.5	0.14	1.0
UAE	0.4	0.5	0.27	0.6	0.4	0.6	0.26	0.6

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 2). The target IOD values in Table 2 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 12 (cont'd)

	TIP4P _{EW}			
	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN
Be ²⁺	-572.8	-153.2	0.87	1.0
Cu ²⁺	-481.1	-61.5	1.57	4.0
Ni ²⁺	-472.4	-52.8	1.60	4.0
Pt ²⁺	-468.0	-48.4	1.61	4.0
Zn ²⁺	-467.4	-47.8	1.61	4.0
Co ²⁺	-458.6	-39.0	1.64	4.0
Pd ²⁺	-456.1	-36.5	1.65	4.0
Ag ²⁺	-445.2	-25.6	1.69	4.1
Cr ²⁺	-442.3	-22.7	1.84	5.6
Fe ²⁺	-440.4	-20.8	1.88	6.0
Mg ²⁺	-436.5	-16.9	1.89	6.0
V ²⁺	-435.6	-16.0	1.90	6.0
Mn ²⁺	-419.9	-0.3	1.96	6.0
Hg ²⁺	-419.9	-0.3	1.96	6.0
Cd ²⁺	-419.6	0.0	1.96	6.0
Yb ²⁺	-361.3	58.3	2.19	6.1
Ca ²⁺	-359.8	59.8	2.20	6.2
Sn ²⁺	-356.2	63.4	2.23	6.5
Pb ²⁺	-339.7	79.9	2.39	7.7
Eu ²⁺	-331.0	88.6	2.45	7.9
Sr ²⁺	-329.5	90.1	2.46	8.0
Sm ²⁺	-328.8	90.8	2.46	8.0
Ba ²⁺	-299.1	120.5	2.67	8.8
Ra ²⁺	-299.1	120.5	2.67	8.8
Average Error	0.1	0.2	-0.36	-0.8
Standard Deviation	0.5	0.5	0.17	1.1
UAE	0.5	0.4	0.36	0.9

Table 13. Final optimized IOD parameter set of divalent metal ions.^a

	$R_{\min}/2$ (Å)	ϵ (kcal/mol)
Be ²⁺	1.168	0.00063064
Cu ²⁺	1.409	0.01721000
Ni ²⁺	1.373	0.01179373
Zn ²⁺	1.395	0.01491700
Co ²⁺	1.404	0.01636246
Cr ²⁺	1.388	0.01386171
Fe ²⁺	1.409	0.01721000
Mg ²⁺	1.395	0.01491700
V ²⁺	1.476	0.03198620
Mn ²⁺	1.467	0.02960343
Hg ²⁺	1.575	0.06751391
Cd ²⁺	1.506	0.04090549
Ca ²⁺	1.608	0.08337961
Sn ²⁺	1.738	0.16500296
Sr ²⁺	1.753	0.17618319
Ba ²⁺	1.913	0.31060194

a. The parameters are same for the four different water models.

Table 14. The simulated HFE, IOD and CN values of the IOD parameter set for divalent metal ions.^a

	TIP3P				SPC/E			
	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN
Be ²⁺	-469.8	-103.9	1.66	4.0	-472.0	-108.1	1.66	4.0
Cu ²⁺	-395.2	-29.3	2.10	6.0	-395.4	-31.5	2.10	6.0
Ni ²⁺	-407.6	-41.7	2.05	6.0	-408.2	-44.3	2.05	6.0
Zn ²⁺	-400.0	-34.1	2.08	6.0	-400.0	-36.1	2.08	6.0
Co ²⁺	-397.0	-31.1	2.09	6.0	-397.1	-33.2	2.09	6.0
Cr ²⁺	-402.8	-36.9	2.07	6.0	-403.2	-39.3	2.07	6.0
Fe ²⁺	-394.8	-28.9	2.10	6.0	-395.3	-31.4	2.10	6.0
Mg ²⁺	-400.3	-34.4	2.08	6.0	-400.5	-36.6	2.08	6.0
V ²⁺	-372.8	-6.9	2.21	6.6	-372.2	-8.3	2.19	6.1
Mn ²⁺	-375.6	-9.7	2.18	6.3	-375.4	-11.5	2.18	6.0
Hg ²⁺	-350.1	15.8	2.41	7.9	-346.4	17.5	2.40	7.6
Cd ²⁺	-365.9	0.0	2.29	7.1	-363.9	0.0	2.28	6.8
Ca ²⁺	-342.3	23.6	2.45	8.0	-340.6	23.3	2.45	7.9
Sn ²⁺	-317.6	48.3	2.61	8.6	-314.8	49.1	2.61	8.2
Sr ²⁺	-314.2	51.7	2.63	8.6	-311.8	52.1	2.62	8.3
Ba ²⁺	-288.2	77.7	2.82	9.2	-285.1	78.8	2.82	9.0
Average Error	51.1	-2.5	-0.01	0.4	51.9	-3.7	-0.01	0.3
Standard Deviation	25.2	25.2	0.00	0.8	24.3	24.3	0.00	0.7
UAE	51.1	20.0	0.01	0.4	51.9	19.7	0.01	0.3

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 2). The target IOD values in Table 2 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 14 (cont'd)

	TIP4P				TIP4P _{EW}			
	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN
Be ²⁺	-435.0	-91.9	1.68	4.4	-446.8	-97.3	1.68	4.0
Cu ²⁺	-368.7	-25.6	2.11	6.0	-378.6	-29.1	2.12	6.0
Ni ²⁺	-380.9	-37.8	2.07	6.0	-390.4	-40.9	2.07	6.0
Zn ²⁺	-373.5	-30.4	2.10	6.0	-382.0	-32.5	2.10	6.0
Co ²⁺	-370.7	-27.6	2.11	6.0	-379.6	-30.1	2.11	6.0
Cr ²⁺	-376.0	-32.9	2.09	6.0	-385.0	-35.5	2.09	6.0
Fe ²⁺	-369.3	-26.2	2.11	6.0	-378.4	-28.9	2.12	6.0
Mg ²⁺	-373.9	-30.8	2.09	6.0	-382.9	-33.4	2.10	6.0
V ²⁺	-349.5	-6.4	2.26	7.0	-356.2	-6.7	2.22	6.3
Mn ²⁺	-351.6	-8.5	2.22	6.6	-359.9	-10.4	2.20	6.1
Hg ²⁺	-329.4	13.7	2.42	8.0	-335.1	14.4	2.42	7.9
Cd ²⁺	-343.1	0.0	2.31	7.4	-349.5	0.0	2.31	7.1
Ca ²⁺	-323.3	19.8	2.46	8.0	-329.2	20.3	2.47	8.0
Sn ²⁺	-299.8	43.3	2.63	8.8	-305.2	44.3	2.63	8.6
Sr ²⁺	-297.7	45.4	2.65	8.8	-301.7	47.8	2.65	8.7
Ba ²⁺	-273.3	69.8	2.84	9.4	-276.9	72.6	2.84	9.2
Average Error	74.8	-1.6	0.01	0.5	67.2	-2.8	0.01	0.4
Standard Deviation	29.3	29.3	0.01	0.9	27.5	27.5	0.00	0.8
UAE	74.8	22.3	0.01	0.5	67.2	21.5	0.01	0.4

Table 15. Final optimized CM parameter set of divalent metal ions for different water models.

	TIP3P		SPC/E		TIP4P _{EW}	
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)
Be ²⁺	0.956	0.00000395	0.961	0.00000460	0.918	0.00000116
Cu ²⁺	1.218	0.00148497	1.223	0.00160860	1.195	0.00101467
Ni ²⁺	1.255	0.00262320	1.253	0.00254709	1.221	0.00155814
Pt ²⁺	1.266	0.00307642	1.272	0.00334975	1.251	0.00247282
Zn ²⁺	1.271	0.00330286	1.276	0.00354287	1.252	0.00250973
Co ²⁺	1.299	0.00483892	1.305	0.00523385	1.288	0.00417787
Pd ²⁺	1.303	0.00509941	1.305	0.00523385	1.288	0.00417787
Ag ²⁺	1.336	0.00770969	1.337	0.00780282	1.323	0.00657749
Cr ²⁺	1.346	0.00868178	1.348	0.00888732	1.333	0.00743559
Fe ²⁺	1.353	0.00941798	1.354	0.00952704	1.343	0.00838052
Mg ²⁺	1.360	0.01020237	1.360	0.01020237	1.353	0.00941798
V ²⁺	1.364	0.01067299	1.365	0.01079325	1.353	0.00941798
Mn ²⁺ /Hg ²⁺	1.407	0.01686710	1.406	0.01669760	1.401	0.01586934
Cd ²⁺	1.412	0.01773416	1.412	0.01773416	1.406	0.01669760
Yb ²⁺	1.642	0.10185975	1.634	0.09731901	1.654	0.10888937
Ca ²⁺	1.649	0.10592870	1.635	0.09788018	1.657	0.11068733
Sn ²⁺	1.666	0.11617738	1.651	0.10710756	1.670	0.11866330
Pb ²⁺	1.745	0.17018074	1.731	0.15989650	1.758	0.17997960
Eu ²⁺	1.802	0.21475916	1.786	0.20184160	1.823	0.23213110
Sr ²⁺	1.810	0.22132374	1.794	0.20826406	1.827	0.23548950
Sm ²⁺	1.819	0.22878796	1.800	0.21312875	1.838	0.24480038
Ba ²⁺ /Ra ²⁺	2.019	0.40664608	1.980	0.37126402	2.050	0.43454345

Table 16. The simulated HFE, IOD and CN values of the CM parameter set for divalent metal ions.^a

	TIP3P				SPC/E			
	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN
Be ²⁺	-547.3	-153.7	1.21	2.0	-547.5	-153.0	1.22	2.0
Cu ²⁺	-455.0	-61.4	1.88	6.0	-454.6	-60.1	1.90	6.0
Ni ²⁺	-447.9	-54.3	1.92	6.0	-448.4	-53.9	1.92	6.0
Pt ²⁺	-443.7	-50.1	1.93	6.0	-443.8	-49.3	1.94	6.0
Zn ²⁺	-442.0	-48.4	1.93	6.0	-441.6	-47.1	1.95	6.0
Co ²⁺	-433.0	-39.4	1.96	6.0	-431.9	-37.4	1.98	6.0
Pd ²⁺	-431.5	-37.9	1.97	6.0	-432.2	-37.7	1.98	6.0
Ag ²⁺	-420.8	-27.2	2.00	6.0	-420.4	-25.9	2.01	6.0
Cr ²⁺	-417.3	-23.7	2.02	6.0	-417.0	-22.5	2.03	6.0
Fe ²⁺	-414.6	-21.0	2.02	6.0	-415.3	-20.8	2.03	6.0
Mg ²⁺	-412.1	-18.5	2.03	6.0	-412.8	-18.3	2.04	6.0
V ²⁺	-410.7	-17.1	2.04	6.0	-411.4	-16.9	2.05	6.0
Mn ²⁺	-396.1	-2.5	2.09	6.0	-396.5	-2.0	2.10	6.0
Hg ²⁺	-396.1	-2.5	2.09	6.0	-396.5	-2.0	2.10	6.0
Cd ²⁺	-393.6	0.0	2.10	6.0	-394.5	0.0	2.10	6.0
Yb ²⁺	-335.7	57.9	2.48	8.0	-335.3	59.2	2.48	8.0
Ca ²⁺	-334.6	59.0	2.49	8.0	-334.5	60.0	2.48	8.0
Sn ²⁺	-331.2	62.4	2.51	8.1	-331.5	63.0	2.50	8.0
Pb ²⁺	-316.2	77.4	2.62	8.7	-315.7	78.8	2.60	8.2
Eu ²⁺	-305.8	87.8	2.69	8.9	-305.2	89.3	2.67	8.5
Sr ²⁺	-304.9	88.7	2.70	8.9	-304.5	90.0	2.68	8.6
Sm ²⁺	-303.0	90.6	2.71	8.9	-302.7	91.8	2.69	8.6
Ba ²⁺	-273.1	120.5	2.94	9.7	-274.7	119.8	2.90	9.2
Ra ²⁺	-273.1	120.5	2.94	9.7	-274.7	119.8	2.90	9.2
Average Error	25.1	-0.8	-0.13	0.0	25.0	0.0	-0.12	0.0
Standard Deviation	0.4	0.4	0.14	0.8	0.6	0.6	0.13	0.7
UAE	25.1	0.8	0.15	0.3	25.0	0.5	0.14	0.3

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 2). The target IOD values in Table 2 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 16 (cont'd)

	TIP4P _{EW}			
	HFE (kcal/mol)	Relative HFE (M ²⁺ -Cd ²⁺) (kcal/mol)	IOD (Å)	CN
Be ²⁺	-532.2	-153.6	1.17	2.0
Cu ²⁺	-439.9	-61.3	1.88	6.0
Ni ²⁺	-432.4	-53.8	1.91	6.0
Pt ²⁺	-427.8	-49.2	1.94	6.0
Zn ²⁺	-427.8	-49.2	1.94	6.0
Co ²⁺	-416.8	-38.2	1.97	6.0
Pd ²⁺	-416.8	-38.2	1.97	6.0
Ag ²⁺	-405.1	-26.5	2.01	6.0
Cr ²⁺	-403.1	-24.5	2.02	6.0
Fe ²⁺	-400.0	-21.4	2.03	6.0
Mg ²⁺	-396.4	-17.8	2.05	6.0
V ²⁺	-395.6	-17.0	2.05	6.0
Mn ²⁺	-380.7	-2.1	2.11	6.0
Hg ²⁺	-380.7	-2.1	2.11	6.0
Cd ²⁺	-378.6	0.0	2.11	6.0
Yb ²⁺	-320.1	58.5	2.52	8.0
Ca ²⁺	-319.0	59.6	2.53	8.0
Sn ²⁺	-316.4	62.2	2.54	8.1
Pb ²⁺	-301.3	77.3	2.66	8.7
Eu ²⁺	-290.6	88.0	2.74	8.9
Sr ²⁺	-289.6	89.0	2.74	8.9
Sm ²⁺	-288.1	90.5	2.75	9.0
Ba ²⁺	-258.4	120.2	3.00	9.9
Ra ²⁺	-258.4	120.2	3.00	9.9
Average Error	40.3	-0.6	-0.11	0.1
Standard Deviation	0.5	0.5	0.16	0.8
UAE	40.3	0.6	0.16	0.3

Table 17. Final optimized HFE parameter set of monovalent ions for different water models.

	TIP3P		SPC/E		TIP4P _{EW}	
	R _{min} /2 (Å)	ε (kcal/mol)	R _{min} /2 (Å)	ε (kcal/mol)	R _{min} /2 (Å)	ε (kcal/mol)
Li ⁺	1.267	0.00312065	1.258	0.00274091	1.226	0.00168686
Na ⁺	1.475	0.03171494	1.454	0.02639002	1.432	0.02154025
K ⁺	1.719	0.15131351	1.683	0.12693448	1.669	0.11803919
Rb ⁺	1.834	0.24140216	1.792	0.20665151	1.767	0.18689752
Cs ⁺	1.988	0.37853483	1.953	0.34673208	1.936	0.33132862
Tl ⁺	1.703	0.14021803	1.668	0.11741683	1.646	0.10417397
Cu ⁺	1.201	0.00112300	1.192	0.00096394	1.156	0.00050520
Ag ⁺	1.341	0.00818431	1.330	0.00716930	1.294	0.00452863
NH ₄ ⁺	1.779	0.19628399	1.743	0.16869420	1.707	0.14295367
H ₃ O ⁺	1.337	0.00780282	1.327	0.00691068	1.292	0.00440914
F ⁻	1.783	0.19945255	1.819	0.22878796	1.842	0.24821230
Cl ⁻	2.252	0.60293097	2.308	0.64367011	2.321	0.65269755
Br ⁻	2.428	0.72070940	2.470	0.74435812	2.520	0.77034233
I ⁻	2.724	0.85418187	2.770	0.86877007	2.819	0.88281946

Table 18. The simulated HFE, IOD and CN values of the HFE parameter set for monovalent metal ions.^a

	TIP3P			SPC/E		
	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN
Li ⁺	-113.1	1.95	4.1	-113.5	1.94	4.0
Na ⁺	-88.2	2.36	5.8	-87.9	2.33	5.6
K ⁺	-71.2	2.73	7.0	-70.7	2.69	6.7
Rb ⁺	-65.9	2.90	8.0	-66.1	2.85	7.5
Cs ⁺	-60.5	3.10	9.2	-60.2	3.09	8.9
Tl ⁺	-72.0	2.72	6.6	-71.5	2.67	6.6
Cu ⁺	-125.4	1.84	4.0	-126.1	1.83	4.0
Ag ⁺	-102.7	2.14	5.3	-102.6	2.10	4.9
NH ₄ ⁺	-68.4	2.82	7.7	-68.0	2.77	7.1
H ₃ O ⁺	-103.3	2.13	5.2	-102.8	2.10	4.9
F ⁻	-119.1	2.71	6.9	-119.7	2.74	6.5
Cl ⁻	-89.5	3.30	8.9	-89.1	3.34	7.9
Br ⁻	-83.1	3.48	8.2	-83.0	3.52	7.5
I ⁻	-74.9	3.79	8.8	-74.5	3.81	9.5
Average Error	-0.2	-0.05	--	-0.1	-0.06	--
Standard Deviation	0.5	0.20	--	0.3	0.22	--
UAE	0.4	0.13	--	0.3	0.15	--

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 1). The target IOD values in Table 1 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 18 (cont'd)

	TIP4P _{EW}		
	HFE (kcal/mol)	IOD (Å)	CN
Li ⁺	-114.0	1.90	4.0
Na ⁺	-88.0	2.32	5.8
K ⁺	-70.8	2.68	6.6
Rb ⁺	-65.8	2.83	7.4
Cs ⁺	-59.8	3.05	8.5
Tl ⁺	-71.7	2.65	6.4
Cu ⁺	-125.5	1.79	4.0
Ag ⁺	-103.2	2.05	4.8
NH ₄ ⁺	-68.8	2.74	6.8
H ₃ O ⁺	-103.6	2.05	4.7
F ⁻	-120.2	2.77	6.3
Cl ⁻	-89.5	3.35	7.4
Br ⁻	-82.4	3.56	7.6
I ⁻	-74.3	3.87	9.2
Average Error	-0.3	-0.11	--
Standard Deviation	0.3	0.28	--
UAE	0.3	0.22	--

Table 19. Final optimized IOD parameter set of monovalent ions for all the three water models.

Ions	$R_{\min}/2$ (Å)	ϵ (kcal/mol)
Li^+	1.315	0.00594975
Na^+	1.465	0.02909167
K^+	1.745	0.17018074
Rb^+	1.820	0.22962229
Cs^+	2.000	0.38943250
Tl^+	1.870	0.27244486
Cu^+	1.214	0.00139196
Ag^+	1.500	0.03899838
NH_4^+	1.790	0.20504355
H^+ (Zundel cation)	0.925	0.00000147
H^+ (Eigen cation)	0.841	0.0000000661
H_3O^+	1.720	0.15202035
F^-	1.739	0.16573832
Cl^-	2.162	0.53154665
Br^-	2.331	0.65952968
I^-	2.590	0.80293907

Table 20. The simulated HFE, IOD and CN values of the IOD parameter set for monovalent metal ions.^a

	TIP3P			SPC/E		
	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN
Li ⁺	-113.1	1.95	4.1	-113.5	1.94	4.0
Na ⁺	-88.2	2.36	5.8	-87.9	2.33	5.6
K ⁺	-71.2	2.73	7.0	-70.7	2.69	6.7
Rb ⁺	-65.9	2.90	8.0	-66.1	2.85	7.5
Cs ⁺	-60.5	3.10	9.2	-60.2	3.09	8.9
Tl ⁺	-72.0	2.72	6.6	-71.5	2.67	6.6
Cu ⁺	-125.4	1.84	4.0	-126.1	1.83	4.0
Ag ⁺	-102.7	2.14	5.3	-102.6	2.10	4.9
NH ₄ ⁺	-68.4	2.82	7.7	-68.0	2.77	7.1
H ₃ O ⁺	-103.3	2.13	5.2	-102.8	2.10	4.9
F ⁻	-119.1	2.71	6.9	-119.7	2.74	6.5
Cl ⁻	-89.5	3.30	8.9	-89.1	3.34	7.9
Br ⁻	-83.1	3.48	8.2	-83.0	3.52	7.5
I ⁻	-74.9	3.79	8.8	-74.5	3.81	9.5
Average Error	-0.2	-0.05	--	-0.1	-0.06	--
Standard Deviation	0.5	0.20	--	0.3	0.22	--
UAE	0.4	0.13	--	0.3	0.15	--

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 1). The target IOD values in Table 1 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 20 (cont'd)

	TIP4P _{EW}		
	HFE (kcal/mol)	IOD (Å)	CN
Li ⁺	-114.0	1.90	4.0
Na ⁺	-88.0	2.32	5.8
K ⁺	-70.8	2.68	6.6
Rb ⁺	-65.8	2.83	7.4
Cs ⁺	-59.8	3.05	8.5
Tl ⁺	-71.7	2.65	6.4
Cu ⁺	-125.5	1.79	4.0
Ag ⁺	-103.2	2.05	4.8
NH ₄ ⁺	-68.8	2.74	6.8
H ₃ O ⁺	-103.6	2.05	4.7
F ⁻	-120.2	2.77	6.3
Cl ⁻	-89.5	3.35	7.4
Br ⁻	-82.4	3.56	7.6
I ⁻	-74.3	3.87	9.2
Average Error	-0.3	-0.11	--
Standard Deviation	0.3	0.28	--
UAE	0.3	0.22	--

Table 21. Estimated HFE parameter set of highly charged metal ions for different water models.

	TIP3P		SPC/E		TIP4P _{EW}	
	R _{min} /2 (Å)	ε (kcal/mol)	R _{min} /2 (Å)	ε (kcal/mol)	R _{min} /2 (Å)	ε (kcal/mol)
Al ³⁺	0.981	0.00000832	0.991	0.00001107	0.876	0.00000026
Fe ³⁺	1.082	0.00011017	1.091	0.00013462	0.984	0.00000907
Cr ³⁺	1.188	0.00089969	1.196	0.00103208	1.096	0.00015019
In ³⁺	1.202	0.00114198	1.209	0.00128267	1.110	0.00020260
Tl ³⁺	1.206	0.00122067	1.213	0.00136949	1.114	0.00022027
Y ³⁺	1.454	0.02639002	1.459	0.02759452	1.375	0.01205473
La ³⁺	1.628	0.09399072	1.629	0.09454081	1.553	0.05807581
Ce ³⁺	1.595	0.07688443	1.597	0.07786298	1.519	0.04525501
Pr ³⁺	1.568	0.06441235	1.571	0.06573030	1.492	0.03655251
Nd ³⁺	1.548	0.05605698	1.551	0.05726270	1.471	0.03064622
Sm ³⁺	1.522	0.04630154	1.526	0.04772212	1.445	0.02431873
Eu ³⁺	1.503	0.03994409	1.507	0.04122946	1.425	0.02014513
Gd ³⁺	1.495	0.03745682	1.499	0.03868661	1.417	0.01863432
Tb ³⁺	1.481	0.03336723	1.485	0.03450196	1.403	0.01619682
Dy ³⁺	1.468	0.02986171	1.472	0.03091095	1.389	0.01400886
Er ³⁺	1.431	0.02133669	1.436	0.02236885	1.350	0.00909668
Tm ³⁺	1.421	0.01937874	1.426	0.02034021	1.340	0.00808758
Lu ³⁺	1.421	0.01937874	1.426	0.02034021	1.340	0.00808758
Hf ⁴⁺	1.087	0.00012321	1.098	0.00015685	0.977	0.00000741
Zr ⁴⁺	1.139	0.00036479	1.149	0.00044254	1.031	0.00003240
Ce ⁴⁺	1.353	0.00941798	1.360	0.01020237	1.257	0.00270120
U ⁴⁺	1.209	0.00128267	1.218	0.00148497	1.105	0.00018227
Pu ⁴⁺	1.273	0.00339720	1.281	0.00379705	1.172	0.00067804
Th ⁴⁺	1.463	0.02858630	1.468	0.02986171	1.370	0.01141046

Table 22. Estimated IOD parameter set of highly charged metal ions for different water models.

	TIP3P		SPC/E		TIP4P _{EW}	
	R _{min} /2 (Å)	ε (kcal/mol)	R _{min} /2 (Å)	ε (kcal/mol)	R _{min} /2 (Å)	ε (kcal/mol)
Al ³⁺	1.297	0.00471279	1.296	0.00465074	1.285	0.00401101
Fe ³⁺	1.386	0.01357097	1.386	0.01357097	1.375	0.01205473
Cr ³⁺	1.344	0.00848000	1.343	0.00838052	1.333	0.00743559
In ³⁺	1.461	0.02808726	1.461	0.02808726	1.450	0.02545423
Tl ³⁺	1.513	0.04321029	1.513	0.04321029	1.502	0.03962711
Y ³⁺	1.602	0.08034231	1.602	0.08034231	1.590	0.07447106
La ³⁺	1.718	0.15060822	1.718	0.15060822	1.707	0.14295367
Ce ³⁺	1.741	0.16721338	1.741	0.16721338	1.729	0.15845086
Pr ³⁺	1.733	0.16134811	1.734	0.16207614	1.722	0.15343866
Nd ³⁺	1.681	0.12564307	1.681	0.12564307	1.669	0.11803919
Sm ³⁺	1.659	0.11189491	1.659	0.11189491	1.647	0.10475707
Eu ³⁺	1.666	0.11617738	1.666	0.11617738	1.655	0.10948690
Gd ³⁺	1.623	0.09126804	1.623	0.09126804	1.612	0.08544204
Tb ³⁺	1.630	0.09509276	1.630	0.09509276	1.619	0.08912336
Dy ³⁺	1.609	0.08389240	1.609	0.08389240	1.597	0.07786298
Er ³⁺	1.602	0.08034231	1.602	0.08034231	1.590	0.07447106
Tm ³⁺	1.602	0.08034231	1.602	0.08034231	1.590	0.07447106
Lu ³⁺	1.588	0.07351892	1.588	0.07351892	1.577	0.06841702
Hf ⁴⁺	1.499	0.03868661	1.501	0.03931188	1.483	0.03393126
Zr ⁴⁺	1.519	0.04525501	1.521	0.04595090	1.503	0.03994409
Ce ⁴⁺	1.684	0.12758274	1.689	0.13084945	1.667	0.11679623
U ⁴⁺	1.684	0.12758274	1.689	0.13084945	1.667	0.11679623
Pu ⁴⁺	1.662	0.11371963	1.666	0.11617738	1.645	0.10359269
Th ⁴⁺	1.708	0.14364160	1.713	0.14710519	1.690	0.13150785

Table 23. Comparison between the $R_{\min}/2$ parameters from the HFE parameter set for the TIP3P water model and the VDW radii calculated from QMSP method.

Ions	$R_{\min}/2$ of HFE parameter set determined for the TIP3P water model (Å)	VDW radius determined by QMSP method ^a (Å)	Absolute Percentage Error ^b
Na ⁺	1.475	1.352	9.1%
K ⁺	1.719	1.671	2.9%
Rb ⁺	1.834	1.801	1.8%
Cs ⁺	1.988	1.997	0.5%
Mg ²⁺	1.284	1.180	8.8%
Ca ²⁺	1.528	1.480	3.2%
Sr ²⁺	1.672	1.625	2.9%
Ba ²⁺	1.842	1.802	2.2%
Al ³⁺	0.981	1.046	6.2%
Y ³⁺	1.454	1.481	1.8%
La ³⁺	1.628	1.642	0.9%
F ⁻	1.783	1.909	6.6%
Cl ⁻	2.252	2.252	0.0%
Br ⁻	2.428	2.298	5.7%
I ⁻	2.724	2.548	6.9%

a. From Stokes.⁸⁸ b. Yielding an average absolute percentage error of 4.0%.

