J|A|C|S
ARTICLES
Published on Web 02/15/2003

First−Second Shell Interactions in Metal Binding Sites in Proteins: A PDB Survey and DFT/CDM Calculations

Todor Dudev,[†] Yen-lin Lin,[‡] Minko Dudev,[†] and Carmay Lim*[,†,‡]

Contribution from the Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan R.O.C., and Department of Chemistry, National Tsing Hua University, Hsinchu 300, Taiwan R.O.C.

Received July 16, 2002; Revised Manuscript Received October 29, 2002; E-mail: carmay@gate.sinica.edu.tw

Abstract: The role of the second shell in the process of metal binding and selectivity in metalloproteins has been elucidated by combining Protein Data Bank (PDB) surveys of Mg, Mn, Ca, and Zn binding sites with density functional theory/continuum dielectric methods (DFT/CDM). Peptide backbone groups were found to be the most common second-shell ligand in Mg, Mn, Ca, and Zn binding sites, followed (in decreasing order) by Asp/Glu, Lys/Arg, Asn/Gln, and Ser/Thr side chains. Aromatic oxygen- or nitrogen-containing side chains (Tyr, His, and Trp) and sulfur-containing side chains (Cys and Met) are seldom found in the second coordination layer. The backbone and Asn/Gln side chain are ubiquitous in the metal second coordination layer as their carbonyl oxygen and amide hydrogen can act as a hydrogen-bond acceptor and donor, respectively, and can therefore partner practically every first-shell ligand. The second most common outer-shell ligand, Asp/Glu, predominantly hydrogen bonds to a metal-bound water or Zn-bound histidine and polarizes the H−O or H−N bond. In certain cases, a second-shell Asp/Glu could affect the protonation state of the metal ligand. It could also energetically stabilize a positively charged metal complex more than a neutral ligand such as the backbone and Asn/Gln side chain. As for the first shell, the second shell is predicted to contribute to the metal selectivity of the binding site by discriminating between metal cations of different ionic radii and coordination geometries. The first-shell−second-shell interaction energies decay rapidly with increasing solvent exposure of the metal binding site. They are less favorable but are of the same order of magnitude as compared to the respective metal−first-shell interaction energies. Altogether, the results indicate that the structure and properties of the second shell are dictated by those of the first layer. The outer shell is apparently designed to stabilize/protect the inner-shell and complement/enhance its properties.

Introduction

About half of all proteins contain metal cations[1] and most members of the ribozyme family cannot function without metal cofactors.[2] Metal cations tend to bind to a protein cavity or crevice that is characterized by a low dielectric constant.[3,4] Among all biogenic metal cations Mg(II), Ca(II), Zn(II), and Mn(II) are found most often bound to protein residues.[5−15] These metals play either a predominantly catalytic role or serve only a structural role. Mg, Ca, and Mn, being "hard" cations, prefer to bind directly to oxygen-containing protein ligands such as Asp, Glu, Asn, Gln and backbone carbonyl groups.[8,9,13,16−18] The "border-line" Zn shows stronger preference toward nitrogen and sulfur-containing ligands such as His and Cys, respectively, though in many catalytic binding sites it is found coordinated to Asp/Glu as well.[10,12−16,19,20]

Most studies on metal binding in proteins have focused mainly on the metal−first-shell interactions.[3,4,6,9,10,14,16−57] First-

[†] Academia Sinica.
[‡] National Tsing Hua University.

(1) Thomson, A. J.; Gray, H. B. *Curr. Opin. Struct. Biol.* **1998**, *2*, 155.
(2) Hanna, R.; Doudna, J. A. *Curr. Opin. Chem. Biol.* **2000**, *4*, 166.
(3) Yamashita, M. M.; Wesson, L.; Eisenman, G.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5648.
(4) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2000**, *104*, 3692.
(5) Bertini, I.; Sigel, A.; Sigel, H., Eds.; *Handbook on Metalloproteins*; Marcel Dekker: New York, 2001.
(6) Cowan, J. A. *Biological Chemistry of Magnesium*; VCH: New York, 1995.
(7) Cowan, J. A. *Chem. Rev.* **1998**, *98*, 1067.
(8) Pidcock, E.; Moore, G. R. *J. Biol. Inorg. Chem.* **2001**, *6*, 479.
(9) Dismukes, G. C. *Chem. Rev.* **1996**, *96*, 2909.
(10) Lipscomb, W. N.; Strater, N. *Chem. Rev.* **1996**, *96*, 2375.
(11) Berg, J. M.; Godwin, H. A. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 357.
(12) Christianson, D. W. *Adv. Prot. Chem.* **1991**, *42*, 281.
(13) Christianson, D. W.; Cox, J. D. *Annu. Rev. Biochem.* **1999**, *68*, 33.
(14) Coleman, J. E. *Annu. Rev. Biochem.* **1992**, *61*, 897.
(15) Alberts, I. L.; Nadassy, K.; Wodak, S. J. *Protein Sci.* **1998**, *7*, 1700.
(16) Jernigan, R.; Raghunathan, G.; Bahar, I. *Curr. Op. Struct. Biol.* **1994**, *4*, 256.
(17) Bock, C. W.; Katz, A. K.; Markham, G. D.; Glusker, J. P. *J. Am. Chem. Soc.* **1999**, *121*, 7360.
(18) Dudev, T.; Cowan, J. A.; Lim, C. *J. Am. Chem. Soc.* **1999**, *121*, 7665.
(19) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647.
(20) Lippard, S. J.; Berg, J. M. *Principles of Bioinorganic Chemistry*; University Science Books: Mill Valley, CA, 1994.
(21) Falke, J. J.; Snyder, E. E.; Thatcher, K. C.; Voertler, C. S. *Biochemistry* **1991**, *30*, 8690.
(22) Krauss, M.; Garmer, D. R. *J. Am. Chem. Soc.* **1991**, *113*, 642.
(23) Klobukowski, M. *Can. J. Chem.* **1992**, *70*, 589.
(24) Black, C. B.; Huang, H. W.; Cowan, J. A. *Coord. Chem. Rev.* **1994**, *135*/*136*, 165.
(25) Garmer, D. R.; Gresh, N. *J. Am. Chem. Soc.* **1994**, *116*, 3556.
(26) Ryde, U. *Int. J. Quantum Chem.* **1994**, *52*, 1229.
(27) Bock, C. W.; Katz, A. K.; Glusker, J. P. *J. Am. Chem. Soc.* **1995**, *117*, 3754.
(28) Glendening, E. D.; Feller, D. *J. Phys. Chem.* **1996**, *100*, 4790.
(29) Gresh, N.; Garmer, D. R. *J. Comput. Chem.* **1996**, *17*, 1481.

10.1021/ja0209722 CCC: $25.00 **© 2003 American Chemical Society**

shell ligands play crucial roles in contributing to the metal complex stability and in determining the selectivity of the binding site.[18,21,52,58,59] Furthermore, they may also play a role in altering the metal coordination geometry or the ligand binding mode (mono/bidentate).[47,60,61]

Recently, several studies have highlighted the role of the second shell in the process of protein−metal recognition and metalloprotein function. In particular, analyses of Zn finger structures have shown the importance of the second layer packing (consisting of backbone (bkb) peptide groups or Lys/Arg side chains) in shielding the negatively charged Zn finger cores.[62] Subsequent calculations confirmed this finding and showed that second-shell ligands contribute to the energetic stabilization of the metal complex.[63] Furthermore, an outer-shell carboxylate in buried Zn binding sites has been proposed to act as a proton acceptor rather than a hydrogen-bond acceptor for an inner-shell His.[44,48,64] Gas-phase ab initio calculations predict that the Zn^{2+}−His^0−Asp/Glu^- triad will isomerize to Zn^{2+}−His^- −Asp/Glu(H)^0.[44,47,48] Experimental studies on human carbonic anhydrase II,[64−70] metallophosphatases calcineurin,[71] human serum transferrin,[72] EcoRV restriction endonuclease,[73] and designed metal binding sites[74] suggest that the second-shell ligands play a role in (1) orienting the first-shell partners at proper positions to enhance the affinity of the binding site for the metal and (2) fine-tuning the pK_a and reactivity of the first-shell catalytic water. In the case of type I Cu proteins, Karlin and co-workers have proposed that the second shell provides flexibility in the metal environment that may help electron transfer.[75−77]

Although the above studies reveal certain roles of the second-shell ligands in protein−metal recognition, there are no systematic studies (to the best of our knowledge) and many questions remain. For example, what is the most common second-shell partner corresponding to the first-shell ligands in the metal binding sites? Which second shell residues contribute most to the stability of the metal complex? How strongly does the dielectric medium affect the first-shell−second-shell interaction energies? How do the first-shell−second-shell interaction energies compare with the respective metal−first-shell interaction energies? Does the second shell contribute to the metal selectivity of the binding site? We address these questions here for Mg(II), Ca(II), Zn(II), and Mn(II) binding sites using a combination of Protein Data Bank (PDB) surveys with density functional theory/continuum dielectric methods (DFT/CDM), as outlined in the next section.

Methods

Database Survey. The PDB[78] was surveyed for <3.0 Å X-ray and NMR structures of proteins containing Mg, Mn, Ca, and Zn that play either a structural or catalytic role. Multinuclear binding sites and those containing other cofactors such as phosphate or sulfate groups were excluded. The protein sequences were aligned using the Modeller 4 program,[79] and those with sequence identity higher than 30% were considered to belong to the same protein family. Only one representative from each protein family, namely, the structure solved at the highest resolution, was included in the survey. The PDB entries used in the survey are given in Supporting Information.

