

# Prediction of Complexation Properties of Crown Ethers Using Computational Neural Networks

ANDREI A. GAKH,\* BOBBY G. SUMPTER, DONALD W. NOID,  
RICHARD A. SACHLEBEN\* and BRUCE A. MOYER

*Chemical and Analytical Sciences Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6197, U.S.A.*

(Received: 9 January 1996; in final form: 3 July 1996)

**Abstract.** A computational neural network method was used for the prediction of stability constants of simple crown ether complexes. The essence of the method lies in the ability of a computer neural network to recognize the structure–property relationships in these host–guest systems. Testing of the computational method has demonstrated that stability constants of alkali metal cation ( $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Cs}^+$ )–crown ether complexes in methanol at 25 °C can be predicted with an average error of  $\pm 0.3$  log  $K$  units based on the chemical structure of the crown ethers alone. The computer model was then used for the preliminary analysis of trends in the stabilities of the above complexes.

**Key words:** Crown ethers, complexes, stability constants, structure-property relationships, computational neural networks.

## 1. Introduction

The application of computational neural networks in solving chemical problems has been receiving growing attention from chemists as a powerful tool of approximation, computation and pattern recognition [1–4]. Such applications have been particularly successful in the prediction of physical properties of organic compounds [5–9].

In our previous papers we demonstrated one of the possible schemes of neural network computing by predicting properties of hydrocarbons and Freons<sup>R</sup> [7,8]. The main idea was to encode structural parameters of the organic molecules by using the set of topological invariants, and then to use these invariants as input for neural network computing [7]. Another approach entailed direct utilization of structural features of compounds as a source of the structural information for the neural networks [8]. Both approaches have demonstrated good results for the prediction of such basic physical characteristics as boiling point, density, heat capacity, etc. [7,8]. Our current interest is to expand this promising method for prediction of more complex phenomena, such as host-guest interactions, a cornerstone of supramolecular chemistry [10]. These interactions control the selectivity and efficiency of complexants and extractants, essential compounds for advanced

---

\* Authors for correspondence.

separation technologies, drug delivery systems, and modern analytical methods used in molecular biology. Host-guest interactions also play a vital role in living organisms, being responsible for ion transport through cell membranes.

Computational chemistry is increasingly used in the study of crown ethers, but most of the known methods entail exceedingly time-consuming and expensive computational procedures such as *ab initio* calculations and detailed simulations [11–15]. To gain a simple and more practical measure of the predictability of chemical processes in general, some workers have turned to statistical methods [16]. The approach employs linear regression to relate a dependent thermodynamic variable to other physical and chemical properties of the reacting components. Only recently have such attempts been made in the case of complexation of cations by crown ethers [17]. The computational method utilized molecular mechanics and molecular dynamic simulations for the crown ether, crystal radius and electronegativity for the metal cation, together with various solvent properties such as dielectric constant and solubility parameters. Fourteen major parameters were used in multiple linear regression analysis to give standard errors ranging from 0.36 to 1.42 log  $K$  units [17].

Here we present a computational scheme for the prediction of complexation constants of simple crown ethers with alkali metal cations in a single solvent (methanol) at 25 °C. The computational method of neural network computing allows us to create computer models of host-guest interactions in these systems. Only the chemical information contained in the name of the crown ether is employed. No other properties of the crown ether, either computed or measured, are needed. The computer models are useful for the inspection of trends in complexation phenomena. They offer a fast measure of predictability in the absence of extensive experimental or computed data on crown ether properties.

## 2. Results and Discussion

Simple crown ethers consist of repeating structural units which can be conveniently encoded directly from their systematic chemical names to numeric vectors suitable for neural network input. For our initial study, we chose monocyclic crown ethers consisting only of aliphatic 1,2-dioxyethylene and aromatic 1,2-dioxyphenylene fragments. Since the exact positions of these units are usually omitted in the chemical names of these compounds, in our computations we include only ring size, overall number of oxygen atoms, and number of aromatic rings. Thus, for example, the input vector of benzo-18-crown-6 (B18C6) is 1 18 6 where the first digit represents the number of benzene rings (benzo in the name), the second digit represents ring size (-18- in the name), and the final digit represents the number of oxygen atoms in the ring. An example of a sample numeric input vector preparation is presented in Figure 1.

