CROATICA CHEMICA ACTA
CCACAA, ISSN 0011-1643, e-ISSN 1334-417X
Croat. Chem. Acta **82** (3) (2009) 633–639.
CCA-3356

Original Scientific Paper

Estimation of Stability Constants with Connectivity Index: Development of Bivariate and Multivariate Linear Models for Copper(II) Chelates with Oligopeptides

Ante Miličević and Nenad Raos*

Institute for Medical Research and Occupational Health, Ksaverska c. 2, P. O. Box 291, HR-10000 Zagreb, Croatia

RECEIVED FEBRUARY 8, 2008; REVISED AUGUST 7, 2008; ACCEPTED AUGUST 11, 2008

Abstract. The equilibrium constants K_1, K_a and β_1 of 25 copper(II) complexes with di- ($N = 15$), tri- ($N = 5$), tetra- ($N = 2$), and pentapeptides ($N = 3$) were estimated by using linear models based on the valence connectivity index of the 3rd order ($^3\chi^v$). For the stability constant K_1 two kinds of models were developed: bivariate models, which divide a molecule into *N*- and *C*-terminal segments, and multivariate models, which treat the connectivity index of each chelate ring as a separate variable. The models proved equally successful, yielding fair reproduction of experimental log K_1 ($S.E._{cv} = 0.16 – 0.19$, max. error = 0.4). Overall deprotonation constant K_a was estimated from the $^3\chi^v$, calculated for the whole molecule of the complex, with the addition of two indicator variables to distinguish various classes of ligands. In this way, the overall stability constant ($\log \beta_1 = \log K_1 – pK_a$) was reproduced with the error < 0.81 (S.E. = 0.31).

Keywords: topological indices, linear regression models, indicator variable, dipeptides, tripeptides, tetrapeptides, pentapeptides

INTRODUCTION

The application of topological indices to estimate the stability constants of coordination compounds is a novel approach, despite a very wide range of application of these indices[1–3] in all fields of chemistry, starting from the modelling of physicochemical parameters, QSPR,[4–6] and biological activities, QSAR,[7,8] to the description of aromatic systems.[9,10] Models using topological indices are very simple, both conceptually and computationally, and perhaps the first point makes a chemist suspicious, for it seems illusory to properly estimate such a complex property as the stability constants of coordination compounds[11] by such simple means.

In our laboratory, we initially applied topological indices on copper(II) complexes with *N*-alkylated glycines[12] as a part of our studies of copper(II) complexes with *N*-alkylated and *N,N*-dialkylated amino acids.[13–17] Later, we extended their use to copper(II) chelates with α-amino acids and their mixed complexes.[18,19] We also applied our models on copper(II) and nickel(II) chelates with fructose-amino acids,[20] copper(II) chelates with diamines and triamines,[21] and, in the last report, on copper(II) chelates with dipeptides.[22] We also succeeded, by applying topological indices, in estimating the stability of one class of compounds by using models developed for another.[20,21] It turned out that the best results were obtained by using the valence connectivity index of the 3rd order ($^3\chi^v$).[18,19] Moreover, the results obtained with the models based on this connectivity index were in no way inferior to the results obtained by the more demanding overlapping spheres method.[21,23–25] Our models based on $^3\chi^v$ index reproduced logarithm of experimental constants usually with an error of less than 0.3.

In our previous report[22] we developed models for copper(II) complexes with dipeptides. It turned out that a fair estimation of overall stability constant (β_1) could be obtained if it was divided into its constituting constants K_1 and K_a.[26,27] Namely, equilibrium constant of reaction (where L^- denotes ligand, *e.g.* peptide, and M stays for central, metal atom),

$$M^{2+} + L^- \rightleftarrows ML^+ \qquad (1)$$

is denoted by K_1. The bonded ligand is further deprotonated in one (dipeptides), two (tripeptiedes) or three steps (higher peptides),

* Author to whom correspondence should be addressed. (E-mail: raos@imi.hr)

$$ML^{+} \rightleftarrows MLH_{-1} + H^{+} \quad (2)$$

$$MLH_{-1} \rightleftarrows MLH_{-2}^{-} + H^{+} \quad (3)$$

$$MLH_{-2} \rightleftarrows MLH_{-3}^{2-} + H^{+} \quad (4)$$

with constants K_{a1} (Eq. 2), K_{a2} (Eq. 3), and K_{a3} (Eq. 4). Obviously, $\beta_1 = K_1 K_a$, where K_a is the overall deprotonation constant, $K_a = \prod_i K_{ai}$.