Defining First- and Second-Shell Ligands in the Metal Binding Sites. Analysis of high-resolution X-ray structures of small metal complexes in the Cambridge Structural Database has shown that the first-shell M−O, M−N, and M−S bond distances (M = Mg, Mn, Ca, Zn) do not exceed 2.6 Å.[80] To account for the lower resolution of some of the PDB structures, a slightly larger cutoff of 2.9 Å was used to locate the first-shell ligands, which were defined as residues with a donor atom (e.g., N, S, or O) within 2.9 Å from the metal. The heavy atoms of the first-shell residues were then selected as centers to search the second-shell ligands using a cutoff of 3.5 Å.[81] These second-shell ligands were verified to form hydrogen bonds with first-shell residues using the WHATIF program.[82] First-shell water molecules, which may play an important role in catalytic processes, were included in the survey whenever they were seen in the X-ray structures. However, second-

(30) Katz, A. K.; Glusker, J. P.; Beebe, S. A.; Bock, C. W. *J. Am. Chem. Soc.* **1996**, *118*, 5752.
(31) Lee, S.; Kim, J.; Park, J. K.; Kim, K. S. *J. Phys. Chem.* **1996**, *100*, 14329.
(32) Ryde, U. *J. Comput. Aided Mol. Des.* **1996**, *10*, 153.
(33) Ryde, U. *Eur. Biophys. J.* **1996**, *24*, 213.
(34) Yliniemela, A.; Uchimaru, T.; Hirose, T.; Baldwin, B. W.; Tanabe, K. *THEOCHEM* **1996**, *369*, 9.
(35) Deerfield, I. D. W.; Pedersen, L. G. *J. Mol. Struct.* (*THEOCHEM*) **1997**, *419*, 221.
(36) Hartmann, M.; Clark, T.; Eldik, R. V. *J. Am. Chem. Soc.* **1997**, *119*, 7843.
(37) Miura, T.; Satoh, T.; Takeuchi, H. *Biochim. Biophys. Acta* **1998**, *1384*, 171.
(38) Pavlov, M.; Siegbahn, P. E. M.; Sandstrum, M. *J. Phys. Chem. A* **1998**, *102*, 219.
(39) Rulisek, L.; Vondrasek, J. *J. Inorg. Biochem.* **1998**, *71*, 115.
(40) Peschke, M.; Blades, A. T.; Kebarle, P. *J. Phys. Chem. A* **1998**, *102*, 9978.
(41) Sponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951.
(42) Topol, I. A.; Casas-Finet, J. R.; Gussio, R.; Burt, S. K.; Erickson, J. W. *J. Mol. Struct.* (*THEOCHEM*) **1998**, *423*, 13.
(43) Dudev, T.; Lim, C. *J. Phys. Chem. A* **1999**, *103*, 8093.
(44) El Yazal, J.; Pang, Y.-P. *J. Phys. Chem. B* **1999**, *103*, 8773.
(45) Rodriguez-Cruz, S. E.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1999**, *121*, 1986.
(46) Tiraboschi, G.; Roques, B.-P.; Gresh, N. *J. Comput. Chem.* **1999**, *20*, 1379.
(47) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2000**, *122*, 11146.
(48) El Yazal, J.; Roe, R. R.; Pang, Y.-P. *J. Phys. Chem. B* **2000**, *104*, 6662
(49) Peschke, M.; Blades, A. T.; Kebarle, P. *J. Am. Chem. Soc.* **2000**, *122*, 10440.
(50) Rulisek, L.; Havlas, Z. *J. Am. Chem. Soc.* **2000**, *122*, 10428
(51) Topol, I. A.; Nemukhin, A. V.; Chao, M.; Iyer, L. K.; Tawa, G. J.; Burt, S. K. *J. Am. Chem. Soc.* **2000**, *122*, 7087.
(52) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2001**, *105*, 4446.
(53) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2001**, *105*, 10709.
(54) Munoz, J.; Sponer, J.; Hobza, P.; Orozco, M.; Luque, F. J. *J. Phys. Chem. B* **2001**, *105*, 6051.
(55) Deubel, D. V. *J. Am. Chem. Soc.* **2002**, *124*, 5834.
(56) Rulisek, L.; Havlas, Z. *J. Phys. Chem. A* **2002**, *106*, 3855.
(57) Yang, W.; Lee, H.-W.; Hellinga, H.; Yang, J. J. *Proteins*: *Struct. Func. Genet.* **2002**, *47*, 344.
(58) Falke, J. J.; Drake, S. K.; Hazard, A. L.; Peersen, O. B. *Q. Rev. Biophys.* **1994**, *27*, 219.
(59) Drake, S. K.; Lee, K. L.; Falke, J. J. *Biochemistry* **1996**, *35*, 6697.
(60) Ryde, U. *Biophys. J.* **1999**, *77*, 2777.
(61) Ozawa, T.; Fukuda, M.; Nara, M.; Nakamura, A.; Komine, Y.; Kohama, K.; Umezawa, Y. *Biochemistry* **2000**, *39*, 14495.
(62) Maynard, A. T.; Covell, D. G. *J. Am. Chem. Soc.* **2001**, *123*, 1047.
(63) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2002**, *124*, 6759.
(64) Christianson, D. W.; Alexander, R. S. *J. Am. Chem. Soc.* **1989**, *111*, 6412.
(65) Lesburg, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1995**, *117*, 6838.
(66) Kiefer, L. L.; Paterno, S. A.; Fierke, C. A. *J. Am. Chem. Soc.* **1995**, *117*, 6831.
(67) Huang, C.-C.; Lesburg, C. A.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1996**, *35*, 3439.
(68) Cox, E. H.; Hunt, J. A.; Compher, K. M.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **2000**, *39*, 13687.
(69) McCall, K. A.; Huang, C.-C.; Fierke, C. A. *J. Nutr.* **2000**, *130*, 1437S.
(70) DiTusa, C. A.; McCall, K. A.; Christensen, T.; Mahapatro, M.; Fierke, C. A.; Toone, E. J. *Biochemistry* **2001**, *40*, 5345.
(71) Mertz, P.; Yu, L.; Sikkink, R.; Rusnak, F. *J. Biol. Chem.* **1997**, *272*, 21296.
(72) He, Q.-Y.; Mason, A. B.; Woodworth, R. C.; Tam, B. M.; MacGillivray, R. T. A.; Grady, J. K.; Chasteen, N. D. *J. Biol. Chem.* **1998**, *273*, 17018.
(73) Vipond, I. B.; Moon, B.-J.; Halford, S. E. *Biochemistry* **1996**, *35*, 1712.
(74) Marino, S. F.; Regan, L. *Chem., & Biol.* **1999**, *6*, 649.
(75) Karlin, S.; Zhu, Z.-Y. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14231.
(76) Karlin, S.; Zhu, Z.-Y.; Karlin, K. D. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14225.
(77) Karlin, K. D.; Zhu, Z.-Y.; Karlin, S. *J. Biol. Inorg. Chem.* **1998**, *3*, 172.
(78) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *122*, 535.
(79) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
(80) Harding, M. M. *Acta Crystallogr.* **1999**, *D55*, 1432.
(81) McDonald, I.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777.
(82) Vriend, G. *J. Mol. Graph.* **1990**, *8*, 52.

shell water molecules were excluded from the survey because they were often not seen in the X-ray structures, and in some binding sites close to the protein surface the second shell incorporates numerous waters (apparently without a structural role) from the bulk solvent. These two factors would lead to biased statistics and thus artifactual findings.

Models Used. Formate ($HCOO^-$) was used as a model for deprotonated Asp and Glu side chains, while formamide ($HCONH_2$) was employed as a model for Asn and Gln side chains as well as backbone peptide groups. Since Mg and Mn complexes in proteins are predominantly octahedral,[16,17,83,84] we considered clusters of the type $[M(H_2O)_{6-i}L_i]\cdot Q$ where M denotes Mg or Mn, L is an inner-shell $HCOO^-$ or $HCONH_2$ ligand, Q is an outer-shell H_2O, $HCOO^-$, or $HCONH_2$ ligand, $i = 0$ or 1, and a "·" is used to distinguish the inner shell from the outer shell. For Ca, the complexes in proteins seem to prefer heptacoordination[16] so they were modeled as $[Ca(H_2O)_{7-i}L_i]\cdot Q$. For Zn, two types of complexes were modeled: octahedral $[Zn(H_2O)_{6-i}L_i]\cdot Q$ complexes as found in aqueous solution[84] and tetrahedral complexes $[Zn(H_2O)_{4-i}L_i]\cdot(H_2O)_2Q$[16] with two water molecules in the Zn second shell, as found in most Zn binding sites.

Mn, unlike the other three cations included in this study, is an open shell ion with a d^5 outer electron configuration. It has three possible ground-state spin arrangements, low-spin doublet, low-spin quartet, and high-spin sextuplet, depending on the strength and symmetry of the ligand field. Experimental studies show that octahedral Mn complexes, with very few exceptions, are high-spin complexes.[83] Accordingly, a high-spin configuration for Mn was considered here.

DFT Calculations. These employed Becke's three-parameter hybrid method[85] in conjunction with the Lee, Yang, and Parr correlation functional[86] and the 6-31++G(2d,2p) basis set. In previous studies we had calibrated the B3LYP/6-31++G(2d,2p) calculations with respect to available experimental data and showed that they are well suited for evaluating the geometries and interaction free energies of complexes between divalent cations and oxygen- and nitrogen-containing ligands.[47,52] Consequently, *full* geometry optimization for each complex was carried out at the B3LYP/6-31++G(2d,2p) level using the Gaussian 98 program.[87] The optimized coordinates are given in Supporting Information. Vibrational frequencies were then computed at the same level of theory/basis to verify that each complex was at the minimum of its potential energy surface. *No* imaginary frequency was found in any of the complexes. After the frequencies were scaled by an empirical factor of 0.9613,[88] the zero-point energy (ZPE), thermal energy (E_T), work (PV), and entropy (S) corrections were evaluated using standard statistical mechanical formulas.[89] The differences ΔE_{elec}, ΔZPE, ΔE_T, ΔPV, and ΔS between the products and reactants were used to compute the reaction free energy (see eqs 3 to 7 below) at room temperature, $T = 298.15$ K, according to the following expression:

$$\Delta G^1 = \Delta E_{elec} + \Delta ZPE + \Delta E_T + \Delta PV - T\Delta S \qquad (1)$$

Continuum Dielectric Calculations. The reaction free energy in a given environment characterized by a dielectric constant $\epsilon = x$ can be calculated according to the thermodynamic cycle shown in Scheme 1. ΔG^1 is the gas-phase free energy computed using eq 1. $\Delta G_{solv}{}^x$ is the free energy for transferring a molecule in the gas phase to a continuous solvent medium characterized by a dielectric constant, x. By solving Poisson's equation using finite difference methods[90,91] to estimate $\Delta G_{solv}{}^x$ (see below), the reaction free energy in an environment modeled by dielectric constant x, ΔG^x, can be computed from:

Scheme 1

$$\begin{array}{ccc} & \Delta G^1 & \\ \text{Reactants } (\varepsilon=1) & \rightarrow & \text{Products } (\varepsilon=1) \\ \Delta G_{solv}{}^x(\text{Reactants}) \downarrow & & \downarrow \Delta G_{solv}{}^x(\text{Products}) \\ \text{Reactants } (\varepsilon=x) & \rightarrow & \text{Products } (\varepsilon=x) \\ & \Delta G^x & \end{array}$$

$$\Delta G^x = \Delta G^1 + \Delta G_{solv}{}^x(\text{products}) - \Delta G_{solv}{}^x(\text{reactants}) \qquad (2)$$

The continuum dielectric calculations employed a 71 × 71 × 71 lattice with an initial grid spacing of 1.0 Å, refined with a spacing of 0.25 Å, ab initio geometries, and natural bond orbital (NBO) atomic charges.[92] The low dielectric region of the solute was defined as the region inaccessible to contact by a 1.4 Å radius sphere rolling over the molecular surface. This region was assigned a dielectric constant of two ($\epsilon_{in} = 2$) to account for the electronic polarizability of the solute. The molecular surface was defined by *effective solute radii*, which were obtained by adjusting the CHARMM (version 22)[93] van der Waals radii to reproduce the experimental hydration free energies of the metal cations and ligands. The "effective" radius of Mn (1.55 Å) was optimized in this study to reproduce its experimental hydration free energy of −437.8 kcal/mol.[94] The other solute radii had been optimized in our previous studies[4,47,52] and are as follows (in Å): $R_{Zn} = 1.40$, $R_{Mg} = 1.50$, $R_{Ca} = 1.75$, $R_O(HCOO^-) = 1.65$, $R_O(H_2O) = 1.69$, $R_O(HCONH_2) = 1.79$, $R_N = 1.7$, $R_C = 1.9$, $R_H(H_2O) = 1.0$, $R_H(C, N) = 1.468$. These effective solute radii reproduce (to within 2.5%) the experimental hydration free energies for the metal cations, $HCOO^-$, and $HCONH_2$.[4,47,52] Buried or partially buried metal binding sites were characterized by an external dielectric constant ϵ_{out} equal to 2 or 4, respectively.[90,95] Thus, Poisson's equation was solved with ϵ_{out} equal to 1, 2, or 4 and $\epsilon_{in} = 2$. The difference between the computed electrostatic potentials in a given dielectric medium ($\epsilon = x$) and in the gas phase ($\epsilon = 1$) yielded the solvation free energy $\Delta G_{solv}{}^x$ of the metal complex.

