Competitive Binding in Magnesium Coordination Chemistry: Water versus Ligands of Biological Interest

Todor Dudev,[†,‡] J. A. Cowan,[§] and Carmay Lim*[,†,||]

Contribution from the Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan, Republic of China, Department of Chemistry, The Ohio State University, 100 West 18th Avenue, Columbus, Ohio 43210, and Department of Chemistry, National Tsing Hua University, Hsinchu 300, Taiwan, R.O.C

Received December 29, 1998. Revised Manuscript Received April 30, 1999

Abstract: Density functional theory and continuum dielectric methods have been employed to evaluate the free energy of successive aqua-substitution reactions: $[Mg(H_2O)_{6-m-n}L_mL'_n]^{2+mz+ny} + L^z \rightarrow [Mg(H_2O)_{5-m-n}L_{m+1}L'_n]^{2+(m+1)z+ny} + H_2O$. The ligands L^z or L'^y under consideration are simple organic molecules that model the amino acid residues that are most commonly found as protein ligands to divalent magnesium. In addition, the Protein Data Bank was surveyed for 3-dimensional protein structures of magnesium-binding sites containing only amino acid ligands. The results obtained were used to delineate the most thermodynamically preferable inner-sphere coordination environment for magnesium. The results also suggest an explanation for two phenomena: (i) the observed inner-sphere binding mode of Mg^{2+} to proteins and (ii) the unique role of Mg^{2+} as a carrier of water molecules that mediate enzymatic hydrolysis reactions.

Introduction

Due to the special role that magnesium plays in biological systems considerable attention has been paid to studies of its chemistry and biochemistry. Although a wealth of information has been accumulated on the subject of magnesium-dependent enzymes and proteins,[1] many questions remain concerning the molecular aspects of this chemistry. In particular, it is not clear why magnesium is an indispensable element for many vital processes in cellular biochemistry. Is it due to the relatively high charge density of magnesium,[1] or does it simply reflect the relatively high intracellular concentration of magnesium relative to other divalent metals?[2] It is also not clear why magnesium tends to bind to nucleic acids indirectly via a water molecule (in an "outer-sphere" coordination mode) whereas it tends to bind to proteins directly (in an "inner-sphere" mode).[1,3] That is, the factors that define the energetically most favorable ligand coordination set are ill-defined.

Divalent magnesium is a "hard" ion and thus prefers binding to "hard" oxygen-containing ligands such as carboxylates, phosphates, hydroxyls, carbonyls, and water. Its complexes usually possess octahedral symmetry in the first coordination shell. Among these the best characterized (both experimentally and theoretically) are the magnesium−water complexes which appear to be quite stable.[4] The free energy penalty for removing a water ligand from the outer sphere of the cluster has been measured to be 5−12 kcal/mol,[5,6] whereas a much larger value of 40−80 kcal/mol has been anticipated for removal of inner-sphere waters.[5] Recent high-level ab initio calculations on gas-phase Mg^{2+}−aqua complexes have shown that the free energy for displacement of a water from the inner coordination sphere of the magnesium cation ranges from 20 to 80 kcal/mol,[7] and strongly depends on the extent of hydration of the cation. The smaller the number of water molecules around the cation, the higher the free energy penalty for water removal.

The most dramatic free energy variations appear in the first coordination shell. Binding of Mg^{2+} in an inner-sphere fashion to a given nonaqua ligand L can occur only if the free energy for each substitution reaction is negative:

$$[Mg(H_2O)_6]^{2+} + L^z \rightarrow [Mg(H_2O)_5L]^{2+z} + H_2O \quad (1)$$

$$[Mg(H_2O)_5L]^{2+z} + L^z \rightarrow [Mg(H_2O)_4L_2]^{2+2z} + H_2O \quad (2)$$

$$[Mg(H_2O)L_5]^{2+5z} + L^z \rightarrow [MgL_6]^{2+6z} + H_2O \quad (3)$$

In these equations z is the total charge on the ligand. Generally, the charge, electron density distribution, polarizability, size, and rigidity of the ligand govern the competitiveness of L relative to water for binding to the cation.

Ab initio calculations have been employed in studying the properties of various magnesium complexes.[7−18] Among these

* Address all correspondence to this author. E-mail: carmay@gate.sinica.edu.tw. Phone: 886-2-2789-9144 or 26523031. FAX: 886−2−2788−7641.

† Academia Sinica.

‡ On leave from the Department of Chemistry, University of Sofia, Bulgaria.

§ The Ohio State University.

|| National Tsing Hua University.

(1) Cowan, J. A., Ed. *Biological Chemistry of Magnesium*; VCH: New York, 1995.

(2) Ebel, H.; Gunther, T. *J. Clin. Chem. Clin. Biochem.* **1980**, *18*, 257.

(3) Black, C. B.; Huang, H. W.; Cowan, J. A. *Coord. Chem. Rev.* **1994**, *135/136*, 165.

(4) Burgess, M. A. *Metal ions in solution*; Ellis Horwood: Chichester, 1978.

(5) Blades, A. T.; Jayaweera, P.; Ikonomou, M. G.; Kebarle, P. *J. Chem. Phys.* **1990**, *92*, 5900.

(6) Peschke, M.; Blades, A. T.; Kebarle, P. *J. Phys. Chem. A* **1998**, *102*, 9978.

(7) Markham, G. D.; Glusker, J. P.; Bock, C. L.; Trachtman, M.; Bock, C. W. *J. Phys. Chem.* **1996**, *100*, 3488.

(8) Yliniemela, A.; Uchimaru, T.; Hirose, T.; Baldwin, B. W.; Tanabe, K. *J. Mol. Struct. (THEOCHEM)* **1996**, *369*, 9.

(9) Garmer, D. R.; Gresh, N. *J. Am. Chem. Soc.* **1994**, *116*, 3556. Gresh, N.; Garmer, D. R. *J. Comput. Chem.* **1996**, *17*, 1481.

10.1021/ja984470t CCC: $18.00
Published on Web 08/10/1999

are magnesium−aqua complexes,[7−10,12,14−16,18] clusters containing one[8−11] or two[8] formates, and complexes with a single neutral ligand such as formamide,[8,9] carbonyls,[13] alcohols, and thioalcohols.[9,13] In addition, Krauss and Stevens have evaluated the binding energies for a series of hexacoordinated magnesium complexes containing one formate and varying numbers of water and formamide ligands.[11] In most cases only the binding energy[8−12,15] or enthalpy[12,16,18] was calculated, and the entropic contribution to metal binding was not taken into account. However, no systematic study on the thermodynamics of the water substitution reactions, eqs 1−3, appears to have been performed to date.

In this work density functional theory (DFT) was employed to evaluate the free energies for successive water substitution reactions in octahedral Mg^{2+} complexes [eqs 1−3]. A series of clusters $[Mg(H_2O)_{6-n}L_n]^{2+nz}$ ($n = 1$ to 6, $z = 0$ or -1) having up to six water molecules replaced by heavier ligands L were examined. The ligands L are simple organic molecules that model the amino acid residues most commonly found coordinated to magnesium in proteins. These are (1) formate, $HCOO^-$ (for negatively charged aspartic and glutamic acid side chains), (2) formic acid, HCOOH (for neutral aspartic and glutamic acid side chains), (3) formamide, $HCONH_2$ (for asparagine and glutamine side chains, as well as the backbone amide group), and (4) methanol, CH_3OH (for serine and threonine side chains). In addition, mixed complexes containing water and two other types of ligands, formate and formamide, $[Mg(H_2O)_{6-m-n}(HCONH_2)_m(HCOO)_n]^{2+nz}$ ($m = 1-5$; $n = 1, 2, 3$) were studied. Subsequently, continuum dielectric theory was used to estimate the substitution free energies for eqs 1−3 in a protein environment by modeling the latter with a dielectric constant equal to 2 or 4. The Protein Data Bank (PDB)[19] was surveyed for 3-dimensional structures of magnesium-binding sites in proteins. Since only amino acid ligands were modeled in the present work, the survey excluded sites in which magnesium was coordinated to non-amino-acid residues, such as porphyrin rings in bacteriochlorophyll and phosphate groups in nucleic acids.

