Theoretical Studies of Metal Ion Selectivity. 1. DFT Calculations of Interaction Energies of Amino Acid Side Chains with Selected Transition Metal Ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+})

Lubomír Rulíšek* and Zdeněk Havlas

Contribution from the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, and Center for Complex Molecular Systems and Biomolecules, Flemigovo nám. 2, 166 10 Prague 6, Czech Republic

Received April 11, 2000. Revised Manuscript Received August 1, 2000

Abstract: The interaction energies of functional groups representing the side chains of amino acid residues with Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+} cations were computed with DFT/B3LYP method. Four coordination geometries, which are most frequently encountered in the metal-binding sites of metalloproteins and smaller-molecule crystal structures (octahedral, square planar, tetrahedral, and linear), were considered for each metal ion. The computational strategy consisted of several steps. First, the affinities of studied metal ions for $(H_2O)_n$ site, pre-organized in particular coordination geometry, have been evaluated. Second, the interaction energy of a single functional group with the transition metal ion of interest has been calculated, while the remaining coordination bonds were saturated with water molecules. Third, and finally, the effect of elongation of the amino acid side chain has been calculated. Together, it gives an insight into the molecular structure of metal-binding sites of metalloproteins and provides an accurate quantification of the affinity and selectivity of amino acid side chains for the studied metal ions. These two quantities play a key role in the metal-binding properties of proteins and peptides. The important implications in an area of bioinorganic chemistry are discussed as well.

I. Introduction

Interactions of metals with biomolecules belong to one of the most studied fields in bioinorganic chemistry. The area of applications of new knowledge ranges from medicinal chemistry[1] through "classical" organometallic chemistry to environment protection (metal binding biomass).[2] The role of metal ions in the structure and function of proteins, nucleic acids, and peptide hormones is fundamental, yet often unknown, at the atomic and electronic level. Nevertheless, the recent achievements in quantum bioinorganic chemistry fill this gap significantly (vide infra).

Of special interest is the metal ion selectivity, defined as the different affinity of specific ligand for different metal ions, which often plays a key role in the function and properties of metal-containing biomolecules. It is a difficult task to evaluate this quantity accurately, because it is determined by the subtle variations in the molecular structures and interaction energies on the background of dominant electrostatic interactions between the ionic systems.

The explanations of factors determining the specificity of metal ion uptake are often based on the qualitative or semiquantitative theories or principles, such as the HSAB (hard and soft acids and bases) principle of Parr and Pearson,[3] the Irving−Williams series of stability constants,[4] and the abundance of transition metals (TM) coordination geometries in the experimentally determined molecular structures.[5,6] Besides, there have been attempts to model TM centers by molecular mechanics[7] and the achievements have been reviewed recently by Comba.[8] However, it should be stressed that TM systems (especially those with partially filled d shells) are challenging even for the sophisticated quantum chemical theories, which implies that one cannot expect to obtain the accurate description of their properties with the force field approach. Therefore, we think that these studies also belong to the category of semiquantitative theories.

The quantitative and accurate modeling of TM centers can be carried out by using high-level quantum chemical methods. In the past decade, a huge amount of papers dealing with the theoretical calculations of miscellaneous TM systems have been published and we refer the reader to the recent reviews.[9−14] Many useful references can be also found in our paper published recently.[15] In the next paragraphs, we confine ourselves to the theoretical studies pertinent to this work.

* Corresponding author. Telephone and Fax: +420-2-20183292. E-mail: lubos@uochb.cas.cz.

(1) *Handbook of Metal−Ligand Interactions in Biological Fluids*; Berthon, G., Ed.; Marcel Dekker: New York, 1995; Vol. 2.

(2) Sousa, C.; Cebolla, A.; de Lorenzo, V. *Nature Biotechnol.* **1996**, *14*, 1017.

(3) (a) Pearson, R. G. *J. Am. Chem. Soc.* **1963**, *85*, 3533. (b) Parr, R. G.; Pearson, R. G. *J. Am. Chem. Soc.* **1983**, *105*, 7512.

(4) (a) Sigel, H.; McCormick, D. B. *Acc. Chem. Res.* **1970**, *3*, 201. (b) Martin, R. B. *J. Chem. Educ.* **1987**, *64*, 402.

(5) Glusker, J. P. *Adv. Protein Chem.* **1991**, *42*, 1.

(6) Rulíšek L., Vondrášek, J. *J. Inorg. Biochem.* **1998**, *71*, 115.

(7) (a) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 4759. (b) Comba, P.; Hambley, T. W.; Strohle, M. *Helv. Chim. Acta* **1995**, *78*, 2042.

(8) Comba, P. *Coord. Chem. Rev.* **1999**, *185−186*, 81.

(9) Ziegler, T. *Chem. Rev.* **1991**, *91*, 651. Ziegler, T. *Can. J. Chem.* **1995**, *73*, 743.

(10) Veillard, A. *Chem. Rev.* **1991**, *91*, 743.

(11) Cory, M. G.; Zerner, M. C. *Chem. Rev.* **1991**, *91*, 813

(12) Deeth, R. J. *Coord. Chem. (Series: Structure and Bonding)* **1995**, *82*, 1.

(13) Chermette, H. *Coord. Chem. Rev.* **1998**, *178−180*, 699.

10.1021/ja001265g CCC: $19.00 © 2000 American Chemical Society
Published on Web 10/10/2000

The work in this field can be approximately divided into two categories:

(a) Small molecule(s) interacting with bare TM ions (the results can be directly compared to the accurate data from gas-phase experiments) or TM ions with nonempty coordination sphere (more closely modeling the solution chemistry). To mention some representative examples, interactions of various TM ions with water molecule(s),[16] nitric oxide,[17] thioethers,[18] ammonia,[19] hydrogen(s),[20] oxygen(s),[21] or halogens[22] have been studied. In most cases, at least satisfactory agreement between the theory and experiment has been achieved. It is especially true for the complexes with ionic character (with the smaller amount of electron transfer between metal ion and ligands) whose properties can be calculated to a higher level of accuracy than those of organometallic compounds. Closely related to our work is the study of de Bruin et al.,[23] and of Hoyau and Ohanessian.[24] In the former one, the authors computed the geometry and electronic structure of bis-(glycinato)$-Cu^{II}\cdot 2H_2O$ complex, using DFT/B3LYP method. They reproduced the higher stability of *trans* isomer, in agreement with the experimental findings. In the latter one, the absolute gas-phase affinities of glycine, serine, and alanine with closed-shell Cu^+ ion have been studied theoretically. The authors concentrated on the analysis of various conformers and the accurate evaluation of their energy differences.

(b) The particular model system derived from the metal-containing biomolecules, mimicking, for example, the metal-binding site of metalloprotein,[25] the part of DNA molecule interacting with metal ion,[26] or porphyrine systems.[27] The quantum chemical studies of such systems, which constitute the core of the quantum bioinorganic chemistry, yield the quantities, which may be inaccessible or complementary to experiments, such as reaction energies, transition-state barriers, or electronic properties of the species of interest. The recent results from this field are encouraging.

In the recently published work,[15] we endeavored to establish the accurate and efficient computational scheme for the type of calculations that we carried out in this work, that is the calculations of molecular energies and structures of $[MeX_n]^{c+}$ complexes, where X is a neutral or charged functional group and the whole complex possesses ionic character. It has been shown that DFT calculations (with B3LYP functional) yield both the accurate reaction energies for the substitution reaction on the metal ion and accurate molecular structures, both for nonsubstituted $[Me(H_2O)_n]^{2+}$ and monosubstituted $[Me(H_2O)_{n-1}X]^{2+}$ complexes, provided the basis set of at least triple-ζ quality with polarization and diffuse functions on all atoms is used. However, we have presumed that the ground states of TM complexes are satisfactorily described by a single electron configuration (Slater determinant).

In the subsequent study,[28] we have analyzed complexes that do not belong to the above category, that is the TM systems with the degenerate or quasi-degenerate ground electronic state and each of its components possessing multireference character. It has been shown, on the example of octahedral Co^{2+} complexes, that a breaking of the ligand field symmetry caused by the substitution of water molecule in $[Me(H_2O)_n]^{2+}$ complex split the originally degenerate molecular states. As a consequence, the coefficient of the leading configuration in the CAS SCF wave function approaches unity, and the usage of single-reference methods is justified. Thus, the computational scheme described above can be used for the calculations of this class of systems as well. Besides, several potential energy surfaces (PES) near the equilibrium geometry of $[Co(H_2O)_5X]^{2+}$ complexes have been calculated, which enables the estimation of force constants of metal$-$ligand bonds.

The aim of this work is to provide a thorough and comparative study of the interactions of selected TM ions with the functional groups representing amino acid side chains, using the above-mentioned computational scheme. In a broader aspect, we believe that methods of theoretical chemistry may ultimately lead to a better understanding of the processes governing the mechanisms of a recognition and coordination of metal ions in biological macromolecules and attempt to add a piece of knowledge in this effort.

The study consists of the following steps:

(i) The calculation of affinities of the studied metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) for $(H_2O)_n$ site, pre-organized in a certain coordination geometry (linear, tetrahedral,

(14) (a) Niu, S.; Hall, M. B. *Chem. Rev.* **2000**, *100*, 353. (b) Loew, G. H.; Harris, D. L. *Chem. Rev.* **2000**, *100*, 407. (c) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421. (d) Frenking, G.; Fröhlich, N. *Chem. Rev.* **2000**, *100*, 717. (e) Hush, N. S.; Reimers, J. R. *Chem. Rev.* **2000**, *100*, 775.

(15) Rulíšek, L.; Havlas, Z. *J. Phys. Chem. A* **1999**, *103*, 1634.

(16) (a) Åkesson, R.; Pettersson, L. G. M.; Sandström, M.; Siegbahn, P. E. M.; Wahlgren, U. *J. Phys. Chem.* **1992**, *96*, 10773. (b) Åkesson, R.; Pettersson, L. G. M.; Sandström, M.; Siegbahn, P. E. M.; Wahlgren, U. *J. Phys. Chem.* **1993**, *97*, 3765. (c) Åkesson, R.; Pettersson, L. G. M.; Sandström, M.; Wahlgren, U. *J. Am. Chem. Soc.* **1994**, *116*, 8691. (d) Åkesson, R.; Pettersson, L. G. M.; Sandström, M.; Wahlgren, U. *J. Am. Chem. Soc.* **1994**, *116*, 8705 (e) Pavlov, M.; Siegbahn, P. E. M.; Sandström, M. *J. Phys. Chem. A* **1998**, *102*, 219. (f) Magnusson, E.; Moriarty, N. W. *J. Comput. Chem.* **1993**, *14*, 961. (g) Irigoras, A.; Ugalde, J. M.; Lopez, X.; Sarasola, C. *Can. J. Chem.* **1996**, *74*, 1824. (h) Irigoras, A.; Fowler, J. E.; Ugalde, J. M. *J. Am. Chem. Soc.* **1999**, *121*, 574. (i) Irigoras, A.; Fowler, J. E.; Ugalde, J. M. *J. Am. Chem. Soc.* **1999**, *121*, 8549. (j) Johnson, D. A.; Nelson, P. G. *Inorg. Chem.* **1995**, *34*, 5666. (k) Papai, I. *J. Chem. Phys.* **1995**, *103*, 1860. (l) Adamo, C.; Lelj, F. *J. Mol. Struct. (THEOCHEM)* **1997**, *389*, 83. (m) Fournier, R. *Theor. Chim. Acta* **1995**, *91*, 129. (n) Gilson, H. S. R.; Krauss, M. *J. Phys. Chem. A* **1998**, *102*, 6525. (o) Waizumi, K.; Ohtaki, H.; Masuda, H.; Fukushima, N.; Watanabe, Y. *Chem. Lett.* **1992**, 1489. (p) Li, J.; Fisher, C. L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35*, 4694.

(17) (a) Niu, S.; Hall, M. B. *J. Phys. Chem A* **1997**, *101*, 1360. (b) Thomas, J. L. C.; Bauschlicher, C. W., Jr.; Hall, M. B. *J. Phys. Chem. A* **1997**, *101*, 8530.

