

Theoretical Studies of Metal Ion Selectivity. 3. A Theoretical Design of the Most Specific Combinations of Functional Groups Representing Amino Acid Side Chains for the Selected Metal Ions (Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+})[†]

Lubomír Rulíšek* and Zdeněk Havlas

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

Received: September 9, 2002; In Final Form: January 10, 2003

Theoretical efforts to devise the most selective combinations of amino acid (AA) side chains (i.e., metal-binding sites) for six studied metal ions— Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+} —have been accomplished in this work. Utilizing the results of the preceding two articles of the series, the complexation energies of these ions have been calculated for each combination of simple functional groups representing AA side chains in four coordination geometries (linear, tetrahedral, square planar, and octahedral) according to the *approximate* formula derived recently. To eliminate the errors of 5–15 kcal mol⁻¹ inherent in approximate calculations, 10 potential candidates for the most specific site for each ion have been subject to a rigorous quantum chemical investigation. It resulted in the calculations of their complexation energies in 50 different sites and the final assignment of three or four most specific sites for each of them. As the calculated results were obtained by the “gas-phase model chemistry”, a part of discussion is devoted to the accuracy and limitations of the above predictions.

I. Introduction

In recent years, there has been a rapid progress in the field of quantum bioinorganic chemistry. It has been stimulated by the beliefs that the combination of theoretical and experimental efforts will give answers to some key questions pertinent to the coordination of metal ions in biomolecules. One of the most important is: “How does the electronic structure of particular metal ion correlate with the structural or functional role it has in the metalloproteins or DNA/RNA molecules”? Equivalently, we may ask if it is the specificity and the uniqueness of the given metal ion that predetermines its abundance in the key enzymes of metabolism or on the contrary, the combination of the external factors, such as its bioavailability.^{1,2} On the way toward answering these questions, numerous studies dealing with the smaller systems modeling the structure or redox properties of active sites in metalloproteins,³ energetics of biocatalysis,⁴ or different binding modes of metal ions in DNA/RNA fragments⁵ can be found. These studies are complemented by the highly accurate studies performed for (usually) the isolated cation–molecule systems,⁶ where comparison with the gas-phase experimental data⁷ is straightforward.

Because these achievements are well documented in excellent reviews written recently,⁸ let us focus on the studies dealing with the central theme of this work: metal ion selectivity. Within this concept, the inherent properties of metal ions are investigated and correlated with the structure of their counterparts, metal-binding sites, to obtain a better understanding of their function in biomolecules. The metal ion selectivity (or, equivalently, the specificity of the site) is quantified by the stability constants of the resulting complex [metal + site]. To illustrate the typical problems encountered in its unraveling, we may point

out to the reader a work of Vrettos et al.,⁹ in which authors experimentally quantify the ion selectivity of the Ca^{2+} site in photosystem II. Theoretical attempts of its quantification for nontransition metal ions, Mg^{2+} and Zn^{2+} , can be found in the work of Dudev and Lim¹⁰ and Garmer et al.¹¹ In the former one, the changes in Gibbs energies, calculated for the $\text{Mg}^{2+}/\text{Zn}^{2+}$ exchange reactions in different environments (characterized by varying dielectric constants), have been used as the function for the assessment of the specificity of the sites comprised the simple models of most common AA side chains. Together with the theoretical evaluations of preference for the particular coordination geometry, they provided some insights into the origin of Zn^{2+} vs Mg^{2+} selectivity.

We contributed to the progress in this field by the careful evaluation of the structural and energetical parameters of the binding of selected transition metal (TM) ions to amino acids and proteins. From the careful analysis of the experimental structures of metalloproteins and their comparison with the smaller molecule crystal structures,¹² that updated the older work of Glusker,¹³ the preferred coordination geometries and the binding modes of the studied TM ions (Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+}) in metalloproteins were determined and the abundance of amino acid side chains directly bound to the metal ion evaluated. Then, the accurate computational strategy for the calculations of various molecular properties of ionic complexes of TM ions has been proposed and tested both for the systems with single-reference¹⁴ and multireference ground electronic states.¹⁵ It has been concluded that the DFT/B3LYP method yields the interaction energies with mean absolute deviation of less than 0.5 kcal mol⁻¹ from the reference CCSD(T) values, but at least a triple- ζ basis set with diffuse and polarization functions on all atoms should be used.

In this series of papers, the above findings were further exploited for the investigation of the energetical and structural aspects of metal–biomolecule interactions. In the first article,¹⁶

[†] Part 1: *J. Am. Chem. Soc.* **2000**, *122*, 10428. Part 2: *J. Phys. Chem. A* **2002**, *106*, 3855.

* To whom correspondence should be addressed. Phone and Fax: +420-2-20183292. E-mail: lubos@uochb.cas.cz.

the interaction energies of all N-, O-, and S-containing (metal-binding) amino acid side chains (including an oxygen and a deprotonated nitrogen of peptide bond) with six studied TM ions have been calculated in four coordination geometries (octahedral—OH, tetrahedral—TH, square planar—SQ, and linear—Lin). The interaction energies have been defined as the energies of substitution of a single functional group X for one H₂O molecule in the perhydrated $[M(H_2O)_n]^{2+}$ complex. It has been demonstrated that the relative affinities of TM ions for AA side chains are determined by the structure of the first coordination shell that is by the local environment of metal. The calculated interaction energies were in agreement with the qualitative theories of coordination chemistry, such as the HSAB principle,¹⁷ the Irving—Williams series of stability constants (IW series),¹⁸ and an observed preference for the particular coordination geometry.^{12,13} Moreover, the calculated data have been used for the quantitative definition of selectivity factors for each metal-binding residue.

In the second article,¹⁹ we have investigated the cooperative effect that is the nonadditivity in the calculated interaction energies (defined as above) accompanying substitution of the second functional group for H₂O. It has been shown that the computed values of the cooperative effect can be used (together with interaction energies) for a reasonably accurate predictions of the complexation energies in general $[MX_1...X_n]^{2+}$ metal-binding sites (with the error of 5–15 kcal mol⁻¹, usually less than 3% in relative scale). Besides, it has been shown that the complexation energies (either calculated or evaluated approximately) are in a good agreement with the experimentally determined stability constants of two metalloenzymes—carbonic anhydrase²⁰ and carboxypeptidase.²¹ The above two studies yielded the underlying theoretical data for the discussions concerning the specificities of metal-binding sites in metalloproteins or complemented the experimental measurements of stability constants of the series of TM ions in a given site, especially in the cases where three-dimensional structure is unknown.

In the present paper, we would like to complete this work by a theoretical design of the most specific sites for each of six TM ions, using two factors determining metal ion selectivity: preferred coordination geometry and a carefully selected combination of functional groups representing side chains of amino acids (H₂O, H₂S, HCHO, CH₃NCH₂, NH₃, OH⁻, SH⁻, and HCOO⁻). The study consists of the following two steps:

(i) The selection of the most specific sites, based upon the *approximate* complexation energies (using the empirical formula derived recently¹⁹ together with the published values of the interaction energies and the cooperative effects) for all possible combinations of model functional groups from the above set. This preselective step is necessary, because the full-scale quantum chemical study of all combinations of AA side chains, even using the above reduced set, would have resulted in more than 40 000 systems to be calculated.

(ii) Quantum chemical calculations of the complexation energies for 10 best candidates for each TM ion. These *calculated* complexation energies are then used as the criterion for the assessment of the candidates. Because it has been shown that the errors in the *approximate* complexation energies are fairly small (vide supra), we expect the probability of not including any specific site in the final selection as being quite low.

Though we are aware that this study is only one step further in the efforts to design highly specific metal-binding sites, one can think of several practical applications, once such sites have been designed: the removal of metals from the polluted en-

vironment, either by bacterial strains²² or other biotechnological techniques;²³ a design of new biosensors;²⁴ redesign of proteins (maquettes);²⁵ or possible new chelating compounds in medicinal chemistry.²⁶ The possible applications also justify the selection of TM ions, which has been explained in preceding articles in more details.

As has been mentioned in previous papers of the series, the gas-phase results should be transferred to the condensed phase with an extreme care, which holds true also for the quantification of the metal-ion selectivity, because the environmental effects (e.g., protein bulk and solvation) may play a significant role in the “real (bio)chemical systems”.

II. Computational Details

All of the calculations were performed with Gaussian 98 program suite.²⁷

All of the calculations were performed in the framework of density functional theory (DFT). The three-parameter functional developed by Becke,²⁸ which combines the Becke’s gradient-corrected exchange functional and the Lee, Yang, and Parr and Vosko, Wilk, and Nusair correlation functionals²⁹ with part of the exact Hartree—Fock exchange energy, has been employed (denoted as B3LYP).

Two basis sets have been used throughout the calculations, denoted as BS1 and BS2. BS1 has been a 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 non-hydrogen elements.

BS2 consisted of the triple- ζ (TZ) basis set of Wachters and Hay³⁰ for the first row transition metals (Co, Ni, Cu, and Zn) and standard 6-311G for other elements (H, C, N, O, and S).³¹ It was augmented by the same diffuse functions as BS1 with further addition of s functions for hydrogens and the following sets of polarization functions: 2fg for TMs, 2df for other heavy atoms, and 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) and 6-311++G(2df,2pd) (BS2) keywords.

For Cd²⁺ and Hg²⁺, effective core potentials (ECP) of Stevens and co-workers³² have been used (denoted SBKJ). To achieve the consistency with the above-described 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 [$\alpha_d(\text{Cd}) = 0.075$, $\alpha_d(\text{Hg}) = 0.040$] and f [$\alpha_f(\text{Cd}) = 0.775$, $\alpha_f(\text{Hg}) = 0.690$] and 2fg [$\alpha_{1f}(\text{Cd}) = 2.0$, $\alpha_{2f}(\text{Cd}) = 0.3$, $\alpha_g(\text{Cd}) = 0.775$; $\alpha_{1f}(\text{Hg}) = 1.35$, $\alpha_{2f}(\text{Hg}) = 0.35$, $\alpha_g(\text{Hg}) = 0.69$] sets of polarization functions, corresponding to BS1 and BS2, respectively.

The computational scheme consisted of several steps:

First, the optimizations of molecular geometries of Zn²⁺ complexes have been carried out at the B3LYP/BS1 level with angles at the metal center fixed at the values corresponding to the given coordination geometry and all other internal coordinates optimized.

