

Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications

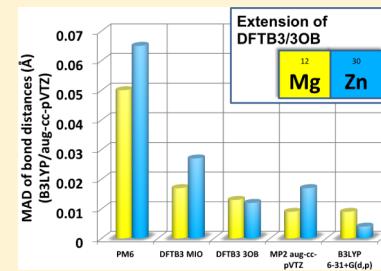
Xiya Lu,[†] Michael Gaus,[†] Marcus Elstner,[‡] and Qiang Cui*,[†]

[†]Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin—Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

[‡]Institute of Physical Chemistry, Karlsruhe Institute of Technology, Kaiserstr. 12, 76131 Karlsruhe, Germany

S Supporting Information

ABSTRACT: We report the parametrization of the approximate density functional theory, DFTB3, for magnesium and zinc for chemical and biological applications. The parametrization strategy follows that established in previous work that parametrized several key main group elements (O, N, C, H, P, and S). This 3OB set of parameters can thus be used to study many chemical and biochemical systems. The parameters are benchmarked using both gas-phase and condensed-phase systems. The gas-phase results are compared to DFT (mostly B3LYP), *ab initio* (MP2 and G3B3), and PM6, as well as to a previous DFTB parametrization (MIO). The results indicate that DFTB3/3OB is particularly successful at predicting structures, including rather complex dinuclear metalloenzyme active sites, while being semiquantitative (with a typical mean absolute deviation (MAD) of ~3–5 kcal/mol) for energetics. Single-point calculations with high-level quantum mechanics (QM) methods generally lead to very satisfying (a typical MAD of ~1 kcal/mol) energetic properties. DFTB3/MM simulations for solution and two enzyme systems also lead to encouraging structural and energetic properties in comparison to available experimental data. The remaining limitations of DFTB3, such as the treatment of interaction between metal ions and highly charged/polarizable ligands, are also discussed.



INTRODUCTION

Magnesium and zinc are two of the most common metal cofactors of numerous metalloenzymes and other biomolecules.^{1–3} The binding of Mg²⁺ to DNA and RNA helps to maintain their structures by stabilizing highly negatively charged phosphates. ATP, the main source of energy in cells, is bound to Mg²⁺ in its biologically active form.⁴ In addition, a large number of enzymes involved in the biochemistry of nucleic acids bind Mg²⁺ for catalytic activity; for example, Mg²⁺ is essential for DNA replication and phosphoryl transfers.⁵ Zinc is the second most abundant cation in the human body, so Zn²⁺ has important biological functions in numerous biomolecular systems^{6,7} such as zinc-finger class of transcription factors, signaling proteins, transport/storage proteins, as well as enzymes. The zinc ion contributes to catalysis in more than 300 enzymes and serves a structural role in many proteins and enzymes.^{8–10}

To describe the structural properties of Mg²⁺ and Zn²⁺ in solution and biomolecules, molecular mechanical (MM) force field parameters have been developed.^{11–15} Some of the advanced force field models such as SIBFA,¹⁶ AMOEBA,^{17,18} and Drude oscillator models¹⁹ treat electronic polarization and charge transfer²⁰ terms explicitly and have been demonstrated to describe potential energy surfaces in close agreement with quantum mechanics (QM) calculations.^{21,22} For chemical reactions that involve these metal ions, however, MM models are no longer appropriate. Even for structural properties, when

the binding mode of ligands is flexible,^{23–25} it is challenging to describe the metal site with a MM model.

An appropriate modeling approach in this context is the QM/MM framework, where the metal site of interest is treated with QM while the rest of the enzyme/solution is described by an empirical force field. For relatively rigid metal sites, *ab initio* QM and QM/MM methods based on either reaction path or free energy simulations have been demonstrated to be powerful for mechanistic analysis.^{26–30} The cost of these calculations for flexible systems, however, remains very high. For instance, one key feature of zinc coordination is its flexibility:^{23–25} a zinc ion can adopt multiple binding modes with the coordination number ranging from 4 to 6, underlining the importance of developing efficient QM(/MM) methods that complement *ab initio* based methods. The dynamical nature of the metal binding mode is particularly important to enzymes that feature solvent accessible active sites.^{31–33} For these applications, semiempirical QM-based QM/MM simulations can potentially strike the balance between computational accuracy and conformational sampling.

The traditional semiempirical QM methods based on the Neglect-of-Diatom-Differential-Overlap (NDDO) approxima-

Special Issue: William L. Jorgensen Festschrift

Received: July 1, 2014

Revised: August 25, 2014

Published: September 2, 2014

tion, such as MNDO,^{34,35} MNDO/d,³⁶ AM1,³⁷ and PM3,³⁸ however, give in general rather poor geometries for Mg/Zn-containing molecules. For example, metal–ligand lengths computed from AM1,^{39,40} PM3, MNDO, and MNDO/d⁴⁰ deviate by roughly 0.1 Å on average in comparison to B3LYP/6-311++G** for small Mg complexes.⁴¹ Recently developed AM1/d parameters⁴² for magnesium provide a clear improvement in accuracy compared to AM1 by incorporating d orbitals. According to a recent test,⁴³ the mean absolute deviations (MADs) of metal–ligand distances for AM1,⁴⁴ PM3, PM3-(ZnB),⁴⁵ MNDO/d,⁴⁶ and PM6⁴⁷ are more than 0.05 Å compared to CCSD(T) values for small zinc-containing compounds. The importance of a reliable treatment for zinc sites to the description of enzyme catalysis is illustrated by our recent study of alkaline phosphatase (AP),^{48,59} where we showed that previous AM1/d-PhoT/MM calculations^{50,51} likely led to incorrect transition states due to distortions of the bimetallic zinc motif.

The Self-Consistent-Charge Density-Functional Tight-Binding (SCC-DFTB) method provides a promising alternative to NDDO approaches.^{52–54} It is derived by expanding the DFT total energy up to second order around a reference density, which is usually taken to be the sum of atomic charge densities. The computational speed of SCC-DFTB is comparable to NDDO, and it has been rather extensively benchmarked for reaction energies, geometries, and vibrational frequencies for a large number of small molecules.^{55–61} The fact that SCC-DFTB is formulated in a DFT framework suggests that it is potentially more reliable for metal ions than NDDO-based methods, which are based on Hartree–Fock. Specifically for Mg⁴¹ and Zn,^{62,63} SCC-DFTB/(MM)^{64–66} has been successfully applied to a rather diverse set of biological and chemical applications that include these ions.^{48,49,67–74} More recently, DFTB has been extended to include third-order contributions, termed as DFTB3.⁷⁵ Compared to the second-order formulation, DFTB3 considers the charge dependence of the atomic Hubbard parameter and therefore provides an improved description of chemical properties such as proton affinity,^{53,75,76} which is important in many biological applications. With the most recent parametrization referred to as 3OB,^{77,78} the electronic properties and therefore atomization energies have been substantially improved; as a result, the DFTB3 approach is shown to have comparable performance as DFT with double- ζ -plus-polarization basis sets for many systems of biological and chemical interest, although certain limitations remain.^{78,79}

Considering the recent success of DFTB3/3OB, here we extend the 3OB parametrization of DFTB3 to magnesium and zinc, again focusing on chemical and biological applications. In the next section, we briefly summarize the DFTB3 methodology, the procedures for parametrization, and simulation details for benchmark calculations. We then present data from multiple levels of test sets in the gas phase, which are followed by additional benchmark calculations in both solution and enzymes. Finally, we summarize this work with several conclusions.

COMPUTATIONAL METHODS

Theory. The DFTB3 parametrization of magnesium and zinc follows the protocol outlined in the previous publications^{77,80} and extends the 3OB parameter set that so far includes elements C, H, N, O,⁷⁷ S, and P.⁷⁸ We first briefly review the formulation of DFTB3.^{54,75}

The total energy of DFTB3 is given by

$$\begin{aligned} E^{\text{DFTB3}} &= E^{\text{H}0} + E^\gamma + E^\Gamma + E^{\text{rep}} \\ &= \sum_{iab} \sum_{\mu \in a} \sum_{\nu \in b} n_i c_{\mu i} c_{\nu i} H_{\mu\nu}^0 + \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab} \\ &\quad + \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab} + \frac{1}{2} \sum_{ab} V_{ab}^{\text{rep}} \end{aligned} \quad (1)$$

Here $H_{\mu\nu}^0$ is the effective Kohn–Sham Hamiltonian evaluated at the atomic reference density ρ_0 , which is determined by solving the Kohn–Sham equations for the atom in the presence of an additional harmonic confining potential with a confining radius r^{wf} for the atomic basis set and a different radius r^{dens} for the density. Off-diagonal $H_{\mu\nu}^0$ integrals are computed in a two-center approximation together with the PBE⁸¹ exchange-correlation functional and an atomic basis set consisting of Slater functions. The Slater functions are defined by l_{\max} as the highest orbital angular momentum included, n_{\max} and $\alpha_0, \alpha_1, \dots, \alpha_4$, the exponents of the Slater functions. In principle, basis functions for each orbital angular momentum can be compressed differently; in this work, however, s and p orbitals are confined with the same compression radius for Mg, and s, p, and d orbitals (compression radius for d-orbitals is defined by the parameter r^{wfd}) are confined with the same compression radius for Zn. The diagonal elements $H_{\mu\mu}^0$ are equal to the calculated atomic eigenvalues ϵ_x ($x = s, p$, or d orbital). No confinement of the orbitals is applied here in order to yield the correct dissociation limit. We optimize ϵ_p as discussed in the following section.

The atomic Hubbard parameter U_a in the second-order term of the energy E^γ describes the electron–electron interaction within one atom and enters the γ -function, which interpolates the on-site and the long-range electronic interaction of atom pairs. The Hubbard parameters are calculated from DFT as the first derivative of the highest occupied orbital with respect to its occupation number. For the third-order term E^Γ , an additional parameter per element U^d , the derivative of the Hubbard parameter with respect to charge has to be determined. For CHNO it was calculated as the second derivative of the highest occupied orbital with respect to its occupation number;⁷⁷ for Mg and Zn we find a significant advantage to optimize it.

Reference Systems and Parameter Fitting. As emphasized earlier,⁵⁴ most parameters in the DFTB3 approach are directly computed based on atoms or atom pairs at the DFT (PBE) level, while some parameters are optimized based on specific reference systems. In this way, accuracy appropriate for chemical applications can be obtained without significantly reducing the transferability of the DFTB3 model. The parametrization for Mg/Zn used herein follows that specified previously for the parametrization of CHNOPS;^{77,78} two distinct groups of parameters have to be optimized, the electronic parameters appearing in the first three terms of the DFTB3 total energy (eq 1) and the repulsive potential.

The electronic parameters are summarized in Table 1. The following parameters are optimized to improve the overall performance:

- r^{wf} and r^{dens} : The general procedure was to start from the electronic parameters used in the MIO set^{75,77} (which are $r_{\text{Mg}}^{\text{wf}} = 5.0$, $r_{\text{Mg}}^{\text{dens}} = 12.0$, $r_{\text{Zn}}^{\text{wf}} = 4.9$, $r_{\text{Zn}}^{\text{dens}} = 10.0$). r^{wf} and r^{dens} are fitted in a “brute force” manner to achieve good performance on bond (metal–ligand), angle, vibrational frequencies, as well as reaction energies.

- ϵ_p : The eigenvalues are in principle computed directly from atomic DFT calculations and enter as diagonal elements into

Table 1. Overview of Electronic Parameters (in Atomic Units if Not Unitless)^a

parameter	Mg	Zn
l_{\max}	1	2
n_{\max}	2	2
α_0	0.50	0.50
α_1	1.11	1.39
α_2	2.45	3.87
α_3	5.42	10.78
α_4	12.00	30.00
r^{wf}	5.5	4.2
r^{wfd}	0.0	4.2
r^{dens}	14.0	9.0
ϵ_s	-0.17267741	-0.21408495
ϵ_p	-0.02	0.02
ϵ_d	-	-0.38831893
E^{spin}	0.0	0.0
U	0.2246	0.2663
U^d	-0.02	-0.03

^aAs described in the text, U , E^{spin} , $\epsilon_{s,d}$, n_{\max} , and α_i are consistent with the standard choices of DFTB parametrization and not subject to optimizations. By contrast, r^{wf} , r^{wfd} , r^{dens} , ϵ_p , and U^d are adjusted based on properties of molecules in the fitting set.

the Hamiltonian matrix without the additional confining potential. Exceptions are made here to magnesium and zinc since the p orbitals are often not heavily involved in the coordination of Mg^{2+} and Zn^{2+} . ϵ_p is calculated to be -0.04914 au for Mg and -0.04261 au for Zn. ϵ_p are fitted in the way that p orbitals are less involved in molecular bonding, so as to improve sequential ligation energies and proton affinities.

• U^d : The Hubbard derivative is generally computed from atomic DFT calculations. In the case of Mg and Zn, they are -0.0496 au and -0.0544 au, respectively. We found a significant improvement of proton affinities after optimizing U^d for Mg and Zn. It is important to note that fitting of the Hubbard derivatives does not significantly alter most properties of neutral molecules like equilibrium geometries, so with all the other parameters held fixed, U^d is tuned as the last step in the parametrization procedure.

Although the repulsive potential, which contains nuclear–nuclear repulsion and double-counting terms that depend on the reference density, can be directly computed, practical application to molecular systems with reasonable accuracy requires fitting the repulsive potential. The repulsive potential between atom type A and type B (V_{ab}^{rep} in eq 1) is described by a spline function in the covalent binding region. A cutoff is introduced at which the function and first three derivatives reach zero. The remaining degrees of freedom for the spline function are fitted to high-level *ab initio* values of geometries, reaction energies, and vibrational stretching frequencies. The overall quality of the repulsive potential is sensitive to the chosen division points. Therefore, establishing reliable repulsive potential parameters remains a time-consuming step despite the progress of partially automatized fitting procedures.^{80,82}

The repulsive potentials are fitted to several properties as summarized in Table 2. Reference values are calculated from *ab initio* methods that are known to produce results close to experimental data. For magnesium, optimized geometries of small closed-shell molecules are calculated from B3LYP^{83–85}/aug-cc-pVTZ, and reaction energies are calculated from the G3B3^{86–88} method. For zinc, both optimized geometries and

Table 2. Parameters Defining the Repulsive Potentials^a

reference geometries for force equations			
$[\text{MgH}_2]$, $[\text{Mg}(\text{CH}_3)_2]$, $[\text{MgO}]$, $[\text{Mg}(\text{H}_2\text{O})_2]^{2+}$, $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$, $[\text{Mg}(\text{NH}_3)_4]^{2+}$, $[\text{MgS}]$, $[\text{Mg}(\text{SH}_2)_3(\text{SH})]^+$, $[\text{Mg}(\text{PH}_3)_2]$, $[\text{Mg}(\text{PH}_3)_4]^{2+}$, $[\text{ZnH}_2]$, $[\text{Zn}(\text{CH}_3)]^+$, $[\text{Zn}(\text{CO})]^{2+}$, $[\text{Zn}(\text{H}_2\text{O})_3]^{2+}$, $[\text{Zn}(\text{OH})_2]$, $[\text{Zn}(\text{H})](\text{NH}_3)$, $[\text{Zn}(\text{NH}_3)_4]^{2+}$, $[\text{Zn}(\text{SH})]^+$, $[\text{Zn}(\text{SH}_2)_3]$, $[\text{Zn}(\text{SH}_2)_3(\text{SH})]^+$, $[\text{Zn}(\text{PH}_3)]^+$, $[\text{Zn}(\text{PH}_3)_3(\text{PH}_2)]^+$			
reference potential energy differences in kcal/mol			
$[\text{Mg}(\text{H}_2\text{O})_2]^{2+}$	→	$[\text{Mg}(\text{H}_2\text{O})]^{2+} + \text{H}_2\text{O}$	73.3
$[\text{Mg}(\text{H}_2\text{O})_3]^{2+}$	→	$[\text{Mg}(\text{H}_2\text{O})_2]^{2+} + \text{H}_2\text{O}$	59.2
$[\text{Mg}(\text{CO})_4]^{2+}$	→	$[\text{Mg}(\text{CO})_3]^{2+} + \text{CO}$	30.9
$[\text{Mg}(\text{NH}_3)]^{2+}$	→	$\text{Mg}^{2+} + \text{NH}_3$	97.7
$[\text{Mg}(\text{NH}_3)_2]^{2+}$	→	$[\text{Mg}(\text{NH}_3)]^{2+} + \text{NH}_3$	83.9
$[\text{Mg}(\text{NH}_3)_3]^{2+}$	→	$[\text{Mg}(\text{NH}_3)_2]^{2+} + \text{NH}_3$	63.1
$[\text{Mg}(\text{NH}_3)_4]^{2+}$	→	$[\text{Mg}(\text{NH}_3)_3]^{2+} + \text{NH}_3$	50.1
$[\text{Mg}(\text{NH}_3)_5]^{2+}$	→	$[\text{Mg}(\text{NH}_3)_4]^{2+} + \text{NH}_3$	28.6
$[\text{Mg}(\text{NH}_3)_6]^{2+}$	→	$[\text{Mg}(\text{NH}_3)_5]^{2+} + \text{NH}_3$	26.3
$[\text{Mg}(\text{SH}_2)_2]^{2+}$	→	$[\text{Mg}(\text{SH}_2)]^{2+} + \text{SH}_2$	62.4
$[\text{Mg}(\text{SH}_2)_3]^{2+}$	→	$[\text{Mg}(\text{SH}_2)_2]^{2+} + \text{SH}_2$	43.9
$[\text{Mg}(\text{SH}_2)_4]^{2+}$	→	$[\text{Mg}(\text{SH}_2)_3]^{2+} + \text{SH}_2$	34.5
$[\text{Mg}(\text{SH}_2)_5]^{2+}$	→	$[\text{Mg}(\text{SH}_2)_4]^{2+} + \text{SH}_2$	20.0
$[\text{Mg}(\text{PH}_3)_2]^{2+}$	→	$[\text{Mg}(\text{PH}_3)]^{2+} + \text{PH}_3$	66.9
$[\text{Mg}(\text{PH}_3)_4]^{2+}$	→	$[\text{Mg}(\text{PH}_3)_3]^{2+} + \text{PH}_3$	34.8
$[\text{Zn}(\text{H}_2\text{O})]^{2+}$	→	$\text{Zn}^{2+} + \text{H}_2\text{O}$	104.3
$[\text{Zn}(\text{H}_2\text{O})_4]^{2+}$	→	$[\text{Zn}(\text{H}_2\text{O})_3]^{2+} + \text{H}_2\text{O}$	43.0
$[\text{Zn}(\text{CO})]^{2+}$	→	$\text{Zn}^{2+} + \text{CO}$	81.6
$[\text{Zn}(\text{CO})_2]^{2+}$	→	$[\text{Zn}(\text{CO})]^{2+} + 2\text{CO}$	149.8
$[\text{Zn}(\text{CO})_3]^{2+}$	→	$[\text{Zn}(\text{CO})]^{2+} + 3\text{CO}$	190.1
$[\text{Zn}(\text{SH}_2)]^{2+}$	→	$\text{Zn}^{2+} + \text{SH}_2$	126.1
$[\text{Zn}(\text{SH}_2)_2]^{2+}$	→	$\text{Zn}^{2+} + 2\text{SH}_2$	210.5
$[\text{Zn}(\text{PH}_3)_4]^{2+}$	→	$[\text{Zn}(\text{PH}_3)_3]^{2+} + \text{PH}_3$	30.5
potential	division points (au)	additional equations (au)	
C–Mg	3.6, 4.0, 4.5, 5.1	$V''(3.959) = 0.073$	
H–Mg	2.8, 3.0, 3.3, 3.9	$V''(3.214) = 0.063$	
N–Mg	3.2, 3.9, 4.2, 5.5	$V''(3.581) = 0.15$, $V''(3.906) = 0.065$	
O–Mg	2.8, 3.8, 4.5, 5.1	$V''(3.330) = 0.23$, $V''(3.660) = 0.070$	
P–Mg	4.4, 4.7, 5.3, 5.7	$V''(4.851) = 0.025$	
S–Mg	3.8, 4.5, 5.3, 5.8	$V''(4.344) = 0.05$	
C–Zn	3.5, 3.7, 3.9, 4.2	$V''(3.698) = 0.19$	
H–Zn	2.4, 2.8, 3.0, 4.0	$V''(2.914) = 0.10$	
N–Zn	3.0, 3.6, 4.4, 5.0	$V''(3.677) = 0.09$	
O–Zn	2.8, 3.5, 3.8, 4.9		
P–Zn	4.0, 4.3, 4.7, 5.2	$V''(4.482) = 0.075$	
S–Zn	3.5, 4.2, 4.7, 5.7		

