J|A|C|S
ARTICLES

Published on Web 12/03/2010

Electronic Properties and Desolvation Penalties of Metal Ions Plus Protein Electrostatics Dictate the Metal Binding Affinity and Selectivity in the Copper Efflux Regulator

Li Rao,[†,§] Qiang Cui,*[,‡] and Xin Xu*[,†,§]

Department of Chemistry, Xiamen University, Xiamen, P. R. China, and Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

Received May 2, 2010; E-mail: xxchem@fudan.edu.cn; cui@chem.wisc.edu

Abstract: We carry out a theoretical analysis of factors that dictate the binding affinity and selectivity of the copper efflux regulator (CueR) toward different metal ions (Cu^+, Ag^+, Au^+, Zn^{2+}, and Hg^{2+}). In addition to a simplified active-site model, we have established a computational framework based on quantum mechanical/molecular mechanical (QM/MM) and Poisson−Boltzmann approaches that allows us, for the first time, to systematically analyze the protein contribution to transition metal binding affinity and selectivity. We find that the QM/MM model leads to relative binding affinities that are consistent with observations from transcription induction experiments, while an active-site model does not, which highlights the importance of explicitly considering the protein environment for a thorough understanding of metal binding properties of metalloproteins. Regarding the trends in binding affinity, our analysis highlights both intrinsic properties of a metal ion and protein contributions. Specifically, the softness and desolvation penalty of a metal ion make large contributions to the binding affinity; for example, we find that the large desolvation penalty for Zn^{2+} rather than any stereoelectronic factor (e.g., linear vs tetrahedron coordination) is the key reason that Zn^{2+} binds much more weakly than Hg^{2+} to CueR. Moreover, our results explicitly demonstrate that the electrostatic environment of CueR is well-tuned to favor the binding of coinage metal ions over divalent ions. Finally, our analyses highlight the importance of considering the proper solution reference (i.e., the metal ion bound to buffer ligands vs water molecules) when discussing the binding affinity of metal ions to proteins.

1. Introduction

Metal ions are essential cofactors to many biomolecules, especially metalloenzymes that catalyze complex chemical transformations.[1] For example, a recent survey[2] indicated that, among 1371 different enzymes for which three-dimensional structures are available, ∼47% contain metal ions, with 41% hosting metals at the catalytic site. Since the catalytic reactivities of different metals are often distinct, it is important that the correct metal ion(s) is incorporated in the catalytic site; similarly, metal-sensing transcription factors are evolved to be activated by specific metal ions. Clearly, revealing the underlying mechanism for metal binding selectivity is a crucial step toward understanding how the function of metalloproteins is regulated in cells[3−5] as well as the design of novel proteins that recognize specific metal ions.[6,7]

As summarized in several recent reviews,[3−5] the *in vivo* metal ion selectivity is often the result of several factors that include structural properties of the active site ("affinity") and global conformations ("allostery") as well as the metal availability in the cellular pool ("access"). Among those, the local structural properties are most clearly defined in physical terms and have been analyzed in several experimental and computational studies. For example, bacterial homeostasis and resistance systems have been described[8−11] for Ag^+, AsO_2^-, AsO_4^{3-}, Cd^{2+}, Co^{2+}, CrO_4^{2-}, $Cu^{2+/+}$, Hg^{2+}, Ni^{2+}, Pb^{2+}, TeO_3^{2-}, Zn^{2+}, and Au^+; metallochaperones have been characterized for Cu^+, Ni^{2+}, and heme.[8,12−14] Some of the proteins have been characterized

[†] Xiamen University.
[‡] University of Wisconsin, Madison.
[§] Current address: Department of Chemistry, Fudan University, Shanghai, P. R. China.

(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. *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218.
(3) Finney, L. A.; O'Halloran, T. V. *Science* **2003**, *300*, 931–936.
(4) Waldron, K. J.; Rutherford, J. C.; Ford, D.; Robinson, N. J. *Nature* **2009**, *460*, 823–830.
(5) Waldron, K. J.; Robinson, N. J. *Nat. Rev. Microbiol.* **2009**, *6*, 25–35.
(6) Chen, P. R.; Wasinger, E. C.; Zhao, J.; van der Lelie, D.; Chen, L. X.; He, C. *J. Am. Chem. Soc.* **2007**, *129*, 12350–12351.
(7) Wegner, S. V.; Boyaci, H.; Chen, H.; Jensen, M. P.; He, C. *Angew. Chem., Int. Ed.* **2009**, *48*, 2339–2341.
(8) Huffman, D. L.; O'Halloran, T. V. *Annu. Rev. Biochem.* **2001**, *70*, 677–701.
(9) Tottey, S.; Harvie, D. R.; Robinson, N. J. *Acc. Chem. Res.* **2005**, *38*, 775–783.
(10) Kim, B. E.; Nevitt, T.; Thiele, D. J. *Nat. Chem. Biol.* **2008**, *4*, 176–185.
(11) Jian, X.; Wasinger, E. C.; Lockard, J. V.; Chen, L. X.; He, C. *J. Am. Chem. Soc.* **2009**, *131*, 10869–10871.
(12) Musiani, F.; Zambelli, B.; Stola, M.; Ciurli, S. *J. Inorg. Biochem.* **2004**, *98*, 803–813.
(13) Bulteau, A. L.; O'Neill, H. A.; Kennedy, M. C.; Ikeda-Saito, M.; Isaya, G.; Szweda, L. I. *Science* **2004**, *305*, 242–245.

10.1021/ja103742k

structurally using X-ray crystallography[15] or NMR,[16,17] which provided hints regarding the mechanism for metal selectivity. Most often, the selectivity is *qualitatively* rationalized in terms of the "pre-organized" coordination environment of the metal ion; e.g., an unusual linear, two-coordinated geometry is invoked[18] to explain the high selectivity of the copper efflux regulator (CueR) in *Escherichia coli* toward Cu^+ over Zn^{2+}. However, it is also realized that many metal sensing/trafficking proteins are structurally very flexible,[19] as required by their function; thus, the importance of the coordination geometry can be significantly overestimated.

A series of computational studies has been carried out by Lim and co-workers[20,21] in the past few years to address binding selectivity of metal ions such as Zn^{2+} vs Hg^{2+},[22] Mg^{2+} vs Ca^{2+},[23] and Al^{3+} vs Ln^{3+} [24] as well as Na^+ vs K^+.[25,26] Useful insights have been obtained from these studies, which highlighted the importance of several ligand properties, including charge, polarizability, charge-donating ability, and denticity. However, these studies[22−24] have used relatively small active-site models and replaced most of the protein environment by an implicit dielectric continuum. Therefore, only limited insights were available regarding contributions from residues beyond the first coordination shell of the metal ion; it was also not possible to evaluate the role of protein flexibility and/or active-site rigidity with these models. In the context of ion selectivity of ion channels, which has been largely analyzed with classical models,[27,28] studies start to emerge in which one employs a quantum mechanical description of the ions and their ligands[29] and a classical description of the protein environment.[30] An accurate estimate of the binding free energy and selectivity, however, remains challenging due to the need for ample sampling[31] or approximations in the treatment of the ion environment.[28,32]

The goal of our study is to establish effective hybrid quantum mechanical/molecular mechanical (QM/MM[33−41]) methodologies to analyze binding selectivity of transition metal ions to proteins, so that contributions from both the first coordination shell of the metal ion and other features of the protein environment can be systematically evaluated. As discussed in more detail in section 2, this is not a trivial task to undertake because the cost of high-level QM methods has to be balanced with an effective treatment of electrostatic interactions and conformational flexibility of a solvated protein. As an important step toward this direction, we have established a QM/MM protocol based on a specific thermodynamic cycle and a Poisson−Boltzmann (PB) solvation model. We then apply this protocol to carry out the *first* systematic analysis of binding selectivity of CueR to several commonly encountered metal ions: Cu^+, Ag^+, Au^+, Zn^{2+}, and Hg^{2+}.

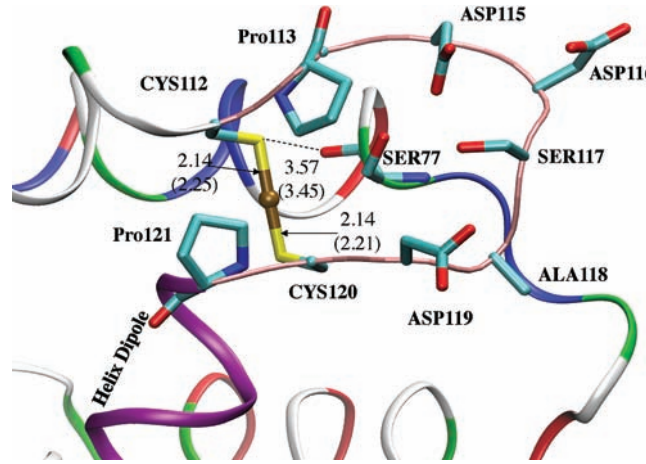

Figure 1. Active site of CueR based on the crystal structure[18] that contains a bound Cu^+, which is shown as a brown sphere bound to two Cys residues. Nearby amino acids and the helix dipole that has been proposed to stabilize the electrostatics of the binding site are indicated. A few distances from the two binding sites (with and without parentheses) in the crystal structure are also included. Note that the Cys ligands of the metal ion are stabilized by hydrogen-bonding interactions from several main-chain NH groups, which are not shown explicitly for clarity.