The DFT and continuum dielectric calculations as well as the PDB survey are described in the Methods section. The ab initio free energies were calibrated against available experimental data and the individual contributions to the binding energies of magnesium complexes were identified (see Results). Furthermore, the changes in the enthalpy, entropy, and free energy for the water-substitution reactions in the gas phase were computed. The free energy changes for eqs 1−3 in a protein environment, modeled by a dielectric constant equal to 2 or 4, were also evaluated. The results from the DFT and continuum dielectric calculations, in combination with those from the PDB survey, were used to delineate the most thermodynamically preferable, first-coordination shell for divalent magnesium (see Discussion). The key results of this work are highlighted in the Conclusions section.

Methods

DFT Calculations. Due to the explicit treatment of electron correlation at a relatively low computational cost, DFT methods have gained popularity over the past decade. DFT calculations have been shown to be especially well suited for evaluating the physicochemical properties of various metal−ligand clusters.[20] The DFT calculations employed Becke's three-parameter hybrid method[21] in conjunction with the Lee, Yang, and Parr correlation functional.[22] Since including polarization and especially diffuse functions on heavy atoms is crucial for the proper treatment of metal ion−ligand interactions,[12,18,23] the 6-31+G* basis set was used. This basis set represents a reasonable compromise between performance and computational cost for the present calculations.

Full geometry optimization for the magnesium complexes was carried out at the B3LYP/6-31+G* level using the Gaussian 94 program.[24] Vibrational frequencies were computed to verify that each cluster was at the minimum of its potential energy surface. No imaginary frequency was found in any of the clusters. The zero point energy (ZPE), thermal energy (E_{TRV}), and entropy (S_{TRV}) were evaluated with the frequencies scaled by an empirical factor of 0.9613[25] using standard statistical mechanical formulas.[26] The differences in the electronic energy ΔE_{elec}, ΔZPE, ΔE_{TRV}, and ΔS_{TRV} between the products and reactants were then employed to compute the free energies for the water−ligand exchange reactions [eqs 1−3] at room temperature, $T = 298.15$ K, according to the following expression:

$$\Delta G_{subst}{}^1 = \Delta E_{elec} + \Delta ZPE + \Delta E_{TRV} - T\Delta S_{TRV} \qquad (4)$$

Energy Decomposition Analysis. The binding energy of some complexes was decomposed into different fragments employing the reduced variational space (RVS) scheme of Stevens and Fink.[27] The RVS analysis partitions the binding energy (BE) into electrostatic (ES), exchange (EX), polarization (PL), and charge transfer (CT) contributions. The RVS energy decomposition analysis has been implemented in the GAMESS program package.[28] Calculations were performed at the HF/6-31+G*//B3LYP/6-31+G* level of theory. Since the direct mode is not available for this analysis, hundreds of millions of two-electron integrals have to be stored on a hard disk for each run. Therefore, the RVS energy decomposition analyses were limited to smaller-size clusters.

Continuum Dielectric Calculations. The successive substitution free energies for eqs 1−3 in a protein environment was evaluated using the following thermodynamic cycle:

$$\begin{array}{ccc} & \Delta G_{subst}{}^1 & \\ A(\varepsilon=1) & \rightarrow & B(\varepsilon=1) \\ \Delta G_{solv}{}^x(A)\downarrow & & \downarrow \Delta G_{solv}{}^x(B) \\ A(\varepsilon=x) & \rightarrow & B(\varepsilon=x) \\ & \Delta G_{subst}{}^x & \end{array}$$

$\Delta G_{subst}{}^1$ is the gas-phase substitution free energy computed using eq 4.

(10) Floris, M. F.; Persico, M.; Tani, A.; Tomasi, J. *Chem. Phys.* **1995**, *195*, 207.
(11) Krauss, M.; Stevens, W. J. *J. Am. Chem. Soc.* **1990**, *112*, 1460.
(12) Glendening, E. D.; Feller, D. *J. Phys. Chem.* **1996**, *100*, 4790.
(13) Partridge, H.; Bauschlicher, C. W., Jr. *J. Phys. Chem.* **1992**, *96*, 8827.
(14) Marcos, E. S.; Pappalardo, R. R.; Rinaldi, D. *J. Phys. Chem.* **1991**, *95*, 8928.
(15) Klobukowski, M. *Can. J. Chem.* **1992**, *70*, 589.
(16) Katz, A. K.; Glusker, J. P.; Beebe, S. A.; Bock, C. W. *J. Am. Chem. Soc.* **1996**, *118*, 5752.
(17) Sponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951.
(18) Pavlov, M.; Siegbahn, P. E. M.; Sandström, M. *J. Phys. Chem. A* **1998**, *102*, 219.
(19) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimannouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535−542.
(20) Ziegler, T. *Can. J. Chem.* **1995**, *73*, 743.
(21) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(22) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *B37*, 785.
(23) Hartmann, M.; Clark, T.; van Eldik, R. *J. Am. Chem. Soc.* **1997**, *119*, 7843.
(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision D.4*; Gaussian, Inc.: Pittsburgh, PA, 1994.

ΔG_{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[29,30] to yield ΔG_{solv}^{x} (see below), ΔG_{subst}^{x}, the substitution free energy in a protein environment modeled by dielectric constant *x*, can be computed from:

$$\Delta G_{subst}^{x} = \Delta G_{subst}^{1} + \Delta G_{solv}^{x}(B) - \Delta G_{solv}^{x}(A) \qquad (5)$$

The continuum dielectric calculations employed a 71 × 71 × 71 lattice centered on the magnesium with a grid spacing of 0.25 Å. The low-dielectric region of the solute was defined as the region inaccessible to contact by a 1.4 Å sphere rolling over the molecular surface. This region was assigned a dielectric constant of two 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[31] (version 22) van der Waals radii to reproduce the experimental *hydration* free energies of Mg^{2+} (−455 kcal/mol[4]), H_2O (−6.3 kcal/mol[32]), CH_3OH (−5.1 kcal/mol[32]), CH_3COOH (−6.7 kcal/mol[32]), CH_3CONH_2 (−9.7 kcal/mol[32]), and CH_3COO^- (−82 kcal/mol[30]). The following atomic radii (in Å) were employed in the calculations: Mg, 1.50; C,1.80; O,1.77; O(COO^-), 1.60; N, 1.77; H, 1.15; and H(N), 1.00. The ab initio geometries and ChelpG[33] partial atomic charges, which were assumed to be the same in a vacuum and solution, were employed in the continuum dielectric calculations. The dielectric constant of a protein is generally assumed to range between 2 and 4,[34] thus Poisson's equation was solved with an external dielectric constant equal to 2 or 4. The difference between the electrostatic potential calculated for the low-dielectric media of the protein ($\epsilon = 2$ or 4) and vacuum ($\epsilon = 1$) yielded the electrostatic solvation free energy ΔG_{solv} of the magnesium complex.

Database Survey. The PDB[19] was surveyed for X-ray and NMR structures of magnesium-bound proteins, in which magnesium was coordinated to amino acid residues. Only four protein NMR structures in the PDB were found to contain magnesium. However, the magnesium in the NMR structures was coordinated to the phosphate groups of guanine diphosphate, hence these structures were excluded from the analyses. A *single* representative of a given family of proteins (namely, the structure solved at the highest resolution) was included in the survey. In addition, the X-ray structures surveyed satisfy the following criterion: (i) the resolution is below 3.0 Å; (ii) apart from Mg^{2+}, no other cofactors such as phosphate or sulfate groups are present at a given binding site; and (iii) magnesium plays a structural or catalytic role.

In the fully optimized DFT complexes, the Mg−O bond distance ranged from 2.0 to 2.2 Å. However, in magnesium-bound proteins, the Mg−O bond distance may exceed 2.2 Å due to the lower resolution (>2 Å) of several of the structures surveyed. Thus, ligands having a Mg−O bond distance shorter than 2.8 Å were considered to participate in inner-sphere binding to the metal. Water oxygen atoms from the Mg^{2+} first coordination shell were not always seen due to the lower resolution of some crystal structures. Therefore, the results were based on the protein ligands in the respective binding sites, assuming that water molecules completed the rest of the first coordination shell and that the coordination number of magnesium is six.