Results

PDB Survey of First-Shell Ligands. The percentage frequency distributions of the first-shell ligands in Mg, Ca, Mn, and Zn binding sites are given in Figure 1. The backbone (BKB) peptide groups coordinate to the metal cation via the carbonyl oxygen atoms. As expected, the "border-line" Zn gives a pattern of ligand distribution that is distinct from the "hard" Mg, Mn, and Ca cations. Figure 1 shows that His and Cys are by far the most preferred inner-shell ligands for Zn, whereas Asp/Glu,

(83) Cotton, F. A.; Wilkinson, G. *Advanced Inorganic Chemistry*; John Wiley & Sons: New York, 1980.
(84) Marcus, Y. *Chem. Rev.* **1988**, *88*, 1475
(85) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(86) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.
(87) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, M. C.; Strain, K. N.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, *Revision A.5*,; Gaussian, Inc.: Pittsburgh, 1998.
(88) Wong, M. W. *Chem. Phys. Lett.* **1996**, *256*, 391.
(89) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976.
(90) Gilson, M. K.; Honig, B. *Biopolymers* **1986**, *25*, 2097.
(91) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610.
(92) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.
(93) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(94) Burgess, M. A. *Metal ions in solution*; Ellis Horwood: Chichester, England, 1978.
(95) Harvey, S. C.; Hoekstra, P. *J. Phys. Chem.* **1972**, *76*, 2987.

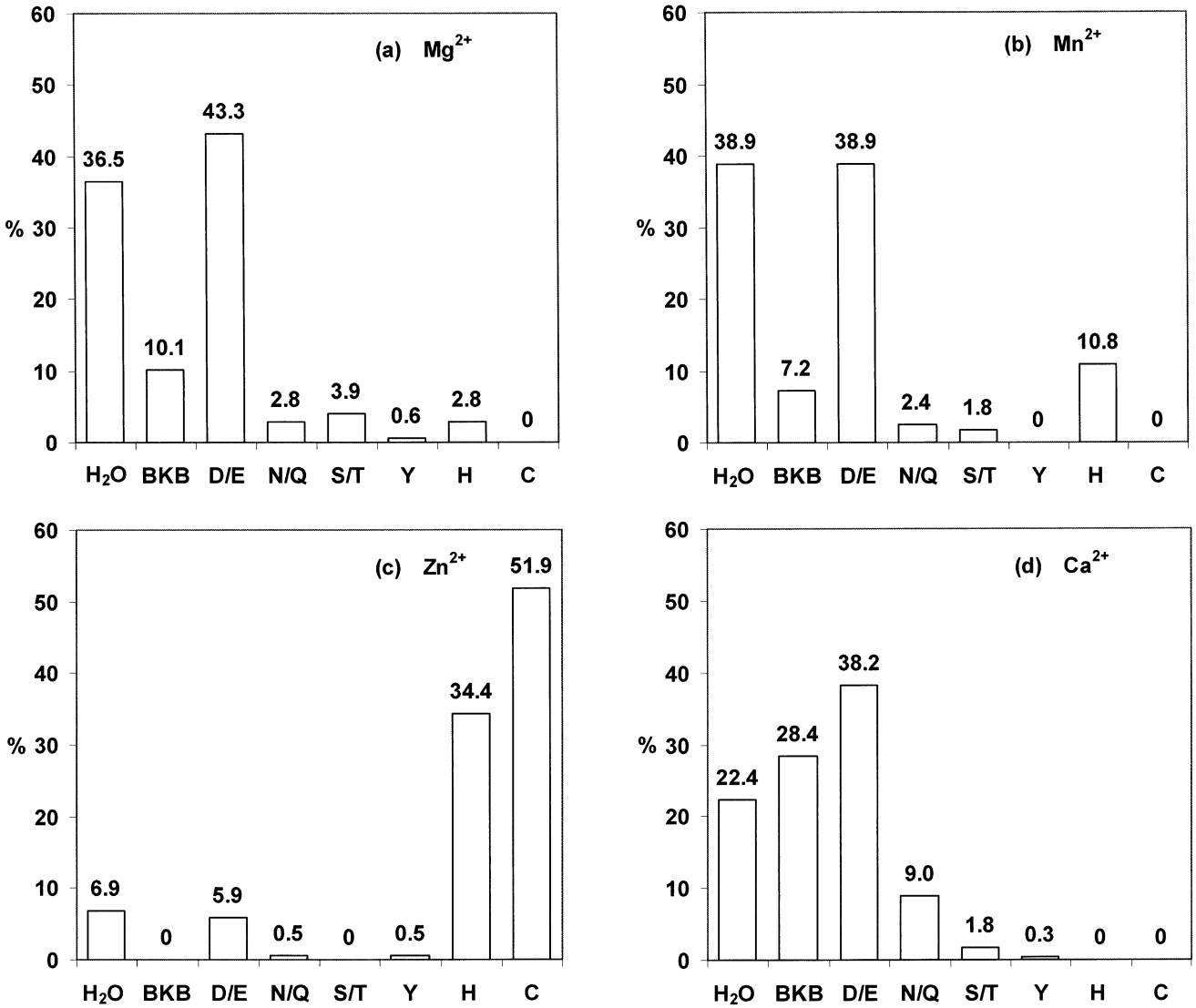


Figure 1. Percentage frequency distribution of first-shell ligands observed in the 3-dimensional structures of proteins bound to (a) Mg, (b) Mn, (c) Zn, and (d) Ca. The total numbers of first-shell ligands bound to Mg, Mn, Zn, and Ca are 194, 260, 505, and 437, respectively.

water, and CONH-containing ligands (BKB and Asn/Gln) dominate the Mg, Mn, and Ca binding sites. Not only is the ligand type common for Mg and Mn but also the occurrence frequency of these "common" ligands. The key difference between the first-shell ligand distributions of Mg and Mn is the affinity for His, which is significantly higher for Mn (10.8%) than for Mg (2.8%). For Ca, however, the relative occurrence frequencies of the "common" ligands differ from those of the other two "hard" cations. Whereas Mg and Mn strongly prefer water to ligands with amide groups, such a preference is not found for Ca, which, on the contrary, has a higher percentage of first-shell peptide backbone and Asn/Gln side chains (37.4%) than water (22.4%).

PDB Survey of Second-Shell Ligands. The percentage frequency distributions of the second-*shell* ligands in Mg, Ca, Mn, and Zn binding sites are given in Figure 2. For each metal, the number of ligands in the first shell is greater than that in the second shell: the ratio between the number of first- and second-shell ligands is 1.2−1.7 for the "hard" cations, and 2.1 for Zn. This indicates that not every first-shell ligand has a second-shell partner from the protein matrix. Despite the significant differences observed in the first-layer distribution of Mg, Ca, Mn, and Zn (see above), the overall second-shell distributions for the four metals appear to be similar. Backbone peptide groups are dominant (52−66%), followed by the side chains of Asp/Glu (14−26%), Lys/Arg (3−13%), Asn/Gln(4−10%), and Ser/Thr (2−5%). The aromatic hydrogen-bond donor/acceptor side chains (Tyr, His, and Trp) are seldom found in the metal second shell (<4%), while even weaker hydrogen-bond donor/acceptor sulfur-containing side chains (Cys and Met) were not found within 3.5 Å of the first-shell heavy atoms (see Methods). As observed for the frequency distributions of the first-shell ligands, Mg and Mn give very similar patterns. However, they differ from Ca, which exhibits a stronger preference than Mg/Mn for Asp/Glu compared to Lys/Arg in its second shell (the ratio of Asp/Glu to Lys/Arg is ∼8 for Ca compared to ∼2 for Mg and Mn).

(1) M(II)−Asp/Glu Second Shell. Figures 3, 4, and 5 show the percentage frequency distributions of second-shell ligands around common first-shell ligands in Mg, Mn, Ca, and Zn proteins; viz., Asp/Glu side chains, backbone carbonyls, and water molecules (see above and Figure 1). The most common