The reason that CueR is selected is that it is one of the best studied metal-sensing transcription factors in terms of both structure and binding affinity/selectivity.[18,42] *In vivo* assays[18] showed that CueR responds to monovalent coinage metal ions (Cu^+, Ag^+, Au^+) but not divalent metal ions like Zn^{2+} and Hg^{2+}. A high-resolution X-ray structure of CueR revealed a linear, two-coordinated geometry of Cu^+ (see Figure 1), which was used to explain the low affinity of Zn^{2+}, an ion that prefers a tetrahedral coordination.[1] On the other hand, we note that the binding site is in a loop region (Figure 1), which can be sufficiently flexible to accommodate an ion (e.g., Zn^{2+}) that prefers a tetrahedron type of coordination, possibly involving water molecules as some of the ligands. The electrostatic interaction between the conserved Lys81 and the metal binding site in CueR has been invoked[18] to explain the discrimination against Hg^{2+}, which also prefers a linear coordination; however, Lys81 is solvent exposed, and therefore its role in electrostatic stabilization might be rather small. Finally, as typical for metal binding processes, the protonation state of active-site residues

(14) Schulz, H.; Hennecke, H.; Thony-Meyer, L. *Science* **1998**, *281*, 1197–1200.
(15) Rosenzweig, A. C. *Acc. Chem. Res.* **2001**, *34*, 119–128.
(16) Wimmer, R.; Herrmann, T.; Solioz, M.; Wuthrich, K. *J. Biol. Chem.* **1999**, *274*, 22597–22603.
(17) Banci, L.; Bertini, I.; Cantini, F.; Gonnelli, L.; Hadjiliadis, N.; Pierattelli, R.; Rosato, A.; Voulgaris, P. *Nat. Chem. Biol.* **2006**, *2*, 367–368.
(18) Changela, A.; Chen, K.; Xue, Y.; Holschen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragon, A. *Science* **2003**, *301*, 1383–1387.
(19) Barondeau, D. P.; Getzoff, E. D. *Curr. Opin. Struct. Biol.* **2004**, *14*, 765–774.
(20) Dudev, T.; Lim, C. *Chem. Rev.* **2006**, *103*, 773–787.
(21) Dudev, T.; Lim, C. *Annu. Rev. Biochem.* **2008**, *37*, 97–116.
(22) Tai, H.; Lim, C. *J. Phys. Chem. A* **2006**, *110*, 452–462.
(23) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2004**, *108*, 4546–4557.
(24) Dudev, T.; Chang, L.; Lim, C. *J. Am. Chem. Soc.* **2005**, *127*, 4091–4103.
(25) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2010**, *132*, 2321–2332.
(26) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2009**, *131*, 8092–8101.
(27) Roux, B.; Allen, T.; Berneche, S.; Im, W. *Q. Rev. Biophys.* **2004**, *37*, 15–103.
(28) Roux, B. *Biophys. J.* **2010**, *98*, 2877–2885.
(29) Varma, S.; Sabo, D.; Rempe, S. B. *J. Mol. Biol.* **2007**, *376*, 13–22.
(30) Bucher, D.; Guidoni, L.; Carloni, P.; Rothlisberger, U. *Biophys. J.* **2010**, *98*, L47–L49.
(31) Yu, H. B.; Roux, B. *Biophys. J.* **2009**, *97*, L15–L17.
(32) Roux, B.; Yu, H. *J. Chem. Phys.* **2010**, *132*, 234101.
(33) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
(34) Monard, G.; Merz, K. M., Jr. *Acc. Chem. Res.* **1999**, *32*, 904–911.
(35) Shurki, A.; Warshel, A. *Adv. Protein Chem.* **2003**, *66*, 249.
(36) Zhang, Y. K. *Theor. Chem. Acc.* **2006**, *116*, 43–50.
(37) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467.
(38) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.
(39) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, N.; Prat-Resina, X.; Konig, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2006**, *110*, 6458–6469.
(40) Hu, H.; Yang, W. T. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.
(41) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
(42) Hobman, J. L. *Mol. Microbiol.* **2007**, *63*, 1275–1278.

may change upon metal coordination. For CueR, both ligands are Cys residues; while some authors proposed that both Cys residues are deprotonated in the metal-bound CueR,[18] others argued[43] that only one Cys is deprotonated. Therefore, CueR is ideally suited for computational analysis because there are both solid experimental background and ambiguities regarding the selectivity mechanism.

The paper is organized as follows. In section 2, we summarize the QM/MM protocol for estimating the metal binding affinity and analysis of various contributions. As a comparison to QM/MM calculations, we also carry out an analysis using active-site models in which the majority of the protein is treated implicitly with a dielectric continuum; the computational details are also summarized in section 2. The results of both implicit protein and QM/MM calculations are presented and discussed in section 3. Finally, we draw a few conclusions in section 4.

2. Computational Methods

As alluded to in the Introduction, it is a significant challenge to carry out quantitative calculations for the absolute binding affinity of a transition metal ion to a protein. First, the electronic structure of transition metal ions is generally complex, and an accurate description requires high-level QM methods,[44] such as density functional theory and correlated *ab initio* approaches, for treating the metal and at least its immediate ligands. Second, the binding of the metal ion is often coupled to significant structural rearrangements in the protein at both the local and global scales,[4,5,15,45] which are not straightforward to sample with typical atomistic molecular dynamics simulations, especially for the global transitions. Finally, the protonation state of active-site residues may change upon metal incorporation, which further complicates the binding calculations.

Considering these nontrivial technical challenges, we have chosen to avoid a "brute-force" methodology based on, for example, free energy perturbation[46,47] calculations as the *first* study. Instead, we follow a protocol that explicitly includes all protein atoms but involves limited conformational sampling of the protein/solvent. The assumptions are that (i) structural fluctuations in the protein do not influence the metal−protein interaction in a significant manner and (ii) there is no major structural transition in the protein coupled to metal binding. The first assumption appears to be reasonable for CueR based on calculation results using different snapshots from room-temperature MD simulations; this is probably because transition metal ions bind more tightly to the protein with a well-defined coordination sphere, in contrast to the binding of Na^+/K^+ to ion channels, for which larger fluctuations are expected.[31] The second assumption is harder to justify, although its impact on the *relative* binding affinity of CueR to different metals, which is the key quantity of interest here, is expected to be small (see additional discussion below).

In the following, we first present the QM/MM protocol. Next, we summarize the computational details for the active-site models in which the majority of the protein is represented with a dielectric continuum with different dielectric constant, ε; with $\varepsilon = 78.5$, the active-site model can be regarded as a model for metal binding to glutathione (GSH), the most abundant thiol in the cell. Similar computational methods are used for computing the binding affinity of the metal ion to dithiothreitol (DTT), as a representative reference ligand in the buffer solution commonly used in experimental studies of metal binding to proteins.

$$\begin{array}{ccc} M^{m+}_{(g)} + P_{(g)} & \xrightarrow{\Delta G^{(g)}_{bind}} & M^{m+}\cdot P'_{(g)} + nH^+ \\ \Delta G^{M^{m+}}_{slv}\downarrow \quad \Delta G^{P}_{slv}\downarrow & & \Delta G^{M^{m+}\cdot P'}_{slv}\downarrow \quad \downarrow n\Delta G^{H^+}_{slv} \\ M^{m+}_{(aq)} + P_{(aq)} & \xrightarrow[\Delta G^{(aq)}_{bind}]{} & M^{m+}\cdot P'_{(aq)} + nH^+_{(aq)} \end{array}$$

Figure 2. Thermodynamic cycle for the binding of a metal ion (M^{m+}) to a protein (P); the notations highlight the fact that metal binding is likely coupled to changes in the protein structure (P′ vs P) and protonation state of active-site residues (release of nH^+). In the actual calculations, the gas-phase binding affinity ($\Delta G^{(g)}_{bind}$) is computed on the basis of QM/MM calculations, and the solvation components ($\Delta G^{M^{m+}}_{slv}$, ΔG^{P}_{slv}, $\Delta G^{M^{m+}\cdot P'}_{slv}$) are estimated on the basis of PB calculations. Since we are mainly interested in *relative* binding affinity of different metal ions to CueR, potentially large structural transitions in the active site during the binding process and entropic components of the binding free energy have not been considered (see main text in section 3.2 for more detailed discussions).

2.1. QM/MM Calculations with an Explicit Protein Environment. Our QM/MM protocol for computing metal binding affinity is based on the thermodynamic cycle outlined in Figure 2, which clearly highlights the fact that metal binding is likely coupled to structural changes in the protein (P′ vs P) and protonation state change of active-site residues (release of nH^+). The quantity of interest for each metal ion (M^{m+}) is the binding free energy in solution, $\Delta G^{(aq)}_{bind}$, which, according to the thermodynamic cycle, can be written as the sum of several contributions:

$$\Delta G^{(aq)}_{bind} = \Delta G^{(g)}_{bind} + \Delta G^{M^{m+}\cdot P'}_{slv} + n\Delta G^{H^+}_{slv} - \Delta G^{M^{m+}}_{slv} - \Delta G^{P}_{slv} \quad (1)$$

In our protocol, we compute these contributions individually using different methods, which are discussed in details below; the only exception is $\Delta G^{H^+}_{slv}$, for which we take the experimental value of −272.54 kcal/mol as adjusted for pH 7.[48] As to the value of n (number of protons released during metal binding), we test several possibilities: $n = 0$ (i.e., both Cys112 and 120 remain protonated after metal binding), $n = 1$ (either Cys112 or Cys120 becomes deprotonated after metal binding), and $n = 2$ (both Cys112 and Cys120 become deprotonated).