Second, all other systems were assumed to adopt the geometry of the optimized Zn²⁺ complexes, and only *n* metal—ligand distances (*n* = 4 and 6) were optimized at the B3LYP/BS1 level. This approximation has been shown to cause errors of less than 0.1 kcal mol⁻¹ in the values of interaction energies of simple ligands.

Third, the single-point energy calculations of all of the studied structures have been carried out at the B3LYP/BS2 level to obtain the final molecular energies of $[MX_n]^{2+}$ complexes.

Fourth, the metal ion at the optimized geometry has been replaced by the corresponding ghost atom B_{QM} and a single-point energy calculated for B_{QM}(X_n) systems.

A short note should be added about the described scheme. By introducing the coordination geometry constraints, we are fully aware that the optimized structure may not be necessarily the global minimum of the complex. However, we must stress that this constraint belongs to the chosen chemical model, because the target structure represented by our model systems is the metalloprotein in the specific coordination geometry. Then, the ligands may be kept at the given coordination geometry by the structural constraints in the biomolecule that may prevent its collapsing into the minimum found for the small model systems.

Throughout the paper, the interaction (complexation) energy of functional groups X₁, ..., X_n coordinated to the metal ion M in a given geometry is defined as

$$E_{\text{int}}(\text{M}, \text{X}_1, \dots, \text{X}_n) = E([\text{MX}_1, \dots, \text{X}_n]^{2+}) - E(\text{B}_{\text{QM}}\text{X}_1, \dots, \text{X}_n) - (E([\text{M}(\text{H}_2\text{O})_n]^{2+}) - E(\text{B}_{\text{QM}}(\text{H}_2\text{O})_n)) \quad (1)$$

where $n = 4$ (TH and SQ) and 6 (OH coordination geometry). According to this formula, the computed interaction energy includes the correction for the nonbonding interactions between ligands and a part of basis set superposition error (BSSE).

As has been described in the preceding articles of the series, the studied TM ions have high-spin ground states in the class of studied compounds. Therefore, the calculated values refer to the quartet state of Co²⁺, triplet of Ni²⁺, and doublet of Cu²⁺ complexes.

III. Results and Discussion

IIIa. Calculations of Approximate Complexation Energies.

In preceding article of the series, it has been shown that the complexation energies of TM ions in simple metal-binding sites of a general formula [MX₁, ..., X_n]²⁺, where X_i is any of {H₂O, H₂S, HCHO, CH₃NCH₂, NH₃, OH⁻, SH⁻, and HCOO⁻} and $n = 4$ (TH and SQ) 6 (OH) can be calculated according to the formula

$$E_{\text{int}}(\text{M}, \text{X}_1, \dots, \text{X}_n, \text{est}) = \sum_i E_{\text{int}}(\text{M}, \text{X}_i) + \frac{1}{2} \sum_{i < j} \Delta E_{\text{coop}}(\text{M}, \text{X}_i, \text{X}_j) \quad (2)$$

The values of $E_{\text{int}}(\text{M}, \text{X}_i)$ and $\Delta E_{\text{coop}}(\text{M}, \text{X}_i, \text{X}_j)$ have been published for all seven model functional groups (H₂S through HCOO⁻) in four coordination geometries (Lin, TH, SQ, and OH), whereas they are by a definition equal to zero for X_i = H₂O or X_j = H₂O (see eq 1 above and eq 2 of ref 19).

Therefore, a program was written that calculates the approximate complexation energies of six studied TM ions for each chemically unique site in TH, SQ, and OH geometries according to the above formula. For eight different functional groups, there are 330, 666, and 8436 chemically unique complexes for TH, SQ, and OH, respectively. We have used only one restriction for these sites regarding the total formal charge of the ligands, which should be less or equal to minus three; that is the combinations containing four, five, or six negative ligands—OH⁻, SH⁻, and HCOO⁻—have been rejected (owing to the absence of such sites in metalloproteins¹² and the calculated instability of the sites containing more than three negatively charged formates⁶ⁱ). It reduced the total number of

sites submitted for subsequent processing to 315, 645, and 7030 (TH, SQ, and OH). For each site, two selectivity factors (maximum and average) were defined according to the formulas

$$SF_{\text{Max}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) = E'_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) - \min(E'_{\text{int}}(\text{M}_j, \text{X}_1, \dots, \text{X}_n), j = 1, \dots, 6; j \neq i) \quad (3)$$

$$SF_{\text{Avg}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) = E'_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) - \text{avg}(E'_{\text{int}}(\text{M}_j, \text{X}_1, \dots, \text{X}_n), j = 1, \dots, 6; j \neq i) \quad (4)$$

$$E'_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) = E_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) - \bar{E}(\text{M}_i) \quad (5)$$

where M_i denotes *i*th TM ion ($i = 1, \dots, 6$) and $\text{avg}(x_i, i = \dots)$ is the arithmetical average of values { x_i }. E'_{int} is the relative interaction energy, resulting from the subtraction of average interaction energy of TM ion M_i (averaged over all values of $E_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n)$ in particular coordination geometry) that is the property inherent to each TM ion and related to the experimentally known IW series. The similar procedure has been used (and discussed in more details) for the evaluation of selectivity factors of each AA side chain in the previous paper.¹⁶ This subtraction is necessary for definition of SF_{Max} to correctly reproduce what is intuitively meant by a selectivity of the given site (whereas SF_{Avg} values would be only shifted by the constant amount, if E_{int} were used for the definition). Selectivity factors, $SF(\text{M}_i, \text{X}_1, \dots, \text{X}_n)$, defined by eqs 3 and 4, measure to what extent is the affinity of the particular TM ion M_i for a given site higher than of the best ($SF_{\text{Max}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n)$), or the average ($SF_{\text{Avg}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n)$) of remaining TM ions. The smaller (or more negative) is the value of SF , the higher is the specificity of the given site for ion M_i.

In the next step of the computational analysis, the preferred coordination geometries were assigned: Co²⁺, Ni²⁺, octahedral; Cu²⁺, square planar; Zn²⁺, Cd²⁺, tetrahedral; Hg²⁺, linear. They have been derived from the analysis of metalloproteins and smaller molecules crystal structures¹² and also theoretically justified in previous work.¹⁶ Then, the calculated approximate complexation energies and selectivity factors for 7030 sites in octahedral geometry were sorted according to the $SF_{\text{Max}}(\text{Co}^{2+}, \text{X}_1, \dots, \text{X}_6)$ and $SF_{\text{Avg}}(\text{Co}^{2+}, \text{X}_1, \dots, \text{X}_6)$ to obtain the candidates for the most specific Co²⁺ sites. The same procedure has been repeated for $SF_{\text{Max}}(\text{Ni}^{2+}, \text{X}_1, \dots, \text{X}_6)$, $SF_{\text{Avg}}(\text{Ni}^{2+}, \text{X}_1, \dots, \text{X}_6)$ in OH geometry; $SF_{\text{Max}}(\text{Cu}^{2+}, \text{X}_1, \dots, \text{X}_4)$, $SF_{\text{Avg}}(\text{Cu}^{2+}, \text{X}_1, \dots, \text{X}_4)$ for 645 complexes in SQ geometry; $SF_{\text{Max}}(\text{Cd}^{2+}, \text{X}_1, \dots, \text{X}_4)$, $SF_{\text{Avg}}(\text{Cd}^{2+}, \text{X}_1, \dots, \text{X}_4)$, $SF_{\text{Max}}(\text{Zn}^{2+}, \text{X}_1, \dots, \text{X}_4)$, $SF_{\text{Avg}}(\text{Zn}^{2+}, \text{X}_1, \dots, \text{X}_4)$ for 315 complexes in TH geometry; and $SF_{\text{Max}}(\text{Hg}^{2+}, \text{X}_1, \text{X}_2)$, $SF_{\text{Avg}}(\text{Hg}^{2+}, \text{X}_1, \text{X}_2)$ for 28 complexes in Lin geometry. For mercury(II), where the values of $E_{\text{int}}(\text{M}, \text{X}_1, \text{X}_2)$ have already represented the calculated complexation energies, the final selection of the most specific sites has been done at this point (and will be presented in the next subchapter together with the other results), whereas for each of five remaining TM ions, 10 candidates for the subsequent quantum chemical calculations were selected according to three criteria:

(i) The site should not contain more than four identical ligands, preferably less. The reason for this restriction is that such sites are rare, if nonexistent in metalloproteins.¹² Besides, it also reduces possible errors, which might be present in the values of $E_{\text{int}}(\text{M}, \text{X})$ and $\Delta E_{\text{coop}}(\text{M}, \text{X}, \text{Y})$ and could sum up to the nonnegligible value (owing to the equal sign for the one type of ligand) in the approximate complexation energy.

(ii) In octahedral geometry, the potential candidate should differ from the already selected one in at least two ligands, or by one ligand and the different configuration of X₁, ..., X₆

TABLE 1: Calculated Relative IW Factors, $\bar{E}_{\text{Gm}}(\text{M}_i) - \bar{E}_{\text{Gm}}(\text{Cd})$, for Each Metal Ion in Four Studied Coordination Geometries^a

coordination geometry (CG)	Co ²⁺	Ni ²⁺	Cu ²⁺	Zn ²⁺	Cd ²⁺	Hg ²⁺
linear	-15.53	-14.10	-32.22	-9.20	0.00	-13.65
tetrahedral	-4.76	-10.38	-18.38	-5.46	0.00	-11.68
sq. planar	-5.99	-9.49	-12.97	-5.89	0.00	-14.71
octahedral	-2.42	-6.48	-15.80	-3.21	0.00	-13.06

^a All values are in kcal mol⁻¹.

(cis-, trans-). The reason is to bring more variety into the set of selected sites, because it is presumed that similar complexes will exhibit the similar level (degree) of specificity.

(iii) The total molecular charge of the ligands in selected sites should be distributed over four possibilities (0, -1, -2, -3). Usually, the pattern (2:3:3:2) has been used (i.e., two complexes with neutral ligands, three with one, three with two, and two with three negatively charged ligands). Again, the reason is to bring more variety into the set of potential candidates. The notable exception was the set of cadmium(II) tetrahedral complexes, where negatively charged complexes had very unfavorable selectivity factors (in agreement with HSAB theory), and the (6:3:1:0) pattern has been used instead.