Stability constants of various crown ether complexes with alkali metal cations (output data for training and testing) were used as reported previously without any

Table I. Experimental and predicted stability constants of crown ether complexes ( $\text{Na}^+$ ).

Compound	log K for $\text{Na}^+$ in MeOH at 25 °C		Set (l-learning; t-testing)
	Experimental	Predicted	
12C4	2.05	1.95	l
	1.75		
	2.10		
	1.20		
15C5	3.42	3.37	t
	3.43		
	3.23		
	3.29		
	3.31		
	3.27		
	3.39		
B15C5	2.89	3.00	l
	3.06		
	3.03		
	3.10		
	2.99		
	3.37		
	2.94		
	2.91		
	3.05		
	3.38		
	2.97		
18C6	4.32	4.38	l
	4.42		
	4.46		
	4.30		
	4.32		
	3.92		
	4.65		
	4.34		
B18C6	4.21	4.37	t
	4.53		
B <sub>2</sub> 18C6	4.37	4.27	l
	4.18		
21C7	1.73	2.41	l
	2.46		

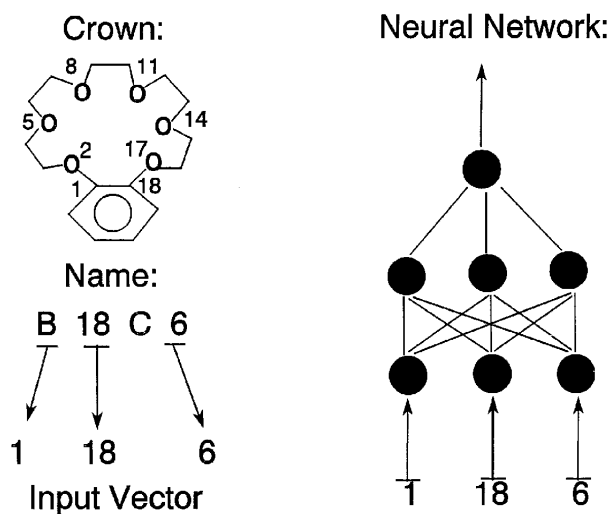


Figure 1. Diagram of the neural network used for predicting the stability of crown ether complexes and sample input file preparation. Bias connections in the neural network scheme are omitted for clarity.

Table I. Continued.

Compound	log K for Na <sup>+</sup> in MeOH at 25 °C		Set (l-learning; t-testing)
	Experimental	Predicted	
B <sub>2</sub> 24C8	2.35	2.40	l
B <sub>2</sub> 27C9	1.88	1.92	t
B <sub>2</sub> 27C9	1.90	1.91	l
B <sub>3</sub> 0C10	1.93	1.91	l
B <sub>2</sub> 30C10	1.98	1.91	l
	2.80		
B <sub>3</sub> 3C11	1.91	1.91	t
B <sub>2</sub> 33C11	1.82	1.90	l

pre-selection [18,19]. In the case of conflicting data, all sets of available constants were taken as input. Since the temperature and solvent also affect the stability constants, we used only data obtained at standard temperature (25 °C) and in one solvent (methanol) [20].

The optimized neural network (optimization procedure included leave-*k*-out and other methods [1–4,7]) consists of three input neurons, one hidden layer (three neurons), and one output neuron (see Figure 1). For simplicity, bias was omitted on the picture. A full data set, including chemical name abbreviations and the experimental and predicted stability constants of crown ethers, is presented in Tables I–III.

Table II. Experimental and predicted stability constants of crown ether complexes ( $K^+$ ).

Compound	log K for $K^+$ in MeOH at 25 °C		Set (l-learning; t-testing)
	Experimental	Predicted	
12C4	1.73	1.66	l
	1.70		
	1.55		
	1.30		
15C5	3.85	3.40	t
	3.60		
	3.43		
	3.86		
	3.63		
	3.38		
	3.30		
	3.35		
	3.36		
	3.34		
B15C5	3.93	3.03	l
	2.71		
	2.96		
	3.05		
	2.90		
18C6	6.20	5.85	l
	6.07		
	6.29		
	6.09		
	6.00		
	6.02		
	6.32		
	6.15		
	6.08		
	6.30		
B18C6	5.29	5.47	t
	5.05		
	5.20		
B <sub>2</sub> 18C6	5.00	4.95	l
	4.80		
21C7	4.22	4.27	l
B <sub>2</sub> 21C7	4.19	4.13	t
B <sub>2</sub> 24C8	3.45	4.02	l
	3.57		

Table II. Continued.