The constants K_1 and K_a were separately estimated. Moreover, in our last report[22] K_1 was calculated by partitioning the initial sets into subsets according to structural similarity or, in another way, by using a bivariate function to separate the influence of two chelate rings of the complex.

The aim of this study was to develop further the proposed bivariate model applying it on copper(II) chelates with tripeptides, tetrapeptides, and pentapeptides. For the sake of comparison, we choose peptides consisted of the same naturally occurring amino acids as the set of dipeptides used previously.[22]

EXPERIMENTAL

Calculation of Topological Indices

Topological indices were calculated with a program system DRAGON 2.1, written by Todeschini and coworkers,[28] which is capable of yielding 262 topological indices in a single run, along with many other molecular descriptors. The connectivity matrix[29] was constructed with the aid of Online SMILES Translator and Structure File Generator.[30]

All calculations were carried out with the connectivity index ${}^3\chi^v$ (the valence molecular connectivity index of the 3rd order):[7,31–35]

$${}^3\chi^v = \sum_{\text{path}} \left[\delta(i)\delta(j)\delta(k)\delta(l)\right]^{-0.5} \quad (5)$$

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. Valence value, $\delta(i)$, of a vertex i is defined by:

$$\delta(i) = \left[Z^v(i) - H(i)\right] / \left[Z(i) - Z^v(i) - 1\right] \quad (6)$$

where $Z^v(i)$ 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. For instance, δ values for primary, secondary, tertiary, and quaternary carbon atoms are 1, 2, 3, and 4, respectively; for oxygen in OH group it is 5, and for NH_2 group $\delta(N) = 3$. 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 in the path length, i.e. in the number of δ's in the summation term, Eq. (5).

All the calculations were done on the valence connectivity graph derived from the formula of MLH_{-1} (dipeptides), MLH_{-2}^{-} (tripeptides), or MLH_{-3}^{2-} (tetrapeptides and pentapeptides), Figure 1. Details of calculations are given in Supplement.

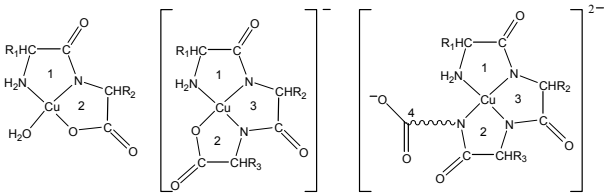

Figure 1. Structure of Cu^{II} complexes with dipeptides (MLH_{-1}), tripeptides (MLH_{-2}^{-}), and higher peptides (MLH_{-3}^{2-}). Numbers 1, 2, and 3 denote the i-th chelate ring, and number 4 denotes the C-terminal residue.

Regression Calculations

Regression calculations, including the leave-one-out (LOO) procedure of cross validation, cv, were done using the CROMRsel program.[36] The standard errors of estimate is defined as:

$$\text{S.E.} = \sqrt{\sum_i \frac{\Delta X_i^2}{N}} \quad (7)$$

where ΔX and N denotes residuals and number of reference points, respectively. The same formula was used for the calculation of standard error of cross validation estimate, S.E._{cv}, where ΔX denotes cv residuals.

Regression Functions

In this study we used two kinds of regression functions. The first one, applied previously for dipeptide complexes,[22]

$$y = b_1 + \sum a_i x_i \qquad i = 1, N \quad (8)$$

is a multiple linear function with regression parameters a_i and b_1. Y stands for log K_1, x_i for the connectivity index ${}^3\chi^v$ of the i-th chelate ring (Figure 1), and N denotes the number of chelate rings (amino-acid residues) in the molecule. (For tetrapeptides and pentapeptides

N = 4, and x_4 is $^3\chi^v$ index of the terminal residue(s) of the peptide chain.)

The second function, with regression parameters b_2, c_1, and c_2,

$$y = b_2 + c_1(x_1 + x_3) + c_2(x_2 + x_3 + x_4) \quad (9)$$

is introduced here for the first time. The variables y and x have the same meaning as in function (8). Obviously, x_4 = 0 for tripeptides, and $x_3 = x_4 = 0$ for dipeptides, in which case Eq. (9) is reduced to Eq. (8). Essentially, function (9) treats all peptide complexes as dipeptide complexes. Namely, x_1 and x_2 belong to terminal rings in all the complexes "modified" by an additional ring (x_3), and chain (x_4), Figure 1.