^aSee text for the methods used as references for different properties.

reaction energies are calculated from B3LYP/aug-cc-pVTZ. Additional equations for the fitting process are prepared using a few vibrational stretching frequencies determined from BLYP/aug-cc-pVTZ (unscaled) using the harmonic approximation. B3LYP with good quality basis set has been shown to describe structures of Zn complexes in good agreement with X-ray⁸⁹ and CCSD(T)^{43,90} data. However, heats of formation (HOF) predicted from B3LYP/aug-cc-pVTZ deviate ~10 kcal/mol on average from experimental data of eight small complexes even though CCSD(T)/aug-cc-pVTZ does not do much better with a mean unsigned error of ~7 kcal/mol.⁹¹ B97-1,⁹² B97-2,⁹³ and B98^{94,95} have also been tested because they have been reported to be reliable for reproducing experimental data for small transition metal complexes;⁹⁶ geometries from the B97 family functionals and B3LYP are very similar, and reaction energies differ by ~2 kcal/mol (data shown in the Supporting Information). Despite shortcomings in HOF, we have chosen

to use the B3LYP functional because it is computationally less intensive so that large sets of consistent reference data are accessible, and it systematically gives reliable results for binding energies and proton affinities, which are of major importance to our work. Since the systems studied here are relatively small in size, dispersion interactions are not expected to impact the structural and energetic properties of interest. For a fairly large active site model for a zinc enzyme, we do discuss briefly the impact of dispersion. For even larger cases, we note that empirical dispersion models have been developed and tested for both SCC-DFTB^{97,98} and DFTB3.⁹⁹

Condensed-Phase Benchmark Calculations. To supplement the gas-phase calculations, simulations are carried out with DFTB3/MM to investigate if the parametrized model works in an explicit condensed-phase environment. They are important tests because the ultimate goal is to use DFTB3/3OB in enzyme simulations.

Mg²⁺ and Zn²⁺ Ion Solvation and pK_a Calculations. Both MM and QM/MM simulations are performed using CHARMM.¹⁰⁰ The generalized solvent boundary potential (GSBP)^{101,102} is used to treat long-range electrostatic interactions in MD simulations. The system is set up with an inner region of 20 Å. All molecules are located in the inner region, and the outer region is described with a constant dielectric continuum with $\epsilon_w = 80$. To be consistent with the GSBP protocol, electrostatic interactions among inner region atoms are treated with the extended electrostatic model¹⁰³ in which interactions beyond 12 Å are modeled with multipolar expansions, including the dipolar and quadrupolar terms. The reaction field matrix \mathbf{M} is evaluated using 400 spherical harmonics for GSBP simulations. The ion is fixed at the center of the sphere. Water molecules are described with the modified version of TIP3P.¹⁰⁴ In the classical MM simulations, standard CHARMM force field¹⁰⁵ parameters are used for Mg²⁺ and Zn²⁺. The QM/MM calculations employ the same water sphere as in MM simulations. The QM region includes the metal ion together with the nearest 6 water molecules (referred to as small QM region) or the nearest 25 water molecules (referred to as large QM region) which include the first and the second solvation shells. All the remaining water molecules are represented by the TIP3P water model. In the large QM region simulations, to ensure the QM water molecules are always closest to the metal ion, Flexible Inner Region Ensemble Separator (FIRES¹⁰⁶) restraining potential is imposed to any outer sphere MM water molecules that are closer to the ion than the most distant inner sphere QM water molecule. The FIRES potential takes the form $E_{\text{FIRES}} = (1/2)\sum k_{\text{FIRES}}(r_i - R_{\text{inner}})^2$ where R_{inner} is defined as the distance between the metal ion and the most distant inner QM water molecule from it. k_{FIRES} is set to be 100 kcal/(mol·Å²). FIRES restraints are not necessary for the small QM region simulations because water molecules in the first solvation shell are strongly bound to the metal ion. All bonds involving hydrogen are constrained with the SHAKE algorithm,¹⁰⁷ allowing a 1 fs time step for MD propagation. Simulations are equilibrated for 500 ps followed by a production run of 500 ps.

To further evaluate the performance of DFTB3/3OB for Mg and Zn from the energetics perspective, we perform calculations of relative pK_a of Mg²⁺/Zn²⁺-bound water in a water sphere using the thermodynamic integration (TI)^{108,109} approach within the dual topology single coordinate (DTSC) scheme; absolute pK_a results are discussed in the Supporting Information.^{66,110} In this approach, the dominant contribution

to the total free energy of deprotonation is from the electrostatic free energy change ($\Delta F_{M^{2+} \cdot (H_2O)_5(HOH/D)}$) associated with converting the acidic proton to a dummy atom (D), i.e., transforming $M^{2+} \cdot (H_2O)_6$ to $M^{2+} \cdot (H_2O)_5(HOD)^-$, where M represents the metal ion, Mg²⁺ or Zn²⁺. The corresponding free energy derivative is given by

$$\frac{\partial \Delta F_{M^{2+} \cdot (H_2O)_5(HOH/D)}}{\partial \lambda} = \langle U_{\text{elec}}^{M^{2+} \cdot (H_2O)_5(HOD)^-} - U_{\text{elec}}^{M^{2+} \cdot (H_2O)_6} \rangle_{\lambda} \quad (2)$$

which represents the QM/MM energy difference averaged for a specific coupling parameter λ using the same set of coordinates for both protonation states. The total electrostatic energy contribution ($\Delta F_{M^{2+} \cdot (H_2O)_5(HOH/D)}$) is determined by integrating the free energy derivatives ($\partial F / \partial \lambda$) over λ from 0 to 1. Estimation of the absolute pK_a is in principle possible with the DTSC-TI protocol, but the solvation free energy of a proton has to be taken into account. To avoid the difficulty of getting accurate experimental/calculated values for this term, we compute the relative pK_a between $Mg^{2+} \cdot (H_2O)_6$ and $Zn^{2+} \cdot (H_2O)_6$ in water so that other contributions are canceled out to a great extent, such as the zero-point energy difference between the protonated and deprotonated states and van der Waals interactions involving the acidic proton.^{66,110} Since $\Delta F_{M^{2+} \cdot (H_2O)_5(HOH/D)}$ is electrostatic in nature, we expect that the free energy derivative depends largely linearly on the coupling parameter λ in this simple system; this is supported by actual calculations (see below). We carry out pK_a calculations with both small QM region and large QM region as defined above. To avoid dissociation of hydroxide from the metal ion in the fully deprotonated state ($\lambda = 1$ window $M^{2+} \cdot (H_2O)_5(HOD)^-$), a NOE potential is added to the distance between M^{2+} and OH⁻. The NOE potential takes the form

$$E_r = 0.0 \quad r_{\min} < r < r_{\max} \\ = 0.5 \cdot k_{\max} \cdot (r - r_{\max})^2 \quad r_{\max} < r \quad (3)$$

in which r_{\min} and r_{\max} set the interval between which the restraining potential is zero; they are taken to be 1.8 and 2.5 Å, respectively. k_{\max} is set to be 1000 kcal/(mol·Å²). To prevent the dummy atom D from being too close to the metal site in the fully deprotonated state, a weak angular constraint is applied to the M–O–D angle to keep it no smaller than 120°.

Myosin-II Enzyme Setup: Test for a Mg Site. The enzyme model is built from a high-resolution X-ray structure for the Mg-ADP-vanadate bound myosin¹¹¹ (PDB code 1VOM), which corresponds to the hydrolyzing state of the *D. discoideum* myosin-II motor domain.^{112,113} As described in our earlier work,^{68,114,115} starting from the PDB structure, the ADP-vanadate is replaced by ATP or ADP-Pi. The QM region includes the triphosphate and part of the ribose group of ATP, three water molecules (include the lytic water) in the active site, Ser 181, Ser 236, the Mg²⁺ ion, and all its ligands (Thr 186 and Ser 237 and two water molecules). Only side chains of protein residues are included in the QM region, and link atoms are introduced to saturate the valence of the QM boundary atoms. ATP is assumed to be fully deprotonated, and the titratable groups are kept in their normal protonation states (i.e., Lys, Arg are protonated, Asp, Glu are deprotonated) which are consistent with a simple estimate of pK_a using the Poisson–Boltzmann (PB) approach.¹¹⁶ GSBP is used as the simulation protocol. The system is divided into a 24 Å spherical

Table 3. Mean and Maximum Absolute Deviation of Metal–Ligand Lengths and Angles in Comparison to B3LYP/aug-cc-pVTZ for the Magnesium Test Set

property ^a	<i>N</i> ^b	MP2 ^c	B3LYP ^d	PM6	DFTB3/MIO	DFTB3/3OB
<i>r</i> (Å)	120	0.009	0.009	0.050	0.017	0.013
<i>r</i> _{max} (Å)		0.051	0.024	0.284	0.141	0.072
<i>a</i> (deg)	117	0.9	0.5	3.4	3.6	2.9
<i>a</i> _{max} (deg)		23.7	10.2	38.7	41.3	45.3

^aMax stands for maximum absolute deviation. ^bNumber of comparisons. ^cBasis set aug-cc-pVTZ. ^dBasis set 6-31+G(d,p).

inner region centered at the Mg ion. Newtonian equations-of-motion are solved for the MD region (within 20 Å), and Langevin equations-of-motion are solved for the buffer region (20–24 Å) with a temperature bath of 300 K. Protein atoms in the buffer region are harmonically constrained with force constants determined from the crystallographic B-factors.¹¹⁷ Protein atoms in the MM region are described by the all-atom CHARMM22 force field,¹⁰⁵ and water molecules are described with the TIP3P model. All bonds involving hydrogen are constrained using SHAKE, and the time step is set to 1 fs. All the water molecules in the inner region are subject to a weak GEO type of restraining potential to keep them inside the inner sphere with the MMFP module in CHARMM. The static field due to outer-region atoms, $\phi_s^{(io)}$, is evaluated with the linear PB equations. The reaction field matrix M is evaluated with 625 spherical harmonics. In the PB calculations, the protein dielectric constant is set as 1; the water dielectric constant is set as 80; and 0.0 M salt concentration is used. Our previous¹¹⁸ and recent¹¹⁹ analyses indicated that the GSBP protocol is reliable for a site well shielded from the bulk solvent.

To study the structural flexibility in the active site of myosin before and after ATP hydrolysis, potential of mean force (PMF) simulations have been carried out. The reaction coordinate is defined as the distance between Mg²⁺ and atom O in Ser 237 (see Figure 5). The umbrella sampling approach¹²⁰ is used to restrain the system along the reaction coordinate with a force constant of 150 kcal/mol·Å⁻². Thirty-two windows are used for PMF, and a 500 ps simulation is performed for each window. Convergence of PMF is monitored by examining the overlap of reaction coordinate distributions sampled in different windows. Weighted histogram analysis (WHAM)¹²¹ is used to postanalyze the probability distribution to obtain the PMF along the reaction coordinate. Errors are calculated from block average.

Alkaline Phosphatase Enzyme Setup: Test for a Bimetallic Zinc Site. The enzyme model is constructed based on the X-ray structures for the *E. coli* Alkaline Phosphatase (AP) mutant R166S with bound inorganic phosphate at 2.05 Å resolution (PDB code 3CMR¹²²). The substrate methyl *p*-nitrophenyl phosphate (MpNPP⁻) is first “mutated” to the orientation with the –OMe group oriented toward the magnesium ion (denoted as α orientation) starting from the PDB structure. All basic and acidic amino acids are kept in their physiological protonation states except for Ser102 in AP, which is accepted to be the nucleophile and deprotonated in the reactive complex. Water molecules are added following the standard protocol of superimposing the system with a water droplet of 25 Å radius centered at Zn²⁺ (see Figure 6 for atomic labels) and removing water molecules within 2.8 Å from any heavy atoms resolved in the crystal structure. GSBP is used to treat long-range electrostatic interactions in MD simulations. In the QM/MM simulations, as described in details in our previous work,^{48,49} the QM region includes the two zinc ions and their

six ligands (Asp51, Asp369, His370, Asp327, His412, His331), Ser102, and the substrate MpNPP⁻. Only side chains of protein residues are included in the QM region, and link atoms are added between C_α and C_β atoms. A NOE potential is added to the C–O bonds in Asp51 in AP to prevent artificial polarization.⁴⁸ The integration time step is 1 fs, and all bonds involving hydrogens are constrained using SHAKE. The DFTB3/MM results are also compared to MM simulations using a conventional nonbonded zinc model¹⁴ (referred to as a Coulomb scheme) or short–long effective functions (SLEF1)²⁵ model developed by Zhang and co-workers. Protein atoms in the MM region are described by the all-atom CHARMM22 force field, and water molecules are described with the TIP3P model.

To further benchmark the applicability of DFTB3/3OB to the reaction of interest, we also study an active site model that includes all QM atoms in the QM/MM enzyme model. The C_β carbons are fixed at their positions in the crystal structure during geometry optimization; the positions of the link atoms used to saturate the valence of the C_β atoms in the active site model are also fixed. The reactant (Michaelis) complex and transition state are located for MpNPP⁻ (methyl *p*-nitrophenyl phosphate), MmNPP⁻ (methyl *m*-nitrophenyl phosphate), and MPP⁻ (methyl phenyl phosphate) using DFTB3 and B3LYP with the 6-31+G(d,p) basis set. The minimum energy path (MEP) calculations are carried out by one-dimensional adiabatic mapping at the DFTB3 level; the reaction coordinate is the antisymmetric stretch involving the breaking and forming P–O bonds. Subsequently, the saddle point is further refined by conjugated peak refinement (CPR¹²³) to obtain precise transition state structure. Single-point energy calculations at DFTB3 and B3LYP geometries are then carried out using B3LYP, M06,¹²⁴ PBE, and MP2 methods using a larger basis set of 6-311++G(d,p). The D3¹²⁵ dispersion corrections are added for B3LYP, M06, and PBE methods.