The final choice of the best candidates can be found in the first columns of Tables 2–6.

IIIb. DFT/B3LYP Calculations of Complexation Energies for the Most Specific Sites. Ten candidates for “most specific site” for each TM ion except Hg²⁺ (totally 50 sites), chosen according to the above procedure, were subject to a quantum chemical investigation according to the described computational scheme. It resulted in the calculations of complexation energies for 300 systems (each ion, including Hg²⁺, interacting with each site).³³ The results are summarized in Tables 2–6. The estimated complexation energies, used for the selection of the particular site in the preceding step, are listed as well.

Before discussing the results, we focus on the quantification of the theoretical Irving–Williams factors (IW factors), defined by the average interaction energy of $\bar{E}_{\text{Gm}}(\text{M}_i)$, where Gm = Lin, TH, SQ, and OH. Such a quantity has been already used for the definition of selectivity of AA side chains,¹⁶ because it enables us to separate the contribution inherent to TM ion and the part determined by the particular site $\{X_1, \dots, X_n\}$ from values of $E_{\text{int}}(\text{M}, X_1, \dots, X_n)$ and introduces the concept of the site specificity. For the purpose of the accurate evaluation of the specificity of the studied sites, the values of $\bar{E}_{\text{Gm}}(\text{M}_i)$ were refined to a highest possible degree. For their evaluation, all of the calculated interaction energies have been used (mono-, di-, and per-substituted species from the previous work, and all the complexes calculated here, totally 36, 55, 77, and 84 complexes in Lin, TH, SQ, and OH geometries, respectively). The resulting values are listed in Table 1 in a relative scale, that is $\bar{E}'_{\text{Gm}}(\text{M}_i) = \bar{E}_{\text{Gm}}(\text{M}_i) - \bar{E}_{\text{Gm}}(\text{Cd})$, because cadmium(II) exhibits the lowest strength of interaction in all four geometries.

Then, the values of $\bar{E}'_{\text{Gm}}(\text{M}_i)$ were subtracted from the calculated $E_{\text{int}}(\text{M}, X_1, \dots, X_n)$, and thus, we have arrived to the most important energetical characteristic of each site, *relative interaction energy*, denoted as $E'_{\text{int}}(\text{M}, X_1, \dots, X_n)$ and listed in Tables 2–6 together with the actual selectivity factors. The adopted approach is a direct analogy of what has been described in the previous section, with the difference that the *calculated* (not only *estimated*) values are used for the evaluation of E'_{int} and SF^* s.

The results in Tables 2–6 will be discussed separately for each set of M²⁺-specific sites.

Co²⁺-Specific Sites (OH, Table 2). Co²⁺-specific sites are the most difficult to devise owing to higher differences between calculated and estimated values of E_{int} (in comparison with remaining sets of M²⁺-specific sites) that are increasing the probability of not including a specific site into the set of 10 candidates. It may be attributed to the computational difficulties accompanying the quantum chemical calculations of Jahn–Teller unstable octahedral $[\text{CoX}_6]^{2+}$ complexes with quasi-degenerate ground electronic states,¹⁵ possibly resulting in errors higher than declared 0.5–1 kcal mol⁻¹ for each value of $E_{\text{int}}(\text{Co}, X_i)$ and $\Delta E_{\text{coop}}(\text{Co}, X_i, X_j)$. It might be the case for the $\text{CH}_3\text{N}=\text{CH}_2$ (Imi) group as the model of histidine side chain, for which the values of $\Delta E_{\text{coop}}(\text{Co}, \text{Imi}, \text{Imi})$ and $\Delta E_{\text{coop}}(\text{Co}, \text{Imi}, X) (\approx -4 \text{ kcal mol}^{-1})$ can be slightly overestimated and, as a consequence, the binding sites containing imidazole groups “unrealistically” preferred. Besides, there are difficulties to discern ions of Co²⁺ and Ni²⁺, because they exhibit quite similar coordination behavior.³⁴ This fact is also reproduced by the calculations: $E'_{\text{int}}(\text{Co}) - E'_{\text{int}}(\text{Ni})$ is usually 1–2 kcal mol⁻¹ for octahedral sites (see Tables 2 and 3). Despite the above difficulties, SF_{Avg} 's for 9 out of 10 sites are substantially less than zero, which means that the calculated sites belong to the specific ones for Co²⁺. With respect to the values of $SF_{\text{Max}}(\text{Co})$ and $SF_{\text{Avg}}(\text{Co})$, we have chosen sites **5**: $\{(\text{H}_2\text{CO})_3, (\text{H}_2\text{O})_2, \text{OH}^-\}$, **9**: $\{(\text{Imi})_3, (\text{HCOO}^-)_3\}$, and **10**: $\{(\text{NH}_3)_2, (\text{HCOO}^-)_2, \text{H}_2\text{S}, \text{OH}^-\}$ as the most Co²⁺-specific. They are depicted in Figure 1, together with the optimized metal–ligand distances that determine the optimum size of the preformed cavity in a more complex molecule (e. g., the metal-binding site in metalloprotein). At this point, it can be reminded that H₂O represents hydroxyl-containing side chains (Ser, Thr, and Tyr); H₂S thiol or thioether groups of Met and Cys; H₂CO carbonyl oxygen (peptide bond oxygen, Gln, and Asn); Imi imidazole ring of His, NH₃ side chain of Lys; OH⁻ alcoholates (Ser⁻, Thr⁻, and Tyr⁻), SH⁻ thiolate group of Cys⁻, and HCOO⁻ carboxylate residues (Asp⁻ and Glu⁻).

Ni²⁺-Specific Sites (OH, Table 3). Ni²⁺-specific sites have favorable selectivity factors. SF_{Max} 's are, with exception of site nos. 5 and 10, negative (−1 to −4.5 kcal mol⁻¹), which means they are much better for Ni²⁺ in comparison with the “average” Ni²⁺ site than they are for remaining TM ions (in comparison with their “average” M²⁺ sites). SF_{Avg} 's are between −5 to −11 kcal mol⁻¹ (with exception of site no. 2). Most usually, Co²⁺ (or Cu²⁺) is the ion with the closest values of *relative* complexation energies, E'_{int} . However, the octahedral geometry is less favorable for the copper(II) ion, and therefore, when deciding between two sites with approximately equal selectivity factors, a preference should be given to the site that discriminates between Ni²⁺ and Co²⁺ rather than Cu²⁺ (which has been the case of site no. 5 in Co²⁺-specific sites in previous section). We have denoted the following three sites as most Ni²⁺-specific: **1**: $\{(\text{Imi})_3, (\text{NH}_3)_3\}$, **5**: $\{(\text{Imi})_3, (\text{H}_2\text{O})_2, \text{HCOO}^-\}$, and **7**: $\{(\text{NH}_3)_2, (\text{HCOO}^-)_2, \text{Imi}, \text{H}_2\text{O}\}$. They are depicted at the bottom of Figure 1.

Cu²⁺-Specific Sites (SQ, Table 4). Cu²⁺-specific sites seem to be the most M²⁺-specific ones. With a couple of them (1, 4, 5, 6, 7, and 10), SF_{Max} 's range from −4.0 to −10.3 kcal mol⁻¹ and SF_{Avg} 's from −7.4 to −13.0 kcal mol⁻¹. This specificity is further enhanced by the preference for square planar geometry, typical for the Cu²⁺ ion. The following three sites are denoted as the most selective: **1**: $\{(\text{Imi})_3, \text{NH}_3\}$, **5**: $\{\text{Imi}, \text{NH}_3, \text{OH}^-, \text{SH}^-\}$, and **6**: $\{(\text{Imi})_2, \text{OH}^-, \text{SH}^-\}$. They are depicted at Figure 2.

Zn²⁺-Specific Sites (TH, Table 5). Zn²⁺-specific sites should be generally quite selective with respect to Ni²⁺ and Cu²⁺ ions

TABLE 2: Estimated, $E_{\text{int}}(\text{est})$,^a Calculated, $E_{\text{int}}(\text{calc})$,^b and Relative, E'_{int} ,^c Values of Complexation Energies (for Their Definition, See the Footnotes of Table 2) Together with the Selectivity Factors for 10 Candidates for the Most Co^{2+} -Specific Metal-Binding Sites (Octahedral Coordination Geometry)^d

