

# ESTIMATION OF STABILITY CONSTANTS OF COORDINATION COMPOUNDS USING MODELS BASED ON TOPOLOGICAL INDICES

Nenad RAOS and Ante MILIČEVIĆ

*Institute for Medical Research and Occupational Health, Zagreb, Croatia*

Received in October 2008

Accepted in October 2008

The principal aim of our research was to develop theoretical models for estimating stability constants of coordination compounds, focusing on chelates of heavy metals, Cu(II) and Ni(II), with bioligands (amines, amino acids and peptides). We tried to solve the problem using graph-theoretical models. We also encountered a specific problem for coordination compounds – the choice of the constitutional formula, that is graph, of the complex species.

**KEY WORDS:** *copper(II) chelates, nickel(II) chelates, valence connectivity index*

There are two kinds of models in chemistry. The first treat molecule as a system of electrons and nuclei, and solve problems in a typical physical way, by solving equations of motion of these particles. This kind of modelling includes quantum-chemical method, and conceptually similar models that stem from statistical mechanics.

The second kind is based on the chemical way of thinking, namely on the comparison of related systems. The fundamental presumption in chemistry is that all properties of a matter correlate with its molecular structure, and that, consequently, molecules of similar structures have similar properties. This has led to a number of heuristic methods (QSPR, QSAR, QSTR, SAR, and SPC).<sup>1</sup> The common denominator of all these methods is the correlation of physicochemical properties (QSPR) or biological activity (QSAR) and molecular descriptors. There are many descriptors in use (1, 2), but one class of descriptors has proven

very simple and useful in predicting many molecular properties. These are called *topological indices*.

The first topological index was Wiener's "path number", now known as Wiener index or Wiener number ( $W$ ), which was used in 1940's to estimate the boiling points of alkanes (3). From these times the number of topological indices has grown enormously, as well as their application. In 1983, Trinajstić presented about two dozen topological indices (4), and at present there are more than a thousand topological indices of various complexity (5). They were successfully used for predicting many properties like density, viscosity, boiling temperature, solubility, refractive index, molar heat capacity, and standard Gibbs energy of formation (6-9).

Despite wide application, nobody has systematically used topological indices to estimate the stability of coordination compounds. The reason may be that topological analysis traditionally belongs to organic chemistry or that coordination chemists have always been skeptical about molecular graph theory, which seems to be "too primitive" to cope with such a

<sup>1</sup> QSPR, quantitative structure-property relationship; QSAR, quantitative structure-activity relationship; QSTR, quantitative structure-toxicity relationship; SAR, structure-activity relationship; SPC, structure-property correlation.

complex problem. Moreover, skepticism of graph theory among quantum chemistry-oriented chemists is notorious, even in its application to well-defined problems (10), so there is small wonder that we had to overcome many obstacles to publish our results.

## PROBLEMS

Stability constants of coordination compounds ( $K_n$ ,  $\beta_n$ )<sup>2</sup> depend on many factors, which generally should be classified as electronic (strength of metal-ligand bonds), stereochemical (the shape of coordination polyhedron), sterical (sterical interactions between atomic groups in a molecule), and solvation interactions (interactions of a ligand, metal, and complex with solute). In addition to qualitative rules such as Irving-Williams order (11) and hard-soft acid-base (HSAB) principle (12) that mostly express electronic influences, computation of sterical effects has widely been relying on molecular mechanics (MM). We used molecular-mechanics calculations to estimate entioselectivity effects for copper(II) chelates with *N*-alkylated amino acids (13, 14) and to estimate the diastereoselectivity effect for the complexes of epimeric amino acids threonine and isoleucine (15, 16). Molecular mechanics was also applied for the calculation of difference in the stability of related chelates (17, 18).

