Scientific paper

Estimation of Stability Constants of Cadmium(II) *bis*-Complexes with Amino Acids by Model 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: 14-05-2010

Abstract

Linear model for estimation of the second, K_2, and overall, β_2, stability constant of cadmium(II) binary and ternary *bis*-complexes with five aliphatic α-amino acids based on valence connectivity index of the 3rd order ($^3\chi^v$) was developed. Set of amino acids included glycine, alanine, 2-aminobutanoic, 2-aminopentanoic (norvaline) and 2-aminohexanoic acid (norleucine), which by bonding to the cadmium(II) gave 25 K_2 and 15 β_2 values. For estimation of log β_2, the model gave r = 0.940, and the S.E.$_{cv}$ = 0.10, and for the two subsets of log K_2 constants the model yielded r = 0.936 and 0.842, and S.E.$_{cv}$ = 0.09. The complex CdGG was excluded from all regressions.

Keywords: Stability of coordination compounds, Theoretical models, Topological indices

1. Introduction

Cadmium is due to its high production and environmental pollution ubiquitous in the biosphere.[1,2] It is toxic metal with deleterious effects on kidney, blood and blood vessels, liver, and bones, not to mention teratogenic effects.[3] Cadmium(II) binds preferently to nucleic acids and their components,[4,5] but its toxic effects are greatly modified by binding to metallothionein.[6,7] Despite its significant biological and environmental role, not many papers were published on cadmium(II) complexes with amino acids.

From the viewpoint of coordination chemistry Cd^{2+} is very plastic cation, meaning that distorted coordination geometry is frequent for complexes of cadmium(II) which have coordination number in the range from 2 to 8.[8,9,10] Although cadmium(II) is considerably «softer» cation than Cu^{2+} and Zn^{2+}, it binds oxalate and diamine with similar affinity as Zn^{2+}, but substantially weaker than Cu^{2+}.[7,8,10] Consequently, cadmium(II) stability constants for complexes with amino acids and peptides are smaller than copper(II) constants and very close to the constants of zinc(II).[11–19] Despite poor stability, cadmium(II) complexes with amino acids show antimicrobial activity,[20] and seem to participate in cadmium(II) toxicity.[21,22]

Because of biological and environmental significance of cadmium, we decided to check our models for estimation of stability constants of coordination compounds, originally developed on copper(II) and nickel(II) chelates, on cadmium(II) complexes with amino acids. The main difficulties in applying topological indices for this purpose[25] stem from the fact that the constitutional formula (*i.e.* molecular graph) of a coordination compound is not as well defined as of an organic compound, and that graph theory can not deal with conformers and »classical« stereoisomers. Thus, the valence connectivity index of the 3rd order ($^3\chi^v$) was calculated for different molecular species[23,24] and correlated to log K_1, log K_2, and log β_2 of copper(II) and nickel(II) chelates with amino acids and their derivatives, diamines, triamines, and peptides (dipeptides to pentapeptides).[25] It was possible to obtain a good agreement between experiment and theory, and, moreover, to judge the quality of experimentally determined stability constants.[26] Fair estimates of the stability constants were obtained even from the regression functions developed on different class of compounds.[27,28] Also, by introduction of an indicator variable we succeeded to obtain a common model for estimation of the stability constants of copper(II) and nickel(II) chelates.[29]

In this paper we dealt with the estimation of the second, K_2, and overall stability constants, β_2, of cadmium(II) binary and ternary *bis*-complexes with five aliphatic α-amino acids (glycine, alanine, 2-aminobutanoic, 2-aminopentanoic (norvaline) and 2-aminohexanoic acid

(norleucine). Both set of constants, K_2 (N = 25) and β_2 (N = 15), were determined at the same experimental conditions (T = 298 K, I ($LiClO_4$) = 3 mol L^{-1}).[10,16,17]

2. Methods

2. 1. Calculation of Topological Indices

We calculated topological indices using a program system E-DRAGON, developed by R. Todeschini and coworkers,[30] which is capable of yielding 119 topological indices in a single run, along with many other molecular descriptors.[31,32] Connectivity matrices were constructed with the aid of *Online SMILES Translator and Structure File Generator*.[33]

All models were developed by using ${}^3\chi^v$ index (the valence molecular connectivity index of the 3rd order), which was defined as:[34–36]

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

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

$$ \delta(i) = [Z^v(i) - H(i)]/[Z(i) - Z^v(i) - 1] \qquad (2) $$

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 the OH group it equals 5, and for NH_2 group δ(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 between each other by the path length, *i.e.* the number of δ´s in the summation term, Eq. 1.