Table 24. The activity derivatives for six different ionic solutions in the SPC/E water model (under ~0.3 M condition).

	JC	H2	HFE	IOD	12-6-4	Experimental values ^a
NaCl	1.08	1.00	1.00	1.03	1.01	0.93
KCl	1.11	1.13	1.03	1.07	0.91	0.90
NaBr	1.07	1.16	1.06	1.00	1.06	0.94
KF	0.87	0.76	0.84	0.87	0.86	0.92
NaI	0.93	0.98	0.87	0.90	0.93	0.97
CsI	0.87	1.09	1.03	0.97	0.92	0.86
Average Error	0.07	0.10	0.05	0.05	0.03	--
Standard Deviation	0.08	0.09	0.04	0.04	0.04	--
UAE	0.10	0.15	0.11	0.09	0.06	--

a. From the book of Robinson and Stokes.²³³

Table 25. The average IOD errors for the HFE parameter set of the mono-, di-, tri- and tetravalent cations.

		M(I)	M(II)	M(III)	M(IV)
TIP3P	Average IOD Error	-0.14 (-5.3%)	-0.27 (-12.4%)	-0.29 (-12.8%)	-0.58 (-25.0%)
	IOD Error Standard Deviation	0.20 (7.2%)	0.14 (7.8%)	0.14 (7.4%)	0.15 (7.0%)
SPC/E	Average IOD Error	-0.17 (-6.5%)	-0.26 (-12.3%)	-0.28 (-12.4%)	-0.57 (-24.5%)
	IOD Error Standard Deviation	0.20 (7.2%)	0.14 (7.6%)	0.13 (7.2%)	0.14 (6.6%)
TIP4P _{EW}	Average IOD Error	-0.25 (-10.1%)	-0.36 (-16.8%)	-0.41 (-18.1%)	-0.74 (-32.0%)
	IOD Error Standard Deviation	0.22 (9.7%)	0.17 (10.2%)	0.16 (9.2%)	0.17 (8.2%)

Table 26. The average HFE errors for the IOD parameter sets of the mono-, di-, tri- and tetravalent cations.

		M(I)	M(II)	M(III)	M(IV)
TIP3P	Average HFE Error	17.2 (-10.0%)	51.1 (-11.5%)	82.7 (-9.3%)	244.3 (-15.7%)
	HFE Error Standard Deviation	26.9 (12.8%)	25.2 (4.5%)	42.7 (3.6%)	62.7 (3.3%)
SPC/E	Average HFE Error	18.4 (-11.6%)	51.9 (-11.7%)	81.8 (-9.2%)	244.9 (-15.8%)
	HFE Error Standard Deviation	26.1 (12.1%)	24.3 (4.3%)	41.6 (3.5%)	61.8 (3.2%)
TIP4P _{EW}	Average HFE Error	22.3 (-14.3%)	67.2 (-15.2%)	108.0 (-12.3%)	283.0 (-18.2%)
	HFE Error Standard Deviation	30.1 (12.6%)	27.5 (4.5%)	46.4 (3.7%)	65.1 (3.3%)

Table 27. Final optimized 12-6-4 parameter set of divalent metal ions for different water models.

	TIP3P			SPC/E			TIP4P _{EW}		
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	κ (Å ⁻²)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	κ (Å ⁻²)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	κ (Å ⁻²)
Be ²⁺	1.203	0.00116124	10.20	1.205	0.00120058	9.800	1.205	0.00120058	11.650
Cu ²⁺	1.476	0.03198620	1.789	1.482	0.03364841	1.758	1.475	0.03171494	2.000
Ni ²⁺	1.431	0.02133669	1.742	1.424	0.01995146	1.714	1.430	0.02113456	2.035
Zn ²⁺	1.455	0.02662782	1.623	1.454	0.02639002	1.588	1.450	0.02545423	1.877
Co ²⁺	1.458	0.02735051	1.442	1.457	0.02710805	1.410	1.455	0.02662782	1.688
Cr ²⁺	1.431	0.02133669	1.120	1.424	0.01995146	1.096	1.425	0.02014513	1.440
Fe ²⁺	1.457	0.02710805	1.128	1.450	0.02545423	1.095	1.450	0.02545423	1.386
Mg ²⁺	1.437	0.02257962	1.046	1.429	0.02093385	0.987	1.436	0.02236885	1.362
V ²⁺	1.494	0.03715368	1.080	1.502	0.03962711	1.060	1.495	0.03745682	1.280
Mn ²⁺	1.485	0.03450196	0.851	1.495	0.03745682	0.828	1.485	0.03450196	1.067
Hg ²⁺	1.641	0.10128575	0.741	1.641	0.10128575	0.751	1.632	0.09620220	0.855
Cd ²⁺	1.535	0.05102457	0.811	1.541	0.05330850	0.819	1.531	0.04953859	0.995
Ca ²⁺	1.642	0.10185975	0.223	1.634	0.09731901	0.230	1.633	0.09675968	0.325
Sn ²⁺	1.777	0.19470705	0.275	1.778	0.19549490	0.286	1.765	0.18535099	0.338
Sr ²⁺	1.777	0.19470705	0.121	1.778	0.19549490	0.137	1.763	0.18380968	0.175
Ba ²⁺	1.936	0.33132862	0.062	1.937	0.33223312	0.072	1.924	0.32049456	0.096

Table 28. Computed C_4 values for the 12-6-4 parameter set of divalent metal ions for different water models.^a

	TIP3P	SPC/E	TIP4P _{EW}
Be ²⁺	186.5	188.1	228.5
Cu ²⁺	290.9	304.4	339.2
Ni ²⁺	212.8	205.2	259.2
Zn ²⁺	231.6	231.2	272.3
Co ²⁺	209.7	209.2	252.8
Cr ²⁺	136.8	131.2	177.4
Fe ²⁺	163.0	155.4	201.1
Mg ²⁺	132.9	122.2	180.5
V ²⁺	195.7	206.6	244.8
Mn ²⁺	146.1	154.9	192.3
Hg ²⁺	288.8	300.2	335.2
Cd ²⁺	185.6	198.8	233.7
Ca ²⁺	87.3	89.0	128.0
Sn ²⁺	187.9	201.1	231.4
Sr ²⁺	82.7	96.3	118.9
Ba ²⁺	71.9	85.8	112.5

a. In Å⁴·kcal/mol.

Table 29. The simulated HFE, IOD and CN values of the 12-6-4 parameter set for divalent metal ions.^a

	TIP3P			SPC/E			TIP4P _{EW}		
	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN
Be ²⁺	-572.8	1.64	4.4	-572.6	1.65	4.3	-572.4	1.64	4.0
Cu ²⁺	-479.8	2.11	6.2	-481.4	2.11	6.0	-480.0	2.11	6.1
Ni ²⁺	-473.1	2.06	6.0	-473.1	2.06	6.0	-472.7	2.06	6.0
Zn ²⁺	-467.0	2.09	6.0	-468.2	2.09	6.0	-467.4	2.09	6.0
Co ²⁺	-457.5	2.10	6.0	-456.8	2.10	6.0	-456.7	2.10	6.0
Cr ²⁺	-441.4	2.08	6.0	-442.0	2.08	6.0	-442.1	2.08	6.0
Fe ²⁺	-439.5	2.11	6.0	-439.8	2.11	6.0	-439.7	2.11	6.0
Mg ²⁺	-436.6	2.09	6.0	-436.7	2.09	6.0	-437.9	2.09	6.0
V ²⁺	-435.9	2.21	7.0	-435.5	2.21	6.8	-436.1	2.22	7.0
Mn ²⁺	-421.2	2.19	6.8	-420.3	2.20	6.5	-420.9	2.20	6.7
Hg ²⁺	-420.0	2.41	8.0	-420.6	2.41	8.0	-421.6	2.41	8.0
Cd ²⁺	-419.4	2.30	7.8	-419.4	2.30	7.5	-420.1	2.30	7.7
Ca ²⁺	-360.6	2.46	8.0	-360.6	2.46	8.0	-359.6	2.46	8.0
Sn ²⁺	-356.1	2.62	9.0	-356.1	2.62	8.8	-356.8	2.62	8.9
Sr ²⁺	-330.4	2.64	8.9	-330.3	2.64	8.7	-329.4	2.64	8.9
Ba ²⁺	-299.4	2.83	9.3	-298.6	2.83	9.2	-298.2	2.83	9.5
Average Error	0.1	0.00	0.6	0.0	0.00	0.5	0.0	0.00	0.6
Standard Deviation	0.5	0.01	0.9	0.6	0.01	0.9	0.5	0.01	0.9
UAE	0.4	0.00	0.6	0.4	0.00	0.5	0.4	0.00	0.6

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 2). The target IOD values in Table 2 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 30. Final optimized 12-6-4 parameter set of monovalent ions for different water models.

	TIP3P			SPC/E		
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	C_4 (kcal/mol•Å ⁴)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	C_4 (kcal/mol•Å ⁴)
Li ⁺	1.325	0.00674244	27	1.327	0.00691068	33
Na ⁺	1.473	0.03117732	0	1.472	0.03091095	6
K ⁺	1.758	0.17997960	8	1.760	0.18150763	19
Rb ⁺	1.831	0.23886274	0	1.826	0.23464849	7
Cs ⁺	2.008	0.39668797	2	2.004	0.39306142	12
Tl ⁺	1.893	0.29273756	50	1.889	0.28918714	61
Cu ⁺	1.217	0.00146124	7	1.218	0.00148497	9
Ag ⁺	1.533	0.05027793	83	1.536	0.05140063	92
NH ₄ ⁺	1.802	0.21475916	4	1.797	0.21069138	13
H ⁺ (Zundel cation)	0.992	0.00001138	108	0.987	0.00000988	106
H ⁺ (Eigen cation)	0.871	0.00000022	51	0.870	0.00000021	51
H ₃ O ⁺	1.774	0.19235093	190	1.773	0.19156806	205
F ⁻	1.725	0.15557763	-27	1.726	0.15629366	-53
Cl ⁻	2.150	0.52153239	-38	2.153	0.52404590	-55
Br ⁻	2.314	0.64785703	-39	2.324	0.65475744	-51
I ⁻	2.567	0.79269938	-45	2.579	0.79809803	-51

Table 30 (cont'd)

	TIP4P _{EW}		
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	C_4 (kcal/mol•Å ⁴)
Li ⁺	1.313	0.00580060	36
Na ⁺	1.459	0.02759452	9
K ⁺	1.751	0.17467422	24
Rb ⁺	1.817	0.22712223	13
Cs ⁺	1.997	0.38670945	16
Tl ⁺	1.883	0.28387745	65
Cu ⁺	1.214	0.00139196	21
Ag ⁺	1.522	0.04630154	94
NH ₄ ⁺	1.791	0.20584696	20
H ⁺ (Zundel cation)	0.997	0.00001309	126
H ⁺ (Eigen cation)	0.876	0.00000026	64
H ₃ O ⁺	1.770	0.18922704	209
F ⁻	1.728	0.15773029	-67
Cl ⁻	2.154	0.52488228	-66
Br ⁻	2.326	0.65612582	-68
I ⁻	2.585	0.80075128	-62

Table 31. The simulated HFE, IOD and CN values of the 12-6-4 parameter set for monovalent ions.^a

	TIP3P			SPC/E			TIP4P _{EW}		
	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN
Li ⁺	-113.4	2.09	5.3	-113.2	2.09	5.1	-112.9	2.09	5.4
Na ⁺	-88.2	2.36	5.7	-88.2	2.35	5.7	-87.9	2.35	5.9
K ⁺	-70.4	2.78	7.8	-70.5	2.79	7.2	-70.5	2.79	7.5
Rb ⁺	-66.0	2.90	7.6	-65.9	2.89	7.8	-65.7	2.89	7.5
Cs ⁺	-60.2	3.13	9.4	-60.4	3.14	8.9	-60.3	3.13	9.3
Tl ⁺	-71.4	2.96	8.5	-71.4	2.97	8.9	-71.6	2.96	8.5
Cu ⁺	-125.0	1.86	4.0	-125.0	1.86	4.0	-124.5	1.86	4.0
Ag ⁺	-102.0	2.41	6.0	-102.8	2.41	6.0	-102.2	2.41	6.0
NH ₄ ⁺	-67.8	2.86	7.4	-67.9	2.85	7.4	-68.4	2.86	7.5
H ⁺ (Zundel cation)	-250.9	1.24	2.0	-251.7	1.23	2.0	-250.7	1.25	2.0
H ⁺ (Eigen cation)	-251.2	1.01	1.0	-251.3	1.01	1.0	-250.6	1.02	1.0
H ₃ O ⁺	-102.8	2.75	8.4	-104.2	2.75	8.4	-103.3	2.75	8.2
F ⁻	-119.3	2.63	6.4	-119.8	2.63	6.2	-119.5	2.63	6.0
Cl ⁻	-89.1	3.18	7.6	-88.5	3.18	7.3	-89.4	3.18	6.9
Br ⁻	-82.6	3.37	7.7	-82.1	3.37	7.3	-81.8	3.37	7.0
I ⁻	-74.5	3.64	8.5	-74.7	3.64	7.7	-74.1	3.64	8.5
Average Error	0.1	0.00	--	-0.1	0.00	--	0.2	0.00	--
Standard Deviation	0.4	0.01	--	0.5	0.01	--	0.5	0.01	--
UAE	0.3	0.00	--	0.4	0.00	--	0.4	0.00	--

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 1). The target IOD values in Table 1 (without considering the error bars) were treated as the reference for the simulated IOD values.

Table 32. Final optimized 12-6-4 parameter set of highly charged metal ions for different water models.

	TIP3P			SPC/E		
	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	C_4 (kcal/mol•Å ⁴)	$R_{\min}/2$ (Å)	ϵ (kcal/mol)	C_4 (kcal/mol•Å ⁴)
Al ³⁺	1.369	0.01128487	399	1.375	0.01205473	406
Fe ³⁺	1.443	0.02387506	428	1.450	0.02545423	442
Cr ³⁺	1.415	0.01827024	258	1.414	0.01809021	254
In ³⁺	1.491	0.03625449	347	1.487	0.03507938	349
Tl ³⁺	1.571	0.06573030	456	1.569	0.06484979	455
Y ³⁺	1.630	0.09509276	216	1.624	0.09180886	209
La ³⁺	1.758	0.17997960	152	1.763	0.18380968	165
Ce ³⁺	1.782	0.19865859	230	1.786	0.20184160	242
Pr ³⁺	1.780	0.19707431	264	1.782	0.19865859	272
Nd ³⁺	1.724	0.15486311	213	1.735	0.16280564	235
Sm ³⁺	1.711	0.14571499	230	1.703	0.14021803	224
Eu ³⁺	1.716	0.14920231	259	1.721	0.15272873	273
Gd ³⁺	1.658	0.11129023	198	1.646	0.10417397	186
Tb ³⁺	1.671	0.11928915	235	1.666	0.11617738	227
Dy ³⁺	1.637	0.09900804	207	1.637	0.09900804	206
Er ³⁺	1.635	0.09788018	251	1.629	0.09454081	247
Tm ³⁺	1.647	0.10475707	282	1.633	0.09675968	262
Lu ³⁺	1.625	0.09235154	249	1.620	0.08965674	247
Hf ⁴⁺	1.600	0.07934493	827	1.592	0.07543075	810
Zr ⁴⁺	1.609	0.08389240	761	1.609	0.08389240	760
Ce ⁴⁺	1.766	0.18612361	706	1.761	0.18227365	694
U ⁴⁺	1.792	0.20665151	1034	1.791	0.20584696	1043
Pu ⁴⁺	1.752	0.17542802	828	1.750	0.17392181	828
Th ⁴⁺	1.770	0.18922704	512	1.773	0.19156806	513

Table 32 (cont'd)

	TIP4P _{EW}		
	R _{min} /2 (Å)	ε (kcal/mol)	C ₄ (kcal/mol•Å ⁴)
Al ³⁺	1.377	0.01232018	488
Fe ³⁺	1.448	0.02499549	519
Cr ³⁺	1.408	0.01703790	322
In ³⁺	1.486	0.03478983	425
Tl ³⁺	1.564	0.06268139	535
Y ³⁺	1.624	0.09180886	294
La ³⁺	1.755	0.17769767	243
Ce ³⁺	1.776	0.19392043	315
Pr ³⁺	1.774	0.19235093	348
Nd ³⁺	1.720	0.15202035	297
Sm ³⁺	1.706	0.14226734	314
Eu ³⁺	1.711	0.14571499	345
Gd ³⁺	1.652	0.10769970	280
Tb ³⁺	1.665	0.11556030	313
Dy ³⁺	1.639	0.10014323	298
Er ³⁺	1.628	0.09399072	328
Tm ³⁺	1.638	0.09957472	356
Lu ³⁺	1.617	0.08806221	331
Hf ⁴⁺	1.599	0.07884906	956
Zr ⁴⁺	1.610	0.08440707	895
Ce ⁴⁺	1.761	0.18227365	835
U ⁴⁺	1.791	0.20584696	1183
Pu ⁴⁺	1.753	0.17618319	972
Th ⁴⁺	1.758	0.17997960	625

Table 33. The simulated HFE, IOD and CN values of the 12-6-4 parameter set for highly charged metal ions.^a

	TIP3P			SPC/E		
	HFE (kcal/mol)	IOD (Å)	CN	HFE (kcal/mol)	IOD (Å)	CN
Al ³⁺	-1082.0	1.87	6.0	-1081.3	1.88	6.0
Fe ³⁺	-1019.4	2.02	6.9	-1019.2	2.02	6.8
Cr ³⁺	-957.8	1.95	6.0	-957.8	1.96	6.0
In ³⁺	-951.5	2.15	8.0	-950.4	2.15	8.0
Tl ³⁺	-948.1	2.22	8.0	-948.5	2.23	8.0
Y ³⁺	-824.9	2.36	9.0	-824.6	2.36	9.0
La ³⁺	-750.7	2.53	9.7	-752.1	2.52	9.2
Ce ³⁺	-765.2	2.55	9.9	-765.1	2.55	9.7
Pr ³⁺	-775.4	2.54	9.9	-776.6	2.54	9.7
Nd ³⁺	-783.6	2.46	9.0	-784.3	2.47	9.0
Sm ³⁺	-795.3	2.44	9.0	-794.8	2.44	9.0
Eu ³⁺	-802.1	2.44	9.0	-803.4	2.45	9.0
Gd ³⁺	-806.5	2.39	9.0	-807.2	2.39	9.0
Tb ³⁺	-813.5	2.40	9.0	-812.2	2.40	9.0
Dy ³⁺	-818.4	2.37	9.0	-819.0	2.37	9.0
Er ³⁺	-834.8	2.36	9.0	-834.9	2.36	9.0
Tm ³⁺	-840.7	2.37	9.0	-840.2	2.36	9.0
Lu ³⁺	-839.3	2.34	9.0	-840.4	2.34	9.0
Average Error	0.1	0.00	--	0.0	0.00	--
Standard Deviation	0.6	0.01	--	0.5	0.00	--
UAE	0.5	0.00	--	0.4	0.00	--
Hf ⁴⁺	-1663.9	2.16	10.0	-1663.9	2.16	8.0
Zr ⁴⁺	-1622.7	2.19	9.9	-1622.9	2.19	9.8
Ce ⁴⁺	-1462.2	2.42	10.0	-1462.1	2.42	10.0
U ⁴⁺	-1566.6	2.41	10.0	-1566.0	2.41	10.0
Pu ⁴⁺	-1519.4	2.39	10.0	-1520.3	2.39	10.0
Th ⁴⁺	-1389.3	2.44	10.0	-1388.3	2.45	10.0
Average Error	0.6	0.00	--	0.8	0.00	--
Standard Deviation	0.4	0.01	--	0.8	0.00	--
UAE	0.6	0.00	--	0.9	0.00	--

a. All the average errors and standard deviations were obtained by treating the corresponding target values as the reference (see Table 3).

Table 33 (cont'd)

	TIP4P _{EW}		
	HFE (kcal/mol)	IOD (Å)	CN
Al ³⁺	-1080.8	1.88	6.0
Fe ³⁺	-1020.2	2.03	6.8
Cr ³⁺	-957.5	1.95	6.0
In ³⁺	-952.2	2.15	7.9
Tl ³⁺	-949.7	2.22	8.0
Y ³⁺	-824.9	2.36	9.0
La ³⁺	-752.4	2.52	9.4
Ce ³⁺	-764.6	2.55	9.8
Pr ³⁺	-775.9	2.54	9.8
Nd ³⁺	-783.4	2.46	9.0
Sm ³⁺	-795.4	2.44	9.0
Eu ³⁺	-802.8	2.45	9.0
Gd ³⁺	-807.6	2.39	9.0
Tb ³⁺	-812.9	2.40	9.0
Dy ³⁺	-819.3	2.38	9.0
Er ³⁺	-836.2	2.36	9.0
Tm ³⁺	-840.6	2.36	9.0
Lu ³⁺	-840.9	2.34	9.0
Average Error	-0.3	0.00	--
Standard Deviation	0.6	0.00	--
UAE	0.6	0.00	--
Hf ⁴⁺	-1663.3	2.16	8.0
Zr ⁴⁺	-1622.6	2.19	9.9
Ce ⁴⁺	-1462.0	2.42	10.0
U ⁴⁺	-1569.3	2.42	10.0
Pu ⁴⁺	-1520.4	2.40	10.0
Th ⁴⁺	-1388.0	2.44	10.0
Average Error	0.4	0.00	--
Standard Deviation	1.2	0.01	--
UAE	1.0	0.00	--

Table 34. Polarizabilities used in the 12-6-4 model for different atom types in the AMBER FF.

Atom type in AMBER FF	Polarizability ($\text{\AA}^3$)	Atom type in AMBER FF	Polarizability ($\text{\AA}^3$)
HO/ H1/ H2/ H4/ HA/ H/ HC/ H5	0.387 ^a	P	1.538 ^a
CM/CA/C/CK/CB/CQ	1.352 ^a	MG	0.048 ^b
CT	1.061 ^a	Cl-	1.910 ^c
N*/N2/NC/NB/NA	1.090 ^a	OW	1.444 ^d
O/O2	0.569 ^a	EP/HW	0.000
OS	0.637 ^a		
OH	0.637 ^a		

a. Referenced or adopted from Miller.²³⁴ b. Calculated at the B3LYP/6-311++G(2d,2p) level of theory. c. From Applequist *et al.*²³⁵ d. From Eisenberg and Kauzmann.²⁰⁵

Table 35. The IODs of the Mg^{2+} and Cl^- ions at different concentrations.^a

	0.25 M	0.5 M	1.0 M
Original AMBER FF parameter	2.01/3.16	2.00/3.16	2.00/3.16
12-6 CM set of parameter	2.07/3.17	2.08/3.17	2.08/3.17
12-6-4 model	2.11/3.21	2.11/3.21	2.11/3.21

a. The units are Å while the experimental results are 2.09 ± 0.04 Å for Mg^{2+} and 3.18 ± 0.06 Å for Cl^- (from Marcus¹²³).

Table 36. Comparison of the structural properties predicted by the three different Mg^{2+} parameters used in the simulations of a Mg^{2+} -nucleic acid system.

	IOD (Å)	Mg^{2+} -backbone phosphate distance (Å)
Original AMBER FF parameter	2.02	3.30
12-6 CM set of parameter	2.05	3.35
12-6-4 model in present work	2.09	3.43
Experimental values	$2.09 \pm 0.04^{\text{a}}$	3.6^{b}

a. From Marcus.¹²³ b. From Caminiti.²³⁶

Table 37. Polarizability values of the AMBER atom types used in the 12-6-4 nonbonded model simulation of a metalloprotein system.^a

Atom type	Polarizability	Atom type	Polarizability	Atom type	Polarizability
HW	0.000	O2	0.569	CB	1.352
FE	0.264	OH	0.637	CC	1.352
H	0.387	CT	1.061	CN	1.352
H1	0.387	N	1.090	CR	1.352
H4	0.387	N2	1.090	CV	1.352
H5	0.387	N3	1.090	CW	1.352
HA	0.387	NA	1.090	OW	1.444
HC	0.387	NB	1.090	S	3.000
HO	0.387	C	1.352	SH	3.000
HP	0.387	C*	1.352		
O	0.569	CA	1.352		

a. Polarizability of Fe^{3+} (with atom type FE) was calculated at the B3LYP/6-311++G(2d,2p) level of theory using Gaussian 09 Revision C.01,²³⁷ polarizability of the water oxygen (OW) was taken from Eisenberg and Kauzmann,²⁰⁵ while polarizabilities of the other atom types were adopted from Miller.²³⁴

Table 38. Small model capping schemes for coordinated amino acid residues.

Atom bound to central metal ion	Original residue	Small model
Sidechain atom in an amino acid	NH-CHR-CO	CH ₃ R
Backbone O atom in a non-terminal residue	NH-CHR-CO	CH ₃ -CO or CH ₂ R-CO (if this residue also has a sidechain atom bound to the metal ion) *C-terminal converted to NH-CH ₃ (if it doesn't bond to the metal ion), or to NH-CH ₂ R (if it has a sidechain atom bound to the metal ion)
Backbone N atom in a non-terminal residue	NH-CHR-CO	NH-CH ₃ or NH-CH ₂ R (if this residue also has a sidechain atom bound to the metal ion) *N-terminal converted to CH ₃ -CO (if it doesn't bonds to the metal ion), or to CH ₂ R-CO (if it has a sidechain atom bound to the metal ion)
Backbone CO ₂ ⁻ in the C-terminal residue	NH-CHR-CO ₂ ⁻	CH ₃ -CO ₂ ⁻
Backbone NH ₂ in the N-terminal residue	NH ₂ -CHR-CO	NH ₂ -CH ₃
Backbone N and O in one residue	NH-CHR-CO	NH-CH ₂ -CO *N-terminal modeled as CH ₃ -CO (or to CH ₂ R-CO if it has a sidechain atom bound to the metal ion) and the C-terminal to NH-CH ₃ (or to NH-CH ₂ R if it has a sidechain atom bound to the metal ion)

Table 39. The RMSD values between the CSD, QM optimized and MM minimized structures.^a

	CSD structure	QM optimized	MM minimized
CSD structure	--	0.280(0.219)	0.377(0.327)
QM optimized	0.280(0.219)	--	0.201(0.193)
MM minimized	0.377(0.327)	0.201(0.193)	--

a. Units are Å. The values outside brackets are for all atoms while the values inside brackets are only for heavy atoms. The CSD structure is from residue 1 in CSD entry FAJYAR01.

APPENDIX B: FIGURES

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															
Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu				
Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lw				

Figure 1. Atomic ions that were parameterized for the 12-6 model in current work. These ions are shown with blue background. The monovalent anions, monovalent cations, divalent cations, trivalent cations, and tetravalent cations are shown in light green, yellow, orange, red, dark red, respectively. The ions that have multiple oxidation states are shown in white, which are $\text{Cr}^{2+}/\text{Cr}^{3+}$, $\text{Fe}^{2+}/\text{Fe}^{3+}$, $\text{Cu}^{+}/\text{Cu}^{2+}$, $\text{Ag}^{+}/\text{Ag}^{2+}$, $\text{Ce}^{3+}/\text{Ce}^{4+}$, $\text{Sm}^{2+}/\text{Sm}^{3+}$, $\text{Eu}^{2+}/\text{Eu}^{3+}$, and $\text{Tl}^{+}/\text{Tl}^{3+}$. Besides these atomic ions, the H_3O^{+} and NH_4^{+} ions were also parameterized in current work.