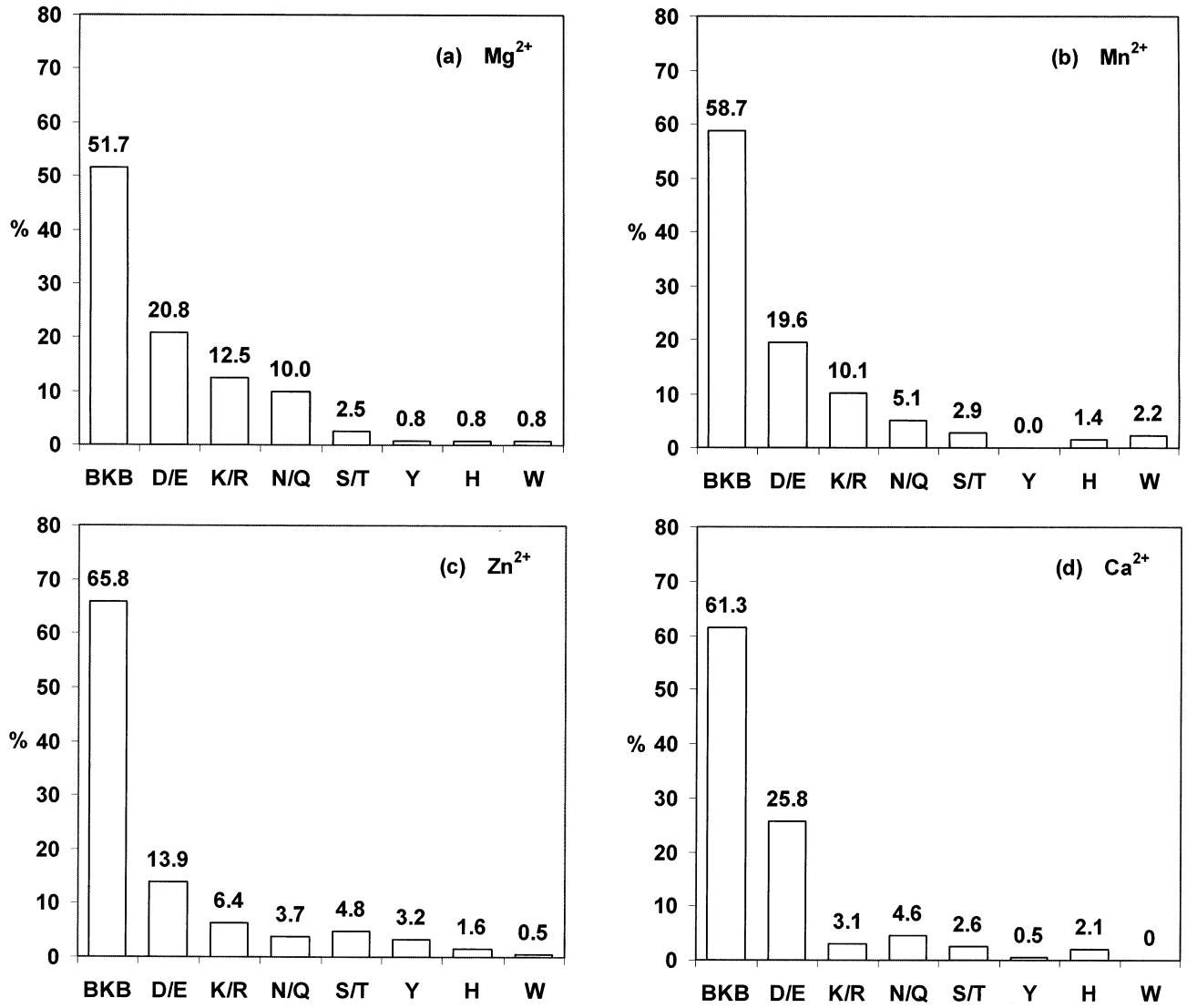
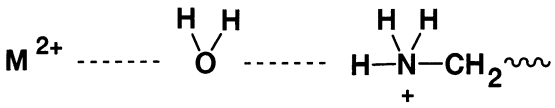


Figure 2. Percentage frequency distribution of second-shell ligands observed in the 3-dimensional structures of proteins bound to (a) Mg, (b) Mn, (c) Zn, and (d) Ca. The total numbers of second-shell ligands bound to Mg, Mn, Zn, and Ca are 131, 190, 220, and 283, respectively.

second-shell partner for first-shell Asp/Glu side chains appears to be the backbone amide group (36−67%, Figure 3), which is found hydrogen bonded to the metal-free carboxylate oxygen. The latter forms salt bridges with outer-shell Lys/Arg side chains in some sites, particularly Mg, Mn, and Zn binding ones. As expected, due to Coulombic repulsion between the negatively charged first- and second-shell Asp/Glu side chains as well as the lack of a hydrogen donor, no second-shell Asp/Glu carboxylates were found hydrogen bonded to a first-shell Asp/Glu.

(2) M(II)−Backbone Second Shell. Inner-layer backbone groups are more selective than first-shell Asp/Glu side chains for the type of second-shell ligands. They interact (via the N−H group) predominantly with the carbonyl oxygen of the backbone or Asn/Gln side chains in the second shell (Figure 4). Interestingly, few second-shell partners are found for the first-shell backbone groups in Ca binding sites. Comparison between Figure 3d and Figure 4c shows that the second-shell backbone groups in Ca proteins prefer to interact with the first-shell *carboxylates* than with the first-shell backbone groups.

(3) M(II)−Water Second Shell. Metal-bound water molecules prefer charged residues to neutral ones and are most commonly found hydrogen bonded to acidic Asp/Glu carboxylates (32−56%, Figure 5). This is not surprising since the triad M−H_2O−Asp/Glu has been found in many metalloenzymes where it is thought to ionize/polarize the catalytic water engaged in a subsequent hydrolytic reaction.[10,13] The first-shell water can also interact with the positively charged Lys/Arg side chains as shown:



Water is often found hydrogen bonded to the backbone carbonyl oxygen (20−28%) in Mg, Mn, and Ca binding sites, but they are not seen hydrogen bonded to the backbone carbonyl oxygen in Zn proteins (Figure 5). Neutral polar residues such as Ser/Thr or Tyr are not frequent second-shell partners of the metal-bound water (3−6% and 0−11%, respectively). Nevertheless, they may be essential for the catalytic mechanism of some enzymes; e.g., the second-shell Thr 199 in carbonic anhydrase I is thought to orient the Zn-bound water molecule for nucleophilic attack of CO_2 and for transition state stabiliza-

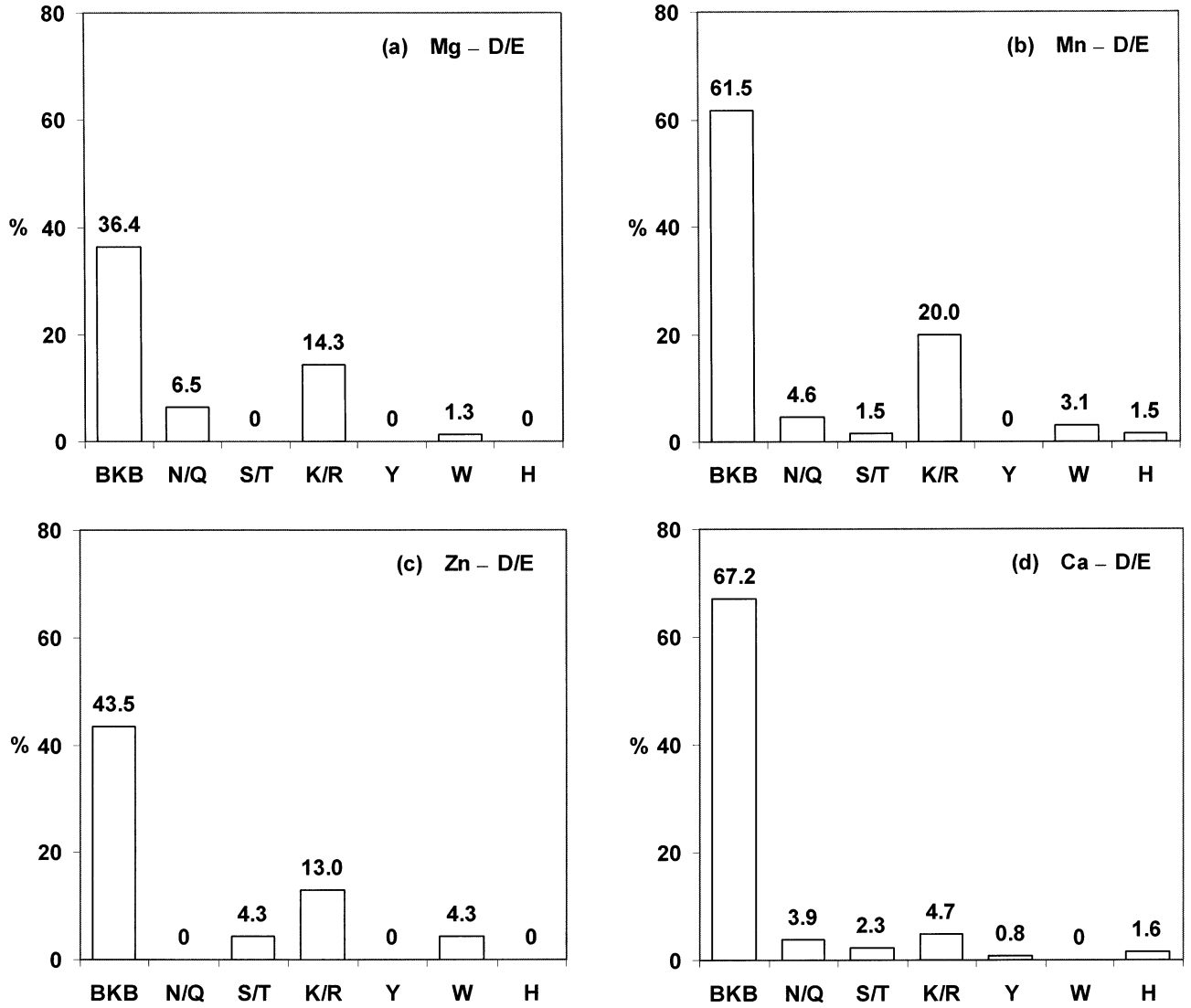
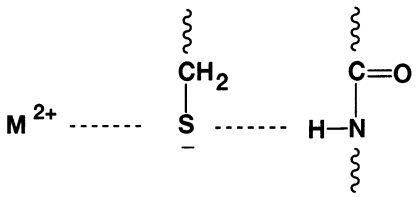


Figure 3. Percentage frequency distribution of second-shell ligands observed in the 3-dimensional structures of proteins bound to first-shell Asp/Glu in (a) Mg, (b) Mn, (c) Zn, and (d) Ca binding sites. For a given metal, the percentage is the number of second-shell ligand of type X divided by the total number of first-shell Asp/Glu (77 for Mg, 65 for Mn, 23 for Zn, and 128 for Ca).

tion.[65−67] This suggests that the second-shell partner of the metal-bound water is not only governed by thermodynamic considerations but may also be determined by the catalytic mechanism of the reaction.

(4) Zn(II)−Cys Second Shell. Figure 6 shows the percentage frequency distributions of second-shell ligands around Zn-specific first-shell ligands; viz., Cys and His (see above and Figure 1). The Zn-bound Cys prefers binding to the backbone or Asn/Gln side chain amide (Figure 6a).



In some cases the cysteine core is stabilized by salt bridges with outer-layer Lys/Arg.