The ${}^3\chi^v$ indices for cadmium(II) *mono*- and *bis*-complexes were calculated from the graph representations of the *aqua* complexes with two water molecules (Fig. 1), assuming that Cd(II) in *mono*-complexes is tetracoordinated, and in *bis*-complexes hexacoordinated, as for copper(II) chelates.[23,37] This is supported by X-ray structures

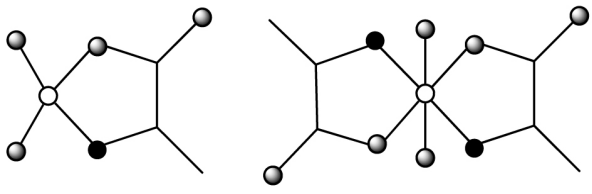

Figure 1. The graph representations for cadmium(II) *mono*- (CdL) and *bis*-complex (CdLA) with alanine. Heteroatoms are marked with ○(Cd), ●(N), and ●(O).

of cadmium(II) *bis*-complexes with glycine[38] and alanine,[39] and some mixed Cd(II) *bis*-complexes.[40,41,42] Moreover, the alternative assumption, that both, CdL and CdLA complexes are tetracoordinated yielded bad results.

2. 2. Regression Calculations

Regression calculations, including the leave-one-out procedure of cross validation, *cv*, were done using the CROMRsel program.[43] The standard error of cross validation estimate is defined as:

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

where ΔX and N denotes *cv* residuals and the number of reference points, respectively.

3. Results and Discussion

3. 1. Estimation of the Overall Stability Constant β_2

Firstly we estimated the overall stability constant, β_2:

$$ \text{M} + \text{L} + \text{A} \overset{\beta_2}{\rightleftarrows} \text{MLA} \qquad (4) $$

where M denotes Cd^{2+}, and L and A denote α-amino acids. Model was developed on 15 cadmium(II) binary and ternary *bis*-complexes with five aliphatic α-amino acids (Table 1). Beside naturally occurring glycine (G) and alanine (A), there were 2-aminobutanoic (B), 2-aminopentanoic (P) and 2-aminohexanoic acid (H) in the set.

Although at first it seemed there is no correlation between log β_2 and ${}^3\chi^v$(CdLA), after closer inspection of Fig. 2 we noticed some kind of the order. As β_2 constant is independent on sequence of ligand bonding to the metal, complexes were such named that the first ligand (L) was always smaller or equal to the second ligand (A). Thus we were able to see from Fig. 2 that stability constant β_2 of *bis*-complexes with identical second ligand (A) depends on the first ligand (L), *i.e.* log β_2 linearly decreases along the homologous series of the first amino acid (with an exception of PH in the hexanoate series). Also, for the mixed complexes with identical first ligand, log β_2 linearly increases along the homologous series of the second amino acid, with an exception of GG and GA in the glycine series. From Fig. 2 it is evident that slopes of the descending lines drop and the slopes of the ascending lines rise along the homologous series.

Subsequently, we developed a linear function based on the ascending lines:

$$ \log \beta_2 = a_1[{}^3\chi^v(\text{CdLA})] + a_2[{}^3\chi^v(\text{CdL})] + a_3[{}^3\chi^v(\text{CdA})/{}^3\chi^v(\text{CdL})] + b \qquad (5) $$

Table 1. Experimental overall, β_2, and the second, K_2, stability constants and $^3\chi^v$ indices calculated for Cd(II) aminoacidates