Compound	log $K$ for $K^+$ in MeOH at 25 °C		Set (l-learning; t-testing)
	Experimental	Predicted	
B27C9	4.00	3.94	t
B <sub>2</sub> 27C9	4.36	4.01	l
B30C10	4.28	3.94	l
B <sub>2</sub> 30C10	4.47	4.00	l
	4.62		
	4.23		
	4.94		
B33C11	3.57	3.93	l
B <sub>2</sub> 33C11	3.79	4.00	t

Our neural network performance data are presented in Figure 2. For each alkali metal cation the computer neural network was trained and tested separately using ‘learning’ datasets (marked as ‘l’ in Tables I–III) and ‘testing’ datasets (marked as ‘t’ in Tables I–III). Three input neurons in the neural network have a linear transfer function, and neurons of the ‘hidden’ layer have a hyperbolic tangent transfer function (except for  $Cs^+$ : sigmoid) using the default parameters established by the ‘Backpropagation Builder’ included in the program package [21]. The best results were achieved using optimized learning rate, momentum, and transfer function. In particular, for the prediction of complexation constants of  $Na^+$  the learning rate was 0.3 for ‘hidden’ layer, 0.15 for output layer, and momentum was 0.6. The learning rates for  $K^+$  and  $Cs^+$  were 0.2/0.1 and 0.3/0.15, with momentums 0.4 and 0.6, respectively. Detailed optimization procedures for these two parameters were reported elsewhere [1,4–7]. In our case, optimization was trivial because of the small size of the neural network (only seven neurons in total). This small neural network could be trained successfully even with a nonoptimum combination of learning rate and momentum. We started the optimization procedure with relatively large momentum and small learning rate and gradually changed them until the optimal parameters were reached. During the training procedure, the neural network performance was monitored using RMS on-screen graphic indicator, and the training process was terminated when no further progress in accuracy was made (86 000–123 000 epochs, 16 cycles each). The trained neural network was then tested using a separate ‘testing’ dataset to validate the results. Each training procedure required 20–30 min of real one-task computing time on an IBM 486 (50 MHz) personal computer [21]. It was found that moderate fluctuations of the learning rate and momentum from the optimal values could increase the overall training time of the neural network but had little impact on the predicted values, indicating the stability of the chosen computer model.

The average error of the training sets was  $\pm 0.27 \log K$  units. The average error of the testing sets was slightly higher ( $\pm 0.34 \log K$  units), an indication that the neural

Table III. Experimental and predicted stability constants of crown ether complexes ( $\text{Cs}^+$ ).

Compound	log $K$ for $\text{Cs}^+$ in MeOH at 25 °C		Set (1-learning; t-testing)
	Experimental	Predicted	
12C4	1.60	1.33	l
	1.65		
15C5	2.78	2.75	t
	2.81		
	2.66		
	3.58		
	1.89		
B15C5	3.52	2.35	l
	2.15		
	1.91		
	2.20		
	2.21		
B <sub>2</sub> 15C5	1.96	1.96	l
18C6	4.44	3.97	t
	4.37		
	4.49		
	4.70		
B18C6	3.95	3.66	l
	4.04		
B <sub>2</sub> 18C6	3.66	3.37	l
	3.20		
21C7	5.01	4.57	l
B <sub>2</sub> 21C7	4.25	4.10	t
	4.20		
B <sub>2</sub> 24C8	3.85	4.35	l
	3.84		
B27C9	4.57	4.57	l
B <sub>2</sub> 27C9	4.35	4.36	l
	3.67		
B30C10	4.35	4.52	t
B <sub>2</sub> 30C10	4.38	4.33	l
	4.00		
	4.79		
B33C11	3.88	4.45	t
B <sub>2</sub> 33C11	4.72	4.27	l

networks are successfully recognizing structure–property relationships, not simply memorizing the training set. Successful construction of neural networks which represent computer models of complexation phenomena allowed us to investigate some tendencies in the stability constants of alkali metal–crown ether complexes.