(18) Jacobsen, H.; Kraatz, H.-B.; Ziegler, T.; Boorman, P. M. *J. Am. Chem. Soc.* **1992**, *114*, 7851.

(19) (a) Langhoff, S. R.; Bauschlicher, C. W., Jr.; Partridge, H.; Sodupe, M. *J. Phys. Chem.* **1991**, *95*, 10677. (b) Newton, M. D. *J. Phys. Chem.* **1991**, *95*, 30. (c) Fournier, R. *J. Chem. Phys.* **1995**, *102*, 5396. (d) Taketsugu, T.; Gordon, M. S. *J. Chem. Phys.* **1997**, *106*, 8504.

(20) (a) Ziegler, T.; Li. J. *Can. J. Chem.* **1994**, *72*, 783. (b) Barone, V.; Adamo, C. *Int. J. Quantum Chem.* **1997**, *61*, 443.

(21) (a) Broclawik, E.; Salahub, D. R. *J. Mol. Catal.* **1993**, *82*, 117. (b) Deeth, R. J. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 3745.

(22) Bray, M. R.; Deeth, R. J.; Paget, V. J.; Sheen, P. D. *Int. J. Quantum Chem.* **1996**, *61*, 85.

(23) de Bruin, T. J. M.; Marcelis, A. T. M.; Zuilhof, H.; Sudholter, E. J. R. *Phys. Chem. Chem. Phys.* **1999**, *1*, 4157.

(24) Hoyau, S.; Ohanessian, G. *J. Am. Chem. Soc.* **1997**, *119*, 2016.

(25) (a) Garmer, D. R.; Krauss, M. *J. Am. Chem. Soc.* **1992**, *114*, 6487. (b) Ryde, U.; Hemmingsen, L. *J. Biol. Inorg. Chem.* **1997**, *2*, 567. (c) Pierloot, K.; De Kerpel, J. O. A.; Ryde, U.; Olsson, M. H. M.; Roos, B. O. *J. Am. Chem. Soc.* **1998**, *120*, 13156. (d) De Kerpel, J. O. A.; Pierloot, K.; Ryde, U. *J. Phys. Chem. B* **1999**, *103*, 8375. (e) Siegbahn, P. E. M.; Crabtree, R. H. *J. Am. Chem. Soc.* **1999**, *121*, 117. (f) Siegbahn, P. E. M.; Eriksson, L.; Himo, F.; Pavlov, M. *J. Phys. Chem. B* **1998**, *102*, 10622. (g) Eriksson, L.; Himo, F.; Siegbahn, P. E. M.; Babcock, G. T. *J. Phys. Chem. A* **1997**, *101*, 9496. (h) Li, J.; Fisher, C. L.; Konecny, R.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1999**, *38*, 929. (i) Konecny, R.; Li, J.; Fisher, C. L.; Dillet, V.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1999**, *38*, 940. (j) Li, J.; Nelson, M. R.; Peng, C. Y.; Bashford, D.; Noodleman, L. *J. Phys. Chem. A* **1998**, *102*, 6311.

(26) (a) Šponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951. (b) Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1999**, *103*, 2528.

(27) (a) Yamamoto, S.; Kashiwagi, H. *Chem. Phys. Lett.* **1993**, *205*, 306. (b) Zwaans, R.; van Lenthe, J. H.; den Boer, D. H. W. *J. Mol. Struct. (THEOCHEM)* **1996**, *367*, 15.

(28) Rulíšek, L.; Havlas, Z. *J. Chem. Phys.* **2000**, *112*, 149.

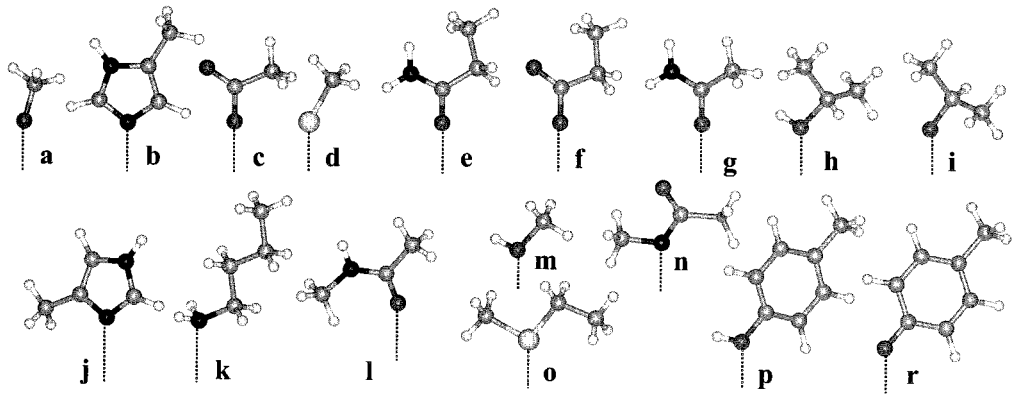

Figure 1. The functional groups used as the models of amino acid side chains (**a**) CH_3O^- (Ser^-), (**b**) 5-methylimidazole (His, binding through Nϵ), (**c**) CH_3COO^- (Asp^-), (**d**) CH_3S^- (Cys^-), (**e**) $CH_3CH_2CONH_2$ (Gln), (**f**) $CH_3CH_2COO^-$ (Glu^-), (**g**) CH_3CONH_2 (Asn), (**h**) $(CH_3)_2CHOH$ (Thr), (**i**) $(CH_3)_2CHO^-$ (Thr^-), (**j**) 4-methylimidazole (His, binding through Nδ), (**k**) $CH_3(CH_2)_3NH_2$ (Lys), (**l**) $CH_3CONHCH_3$, (peptide bond oxygen), (**m**) CH_3OH (Ser), (**n**) $CH_3CONCH_3^-$ (deprotonized peptide bond nitrogen), (**o**) $CH_3CH_2SCH_3$ (Met), (**p**) $CH_3C_6H_4OH$ (Tyr), (**r**) $CH_3C_6H_4O^-$ (Tyr^-).

square planar, octahedral) with the aim to reproduce the experimentally found preferences of the studied TM ions for the specific environment.

(ii)The calculation of equilibrium geometries and interaction energies of model functional groups representing amino acid side chains, peptide bond oxygen, and deprotonated peptide bond nitrogen with the studied TM ions. For each of them, all four coordination geometries will be considered, and the nonbond interactions between ligands accounted for. The results will be analyzed with respect to the metal ion selectivity of the studied functional groups. It should be mentioned that coordination compounds of the studied TM ions may assume other coordination geometries as well, but these four have been found to be most abundant in the metal-binding sites of metalloproteins.[6]

(iii)The analysis of the effect of an amino acid side chain elongation, revealing what properties calculated for the model systems are conserved in the metal binding sites of metalloproteins.

All of the results will be compared to qualitative and semiquantitative experimental findings, and the above-discussed basic concepts of coordination chemistry, because, to the best of our knowledge, the direct experimental data for the studied type of interactions are not available.

Due to a large amount of the calculations that have to be performed, we have chosen the simplest models for the side chains of amino acid residues. Each side chain is terminated by hydrogen atom, replacing the C_α atom of the peptide or protein backbone. We have taken into account the metal-coordinating amino acid residues, that is, side chains containing oxygen, nitrogen, or sulfur as the donor atom. As it has been found by the analysis of experimental structures,[6] almost all coordination of TM ions in metalloproteins occurs via side chains; however, in some rare cases, the peptide oxygen or deprotonated peptide nitrogen has been observed to participate in the coordination bond. For the sake of completeness, those two have been included in the study as well.

The model functional groups are depicted at Figure 1.

They are divided into two groups: aprotic and protic. To the former one belong side chains of *Asn*, *Gln*, *Met*, and *peptide bond oxygen*, to the latter one the rest of them: *Ser*, *Asp*, *Cys*, *His*, *Thr*, *Glu*, *Lys*, *Tyr*, *peptide bond nitrogen*. The very important question is whether these functional groups are protonated or deprotonated when coordinated to the TM ion of interest, or, in physicochemical terms, what are their pK_a values. The determination of protonation states of amino acid side chains in the metal-binding sites of metalloproteins is a very difficult task, and even the thorough discussion of the problem is beyond the scope of this work. An attempt to evaluate them by means of theoretical chemistry will be subject of a subsequent paper. In the meantime, we refer the reader to the work of Noodleman et al.[25h−j] for latest developments in this area.

In this work, we assumed (depending on the coordination geometry) the following protonation states of protic functional groups:

•*octahedral: Cys* deprotonated (CH_3S^-); *Ser*, *Thr* protonated ($-OH$); *Tyr* both protonated and deprotonated ($CH_3C_6H_4OH$, $CH_3C_6H_4O^-$),

•*tetrahedral, square planar: Cys*, *Tyr* deprotonated; *Ser*, *Thr* both protonated and deprotonated,

•*linear: Cys*, *Ser*, *Thr*, *Tyr* deprotonated.

As for the remaining, *Asp* and *Glu* have been calculated in the deprotonated state ($-COO^-$) throughout, because they have p$K_a \approx 4$ in noncoordinated form, *Lys* has been deprotonated ($CH_3(CH_2)_3NH_2$) since it cannot bind as ammonium cation, and *His* has been considered in its neutral form (i.e., with only one nitrogen protonated). The coordination properties of *peptide bond nitrogen* have been reviewed by Martin and Sigel.[29] It can bind to TM ions only in deprotonated form. From the studied TM ions, only copper(II) is known to form tetracoordinate complexes with *peptide bond nitrogen* as the ligand (none hexacoordinate complexes are known). Nevertheless, it has been taken into account as the ligand for all TM ions in all coordination geometries except octahedral.

The above considerations are based on the perusal of Martell tables of stability constants,[30] known experimental pK_a constants[31] of aqua complexes, and known pK_a constants of the free functional groups. In all cases, they refer to pH = 7. As can be seen above, in the case of an uncertainty (i.e., estimated pK_a constants of the functional group bound to metal ion is about 7), we computed both protonated and deprotonated forms. Six representatives of the studied systems are depicted at Figure 2.

At the end, a note should be added about the choice of six TM ions studied here. They were selected because of two reasons. First, their metalloprotein complexes belong to the most abundant. Second, these metals are the major pollutants of the environment. Cadmium and mercury are highly toxic; cobalt, nickel, and copper are undesirable in the environment as well.

(29) Sigel, H.; Martin, R. B. *Chem. Rev.* **1982**, *82*, 385.

(30) Smith, R. M.; Martell, A. E. *Critical Stability Constants*; Plenum Press: New York, 1974−1989; Vols. 1−6.

(31) Baes, C. F., Jr.; Mesmer, R. E. *The Hydrolysis of Cations*; John Wiley & Sons: New York, 1976.

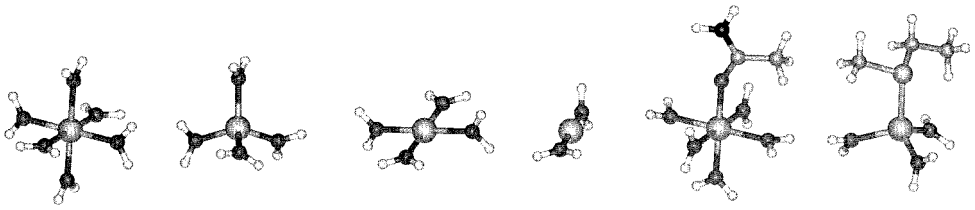

Figure 2. Nonsubsituted $[Me(H_2O)_n]^{2+}$ complexes in four different coordination geometries: (**a**) octahedral ($n = 6$), (**b**) tetrahedral ($n = 4$), (**c**) square planar ($n = 4$), (**d**) linear ($n = 2$) and two examples of the monosubstituted complexes: (**e**) $[Zn(H_2O)_5(CH_3CONH_2)]^{2+}$, (**f**) $[Zn(H_2O)_3(CH_3CH_2SCH_3)]^{2+}$.

The same holds true for zinc (biogenic element), which becomes also toxic in higher concentrations. The development of methods for the selective binding of these metals may ultimately lead to their successful removal from the environment. Presuming that this problem can be approached also theoretically, the molecular design of such specific sites should necessarily start with the type of information we are trying to provide here.