(Complexes (CdLA or CdAL)*	log β_2	log K_2	$^3\chi^v$(CdLA)	$^3\chi^v$(CdL)	$^3\chi^v$(CdA)
CdGG	7.49	3.48	6.933	2.436	2.436
CdGA[a]	7.47	3.46	7.292	2.436	2.921
CdGB[a]	7.34	3.33	7.496	2.436	3.125
CdGP[a]	7.41	3.40	7.656	2.436	3.285
CdGH[a]	7.43	3.42	7.925	2.436	3.554
CdAA[ab]	6.93	3.24	7.651	2.921	2.921
CdAB[a]	7.13	3.44	7.855	2.921	3.125
CdAP[a]	7.15	3.46	8.015	2.921	3.285
CdAH[a]	7.24	3.55	8.284	2.921	3.554
CdBB[ab]	6.88	3.24	8.059	3.125	3.125
CdBP[a]	7.05	3.41	8.219	3.125	3.285
CdBH[a]	7.2	3.56	8.488	3.125	3.554
CdPP[ab]	7.01	3.29	8.379	3.285	3.285
CdPH[a]	7.31	3.59	8.648	3.285	3.554
CdHH[ab]	7.03	3.29	8.917	3.554	3.554
CdAG[b]		3.78	7.292	2.436	2.921
CdBG[b]		3.70	7.496	2.436	3.125
CdBA[b]		3.49	7.855	2.921	3.125
CdPG[b]		3.69	7.656	2.436	3.285
CdPA[b]		3.43	8.015	2.921	3.285
CdPB[b]		3.33	8.219	3.125	3.285
CdHG[b]		3.69	7.925	2.436	3.554
CdHA[b]		3.50	8.284	2.921	3.554
CdHB[b]		3.46	8.488	3.125	3.554
CdHP[b]		3.57	8.648	3.285	3.554

* L always denotes smaller ligand and A the bigger one.
[a] Complexes in the set for estimation of β_2, and in the second subset for estimation of K_2.
[b] Complexes in the first subset for estimation of K_2.

The regression gave $r = 0.940$, and the S.E.$_{cv} = 0.10$ (Table 2, Fig. 3). The complex CdGG was not included into regression because it does not belong to any family of lines.

The $^3\chi^v$(CdA) index was included in the function because ascending lines show linear dependence of β_2 on the variation of the second ligand. Also, $^3\chi^v$(CdA) was divided with $^3\chi^v$(CdL) to compensate different slope of the lines, which depends on the first ligand (Fig. 2). On the other hand, the second term of Eq. 5, $^3\chi^v$(CdL), was introduced as a compensation for difference in intercepts of the ascending lines (Fig. 2). They decrease in the homologous series of the first ligand.

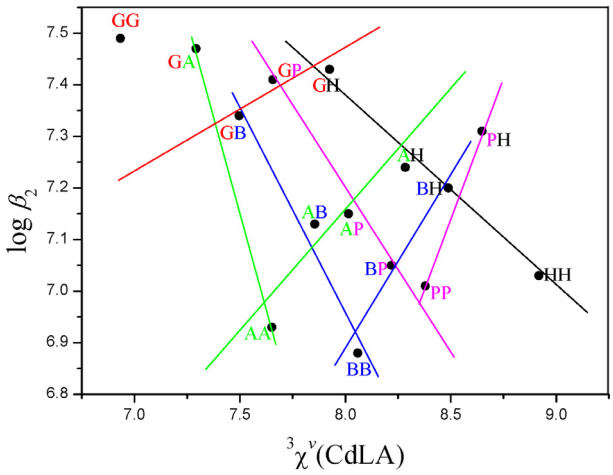

Figure 2. Plot of log β_2 *vs.* $^3\chi^v$(CdLA) reveals sequences of linear dependences. The ligand and the line that indicate linear dependence of that ligand series are coloured with the same colour.

Table 2. Linear regressions (Eq. 5) for the estimation of the overall, β_2, and second, K_2, stability constants

Regr. No.	*N*	Dependent variable	Slope (S.E.) Independent variable			Intercept (S.E.)	*r*	S.E.	S.E.$_{cv}$
			$^3\chi^v$ (MLA)	$^3\chi^v$ (ML)	$^3\chi^v$ (MA) / $^3\chi^v$ (ML)				
1	14	log β_2	3.33(89)	–6.6(18)	–8.0(24)	8.85(92)	0.940	0.06	0.10
2	14	log K_2	3.08(89)	–7.7(24)	–6.3(18)	5.71(93)	0.936	0.06	0.09
3	14	log K_2	3.50(86)	–6.8(17)	–8.9(23)	5.06(89)	0.842	0.06	0.09

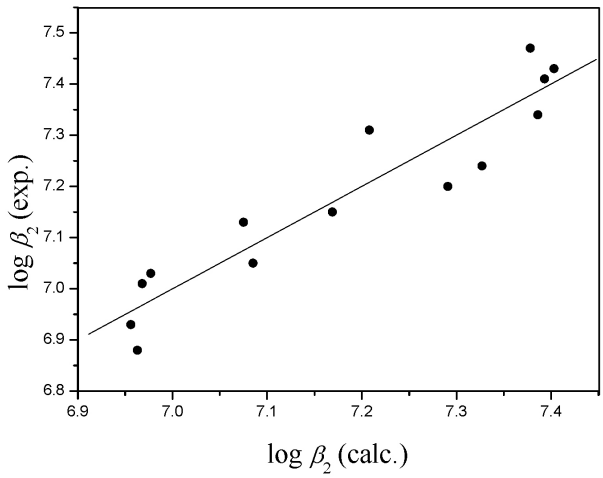

Figure 3. Experimental *vs.* calculated values of log β_2 for cadmium(II) complexes with five aliphatic α-amino acids (N = 14), Table 2, No. 1; r = 0.940, S.E.$_{cv}$ = 0.10.