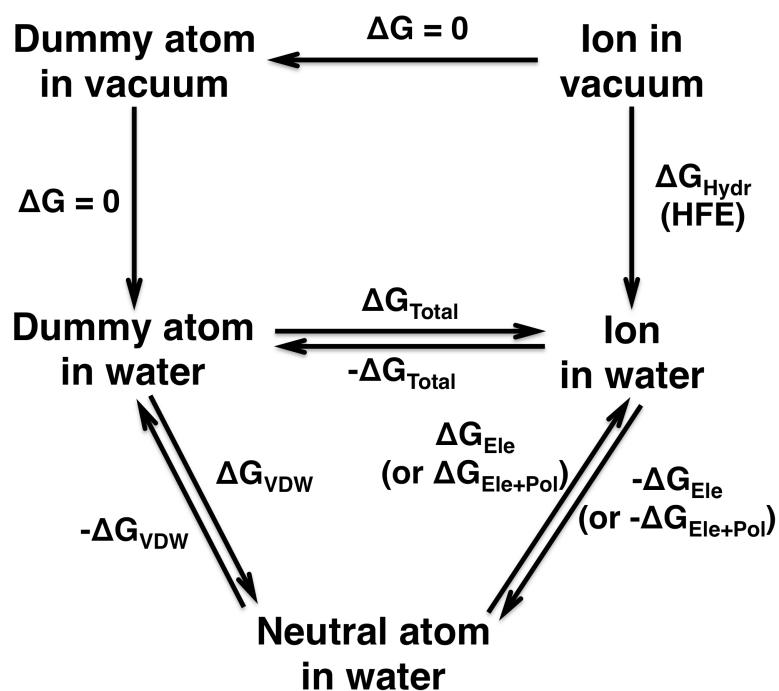


Figure 2. Thermodynamics cycle used to determine the HFE values.

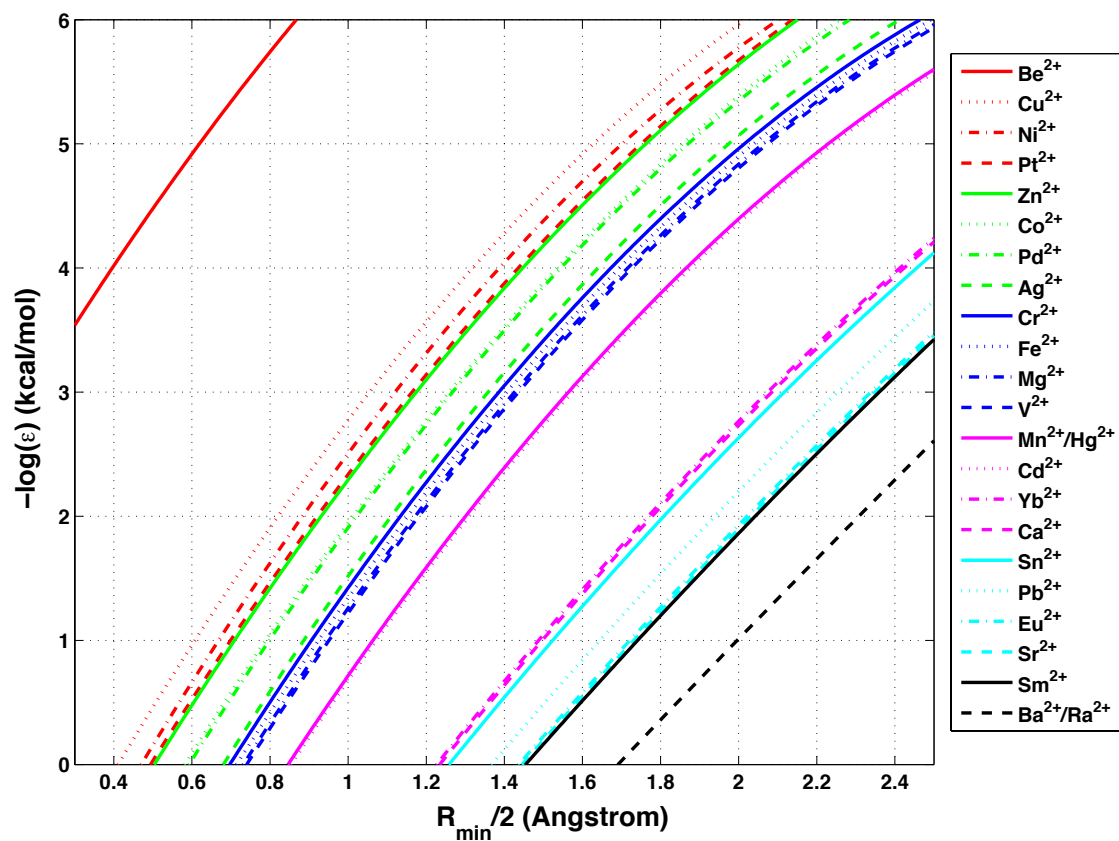


Figure 3. HFE fitting curves for 24 M(II) metal ions for the TIP3P water model.

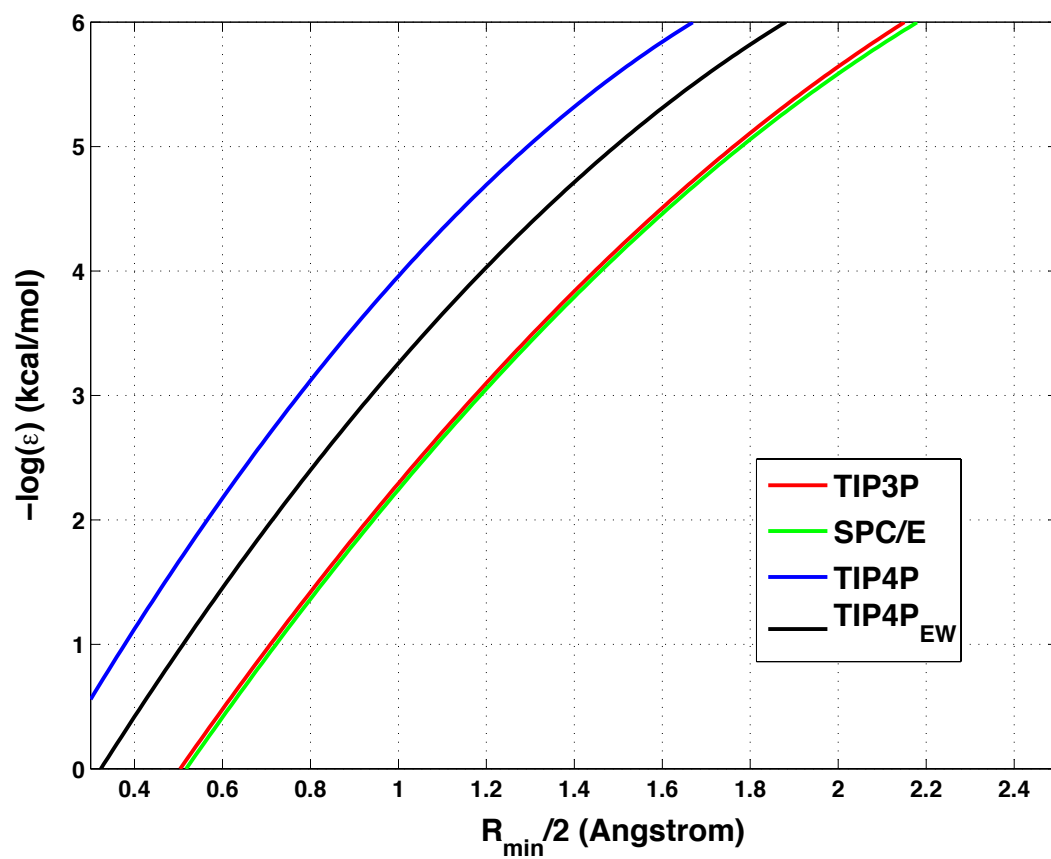


Figure 4. HFE fitting curves for Zn^{2+} in four different water models.

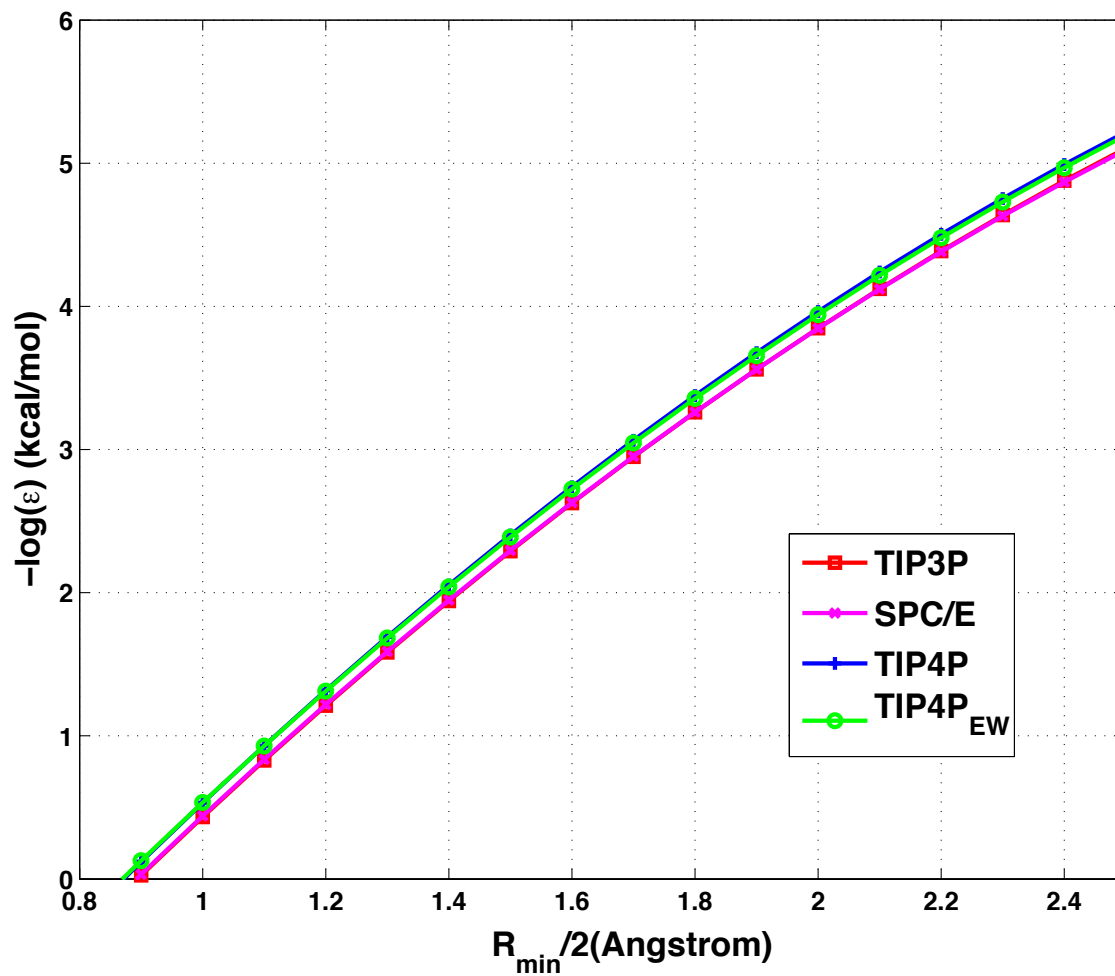


Figure 5. IOD fitting curves for Zn^{2+} in four different water models.

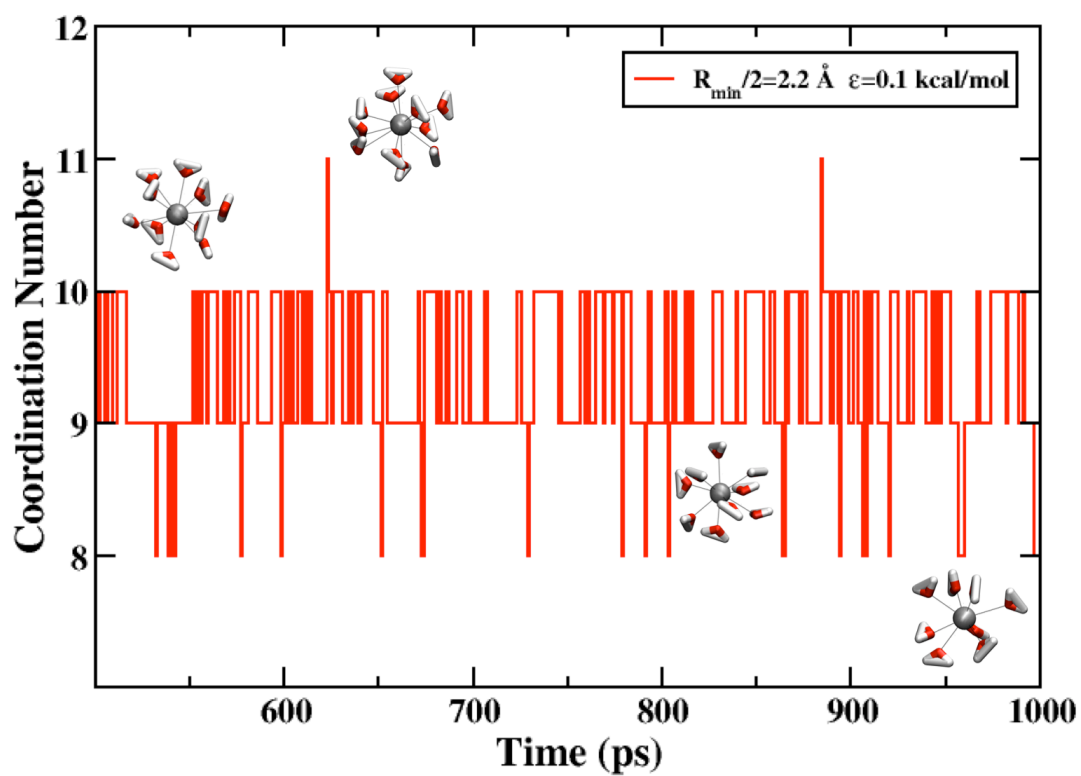


Figure 6. CN switching during a MD simulation when $R_{\min}/2 = 2.2 \text{ \AA}$ and $\epsilon = 0.1 \text{ kcal/mol}$ for a M(II) ion in the TIP3P water model.

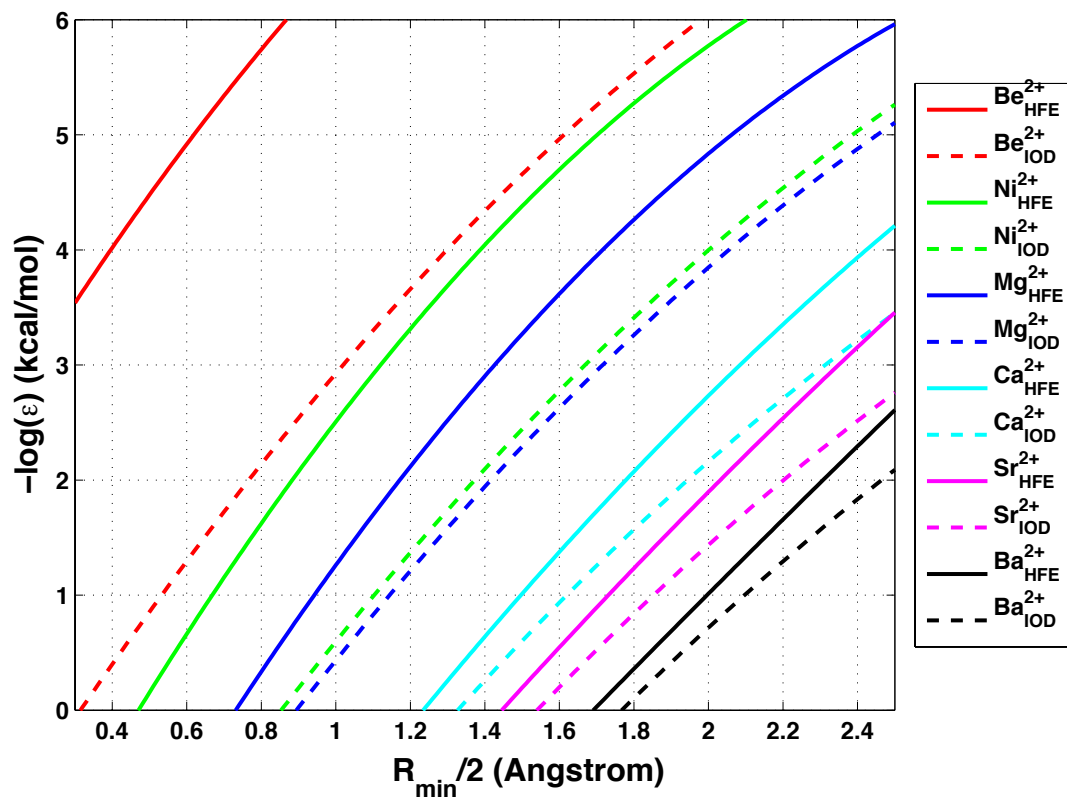


Figure 7. HFE and IOD fitting curves of six representative M(II) metal ions in the TIP3P water model.

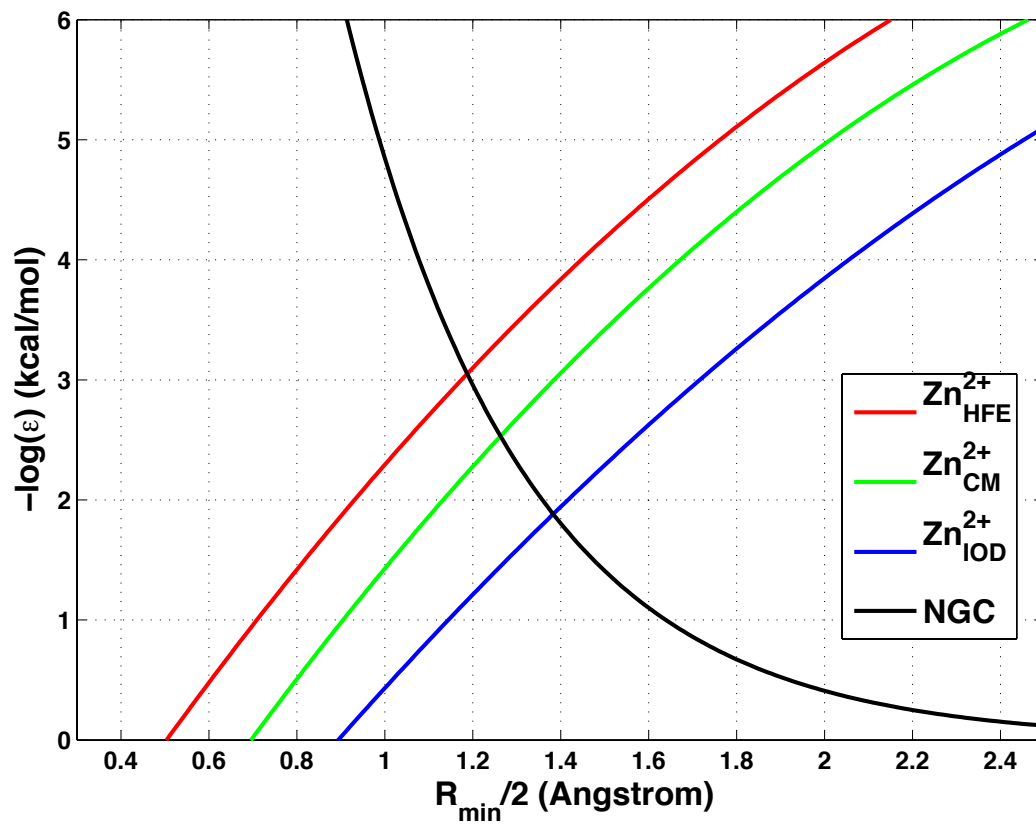


Figure 8. Determination of the three parameter sets for Zn^{2+} in the TIP3P water model.

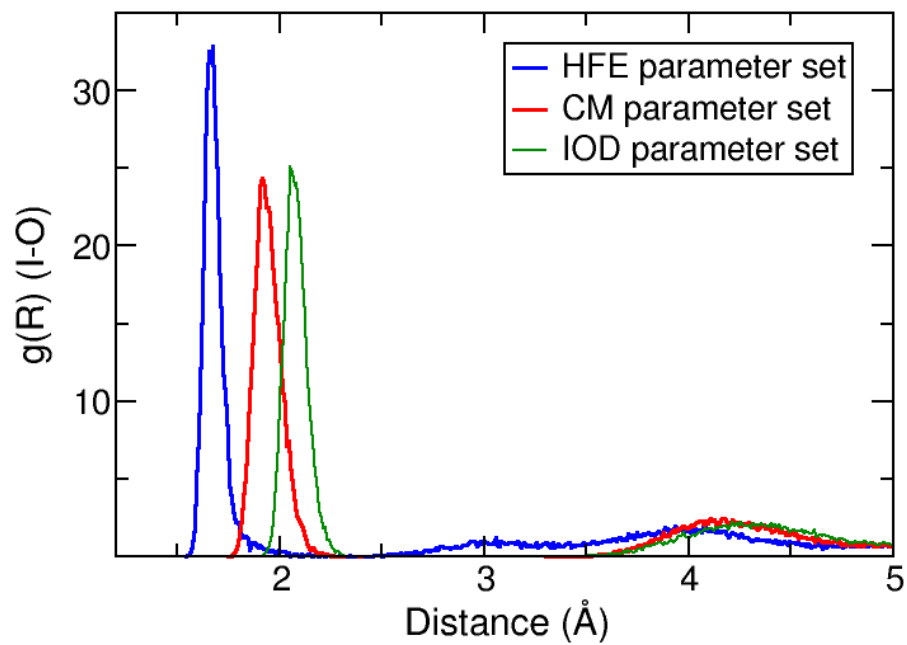


Figure 9. Radial distribution functions of the three parameter sets for Zn^{2+} in the TIP3P water model.

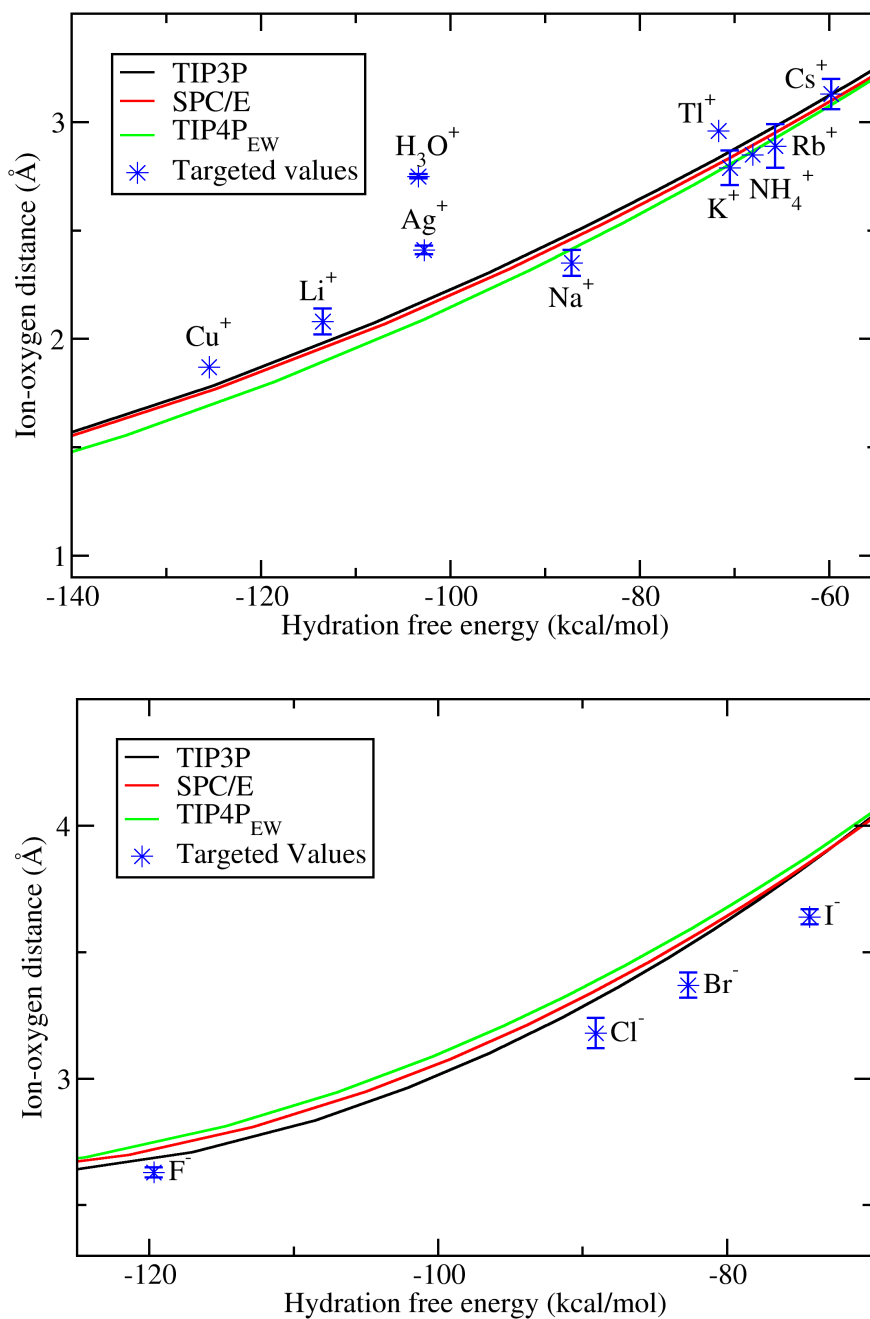


Figure 10. Fitting curves between the HFE and IOD values for the positive (upper) and negative (lower) monovalent ions in the three water models together with the target values of the ions investigated in the present work.

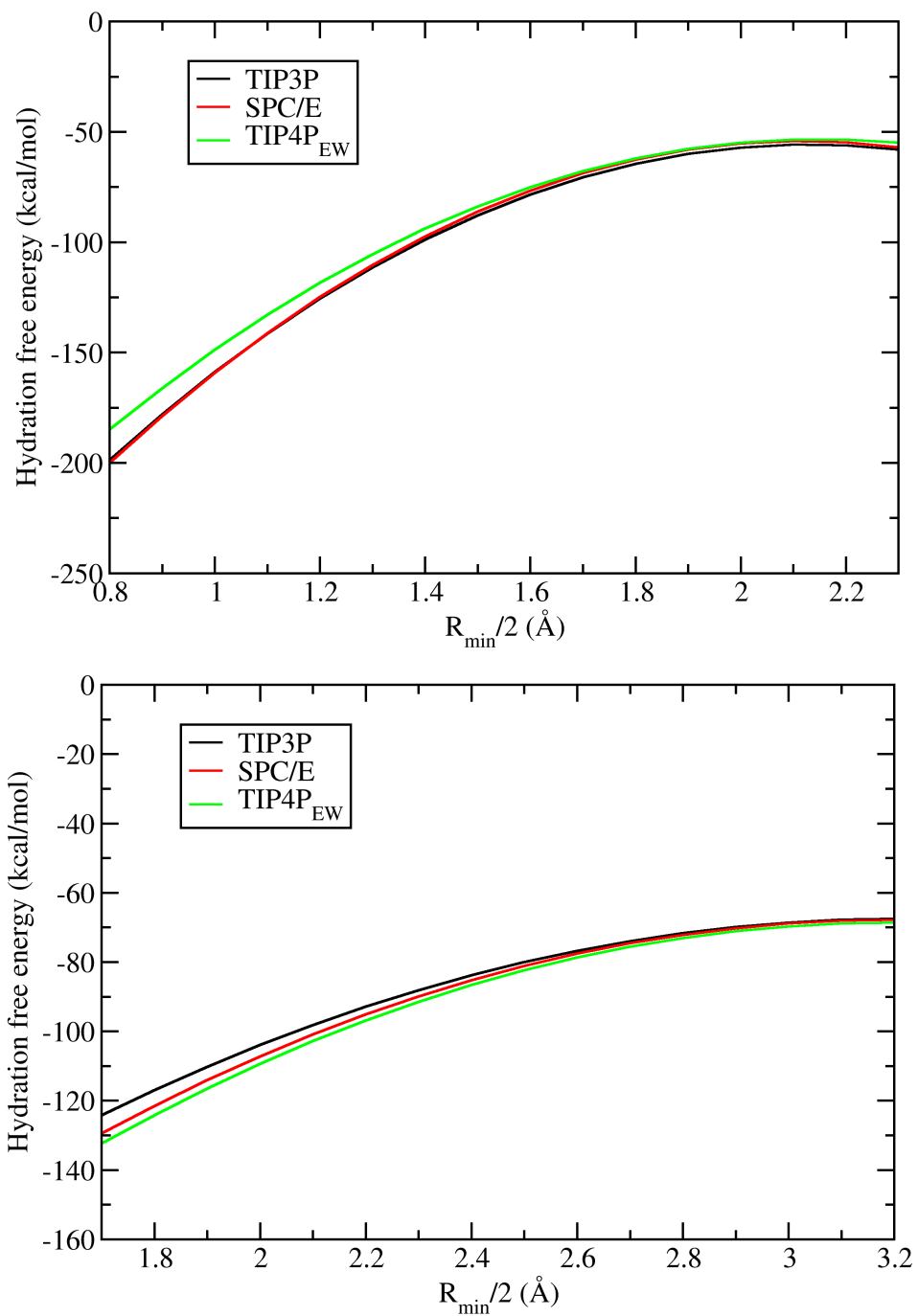


Figure 11. Fitting curves between the HFE and $R_{\min}/2$ values for the positive (upper) and negative (lower) monovalent ions in the three water models.