2.1.1. $\Delta G^{(g)}_{bind}$. The binding energy of the metal ion to the protein in a gas-phase setting is calculated using the following QM/MM protocol. Starting with the crystal structure for the metal-bound system (PDB code 1Q05[18] with Cu^+ bound), the missing residues (Asp115−Asp119 in chain A and Arg75 in chain B) are first generated by using structural information from complementary chains (i.e., using coordinates for chain B/A for missing residues in chain A/B); hydrogen atoms are added using the HBUILD facility in CHARMM.[49] The structure is then overlaid with a water sphere of 30 Å radius centered at the copper ion in chain B, and the system is equilibrated with a relatively short, classical MD simulation of 300 ps using the CHARMM22 force field,[50] during which the metal ion and its two Cys ligands are fixed while other protein/solvent atoms within 20 Å from the metal ion are allowed to move; the rest of the protein is fixed at its crystal structure positions. The snapshot with the lowest potential energy is used as the starting structure for QM/MM geometry optimization; to evaluate the effects due to structural fluctuations, five additional snapshots are used as initial structures for Cu^+ and Hg^{2+}. The QM region (68 atoms, 4 link atoms, see Figure 1) contains the metal ion, all atoms (including side chains and main chains) of the two Cys ligands (Cys112, Cys120) except for the N−H groups, all atoms of Pro113, Pro121,

(43) Hobman, J. L.; Wilkie, J.; Brown, N. L. *BioMetals* **2005**, *18*, 429–436.
(44) Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757–10816.
(45) Schreiter, E. R.; Sintchak, M. D.; Guo, Y. Y.; Chivers, P. T.; Sauer, R. T.; Drennan, C. L. *Nat. Struct. Biol.* **2003**, *10*, 794–799.
(46) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: San Diego/London, 2002.
(47) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430–437.
(48) Liptak, M.; Gross, K.; Seybold, P.; Feldgus, S.; Shields, G. *J. Am. Chem. Soc.* **2002**, *124*, 6421–6427.
(49) Brooks, B.; et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
(50) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

and Ile122, and the N−H groups of Gly114 and Ile123. The key motivation for making this selection is to describe all hydrogen-bonding (NH groups of Gly114, Ile122, and Ile123) interactions with the sulfur atoms in Cys112/120 at the QM level and to leave the QM/MM boundary relatively far from the metal ion. As discussed in more detail in the Supporting Information, this choice of the QM region is expected to be adequate for the current purpose. The rest are MM atoms treated using the CHARMM22 force field. For geometry optimization, the QM level is B3LYP,[51,52] with the 6-31+G(d) basis for main-group elements and Lanl2dz (basis and effective core potential) for the metal; for Hg^{2+}, the def2-SV(P)[53] set of effective core potential and basis is used.[54] For more accurate binding energy calculations, larger basis sets are used, which contain 6-311++G(2d,2p) for main-group elements and SDD (Stuttgart/Dresden effective core potentials and basis sets) for the metal.[54−57] To be consistent with the thermodynamic cycle used for binding calculations (Figure 2), although explicit water molecules are included in the structural equilibration and minimization to avoid unphysical structural distortions, they are deleted in the larger-basis-set QM energy calculations. For more discussions on the choice of the QM level and basis set, please see the Supporting Information.

To calculate the binding energy, one has to subtract the total QM/MM energies of the metal-bound system and the *apo* protein (following eq 1). Since these are large numerical values, for each specific protonation state for the Cys ligands in the metal-bound state (i.e., a specific *n* value), we start the geometry minimization for the *apo* structure from the optimized metal-bound protein (after removing the metal ion); i.e., although the *apo* protein always has neutral Cys112/Cys120, a slightly different structure is obtained, depending on the protonation pattern of Cys112/Cys120 in the metal-bound protein. In this way, we minimize the contribution from structural differences between the *apo* and metal-bound states; as discussed above, by computing such "vertical binding affinity", we assume that the structural reorganization associated with metal binding is minimal in the CueR system, which is most meaningful for probing *relative* binding affinities of the protein to different ions. Test calculations confirm that, with this protocol, the calculated binding energy does not vary dramatically if different snapshots from the equilibration are used to initiate the QM/MM optimizations (see section 3.2). In a similar spirit, to better compare the binding affinity of different metal ions, for each protonation state, the optimized Cu^+-bound structure is used as the starting point for other metal-bound systems (Ag^+, Au^+, Zn^{2+}, and Hg^{2+}). This is justified by the observation that the crystal structures of CueR[18] with different metal ions bound are very similar, especially in the active-site region; moreover, the QM/MM calculations lead to active-site structures very similar to the corresponding crystal structures (see Supporting Information and discussions in section 3.2).

For the divalent ions, it is possible that an alternative coordination sphere (e.g., tetrahedral for Zn^{2+}) is energetically more favorable than the linear coordination seen for $Cu^+/Ag^+/Au^+$ in the crystal structures.[18] The different coordination sphere can involve additional water molecule(s) and/or slight distortion of nearby residues. Therefore, for Zn^{2+}, additional QM/MM optimization/single-point calculations are carried out following 1 ns of classical MD simulations in which the Zn^{2+} is represented using the nonbonded parameter developed by Stote and co-workers.[58]

The QM/MM calculations are carried out using the CHEMSHELL package,[59] and optimizations are done with the HLDCopt module. The convergence threshold for the optimization is set to 0.00135 au for gradient in the HLDC optimizer. The classical MD simulations are carried out with CHARMM.[49] For the entropic and vibrational enthalpic components of metal binding, which in the current case is likely dominated by the loss of translational entropy of the metal ion and zero-point corrections associated with proton release from the Cys ligands, respectively, we take the corresponding terms from the active-site model (see below). Although clearly an approximation, this is expected to be adequate for the current analysis; it also makes it straightforward to compare the active-site and QM/MM models.

2.1.2. $\Delta G_{slv}^{M^{m+}\cdot P'}$ and ΔG_{slv}^{P}. To compute the solvation free energies for the metal-bound and *apo* proteins, PB calculations[60] are carried out for the corresponding QM/MM minimized structures (with the explicit water molecules removed). The partial charges are the CHARMM charges, and the radii are those optimized by Roux and co-workers;[61] for the metal ions, the experimental ionic radii were used, although test calculations using different radii suggest that the results are not sensitive to the metal radii, which is expected given that the binding site is well buried. To be consistent with the QM/MM calculations, a dielectric constant of 1.0 is used for the protein in the PB calculations, although using a value of 2.0 does not change trends in the computed binding affinities (see section 3.2); a value of 78.54 is used for water. The calculations have been carried out using the APBS package.[62]

2.1.3. $\Delta G_{slv}^{M^{m+}}$. To compute $\Delta G_{slv}^{M^{m+}}$ more accurately, we have included a few explicit water molecules as the first solvation shell; specifically, six water molecules are used. Accordingly, $\Delta G_{slv}^{M^{m+}}$ is calculated on the basis of the complex formed with the metal ion and a water hexamer:

$$\Delta G_{slv}^{M^{m+}} \approx \Delta G^{IEFPCM}_{M^{m+}\cdot(H_2O)_6} - \Delta G^{IEFPCM}_{(H_2O)_6} - [G^{(g)}_{M^{m+}\cdot(H_2O)_6} - G^{(g)}_{M^{m+}} - G^{(g)}_{(H_2O)_6}] \quad (2)$$

where the superscript IEFPCM indicates that the solvation free energy of the corresponding species is calculated using the IEFPCM approach[63−65] and B3LYP functional implemented in Gaussian03;[66] the UFF atomic radii[67] are used. The basis sets and effective core potentials used in the geometry optimization and single-point energy calculations are the same as those used in the QM/MM calculations. We note that our scheme is the same used by Merz and co-workers[68] as well as by Goddard and co-workers,[69,70] in which the free energy of cluster formation in solution is zero except for a standard-state

(51) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(52) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(53) Andrae, D.; Haeussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123–141.
(54) Haeussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Mol. Phys.* **1993**, *78*, 1211–1224.
(55) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1987**, *86*, 866–872.
(56) Andrae, D.; Haeussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chem. Acc.* **1990**, *77*, 123–141.
(57) Schwerdtfeger, P.; Dolg, M.; Schwarz, W. H. E.; Bowmaker, G. A.; Boyd, P. D. W. *J. Chem. Phys.* **1989**, *91*, 1762–1774.
(58) Stote, R. H.; Karplus, M. *Proteins: Struct., Funct. Genet.* **1995**, *23*, 12–31.
(59) Sherwood, P.; et al. *J. Mol. Struct. Theochem* **2003**, *632*, 1–28.
(60) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509.
(61) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.
(62) Baker, N.; Sept, D.; Joseph, S.; Holst, M.; McCammon, J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037–10041.
(63) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151–5158.
(64) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.
(65) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253–260.
(66) Frisch, M. J.; et al. *Gaussian 03*, revision E.01; Gaussian, Inc.: Pittsburgh, PA, 2003.
(67) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.
(68) da Silva, E. F.; Svendsen, H. F.; Merz, K. M., Jr. *J. Phys. Chem. A* **2009**, *113*, 6404–6409.
(69) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A., III. *J. Phys. Chem. B* **2008**, *112*, 9709–9719.
(70) Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A., III. *J. Phys. Chem. A* **2008**, *112*, 9104–9112.

