Consequences of Overfitting the van der Waals Radii of Ions

Madelyn Smith,[†] Zhen Li,[‡] Luke Landry,[†] Kenneth M. Merz Jr.,^{‡§*} and Pengfei Li^{†*}

 [†]Department of Chemistry and Biochemistry, Loyola University Chicago, Chicago, Illinois 60660, United States
[‡]Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States
[§]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States

Emails of corresponding authors: <u>kmerz1@gmail.com</u> and <u>pli4@luc.edu</u>

Abstract

Atomic radii play important role in scientific research. The covalent radii of atoms, ionic radii of ions, and van der Waals (VDW) radii of neutral atoms can all be derived from crystal structures. However, the VDW radii of ions are a challenge to determine because the atomic distances in crystal structures were determined by a combination of the VDW interactions plus the electrostatic interactions, making it unclear how to define the VDW sphere of ions in such an environment. In the present study, we found that the VDW radii, which were determined based on the 0.0015 au electron density contour through a wavefunction analysis on atoms, have excellent agreement with the VDW radii of noble gas atoms determined experimentally. Based on this criterion, we calculated the VDW radii for various atomic ions across the periodic table, providing a systematic set of VDW radii of ions. Previously we have shown that the 12-6 Lennard-Jones nonbonded model could not simultaneously reproduce the hydration free energy (HFE) and ion-oxygen distance (IOD) for an atomic ion when its charge is +2 or higher. Because of this, we developed the 12-6-4 model to reproduce both properties at the same time by explicitly considering the ioninduced dipole interactions. However, recent studies showed it was possible to use the 12-6 model to simulate both properties simultaneously when an ion has the $R_{min}/2$ parameter (*i.e.*, the VDW radius) close to the Shannon ionic radius. In the present study, we show that such a "success" is due to an unphysical overfitting, as the VDW radius of an ion should be significantly larger than its ionic radius. Through molecular dynamics simulations, we show that such overfitting causes significant issues when transferring the parameters from ion-water systems to ion-ligand and metalloprotein systems. In comparison, the 12-6-4 model shows significant improvement in comparison to the overfitted 12-6 model, showing excellent transferability across different systems. In summary, although both the 12-6-4 and 12-6 models could reproduce HFE and IOD for an ion, the 12-6-4 model accomplishes such a task based on the consideration of the physics involved, while the 12-6 model accomplishes this through overfitting, which brings significant transferability issues when simulating other systems. Hence, we strongly recommend the use of the 12-6-4 model (or even more sophisticated models) instead of overfitted 12-6 models when simulating complex systems such as metalloproteins.

Introduction

Atomic radii are a fundamental concept in chemistry. Different atomic radii have been defined, including covalent radii,¹ ionic radii,^{2.4} metallic radii,⁵ and van der Waals (VDW) radii.⁶⁻⁷ In these radii, the atoms are usually represented by spheres. Although atoms can display anisotropy while forming a chemical bond, the sphere approximation is still the most widely used approximation due to its simplicity. The concept of atomic radii has been employed extensively in scientific research. They are usually used to determine the chemical bonds between atoms, the sizes of molecules, and the interaction strength between molecules.⁸⁻⁹ For example, covalent radii and metallic radii can be used to determine the covalent and metallic bonds and provide equilibrium bond lengths for molecular simulations involving such bonds. The VDW volume describes the volume occupied by a molecule. Likewise, the van der Waals radius, is the contact distance between atoms/molecules.¹⁰ VDW radii have been used for various applications, such as determining hydrogen-bonding,¹¹ calculating the hydration free energy,¹² as well as in crystal chemistry and stereochemistry.¹³ Furthermore, a reasonable set of VDW radii is quite important for one to get reasonable atomic distances when modeling molecular systems in computational chemistry. VDW radii are also needed in the ESP¹⁴ and CHELPG¹⁵ fitting schemes for choosing the special extent of the electrostatic potential in order to fit atomic partial charges. Because of this importance, we believe it is critical to have physically meaningful parameters for the VDW radii in molecular modeling.

The covalent radii are for atoms having covalent bonds with each other.¹ One can determine whether there is a covalent bond between two atoms by comparing their atomic distance with the sum of their covalent radii (usually within a certain tolerance). Metallic radii are defined similarly but for metals forming metallic bonds, which is intrinsically different from covalent bonds.² The ionic radii are defined for ions in ionic crystal structures. The sum of ionic radii are used to determine the atomic distance between the neighboring counterions in the crystal lattice.³ These three different radii are used for covalent, metallic, and ionic bonds. One can say that the covalent interaction determines the covalent bonds being termed as the bonded interactions in molecular modeling. However, the ionic radii are not determined by only the VDW or electrostatic interactions, but by the sum of them, which is termed as the nonbonded interactions in molecular modeling. In ionic crystal structures, the balance between VDW and electrostatic interactions is o important that without the VDW interactions, ions with opposite charges will collapse because of the attractive force. Similarly, without the electrostatic attraction, counterions will have much larger distances from each other.¹⁶

The same issue exists for the VDW radii of ions. Specifically, it is quite challenging to determine the VDW radii of ions based on crystal structures because the electrostatic and VDW interactions co-exist with one another. The nonbonded interactions have both components by nature, where their decomposition is difficult for systems having atoms with partial charges. Fortunately, systems that only have VDW interactions exist, which are the noble gas atoms. Scientists determined their radii experimentally,¹² and these radii are intrinsically VDW radii. However, for atoms other than noble gas atoms, experimental determination of their VDW radii is not straightforward and usually needs extra assumptions. Inconsistencies in assumptions and methods among different studies leads to a range VDW radii. For example, one may assume atoms are spheres,⁶⁻⁷ another may not

but include the anisotropic effect.¹⁷ Oxygen atoms could be used as a probe, while nitrogen or other elements may also be used.¹⁸

Besides experimental approaches, theoretical calculation is another way to determine the VDW radii of atoms and ions. For example, Stokes calculated the VDW radii for sixteen cations and anions which shared the same electronic structures as the noble gas atoms.¹² He obtained the VDW radii of these ions based on the VDW radii of noble gas atoms and the quantum scaling principle (OMSP), which states that for ions and atoms which are isoelectronic, any characteristic radius associated with a certain electron is inversely proportional to the effective nuclear charge.¹² In 1967. Bader et al. demonstrated that an outer contour of the electron density of 0.002 au could be used to define the van der Waals envelope.¹⁹ In 1977, Boyd determined the atomic sizes of H to Xe by empirically scaling the radii of the 0.0001 au contour.²⁰ In following work, rather than using a radii which was then scaled back, Bader et al. determined the atomic and molecular volumes based a 0.001 au and 0.002 au criteria directly.²¹ Afterwards, Rahm et al. proposed that 0.001 au is an appropriate density cutoff for determining VDW radii and they determined the VDW radii of atoms, monovalent charged cations, and monovalent anions of elements 1-96.²² Their predictions agreed well with the VDW radii of noble gas atoms that were determined by Vogt and Alvarez.²³ They noted that this method can complement experimental data, especially when experimental data is not available.