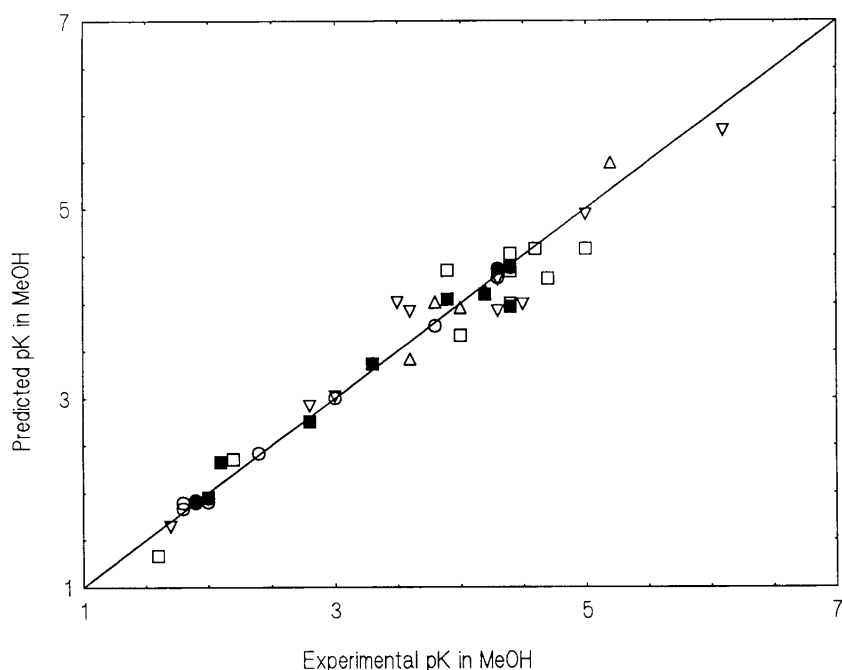


Figure 2. Experimental versus neural network predicted values of the stability of crown ether complexes: (a)  $\text{Na}^+$ : learning (training) set – open circles; testing (validation) set – filled circles, (b)  $\text{K}^+$ : learning (training) set – open triangles with down vertex; testing (validation) set – open triangles with up vertex, (c)  $\text{Cs}^+$ : learning (training) set – open squares; testing (validation) set – filled squares.

For that purpose we prepared artificial input files containing all combinations of ring size and number of benzene units and depicted the results of neural network predictions in two-dimensional plots.

### 2.1. RING SIZE

It is well documented that ring size has a major effect on the stability constants of crown ethers [18–19]. To make the overall tendencies (stability constant versus ring size) more easy to interpret, we first calculated stability constants of crown ethers with the general formula  $(-\text{O}-\text{CH}_2\text{CH}_2-)_n$  and plotted the results for the three alkali metal cations,  $\text{Na}^+$ ,  $\text{K}^+$ , and  $\text{Cs}^+$ , versus  $n$  (Figure 3).

The depicted results are in general agreement with the well known correlations between the ring size and the ability of crown ethers to complex alkali metal cations [18–19, 22]. The main feature of these correlations is the existence of maxima on the ring size – stability constant curve which represent the structure of the crown ether best suited for the complexation of the specific metal cation. For sodium and potassium, the optimum binding is obtained with 18C6; smaller and larger crown