Function (9) has two advantages over function (8). First, it needs less regression parameters, and, second, it enables estimation of stability constants of peptide types not included in the regression, *i.e.* tripeptides and oligopeptides from dipeptides and *vice versa*.

Obviously, functions (8) and (9) are equivalent if conditions,

$$a_1 = c_1 \quad (10)$$

$$a_2 = a_4 = c_2 \quad (11)$$

$$a_3 = c_1 + c_2 \quad (12)$$

are fulfilled.

We also used the following three-parameter function:

$$y = b_3 + d_1X_1 + d_2X_2 \quad (13)$$

where X_1 and X_2 are $^3\chi^v$ indices of the first (*N*-terminal) and the second (*C*-terminal) segment of a molecule, respectively, as defined in Eq. (9). Thus, Eq. (9) should be regarded as an approximation of Eq. (13).

RESULTS AND DISCUSSION

Estimation of Stability Constant K_1

Altogether, we tested six models for the estimation of stability constant K_1 (Table 1). Models 1 and 2 are based on Eq. (8), models 1a and 2a on Eq. (9) and models 1b and 2b on Eq. (13). These models were applied on the two sets of copper(II) chelates with peptides (Table 2). The first set consisted of 20 chelates with dipeptides and tripeptides, and the second set was identical to the first one, save for an addition of two tetrapeptide and three pentapeptide complexes (N = 25).

All the models produced results of essentially the same quality; S.E. = 0.13–0.15, $S.E._{cv}$ = 0.16–0.19. Although having less regression parameters, models 1a, 1b, 2a, and 2b, gave no worse results, either in terms of S.E. or $S.E._{cv}$; moreover, the bivariate models 2a and 2b produced better results than the tetravariate Model 2 on the same set of data. All the models fulfilled the conditions of identity, Eqs. (10) – (12), of course within the limits of S.E. A comparison of regression parameters reveals that the chelate ring No. 2 (Figure 1) has the least influence on the stability of the complex, which is in accordance with previous theoretical as well as experimental results.[22]

The comparison of the estimates (Table 2) also supports the conclusion that the differences between the models are small. The differences between the estimates were mostly less than 0.1, and only one (for GGGGA) was higher than 0.2, reaching 0.27. All the estimates gave the difference from the logarithm of experimental constant of less than 0.3; the only exception being the constants for GF, GAG, and AGGGG.

Estimation of Deprotonation Constant K_a

The deprotonation constant K_a (see Introduction) was estimated as described elsewhere,[22] *i.e.* by correlating it to the $^3\chi^v$ index of the chelate, with an addition of indicator variables to enable grouping of the complexes. However, we introduced two modifications to the original models in order to include tripeptides and oligopeptides. First, overall deprotonation constant K_a was defined as the product of the stepwise constants, Eqs. (2)–(4), and, second, a new indicator variable, In_2, was introduced to distinguish between dipeptides ($In_2 = 0$), tripeptides ($In_2 = 1$), and higher peptides ($In_2 = 2$), see Figure 2 and Table 3.

Most of the estimates for dipeptides and tripeptides in both regressions are quite acceptable, *i.e.* the error of estimate is less than 0.3 (Table 4). The worst reproduction in this group of complexes was obtained for FG (error of 0.4 and 0.74 pK units, for Models 3 and 4, respectively). However, with tetrapeptides and pentapeptides, two pK_a values (for GGGG and GGGGG) were nearly perfectly reproduced, two (AGGGG and GGGGA) were reproduced poorly (with an error of 0.36 pK_a units), and the pK_a(AGGG) was reproduced with an error equal to one log unit. However, the estimates obtained by Model 3 (Table 4) were comparable to the estimates obtained by the analogous model for dipeptides,[22] which yielded $S.E._{cv}$ = 0.25 (Model 7), and the error of mean of both models, S.E. = 0.33, is equal to the error of mean for all models referred to in previous

Table 1. Linear regressions of log K_1 on the connectivity index $^3\chi^v$ of Cu^{II} complexes with peptides (n = 2–5), *cf.* Table 2