There are not many molecular descriptors in use in coordination chemistry. Usually, the researchers try to correlate stability constants to each other and to the protonation constants of the ligands (19, 20). We used overlapping-sphere (OS) volumes as molecular descriptors for copper(II) and nickel(II) chelates (21-23). Stability constants of complexes of 63 ligands and 14 lantanoides were estimated using many descriptors, but, unfortunately, they were not used systematically (24).

Difficulties in applying topological indices to estimate stability constants 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. The exact structure of a complex species in a solution is seldom known, as is the nature of chemical bonding. Even in the solid state, a molecule of a coordination compound has “weak” and “dubious” bonds and many sterical interactions

that are not easy to define. There is also a problem of coordination polyhedron plasticity, leading to distortion isomerism (25), not to mention problems with conformers and “classical” stereoisomers. As graph-theoretical methods do not explicitly deal with 3D structures, theoretical results are much more sensitive to sterical effects than in the case of organic compounds, for which these effects are less complex.

## SOLUTIONS

The use of topological indices in QSPR analysis of organic compounds follows these steps: 1) construct molecular graphs, 2) correlate indices derived from graphs with experimental values, 3) choose the best indices, and 4) try to improve the results doing multiple regression by combining the indices.

This simple scheme hardly works for coordination compounds because it is not at first clear which molecular species should be used as a basis for the graph. Also, due to paucity of experimental data, it is not always allowed to apply multiple regression or use any function but linear. As the number of suitable graph-theoretical indices is limited (see below), the only real possibility to improve the model is to modify the constitutional formula, *i.e.* graph, of the complex species. In doing so, the researcher has to bear in mind that the graph always has to truthfully reflect real interactions in a molecule, and not serve as a mere tool for improving statistical parameters in a regression.

### *Influence of constitutional formula*

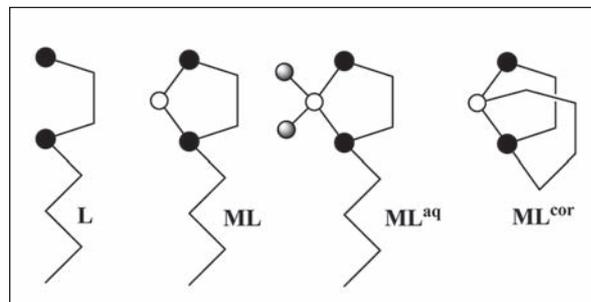
In our first applications of the graph-theoretical method for the estimation of stability constants of copper(II) chelates with amino acids and their *N*-alkylated derivatives (26, 27), we came to two important conclusions: (1) the best topological index is the valence-connectivity index of the 3<sup>rd</sup> order ( ${}^3\chi^v$ )<sup>3</sup> and (2) the results much depend on the molecular species for which topological index is calculated.

This led to a systematic analysis of the form of molecular graphs (28) and development of models based on four representations: (1) graph of the ligand (L), (2) graph of the complex (ML), (3) graph of the

<sup>2</sup>  $K_n = [ML_n]/[ML_{n-1}][L]$ ,  $\beta_n = [ML_n]/[M][L]^n$ .

<sup>3</sup>  ${}^3\chi^v = \sum[\delta(i)\delta(j)\delta(k)\delta(l)]^{-0.5}$ , where  $\delta(i)$  *etc.* are weights of vertices (atoms) making up the path of three consecutive bonds.

diaqua-complex ( $ML^{aq}$ ), and (4) graph of the complex with a secondary bonding of the central atom to the side chain ( $ML^{cor}$ ) (Figure 1).



**Figure 1** Four kinds of graphs used to estimate stability constants of coordination compounds. Example:  $M=Cu(II)$ ,  $L=N$ -butyldiaminoethane (28)

A comparison of the correlation coefficients of regressions of the first ( $\log K_1$ ) and overall stability constant ( $\log \beta_2$ ) for three sets of  $Cu(II)$  complexes with diamines,  $N$ -alkylated glycines, and amino acids shows that a more realistic representation of molecular structure yields better  $r$  values (Figure 2). Generally, a model based on the graph of the ligand ( $L$ ) proved worst, and  $ML^{aq}$  and  $ML^{cor}$  models proved equivalent. However,  $cor$  models should only be used for amino-acid chelates, as these models better describe interactions between side chains and/or their interactions with the central atom.