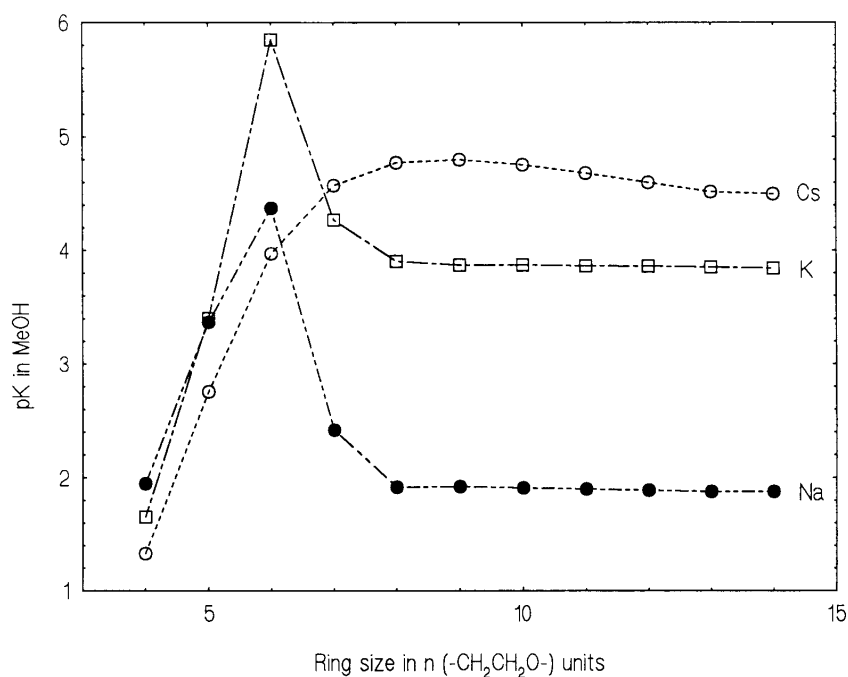


Figure 3. Neural network predicted values of the stability of model unsubstituted crown ethers complexes versus ring size (in  $-\text{CH}_2\text{CH}_2\text{O}-$ ) units: (a)  $\text{Na}^+$ : filled circles, (b)  $\text{K}^+$ : open squares, (c)  $\text{Cs}^+$ : open circles.

ethers bind more weakly. For Cs, a maximum is reached with 24C8; however, the decrease in binding strength with increasing ring size is small.

Further examination of the results produced by the neural network computational model indicates that the binding strengths reach a plateau for each of the metal ions at large ring sizes. In the case of sodium, this plateau is reached (following the maximum at  $n = 6$ ) for  $n = 8$  (24C8) at a value of  $\log K \approx 2$ . Potassium also reaches a plateau at  $n = 8$ , but with  $\log K \approx 4$ . Cesium exhibits a slow steady increase of the stability constants with increasing  $n$ , reaching a plateau in the diffuse region  $n = 7-9$  with  $\log K \approx 4.5$ ; it then shows a slight decrease for larger rings.

The observation of these plateau regions suggests a systematic trend in the coordination of the 1,2-dioxyethylene units with each of the above cations. The plateau region for sodium is virtually flat at the level of  $\log K \approx 2$  indicating a lower overall preference of the 1,2-dioxyethylene unit for sodium as compared to potassium and cesium, where the plateau regions occur at  $\log K \approx 4-5$ . It may be noted that this preference corresponds to the results of molecular mechanics calculations [23,24].