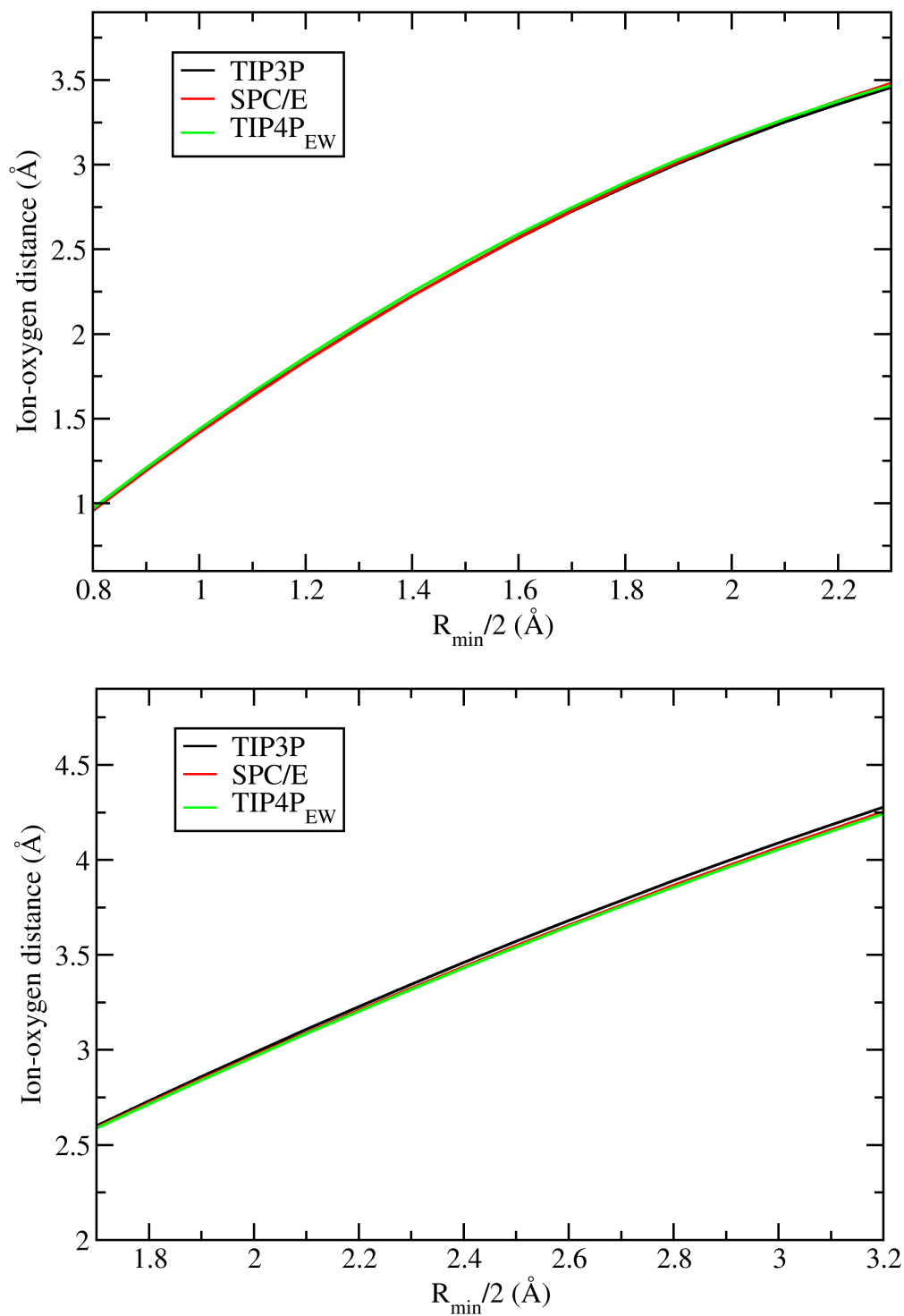


Figure 12. Fitting curves between the IOD and $R_{\min}/2$ values for the positive (upper) and negative (lower) monovalent ions in the three water models.

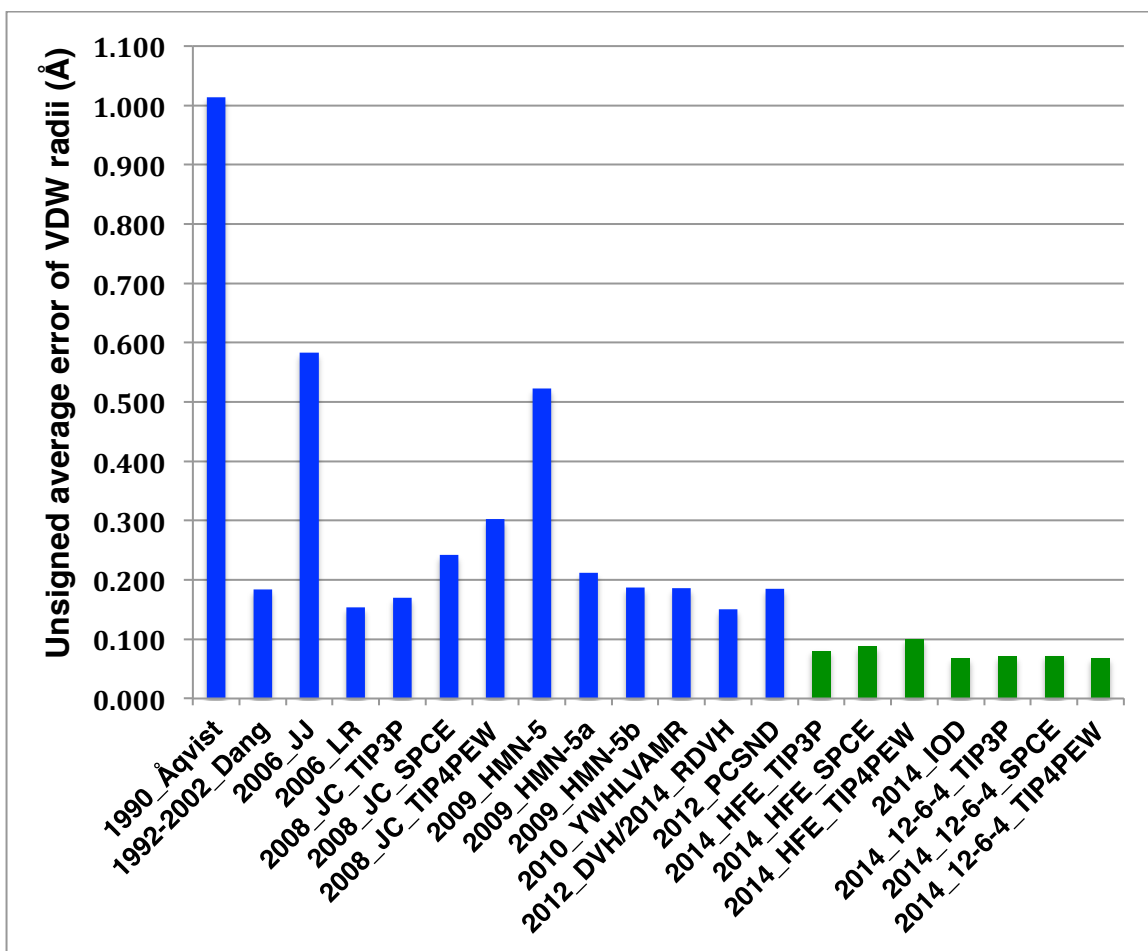


Figure 13. Unsigned average errors of the VDW radii of different parameter sets using QMSP calculated values as reference. The parameter sets developed herein are shown in the seven green columns on the right side of the figure.



Figure 14. PDB entry 4BV1. Waters are not shown in the figure, the ferric ions are shown as sliver spheres. This picture was created using VMD.¹

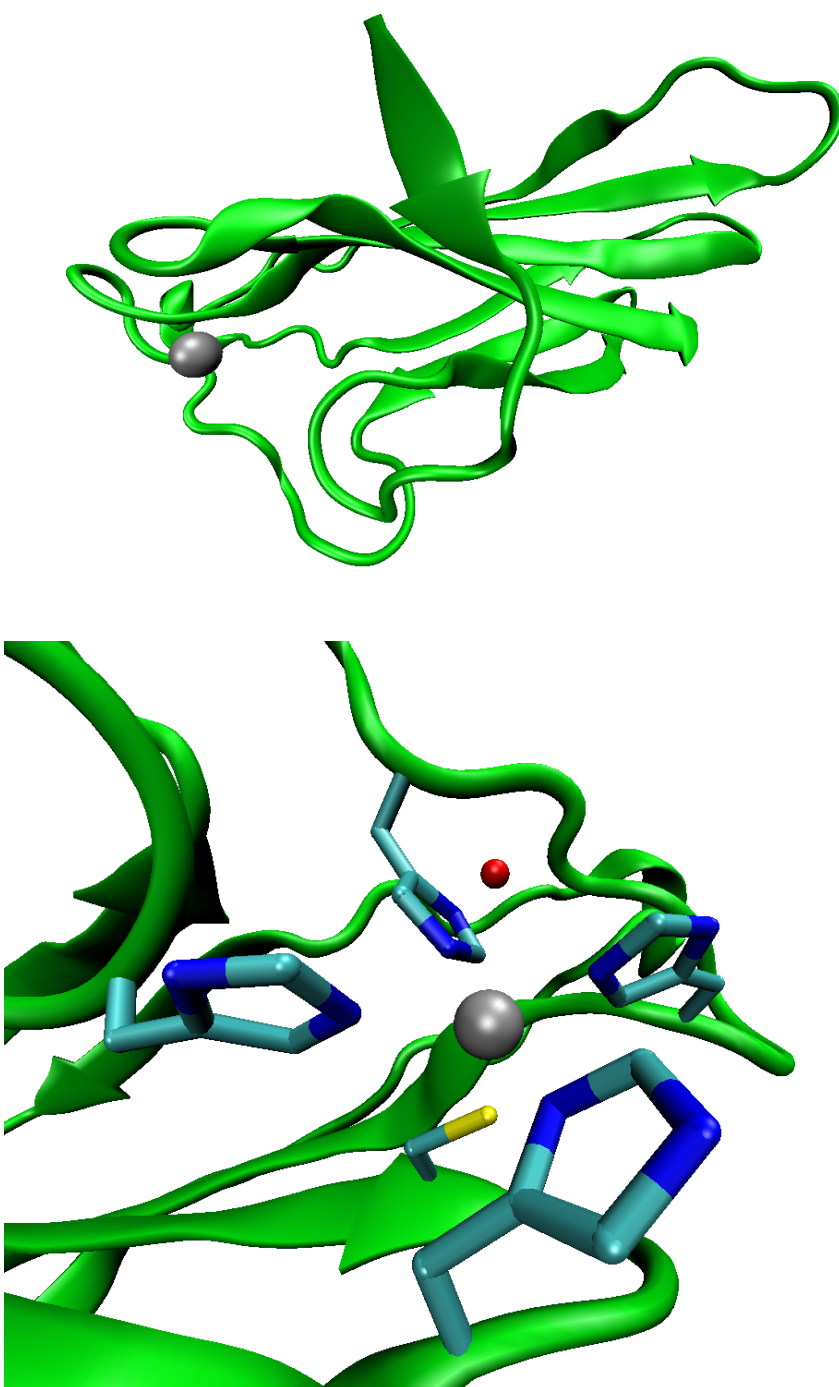


Figure 15. Chain C in PDB entry 4BV1 (upper) and a close-up figure of the metal site in Chain C (lower). The ferric ion is represented as a silver sphere and it is coordinated by one Cys, four His and one water molecule. The figures were made using VMD.¹

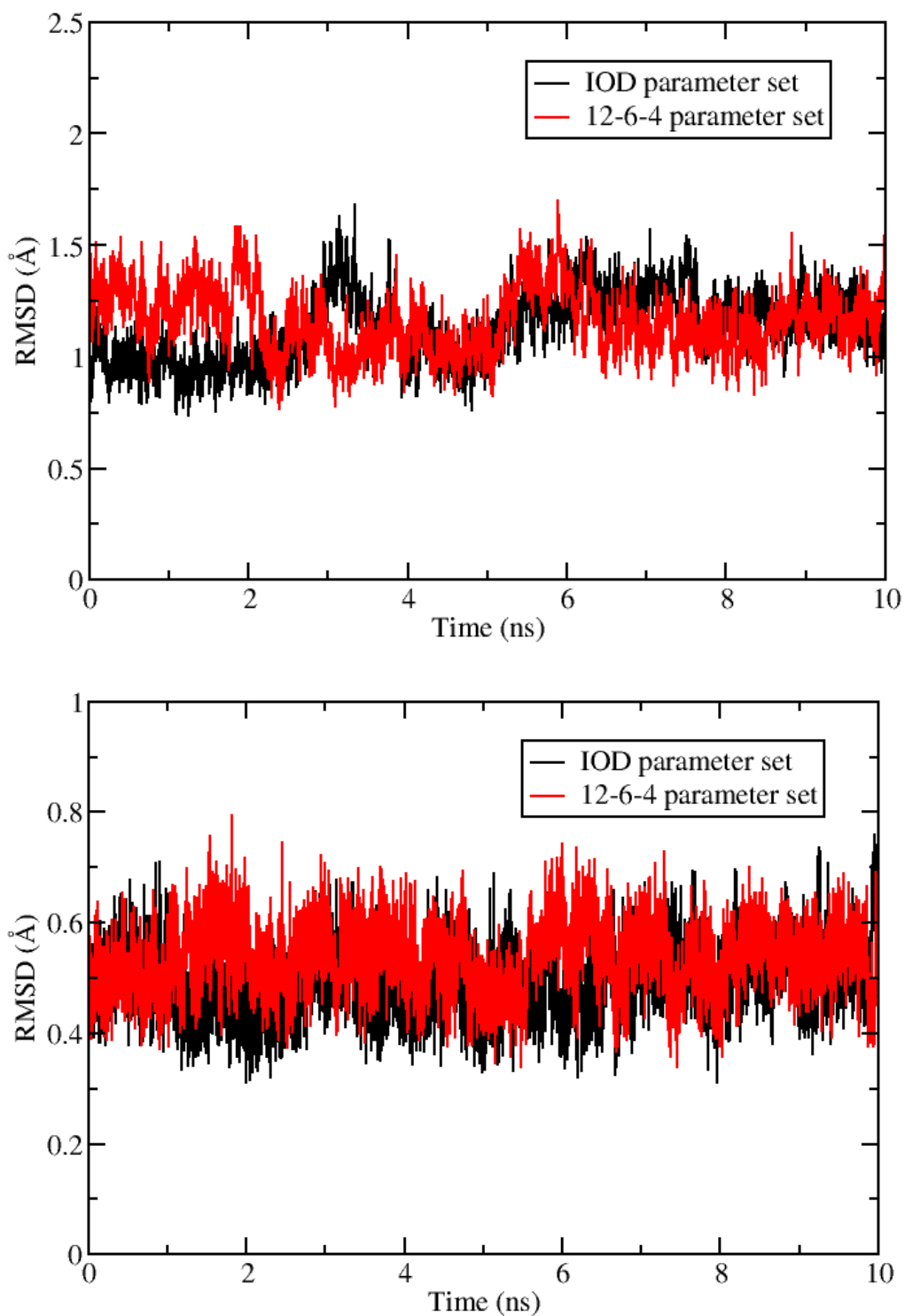


Figure 16. RMSDs of the backbone CA, C, N atoms (upper) and the metal site heavy atoms (lower) for the simulations with the IOD and 12-6-4 parameter sets using the initial structure (experimental structure) as reference.

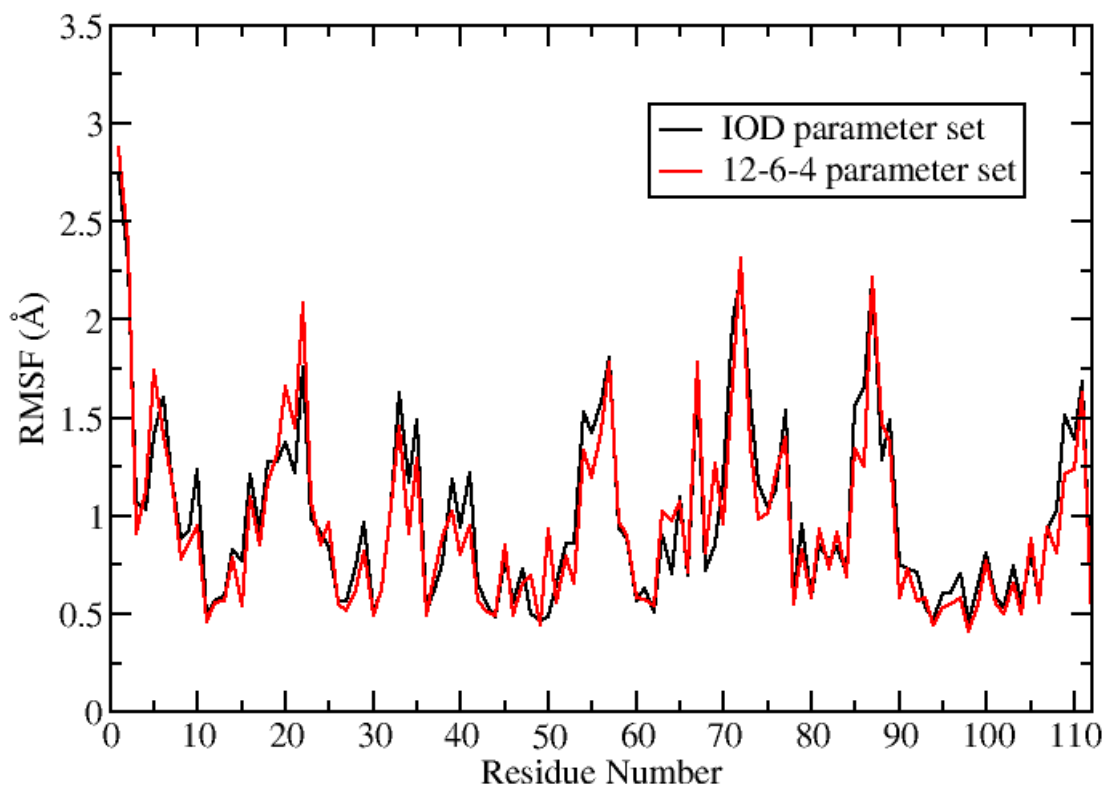


Figure 17. RMSFs of the heavy atoms of the protein residues in the simulations using the IOD and 12-6-4 parameter sets.

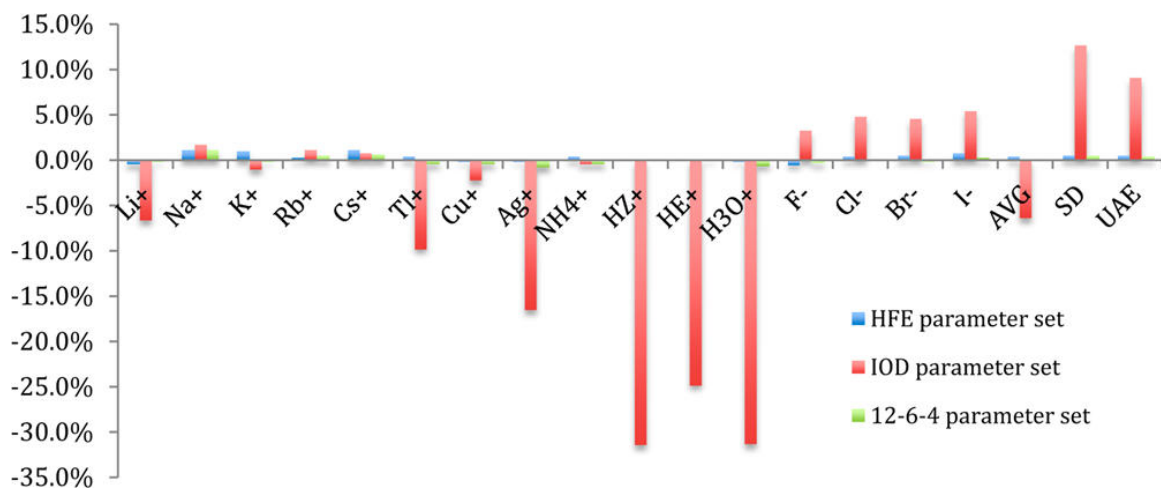


Figure 18. HFE percent errors for different parameter sets of monovalent ions in the TIP3P water model. HZ^+ and HE^+ represent the H^+ Zundel and Eigen ions respectively. Since we did not design HFE set of parameters for the HZ^+ and HE^+ ions, they are not shown in the figure.

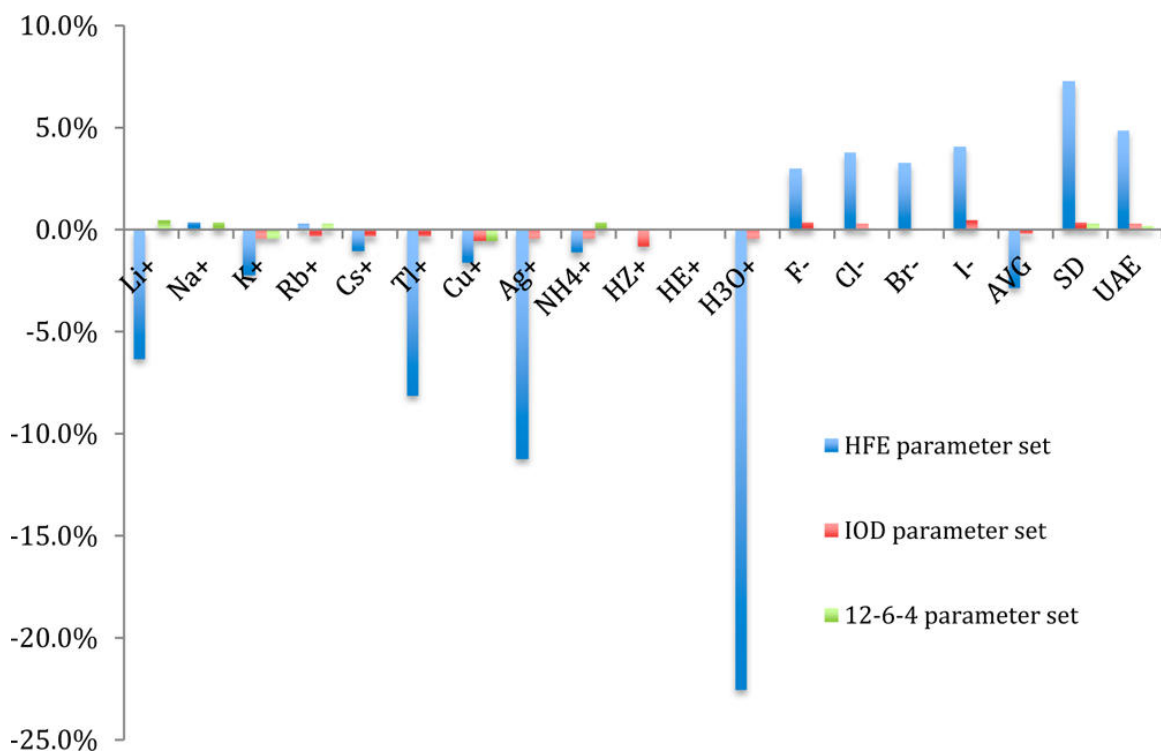


Figure 19. IOD percent errors for different parameter sets of monovalent ions in the TIP3P water model. HZ^+ and HE^+ represent the H^+ Zundel and Eigen ions respectively. Since we did not design IOD set of parameters for the HZ^+ and HE^+ ions, they are not shown in the figure.

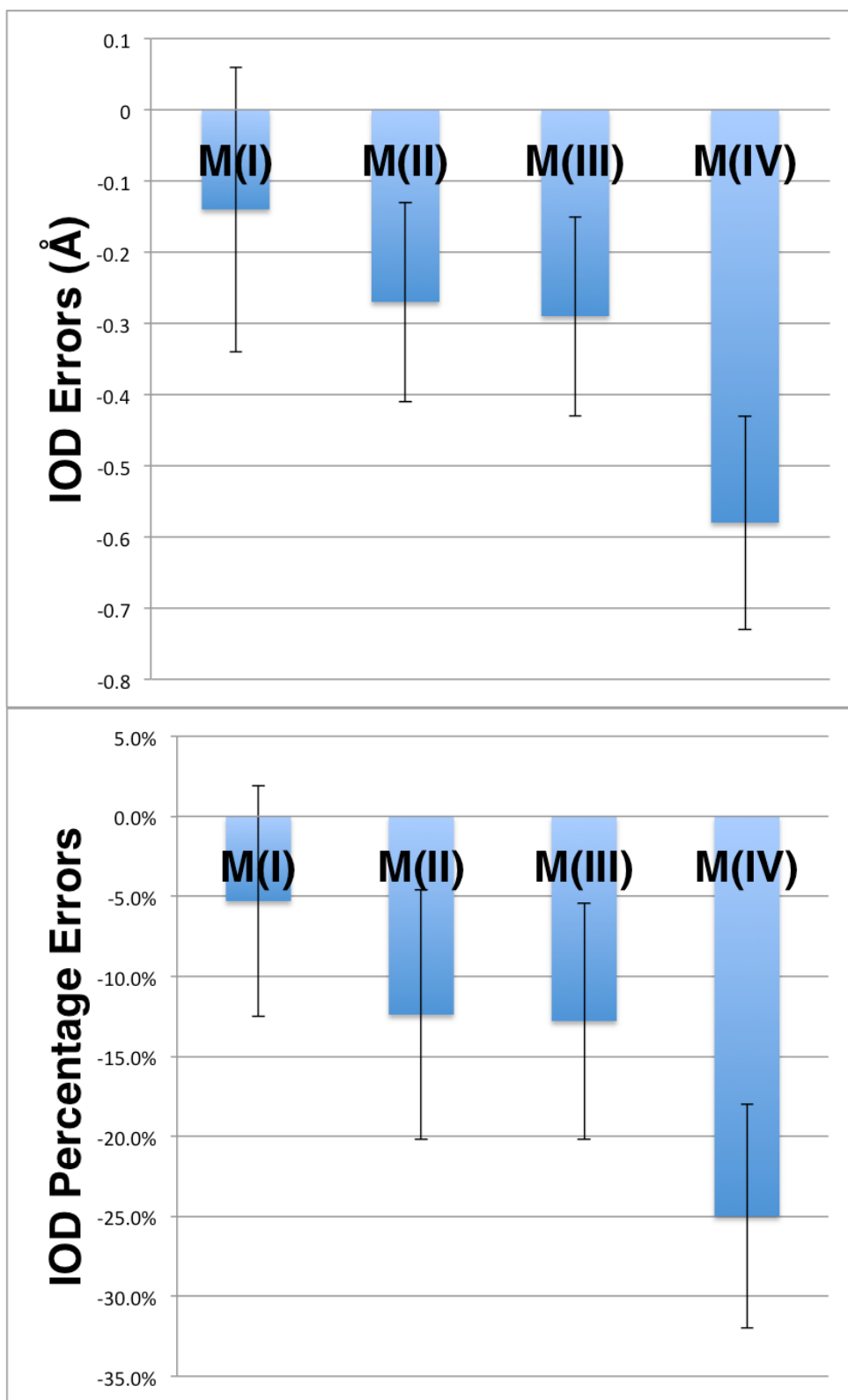


Figure 20. The average errors (upper) and average percentage errors (lower) of the IODs for the HFE parameter set for the mono-, di-, tri- and tetravalent cations in the TIP3P water model.

