

# Theoretical Model for the Prediction of the Stability of $\text{Co}^{2+}$ , $\text{Ni}^{2+}$ , $\text{Cu}^{2+}$ , $\text{Zn}^{2+}$ , and $\text{Cd}^{2+}$ mono-complexes with monocarboxylic acids based on ${}^3\chi^v$ connectivity index

Ante Miličević\* and Nenad Raos

Institute for Medical Research and Occupational Health, Ksaverska c. 2, HR-10001 Zagreb, Croatia

\* Corresponding author: E-mail: antem@imi.hr

Received: 30-07-2012

## Abstract

The quadratic model for the prediction of stability constants of transition metal ( $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$ ) complexes with four monocarboxylic amino acids (methanoic, ethanoic, propanoic, and butanoic) was developed. The model yielded regression coefficient  $r = 0.996$ , and standard error S.E. = 0.05 ( $N = 20$ ). As a test of goodness of the model, we predicted  $\log K_1$  of three  $\text{Co}^{2+}$  complexes (with methanoic, propanoic and butanoic acid) from the training set ( $N = 17$ ), consisted of the constants of other metal complexes and  $\log K_1$  of  $[\text{CoAc}]^+$ . The model yielded predictions with the S.E. = 0.08.

**Keywords:** Stability constants, Transition metals, Carboxylate complexes, Topological indices

## 1. Introduction

The principal difficulty in modeling of the stability constants of coordination compounds using connectivity indices<sup>1,2</sup> is the lack of experimental data. Stability constants were determined by different methods, at different experimental conditions (temperature, ionic strength, composition of water-organic mixtures, etc.), and even if it is possible to assemble a homogenous data set, the number of constants is frequently not enough for a proper statistical verification.

Researchers usually determine stability constants of one metal with many ligands, or many metals with a few ligands. This led us to develop two kinds of models. Copper(II) chelates with amino acids, diamines and oligopeptides, for example, were suitable for the first kind of models, one metal – many ligands,<sup>3–5</sup> and the complexes of lanthanides for the second, many metals – few ligands.<sup>6</sup> Also, we proposed models (for  $\log K_1$  and  $\log \beta_2$ , respectively) for the complexes of five metals ( $\text{Cu}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Fe}^{2+}$ , and  $\text{Mn}^{2+}$ ) with four amino acids (glycine, alanine, valine, and leucine).<sup>7</sup> Three or four stability constants for particular metal are obviously too few for any regression, but despite this statistical „impossibility“ we obtained, by common regression on all the metals, fair

agreement with the experiment (S.E. = 0.05 for  $\log K_1$  and S.E. = 0.11 for  $\log \beta_2$ ).<sup>7</sup>

This paper is a continuation of our research on the complexes of monocarboxylic acids with lanthanides<sup>6</sup> and cadmium(II).<sup>8</sup> Its aim is to develop a common model for the divalent ions of five transition metals (cobalt, nickel, copper, zinc, and cadmium) with the first four acids from the homologous series (methanoic, ethanoic, propanoic, and butanoic). To secure the homogeneity of data, all the constants were extracted from the single paper,<sup>9</sup> belonging to the suite of papers on the complexes of monocarboxylato and hydroxyacidato complexes.<sup>10–14</sup> Stability constants  $K_1$  were measured at the same conditions,  $t = 25$  °C,  $I$  ( $\text{NaClO}_4$ ) = 2 mol dm<sup>-3</sup>, potentiometrically for  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$ ,<sup>9</sup> and spectrophotometrically for  $\text{Cu}^{2+}$ .<sup>10</sup>

## 2. Methods

### 2. 1. Calculation of Topological Indices

We calculated topological indices using a program system E-DRAGON, developed by R. Todeschini and coworkers, which is capable of yielding 119 topological indices in a single run, along with many other molecular descriptors.<sup>15,16</sup> Connectivity matrices were constructed

with the aid of the *Online SMILES Translator and Structure File Generator*.<sup>17</sup>

All models were developed by using the  ${}^3\chi^v$  index (the valence molecular connectivity index of the 3rd order), which was defined as:<sup>18–21</sup>