Model No. (n)[a]	N	Variables[b]	Slope (S.E.)				Intercept (S.E.)	r	S.E.	S.E.$_{cv}$
			Var1	Var2	Var3	Var4				
1 (2,3)	20	x_1, x_2, x_3	–0.605(78)	0.143(72)	–0.374(56)		6.30(23)	0.950	0.13	0.17
1a (2,3)	20	x_1+x_3, x_2+x_3	–0.555(45)	0.185(49)			6.15(12)	0.948	0.13	0.16
1b (2,3)	20	X_1, X_2	–0.483(43)	0.244(49)			5.92(10)	0.940	0.14	0.17
2 (2,3,4,5)	25	x_1, x_2, x_3, x_4	–0.550(81)	0.167(77)	–0.349(57)	0.36(15)	6.16(24)	0.929	0.14	0.19
2a (2,3,4,5)	25	x_1+x_3, $x_2+x_3+x_4$	–0.527(47)	0.214(44)			6.04(11)	0.923	0.15	0.17
2b (2,3,4,5)	25	X_1, X_2	–0.475(41)	0.239(35)			5.911(83)	0.929	0.14	0.16

[a] n denotes number of amino-acid residues in peptide.
[b] See Eqs. (8), (9), (13)

Table 2. Cross-validation estimates of log K_1 for copper(II) complexes with peptides from the regression models given in Table 1

Peptide[a]	Experimental log K_1 [N][b]	Models					
		1	2	1a	2a	1b	2b
GG	5.51[6]	5.68	5.63	5.60	5.56	5.53	5.52
GA	5.73[3]	5.70	5.67	5.68	5.65	5.64	5.63
AG	5.46[1]	5.29	5.28	5.28	5.26	5.25	5.25
GV	5.67[3]	5.78	5.76	5.77	5.77	5.76	5.76
VG	4.87[1]	5.10	5.12	5.11	5.09	5.11	5.11
GnV	5.88[1]	5.75	5.73	5.74	5.73	5.73	5.72
GnL	5.92[1]	5.78	5.77	5.79	5.79	5.8	5.79
GF	5.59[2]	5.94	5.94	5.95	5.96	5.99	5.96
FG	4.66[1]	4.66	4.73	4.71	4.72	4.78	4.79
AA	5.37[1]	5.39	5.39	5.40	5.39	5.40	5.40
VV	5.20[1]	5.20	5.24	5.24	5.27	5.31	5.31
LL	5.24[1]	5.12	5.18	5.20	5.23	5.28	5.28
GY	5.90[1]	5.87	5.87	5.90	5.92	5.95	5.93
GL	6.00[2]	5.75	5.74	5.75	5.75	5.75	5.74
LG	5.05[2]	4.99	5.02	5.00	5.00	5.02	5.03
GGA	5.08[1]	5.22	5.22	5.19	5.25	5.12	5.12
GAG	5.18[1]	4.83	4.85	4.83	4.94	4.79	4.83
AGG	4.81[1]	4.81	4.84	4.80	4.86	4.78	4.79
AAA	4.65[1]	4.72	4.78	4.73	4.84	4.83	4.81
GGG	4.96[4]	5.17	5.15	5.10	5.14	5.00	5.00
GGGG	5.12[3]		5.24		5.24		5.28
AGGG	5.08[1]		4.93		4.94		5.05
GGGGG	5.32[1]		5.42		5.33		5.37
AGGGG	5.40[1]		5.01		5.00		5.12
GGGGA	5.35[1]		5.67		5.40		5.46
S.E.$_{cv}$		0.17	0.19	0.16	0.17	0.17	0.16

[a] Amino-acid residues were denoted with a standard one letter code with the exception of nV (norvaline) and nL (norleucine)
[b] Arithmetic mean of N experimental values, Refs. 26, 37–50; All the constants were measured at the same temperature (θ = 25 °C) and two ionic strength (I_c = 0.1 and 0.16 mol L^{-1})

Table 3. Linear regressions of pK_a on the connectivity index $^3\chi^v$ of Cu^{II} complexes with peptides (n = 2–5), *cf.* Table 4

Model No. (n)[(a)]	N	Slope (S.E.)			Intercept (S.E.)	r	S.E.	S.E.$_{cv}$
		$^3\chi^v$	In_1[(b)]	In_2[(b)]				
3 (2,3)	20	–0.80(10)	0.90(12)	9.17(20)	7.35(46)	0.999	0.18	0.22
4 (2,3,4,5)	25	–0.89(19)	1.14(22)	10.05(30)	7.55(87)	0.999	0.36	0.42

[(a)] n denotes number of amino-acid residues in peptide.
[(b)] In_1 = 0 for complexes with G or A, In_1 = 1 for complexes with any other amino acid as the second residue, In_2 = 0 for complexes with dipeptides, In_2 = 1 for complexes with tripeptides, In_2 = 2 for complexes with higher peptides.