ligands (selectivity factors)	energies	Co^{2+}	Ni^{2+}	Cu^{2+}	Zn^{2+}	Cd^{2+}	Hg^{2+}
1: {(H ₂ CO) ₃ ,(Imi) ₃ }	$E_{\text{int}}(\text{calc})$	-40.9	-46.5	-50.8	-37.2	-36.7	-45.6
$SF_{\text{Max}}(1) = 1.5$	$E_{\text{int}}(\text{est})$	-61.4	-44.6	-54.5	-40.6	-40.7	-51.0
$SF_{\text{Avg}}(1) = -2.8$	E'_{int}	-38.5	-40.0	-35.0	-33.9	-36.7	-32.5
2: {(H ₂ CO) ₂ ,(Imi) ₂ ,H ₂ S ₂ }	$E_{\text{int}}(\text{calc})$	-9.6	-13.9	-25.8	-5.8	-10.7	-24.3
$SF_{\text{Max}}(2) = 4.0$	$E_{\text{int}}(\text{est})$	-23.4	-11.6	-20.0	-5.7	-11.3	-27.7
$SF_{\text{Avg}}(2) = 1.2$	E'_{int}	-7.2	-7.5	-10.0	-2.6	-10.7	-11.2
3: {(Imi) ₄ ,H ₂ CO,OH ⁻ }	$E_{\text{int}}(\text{calc})$	-254.3	-255.8	-272.8	-248.8	-242.4	-260.5
5.2	$E_{\text{int}}(\text{est})$	-261.9	-247.7	-259.5	-244.8	-238.6	-256.9
-3.5	E'_{int}	-251.9	-249.4	-257.0	-245.6	-242.4	-247.5
4: {(Imi) ₃ ,H ₂ CO ₂ ,HCOO ⁻ }	$E_{\text{int}}(\text{calc})$	-214.2	-218.9	-227.2	-210.4	-205.7	-216.7
0.6	$E_{\text{int}}(\text{est})$	-245.5	-233.2	-240.1	-227.3	-221.3	-232.9
-3.7	E'_{int}	-211.8	-212.4	-211.4	-207.2	-205.7	-203.6
5: {(H ₂ CO) ₃ ,H ₂ O ₂ ,OH ⁻ }	$E_{\text{int}}(\text{calc})$	-223.5	-223.3	-238.4	-220.0	-213.6	-226.9
1.6	$E_{\text{int}}(\text{est})$	-219.2	-219.2	-231.3	-217.1	-211.7	-226.0
-4.3	E'_{int}	-221.1	-216.8	-222.6	-216.7	-213.6	-213.9
6: {(Imi) ₃ ,H ₂ CO,OH ⁻ ,HCOO ⁻ }	$E_{\text{int}}(\text{calc})$	-409.6	-412.3	-431.9	-407.2	-397.9	-413.7
8.8	$E_{\text{int}}(\text{est})$	-441.3	-430.3	-437.7	-424.8	-413.1	-428.9
-2.3	E'_{int}	-407.2	-405.8	-416.1	-404.0	-397.9	-400.6
7: {(Imi) ₂ ,H ₂ CO ₂ ,HCOO ⁻ ,OH ⁻ }	$E_{\text{int}}(\text{calc})$	-411.5	-414.0	-432.2	-405.3	-395.5	-405.7
7.3	$E_{\text{int}}(\text{est})$	-430.9	-422.0	-432.1	-416.5	-405.1	-419.7
-6.2	E'_{int}	-409.1	-407.6	-416.4	-402.1	-395.5	-392.6
8: {(Imi) ₂ ,H ₂ O ₂ ,HCOO ⁻ ,OH ⁻ }	$E_{\text{int}}(\text{calc})$	-409.4	-413.1	-429.8	-407.1	-396.5	-405.9
7.0	$E_{\text{int}}(\text{est})$	-422.7	-425.7	-437.9	-417.9	-406.6	-422.3
-4.3	E'_{int}	-407.0	-406.6	-414.0	-403.9	-396.5	-392.8
9: {(Imi) ₃ ,HCOO ⁻ ₃ }	$E_{\text{int}}(\text{calc})$	-537.5	-539.7	-544.0	-533.8	-522.0	-523.8
-1.9	$E_{\text{int}}(\text{est})$	-593.4	-586.8	-583.1	-579.4	-561.4	-568.2
-10.1	E'_{int}	-535.1	-533.2	-528.2	-530.6	-522.0	-510.8
10: {(NH ₃) ₂ ,HCOO ⁻ ₂ ,H ₂ S,OH ⁻ }	$E_{\text{int}}(\text{calc})$	-570.6	-574.4	-577.5	-568.2	-556.7	-563.0
-0.3	$E_{\text{int}}(\text{est})$	-584.6	-592.5	-597.6	-585.4	-568.6	-580.8
-8.0	E'_{int}	-568.2	-567.9	-561.7	-565.0	-556.7	-549.9

^a $E_{\text{int}}(\text{M}, \text{X}_1, \dots, \text{X}_n, \text{est}) = \sum_i E_{\text{int}}(\text{M}, \text{X}_i) + 1/2 \sum_{i < j} \Delta E_{\text{coop}}(\text{M}, \text{X}_i, \text{X}_j)$. ^b $E_{\text{int}}(\text{M}, \text{X}_1, \dots, \text{X}_n, \text{calc}) = E([\text{MX}_1 \dots \text{X}_n]^{2+}) - E(\text{BqM}, \text{X}_1 \dots \text{X}_n) - (E([\text{M}(\text{H}_2\text{O})_n]^{2+}) - E(\text{BqM}(\text{H}_2\text{O})_n))$. ^c $E'_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) = E_{\text{int}}(\text{M}_i, \text{X}_1, \dots, \text{X}_n) - \bar{E}(\text{M}_i)$. The values of $\bar{E}(\text{M}_i)$ are taken from Table 1. ^d All values are in kcal mol⁻¹.

because of zinc(II) preference for tetrahedral coordination. Soft metal ions, Cd^{2+} and Hg^{2+} , do not prefer to bind to the sites with negatively charged ligands, as is well reproduced by the calculations. The most difficult ion to differentiate from is cobalt(II), for which tetrahedral coordination is also quite favorable and these two ions have similar stability constants in their complexes, according to IW series. As a consequence, SF_{Max} 's are close to zero (from -2.2 to 2.3 kcal mol⁻¹), whereas SF_{Avg} 's are substantially more negative. Three most Zn^{2+} -specific sites—**2**:{(NH₃)₂,H₂O,HCOO⁻}, **5**:{(NH₃)₂,HCOO⁻₂}, and **10**:{(OH⁻)₃,NH₃}—are depicted at Figure 3.

Cd²⁺-Specific Sites (TH, Table 6). Cd^{2+} -specific sites have more favorable selectivity factors than Zn^{2+} -specific tetrahedral sites. In agreement with the principles of coordination chemistry, neutral ligands are preferred by the Cd^{2+} ion. It can be seen that SF_{Avg} 's range from -7.1 to -10.7 kcal mol⁻¹ for sites containing only neutral ligands (nos. 1–6), whereas they are only -1.3 to +2.2 kcal mol⁻¹ for sites with negatively charged ligands.

It can be seen that in all sites the Hg^{2+} ion has approximately equal values of E'_{int} , and therefore, it is difficult to differentiate these two soft TM ions. This phenomenon is most striking in the case of site **6**:{(H₂S)₂,H₂O₂}, for which values of $E'_{\text{int}}(\text{Cd})$ and $E'_{\text{int}}(\text{Hg})$ are by approximately 15 kcal mol⁻¹ more favorable than for the first row TM ions. Finally, we have chosen three sites—**1**:{(H₂O)₂,H₂S,H₂CO}, **2**:{(H₂CO)₂,H₂O,H₂S}, and **3**:{(H₂CO)₃,H₂S}—as the most Cd^{2+} -specific. They are depicted at the bottom of Figure 3.

Hg²⁺-Specific Sites (Lin). Hg^{2+} -specific sites can be selected on the basis of the values of $E_{\text{int}}(\text{M}, \text{X}_1, \text{X}_2)$ that were calculated in the previous work¹⁹ and used for the definition of cooperative effect. Though the number of linear complexes with all possible combinations of eight functional groups is rather limited (36 systems), this geometry has an advantage that all values have been calculated (no estimated data are used in the selection of candidates). The most Hg^{2+} -specific sites—**1**:{(Imi)₂; $SF_{\text{Max}} = -6.9$ kcal mol⁻¹, $SF_{\text{Avg}} = -12.1$ kcal mol⁻¹}, **2**:{(H₂S)₂; -5.7, -22.4 kcal mol⁻¹}, **3**:{(SH)₂; -6.6, -11.4 kcal mol⁻¹}, and **4**:{(Imi,H₂S; -6.2, -16.5 kcal mol⁻¹)}—are depicted at Figure 4. The specificity of the sites should be further enhanced by the uniqueness of the preference of linear coordination geometry by Hg^{2+} ion.

Estimate of a Quality of the Candidates' Selection. The probability of overlooking a specific site monotonically depends on the differences between calculated and estimated values of $E_{\text{int}}(\text{M}, \text{X}_1, \dots, \text{X}_n)$. Therefore, two characteristics were evaluated for each site, ΔE_{MAD} and ΔE_{MaxDev} , defined as

$$\Delta E_{\text{MAD}} = \text{avg}_{i=1, \dots, 6} |(E_{\text{int}}(\text{M}_i, \text{calc}) - E_{\text{int}}(\text{M}_i, \text{est})) - \text{avg}_{j=1, \dots, 6} (E_{\text{int}}(\text{M}_j, \text{calc}) - E_{\text{int}}(\text{M}_j, \text{est}))| \quad (6a)$$

$$\Delta E_{\text{MaxDev}} = |\max_{i=1, \dots, 6} (E_{\text{int}}(\text{M}_i, \text{calc}) - E_{\text{int}}(\text{M}_i, \text{est})) - \min_{i=1, \dots, 6} (E_{\text{int}}(\text{M}_i, \text{calc}) - E_{\text{int}}(\text{M}_i, \text{est}))| \quad (6b)$$

The former one, mean absolute deviation, is a standard estimate

TABLE 3: Estimated, Calculated, and Relative Values of Complexation Energies (for Their Definition, See the Footnotes of Table 2) Together with the Selectivity Factors for 10 Candidates for the Most Ni²⁺-Specific Metal-Binding Sites (Octahedral Coordination Geometry)^a