RESULTS AND DISCUSSIONS

In the following subsections, we present benchmark calculations for DFTB3/3OB regarding geometries, metal–ligand dissociation energies, and proton affinities for Mg/Zn containing molecules and compare the results with commonly used *ab initio* methods in the gas phase. We also report DFTB3 results when the old set of MIO parameters is used:⁷⁵ DFTB3/MIO use parameters as defined in ref 75, U^d = −0.23, −0.16, −0.13, −0.19, −0.14, and −0.09 au for C, H, N, O, P, and S and ζ = 4.2; U^d for Mg/Zn is fitted to −0.1 to achieve the best sequential ligand dissociation energies and proton affinities. Unless noted otherwise, energies are given for geometries that are optimized at the respective level of theory. For PM6, we note that the choice of heat of formation for protons is essential to proton affinities; we take −54.0 kcal/mol for protons as suggested by the MOPAC program. To compare the potential energy surfaces between different computation levels, zero-

Table 4. Sequential Ligand Dissociation Energies of Magnesium -Containing Molecules Compared to G3B3^a

molecule ^b	G3B3	MP2 ^c	B3LYP ^c	PM6	DFTB3/MIO	DFTB3/3OB	G3B3//DFTB3 ^d
[Mg(H ₂ O) ₂] ²⁺	73.3	-3.5	-0.1	-10.9	-3.4	+0.5	+0.6
[Mg(H ₂ O) ₃] ²⁺	59.2	-2.0	-1.1	-11.7	-1.5	+1.4	-0.3
[Mg(H ₂ O) ₄] ²⁺	49.1	-2.3	-3.0	-11.0	-3.3	-1.9	+1.2
[Mg(H ₂ O) ₅] ²⁺	34.7	-2.2	-5.0	-7.4	-6.7	-6.4	-0.2
[Mg(H ₂ O) ₆] ²⁺	32.9	-2.5	-5.9	-11.0	-6.8	-6.4	+0.4
[Mg(CO) ₂] ²⁺	46.3	-0.9	-0.3	+10.0	-2.9	-3.8	-0.6
[Mg(CO) ₃] ²⁺	36.1	+0.5	-0.9	+1.3	-0.3	-2.3	-0.7
[Mg(CO) ₄] ²⁺	30.9	-0.1	-2.8	-1.5	-0.1	-2.6	-0.9
[Mg(NH ₃) ₂] ²⁺	83.9	-3.4	-0.3	-4.4	-13.8	-6.4	+0.5
[Mg(NH ₃) ₃] ²⁺	63.1	-1.8	-2.0	-9.2	-7.5	-1.4	+0.2
[Mg(NH ₃) ₄] ²⁺	50.1	-1.8	-3.7	-8.8	-3.9	+1.2	+0.2
[Mg(NH ₃) ₅] ²⁺	28.6	-1.0	-5.3	-2.9	-5.8	-3.5	-0.2
[Mg(NH ₃) ₆] ²⁺	26.3	-1.5	-6.2	-5.5	-5.6	-2.5	-0.2
[Mg(SH ₂) ₂] ²⁺	62.4	-3.0	-0.3	-7.0	-7.9	-2.9	-1.2
[Mg(SH ₂) ₃] ²⁺	43.9	-1.1	-2.6	-12.1	-4.5	+0.1	-0.6
[Mg(SH ₂) ₄] ²⁺	34.5	-0.8	-4.4	-13.3	-2.8	+1.2	+1.7
[Mg(SH ₂) ₅] ²⁺	20.0	+0.2	-6.1	-12.6	+0.0	+1.1	-0.8
[Mg(SH ₂) ₆] ²⁺	20.7	-0.5	-7.6	-14.0	-1.8	-0.1	+0.5
[Mg(PH ₃) ₂] ²⁺	66.9	-2.3	-0.5	-2.4	-8.6	-0.1	+0.7
[Mg(PH ₃) ₃] ²⁺	44.5	-0.8	-2.9	-18.6	-5.4	+2.3	+0.7
[Mg(PH ₃) ₄] ²⁺	34.8	-0.6	-4.5	-17.9	-3.8	+3.2	+0.6
[Mg(OH) ₂]	243.1	-7.2	-4.6	-7.1	-21.0	-14.0	+0.1
[Mg(NH ₂) ₂]	237.9	-7.3	-6.6	-4.6	-9.4	-12.0	+0.9
[Mg(SH) ₂]	201.2	-3.1	-5.0	-30.3	-16.1	-15.8	+0.1
[Mg(PH ₂) ₂]	188.8	-1.8	-5.2	+11.0	-20.9	-13.3	+0.6
MAD		2.1	3.5	9.9	6.6	4.3	0.6
MAX		7.3	7.6	30.3	21.0	15.8	1.7

^aEnergies except PM6 are calculated at 0 K excluding zero-point energy and thermal corrections. All numbers are given in kcal/mol. ^bThe nondissociated is listed. ^cBasis set aug-cc-pVTZ. ^dG3B3 single-point calculations on top of DFTB3/3OB geometries.

point energies (ZPEs) are not included. For the calculations of noncovalent interactions, dispersion interactions are not included because most test molecules discussed here are small, although we are aware of the importance of dispersion in general. We explore explicitly the effects of dispersion to a nontrivial active site model of AP that involves zinc. For DFTB3 calculations, we have used our in-house DFTB code, and all PM6, DFT, and MP2 calculations are performed using the Gaussian09 software package.¹²⁶ Following the gas-phase benchmarks, we show results on Mg²⁺/Zn²⁺-water structural properties and relative pK_a in solution. Finally we apply DFTB3/3OB to enzyme systems to further test the new parameters.

Magnesium. Geometries. The “small” magnesium test set is composed of 55 molecules, which are simplified biologically relevant ligated magnesium species of the form Mg²⁺[X]_n with X = H₂O, NH₃, SH₂, PH₃, CO for n = 1–4/6 and X = OH[−], SH[−], NH₂[−], and PH₂[−] for n = 1–2. Table 3 summarizes the results, and further details are included in the Supporting Information. Good agreement is seen between MP2 and B3LYP with a small or large basis set. Of the semiempirical methods, geometries are described very well by DFTB3 models and particularly well by DFTB3/3OB. The MADs in metal-ligand distances are 0.013, 0.017, and 0.050 Å for DFTB3/3OB, DFTB/MIO, and PM6, respectively. 3OB improves over MIO, although the magnitude of improvement is modest. The largest metal-ligand distance error of 0.072 Å from DFTB3/3OB is found in Mg–S distance in [Mg(SH₂)₂]²⁺, which is likely due to the minimal basis nature of DFTB3 for a polarizable ligand. Other deviations in metal-ligand distances are substantially

smaller in DFTB3/3OB. Molecules containing S or P are described poorly by PM6 within our test case;⁷⁸ PM6 systematically underestimates Mg–S distance by 0.2 Å and Mg–P distance by 0.1 Å. For angles, DFTB3/3OB differs from the reference data by a MAD of 2.9°, which is of the same order with other semiempirical methods. The largest angle derivation of 45.3° is for angle Mg–O–H in [Mg(OH)]⁺, for which DFTB3/3OB predicts a linear rather than bent structure; interestingly, MP2 also predicts a structure almost linear, while PM6 is close to the B3LYP result (see Supporting Information). Future development of DFTB3 to include multipoles may improve the geometry in this case. It is worth pointing out that the Mg–S–H angle is generally overestimated by ~13° in DFTB3/3OB. More statistics including MAD for different metal–ligand types are given in the Supporting Information.

Sequential Metal–Ligand Dissociation Energies. We define sequential ligand dissociation energies (sBDEs) corresponding to the reaction



where M is Mg²⁺/Zn²⁺ and L is a neutral or charged ligand.

Table 4 shows the sequential ligand dissociation energies for compounds containing Mg²⁺. Values obtained for DFTB3/3OB and other methods are compared to those obtained using G3B3. The differences between B3LYP, MP2, and G3B3 are relatively small; however, complexes with charged ligands give several deviations greater than 5 kcal/mol. DFTB3/3OB shows a MAD of 2.4 kcal/mol for complexes with neutral ligands and 13.8 kcal/mol for complexes with charged ligands. Compared

Table 5. Gas-Phase Proton Affinities of Magnesium-Containing Molecules Compared to G3B3^a

molecule ^b	G3B3	MP2 ^c	B3LYP ^c	PM6	DFTB3/MIO	DFTB3/3OB	G3B3//DFTB3 ^d
[Mg(H ₂ O) ₁] ²⁺	95.6	+0.6	-3.5	+0.1	-5.5	-13.9	-0.8
[Mg(H ₂ O) ₂] ²⁺	115.6	+0.0	-1.4	-5.4	-2.2	-9.0	-0.9
[Mg(H ₂ O) ₃] ²⁺	133.1	+0.6	+0.2	-7.1	+3.0	-1.2	-0.4
[Mg(H ₂ O) ₄] ²⁺	149.2	-0.6	-0.1	-11.4	+4.8	+2.3	+0.8
[Mg(H ₂ O) ₅] ²⁺	160.3	-2.2	-1.9	-15.7	+3.2	+2.0	+1.1
[Mg(H ₂ O) ₆] ²⁺	170.5	-3.0	-2.2	-24.3	+0.2	-0.2	+0.8
[Mg(NH ₃) ₁] ²⁺	120.2	+1.5	-3.5	+2.6	-10.5	-15.7	+0.0
[Mg(NH ₃) ₂] ²⁺	144.1	+0.8	-0.4	+1.5	-10.2	-14.1	+0.2
[Mg(NH ₃) ₃] ²⁺	163.9	+0.5	+0.3	+1.9	-10.6	-13.6	+0.5
[Mg(NH ₃) ₄] ²⁺	181.3	-0.1	+0.6	+1.0	-6.4	-7.5	+1.0
[Mg(NH ₃) ₅] ²⁺	190.5	-0.4	+0.9	+4.9	-5.2	-6.1	+1.4
[Mg(NH ₃) ₆] ²⁺	200.7	+1.8	-0.5	+5.0	-6.2	-5.4	+2.3
[Mg(SH ₂) ₁] ²⁺	83.4	+2.1	-1.8	+2.4	-6.7	-0.7	-3.6
[Mg(SH ₂) ₂] ²⁺	107.0	+0.8	+0.5	+4.6	-6.2	+0.8	-2.9
[Mg(SH ₂) ₃] ²⁺	123.1	+0.4	+1.5	+3.3	-6.2	+2.1	-3.5
[Mg(SH ₂) ₄] ²⁺	135.4	+0.0	+2.0	+1.0	-5.7	+4.2	-1.7
[Mg(SH ₂) ₅] ²⁺	143.1	-0.2	+2.0	-5.7	-6.9	+3.9	-3.9
[Mg(SH ₂) ₆] ²⁺	151.1	-0.6	+1.8	-12.4	-8.7	+3.5	-3.0
[Mg(PH ₃) ₁] ²⁺	96.8	+4.1	-1.8	+12.6	-12.8	-1.1	+0.6
[Mg(PH ₃) ₂] ²⁺	129.4	+1.7	+0.4	+10.2	-15.3	-0.4	+1.4
[Mg(PH ₃) ₃] ²⁺	150.1	+1.0	+1.2	+8.2	-17.6	-1.7	+1.4
[Mg(PH ₃) ₄] ²⁺	164.2	+0.7	+1.7	+3.0	-18.1	-0.7	+1.2
MAD		1.1	1.4	6.6	7.8	5.0	1.5
MAX		4.1	3.5	24.3	18.1	15.7	3.9

^aEnergies except PM6 are calculated at 0 K excluding zero-point energy and thermal corrections. All numbers are given in kcal/mol. ^bThe protonated form is listed. ^cBasis set aug-cc-pVTZ. ^dG3B3 single-point calculations on top of DFTB3/3OB geometries.

to DFTB3/MIO, DFTB3/3OB improves substantially for complexes involving P, S elements and charged ligands. For the other semiempirical method PM6, the deviation is 9.9 kcal/mol compared to G3B3. The large error is anticipated because geometries containing S and P are often poor. Note that we compare potential energy differences for all the methods except PM6 which intrinsically considers heats of formation. We have also carried out single-point G3B3 calculations at DFTB3/3OB geometries, and the MAD for Mg²⁺ compounds with neutral and charged ligands is less than 1 kcal/mol. These results highlight the similarities between DFTB3 and B3LYP optimized structures, suggesting that DFTB3/3OB is of particular value for the study of biologically relevant magnesium-containing molecules.

Proton Affinities. Table 5 lists the proton affinities for 22 biological relevant molecules Mg²⁺(XH_m)_n where X = O, N, S, and P obtained from G3B3, DFTB3/3OB, and other methods. MADs for MP2 and B3LYP are generally small (\sim 1 kcal/mol). The discrepancy between DFTB3/3OB and G3B3 is 5 kcal/mol, whereas DFTB3/MIO and PM6 yield slightly larger MADs of 7.8 and 6.6 kcal/mol, respectively. We note that G3B3 single points at DFTB3/3OB optimized structures lead to small MAD of 1.5 kcal/mol and MAX of 3.9 kcal/mol (due to the unsatisfactory deprotonated [Mg(SH₂)_n]²⁺ structures) compared to the original G3B3, again highlighting the good quality of DFTB3/3OB structures. Future improvements should focus on the nitrogen-containing species, an issue we noted in previous studies.^{75,77}

Large Test Set. To further test DFTB3/3OB parameters we have compiled a “large test set” that models biologically relevant ligands of magnesium with larger molecules than those discussed above. The set mainly consists of model magnesium compounds that are simplified active sites of crystal structures

collected in a survey.¹²⁷ Although Mg²⁺ is generally 6-coordinated, complexes with coordination number of 4 and 5 are also investigated. All complexes are constructed in the way that all ligands are coordinated to the central magnesium ion rather than via hydrogen bonds. CH₃COO⁻/HCOO⁻ is used to model the Asp/Glu side chain, CH₃OH to represent the Ser/Thr side chain, and OCHNH₂ to represent carbonyls. In total, the set contains 58 magnesium compounds (see Supporting Information).

Metal–ligand distances calculated with DFTB3/3OB deviate only by 0.018 Å on average and 0.125 Å at most (Table 6). The

Table 6. Mean and Maximum Absolute Deviation of Metal–Ligand Lengths and Angles in Comparison to B3LYP/aug-cc-pVTZ for the Magnesium Large Test Set

property ^a	N ^b	B3LYP ^c	PM6	DFTB3/MIO	DFTB3/3OB
r (Å)	329	0.011	0.063	0.041	0.018
r _{max} (Å)		0.033	1.414	0.481	0.125
a (deg)	658	0.7	11.2	4.5	3.1
a _{max} (deg)		17.4	64.0	37.5	30.8

^aMax stands for maximum absolute deviation. ^bNumber of comparisons. ^cBasis set 6-31+G(d,p).

case with the largest error is the Mg–S distance in [Mg(SH₂)₄(SH)]⁻. Besides this outlier, all other metal–ligand length errors are substantially smaller within 0.1 Å. DFTB3/3OB outperforms DFTB3/MIO in the large test set. B3LYP with the small basis set 6-31+G(d,p) deviates on average by 0.011 Å, which is comparable with DFTB3/3OB, and the maximum deviation is 0.033 Å, which is much better than 3OB. PM6 strongly underestimates Mg–O distance by 0.04 Å. More statistics including mean absolute deviations for different

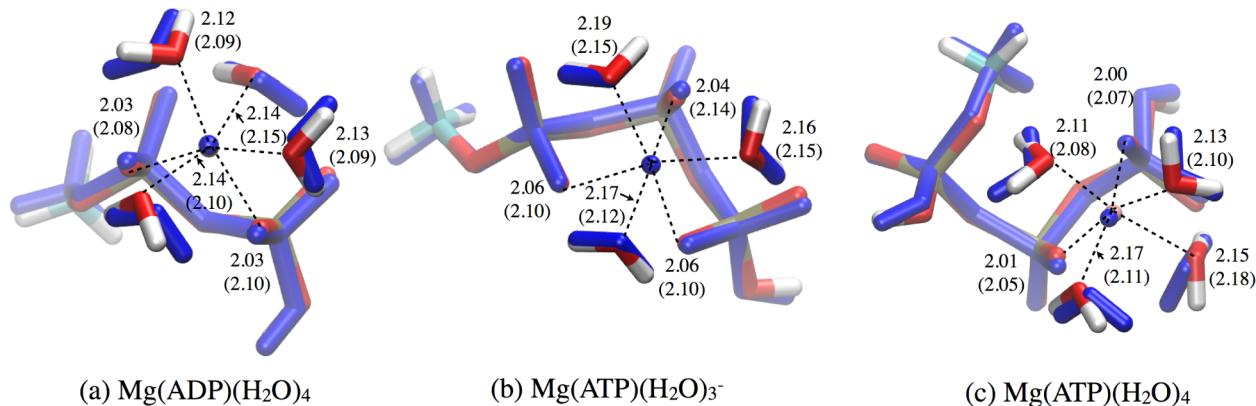


Figure 1. Optimized structures of Mg-AT(D)P-water model molecules at different QM levels. Distances are given in angstroms. The numbers without parentheses were obtained at the B3LYP aug-cc-pVTZ level; those in the parentheses are DFTB3 values. The B3LYP optimized structures are depicted in colors, whereas DFTB3 optimized structures are colored in blue.

Table 7. Error Statistics for the Ligand Dissociation Energies and Proton Affinities of Magnesium Containing Large Molecules Compared to G3B3^a

	B3LYP ^b	PM6	DFTB3/MIO	DFTB3/3OB	G3B3//DFTB3 ^c
Ligand Dissociation Energies					
MAD	5.3	5.1	3.5	4.6	0.8
MAX	7.9	13.0	6.6	7.7	2.2
Proton Affinities					
MAD	1.9	17.2	2.7	2.6	0.9
MAX	4.8	25.0	8.0	6.1	2.3

^aFor detailed data, see Supporting Information. ^bBasis set aug-cc-pVTZ. ^cG3B3 single-point calculations on top of DFTB3/3OB geometries.

Table 8. Mean and Maximum Absolute Deviation of Metal–Ligand Lengths and Angles in Comparison to B3LYP/aug-cc-pVTZ for the Zinc Test Set

property ^a	N ^b	MP2 ^c	B3LYP ^d	PM6	DFTB3/MIO ^e	DFTB3/3OB
r (Å)	163	0.017	0.004	0.065	0.027	0.012
r _{max} (Å)		0.077	0.011	0.553	0.139	0.091
a (deg)	180	0.6	0.2	4.2	3.0	2.3
a _{max} (deg)		4.3	4.2	64.0	16.2	14.6

^aMax stands for maximum absolute deviation. ^bNumber of comparisons. ^cBasis set aug-cc-pVTZ. ^dBasis set 6-31+G(d,p). ^eMolecules containing PH₂⁻ or H⁻ do not converge and are excluded from statistics.

metal–ligand types are given in the Supporting Information. For the angle, the MADs are 3.1, 4.5, 0.7, and 11.2°, respectively, for DFTB3/3OB, DFTB3/MIO, B3LYP/6-31+G-(d,p), and PM6. In the O–Mg–O angles calculated with DFTB3/3OB, about 6% out of 505 angles deviate more than 10°; the large errors stem mainly from erroneous orientation of hydroxide.

Due to the relatively high affinity of ATP/ADP to Mg²⁺ and the importance of this metal chelate in the ATP-binding enzyme, we have constructed model molecules involving Mg and phosphate/model AT(D)P as an additional test (Figure 1). The precise binding mode of Mg²⁺ to ATP remains unclear despite extensive spectroscopic studies. NMR studies proposed several models for the Mg-ATP complex: β-monodentate,¹²⁸ β,γ-bidentate,¹²⁹ α,β,γ-tridentate,^{130,131} and a mixture of α,β-, β,γ-, and α,γ-bidentate.¹³² Raman and infrared spectra suggested a mixture of β,γ-bidentate and α,β,γ-tridentate.¹³³ Here two stable binding modes are tested, one with Mg coordinating β- and γ-phosphates and the other with Mg coordinating α-, β-, and γ-phosphates.¹³⁴ Geometries are described very well by DFTB3/3OB. The Mg–O distance agrees on average within 0.04 Å with B3LYP/aug-cc-pVTZ

calculations. The O–Mg–O, O–P–O, and P–O–P angles differ on average 5, 1, and 7°, respectively.

Ligand dissociation energies are compared in Table 7 (for detailed data, see Supporting Information). One neutral oxygen ligand is removed from the magnesium complexes selected from the large test set. Because the reactions are chosen somewhat arbitrarily, the statistics given in Table 7 have to be considered with care. MADs are reasonably small for all semiempirical methods; the MAD for DFTB3/3OB is 4.6 kcal/mol, even smaller than B3LYP/aug-cc-pVTZ with 5.3 kcal/mol. Interestingly, DFTB3/3OB and B3LYP turn out to systematically underestimate the dissociation energy of the oxygen ligand; the errors seem to be inherited from the sBDE reactions where one water molecule is taken away from [Mg(H₂O)₅]²⁺ or [Mg(H₂O)₆]²⁺. Further fitting U^d or ε_p cannot improve these reaction energies significantly. To demonstrate the good performance of DFTB3 for structural properties, we have carried out single-point G3B3 energies at DFTB3 geometries, and the MAD is only 0.8 kcal/mol.