II. Computational Details

All of the calculations were performed with the Gaussian 98 program suite.[32]

Almost all the calculations were performed in the framework of density functional theory (DFT). The three-parameter functional developed by Becke,[33] which combines the Becke's gradient-corrected exchange functional and the Lee−Yang−Parr and Vosko−Wilk−Nusair correlation functionals[34] with part of the exact Hartree−Fock exchange energy, has been employed (denoted as B3LYP).

To ascertain the multiplicity of the ground states of several of the studied systems, the complete active space self-consistent field (CAS SCF) method[35] (with five MOs, composed mainly from d orbitals of metal ion and corresponding number of electrons, in active space) has been used.

Three basis sets have been used throughout the calculations, denoted as BS1, BS2, and BS3. BS1 has been 6-31G basis set stored internally in Gaussian 98 both for the first and second row atoms and the first row transition metals. It was further augmented by diffuse functions: (s,2p,d) set for TMs; sp functions for other heavy elements and the single set of polarization functions: f for TMs, and d for other heavy elements.

BS2 consisted of the triple-ζ (TZ) basis set of Watchers and Hay[36] for the first-row transition metals (Co, Ni, Cu, Zn), and standard 6-311G for other elements (H, C, N, O, S).[37] It was augmented by diffuse and polarization functions in the same way as BS1.

BS3 has been derived from BS2, by addition of diffuse *s* functions for hydrogens, and different sets of polarization functions: 2fg for TMs, 2df for other heavy atoms, 2pd for hydrogens. The exponents of all diffuse and polarization functions were used as implemented in Gaussian 98, and the described basis sets have been approached via 6-31+G(d) (BS1), 6-311+G(d) (BS2), and 6-311++G(2df,2pd) keywords.

For Cd^{2+} and Hg^{2+}, effective core potentials (ECP) of Stevens and co-workers[38] have been used (denoted SBKJ). To achieve the consistency with the above-described BS2 and BS3 basis sets used for the first row TMs, the original valence basis set was further augmented with the following uncontracted GTO basis functions: diffuse d functions (α_d(Cd) = 0.075, α_d(Hg) = 0.040); and f (α_f(Cd) = 0.775, α_f(Hg) = 0.690) and 2fg (α_{1f} (Cd) = 2.0, α_{2f} (Cd) = 0.3, α_g (Cd) = 0.775, α_{1f} (Hg) = 1.35, α_{2f} (Hg) = 0.35, α_g (Hg) = 0.69) sets of polarization functions, corresponding to BS2 and BS3, respectively.

The computational scheme consisted of several steps:

First, the optimization of molecular geometries of the selected systems has been carried out at the B3LYP/BS2 level with the angles at the metal centers fixed at the values corresponding to the given coordination geometry and all other internal coordinates optimized. The selected systems were: $[Ni(H_2O)_5X]^{2+}$ in octahedral, $[Zn(H_2O)_3X]^{2+}$, $[Cd(H_2O)_3X]^{2+}$ in tetrahedral, $[Cu(H_2O)_3X]^{2+}$ in square planar, and $[Hg(H_2O)X]^{2+}$ in linear coordination geometry. According to the previous work,[6] these geometries could be considered as the preferred ones by the studied TM ions.

Second, all other systems were assumed to adopt the geometry of the optimized complexes mentioned above (note that we have at least one optimized system for each of the studied coordination geometries) and only *n* metal−ligand distances ($n = 2, 4, 6$) optimized at the B3LYP/BS1 level. The only exceptions were octahedral Co^{2+} and Cu^{2+} complexes, which are in principle Jahn−Teller unstable, as a consequence of the degenerate ground state in ideal O_h ligand field symmetry. Therefore, they have been assumed to adopt the same geometries as the corresponding $[Ni(H_2O)_5X]^{2+}$ systems, with all six metal−ligand distances increased by the experimentally and theoretically found differences between the ionic radii of Co^{2+}, Cu^{2+}, and Ni^{2+}, in octahedral coordination which are +0.04 Å (Co^{2+}) and +0.03 Å (Cu^{2+}).[39] This is a plausible approximation that has been tested in the previous work on model $[Co(H_2O)_5X]^{2+}$ complexes[28] and that is a posteriori justified by the calculated data. The simplification of the computational scheme (descending to BS1 basis set and the optimization of a limited number of internal coordinates) reduced CPU time approximately 10 times with virtually no loss in accuracy, which will be demonstrated below.

Third, the single-point energy calculations of all the studied structures have been carried out at the B3LYP/BS3 level to obtain the final molecular energy of $[Me(H_2O)_{n-1}X]^{2+}$ complex.

Fourth, the metal ion at the optimized geometry was replaced by the corresponding ghost atom Bq and single-point energy calculated for $Bq(H_2O)_{n-1}X$ system.

Now, three important points, related to the optimization of molecular geometry should be discussed.

(i) The use of optimized molecular geometry of $[Me^{I}(H_2O)_{n-1}X]^{2+}$ complex for $[Me^{II}(H_2O)_{n-1}X]^{2+}$ and optimization of only *n* metal−

(32) 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, K. N.; Strain, M. C.; 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.6; Gaussian, Inc.: Pittsburgh, PA, 1998.

(33) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(34) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(35) (a) Dalgaard, E.; Jørgensen, P. *J. Chem. Phys.* **1978**, *69*, 3833. (b) Hegarty, D.; Robb, M. A. Mol. Phys. **1979**, *38*, 1795. (c) Siegbahn, P. E. M.; Almlöf, J.; Heiberg, A.; Roos, B. O. *J. Chem. Phys.* **1981**, *74*, 2384.

(36) (a) Watchers, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033. (b) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377. (c) Raghavachari, K.; Trucks, G. W. *J. Chem. Phys.* **1989**, *91*, 1062.

(37) (a) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639. (b) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. **1980**, *72*, 650.

(38) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612.

(39) Marcus, Y. *Chem. Rev.* **1988**, *88*, 1475.

ligand distances. Although they undoubtedly represent the most important degrees of freedom in the complex, we still considered it important to validate this approximation. We made advantage of the fact that Zn^{2+} and Cd^{2+} ions have been fully optimized in tetrahedral coordination geometry at the B3LYP/BS2 level. Therefore, we substituted Zn for Cd in optimized $[Zn(Asn)(H_2O)_3]^{2+}$, $[Zn(Ser)(H_2O)_3]^{2+}$ complexes, and Cd for Zn in $[Cd(H_2O)_4]^{2+}$, $[Cd(Met)(H_2O)_3]^{2+}$ and optimized only four metal−ligand distances (at the B3LYP/BS2 level) in each of them. Then the single-point energy calculation has been carried out at the B3LYP/BS3 level. In all cases, the computed values of molecular energies differed by less than 0.1 kcal·mol^{-1} from the ones obtained after the full optimization of these complexes. Since the difference in molecular geometries of the complexes of TM ions belonging to the different rows of periodic table is presumably greater than the ones belonging to the same row, we consider the above calculations as the sufficient validation of the adopted approximation.

(ii) The use of BS1 basis set (instead of BS2) for the optimization of *n* metal−ligand distances. To justify this approximation, the metal−ligand distances of four model complexes: $[Cu(HCHO)(H_2O)_3]^{2+}$, $[Cu(NH_3)(H_2O)_3]^{2+}$ in square planar and $[Zn(H_2O)_4]^{2+}$, $[Zn(Met)(H_2O)_3]^{2+}$ in tetrahedral coordination geometry have been optimized both at the B3LYP/BS1 and B3LYP/BS2 levels. Both optimizations yielded almost equivalent metal−ligand distances (with average $|\Delta d| = 0.003$ Å, and maximum $|\Delta d| = 0.012$ Å) and almost negligible differences in molecular energies calculated at the BS1 and BS2 optimized geometries at the B3LYP/BS3 level (less than 0.1 kcal·mol^{-1}).

(iii) The use of restricted optimization procedure with the angles at the metal center fixed exactly at the values corresponding to the particular coordination environment. This restriction belongs to the chosen theoretical model rather than approximation (and virtually does not save a computational effort). Still, we have compared the results of the restricted optimization with the fully optimized octahedral nickel(II) complexes: $[Ni(H_2O)_5(CH_3COOH)]^{2+}$, $[Ni(H_2O)_5(His)]^{2+}$, $[Ni(H_2O)_5(CH_3NH_2)]^{2+}$. The full optimization decreased the computed values of interaction energy by almost the same amount for all three complexes: 2.2 kcal·mol^{-1}, 2.1 kcal·mol^{-1}, 2.6 kcal·mol^{-1}, respectively, and it led to the average deviation of 2.2° (with maximum of 9.1°) in the values of L−M−L angles from their ideal values.

With respect to the principal aim of the work, which is the evaluation of metal ion selectivity, we presume that the described approximations are well below the error of the theoretical method itself. Moreover, systematic errors are further corrected by the fact that the calculations have been carried out for the series of molecules containing chemically similar ligands.

Throughout the paper, the interaction energy of a single amino acid residue X with the metal ion Me in the given coordination geometry is defined as:

$$E_{int}(Me,X) = E([Me(H_2O)_{n-1}X]^{2+}) - E(Bq(H_2O)_{n-1}X) - (E([Me(H_2O)_n]^{2+}) - E(Bq(H_2O)_n)) \quad (1)$$

where $n = 2$ (linear coordination geometry), 4 (square planar, tetrahedral), and 6 (octahedral). According to this formula, the computed interaction energy has been corrected for the nonbond interactions between neighboring ligands and for a part of basis set superposition error (BSSE).

III. Results and Discussion

Ground-State Multiplicities of the Studied Transition Metal Ions. Three of the studied TM ions − Zn^{2+}, Cd^{2+}, Hg^{2+} − are d^{10} ions, and therefore their complexes are closed-shell systems with the singlet ground state. Cu^{2+} is d^9 ion, and its complexes have the doublet ground state. A more complicated situation is with Co^{2+} and Ni^{2+} ions, which may exist in their complexes both in high-spin and low-spin states (quartet and doublet for Co^{2+}, triplet and singlet for Ni^{2+}). The ordering of spin states depends on the strength and symmetry of the ligand field (LF).[40] From this point of view, the ligands studied in this work are of the same chemical character as water molecules. Besides, they contribute only by 1/2, 1/4, and 1/6 to the total ligand field acting on TM ion in $[Me(H_2O)_{n-1}X]^{2+}$ complexes. It implies that the energy differences between the lowest electronic states of different multiplicities should be approximately the same as in $[Me(H_2O)_n]^{2+}$ species. The hexahydrates of Co^{2+}, and Ni^{2+} are well-established high-spin complexes, and the same holds true for their tetrahedral $[Me(H_2O)_4]^{2+}$ complexes.[40] For example, the lowest doublet states of octahedral $[Co(H_2O)_6]^{2+}$ and tetrahedral $[Co(H_2O)_4]^{2+}$ are by 16 000 and 20 000 cm^{-1} higher in energy than the quartet ground states, respectively.[16n] These values correspond to the energy difference between the ^4F and ^2G atomic states (ΔE_{LS-HS}) of Co^{2+} ion, which is experimentally found to be 16 543 cm^{-1}.[41]

We presume, that the LF induced by two "H_2O-like" ligands in linear coordination geometry is too weak to shift atomic ΔE_{LS-HS} significantly. To ascertain this point, we carried out CAS SCF calculations (using BS2 basis set) for model $[Me(H_2O)_2]^{2+}$, $[Me(OH)_2]$, $[Me(H_2O)(NH_3)]^{2+}$, and $[Me(H_2O)(CH_3SH)]^{2+}$ systems (Me = Co, Ni) in both high-spin and low-spin states. In all cases, the high spin states have been unambiguously identified as the ground electronic states with the lowest doublets lying by 16 000−17 200 cm^{-1} (for Co^{2+} species) and lowest singlets (for Ni^{2+}) by 11 200−13 200 cm^{-1} higher in energy.