Table 1. Experimental and Calculated Relative Free Energies $\Delta\Delta G_i$ (=$\Delta G_i - \Delta G_{Mg-methanol}$) of Mg^+−L Complex Formation (in kcal/mol)

ligand L	$\Delta\Delta G^{exp}$ ref 35	$\Delta\Delta G^{calc}$ ref 13	$\Delta\Delta G^{calc}$ this work	$\Delta\Delta G^{calc}$(BSSE) this work
CH_3OH	0	0	0	0
CH_3CH_2OH	−1.77	−2.14	−2.14	−2.23
$CH_3CH_2CH_2OH$	−2.95	−3.46	−3.32	−3.39
CH_3CHO	−2.28	−2.04	−3.22	−3.75
CH_3CH_2CHO	−3.67	−3.29	−4.62	−5.17
$CH_3CH_2CH_2CHO$	−4.60	−4.03	−5.42	−5.96

Results

Calibration of the DFT Calculations. The B3LYP/6-31+G* ab initio calculations were calibrated against available experimental data. Operti, Tews, and Freiser have measured the free energies of formation of a series of Mg^+−single ligand complexes in the gas phase.[35] They examined the complexation of singly charged Mg with simple organic ligands such as alcohols, ethers, and aldehydes and reported free energies relative to that of the Mg^+−methanol complex. In Table 1 the B3LYP/6-31+G* relative free energies $\Delta\Delta G_i$ (= $\Delta G_i - \Delta G_{Mg-methanol}$) of Mg^+ complexed with methanol, ethanol, *n*-propanol, ethanal, propanal, and *n*-butanal are compared with the experimental data[35] and with another set of ab initio free energies obtained by Partridge and Bauschlicher at the MP2/TZ2P level.[13] Table 1 shows reasonably good agreement between our results and the experimental data. The B3LYP/6-31+G* free energies reproduce the order of the experimental free energy differences: $|\Delta\Delta G|_{Mg-ethanol} < |\Delta\Delta G|_{Mg-ethanal} < |\Delta\Delta G|_{Mg-n-propanol} < |\Delta\Delta G|_{Mg-n-propanal} < |\Delta\Delta G|_{Mg-n-butanal}$. Note that this is not the case with the MP2/TZ2P calculations,[13] which predict $|\Delta\Delta G|_{Mg-ethanal} < |\Delta\Delta G|_{Mg-ethanol} < |\Delta\Delta G|_{Mg-n-propanal} < |\Delta\Delta G|_{Mg-n-propanol} < |\Delta\Delta G|_{Mg-n-butanal}$. The B3LYP/6-31+G* free energies are systematically overestimated relative to the experimental values but the errors, though overall larger than those of the MP2/TZ2P calculations, do not exceed 1 kcal/mol, and are less than 0.4 kcal/mol for the Mg^+−alcohol complexes.

The approximate method of Boys and Bernardi,[36] known as the counterpoise method, was used to evaluate the basis set superposition error (BSSE). Although the absolute BSSEs for the individual complexes are between 0.5 and 1.1 kcal/mol, the BSSEs relative to that of the Mg^+−methanol complex are smaller ranging from 0.07 to 0.55 kcal/mol. Inspection of the corrected free energies in Table 1 shows that including the BSSE term in the calculations increases the overestimation of the theoretical $\Delta\Delta G_i$ relative to the experimental values. Thus, taking into account the basis set superposition error does not appear to improve the quality of the calculations, a tendency also observed by other workers.[12,16,18] This observation and the fact that it is the free energy difference for the water substitution reactions, eqs 1−3, that is of interest prompted us to omit the BSSE corrections in the results presented below.

Binding Energy Components of Mg^{2+}−Single Ligand Complexes. Table 2 summarizes the ab initio *absolute* dipole moments and mean polarizabilities for the H_2O, CH_3OH, HCOOH, $HCONH_2$, and $HCOO^-$ ligands and the RVS decomposition terms for the respective Mg^{2+}−L complexes. The latter results show that the electrostatic interaction between the metal and the ligand accounts for most of the complex binding energy.

(25) Wong, M. W. *Chem. Phys. Lett.* **1996**, *256*, 391.
(26) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976.
(27) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15.
(28) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
(29) Gilson, M. K.; Honig, B. H. *Proteins* **1988**, *4*, 7.
(30) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 5610.
(31) Brooks, B. R.; Bruccoleri, R. D.; Olafson, B. O.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(32) Cabani, S.; Gianni, P.; Mollica V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563. Ben-Naim, A. Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.
(33) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.
(34) Harvey, S. C.; Hoekstra, P. *J. Phys. Chem.* **1972**, *76*, 2987. Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 2097.
(35) Operti, L.; Tews, E. C.; Freiser, B. S. *J. Am. Chem. Soc.* **1988**, *110*, 3847.
(36) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

Table 2. Ab Initio Dipole Moments (in D) and Mean Polarizabilities (in au[3]) for H_2O, CH_3OH, HCOOH, $HCONH_2$,and $HCOO^-$, as Well as the Electrostatic (ES), Exchange (EX), Polarization (PL), and Charge Transfer (CT) Contributions to the Binding (BE) Energies (in kcal/mol) for the Respective Mg^{2+}−Single Ligand complexes

	H_2O	CH_3OH	HCOOH	$HCONH_2$	$HCOO^-$
$\|\mu\|_L$	2.25	1.94	1.52	4.13	
$\bar{\alpha}_L$	6.90	17.97	20.26	24.99	29.54
ES	−79.2	−88.4	−85.0	−109.1	−357.7
EX	26.0	30.8	28.9	33.7	54.5
PL_L	−25.0	−38.0	−48.1	−54.4	−52.3
$CT_{L\rightarrow Mg}$	−2.9	−3.4	−3.5	−4.0	−12.6
$PL_{Mg} + CT_{Mg\rightarrow L}$	−0.1	0.0	−0.8	0.0	0.0
BE	−81.2	−99.0	−108.5	−133.8	−368.1

Table 3. Results from RVS Energy Decomposition Analysis for Complexes of Mg^{2+} with Water (w), Formamide (am), and Formate (f^-) Ligands (all energies in kcal/mol)

	w-Mg^{2+}-f^-	am-Mg^{2+}-f^-	w-Mg^{2+}-$(f^-)_2$	am-Mg^{2+}-$(f^-)_2$
ES	−417.7	−432.4	−617.1	−625.3
EX	71.0	72.2	85.5	89.8
PL_f	−47.9	−45.9	⟨−27.5⟩	⟨−26.8⟩
PL_w	−16.2		−8.2	
PL_{am}		−33.4		−14.3
$CT_{f\rightarrow Mg}$	−11.2	−10.8	⟨−6.7⟩	⟨−6.6⟩
$CT_{w\rightarrow Mg}$	−1.0		0.1	
$CT_{am\rightarrow Mg}$		−2.2		−2.1
BE	−423.1	−452.6	−608.3	−618.7

As expected, the electrostatic metal−ligand interaction energy is most favorable for complexes with negatively charged ligands such as the magnesium−formate complex. The positive exchange term (EX) is offset by the favorable polarization contribution from the ligand (PL_L). For the *neutral* ligands, the PL_L term becomes more favorable with increasing mean polarizability and accounts for 31−44% of the overall binding energy. Thus polarization of the ligand plays an important role in the formation of Mg^{2+} complexes with neutral nonaqua ligands. In contrast, charge transfer from the ligand to magnesium contributes less than 4% of the total binding energy, while polarization and charge-transfer effects involving the magnesium cation are negligible. Since the magnesium−aqua complex has the least favorable binding energy in the series (−81 kcal/mol for water vs −99, −109, −134, and −368 kcal/mol for methanol, formic acid, formamide, and formate, respectively), Mg^{2+} will prefer to bind to the heavier ligands rather than water in the gas phase.