The 12-6 Lennard-Jones (LJ) nonbonded model is widely used to simulate ions. It consists of Coulomb and LJ terms representing the electrostatic and VDW interactions, respectively. Our previous study indicated that one could find numerous parameters in the 12-6 LJ model to reproduce the same physical property.²⁴ To solve this issue, we fitted a curve based on the VDW parameters of noble gas atoms to correlate the $R_{min}/2$ and ε values in the LJ equation.²⁴ By using this noble gas curve (NGC), we parameterized the LJ parameters for various ions.²⁴⁻²⁶ The parameters agreed well with the VDW radii calculated based on the QMSP method, and showed excellent transferability to other systems such as metalloproteins and salt solutions.²⁴⁻²⁶ In addition, we found that for atomic ions with charge of +2 or higher, one cannot determine a set of 12-6 LJ parameters that are in a physical range and can reproduce the hydration free energy (HFE) and ion-oxygen distance (IOD) values simultaneously.²⁴⁻²⁶ And we attributed this to overlooking the ion-induced dipole interaction in the 12-6 nonbonded model. Based on which we developed the 12-6-4 model to solve this issue.²⁷ The 12-6-4 model showed significant improvement over the 12-6 model: it could well reproduce both properties at the same time and showed excellent transferability to other systems.²⁵⁻²⁷

Recently, Zhang *et al.* designed a new set of VDW parameters for the 12-6 model (z12-6) to reproduce the HFE and IOD values at the same time.²⁸⁻³⁰ However, the $R_{min}/2$ parameters they obtained were quite close to the Shannon ionic radii, which is significantly smaller than the VDW radii. The sum of the Shannon ionic radii gives the distance between counter-ions in crystal structures, which are determined by not only the VDW interactions but also the strong electrostatic attractions between the counter-ions. Moreover, the ε values they obtained were quite large (*e.g.*, can be thousands of kcal/mol for highly charged metal ions), which is contraindicated given the nature of the dispersion interaction as well. Hence, although their parameters were able to simultaneously reproduce the HFE and IOD values for an ion with charge of +2 or higher, we

believe their parameters result from overfitting, and could cause significant issues in molecular modeling.

In the present study, we tested different schemes to derive the VDW radii of atoms and ions based on the electron density distributions. Our results suggested that atomic radii determined based on the 0.0015 au electron iso-surface gives an excellent agreement with the experimentally determined VDW radii for noble gas atoms. Based on this we calculated the VDW radii of various atoms and ions, and these results showed excellent agreement with our previously determined LJ parameters for atomic ions. However, these calculated VDW radii deviated significantly from the VDW parameter set from Zhang et al.²⁸⁻³⁰ In addition, molecular dynamics (MD) simulations show the Zhang *et al.* parameter set could introduce significant artifacts when simulating ion-ligand and metalloprotein systems. In comparison, our 12-6-4 model showed significantly better performance for simulating these systems. Hence, the Zhang *et al.* parameters should be used with significant care because of serious transferability issues.

In sum, through quantum calculations, this study provides a consistent set of VDW radii for ions across the periodic table for their most common oxidation states and different spin states, which has not been reported to the best of our knowledge. Moreover, our study also indicated that physical VDW radii are critical for obtaining meaningful results in MD simulations of ion-containing systems, while unphysical/overfitted parameters have limited transferability and should be used with caution.

Computational Methods

Quantum Mechanical Calculations In the present study, the quantum calculations were carried out using the Gaussian16 program (version C.01).³¹ Afterwards the VDW radii of atoms and ions were obtained based on the wavefunction analysis using the Multiwfn program (version 3.7).³² Following the previous work done by Rahm et al.,²² we assume atoms and ions are spheres. Specifically, the radius of a particle was computed based on the contour covered volume with $r = \sqrt[3]{3V/4\pi}$, which was defined as the region encompassed by a specific electron density cutoff.²² The Mento Carlo method was used for calculating the contour covered volume, which incorporated randomness for the calculations. In the Mento Carlo procedure, $N=100*2^9$ particles were randomly distributed in a box. The *k* factor was set to 1.7 to define the box. These are the recommended values from the manual. Further details of the procedure and these parameters can be found in the Multiwfn manual. Our tests indicated the uncertainty (one σ value of the Gaussian distribution) of the calculated radii is around or within 0.01 Å.

The 12-6 and 12-6-4 models The 12-6 LJ and 12-6-4 LJ nonbonded models for ions are employed in the present study. The potential equations between the metal ions and other particles in the 12-6 model (Eq. 1) and the 12-6-4 model (Eq. 2) are shown below:

$$U_{12-6}(r_{ij}) = \frac{e^2 Q_i Q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}$$

= $\frac{e^2 Q_i Q_j}{r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right]$
= $\frac{e^2 Q_i Q_j}{r_{ij}} + 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$ (1)

$$U_{12-6-4}(r_{ij}) = \frac{e^2 Q_i Q_j}{r_{ij}} + \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} - \frac{C_4^{ij}}{r_{ij}^4}$$
$$= \frac{e^2 Q_i Q_j}{r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right] - \frac{C_4^{ij}}{r_{ij}^4}$$
$$= \frac{e^2 Q_i Q_j}{r_{ij}} + 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] - \frac{C_4^{ij}}{r_{ij}^4}$$
(2)