Even more difficult situation is with square planar Co^{2+} and Ni^{2+} complexes. Applying the quantitative crystal field (CF) theory, it is expected that energy of $d_{x^2-y^2}$ orbital substantially rises as a consequence of strong repulsion between electrons of metal and ligands, and low-spin state becomes the ground state of a molecule. The effect is most profound for Ni^{2+} (further stabilization of a closed-shell singlet over an open-shell state), and therefore, it can be found in the inorganic chemistry textbooks[40] that all square planar complexes of Ni^{2+} are low-spin. However, we must keep in mind that this evidence should be rigorously translated as "all the existing square planar complexes of Ni^{2+} are low-spin" and that the studied complexes with the geometry fixed at the square planar arrangement are only theoretical models and may not be even global minima on the PES of the complex (in many cases, tetrahedral coordination may be preferred). The same reasoning could be repeated for Co^{2+} in square planar coordination, and therefore, ΔE_{LS-HS} has been calculated for model compounds of $[Me(H_2O)_4]^{2+}$, $[Me(H_2O)_2(OH)_2]$, $[Me(H_2O)_3(NH_3)]^{2+}$, and $[Me(H_2O)_3(CH_3SH)]^{2+}$ (Me = Co, Ni) in square planar coordination geometry, at the CAS SCF/BS2 level. This method is considered to be more appropriate for this type of calculation, since hybrid DFT methods may yield incorrect estimates of energy gaps between states with different spin multiplicity.[42]

Surprisingly, we have found that all four model complexes of Co^{2+} and Ni^{2+} have high-spin ground states and the lowest doublets (for Co^{2+}) are by 11 900−13 000 cm^{-1} and the lowest singlets (for Ni^{2+}) by 9500−12 100 cm^{-1} higher in energy. Despite a considerable shift in ΔE_{LS-HS} (4000−7000 cm^{-1}) in comparison with the bare Me^{2+} ions, the ligand field of the studied functional groups is too weak to cause the low-spin state to become the ground state of the molecule. The leading electronic configurations in CAS SCF wave functions (with coefficients of 0.94−0.99) could be approximately described

(40) Cotton, F. A.; Wilkinson, G. *Advanced Inorganic Chemistry*, 4th ed.; Wiley: NewYork, 1980.
(41) NBS atomic energy tables.
(42) Yanagisawa, S.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2000**, *112*, 545.

Table 1. Affinities E_{Gm}(Me) of Studied TM Ions for a Given Coordination Geometry, as Defined by Eq 2 (All Values in kcal·mol^{-1})

coordination geometry	Co^{2+}	Ni^{2+}	Cu^{2+}	Zn^{2+}	Cd^{2+}	Hg^{2+}
octahedral	−358.9	−414.1	−373.5	−364.0	−305.0	−303.1
tetrahedral	−293.9	−332.6	−305.0	−301.6	−244.1	−246.6
square planar	−286.0	−336.2	−320.8	−292.4	−239.9	−245.2
linear	−170.8	−223.7	−197.2	−194.6	−154.4	−172.2

as $(d_{xz})^2(d_{yz})^2(d_{z^2})^2(d_{xy})^1(d_{x^2-y^2})^0$ for the lowest doublet state of $[CoX_4]^{2+}$ square planar complexes, $(d_{xz})^2(d_{yz})^2(d_{z^2})^1(d_{xy})^1(d_{x^2-y^2})^1$ for quartet state of $[CoX_4]^{2+}$, $(d_{xz})^2(d_{yz})^2(d_{z^2})^2(d_{xy})^2(d_{x^2-y^2})^0$ for singlet state of $[NiX_4]^{2+}$, and $(d_{xz})^2(d_{yz})^2(d_{z^2})^2(d_{xy})^1(d_{x^2-y^2})^1$ for triplet state of $[NiX_4]^{2+}$.

Therefore, all the complexes of Co^{2+} and Ni^{2+} ions calculated in this work have been assumed to have high-spin ground states and we have endeavored to bring enough evidence for this fact in the previous paragraphs.

The Preferred Coordination Geometries: A Theoretical Prediction. We define the affinity of a TM ion for a coordination geometry Gm by the equation:

$$E_{Gm}(Me) = E([Me(H_2O)_n]^{2+}) - E(Bq_{Me}(H_2O)_n) - E(Me^{2+}(Bq_{H_2O})_n) \quad (2)$$

where Gm = OH, TH, SQ, Lin (octahedral, tetrahedral, square planar, linear); $n = 2, 4, 6$ (coordination number); Bq_Y denotes ghost atom(s) with basis functions left from atom (molecule) Y. Equation 2 represents the BSSE-corrected complexation (interaction) energy of Me^{2+} with $(H_2O)_n$ cluster. For the definition of E_{Gm}(Me), the water molecule has been chosen as the appropriate reference ligand because of several reasons: (i) it is possibly the simplest neutral ligand forming complexes with the ionic character; (ii) it models solvated TM ion; (iii) the interaction energy of a particular amino acid side chain is defined as the energy of substitution of one water molecule in per-hydrated complex (with the minor corrections for nonbond interactions of ligands, see eq 1). The simple addition of E_{Gm}(Me) and E_{int}(Me,X) then yields the affinity of metal ion Me for $(H_2O)_{n-1}$(X) cluster and opens the way for a different exploitation of the presented results, based upon the absolute affinities for the metal-binding sites rather than relative ones (which is the main subject of this work).

The calculated values of E_{Gm}(Me) are summarized in Table 1.

To demonstrate more clearly their significance, we define the relative affinity, E'_{Gm}(Me), as:

$$E'_{Gm}(Me) = E_{Gm}(Me) - {}^1/_4(E_{OH}(Me) + E_{TH}(Me) + E_{SQ}(Me) + E_{Lin}(Me)) \quad (3)$$

By eq 3, E_{Gm}(Me) values have been shifted by addition of the constant dependent on the TM ion only, such as the average of relative affinities E'_{Gm}(Me) over four considered coordination geometries is zero and the gas-phase stability of bare TM ion—the value without a straightforward chemical significance—separated from E_{Gm}(Me). We consider the relative affinities to be more illustrative.

The calculated values of E'_{Gm}(Me) for each metal ion are plotted in Figure 3.

Despite the simplicity of the model which (besides using water molecules as the reference ligands) takes into account only the first solvation layer, that is, ligands directly bound to the metal ion, the results are in good agreement with the empirical evidence. According to the values of E'_{Gm}(Me), octahedral coordination geometry is favorable for cobalt(II) and nickel(II), tetrahedral for cobalt(II) and zinc(II), square planar for copper(II), and linear for soft metal ions—cadmium(II) and mercury(II). Similar results have been obtained by the statistical analysis of experimental structures deposited in CSD and PDB.[6] A good correlation between theory and experimental evidence obtained with this simple model proves an important and encouraging fact that many seemingly complex properties of TM systems are determined by the character and arrangement of the first coordination shell.

On the other hand, the variety of coordination geometries known for each of the studied TM ions and the above calculated energy differences (see Figure 3), which are within the range of the conformational changes in more complex ligands prevent us from doing rigorous conclusions out of the calculated data.

The Interactions of Amino Acid Side Chains with TM Ions. The interaction energies of amino acid side chains, calculated according to eq 1 are summarized in Tables 2−5. Accompanying information, the interatomic distances between TM ion and a donor atom of amino acid side chain—d_{MeX}—and mean interatomic distances between metal and $(n - 1)$ water molecules—d_{MeO}(mean)—are listed in Table 6.

As can be seen in Tables 2−5, there are some general trends and evidences common to all coordination geometries, and they will be used as a starting point of the discussion.

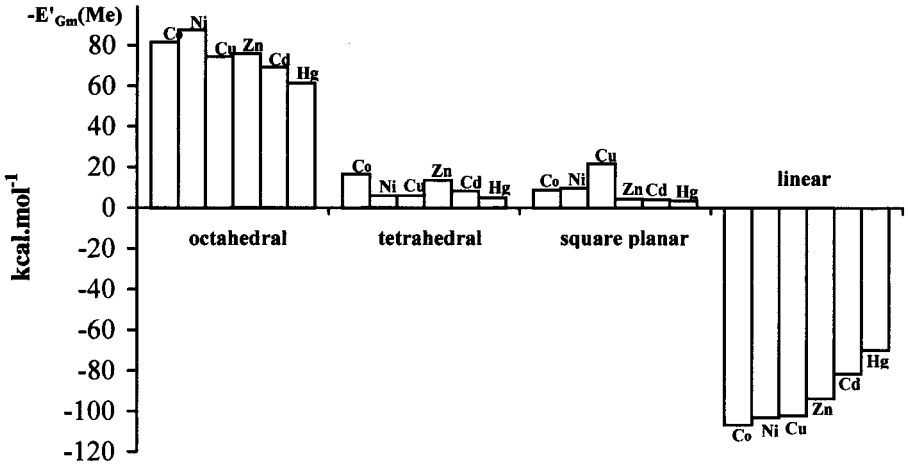

Figure 3. The calculated values of the relative affinities E'_{Gm}(Me), as defined by eq 3. Note that negative of E'_{Gm}(Me) is used for definition of *Y* axis. The more positive (or less negative) is *Y* value, the higher is the affinity of TM ion for that particular coordination geometry.

Table 2. Interaction Energies,[a] E_{int}(Me,X), of Amino Acid Side Chains (AA) with Studied TM Ions in Octahedral Coordination Geometry, Defined by Eq 1, and Relative Interaction Energies E_{rel}(Me,X)—Listed in Lower Part of Table in Italics—Defined by Eq 4 (All Values in kcal·mol^{-1})

TM\AA	Asn	Asp$^-$	Cys$^-$	Gln	Glu$^-$	His$_{N\delta}$	His$_{N\epsilon}$	Lys	Met	PeptO	Ser	Thr	Tyr	Tyr$^-$	$\bar{E}_{int}$(Me)
Co^{2+}	−25.3	−200.4	−198.2	−28.2	−200.6	−30.1	−33.0	−19.2	−0.3	−28.5	−3.4	−5.4	−8.4	−197.8	***−69.9***
Ni^{2+}	−26.5	−204.5	−202.0	−29.4	−204.8	−32.5	−36.1	−22.8	−3.1	−30.0	−4.8	−11.3	−9.9	−200.6	***−72.7***
Cu^{2+}	−28.1	−209.1	−216.8	−31.1	−209.7	−37.3	−41.6	−30.1	−11.5	−32.1	−5.5	−12.8	−14.3	−212.4	***−78.0***
Zn^{2+}	−25.6	−200.6	−197.5	−28.5	−200.8	−29.9	−33.7	−19.7	−0.3	−28.9	−3.9	−10.3	−7.6	−195.8	***−70.2***
Cd^{2+}	−24.1	−193.9	−198.6	−26.8	−194.0	−28.7	−32.4	−19.5	−4.1	−27.1	−3.8	−9.6	−6.2	−187.2	***−68.2***
Hg^{2+}	−24.2	−199.1	−221.9	−26.9	−199.3	−34.8	−38.8	−26.6	−13.5	−27.6	−4.1	−10.1	−7.0	−195.6	***−73.5***
E_{int}(X)	***−25.6***	***−201.2***	***−205.8***	***−28.4***	***−201.5***	***−32.2***	***−35.9***	***−23.0***	***−5.5***	***−29.0***	***−4.2***	***−9.9***	***−9.0***	***−198.2***	
Co^{2+}	*44.6*	*−130.5*	*−128.3*	*41.7*	*−130.7*	*39.8*	*36.9*	*50.7*	*69.6*	*41.4*	*66.5*	*64.5*	*61.5*	*−127.9*	
Ni^{2+}	*46.2*	*−131.8*	*−129.3*	*43.3*	*−132.1*	*40.2*	*36.6*	*49.9*	*69.6*	*42.7*	*67.9*	*61.4*	*62.8*	*−127.9*	
Cu^{2+}	*49.9*	*−131.1*	*−138.8*	*46.9*	*−131.7*	*40.7*	*36.4*	*47.9*	*66.5*	*45.9*	*72.5*	*65.2*	*63.7*	*−134.4*	
Zn^{2+}	*44.6*	*−130.4*	*−127.3*	*41.7*	*−130.6*	*40.3*	*36.5*	*50.5*	*69.9*	*41.3*	*66.3*	*59.9*	*62.6*	*−125.6*	
Cd^{2+}	*44.1*	*−125.7*	*−130.4*	*41.4*	*−125.8*	*39.5*	*35.8*	*48.7*	*64.1*	*41.1*	*64.4*	*58.6*	*62.0*	*−119.0*	
Hg^{2+}	*49.3*	*−125.6*	*−148.4*	*46.6*	*−125.8*	*38.7*	*34.7*	*46.9*	*60.0*	*45.9*	*69.4*	*63.4*	*66.5*	*−122.1*	

[a] The smaller (more negative) values indicate the higher affinity of a substituting functional group for metal (compared to H_2O).