(5) Zn(II)−His Second Shell. In contrast to the Zn-bound cysteines, which are mainly found in proteins where Zn serves only a structural role ("structural Zn" proteins), Zn-bound His are found in both "structural Zn" proteins, and in proteins where Zn plays a predominantly catalytic role ("catalytic Zn" proteins). The adequate numbers of Zn-bound His in both "structural Zn" and "catalytic Zn" proteins allowed a statistical analysis of the second-shell interactions in these two types of binding sites. In structural Zn binding sites (Figure 6b), the first-shell His is predominantly hydrogen bonded (via its metal-free N−H group) to either a backbone carbonyl oxygen or a Ser/Thr side chain oxygen, but in catalytic Zn binding sites (Figure 6c), it is predominantly hydrogen bonded to a second-shell Asp/Glu carboxylate oxygen. The Zn-His-Asp/Glu triad appears characteristic of "catalytic Zn" binding sites, and its significance has been analyzed in previous works.[44,48,64−68,74] The second-shell carboxylate has been proposed to fine tune the positioning and metal affinity of its first-shell partner.[64−68,74]

Second-Shell Formation. From the PDB survey results in Figures 1−6, we can discern two types of second-shell formation. In the first case, the water molecule remains bound to the metal in the preformed protein binding site and subsequently interacts with a second-shell protein ligand (Figure 5). In the

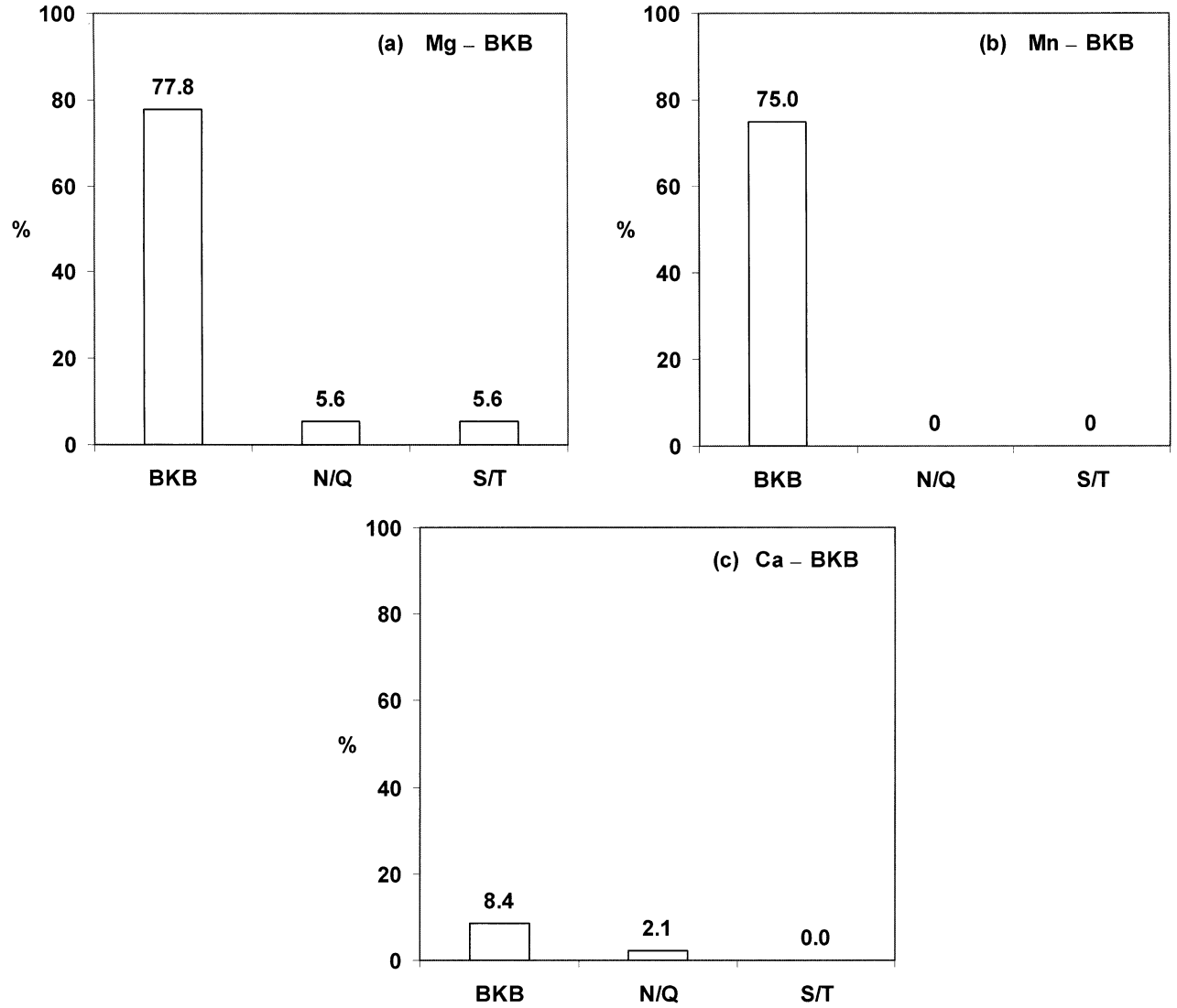


Figure 4. Percentage frequency distribution of second-shell ligands observed in the 3-dimensional structures of proteins bound to first-shell backbone groups in (a) Mg, (b) Mn, and (c) Ca binding sites. For a given metal, the percentage is the number of second-shell ligand of type X divided by the total number of first-shell backbone groups (18 for Mg, 12 for Mn, and 95 for Ca). No such distribution is presented for Zn binding sites because no backbone groups were found coordinated to mononuclear Zn (see Figure 1c).

second case, both the first- and second-shell ligands are provided by the protein (Figures 3, 4, and 6). We focus on the thermodynamics, but not the mechanism, of the second-shell formation process, so only the initial and final stages of the second-shell formation process were considered in this work. In interpreting the results below we emphasize the *trend* (sign) of the free energy changes rather than their absolute values.

First-Shell Water, Second-Shell Protein Ligands. Metal-bound water molecules are most commonly found hydrogen bonded to carboxylate side chains present in the second shell (see above and Figure 5). This process was modeled by a hydrated metal ion from aqueous solution finding a protein cavity (characterized by a dielectric constant equal to 2 for buried sites or 4 for partially buried sites), and subsequently, exchanging its second-shell water for a formate; i.e.,

$$\{[M(H_2O)_n]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[M(H_2O)_n]\cdot HCOO\}^{+} + H_2O \quad (3)$$

The number of water molecules in the first shell, n, is 6 for M = Mg^{2+}, Mn^{2+}, or Zn^{2+} and 7 for Ca^{2+}. The $\{[M(H_2O)_6]\cdot H_2O\}^{2+}$ (M = Mg, Mn, Zn) complex with a second-shell water molecule and the corresponding $\{[M(H_2O)_6]\cdot HCOO\}^{+}$ complexes with an outer-shell formate bound to inner-shell water molecules are illustrated in Figures 7a and 7b, respectively. Since Zn prefers to be tetracoordinated in proteins, the water−formate exchange was also modeled as

$$\{[Zn(H_2O)_6]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[Zn(H_2O)_4]\cdot (H_2O)_2 HCOO\}^{+} + H_2O \quad (4)$$

where the Zn coordination number in the product is four with two water molecules and a formate in the outer coordination layer (Figure 7c). The computed gas-phase enthalpies and free energies in various dielectric media for reactions 3 and 4 are listed in Table 1.

All the reactions in Table 1 are enthalpy-driven as the entropy term, $T\Delta S$[1], contributes <3% to the *gas-phase* free energy; therefore, a second-shell Asp/Glu is unlikely to make a significant entropic contribution to the cationic metal complex

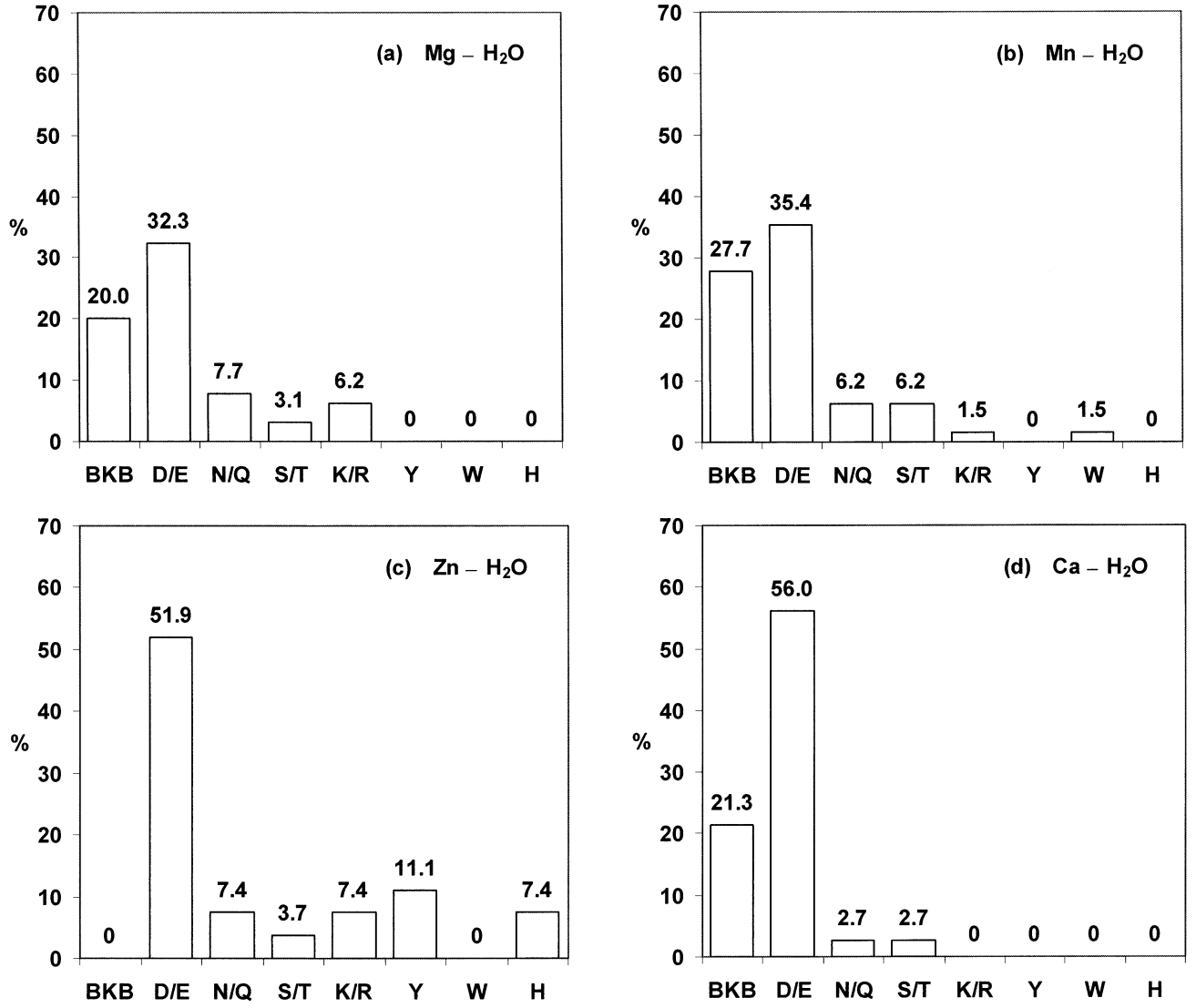


Figure 5. Percentage frequency distribution of second-shell ligands observed in the 3-dimensional structures of proteins bound to first-shell water in (a) Mg, (b) Mn, (c) Zn, and (d) Ca binding sites. For a given metal, the percentage is the number of second-shell ligand of type X divided by the total number of first-shell water molecules (65 for Mg, 65 for Mn, 27 for Zn, and 75 for Ca).

stability in a buried cavity. Due to the strong electrostatic interaction between the negatively charged formate and the positively charged metal hydrate, the ΔG^x ($x = 1$, 2, or 4) for the first five reactions in Table 1 are negative, implying that second-shell water−formate exchange is favorable in the gas phase and in a protein cavity. The free energy of the second-shell water−formate exchange decreases rapidly in going from the gas phase ($|\Delta G^1| = 171-194$ kcal/mol) to buried ($|\Delta G^2| = 77-93$ kcal/mol) and partially buried ($|\Delta G^4| = 29-42$ kcal/mol) binding sites. However, it does not appear to be sensitive to metal cations with the same coordination geometry and similar ionic radii, as the ΔG^x ($x = 1$, 2, or 4) for the octahedral complexes of Mg, Zn, and Mn, whose *ionic* radii are 0.72, 0.75, and 0.83 Å, respectively,[84] differ by only 1−3 kcal/mol. Decreasing the Zn coordination number from six to four results in an additional free energy gain (of 5−9 kcal/mol, compare reactions 3 and 4 in Table 1). The *hepta*coordinated Ca complexes (Table 1, reaction 5) have the smallest absolute exchange free energies in the series: the $|\Delta G^x|$ values ($x = 1$, 2, or 4) are smaller than those for the *hexa*coordinated Mg, Mn, and Zn complexes (by 6−14 kcal/mol) and are much smaller than those for the *tetra*coordinated Zn complex (by 13−23 kcal/mol).