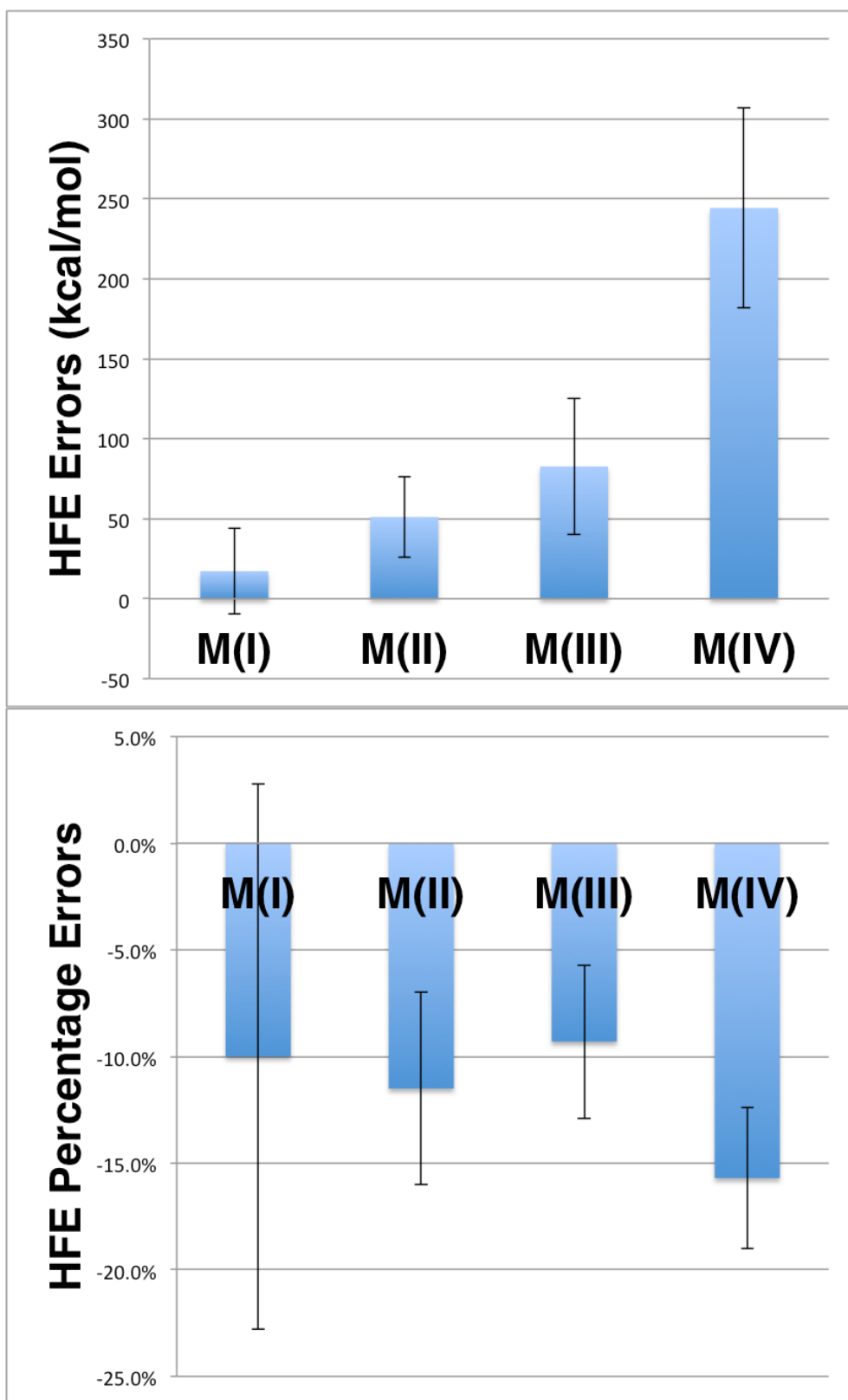


Figure 21. The average errors (upper) and average percentage errors (lower) of the HFEs for the IOD parameter set for the mono-, di-, tri- and tetra-valent cations in the TIP3P water model.

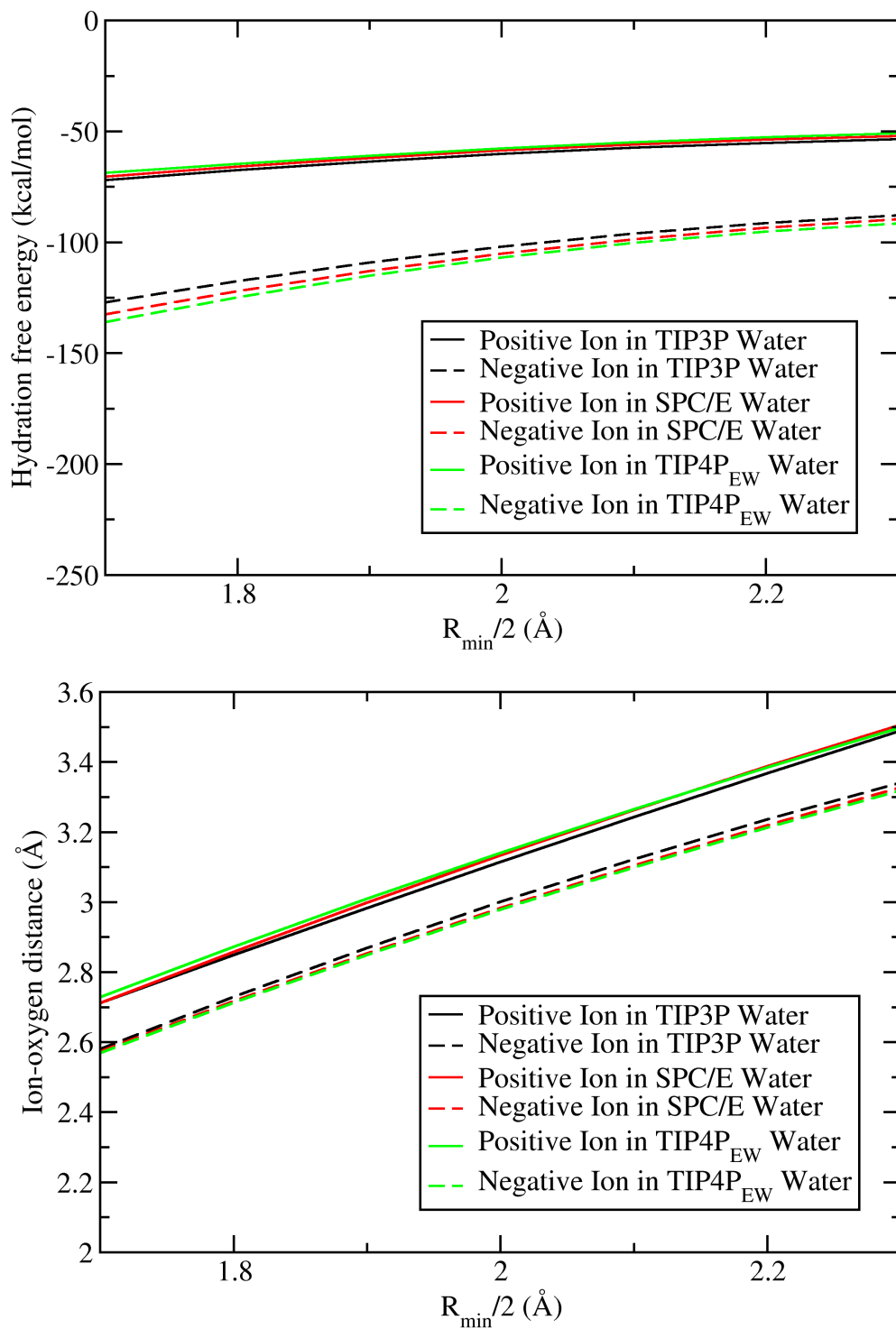


Figure 22. Fitting curves between the HFE and $R_{\min}/2$ values (upper) and fitting curves between the IOD and $R_{\min}/2$ values (lower) for the positive and negative monovalent ions in different water models.

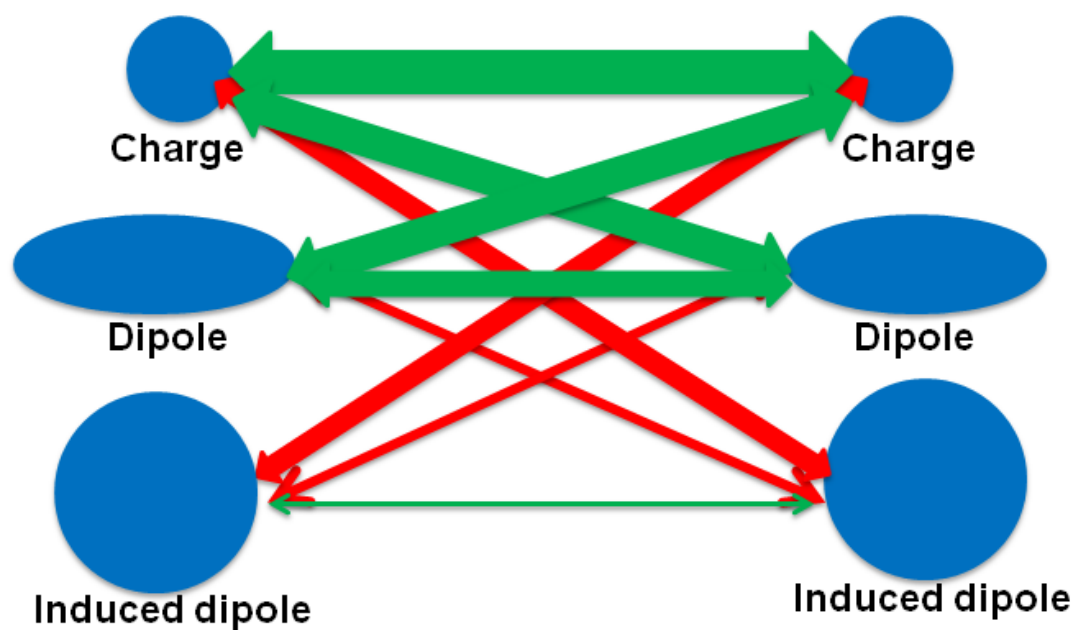


Figure 23. Scheme representing intermolecular interactions: the green double-headed arrow and red double-headed arrow represent the interactions that are included and not included in the 12-6 nonbonded model, respectively.

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															
			Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	
			Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lw	

Figure 24. Atomic ions that were parameterized for the 12-6-4 model in current work. These ions are shown with blue background. The monovalent anions, monovalent cations, divalent cations, trivalent cations, and tetravalent cations are shown in light green, yellow, orange, red, dark red, respectively. The ions that have multiple oxidation states are shown in white, which are $\text{Cr}^{2+}/\text{Cr}^{3+}$, $\text{Fe}^{2+}/\text{Fe}^{3+}$, $\text{Cu}^{+}/\text{Cu}^{2+}$, $\text{Ce}^{3+}/\text{Ce}^{4+}$, and $\text{Tl}^{+}/\text{Tl}^{3+}$. Besides these atomic ions, the H_3O^{+} and NH_4^{+} ions were also parameterized in current work.

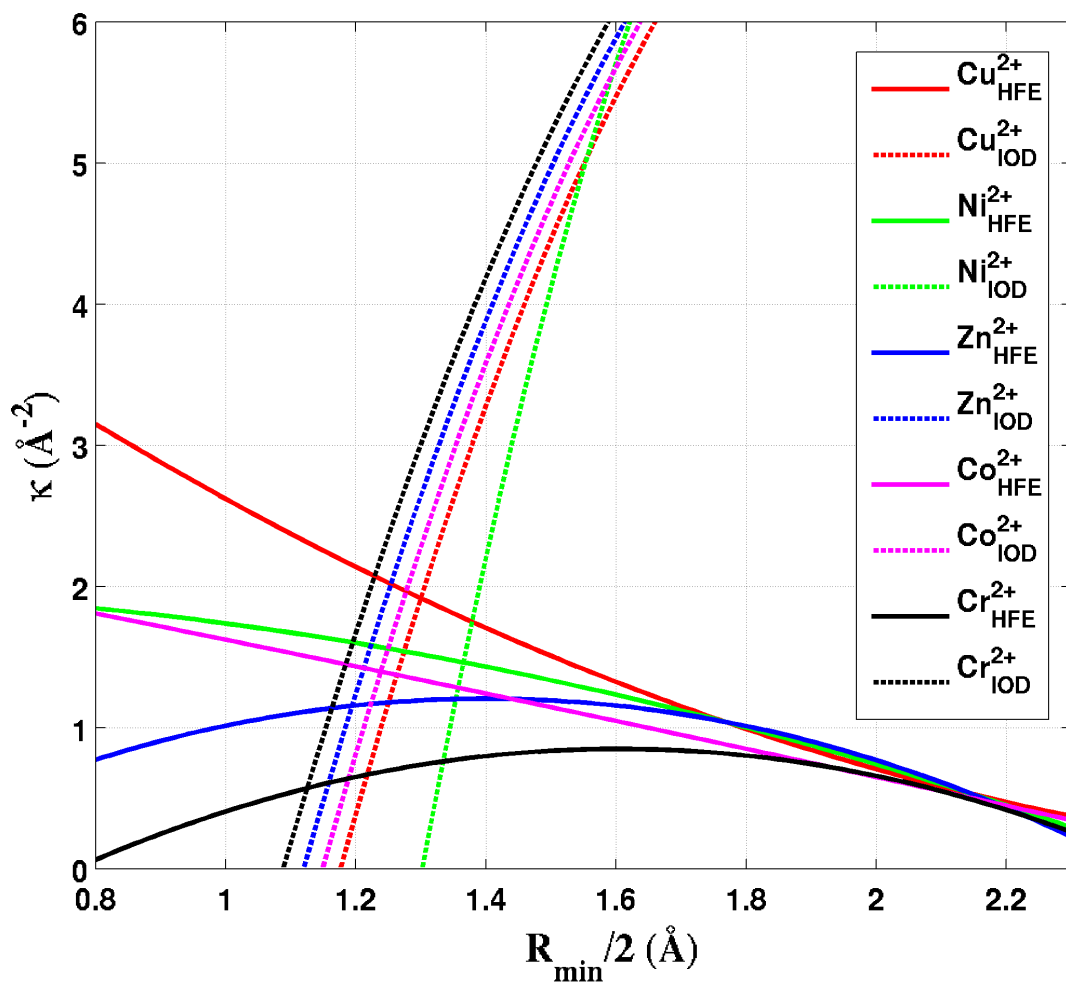


Figure 25. HFE and IOD fitting curves for the Cu^{2+} , Ni^{2+} , Zn^{2+} , Co^{2+} , and Cr^{2+} ions in the TIP3P water model.

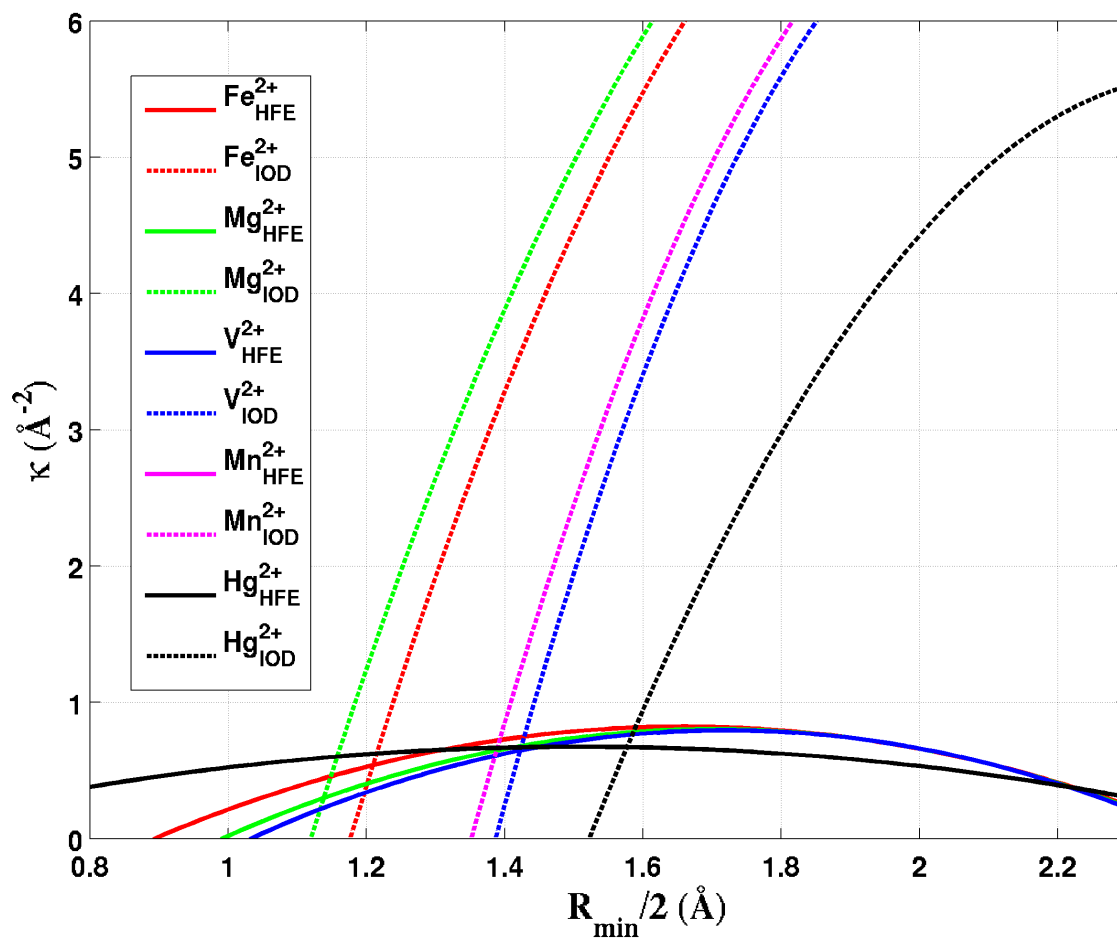


Figure 26. HFE and IOD fitting curves for the Fe^{2+} , Mg^{2+} , V^{2+} , Mn^{2+} , and Hg^{2+} ions in the TIP3P water model.

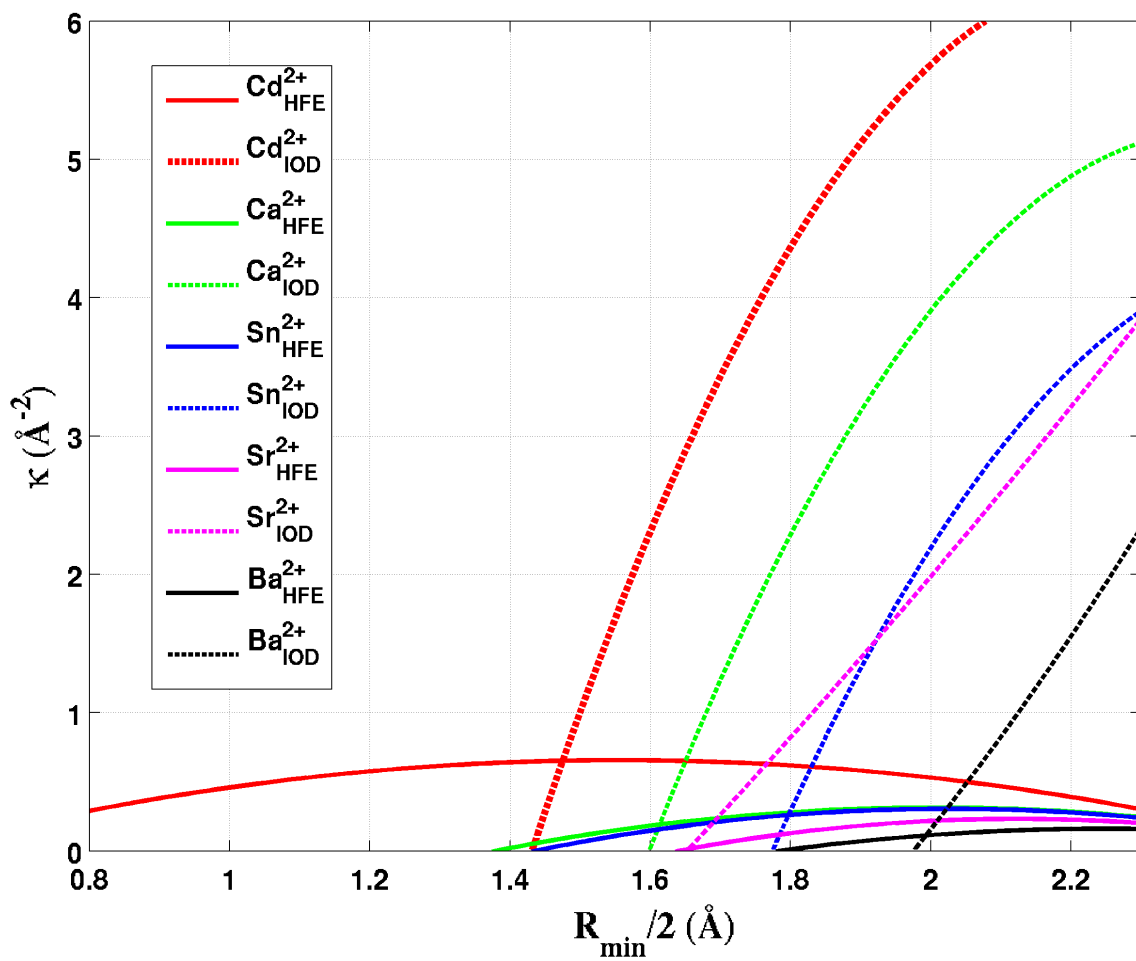


Figure 27. HFE and IOD fitting curves for the Cd^{2+} , Ca^{2+} , Sn^{2+} , Sr^{2+} , and Ba^{2+} ions in the TIP3P water.

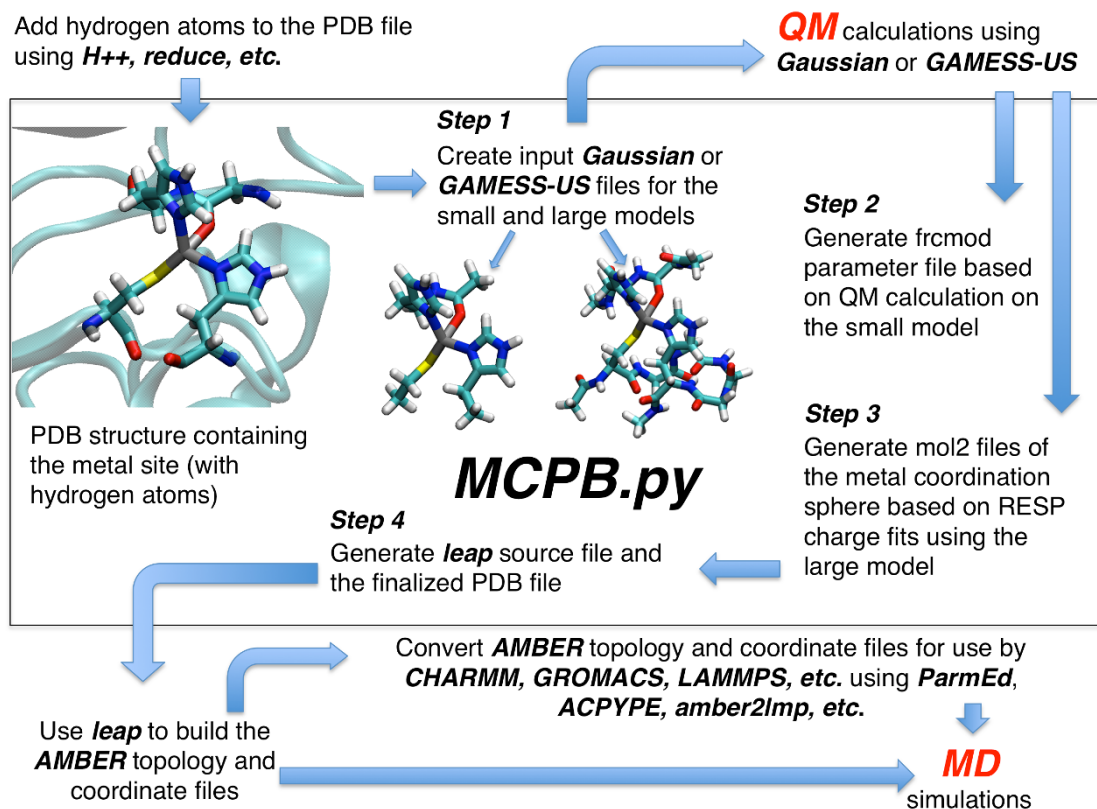


Figure 28. Workflow of the MCPB.py program.

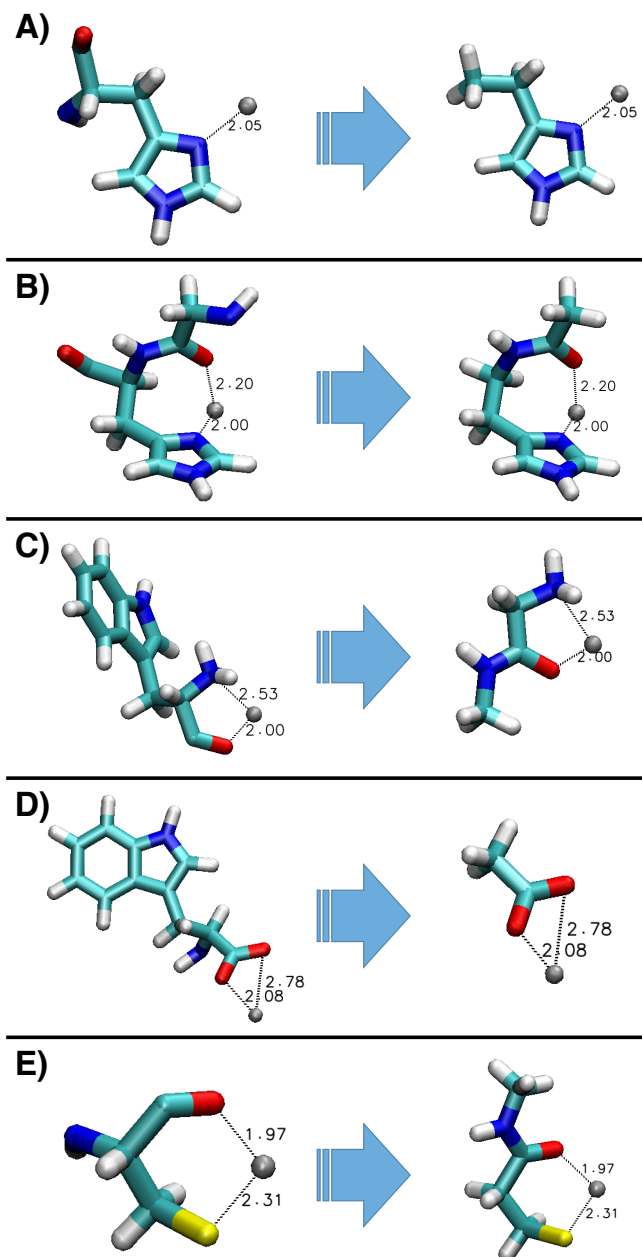


Figure 29. Examples of capping different metal bound amino acid residues in the small model. Structures on the left are the coordinated amino acids in the protein, and the structures on the right are the capped residues in the small models built by MCPB.py. Panels A and B are for PDB entry 1E67, panels C, D, and E are for PDB entries 1AK0, 1Y79, and 2BZS, respectively. Zinc coordination is shown as dotted lines with the bond lengths given in Å.