Table 1. Calculated Metal Ion Solvation Energies (in kcal/mol)[a]

	computed ΔG_{hyd}	experimental ΔG_{hyd}	
Cu^+	−129.5	N/A	−125.5
Ag^+	−102.4	−107.0	−102.8
Au^+	−132.6	N/A	−137.4
Zn^{2+} [b]	−493.2 (−474.0)	−484.6	−467.3
Hg^{2+}	−435.5	−436.3	−420.7

[a] The hydration free energies of metal ions are given by the free energy of solvated Metal$-(H_2O)_6$ cluster and $(H_2O)_6$ following the same method and energy corrections used by Merz et al.[68] (also see eq 2). The theory is B3LYP/SDD, 6-311++G(2d,2p)//B3LYP/Lanl2dz (def2-SV(P) for Hg), 6-31+G(d) level of methodology. Solvation correction is obtained by using IEFPCM. Zero-point correction and thermal and entropic corrections for the gas-phase contributions are obtained with rigid-rotor harmonic oscillator models. For the experimental values, the first column are from ref 71, and the second column from ref 75; note the significant deviations, especially for the divalent ions. [b] Data in parentheses are obtained by using Lanl2dz instead of SDD for Zn^{2+} in the energy calculations. See Supporting Information regarding he discussion of Lanl2dz vs SDD for Zn^{2+}-related calculations.

correction factor for a water cluster in bulk water, as clearly discussed by Bryantsev et al.;[69] we also adopt the same standard-state correction as in refs 68 and 69. Both groups found that using the water-cluster-based thermodynamic cycle gives consistent solvation free energies for ionic species,[68,69] although the optimal number of water molecules to include in the cluster is not clear. In this study, we use a water hexamer since a similar size water cluster was found to give reliable results in the work of Merz and co-workers.[68]

As shown in Table 1, such calculated solvation free energies for the coinage ions of interest are in fairly good agreement with experimental values. For the divalent ions, especially for Zn^{2+}, the difference is more notable. We note that the experimental solvation free energies for metal ions have been estimated from the solvation free energies of relevant salts[71] and therefore have considerable uncertainties up to tens of kcal/mol. For example, even for the "simple" and prevalent Na^+, the solvation energy has been experimentally estimated to be −98.2 kcal/mol by Lim,[72] −88.6 kcal/mol by Schmid,[73] and −105.1 kcal/mol by Tissandier;[74] for the divalent ions, the two sets of values from refs 71 and75 differ by almost 20 kcal/mol, which highlights the difficulty of absolute binding calculations for these ions. For our work, we use the values from ref 71 as reference for gauging the method for the divalent ions because they were used in previous active-site model analyses of divalent ion bindings.[22,76] The most important point for this work is that the calculations agree with experiments in terms of the relative trends within each series. Moreover, the solvation free energy of the metal ion is canceled out when we consider the protein contribution to binding affinity and selectivity by comparing active-site and QM/MM models, which is a major motivation of this work.

2.2. Active-Site Models with an Implicit Protein Environment. To reveal whether the protein environment makes a significant contribution to the binding affinity, and more importantly, binding selectivity, of metal ions, an active-site model is also studied. The binding affinity can still be cast into the form of eq 1, except that the "protein" is simplified into two ethyl mercaptans (C_2H_5SH); similar to the full enzyme case, different protonation states are considered for the metal-bound state. In the "apo" state, to be consistent with the QM/MM calculations, a dimer consisting of two ethyl mercaptans hydrogen-bonded together is used; this choice (as opposed to using two infinitely separated ethyl mercaptans) makes the entropic component of binding much more consistent between the active-site and QM/MM models. The $\Delta G^{(g)}_{bind}$ is calculated using pure QM (B3LYP with the same basis sets as the full protein calculations), and the solvation free energies for the "bound" and "apo" states are calculated using B3LYP and IEFPCM; entropic and thermal vibrational contributions to the binding affinity are estimated using a rigid-rotor harmonic oscillator model[77] in gas-phase frequency calculations. To approximately analyze the contribution of binding site polarity, similar to the earlier work of Lim and co-workers,[21] several dielectric constants ranging from 4 to 20 have been used in the calculations of $\Delta G^{M^{m+}\cdot P'}_{slv}$ and ΔG^{P}_{slv}. In addition, calculations are done with $\varepsilon = 78.5$; in this case, the active-site model can be regarded as a model for metal binding to GSH, the most abundant thiol in the cell. All calculations for the active-site model are done with Gaussian03.[66]

3. Results and Discussion

In this section, we first discuss results using the active-site models, which serve as a useful reference. Next, we present results from QM/MM calculations with the explicit protein environment. Comparison between the two sets of results helps distinguish intrinsic differences between different metal ions and contributions from the protein to binding affinity and selectivity. We emphasize again that, since we are mostly interested in relative trends, the entropic contributions to binding affinity have not been included here; by convention, throughout this work, negative (positive) values indicate favorable (unfavorable) binding.

3.1. Active-Site Models. With the active-site models, the general trend is that the monovalent ions (Cu^+, Ag^+, and Au^+) have higher binding affinities with one of the Cys residues ionized in the bound state than when both Cys remain charge neutral, while the binding affinities are similar if both Cys are ionized in the bound state (see Table 2). This trend is most clear if the active site is in a nonpolar (low dielectric constant) environment. For example, with $\varepsilon = 4$, the calculated binding affinity of Cu^+, with one Cys deprotonated, is −19.3 kcal/mol; the values are −3.5 and −12.5 kcal/mol with two neutral and deprotonated Cys residues, respectively. As the active site becomes more polar, the dependence on the protonation state of Cys becomes less significant. With $\varepsilon = 20$, for example, the calculated binding affinity of Cu^+ is −13.6, −21.7, and −22.8 kcal/mol with 0, 1, and 2 Cys deprotonated, respectively.

The divalent metal ions (Zn^{2+} and Hg^{2+}), by contrast, strongly prefer a bound state in which both Cys ligands are deprotonated. For example, even with $\varepsilon = 20$, the calculated binding affinity of Hg^{2+} is −21.0, −48.8, and −72.1 kcal/mol with 0, 1, and 2 Cys deprotonated, respectively. Therefore, for both monovalent and divalent ions, the binding affinity is the highest when the bound state is charge neutral, although the preference is much stronger with the divalent ions. As shown by the individual components of binding affinity in Table 2, the intrinsic binding affinity between a metal ion and Cys ligands in the gas phase clearly increases as more Cys residues are deprotonated. This is even true for monovalent ions; e.g., the binding affinity between Cu^+ and two neutral Cys ligands is −95.2 kcal/mol, and the value increases to −255.4 kcal/mol when both Cys ligands are deprotonated. However, since the total binding

(71) Burgess, M. A. *Metal Ions in Solution*; Ellis Horwood: Chichester, England, 1978.
(72) Babu, C. S.; Lim, C. *J. Phys. Chem. B* **1999**, *103*, 7958–7968.
(73) Schmid, R.; Miah, A. M.; Sapunov, V. N. *Phys. Chem. Chem. Phys.* **1999**, *2*, 97–1028.
(74) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V. *J. Phys. Chem. A* **1998**, *102*, 7787–7794.
(75) Marcus, Y. *Biophys. Chem.* **1994**, *51*, 111–127.
(76) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2006**, *128*, 1553–1561.
(77) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1973.

Table 2. Calculated Binding Affinities with the Active-Site Model (in kcal/mol)

	gas-phase affinity[a]	solvation contribution[b]	total binding energy[c]
		Thiol−Thiol ($n = 0$)	
Cu^+	−95.2	91.7(85.5)[81.6]	−3.5(−9.7)[−13.6]
Ag^+	−71.8	63.4(56.6)[52.2]	−8.4(−15.3)[−19.6]
Au^+	−109.7	95.2(89.0)[85.1]	−14.5(−20.7)[−24.6]
Zn^{2+}	−248.6	342.5(313.2)[294.0]	94.0(64.6)[45.4]
$Zn^{2+}\cdot W_2$[d]	−287.5	363.2(340.5)[326.0]	75.7 (53.0)[38.6]
Hg^{2+}	−250.3	280.8(249.9)[229.3]	30.5(−0.4)[−21.0]
		Thiolate−Thiol ($n = 1$)	
Cu^+	132.6/*−215.0*	−151.8(−153.3)[−154.3]	−19.3(−20.7)[−21.7]
Ag^+	155.7/*−191.8*	−179.8(−181.7)[−183.0]	−24.0(−25.9)[−27.3]
Au^+	112.5/*−235.1*	−147.5(−148.7)[−149.6]	−35.0(−36.2)[−37.1]
Zn^{2+}	−138.1/*−485.7*	172.0(161.7)[154.6]	33.8(23.6)[16.5]
$Zn^{2+}\cdot W_2$[d]	−152.1/*−499.6*	179.5(172.2)[167.3]	27.5(20.2)[15.3]
Hg^{2+}	−136.4/*−483.9*	108.7(96.3)[87.6]	−27.7(−40.1)[−48.8]
		Thiolate−Thiolate ($n = 2$)	
Cu^+	439.8/*−255.4*	−452.3(−458.6)[−462.6]	−12.5(−18.8)[−22.8]
Ag^+	463.9/*−231.2*	−478.9(−485.3)[−489.3]	−15.0(−21.4)[−25.4]
Au^+	418.7/*−276.4*	−447.9(−454.0)[−457.9]	−29.2(−35.3)[−39.2]
Zn^{2+}	61.1/*−634.0*	−63.5(−66.4)[−68.7]	−2.4(−5.3)[−7.6]
$Zn^{2+}\cdot W_1$[d]	62.9/*−632.2*	−56.3(−59.1)[−61.1]	6.6(3.8)[1.8]
$Zn^{2+}\cdot W_2$[d]	66.5/*−628.6*	−60.5(−62.0)[−63.2]	6.0(4.5)[3.3]
Hg^{2+}	63.7/*−631.4*	−127.0(−132.0)[−135.8]	−63.3(−68.3)[−72.1]