ligands (selectivity factors)	energies	Co ²⁺	Ni ²⁺	Cu ²⁺	Zn ²⁺	Cd ²⁺	Hg ²⁺
1: {(Imi) ₃ , (NH ₃) ₃ }	<i>E</i> _{int} (calc)	-63.1	-71.5	-74.8	-59.4	-57.6	-68.0
<i>SF</i> _{Max} (1) = -4.4	<i>E</i> _{int} (est)	-55.1	-73.3	-76.7	-63.6	-62.0	-78.0
<i>SF</i> _{Avg} (1) = -7.3	<i>E'</i> _{int}	-60.6	-65.0	-59.0	-56.2	-57.6	-55.0
2: {(Imi) ₃ , (H ₂ O) ₂ , NH ₃ }	<i>E</i> _{int} (calc)	-45.9	-52.1	-60.3	-44.1	-43.7	-55.0
-1.1	<i>E</i> _{int} (est)	-47.1	-55.5	-64.3	-46.9	-46.5	-60.4
-2.8	<i>E'</i> _{int}	-43.4	-45.6	-44.5	-40.9	-43.7	-41.9
3: {(NH ₃) ₃ , (Imi) ₂ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-237.4	-245.3	-253.2	-234.4	-227.0	-235.4
-1.5	<i>E</i> _{int} (est)	-235.0	-251.6	-255.4	-243.5	-236.0	-250.0
-8.3	<i>E'</i> _{int}	-235.0	-238.9	-237.4	-231.1	-227.0	-222.3
4: {H ₂ O, H ₂ S, H ₂ CO, Imi, NH ₃ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-203.0	-208.7	-215.8	-200.9	-196.2	-205.2
-1.6	<i>E</i> _{int} (est)	-205.1	-210.3	-217.1	-203.1	-199.9	-212.9
-4.9	<i>E'</i> _{int}	-200.6	-202.2	-200.0	-197.7	-196.2	-192.2
5: {(Imi) ₃ , (H ₂ O) ₂ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-220.7	-227.1	-232.6	-217.4	-210.5	-217.8
-2.3	<i>E</i> _{int} (est)	-230.6	-234.6	-242.5	-226.1	-220.1	-233.1
-7.7	<i>E'</i> _{int}	-218.3	-220.6	-216.8	-214.2	-210.5	-204.8
6: {(H ₂ O) ₂ , (HCOO ⁻) ₂ , Imi, NH ₃ }	<i>E</i> _{int} (calc)	-390.1	-396.3	-404.5	-387.9	-375.7	-380.9
-1.1	<i>E</i> _{int} (est)	-396.7	-406.2	-412.8	-398.3	-386.3	-396.8
-8.9	<i>E'</i> _{int}	-387.6	-389.8	-388.7	-384.7	-375.7	-367.8
7: {(NH ₃) ₂ , (HCOO ⁻) ₂ , Imi, H ₂ O}	<i>E</i> _{int} (calc)	-397.6	-403.9	-408.1	-395.0	-383.2	-388.7
-2.3	<i>E</i> _{int} (est)	-405.9	-416.4	-421.4	-408.8	-396.0	-407.7
-9.8	<i>E'</i> _{int}	-395.1	-397.4	-392.3	-391.8	-383.2	-375.7
8: {(H ₂ O) ₂ , (NH ₃) ₂ , HCOO ⁻ , OH ⁻ }	<i>E</i> _{int} (calc)	-424.1	-426.7	-441.4	-422.7	-410.4	-426.9
5.4	<i>E</i> _{int} (est)	-423.6	-430.5	-442.3	-426.3	-413.4	-429.7
-2.0	<i>E'</i> _{int}	-421.6	-420.2	-425.6	-419.5	-410.4	-413.8
9: {(H ₂ O) ₃ , (HCOO ⁻) ₃ }	<i>E</i> _{int} (calc)	-548.2	-552.7	-557.0	-546.2	-530.0	-530.7
-0.4	<i>E</i> _{int} (est)	-555.9	-566.0	-567.6	-557.3	-539.5	-545.4
-10.7	<i>E'</i> _{int}	-545.8	-546.2	-541.2	-543.0	-530.0	-517.7
10: {(NH ₃) ₂ , (HCOO ⁻) ₂ , H ₂ S, OH ⁻ }	<i>E</i> _{int} (calc)	-570.6	-574.4	-577.5	-568.2	-556.7	-563.0
0.3	<i>E</i> _{int} (est)	-584.6	-592.5	-597.6	-585.4	-568.6	-580.8
-7.6	<i>E'</i> _{int}	-568.2	-567.9	-561.7	-565.0	-556.7	-549.9

^a All values are in kcal mol⁻¹.**TABLE 4: Estimated, Calculated, and Relative Values of Complexation Energies (for Their Definition, See the Footnotes of Table 2) Together with the Selectivity Factors for 10 Candidates for the Most Cu²⁺-Specific Metal-Binding Sites (Square Planar Coordination Geometry)^a**

ligands (sel. factors)	energies	Co ²⁺	Ni ²⁺	Cu ²⁺	Zn ²⁺	Cd ²⁺	Hg ²⁺
1: {(Imi) ₃ , NH ₃ }	<i>E</i> _{int} (calc)	-72.1	-74.0	-90.0	-70.9	-66.7	-80.2
<i>SF</i> _{Max} (1) = -10.3	<i>E</i> _{int} (est)	-76.2	-85.4	-96.9	-78.8	-73.7	-92.5
<i>SF</i> _{Avg} (1) = -11.5	<i>E'</i> _{int}	-66.1	-64.5	-77.0	-65.1	-66.7	-65.4
2: {(Imi) ₂ , NH ₃ , OH ⁻ }	<i>E</i> _{int} (calc)	-293.8	-300.7	-305.2	-290.2	-279.7	-297.9
-1.0	<i>E</i> _{int} (est)	-309.7	-315.6	-325.3	-307.2	-295.5	-317.9
-6.9	<i>E'</i> _{int}	-287.8	-291.2	-292.2	-284.4	-279.7	-283.2
3: {(Imi) ₂ , H ₂ S, OH ⁻ }	<i>E</i> _{int} (calc)	-272.9	-279.0	-283.2	-267.7	-260.9	-278.6
-0.7	<i>E</i> _{int} (est)	-288.7	-293.6	-303.4	-284.9	-277.1	-301.5
-5.6	<i>E'</i> _{int}	-266.9	-269.5	-270.2	-261.9	-260.9	-263.9
4: {(NH ₃) ₃ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-257.8	-261.2	-271.4	-255.4	-244.6	-255.1
-6.6	<i>E</i> _{int} (est)	-266.7	-270.7	-281.5	-267.5	-254.2	-267.8
-10.8	<i>E'</i> _{int}	-251.8	-251.7	-258.5	-249.5	-244.6	-240.4
5: {Imi, NH ₃ , OH ⁻ , SH ⁻ }	<i>E</i> _{int} (calc)	-469.9	-477.1	-486.3	-465.1	-454.6	-470.9
-5.7	<i>E</i> _{int} (est)	-501.0	-507.5	-518.7	-497.9	-486.2	-514.6
-13.0	<i>E'</i> _{int}	-463.9	-467.6	-473.3	-459.2	-454.6	-456.2
6: {(Imi) ₂ , OH ⁻ , SH ⁻ }	<i>E</i> _{int} (calc)	-471.2	-475.8	-485.6	-465.1	-454.8	-472.6
-6.3	<i>E</i> _{int} (est)	-502.1	-509.1	-519.5	-498.9	-487.0	-516.1
-12.0	<i>E'</i> _{int}	-465.2	-466.4	-472.7	-459.2	-454.8	-457.9
7: {(NH ₃) ₂ , SH ⁻ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-444.0	-451.2	-458.7	-444.7	-433.6	-454.0
-4.0	<i>E</i> _{int} (est)	-464.5	-471.3	-480.2	-465.5	-451.9	-477.6
-7.4	<i>E'</i> _{int}	-438.0	-441.7	-445.7	-438.8	-433.6	-439.3
8: {(OH ⁻) ₂ , H ₂ CO, SH ⁻ }	<i>E</i> _{int} (calc)	-657.2	-663.7	-665.3	-657.1	-641.4	-661.8
1.9	<i>E</i> _{int} (est)	-711.9	-716.0	-729.4	-708.1	-688.3	-716.5
-3.3	<i>E'</i> _{int}	-651.2	-654.2	-652.3	-651.2	-641.4	-647.0
9: {(SH ⁻) ₂ , Imi, OH ⁻ }	<i>E</i> _{int} (calc)	-627.2	-636.9	-642.1	-626.7	-616.8	-639.9
-1.7	<i>E</i> _{int} (est)	-688.5	-698.4	-707.0	-685.2	-671.0	-701.8
-6.8	<i>E'</i> _{int}	-621.2	-627.4	-629.1	-620.8	-616.8	-625.2
10: {NH ₃ , OH ⁻ , SH ⁻ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-646.0	-649.6	-657.8	-645.2	-628.4	-638.8
-4.8	<i>E</i> _{int} (est)	-692.6	-699.3	-708.3	-690.9	-669.1	-693.5
-10.5	<i>E'</i> _{int}	-640.0	-640.1	-644.8	-639.3	-628.4	-624.1

^a All values are in kcal mol⁻¹.

of the quality of a mutual correspondence of different sets of data, whereas the latter one, ΔE_{MaxDev} (maximum absolute

deviation), has been introduced in the previous article¹⁹ as a very sensitive gauge of the accuracy of the estimated values of

TABLE 5: Estimated, Calculated, and Relative Values of Complexation Energies (for Their Definition, See the Footnotes of Table 2) Together with the Selectivity Factors for 10 Candidates for the Most Zn²⁺-Specific Metal-Binding Sites (Tetrahedral Coordination Geometry)^a