H																	He
Li	Be											B	C	N	O	F	Ne
Na	Mg											Al	Si	P	S	Cl	Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															
Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu				
Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr				

Figure 30. Metal ions currently supported by the MCPB.py program. The metals with a blue background use the VDW parameters from section 2 while the metals with green background use the VDW parameters adapted from UFF.²

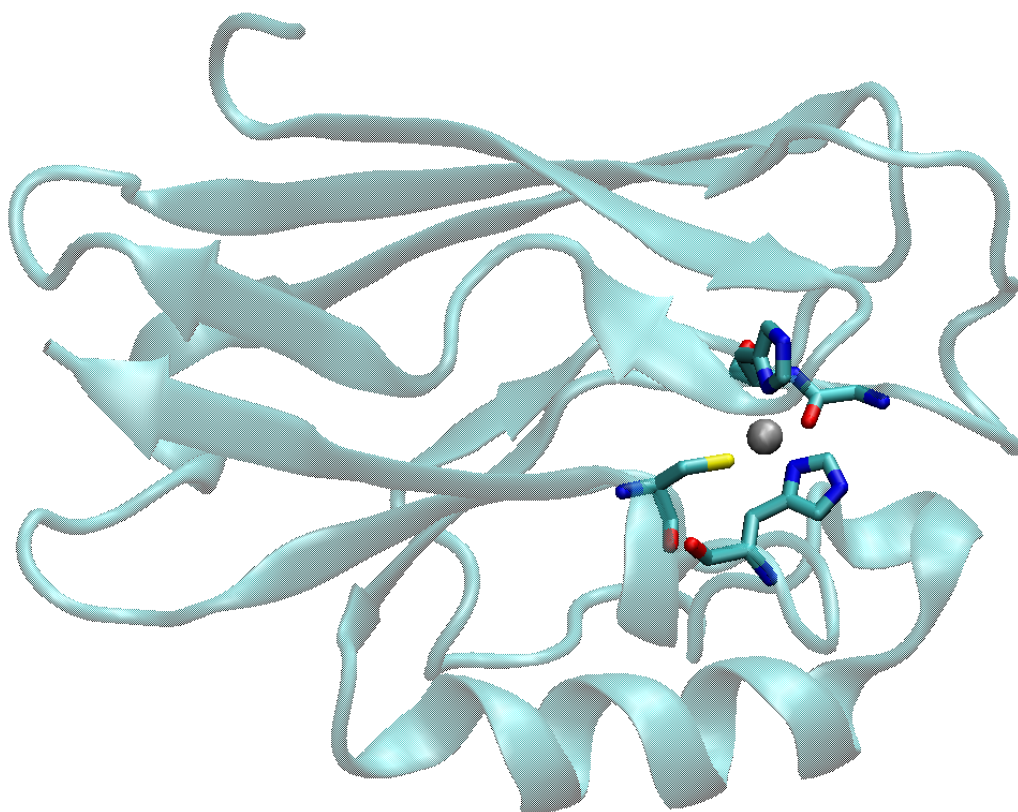


Figure 31. Structure of chain A from PDB entry 1E67 with the zinc ion represented by a VDW sphere and the metal site residues indicated by sticks. This figure was made using VMD.¹

```
original_pdb 1E67_fixed_H.pdb  
group_name 1E67  
cut_off 2.7  
ion_ids 1931  
ion_mol2files ZN.mol2
```

Figure 32. Content of the MCPB.py input file for the construction of the metal site found in chain A of PDB entry 1E67.

```
MCPB.py -i 1E67.in -s 1
g03 < 1E67_small_opt.com > 1E67_small_opt.log
g03 < 1E67_small_fc.com > 1E67_small_fc.log
formchk 1E67_small_opt.chk
g03 < 1E67_large_mk.com > 1E67_large_mk.log
MCPB.py -i 1E67.in -s 2
MCPB.py -i 1E67.in -s 3
MCPB.py -i 1E67.in -s 4
tleap -s -f 1E67_tleap.in > 1E67_tleap.out
```

Figure 33. Commands used for the construction of the metal site found in chain A of PDB entry 1E67 using MCPB.py. The second to the fifth lines are for the QM calculations using Gaussian03, while the last line is to build the topology and coordinate files using tleap.

BOND			
Y1-M1	27.1	2.0672	Created by Seminario method using MCPB.py
Y2-M1	61.8	2.0539	Created by Seminario method using MCPB.py
Y3-M1	96.5	2.2398	Created by Seminario method using MCPB.py
Y4-M1	70.6	2.0312	Created by Seminario method using MCPB.py
C -Y1	570.0	1.229	JCC,7,(1986),230; AA,CYT,GUA,THY,URA
CC-Y2	410.0	1.394	JCC,7,(1986),230; HIS
CC-Y4	410.0	1.394	JCC,7,(1986),230; HIS
CT-Y3	237.0	1.810	changed from 222.0 based on methanethiol nmodes
Y2-CR	488.0	1.335	JCC,7,(1986),230; HIS
Y4-CR	488.0	1.335	JCC,7,(1986),230; HIS
ANGL			
C -Y1-M1	17.95	137.78	Created by Seminario method using MCPB.py
CC-Y2-M1	53.19	129.51	Created by Seminario method using MCPB.py
CC-Y4-M1	62.23	127.49	Created by Seminario method using MCPB.py
CT-Y3-M1	82.67	103.87	Created by Seminario method using MCPB.py
M1-Y2-CR	47.72	123.07	Created by Seminario method using MCPB.py
M1-Y4-CR	65.46	124.24	Created by Seminario method using MCPB.py
Y1-M1-Y2	35.60	93.07	Created by Seminario method using MCPB.py
Y1-M1-Y3	16.02	117.34	Created by Seminario method using MCPB.py
Y1-M1-Y4	41.15	94.09	Created by Seminario method using MCPB.py
Y2-M1-Y3	34.92	120.93	Created by Seminario method using MCPB.py
Y2-M1-Y4	32.58	107.73	Created by Seminario method using MCPB.py
Y3-M1-Y4	19.26	118.04	Created by Seminario method using MCPB.py
CC-Y2-CR	70.0	117.00	AA his
CC-Y4-CR	70.0	117.00	AA his
CT-CC-Y2	70.0	120.00	AA his
CT-CC-Y4	70.0	120.00	AA his
CW-CC-Y2	70.0	120.00	AA his
CW-CC-Y4	70.0	120.00	AA his
CX-C -Y1	80.0	120.40	(was CT-C-0)
CX-CT-Y3	50.0	108.60	AA cys (was CT-CT-SH)
N -C -Y1	80.0	122.90	AA general
Y2-CR-H5	50.0	120.00	AA his
Y2-CR-NA	70.0	120.00	AA his
Y3-CT-H1	50.0	109.50	AA cyx changed based on NMA nmodes
Y4-CR-H5	50.0	120.00	AA his
Y4-CR-NA	70.0	120.00	AA his

Figure 34. Bond parameters (above) and angle parameters (below) for the metal site in chain A from the 1E67 PDB structure. Here M1 is the atom type of the zinc ion, while Y1, Y2, Y3, Y4 are the atom types for the four atoms bound to the zinc ion, which are the GLY45 backbone oxygen atom, the d nitrogen atom in HIE46, the sidechain sulfur atom in CYM112 and the d nitrogen atom in HIE117, respectively. Here HIE and CYM are the “AMBER style” residue names: HIE means a His residue which has the e nitrogen protonated, CYM indicates a Cys residue which is negatively charged (*i.e.* with a CH₂S⁻ sidechain group).

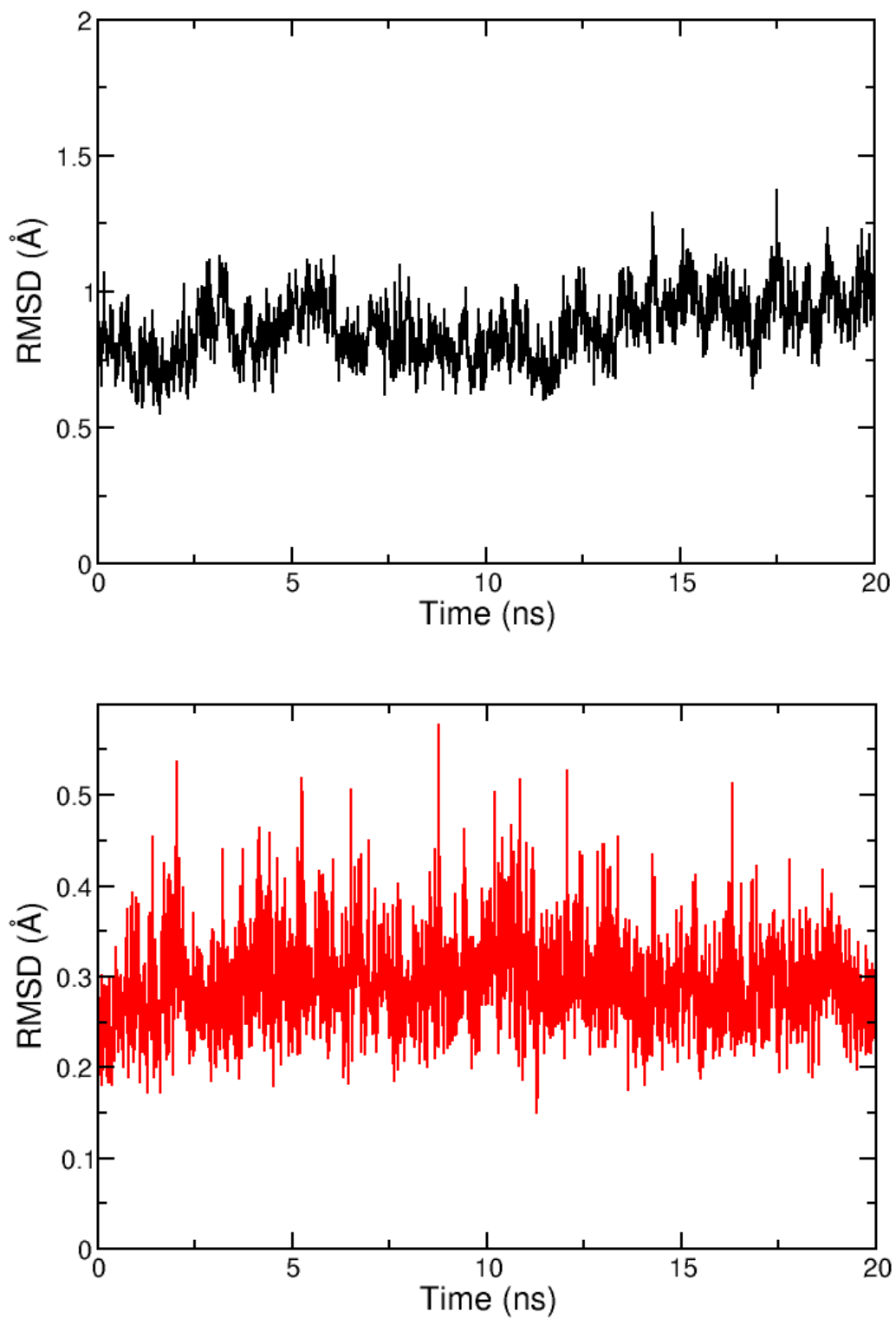


Figure 35. RMSD values of the protein backbone N, CA and C atoms (upper) and the heavy atoms in the metal site (lower) over 20 ns of simulation. The RMSD values were calculated using CPPTRAJ.³

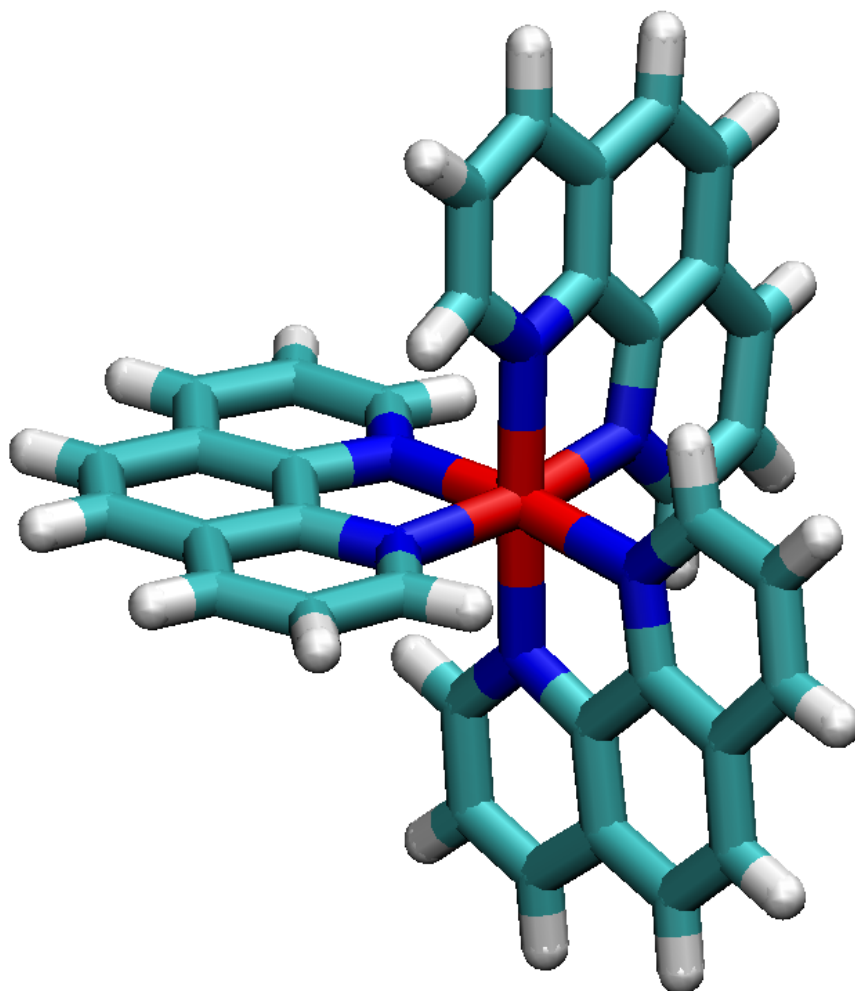


Figure 36. Structure of residue 1 from CSD entry FAJYAR01. This figure was made using VMD.¹

```
original_pdb FAJYAR01.pdb
group_name FAJYAR01
ion_ids 1
ion_mol2files OS.mol2
naa_mol2files RES.mol2
frcmod_files RES.frcmod
```

Figure 37. Content of the MCPB.py input file for the construction of the metal site found in residue 1 of CSD entry FAJYAR01.

```
MCPB.py -i FAJYAR01.in -s 1
g03 < FAJYAR01_small_opt.com > FAJYAR01_small_opt.log
g03 < FAJYAR01_small_fc.com > FAJYAR01_small_fc.log
formchk FAJYAR01_small_opt.chk
g03 < FAJYAR01_large_mk.com > FAJYAR01_large_mk.log
MCPB.py -i FAJYAR01.in -s 2
MCPB.py -i FAJYAR01.in -s 3
MCPB.py -i FAJYAR01.in -s 4
tleap -s -f FAJYAR01_tleap.in > FAJYAR01_tleap.out
```

Figure 38. Commands used for the construction of the metal site found in residue 1 of CSD entry FAJYAR01 using MCPB.py. The second to the fifth lines are for the QM calculations using Gaussian03, while the last line is to build the topology and coordinate files using tleap.

BOND					
M1-Y1	80.2	2.1026	Created by Seminario method using MCPB.py		
M1-Y2	76.1	2.1020	Created by Seminario method using MCPB.py		
M1-Y3	81.4	2.1024	Created by Seminario method using MCPB.py		
M1-Y4	78.6	2.1025	Created by Seminario method using MCPB.py		
M1-Y5	76.4	2.1024	Created by Seminario method using MCPB.py		
M1-Y6	77.5	2.1024	Created by Seminario method using MCPB.py		
Y1-ca	483.1	1.3420	SOURCE3	104	0.0076
Y3-ca	483.1	1.3420	SOURCE3	104	0.0076
Y5-ca	483.1	1.3420	SOURCE3	104	0.0076
ca-Y2	483.1	1.3420	SOURCE3	104	0.0076
ca-Y4	483.1	1.3420	SOURCE3	104	0.0076
ca-Y6	483.1	1.3420	SOURCE3	104	0.0076
ANGL					
M1-Y1-ca	144.27	121.08	Created by Seminario method using MCPB.py		
M1-Y2-ca	144.88	121.08	Created by Seminario method using MCPB.py		
M1-Y3-ca	146.06	121.08	Created by Seminario method using MCPB.py		
M1-Y4-ca	145.17	121.08	Created by Seminario method using MCPB.py		
M1-Y5-ca	142.46	121.08	Created by Seminario method using MCPB.py		
M1-Y6-ca	151.58	121.08	Created by Seminario method using MCPB.py		
Y2-M1-Y1	197.44	78.94	Created by Seminario method using MCPB.py		
Y3-M1-Y1	153.76	96.01	Created by Seminario method using MCPB.py		
Y3-M1-Y2	219.58	173.03	Created by Seminario method using MCPB.py		
Y4-M1-Y1	140.45	89.31	Created by Seminario method using MCPB.py		
Y4-M1-Y2	152.39	96.09	Created by Seminario method using MCPB.py		
Y4-M1-Y3	200.41	78.93	Created by Seminario method using MCPB.py		
Y5-M1-Y1	157.32	96.03	Created by Seminario method using MCPB.py		
Y5-M1-Y2	138.66	89.37	Created by Seminario method using MCPB.py		
Y5-M1-Y3	155.74	95.99	Created by Seminario method using MCPB.py		
Y5-M1-Y4	220.28	173.01	Created by Seminario method using MCPB.py		
Y6-M1-Y1	217.79	173.01	Created by Seminario method using MCPB.py		
Y6-M1-Y2	147.77	96.04	Created by Seminario method using MCPB.py		
Y6-M1-Y3	134.31	89.38	Created by Seminario method using MCPB.py		
Y6-M1-Y4	152.07	96.10	Created by Seminario method using MCPB.py		
Y6-M1-Y5	194.35	78.94	Created by Seminario method using MCPB.py		
Y1-ca-ca	69.16	122.63	SOURCE3	83	1.1249
Y1-ca-h4	51.82	115.94	SOURCE3	52	0.7370
Y3-ca-ca	69.16	122.63	SOURCE3	83	1.1249
Y3-ca-h4	51.82	115.94	SOURCE3	52	0.7370
Y5-ca-ca	69.16	122.63	SOURCE3	83	1.1249
Y5-ca-h4	51.82	115.94	SOURCE3	52	0.7370
ca-Y1-ca	68.59	115.86	SOURCE3	46	1.1645
ca-Y2-ca	68.59	115.86	SOURCE3	46	1.1645
ca-Y3-ca	68.59	115.86	SOURCE3	46	1.1645
ca-Y4-ca	68.59	115.86	SOURCE3	46	1.1645
ca-Y5-ca	68.59	115.86	SOURCE3	46	1.1645
ca-Y6-ca	68.59	115.86	SOURCE3	46	1.1645
ca-ca-Y2	69.16	122.63	SOURCE3	83	1.1249
ca-ca-Y4	69.16	122.63	SOURCE3	83	1.1249
ca-ca-Y6	69.16	122.63	SOURCE3	83	1.1249
h4-ca-Y2	51.82	115.94	SOURCE3	52	0.7370
h4-ca-Y4	51.82	115.94	SOURCE3	52	0.7370
h4-ca-Y6	51.82	115.94	SOURCE3	52	0.7370

Figure 39. Bond parameters (above) and angle parameters (below) for the metal site of residue 1 from CSD entry FAJYAR01. Here M1 is the atom type of the Os²⁺ ion, while Y1 to Y6 are the atom types for the six metal bound nitrogen atoms.

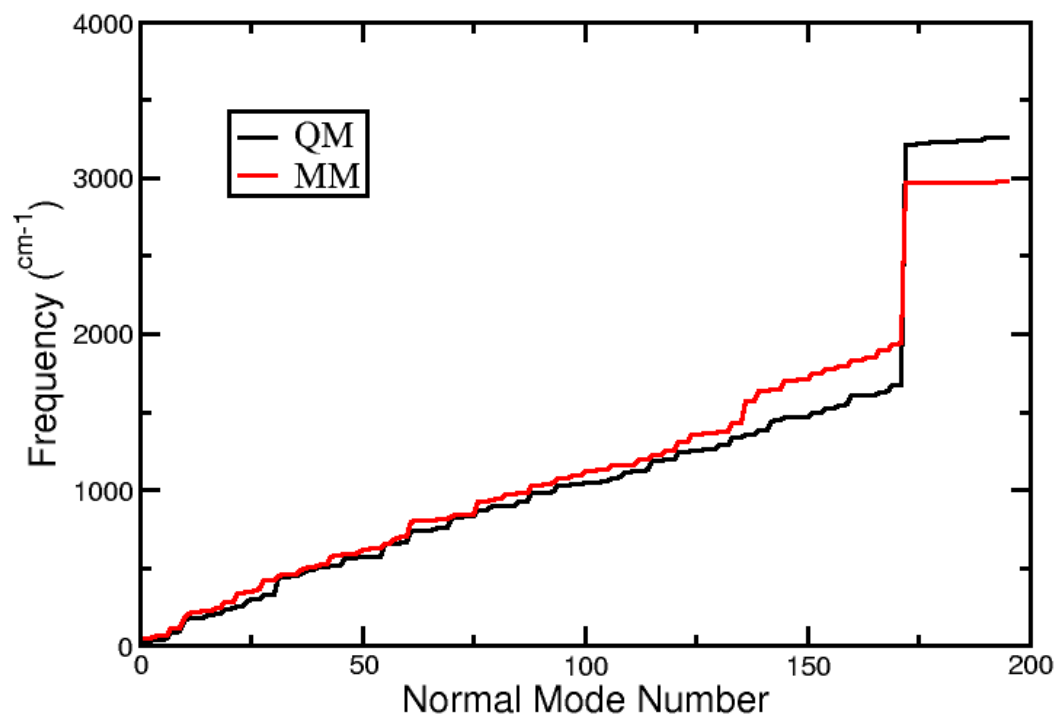


Figure 40. Comparison of the QM and MM calculated normal modes.

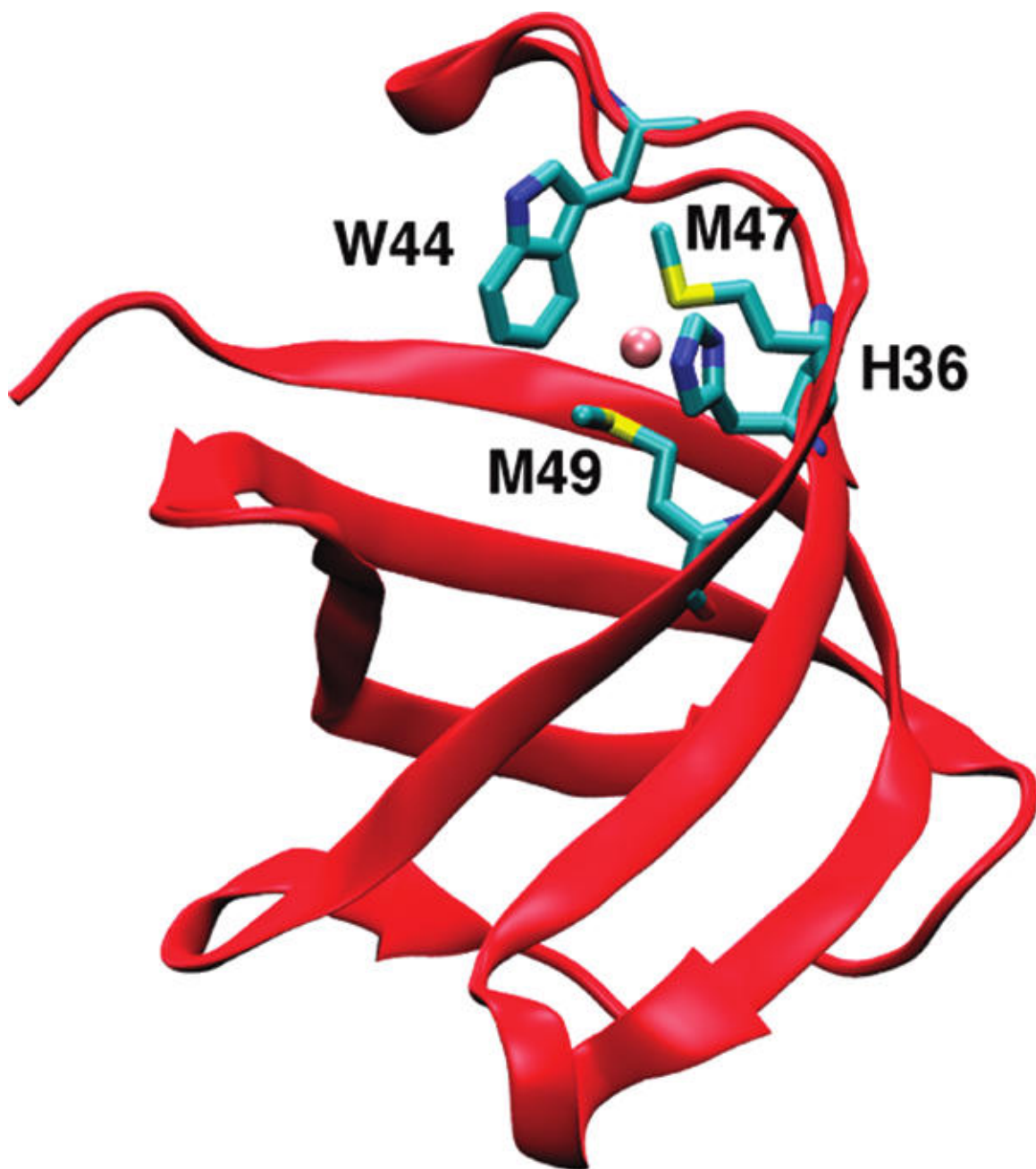


Figure 41. Cartoon representation of $\text{Cu(I)}\cdot\text{CusF}$ (PDB entry 2VB2). Cu^+ is shown as a pink sphere, and the metal binding residues are represented as sticks.