Herein, the first term is the electrostatic term, where *e* is the proton charge. Q_i and Q_j represent the partial charges of the atoms *i* and *j*, with the partial charge of the atomic ion it's oxidation state. r_{ij} is the distance between atoms *i* and *j*. In terms of the VDW term, $R_{min,ij}$ is the distance where the VDW interaction between atoms *i* and *j* has an energy minimum. σ_{ij} is the distance where the VDW interaction between these two atoms equals zero. ε_{ij} is the well depth of the VDW interaction. For the VDW terms, the Lorentz-Berthelot combining rules (Eqs. 3 and 4) were used, which is consistent with the AMBER³³ and CHARMM³⁴ force fields, as well as our previous studies^{24-27, 35-37} and in the studies of Zhang et al.²⁸⁻³⁰

$$R_{min,ij} = \frac{R_{min,ii} + R_{min,jj}}{2} = R_{min,i} + R_{min,j}$$
(3)

$$\varepsilon_{ij} = \sqrt{\left(\varepsilon_i \times \varepsilon_j\right)} \tag{4}$$

Herein $R_{min,ii}$ and $R_{min,jj}$ are the VDW diameters of atoms *i* and *j*, while $R_{min,i}$ and $R_{min,j}$ are the VDW radii of atom *i* and *j*. The difference between the 12-6 and 12-6-4 models is that the 12-6-4 model has an extra r^4 term which represents the ion-induced dipole interactions.²⁵⁻²⁷ In the present study, the C_4 term between a divalent ion and the water oxygen in the OPC water model³⁸ (represented by C_4^{m-w}) was taken from our previous study.³⁵ Physically, C_4 is proportional to the polarizability of the particle which is interacting with the metal ion.²⁷ Hence, the C_4 terms between the ion and atoms which are not oxygen in water (denoted as *l*) are calculated using Eq. 5. Herein α_w is the polarizability of the water oxygen, while the polarizability of *l* is represented by α_l .

$$C_4^{m-l} = C_4^{m-w} \times \frac{\alpha_l}{\alpha_w} \tag{5}$$

If the system has multiple ions, the C_4 term between each two ions (denoted as *i* and *j*) is calculated by using Eq. 6. Herein α_i and α_j are the polarizabilities of ions *i* and *j*, respectively.

$$C_4^{i-j} = C_4^{i-w} \times \frac{\alpha_j}{\alpha_w} + C_4^{j-w} \times \frac{\alpha_i}{\alpha_w}$$
(6)

In the present study, the polarizabilities of atoms and ions are from the \$AMBERHOME/dat/leap/parm/lj_1264_pol.dat file in the AMBER 2021 software package.³⁹

Molecular Dynamics Simulations In terms of simulating the ion-imidazole and ion-acetate systems in the aqueous phase, the OPC water model³⁸ was used to model the solvent box. Partial charges of the imidazole and acetate are given in Figure S1 and Table S1, which are the same as in the previous study.⁴⁰ The AMBER ff14SB force field⁴¹ was used to model the protein, imidazole, and acetate, with the additional parameters for imidazole listed in the SI. In terms of the simulations of the metalloprotein systems, the PDB entry $1F9Z^{42}$ was used for modeling the Ni²⁺-containing glyoxalase Glx I and the PDB entry $2YRH^{43}$ was employed for modeling the zinc finger protein 473. For these simulations, the system was minimized, heated, equilibrated, and followed by a 100 ns production simulation. In these simulations, both the Zhang et al. 12-6 parameter set for Zn²⁺ and Ni²⁺ in conjunction with the OPC water model³⁰ and the Li/Merz 12-6-4 parameter set for Zn²⁺ and Ni²⁺ in conjunction with the OPC water model³⁵ were investigated.

Results and Discussions

In 2016, Rahm et al. determined the atomic and ionic radii of elements 1-96 based on the 0.001 au electron contour.²² In their work they determined the radii for atoms and the corresponding singly charged cations and anions. However, many ions have oxidation states larger than +1 and could have a multiplicity larger than 2. In the present study, following their work, we have determined the van der Waals (VDW) radii for various elements across the periodic table and their associated ions based on the electron density contour. Specifically, in our study, we found the 0.0015 au contour would be a better criterion for evaluating the VDW volumes/radii, and we have calculated the VDW radii for atoms and ions with different oxidation states and spin states based on this criterion. This work provides a systematic set of the VDW radii of atomic cations and anions.

To test the sensitivity of the calculated radii to the levels of theory we benchmarked the results by varying the method (Table S2) and the basis set (Table S3), respectively. In these benchmarks, we

used the electron density of 0.0015 au as the cutoff because it showed the best agreement when compared to the experimentally determined VDW radii of noble gas atoms (discussed below). In particular, we used the Hartree-Fock,⁴⁴ B3LYP,⁴⁵⁻⁴⁶ PBE0,⁴⁷⁻⁴⁸ MP2,⁴⁹ and MP4SDQ⁵⁰ methods along with the aug-cc-pVTZ basis set⁵¹⁻⁵² in our tests (Table S2A). Our results showed these methods provided similar radii when using the same basis set. Although the Hartree-Fock method does not include electron correlation, it showed similar results to other levels of theory, which agrees with previous work.²² Comparing to cations, the electron cloud of anions are more dispersed, causing their sizes might be more likely depend on the density functional. Because of this reason, we also benchmarked different density functionals for calculating the VDW radii of anions (Table S2B). Specifically, the B3LYP,⁴⁵⁻⁴⁶ PBE0,⁴⁷⁻⁴⁸ BLYP,^{46, 53} and PBE⁴⁸ functionals were tested in conjugation with the SDD (Stuttgart/Dresden effective core potential (ECP)) basis set.³¹ The results indicated that the VDW radii calculated in the present way are not particularly sensitive to the functional choice. This may be because we are dealing with atomic anions, which are very small systems. Finally, the B3LYP density functional was chosen because it has been one of the most widely used functional due to the excellent performance and relatively low computational cost. In addition, we benchmarked the results by B3LYP with various basis sets, ranging from the Pople basis sets to the correlation consistent basis sets and to the SDD basis set (Table S3). We also found that the results were relatively insensitive to the basis set chosen. Again, this may be because the investigated systems (atoms and atomic ions) are very small, hence the basis set effect is also small. Finally, we chose the B3LYP/SDD level of theory to carry out the calculations, because the SDD basis set³¹ covers almost the entire periodic table and estimates the realistic effect for heavy atoms (using ECPs on the atoms beyond Ar).