It is of interest to estimate how the free energies for the outer-water−formate exchange reactions compare with those for the inner-water−formate exchange. The computed values in the lower half of Table 1 show that the inner-water−formate exchange is more favorable than the outer-water−formate exchange in the gas phase (by 12−16 kcal/mol) and to a lesser extent (by 5−11 kcal/mol) in a buried or partially buried cavity. Generally, the inner-shell exchange reactions follow the same trend of changes found for the respective outer-shell reactions (see above): the absolute free energy decreases rapidly (but remains negative) with increasing dielectric constant and increasing coordination number of the core metal.

First- and Second-Shell Ligands Are Provided by the Protein. As discussed above, backbone peptide groups and Asn/Gln side chains comprising part of the first shell in Mg, Mn, and Ca binding sites interact predominantly with the carbonyl oxygen of the backbone or Asn/Gln side chains in the outer

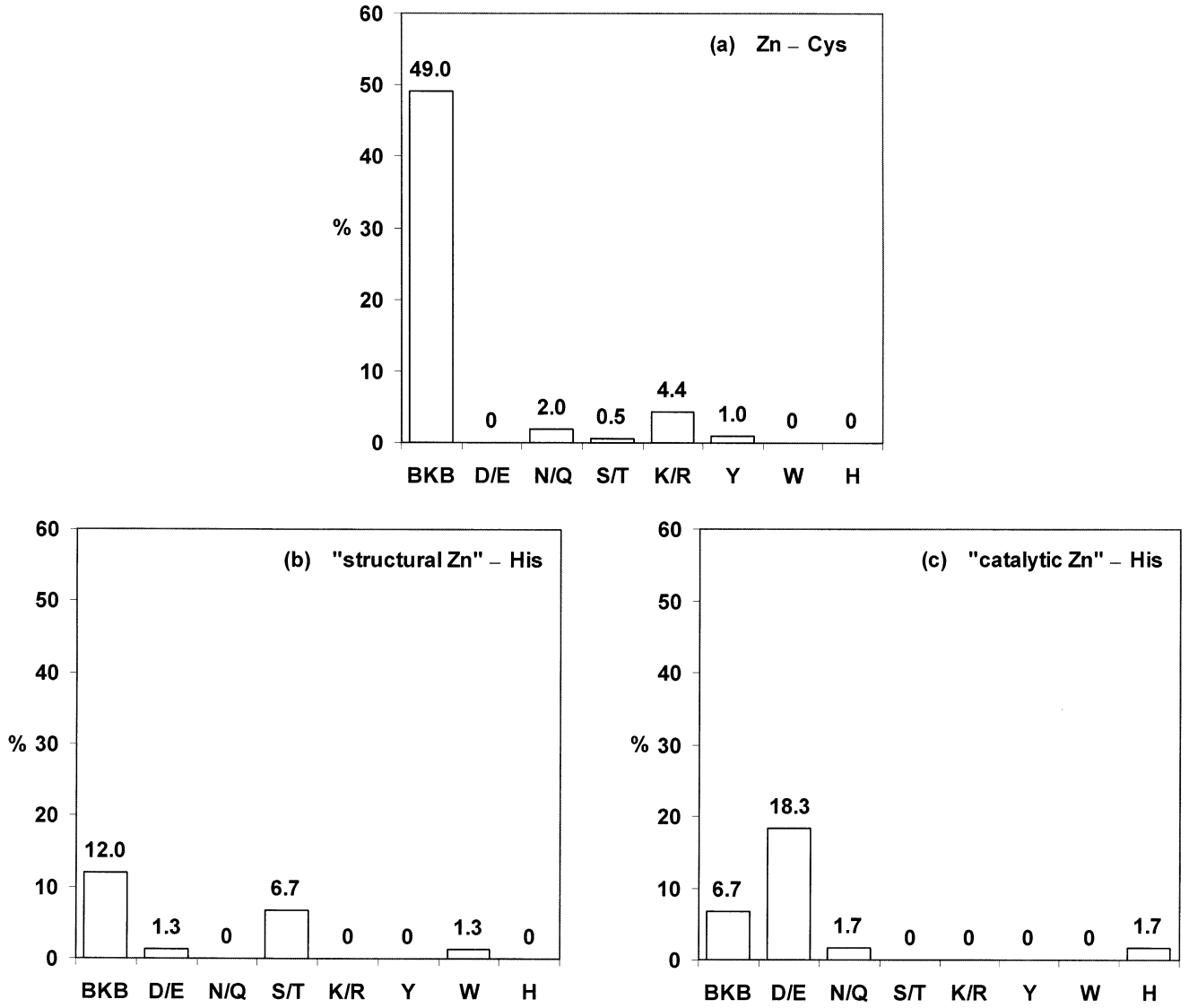


Figure 6. Percentage frequency distribution of second-shell ligands observed in the 3-dimensional structures of Zn proteins bound to (a) first-shell Cys (204), (b) first-shell His in structural sites (75), and (c) first-shell His in catalytic sites (60). The percentage is the number of second-shell ligand of type X divided by the total number of first-shell Cys /His molecules.

layer (Figure 4). We discern two types of reactions. In the first case the first- and second-shell backbone groups or Asn/Gln side chains are not hydrogen bonded in the metal-free protein, but become hydrogen bonded after the metal is bound in a flexible binding site. This process is modeled thermodynamically by the following reaction:

$$[M(H_2O)_n]^{2+} + 2(HCONH_2) \rightarrow \{[M(H_2O)_{n-1}HCONH_2]\cdot HCONH_2\}^{2+} + H_2O \quad (5)$$

where M = Mg, Mn, or Ca and $n = 6$ or 7. In the second case the first- and second-shell backbone groups or Asn/Gln side chains are hydrogen bonded in the metal-free and metal-bound protein. This process is modeled as

$$[M(H_2O)_n]^{2+} + (HCONH_2\cdots HCONH_2) \rightarrow \{[M(H_2O)_{n-1}HCONH_2]\cdot HCONH_2\}^{2+} + H_2O \quad (6)$$

The fully optimized $\{[M(H_2O)_{n-1}(HCONH_2)]\cdot HCONH_2\}^{2+}$ structures are shown in Figure 8, while the computed gas-phase enthalpies and free energies in various dielectric media for reactions 5 and 6 are listed in Table 2.

The gas-phase entropic term, ΔS^1, have opposite signs for the two types of reactions considered. It is unfavorable (negative) for eq 5, but is favorable (positive) for eq 6 where the two formamides are hydrogen bonded prior to metal binding. The unfavorable entropic changes for eq 5 are compensated by corresponding enthalpy changes that are more favorable than those for eq 6. Consequently, the gas-phase free energies, ΔG^1, for eq 5 are only slightly less favorable (by ~2 kcal/mol) than those evaluated for eq 6. However, because of the greater desolvation penalty for the free formamides relative to the hydrogen-bonded formamides, the ΔG^x ($x = 2$ or 4) for eq 5 become even less favorable than those evaluated for eq 6 (by 4−6 kcal/mol) so much so that in a partially solvent exposed cavity ($\epsilon \geq 4$), reaction 5 is predicted not to take place (positive ΔG^4 for the first three reactions in Table 2), whereas reaction 6 remains favorable under the same conditions.

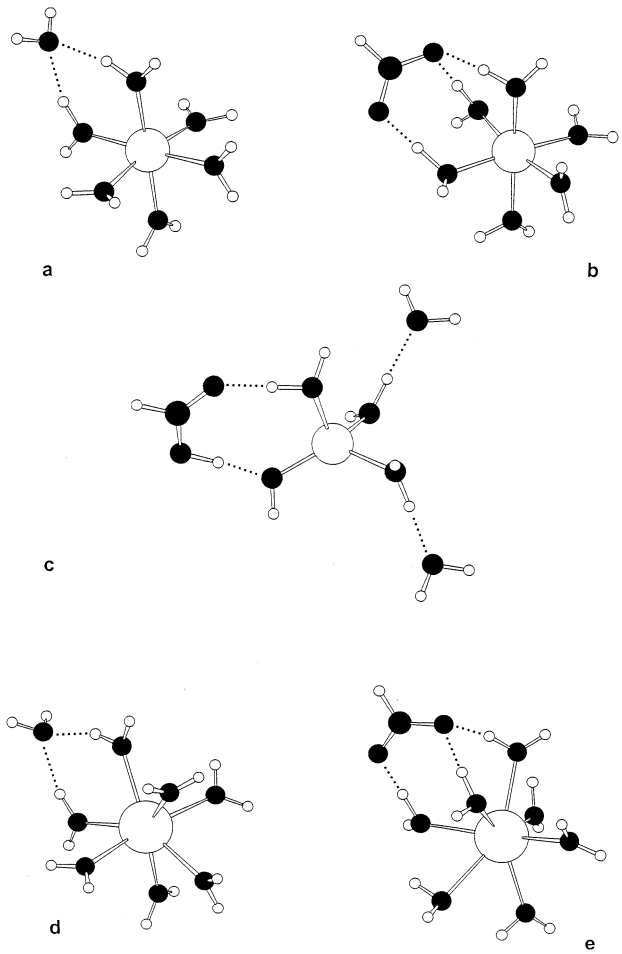


Figure 7. Ball-and-stick diagram of (a) $\{[M(H_2O)_6]\cdot H_2O\}^{2+}$ (M = Mg, Mn, Zn); (b) $\{[M(H_2O)_6]\cdot HCOO\}^{+}$ (M = Mg, Mn, Zn); (c) $\{[Zn(H_2O)_4]\cdot(H_2O)_2HCOO\}^{+}$; (d) $\{[Ca(H_2O)_7]\cdot H_2O\}^{2+}$; and (e) $\{[Ca(H_2O)_7]\cdot HCOO\}^{+}$.