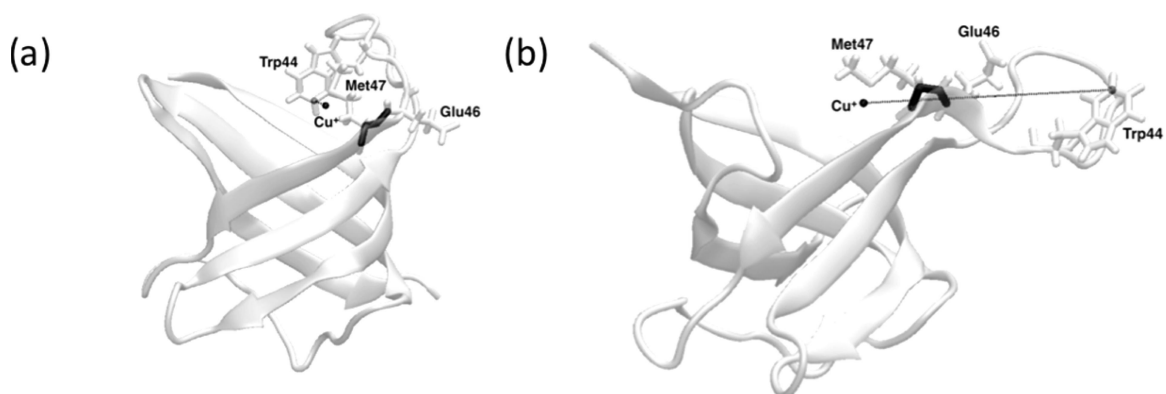


Figure 42. Cu^+ –Trp44 distance and Glu46–Met47 Φ dihedral coordinates (black lines) for our two-dimensional PMF MD simulations shown in (a) a closed structure and (b) a wide-open structure of $\text{Cu(I)}\cdot\text{CusF}$.

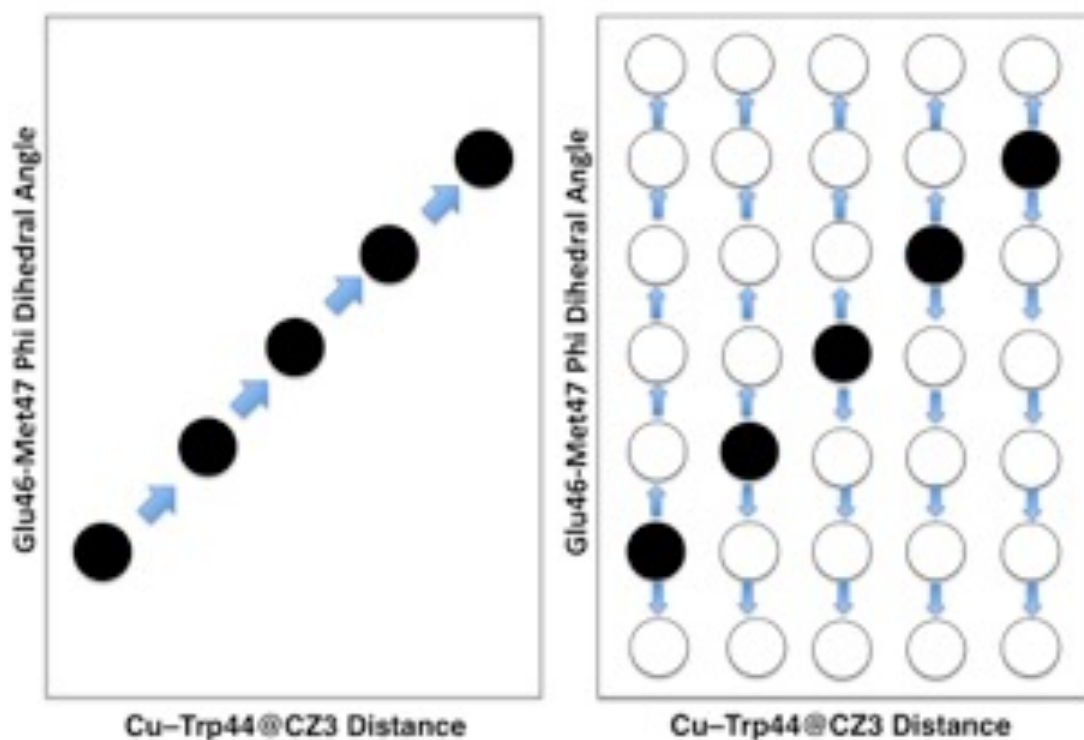


Figure 43. Scheme employed to generate the starting conformations for the two-dimensional PMF calculation. (Left) Initial conformations were generated by sampling along the diagonal from the closed to the wide-open conformations. Arrows indicate the direction of sampling along the reaction coordinate. Equilibrated conformation from a window was used as the starting coordinate of the adjacent window as we moved along the diagonal. (Right) Scheme employed to sample conformations starting from windows along the diagonal (solid circles) to the hollow circles.

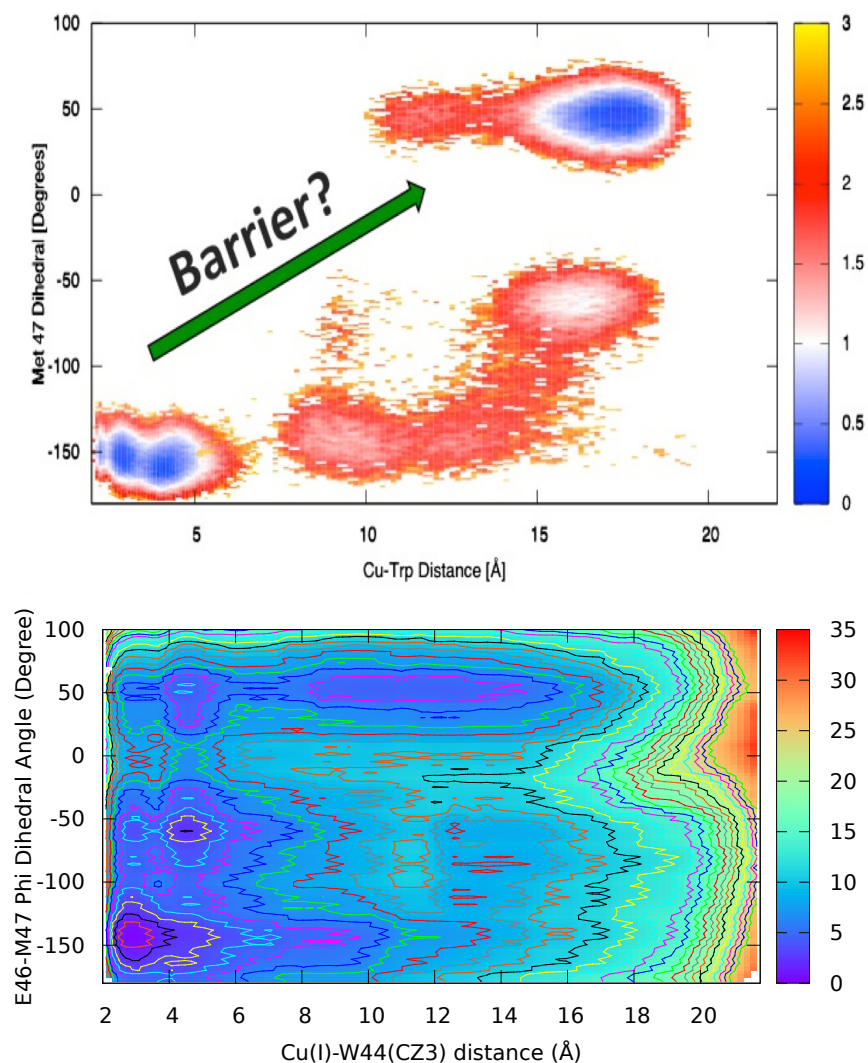


Figure 44. (Upper) A two-dimensional free energy landscape calculated for the Cu-Trp44 distance *versus* the Glu46-Met47 Φ dihedral from our classical MD simulations of Cu(I)•CusF. Cu(I)•CusF is unable to make the transition from the closed to the open conformations. The closed conformations are represented by smaller Cu⁺-Trp distances, while the open conformations are represented by larger Cu⁺-Trp44 distances. (Lower) PMF profile based on the umbrella sampling simulations in which the Cu⁺-Trp44 distance and Glu46-Met47 Φ dihedral were gradually modified to force the transition from the closed conformation to the wide-open conformation of Cu(I)•CusF. Error bars were calculated using a bootstrapping algorithm implemented in Alan Grossfield's WHAM code (<http://membrane.urmc.rochester.edu/content/wham>). Error bars were calculated along each axis of the 2D PMF profile. The error bars along the dihedral dimension are between 0.002 and 0.040 kcal/mol with an average of 0.022 kcal/mol. The error bars along the distance dimension are in the range of 0.001 to 0.128 kcal/mol with an average of 0.012 kcal/mol.

APPENDIX C: COPYRIGHT NOTICE

Figure 1 and Figure 24 of this dissertation are copyrighted by PENGFEI LI 2016.

Figure 41 is reprinted with permission from ref 111. Copyright 2011 American Chemical Society.

The remaining parts of this dissertation (including its supporting information) are adapted with permissions from several publications listed below:

- (1) Adapted with permission from ref 143. Copyright 2015 American Chemical Society.
- (2) Adapted with permission from ref 144. Copyright 2013 American Chemical Society.
- (3) Adapted with permission from ref 145. Copyright 2014 American Chemical Society.

The paper can be accessed from webpage:

<http://pubs.acs.org/doi/abs/10.1021/jp505875v>. This is an unofficial adaptation of an article that appeared in an ACS publication. ACS has not endorsed the content of this adaptation or the context of its use.

- (4) Adapted with permission from ref 174. Copyright 2014 American Chemical Society.
- (5) Adapted with permission from ref 219. Copyright 2016 American Chemical Society.
- (6) Adapted with permission from ref 230. Copyright 2016 American Chemical Society.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Humphrey, W.; Dalke, A.; Schulten, K., Vmd: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, 14, 33-38.
2. Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skiff, W. M., Uff, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, 114, 10024-10035.
3. Roe, D. R.; Cheatham III, T. E., Ptraj and Cpptraj: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, 9, 3084-3095.
4. Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz Jr, K. M., Structural Survey of Zinc-Containing Proteins and Development of the Zinc Amber Force Field (Zaff). *J. Chem. Theory Comput.* **2010**, 6, 2935-2947.
5. Tus, A.; Rakipović, A.; Peretin, G.; Tomić, S.; Šikić, M., Biome: Biologically Relevant Metals. *Nucleic Acids Res.* **2012**, 40, W352-W357.
6. Waldron, K. J.; Robinson, N. J., How Do Bacterial Cells Ensure That Metalloproteins Get the Correct Metal? *Nat. Rev. Microbiol.* **2009**, 7, 25-35.
7. Dupureur, C. M., Roles of Metal Ions in Nucleases. *Curr. Opin. Chem. Biol.* **2008**, 12, 250-255.
8. Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G.; Thornton, J., Metal Ions in Biological Catalysis: From Enzyme Databases to General Principles. *J Biol Inorg Chem* **2008**, 13, 1205-1218.
9. Chaturvedi, U. C.; Shrivastava, R., Interaction of Viral Proteins with Metal Ions: Role in Maintaining the Structure and Functions of Viruses. *FEMS Immunol. Med. Microbiol.* **2005**, 43, 105-114.
10. Rosenzweig, A. C., Metallochaperones: Bind and Deliver. *Chem. Biol.* **2002**, 9, 673-677.
11. Thomson, A. J.; Gray, H. B., Bio-Inorganic Chemistry. *Curr. Opin. Chem. Biol.* **1998**, 2, 155-158.
12. Berridge, M. J.; Bootman, M. D.; Lipp, P., Calcium - a Life and Death Signal. *Nature* **1998**, 395, 645-648.
13. Christianson, D. W., Structural Chemistry and Biology of Manganese Metalloenzymes. *Prog. Biophys. Mol. Biol.* **1997**, 67, 217-252.

14. Lipscomb, W. N.; Sträter, N., Recent Advances in Zinc Enzymology. *Chem. Rev.* **1996**, 96, 2375-2434.
15. Linder, M. C.; Hazegh-Azam, M., Copper Biochemistry and Molecular Biology. *Am. J. Clin. Nutr.* **1996**, 63, 797S-811S.
16. Holm, R. H.; Kennepohl, P.; Solomon, E. I., Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **1996**, 96, 2239-2314.
17. Coleman, J. E., Zinc Proteins: Enzymes, Storage Proteins, Transcription Factors, and Replication Proteins. *Annu. Rev. Biochem.* **1992**, 61, 897-946.
18. Berg, J. M., Zinc Finger Domains: Hypotheses and Current Knowledge. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, 19, 405-421.
19. Jacoby, M., How Computers Are Helping Chemists Discover Materials That Don't yet Exist. *Chem. Eng. News* **2015**, 93, 8-11.
20. Dudev, T.; Lim, C., Principles Governing Mg, Ca, and Zn Binding and Selectivity in Proteins. *Chem. Rev.* **2003**, 103, 773-788.
21. Dudev, T.; Lim, C., Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, 114, 538-556.
22. Valenzano, L.; Civalieri, B.; Sillar, K.; Sauer, J., Heats of Adsorption of Co and Co₂ in Metal–Organic Frameworks: Quantum Mechanical Study of Cpo-27-M (M= Mg, Ni, Zn). *J. Phys. Chem. C* **2011**, 115, 21777-21784.
23. Pang, Y.-P., Successful Molecular Dynamics Simulation of Two Zinc Complexes Bridged by a Hydroxide in Phosphotriesterase Using the Cationic Dummy Atom Method. *Proteins: Struct., Funct., Bioinf.* **2001**, 45, 183-189.
24. Yang, Q.; Zhong, C., Molecular Simulation of Carbon Dioxide/Methane/Hydrogen Mixture Adsorption in Metal-Organic Frameworks. *J. Phys. Chem. B* **2006**, 110, 17776-17783.
25. Wu, R.; Hu, P.; Wang, S.; Cao, Z.; Zhang, Y., Flexibility of Catalytic Zinc Coordination in Thermolysin and Hdac8: A Born-Oppenheimer Ab Initio Qm/Mm Molecular Dynamics Study. *J. Chem. Theory Comput.* **2009**, 6, 337-343.
26. Hofer, T. S.; Weiss, A. K.; Randolph, B. R.; Rode, B. M., Hydration of Highly Charged Ions. *Chem. Phys. Lett.* **2011**, 512, 139-145.
27. Sokol, A. A.; Bromley, S. T.; French, S. A.; Catlow, C. R. A.; Sherwood, P., Hybrid Qm/Mm Embedding Approach for the Treatment of Localized Surface States in Ionic Materials. *Int. J. Quantum Chem.* **2004**, 99, 695-712.

28. Yu, D.; Yazaydin, A. O.; Lane, J. R.; Dietzel, P. D.; Snurr, R. Q., A Combined Experimental and Quantum Chemical Study of Co₂ Adsorption in the Metal–Organic Framework Cpo-27 with Different Metals. *Chem. Sci.* **2013**, 4, 3544-3556.
29. Duan, Y.; Kollman, P. A., Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution. *Science* **1998**, 282, 740-744.
30. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y., Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, 330, 341-346.
31. Jorgensen, W. L.; Ravimohan, C., Monte Carlo Simulation of Differences in Free Energies of Hydration. *J. Chem. Phys.* **1985**, 83, 3050-3054.
32. Jorgensen, W. L.; Duffy, E. M., Prediction of Drug Solubility from Monte Carlo Simulations. *Bioorg. Med. Chem. Lett.* **2000**, 10, 1155-1158.
33. Hoops, S. C.; Anderson, K. W.; Merz, K. M., Force Field Design for Metalloproteins. *J. Am. Chem. Soc.* **1991**, 113, 8262-8270.
34. Roberts, B. P.; Miller III, B. R.; Roitberg, A. E.; Merz Jr, K. M., Wide-Open Flaps Are Key to Urease Activity. *J. Am. Chem. Soc.* **2012**, 134, 9934-9937.
35. Chakravorty, D.; Wang, B.; Lee, C.; Guerra, A.; Giedroc, D.; Merz, K., Jr., Solution Nmr Refinement of a Metal Ion Bound Protein Using Metal Ion Inclusive Restrained Molecular Dynamics Methods. *J. Biomol. NMR* **2013**, 56, 125-137.
36. Ucisik, M. N.; Chakravorty, D. K.; Merz Jr, K. M., Structure and Dynamics of the N-Terminal Domain of the Cu (I) Binding Protein Cusb. *Biochemistry* **2013**, 52, 6911-6923.
37. Åqvist, J., Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, 94, 8021-8024.
38. Stote, R. H.; Karplus, M., Zinc Binding in Proteins and Solution: A Simple but Accurate Nonbonded Representation. *Proteins: Struct., Funct., Bioinf.* **1995**, 23, 12-31.
39. Joung, I. S.; Cheatham, T. E., Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, 112, 9020-9041.
40. Merz, K. M., Carbon Dioxide Binding to Human Carbonic Anhydrase Ii. *J. Am. Chem. Soc.* **1991**, 113, 406-411.
41. Chakravorty, D. K.; Wang, B.; Ucisik, M. N.; Merz, K. M., Insight into the Cation–Π Interaction at the Metal Binding Site of the Copper Metallochaperone Cusf. *J. Am. Chem. Soc.* **2011**, 133, 19330-19333.

42. Saxena, A.; Sept, D., Multisite Ion Models That Improve Coordination and Free Energy Calculations in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, 9, 3538-3542.
43. Duarte, F.; Bauer, P.; Barrozo, A.; Amrein, B. A.; Purg, M.; Åqvist, J.; Kamerlin, S. C. L., Force-Field Independent Metal Parameters Using a Non-Bonded Dummy Model. *J. Phys. Chem. B* **2014**, 118, 4351-4362.
44. Hu, L.; Ryde, U., Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, 7, 2452-2463.
45. Hay, B. P., Methods for Molecular Mechanics Modeling of Coordination Compounds. *Coord. Chem. Rev.* **1993**, 126, 177-236.
46. Sakharov, D. V.; Lim, C., Force Fields Including Charge Transfer and Local Polarization Effects: Application to Proteins Containing Multi/Heavy Metal Ions. *J. Comput. Chem.* **2009**, 30, 191-202.
47. Yu, H.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D.; Roux, B. t., Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, 6, 774-786.
48. Zhang, J.; Yang, W.; Piquemal, J.-P.; Ren, P., Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential. *J. Chem. Theory Comput.* **2012**, 8, 1314-1324.
49. Soniat, M.; Rick, S. W., The Effects of Charge Transfer on the Aqueous Solvation of Ions. *J. Chem. Phys.* **2012**, 137, 044511.
50. Ponomarev, S. Y.; Click, T. H.; Kaminski, G. A., Electrostatic Polarization Is Crucial in Reproducing Cu(I) Interaction Energies and Hydration. *J. Phys. Chem. B* **2011**, 115, 10079-10085.
51. Molina, J. J.; Lectez, S.; Tazi, S.; Salanne, M.; Dufreche, J.-F.; Roques, J.; Simoni, E.; Madden, P. A.; Turq, P., Ions in Solutions: Determining Their Polarizabilities from First-Principles. *J. Chem. Phys.* **2011**, 134, 014511.
52. Jorgensen, W. L., Quantum and Statistical Mechanical Studies of Liquids. 10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers. Application to Liquid Water. *J. Am. Chem. Soc.* **1981**, 103, 335-340.
53. Dang, L. X., Development of Nonadditive Intermolecular Potentials Using Molecular Dynamics: Solvation of Li⁺ and F⁻ Ions in Polarizable Water. *J. Chem. Phys.* **1992**, 96, 6970-6977.
54. Dang, L. X.; Garrett, B. C., Photoelectron Spectra of the Hydrated Iodine Anion from Molecular Dynamics Simulations. *J. Chem. Phys.* **1993**, 99, 2972-2977.

55. Dang, L. X., Free Energies for Association of Cs⁺ to 18-Crown-6 in Water. A Molecular Dynamics Study Including Counter Ions. *Chem. Phys. Lett.* **1994**, 227, 211-214.
56. Smith, D. E.; Dang, L. X., Computer Simulations of NaCl Association in Polarizable Water. *J. Chem. Phys.* **1994**, 100, 3757-3766.
57. Dang, L. X., Mechanism and Thermodynamics of Ion Selectivity in Aqueous Solutions of 18-Crown-6 Ether: A Molecular Dynamics Study. *J. Am. Chem. Soc.* **1995**, 117, 6954-6960.
58. Chang, T.-M.; Dang, L. X., Detailed Study of Potassium Solvation Using Molecular Dynamics Techniques. *J. Phys. Chem. B* **1999**, 103, 4714-4720.
59. Dang, L. X., Computational Study of Ion Binding to the Liquid Interface of Water. *J. Phys. Chem. B* **2002**, 106, 10388-10394.
60. Peng, T.; Chang, T.-M.; Sun, X.; Nguyen, A. V.; Dang, L. X., Development of Ions-Tip4p-Ew Force Fields for Molecular Processes in Bulk and at the Aqueous Interface Using Molecular Simulations. *J. Mol. Liq.* **2012**, 173, 47-54.
61. Peng, Z.; Ewig, C. S.; Hwang, M.-J.; Waldman, M.; Hagler, A. T., Derivation of Class II Force Fields. 4. Van Der Waals Parameters of Alkali Metal Cations and Halide Anions. *J. Phys. Chem. A* **1997**, 101, 7243-7252.
62. Jensen, K. P.; Jorgensen, W. L., Halide, Ammonium, and Alkali Metal Ion Parameters for Modeling Aqueous Solutions. *J. Chem. Theory Comput.* **2006**, 2, 1499-1509.
63. Lamoureux, G.; Roux, B., Absolute Hydration Free Energy Scale for Alkali and Halide Ions Established from Simulations with a Polarizable Force Field. *J. Phys. Chem. B* **2006**, 110, 3308-3322.
64. Babu, C. S.; Lim, C., Empirical Force Fields for Biologically Active Divalent Metal Cations in Water†. *J. Phys. Chem. A* **2005**, 110, 691-699.
65. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, 79, 926-935.
66. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, 91, 6269-6271.
67. Horn, H. W.; Swope, W. C.; Pitner, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T., Development of an Improved Four-Site Water Model for Biomolecular Simulations: Tip4p-Ew. *J. Chem. Phys.* **2004**, 120, 9665-9678.

68. Horinek, D.; Mamatkulov, S. I.; Netz, R. R., Rational Design of Ion Force Fields Based on Thermodynamic Solvation Properties. *J. Chem. Phys.* **2009**, 130, 124507.
69. Deublein, S.; Vrabec, J.; Hasse, H., A Set of Molecular Models for Alkali and Halide Ions in Aqueous Solution. *J. Chem. Phys.* **2012**, 136, 084501.
70. Reiser, S.; Deublein, S.; Vrabec, J.; Hasse, H., Molecular Dispersion Energy Parameters for Alkali and Halide Ions in Aqueous Solution. *J. Chem. Phys.* **2014**, 140, 044504.
71. Reif, M. M.; Hunenberger, P. H., Computation of Methodology-Independent Single-Ion Solvation Properties from Molecular Simulations. Iv. Optimized Lennard-Jones Interaction Parameter Sets for the Alkali and Halide Ions in Water. *J. Chem. Phys.* **2011**, 134, 144104.
72. Åqvist, J., Comment on "Transferability of Ion Models". *J. Phys. Chem.* **1994**, 98, 8253-8255.
73. Marrone, T. J.; Merz, K. M., Transferability of Ion Models. *J. Phys. Chem.* **1993**, 97, 6524-6529.
74. Marrone, T. J.; Merz, K. M., Reply to Comment on "Transferability of Ion Models". *J. Phys. Chem.* **1994**, 98, 8256-8257.
75. Fyta, M.; Netz, R. R., Ionic Force Field Optimization Based on Single-Ion and Ion-Pair Solvation Properties: Going Beyond Standard Mixing Rules. *J. Chem. Phys.* **2012**, 136, 124103.
76. Toukmaji, A.; Sagui, C.; Board, J.; Darden, T., Efficient Particle-Mesh Ewald Based Approach to Fixed and Induced Dipolar Interactions. *J. Chem. Phys.* **2000**, 113, 10913-10927.
77. Petersen, H. G., Accuracy and Efficiency of the Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, 103, 3668-3679.
78. Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A., Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, Rna, and Proteins. *J. Am. Chem. Soc.* **1995**, 117, 4193-4194.
79. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An Nlog(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, 98, 10089-10092.
80. Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F., Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, 222, 529-539.

81. Steinbrecher, T.; Mobley, D. L.; Case, D. A., Nonlinear Scaling Schemes for Lennard-Jones Interactions in Free Energy Calculations. *J. Chem. Phys.* **2007**, 127, 214108.
82. Kollman, P., Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, 93, 2395-2417.
83. Hummer, G.; Szabo, A., Calculation of Free-Energy Differences from Computer Simulations of Initial and Final States. *J. Chem. Phys.* **1996**, 105, 2004-2010.
84. Mitchell, M. J.; McCammon, J. A., Free Energy Difference Calculations by Thermodynamic Integration: Difficulties in Obtaining a Precise Value. *J. Comput. Chem.* **1991**, 12, 271-275.
85. Simonson, T.; Carlsson, J.; Case, D. A., Proton Binding to Proteins: Pka Calculations with Explicit and Implicit Solvent Models. *J. Am. Chem. Soc.* **2004**, 126, 4167-4180.
86. Mezei, M., The Finite Difference Thermodynamic Integration, Tested on Calculating the Hydration Free Energy Difference between Acetone and Dimethylamine in Water. *J. Chem. Phys.* **1987**, 86, 7084-7088.
87. Straatsma, T. P.; Berendsen, H. J. C., Free Energy of Ionic Hydration: Analysis of a Thermodynamic Integration Technique to Evaluate Free Energy Differences by Molecular Dynamics Simulations. *J. Chem. Phys.* **1988**, 89, 5876-5886.
88. Stokes, R., The Van Der Waals Radii of Gaseous Ions of the Noble Gas Structure in Relation to Hydration Energies. *J. Am. Chem. Soc.* **1964**, 86, 979-982.
89. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, 117, 5179-5197.
90. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. a., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, 102, 3586-3616.
91. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the Opls All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, 118, 11225-11236.
92. van Gunsteren, W. F.; Daura, X.; Mark, A. E. Gromos Force Field. In *Encyclopedia of Computational Chemistry*; John Wiley & Sons, Ltd: 2002.
93. Panteva, M. T.; Giambaşu, G. M.; York, D. M., Comparison of Structural, Thermodynamic, Kinetic and Mass Transport Properties of Mg²⁺ Ion Models Commonly Used in Biomolecular Simulations. *J. Comput. Chem.* **2015**, 36, 970-982.

94. Vedani, A.; Dobler, M.; Dunitz, J. D., An Empirical Potential Function for Metal Centers: Application to Molecular Mechanics Calculations on Metalloproteins. *J. Comput. Chem.* **1986**, 7, 701-710.
95. Nakagawa, I.; Shimanouchi, T., Infrared Absorption Spectra of Aquo Complexes and the Nature of Co-Ordination Bonds. *Spectrochimica Acta* **1964**, 20, 429-439.
96. Sacconi, L.; Sabatini, A.; Gans, P., Infrared Spectra from 80 to 2000 Cm⁻¹ of Some Metal-Ammine Complexes. *Inorg. Chem.* **1964**, 3, 1772-1774.
97. Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E., Systematic Ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives. *J. Am. Chem. Soc.* **1979**, 101, 2550-2560.
98. Seminario, J. M., Calculation of Intramolecular Force Fields from Second - Derivative Tensors. *Int. J. Quantum Chem.* **1996**, 60, 1271-1277.
99. Nilsson, K.; Lecerof, D.; Sigfridsson, E.; Ryde, U., An Automatic Method to Generate Force-Field Parameters for Hetero-Compounds. *Acta Crystallographica Section D: Biological Crystallography* **2003**, 59, 274-289.
100. Lin, F.; Wang, R., Systematic Derivation of Amber Force Field Parameters Applicable to Zinc-Containing Systems. *J. Chem. Theory Comput.* **2010**, 6, 1852-1870.
101. Chakravorty, D. K.; Wang, B.; Lee, C. W.; Giedroc, D. P.; Merz Jr, K. M., Simulations of Allosteric Motions in the Zinc Sensor Czra. *J. Am. Chem. Soc.* **2011**, 134, 3367-3376.
102. Chakravorty, D. K.; Wang, B.; Ucisik, M. N.; Merz Jr, K. M., Insight into the Cation- Π Interaction at the Metal Binding Site of the Copper Metallochaperone Cusf. *J. Am. Chem. Soc.* **2011**, 133, 19330-19333.
103. Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Madej, B.; Merz Jr, K. M.; Monard, G.; Needham, P.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Salomon-Ferrer, R.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; York, D. M.; Kollman, P. A. *Amber 2015*, University of California, San Francisco: 2015.
104. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; J. A. Montgomery, J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J.