Binding Energy Components for Mg^{2+}−Mixed-Ligand Complexes. Table 3 summarizes the RVS decomposition results for mixed complexes of Mg^{2+} with water, formamide, and formate: $[Mg(H_2O)(HCOO)]^+$, $[Mg(HCONH_2)(HCOO)]^+$, $[Mg(H_2O)(HCOO)_2]^0$, and $[Mg(HCONH_2)(HCOO)_2]^0$. Binding to the negatively charged formate(s) decreases the positive charge on the metal ion, and so the polarization contribution to the binding energy is expected to decrease in magnitude accordingly. Indeed, the magnitude of PL_{H_2O} decreases from 25.0 kcal/mol in the Mg−water complex (Table 2) to 16.2 and 8.2 kcal/mol in the Mg−water−monoformate and Mg−water−diformate complexes, respectively (Table 3). The same trend is observed for the polarization contribution from the formamide ligand, PL_{HCONH_2}, where the respective numbers are 54.4, 33.4, and 14.3 kcal/mol for the Mg−formamide, Mg−formamide−monoformate and Mg−formamide−diformate complexes, respectively. Note that for the dicationic, monocationic, and neutral complexes, the differences between PL_{H_2O} and PL_{HCONH_2} are 29.4, 17.2, and 6.1 kcal/mol, respectively, while the respective differences between ES_{H_2O} and ES_{HCONH_2} are 29.9, 14.7, and 8.2 kcal/mol. Consequently, the total binding energy difference between water and formamide complexes diminishes with increasing negative charge density around the metal cation (Tables 2 and 3).

Table 4. Calculated Thermodynamical Parameters for the Water Substitution Reactions in $[Mg(H_2O)_{6-n}(L)_n]^{2+}$ ($n = 0$, ..., 5) Complexes[a]

reaction	ΔH_{subst}^1, kcal/mol	$T\Delta S_{subst}^1$, kcal/mol	ΔG_{subst}^1, kcal/mol	ΔG_{subst}^2, kcal/mol	ΔG_{subst}^4, kcal/mol
$[Mg(H_2O)_{6-n}(CH_3OH)_n]^{2+} + CH_3OH \rightarrow [Mg(H_2O)_{5-n}(CH_3OH)_{n+1}]^{2+} + H_2O$					
$n = 0$	−3.3	+1.0	−4.3	−1.5	−0.3
$n = 1$	−2.9	−1.2	−1.7	+1.0	+2.2
$n = 2$	−3.1	−0.3	−2.8	−0.1	+1.0
$n = 3$	−2.5	−1.2	−1.3	+0.8	+1.8
$n = 4$	−2.7	−0.6	−2.1	+0.3	+1.3
$n = 5$	−2.4	−1.5	−0.9	+0.8	+1.8
$[Mg(H_2O)_{6-n}(HCOOH)_n]^{2+} + HCOOH \rightarrow [Mg(H_2O)_{5-n}(HCOOH)_{n+1}]^{2+} + H_2O$					
$n = 0$	−2.5	+1.5	−4.0	+0.6	+2.5
$n = 1$	−1.9	−0.1	−1.8	+2.4	+4.2
$n = 2$	−3.4	−1.1	−2.3	+1.0	+2.3
$n = 3$	−1.8	−1.0	−0.8	+2.1	+3.3
$n = 4$	−0.8	+0.4	−1.2	+1.4	+2.6
$n = 5$	−1.7	−0.8	−0.9	+1.4	+2.3
$[Mg(H_2O)_{6-n}(HCONH_2)_n]^{2+} + HCONH_2 \rightarrow [Mg(H_2O)_{5-n}(HCONH_2)_{n+1}]^{2+} + H_2O$					
$n = 0$	−16.5	+1.4	−17.9	−10.6	−7.0
$n = 1$	−13.3	−2.0	−11.3	−5.4	−2.8
$n = 2$	−13.1	−1.2	−11.9	−6.8	−4.3
$n = 3$	−10.7	+0.8	−11.5	−7.4	−5.4
$n = 4$	−10.9	−2.3	−8.6	−4.9	−3.3
$n = 5$	−8.8	−1.7	−7.1	−3.4	−1.5
$[Mg(H_2O)_{6-n}(HCOO)_n]^{2-n} + HCOO^- \rightarrow [Mg(H_2O)_{5-n}(HCOO)_{n+1}]^{1-n} + H_2O$					
$n = 0$	−197.1	0.5	−197.6	−93.3	−41.3
$n = 1$	−117.9	−2.6	−115.3	−54.3	−23.4
$n = 2$	−35.0	−3.0	−32.0	−15.4	−6.8
$n = 3$	+42.9	0.0	+42.9	+20.0	+8.6

[a] $-\Delta G_{solv}^2$ ($-\Delta G_{solv}^4$) values for $Mg(H_2O)_6$, water, methanol, formic acid, formamide, and formate are 97.9 (148.8), 2.0 (3.7), 1.4 (2.6), 2.1 (3.9), 3.2 (5.6), and 38.7 (59.0) kcal/mol, respectively. $-\Delta G_{solv}^2$ ($-\Delta G_{solv}^4$) values for the methanol, HCOOH, and formamide complexes range from 94.5 (143.7) to 79.9 (121.3), 93.4 (142.5) to 78.6 (121.8), and 91.8 (139.8) to 75.3 (116.2) kcal/mol, respectively. $-\Delta G_{solv}^2$ ($-\Delta G_{solv}^4$) values for $[Mg(H_2O)_{5-n}(HCOO)_{n+1}]^{1-n}$, $n = 0-3$, are 30.3 (47.8), 6.0 (11.2), 26.1 (41.3), and 85.7 (130.9) kcal/mol, respectively.

Thermodynamics of Water−Ligand Exchange in $[Mg(H_2O)_{6-n}(L^z)_n]^{2+nz}$ ($n = 0-5$) Complexes. Changes in the gas-phase enthalpy, entropy, and free energy for the exchange of water with methanol, formic acid, formamide, and formate are listed in Table 4. The corresponding substitution free energies for a protein environment characterized by $\epsilon = 2$ and 4 are presented in the last two columns of Table 4. The errors in these free energies are expected to be at least as large as the errors in the respective gas-phase numbers (∼1 kcal/mol). Furthermore, the error in ΔG_{subst}^2 or ΔG_{subst}^4 for a water ↔ neutral ligand exchange (where the solvation free energy of each reactant is similar to that of the corresponding product) is expected to be smaller than that for a water ↔ formate anion exchange (where the net charge on each reactant differs from the net charge on the corresponding product). Consideration of the errors in the computed free energies has been taken into account in interpreting the results.

(a) L = CH_3OH. For each successive replacement of water with methanol in the gas phase, the enthalpic contribution is favorable (negative ΔH_{subst}) and the magnitude of ΔH_{subst} is

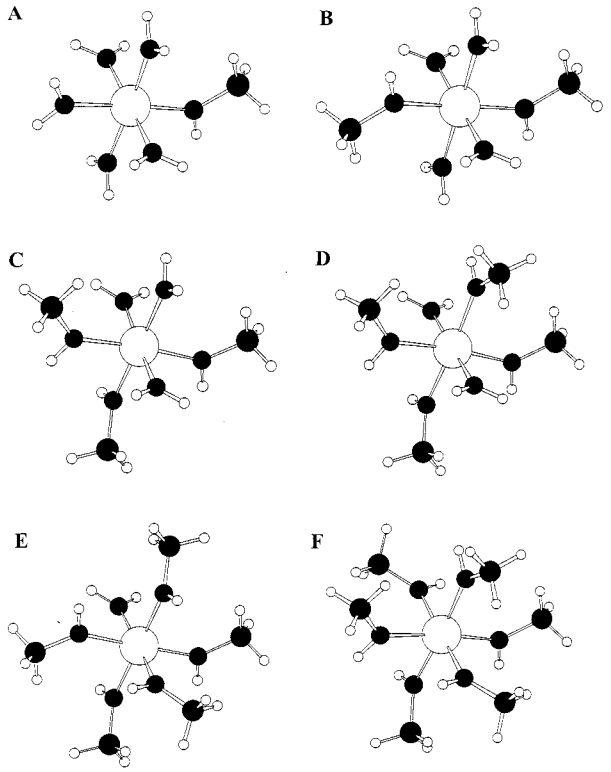

Figure 1. Ball and stick diagram of the lowest energy $[Mg(H_2O)_{6-n}(CH_3OH)_n]^{2+}$ (n = 1, ..., 6) complexes.