Proton affinities are also compiled in Table 7. The MAD for DFTB3/3OB is 2.6 kcal/mol, similar to B3LYP/aug-cc-pVTZ with 1.9 kcal/mol. Despite the structural difference for the

Table 9. Sequential Ligand Dissociation Energies of Zinc Containing Molecules Compared to B3LYP/aug-cc-pVTZ^a

molecule ^b	B3LYP ^c	MP2 ^c	B3LYP ^d	PM6	DFTB3/MIO	DFTB3/3OB	B3LYP//DFTB3 ^e
[Zn(H ₂ O) ₂] ²⁺	89.7	-2.9	+2.0	-15.7	-5.2	+4.3	+0.2
[Zn(H ₂ O) ₃] ²⁺	56.4	+1.2	+1.5	+1.4	-1.1	-2.2	-0.1
[Zn(H ₂ O) ₄] ²⁺	43.0	+2.2	+1.9	+3.2	-2.0	-0.9	+0.8
[Zn(H ₂ O) ₅] ²⁺	25.3	+3.2	+2.6	+1.8	-3.8	-1.7	-0.4
[Zn(H ₂ O) ₆] ²⁺	23.4	+3.7	+2.3	-3.2	-3.9	-1.5	+0.3
[Zn(CO) ₂] ²⁺	68.2	+0.4	-0.4	+28.0	-5.8	-3.7	+0.4
[Zn(CO) ₃] ²⁺	40.3	+3.6	+0.1	+15.3	+1.2	-7.3	+0.3
[Zn(CO) ₄] ²⁺	30.6	+4.5	+0.3	+16.9	+3.9	-4.6	-0.2
[Zn(NH ₃) ₂] ²⁺	108.3	-0.6	+1.9	-8.1	-7.8	-4.3	+0.4
[Zn(NH ₃) ₃] ²⁺	59.7	+2.9	+2.4	+8.0	+0.2	-2.6	+0.3
[Zn(NH ₃) ₄] ²⁺	43.8	+3.6	+2.0	+9.4	+4.1	+2.8	+0.1
[Zn(NH ₃) ₅] ²⁺	15.3	+3.5	+2.1	+1.7	-2.7	+3.2	-0.2
[Zn(NH ₃) ₆] ²⁺	13.1	+4.3	+1.8	+1.4	-2.3	+4.1	-0.4
[Zn(SH ₂) ₂] ²⁺	84.4	+0.0	+0.5	-5.8	-2.0	+1.0	-1.4
[Zn(SH ₂) ₃] ²⁺	42.9	+3.9	+0.5	-1.6	+3.5	+1.8	-0.1
[Zn(SH ₂) ₄] ²⁺	29.1	+5.2	+1.2	+0.4	+4.8	+4.9	+2.0
[Zn(PH ₃) ₂] ²⁺	89.4	+1.9	+1.1	+55.7	-7.1	+1.2	+0.7
[Zn(PH ₃) ₃] ²⁺	43.0	+4.9	+0.8	+51.9	+0.6	+3.7	+0.6
[Zn(PH ₃) ₄] ²⁺	30.5	+7.0	+0.9	+65.9	+3.0	+5.5	+0.5
[Zn(NH=CH ₂) ₂] ²⁺	111.9	+0.5	+0.9	-2.2	-1.7	+0.1	+0.1
[Zn(NH=CH ₂) ₃] ²⁺	60.1	+4.4	+1.2	+10.5	+1.3	-1.5	-0.4
[Zn(NH=CH ₂) ₄] ²⁺	43.5	+5.9	+1.1	+9.8	+3.7	+2.8	+0.4
[Zn(SH-CH ₃) ₂] ²⁺	92.2	+2.1	+0.1	-8.2	-5.8	-2.8	-0.3
[Zn(SH-CH ₃) ₃] ²⁺	46.0	+6.4	+0.5	-1.2	+0.5	-0.4	+0.4
[Zn(SH-CH ₃) ₄] ²⁺	31.1	+8.1	+1.1	-0.1	+2.3	+2.9	+0.3
[Zn(OH) ₂]	253.8	+1.9	+2.7	+12.0	-15.1	-3.4	+0.5
[Zn(NH ₂) ₂]	248.6	+4.9	+2.6	+15.9	+11.5	-4.5	+3.1
[Zn(SH) ₂]	213.2	+6.1	+0.2	-22.1	-6.7	-9.1	-0.2
[Zn(PH ₂) ₂]	199.3	+7.1	+0.8	+108.8	f	-11.7	+0.9
[Zn(OCH ₃) ₂]	235.4	+7.6	+2.2	+6.9	-10.0	-0.5	+0.2
[Zn(NCH ₂) ₂]	220.6	+10.9	+0.6	+25.7	+13.8	+8.0	+1.1
[Zn(SCH ₃) ₂]	207.4	+8.2	+0.5	-18.7	-7.8	-7.5	+0.0
MAD		4.2	1.3	16.8	4.7	3.6	0.5
MAX		10.9	2.7	108.8	15.1	11.7	3.1

^aEnergies except PM6 are calculated at 0 K excluding zero-point energy and thermal corrections. All numbers are given in kcal/mol. ^bThe nondissociated is listed. ^cBasis set aug-cc-pVTZ. ^dBasis set 6-31+G(d,p). ^eB3LYP/aug-cc-pVTZ single-point calculations on top of DFTB3/3OB geometries. ^fMolecule does not converge and is excluded from the statistics.

deprotonated structures mentioned above, the G3B3 single-point calculations at DFTB3/3OB geometries show excellent statistics for proton affinities, demonstrating the strength of DFTB3/3OB in reproducing B3LYP/aug-cc-pVTZ geometries or at least reproducing relative energies for multiple points on the potential energy surface. Similar to findings from our previous studies,⁷⁸ PM6 appears to be less reliable for proton affinities with large errors.

Zinc. Geometries. Typical biological ligands for Zn²⁺ include water, cysteine, histidine, and glutamic/aspartic acids. Therefore, we performed calculations for a series of complexes that involve the zinc ion and small molecule models for these ligands. Similar to what has been done for magnesium, we have compiled a test set in the form of Zn²⁺[X]_n with X = H₂O, NH₃, SH₂, PH₃, CO, NH=CH₂, and SH-CH₃ for n = 1–4/6 and X = OH⁻, SH⁻, NH₂⁻, PH₂⁻, NCH₂⁻, and SCH₃⁻ for n = 1–2. Here two extra ligands NH=CH₂ and SH-CH₃ (and their deprotonated form) are added to model histidine and cysteine side chains. The structural comparisons are given in Table 8.

The DFTB3/3OB results are tested against B3LYP/aug-cc-pVTZ for geometries and energetics. Overall good geometries

are obtained from DFTB3/3OB with metal-ligand distance MAD of 0.012 Å and angle MAD of 2.3°, which are notable improvements over DFTB3/MIO. The largest deviation 0.091 Å of metal-ligand length in DFTB3/3OB is for [Zn(N=CH₂)⁺]. The distance between Zn and N is underestimated by ~0.1 Å if r(ZnN) is shorter than 1.95 Å due to the imperfect Zn–N repulsive potential at the short-range. Similar to cases for magnesium, angles Zn–S–H are uniformly shifted up by ~10° in complexes containing ligand SH₂ or SH-CH₃ in DFTB3/3OB optimized structures. B3LYP/6-31+G(d,p) is structurally very similar to B3LYP/aug-cc-pVTZ with a distance MAD of merely 0.004 Å. Nevertheless, MP2/aug-cc-pVTZ is rather different from B3LYP/aug-cc-pVTZ with a MAD of 0.017 Å; metal-ligand distances between zinc and C, N, O, P, S are uniformly underestimated by 0.02–0.03 Å. Molecular geometries are described rather poorly by PM6 in our test set; for example, the metal-ligand lengths between zinc and phosphorus deviate on average 0.3 Å. More statistics for different metal-ligand types are given in the Supporting Information.

Sequential Ligand Dissociation Energies. Sequential ligand dissociation energies for zinc species with neutral and

Table 10. Gas-Phase Proton Affinities of Zinc Containing Molecules Compared to B3LYP/aug-cc-pVTZ^a

molecule ^b	B3LYP ^c	MP2 ^c	B3LYP ^d	PM6	DFTB3/MIO	DFTB3/3OB	B3LYP//DFTB3 ^e
[Zn(H ₂ O) ₁] ²⁺	61.7	+5.4	-0.6	+7.8	-1.2	-7.7	+0.8
[Zn(H ₂ O) ₂] ²⁺	94.2	+0.6	-0.3	-2.5	-1.1	-2.8	+0.5
[Zn(H ₂ O) ₃] ²⁺	121.3	-0.7	-0.5	-9.6	+1.9	-1.2	+0.1
[Zn(H ₂ O) ₄] ²⁺	141.5	-1.6	-0.7	-16.9	+2.3	-1.3	+0.7
[Zn(NH ₃) ₁] ²⁺	86.9	+4.1	+0.1	+6.5	-13.6	-10.8	+1.8
[Zn(NH ₃) ₂] ²⁺	127.3	-0.8	+0.7	-2.9	-18.5	-10.7	+0.6
[Zn(NH ₃) ₃] ²⁺	155.0	-1.7	+1.0	-9.7	-18.0	-10.6	+0.2
[Zn(NH ₃) ₄] ²⁺	178.1	-2.2	+1.0	-4.8	-10.3	-8.4	+0.7
[Zn(SH ₂) ₁] ²⁺	56.6	+1.6	-1.7	+16.1	-13.0	-5.4	-4.0
[Zn(SH ₂) ₂] ²⁺	93.2	-2.0	-1.5	+12.5	-10.6	-0.2	-4.0
[Zn(SH ₂) ₃] ²⁺	115.0	-2.6	-2.0	+7.7	-9.8	-0.5	-4.0
[Zn(SH ₂) ₄] ²⁺	131.1	-3.0	-2.2	+6.0	-9.1	-1.0	-2.1
[Zn(PH ₃) ₁] ²⁺	66.9	+3.2	-0.7	+21.8	f	-1.5	+0.3
[Zn(PH ₃) ₂] ²⁺	115.6	-0.2	-0.3	-5.7	f	-0.9	+0.8
[Zn(PH ₃) ₃] ²⁺	140.8	-1.0	-0.1	-30.0	f	-2.7	+0.7
[Zn(PH ₃) ₄] ²⁺	158.0	-1.6	+0.1	+14.3	-18.1	-3.2	+1.0
[Zn(NH=CH ₂) ₁] ²⁺	89.1	+6.1	-0.9	+2.3	-16.2	-2.8	+1.1
[Zn(NH=CH ₂) ₂] ²⁺	140.9	-0.4	-0.2	-12.6	-14.1	-2.1	+0.4
[Zn(NH=CH ₂) ₃] ²⁺	168.6	-1.5	+0.1	-18.1	-14.2	-2.2	+0.0
[Zn(NH=CH ₂) ₄] ²⁺	188.1	-2.1	-0.1	-14.3	-14.8	-1.6	+2.4
[Zn(SH-CH ₃) ₁] ²⁺	74.2	+0.3	-2.5	+12.1	-4.6	-8.3	-2.5
[Zn(SH-CH ₃) ₂] ²⁺	114.9	-3.0	-2.3	+5.6	-10.0	-4.2	-2.1
[Zn(SH-CH ₃) ₃] ²⁺	136.7	-4.1	-2.4	+1.1	-10.5	-4.8	-1.7
[Zn(SH-CH ₃) ₄] ²⁺	152.6	-6.0	-2.2	+0.3	-2.6	-5.9	-1.4
MAD		2.3	1.0	10.1	10.2	4.2	1.4
MAX		6.1	2.5	30.0	18.5	10.8	4.0

^aEnergies except PM6 are calculated at 0 K excluding zero-point energy and thermal corrections. All numbers are given in kcal/mol. ^bThe protonated form is listed. ^cBasis set aug-cc-pVTZ. ^dBasis set 6-31+G(d,p). ^eB3LYP/aug-cc-pVTZ single-point calculations on top of DFTB3/3OB geometries. fMolecule does not converge and is excluded from the statistics.

negatively charged ligands are summarized in Table 9. DFTB3/3OB shows a MAD of 3.6 kcal/mol compared to B3LYP/aug-cc-pVTZ. Again, deviations close to or greater than 10 kcal/mol are found for charged ligands. The difference between the small and large basis set in B3LYP is rather small, and even charged ligands produce errors no larger than 2.7 kcal/mol, indicating B3LYP with small basis set may be useful in *ab initio* QM/MM calculations. There is a slightly large discrepancy between MP2 and B3LYP; it may be attributed to the differences in structures between these two methods as discussed above. PM6 shows large deviations for sBDE with a MAD of 16.8 kcal/mol. With the unphysical Zn-P geometries removed from statistics, the MAD for PM6 drops to 9.1 kcal/mol, still substantially larger than that for DFTB3. We have also performed single-point calculations of B3LYP/aug-cc-pVTZ on top of DFTB3/3OB geometries. The results are remarkably good with a MAD of only 0.5 kcal/mol.

Proton Affinities. Table 10 shows the protonation energies calculated for 24 biologically relevant ligand model molecules obtained using B3LYP, MP2, and semiempirical methods. As expected, B3LYP/6-31+G(d,p) performs very well compared to B3LYP with a large basis set. The MAD for DFTB3 is 4.2 kcal/mol, which is substantially better than DFTB3/MIO and PM6; the errors are small enough to render this approach useful. High-level single-point correction at DFTB3/3OB geometries once again delivers excellent proton affinities except the ones involving the SH⁻ ligand whose geometries are less well described as discussed above.

Mixed Ligands. One goal to develop DFTB3/3OB is to describe enzyme active sites. Zn-containing structures are

mainly tetrahedrally bound in metalloenzymes. The most abundant primary ones are CCCC, zinc ion with four coordinating cysteines, followed by CCCH, CCHH, CHHH, HHHH, HHHO, HHOO, HOOO, HHHD, and HHDD, where O stands for water and the remaining are 1-letter amino acid codes.¹² In this subsection, we compare geometries and energetics for 43 zinc complexes with mixed types of simple ligands (including the deprotonated form) to model representative active sites in proteins. Here, H₂O is used to represent the Asp/Glu side chain, NH₃ to represent the His side chain, and H₂S to represent the Cys side chain. More complicated/realistic ligands are discussed in the next subsection on “Large Test Set”. Table 11 summarizes the geometrical properties. The MADs in comparison to B3LYP/aug-cc-pVTZ for 161 metal-ligand lengths are 0.005, 0.019, 0.053, and 0.118 Å for B3LYP/6-31+G(d,p), DFTB3/3OB, DFTB3/MIO, and PM6, respectively. Metal-ligand angles on

Table 11. Mean and Maximum Absolute Deviation of Metal-Ligand Lengths and Angles in Comparison to B3LYP/aug-cc-pVTZ for Zinc Mixed Ligand Test Set

property ^a	N ^b	B3LYP ^c	PM6	DFTB3/MIO ^d	DFTB3/3OB
r (Å)	161	0.005	0.118	0.053	0.019
r _{max} (Å)		0.027	0.514	0.179	0.070
a (deg)	231	0.9	6.7	3.0	2.6
a _{max} (deg)		22.2	57.2	21.6	24.7

^amax stands for maximum absolute deviation. ^bnumber of comparisons. ^cbasis set 6-31+G(d,p). ^dMolecules containing PH₂⁻ do not converge and are excluded from statistics.

Table 12. Error Statistics for the Ligand Dissociation Energies and Proton Affinities of Zinc Containing Mixed Ligand Molecules Compared to B3LYP/aug-cc-pVTZ^a

	B3LYP ^b	PM6	DFTB3/MIO	DFTB3/3OB	B3LYP//DFTB3 ^c
Ligand Dissociation Energies					
MAD	1.9	8.2	3.1	3.2	0.9
MAX	2.8	31.6	6.8	6.5	2.6
Proton Affinities					
MAD	0.9	10.4	5.5	3.8	0.9
MAX	1.9	21.7	11.3	7.1	2.4

^aFor detailed data, see Supporting Information. ^bBasis set 6-31+G(d,p). ^cB3LYP/aug-cc-PVTZ single-point calculations on top of DFTB3/3OB geometries.

average deviate by 0.9, 2.6, 3.0, and 6.7°. Again, the 3OB set is a major improvement over the MIO set in metal–ligand lengths. PM6 is not able to give reliable geometries due to a large number of unreasonable metal–ligand distances between zinc and other main group elements.

DFTB3/3OB energetics are predicted with similar accuracy as for complexes with the same type of ligand as shown in Table 12 (for detailed data, see Supporting Information). Once again, the reactions are collected somewhat arbitrarily so the error analysis needs to be treated with care. B3LYP/6-31+G(d,p) gives small error compared to B3LYP with large basis set, demonstrating basis set superposition error is not severe here. DFTB3 gives rather good results in comparison with the reference (MAD of ~3 kcal/mol), especially after single-point energies using B3LYP/aug-cc-pVTZ (MAD of 0.9 kcal/mol), demonstrating that the 3OB set is transferable from one type ligand to complicated coordinating environment with multiple types of ligands.

Large Test Set. To further test the 3OB parameter set, we have chosen larger molecules to model the biologically relevant zinc coordinating environment. In this “large test set”, CH₃COO⁻/HCOO⁻ is used to model the Asp/Glu acid side chain, imidazole/NCH₂CH₃ to represent His, and SHCH₃ to model the Cys side chain. Cys can adopt both protonated and deprotonated form in the active site so that a random mixture of SHCH₃ and SCH₃⁻ is tested in modeling the Cys containing active site. In total we have compiled 51 zinc molecules. Geometries are compared to B3LYP/aug-cc-pVTZ optimized structures. B3LYP with the small basis set deviates on average by only 0.005 Å, and the maximal deviation is only 0.063 Å (Table 13). The MAD for metal–ligand length from DFTB3/

estimated by 0.141 and 0.103 Å, respectively. All others errors are smaller than 0.1 Å. The largest angle deviation 25.8° occurs in angle Zn–N–C in [Zn(C₃N₂H₃)]⁺ (i.e., Zn²⁺ bound to a single deprotonated imidazole), which is not encountered often in biological systems.

Ligand dissociation energies are compiled in Table 14 (for detailed data, see Supporting Information). DFTB3 shows that removal of neutral imidazole may cause an error of more than 10 kcal/mol, indicating interactions between Zn–sp₂ hybridized N are problematic. Nevertheless, when calculating single-point energies using B3LYP/aug-cc-pVTZ the errors turn out to be small with the MAD of merely 0.5 kcal/mol. Proton affinities are also compared in Table 14. DFTB3/3OB shows excellent results with a MAD of 3.3 kcal/mol, which is fairly remarkable in this set of rather complicated molecules; single-point B3LYP calculations lead to an even smaller MAD of 0.6 kcal/mol.