J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision E. 01*, Gaussian Inc.: Wallingford, CT, 2004.

105. Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S., General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, 14, 1347-1363.

106. Sousa da Silva, A. W.; Vranken, W. F., Acpype - Antechamber Python Parser Interface. *BMC Res. Notes* **2012**, 5, 1-8.

107. Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., Charmm: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, 30, 1545-1614.

108. van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J., Gromacs: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, 26, 1701-1718.

109. Ma, Z.; Jacobsen, F. E.; Giedroc, D. P., Coordination Chemistry of Bacterial Metal Transport and Sensing. *Chem. Rev.* **2009**, 109, 4644-4681.

110. Su, C.-C.; Long, F.; Zimmermann, M. T.; Rajashankar, K. R.; Jernigan, R. L.; Edward, W. Y., Crystal Structure of the CusA Heavy-Metal Efflux Complex of Escherichia Coli. *Nature* **2011**, 470, 558-562.

111. Chakravorty, D. K.; Wang, B.; Ucisik, M. N.; Merz Jr, K. M., Insight into the Cation- π Interaction at the Metal Binding Site of the Copper Metallochaperone CusF. *J. Am. Chem. Soc.* **2011**, 133, 19330-19333.

112. Kim, E.-H.; Nies, D. H.; McEvoy, M. M.; Rensing, C., Switch or Funnel: How Rnd-Type Transport Systems Control Periplasmic Metal Homeostasis. *J. Bacteriol.* **2011**, 193, 2381-2387.

113. Chacón, K. N.; Mealman, T. D.; McEvoy, M. M.; Blackburn, N. J., Tracking Metal Ions through a Cu/Ag Efflux Pump Assigns the Functional Roles of the Periplasmic Proteins. *Proceedings of the National Academy of Sciences* **2014**, 111, 15373-15378.

114. Mealman, T. D.; Zhou, M.; Affandi, T.; Chacón, K. N.; Aranguren, M. E.; Blackburn, N. J.; Wysocki, V. H.; McEvoy, M. M., N-Terminal Region of CusB Is Sufficient for Metal Binding and Metal Transfer with the Metallochaperone CusF. *Biochemistry* **2012**, 51, 6767-6775.

115. Marcus, Y., Thermodynamics of Solvation of Ions. Part 5.-Gibbs Free Energy of Hydration at 298.15 K. *J. Chem. Soc., Faraday Trans.* **1991**, 87, 2995-2999.
116. Schmid, R.; Miah, A. M.; Sapunov, V. N., A New Table of the Thermodynamic Quantities of Ionic Hydration: Values and Some Applications (Enthalpy-Entropy Compensation and Born Radii). *Phys. Chem. Chem. Phys.* **2000**, 2, 97-102.
117. Tawa, G. J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A., Calculation of the Aqueous Solvation Free Energy of the Proton. *J. Chem. Phys.* **1998**, 109, 4852-4863.
118. Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R., The Proton's Absolute Aqueous Enthalpy and Gibbs Free Energy of Solvation from Cluster-Ion Solvation Data. *J. Phys. Chem. A* **1998**, 102, 7787-7794.
119. Fawcett, W. R., Thermodynamic Parameters for the Solvation of Monatomic Ions in Water. *J. Phys. Chem. B* **1999**, 103, 11181-11185.
120. Zhan, C.-G.; Dixon, D. A., Absolute Hydration Free Energy of the Proton from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2001**, 105, 11534-11540.
121. Palascak, M. W.; Shields, G. C., Accurate Experimental Values for the Free Energies of Hydration of H⁺, OH⁻, and H₃O⁺. *J. Phys. Chem. A* **2004**, 108, 3692-3694.
122. Grossfield, A.; Ren, P.; Ponder, J. W., Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **2003**, 125, 15671-15682.
123. Marcus, Y., Ionic Radii in Aqueous Solutions. *Chem. Rev.* **1988**, 88, 1475-1498.
124. Persson, I.; Jalilehvand, F.; Sandström, M., Structure of the Solvated Thallium (I) Ion in Aqueous, Dimethyl Sulfoxide, N, N'-Dimethylpropyleneurea, and N, N'-Dimethylthioformamide Solution. *Inorg. Chem.* **2002**, 41, 192-197.
125. Shannon, R., Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1976**, 32, 751-767.
126. Vchirawongkwin, V.; Hofer, T. S.; Randolph, B. R.; Rode, B. M., Tl (I) - the Strongest Structure - Breaking Metal Ion in Water? A Quantum Mechanical/Molecular Mechanical Simulation Study. *J. Comput. Chem.* **2007**, 28, 1006-1016.
127. Burda, J. V.; Pavelka, M.; Šimánek, M., Theoretical Model of Copper Cu (I)/Cu (II) Hydration. Dft and Ab Initio Quantum Chemical Study. *J. Mol. Struct.: THEOCHEM* **2004**, 683, 183-193.

128. Detellier, C.; Laszlo, P., Role of Alkali Metal and Ammonium Cations in the Self-Assembly of the 5'-Guanosine Monophosphate Dianion. *J. Am. Chem. Soc.* **1980**, 102, 1135-1141.
129. Pauling, L., *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*. Cornell University Press: 1960; Vol. 18.
130. Jorgensen, W. L.; Gao, J., Monte Carlo Simulations of the Hydration of Ammonium and Carboxylate Ions. *J. Phys. Chem.* **1986**, 90, 2174-2182.
131. Brugé, F.; Bernasconi, M.; Parrinello, M., Ab Initio Simulation of Rotational Dynamics of Solvated Ammonium Ion in Water. *J. Am. Chem. Soc.* **1999**, 121, 10883-10888.
132. Meraj, G.; Chaudhari, A., Hydrogen Bonded Hydrated Hydronium and Zündel Ion Complexes. *J. Mol. Liq.* **2014**, 190, 1-5.
133. Sobolewski, A. L.; Domcke, W., Ab Initio Investigation of the Structure and Spectroscopy of Hydronium-Water Clusters. *J. Phys. Chem. A* **2002**, 106, 4158-4167.
134. Markus Blauth, C.; Pribil, A. B.; Randolph, B. R.; Rode, B. M.; Hofer, T. S., Structure and Dynamics of Hydrated Ag⁺: An Ab Initio Quantum Mechanical/Charge Field Simulation. *Chem. Phys. Lett.* **2010**, 500, 251-255.
135. Blumberger, J.; Bernasconi, L.; Tavernelli, I.; Vuilleumier, R.; Sprik, M., Electronic Structure and Solvation of Copper and Silver Ions: A Theoretical Picture of a Model Aqueous Redox Reaction. *J. Am. Chem. Soc.* **2004**, 126, 3928-3938.
136. Sham, T., Electronic Structure of Hydrated Ce⁴⁺ Ions in Solution: An X-Ray-Absorption Study. *Phys. Rev. B* **1989**, 40, 6045-6051.
137. Ankudinov, A.; Conradson, S.; de Leon, J. M.; Rehr, J., Relativistic Xanes Calculations of Pu Hydrates. *Phys. Rev. B* **1998**, 57, 7518-7525.
138. Moll, H.; Denecke, M.; Jalilehvand, F.; Sandström, M.; Grenthe, I., Structure of the Aqua Ions and Fluoride Complexes of Uranium (Iv) and Thorium (Iv) in Aqueous Solution an Exafs Study. *Inorg. Chem.* **1999**, 38, 1795-1799.
139. Hagfeldt, C.; Kessler, V.; Persson, I., Structure of the Hydrated, Hydrolysed and Solvated Zirconium (Iv) and Hafnium (Iv) Ions in Water and Aprotic Oxygen Donor Solvents. A Crystallographic, Exafs Spectroscopic and Large Angle X-Ray Scattering Study. *Dalton Trans.* **2004**, 2142-2151.
140. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, 26, 1668-1688.

141. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, 23, 327-341.
142. Miyamoto, S.; Kollman, P. A., Settle: An Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, 13, 952-962.
143. Li, P.; Song, L. F.; Merz Jr, K. M., Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theory Comput.* **2015**, 11, 1645-1657.
144. Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz Jr, K. M., Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, 9, 2733-2748.
145. Li, P.; Song, L. F.; Merz Jr, K. M., Parameterization of Highly Charged Metal Ions Using the 12-6-4 Lj-Type Nonbonded Model in Explicit Water. *J. Phys. Chem. B* **2015**, 119, 883-895.
146. Eastman, P.; Pande, V., Openmm: A Hardware-Independent Framework for Molecular Simulations. *Comput. Sci. Eng.* **2010**, 12, 34-39.
147. Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shukla, D., Openmm 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *J. Chem. Theory Comput.* **2012**, 9, 461-469.
148. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., H++: A Server for Estimating Pk_as and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, 33, W368-W371.
149. Stone, J. E.; Hardy, D. J.; Ufimtsev, I. S.; Schulten, K., Gpu-Accelerated Molecular Modeling Coming of Age. *J. Mol. Graphics Model.* **2010**, 29, 116-125.
150. Yang, J.; Wang, Y.; Chen, Y., Gpu Accelerated Molecular Dynamics Simulation of Thermal Conductivities. *J. Comput. Phys.* **2007**, 221, 799-804.
151. Stone, J. E.; Phillips, J. C.; Freddolino, P. L.; Hardy, D. J.; Trabuco, L. G.; Schulten, K., Accelerating Molecular Modeling Applications with Graphics Processors. *J. Comput. Chem.* **2007**, 28, 2618-2640.
152. Rodrigues, C. I.; Hardy, D. J.; Stone, J. E.; Schulten, K.; Hwu, W.-M. W., In *Proceedings of the 5th conference on Computing frontiers*; ACM: Ischia, Italy, 2008, pp 273-282.
153. Preis, T.; Virnau, P.; Paul, W.; Schneider, J. J., Gpu Accelerated Monte Carlo Simulation of the 2d and 3d Ising Model. *J. Comput. Phys.* **2009**, 228, 4468-4477.

154. Liu, W.; Schmidt, B.; Voss, G.; Müller-Wittig, W., Accelerating Molecular Dynamics Simulations Using Graphics Processing Units with Cuda. *Comput. Phys. Commun.* **2008**, 179, 634-641.
155. Liu, W.; Schmidt, B.; Voss, G.; Müller-Wittig, W. Molecular Dynamics Simulations on Commodity Gpus with Cuda
High Performance Computing – Hipc 2007. In Aluru, S.; Parashar, M.; Badrinath, R.; Prasanna, V., Eds.; Springer Berlin / Heidelberg: 2007; Vol. 4873, pp 185-196.
156. Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; Legrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande, V. S., Accelerating Molecular Dynamic Simulation on Graphics Processing Units. *J. Comput. Chem.* **2009**, 30, 864-872.
157. Netz, R. R.; Horinek, D., Progress in Modeling of Ion Effects at the Vapor/Water Interface. *Annu. Rev. Phys. Chem.* **2012**, 63, 401-418.
158. Jones, J. E., On the Determination of Molecular Fields. Ii. From the Equation of State of a Gas. *Proc. R. Soc. London, A* **1924**, 106, 463-477.
159. Radëtisig, A.; Smirnov, B., *Reference Data on Atoms, Molecules, and Ions*. Springer-Verlag (Berlin and New York): 1985.
160. Ramaniah, L. M.; Bernasconi, M.; Parrinello, M., Ab Initio Molecular-Dynamics Simulation of K⁺ Solvation in Water. *J. Chem. Phys.* **1999**, 111, 1587-1591.
161. Soper, A., Orientational Correlation Function for Molecular Liquids: The Case of Liquid Water. *J. Chem. Phys.* **1994**, 101, 6888-6901.
162. Ohtaki, H.; Radnai, T., Structure and Dynamics of Hydrated Ions. *Chem. Rev.* **1993**, 93, 1157-1204.
163. Kuznetsov, A.; Zueva, E.; Masliy, A.; Krishtalik, L., Redox Potential of the Rieske Iron-Sulfur Protein Quantum-Chemical and Electrostatic Study. *Biochim. Biophys. Acta* **2010**, 1797, 347-359.
164. Moučka, F.; Nezbeda, I.; Smith, W. R., Molecular Simulation of Aqueous Electrolytes: Water Chemical Potential Results and Gibbs-Duhem Equation Consistency Tests. *J. Chem. Phys.* **2013**, 139, 124505.
165. Rajamani, S.; Ghosh, T.; Garde, S., Size Dependent Ion Hydration, Its Asymmetry, and Convergence to Macroscopic Behavior. *J. Chem. Phys.* **2004**, 120, 4457-4466.
166. Mukhopadhyay, A.; Fenley, A. T.; Tolokh, I. S.; Onufriev, A. V., Charge Hydration Asymmetry: The Basic Principle and How to Use It to Test and Improve Water Models. *J. Phys. Chem. B* **2012**, 116, 9776-9783.

167. Mukhopadhyay, A.; Aguilar, B. H.; Tolokh, I. S.; Onufriev, A. V., Introducing Charge Hydration Asymmetry into the Generalized Born Model. *J. Chem. Theory Comput.* **2014**, 10, 1788-1794.
168. Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A., Towards Accurate Solvation Dynamics of Divalent Cations in Water Using the Polarizable Amoeba Force Field: From Energetics to Structure. *J. Chem. Phys.* **2006**, 125, 054511.
169. Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P., Simulation of Ca²⁺ and Mg²⁺ Solvation Using Polarizable Atomic Multipole Potential. *J. Phys. Chem. B* **2006**, 110, 18553-18559.
170. Dang, L. X., Development of Nonadditive Intermolecular Potentials Using Molecular Dynamics: Solvation of Li[Sup +] and F[Sup -] Ions in Polarizable Water. *J. Chem. Phys.* **1992**, 96, 6970-6977.
171. Sakharov, D. V.; Lim, C., Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **2005**, 127, 4921-4929.
172. Wu, R.; Lu, Z.; Cao, Z.; Zhang, Y., A Transferable Nonbonded Pairwise Force Field to Model Zinc Interactions in Metalloproteins. *J. Chem. Theory Comput* **2011**, 7, 433-443.
173. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A., A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The Resp Model. *J. Phys. Chem.* **1993**, 97, 10269-10280.
174. Li, P.; Merz Jr, K. M., Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J. Chem. Theory Comput.* **2014**, 10, 289-297.
175. Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C., Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-Chain Amide Orientation. *J. Mol. Biol.* **1999**, 285, 1735-1747.
176. Thompson, W. H.; Hynes, J. T., Frequency Shifts in the Hydrogen-Bonded Oh Stretch in Halide-Water Clusters. The Importance of Charge Transfer. *J. Am. Chem. Soc.* **2000**, 122, 6278-6286.
177. Lee, H. M.; Kim, D.; Kim, K. S., Structures, Spectra, and Electronic Properties of Halide-Water Pentamers and Hexamers, X-(H₂O)₅, 6 (X= F, Cl, Br, I): Ab Initio Study. *J. Chem. Phys.* **2002**, 116, 5509-5520.
178. Pavlov, M.; Siegbahn, P. E.; Sandström, M., Hydration of Beryllium, Magnesium, Calcium, and Zinc Ions Using Density Functional Theory. *J. Phys. Chem. A* **1998**, 102, 219-228.

179. Irving, H.; Williams, R., 637. The Stability of Transition-Metal Complexes. *J. Chem. Soc. (Resumed)* **1953**, 3192-3210.
180. Gorelsky, S. I.; Basumallick, L.; Vura-Weis, J.; Sarangi, R.; Hodgson, K. O.; Hedman, B.; Fujisawa, K.; Solomon, E. I., Spectroscopic and Dft Investigation of $[M\{Hb(3,5-Ipr2pz)3\}(Sc6f5)]$ ($M=Mn, Fe, Co, Ni, Cu$, and Zn) Model Complexes: Periodic Trends in Metal-Thiolate Bonding. *Inorg. Chem.* **2005**, 44, 4947-4960.
181. Tai, H.-C.; Lim, C., Computational Studies of the Coordination Stereochemistry, Bonding, and Metal Selectivity of Mercury. *J. Phys. Chem. A* **2006**, 110, 452-462.
182. Beattie, J. K.; Best, S. P.; Skelton, B. W.; White, A. H., Structural Studies on the Caesium Alums, $Csmiii[So4]2 \cdot 12H_2O$. *J. Chem. Soc., Dalton Trans.* **1981**, 2105-2111.
183. Bard, A. J.; Parsons, R.; Jordan, J., *Standard Potentials in Aqueous Solution*. CRC press: 1985; Vol. 6.
184. Moin, S. T.; Hofer, T. S.; Pribil, A. B.; Randolph, B. R.; Rode, B. M., A Quantum Mechanical Charge Field Molecular Dynamics Study of Fe^{2+} and Fe^{3+} Ions in Aqueous Solutions. *Inorg. Chem.* **2010**, 49, 5101-5106.
185. Hofer, T. S.; Randolph, B. R.; Rode, B. M., Al (Iii) Hydration Revisited. An Ab Initio Quantum Mechanical Charge Field Molecular Dynamics Study. *J. Phys. Chem. B* **2008**, 112, 11726-11733.
186. Canaval, L. R.; Sakwarathorn, T.; Rode, B. M.; Messner, C. B.; Lutz, O. M.; Bonn, G. n. K., Erbium (Iii) in Aqueous Solution: An Ab Initio Molecular Dynamics Study. *J. Phys. Chem. B* **2013**, 117, 15151-15156.
187. Jalilehvand, F.; Spångberg, D.; Lindqvist-Reis, P.; Hermansson, K.; Persson, I.; Sandström, M., Hydration of the Calcium Ion. An Exafs, Large-Angle X-Ray Scattering, and Molecular Dynamics Simulation Study. *J. Am. Chem. Soc.* **2000**, 123, 431-441.
188. Martin, B.; Richardson, F. S., Lanthanides as Probes for Calcium in Biological Systems. *Q. Rev. Biophys.* **1979**, 12, 181-209.
189. Cotton, S., *Lanthanide and Actinide Chemistry*. John Wiley & Sons: 2007; Vol. 27.
190. Habenschuss, A.; Spedding, F. H., The Coordination (Hydration) of Rare Earth Ions in Aqueous Chloride Solutions from X Ray Diffraction. I. $TbCl_3$, $DyCl_3$, $ErCl_3$, $TmCl_3$, and $LuCl_3$. *J. Chem. Phys.* **1979**, 70, 2797-2806.
191. Habenschuss, A.; Spedding, F. H., The Coordination (Hydration) of Rare Earth Ions in Aqueous Chloride Solutions from X - Ray Diffraction. Iii. $SmCl_3$, $EuCl_3$, and Series Behavior. *J. Chem. Phys.* **1980**, 73, 442-450.

192. Habenschuss, A.; Spedding, F. H., The Coordination (Hydration) of Rare Earth Ions in Aqueous Chloride Solutions from X - Ray Diffraction. Ii. LaCl_3 , PrCl_3 , and NdCl_3a). *J. Chem. Phys.* **1979**, 70, 3758-3763.
193. Yamaguchi, T.; Nomura, M.; Wakita, H.; Ohtaki, H., An Extended X - Ray Absorption Fine Structure Study of Aqueous Rare Earth Perchlorate Solutions in Liquid and Glassy States. *J. Chem. Phys.* **1988**, 89, 5153-5159.
194. Kanno, H.; Hiraishi, J., Raman Spectroscopic Evidence for a Discrete Change in Coordination Number of Rare Earth Aquo-Ions in the Middle of the Series. *Chem. Phys. Lett.* **1980**, 75, 553-556.
195. D'Angelo, P.; Zitolo, A.; Migliorati, V.; Chillemi, G.; Duvail, M.; Vitorge, P.; Abadie, S.; Spezia, R., Revised Ionic Radii of Lanthanoid (Iii) Ions in Aqueous Solution. *Inorg. Chem.* **2011**, 50, 4572-4579.
196. D'Angelo, P.; Zitolo, A.; Migliorati, V.; Persson, I., Analysis of the Detailed Configuration of Hydrated Lanthanoid (Iii) Ions in Aqueous Solution and Crystalline Salts by Using K- and L3-Edge Xanes Spectroscopy. *Chem. Eur. J.* **2010**, 16, 684-692.
197. Taylor, D. M., The Bioinorganic Chemistry of Actinides in Blood. *J. Alloys Compd.* **1998**, 271, 6-10.
198. Canaval, L. R.; Weiss, A. K.; Rode, B. M., Structure and Dynamics of the Th^{4+} -Ion in Aqueous Solution—an Ab Initio Qmcf-Md Study. *Comp. Theor. Chem.* **2013**, 1022, 94-102.
199. Kirm, M.; Krupa, J.; Makhov, V.; Negodin, E.; Zimmerer, G.; Gesland, J., $6d5f$ and $5f^2$ Configurations of U^{4+} Doped into LiYF_4 and YF_3 Crystals. *J. Lumin.* **2003**, 104, 85-92.
200. Frick, R. J.; Pribil, A. B.; Hofer, T. S.; Randolph, B. R.; Bhattacharjee, A.; Rode, B. M., Structure and Dynamics of the U^{4+} Ion in Aqueous Solution: An Ab Initio Quantum Mechanical Charge Field Molecular Dynamics Study. *Inorg. Chem.* **2009**, 48, 3993-4002.
201. Odoh, S. O.; Bylaska, E. J.; de Jong, W. A., Coordination and Hydrolysis of Plutonium Ions in Aqueous Solution Using Car-Parrinello Molecular Dynamics Free Energy Simulations. *J. Phys. Chem. A* **2013**, 117, 12256-12267.
202. Wilson, R. E.; Skanthakumar, S.; Burns, P. C.; Soderholm, L., Structure of the Homoleptic Thorium (Iv) Aqua Ion $[\text{Th}(\text{H}_2\text{O})_{10}]\text{Br}_4$. *Angew. Chem.* **2007**, 119, 8189-8191.
203. Stephens, P.; Jollie, D.; Warshel, A., Protein Control of Redox Potentials of Iron-Sulfur Proteins. *Chem. Rev.* **1996**, 96, 2491-2514.
204. Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H., Heme-Containing Oxygenases. *Chem. Rev.* **1996**, 96, 2841-2888.

205. Eisenberg, D. S.; Kauzmann, W., *The Structure and Properties of Water*. Oxford University Press: 1969.
206. Dyke, T. R.; Muentner, J. S., Electric Dipole Moments of Low J States of H₂O and D₂O. *J. Chem. Phys.* **1973**, 59, 3125-3127.
207. Gubskaya, A. V.; Kusalik, P. G., The Total Molecular Dipole Moment for Liquid Water. *J. Chem. Phys.* **2002**, 117, 5290-5302.
208. Silvestrelli, P. L.; Parrinello, M., Structural, Electronic, and Bonding Properties of Liquid Water from First Principles. *J. Chem. Phys.* **1999**, 111, 3572-3580.
209. Wu, J. C.; Piquemal, J.-P.; Chaudret, R.; Reinhardt, P.; Ren, P., Polarizable Molecular Dynamics Simulation of Zn (II) in Water Using the Amoeba Force Field. *J. Chem. Theory Comput.* **2010**, 6, 2059-2070.
210. Ren, P.; Ponder, J. W., Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, 107, 5933-5947.
211. Sangster, M.; Atwood, R., Interionic Potentials for Alkali Halides. II. Completely Crystal Independent Specification of Born-Mayer Potentials. *J. Phys. C* **1978**, 11, 1541-1555.
212. Wagman, D. D.; Evans, W. H.; Parker, V. B.; Schumm, R. H.; Halow, I. *The NBS Tables of Chemical Thermodynamic Properties. Selected Values for Inorganic and C1 and C2 Organic Substances in SI Units*; DTIC Document: 1982.
213. Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A., On the Structure and Thermodynamics of Solvated Monoatomic Ions Using a Hybrid Solvation Model. *J. Chem. Phys.* **1999**, 111, 10998-11014.
214. Mejias, J. A.; Lago, S., Calculation of the Absolute Hydration Enthalpy and Free Energy of H⁺ and OH⁻. *J. Chem. Phys.* **2000**, 113, 7306-7316.
215. Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B., Hydration Structure and Free Energy of Biomolecularly Specific Aqueous Dications, Including Zn²⁺ and First Transition Row Metals. *J. Am. Chem. Soc.* **2004**, 126, 1285-1289.
216. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, 26, 1781-1802.
217. Plimpton, S., Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, 117, 1-19.
218. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Model.* **2006**, 25, 247-260.

219. Li, P.; Merz Jr, K. M., Mcpb.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, 56, 599-604.
220. Nar, H.; Huber, R.; Messerschmidt, A.; Filippou, A. C.; Barth, M.; Jaquinod, M.; van de Kamp, M.; Canters, G. W., Characterization and Crystal Structure of Zinc Azurin, a by-Product of Heterologous Expression in Escherichia Coli of Pseudomonas Aeruginosa Copper Azurin. *Eur. J. Biochem.* **1992**, 205, 1123-1129.
221. Besler, B. H.; Merz Jr, K. M.; Kollman, P. A., Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, 11, 431-439.
222. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision C. 01*, Gaussian Inc.: Wallingford, CT, 2009.
223. Merz Jr, K. M., Carbon Dioxide Binding to Human Carbonic Anhydrase Ii. *J. Am. Chem. Soc.* **1991**, 113, 406-411.
224. Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz Jr, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *Amber 14*, University of California, San Francisco: 2014.
225. Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, 8, 1542-1555.
226. Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, 9, 3878-3888.
227. Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J., Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, 81, 3684-3690.

228. Demadis, K. D.; Dattelbaum, D. M.; Kober, E. M.; Concepcion, J. J.; Paul, J. J.; Meyer, T. J.; White, P. S., Vibrational and Structural Mapping of [Os(Bpy)₃]^{3+/2+} and [Os(Phen)₃]^{3+/2+}. *Inorg. Chim. Acta* **2007**, 360, 1143-1153.
229. Xue, Y.; Davis, A. V.; Balakrishnan, G.; Stasser, J. P.; Staehlin, B. M.; Focia, P.; Spiro, T. G.; Penner-Hahn, J. E.; O'Halloran, T. V., Cu (I) Recognition Via Cation- π and Methionine Interactions in Cusf. *Nat. Chem. Biol.* **2008**, 4, 107-109.
230. Chakravorty, D. K.; Li, P.; Tran, T. T.; Bayse, C. A.; Merz Jr, K. M., Metal Ion Capture Mechanism of a Copper Metallochaperone. *Biochemistry* **2016**, 55, 501-509.
231. Miyanaga, T.; Watanabe, I.; Ikeda, S., Amplitude in Exafs and Ligand Exchange Reaction of Hydrated 3d Transition Metal Complexes. *Chem. Lett.* **1988**, 17, 1073-1076.
232. Smirnov, P. R.; Trostin, V. N., Structural Parameters of Close Surroundings of Sr²⁺ and Ba²⁺ Ions in Aqueous Solutions of Their Salts. *Russ. J. Gen. Chem.* **2011**, 81, 282-289.
233. Robinson, R. A.; Stokes, R. H., *Electrolyte Solutions*. Courier Dover Publications: 2002.
234. Miller, K. J., Additivity Methods in Molecular Polarizability. *J. Am. Chem. Soc.* **1990**, 112, 8533-8542.
235. Applequist, J.; Carl, J. R.; Fung, K.-K., Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* **1972**, 94, 2952-2960.
236. Caminiti, R., Complex Formation and Phosphate-H₂O Interactions in a Concentrated Aqueous Mg(H₂PO₄)₂ Solution. *J. Mol. Liq.* **1984**, 28, 191-204.
237. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Jr., J. A. M.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision C.01 *Gaussian, Inc., Wallingford CT* **2010**.