Table 3. Interaction Energies,[a] E_{int}(Me,X), of Amino Acid Side Chains with Studied TM Ions in Tetrahedral Coordination Geometry, Defined by Eq 1, and Relative Interaction Energies E_{rel}(Me,X)—Listed in Lower Part of Table in Italics—Defined by Eq 4 (All Values in kcal·mol^{-1})

TM\AA	Asn	Asp$^-$	Cys$^-$	Gln	Glu$^-$	His$_{N\delta}$	His$_{N\epsilon}$	Lys	Met	PeptO	PeptN$^-$	Ser	Ser$^-$	Thr	Thr$^-$	Tyr$^-$	$\bar{E}_{int}$(Me)
Co^{2+}	−38.1	−230.3	−235.3	−41.8	−230.8	−45.3	−48.1	−30.7	−16.4	−43.1	−230.8	−9.0	−258.7	−19.3	−255.8	−231.3	***−122.8***
Ni^{2+}	−40.7	−234.9	−245.2	−44.6	−235.6	−48.2	−50.9	−34.7	−22.0	−46.1	−237.9	−9.4	−266.7	−20.0	−263.9	−238.8	***−127.5***
Cu^{2+}	−44.3	−244.0	−264.7	−48.8	−245.2	−54.2	−57.6	−45.3	−34.3	−50.3	−255.3	−10.7	−279.2	−22.6	−274.8	−258.8	***−136.9***
Zn^{2+}	−37.8	−230.5	−236.0	−41.4	−231.0	−46.2	−48.9	−31.3	−17.2	−42.6	−228.7	−8.6	−254.0	−19.1	−250.2	−227.3	***−121.9***
Cd^{2+}	−34.0	−233.0	−236.3	−37.3	−232.7	−42.9	−45.3	−29.7	−20.1	−38.4	−222.7	−7.4	−247.0	−16.6	−243.0	−218.1	***−119.0***
Hg^{2+}	−34.6	−238.1	−257.7	−37.9	−238.0	−49.4	−51.6	−37.5	−32.5	−39.5	−236.8	−8.0	−257.9	−17.9	−254.5	−228.1	***−126.2***
$\bar{E}_{int}$(X)	***−38.3***	***−235.1***	***−245.9***	***−42.0***	***−235.6***	***−47.7***	***−50.4***	***−34.9***	***−23.8***	***−43.3***	***−235.4***	***−8.9***	***−260.6***	***−19.3***	***−257.0***	***−233.7***	
Co^{2+}	*84.7*	*−107.5*	*−112.5*	*81.0*	*−108.0*	*77.5*	*74.7*	*92.1*	*106.4*	*79.7*	*−108.0*	*113.8*	*−135.9*	*103.5*	*−133.0*	*−108.5*	
Ni^{2+}	*86.8*	*−107.4*	*−117.7*	*82.9*	*−108.1*	*79.3*	*76.6*	*92.8*	*105.5*	*81.4*	*−110.4*	*118.1*	*−139.2*	*107.5*	*−136.4*	*−111.3*	
Cu^{2+}	*92.6*	*−107.1*	*−127.8*	*88.1*	*−108.3*	*82.7*	*79.3*	*91.6*	*102.6*	*86.6*	*−118.4*	*126.2*	*−142.3*	*114.3*	*−137.9*	*−121.9*	
Zn^{2+}	*84.1*	*−108.6*	*−114.1*	*80.5*	*−109.1*	*75.7*	*73.0*	*90.6*	*104.7*	*79.3*	*−106.8*	*113.3*	*−132.1*	*102.8*	*−128.3*	*−105.4*	
Cd^{2+}	*85.0*	*−114.0*	*−117.3*	*81.7*	*−113.7*	*76.1*	*73.7*	*89.3*	*98.9*	*80.6*	*−103.7*	*111.6*	*−128.0*	*102.4*	*−124.0*	*−99.1*	
Hg^{2+}	*91.6*	*−111.9*	*−131.5*	*88.3*	*−111.8*	*76.8*	*74.6*	*88.7*	*93.7*	*86.7*	*−110.6*	*118.2*	*−131.7*	*108.3*	*−128.3*	*−101.9*	

[a] The smaller (more negative) values indicate the higher affinity of a substituting functional group for metal (compared to H_2O).

Table 4. Interaction Energies,[a] E_{int}(Me,X), of Amino Acid Side Chains with Studied TM Ions in Square Planar Coordination Geometry, Defined by Eq 1, and Relative Interaction Energies E_{rel}(Me,X)—Listed in Lower Part of Table in Italics—Defined by Eq 4 (All Values in kcal·mol^{-1})

TM\AA	Asn	Asp$^-$	Cys$^-$	Gln	Glu$^-$	His$_{N\delta}$	His$_{N\epsilon}$	Lys	Met	PeptO	PeptN$^-$	Ser	Ser$^-$	Thr	Thr$^-$	Tyr$^-$	$\bar{E}_{int}$(Me)
Co^{2+}	−38.7	−233.7	−234.6	−42.6	−234.5	−44.2	−46.7	−16.9	−14.5	−44.2	−233.5	−8.9	−262.4	−19.6	−254.9	−232.7	***−122.7***
Ni^{2+}	−38.1	−234.1	−235.9	−42.0	−235.0	−46.9	−49.3	−25.2	−17.3	−43.5	−233.1	−8.7	−266.5	−19.2	−256.1	−235.9	***−124.2***
Cu^{2+}	−40.9	−239.4	−247.0	−44.8	−240.5	−51.5	−53.8	−37.6	−25.6	−46.6	−241.4	−9.7	−266.8	−21.7	−261.1	−244.8	***−129.6***
Zn^{2+}	−37.9	−231.0	−234.1	−41.6	−231.6	−46.2	−48.3	−29.9	−16.0	−43.1	−228.7	−8.5	−256.1	−19.6	−249.8	−224.4	***−121.7***
Cd^{2+}	−34.2	−219.6	−234.3	−37.7	−220.1	−43.1	−45.0	−29.3	−19.3	−39.0	−222.8	−7.3	−247.0	−17.3	−241.5	−215.3	***−117.1***
Hg^{2+}	−36.4	−225.5	−260.1	−40.1	−226.4	−52.1	−53.7	−40.9	−33.9	−41.9	−243.7	−7.7	−261.9	−18.7	−258.6	−230.8	***−127.0***
E_{int}(X)	***−37.7***	***−230.5***	***−241.0***	***−41.5***	***−231.4***	***−47.3***	***−49.5***	***−30.0***	***−21.1***	***−43.1***	***−233.9***	***−8.5***	***−260.1***	***−19.4***	***−253.7***	***−230.7***	
Co^{2+}	*84.0*	*−111.0*	*−111.9*	*80.1*	*−111.8*	*78.5*	*76.0*	*105.8*	*108.2*	*78.5*	*−110.8*	*113.8*	*−139.7*	*103.1*	*−132.2*	*−110.0*	
Ni^{2+}	*86.1*	*−109.9*	*−111.7*	*82.2*	*−110.8*	*77.3*	*74.9*	*99.0*	*106.9*	*80.7*	*−108.9*	*115.5*	*−142.3*	*105.0*	*−131.9*	*−111.7*	
Cu^{2+}	*88.7*	*−109.8*	*−117.4*	*84.8*	*−110.9*	*78.1*	*75.8*	*92.0*	*104.0*	*83.0*	*−111.8*	*119.9*	*−137.2*	*107.9*	*−131.5*	*−115.2*	
Zn^{2+}	*83.8*	*−109.3*	*−112.4*	*80.1*	*−109.9*	*75.5*	*73.4*	*91.8*	*105.7*	*78.6*	*−107.0*	*113.2*	*−134.4*	*102.1*	*−128.1*	*−102.7*	
Cd^{2+}	*82.9*	*−102.5*	*−117.2*	*79.4*	*−103.0*	*74.0*	*72.1*	*87.8*	*97.8*	*78.1*	*−105.7*	*109.8*	*−129.9*	*99.8*	*−124.4*	*−98.2*	
Hg^{2+}	*90.6*	*−98.5*	*−133.1*	*86.9*	*−99.4*	*74.9*	*73.3*	*86.1*	*93.1*	*85.1*	*−116.7*	*119.3*	*−134.9*	*108.3*	*−131.6*	*−103.8*	

[a] The smaller (more negative) values indicate the higher affinity of a substituting functional group for metal (compared to H_2O).

(a) The interaction energies of negatively charged residues (deprotonated amino acid side chains) are by an order of magnitude greater than those of the neutral species. This trivial fact is a consequence of the gas-phase calculation of molecular complexes with different charges: in one case, the interaction energy of dipositive ion with neutral ligand is computed, while in the other, the interaction of dipositive and negative charge makes the dominant contribution to the overall interaction energy. However, since we are primarily interested in the differences between the studied TM ions (columns of Tables 2−5) and they are all of the same charge, the described phenomenon is not an objection to the meaningful analysis of results.

(b) In the last column of Tables 2−5 is the average of E_{int}(Me,X) over all the functional groups, denoted as $\bar{E}_{int}$(Me). The more negative its value, the greater is the overall affinity of the particular TM ion for the amino acid side chain. The value of $\bar{E}_{int}$(Me) is highest for Cu^{2+}, then follows Ni^{2+}, and approximately at the same level are Zn^{2+} and Co^{2+}. As for the heavier TM ions, Cd^{2+} is invariably last in this series, while $\bar{E}_{int}$(Hg) is approximately equal to that of Ni^{2+} in most coordination geometries. These results are in a very good agreement with the Irving−Williams (IW) series of stability constants[4] (applying to the first row TM ions) which are, in turn, determined by the bond strength between the TM ion and

Table 5. Interaction Energies,[a] E_{int}(Me,X), of Amino Acid Side Chains with Studied TM Ions in Linear Coordination Geometry, Defined by Eq 1, and Relative Interaction Energies E_{rel}(Me,X)—Listed in Lower Part of Table in Italics—Defined by Eq 4 (All Values in kcal·mol^{-1})

TM\AA	Asn	Asp$^-$	Cys$^-$	Gln	Glu$^-$	His$_{N\delta}$	His$_{N\epsilon}$	Lys	Met	PeptO	PeptN$^-$	Ser$^-$	Thr$^-$	Tyr$^-$	$\bar{E}_{int}$(Me)
Co^{2+}	−71.1	−299.1	−304.5	−76.5	−296.8	−81.0	−84.1	−59.1	−50.1	−78.7	−299.8	−322.3	−320.1	−305.0	***−189.2***
Ni^{2+}	−64.5	−297.7	−309.7	−70.3	−297.9	−81.3	−85.1	−56.2	−44.6	−74.0	−310.7	−324.4	−322.5	−308.4	***−189.1***
Cu^{2+}	−79.6	−316.4	−341.4	−86.6	−317.4	−106.9	−111.2	−83.5	−72.8	−88.9	−339.0	−344.6	−344.0	−341.5	***−212.4***
Zn^{2+}	−60.1	−283.2	−293.9	−65.4	−282.7	−74.8	−75.7	−51.8	−48.2	−67.8	−287.2	−304.6	−304.7	−276.7	***−176.9***
Cd^{2+}	−53.7	−273.0	−287.4	−58.5	−272.3	−68.4	−69.3	−49.0	−47.7	−60.7	−275.5	−292.2	−292.6	−267.9	***−169.2***
Hg^{2+}	−59.6	−278.9	−309.5	−64.8	−278.4	−80.9	−81.8	−61.9	−63.7	−67.5	−295.1	−308.0	−308.7	−285.5	***−181.7***
E_{int}(X)	***−64.6***	***−291.2***	***−306.8***	***−69.4***	***−289.5***	***−82.2***	***−84.4***	***−60.1***	***−54.3***	***−71.7***	***−299.8***	***−315.5***	***−315.3***	***−297.5***	
Co^{2+}	*118.1*	*−109.9*	*−115.3*	*112.7*	*−107.6*	*108.2*	*105.1*	*130.1*	*139.1*	*110.5*	*−110.6*	*−133.1*	*−130.9*	*−115.8*	
Ni^{2+}	*124.6*	*−108.6*	*−120.6*	*118.8*	*−108.8*	*107.8*	*104.0*	*132.9*	*144.5*	*115.1*	*−121.6*	*−135.3*	*−133.4*	*−119.3*	
Cu^{2+}	*132.8*	*−104.0*	*−129.0*	*125.8*	*−105.0*	*105.5*	*101.2*	*128.9*	*139.6*	*123.5*	*−126.6*	*−132.2*	*−131.6*	*−129.1*	
Zn^{2+}	*116.8*	*−106.3*	*−117.0*	*111.5*	*−105.8*	*102.1*	*101.2*	*125.1*	*128.7*	*109.1*	*−110.3*	*−127.7*	*−127.8*	*−99.8*	
Cd^{2+}	*115.5*	*−103.8*	*−118.2*	*110.7*	*−103.1*	*100.8*	*99.9*	*120.2*	*121.5*	*108.5*	*−106.3*	*−123.0*	*−123.4*	*−98.7*	
Hg^{2+}	*122.1*	*−97.2*	*−127.8*	*116.9*	*−96.7*	*100.8*	*99.9*	*119.8*	*118.0*	*114.2*	*−113.4*	*−126.3*	*−127.0*	*−103.8*	

[a] The smaller (more negative) values indicate the higher affinity of a substituting functional group for metal (compared to H_2O).

ligand, presuming that the complexes have the same coordination number, geometry and similar character of ligands. More of this principle will be discussed below.