ligands (sel. factors)	energies	Co ²⁺	Ni ²⁺	Cu ²⁺	Zn ²⁺	Cd ²⁺	Hg ²⁺
1: {(Imi) ₂ , (NH ₃) ₂ }	<i>E</i> _{int} (calc)	-73.6	-77.2	-82.0	-75.0	-69.2	-83.0
<i>SF</i> _{Max} (1) = 1.9	<i>E</i> _{int} (est)	-76.1	-87.0	-92.6	-78.7	-73.4	-90.5
<i>SF</i> _{Avg} (1) = -1.5	<i>E'</i> _{int}	-68.8	-66.8	-63.6	-69.5	-69.2	-71.4
2: {(NH ₃) ₂ , H ₂ O, HCOO ⁻ }	<i>E</i> _{int} (calc)	-247.6	-251.6	-258.2	-249.0	-239.1	-249.2
-0.7	<i>E</i> _{int} (est)	-250.8	-258.5	-267.6	-253.2	-243.9	-256.8
-3.4	<i>E'</i> _{int}	-242.8	-241.2	-239.8	-243.5	-239.1	-237.5
3: {(H ₂ O) ₂ , NH ₃ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-234.9	-240.1	-247.2	-236.3	-227.7	-236.4
-0.7	<i>E</i> _{int} (est)	-234.9	-240.1	-247.2	-236.3	-227.7	-236.4
-2.6	<i>E'</i> _{int}	-230.1	-229.7	-228.8	-230.8	-227.7	-224.7
4: {H ₂ S, Imi, NH ₃ , HCOO ⁻ }	<i>E</i> _{int} (calc)	-245.2	-251.7	-255.5	-244.4	-239.4	-253.0
2.3	<i>E</i> _{int} (est)	-246.7	-256.5	-263.3	-249.3	-242.2	-260.5
0.9	<i>E'</i> _{int}	-240.4	-241.3	-237.1	-239.0	-239.4	-241.3
5: {(NH ₃) ₂ , (HCOO ⁻) ₂ }	<i>E</i> _{int} (calc)	-447.2	-450.9	-454.8	-448.2	-431.3	-439.1
-0.3	<i>E</i> _{int} (est)	-458.1	-466.3	-472.0	-460.9	-443.6	-457.0
-7.0	<i>E'</i> _{int}	-442.4	-440.6	-436.5	-442.7	-431.3	-427.5
6: {(H ₂ S) ₂ , (HCOO ⁻) ₂ }	<i>E</i> _{int} (calc)	-394.4	-399.1	-405.9	-394.8	-386.5	-399.7
0.3	<i>E</i> _{int} (est)	-413.0	-420.2	-431.0	-415.0	-406.3	-425.0
-1.3	<i>E'</i> _{int}	-389.6	-388.8	-387.6	-389.4	-386.5	-388.0
7: {(OH ⁻) ₂ , Imi, NH ₃ }	<i>E</i> _{int} (calc)	-513.5	-518.6	-525.5	-512.7	-493.2	-503.6
1.5	<i>E</i> _{int} (est)	-531.1	-538.2	-550.7	-533.0	-512.8	-531.4
-5.4	<i>E'</i> _{int}	-508.8	-508.3	-507.2	-507.3	-493.2	-491.9
8: {(HCOO ⁻) ₂ , H ₂ S, NH ₃ }	<i>E</i> _{int} (calc)	-420.2	-425.3	-430.5	-420.7	-408.1	-418.6
0.1	<i>E</i> _{int} (est)	-435.4	-442.9	-450.8	-437.7	-424.8	-440.9
-3.8	<i>E'</i> _{int}	-415.4	-414.9	-412.1	-415.3	-408.1	-406.9
9: {(OH ⁻) ₂ , Imi, SH ⁻ }	<i>E</i> _{int} (calc)	-669.0	-675.4	-679.4	-669.0	-650.5	-662.9
1.5	<i>E</i> _{int} (est)	-705.5	-715.9	-725.5	-713.4	-692.7	-717.6
-5.1	<i>E'</i> _{int}	-664.3	-665.0	-661.1	-663.5	-650.5	-651.2
10: {(OH ⁻) ₃ , NH ₃ }	<i>E</i> _{int} (calc)	-701.1	-721.8	-717.5	-719.1	-695.7	-701.0
-2.2	<i>E</i> _{int} (est)	-742.7	-757.2	-770.7	-754.3	-726.2	-744.1
-15.2	<i>E'</i> _{int}	-696.4	-711.4	-699.1	-713.6	-695.7	-689.4

^a All values are in kcal mol⁻¹.**TABLE 6: Estimated, Calculated, and Relative Values of Complexation Energies (for Their Definition, See the Footnotes of Table 2) Together with the Selectivity Factors for 10 Candidates for the Most Cd²⁺-Specific Metal-Binding Sites (Tetrahedral Coordination Geometry)^a**

ligands (selectivity factors)	energies	Co ²⁺	Ni ²⁺	Cu ²⁺	Zn ²⁺	Cd ²⁺	Hg ²⁺
1: {(H ₂ O) ₂ , H ₂ S, H ₂ CO}	<i>E</i> _{int} (calc)	0.6	-2.7	-7.0	-0.1	-3.6	-12.5
<i>SF</i> _{Max} (1) = -2.8	<i>E</i> _{int} (est)	0.6	-2.7	-7.0	-0.1	-3.6	-12.5
<i>SF</i> _{Avg} (1) = -9.4	<i>E'</i> _{int}	5.4	7.7	11.4	5.4	-3.6	-0.8
2: {(H ₂ CO) ₂ , H ₂ O, H ₂ S}	<i>E</i> _{int} (calc)	-3.7	-7.3	-13.5	-3.6	-7.0	-15.9
-2.8	<i>E</i> _{int} (est)	-4.0	-7.7	-14.1	-4.1	-7.5	-17.0
-8.3	<i>E'</i> _{int}	1.1	3.1	4.9	1.8	-7.0	-4.2
3: {(H ₂ CO) ₃ , H ₂ S}	<i>E</i> _{int} (calc)	-9.0	-11.1	-14.8	-7.4	-10.1	-19.7
-2.1	<i>E</i> _{int} (est)	-8.0	-10.4	-16.6	-7.8	-11.1	-21.5
-7.9	<i>E'</i> _{int}	-4.3	-0.7	3.6	-1.9	-10.1	-8.0
4: {(H ₂ CO) ₂ , H ₂ S, NH ₃ }	<i>E</i> _{int} (calc)	-19.6	-23.7	-27.3	-19.9	-21.8	-33.6
0.1	<i>E</i> _{int} (est)	-19.3	-23.9	-29.5	-20.9	-23.4	-37.0
-7.1	<i>E'</i> _{int}	-14.9	-13.3	-8.9	-14.4	-21.8	-21.9
5: {H ₂ O, H ₂ S, H ₂ CO, NH ₃ }	<i>E</i> _{int} (calc)	-14.9	-19.7	-25.0	-15.7	-18.1	-29.2
-0.6	<i>E</i> _{int} (est)	-15.3	-20.6	-27.6	-17.1	-19.9	-32.9
-7.4	<i>E'</i> _{int}	-10.1	-9.3	-6.7	-10.3	-18.1	-17.5
6: {(H ₂ O) ₂ , (H ₂ S) ₂ }	<i>E</i> _{int} (calc)	10.0	5.4	-2.4	7.5	0.7	-14.2
3.2	<i>E</i> _{int} (est)	10.0	5.4	-2.4	7.5	0.7	-14.2
-10.7	<i>E'</i> _{int}	14.8	15.8	16.0	13.0	0.7	-2.5
7: {(H ₂ O) ₂ , H ₂ S, HCOO ⁻ }	<i>E</i> _{int} (calc)	-211.8	-216.9	-226.1	-212.8	-208.7	-220.2
-0.2	<i>E</i> _{int} (est)	-211.8	-216.9	-226.1	-212.8	-208.7	-220.2
-1.3	<i>E'</i> _{int}	-207.0	-206.5	-207.7	-207.3	-208.7	-208.5
8: {H ₂ O, HCOO ⁻ , H ₂ S, NH ₃ }	<i>E</i> _{int} (calc)	-223.7	-230.5	-239.0	-224.6	-219.2	-232.0
1.5	<i>E</i> _{int} (est)	-228.6	-235.9	-246.8	-230.9	-225.9	-241.9
0.7	<i>E'</i> _{int}	-219.0	-220.2	-220.6	-219.2	-219.2	-220.3
9: {H ₂ CO, HCOO ⁻ , H ₂ S, NH ₃ }	<i>E</i> _{int} (calc)	-230.9	-236.2	-242.8	-229.3	-226.0	-238.6
0.9	<i>E</i> _{int} (est)	-231.3	-238.2	-246.7	-233.1	-227.8	-244.0
-0.6	<i>E'</i> _{int}	-226.1	-225.8	-224.4	-223.9	-226.0	-226.9
10: {(H ₂ S) ₂ , (HCOO ⁻) ₂ }	<i>E</i> _{int} (calc)	-394.4	-399.1	-405.9	-394.8	-386.5	-399.7
3.2	<i>E</i> _{int} (est)	-413.0	-420.2	-431.0	-415.0	-406.3	-425.0
2.2	<i>E'</i> _{int}	-389.6	-388.8	-387.6	-389.4	-386.5	-388.0

^a All values are in kcal mol⁻¹.

complexation energies. The calculated values of ΔE_{MAD} and ΔE_{MaxDev} are listed in Table 7.

It can be seen that the average values of ΔE_{MAD} are quite small, 0.8–3.9 kcal mol⁻¹, whereas these of ΔE_{MaxDev} are 2.7–

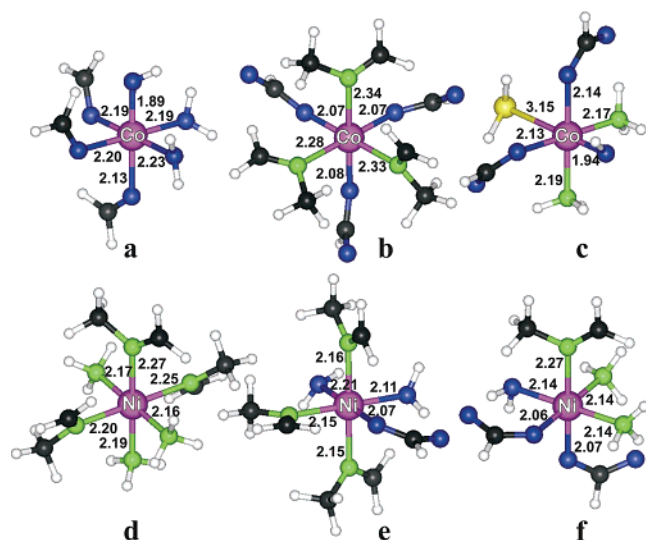


Figure 1. Most specific metal-binding sites, as predicted by theoretical calculations, for Co^{2+} and Ni^{2+} (both of them preferring octahedral coordination): (a) $[\text{Co}(\text{H}_2\text{CO})_3(\text{H}_2\text{O})_2(\text{OH})]^+$, (b) $[\text{Co}(\text{Imi})_3(\text{HCOO})_3]^-$, (c) $[\text{Co}(\text{NH}_3)_2(\text{HCOO})_2(\text{H}_2\text{S})(\text{OH})]^-$, (d) $[\text{Ni}(\text{Imi})_3(\text{NH}_3)_3]^{2+}$, (e) $[\text{Ni}(\text{Imi})_3(\text{H}_2\text{O})_2(\text{HCOO})]^+$, and (f) $[\text{Ni}(\text{NH}_3)_2(\text{HCOO})_2(\text{Imi})(\text{H}_2\text{O})]$. The metal–ligand distances are in Å.

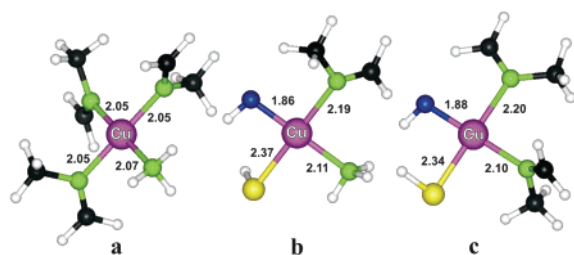


Figure 2. Most specific metal-binding sites for Cu^{2+} (preferring square planar coordination): (a) $[\text{Cu}(\text{Imi})_3(\text{NH}_3)]^{2+}$, (b) $[\text{Cu}(\text{Imi})(\text{NH}_3)(\text{OH})(\text{SH})]$, and (c) $[\text{Cu}(\text{Imi})_2(\text{OH})(\text{SH})]$. The metal–ligand distances are in Å.