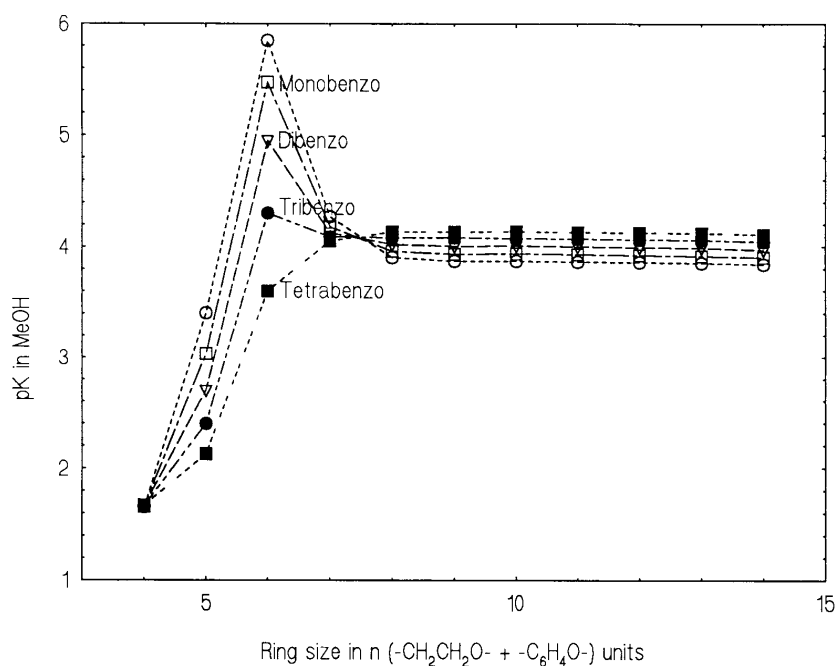


Figure 4. Neural network predicted values of the stability of  $K^+$  complexes of model crown ethers versus ring size (in  $-CH_2CH_2O-$  +  $-C_6H_4O-$ ) units for different numbers of 1,2-dioxyphenylene fragments.

## 2.2. NUMBER OF BENZENE RINGS

All of the previous analyses were made for the systems having no substituents on the macrocyclic ring. Addition of benzene rings, as 1,2-dioxyphenylene units, is expected to result in changes in the complexing ability of crown ethers. Such perturbations can be clearly seen in Figure 4, where five different curves represent the stabilities of the potassium complexes of benzo-crown ethers as a function of ring size (in  $n$  representing the overall number of 1,2-dioxyphenylene and 1,2-dioxyethylene fragments). The addition of each 1,2-dioxyphenylene fragment results in a steady decrease in binding strength of the crown ether in the region of strongest binding, that is, the 'peak' or 'hole-size' match region [22]. With four 1,2-dioxyphenylene fragments the 'match' phenomena disappear and a plateau of maximum binding occurs at  $n \geq 8$ .

The same trends were observed for  $Na^+$  ion complexation. However, the overall dependence of  $\log K$  on the numbers of 1,2-dioxyphenylene fragments is much weaker than for  $K^+$ , with the tendency of shifting the maxima to the higher ring sizes (Figure 5).

Somewhat different were the results obtained for  $Cs^+$  ion complexation trends (Figure 6). In this case the increase in the number of 1,2-dioxyphenylene fragments results in a smooth decrease of complexation ability, more pronounced for smaller

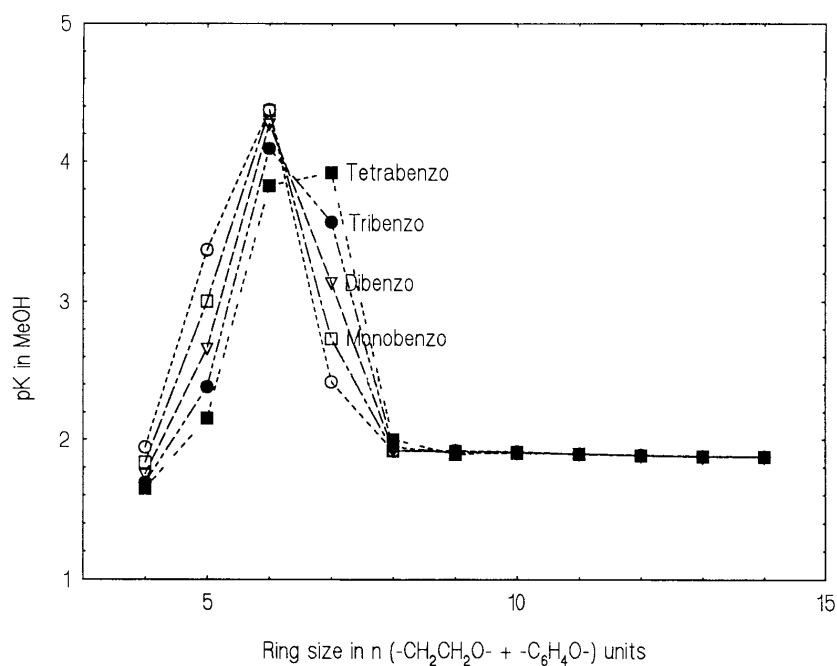


Figure 5. Neural network predicted values of the stability of  $\text{Na}^+$  complexes of model crown ethers versus ring size (in  $-\text{CH}_2\text{CH}_2\text{O}- + -\text{C}_6\text{H}_4\text{O}-$ ) units for different numbers of 1,2-dioxyphenylene fragments.

rings than for bigger ones. In addition, each decrease becomes sequentially smaller. Thus, for example, there is practically little or no difference between tribenzo- and tetrabenzocrowns having  $n > 6$ .