3. 2. Estimation of the Second Stability Constant K_2

The values of the second stability constant, K_2:

$$\mathrm{ML(A) + A(L)} \overset{K_2}{\rightleftharpoons} \mathrm{MLA(MAL)} \qquad (6)$$

of the cadmium(II) complexes (Table 1), with the same ligands as in the case of the β_2 constant, were also estimated by the Eq. 5. However, because of the dependence of K_2 on the sequence of ligand bonding to the Cd(II), there were 25 K_2 values which we divided into two subsets (Table 1). The dependence of log K_2 on $^3\chi^v$(CdLA) (Fig. 4) shows that for every first ligand series of *bis*-complexes the binary complex has the minimum K_2 value. The exception was again the glycine series. Thus we divided the data to make the first subset consisted of complexes preceding the minimum and the second subset consisted of complexes to go after the minimum (Fig. 4). In the second subset were complexes with the first ligand smaller or equal to second, same as in the case of β_2 constant, and the opposite was true for the first subset (Table 1). For both subsets the pattern similar to log β_2 pattern (Fig. 2) was observed, although with somewhat larger deviations for the second set. Consequently, Eq. 5 was also used to estimate log K_2. The only difference was that for K_2 constant, L denoted the smaller and A the bigger ligand, regardless the order of their bonding to the cadmium(II) (Eq. 6).

For the first subset (N = 14) the regression gave r = 0.936, and S.E.$_{cv}$ = 0.09 (Table 2, Fig. 5), and for the second subset (N = 14) r = 0.842, and S.E.$_{cv}$ = 0.09

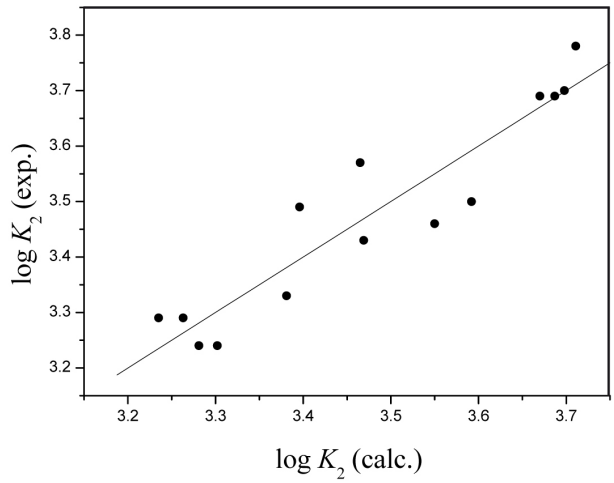

Figure 5. Experimental *vs.* calculated values of log K_2 for the first subset of cadmium(II) complexes (Table 1) with five aliphatic α-amino acids (N = 14), Table 2, No. 2; r = 0.936, S.E.$_{cv}$ = 0.09.

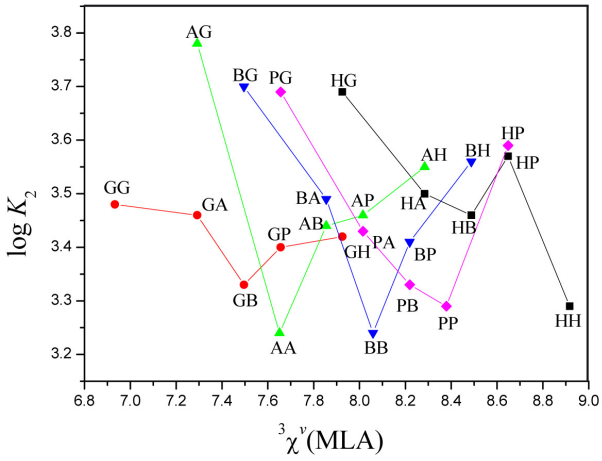

Figure 4. Plot of log K_2 *vs.* $^3\chi^v$(CdLA). The ligand and the curve that connect the complexes of that ligand series are coloured with the same colour.