(c) In the middle row of Tables 2−5, the average of E_{int}(Me,X) over six studied TM ions, denoted as $\bar{E}_{int}$(X), is evaluated. Using this quantity, the following approximate order of amino acid side chains according to their affinity to the studied TM ions can be derived from the calculated data: *neutral amino acid side chains: His(Nϵ) > His(Nδ)* > carbonyl oxygen (*Asn*, *Gln*, *PeptO*) > *Lys* > *Tyr* (calculated only in OH geometry) > *Thr*, *Met* (most variable with respect to TM ion and coordination geometry) > *Ser*; *deprotonated amino acid side chains: Ser*$^-$ > *Thr*$^-$ > *Cys*$^-$ > *PeptN*$^-$ > *Glu*$^-$, *Asp*$^-$, *Tyr*$^-$. The calculations reproduce the empirical evidence that *His* is the most common binding residue in metalloproteins and even correlate with the fact that binding through *His(Nϵ)* prevails over *His(Nδ)* which has been explained by its better sterical accessibility.[6] However, as follows from the performed calculations, it is also energetically more favorable, contrary to a situation in free histidine, in which Nδ is deprotonated first, or equivalently, 4-methylimidazole is more stable than 5-methylimidazole. The order in the series of deprotonated amino acid side chain affinities should be modified by the estimates of the energetic cost of deprotonation of neutral residues. It can be done from the differences in their gas-phase deprotonation energies or known pK_a constants of their free forms. We have used the former quantity, because it is consistent with our model (since we calculate E_{int}, not G_{int}). The calculations have been carried out at the B3LYP/6-311++G(d,p)//B3LYP/6-311+G(d) level and the following values of $\Delta\Delta E_{deprot} = \Delta E_{deprot}(H_2O) - \Delta E_{deprot}(XH)$ obtained: *Asp* (−43.1 kcal·mol^{-1}), *Cys* (−35.1 kcal·mol^{-1}), *Glu* (−44.3 kcal·mol^{-1}), *Peptide nitrogen* (−26.1 kcal·mol^{-1}), *Ser* (−8.1 kcal·mol^{-1}), *Thr* (−13.8 kcal·mol^{-1}), *Tyr* (−41.7 kcal·mol^{-1}). Therefore, the modified order runs approximately as: *Cys*$^-$ > *Asp*$^-$, *Glu*$^-$ > *Tyr*$^-$ > *PeptN*$^-$, *Ser*$^-$, *Thr*$^-$. It perfectly correlates with the evidence that *Cys*$^-$, *Asp*$^-$, *Glu*$^-$, *Tyr*$^-$ are the second, third, fourth, and fifth residues according to their abundance in the metal-binding sites.[6]

(d) The interaction energies E_{int}(Me,X) are always negative, which means that all these functional groups are capable of substituting water molecule from per-hydrated complex, if the effect of the molecular environment is neglected (e.g., conformational changes in more complex ligands, solvation and entropy effects). It is not a very surprising result for the ligands binding via oxygen atoms, because water is smallest of all of them and does not contain the functional groups that can act as electron donors for oxygen atoms and strengthen the bond to a metal ion (for a more detailed discussion, see also the next section). On the other hand, we presume that no such statement can be a priori made, if also ligands with nitrogen and sulfur are included in the comparison. In this respect, we consider it a noteworthy finding.

(e) When four considered coordination geometries are compared with each other, the values of $-E_{int}$(Me,X) are highest for linear, followed by tetrahedral and square planar, and smallest for octahedral coordination geometry. It is a consequence of the above-discussed greater affinity of the studied ligands for TM ions in comparison with water. As has been mentioned above, the interaction between TM ion and the particular functional group contributes approximately by $^1/_n$ (where n is the coordination number) to the overall interaction energy. The effective interaction between the TM ion and a given ligand is therefore weakest for the octahedral and strongest for linear geometry. Since all the functional groups are, according to computed interaction energies, better ligands than water, it can be expected that their affinities toward the studied TM ions will increase in the same direction. As can be seen in Tables 2−5, this fact has been well reproduced by the calculations.

Before exploiting more information from Tables 2−5 and addressing the key point of the paper—metal ion selectivity—several observations concerning the calculated equilibrium bond distances will be made.

The optimized bond distances enable us to calculate the mutual ratio of ionic radii of the studied TM ions in their coordination compounds. It must be kept in mind that they represent gas-phase values obtained from the accurate quantum chemical calculations. Nevertheless, we believe that it is a very instructive parameter and gives the complementary information to the ionic radii obtained from the experimental crystal structures (where the crystal packing forces cause that the crystal geometry differs from the geometry in solution). Averaging over all coordination geometries and studied functional groups, the following ratio (dimensionless) has been obtained:

$$r(Co^{2+}):r(Ni^{2+}):r(Cu^{2+}):r(Zn^{2+}):r(Cd^{2+}):r(Hg^{2+}) = 1.01:1:1:1.01:1.12:1.15$$

Furthermore, it should be noted that bond distances between TM ions and negatively charged ligands are 0.1−0.2 Å shorter than those of their neutral counterparts. It is a consequence of a stronger interaction between systems with opposite charges. It leads to a weakening of the remaining (n − 1) coordination bonds between TM ion and water molecules. The mean interatomic distance between these two increases by 0.05−0.1 Å.

Table 6. Equilibrium Metal−Ligand Bond Distances between the Donor Atom of the Amino Acid Side Chain and TM Ion and Mean Interatomic Distances between TM Ion and Water Molecules (in Italics)[a]

coordination geometry	metal ion	Asn	Asp^-	Cys^-	H_2O	$His_{N\delta}$	$His_{N\epsilon}$	Lys	Met	PeptO	$PeptN^-$	Ser	Ser^-	Thr	Thr^-	Tyr	Tyr^-
octahedral	Co^{2+}	2.081	2.038	2.402	2.117	2.132	2.100	2.151	2.542	2.073		2.113		2.097		2.131	1.994
		2.126	*2.135*	*2.164*		*2.141*	*2.146*	*2.144*	*2.135*	*2.127*		*2.121*		*2.124*		*2.120*	*2.151*
	Ni^{2+}	2.040	1.998	2.357	2.087	2.092	2.060	2.111	2.495	2.033		2.073		2.056		2.091	1.954
		2.086	*2.095*	*2.124*		*2.101*	*2.105*	*2.104*	*2.095*	*2.087*		*2.081*		*2.084*		*2.080*	*2.111*
	Cu^{2+}	2.071	2.028	2.390	2.112	2.122	2.090	2.141	2.529	2.063		2.103		2.087		2.121	1.984
		2.116	*2.125*	*2.154*		*2.131*	*2.136*	*2.134*	*2.125*	*2.117*		*2.111*		*2.114*		*2.110*	*2.141*
	Zn^{2+}	2.049	2.005	2.336	2.124	2.105	2.069	2.124	2.534	2.040		2.108		2.081		2.127	1.942
		2.141	*2.151*	*2.220*		*2.167*	*2.173*	*2.171*	*2.154*	*2.144*		*2.133*		*2.137*		*2.128*	*2.172*
	Cd^{2+}	2.250	2.155	2.485	2.328	2.291	2.260	2.336	2.683	2.240		2.307		2.283		2.318	2.112
		2.342	*2.362*	*2.436*		*2.367*	*2.371*	*2.365*	*2.362*	*2.345*		*2.334*		*2.338*		*2.331*	*2.383*
	Hg^{2+}	2.286	2.168	2.423	2.391	2.227	2.194	2.274	2.576	2.267		2.351		2.317		2.364	2.127
		2.417	*2.465*	*2.619*		*2.484*	*2.491*	*2.478*	*2.489*	*2.425*		*2.404*		*2.412*		*2.407*	*2.526*
tetrahedral	Co^{2+}	1.901	1.889	2.206	2.005	1.973	1.964	2.043	2.371	1.895	1.884	1.970	1.798	1.950	1.803		1.835
		2.027	*2.032*	*2.079*		*2.032*	*2.033*	*2.023*	*2.026*	*2.031*	*2.085*	*2.012*	*2.080*	*2.018*	*2.081*		*2.071*
	Ni^{2+}	1.886	1.876	2.180	2.004	1.942	1.935	2.001	2.328	1.879	1.862	1.947	1.782	1.930	1.783		1.846
		2.030	*2.047*	*2.100*		*2.045*	*2.046*	*2.039*	*2.042*	*2.036*	*2.119*	*2.021*	*2.092*	*2.028*	*2.091*		*2.093*
	Cu^{2+}	1.862	1.864	2.205	1.998	1.937	1.939	1.954	2.309	1.855	1.921	1.890	1.818	1.888	1.826		1.907
		2.053	*2.074*	*2.150*		*2.080*	*2.083*	*2.053*	*2.065*	*2.059*	*2.149*	*2.037*	*2.115*	*2.046*	*2.120*		*2.142*
	Zn^{2+}	1.908	1.898	2.226	1.997	1.974	1.965	2.039	2.374	1.903	1.902	1.973	1.821	1.953	1.821		1.849
		2.016	*2.023*	*2.078*		*2.028*	*2.029*	*2.022*	*2.031*	*2.019*	*2.064*	*2.004*	*2.059*	*2.010*	*2.059*		*2.043*
	Cd^{2+}	2.131	2.218	2.415	2.229	2.180	2.176	2.261	2.570	2.125	2.098	2.204	2.033	2.178	2.036		2.061
		2.250	*2.291*	*2.342*		*2.268*	*2.265*	*2.259*	*2.272*	*2.254*	*2.319*	*2.236*	*2.318*	*2.243*	*2.320*		*2.296*
	Hg^{2+}	2.177	2.280	2.419	2.282	2.181	2.178	2.267	2.555	2.169	2.109	2.252	2.074	2.219	2.076		2.127
		2.314	*2.386*	*2.489*		*2.350*	*2.346*	*2.334*	*2.363*	*2.320*	*2.453*	*2.291*	*2.444*	*2.305*	*2.448*		*2.445*
square planar	Co^{2+}	1.915	1.904	2.214	2.032	2.001	1.999	2.054	2.423	1.925	1.876	2.023	1.796	1.990	1.800		1.824
		2.066	*2.047*	*2.145*		*2.062*	*2.060*	*2.078*	*2.067*	*2.054*	*2.138*	*2.029*	*2.114*	*2.038*	*2.120*		*2.106*
	Ni^{2+}	1.919	1.890	2.231	1.993	1.981	1.983	2.013	2.394	1.913	1.897	2.005	1.794	1.977	1.807		1.856
		2.026	*2.020*	*2.097*		*2.026*	*2.014*	*2.056*	*2.047*	*2.029*	*2.091*	*1.985*	*2.072*	*2.001*	*2.077*		*2.073*
	Cu^{2+}	1.912	1.900	2.292	1.967	1.954	1.953	1.988	2.410	1.906	1.927	1.969	1.843	1.945	1.851		1.926
		1.986	*1.991*	*2.049*		*1.994*	*1.992*	*2.016*	*2.016*	*1.990*	*2.073*	*1.965*	*2.047*	*1.973*	*2.046*		*2.097*
	Zn^{2+}	1.939	1.915	2.244	2.035	1.994	1.994	2.013	2.405	1.930	1.899	2.037	1.811	1.998	1.821		1.827
		2.056	*2.052*	*2.172*		*2.076*	*2.073*	*2.114*	*2.100*	*2.059*	*2.154*	*2.029*	*2.130*	*2.039*	*2.134*		*2.102*
	Cd^{2+}	2.159	2.079	2.422	2.246	2.197	2.197	2.238	2.589	2.149	2.101	2.246	2.019	2.207	2.032		2.047
		2.267	*2.274*	*2.391*		*2.288*	*2.285*	*2.309*	*2.312*	*2.272*	*2.368*	*2.243*	*2.350*	*2.255*	*2.354*		*2.343*
	Hg^{2+}	2.175	2.097	2.386	2.294	2.136	2.138	2.179	2.511	2.157	2.071	2.293	2.017	2.227	2.025		2.124
		2.339	*2.365*	*2.557*		*2.402*	*2.396*	*2.416*	*2.435*	*2.351*	*2.528*	*2.293*	*2.494*	*2.319*	*2.504*		*2.541*
linear	Co^{2+}	1.815	1.827	2.130	1.930	1.885	1.889	1.971	2.273	1.803	1.851		1.721		1.736		1.756
		1.933	*1.972*	*2.026*		*1.975*	*1.973*	*1.935*	*1.971*	*1.953*	*1.963*		*1.990*		*2.003*		*1.977*
	Ni^{2+}	1.777	1.800	2.118	1.932	1.910	1.904	1.945	2.246	1.776	1.794		1.713		1.719		1.742
		1.927	*1.920*	*1.980*		*1.931*	*1.929*	*1.928*	*1.955*	*1.926*	*1.960*		*1.954*		*1.958*		*1.962*
	Cu^{2+}	1.787	1.827	2.148	1.882	1.912	1.910	1.913	2.249	1.805	1.856		1.797		1.804		1.812
		1.890	*1.921*	*1.945*		*1.908*	*1.905*	*1.896*	*1.925*	*1.897*	*1.924*		*1.921*		*1.923*		*1.919*
	Zn^{2+}	1.799	1.842	2.144	1.868	1.861	1.859	1.945	2.274	1.797	1.815		1.754		1.762		1.781
		1.885	*1.928*	*1.967*		*1.891*	*1.888*	*1.889*	*1.914*	*1.886*	*1.962*		*1.926*		*1.933*		*1.961*
	Cd^{2+}	2.034	2.080	2.349	2.106	2.071	2.070	2.163	2.485	2.029	2.027		1.977		1.986		2.050
		2.122	*2.178*	*2.208*		*2.122*	*2.119*	*2.120*	*2.148*	*2.123*	*2.180*		*2.177*		*2.187*		*2.262*
	Hg^{2+}	2.026	2.118	2.346	2.088	2.042	2.040	2.124	2.442	2.022	2.022		1.985		1.998		2.130
		2.111	*2.196*	*2.242*		*2.117*	*2.112*	*2.114*	*2.162*	*2.113*	*2.198*		*2.186*		*2.204*		*2.365*