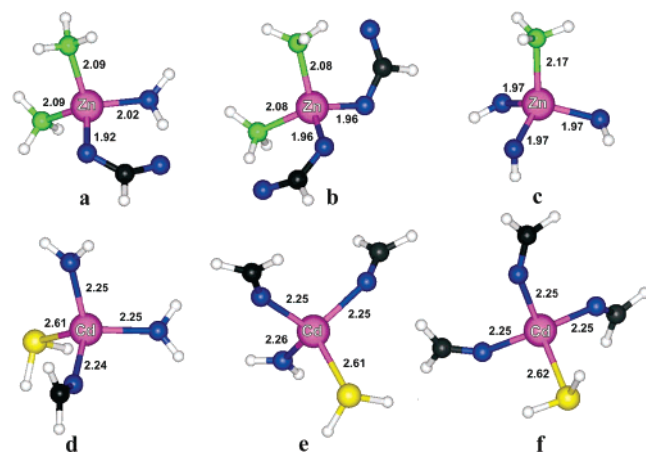


Figure 3. Most specific metal-binding sites for Zn^{2+} and Cd^{2+} (both of them preferring tetrahedral coordination): (a) $[\text{Zn}(\text{NH}_3)_2(\text{H}_2\text{O})(\text{HCOO})]^+$, (b) $[\text{Zn}(\text{NH}_3)_2(\text{HCOO})_2]$, (c) $[\text{Zn}(\text{OH})_3(\text{NH}_3)]^-$, (d) $[\text{Cd}(\text{H}_2\text{O})_2(\text{H}_2\text{S})(\text{H}_2\text{CO})]^{2+}$, (e) $[\text{Cd}(\text{H}_2\text{CO})_2(\text{H}_2\text{O})(\text{H}_2\text{S})]^{2+}$, and (f) $[\text{Cd}(\text{H}_2\text{CO})_3(\text{H}_2\text{S})]^{2+}$. The metal–ligand distances are in Å.

16.6 kcal mol⁻¹. It not only agrees with the previous findings but, in most cases, the deviations are even smaller than expected. As has been mentioned above, the agreement is poorer for Co^{2+} -specific sites ($\Delta\bar{E}_{\text{MAD}} = 3.9$; $\Delta\bar{E}_{\text{MaxDev}} = 16.6$ kcal mol⁻¹), because of the discrepancies discussed above. For Ni^{2+} -, Cu^{2+} -, and Zn^{2+} -specific sites, the values of $\Delta\bar{E}_{\text{MaxDev}}$ are usually below

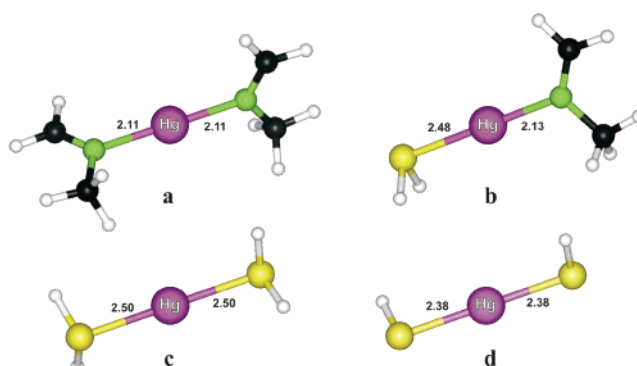


Figure 4. Most specific metal-binding sites for Hg^{2+} (preferring linear coordination): (a) $[\text{Hg}(\text{Imi})_2]^{2+}$, (b) $[\text{Hg}(\text{Imi})(\text{H}_2\text{S})]^{2+}$, (c) $[\text{Hg}(\text{H}_2\text{S})_2]^{2+}$, and (d) $[\text{Hg}(\text{SH})_2]$. The metal–ligand distances are in Å.

TABLE 7: Mean Absolute Deviations ($\Delta\bar{E}_{\text{MAD}}$) and Maximum Absolute Deviations ($\Delta\bar{E}_{\text{MaxDev}}$, in Parentheses) between Computed and Estimated Complexation Energies for Each of the 50 Studied Sites, Defined by Eqs 6a and 6b^a

site no.	Co^{2+} sites	Ni^{2+} sites	Cu^{2+} sites	Zn^{2+} sites	Cd^{2+} sites
1	4.9 (22.4)	3.8 (18.0)	2.4 (8.2)	2.9 (8.1)	0.0 (0.0)
2	4.7 (19.6)	1.0 (4.2)	1.8 (5.2)	2.0 (6.2)	0.2 (0.8)
3	4.3 (20.9)	4.4 (17.0)	2.5 (8.3)	0.0 (0.0)	1.0 (2.8)
4	4.5 (18.4)	1.7 (6.4)	1.3 (3.8)	1.9 (6.3)	1.1 (3.7)
5	1.6 (6.2)	1.7 (7.8)	3.3 (13.3)	2.4 (7.0)	0.9 (3.3)
6	5.2 (25.9)	2.0 (9.3)	3.0 (12.6)	2.3 (6.7)	0.0 (0.0)
7	4.5 (19.5)	2.1 (10.7)	1.2 (5.3)	3.2 (10.2)	0.0 (0.0)
8	2.2 (8.3)	1.4 (4.3)	3.9 (17.2)	2.1 (7.1)	1.3 (5.0)
9	4.3 (16.8)	1.9 (7.0)	2.7 (10.7)	4.3 (18.2)	1.5 (5.0)
10	2.4 (8.2)	2.4 (8.2)	3.7 (14.0)	6.1 (22.7)	2.3 (6.7)
$\bar{E}_{\text{MAD}}(\bar{E}_{\text{MaxDev}})$	3.9 (16.6)	2.2 (9.3)	2.6 (9.9)	2.7 (9.2)	0.8 (2.7)

^a All values are in kcal mol⁻¹.

10 kcal mol⁻¹. The best agreement between calculated and estimated values of complexation energies is achieved for Cd^{2+} -specific complexes ($\Delta\bar{E}_{\text{MAD}} = 0.8$; $\Delta\bar{E}_{\text{MaxDev}} = 2.7$ kcal mol⁻¹). Partially, it is a consequence of the fact that the six sites were neutral, three sites with (1-) and only one with (2-) charge of ligands. The more neutral the site is, the smaller the value of $\Delta E_{\text{coop}}(\text{M}, \text{X}_i, \text{X}_j)$ is, which probably results in more accurate estimates of $E_{\text{int}}(\text{M}, \text{X}_1, \dots, \text{X}_n)$.

We consider the overall agreement between calculated and estimated values of complexation energies as satisfying and therefore and presume that the probability of missing any specific site in the final selection of 10 candidates is low.

IIIc. “Gas-Phase Chemistry” Predictions vs Experimental Stability Constants. Because there are always doubts concerning the relevance and quality of theoretical “gas-phase chemistry” predictions, it is necessary to formulate a few concluding remarks about this series of papers.

We have tried to conceive the most specific sites for each of six studied TM ions, utilizing the results of the theoretical evaluation of two important factors: coordination geometry (including optimum metal–ligand distances) and the different affinity of these ions toward the AA side chains. Because the computational scheme has been carefully tested, the uncertainty rests mainly with the ability of the small gas-phase systems to model the “real” metal-binding sites in condensed phase. We are not yet ready to ultimately answer this question, but rather, we give the reader several examples of the agreement between calculated data^{16,19} and the experimental observations. First, the preference for the particular coordination geometry has been theoretically reproduced and quantified, using the values of complexation energies of studied metal ions in $[\text{M}(\text{H}_2\text{O})_n]^{2+}$

complexes. Second, the calculated interaction energies obeyed well-known IW series of stability constants. Third, the average affinities of AA residues correlated with the abundance of these residues in metal-binding sites of metalloproteins (implying that the process of metal uptake is thermodynamically driven). Fourth (only theoretical results), the invariance of the relative complexation energies with the growing size of model functional groups representing a metalloprotein site has been established. Finally, and most importantly, the correct order of stability constants for six studied TM ions in metal-binding sites of two metalloproteins has been reproduced by the calculated values of complexation energies (with one exception of Co^{2+} ion).

The calculations could be further refined to take into account the external factors, such as the effect of the surrounding environment or solvation of the site. Unfortunately, we think that theoretical chemistry is not able to address them with the accuracy comparable to gas-phase calculations of complexation energies, and the presented results stand at the today's limits of computational chemistry. We believe that by applying the relative scale, choosing a reasonable reference states and studying the ions with the same charge, many important physical features of the target metal-binding sites are captured by gas-phase calculations. This is in agreement with our previous findings concerning the elongation of the functional groups modeling AA side chains¹⁶ or the enlargement of the whole binding site¹⁹ and with the findings of other authors (see comparisons of the enthalpies and Gibbs energies in different environments in papers of Dudev and Lim^{6k,10}).

It must be stressed, however, that this work presents only the theoretical results, and their careful usage is highly recommended. Nineteen combinations of simple functional groups representing amino acid side chains have been devised (3–4 for each metal ion). If the value of complexation energy is the factor playing a key role in determining the specificity of the site, then these sites should be considered as the potential candidates for the experimental testing (introducing them into proteins by iterative redesign, or merging the isolated amino acid side chains ligated at the vertexes of a given coordination polyhedron into a single peptide chains). On the other hand, if the major reorganization of the coordination shell (including the transition in the coordination geometry) or a change of the metal ion oxidation states occurs during the enzymatic reaction, the data presented here can be of only limited value.

IV. Conclusions

The presented work is concluding the series of papers dealing with the specificity of the binding of selected TM ions to amino acid side chains. Exploiting the previously published results in a straightforward way, we have selected 10 candidates for the most specific site for each of studied TM ions (Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+}). The criterion for the selection of the given site were the selectivity factors, defined as to satisfy the intuitive idea of the site specificity as the maximum strength of interaction that is possible to achieve by employing the best of all combinations of ligands. At the first step, selectivity factors were defined using the approximate interaction energies. The selected candidates were then subject to quantum chemical calculations of the interaction (complexation) energies, and refined selectivity factors served as the criterion for the final selection of three or four most selective sites for each ion. The sites are graphically depicted at Figures 1–4 together with the optimum distances that determine the optimum size of the cavity formed by (hypothetical) complex ligands, such as peptides or proteins. It can be mentioned that the construction of the

fragments linking the calculated positions of AA side chains in most specific sites is in progress.