Table 1. Enthalpies (ΔH^1) and Free Energies (ΔG^x)of Water−Formate Exchange for Media of Different Dielectric Constant x[a]

reaction	ΔH^1	ΔG^1	ΔG^2	ΔG^4
$\{[Mg(H_2O)_6]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[Mg(H_2O)_6]\cdot HCOO\}^{+} + H_2O$	−188.0	−183.2	−85.4	−36.6
$\{[Mn(H_2O)_6]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[Mn(H_2O)_6]\cdot HCOO\}^{+} + H_2O$	−186.6	−182.2	−84.0	−35.0
$\{[Zn(H_2O)_6]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[Zn(H_2O)_6]\cdot HCOO\}^{+} + H_2O$	−190.0	−185.2	−86.2	−36.8
$\{[Zn(H_2O)_6]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[Zn(H_2O)_4]\cdot(H_2O)_2HCOO\}^{+} + H_2O$	−195.4	−193.9	−92.5	−41.7
$\{[Ca(H_2O)_7]\cdot H_2O\}^{2+} + HCOO^- \rightarrow \{[Ca(H_2O)_7]\cdot HCOO\}^{+} + H_2O$	−176.2	−171.2	−76.5	−28.7
$[Mg(H_2O)_6]^{2+} + HCOO^- \rightarrow [Mg(H_2O)_5HCOO]^{+} + H_2O$	−198.7	−197.6	−93.5	−41.7
$[Mn(H_2O)_6]^{2+} + HCOO^- \rightarrow [Mn(H_2O)_5HCOO]^{+} + H_2O$	−198.7	−196.1	−92.5	−40.7
$[Zn(H_2O)_6]^{2+} + HCOO^- \rightarrow [Zn(H_2O)_5HCOO]^{+} + H_2O$	−201.9	−199.9	−95.8	−44.2
$[Zn(H_2O)_6]^{2+} + HCOO^- \rightarrow \{[Zn(H_2O)_3HCOO]\cdot(H_2O)_2\}^{+} + H_2O$	−207.3	−209.5	−103.9	−51.5
$[Ca(H_2O)_7]^{2+} + HCOO^- \rightarrow [Ca(H_2O)_6HCOO]^{+} + H_2O$	−187.4	−182.8	−83.4	−33.3

[a] All energies in kcal/mol; $x = 1$ corresponds to gas-phase values whereas $x = 2$ or 4 represents buried or partially buried metal binding sites, respectively.

To assess the contribution of the second-shell backbone binding to the net free energy for reaction 5 or 6, the free energy for exchanging a metal-bound water for a formamide in the absence of a second shell formamide was computed; i.e.,

$$[M(H_2O)_n]^{2+} + HCONH_2 \rightarrow [M(H_2O)_{n-1}HCONH_2]^{2+} + H_2O \quad (7)$$

The ΔH^1 for eq 7 (last three reactions in Table 2) are less favorable than those for eq 5 and eq 6, indicating that a second-shell formamide energetically stabilizes a dicationic metal complex in the gas phase. The ΔS^1 for eq 7 are also less favorable than those for eq 6 but are more favorable than those for eq 5, suggesting that a second-shell formamide hydrogen bonded to a first-shell formamide prior to metal binding could make an entropic contribution and enhance dicationic metal complex stability in the gas phase. In a *buried* metal binding site, a second-shell formamide could stabilize a dicationic metal complex, as evidenced by the more negative ΔG^1 and ΔG^2 for eq 5 or eq 6 compared to those for eq 7 (Table 2). In a *partially solvent-exposed* metal binding site, a second-shell formamide

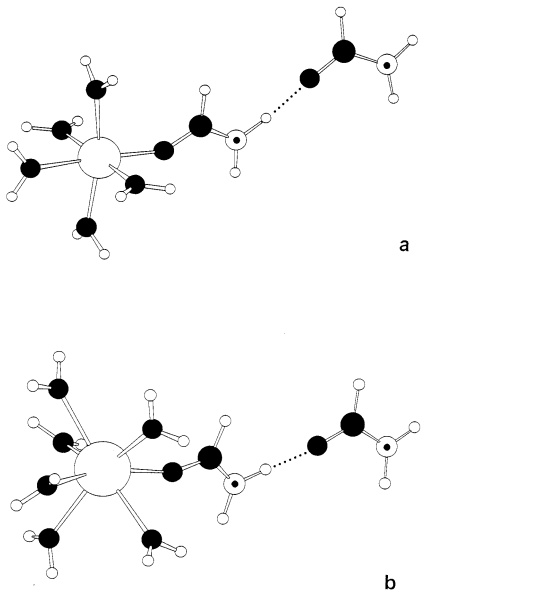


Figure 8. Ball-and-stick diagram of (a) $\{[M(H_2O)_5\,HCONH_2]\cdot HCONH_2\}^{2+}$ (M = Mg, Mn); and (b) $\{[Ca(H_2O)_6\,HCONH_2]\cdot HCONH_2\}^{2+}$.

Table 2. Enthalpies (ΔH^1) and Free Energies (ΔG^x) of Formamide Binding to Mg, Mn, and Ca Complexes for Media of Different Dielectric Constant x[a]

reaction	ΔH^1	ΔG^1	ΔG^2	ΔG^4
$[Mg(H_2O)_6]^{2+} + 2(HCONH_2) \rightarrow \{[Mg(H_2O)_5HCONH_2]\cdot HCONH_2\}^{2+} + H_2O$	−39.0	−32.1	−10.6	0.5
$[Mn(H_2O)_6]^{2+} + 2(HCONH_2) \rightarrow \{[Mn(H_2O)_5HCONH_2]\cdot HCONH_2\}^{2+} + H_2O$	−39.4	−30.0	−8.3	2.9
$[Ca(H_2O)_7]^{2+} + 2(HCONH_2) \rightarrow \{[Ca(H_2O)_6HCONH_2]\cdot HCONH_2\}^{2+} + H_2O$	−33.8	−25.2	−6.4	3.3
$[Mg(H_2O)_6]^{2+} + HCONH_2\cdots\cdots HCONH_2 \rightarrow \{[Mg(H_2O)_5HCONH_2]\cdot HCONH_2\}^{2+} + H_2O$	−29.7	−33.8	−15.2	−6.2
$[Mn(H_2O)_6]^{2+} + HCONH_2\cdots\cdots HCONH_2 \rightarrow \{[Mn(H_2O)_5HCONH_2]\cdot HCONH_2\}^{2+} + H_2O$	−30.1	−31.7	−12.9	−3.8
$[Ca(H_2O)_7]^{2+} + HCONH_2\cdots\cdots HCONH_2 \rightarrow \{[Ca(H_2O)_6HCONH_2]\cdot HCONH_2\}^{2+} + H_2O$	−24.4	−27.0	−11.1	−3.5
$[Mg(H_2O)_6]^{2+} + HCONH_2 \rightarrow [Mg(H_2O)_5HCONH_2]^{2+} + H_2O$	−17.8	−19.1	−8.3	−3.1
$[Mn(H_2O)_6]^{2+} + HCONH_2 \rightarrow [Mn(H_2O)_5HCONH_2]^{2+} + H_2O$	−18.1	−17.5	−6.7	−1.1
$[Ca(H_2O)_7]^{2+} + HCONH_2 \rightarrow [Ca(H_2O)_6HCONH_2]^{2+} + H_2O$	−15.4	−15.0	−5.6	−0.5

[a] See footnote to Table 1.

Table 3. Selected Structural and Electronic Parameters for Zn and Ca Complexes Evaluated at the B3LYP/6-31++G(2d,2p) Level

	$R_{M-O(water)}$ (Å)	$R_{M-O(HCOO)}$ (Å)	R_{O-H} (Å)	q_{O-H}[a] (e)	$\|p_{H_2O}\|$[b] (Å e)	q_M (e)
H_2O			0.963	−1.48	0.58	
$[Zn(H_2O)_4]^{2+}$	1.996		0.970	−1.67	0.80	1.730
$[Zn(H_2O)_6]^{2+}$	2.123		0.967	−1.62	0.75	1.684
$[Ca(H_2O)_7]^{2+}$	2.479		0.966	−1.60	0.68	1.849
Inner-shell formate						
$\{[Zn(H_2O)_3HCOO]\cdot(H_2O)_2\}^{+}$		1.917				1.662
$[Zn(H_2O)_5HCOO]^{+}$		1.991				1.639
$[Ca(H_2O)_6HCOO]^{+}$		2.375				1.841
Outer-shell formate						
$\{[Zn(H_2O)_4]\cdot(H_2O)_2\ HCOO\}^{+}$	1.862[c]	3.732	1.526[c]	−1.83[c]		1.684
$\{[Zn(H_2O)_6]\cdot HCOO\}^{+}$	2.011[c]	3.740	1.078[c]	−1.66[c]		1.654
$\{[Ca(H_2O)_7]\cdot HCOO\}^{+}$	2.404[c]	4.131	1.052[c]	−1.67[c]		1.843

[a] Bond charge, $q_{O-H} = q_O - q_H$, is defined as the difference between the respective atomic charges, where q_i are NBO atomic charges. [b] $|p_{H_2O}| = [(\Sigma q_i x_i)^2 + (\Sigma q_i y_i)^2 + (\Sigma q_i z_i)^2]^{1/2}$, where x_i, y_i, and z_i are the atomic Cartesian coordinates. [c] Quantities refer to the water molecule involved in outer-shell formate binding.

could also stabilize a dicationic metal complex only if it is hydrogen bonded to the first-shell formamide prior to metal binding (ΔG^4 for eq 6 is more negative than ΔG^4 for the last three reactions in Table 2).

Although the gas-phase free energies, ΔG^1, in Table 2 are negative, they are about an order of magnitude smaller than the water−formate exchange ΔG^1 in Table 1. Furthermore, even though all the reactions in Table 2, like those in Table 1, are enthalpy-driven, the $T\Delta S^1$ contributions to the *gas-phase* free energy, especially for the first three reactions (18−24%), are significantly greater than those in Table 1 (1−3%). As observed in Table 1, Ca binding to formamides is the least preferred in the series.

Factors Governing First- and Second-Shell Interactions. The trends observed in Tables 1 and 2 can be rationalized in terms of charge−charge, charge−dipole, and charge-transfer interactions, which make the major contributions to the metal complex stability.[25,43] The Coulombic interaction energy, $\Delta E_{qq} = -q_M q_L / R_{ML}$, depends on the charges on the metal (q_M) and ligand atoms (q_L) as well as inversely on the metal−ligand atom distance (R_{ML}). The charge−dipole interaction energy,[96] $\Delta E_{qp} = -(q_M p_L \cos\theta / r^2 + 1/2 q_M^2 \alpha_L / r^4)$, depends also on the metal charge as well as the dipole moment (p_L) and polarizability (α_L) of the neutral ligand, the distance between the charge and the dipole moment center (r), and the angle between the r and p axes (θ). The water polarizability α_{H_2O} generally depends on the molecular volume and delocalization of the molecular charges, so changes in the water O−H bond length (R_{O-H}) and bond charge (q_{O-H}) may be considered as indirect measures of α_{H_2O} changes. The structural and electronic parameters comprising the charge−charge and charge−dipole interaction energies for several representative metal clusters are listed in Table 3.