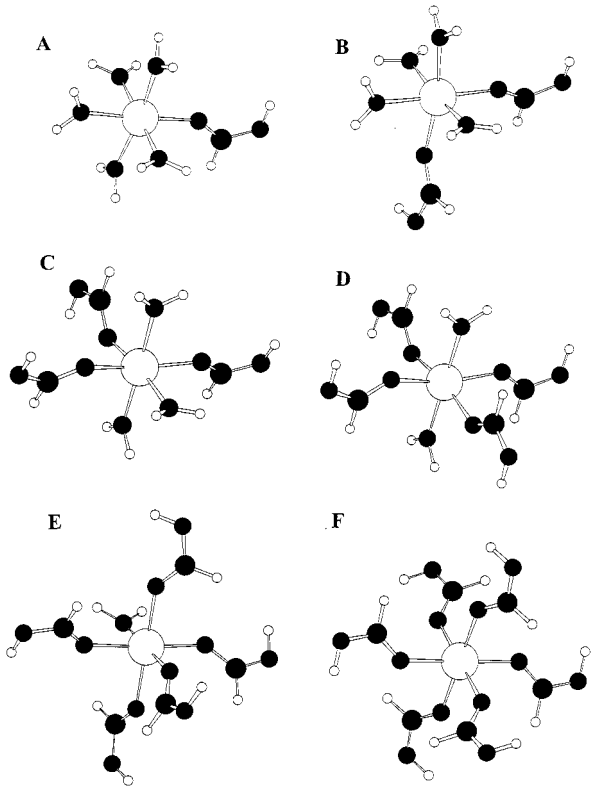

Figure 2. Ball and stick diagram of the lowest energy $[Mg(H_2O)_{6-n}(HCOOH)_n]^{2+}$ (n = 1, ..., 6) complexes.

greater than the corresponding $T\Delta S_{subst}$ term. In contrast, the entropic contribution is unfavorable (negative ΔS_{subst}) except for the first substitution reaction where $T\Delta S_{subst}$ is positive (1.0 kcal/mol). Replacing a water in $[Mg(H_2O)_6]^{2+}$ with methanol results in a loss of symmetry (the octahedral T_h symmetry reduces to C_s) and hence, a less compact complex. The entropy term $T\Delta S_{subst}$ increases from −56.9 kcal/mol for $[Mg(H_2O)_6]^{2+}$ to −55.9 kcal/mol for $[Mg(H_2O)_5(CH_3OH)]^{2+}$ and, as a result of this increase, the first water substitution reaction shows the largest free energy gain ($\Delta G_{subst}^1 = -4.3$ kcal/mol). The next four water ↔ methanol exchange reactions are also predicted to be thermodynamically favorable in the gas phase ($\Delta G_{subst}^1 = -1.3$ to −2.8 kcal/mol), but the last water ↔ methanol exchange may not occur since the computed ΔG_{subst}^1 of −0.9 kcal/mol is within the error range of the present calculations. The finding that methanol can displace a Mg^{2+}-bound water in the gas phase correlates with the observation that the Mg−O(CH_3OH) bond distances are shorter than the Mg−O(H_2O) distances in the $[Mg(H_2O)_{6-n}(CH_3OH)_n]^{2+}$ (n = 1, ..., 6) complexes, whose fully optimized structures are illustrated in Figure 1. However, in a protein medium characterized by ϵ = 4, methanol may not be able to replace a Mg^{2+}-bound water without additional stabilizing interactions from the protein matrix.

(b) L = HCOOH. Formic acid has a smaller molecular dipole moment, but a larger mean polarizability, compared to methanol (Table 2). The magnitude of ΔH_{subst} for each water ↔ formic acid gas-phase exchange (except the third one) is slightly less than the corresponding $|\Delta H_{subst}|$ for water ↔ methanol exchange. However, the magnitudes of ΔH_{subst} and $T\Delta S_{subst}$ for most of the HCOOH substitution reactions are comparable, thus, the entropy term makes a significant contribution to the net free energy for replacing water with formic acid in the gas phase. Although the successive exchange of the first three water molecules for formic acid is predicted to be thermodynamically favorable in the gas phase ($\Delta G_{subst}^1 = -1.8$ to −4.0 kcal/mol), the same cannot be said about the subsequent water substitution reactions, whose free energies ($\Delta G_{subst}^1 = -0.8$ to −1.2 kcal/mol) fall within the accuracy limits of the present calculations. Interestingly, the fully optimized $[Mg(H_2O)_{6-n}(HCOOH)_n]^{2+}$ (n = 1, ..., 6) structures in Figure 2 show that the second formic acid ligand in the $[Mg(H_2O)_4(HCOOH)_2]^{2+}$ complex prefers to occupy a position cis (rather than trans) to the first ligand, in contrast to the $[Mg(H_2O)_4(CH_3OH)_2]^{2+}$ complex (Figure 1B). The ΔG_{subst}^2 and ΔG_{subst}^4 values for the water ↔ HCOOH exchange reactions are all positive due to the unfavorable solvation of magnesium complexes containing a greater number of HCOOH ligands.

(c) L = $HCONH_2$. Formamide has both a larger molecular dipole moment and mean polarizability, relative to methanol (Table 2). Consequently, the enthalpies for the successive replacement of water with formamide in the gas phase are much more favorable than the respective ΔH_{subst} for water ↔ methanol exchange. On the other hand, the $T\Delta S_{subst}$ terms are similar in magnitude to the corresponding values in the methanol complexes, and hence the enthalpy term makes the dominant contribution to ΔG_{subst}^1. As in the case of water exchange with methanol, the first water substitution reaction has a positive entropy change ($T\Delta S_{subst}$ = 1.4 kcal/mol) and the highest free energy gain among the six reactions in the gas phase (−17.9 kcal/mol). Note that even the last water substitution reaction in the gas phase has a significant free energy gain (−7.1 kcal/mol), indicating that favorable electrostatic and polarization forces can overcome repulsive steric crowding in the $[Mg(HCONH_2)_6]^{2+}$ complex. The structures of the $[Mg(H_2O)_{6-n}(HCONH_2)_n]^{2+}$ (n = 2 or 4) complexes are similar to those of the methanol complexes (Figure 1) in that the heavy ligands are opposite one another. Unlike the water ↔ methanol and water ↔ formic-acid exchange reactions, ΔG_{subst}^2 and ΔG_{subst}^4 remain negative for the entire series of water−formamide substitutions.

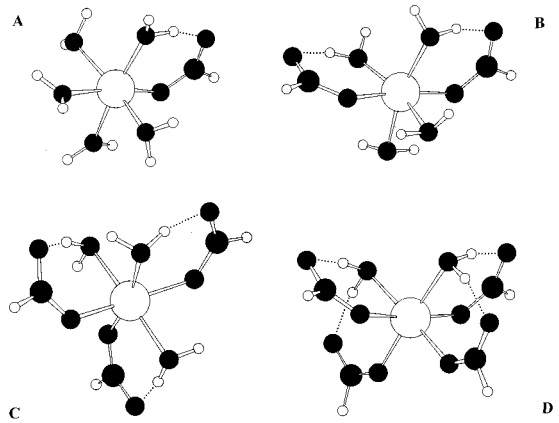

Figure 3. Ball and stick diagram of the lowest energy $[Mg(H_2O)_{6-n}(HCOO^-)_n]^{2-n}$ (n = 1, ..., 4) complexes. A dotted line denotes a hydrogen bond, which is defined by a hydrogen to oxygen distance <2.0 Å and deviation of the OH···O angle of <30° from linearity.

(d) L = HCOO$^-$. The formate complexes exhibit the most dramatic energy changes in the series under study due to the strong Coulombic forces governing the ligand−metal and ligand−ligand interactions. In contrast to the substitution of water with neutral formic acid, the magnitude of $T\Delta S_{subst}$ is negligible compared to that of the respective ΔH_{subst}, which dictates the free energy change, $\Delta G_{subst}{}^1$. The successive displacement of the first three water molecules with formate in the gas phase is highly favorable. The first substitution gains a free energy of −197.6 kcal/mol, the second one, −115.3 kcal/mol, and the third one, −32.0 kcal/mol. Two factors contribute to the stability of these isolated gas-phase complexes: (1) the strong attractive Coulombic forces between Mg^{2+} and the $HCOO^-$ ligands and (2) the hydrogen bonds formed between each formate and a neighboring water molecule (see Figure 3). The attack of a fourth formate on the negatively charged complex containing three formates is thermodynamically unfavorable: $\Delta G_{subst}{}^1$ becomes positive (42.9 kcal/mol). Hence, subsequent water-substitution reactions in complexes with four or five formates were not considered.