Dinuclear Mg/Zn Active Sites. An important catalytic motif in metalloenzymes is the bimetallic (or dinuclear) Mg/Zn site.^{135,136} In this bimetallic motif, metal ions are commonly coordinated through a monatomic bridging ligand like water/hydroxide or a polyatomic group like carboxylate in a bidentate fashion. The two-metal-ion mechanism⁵ was proposed for this type of motif, where one metal ion actively participates in facilitating the dissociation of the leaving group and stabilizing the pentavalent intermediate while the second metal ion holds an important structural role during catalysis. The advantage of bimetallic centers is the charge delocalization so as to lower the barrier of binding large substrates and generating a nucleophile.¹³⁷ To assess DFTB3/3OB for these important catalytic motifs, we have constructed four model compounds based on bimetallic Mg/Zn centers in enzymes (Figure 2). In our models, certain ligands are protonated to prevent a high net charge in the gas phase.

Figure 2(a) is a simplified active site structure taken from pyruvate kinase crystal structure (PDB ID 1A49¹³⁸). Pyruvate kinase is an enzyme of the glycolytic pathway and phosphorylates ADP to generate ATP and pyruvate. This enzyme can also utilize other divalent metal ions like Mn²⁺, Ni²⁺, and Zn²⁺. Figure 2(b) is simplified from the polymerase η -DNA-dATP complex crystal structure (PDB ID 4ECQ¹³⁹) which is involved in DNA synthesis. The formation of a phosphodiester bond is recently captured by time-resolved X-ray crystallography.¹³⁹ In the (a)(b) active sites, one Mg²⁺ is tridentated to α -, β -, and γ -oxygens in ATP. Figure 2(c) is modified from the type II topoisomerase (topoII)-DNA cleavage complex crystal structure (PDB ID 3L4K¹⁴⁰). This enzyme cleaves and ligates DNA using tyrosine as the nucleophilic agent. Following the nucleophilic attack toward DNA backbones, a covalent phosphotyrosyl bond that links

Table 13. Mean and Maximum Absolute Deviation of Metal–Ligand Lengths and Angles in Comparison to B3LYP/aug-cc-pVTZ for the Zinc Large Test Set

property ^a	N ^b	B3LYP ^c	PM6	DFTB3/MIO	DFTB3/3OB
r (Å)	173	0.005	0.128	0.053	0.024
r _{max} (Å)		0.063	0.955	0.177	0.141
a (deg)	236	0.6	5.3	4.0	3.3
a _{max} (deg)		5.1	63.3	25.8	25.8

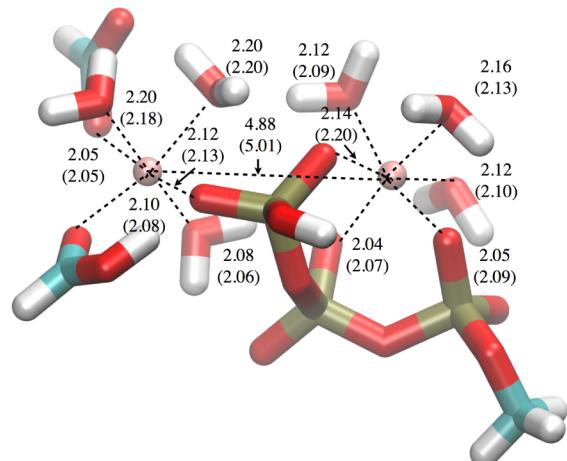
^aMax stands for maximum absolute deviation. ^bNumber of comparisons. ^cBasis set 6-31+G(d,p).

3OB optimized structures is 0.024 Å. MAD of metal–ligand lengths Zn–O is 0.033 Å and Zn–S is 0.027 Å; they appear to be worse than the small test set but are good enough for modeling enzyme active sites. The largest deviations for metal–ligand lengths are found in Zn–O distances in [Zn(Im)(H₂O)(SCH₃)₂] and [Zn(H₂O)(SHCH₃)(SCH₃)₂], which are under-

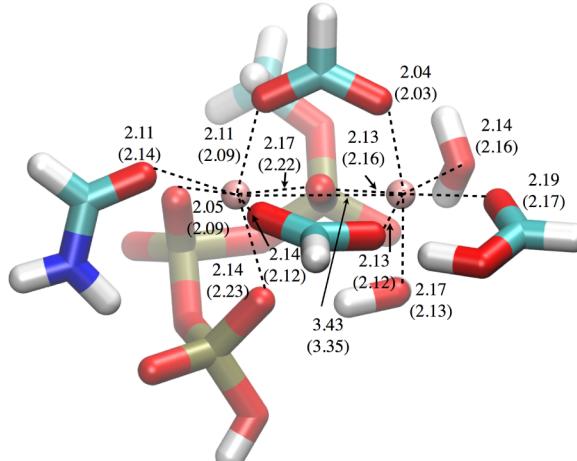
Table 14. Error Statistics for the Ligand Dissociation Energies and Proton Affinities of Zinc Containing Large Molecules Compared to B3LYP/aug-cc-pVTZ^a

	B3LYP ^b	PM6	DFTB3/MIO	DFTB3/3OB	B3LYP//DFTB3 ^c
Ligand Dissociation Energies					
MAD	1.3	5.9	3.6	3.4	0.5
MAX	2.6	20.2	14.4	13.8	1.3
Proton Affinities					
MAD	1.1	10.1	4.9	3.3	0.6
MAX	2.1	19.3	12.7	6.9	2.6

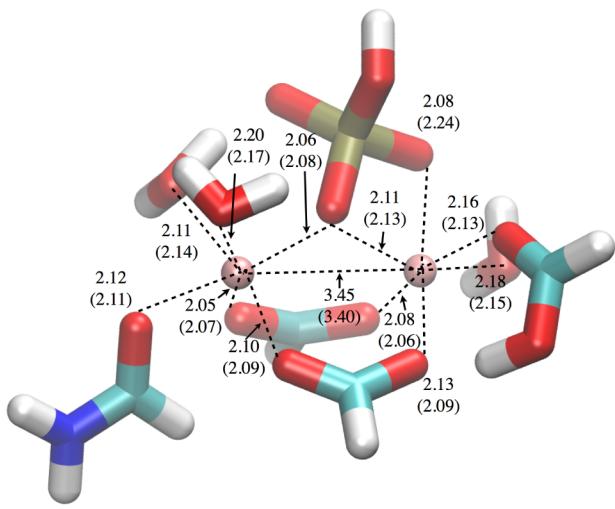
^aFor detailed data, see Supporting Information. ^bBasis set 6-31+G(d,p). ^cB3LYP/aug-cc-PVTZ single-point calculations on top of DFTB3/3OB geometries.



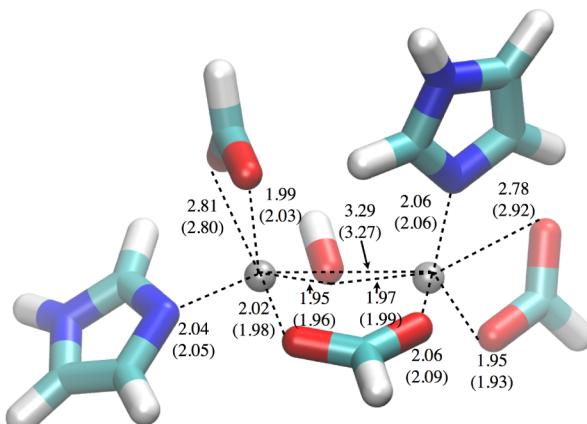
(a) $\text{Mg}_2(\text{ATP})(\text{H}_2\text{O})_6(\text{HCOO})_2$



(b) $\text{Mg}_2(\text{ATP})(\text{H}_2\text{O})_2(\text{HCOO})_3(\text{OCHNH}_2)^-$



(c) $\text{Mg}_2(\text{H}_2\text{O})_3(\text{HCOO})_3(\text{OCHNH}_2)_2(\text{Phos})$



(d) $\text{Zn}_2(\text{Im})_2(\text{OH})(\text{HCOO})_3$

Figure 2. Optimized structures of dinuclear $\text{Mg}^{2+}/\text{Zn}^{2+}$ complexes of the simplified protein active site (a) PDB ID 1A49; (b) PDB ID 4ECQ; (c) PDB ID 3L4K; (d) PDB ID 1AMP. Distance are given in angstroms. The numbers without parentheses are obtained at the B3LYP 6-31+G(d,p) level; those with parentheses are DFTB3 values.

topoII and partial DNA is formed. It was long believed that a catalytic reaction could be aided via a single Mg^{2+} . More recently, it was found that the canonical two-metal-ion mechanism may be possible.^{140,141} Therefore, two Mg^{2+} ions are modeled in the reactant state. Figure 2(d) is the active site in *Aeromonas Protelytica Aminopeptidase* (AAP) (PDB ID 1AMP¹⁴²), which is a prototypical member of the bizinc enzyme family.

As shown in Figure 2, DFTB3 is tested against B3LYP/6-31+G(d,p) optimized structures. The overall structures from DFTB3 reproduce B3LYP structures very well, particularly the metal ion–ligand distances. Mg–Mg/Zn–Zn distances are also excellent with the largest deviation of 0.13 Å. The monodentate or bridging binding modes of carboxylate are correctly described by the 3OB parameter set in comparison to B3LYP. These calculations demonstrate that DFTB3/3OB is

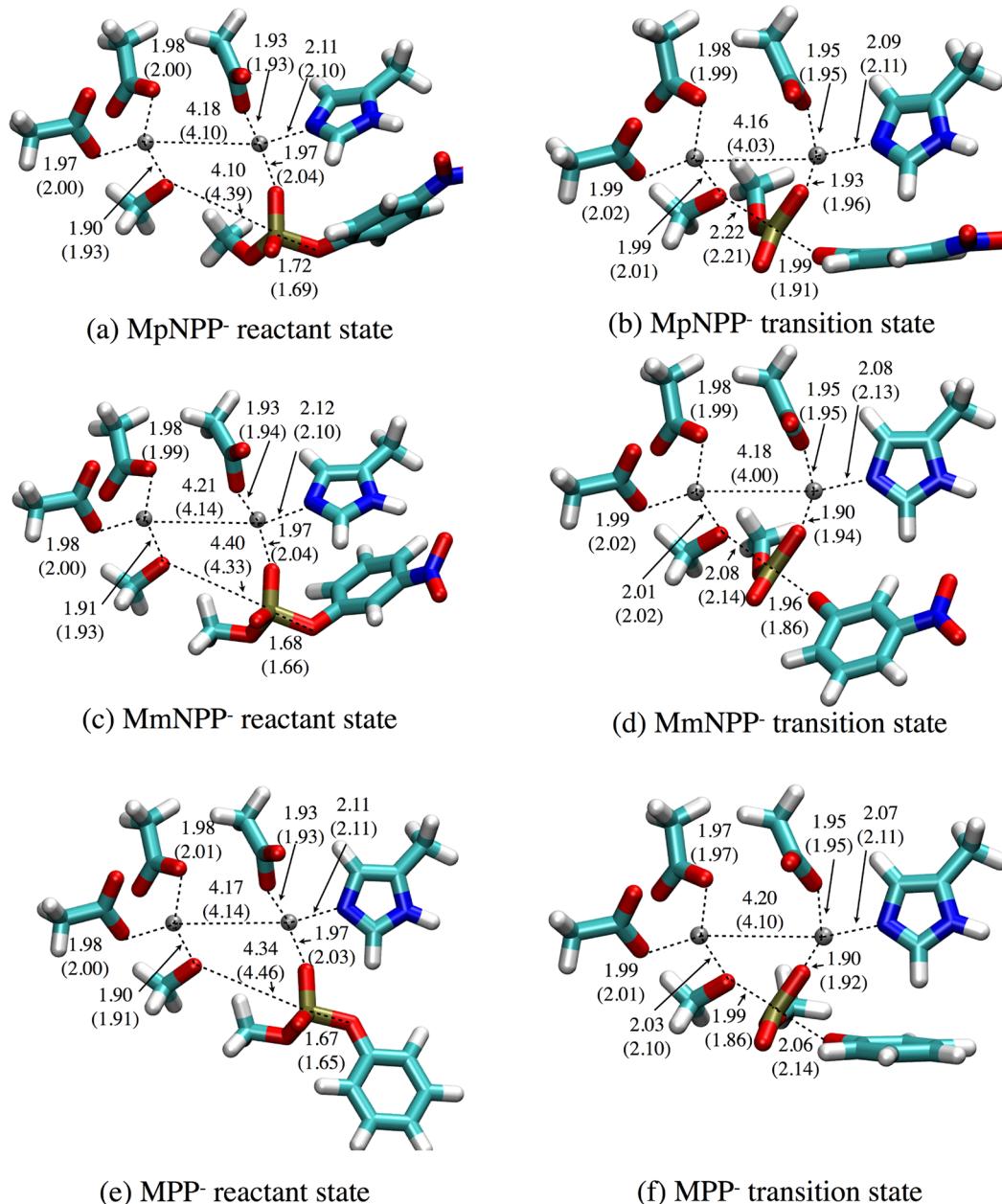


Figure 3. AP active site gas-phase model with MpNPP⁻, MmNPP⁻, and MPP⁻. Distances are given in angstroms. The numbers without parentheses are obtained at the B3LYP 6-31+G(d,p) level; those in the parentheses are DFTB3 values. (a) MpNPP⁻ reactant state; (b) MpNPP⁻ transition state; (c) MmNPP⁻ reactant state; (d) MmNPP⁻ transition state; (e) MPP⁻ reactant state; and (f) MPP⁻ transition state.

useful for modeling bimetallic centers in enzymes with good accuracy.

To further benchmark DFTB3/3OB in describing the transition state and reaction energy barrier, we study an active site model of AP in the gas phase. AP catalyzes the hydrolytic reactions of various phosphates via a two-step mechanism: an oxygen neocleophile first attacks the phosphorus, then a water (hydroxide) replaces the leaving group in the subsequent step.¹⁴³ In our previous work,⁴⁸ we studied the first step for the hydrolysis of a well-studied phosphate diester MpNPP⁻ in R166S AP and showed that the transition states are synchronous in nature. To help understand the intrinsic errors in DFTB3/3OB for the reactions of interest, we investigate the active site model in the gas phase with QM calculations. Specifically we study the

same reaction for two additional phosphate diesters methyl 3-nitrophenyl phosphate (MmNPP⁻) and methyl phenyl phosphate (MPP⁻) and compare the energetics and nature of the transition state to MpNPP⁻. For simplicity, we only study the α orientation of different substrates.

As shown in Figure 3, there is generally good agreement between DFTB3/3OB and B3LYP/6-31+G(d,p) results in terms of geometries; the P–O distances, Zn²⁺–Zn²⁺ distances, and coordination structures are similar at the two QM computational levels for both reactant and transition states. The Zn–O distance error is within 0.03 Å and is a major improvement over MIO parameters which systematically overestimate Zn–O distance by 0.1 Å (MIO data not shown here). The Zn²⁺–Zn²⁺ distance is underestimated in DFTB3/3OB by ~0.1 Å in the reactant and is shrunk ~0.1 Å more in

Table 15. Calculated Activation Barrier (in kcal/mol) for Diester Hydrolysis in the Active Site Model of AP in the Gas Phase^a

substrate	DFTB3	single points at DFTB3 geometries				single points at B3LYP geometries			
		B3LYP	M06	PBE	MP2	B3LYP	M06	PBE	MP2
MpNPP ⁻	8.8	9.2 (4.2 ^b)	5.7 (5.1)	5.7 (3.1)	7.6	12.0 (6.7)	6.8 (6.1)	7.2 (4.3)	8.8
MmNPP ⁻	9.7	9.2 (9.4)	12.1 (12.4)	6.8 (7.5)	13.6	15.3 (12.0)	13.2 (12.8)	11.0 (9.3)	13.2
MPP ⁻	20.0	16.2 (9.1)	12.4 (10.5)	11.6 (7.1)	9.6	20.1 (11.6)	13.0 (10.8)	14.0 (8.6)	10.4

^aThe geometries are optimized at either the DFTB3/3OB or B3LYP/6-31+G(d,p) level. Single-point energy calculations are carried out at the B3LYP, M06, PBE, and MP2 method at two levels of geometries using a large basis set of 6-311++G(d,p). ^bThe numbers with parentheses are obtained with corresponding DFT functional plus D3 dispersion.

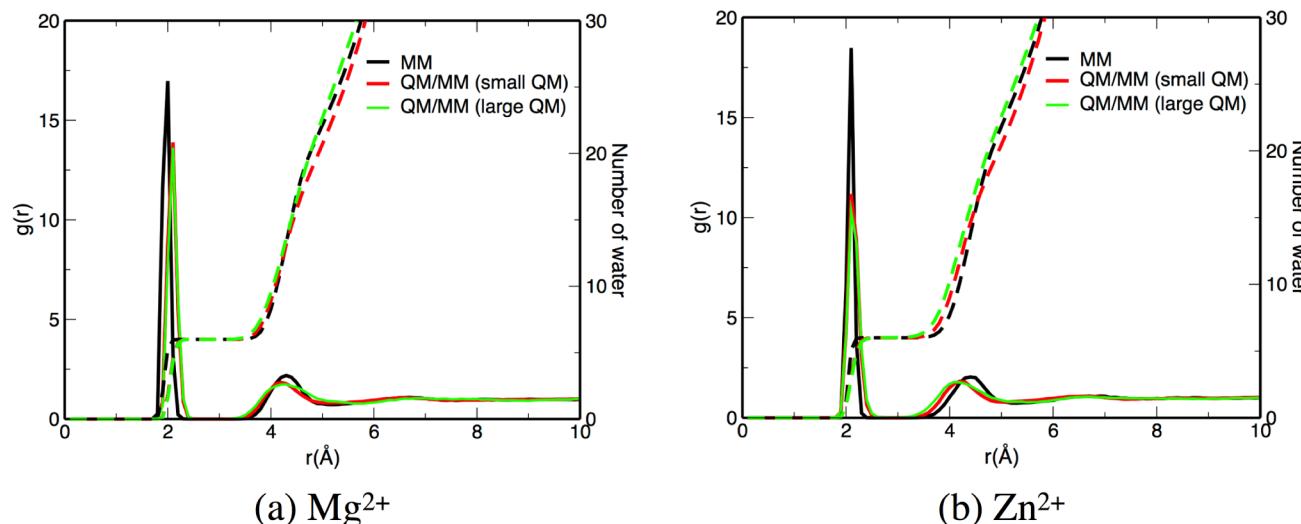


Figure 4. Radial distribution function of water oxygen around $\text{Mg}^{2+}/\text{Zn}^{2+}$ in 20 Å water droplet. The pure MM model is based on the CHARMM force field.¹⁰⁵

the transition state in comparison to B3LYP distances. In terms of the nature of the transition state, both DFTB3 and B3LYP predict a synchronous type of structure with the tightness coordinate (i.e., $\text{P}-\text{O}^{\text{ligand}} + \text{P}-\text{O}^{\text{nucleophile}}$) of 4.0–4.2 Å even though the individual P–O bond is overestimated or underestimated by 0.1 Å in DFTB3, indicating DFTB3 provides a good description for the potential energy surface around the transition state for the three substrates.