### 2.3. EXPERIMENTAL ERRORS ANALYSIS

Finally, inspection of the literature data listed in Tables I–III reveals disagreement among several sources of  $\log K$  for the binding of a given metal ion by the same crown ether. Such disagreement frustrates use of  $\log K$  values for correlations and interpretation, raising the dilemma of which particular data points to use and which to reject. Careful analysis of the individual literature sources may provide some confidence in the choice of the most reliable values, but not always. Thus, it may be suggested that the neural-network approach offers some additional means of deciding on the reliability of reported data. This is because the neural network bases its prediction on the entire series of crown ethers rather than on the small subset of values reported for a single case. Thus, one has a measure of overall consistency. If inconsistency is found, it may motivate attempts to check the data, in which case either the data will be revised or else an unsuspected phenomenon may be occurring.

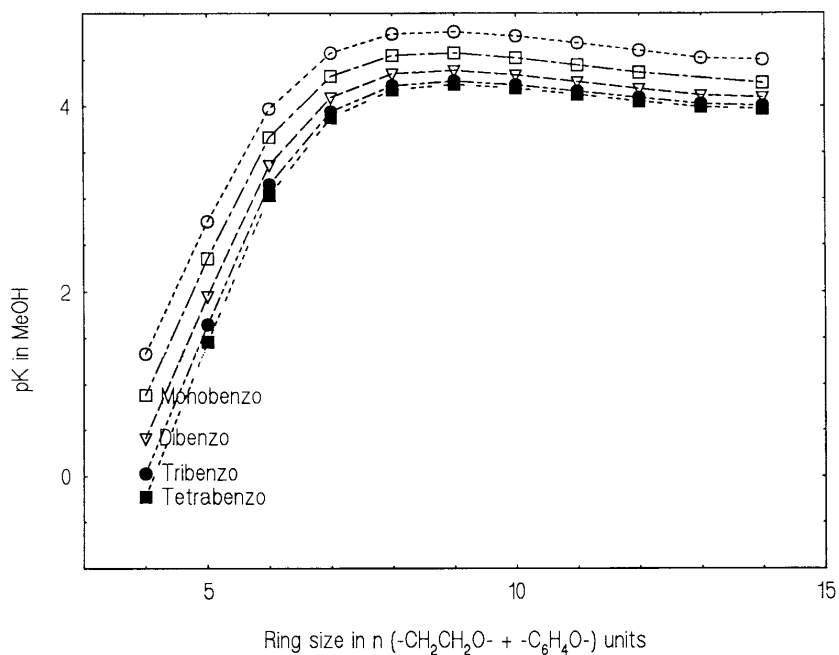


Figure 6. Neural network predicted values of the stability of  $\text{Cs}^+$  complexes of model crown ethers versus ring size (in  $-\text{CH}_2\text{CH}_2\text{O}- + -\text{C}_6\text{H}_4\text{O}-$ ) units for different numbers of 1,2-dioxyphenylene fragments.

### 3. Conclusions

We have shown that neural networks can predict the stability constants of crown ether complexes based on chemical structures encoded in their chemical names. This simple computational scheme has an average accuracy of  $\pm 0.3 \log K$  units and does not require any prior knowledge of structure-property relationships for this class of compounds. In addition, all necessary computations can be performed on a personal computer. The neural network computing thus provides a model of host-guest interaction which can be used for predictive analysis and as an aid in the development of new selective/exclusive complexants and extractants.

### Acknowledgement

This research was sponsored by the Divisions of Chemical Sciences (RAS and BAM), and Material Sciences (AAG, BGS, and DWN), Office of Basic Energy Sciences, U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp., and in part by an appointment to the ORNL Postdoctoral Research Associates Program administered jointly by ORNL and the Oak Ridge Institute for Science and Education (ORISE).