#### Influence of the form of regression function

To estimate stability constants of coordination compounds with the graph-theoretical method, we would usually use univariate linear regression with

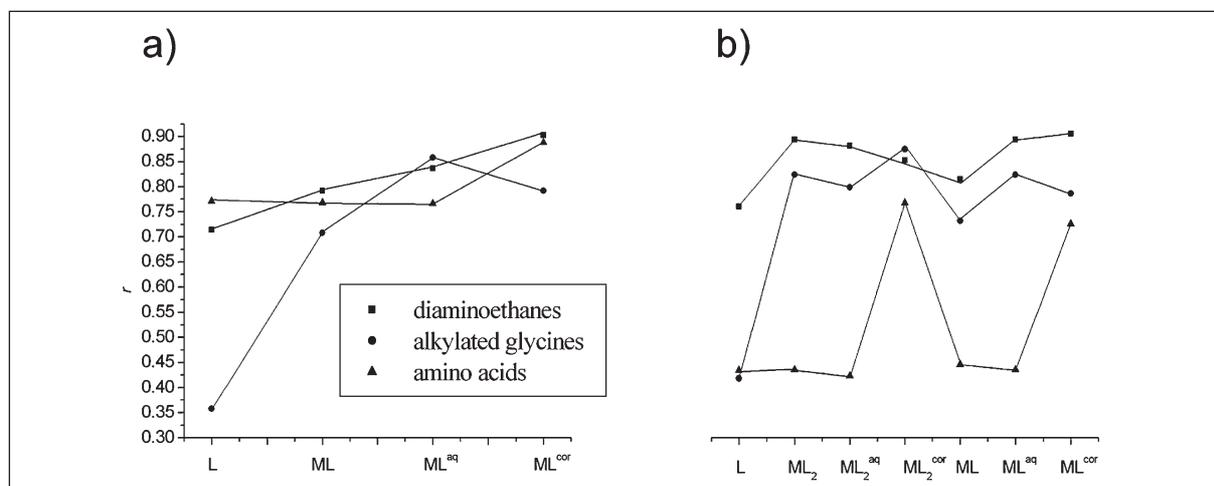
${}^3\chi^v$ . Values of  $\log K_1$  for  $Cu(II)$  chelates with naturally occurring amino acids were correlated to the respective connectivity indices, yielding  $S.E._{cv}$  about 0.11 (28). Moreover, we did correlations on sets consisting of complexes of both  $Cu^{2+}$  and  $Ni^{2+}$  by introducing indicator variable, *i.e.* by imposing the condition that regression lines for two metals have to be parallel (22). These regressions were also used as calibration lines for estimating stability constants for complexes with fructose-amino acid adducts, yielding  $rms(\log K_1)$  of 0.22 and 0.27 for  $Cu(II)$  and  $Ni(II)$  complexes, respectively (29). Using the indicator variable,  $\log K_1$  of eight  $Cu(II)$  complexes with diethylenetriamines and  $\log \beta_2$  of their 18 mixed complexes with amino acids were successfully correlated (30).

Another way to use multiple linear regression is to use  ${}^3\chi^v$  of molecular segments (28). To calculate  $\log \beta_2$  for mixed  $Cu(II)$  chelates with amino acids ( $MLA$ ), we applied the bivariate function with variables  $[{}^3\chi^v(ML) + {}^3\chi^v(MA)]$  and  $[{}^3\chi^v(ML) - {}^3\chi^v(MA)]$ . To model these complexes, we used functions of even higher complexity (31):