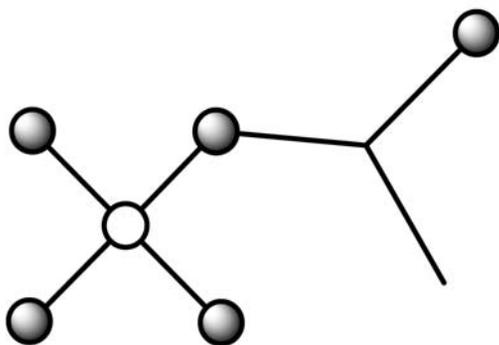
$${}^3\chi^v = \sum_{\text{path}} [\delta(i) \delta(j) \delta(k) \delta(l)]^{-0.5} \quad (1)$$

where  $\delta(i)$ ,  $\delta(j)$ ,  $\delta(k)$ , and  $\delta(l)$  are weights (valence values) of vertices (atoms)  $i$ ,  $j$ ,  $k$ , and  $l$  making up the path of length 3 (three consecutive chemical bonds) in a vertex-weighted molecular graph. The valence value,  $\delta(i)$ , of a vertex  $i$  is defined by:

$$\delta(i) = [Z'(i) - H(i)]/[Z(i) - Z'(i) - 1] \quad (2)$$

where  $Z'(i)$  is the number of valence electrons belonging to the atom corresponding to vertex  $i$ ,  $Z(i)$  is its atomic number, and  $H(i)$  is the number of hydrogen atoms attached to it. It has to be pointed out that  ${}^3\chi^v$  is only a member of the family of valence connectivity indices  ${}^n\chi^v$ , which differ between each other by the path length, *i.e.* the number of  $\delta$ 's in the summation term, Eq. 1.

The  ${}^3\chi^v$  index for all *mono*-complexes were calculated from the graph representations of *aqua*-complexes (Figure 1), assuming that metal ions are tetracoordinated.<sup>1</sup>



**Figure 1.** The graph representation for metal(II) mono-complexes with ethanoic acid. Heteroatoms are marked with ○ (metal, M) and ● (O).

## 2. 2. Regression Calculations

Regression calculations, including the leave-one-out procedure (LOO) of cross validation were done using the CROMRsel program.<sup>22</sup> The standard error of cross validation estimate is defined as:

$$\text{S.E.}_{\text{cv}} = \sqrt{\sum_i \frac{\Delta X_i^2}{N}} \quad (3)$$

where  $\Delta X$  and  $N$  denotes cv residuals and the number of reference points, respectively.

## 3. Results and Discussion

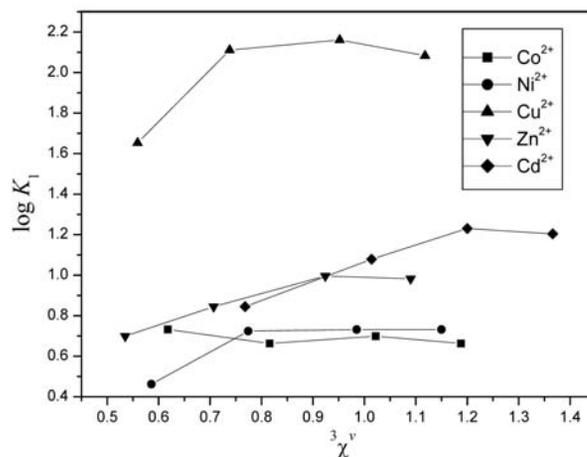
From Table 1 and Figure 2 it can be seen that  $\log K_1$  for all metals studied ( $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$ ), except for  $\text{Co}^{2+}$ , rises from methanoic to propanoic acid complexes, and drops slightly from propanoic to butanoic acid ones. Assuming the value of  $\log K_1$  for cobalt(II) formate complex measured potentiometrically, 0.732, is erroneous, we replaced it with its spectrophotometric value,<sup>10</sup>  $\log K_1 = 0.398$ , and proposed parabolic dependence of  $\log K_1$  on  ${}^3\chi^v$  for all the metals:

$$\log K_1 = a_1[{}^3\chi^v]^2 + a_2[{}^3\chi^v] + b \quad (4)$$

**Table 1.** Measured  $\log K_1$  for  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$  *mono*-complexes with four *monocarboxylic* acids;  $t = 25$  °C,  $I$  ( $\text{NaClO}_4$ ) = 2 mol  $\text{dm}^{-3}$  [Ref. 9]

Carboxylic acids	$\log K_1$ for complexes of				
	$\text{Co}^{2+}$	$\text{Ni}^{2+}$	$\text{Cu}^{2+}$	$\text{Zn}^{2+}$	$\text{Cd}^{2+}$
Methanoic	0.732 <sup>a</sup>	0.462	1.653	0.699	0.845
Ethanoic	0.663	0.724	2.111	0.845	1.079
Propanoic	0.699	0.732	2.161	0.996	1.230
Butanoic	0.663	0.732	2.083	0.982	1.204

<sup>a</sup> In Model 1 replaced by 0.398 [Ref. 9].