As mentioned above, both 0.001 au and 0.002 au have been used as the criteria previously to determine the VDW volumes of atoms and molecules. However, the criterion 0.0015 au has not been used before. Herein, we benchmarked all of these three criteria for calculating the VDW radii of noble gas atoms (He to Xe) and the results are shown in Table S4. We observe that the 0.0015 au set showed the smallest average unsigned errors (AUEs) towards the five of six data sets which were determined experimentally. In comparison, the 0.001 au provided significantly larger VDW radii, yielding AUEs consistently higher when compared to these experimental data sets. Meanwhile, 0.002 au showed significantly smaller VDW radii, also providing higher AUEs towards the experimental data set. The data set of Rahm et al.²² was obtained based on the 0.001 au criterion, yielding the VDW radii of He, Ne, Ar, Kr, Xe as 1.34, 1.56, 1.97, 2.12, and 2.32 Å, which are very close to the 0.001 au set we obtained herein (with an average unsigned difference of ~0.014 Å), although they used a different level of theory (PBE0/ANO-RCC). These results agree with our above calculations that the calculated VDW radii are insensitive to the level of theory one uses.

Using the 0.0015 au criteria, we have determined the VDW radii for atoms and atomic ions across the periodic table. Specifically, we have calculated the VDW radii for various atoms and their atomic ions with the most common oxidation states. These results are provided in Tables S5 and S6, with select VDW radii of ions being shown in Figure 1. Note that for some of the ions (e.g. Ga³⁺, Ge⁴⁺) the SDD basis set does not have any electrons described explicitly (i.e. all the electrons in these ions are described by the effective core potentials), hence we calculated their VDW radii based on the Def2SV basis set,⁵⁴⁻⁵⁵ and then scaled these radii based on a factor which was obtained



by comparing the calculated radii using the SDD and Def2SV basis sets for the neutral atoms of these ions (see Table S7).

Figure 1. Calculated VDW radii of select ions in the periodic table. Only the VDW radii of ions with the most common oxidation states are shown, with the oxidation state specified before the radius. For a certain oxidation state, multiple spin states can be available, herein only the radius of the lowest spin state for each oxidation state is shown. VDW radii of other species can be seen in Tables S5 and S6.

Comparison with the VDW radii calculated based on the quantum scaling principle (QMSP). Stokes calculated the VDW radii of various cations and anions in 1964.¹² He obtained the VDW radii of noble gas atoms based on different experiments, and then he estimated the VDW radii of cations and anions which are isoelectronic to noble gas atoms based on the QMSP, that is "within an isoelectronic sequence of ions or atoms, any characteristic radius r' associated with a given electron is inversely proportional to the effective atomic number (Z-S), where the screening parameter S is constant for the isoelectronic sequence."¹² Because Stokes did not discuss the origins of this principle in the 1964 paper, we hypothesize that its origin comes from the hydrogenic wavefunctions which has the form of $\psi \propto e^{-Z_{\text{eff}}r/na_0}$, where $Z_{\text{eff}} = Z - S$, and *n* is the principal quantum number. However, we note that this is just an approximation, where the Z_{eff} term is not only in the exponential term but also in the normalization factor (the 1s orbital is shown in Eq 7 as an example).

$$\psi(1s) = \frac{1}{\sqrt{\pi}} \left(\frac{Z_{\rm eff}}{a_0}\right)^{3/2} e^{-Z_{\rm eff}r/2a_0}$$
(7)

However, the QMSP method does not consider the effect of Z_{eff} in the normalization factor, which causes the QMSP values deviate from our calculated values (Table S8). Such a deviation is

minimal for the monovalent cations. However, this deviation is significantly larger for anions than cations, and it becomes larger when the ion charge increases.

Comparison with the Shannon ionic radii

In the studies of Zhang et al.,²⁸⁻³⁰ they noted that their $R_{min}/2$ parameters agreed well with the Shannon ionic radii.⁴ We have compared the calculated VDW radii in the present study with the Shannon ionic radii (Table S8). We can see that the calculated VDW radii are significantly larger than the ionic radii for the same species. This agrees with our discussion above that noted that the ionic radii are determined not only by the VDW interactions but also the electrostatic interactions; hence the ionic radii should be significantly smaller than the VDW radii that were determined only by the VDW interactions. This principle should be always valid except in the case of noble gas atoms, for which the ionic radii and VDW radii would agree with each other, because only VDW interactions are available between these atoms.

Comparison with the VDW radii of ions in different parameter sets

To make a direct comparison between z12-6 with the calculated VDW radii in the present study, we give these values in Table S9, along with the VDW radii from the Li/Merz parameter sets. Herein the parameters for the OPC water model were used as an example, while the parameters for other water models follow the same trends. By referring to the calculated VDW radii in the present study as the standard, the average unsigned errors (AUEs) of different parameter sets are shown in Figure 2. From Table S9 and Figure 2, we can see that the z12-6 parameters are significantly smaller than the calculated VDW radii, with the AUE increasing along with the charge on the ions. This means the higher charge of the ion, the larger the error of the VDW radii used by z12-6 parameter sets. For the tetravalent metal ions, the AUE can be as high as ~0.97 Å. In comparison, the Li/Merz parameter sets showed much smaller AUEs for these ions, with the IOD and 12-6-4 parameter sets providing the lowest AUEs. The above comparison indicates that reproduction of both the HFE and IOD values for an atomic ion at the same time, can lead to overfitting which does not ensure a high quality model. Such an observation is consistent with that in z12-6 parameter sets where the ε values increases along with the ion charge. For example, their parameter sets for the OPC water model yield ε values (*i.e.* the well depth or interaction strength) for Zn²⁺, Fe³⁺, and Hf⁴⁺ of 225.8396, 1655.6733, and 3388.4415 kcal/mol, respectively. These values are inconsistent with the well-known weakness of VDW interactions.



Figure 2. Average unsigned errors (AUEs) of the VDW radii in the Li/Merz parameter sets (HFE, IOD, CM, 12-6-4) and the Zhang *et al.* parameter set for atomic ions with different charges in conjunction with the OPC water model. These AUEs were calculated using the calculated VDW radii in the present study as the reference.