Because the calculated data were obtained by employing the “gas-phase model chemistry”, a part of discussion has been devoted to the relation between the calculated complexation energies and stability constants, determining a metal ion uptake in the experiments. There, the limitations of the adopted theoretical model are described.

Extrapolating the current growth of computational power, it might be possible to repeat the presented work (using an automatic generation of the input files, because of the total number of approximately 40 000 chemically unique complexes to be calculated) with the current quantum chemical methodology in 10 or 15 years. Then, the nonzero probability of missing any combination in the preselection process will vanish and the efforts can be concentrated upon the factors neglected here (mainly environmental effects).

Acknowledgment. This work was supported by Project LN00A032 (Center for Complex Molecular Systems and Biomolecules) and Grants 203/01/0832 (GA CR), A4055103/01 (GA AV CR), and Z4055905. CPU time on Origin 2000 at MU Brno (Project MetaCentre) is gratefully acknowledged.

Supporting Information Available: The geometries of all the studied complexes (totally 300 systems) optimized according to the described procedures at the B3LYP/BS1 level and their molecular energies in hartrees computed at the B3LYP/BS2 level (for both $[\text{MX}_n]^{2+}$ and $[\text{BqMX}_n]^{2+}$ systems) are available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Fraústo da Silva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of the Elements*; Clarendon Press: Oxford, U.K., 1993.
- (2) *Handbook of Metal–Ligand Interactions in Biological Fluids*; Berthon, G., Ed.; Marcel Dekker: New York, 1995; Vol. 1.
- (3) (a) Ghosh, P.; Shabat, D.; Kumar, S.; Sinha, S. C.; Grynszpan, F.; Li, J.; Noodleman, L.; Keinan, E. *Nature* **1996**, *382*, 339. (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) Ryde, U. *Biophys. J.* **1999**, *77*, 2777. (f) Noy, D.; Yerushalmi, R.; Brumfeld, V.; Ashur, I.; Scheer, H.; Baldrige, K. K.; Scherz, A. *J. Am. Chem. Soc.* **2000**, *122*, 3937. (g) Lovell, T.; Li, J.; Liu, T.; Case, D. A.; Noodleman, L. *J. Am. Chem. Soc.* **2001**, *123*, 12392. (h) Olsson, M. H. M.; Ryde, U. *J. Am. Chem. Soc.* **2001**, *123*, 7866. (i) Davis, M. I.; Orville, A. M.; Neese, F.; Zaleski, J. M.; Lipscomb, J. D.; Solomon, E. I. *J. Am. Chem. Soc.* **2002**, *124*, 602.
- (4) (a) Garmer, D. R. *J. Phys. Chem. B* **1997**, *101*, 2945. (b) Li, J.; Fisher, C. L.; Konecny, R.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1999**, *38*, 929. (c) Konecny, R.; Li, J.; Fisher, C. L.; Dillet, V.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1999**, *38*, 940. (d) Siegbahn, P. E. M.; Crabtree, R. H. *J. Am. Chem. Soc.* **1999**, *121*, 117. (e) Siegbahn, P. E. M. *Inorg. Chem.* **2000**, *39*, 2923. (f) Himio, F.; Eriksson, L. A.; Maseras, F.; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **2000**, *122*, 8031. (g) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2001**, *6*, 27.
- (5) (a) Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1999**, *103*, 2528. (b) Šponer, J.; Sabat, M.; Gorb, L.; Leszczynski, J.; Lippert, B.; Hobza, P. *J. Phys. Chem. B* **2000**, *104*, 7535.
- (6) (a) Bray, M. R.; Deeth, R. J.; Paget, V. J.; Sheen, P. D. *Int. J. Quantum Chem.* **1996**, *61*, 85. (b) Li, J.; Fisher, C. L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35*, 4694. (c) Hoyau, S.; Ohanessian, G. *J. Am. Chem. Soc.* **1997**, *119*, 2016. (d) Taketsugu, T.; Gordon, M. S. *J. Chem. Phys.* **1997**, *106*, 8504. (e) Niu, S.; Hall, M. B. *J. Phys. Chem. A* **1997**, *101*, 1360. (f) Barone, V.; Adamo, C. *Int. J. Quantum Chem.* **1997**, *61*, 443. (g) Irgoras, A.; Fowler, J. E.; Ugalde, J. M. *J. Am. Chem. Soc.* **1999**, *121*, 8549. (h) de Bruin, T. J. M.; Marcelis, A. T. M.; Zuilhof, H.; Sudholter, E. J. R. *Phys. Chem. Chem. Phys.* **1999**, *1*, 4157. (i) Dudev, T.; Lim, C. *J. Phys. Chem. A* **1999**, *103*, 8093. (j) Bärtsch, S.; Schröder, D.; Schwarz, H. *Helv. Chim. Acta* **2000**, *83*, 827. (k) Dudev, T.; Lim, C. *J. Phys. Chem. B* **2001**, *105*, 10709.
- (7) (a) Rogalewicz, F.; Hoppilliard, Y.; Ohanessian, G. *Int. J. Mass. Spectrom.* **2001**, *206*, 45. (b) Bluhm, B. K.; Shields, S. J.; Bayse, C. A.;

- Hall, M. B.; Russell, D. H. *Int. J. Mass. Spectrom.* **2001**, *204*, 31. (c) Landry-Hum, J.; Bussiere, G.; Daniel, C.; Reber, C. *Inorg. Chem.* **2001**, *40*, 2595.
- (8) (a) Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221. (b) Niu, S.; Hall, M. B. *Chem. Rev.* **2000**, *100*, 353. (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.
- (9) Vrettos, J. S.; Stone, D. A.; Brudvig, G. W. *Biochemistry* **2001**, *40*, 7937.
- (10) (a) Dudev, T.; Cowan, J. A.; Lim, C. J. *Am. Chem. Soc.* **1999**, *121*, 7665. (b) Dudev, T.; Lim, C. J. *Am. Chem. Soc.* **2000**, *122*, 11146. (c) Dudev, T.; Lim, C. J. *J. Phys. Chem. B* **2001**, *105*, 4446.
- (11) Garmer, D. R.; Gresh, N.; Roques, B.-P. *Proteins* **1998**, *31*, 42.
- (12) Rulíšek, L.; Vondrášek, J. J. *Inorg. Biochem.* **1998**, *71*, 115.
- (13) Glusker, J. P. *Adv. Protein Chem.* **1991**, *42*, 1.
- (14) Rulíšek, L.; Havlas, Z. *J. Phys. Chem. A* **1999**, *103*, 1634.
- (15) Rulíšek, L.; Havlas, Z. *J. Chem. Phys.* **2000**, *112*, 149.
- (16) Rulíšek, L.; Havlas, Z. *J. Am. Chem. Soc.* **2000**, *122*, 10428.
- (17) (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.
- (18) (a) Irving, H.; Williams, R. J. P. *Nature* **1948**, *162*, 746. (b) Sigel, H.; McCormick, D. B. *Acc. Chem. Res.* **1970**, *3*, 201.
- (19) Rulíšek, L.; Havlas, Z. *J. Phys. Chem. A* **2002**, *106*, 3855.
- (20) Lindskog, S.; Nyman, P. O. *Biochim. Biophys. Acta* **1964**, *85*, 462.
- (21) Coleman, J. E.; Vallee, B. L. *J. Biol. Chem.* **1961**, *236*, 2244.
- (22) (a) Kotrba, P.; Pospíšil, P.; de Lorenzo, V.; Ruml, T. *J. Recept. Signal Transduct. Res.* **1999**, *19*, 703. (b) Kotrba, P.; Dolečková, L.; de Lorenzo, V.; Ruml, T. *Appl. Environ. Microbiol.* **1999**, *65*, 1092. (c) Sousa, C.; Kotrba, P.; Ruml, T.; Cebolla, A.; de Lorenzo, V. *J. Bacteriol.* **1998**, *180*, 2280. (d) Sousa, C.; Cebolla, A.; de Lorenzo, V. *Nature Biotechnol.* **1996**, *14*, 1017.
- (23) (a) Brown, S. *Nature Biotechnol.* **1997**, *15*, 269. (b) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82. (c) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Natri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779. (d) Kostal, J.; Mulchandani, A.; Chem, W. *Macromolecules* **2001**, *34*, 2257.
- (24) (a) Ramanathan, S.; Ensor, M.; Daunert, S. *Trends Biotechnol.* **1997**, *15*, 500. (b) Hellinga, H. W.; Marvin, J. S. *Trends Biotechnol.* **1998**, *16*, 183. (c) Bontidean, I.; Kumar, A.; Csoregi, E.; Galaev, I. Y.; Mattiasson, B. *Angew Chem. Int. Ed. Engl.* **2001**, *40*, 2676.
- (25) (a) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1999**, *121*, 4952. (b) Chen, X.; Discher, B. M.; Pilloud, D. L.; Gibney, B. R.; Moser, C. C.; Dutton, P. L. *J. Phys. Chem. B* **2002**, *106*, 617.
- (26) *Handbook of Metal-Ligand Interactions in Biological Fluids*; Berthon, G., Ed.; Marcel Dekker: New York, 1995; Vol. 2.
- (27) 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. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.6; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (28) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (29) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (c) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
- (30) (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.
- (31) (a) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639. (b) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
- (32) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612.
- (33) In fact, only 270 complexes were computed, because three sites (18 complexes) contained only two "nonwater" functional groups, and therefore, they have been already studied in the previous work (Part 2); two pairs of sites (no. 10 in Co²⁺-specific with no. 10 in Ni²⁺-specific and no. 6 in Zn²⁺-specific with no. 10 in Cd²⁺-specific sites) have been identical, i.e., potential candidates for both TM ions.
- (34) Bertini, I.; Briganti, F.; Scozzafava, A. *Handbook of Metal-Ligand Interactions in Biological Fluids*; Berthon, G., Ed.; Marcel Dekker: New York, 1995; Vol. 1, p 81.