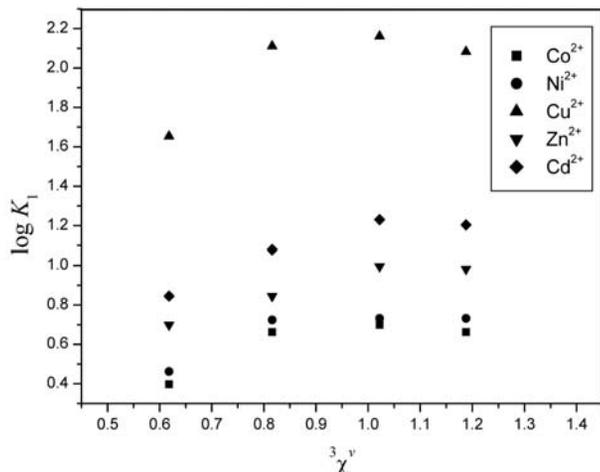


**Figure 2.** Dependence of  $\log K_1$  on  ${}^3\chi^v$  for  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$  complexes with methanoic, ethanoic, propanoic, and butanoic acid.

Because four points are too few to verify a three parameters model, and in order to develop a model for simultaneous prediction of  $\log K_1$  of different metal complexes, we developed, starting from Eq. 4, a common model for all the metals studied. It was done in two steps.

First, we used  ${}^3\chi^v$  index insensitive to a metal, *i.e.* we assumed the same  $\delta$  value (Eq. 2) for all the metals (arbitrary, we took  $\delta$  value of  $\text{Co}^{2+}$ ). Namely, from our previ-

our work<sup>7,23,24</sup> we concluded that the shape of  $\log K_1$  function is the same for all transition metals, depending only on the type of ligands (diamines, amino acids, dipeptides, etc.). In this way we obtained set of parabolas one above the other (Figure 3), with the maxima in the range 0.999–1.171.



**Figure 3.** Dependence of  $\log K_1$  on  ${}^3\chi^\nu$  with the  $\delta$  value equal for all the metals;  $\delta(M) = \delta(\text{Co}^{2+})$  for  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$  complexes with methanoic, ethanoic, propanoic, and butanoic acid.

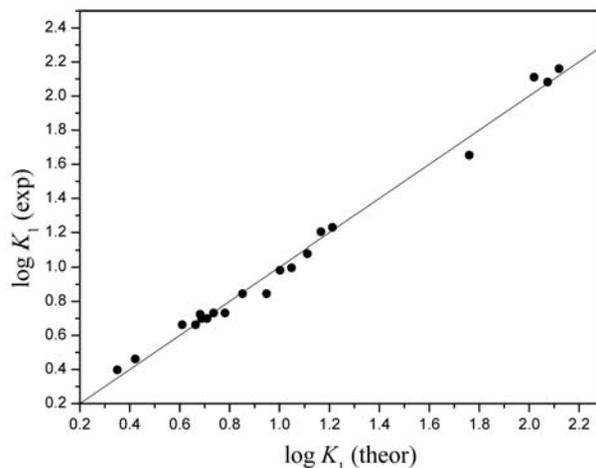
Second, in Eq. 4 we included  $\log K_1^S$ , an additional variable resembling indicator variable,<sup>25</sup> to connect parabolas along the  $\log K_1$  axis. For all the complexes of a metal  $\log K_1^S$  is a constant with the value corresponding to the  $\log K_1$  of a complex with the particular ligand, e.g. methanoic acid. For  $\log K_1^S$  one can use  $\log K_1$  of the complex of any ligand, but the ligand must be the same for every metal.