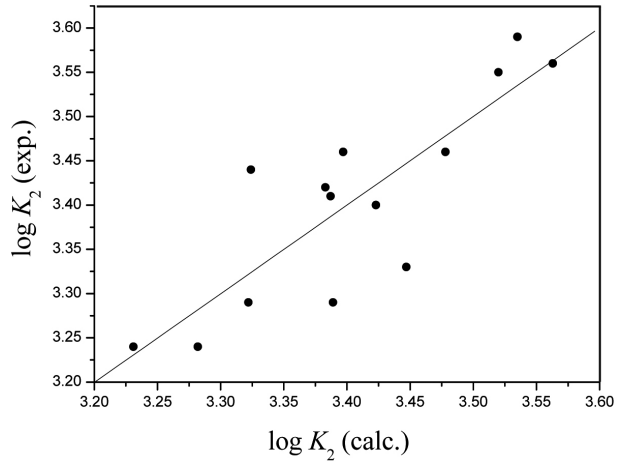

Figure 6. Experimental *vs.* calculated values of log K_2 for the second subset of cadmium(II) complexes (Table 1) with five aliphatic α-amino acids (N = 14), Table 2, No. 3; r = 0.842, S.E.$_{cv}$ = 0.09

(Table 2, Fig. 6) were yielded. From both subsets, as for the estimation of β_2, the complex CdGG was excluded because it does not belong to any series of lines.

4. Conclusion

In this paper, for the first time we developed the regression model for the estimation of stability constants of cadmium(II) complexes. The model is valid for the estimation of second, K_2, and overall, β_2, stability constant of cadmium(II) *bis*-complexes with aliphatic amino acids. It is based on the observation that stability constants of *bis*-complexes with the same smaller ligand linearly increase along the homologuos series of the bigger ligand, and that stability constants of *bis*-complexes with the same bigger ligand linearly decrease along the homologuos series of the smaller ligand.

In further research, we will try to apply this model to other metal complexes to see if it is generally valid.