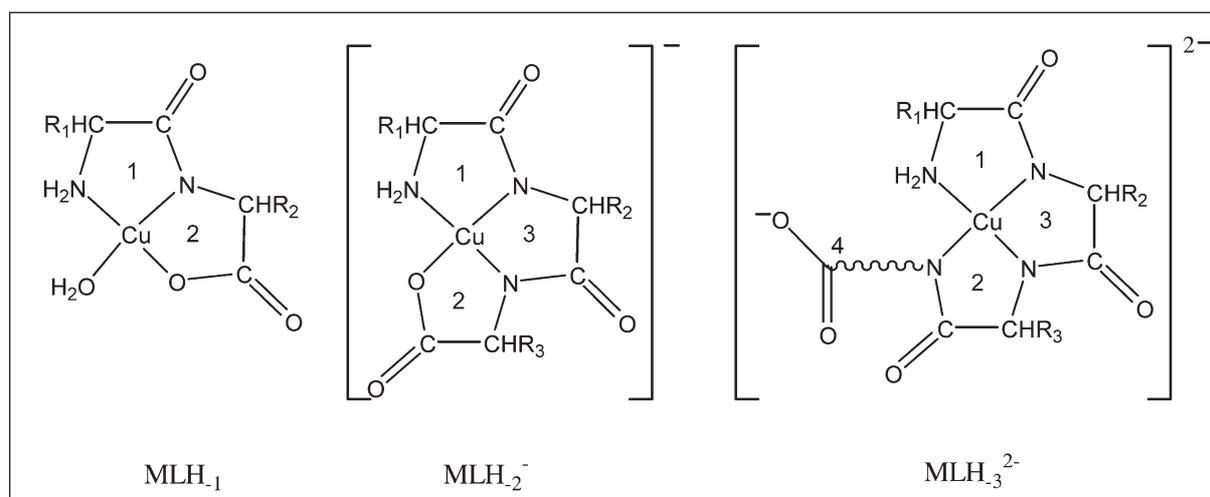
$$\log K_2 = a_1[{}^3\chi^v(MA)] + a_2[{}^3\chi^v(MLA) - {}^3\chi^v(ML) - {}^3\chi^v(MA)] + b \quad (1)$$

$$\log \beta_2 = a_1[{}^3\chi^v(MA)] + a_2[{}^3\chi^v(MLA) - {}^3\chi^v(ML) - {}^3\chi^v(MA)] + a_3[{}^3\chi^v(ML) - {}^3\chi^v(MA)] + b \quad (2)$$

where the second and the third term were introduced for modelling interactions between chelate rings. By virtue of these functions, better estimates for both binary and ternary complexes were obtained, and Eq. 1 made possible direct calculations of  $\log K_2$ .



**Figure 2** Correlation coefficients for the linear regression of a)  $\log K_1$  (a) and b)  $\log \beta_2$  on  ${}^3\chi^v$  indices of various representations of  $Cu(II)$  chelates with diaminoethanes ( $N=14$ ),  $N$ -alkylated glycines ( $N=11$ ) and naturally occurring amino acids ( $N=9$ ) (28)



**Figure 3** 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.

A similar way to develop a multivariate function was tested on Cu(II) and Ni(II) chelates with dipeptides (32) and higher peptides (from tripeptides to pentapeptides) (33). The graph of a molecule was divided into subgraphs, each representing one chelate ring. Valence connectivity index of each chelate ring,  ${}^3\chi^v(r_1)$ ,  ${}^3\chi^v(r_2)$ , etc. (Figure 3) should be treated as an independent variable in linear regression (32, 33). However, this is not practical, for the number of regression parameters depends on the number of amino-acid residues. Thus, we introduced the function:

$$\log K_1 = a_1 [{}^3\chi^v(r_1) + {}^3\chi^v(r_3)] + a_2 [{}^3\chi^v(r_2) + {}^3\chi^v(r_3) + {}^3\chi^v(r_4)] + b \quad (3)$$

where  $r_4$  denotes (4<sup>th</sup> and higher) amino-acid residue(s) which is/are not coordinated to central atom.