The $\Delta G_{subst}{}^2$ and $\Delta G_{subst}{}^4$ follow the same trend of changes found in the gas phase, although their absolute values are smaller. Note that the free energy change for the water−formate substitution reactions depends strongly on the dielectric constant of the medium. In a protein environment where $\epsilon = 2-4$, ΔG_{subst} for the first substitution reaction varies between −93.3 and −41.3 kcal/mol (Table 4), but it becomes positive (+8.8 kcal/mol) in aqueous solution where $\epsilon = 80$, implying that the water−formate exchange reaction in water is thermodynamically unlikely.

Thermodynamics of Water−Formamide Exchange in $[Mg(H_2O)_{6-m-n}(HCONH_2)_m(HCOO)_n]^{2-n}$ (m = 0, ..., 4; n = 1, 2, 3) Complexes. Aqua substitution chemistry in metal complexes is a competition between the water and the larger organic ligand for the metal. In the gas phase, both electrostatic and van der Waals forces govern the interaction of the ligand with the metal cation. For neutral ligands, charge-dipole and charge-induced dipole forces dominate the interaction. On the other hand, for bulkier organic ligands there are stronger repulsive steric interactions between ligands. Water is a small molecule and forms a compact complex. However, it has a lower polarizability compared with the other neutral ligands included in this study (Table 2), and as shown in Table 4, cannot compete successfully for the metal in the gas phase. For bulkier ligands the binding energy can apparently offset the unfavorable steric crowding of the heavy ligands around Mg^{2+}.

Table 5. Calculated Thermodynamical Parameters for the Water Substitution Reactions in $[Mg(H_2O)_{6-m-n}(HCONH_2)_m(HCOO)_n]^{2-n}$ (m = 0−4; n = 1, 2, 3) Complexes[a]

reaction	$\Delta H_{subst}{}^1$, kcal/mol	$T\Delta S_{subst}{}^1$, kcal/mol	$\Delta G_{subst}{}^1$, kcal/mol	$\Delta G_{subst}{}^2$, kcal/mol	$\Delta G_{subst}{}^4$, kcal/mol
$[Mg(H_2O)_{5-m}(HCOO)(HCONH_2)_m]^+ + HCONH_2 \rightarrow [Mg(H_2O)_{4-m}(HCOO)(HCONH_2)_{m+1}]^+ + H_2O$					
$m = 0$	−7.6	−0.9	−6.7	−4.5	−3.5
$m = 1$	−7.4	−2.9	−4.5	−1.7	−0.4
$m = 2$	−8.5	−2.4	−6.1	−3.5	−2.1
$m = 3$	−4.4	−0.5	−3.9	−2.5	−1.9
$m = 4$	+0.9	−0.8	+1.7	+2.6	+2.7
$[Mg(H_2O)_{4-m}(HCOO)_2(HCONH_2)_m]^0 + HCONH_2 \rightarrow [Mg(H_2O)_{3-m}(HCOO)_2(HCONH_2)_{m+1}]^0 + H_2O$					
$m = 0$	−0.9	−1.5	+0.6	−0.3	−1.0
$m = 1$	+2.3	−2.5	+4.8	+3.9	+3.2
$Mg(H_2O)_3(HCOO)_3]^- + HCONH_2 \rightarrow [Mg(H_2O)_2(HCOO)_3(HCONH_2)]^- + H_2O$					
	+9.7	−0.4	+10.1	+9.4	+8.6

[a] $-\Delta G_{solv}{}^2$ ($-\Delta G_{solv}{}^4$) values for $[Mg(H_2O)_{4-m}(HCOO)(HCONH_2)_{m+1}]^+$, m = 0−4, are 29.3 (46.5), 27.7 (44.3), 26.3 (42.2), 26.1 (42.1), and 26.4 (43.0) kcal/mol, respectively. $-\Delta G_{solv}{}^2$ ($-\Delta G_{solv}{}^4$) values for $[Mg(H_2O)_3(HCOO)_2(HCONH_2)]^0$, $[Mg(H_2O)_2(HCOO)_2(HCONH_2)_2]^0$, and $[Mg(H_2O)_2(HCOO)_3(HCONH_2)]^-$ are 8.1 (14.7), 10.2 (18.2), and 28.0 (44.7) kcal/mol, respectively.

For neutral ligands the dipoles are placed in the strong polarizing field of the divalent magnesium cation. However, when the positive charge on magnesium is partially neutralized, for example by coordination to negatively charged ligand(s), the polarization contribution from the ligand rapidly attenuates and becomes less crucial in stabilizing the cluster (see Tables 2 and 3). Furthermore, the distinct contributions of polarization and electrostatics for water versus organic ligands to the binding free energy diminishes as the number of negatively charged ligands around Mg^{2+} increases (see above). How will this affect the competition between water and the heavier ligand for magnesium in the octahedral complexes considered herein? To answer this question a series of mixed water−formate−formamide complexes were examined and the results obtained are presented in Table 5.

In the complexes containing one formate the charge on the magnesium ion is partially neutralized. Thus the enthalpy change for each successive substitution of water with formamide in the monocationic complex (Table 5) is less favorable than the corresponding ΔH_{subst} in the dicationic complexes (Table 4), and becomes slightly positive for the last water ↔ formamide exchange (ΔH_{subst} = 0.9 kcal/mol). Since the $T\Delta S_{subst}$ terms, which are similar in magnitude to those in Table 4, are negative for all the water substitution reactions, the $\Delta G_{subst}{}^1$ values for the monocationic complexes in Table 5 are, like ΔH_{subst}, less favorable compared to those for the respective dicationic complexes. The substitution free energies in both gas phase and $\epsilon = 2$ *or* 4 for the first four reactions in Table 5 are negative, indicating that replacing water with formamide is still a favorable event. However, the $\Delta G_{subst}{}^1$, $\Delta G_{subst}{}^2$, and $\Delta G_{subst}{}^4$ values for the last water ↔ formamide exchange are positive (1.7, 2.6, and 2.7 kcal/mol, respectively), implying that this substitution process may not occur. A probable explanation is that in the Mg^{2+} complex with no waters the repulsive interactions among the six heavy ligands may offset the energy gain resulting from the fifth formamide binding, which, in the circumstance, is not large. The structure of the $[Mg(HCONH_2)_5(HCOO)]^+$ complex is shown in Figure 4A.

The increased negative charge density surrounding the metal ion due to the presence of two or three formates decreases the strength of the external field (magnesium cation) such that the

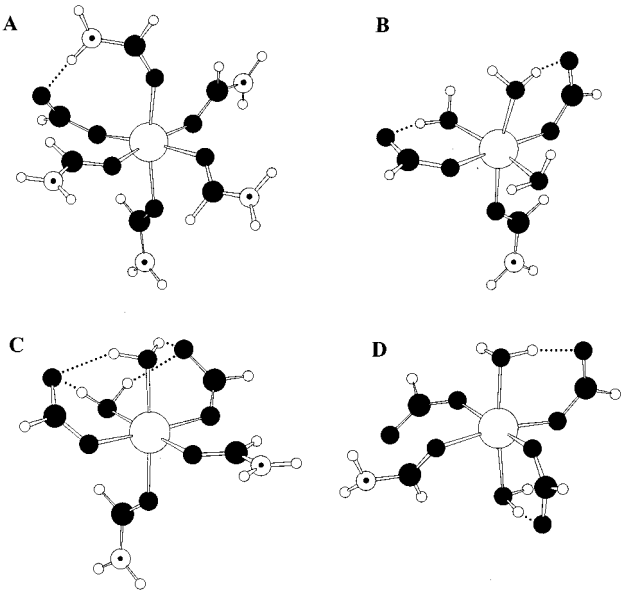

Figure 4. Ball and stick diagram of the $[Mg(HCONH_2)_5(HCOO)]^{1+}$ (A), $[Mg(H_2O)_3(HCONH_2)(HCOO)_2]^0$ (B), $[Mg(H_2O)_2(HCONH_2)_2(HCOO)_2]^0$ (C), and $[Mg(H_2O)_2(HCONH_2)(HCOO)_3]^{1-}$ (D) complexes. A dotted line denotes a hydrogen bond, which is defined by a hydrogen-to-oxygen distance <2.0 Å and deviation of the OH···O angle of <30° from linearity. For the non-hydrogen-bonded water in part C, each water hydrogen is within 2.0 Å of its nearest formate oxygen, but the OH···O angle (146° or 148°) is just outside the hydrogen bond definition.

size of the neutral ligand seems to be the key factor governing the substitution of water for formamide. Table 5 shows that the water molecule, and not the formamide, is now the preferred ligand. While the $[Mg(H_2O)_4(HCOO)_2]^0$ complex is almost indifferent toward exchange of the first water ($\Delta G_{subst}{}^1$, $\Delta G_{subst}{}^2$, and $\Delta G_{subst}{}^4 = 0.6$, -0.3, and -1.0 kcal/mol, respectively), it opposes the second water substitution, as evidenced by positive free energies for this reaction (Table 5). The complex containing three formates and three waters $[Mg(H_2O)_3(HCOO)_3]^-$ does not favor exchange of any of its waters: the $\Delta G_{subst}{}^1$, $\Delta G_{subst}{}^2$, and $\Delta G_{subst}{}^4$ numbers for the first water substitution are 10.1, 9.4, and 8.6 kcal/mol, respectively. The fully optimized geometries of the products for the last three reactions in Table 5 are shown in Figure 6, parts B, C, and D, respectively.