The transferability issue

In both of the Li/Merz parameter sets and the Zhang *et al.* parameter set, the VDW parameters follow the Lorentz-Berthelot combining rules. According to these combining rules, the equilibrium distance of the VDW interaction is equal to the sum of the VDW radii of the two interacting particles: $R_{min,ij} = R_{min,i} + R_{min,j}$, which is based on the hard-sphere approximation, while the well depth of the VDW interaction is equal to the geometric average of the ε values of the two interacting particles $\varepsilon_{ij} = \sqrt{\varepsilon_i \times \varepsilon_j}$. Because Zhang *et al.*'s parameters are unbalanced (*i.e.*, $R_{min}/2$ is too small while ε is too large), we expect them to have transferability issues from water to more complex systems.

To characterize transferability, we use Zn^{2+} and Hf^{4+} as examples, where we compared the z12-6 parameter set with our 12-6-4 parameter set for generating the VDW parameters for different interaction pairs between metal ions and ligating atoms (Table S10). These ligating atoms are the oxygen atom from water, the carboxylate oxygen from ASP/GLU, the δ/ϵ nitrogen from HID/HIE, and sulfur from CYM, which represent the most common coordinating atoms in metalloproteins. For both Zn^{2+} and Hf^{4+} , We found that the z12-6 parameters provide significantly larger A_{ij} and B_{ij} value than the Li/Merz 12-6-4 parameter set, which implies these parameters would not have the same transferability to mixed systems as the Li/Merz parameters.

Transferability to ion-ligand systems in aqueous phase

In order to investigate the transferability of the z12-6 parameters for modeling ion-ligand systems, we calculated the binding free energy between Zn^{2+} and imidazole in a water box. For comparison, we also calculated the binding free energy between Zn^{2+} and imidazole in the aqueous phase based on the 12-6-4 parameter set for Zn^{2+} (the default 12-6-4). Specifically, we performed potential of

mean force (PMF) calculations between Zn^{2+} and the ligating nitrogen atom of imidazole, based on which we determined the binding free energy between Zn^{2+} and imidazole in the aqueous phase. The simulated results are shown in Table 1, together with the experimental value. We can see that the z12-6 parameters significantly underestimated the binding affinity between Zn^{2+} and imidazole by ~20 kcal/mol. In comparison, although the 12-6-4 parameter set also underestimate this binding free energy, the underestimation is ~7 kcal/mol, which is only one third of the error of the z12-6. Furthermore, in a previous study, it was shown that the Zn^{2+} -imidazole binding affinity can be accurately reproduced through optimizing the C₄ term in the 12-6-4 model.⁴⁰ Hence, because the z12-6 parameters were overfitted, with R_{min}/2 is too small and ε too large, it has limited transferability to describe Zn²⁺-imidazole interactions.

Parameter Set	Imidazole	Acetate	Acetate bidentate
	(kcal/mol)	monodentate	(kcal/mol)
		(kcal/mol)	
Default 12-6-4	3.14	1.44	3.03
Zhang et al.	17.73	-11.60	-14.48
Experimental values	-3.7^{a}	-1.2^{b}	> -1.2°

Table 1. List of binding free energies between Zn^{2+} and imidazole and acetate in aqueous phase.

^aBased on the logK of 2.663 from Ref ⁵⁶.

^bBased on the logK of 0.91 from ref⁵⁷

 $^c \rm This value was determined based on that monodentate is preferred for <math display="inline">\rm Zn^{2+}$ with acetate in aqueous solution 58

In order to further investigate the performance of the 12-6 model for describing the divalent ionimidazole interactions in aqueous phase, we performed a parameter scan. Specifically, to keep the $R_{min}/2$ and ε values in physical ranges, we generated the ε parameter based on the $R_{min}/2$ value by using the NGC. Finally, we plotted our results in Figure 3, which illustrates a curve of the $R_{min}/2$ values and the simulated binding affinities. Figure 3 indicates that as the VDW radius increases, the divalent ion-imidazole binding affinity would decrease. However, when $R_{min}/2$ is in the physically meaningful range, it is impossible to reproduce the binding affinity between Zn^{2+} and imidazole based on the current protocol (HID charges on imidazole and the OPC water model), indicating the limitations of the 12-6 model to simulate ion-ligand interactions in aqueous phase. Again, in comparison, it was shown that the Zn^{2+} -imidazole binding affinity can be accurately reproduced through optimizing the C₄ term in the 12-6-4 model.⁴⁰



Figure 3. Parameter scan results of the binding affinity between a divalent metal ion and imidazole in a OPC water box using the 12-6 model.

Acetate is another common ligand for metal ions in metalloproteins. Herein, we also calculated the binding free energy between Zn^{2+} and acetate by using the z12-6 parameter set and the 12-6-4 model and the results are shown in Table 1. Unlike the Zn^{2+} -imidazole interaction, the z12-6 parameter set significantly overestimates the Zn^{2+} -acetate interaction (>10 kcal/mol), incorrectly predicting that the bidentate mode is more favorable than the monodentate one and overestimated the coordination number for Zn^{2+} in the bidentate mode. All of these results indicate that there is a significant overestimation of the Zn^{2+} -carboxylate interactions when using the z12-6 parameter set. In comparison, the 12-6-4 parameters correctly predicted the preference of the monodentate mode over the bidentate mode, showed much better agreement in terms of the binding affinity of the most stable binding mode (wrong sign, but off only by ~2.6 kcal/mol). Furthermore, a previous study has shown that the 12-6-4 model can model the chelate effect after optimizing the C4 parameters.⁵⁹

Moreover, we also performed the parameter scan for a divalent metal ion with acetate. In this scan, the OPC water model was used as well, and the acetate partial charges are shown in Figure S1. Again, the NGC was used to restrain the relationship between $R_{min}/2$ and ε . The scan results are shown in Figure 4. Unlike Zn²⁺-imidazole, we can see that it is possible to adjust the $R_{min}/2$ parameter to reproduce the experimental value, which is -1.2 kcal/mol for the Zn²⁺-acetate binding affinity in aqueous phase.⁵⁷ However, such a $R_{min}/2$ value (1.9 Å) is significantly larger than the VDW radius of Zn²⁺ (calculated as 1.456 Å in the present study). In addition, while reproducing the binding free energy with acetate, the bidentate mode is slightly more favored, which is in contrast to the experimental result where the monodentate form is more favored by Zn^{2+, 58} Again, in comparison, we showed that the 12-6-4 model can reproduce the coordination mode between a divalent metal ion and acetate through refinement of the C₄ parameter.⁴⁰



Figure 4. Parameter scan results of the binding affinity between a divalent metal ion and acetate in a OPC water box. Herein three curves are shown, one is for the monodentate form, one is for the bidentate form, while the third one is for the most stable form. The binding affinities for different binding modes were obtained based on the PMF plot between Zn^{2+} and the carboxylate carbon of the acetate in a water box.