Table 4. Cross-validation estimates of log pK_a for copper(II) complexes with peptides from the regression models given in Table 3

Peptide[(a)]	Experimental pK_a [N][(b)]	Models	
		3	4
GG	4.15[6]	4.54	4.35
GA	4.14[3]	4.06	3.86
AG	4.22[1]	4.08	3.89
GV	4.74[3]	4.69	4.75
VG	3.85[1]	3.74	3.51
GnV	4.73[1]	4.67	4.73
GnL	4.82[1]	4.41	4.44
GF	3.83[2]	4.17	4.14
FG	3.50[1]	3.10	2.76
AA	3.61[1]	3.78	3.54
VV	3.85[1]	4.06	4.03
LL	3.86[1]	3.83	3.76
GY	3.85[1]	4.08	4.04
GL	4.78[2]	4.54	4.58
LG	3.56[2]	3.71	3.45
GGA	11.99[1]	11.78	12.4
GAG	11.94[1]	11.81	12.42
AGG	11.82[1]	11.86	12.45
AAA	11.1[1]	11.19	11.81
GGG	11.99[4]	12.25	12.85
GGGG	21.69[3]		21.66
AGGG	22.1[1]		21.15
GGGGG	21.24[1]		21.28
AGGGG	21.21[1]		20.85
GGGGA	21.31[1]		20.94
S.E.$_{cv}$		0.22	0.42

[(a)] Amino-acid residues were denoted with a standard one letter code with theexception of nV (norvaline) and nL (norleucine)
[(b)] Arithmetic mean of N experimental values, Refs. 26, 37–50; All the constants were measured at the same temperature (θ = 25 °C) and two ionic strength (I_c = 0.1 and 0.16 mol L^{-1})

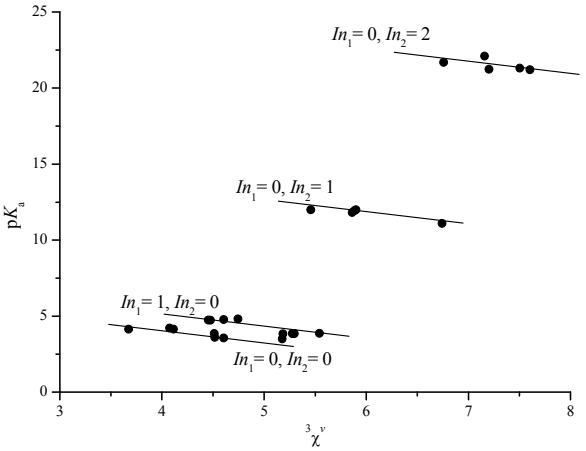

Figure 2. Regression model for pK_a of copper(II) complexes with dipeptides, tripeptides and higher peptides (Model 4, Table 3). In_1 = 0 if the second residue is G or A, In_1 = 1 if the second residue is any other amino-acid residue, In_2 = 0 for dipeptides, In_2 = 1 for tripeptides and In_2 = 2 for higher peptides.

Table 5. The S.E. values for the estimates of the log β_1 (= logK_1 – pK_a) for copper(II) complexes with peptides from the estimates presented in Tables 2 and 4

		log K_1	
		Model 2b	Mean (all models)
pK_a	Model 4	0.43	0.41
	Mean (models 3 and 4)	0.33	0.31

paper,[22] for both copper(II) and nickel(II) complexes. Discussing the results presented in Table 4, one must also bear in mind that Model 4 covers the range of altogether 18.5 pK units.

Estimation of overall stability constant β_1

The overall stability constant was reproduced choosing the most representative data from Tables 2 and 4, with an S.E. in the range 0.31–0.43 (Table 5). Obviously, the

choice of log K_1 is not as critical as the choice of pK_a, for the mean values of Models 3 and 4 gave 0.1 log β units better results in terms of *S.E.* than Model 4. The best combinations (S.E. = 0.31) yielded six (out of 25) estimates with an error higher than 0.3. Higher discrepancy between experiment and theory was recorded for AGGG (0.84), FG (0.64), and GAG (0.50). However, by discarding the complex with AGGG from the data set, S.E. drops from 0.33 to 0.28, and from 0.31 to 0.27, reaching the upper error for dipeptides (0.27).