In this way we obtained model (Eq. 5) that includes carboxylic complexes of all the metals studied:

$$\log K_1 = a_1[{}^3\chi^\nu]^2 + a_2[{}^3\chi^\nu] + a_3[\log K_1^S] + b \quad (5)$$

Using Eq. 5 and  $\log K_1(\text{MForm})$  as  $\log K_1^S$ , we obtained Model 1 (Table 2). The model gave  $r = 0.996$ , S.E. = 0.05 and S.E.<sub>cv</sub> = 0.07 (Figure 4).

To test the prediction power of our model we choose cobalt(II) complexes as the test set ( $N = 3$ ), and  $\log K_1(\text{MAc})$  as  $\log K_1^S$ . The training set consisted of  $\log K_1$



**Figure 4.** Plot of experimental vs. theoretical  $\log K_1$  calculated by the Eq. 5;  $r = 0.996$ , S.E. = 0.05.

values of complexes of other metals and just one, acetato,  $\text{Co}^{2+}$  complex ( $N = 17$ ). Thus we could also verify  $\log K_1$  value of  $\text{Co}^{2+}$  formato complex. The model (Model 2, Table 2) yielded fair predictions of  $\log K_1$ ; 0.426, 0.795, 0.757 for (CoForm), (CoProp) and (CoBut), respectively. As expected, calculated value for cobalt(II)-formato complex is much closer to the value measured spectrophotometrically ( $\Delta \log K_1 = 0.028$ ) than potentiometrically ( $\delta \log K_1 = 0.306$ ). The standard error of the test set ( $N = 3$ ) is 0.08, which is in agreement with S.E. and S.E.<sub>cv</sub> obtained by Models 1 and 2 (Table 2).

## 4. Conclusion

Assuming that dependence of  $\log K_1$  on  ${}^3\chi^\nu$  is parabolic, we developed a common model for predicting of the stability of transition metal complexes with carboxylic acids. The model allows, knowing  $\log K_1$  of just one complex of a metal, to predict stability constants of other complexes of that metal. The model also proved useful in determination of dubiously measured stability constants, as was shown by Co(Form).

In our previous paper on cadmium(II) complexes with monocarboxylic and hydroxycarboxylic acids,<sup>8</sup> we obtained, on a different set of experimental data,<sup>26</sup> a positive linear correlation of  $\log K_1$  on  ${}^3\chi^\nu$  ( $N = 9$ ), but also parabolic correlation if only complexes studied in this paper were taken into account ( $N = 4$ ). The linear correlation in that paper (S.E. = 0.05, S.E.<sub>cv</sub> = 0.06) has been of the

**Table 2.** Regression models for the estimation of the  $\log K_1$  of  $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$  mono-complexes with four monocarboxylic acids

Model	N	Regression coefficients				r	S.E.	S.E. <sub>cv</sub>
		a <sub>1</sub> (S.E.)	a <sub>2</sub> (S.E.)	a <sub>3</sub> (S.E.)	b(S.E.)			
1	20	-2.05(35)	4.24(64)	1.124(28)	-1.94(28)	0.996	0.05	0.07
2	17	-1.99(32)	4.18(88)	0.932(32)	-2.02(39)	0.993	0.06	0.09

same quality as the parabolic correlation in this paper (see Table 2), therefore the linear model cannot be assumed as an approximation of the parabolic. At any rate, linear and parabolic functions were found to be equally suitable for the correlation of stability constants on  ${}^3\chi^v$ ,<sup>1,3–6</sup> but, of course, not so for the same set of complexes.

Tentatively, and more fundamentally, the problem of parabolic vs. linear function could be related to chemical action concept and consequently to Density Functional Theory (DFT).<sup>27</sup> Namely, electronegativity and chemical hardness may be defined as the first and second derivative of the energy of the system, respectively, assuming parabolic form of the energy function. They could in turn be connected with polarizability (electronegativity), and steric effect (chemical hardness), leading to an assumption that the stability of coordination compounds is determined by electronic and steric factors. Unfortunately,  ${}^3\chi^v$  does not discriminate these factors, but their influence is reflected on the form of the regression function. Interesting to note, this way of thinking should be traced in other applications of parabolic functions in QSPR and QSAR analysis, from Hansch pioneering work<sup>28</sup> till nowadays.<sup>29–31</sup>

## 5. Acknowledgement

This work was supported by Croatian Ministry of Science, Technology, Education and Sport (project 022-1770495-2901).