The R_{M-O} distances for the all-aqua metal complexes in Table 3 show that the metal−O(water) distance increases with increasing coordination number and ionic radius of the metal (0.75 Å for Zn, 1.00 Å for Ca). The α_{H_2O} (as reflected by R_{O-H} and q_{O-H}) and $|p_{H_2O}|$ of the metal hydrates are greater than those of metal-free water. Although the R_{O-H} bond length remains relatively constant with increasing R_{M-O} or coordination number, the bond charge, q_{O-H}, and hence α_{H_2O}, as well as $|p_{H_2O}|$ decreases with increasing metal−O(water) distance.

Table 1 shows that the $|\Delta H^1|$ for the *outer*-shell clusters (first five reactions) are smaller than those for the *inner*-shell clusters (last five reactions). This is partly because the outer-shell carboxylate oxygen is further from the metal compared to an inner-shell carboxylate oxygen (3.73 vs 1.92 Å for Zn and 4.13 vs 2.38 Å for Ca, Table 3). Hence, charge−charge interactions for the *outer*-shell clusters are weaker than those for the *inner*-shell clusters.

Tables 1 and 2 show that the formamide complexes are less stable than the formate complexes. This is because the charge−dipole interaction between the metal dication and the first-shell formamide in $[M(H_2O)_{n-1}(HCONH_2)]^{2+}$ is much weaker than the charge−charge interaction in the respective first-shell formate complexes. Furthermore, the interaction between the first- and second-shell formamides in $\{[M(H_2O)_{n-1}(HCONH_2)]\}\cdot HCONH_2\}^{2+}$ is of the dipole−dipole type, which is weaker than the dipole−charge interaction between the first-shell water and second-shell formate in the respective $\{[M(H_2O)_n]\}\cdot HCOO\}^{2+}$ in Table 1.

The ΔG^1 values in Table 1, which are dictated by the enthalpic term, show that exchanging an inner- or outer-shell water molecule for a formate becomes less favorable with increasing coordination number of the core metal. For example, the gas-phase enthalpies for the 4-coordinated Zn, 6-coordinated Zn, and 7-coordinated Ca complexes containing an inner-shell formate are −207, −202, and −187 kcal/mol, respectively (Table 1). The $|\Delta H^1|$ values are correlated with formate−metal charge transfer, which decreases with increasing metal coordination number. The q_{Zn} decreases by 0.068e for the tetrahedral Zn complex (from 1.730e in $[Zn(H_2O)_4]$ to 1.662e in $\{[Zn(H_2O)_3\ HCOO]\}^{+}$), by 0.045 e for the octahedral Zn complex, and by only 0.008 e for the heptacoordinated Ca complex (Table 3). Therefore, the observed decrease in the $|\Delta H^1|$ with increasing metal coordination number is due to decreasing charge transfer from the formate to the metal as well as increasing steric repulsion among the ligands with increasing metal coordination number.

An Outer-Shell Carboxylate Might Act as a Proton Acceptor for an Inner-Shell Water. The water in $[Zn(H_2O)_4]$ is polarized, as evidenced by an increase in R_{O-H} by 0.007 Å and q_{O-H} by 0.19e relative to the metal-free values (Table 3). It is even more polarized when it is hydrogen bonded to a carboxylate group in the Zn outer shell. The effect is so pronounced that it leads to the breaking up of the water O−H bond and proton transfer to $HCOO^-$: the water R_{O-H} bond length increases to 1.53 Å, while the respective bond charge,

(96) Daune, M. *Molecular Biophysics*; Oxford University Press: Oxford, 1999.

q_{O-H}, increases to −1.83e (see Table 3 and Figure 7c). This implies that in a fully buried binding site where there are no interactions stabilizing $HCOO^-$ or destabilizing HCOOH, the $Zn^{2+}-H_2O-HCOO^-$ triad may isomerize to form $Zn^{2+}-HO^--HCOOH$ because the charge−charge−dipole interactions in the latter would be more favorable than the charge−dipole−charge interactions in the former.

Discussion

The PDB survey results in Figures 1 and 2 show that backbone peptide groups, which are by far the most numerous ligands provided by the protein, are weakly represented in the first coordination sphere of Mg, Mn, and Zn binding sites but dominate the second shell. The results in Tables 1 and 2 provide the energetic basis for the observed occurrence frequency difference between first- and second-shell backbone peptide groups. Inner-shell binding of carboxylate oxygen is an order of magnitude more favorable than that of backbone oxygen. Therefore, binding sites containing carboxylate side chains are expected to be much more effective in sequestering the metal cation from the surrounding fluids compared to binding sites comprising of backbone peptide groups. Although the second-shell backbone interactions are not strong (Table 2), the amide group (−CO−NH−) wins over the rest of the second-shell ligands by its universality: with its two potential hydrogen-bonding centers it can participate in a large spectrum of interactions with the inner-shell ligands. The second-shell backbone carbonyl oxygen can act as a hydrogen-bond acceptor for H^O from water, Ser, Thr, and Tyr as well as H^N from His, Asn, Gln, and the backbone amide, whereas the second-shell backbone amide proton can serve as a hydrogen-bond donor for carboxylates ($COO^-\cdots HN$), cysteinates ($S^-\cdots HN$), and the backbone carbonyl group ($CO\cdots HN$). Thus, the second-shell backbone or Asn/Gln side chain can partner practically every first-shell ligand—a unique property shared by no other second-shell ligand. Hence, despite the relatively weak interaction energy, backbone groups contribute to energetically stabilizing a buried metal binding site (see Table 2) and, in the case of cysteine-rich Zn-finger cores, help to shield the inner-shell core from unwanted interactions with the environment.[63]

Aspartates and glutamates are the second-most frequently found ligands in the metal second shell (Figure 2). These predominantly partner first-shell waters (Figure 5) and "catalytic" Zn-bound histidines (Figure 6c), by forming $COO^-\cdots H-O$ and $COO^-\cdots H-N$ bonds, respectively. The results in Tables 1 and 3 suggest two plausible roles for an outer-shell carboxylate group. A second-shell carboxylate significantly stabilizes a positively charged metal complex in a protein cavity, as evidenced by the large, negative free energies evaluated for the cationic outer-shell formate complexes in Table 1. Furthermore, it polarizes the O−H bond of an inner-shell water molecule, considered to be a prerequisite step for a subsequent catalytic reaction,[10,13] as evidenced by the increase in the O−H bond distance (R_{O-H}) and bond charge (q_{O-H}) in going from $[M(H_2O)_n]^{2+}$ to $\{[M(H_2O)_n]\cdot HCOO\}^+$ in Table 3. The R_{O-H} and $|q_{O-H}|$ increase by 0.11 Å and 0.05e for octahedral Zn complexes and by 0.09 Å and 0.07e for the 7-coordinated Ca complexes (Table 3). In totally buried sites where the electrostatic interactions are enhanced, an ionization of the metal-bound water (see Table 3 and Figure 7c) or Zn-bound His[44,48] (see Introduction), accompanied by a proton transfer to the carboxylate, could not be ruled out.

In addition to a second-shell carboxylate (see above), a second-shell backbone or side chain CONH group binding to a positively charged metal complex can also contribute to the net binding free energy, thus enhancing the binding-site affinity for the metal (see Results and Table 2). These findings, as evidenced by the trends of the binding free energy changes in Tables 1 and 2, are in accord with experimental studies showing that second-shell ligands increase the binding-site affinity for the metal.[64−70,72,74] For example, in human carbonic anhydrase II the outer-layer Glu, Gln, Asn, and Thr, which hydrogen bond to the first-shell His_3Water core, increase the binding-site affinity for Zn by a factor of 10 per hydrogen bond.[65,66,70] Similar results have been reported for the role of the second-shell ligands in a Zn(II) binding site in a variant of the IgG binding protein G B1 domain[74] and iron-bound human serum transferrin.[72] The finding that a second-shell Asp/Glu stabilizes a positively charged metal complex more than a neutral ligand such as the backbone and Asn/Gln side chain is also in accord with the experimental finding that the Zn−His−carboxylate triad in human carbonic anhydrase II is more stable than the Zn−His−carboxamide triad.[66]

In analogy to the first shell, the second shell may contribute to the metal selectivity of the respective binding site. Although it is not as selective as the first shell, it appears to be selective toward metal cations that possess different coordination geometry and/or ionic size. The second shell favors cations such as Zn that prefer tetrahedral to octahedral geometry and that can successfully compete with bulkier cations with 7 or more inner-shell ligands for the binding site (Table 1). The second shell, however, seems to be much less discriminative toward metal ions that have the same coordination geometry and similar ionic size, such as octahedral Mg, Mn, and Zn: The free energies for the first−second shell interactions in the respective complexes are close to each other (Tables 1 and 2). Although we could not find any direct experimental data proving or disproving the prediction that second-shell ligands might discriminate between metal cations possessing different coordination geometry and number, indirect support comes from the works of Karlin and associates who have shown that the structure of the second shell depends on, and thus appears to be specific to, the type of the bound metal.[75−77]

The aforementioned discussions suggest that the structure and properties of the metal second coordination sphere are generally correlated with those of the first coordination layer. The second layer is designed to either stabilize/protect the inner shell or to enhance its properties. Generally, the second shell helps in enhancing the binding-site metal affinity and additionally stabilizing the first layer by creating an elaborate network of hydrogen bonds or salt bridges around it. The second-layer partner is chosen in accordance with the chemical properties of the respective first-shell ligand so that favorable interactions can occur via hydrogen bond or salt bridge formation. In some cases (e.g., Zn finger cores) the second layer tightly encapsulates the inner-core structure, thus preventing it from undesired interactions with the environment. In metalloenzymes second-shell ligands may play a role in fine-tuning the positioning of catalytically important inner-shell ligands. In certain catalytic binding sites a second-shell carboxylate may further polarize

(and activate) a first-shell water needed for a subsequent catalytic reaction. The outer shell may also complement the inner-shell metal specificity.

Acknowledgment. We thank Dr. J. Wright for writing the program to analyze the metal binding sites from the PDB.We are grateful to D. Bashford, M. Sommer, and M. Karplus for the program to solve the Poisson equation. This work was supported by the National Science Council, Taiwan (NSC Contract 91-2311-B-001), the Institute of Biomedical Sciences, and the National Center for High-Performance Computing, Taiwan.

Supporting Information Available: PDB entries used in the survey and DFT optimized coordinates for the respective metal complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0209722