This analysis again shows that the z12-6 model has serious transferability issues when simulating ion-ligand systems. Specifically, our simulations indicated that the z12-6 parameters underestimate the Zn^{2+} -imidazole interaction but overestimates the Zn^{2+} -acetate interaction, while the 12-6-4 parameters showed significant improvement for describing these interactions.

Transferability to metalloprotein systems

To further test the transferability of the z12-6 parameters to complex systems, we performed MD simulations for Ni²⁺-containing glyoxalase Glx I (PDB ID: 1F9Z)⁴² and a zinc finger domain from the zinc finger protein 473 (PDB ID: 2YRH) using z12-6 and the 12-6-4 parameter set. Again, the OPC water model was used to simulate the solvent in these simulations.

Figure 5 shows the RMSD values of two Ni²⁺ binding sites of Glx I (with site 1 consisting of HID5, GLU56, HID200, GLU248 and two water molecules, and site 2 consisting of HID74, GLU122, HID131, GLU182 and two water molecules) and the RMSD values of the protein backbone N, C_{α} , and C atoms. By comparing the results based on the z12-6 parameter set and the 12-6-4 parameter set, we found the latter successfully maintained the coordination modes of both metal sites during the simulations, not unlike what we observed previously for an Fe³⁺ containing metalloprotein.³⁶ In contrast, the z12-6 parameters could not maintain the structure of metal site 1, showing a RMSD of ~1.4 Å after 80 ns. To better understand this, a distance analysis was carried out. We found the Ni²⁺-HID5@NE2 distance increased to ~4.5 Å after 80 ns (Figure 6), losing this coordination bond. To better illustrate these structural changes, we depict the last snapshot of site 1 in glyoxalase Glx I from the MD simulations in Figure 7. We can see that the z12-6 parameter set loses the two HID residues complexed to Ni²⁺ while both Glu56 and Glu248 are bidentate to Ni²⁺. This is entirely consistent with our PMF simulations which showed that the z12-6 parameter set underestimates the Zn²⁺-imidazole interaction but overestimate the Zn²⁺-acetate interaction and prefers bidentate coordination over monodentate (although the previous PMF calculations were for Zn²⁺, we would

expect to have similar results for Ni^{2+}). In comparison, the 12-6-4 parameter set successfully maintained the metal coordination mode with two monodentate GLU residues, two HID residues and two water molecules over the MD simulations (Figures 5, 6, and 7).



Figure 5. RMSD values of the metal site residues (Ni²⁺ plus the six ligating residues) in the two Ni²⁺ binding sites and the protein backbone N, C_{α} , C atoms of the glyoxalase Glx I system. The left panel is for the z12-6 parameter set, while the right panel is for the 12-6-4 parameter set.



Figure 6. Distance between HID5@NE2 and Ni²⁺ or between HID200@NE2 and Ni²⁺ in the first metal binding site of Glx I. The left panel is for the z12-6 parameter set, while the right panel is for the 12-6-4 parameter set. The distorted binding pocket leads to enlarged Ni²⁺-N distances when using the z12-6 parameter set.



Figure 7. (Left) The crystal structure of the Ni²⁺ binding site I, along with the last snapshot of this metal site from the MD simulations using (middle) the Zhang et al.'s parameter set or (right) the 12-6-4 parameter set.

Figure S2 shows the RMSD values of the zinc binding site (CYM13, CYM16, HID29, HID33) and of the protein backbone during the MD simulations using the z12-6 parameter set and the 12-6-4 parameter set. The 12-6-4 parameter set showed lower RMSD values when compared to the z12-6 parameter set, with the latter showing an increasing RMSD value after ~10 ns. A distance analysis was performed between Zn^{2+} and HID33@NE2 (Figure S3). We found the fluctuation of this Zn-N distance is significantly larger for the z12-6 parameter set than that of the 12-6-4 parameter set. Figure S4 shows the last snapshot of the metal site for each simulation. We observe that the z12-6 parameter set yielded a five-coordinate structure, with an additional water coming in to bind to Zn^{2+} . This agrees with our PMF simulations the z12-6 parameter set seriously underestimates the Zn^{2+} -imidazole interactions, which could cause water to out-compete the histidine residues in the metal coordination. In contrast, the 12-6-4 parameter set successfully maintained the zinc binding mode formed by two CYM and two HID residues over the timescale of the MD simulations (Figures S2-S4), indicating its better transferability over the z12-6 parameter set.

Overall, our simulations indicate, that even though the z12-6 parameter set reproduces both the experimental HFE and IOD for an ion, it has serious transferability issues caused by overfitting of the model into non-physical parameter space. Specifically, it has unbalanced ion-water and ion-ligand interactions and would likely yield distorted structures when simulating metalloprotein systems. Moreover, it is difficult for the 12-6 model, in general, to balance the ion-water and ion-ligand interactions when using parameters in a more physical range. In comparison, the 12-6-4 model, which is based on physical consideration, has shown excellent performance in simulating ions in aqueous phase and well-behaved transferability to modeling metalloproteins. Previous work also showed with further refinement, the 12-6-4 model can be systematically improved.^{40, 59} Therefore, we would recommend the use of the 12-6-4 model instead of the 12-6 model to model metalloproteins as the former would give a more physical and accurate description along with better transferability.

Conclusions

Atomic radii play fundamental roles in scientific research. There are different atomic radii available: covalent radii, ionic radii, and VDW radii, etc. Although crystal structures have been used to derive the covalent radii and ionic radii of elements as well as the VDW radii of noble gas atoms, it is not straightforward to derive the VDW radii of ions based on the crystal structures. This is because the electrostatic interaction and VDW interaction together determine the atomic distances in such a case, while decomposing them is not straightforward and usually involves assumptions. Determination of the VDW volume of an atom or a molecule based on the electron density contour at a certain cut-off has been used in previous studies. In the present study, we found that using a 0.0015 au electron density cut-off, one can obtain the VDW radii of noble gas atoms which have better agreement with experimentally determined values than using a 0.001 au or 0.002 au cut-off. Based on this criterion, we have calculated the VDW radii of various atomic ions across the periodic table, thereby providing a consistent set of VDW radii of ions.