[a] All distances are in Å.

The calculated ratio of ionic radii tempts us to modify the explanation of the IW series of stability constants. It has been explained by the different ionic radii of the first row TM ions, presuming purely ionic nature of metal−ligand bond. Thus, the greater the ionic radius is, the higher are the stability constants of that particular TM ion (when comparing TM ions with the same charges). As has been demonstrated above, the calculated interaction energies $-\bar{E}_{int}$(Me) correlate exceedingly well with IW series, but the same is not true for the calculated ratio of ionic radii. There is a good correlation between Co^{2+}, Ni^{2+}, and Zn^{2+} ions, but Cu^{2+}, for which the highest values of $-\bar{E}_{int}$(Me) were unequivocally computed, has the same average radii as Ni^{2+} and does not suit this explanation. This effect is most pronounced in octahedral Cu^{2+} complexes, whose structures were derived from optimized Ni^{2+} species by increasing metal−ligand distances by 0.03 Å. Thus, for reverse order of radii the correct order of interaction energies has been obtained. We do not think it may be an artifact of the calculations and note that the reference state is per-hydrated complex, which further corrects errors that might be otherwise present. We rather put forward the hypothesis that the IW series of stability constants is not determined purely by ionic radius and electrostatic nature of coordination bonds in complexes, but the covalent part of metal−ligand bond (although being a small contribution to the overall interaction energy) plays an important role in determining the stability constants. Unfortunately, it cannot be understood in such simple terms as electrostatic interactions of ionic species, since the polarization, exchange, and correlation energy significantly contributes to it. Fortunately, as can be seen above, high-level quantum chemical calculations performed on simple models of TM centers yields results in good agreement with experimental evidence.

We have already addressed two factors that help to a selective binding of studied TM ions to a predesigned site: the preference for a coordination geometry and the optimum metal−ligand distance. The last and perhaps the most important one is related to the differences in the affinities of TM ions for the amino acid side chains, characterized by the values of E_{int}(Me,X). Therefore, after analyzing general trends drawn from this quantity in the former part of the discussion, the attention will be paid to the specific trends that may be observed in Tables 2−5.

To separate the contributions that are common to all amino acid side chains (intrinsic qualities of particular TM ion, determining the above-discussed IW series) from E_{int}(Me,X) and make this quantity more illustrative for purpose of studying metal ion selectivity, we have adopted a similar approach as for the evaluation of the affinities for coordination geometry. Hence, the relative affinity is defined as:

$$E_{rel}(\mathrm{Me,X}) = E_{int}(\mathrm{Me,X}) - \bar{E}_{int}(\mathrm{Me}) \qquad (4)$$

It should be noted that sum of E_{rel}(Me,X) over all functional groups, E_{rel}(Me), is zero. The calculated relative affinities for each amino acid side chain are listed in lower parts of Tables 2−5, and they will serve as the measures of their selectivity for the studied TM ions. In physical terms, the eq 4 defines hypothetical "average amino acid side chain" to be the reference state for the calculation of interaction energy rather than [Me-$(H_2O)_n]^{2+}$ complex.

In the next part, the studied amino acid residues will be discussed separately and the differences between maximum and minimum from {$E_{rel}(Me_1,X)$, ..., $E_{rel}(Me_6,X)$}, denoted as ΔE_{rel}-(Gm), evaluated for each of them (in all four coordination geometries). The values of ΔE_{rel}(Gm) can be considered as the measures of the total selectivity of a given residue (selectivity factors: SF).

Asn (ΔE_{rel}(OH) = 5.8 kcal·mol^{-1}; ΔE_{rel}(TH) = 8.5 kcal·mol^{-1}; ΔE_{rel}(SQ) = 7.7 kcal·mol^{-1}; ΔE_{rel}(Lin) = 17.3 kcal·mol^{-1}) exhibits enhanced affinity toward Co^{2+}, Zn^{2+}, and Cd^{2+} ions, followed by Ni^{2+}, and disfavors Cu^{2+} and Hg^{2+} by approximately 5−10 kcal·mol^{-1}. According to the SFs, carbonyl oxygen (present also in *Gln* and *Peptide oxygen*) could be classified as the moiety with lower to medium selectivity.

Asp$^-$ (6.2 kcal·mol^{-1}; 6.9 kcal·mol^{-1}; 12.5 kcal·mol^{-1}; 12.7 kcal·mol^{-1}) contains negatively charged carboxyl oxygen and according to HSAB principle should not prefer soft metals (Cd^{2+} and Hg^{2+}). With the exception of tetrahedral coordination geometry that we cannot explain, this fact is perfectly reproduced by the calculations. It has virtually equal relative affinity toward the first row TM ions (Co^{2+}, Cu^{2+}, Ni^{2+}, and Zn^{2+}) that decreases by approximately 5 kcal·mol^{-1} for Cd^{2+} and 7 kcal·mol^{-1} for Hg^{2+}. It belongs to the residues with medium selectivity.

Cys$^-$ (21.1; 19.0; 21.4; 13.7) is remarkable for its high total selectivity and high affinity for Hg^{2+} (in accord with HSAB), which is closely followed by Cu^{2+}. Cd^{2+} and Ni^{2+} are disfavored by approximately 10 kcal·mol^{-1} and the smallest affinity is invariably (in all geometries) exhibited for Co^{2+} and Zn^{2+} that lose another 2−3 kcal·mol^{-1} in the values of E_{rel}.

Gln (5.5; 7.8; 7.5; 15.1) as the ligand for the studied TM ions has the same behavior as already described *Asn* residue.

Glu$^-$ (6.3; 5.7; 12.4; 12.1) is almost the same ligand as *Asp*$^-$.

His (2.1; 6.7; 4.2; 6.3) has surprisingly low selectivity toward the studied TM ions, despite its very high average interaction energy−$\bar{E}_{int}$(His)−(vide supra). It should be noted that the difference in the E_{int}(Me,X) of His(Nδ) and His(Nϵ) remains almost constant for all TM ions and all coordination geometries. Therefore, all of the observed trends are valid for both isomers. It slightly prefers Zn^{2+} and Cd^{2+}, but the difference of approximately 2 kcal·mol^{-1} between these two and other TM ions is almost negligible.

Lys (3.8; 4.1; 19.7; 13.1) prefers to bind to Cd^{2+} and Hg^{2+} ions by approximately 3−5 kcal·mol^{-1}. They are followed by Zn^{2+}, and Cu^{2+}, and it has lowest affinity for Co^{2+} and Ni^{2+}. According to SF, it has a low selectivity in octahedral and tetrahedral coordination geometries. However, it discriminates Co^{2+} and Ni^{2+} ions to a large extent (that causes a high value of SF) in square planar and linear geometries.

Met (9.9; 12.7; 15.1; 26.5) exhibits similar tendencies as *Cys*$^-$. It invariably prefers Hg^{2+}, which is followed by Cd^{2+} (with $-E_{rel}$ by approximately 4 kcal·mol^{-1} smaller). By another 5 kcal·mol^{-1} behind is Cu^{2+}, and then follow three remaining TM ions, losing approximately another 2 kcal·mol^{-1} (Co^{2+}, Ni^{2+}, Zn^{2+}). According to the values of ΔE_{rel}, *Met* belongs to residues with high total selectivity.

Peptide oxygen (4.8; 7.4; 7.0; 15.0) contains carbonyl oxygen as donor atom and therefore has almost the same coordination properties as *Asn* and *Gln*, measured by the value of E_{rel}.

Deprotonated peptide nitrogen (n.a.; 14.7; 11.0; 20.3) is a rare ligand in the metal-binding sites of metalloproteins, since only copper(II) ion is likely to deprotonate amide nitrogen of peptide bond in four-coordinate complexes (we presume that other divalent ions may deprotonate it in the sites with linear arrangement). Indeed, the calculations reproduce these exceptional properties of Cu^{2+}. The highest value of $-\bar{E}_{int}$(Me), discussed in the context of the IW series, is further supported by the high values of $-E_{rel}$ compared to the remaining TM ions (only Hg^{2+} in square planar coordination has slightly higher $-E_{rel}$). Since, in the series of equivalent complexes differing by only central TM ion, the deprotonation energy represents the dominant contribution to pK_a value of a bound functional group, we consider the computed values of E_{int}(Me,X) as another demonstration of an agreement between the simple theoretical model and experimental evidence. Theoretically (according to SFs), it has a high total selectivity, which is at the same time the most variable among the studied TM ions. It can be only said that the relative affinity of *PeptN*$^-$ is lowest for Cd^{2+}.