## References

1. J. Zupan and J. Gasteiger: *Anal. Chim. Acta* **248**, 1 (1991).
2. G. M. Maggiora, D. W. Elrod, and R. G. Trenary: *J. Chem. Inf. Comput. Sci.* **32**, 732 (1992).
3. J. Gasteiger and J. Zupan: *Angew. Chem. Int. Ed. Engl.* **32**, 503 (1993).
4. B. G. Sumpter, C. Getino, and D. W. Noid: *Annu. Rev. Phys. Chem.* **45**, 439 (1994).
5. N. Bodor, A. Harget, and M.-J. Huang: *J. Am. Chem. Soc.* **113**, 9480 (1991).
6. L. M. Egolf and P. C. Jurs: *J. Chem. Inf. Comput. Sci.* **33**, 616 (1993), and references therein.
7. A. A. Gakh, E. G. Gakh, B. G. Sumpter, and D. W. Noid: *J. Chem. Inf. Comput. Sci.* **34**, 832 (1994).
8. A. A. Gakh, E. G. Gakh, B. G. Sumpter, D. W. Noid, L. D. Trowbridge, and D. A. Harkins: *J. Fluorine Chem.* **73**, 107 (1995).
9. B. G. Sumpter and D. W. Noid: *Macromol. Theory Simul.* **3**, 363, (1994).
10. J.-M. Lehn: *Angew. Chem. Int. Ed. Engl.* **27**, 89 (1988).
11. M. H. Mazor, J. A. McCammon, and T. P. Lybrand: *J. Am. Chem. Soc.* **111**, 55 (1989).
12. P. D. J. Grootenhius and P. A. Kollman: *J. Am. Chem. Soc.* **111**, 4046 (1989).
13. P. D. J. Grootenhius, J. van Eerden, E. J. R. Sudholter, D. N. Reinhoudt, A. Roos, S. Harkema, and D. Feil: *J. Am. Chem. Soc.* **109**, 4792 (1987).
14. M. A. Thompson, E. D. Glendening, and D. Feller: *J. Phys. Chem.* **98**, 10465 (1994).
15. B. P. Hay, J. R. Rustad, and C. J. Hostetler: *J. Am. Chem. Soc.* **115**, 11158 (1993).
16. M. J. Kamlet, R. M. Doherty, M. H. Abraham, Y. Marcus, and R. W. Taft: *J. Phys. Chem.* **92**, 5244 (1988).
17. Z.-G. Shi and E. A. McCullough, Jr.: *J. Incl. Phenom.* **18**, 9 (1994).
18. R. M. Izatt, J. S. Bradshaw, S. A. Nielsen, J. D. Lamb, and J. J. Christensen: *Chem. Rev.* **85**, 271 (1985).
19. R. M. Izatt, K. Pawlak, J. S. Bradshaw, and R. L. Bruening: *Chem. Rev.* **91**, 1721 (1991).
20. Our preliminary computations indicate that the influence of temperature and solvent can be incorporated in neural network computing by adding two additional inputs (temperature and solvent) in the neural network architecture. For a limited set of solvents (for example, methanol, acetonitrile, water) they can be coded sequentially, 1, 2, and 3. However, for a large set of solvents a more sophisticated coding algorithm is required. These computations lie beyond the scope of the current work.
21. Neural network computational packages (NeuralWorks Professional II/PLUS and NeuralWorks Explorer) are commercially available from NeuralWare, Inc. (Penn Center West, Building IV Pittsburgh, PA 15276-9910) both for PC and mainframe computers. In this study we have used backpropagation-type neural networks included in this program package. For more details of neural network computing see, for example [1-4, 7], and references therein.
22. R. L. Bruening, R. M. Izatt, and J. S. Bradshaw: *Understanding Cation-Macrocyclic Binding Selectivity in Single-Solvent, Extraction and Liquid Membrane Systems by Quantifying Thermodynamic Interactions* (Cation Binding by Macrocycles: Complexation of Cationic Species by Crown Ethers, Eds. Y. Inoue and G. W. Gokel), pp 111-132. Marcel Dekker, New York, NY (1990).
23. B. P. Hay and J. R. Rustad: *J. Am. Chem. Soc.* **116**, 6316 (1994).
24. R. D. Hancock: *Pure Appl. Chem.* **58**, 1445 (1986).