Regarding the reaction energy barrier, Table 15 shows that the qualitative trend is consistent among various computational levels although there are substantial differences in the quantitative values. Previous calculations⁴⁸ of proton affinities for the leaving group indicate that MP2 gives results close to the experimental values, so MP2/6-311++G(d,p) single-point energies are used as the reference here. Numbers span a broad range with different DFT methods without dispersion, and the results become much closer after adding the D3 dispersion; an explicit consideration of dispersion is important for a large active site discussed here. Compared to MP2, DFTB3 gives a reasonable reaction energy for MpNPP⁻ but underestimates the barrier of MmNPP⁻ by ~4 kcal/mol and overestimates the barrier of MPP⁻ by ~9 kcal/mol. However, MP2 single-point energies on top of DFTB3 and B3LYP/6-31+G(d,p) geometries differ by ~1 kcal/mol for all the substrates, indicating DFTB3 produces satisfactory geometries and also highlighting the importance of single-point energy correction with a reliable *ab initio* method at DFTB3 geometries. DFTB3 reaction barriers are close to B3LYP results without D3 correction; this is not surprising because DFTB3 is parametrized based on B3LYP (see Methods) and inherits some errors. M06 behaves

the best among three DFT functionals as compared to MP2, especially with the D3 dispersion included. PBE systematically underestimates the energy barrier. Further optimization of the D3 dispersion model developed for DFTB3⁹⁹ will further improve DFTB3/3OB energies.

Mg²⁺ and Zn²⁺ Ions Solvation Structure and Relative pK_a. The metal ion-O radial distribution function ($g(r)$) and coordination numbers are shown in Figure 4 for the two cations in solution. A well-defined first solvation shell is observed in both cases. The first peak in DFTB3/MM hydrated Mg^{2+} occurs at 2.07 Å, in good agreement with X-ray diffraction data of $2.09 \text{ \AA} \pm 0.04 \text{ \AA}$ ¹⁴⁴ and with previous AIMD simulations in the range of 2.08–2.13 Å.^{145–149} The MM force field predicts the Mg^{2+} –O distance to be shorter than the DFTB3/MM model with a sharper peak at 2.0 Å. The inner coordination sphere remains similar to the gas-phase structure. For the cluster $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ in the gas phase, the Mg^{2+} –O distance is 2.10 Å in both DFTB3 and B3LYP/aug-cc-pVTZ optimized structures. For Zn^{2+} , the first maximum from DFTB3/MM is at 2.11 Å, consistent with X-ray absorption measurements $2.06 \pm 0.02 \text{ \AA}$ ¹⁵⁰ and earlier AIMD results 2.07–2.13 Å.^{145,146,151} The classical MM model gives the first peak at 2.08 Å, although the peak is higher and narrower than in the DFTB3/MM model. The small difference in the location of the first peak in DFTB3/MM may be traced back to the gas-phase model: for the cluster $[\text{Zn}(\text{H}_2\text{O})_6]^{2+}$ in the gas phase, Zn^{2+} –O is 2.16 Å in DFTB3 versus 2.12 Å in B3LYP/aug-cc-pVTZ.

The integral of the first solvation shell for both cations gives the coordination number of 6. The large QM region results predict very similar first solvation shell as with the small QM

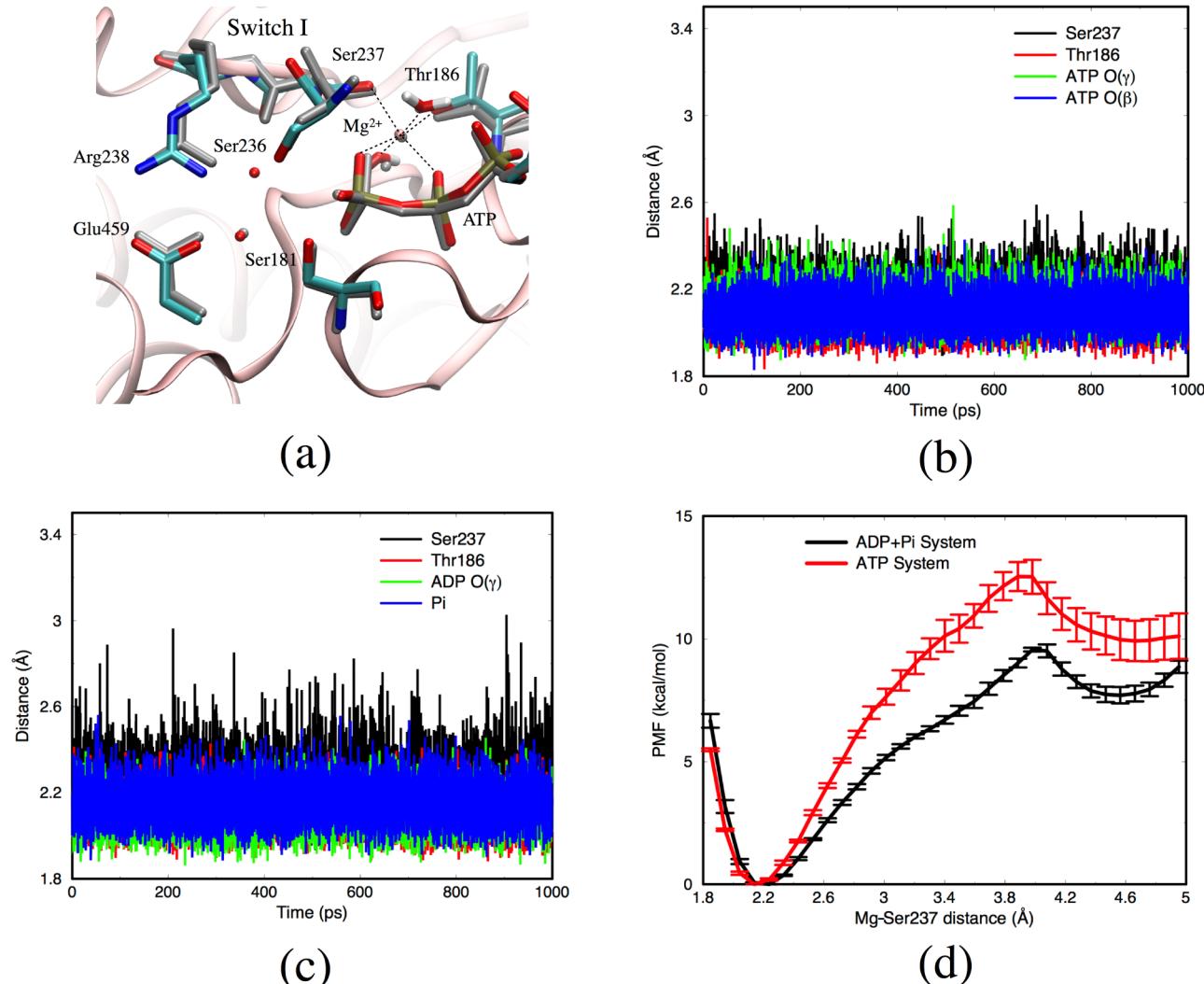


Figure 5. Structural properties of the Myosin active site during equilibrium QM/MM MD simulations with different nucleotide chemical states (ATP or ADP-Pi). (a) Overlay of average DFTB3/MM structure and crystal structure (gray) of Myosin-II motor domain active site in the ATP state. The VO₄ moiety in the crystal structure (PDB code 1VOM) is replaced by a PO₃ phosphate group and a presumptive lytic water molecule. Instantaneous distances between Mg²⁺ and oxygen atoms in its four nonwater ligands at (b) ATP state and (c) ADP-Pi state. Ser237 refers to O^{Ser237} and Thr186 refers to O^{Thr186}. (d) PMF calculations for the ATP/ADP-Pi states. The reaction coordinate is the distance between Mg²⁺ and O^{Ser237}.

region, though they differ slightly for the second solvation shell; the large QM region leads to a broader peak for the second solvation shell than the small QM region and MM model. It is worth noting that in the large QM region simulation of Zn²⁺ water exchanges between the first and second solvation shells (coordination number remains the same), whereas no exchange of water is seen in Mg²⁺ simulations. These observations are in agreement with experimental findings; for Mg²⁺, experimental results suggest that water molecules could reside in the first shell up to a few microseconds,^{152,153} while the water residence time for Zn²⁺ is only hundreds of picoseconds.¹⁵⁴

To further benchmark the energetics of DFTB3/3OB, we use the DTSC-TI approach to calculate the relative pK_a of Mg²⁺/Zn²⁺-bound water with the DFTB3/MM model. Altogether six intermediate states are used to obtain the $\partial F/\partial \lambda$ values. In each window, the data collected during the first 250 ps are discarded as part of the equilibration process, and the remaining 500 ps data are then subject to statistical analysis to obtain the proper average of free energy derivative. The statistical errors associated with the free energy are on the order

of 0.2–0.5 kcal/mol. In all cases, the linear dependence of the free energy derivatives on λ holds well with a correlation coefficient (R^2) higher than 0.99 (see Figure S1 in Supporting Information). Integrating $\partial F/\partial \lambda$ over λ gives a value of 132.9 kcal/mol for Mg²⁺ with the small QM region, 124.1 kcal/mol for Zn²⁺ with the small QM region, 130.0 kcal/mol for Mg²⁺ with the large QM region, and 125.5 kcal/mol for Zn²⁺ with the large QM region, leading to a ΔpK_a of 6.4 pK_a units with the small QM region and 3.3 pK_a units with the large QM region. The latter is in good agreement with the experimental value of 3.2 pK_a units (12.4 for a Mg²⁺ solution and 9.2 for hydrated Zn²⁺).^{145,155} The better performance of the large QM region results highlights the importance of using a QM description for the second solvation shell, which allows coordination number fluctuation in the Zn²⁺ solution, especially at the deprotonated state. We have also calculated the absolute pK_a of Mg²⁺/Zn²⁺-bound water as a further benchmark. Compared to the experimental values, the closest computational estimate from current work is 14.0 for Mg²⁺ and 10.4 for Zn²⁺ (see Supporting Information).

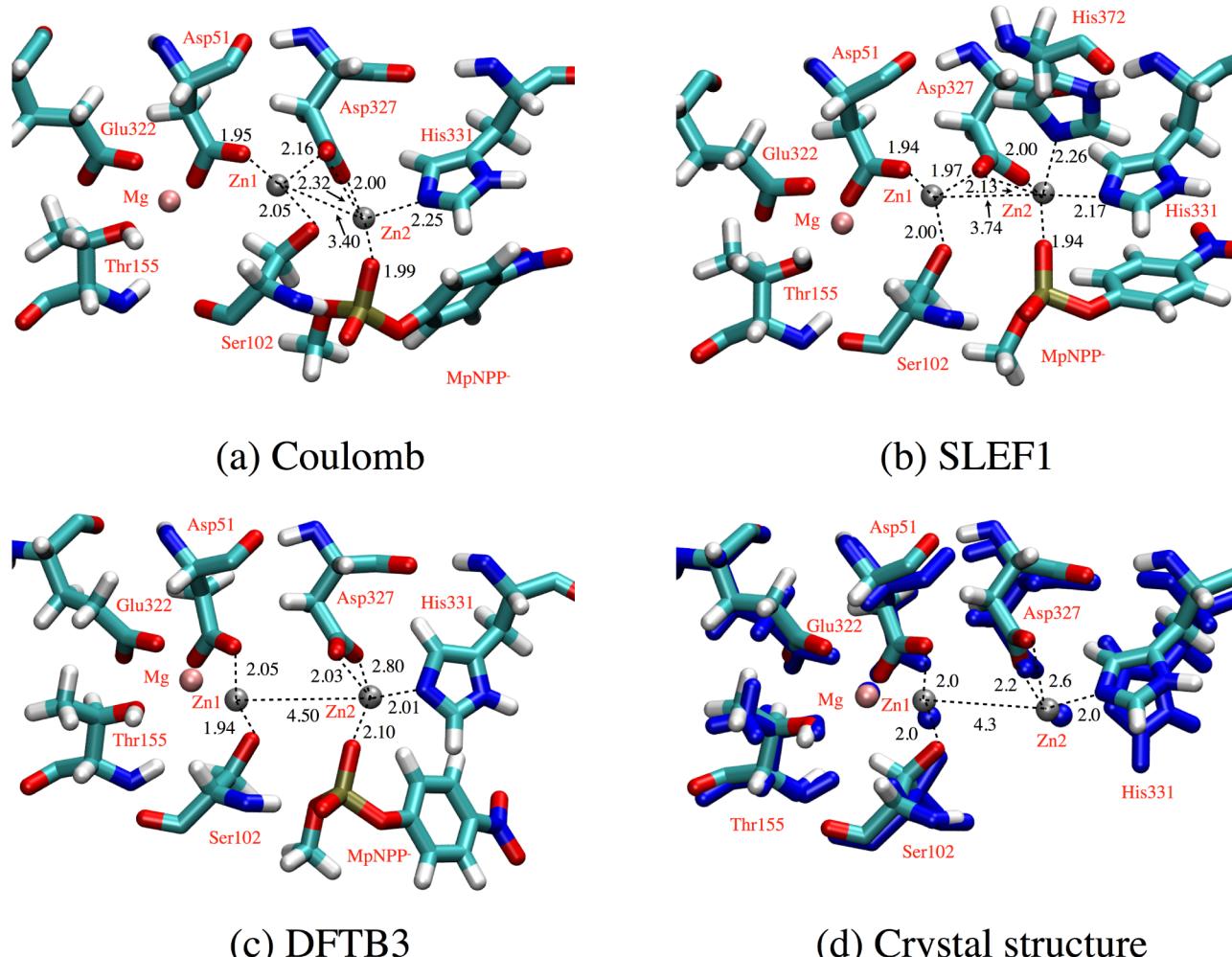


Figure 6. Structural properties of the AP active site during equilibrium QM/MM MD Simulations. A snapshot for the reactant state with (a) Coulomb MM force field, (b) SLEF1 zinc MM force field, (c) DFTB3/3OB, and (d) overlay of crystal and DFTB3/MM structure (blue), with average key distances in angstroms labeled. Asp369, His370, and His412 are omitted for clarity. Substrate inorganic phosphate in the crystal structure (3CMR) is not shown.

Structural Flexibility of the Active Site in Myosin. Myosin-II (hereafter referred to as myosin) is one of the best characterized molecular motors, which couple large-scale conformation transitions with ATP binding and hydrolysis. Despite many previous efforts from both experimental and theoretical studies,^{156–160} the complex nature of “mechanochemical coupling” still remains to be well understood.

As shown in Figure 5(a), binding modes obtained from the DFTB3/MM simulation reproduce the key features of the crystal structure. In the discussion of our earlier work,¹¹⁵ the weakened interaction between Mg^{2+} -nucleotide and Ser237 as reflected by the larger distance fluctuation was proposed to indicate the destabilization of Switch I following ATP hydrolysis. As a test of the DFTB3/3OB magnesium parameters in the enzyme, DFTB3/MM equilibrium simulations are carried out with two nucleotide chemical states (ATP or ADP·Pi), and the distances between Mg^{2+} and four nonwater ligands are monitored. As shown in Figures 5(b) and (c), the distances between Mg^{2+} and ligands are reasonably stable over 1 ns; the most obvious difference observed between the two nucleotide states is the fluctuation of Mg^{2+} –O^{Ser237} distance, for which the average and fluctuation are 2.16 and 0.08 Å in the ATP state and 2.24 and 0.11 Å in the ADP·Pi

state. The magnitude of fluctuation observed in current DFTB3/MM simulations is notably less compared with previous work (MM simulations and SCC-DFTBPR/MM simulation, in which the Mg^{2+} –O^{Ser237} distance reached ~3.5 Å), which led to coordination number shifting from 6 to 5 after ATP hydrolysis.¹¹⁵ To further evaluate the situation, PMFs are calculated for the distance between Mg^{2+} and O^{Ser237} in the ATP and ADP·Pi states. As seen in Figure 5(d), dissociation of Ser237 is indeed energetically more favorable in the ADP·Pi state than the ATP state by a few kcal/mol. Nevertheless, there is a significant energetic cost to fully dissociate the Ser237 side chain from the Mg^{2+} , even in the ADP·Pi state. Further experimental analysis based on infrared spectroscopy can potentially generate additional insights into this issue.¹⁶¹

Force Fields Comparison in AP Active Site. We have benchmarked DFTB3/3OB for the AP active site with a gas-phase model in earlier discussion. As an additional test, DFTB3/MM simulations for the R166S AP enzyme are carried out, and the results are compared to two popular MM force fields for zinc. The comparison of equilibrated reactant structures by Coulomb,¹⁴ SLEF1,²⁵ and DFTB3/MM is shown in Figure 6, with the crystal structure with bound phosphate as a reference. Binding modes obtained from the

DFTB3/MM simulation reproduce the key features of the crystal structure and previous SCC-DFTBPR/MM simulations⁴⁸ very well (shown in Figure 6(c) and (d)). The conventional Coulomb scheme (Figure 6(a)) leads to a much shortened Zn²⁺–Zn²⁺ distance and significant changes in the active site geometry: Ser102 and Asp327 are stabilized by two zinc ions simultaneously; Asp369 (not shown in Figure 6 for clarity) is bidentate with the Zn²⁺ in the Coulomb model but monodentate in the crystal and DFTB3/MM structures. By comparison, coordination geometry is notably improved by the SLEF1 force field (Figure 6(b)), even though the Zn²⁺–Zn²⁺ (3.7 Å) distance is still substantially shorter than that in the crystal structure (4.3 Å). It is perhaps not surprising that SLEF1 does not provide a correct distance between zinc ions as SLEF1 was not parametrized for the binuclear zinc motif and designed to be compatible with the Amber99SB force field rather than with the CHARMM force field. Nevertheless, SLEF1 correctly describes the binding modes of Ser102 and Asp369 to the zinc ions; yet Asp327 still has a strong preference to bridge the two zinc cations due to the short distance between them. Moreover, an extra His372 is pulled closer to bind to Zn²⁺; this might have been caused by the stronger attractive interaction between the zinc ion and nitrogen in the short-range part of the SLEF1 force field. In short, DFTB3/MM produces encouraging results for modeling a challenging bimetallic zinc active site, demonstrating the greater applicability of DFTB3/MM to metalloenzymes even when only structural properties are considered.