PDB Survey. The Protein Data Bank[19] was surveyed for 3-dimensional structures of Mg^{2+}-bound proteins (see Methods). Thirty-two metal-binding sites were found and grouped according to the number of acidic side chains (Asp or Glu) present. The results are presented in Table 6 and Figure 5. All sites contain Asp or Glu. Roughly equal numbers of the metal-binding sites were found to possess one (10 sites), or two (11 sites), or three (9 sites) acidic residues. Only two sites were found with four carboxylates. Most of the sites possessing noncharged ligands contain either an amide group from an asparagine side chain or a backbone carbonyl ligand (13 out of 18 sites) while the remaining neutral ligands found are Ser and Thr (3 sites), His (2 sites), and Tyr (1 site). It should be pointed out that the number of Mg^{2+}-binding sites found in the PDB is relatively small, and thus care has to be paid in interpreting the data presented in Table 6 and Figure 5.

Discussion

The dielectric properties of the solvent medium play a crucial role in determining the mode of magnesium binding. In a vacuum or a low dielectric medium ($\epsilon \leq 4$), a formate can replace a magnesium-bound water (Table 4), but in aqueous solution ($\epsilon = 80$) the water ↔ formate exchange becomes unlikely ($\Delta G_{subst} = +8.8$ kcal/mol,[37] see also Results). These results correlate with the observation that magnesium tends to bind in an inner-sphere mode[1,3] to protein cavities with limited solvent accessibility. In such low-dielectric cavities a negatively charged aspartate or glutamate side chain may replace a magnesium-bound water molecule.

The results obtained for both the gas-phase and protein environment show that the most preferable inner-sphere ligands for divalent magnesium in a low dielectric medium is a few (up to three) deprotonated acidic residues (Table 4). The free energy gain from binding to one, two, or three formates greatly exceeds that from binding to the other (noncharged) ligands such as HCOOH, CH_3OH, and $HCONH_2$. The $[Mg(H_2O)_{6-n}(HCOO^-)_n]^{2-n}$ ($n = 1$ to 3) complexes are stabilized not only by the strong Coulombic interactions between the metal cation and formate anions, but also by the hydrogen bonds formed between the carboxylate oxygens and neighboring water hydrogens (see Figure 3). These results correlate with the observation that all Mg^{2+}-binding sites found in the PDB survey contain at least one acidic amino acid residue: Asp or Glu.

The ab initio and continuum dielectric calculations on the model systems reveal the importance of the ionization potential of the acidic ligand on its ability to bind magnesium. Table 4 shows that protonation of a formate practically abolishes the free energy gain upon water exchange, so that the neutral ligand does not compete effectively with water for the metal dication. Most enzymes containing magnesium as a cofactor are active around pH 8.[38] At this pH, Asp and Glu are likely to be deprotonated (assuming typical pK_a values in proteins around 4.4)[39,40] thus providing an attractive binding target for the incoming metal cation. In cases where constellations of several carboxylate residues lie in close proximity, it is common for the pK_a of one or more side chains to increase.[41]

The calculations indicate an upper limit in the number of deprotonated acidic residues that can be coordinated to a Mg^{2+}–water cluster. Table 4 shows that for complexes containing three (or more) formates, the exchange of a Mg^{2+}-bound water for another formate becomes thermodynamically unfavorable. The dication is able to accommodate up to only three negatively charged ligands in its first coordination shell. This finding, at first glance, appears to be at odds with the four-carboxylate binding pocket (D244, D286, E180, E216; site I) found in the X-ray structure of Mg^{2+}-bound xylose isomerase (PDB entry 1XYA in Table 6), which was crystallized at pH 7.4. However, one of the glutamate ligands (E216) is shared with a second magnesium binding site (1XYA, site II), containing one histidine (H219) and three carboxylates (D254, D256, E216). Since the electron density on E216 is shared, the net negative charge involved in magnesium binding in site I will be less than -4. Alternatively, one of the acidic residues constituting binding site I may possess an abnormally high pK_a value such that it is protonated at the pH of crystallization (7.4), thereby reducing the total negative charge contribution to magnesium binding. Furthermore, E180 in the 1XYA structure has the longest Mg–O distance of all the ligands in Table 6, and this distance is just under the threshold of 2.8 Å for first-shell coordination. There is one other binding site found in the PDB survey that has four

(37) Dudev, T.; Lim, C. To be submitted for publication.
(38) Cowan, J. A. *Chem. Rev.* **1998**, *98*, 1067.
(39) Stryer, L. *Biochemistry*; Freeman: New York, 1988.
(40) Cowan, J. A. *Inorganic Biochemistry: An Introduction*, 2nd ed.; Wiley-VCH: New York, 1997.
(41) Oda, Y.; Yamazaki, T.; Nagayama, K.; Kanaya, S.; Kuroda, Y.; Nakamura, H. *Biochemistry* **1994**, *33*, 5275.

Table 6. Results from the PDB Survey on Magnesium-Bound Proteins

PDB entry	resolution (Å)	protein	ligands
		sites with one carboxylate	
1OBW	1.9	Inorganic Pyrophosphatase	Site I: D70[a]
4PAL	1.8	Parvalbumin	Site I: D53[a]
5EAS	2.2	Arystolochene Synthase	Site I: D301[c]
1CMC	1.8	Met Holo-Repressor	E19,[a] Y104(bkb)[c]
1FRF	2.7	Ni−Fe Hydrogenase	E53,[b] L495(bkb)[c]
1INP	2.3	Inositol Phosphatase	E79,[c] I155(bkb)[b]
1ALO	2.0	Aldehyde Oxyreductase	E651,[c] A649(bkb),[b] G693(bkb)[a]
1DEK	2.0	Deoxynucleoside MP Kinase	E108,[b] Q85,[c] Y42[b]
1BGL	2.5	Beta Galactosidase	SiteII: D193,[a] N18(bkb),[a]V21(bkb),[b] D151(bkb)[a]
1IDO	1.7	I-Domain from Integrin Cr3	E314,[a] S142,[a] S144,[a] T209[a]
		sites with two carboxylates	
1BGL	2.5	Beta Galactosidase	Site I: E416,[a] E461[b]
1DBR	2.4	Phosphorybosyltransferase	D147,[b] E146[a]
1PYS	2.9	Phenylalanyl-Trna Synthetase	E262,[b] E461[b]
1VSD	1.7	ASV Integrase	D64,[a] D121[a]
1ALK	2.0	Alkaline Phosphatase	D51,[a] E322,[a] T155[a]
1AR1	2.7	Cytochrome C Oxidase	D404,[a] E218,[a] H403[a]
1CHN	1.8	CheY	D13,[a] D57,[a] N59(bkb)[a]
1RDD	2.8	Ribonuclease HI	D10,[a] E48,[b] G11(bkb)[b]
1WDC	2.0	Scallop Myosin	D28,[b] D30,[b] F34(bkb)[b]
5EAS	2.2	Arystolochene Synthase	Site II: D444,[b] E452,[c] T448[c]
5ICB	1.5	Bovine Calbindine	D54,[a] D58,[b] N56,[b] E60(bkb)[a]
		sites with three carboxylates	
1DUT	1.9	Fiv Dutp Pyrophosphatase	D64A,[a] D64B,[a] D64C[a]
1E2A	2.3	Enzyme IIA	D81A,[b] D81B,[b] D81C[b]
1EBH	1.9	Enolase	D246,[a] D320,[a] E295[a]
1MDR	2.1	Mandelate Racemase	D195[a], E221[a], E247[a]
1MPM	2.6	Maltoporin Maltose Complex	D78A,[b] D78B,[b] D78C[b]
1OBW	1.9	Inorganic Pyrophosphatase	Site II: D65,[a] D70,[a] D102[a]
1AUK	2.1	Human Arylsulfatase A	D29,[b] D30,[b] D281,[c] N282[c]
1BPM	2.9	LEU Aminopeptidase	D255,[a] D332,[a] E334,[a] D332(bkb)[b]
1XYA	1.8	Xylose Isomerase	Site II: D254,[c] D256,[b] E216,[c] H219[c]
		sites with four carboxylates	
1XYA	1.8	Xylose Isomerase	Site I: D244,[b] D286,[b] E180,[c] E216[b]
4PAL	1.8	Parvalbumin	Site II: D90,[a] D92,[a] D94,[a] E101,[a] M96(bkb)[a]