[a] The gas-phase binding affinity is given by $G^{(g)}_{M^{m+}(C_2H_5S)_2(H)_{2-n}} - G^{(g)}_{(C_2H_5SH)_2} - G^{(g)}_{M^{m+}} + nG^{(g)}_{H^+}$. Values after the slash in italics are modified with the proton affinity of the thiol (i.e., the intrinsic binding energy of the ion to the model Cys in the same protonation state as in the metal-bound state). Zero-point correction and thermal and entropic corrections are obtained by gas-phase frequency analysis with rigid-rotor harmonic oscillator models. [b] The solvation contribution is given by $\Delta G^{M^{m+}(C_2H_5S)_2(H)_{2-n}}_{slv,\varepsilon} + nG^{H^+}_{slv,aq,pH\,7} - \Delta G^{(C_2H_5SH)_2}_{slv,\varepsilon=5} - \Delta G^{M^{m+}}_{slv,aq}$, where n (0, 1, or 2) is the number of proton(s) released during binding and ε is the dielectric constant for the active site; values in parentheses and brackets are for $\varepsilon = 8$ and 20, respectively. The solvation free energy of the gas-phase proton to pH 7 solution is taken to be −272.2 kcal/mol. [c] The total binding energy is the sum of gas-phase binding affinity and solvation contributions for $\varepsilon = 4$, 8, and 20, respectively. [d] One or two water molecules are added in the model to allow triangle or tetrahedron coordination of the Zn^{2+} ion. The *apo* state is taken as a C_2H_5SH dimer and one or two free water molecules.

energy also includes contributions from Cys deprotonation and solvation, different components compensate each other to give the trend mentioned above; i.e., the active-site model prefers charge neutrality for the metal-bound state. Also, we note that, for the charge-neutral bound state, the calculated binding affinity is fairly insensitive to the value of ε; for monovalent ions, the binding affinity changes by only 2 kcal/mol as ε increases from 4 to 20, and even for Hg^{2+}, the binding affinity increases by 8 kcal/mol (out of 60−70 kcal/mol) as ε increases from 4 to 20.

With the most favorable protonation pattern of the bound state, the binding affinity follows the order $Hg^{2+} > Au^+ > Ag^+ \sim Cu^+ > Zn^{2+}$. For the coinage metals, the strongest binding affinity of Au^+ to the Cys residues is consistent with its softer nature compared to that of Cu^+ and Ag^+, owing largely to the more significant relativistic effects on the electronic structure of Au^+; the hardness of Au^+, Cu^+, and Ag^+ is 5.6, 6.28, and 6.96 eV, respectively,[78] which also follows the same trend as the intrinsic gas-phase binding affinity (Table 2) of these metal ions to Cys ligands, regardless of the protonation state of the latter. Although Ag^+ has a weaker gas-phase binding affinity than Cu^+, it has a stronger binding affinity to the active site due to the lower solvation free energy. Similarly, the lowest binding affinity of Zn^{2+} is consistent with its high desolvation free energy penalty. Although Zn^{2+} binds the ligands with a magnitude of interaction similar to that of Hg^{2+} [e.g., 634.0 vs 631.4 kcal/mol to two deprotonated Cys ligands, which is somewhat unexpected since Zn^{2+} has a substantially larger hardness[78] (10.88 eV) than Hg^{2+} (7.7 eV); as discussed in the Supporting Information, the current QM method tends to overestimate the binding affinity of Zn^{2+} to CH_3S^-, which, in fact, further supports our key observation that Zn^{2+} has an unfavorable binding affinity], the much higher desolvation penalty [Table 1, experimental values of 484.6 vs 436.3 kcal/mol for Zn^{2+} and Hg^{2+}, respectively; again, this is consistent with the fact that water has a hardness (9.5 eV[78]) closer to that of Zn^{2+} than to that of Hg^{2+}] leads to a difference of ∼60 kcal/mol in the solvation contribution to binding affinity (see Table 2). As a result, the binding affinity of Hg^{2+} is substantially higher than that of Zn^{2+}; in fact, the calculated binding affinity is either unfavorable or only slightly favorable (less than 10 kcal/mol) for Zn^{2+} for all protonation patterns of the Cys ligands. This conclusion is not changed even when two water molecules are included to make the coordination sphere of Zn^{2+} a tetrahedron, which is the favored coordination environment for Zn^{2+}.[1] This is probably due in part to the electrostatic repulsion between the negatively charged Cys ligands in the tetrahedron coordination configuration; as shown in Figure 3e, the Cys−Zn−Cys angle is 151.5° when both Cys ligands are deprotonated, which is substantially larger than the value of 124.2° when both Cys ligands are charge-neutral. Moreover, there is an entropic penalty associated with sequestering two water molecules from bulk solution.

3.2. QM/MM Calculations. With the full protein model, there are both similarities and differences in the results compared to the active-site models. Structurally, the optimized geometrical parameters in the QM/MM calculations (Figure 4) are very close to those from active-site models (Figure 3a−d), which indicates the lack of significant steric effects in the binding site. These optimized values also compare favorably with the bond distances from the crystal structures for all three coinage metals (summarized in Supporting Information), especially considering the uncertainty in these experimental distances, given the ∼2−2.5 Å resolution of the crystal structures.[18]

For the binding affinities of the coinage metals, the overall trend predicted from the full protein model is similar to that from the active-site models: $Au^+ > Ag^+ \sim Cu^+$, which suggests that the intrinsic binding property and desolvation penalty of these metal ions (e.g., hardness) have not been overcome by the protein environment; this result is qualitatively consistent with transcription induction experiments,[18] which found that all three coinage metals were able to activate transcription through binding to CueR. On the other hand, although the active-site models predict that the bound state with two neutral Cys residues has a fairly significant binding affinity to the coinage metals, the full protein calculations suggest that deprotonation of Cys is essential to metal binding. In fact, the calculated binding affinity is positive (unfavorable) when both Cys remain charge neutral in the bound state. With the limited sampling performed here, the bound state that contains one or two deprotonated Cys residues has large binding affinities for the coinage metals (see Table 3 and the perturbative analysis below). While whether this trend holds with more conformational sampling remains to be analyzed in the future, our calculations highlight that one should not automatically assume that both Cys residues are deprotonated in the metal-bound state. The calculated binding affinity of Cu^+ is −15.8/−21.3 and −24.4 kcal/mol with one

(78) Parr, R. G.; Yang, W. T. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

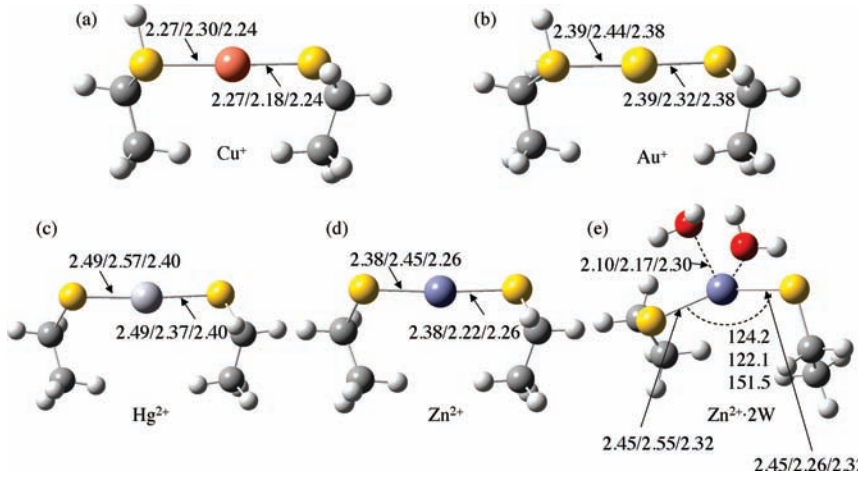

Figure 3. Representative optimized structures for the metal-bound state with the active-site model. Distances and angles are in angstroms and degrees, respectively. For Zn^{2+}, a model is also studied in which two water molecules are included to allow a tetrahedral coordination geometry. For each case, the three values separated by slashes correspond to different protonation patterns with n (0, 1, and 2) model Cys residue(s) deprotonated.