CONCLUSION

The results presented in this report show that it is possible to estimate stability constants of copper(II) chelates with tri-, tetra-, and pentapeptides by using essentially the same models as developed for dipeptide complexes. Bivariate models proved as successful in estimating log K_1 as more complex trivariate and tetravariate models. In addition, bivariate models make it possible to estimate stability constants of one type of peptides from another type.

The results presented in this report are, generally, not so successful as those obtained for dipeptides (S.E.(log β_1) = 0.19 – 0.27),[22] but the difference should be attributed to virtually only one ligand, namely AGGG, which gave very poor estimates of pK_a (*cf.* § *Estimation of overall stability constant* β_1). This also holds true for the group consisting of tri-, tetra-, and pentapeptides (N = 10), which yielded S.E. = 0.40 (S.E. = 0.31 for all peptides), but upon discarding the complex with AGGG, it yielded S.E. = 0.32 (N = 9).

Acknowledgement. This work was supported by the Croatian Ministry of Science, Education and Sports (Project No. 022-1770495-2901).

REFERENCES

1. D. H. Rouvray, *J. Comput. Chem.* **8** (1987) 470–480.
2. M. W. Diudea, and I. Gutman, *Croat. Chem. Acta* **71** (1998) 21–51.
3. N. Trinajstić, *Chemical Graph Theory,* CRC Press, 2nd ed., Boca Raton, Florida, 1992.
4. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20.
5. P. G. Seybold, M. May, and U. A. Begal, *J. Chem. Educ.* **64** (1987) 575–581.
6. B. Lučić, A. Miličević, S. Nikolić, and N. Trinajstić, *Croat. Chem. Acta* **75** (2002) 847–868.
7. L. B. Kier and L. H. Hall, *Molecular Connectivity and Drug Design,* Academic Press, New York, 1976.
8. A. Miličević, S. Nikolić, and N. Trinajstić, *Mol. Diversity* **10** (2006) 95–99.
9. A. Graovac, I. Gutman, and N. Trinajstić, *Topological Approach to the Chemistry of Conjugated Molecules,* Lecture Notes in Chemistry, Vol. 4, Springer-Verlag, Berlin, 1977.
10. A. Miličević, S. Nikolić, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **44** (2004) 415–421.
11. R. D. Hancock and A. E. Martell, *Chem. Rev.* **89** (1989) 1875.
12. B. Grgas, S. Nikolić, N. Paulić, and N. Raos, *Croat. Chem. Acta* **72** (1999) 885–895.
13. N. Raos and V. Simeon, *Croat. Chem. Acta* **56** (1983) 79–85.
14. J. Sabolović and N. Raos, *Polyhedron* **9** (1990) 1277–1286.
15. J. Sabolović and N. Raos, *Polyhedron* **9** (1990) 2419–1427.
16. B. Kaitner, G. Pavlović, N. Paulić, and N. Raos, *J. Coord. Chem.* **43** (1998) 309–319.
17. N. Paulić and N. Raos, *J. Coord. Chem.* **31** (1994) 181–190.
18. S. Nikolić and N. Raos, *Croat. Chem. Acta* **74** (2001) 621–631.
19. A. Miličević and N. Raos, *Polyhedron* **25** (2006) 2800–2808.
20. A. Miličević and N. Raos, *Croat. Chem. Acta* **80** (2007) 557–563.
21. A. Miličević and N. Raos, *Polyhedron* **26** (2007) 3350–3356.
22. A. Miličević and N. Raos, *Polyhedron* **27** (2008) 887–892.
23. A. Miličević and N. Raos, *Arh. Hig. Rada Toksikol.* **60** (2009) 123–128.
24. A. Miličević and N. Raos, *Croat. Chem. Acta* **78** (2005) 525–534.
25. A. Miličević and N. Raos, *Croat. Chem. Acta* **79** (2006) 281–290.
26. G. Brookes and L. D. Pettit, *J. Chem. Soc., Dalton Trans.* (1975) 2106–2112.
27. S. H. Laurie, *Amino acids, peptides and proteins,* in: Comprehensive Coordination Chemistry, vol. 1, Pergamon Press, Oxford, 1987, pp. 739–776.
28. http://www.disat.unimib.it/chm/
29. D. Janežič, A. Miličević, S. Nikolić, and N. Trinajstić, *Graph-Theoretical Matrices in Chemistry,* Mathematical chemistry Monographs, No. 3, University of Kragujevac and Faculty of Science Kragujevac, 2007.
30. http://cactus.nci.nih.gov/services/translate/
31. M. Randić, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.
32. K. Altenburg, *Z. Phys. Chem.* **261** (1980) 389–393.
33. M. Randić, *J. Mol. Graph. Modell.* **20** (2001) 19–35.
34. X. Li and I. Gutman, *Mathematical Aspects of Randić-Type Molecular Structure Descriptors,* Mathematical Chemistry Series, Kragujevac, 2006.
35. L. B. Kier and L. H. Hall, *J. Pharm. Sci.* **65** (1976) 1806–1809.
36. B. Lučić and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **39** (1999) 121–132.
37. B. R. Rabin, *Trans. Faraday Soc.* **52** (1956) 1130–1136.
38. J. L. Biester and P. M.Ruoff, *J. Am. Chem. Soc.* **81** (1959) 6517–6521.
39. G. F. Bryce, J. M. H. Pinkerton, L. K. Steinrauf, and F. R. N. Gurd, *J. Biol. Chem.* **240** (1965) 3829–3836.
40. O. Yamauchi, Y. Hirano, Y. Nakao, and A. Nakahara, *Can. J. Chem.* **47** (1969) 3441–3445.
41. W. L. Koltun, M. Fried, and F. R. N. Gurd, *J. Am. Chem. Soc.* **82** (1960) 233–241.
42. R. F. Pasternack, L. Gipp, and H. Sigel, *J. Am. Chem. Soc.* **94** (1972) 8031–8038.
43. S. P. Datta, R. Leberman, and B. R. Rabin, *Trans. Faraday Soc.* **55** (1959) 2141–2151.
44. A. P. Brunetti, E. J. Burke, M. C. Lim, and G. H. Nancollas, *J. Solution. Chem.* **1** (1972) 153–165.
45. U. I. Salakhutdinov, A. O. Borisova, Yu. V. Granovskii, I. A. Savich, and V. I. Spitsyn, *Proc. Acad. Sci. (U.S.S.R.)* **177** (1967) 1039.
46. W. L. Koltun, R. H. Roth, and F. R. N. Gurd, *J. Biol. Chem.* **238** (1963) 124–131.
47. M. K. Kim and A. E. Martell, *J. Am. Chem. Soc.* **88** (1966) 914–918.
48. H. Hauer, Gary R. Dukes, and D. W. Margerum, *J. Am. Chem. Soc.* **95** (1973) 3515–3522.
49. C. R. Hartzell and F. R. N. Gurd, *J. Biol. Chem.* **244** (1969) 147–153.
50. G. F. Bryce and F. R. N. Gurd, *J. Biol. Chem.* **241** (1966) 1439–1448.