Ser (8.1; 14.6; 10.1; n.a.) is, according to the values of SF, a residue with a higher total selectivity. In comparison with hypothetical "average residue", it prefers to bind to Cd^{2+}, followed by Zn^{2+} and Co^{2+}, then Ni^{2+} and Hg^{2+}, and the last in the series is Cu^{2+}. This order is not surprising, since it is almost the same ligand as water molecule, and thus we obtain an inverse IW series. It should be noted that we are not certain about the protonation states of primary and secondary alcohol groups in four-coordinated complexes and phenol ring in octahedral geometry at pH = 7.

Ser$^-$ (n.a.; 14.3; 12.4; 12.3) has a negatively charged oxygen as donor atom, and in accord with HSAB principle prefers to bind to Ni^{2+}, Cu^{2+}, and Co^{2+}, followed by Zn^{2+} (with approximately 3−5 kcal·mol^{-1} smaller relative affinity) and the last in the series are Hg^{2+} and Cd^{2+}.

Thr (6.6; 11.9; 8.5; n.a.) has similar coordination properties as *Ser*; however, in octahedral geometry, it disfavors Cu^{2+} to a smaller extent and slightly more disfavors Co^{2+}.

Thr$^-$ (n.a.; 13.9; 7.8; 10.0) is almost the same ligand as *Ser*$^-$, with the exception of Ni^{2+} in square planar coordination geometry, which is disfavored by 5 kcal·mol^{-1} in comparison with *Ser*$^-$.

Tyr (5.0; n.a.; n.a.; n.a.) has been computed only in octahedral coordination geometry. According to SF, belongs to residues

with low selectivity. The relative affinity is highest for Hg^{2+}, while all the remaining TM ions are within 2 kcal·mol^{-1} range.

Tyr$^-$ (15.4; 22.8; 17.0; 30.4) has very high total selectivity (SF). It favors Cu^{2+} ion to a great extent (by 5−10 kcal·mol^{-1} in all coordination geometries). It also exhibits tendency to bind to Ni^{2+} and Co^{2+}, while Cd^{2+} is the least likely to be found in the sites containing *Tyr*$^-$.

We believe that a different affinity of amino acid side chains for the studied TM ions (metal ion selectivity) has been clearly demonstrated. It is determined by the electronic structure of amino acid side chains, for example, its polarizibility, charge distribution, and character of donor atom. Nevertheless it should be stressed that the estimates of metal ion selectivity based on E_{rel}(Me,X) only shift the overall equilibrium that is primarily settled by the intrinsic properties of transition metals (of which preference for specific coordination geometry and IW series have been discussed). On the other hand, the differences in relative interaction energies between different TM ions that have been shown to be about 3−5 kcal·mol^{-1} may sum up to give a total of 20−30 kcal·mol^{-1} in octahedral coordination (under the assumption of additivity of E_{rel}) which is encouraging. It also explains why the metal uptake by more complex molecules does not always obey the simple rules derived from the behavior of the smaller ligands.

Elongation of Amino Acid Side Chain. To draw our theoretical models nearer to the target systems—metal-binding sites of metalloproteins—we have to evaluate the influence of further elongation of the amino acid side chains on the interaction energies. Owing to long-range properties of dominating electrostatic interaction, the absolute value of interaction energy of TM ion with a surrounding biomolecule will be different from the E_{int}(Me,X) energies calculated in this work. Nevertheless, we put forward the hypothesis that differences in E_{int}(Me,X) between TM ions which theoretically define metal ion selectivity, are determined by local environment of central TM ion. More precisely, we presume that the selectivity of particular metal-binding site in metalloprotein will be to a large extent determined by the binding residues, while the rest of protein may influence the overall stability of the system or the kinetics of metal uptake.

The thorough discussion of this problem is beyond the scope of this work, but the calculations of eight systems in each coordination geometry have been carried out to address it in an approximate way. We have compared the interaction energies of *Asp*$^-$ with *Glu*$^-$, *Asn* with *Gln*, CH_3NH_2 with *Lys*, and CH_3SCH_3 with *Met*. We define ΔE_{elong} as:

$$\Delta E_{elong}(\mathrm{Me,X}) = E_{int}(\mathrm{Me,X_{long}}) - E_{int}(\mathrm{Me,X_{short}}) \quad (5)$$

where X_{long} = (*Glu*$^-$, *Gln*, *Lys*, and *Met*) and X_{short} = (*Asp*$^-$, *Asn*, CH_3NH_2, and CH_3SCH_3).

The calculated values of ΔE_{elong}(Me,X) are listed in Table 7.

It should be noted that most of the values of ΔE_{elong} are negative which implies the higher affinity of longer functional groups for TM ions. It is caused by the positive inductive effect, that is, by the electron-donating properties of alkyl group. Because the negatively charged carboxyl group is the least polarizable and therefore least influenced by the inductive effect, it explains why the interaction energies of amino acid side chains containing a carboxyl group do not vary with the addition of methyl group. Also for *Gln*(*Asn*) and *Met*(CH_3SCH_3) functional groups (with the exception of $[Ni(Met)(H_2O)]^{2+}$ complex in linear and perhaps $[Co(Met)(H_2O)_5]^{2+}$ in octahedral coordination), the ΔE_{elong} is virtually equal for all six TM ions. The small differences in ΔE_{elong} exhibit the similar trends as described by the IW series (quantified by $-\bar{E}_{int}$(Me) in this work) which is not very surprising. The same holds true for *Lys*(CH_3NH_2) residue, but owing to the addition of propyl group the correlation between IW series and trends in ΔE_{elong} is more pronounced. We think that three complexes with deviations in ΔE_{elong} could be identified: $[Co(Lys)(H_2O)_3]^{2+}$, $[Ni(Lys)(H_2O)_3]^{2+}$ in square planar, and $[Cu(Lys)(H_2O)]^{2+}$ in linear coordination. At present we have no rigorous explanation for the source of the observed deviations.

Table 7. Calculated Differences in the Interaction Energies, ΔE_{elong} ($\Delta E_{elong} = E$(Me,longX) − E(Me,shortX)), Quantifying the Effect of an Elongation of Amino Acid Side Chain, Computed for All Four Coordination Geometries (All Values in kcal·mol^{-1})

coordination geometry	AA s. ch.	Co^{2+}	Ni^{2+}	Cu^{2+}	Zn^{2+}	Cd^{2+}	Hg^{2+}
octahedral	Glu^-	−0.2	−0.3	−0.6	−0.2	−0.1	−0.2
	Gln	−2.9	−2.9	−3.0	−2.9	−2.7	−2.7
	Lys	−5.9	−6.7	−8.0	−6.4	−6.1	−7.3
	Met	−2.6	1.4	0.8	1.4	0.9	1.1
tetrahedral	Glu^-	−0.5	−0.7	−1.2	−0.5	0.3	0.1
	Gln	−3.7	−3.9	−4.5	−3.6	−3.3	−3.3
	Lys	−8.1	−8.9	−12.2	−8.0	−7.6	−8.8
	Met	−1.2	−1.3	−1.8	−1.2	−1.2	−1.5
square planar	Glu^-	−0.8	−0.9	−1.1	−0.6	−0.5	−0.9
	Gln	−3.9	−3.9	−3.9	−3.7	−3.5	−3.7
	Lys	3.2	−3.0	−9.9	−8.2	−7.8	−9.4
	Met	−2.1	−2.7	−2.8	−2.7	−2.6	−3.1
linear	Glu^-	2.3	−0.2	−1.0	0.5	0.7	0.5
	Gln	−5.4	−5.8	−6.9	−5.3	−4.8	−5.2
	Lys	−14.2	−17.0	−26.8	−11.8	−11.4	−12.8
	Met	−4.0	−1.0	−4.6	−4.3	−4.0	−4.2

It has been demonstrated that in most cases the elongation of amino acid side chain does not change the calculated metal ion selectivities and that the functional groups used in this work are satisfactory models for the amino acid residues that are part of protein structure.

IV. Conclusions

In this work, the results of DFT calculations using a large basis set have been presented. We presume that the data itself may turn up to be an invaluable source of information for theoretical chemists studying the systems containing TM ions (e.g., for the construction of more accurate parameters to force fields, development of hybrid QM/MM methods, etc.), structural biochemists (assignment of initial parameters to crystallographic or XAS measurements) or may be utilized in different context than we have done.

We attempted to address three important issues that are invariably met in the discussions of the selectivity of particular molecular or supramolecular systems (e.g., metal-binding sites) for TM ions:

(i) The preference for the specific coordination arrangement of ligands. Although factors other than the first coordination shell ligands may play an important role in determining this quality, we have attempted to characterize it quantitatively by the interaction (complexation) energies of TM ions with preorganized $(H_2O)_n$ site (E_{Gm}(Me) and E'_{Gm}(Me), defined by eqs 2 and 3). The calculations reproduced the experimentally observed preference of Co^{2+}, Ni^{2+} ions for octahedral, Cu^{2+} for square planar, Zn^{2+} for tetrahedral, and Hg^{2+} for linear coordination geometry. The preferred coordination geometry cannot be unambiguously assigned to Cd^{2+} ion, but the calculations suggest that it is likely to be found in linear and tetrahedral coordinations.

(ii) The different affinity of TM ions for a particular donor atom or whole ligand, which has been quantitatively characterized by the interaction energy (E_{int}(Me,X) and E_{rel}(Me,X) defined by eqs 1 and 4) of amino acid side chains, peptide bond oxygen, and peptide bond nitrogen with TM ions. The good correlation between the trends in the computed interaction energies and the IW series of stability constants has been shown. They have been also in a good agreement with the experimentally found abundance of amino acid side chains in the metal-binding sites of metalloproteins and with HSAB principle. Furthermore, the binding of His through Nϵ rather than Nδ has been explained upon the basis of differences in molecular energies of these isomers. The important issue of the metal ion selectivity has been addressed, selectivity factors evaluated for all amino acid side chains, and their preference for specific TM ion discussed.

(iii) The optimum size of the metal-binding site for a given TM ion. It has been quantitatively characterized by the equilibrium metal−ligand distances between the functional groups representing amino acid side chains and TM ions. By averaging the optimized metal−ligand distances, the ratio of ionic radii of studied TM ions has been calculated. We presume that it may be closer to the actual ratio in solution than the values derived from the crystal structures. Furthermore, the ratio of ionic radii leads to a modification in the explanation of the Irving−Williams series of stability constants. Due to the good correlation between the computed interaction energies and the IW series and the poorer correlation between the computed ionic radii and the IW series, we have postulated that the IW series is not determined purely by the ratio of ionic radius and TM ion charge (ionic contribution) and that covalent contribution to the coordination bond (characterized by a charge transfer between TM ion and ligand) plays a non-negligible role in determining the energetic stability of the complex.

As has been mentioned in the Introduction, the computational scheme used in this work has been well-tested by comparison with CCSD(T) method, and the role of the basis set has been thoroughly investigated. Therefore, we assume that the possible extensions of this work should focus on the chemical model rather than the improvements in the computational scheme. In the near future, we would like to discuss the cooperative effect, that is, the simultaneous binding of two and more residues (which is equivalent to substitution of more than one water molecule), and its influence on the metal ion selectivity. There is also an open field in the accurate quantification of the effect of elongation of amino acid residues, which has been treated in the approximate way in this work. In the first step, the capping hydrogen representing the C_α atom in our work may be replaced by the small fragment of peptide backbone (e.g., NH_2CHCHO group) in the model functional groups. Another discussion can concentrate upon the determination of pK_a microconstants (quantifying the deprotonation of a single donor atom of amino acid side chain) which would enable to assign the protonation states in the metal-binding sites at specific pH.

Despite the suggested continuations, we believe that the presented work brings a useful piece of knowledge to the field of bioinorganic chemistry.

Acknowledgment. This work was supported by Grants 203/98/0650 (GA CR), A4055801/1998 (GA AV CR), and Project LN00A032 (Center for Complex Molecular Systems and Biomolecules). CPU time on Origin 2000 at MU, Brno (project MetaCentre) is gratefully acknowledged.

JA001265G