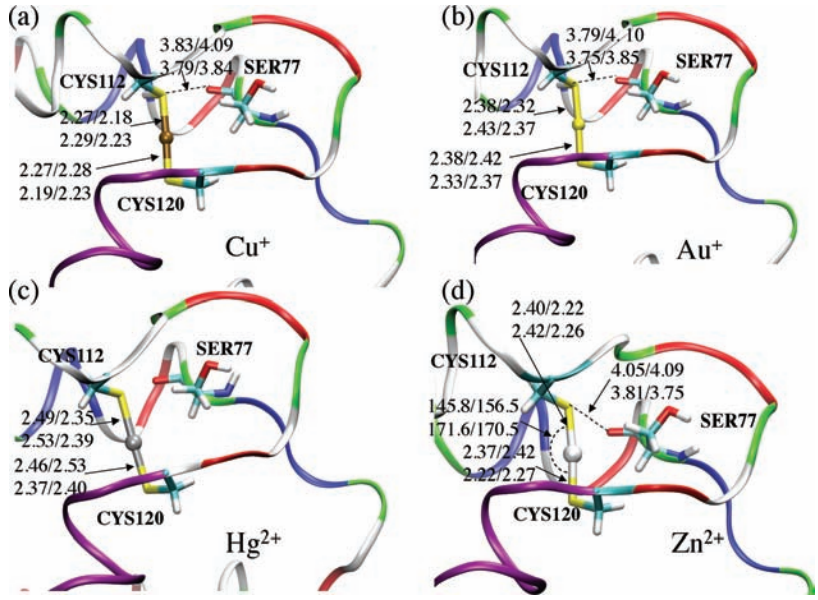

Figure 4. Representative optimized structures for the active site of the metal-bound state with the QM/MM model; note that the region shown is only part of the QM region used in the QM/MM calculations, which also includes main-chain interactions that stabilize the Cys ligands (see Computational Methods). Distances and angles are in angstroms and degrees, respectively. For Zn^{2+}, a model is also studied in which two water molecules are included to allow a tetrahedral coordination geometry. For each case, the four values separated by slashes correspond to different protonation patterns with n (0, 1, and 2) Cys residue(s) deprotonated; for $n = 1$, either Cys112 or Cys120 can be deprotonated. The protein in ribbon form is colored-coded, based on residue type (red, acidic; blue, basic; green, polar; white, nonpolar), except for the helix dipole near the binding site, which is highlighted in purple.

and two deprotonated Cys residues, respectively. In other words, the protein only strongly selects against a net positive charge in the binding site, while being accommodating to charge neutrality or a net negative charge in the binding site (see more discussions below on perturbative analysis); this difference compared to the predictions from active-site models suggests that the protein cannot be described as a dielectric medium for a thorough understanding of metal binding affinity (also see below for the discussion of perturbative analysis).

At a quantitative level, the calculated binding affinities of Cu^+ and Au^+ to CueR compare fairly well to the measured K_d in the recent study by He and co-workers.[11] Converting the measured K_d values to standard binding free energies gives -28.6 and -47.9 kcal/mol for Cu^+ and Au^+, respectively; these can be compared to the current best estimates of -24.4 and -39.3 kcal/mol, respectively, which is encouraging if we consider the uncertainty in the (even experimental) solvation free energy of ions and the limited amount of conformational sampling performed here.

For the divalent ions, the preference for the fully deprotonated Cys ligands is even more apparent with the full protein model. With the active-site model, Hg^{2+} has a favorable binding energy for all protonation patterns of the bound state (with the only exception of two neutral Cys residues and the dielectric of 4). In the protein, however, only the fully deprotonated binding site has a significant binding affinity to Hg^{2+}. Even in this case, the absolute binding affinity is reduced to ~ -42 kcal/mol, which is substantially weaker compared to the active-site model predictions of $\sim -60-72$ kcal/mol. For the zinc ion, similar to the prediction of the active-site model, the calculated binding affinity is unfavorable with all protonation patterns, which is consistent with observation from the transcription induction

Table 3. Calculated Binding Affinities with the QM/MM Model (in kcal/mol)[a]

	gas-phase QM/MM Affinity[b]	solvation contribution[c]	total binding energy[d]
	Thiol−Thiol ($n = 0$)		
Cu^+	−212.2/−117.0	232.6(228.9)/140.9	20.4/23.9
Ag^+	−189.1/−117.2	205.1/141.7	16.0/24.5
Au^+	−226.4/−116.7	232.8/137.6	6.3/20.9
Zn^{2+}	−516.5/−268.0	664.1/321.6	147.6/53.6
Hg^{2+}	−511.2/−260.9	606.5/325.7	95.3/64.8
	Cys112 Thiolate−Cys120 Thiol ($n = 1$)		
Cu^+	124.9/−7.7	−140.7(−142.3)/11.1	−15.8/3.4
Ag^+	148.6/−7.1	−167.6/12.2	−19.0/5.1
Au^+	105.9/−6.6	−136.5/11.0	−30.6/4.4
Zn^{2+}	−258.7/−120.5	322.6/150.6	63.9/30.1
Hg^{2+}	−253.3/−116.9	264.8/156.1	11.5/39.2
	Cys112 Thiol−Cys120 Thiolate ($n = 1$)		
Cu^+	129.3/−3.3	−150.6(−149.3)/1.2	−21.3[e]/−2.1
Ag^+	153.2/−2.5	−177.6/2.2	−24.4/−0.4
Au^+	109.7/−2.8	−148.2/−0.7	−38.5/−3.5
Zn^{2+}	−253.2/−115.0	310.2/138.2	57.0/23.2
Hg^{2+}	−249.8/−113.4	253.9/145.2	4.2/31.9
	Thiolate−Thiolate ($n = 2$)		
Cu^+	524.0/84.2	−548.4(−544.9)/−96.1	−24.4/−11.9
Ag^+	551.0/87.1	−576.0/−97.1	−25.0/−9.9
Au^+	504.8/86.1	−544.1/−96.2	−39.3/−10.1
Zn^{2+}	61.0/−0.1	−50.7/12.8	10.3/12.7
$Zn^{2+}\cdot W_1$[f]	59.7/−3.2	−40.9/15.4	18.8/12.2
Hg^{2+}	65.5/1.8	−108.2/18.8	−42.7[e]/20.6

[a] Values after the slash are the difference between the results from QM/MM and active-site models (Table 2); negative values indicate that the specific component is more favorable in the QM/MM model. [b] The gas-phase binding affinity is given by $G_{M^{m+}\cdot P'}^{QM/MM} - G_{P}^{QM/MM} - G_{M^{m+}}^{gas} + nG_{H^+}^{(g)}$. Zero-point correction and thermal and entropic corrections are from the corresponding active-site model calculations. [c] The solvation contribution is given by $\Delta G_{slv(PB)}^{M^{m+}\cdot P'} + n\Delta G_{slv,aq,pH\ 7}^{H^+} - \Delta G_{slv(PB)}^{P} - \Delta G_{slv,aq}^{M^{m+}}$, where n (0, 1, or 2) is the number of proton(s) released during binding. The dielectric constant of the protein is taken to be 1.0 in PB calculations; values in parentheses are obtained with a dielectric constant of 2.0. [d] Total binding energy is the sum of gas-phase QM/MM binding affinity and solvation contribution (see Figure 2). [e] We test the effect of using different snapshots from the MD run as the initial structures of QM/MM calculations. For five snapshots for the Cu−thiolate−Cys120 case, the changes in calculated binding energy are 3.8, 5.7, 2.6, 6.2, and −3.7 kcal/mol (positive means decrease in binding affinity). For five snapshots for the Hg−thiolate−thiolate case, the changes are −3.7, −3.6, 1.5, −3.3, and −1.4 kcal/mol. [f] One water molecule is included in the binding site of the zinc-bound form of the protein following MD simulation, which together with the backbone carbonyl of Asp 119 allows Zn^{2+} to adopt a tetrahedral coordination (see Figure 5). These groups are included in the QM region. The *apo* state is taken as the *apo* protein plus one free water molecule.

assay.[18] The calculated binding affinity remains unfavorable when water molecules are included as additional ligands to satisfy the tetrahedral coordination configuration of the zinc ion; this is reasonable to explore since MD simulations suggest that water molecules may penetrate into the metal binding site (Figure 5a). Along this line, we note that Zn^{2+} may also form interactions with the backbone carbonyl of Asp119, which is perhaps why the coordination geometry around Zn^{2+} (Figure 5b) is quite different from that in the active-site model (Figure 3e); the Cys−Zn−Cys angle, for example, is substantially more compressed (∼104.8 vs 151.5°), and the Cys−Zn distances are slightly shorter (2.22−2.25 vs 2.30−2.32 Å) in the protein active site, which may also contribute to the even weaker binding affinity of Zn^{2+} to CueR compared to that in the active-site model. Collectively, these results highlight that the high desolvation penalty for zinc remains a key factor for its lack of binding to the protein.

It is most interesting to compare the calculated binding affinity of Hg^{2+} to the coinage metals. First, the most favorable binding affinity predicted for Hg^{2+} (∼ −42 kcal/mol) is still higher than those calculated for the coinage metals (∼ −20−39 kcal/mol), which appears to be contradictory to the fact that CueR can be activated by only the coinage ions but not Hg^{2+};[18] we note that the relative trend in the calculated binding affinity is *not* sensitive to structural fluctuations (see Table 3 for results averaged over several snapshots). However, we noted that the transcription induction experiments are typically done with buffer solutions; thus, the metal ions are likely bound to not water but other ligands. As an example, we use DTT as the reference ligand in buffer solution; DTT was used in the experimental studies of CueR.[18] As shown in Table 4, the calculated binding affinity of the coinage metal ions to DTT (<20 kcal/mol) is fairly low compared to the calculated values for binding to CueR; by contrast, Hg^{2+} binds strongly to DTT, by more than 15 kcal/mol more than to CueR. If we use the active-site model with $\varepsilon = 78.5$ as the model for metal binding to GSH, we see that the QM/MM calculations also lead to the desired trend; i.e., coinage metal ions (especially Cu^+ and Au^+) have binding affinity to CueR stronger than or comparable to that to GSH, while Hg^{2+} binds to GSH much more tightly than to CueR. Therefore, the binding affinities calculated by the QM/MM model are consistent with the experimentally observed behavior of CueR toward different metal ions under both *in vitro* and *in vivo* conditions.