5. Acknowledgement

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

6. References

1. L. Friberg, M. Piscator, G. F. Nordberg, T. Kjeldström, *Cadmium in the Environment*, 2nd Ed., CRC Press, Cleveland, 1974.
2. *Evaluation of Impact of Cadmium on the Health of Man*, Commission of the European Communities, Luxemburg, 1977.
3. L. Friberg, T. Kjellström, G. F. Nordberg, M. Piscator, Cadmium, in: L. Friberg, G. F. Nordberg, V. B. Vouk (Eds), *Handbook on the Toxicology of Metals*, Elsevier, Amsterdam, 1979, pp.355–381.
4. K.B. Jacobson, J. E. Turner, *Toxicology* **1980**, *16*, 1–37.
5. M.A. Shipman, C. Price, A.E. Gibson, M. R. J. Elsegood, W. Clegg, A. Houlton, *Chem. Eur. J.* **2000**, *6*, 4371–4378.
6. J. R. H. Kägi, M. Nordberg, *Metallothionein*, Birkhäuser Verlag, Basel, 1979.
7. E.C. Foulkes (Ed.), Biological roles of metallothionein, in: *Development in Toxicology and Environmental Sciences*, Vol. 9, Elsevier-North Holland, Amsterdam, 1982.
8. O. Andersen, *Environm. Healtth Perspect.* **1984**, *54*, 249–266.
9. D. G. Tuck, *Rev. Inorg. Chem.* **1979**, *1*, 209–243.
10. H. Matsui, *Complex Formation of Cadmium(II) Ion*s, Government Industrial Research Institute, Nagoya, Japan, 1985.
11. D. J. Perkins, *Biochem J.* **1953**, *55*, 649–652.
12. D. J. Perkins, *Biochem J.* **1954**, *57*, 702–704.
13. M. L. S. Simoes Goncalves, M. M. Correia dos Santos, *J. Electroanal. Chem.* **1985**, *187*, 333–348.
14. J. H. Smith, A. M. Cruickshank, J. T. Donoghue, J. F. Pysz, Jr., *Inorg. Chem.* **1962**, *1*, 148–150.
15. F. Khan, A. Khanam, *Eclet. Quim.* **2008**, *33*, 29–36.
16. H. Matsui, H. Ohtaki, *Bull. Chem. Soc. Jpn.* **1982**, *55*, 461–465.
17. H. Matsui, H. Ohtaki, *Bull. Chem. Soc. Jpn.* **1982**, *55*, 2131–2133.
18. P. Leverrier, C. Montigny, M. Garrigos, P. Champeil, *Anal. Biochem.* **2007**, *371*, 215–228.
19. G. Branica-Jurković, V. Simeon, *J. Electroanal. Chem.* **1989**, *266*, 83–91.
20. J. A. Khanam, M.F. Begum, J. Ara, M. Jesmin, A. Abu Taher, S.M. Moshin Ali, *Dhaka Univ. J. Pharm. Sci.* **2006**, *5*, 29–32.
21. L. Bancu, S. Stefan, I. Jitaru, *Rev. Chim.* **2006**, *57*, 582–585.
22. W. Tang, J. Xie, Z.A. Shaikh, *Toxicology* **2006**, *223*, 202–208.
23. A. Miličević, N. Raos, *Polyhedron* **2006**, *25*, 2800–2808.
24. A. Miličević, N. Raos, *Acta Chim. Slov.* **2009**, *56*, 373–378.
25. N. Raos, A. Miličević, *Arh. Hig. Rada Toksikol.* **2009**, *60*, 123–128.
26. N. Raos, G. Branica, A. Miličević, *Croat. Chem. Acta* **2008**, *81*, 511–517.
27. A. Miličević, N. Raos, *Croat. Chem. Acta* **2007**, *80*, 557–563.
28. A. Miličević, N. Raos, *Polyhedron* **2007**, *26*, 3350–3356.
29. A. Miličević, N. Raos, *Polyhedron* **2008**, *27*, 887–892.
30. VCCLAB, Virtual Computational Chemistry Laboratory, http://www.vcclab.org, 2005.
31. R. Todeschini, V. Consonni, Handbook of Molecular Descriptors**,** in: R. Mannhold, H. Kubinyi, H. Timmerman (Eds), *Methods and Principles in Medicinal Chemistry*, Vol. 11, Wiley-VCH, Weinheim, 2000.
32. R. Todeschini, V. Consonni, Molecular Descriptors for Chemoinfomatics, 2nd Ed, in: R. Mannhold, H. Kubinyi, G. Folkers (Eds), *Methods and Principles in Medicinal Chemistry*, Vol. 41, Wiley-VCH, Weinheim, 2009.
33. http://cactus.nci.nih.gov/services/translate/
34. L. B. Kier, L. H. Hall, *J. Pharm. Sci.* **1976**, *65*, 1806–1809.
35. L. B. Kier, L. H. Kall, *Molecular Connectivity in Structure-Activity Analysis*, Willey, New York, 1986.
36. M. Randić, *MATCH Commun. Math. Comput. Chem.* **2008**, *59*, 5–124.
37. A. Miličević, N. Raos, *J. Phys. Chem. A* **2008**, *112*, 7745–7749.
38. P. J. Barrie, A. Gyani, M. Motevalli, P. O'Brien, *Inorg. Chem.* **1993**, *32*, 3862–3867.
39. A. Demaret, F. Abraham, *Acta Cryst. C* **1987**, *43*, 2067–2069.
40. R. J. Flook, H.C. Freeman, F. Huq and J.M. Rosalky, *Acta Cryst. B* **1973**, *29*, 903–906.
41. T. Kobayashi, Y. Inomata, T. Takeuchi, F. S. Howell, *Inorg. Chem. Acta* **1997**, *258*, 109–112.

42. R.V. Krishnakumar, M.P. Rameela, Dr. S. Natarajan, *Cryst. Res. Tech.* **1996**, *31*, 203–207.

43. B. Lučić, N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 121–132.

Povzetek

Razvili smo linearni model za oceno druge in splošne konstante stabilnosti, K_2 in β_2, za kadmijeve(II) binarne in ternarne *bis*-komplekse s petimi alifatskimi α-amino kislinami; model je osnovan na valenčnem indeksu povezanosti tretjega reda ($^3\chi^v$). Niz aminokislin je vseboval glicin, alanin, 2-aminobutanoično, 2-aminopentanoično (norvalin) in 2-aminoheksanoično kislino (norlevcin), ki se vežejo na kadmij(II) in tako določimo 25 K_2 and 15 β_2 vrednosti. Napovedi modela za log β_2 so ocenjene z r = 0.940 in S.E.$_{cv}$ = 0.10, medtem ko so statistični parametri ocene modela za log K_2 dveh delnih nizov r = 0.936 in 0.842, in S.E.$_{cv}$ = 0.09. Kompleks CdGG je bil izločen iz vseh regresijskih modelov.