SAŽETAK

Procjena konstanti stabilnosti iz indeksa povezanosti: razvoj bivarijatnih i multivarijatnih linearnih modela za bakrove(II) kelate s oligopeptidima

Ante Miličević i Nenad Raos

Institut za medicinska istraživanja i medicinu rada,
Ksaverska c. 2, P. O. Box 291, HR-10000 Zagreb, Croatia

Konstante ravnoteže K_1, K_a i β_1 25 bakrovih(II) kompleksa s dipeptidima (N = 15), tripeptidima (N = 5), tetrapeptidima (N = 2) i pentapeptidima (N = 3) procjenjivane su modelima temeljenim na valencijskom indeksu povezanosti trećega reda ($^3\chi^v$). Za procjenu konstante stabilnosti K_1 razvijene su dvije vrste modela: bivarijatni modeli, koji dijele kompleks na *N*-terminalni i *C*-terminalni segment, te multivarijatni modeli, koji uzimaju indeks povezanosti svakog prstena kao zasebnu varijablu. Modeli su se pokazali podjednako uspješnima, te su dobro reproducirali izmjerene vrijednosti log K_1 ($S.E._{cv}$ = 0,16 – 0,19, max. pogreška = 0,4). Ukupna deprotonacijska konstanta K_a procijenjena je iz indeksa $^3\chi^v$, izračunanog za cijelu molekulu kompleksa, uz dvije dodatne indikatorske varijable za razlikovanje klasa peptida. Na taj je način ukupna konstanta stabilnosti (log β_1 = log K_1 – pK_a) procijenjena s pogreškom < 0,84 (S.E. = 0,31).