It is important to note that this is not the case with the active-site model predictions, which lead to calculated binding affinity for Hg^{2+} being too strong compared to that for DTT. Comparing results from the active-site model and the full protein (QM/MM) model, the relative binding affinity of Hg^{2+} vs the coinage metals is reduced from ∼30−50 kcal/mol in the active-site model to only ∼3−18 kcal/mol in the protein. In other words, the protein environment of CueR significantly reduces its ability to bind to Hg^{2+} relative to the coinage metals, which is clearly critical to the function of CueR.

To understand this important trend, we first examine various components of the calculated binding affinity and how they differ between QM/MM and active-site models. As shown in the first column of Table 3, the contribution from direct interaction between the binding site and the MM protein environment (i.e., contributions to gas-phase QM/MM affinity) is fairly small when the binding site in the bound state is charge-neutral; this is true for both monovalent and divalent ions. This observation also suggests that the use of a larger QM region in the QM/MM calculations than in active-site calculations does not contribute significantly to the different binding selectivities computed with the two sets of models. For other protonation patterns of the Cys residues, the protein MM contribution favors a positively charged binding site; for example, the only cases that have positive (unfavorable) protein MM contributions are those with the coinage metal ions and two deprotonated Cys residues. This trend, however, is largely compensated by the solvation contribution, which strongly favors a negatively charged binding site. As a result, the net protein contribution (i.e., the difference between QM/MM and active-site models, shown in the last column of Table 3) for the binding of coinage metal ions is nearly zero (∼ ±1−5 kcal/mol) when the binding site is charge-neutral, and rather negative (favorable, ∼ −10 kcal/mol) when the binding site bears a negative charge (i.e., with two deprotonated Cys). For the divalent ions, the net protein contribution is always substantially positive; the least unfavorable case has both Cys deprotonated, which corresponds to a

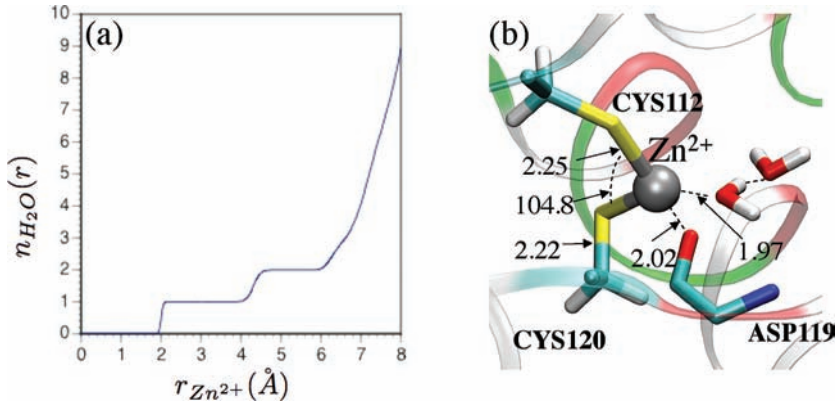

Figure 5. Active-site properties of Zn^{2+}-bound CueR. (a) Integrated number of water molecules near Zn^{2+} from a classical MD simulation (1 ns, see main text). (b) Optimized structure of the active site in which two water molecules are also included in the QM region to allow a tetrahedral coordination sphere around Zn^{2+}. Both Cys112 and Cys120 are deprotonated.

Table 4. Calculated Binding Affinities of Metal Ions with Possible Competing Reagents, DTT and GSH (in kcal/mol)[a]

ions	DTT	GSH
Cu^+	−9.5	−13.0
Ag^+	−9.9	−15.7
Au^+	−17.6	−29.4
Zn^{2+}	−3.6	2.8
Hg^{2+}	−57.8	−62.6

[a] Obtained by the same methodology as for the active-site model; the only difference is that, for GSH, the reference state is two infinitely separated ethyl mercaptans instead of a dimer. All in the most favorable protonation state. For representative structures, see Figure S3 in the Supporting Information.

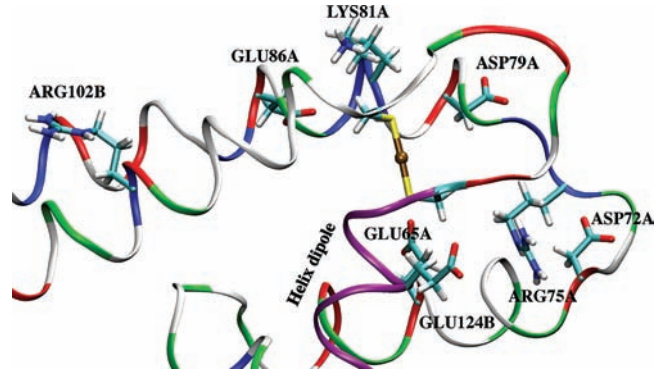

Figure 6. Residues found to be important to metal binding in perturbative analysis and the helix dipole (in purple). The notation "A/B" following the residue number indicates to which of the two chains of CueR that a particular residue belongs.

charge-neutral binding site that still suffers substantial (∼18 kcal/mol for Hg^{2+}) repulsions from the protein compared to an active-site model.

To shed further light on the identity of residues and structural motifs that make the largest contributions to binding, we next turn to results of perturbation analysis, in which partial charges for MM residues are systematically turned off and the difference in the calculated binding affinity is taken as the contribution from the corresponding MM residues; both direct QM/MM and the solvation components are included. A preliminary analysis of the results suggests that, as expected, major contributions are due to charged residues and the helix dipole near the binding site; in the following, we focus on these residues/motifs (for their locations, see Figure 6). For both Cu^+ and Hg^{2+}, the perturbation analysis shows trends consistent with the total binding energy differences (listed in the last column of Table 3); the sums of contributions from those shown in Figure 7 are fairly close to the total differences between the full protein and active-site models, supporting the essential roles of these residues. We note that the hydrogen-bonding interactions from the nearby main-chain NH groups that stabilize the Cys residues are treated at the QM level in the QM/MM calculations and likely make a contribution to the difference between QM/MM and active-site models, such as the different degrees of preference for a specific metal ion to the various protonation patterns of the Cys ligands. However, this contribution is not expected to differ significantly with different metal ions bound to the active site (with a specific protonation state of the Cys residues); thus, we do not consider this contribution to *binding selectivity*.

When the binding site is left with a positive charge (e.g., both Cys remain charge-neutral or Hg^{2+} with one Cys deprotonated), many charged residues and the helix dipole contribute very unfavorably to metal binding; the contributions are most striking for Lys81 and Arg75 from the same chain and for the helix dipole, especially in the case of Hg^{2+}. The unfavorable interactions become much smaller when the binding site is charge-neutral, especially when Cys120 is deprotonated with a bound Cu^+. Interestingly, due to the different orientations of the dipole formed by Cu^+ and the deprotonated Cys112/Cys120, whether certain charged residues contribute favorably or unfavorably to metal binding depends on whether Cys112 or Cys120 is deprotonated (compare Figure 7b and c); this explains why the binding affinity of the coinage metals to CueR tends to be higher when Cys120 is deprotonated than when Cys112 is deprotonated (see Table 3). For the case of Hg^{2+} with both Cys deprotonated, although the net charge of the binding site is zero, interactions from Lys81 and Arg75 still constitute a significant amount of repulsion (∼9.3 kcal/mol). The stronger repulsion from these charged residues compared to the similar cases with the coinage metals is probably because the Hg^{2+} binding site features a large quadrupole moment (two negatively charged Cys surrounding a Hg^{2+}), while a charge-neutral coinage metal binding site is better described with a moderate dipole formed by the monovalent ion and a deprotonated Cys. Finally, with the coinage metal and two deprotonated Cys residues, the binding site bears a negative charge; correspondingly, interactions from nearby residues, again dictated by Lys81, Arg75, and the helix dipole, become favorable for binding by as much as nearly −17 kcal/mol.