## 6. References

1. N. Raos, A. Miličević, *Arh. Hig. Rada Toksikol.* **2009**, *60*, 123–128.
2. N. Raos, A. Miličević, in A. D. Morrison (Ed.): *Handbook of Inorganic Chemistry Research*, Nova Publishers, New York, **2010**, pp. 267–294.
3. A. Miličević, N. Raos, *J. Phys. Chem. A* **2008**, *112*, 7745–7749.
4. A. Miličević, N. Raos, *Chin. J. Chem.* **2011**, *29*, 2421–2427.
5. A. Miličević, N. Raos, *Croat. Chem. Acta* **2009**, *82*, 633–639.
6. A. Miličević, N. Raos, *Chem. Phys. Lett.* **2012**, *528*, 63–67.
7. A. Miličević, G. Branica, N. Raos, *Molecules*, **2011**, *16*, 1103–1112.
8. A. Miličević, N. Raos, *Acta Chim. Slov.* **2012**, *59*, 194–198.
9. I. Filipović, T. Matusinović, B. Meyer, I. Piljac, B. Bach-Druginović, A. Bujak, *Croat. Chem. Acta* **1970**, *42*, 541–549.
10. B. Grabarić, I. Filipović, *Croat. Chem. Acta* **1970**, *42*, 479–492.
11. I. Filipović, I. Piljac, A. Medved, S. Savić, A. Bujak, B. Bach-Druginović, B. Meyer, *Croat. Chem. Acta* **1968**, *40*, 131–142.
12. J. Savić, M. Savić, I. Filipović, *Croat. Chem. Acta* **1972**, *44*, 305–315.
13. S. Nushi, I. Piljac, B. Grabarić, I. Filipović, *Croat. Chem. Acta* **1973**, *45*, 453–456.
14. I. Filipović, I. Piljac, B. Bach-Druginović, I. Kruhac, B. Grabarić, *Croat. Chem. Acta* **1973**, *45*, 447–452.
15. I. V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk, V. V. Prokopenko, *J. Comput. Aid. Mol. Des.* **2005**, *19*, 453–463.
16. VCCLAB, Virtual Computational Chemistry Laboratory, <http://www.vcclab.org>, 2005.
17. <http://cactus.nci.nih.gov/services/translate/>
18. L. B. Kier, L. H. Hall, *J. Pharm. Sci.* **1976**, *65*, 1806–1809.
19. L. B. Kier, L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Wiley, New York, **1986**.
20. L. B. Kier, L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, **1976**.
21. M. Randić, *MATCH Commun. Math. Comput. Chem.* **2008**, *59*, 5–124.
22. B. Lučić, N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 121–132.
23. A. Miličević, N. Raos, *Croat. Chem. Acta* **2005**, *78*, 525–534.
24. A. Miličević, N. Raos, *Polyhedron* **2008**, *27*, 887–892.
25. S. M. Free Jr., J. W. Wilson, *J. Med. Chem.* **1964**, *7*, 395–399.
26. H. Matsui, *Complex Formation of Cadmium(II) Ions*, Government Industrial Research Institute, Nagoya, **1985**.
27. M. V. Putz, *MATCH Commun. Math. Comput. Chem.* **2011**, *66*, 35–63.
28. C. Hansch, *Acc. Chem. Res.* **1969**, *2*, 232–239.
29. L. Tarko, M. V. Putz, *J. Math. Chem.* **2010**, *47*, 487–495.
30. M. Selegean, M. V. Putz, T. Rugea, *Int. J. Mol. Sci.* **2009**, *10*, 3616–3634.
31. M. V. Putz, M. Lazea, A.-M. Putz, C. Duda-Seiman, *Int. J. Mol. Sci.* **2011**, *12*, 9533–9569.

## Povzetek

Razvili smo kvadratni model za napoved konstant stabilnosti za komplekse kovin prehoda ( $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Cd}^{2+}$ ) s štirimi monokarbonsilnimi amino kislinami (metanojsko, etanojsko, propanojsko in butanojsko). Koeficient regresije uporabljenega modela znaša,  $r = 0.996$ , in standardna napaka,  $S.E. = 0.05$  ( $N = 20$ ). Model smo uspešno preverili s testnim izračunom konstant stabilnosti kompleksov  $\text{Co}^{2+}$  z metanojsko, propanojsko in butanojsko kislino