[a] 1.9 < r(Mg-L) ≤ 2.2 Å. [b] 2.2 < r(Mg-L) ≤ 2.5 Å. [c] 2.5 < r(Mg-L) ≤ 2.8 Å; bkb denotes backbone.

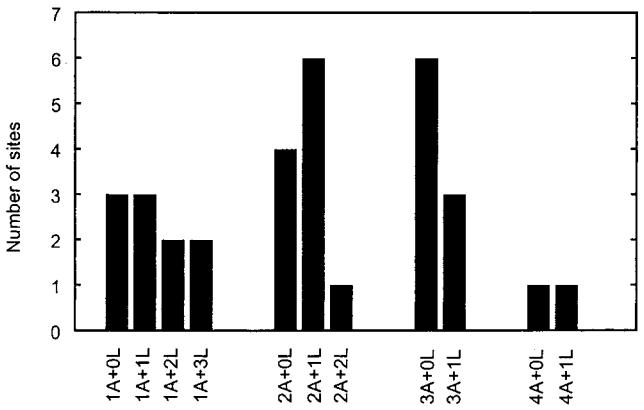

Figure 5. Frequency distribution of magnesium binding sites as a function of the number of acidic side chains (denoted by A) and number of nonacidic amino acid residues (denoted by L).

carboxylate ligands: this is site II in parvalbumin (PDB entry 4PAL in Table 6), which was crystallized at pH 7.0. Again at this pH, one of the four acidic residues constituting site II may be protonated, thus the net negative charge involved in magnesium binding may be less than −4. Furthermore, parvalbumin is a Ca^{2+}-binding protein and, therefore, the structure of the binding site is presumably adapted to the requirements of binding calcium rather than magnesium. The 4PAL X-ray structure corresponds to parvalbumin with Ca^{2+} replaced by Mg^{2+} in the presence of a high concentration of $MgSO_4$.[42]

Although the availability of pocket(s) containing negatively charged carboxylates appears to be important for magnesium binding to proteins, the noncharged amino acid residues also play a role in the complexation event. On one hand, they participate in shaping the size of the binding cavity, which may play a role in selecting the cation. On the other hand, neutral ligands, in particular amides, can contribute to the free energy of magnesium binding. Once Mg^{2+} is bound to a single Asp or Glu, coordination to backbone carbonyls or side chains of a few Asn and Gln residues may further decrease the binding free energy (Table 5). In sites containing two carboxylates, magnesium binding to one uncharged amino acid residue cannot be ruled out, whereas in sites with three carboxylates bound to magnesium, neutral amino acid residues may not be able to dislodge the Mg^{2+}-bound waters (Table 5), without compensating effects from the protein matrix. The latter may be responsible for the observation of three magnesium-binding sites containing three carboxylates and one neutral ligand (Figure 5). On the other hand, such sites may be stable because the total negative charge directly involved in magnesium binding is less than −3. For example, one of the three magnesium-binding sites with three carboxylates and one neutral ligand is site II in xylose isomerase (see above and Table 6), where E216 directly coordinates with two magnesium cations. As another example, magnesium, which replaced the natural cofactor Zn^{2+} in LEU aminopeptidase (1BPM, Table 6) during crystallization,[43] shares

(42) Declercq, J. P.; Tinant, B.; Parello, J.; Rambaud, J. *J. Mol. Biol.* **1991**, *220*, 1017.

two of the three carboxylate ligands (Asp255 and Glu334) with a zinc cation occupying a nearby binding site. In the third magnesium-binding site with three carboxylates and a neutral ligand (1AUK, Table 6), the protein human arylsulfatase A was crystallized at a pH of 5.4[44] so that one of the three aspartic acid side chains may not be fully ionized under these conditions. The theoretical results described herein imply that there is a reverse correlation between the number of negatively charged carboxylates and the number of uncharged residues: the greater the number of carboxylates in the binding pocket, the less (thermodynamically) favorable the binding to amide or alcohol derivatives.

The results in Table 5 also show that in the carboxylate complexes Mg^{2+} will retain at least one solvent water molecule. Thus, the thermodynamically most favorable configurations for the magnesium−carboxylate complexes contain, as a rule, water ligand(s). These are predicted to contain at least one water for magnesium−monocarboxylate clusters, three or four waters for magnesium−dicarboxylate clusters, and three waters for magnesium−tricarboxylate complexes. Note that the last three types of complexes appear to be the most abundant structures found in the PDB survey (Figure 5). The stability of these water-containing magnesium complexes may explain why Nature has given priority to this form of bound magnesium center whenever a metal cofactor is needed for an enzymatic hydrolytic reaction. The importance of the solvation shell in mediating the reaction chemistry of magnesium-dependent nucleases has been noted previously.[45,46] Our results also suggest that complete inner-mode binding of Mg^{2+} to proteins, where all the first-shell water molecules have been replaced by amino acid ligands, is an unlikely event.

In general, the results from the PDB survey correlate with our theoretical findings based on modeling the protein environment with a low dielectric constant. Such a correlation implies that water−protein ligand exchange reactions are likely to occur in protein cavities with low ϵ. This, in turn, implies that magnesium most likely binds to proteins in a stepwise (as opposed to a one-step) process: initially, the hydrated magnesium ion positions itself in a pocket with low ϵ, and subsequently, some of its inner-shell waters are replaced by protein ligands.

Conclusions

Our theoretical calculations on model systems in conjunction with results from a PDB survey have provided new insights on the thermodynamics of magnesium binding to proteins:

1. The dielectric properties of the solvent medium appear to govern the mode (inner/outer sphere) of magnesium binding in biological systems. Water−carboxylate exchange reactions occur only in sites characterized by a low dielectric constant, but are thermodynamically unfavorable in aqueous solution.

2. A hexahydrated magnesium dication has a high affinity for negatively charged acidic residues in protein cavities with a low dielectric constant ($\epsilon \leq 4$). However, there is an upper limit in the number of carboxylates that can be exchanged for the metal-bound water molecules in a low ($\epsilon \leq 4$) dielectric environment: only up to three negatively charged ligands can replace water in the first coordination shell of magnesium.

3. Protonated, acidic residues may not be able to dislodge magnesium-bound waters without compensating interactions from the protein matrix.

4. In carboxylate complexes Mg^{2+} is likely to retain at least one water ligand in its first coordination shell.

Acknowledgment. We are grateful to D. Bashford, M. Sommer, and M. Karplus for the program to solve the Poisson equation. T.D. is supported by the Institute of Biomedical Sciences. J.A.C. is supported by the National Science Foundation (CHE-9706904). C.L. is supported by the Institute of Biomedical Sciences at Academia Sinica, the National Center for High Performance Computing, and the National Science Council, Republic of China (NSC-88-2113-M-001).

JA984470T

(43) King, H.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5006.

(44) Lukatela, G.; Krauss, N.; Theis, K.; Selmer, T.; Gieselmann, V.; von Figura, K.; Saenger, W. *Biochemistry* **1998**, *37*, 3654.

(45) Black, C. B.; Foster, M.; Cowan, J. A. *J. Biol. Inorg. Chem.* **1996**, *1*, 500.

(46) Cowan, J. A. *Inorg. Chim. Acta* **1997**, *275/276*, 24.