## CONCLUSION

Our research shows that despite their “simplicity”, models based on graph-theoretical indices provide a fair ( $rms < 0.3$ , max. error  $< 0.5$ ) estimation of the stability constants of coordination compounds. However, these models are not easy to develop because for each system specific influences have to be modelled. This means that there is no general QSPR model for the estimation of stability constants of coordination compounds from topological indices.

## Acknowledgement

This work was supported by the Ministry of Science, Education and Sports, Republic of Croatia (grant no. 022-1770495-2901 “Development of models for estimation of stability constants”).

## REFERENCES

- Dagani R. Toxicologists seek structure-activity links. *Chem Eng News* 1981;9:26-9.
- van de Waterbeemd H. Quantitative approaches to structure-activity relationship. In: Wermuth CG, editor. *The practice of medical chemistry*. London: Academic Press; 1996. p. 367-89.
- Wiener H. Structural determination of paraffin boiling points. *J Am Chem Soc* 1947;69:17-20.
- Trinajstić N. *Chemical Graph Theory*. 1<sup>st</sup> edition. Boca Raton (FL): CRC Press; 1983.
- Todeschini R, Consonni V. *Handbook of Molecular Descriptors*. Weinheim: Wiley, VCH; 2000.
- Konstantinova EV, Diudea MV. The Wiener polynomial derivatives and other topological indices in chemical research. *Croat Chem Acta* 2000;73:383-403.
- Lučić B, Miličević A, Nikolić S, Trinajstić N. Harary index – twelve years later. *Croat Chem Acta* 2002;75:847-68.
- Balaban AT, Mills D, Ivanciuc O, Basak SC. Reverse Wiener indices. *Croat Chem Acta* 2000;73:923-41.
- Rouvray DH. The modeling of chemical phenomena using topological indices. *J Comput Chem* 1987;8:470-80.
- Randić M. On history of the Randić index and emerging hostility toward chemical graph theory. *MATCH Commun Math Comput Chem* 2008;59:5-124.
- Irving H, Williams RJP. Order of stability of metal complexes. *Nature* 1948;162:746-7.