CONCLUDING REMARKS

Magnesium and zinc play many important roles in metalloenzymes and various chemical systems. An efficient and accurate QM description for Mg/Zn is critical for the use of theoretical methods to study the reaction mechanism of these systems. In this work, we extend the parametrization of the approximate density functional tight binding theory, DFTB3, to magnesium and zinc. The parametrization is performed in a framework consistent with the DFTB3/3OB set for CHNOPS; hence this parameter set can be used to study the chemistry of Mg/Zn moieties in solution and biological systems.

Benchmark calculations in the gas phase and in the condensed phase with a QM/MM framework demonstrate that the current parametrization generally leads to reliable structures and semiquantitative energetics (with typically a MAD of ~3–5 kcal/mol compared to high level *ab initio* calculations). In the gas phase, DFTB3/3OB are tested with a diverse range of molecules that reflect the typical coordination environment of magnesium and zinc in biomolecules, including dinuclear metal sites. We focus on the geometries, ligand dissociation energies, and ligand proton affinities. The results calculated at the DFTB3/3OB level are compared to B3LYP, *ab initio* (G3B3, MP2), and a popular semiempirical method PM6 as well as the older parametrization, MIO. In general, DFTB3/3OB shows substantial improvement over DFTB3/MIO and outperforms PM6 in many aspects within our test sets, particularly in terms of structural properties and ligand proton affinities. Compared to high-level calculations, DFTB3/3OB is very successful at predicting geometries. Large errors can still be found in certain cases, especially with charged ligands. Even for these cases, single-point energy calculations with high level QM methods generally give very reliable energetics, suggesting that changes of structures during typical

processes of interest (e.g., ligation dissociation or ligand deprotonation) are well captured with DFTB3/3OB.

Due to computational efficiency and general robustness for geometry predictions, DFTB3/MM calculations are expected to be effective for studies in condensed-phase systems. For example, our solution benchmark study for the magnesium and zinc ion has reproduced experimental solvation structures and relative pK_a of metal bound water, especially when both the first and second solvation shells are treated as DFTB3/3OB. For the enzyme benchmark, DFTB3/3OB is successful at producing active site structures in myosin and AP relative to available crystal structures; for AP, DFTB3/3OB also leads to semi-quantitative energetics for an active site model, although our calculations using several DFT functionals and MP2 highlight the importance of careful calibration of energetics for large metal sites in enzymes, including consideration of dispersion interactions. These studies have laid the groundwork for using DFTB3/MM simulations with state-of-the-art sampling techniques such as metadynamics¹⁶² and finite-temperature string methods^{163–165} to study the chemical processes in myosin and AP; high-level single-point QM/MM calculations are still required for improving the quantitative nature of the free energy surfaces. Finally, we emphasize that there are still remaining limitations in DFTB3, such as the treatment of interaction between Mg/Zn and highly charged/polarizable ligands (e.g., hydroxide and SH[−]), thus it is essential to continue the formal development of DFTB3, such as adding multipole terms for the charge fluctuations and formulating a better description of polarization and short-range Pauli repulsions.

ASSOCIATED CONTENT

S Supporting Information

Additional details regarding molecular geometries and statistics for benchmark systems are included; absolute pK_a calculations of Mg²⁺/Zn²⁺-bound water are also discussed. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: cui@chem.wisc.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We dedicate this work to the special issue that honors Professor Bill Jorgensen, who has made numerous contributions to the simulation of condensed-phase systems, including his pioneering work in QM/MM simulations and development of semiempirical QM methods. The work has been supported by NIH grant R01-GM106443. Computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by NSF grant number OCI-1053575, are greatly appreciated; computations are also partly supported by the National Science Foundation through a major instrument grant (CHE-0840494).

REFERENCES

- (1) Lippard, S. J.; Berg, J. M. *Principles of Bioinorganic Chemistry*; University Science Books: Mill Valley, CA, 1994.
- (2) Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. Metal Ions in Biological Catalysis: From Enzyme Databases to General Principles. *J. Biol. Inorg. Chem.* 2008, 13, 1205–1218.

- (3) Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J. Metalloproteins and Metal Sensing. *Nature* **2009**, *460*, 823–830.
- (4) Bergman, C.; Kashiwaya, Y.; Veech, R. L. The Effect of pH and Free Mg²⁺ on ATP Linked Enzymes and the Calculation of Gibbs Free Energy of ATP Hydrolysis. *J. Phys. Chem. B* **2012**, *114*, 16137–16146.
- (5) Steitz, T.; Steitz, J. A General 2-Metal-Ion Mechanism For Catalytic RNA. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6498–6502.
- (6) Parkin, G. Synthetic Analogues Relevant to the Structure and Function of Zinc Enzymes. *Chem. Rev.* **2004**, *104*, 699–768.
- (7) Maret, W.; Li, Y. Coordination Dynamics of Zinc in Proteins. *Chem. Rev.* **2009**, *109*, 4682–4707.
- (8) Lipscomb, W. N.; Strater, N. Recent Advances in Zinc Enzymology. *Chem. Rev.* **1996**, *96*, 2375–2433.
- (9) Auld, D. S. Zinc Coordination Sphere in Biochemical Zinc Sites. *BioMetals* **2001**, *14*, 271–313.
- (10) Lee, Y.-M.; Lim, C. Physical Basis of Structural and Catalytic Zn-Binding Sites in Proteins. *J. Mol. Biol.* **2008**, *379*, 545–553.
- (11) Hoops, S. C.; Anderson, K. W.; Merz, K. M., Jr. Force-Field Design for Metalloproteins. *J. Am. Chem. Soc.* **1991**, *113*, 8262–8270.
- (12) Peters, M. B.; Yang, Y.; Wang, B.; Füsti-Molnár, L.; Weaver, M. N.; Merz, 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.
- (13) Babu, C. S.; Lim, C. Empirical Force Fields for Biologically Active Divalent Metal Cations in Water. *J. Phys. Chem. A* **2006**, *110*, 691–699.
- (14) Stote, R. H.; Karplus, M. Zinc-Binding in Proteins and Solution - A Simple but Accurate Nonbonded Representation. *Proteins* **1995**, *23*, 12–31.
- (15) Duarte, F.; Bauer, P.; Barrozo, A.; Amrein, B. A.; Purg, M.; Åqvist, J.; Kamerlin, S. C. L. Force Field Independent Metal Parameters Using a Nonbonded Dummy Model. *J. Phys. Chem. B* **2014**, *118*, 4351–4362.
- (16) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand-Macromolecule Complexes. A Bottom-Up Strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960–1986.
- (17) 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.
- (18) 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.
- (19) Yu, H. B.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; MacKerell, A. D., Jr.; Roux, B. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 774–786.
- (20) 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.* **2008**, *30*, 191–202.
- (21) de Courcy, B.; Piquemal, J.-P.; Gresh, N. Energy Analysis of Zn Polycoordination in a Metalloprotein Environment and of the Role of a Neighboring Aromatic Residue. What Is the Impact of Polarization? *J. Chem. Theory Comput.* **2008**, *4*, 1659–1668.
- (22) Gresh, N.; Piquemal, J.-P.; Krauss, M. Representation of Zn(II) Complexes in Polarizable Molecular Mechanics. Further Refinements of the Electrostatic and Short-Range Contributions. Comparisons with Parallel Ab Initio Computations. *J. Comput. Chem.* **2005**, *26*, 1113–1130.
- (23) Ryde, U. Carboxylate Binding Modes in Zinc Proteins: A Theoretical Study. *Biophys. J.* **1999**, *77*, 2777–2787.
- (24) 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.* **2010**, *6*, 337–343.
- (25) 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.
- (26) Plotnikov, N. V.; Kamerlin, S. C. L.; Warshel, A. Paradynamics: An Effective and Reliable Model for Ab Initio QM/MM Free-Energy Calculations and Related Tasks. *J. Phys. Chem. B* **2011**, *115*, 7950–7962.
- (27) Hu, H.; Yang, W. Free Energies of Chemical Reactions in Solution and in Enzymes with Ab Initio Quantum Mechanics/Molecular Mechanics Methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.
- (28) Wu, R. B.; Cao, Z. X.; Zhang, Y. K. Computational Simulations of Zinc Enzyme: Challenges and Recent Advances. *Prog. Chem.* **2012**, *24*, 1175–1184.
- (29) Wang, L. H.; Yu, X. Y.; Hu, P.; Broyde, S.; Zhang, Y. K. A Water-Mediated and Substrate-Assisted Catalytic Mechanism for Sulfolobus solfataricus DNA Polymerase IV. *J. Am. Chem. Soc.* **2007**, *129*, 4731–4737.
- (30) Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R. Z.; Siegbahn, P. E. M. Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601–3658.
- (31) Jonas, S.; Hollfelder, F. Mapping Catalytic Promiscuity in the Alkaline Phosphatase Superfamily. *Pure Appl. Chem.* **2009**, *81*, 731–742.
- (32) Lassila, J. K.; Zalatan, J. G.; Herschlag, D. Biological Phosphoryl Transfer Reactions: Understanding Mechanism and Catalysis. *Annu. Rev. Biochem.* **2011**, *80*, 669–702.
- (33) Duarte, F.; Amrein, B. A.; Kamerlin, S. C. L. Modeling Catalytic Promiscuity in the Alkaline Phosphatase Superfamily. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11160–11177.
- (34) Dewar, M. J. S.; Thiel, W. Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- (35) Dewar, M. J. S.; Thiel, W. Ground States of Molecules. 39. MNDO Results for Molecules Containing Hydrogen, Carbon, Nitrogen, and Oxygen. *J. Am. Chem. Soc.* **1977**, *99*, 4907–4917.
- (36) Thiel, W.; Voityuk, A. A. Extension of MNDO to d Orbitals: Parameters and Results for the Second-Row Elements and for the Zinc Group. *J. Phys. Chem.* **1996**, *100*, 616–626.
- (37) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (38) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (39) Hutter, M. C.; Reimers, J. R.; Hush, N. S. Modeling the Bacterial Photosynthetic Reaction Center. 1. Magnesium Parameters for the Semiempirical AM1Method Developed Using a Genetic Algorithm. *J. Phys. Chem. B* **1998**, *102*, 8080–8090.
- (40) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods IV: Extension of MNDO, AM1, and PM3 to More Main Group Elements. *J. Mol. Model.* **2004**, *10*, 155–164.
- (41) Cai, Z.; Lopez, P.; Reimers, J. R.; Cui, Q.; Elstner, M. Application of the Computationally Efficient Self-Consistent-Charge Density-Functional Tight-Binding Method to Magnesium-Containing Molecules. *J. Phys. Chem. A* **2007**, *111*, 5743–5750.
- (42) Imhof, P.; Noé, F.; Fischer, S.; Smith, J. C. AM1/d Parameters for Magnesium in Metalloenzymes. *J. Chem. Theory Comput.* **2006**, *2*, 1050–1056.
- (43) Sorkin, A.; Truhlar, D. G.; Amin, E. A. Energies, Geometries, and Charge Distributions of Zn Molecules, Clusters, and Biocenters from Coupled Cluster, Density Functional, and Neglect of Diatomic Differential Overlap Models. *J. Chem. Theory Comput.* **2009**, *5*, 1254–1265.
- (44) Dewar, M. J. S.; Merz, K. M. AM1 Parameters for Zinc. *Organometallics* **1988**, *7*, 522–524.
- (45) Brothers, E. N.; Suarez, D.; Deerfield, D. W.; Merz, K. M. PM3-Compatible Zinc Parameters Optimized for Metalloenzyme Active Sites. *J. Comput. Chem.* **2004**, *25*, 1677–1692.

- (46) Bräuer, M.; Kunert, M.; Dinjus, E.; Klussmann, M.; Döring, M.; Görls, H.; Anders, E. Evaluation of the Accuracy of PM3, AM1 and MNDO/d as Applied to Zinc Compounds. *J. Mol. Struct. (Theochem)* **2000**, *505*, 289–301.
- (47) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (48) Hou, G.; Cui, Q. QM/MM Analysis Suggests That Alkaline Phosphatase (AP) and Nucleotide Pyrophosphatase/Phosphodiesterase Slightly Tighten the Transition State for Phosphate Diester Hydrolysis Relative to Solution: Implication for Catalytic Promiscuity in the AP Superfamily. *J. Am. Chem. Soc.* **2012**, *134*, 229–246.
- (49) Hou, G.; Cui, Q. Stabilization of Different Types of Transition States in a Single Enzyme Active Site: QM/MM Analysis of Enzymes in the Alkaline Phosphatase Superfamily. *J. Am. Chem. Soc.* **2013**, *135*, 10457–10469.
- (50) López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. Theoretical Study of Phosphodiester Hydrolysis in Nucleotide Pyrophosphatase/Phosphodiesterase. Environmental Effects on the Reaction Mechanism. *J. Am. Chem. Soc.* **2010**, *132*, 6955–6963.
- (51) López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. Promiscuity in Alkaline Phosphatase Superfamily. Unraveling Evolution through Molecular Simulations. *J. Am. Chem. Soc.* **2011**, *133*, 12050–12062.
- (52) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- (53) Elstner, M. SCC-DFTB: What is the Proper Degree of Self-Consistency? *J. Phys. Chem. A* **2007**, *111*, 5614–5621.
- (54) Gaus, M.; Cui, Q.; Elstner, M. Density Functional Tight Binding (DFTB): Application to Organic and Biological Molecules. *WIREs Comput. Mol. Sci.* **2014**, *4*, 49–61.
- (55) Kruger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. Validation of the Density Functional Based Tight-Binding Approximation Method for the Calculation of Reaction Energies and Other Data. *J. Chem. Phys.* **2005**, *122*, 114110 1–5.
- (56) Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. Comparison of SCC-DFTB and NDDO-Based Semiempirical Molecular Orbital Methods for Organic Molecules. *J. Phys. Chem. A* **2006**, *110*, 13551–13559.
- (57) Otte, N.; Scholten, M.; Thiel, W. Looking at Self-Consistent-Charge Density Functional Tight Binding from a Semiempirical Perspective. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.
- (58) Kubar, T.; Bodrog, Z.; Gaus, M.; Köhler, C.; Aradi, B.; Frauenheim, T.; Elstner, M. Parametrization of the SCC-DFTB Method for Halogens. *J. Chem. Theory Comput.* **2013**, *9*, 2939–2949.
- (59) Zheng, G.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K.; Lundberg, M.; Elstner, M.; Köhler, C.; Köhler, T.; et al. Parameter Calibration of Transition-Metal Elements for the Spin-Polarized Self-Consistent-Charge Density-Functional Tight-Binding (DFTB) Method: Sc, Ti, Fe, Co, and Ni. *J. Chem. Theory Comput.* **2007**, *3*, 1349–1367.
- (60) Dolgonos, G.; Aradi, B.; Moreira, N. H.; Frauenheim, T. An Improved Self-Consistent-Charge Density-Functional Tight-Binding (SCC-DFTB) Set of Parameters for Simulation of Bulk and Molecular Systems Involving Titanium. *J. Chem. Theory Comput.* **2010**, *6*, 266–278.
- (61) Grundkötter-Stock, B.; Bezugly, V.; Kunstmann, J.; Cuniberti, G.; Frauenheim, T.; Niehaus, T. A. SCC-DFTB Parametrization for Boron and Boranes. *J. Chem. Theory Comput.* **2012**, *8*, 1153–1163.
- (62) Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. Modeling Zinc in Biomolecules with the Self Consistent Charge-Density Functional Tight Binding (SCC-DFTB) Method: Applications to Structural and Energetic Analysis. *J. Comput. Chem.* **2003**, *24*, 565–581.
- (63) Moreira, N. H.; Dolgonos, G.; Aradi, B.; Da Rosa, A. L.; Frauenheim, T. Toward an Accurate Density-Functional Tight-Binding Description of Zinc-Containing Compounds. *J. Chem. Theory Comput.* **2009**, *5*, 605–614.
- (64) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569–585.
- (65) Elstner, M.; Frauenheim, T.; Suhai, S. An Approximate DFT Method for QM/MM Simulations of Biological Structures and Processes. *J. Mol. Struct. (Theochem)* **2003**, *632*, 29–41.
- (66) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; et al. Development of Effective Quantum Mechanical/Molecular Mechanical (QM/MM) Methods for Complex Biological Processes. *J. Phys. Chem. B* **2006**, *110*, 6458–6469.
- (67) Cui, Q.; Elstner, M.; Karplus, M. A Theoretical Analysis of the Proton and Hydride Transfer in Liver Alcohol Dehydrogenase (LADH). *J. Phys. Chem. B* **2002**, *106*, 2721–2740.
- (68) Yang, Y.; Yu, H.; Cui, Q. Extensive Conformational Transitions Are Required to Turn On ATP Hydrolysis in Myosin. *J. Mol. Biol.* **2008**, *381*, 1407–1420.
- (69) Riccardi, D.; Yang, S.; Cui, Q. Proton Transfer Function of Carbonic Anhydrase: Insights from QM/MM simulations. *Biochim. Biophys. Acta* **2010**, *1804*, 342–351.
- (70) Yang, Y.; Miao, Y. P.; Wang, B.; Cui, G. L.; Merz, K. M., Jr. Catalytic Mechanism of Aromatic Prenylation by NphB. *Biochemistry* **2012**, *51*, 2606–2618.
- (71) Chakravorty, D. K.; Wang, B.; Lee, C. W.; Giedroc, D. P.; Merz, K. M., Jr. Simulations of Allosteric Motions in the Zinc Sensor CzrA. *J. Am. Chem. Soc.* **2012**, *134*, 3367–3376.
- (72) Xu, D. G.; Cui, Q.; Guo, H. Quantum Mechanical/Molecular Mechanical Studies of Zinc Hydrolases. *Int. Rev. Phys. Chem.* **2014**, *33*, 1–41.
- (73) Moreira, N. H.; da Rosa, A. L.; Frauenheim, T. Covalent Functionalization of ZnO Surfaces: A Density Functional Tight Binding Study. *Appl. Phys. Lett.* **2009**, *94*, 193109 1–3.
- (74) Holthaus, S. G.; Köppen, S.; Frauenheim, T.; Ciacchi, L. C. Atomistic Simulations of the ZnO(1 $\bar{2}$ 10)/Water Interface: A Comparison between First-Principles, Tight-Binding, and Empirical Methods. *J. Chem. Theory Comput.* **2012**, *8*, 4517–4526.
- (75) Gaus, M.; Cui, Q.; Elstner, M. DFTB3: Extension of the Self-Consistent-Charge Density-Functional Tight-Binding Method (SCC-DFTB). *J. Chem. Theory Comput.* **2011**, *7*, 931–948.
- (76) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. Extension of the Self-Consistent-Charge Tight-Binding-Density-Functional (SCC-DFTB) Method: Third Order Expansion of the DFT Total Energy and Introduction of a Modified Effective Coulomb Interaction. *J. Phys. Chem. A* **2007**, *111*, 10861–10873.
- (77) Gaus, M.; Goez, A.; Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.
- (78) Gaus, M.; Lu, X.; Elstner, M.; Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. *J. Chem. Theory Comput.* **2014**, *10*, 1518–1537.
- (79) Cui, Q.; Elstner, M. Density Functional Tight Binding: Values of Semi-Empirical Methods in an *Ab Initio* Era. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14368–14377.
- (80) Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. Automatized Parametrization of SCC-DFTB Repulsive Potentials: Application to Hydrocarbons. *J. Phys. Chem. A* **2009**, *113*, 11866–11881.
- (81) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (82) Bodrog, Z.; Aradi, B.; Frauenheim, T. Automated Repulsive Parametrization for the DFTB Method. *J. Chem. Theory Comput.* **2011**, *7*, 2654–2664.
- (83) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (84) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