The 12-6 Lennard-Jones equation has been widely used to describe the VDW interactions in force fields. In our previous studies, we showed that the 12-6 nonbonded model of ions could not reproduce both the HFE and IOD values of an ion simultaneously when its VDW parameters are in a physical range and its charge is +2 or higher.²⁴⁻²⁶ To solve this issue, we developed the 12-6-4 model by adding an extra term to represent the ion-induced dipole interaction²⁷ and we showed that the 12-6-4 model could reproduce both the HFE and IOD at the same time, along with excellent transferability to complex systems such as metalloproteins.²⁵⁻²⁷ However, recent studies from Zhang et al. showed that one could simultaneously reproduce the HFE and IOD for an ion with charge of +2 or higher by using the 12-6 model²⁸⁻³⁰ and the VDW radii obtained by them were close to the Shannon ionic radii.⁴ Herein, we showed that our VDW parameters have much better agreement with the QM calculated VDW radii. In contrast, Zhang *et al.*'s parameter set were found to be overfitted and deviated significantly from these physically meaningful values.

In addition, we performed MD simulations based on the z12-6 parameter set and our 12-6-4 parameter set. Our results indicated the z12-6 parameter set showed large deviations when simulating the ion-imidazole and ion-acetate binding affinities. In comparison, the 12-6-4 parameter set showed significantly improved performance. Moreover, the 12-6-4 parameter set retains the metal site structures when simulating the Ni²⁺-containing glyoxalase Glx I and the zinc finger protein 473. However, the z12-6 parameter set provided distorted structures for these metal sites. Furthermore, we found that in general it is difficult for the 12-6 model to describe ion-ligand interactions in the aqueous phase when using physically meaningful parameters. In comparison, the 12-6-4 model could solve this issue after parameter optimization.^{40, 59} Hence, our recommendation is the 12-6-4 model is preferred for highly-charged ions. Indeed, this model gives better answers for the right reason. In contrast, for these systems the 12-6 or z12-6 models should be used with caution when simulating complex systems such as metalloproteins because of the parameter transferability issue.

In summary, this study provides a consistent set of VDW radii for ions, and showed that it is important to have the VDW radii of ions in a physically meaningful range when developing and using force field models for highly charged ions in MD simulations.

Acknowledgement

This work was partially supported by the start-up funds from Loyola University Chicago (to P. L.) and partially supported by the grants from the National Institutes of Health (to K. M. M., Grant Numbers GM044974 and GM066859). We thank Dr. Lin Song for providing the script for the potential-mean-force calculations. P. L., M. S., and L. L. acknowledge the computational support from the Extreme Science and Engineering Discovery Environment (XSEDE).⁶⁰ XSEDE is funded by the National Science Foundation (NSF) with Grant Number ACI-1548562. Specifically, this work used the computational resources on Expanse at the San Diego Supercomputer Center (Allocation Numbers TG-CHE200102 and TG-BIO210105). Z. L. and K. M. M. thank the computational support from the High Performance Computing Center (HPCC) in the Institute of Cyber-Enabled Research (iCER) at Michigan state university (MSU).

Supporting Information

The supporting information is available free of charge at the ACS website: force field parameters; calculated van der Waals radii for atoms and ions; simulation results of the zinc finger system.

ORCIDs:

Madelyn Smith: 0000-0003-2090-8350 Zhen Li: 0000-0001-7947-3887 Luke Landry: 0009-0004-9289-4714 Kenneth M. Merz Jr: 0000-0001-9139-5893 Pengfei Li: 0000-0002-2572-5935

Notes

The authors declare no competing financial interest.

References

- 1. Cordero, B.; Gómez, V.; Platero-Prats, A. E.; Revés, M.; Echeverría, J.; Cremades, E.; Barragán, F.; Alvarez, S., Covalent radii revisited. *Dalton Trans.* **2008**, 2832-2838.
- 2. Pauling, L., The Principles Determining the Structure of Complex Ionic Crystals. J. Am. Chem. Soc. 1929, 51, 1010-1026.
- 3. Johnson, O., Ionic radii for spherical potential ions. Inorg. Chem. 1973, 12, 780-785.
- 4. Shannon, R. D., Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. Sect. A: Found. Crystallogr.* **1976**, *A32*, 751-767.
- 5. Goldschmidt, V., Über atomabstände in metallen. Z. Phys. Chem. 1928, 133, 397-419.
- 6. Bondi, A., van der Waals Volumes and Radii. J. Phys. Chem. 1964, 68, 441-451.
- 7. Bondi, A., van der Waals Volumes and Radii of Metals in Covalent Compounds. J. Phys. Chem. 1966, 70, 3006-3007.
- 8. Pauling, L., *The Nature of the Chemical Bond and the Structure of Molecules and Crystals*. Cornell University Press: 1941.
- 9. Dickinson, S. K., *Ionic, covalent, and metallic radii of the chemical elements*. Air Force Cambridge Research Laboratories, Office of Aerospace Research: 1970.
- 10. Allinger, N., Calculation of molecular structure and energy by force-field methods. *Adv. Phys. Org. Chem.* **1976**, *13*, 1-82.
- 11. Klein, R. A., Modified van der Waals atomic radii for hydrogen bonding based on electron density topology. *Chem. Phys. Lett.* **2006**, *425*, 128-133.
- 12. Stokes, R. H., 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.
- 13. Zefirov, Y. V.; Zorky, P. M., New applications of van der Waals radii in chemistry. *Russ. Chem. Rev.* **1995**, *64*, 415-428.
- 14. Singh, U. C.; Kollman, P. A., An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129-145.
- 15. Breneman, C. M.; Wiberg, K. B., Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **1990**, *11*, 361-373.
- 16. Kékicheff, P.; Spalla, O., Long-Range Electrostatic Attraction between Similar, Charge-Neutral Walls. *Phys. Rev. Lett.* **1995**, *75*, 1851-1854.
- 17. Batsanov, S. S., Van der Waals Radii of Elements. Inorg. Mater. 2001, 37, 871-885.
- 18. Alvarez, S., A cartography of the van der Waals territories. Dalton Trans. 2013, 42, 8617-8636.
- 19. Bader, R. F. W.; Henneker, W. H.; Cade, P. E., Molecular Charge Distributions and Chemical Binding. J. Chem. Phys. 1967, 46, 3341-3363.