12. Pearson RG. Hard and soft acids and bases. *J Am Chem Soc* 1963;85:3533-9.
13. Sabolović J, Raos N. Critical evaluation of empirical force-field models for simulation of plasticity of copper(II) coordination. *Polyhedron* 1990;9:1277-86.
14. Sabolović J, Raos N. Simulation of plasticity of a copper(II) coordination polyhedron with a force field based on coulombic interactions: conformational analysis of copper(II) chelates with  $\alpha$ -amino acids. *Polyhedron* 1990;9:2419-27.
15. Raos N, Niketić SR, Simeon VI. Conformational analysis of copper(II) chelates with epimeric amino acids. *J Inorg Biochem* 1982;16:1-19.
16. Raos N, Simeon V. Conformational analysis and hydration models of copper(II) threoninates and isoleucinates. *J Inorg Biochem* 1983;18:133-51.
17. Brubaker G, Johnson DW. Quantitative estimates of steric effects. Intramolecular strain-energy effects on the stability and dissociation rate constants of polythia ether macrocycle complexes of copper(II). *Inorg Chem* 1984;23:1591-5.
18. Hancock RD, McDougall GJ, Marsicano F. Empirical force field calculations of strain-energy contributions to the thermodynamics of complex formation. 3. Chelate effect in complexes of polyamines. *Inorg Chem* 1979;18:2847-52.
19. Sigel H, Massaud SS, Corfu NA. Comparison of the extent of base backbinding in complexes of divalent metal ions with guanine ( $GMP^{2-}$ ), inosine ( $IMP^{2-}$ ) and adenosine 5'-monophosphate ( $AMP^{2-}$ ). The crucial role of N-7 basicity in metal ion-nucleic base recognition. *J Am Chem Soc* 1994;116:2958-71.
20. Vinokurov EG, Bondar VV. Prediction of stability constants for Cr(III) and Cr(II) complexes. *Russ J Coord Chem* 2003;29:66-72.
21. Raos N. Estimation of stability constants of copper(II) chelates with amino acids by overlapping spheres method. *Croat Chem Acta* 2005;78:175-80.
22. Miličević A, Raos N. Estimation of stability constants of copper(II) and nickel(II) chelates with 1,2-diaminoethanes by the overlapping spheres method. *Croat Chem Acta* 2005;78:525-34.
23. Miličević A, Raos N. Estimation of stability constants of copper(II) bis-chelates by the overlapping spheres method. *Croat Chem Acta* 2006;79:281-90.
24. Svetlitski R, Lomaka A, Karelson M. QSPR modelling of lanthanide-organic complex stability constants. *Sep Sci Technol* 2006;41:197-216.
25. Gažo J. Structure and properties of some copper(II) coordination compounds. Distortion isomerism of Cu(II) compounds. *Pure Appl Chem* 1974;38:279-301.
26. Grgas B, Nikolić S, Paulić N, Raos N. Estimation of stability constants of copper(II) chelates with *N*-alkylated amino acids using topological indices. *Croat Chem Acta* 1999;72:885-95.
27. Nikolić S, Raos N. Estimation of stability constants of mixed amino acid complexes with copper(II) from topological indices. *Croat Chem Acta* 2001;74:621-31.
28. Miličević A, Raos N. Estimation of stability of coordination compounds by using topological indices. *Polyhedron* 2006;25:2800-8.
29. Miličević A, Raos N. Prediction of stability of copper(II) and nickel(II) complexes with fructose-amino acids from the molecular graph models developed on amino acid chelates. *Croat Chem Acta* 2007;80:557-63.
30. A. Miličević A, Raos N. Estimation of stability constants of copper(II) chelates with triamines and their mixed complexes with amino acids by using topological indices and the overlapping spheres method. *Polyhedron* 2007;26:3350-6.
31. Miličević A, Raos N. Influence of chelate ring interactions on copper(II) chelate stability studied by connectivity index functions. *J Phys Chem* 2008;112:7745-9.
32. Miličević A, Raos N. Estimation of stability constants of copper(II) and nickel(II) chelates with dipeptides by using topological indices. *Polyhedron* 2008;27:887-92.
33. Miličević A, Raos N. Estimation of stability constants with connectivity index: development of bivariate and multivariate linear models for copper(II) chelates with oligopeptides. *Croat Chem Acta*, in press.

*Sažetak*

## PROCJENA KONSTANTA STABILNOSTI KOORDINACIJSKIH SPOJEVA S POMOĆU MODELA TEMELJENIH NA TOPOLOŠKIM INDEKSIMA

Prikazani su načini rješavanja problema procjene konstanta stabilnosti bakrovih(II) i niklovi(II) kelata s aminima, aminokiselinama i peptidima primjenom graf-teorijskih modela. Osebujni problem za koordinacijske spojeve je izbor konstitucijske formule, tj. grafa kompleksnoga spoja, pa su stoga iskušani modeli temeljeni na formuli liganda i više oblika konstitucijske formule kompleksa. Pokazalo se da se najbolji rezultati postižu valencijskim indeksom povezanosti trećeg reda,  ${}^3\chi^v$ . Iskušani su modeli temeljeni na kombinacijama indeksa izvedenih iz molekularnih segmenata. Takvi su se modeli pokazali osobito uspješnima za procjenu stabilnosti miješanih kompleksa.

**KLJUČNE RIJEČI:** bakrovi(II) kelati, niklovi(II) kelati, valencijski indeks povezanosti

## CORRESPONDING AUTHOR:

Nenad Raos  
Institute for Medial Research and Occupational Health  
P. O. Box. 291, HR-10001 Zagreb, Croatia  
E-mail: [nraos@imi.hr](mailto:nraos@imi.hr)