- (85) 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* **1988**, *37*, 785–789.
- (86) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.
- (87) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. Assessment of Gaussian-3 and Density Functional Theories for a Larger Experimental Test Set. *J. Chem. Phys.* **2000**, *112*, 7374–7383.
- (88) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Assessment of Gaussian-3 and Density-Functional Theories on the G3/05 Test Set of Experimental Energies. *J. Chem. Phys.* **2005**, *123*, 124107 1–12.
- (89) Cheng, F.; Zhang, R.; Luo, X.; Shen, J.; Li, X.; Gu, J.; Zhu, W.; Shen, J.; Sagi, I.; Ji, R.; et al. Quantum Chemistry Study on the Interaction of the Exogenous Ligands and the Catalytic Zinc Ion in Matrix Metalloproteinases. *J. Phys. Chem. B* **2002**, *106*, 4552–4559.
- (90) Amin, E. A.; Truhlar, D. G. Zn Coordination Chemistry: Development of Benchmark Suites for Geometries, Dipole Moments, and Bond Dissociation Energies and Their Use To Test and Validate Density Functionals and Molecular Orbital Theory. *J. Chem. Theory Comput.* **2008**, *4*, 75–85.
- (91) Weaver, M. N.; Merz, K. M.; Ma, D.; Kim, H. J.; Gagliardi, L. Calculation of Heats of Formation for Zn Complexes: Comparison of Density Functional Theory, Second Order Perturbation Theory, Coupled-Cluster and Complete Active Space Methods. *J. Chem. Theory Comput.* **2013**, *9*, 5277–5285.
- (92) Hamprecht, F. A.; Cohen, A.; Tozer, D. J.; Handy, N. C. Development and Assessment of New Exchange-Correlation Functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271.
- (93) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid Exchange-Correlation Functional Determined from Thermochemical Data and Ab Initio Potentials. *J. Chem. Phys.* **2001**, *115*, 9233–9242.
- (94) Becke, A. D. Density-functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
- (95) Schmider, H. L.; Becke, A. D. Optimized Density Functionals from the Extended G2 Test Set. *J. Chem. Phys.* **1998**, *108*, 9624–9631.
- (96) Tekarli, S. M.; Drummond, M. L.; Williams, T. G.; Cundari, T. R.; Wilson, A. K. Performance of Density Functional Theory for 3d Transition Metal-Containing Complexes: Utilization of the Correlation Consistent Basis Sets. *J. Phys. Chem. A* **2009**, *113*, 8607–8614.
- (97) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. Hydrogen Bonding and Stacking Interactions of Nucleic Acid Base Pairs: A Density-Functional-Theory Based Treatment. *J. Chem. Phys.* **2001**, *114*, 5149–5155.
- (98) Sedlak, R.; Janowski, T.; Pitonak, M.; Rezac, J.; Pulay, P.; Hobza, P. Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 3364–3374.
- (99) Risthaus, T.; Grimme, S. Benchmarking of London Dispersion-Accounting Density Functional Theory Methods on Very Large Molecular Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 1580–1591.
- (100) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (101) Im, W.; Bernèche, S.; Roux, B. Generalized Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys.* **2001**, *114*, 2924–2937.
- (102) Schaefer, P.; Riccardi, D.; Cui, Q. Reliable Treatment of Electrostatics in Combined QM/MM Simulation of Macromolecules. *J. Chem. Phys.* **2005**, *123*, 014905 1–14.
- (103) Stote, R. H.; States, D. J.; Karplus, M. On the Treatment of Electrostatic Interactions in Biomolecular Simulations. *J. Chem. Phys.* **1991**, *88*, 2419–2433.
- (104) 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.
- (105) Mackerell, A. D.; Bashford, D.; Bellott, Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (106) Rowley, C. N.; Roux, B. The Solvation Structure of Na⁺ and K⁺ in Liquid Water Determined from High Level ab Initio Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 3526–3535.
- (107) 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.
- (108) Kollman, P. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.
- (109) Simonson, T.; Archontis, G.; Karplus, M. Free Energy Simulations Come of Age: Protein-Ligand Recognition. *Acc. Chem. Res.* **2002**, *35*, 430–437.
- (110) Li, G.; Cui, Q. pK_a Calculations with QM/MM Free Energy Perturbations. *J. Phys. Chem. B* **2003**, *107*, 14521–14528.
- (111) Smith, C. A.; Rayment, I. X-ray Structure of the Magnesium(II)-ADP-Vanadate Complex of the Dictyostelium discoideum Myosin Motor Domain to 1.9 Å Resolution. *Biochemistry* **1996**, *35*, 5404–5417.
- (112) Geeves, M. A.; Holmes, K. Structural Mechanism of Muscle Contraction. *Annu. Rev. Biochem.* **1999**, *68*, 687–728.
- (113) Rayment, I. The Structural Basis of the Myosin ATPase Activity. *J. Biol. Chem.* **1996**, *271*, 15850–15853.
- (114) Li, G.; Cui, Q. Mechanochemical Coupling in Myosin: A Theoretical Analysis with Molecular Dynamics and Combined QM/MM Reaction Path Calculations. *J. Phys. Chem. B* **2004**, *108*, 3342–3357.
- (115) Yu, H.; Ma, L.; Yang, Y.; Cui, Q. Mechanochemical Coupling in the Myosin Motor Domain. I. Insights from Equilibrium Active-Site Simulations. *PLoS Comput. Biol.* **2007**, *3*, 0199–0213.
- (116) Bashford, D.; Karplus, M. pK_a's of Ionizable Groups in Proteins: Atomic Detail from a Continuum Electrostatic Model. *Biochemistry* **1990**, *29*, 10219–10225.
- (117) Brooks, C. L., III; Karplus, M. Solvent Effects on Protein Motion and Protein Effects on Solvent Motion: Dynamics of the Active Site Region of Lysozyme. *J. Mol. Biol.* **1989**, *208*, 159–181.
- (118) Riccardi, D.; Cui, Q. pK_a Analysis for the Zinc-Bound Water in Human Carbonic Anhydrase II: Benchmark for "Multi-scale" QM/MM Simulations and Mechanistic Implications. *J. Phys. Chem. A* **2007**, *111*, 5703–5711.
- (119) Lu, X.; Cui, Q. Charging Free Energy Calculations Using the Generalized Solvent Boundary Potential (GSBP) and Periodic Boundary Condition: A Comparative Analysis Using Ion Solvation and Oxidation Free Energy in Proteins. *J. Phys. Chem. B* **2013**, *117*, 2005–2018.
- (120) Torrie, G.; Valleau, J. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (121) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (122) O'Brien, P. J.; Lassila, J. K.; Fenn, T. D.; Zalatan, J. G.; Herschlag, D. Arginine Coordination in Enzymatic Phosphoryl Transfer: Evaluation of the Effect of Arg166 Mutations in Escherichia coli Alkaline Phosphatase. *Biochemistry* **2008**, *47*, 7663–7672.
- (123) Fischer, S.; Karplus, M. Conjugate Peak Refinement: An Algorithm for Finding Reaction Paths and Accurate Transition States in Systems with Many Degrees of Freedom. *Chem. Phys. Lett.* **1992**, *194*, 252–261.
- (124) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Non-covalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

- (125) Grimme, S.; E, S.; Antony, J.; Krieg, H. A Consistent and Accurate Ab Initio Parameterization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104 1–19.
- (126) 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. et al. *Gaussian 09*, Revision A.1; Gaussian Inc.: Wallingford, CT, 2009.
- (127) Dudev, T.; Cowan, J. A.; Lim, C. Competitive Binding in Magnesium Coordination Chemistry: Water versus Ligands of Biological Interest. *J. Am. Chem. Soc.* **1999**, *121*, 7665–7673.
- (128) Son, T.; Roux, M.; Ellenberger, M. Interaction of Mg²⁺ Ions with Nucleoside Triphosphates by Phosphorus Magnetic-Resonance Spectroscopy. *Nucleic Acids Res.* **1975**, *2*, 1101–1110.
- (129) Cohn, T. R. J.; Hughes, M. Nuclear Magnetic Resonance Spectra of Adenosine Di- and Triphosphate. II. Effect of Complexing with Divalent Metal Ions. *J. Biol. Chem.* **1962**, *237*, 176–181.
- (130) Bishop, E.; Kimber, S.; Orchard, D.; Smith, B. A ³¹P-NMR Study of Monomagnesium and Dimagnesium Complexes of Adenosine 5'-Triphosphate and Model Systems. *Biochim. Biophys. Acta* **1981**, *635*, 63–72.
- (131) Huang, S. L.; Tsai, M. D. Does the Magnesium(II) Ion Interact with the Alpha-Phosphate of ATP? An Investigation by Oxygen-17 Nuclear Magnetic Resonance. *Biochemistry* **1982**, *21*, 951–959.
- (132) Ramirez, F.; Marecek, J. Coordination of Magnesium with Adenosine 5'-Diphosphate and Triphosphate. *Biochim. Biophys. Acta* **1980**, *589*, 21–29.
- (133) Takeuchi, H.; Murata, H.; Harada, I. Interaction of Adenosine 5'-Triphosphate with Mg²⁺: Vibrational Study of Coordination Sites by Use of ¹⁸O-Labeled Triphosphates. *J. Am. Chem. Soc.* **1988**, *110*, 392–397.
- (134) Liao, J. C.; Sun, S.; Chandler, D.; Oster, G. The Conformational States of Mg-ATP in Water. *Eur. Biophys. J.* **2004**, *33*, 29–37.
- (135) Stamper, C.; Bennett, B.; Edwards, T.; Holz, R. C.; Ringe, D.; Petsko, G. Inhibition of the Aminopeptidase from *Aeromonas proteolytica* by 1-Leucinephosphonic Acid. Spectroscopic and Crystallographic Characterization of the Transition State of Peptide Hydrolysis. *Biochemistry* **2001**, *40*, 7035–7046.
- (136) Englert, M.; Xia, S.; Okada, C.; Nakamura, A.; Tanavde, V.; Yao, M.; Eom, S. H.; Konigsberg, W. H.; Soell, D.; Wang, J. Structural and Mechanistic Insights into Guanylylation of RNA-Splicing Ligase RtcB Joining RNA between 3'-Terminal Phosphate and 5'-OH. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 15235–15240.
- (137) Jedrzejas, M. J.; Setlow, P. Comparison of the Binuclear Metalloenzymes Diphosphoglycerate-Independent Phosphoglycerate Mutase and Alkaline Phosphatase: Their Mechanism of Catalysis via a Phosphoserine Intermediate. *Chem. Rev.* **2001**, *101*, 607–618.
- (138) Larsen, T. M.; Benning, M. M.; Rayment, I.; Reed, G. H. Structure of the Bis(Mg²⁺)-ATP-Oxalate Complex of the Rabbit Muscle Pyruvate Kinase at 2.1 Å Resolution: ATP Binding over a Barrel. *Biochemistry* **1998**, *37*, 6247–6255.
- (139) Nakamura, T.; Zhao, Y.; Yamagata, Y.; Jin Hua, Y.; Yang, W. Watching DNA Polymerase η Make a Phosphodiester Bond. *Nature* **2012**, *487*, 196–201.
- (140) Schmidt, B. H.; Burgin, A. B.; Deweese, J. E.; Osheroff, N.; Berger, J. M. A Novel and Unified Two-Metal Mechanism for DNA Cleavage by Type II and IA Topoisomerases. *Nature* **2010**, *465*, 641–644.
- (141) Palermo, G.; Stenta, M.; Cavalli, A.; Dal Peraro, M.; De Vivo, M. Molecular Simulations Highlight the Role of Metals in Catalysis and Inhibition of Type II Topoisomerase. *J. Chem. Theory Comput.* **2013**, *9*, 857–862.
- (142) Chevrier, B.; Schalk, C.; D'Orchymont, H.; Rondeau, J.; Moras, D.; Tarnus, C. Crystal Structure of *Aeromonas Proteolytica* Aminopeptidase: A Prototypical Member of the Co-Catalytic Zinc Enzyme Family. *Structure* **1994**, *2*, 283–291.
- (143) Coleman, J. E. Structure and Mechanism of Alkaline Phosphatase. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 441–483.
- (144) Marcus, Y. Ionic Radii in Aqueous Solutions. *Chem. Rev.* **1988**, *88*, 1475–1498.
- (145) Bernasconi, L.; Baerends, E. J.; Sprik, M. Long-Range Solvent Effects on the Orbital Interaction Mechanism of Water Acidity Enhancement in Metal Ion Solutions: A Comparative Study of the Electronic Structure of Aqueous Mg and Zn Dications. *J. Phys. Chem. B* **2006**, *110*, 11444–11453.
- (146) Riahi, S.; Roux, B.; Rowley, C. N. QM/MM Molecular Dynamics Simulations of the Hydration of Mg(II) and Zn(II) Ions. *Can. J. Chem.* **2013**, *91*, 552–558.
- (147) Kulik, H. J.; Marzari, N.; Correa, A. A.; Prendergast, D.; Schwegler, E.; Galli, G. Local Effects in the X-ray Absorption Spectrum of Salt Water. *J. Phys. Chem. B* **2010**, *114*, 9594–9601.
- (148) Bhattacharjee, A.; Pribil, A. B.; Randolph, B. R.; Rode, B. M.; Hofer, T. S. Hydration of Mg²⁺ and Its Influence on the Water Hydrogen Bonding Network via Ab Initio QMCF MD. *Chem. Phys. Lett.* **2012**, *536*, 39–44.
- (149) Ikeda, T.; Boero, M.; Terakura, K. Hydration Properties of Magnesium and Calcium Ions from Constrained First Principles Molecular Dynamics. *J. Chem. Phys.* **2007**, *127*, 074503 1–8.
- (150) Kuzmin, A.; Obst, S.; Purans, J. X-ray Absorption Spectroscopy and Molecular Dynamics Studies of Zn²⁺ Hydration in Aqueous Solutions. *J. Phys.: Condens. Matter* **1997**, *9*, 10065–10078.
- (151) Rega, N.; Brancato, G.; Petrone, A.; Caruso, P.; Barone, V. Vibrational Analysis of X-ray Absorption Fine Structure Thermal Factors by Ab Initio Molecular Dynamics: The Zn(II) Ion in Aqueous Solution as a Case Study. *J. Chem. Phys.* **2011**, *134*, 074504 1–10.
- (152) Neely, J.; Connick, R. Rate of Water Exchange from Hydrated Magnesium Ion. *J. Am. Chem. Soc.* **1970**, *92*, 3476–3478.
- (153) Helm, L.; Merbach, A. Water Exchange on Metal Ions: Experiments and Simulations. *Coord. Chem. Rev.* **1999**, *187*, 151–181.
- (154) Salmon, P.; Bellissenfunel, M.; Herdman, G. The Dynamics of Aqueous Zn²⁺ Solutions: A Study Using Incoherent Quasi-Elastic Neutron Scattering. *J. Phys.: Condens. Matter* **1990**, *2*, 4297–4309.
- (155) Douglas, B. E.; McDaniel, D. H.; Alexander, J. J. *Concepts and Models of Inorganic Chemistry*; John Wiley and Sons: New York, 1994.
- (156) Vale, R.; Milligan, R. The Way Things Move: Looking under the Hood of Molecular Motor Proteins. *Science* **2000**, *288*, 88–95.
- (157) Bustamante, C.; Keller, D.; Oster, G. The Physics of Molecular Motors. *Acc. Chem. Res.* **2001**, *34*, 412–420.
- (158) Gao, Y.; Yang, W.; Marcus, R.; Karplus, M. A Model for the Cooperative Free Energy Transduction and Kinetics of ATP Hydrolysis by F₁-ATPase. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11339–11344.
- (159) Strajbl, M.; Shurki, A.; Warshel, A. Converting Conformational Changes to Electrostatic Energy in Molecular Motors: The Energetics of ATP Synthase. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14834–14839.
- (160) McGrath, M. J.; Kuo, I.-F. W.; Hayashi, S.; Takada, S. Adenosine Triphosphate Hydrolysis Mechanism in Kinesin Studied by Combined Quantum-Mechanical/Molecular-Mechanical Metadynamics Simulations. *J. Am. Chem. Soc.* **2013**, *135*, 8908–8919.
- (161) Xia, F.; Rudack, T.; Q. Cui, C. K.; Gerwert, K. Detailed Structure of the H₂PO₄⁻-GDP Intermediate in Ras-GAP Decoded from FTIR Experiments by Biomolecular Simulations. *J. Am. Chem. Soc.* **2012**, *116*, 12522–12534.
- (162) Laio, A.; Gervasio, F. L. Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. *Rep. Prog. Phys.* **2008**, *71*, 126601 1–22.
- (163) W, E.; Ren, W.; Vanden-Eijnden, E. Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier-Crossing Events. *J. Chem. Phys.* **2007**, *126*, 164103 1–8.
- (164) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. String Method in Collective Variables: Minimum Free Energy Paths and Isocommittor Surfaces. *J. Chem. Phys.* **2006**, *125*, 024106 1–15.
- (165) Vanden-Eijnden, E.; Venturoli, M. Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies. *J. Chem. Phys.* **2009**, *130*, 194103 1–17.