- 20. Boyd, R. J., The relative sizes of atoms. J. Phys. B At. Mol. Opt. Phys. 1977, 10, 2283-2291.
- 21. Bader, R. F. W.; Carroll, M. T.; Cheeseman, J. R.; Chang, C., Properties of atoms in molecules: atomic volumes. J. Am. Chem. Soc. 1987, 109, 7968-7979.
- 22. Rahm, M.; Hoffmann, R.; Ashcroft, N. W., Atomic and Ionic Radii of Elements 1 –96. *Eur. J. Chem.* **2016**, *22*, 14625-14632.
- 23. Vogt, J.; Alvarez, S., van der Waals Radii of Noble Gases. Inorg. Chem. 2014, 53, 9260-9266.
- 24. Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, 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.
- 25. Li, P.; Song, L. F.; Merz, 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.
- 26. Li, P.; Song, L. F.; Merz, K. M., Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. J. Chem. Theory Comput. 2015, 11, 1645-1657.
- 27. Li, P.; Merz, K., Taking into account the ion-induced dipole interaction in the nonbonded model of ions. J. Chem. Theory Comput. 2014, 10, 289-297.
- 28. Qiu, Y.; Jiang, Y.; Zhang, Y.; Zhang, H., Rational Design of Nonbonded Point Charge Models for Monovalent Ions with Lennard-Jones 12–6 Potential. J. Phys. Chem. B 2021, 125, 13502-13518.
- 29. Zhang, Y.; Jiang, Y.; Qiu, Y.; Zhang, H., Rational Design of Nonbonded Point Charge Models for Highly Charged Metal Cations with Lennard-Jones 12-6 Potential. *J. Chem. Inf. Model.* **2021**, *61*, 4613-4629.
- 30. Zhang, Y.; Jiang, Y.; Peng, J.; Zhang, H., Rational Design of Nonbonded Point Charge Models for Divalent Metal Cations with Lennard-Jones 12-6 Potential. J. Chem. Inf. Model. 2021, 61, 4031-4044.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
- 32. Lu, T.; Chen, F., Multiwfn: A multifunctional wavefunction analyzer. *Journal of Computational Chemistry* **2012**, *33*, 580-592.
- 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.
- Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-Mccarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* 1998, *102*, 3586-3616.
- 35. Li, Z.; Song, L. F.; Li, P.; Merz, K. M., Systematic Parametrization of Divalent Metal Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. J. Chem. Theory Comput. 2020, 16, 4429-4442.
- Li, Z.; Song, L. F.; Li, P.; Merz, K. M., Parametrization of Trivalent and Tetravalent Metal Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. *Journal of Chemical Theory and Computation* 2021, 17, 2342-2354.
- 37. Sengupta, A.; Li, Z.; Song, L. F.; Li, P.; Merz, K. M., Parameterization of Monovalent Ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB Water Models. J. Chem. Inf. Model. 2021, 61, 869-880.

- 38. Izadi, S.; Anandakrishnan, R.; Onufriev, A. V., Building Water Models: A Different Approach. J. Phys. Chem. Lett. 2014, 5, 3863-3871.
- D.A. Case, H. M. A., K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K., Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman, Amber 2021. 2021.
- 40. Song, L. F.; Sengupta, A.; Merz, K. M., Thermodynamics of Transition Metal Ion Binding to Proteins. J. Am. Chem. Soc. 2020, 142, 6365-6374.
- Maier, J. A. M., C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 2015, 11, 3696-3713.
- He, M. M.; Clugston, S. L.; Honek, J. F.; Matthews, B. W., Determination of the Structure of Escherichia coli Glyoxalase I Suggests a Structural Basis for Differential Metal Activation. *Biochemistry* 2000, 39, 8719-8727.
- 43. Qin, X. R., Hayashi, F., Yokoyama, S., RIKEN Structural Genomics/Proteomics Initiative (RSGI), Solution structure of the C2H2-type zinc finger domain (699-729) from zinc finger protein 473. **2007**.
- 44. Slater, J. C., The Self Consistent Field and the Structure of Atoms. Phys. Rev. 1928, 32, 339-348.
- 45. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys. 1993, 98, 5648-5652.
- 46. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B. Condens. Matter.* **1988**, *37*, 785-789.
- 47. Adamo, C.; Barone, V., Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158-6170.
- 48. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865-3868.
- 49. M. J. Frisch, M. H.-G., and J. A. Pople, Direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166*, 275-80.
- 50. Trucks, G. W.; Watts, J. D.; Salter, E. A.; Bartlett, R. J., Analytical MBPT(4) gradients. *Chem. Phys. Lett.* **1988**, *153*, 490-495.
- 51. Kendall, R. A.; Dunning, T. H.; Harrison, R. J., Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. J. Chem. Phys. 1992, 96, 6796-6806.
- 52. Woon, D. E.; Dunning, T. H., Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. J. Chem. Phys. **1993**, *98*, 1358-1371.
- 53. Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098-3100.
- 54. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.
- 55. Weigend, F., Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057-1065.
- 56. Sjöberg, S., Critical evaluation of stability constants of metal-imidazole and metal-histamine systems (Technical Report). *Pure Appl. Chem.* **1997**, *69*, 1549-1570.
- 57. Bunting, J. W.; Thong, K. M., Stability constants for some 1:1 metal–carboxylate complexes. *Can. J. Chem.* **1970**, *48*, 1654-1656.
- 58. Tackett, J. E., FT-IR Characterization of Metal Acetates in Aqueous Solution. *Appl. Spectmsc.* **1989**, 43, 483-489.

- 59. Sengupta, A.; Seitz, A.; Merz, K. M., Simulating the Chelate Effect. J. Am. Chem. Soc. 2018, 140, 15166-15169.
- 60. Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; Roskies, R.; Scott, J. R.; Wilkins-Diehr, N., XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.* **2014**, *16*, 62-74.

TOC Graphic