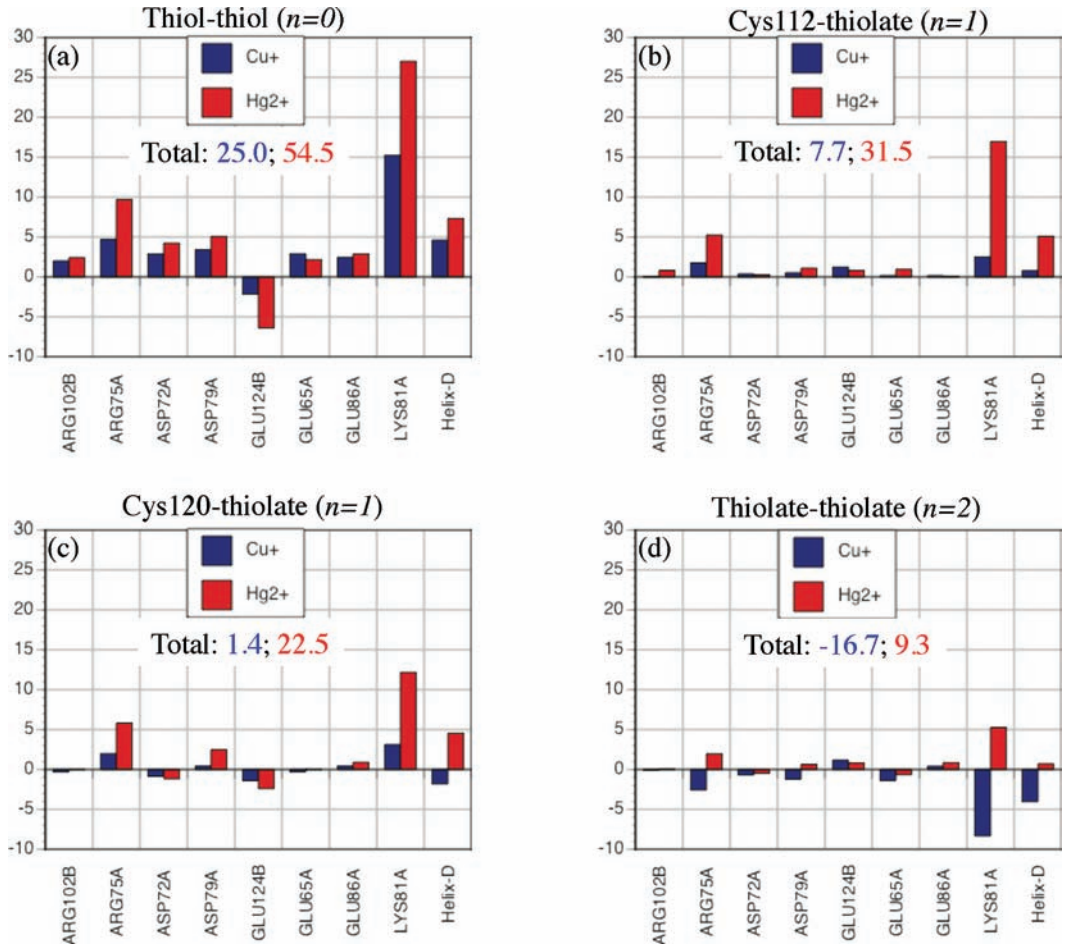

Figure 7. Results of perturbative analysis for the binding of Cu^+ (blue) and Hg^{2+} (red) to CueR with different protonation patterns for the Cys residues (Cys112/120) in the binding site; both QM/MM and solvation components are included in the perturbative analysis. Only results for dominant residues and the helix dipole (see Figure 6) are included. Positive/negative values (in kcal/mol) indicate unfavorable/favorable contributions to binding. "Total" indicates the sum of contributions from each plot, which can be compared to the difference between QM/MM and active-site models (i.e., the protein contribution to binding) listed in the last column of Table 3; they are fairly close, supporting the essential role of these residues/motifs in regulating the relative binding affinity of metal ions.

Therefore, it is clear that the electrostatic environment of CueR has been well tuned to favor coinage metal ions over Hg^{2+}. Even when the binding site is charge-neutral, the protein electrostatic environment disfavors Hg^{2+} rather substantially over the coinage metal ions. Moreover, with the coinage metal ions bound, the binding site can also adopt a negative charge with both Cys deprotonated, which interacts favorably with the protein environment. The selection for such a negatively charged binding site is not unexpected, considering that CueR is a transcription factor whose other ligand is the negatively charged DNA. The fact that the perturbative analysis revealed significant contributions from nearby charged residues and helix dipole highlights once again that important protein contributions would be missing from an active-site model even if the same QM region as in the QM/MM calculations is used.

4. Concluding Remarks: Metal Selectivity in CueR and Related Metalloproteins

Understanding the physical factors that dictate the binding affinity and selectivity of metal ions to metalloproteins is important to many areas of bioinorganic chemistry, such as metal trafficking and rational design of metal-specific sensors, transcription factors, and enzymes. Several factors, such as electrostatics and active-site coordination geometry, are known to contribute, although their relative importance in specific systems is difficult to determine without explicitly evaluating the corresponding contributions to the binding free energy. In this study, we have carried out, to the best of our knowledge, the first computational analysis of transition metal binding affinity and selectivity to a protein with an explicit treatment of the protein environment, using the binding of several transition metal ions (Cu^+, Ag^+, Au^+, Zn^{2+}, and Hg^{2+}) to the transcription factor CueR as an example. The computational framework is based on a combination of QM/MM and PB solvation calculations, which treat both electronic structure of transition metal ions and electrostatics in the protein/solvent environment.

Overall, the binding affinities calculated by using the QM/MM-based models are consistent with transcription induction experiments in that the coinage metal ions are found to bind to CueR with significant affinity, while the divalent ions have weaker affinities to CueR than to other chelating ligands, such as DTT and GSH. This supports the use of the computational methodology established here to analyze factors that dictate binding affinity and selectivity. By contrast, small active-site models do not reproduce all trends, especially concerning the binding affinity of Hg^{2+} and how the protonation pattern of the Cys residues in the active site influences the binding affinity of the coinage metal ions. Perturbative analysis of QM/MM

calculations with a fairly large QM region reveals that charged residues and the helix dipole near the binding site (but beyond the coordination sphere of the metal ion) make substantial contributions to the relative binding affinities of monovalent and divalent ions. Therefore, our study highlights the importance of going beyond active-site models to better understand metal binding in proteins.

Specifically for CueR, analysis of the QM/MM and active-site models has led to the following key insights regarding metal binding affinity and selectivity:

(1) For the coinage metals, the binding affinity to the protein largely follows the intrinsic binding properties, or the softness, of the metal ions and their desolvation penalty, i.e., $Au^+ > Ag^+ \sim Cu^+$. Under *in vivo* conditions, the sensitivity of CueR to Cu^+ is likely due to the higher concentration of Cu^+ than Au^+ and Ag^+ (i.e., "access" [4,5]).

(2) The absolute binding affinity of hydrated Hg^{2+} to CueR is, in fact, larger than that of the coinage metals. However, Hg^{2+} binds more strongly to ligands such as DTT and GSH, which is the reason that its binding to CueR in buffer solution or in cells is unfavorable. This highlights the importance of considering the proper reference when discussing the binding affinity of metal ions to proteins.

(3) The calculated binding affinity of Zn^{2+} is unfavorable with both active-site and protein models, regardless of whether the coordination environment is linear (with two Cys residues) or tetrahedral (with two additional water molecules). Therefore, it is likely that the weak binding of Zn^{2+} to CueR is due largely to the large desolvation penalty for Zn^{2+} compared to other ions (which is consistent with the large hardness value[78] of Zn^{2+}), while coordination geometry is much less important.

(4) The electrostatic environment of the protein is well-tuned to favor the binding of coinage metal ions over divalent ions. In particular, charged residues and the helix dipole near the binding site contribute either favorable or small unfavorable interactions when the binding site with a coinage metal ion is charge-neutral (one Cys deprotonated) or bears a negative charge (both Cys deprotonated). With a divalent ion such as Hg^{2+}, the interaction is more unfavorable, even when the binding site is charge-neutral.

The successful application of our QM/MM-based approach to CueR encourages similar studies of other systems that exhibit interesting binding properties to transition metals, such as those that bind specifically to divalent ions (e.g., ZntR[18]) or more "exotic" ions such as uranyl.[7] A particularly interesting system is the closely related CupR,[11] which has a binding site highly similar to that of CueR but very different relative binding affinities toward Cu^+ and Au^+. It is also important to better understand the limitations of the methodology and therefore identify ways to improve its robustness. For example, the impact of structural fluctuations in the protein has only been briefly explored here by using different initial structures collected from molecular dynamics simulations; a more rigorous test involves comparison to QM/MM-based free energy simulations,[40,79] which have become increasingly more practical due to improvements in both computational hardware and enhanced sampling techniques.[80] Another possibility is to develop effective empirical models for metal binding, which would necessarily include charge transfer[81] and ligand field contributions.[82,83] Once a robust computational methodology for evaluating metal binding affinity is available, rational design of metal-specific proteins will become a realistic possibility.

Acknowledgment. This research has been partially supported by the National Natural Science Foundation of China (Grant Nos. 10774126, 20923004) and the Ministry of Science and Technology (Grant No. 2007CB815206). Q.C. acknowledges support from the National Science Foundation (CHE-0957285).

Supporting Information Available: Discussions of additional benchmark calculations, choice of the QM method and QM region, comparison with crystal structures, optimized structures for the complexes formed between the metal ions and DTT, and complete refs 49, 50, and 66. This material is available free of charge via the Internet at http://pubs.acs.org.

JA103742K

(79) Zhang, Y. K.; Liu, H. Y.; Yang, W. T. *J. Chem. Phys.* **2000**, *112*, 3483–3492.
(80) Zheng, L. Q.; Chen, M. G.; Yang, W. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20227–20232.
(81) Sakharov, D. V.; Lim, C. *J. Comput. Chem.* **2008**, *30*, 191–202.
(82) Deeth, R. J. *Coord. Chem. Rev.* **2001**, *212*, 11–34.
(83) Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B. *J. Am. Chem. Soc.* **2004**, *126*, 1285–1289.