

New Topological Index and Prediction of Phase Transfer Energy for Protonated Amines and Tetraalkylamines Ions

Lu Xu,* Yu-yan Yao, and Han-Ming Wang

Applied Spectroscopy Laboratory, Changchun Institute of Applied Chemistry,
Academia Sinica Changchun 130022, People's Republic of China

Received April 14, 1994[®]

The applications of new topological indices A_{x1} – A_{x3} suggested in our laboratory for the prediction of Gibbs energy values of phase transfer (water to nitrobenzene) of amine ions are described with satisfactory results. Multiple regression analysis and neural network were employed simultaneously in this study.

The essential assumption for quantitative structure–activity relationship (QSAR) or quantitative structure–property relationship (QSPR) studies is that the activities or properties depend on the structure of compounds. Thus, if we have methods to characterize appropriately the structures of compounds, the activities or properties of compounds can be quantitatively predicted in terms of the characterization. For this purpose, method of topological indices, i.e., graph-theoretical indices, is one of the simplest and useful methods. Over 100 indices including a large number of information-theoretical indices have been put forward to date, such as Wiener index W ,¹ Randic index ID ,² Hosoya index Z ,³ Balaban index J ,⁴ and the general a_N index.⁵ One significant development has been the increasing widespread usage of topological indices, a trend which has become of growing importance in recent years. Topological indices A_{x1} – A_{x3} based on the augmented distance-matrices devised recently by our laboratory have been successfully employed to the studies on structure–activity relationships for compounds of alkanes, alcohols, and barbiturates.⁶ In this study, we further utilized these indices to amine compounds for the prediction of Gibbs energy of phase transfer (water to nitrobenzene) with satisfactory results. The program was written in FORTRAN 77 and run on a micro VAX II.

NEURAL NETWORK ALGORITHM

The neural network has a three-layer, fully-connected, feed-forward 4:3:1 architecture. As indicated, the input layer with four nodes accepts the molecular descriptor values which numerically encode the features of each compound. The input signals are weighted as they are transmitted to the nodes of the second layer, the hidden layer. The three hidden layer neurons process the data and send a signal to the neurons of the output layer. The output layer provides the classification. A neural network is trained to relate certain inputs to target outputs. To accomplish this, a variety of neural network learning algorithms can be utilized. In this research, back-propagation (BP) and quasi-Newton methods were used. We found that the quasi-Newton method required fewer training cycles than did the back-propagation algorithm.

1. The Back-Propagation Training Algorithm. A

neuron in the hidden layer computes the weighted sum of inputs, net_j , given by

$$net_j = \sum w_{ij}x_i + \theta_j \quad (1)$$

where w_{ij} denotes the connection weight between node i in the input layer and node j in the hidden layer, x_i is the i th output from the input layer node i , and θ_j is the bias for node j . The output of node j in the hidden layer is calculated using a sigmoidal function

$$h_j = 1/(1 + e^{-\alpha(net_j)}), \quad \alpha > 0 \quad (2)$$

where α is referred to as the shape parameter. Similarly, the value feeding into an output layer node k is

$$net_k = \sum w_{jk}h_j + \theta_k \quad (3)$$

With the output of this node given by

$$o_k = 1/(1 + e^{-\alpha(net_k)}) \quad (4)$$

For training, the BP algorithm computes the sum-squared-error between the network output and the target values

$$E = \sum_p \sum_k (t_{pk} - o_{pk})^2 \quad (5)$$

where the number of training observations is denoted by p . The target value for pattern p is t_{pk} , and the computed value is o_{pk} .

The connection weights and the biases in the network are adjusted sequentially to reduce the error. This adjustment is made from the output layer to the hidden layer using a gradient descent method. In this method, the partial derivative of the error function is used to determine each weight adjustment, Δw_{ij} . If the neuron of interest is contained in the output layer, then the error is calculated from the difference between the output value and the target value multiplied by the derivative of the output value. The error terms for the hidden layer neurons are more complicated because the target values do not exist and must be calculated recursively from neurons already modified. To improve training time of the back-propagation algorithm and avoid the hazards of oscillating or becoming trapped in local minima, a momentum term is added to the weight adjustment

[®] Abstract published in *Advance ACS Abstracts*, August 15, 1994.

Table 1. Gibbs Energy (kJ/mol) of Transfer (Water to Nitrobenzene) and Indices $A_{\times 1}$ – $A_{\times 3}$ of 15 Alkylamines^a

no.	(R ¹ R ² R ³ R ⁴)N ⁺				$A_{\times 1}$	$A_{\times 2}$	$A_{\times 3}$	ΔG_{tr}° (exp.)	ΔG_{tr}° (reg)	ΔG_{tr}° (neural)
	R ¹	R ²	R ³	R ⁴						
1	H	H	H	Er	5.5234	5.8837	4.5238	−18.8	−18.3	−17.7
2	H	H	H	Pr	7.7035	8.4135	9.0300	−15.9	−15.2	−16.1
3 ^b	H	H	H	Bu	9.7652	11.5101	12.1033	−13.9	−11.9	−13.3
4	H	H	Me	Me	5.4997	5.9206	4.4997	−16.6	−18.3	−17.7
5	H	H	Et	Et	9.7292	11.4845	12.1412	−12.8	−11.9	−13.4
6	H	H	Pr	Pr	13.8448	16.7351	18.6670	−7.3	−5.9	−7.3
7 ^b	H	Me	Me	Me	7.8063	12.9951	6.3063	−11.9	−12.9	−11.7
8	H	Et	Et	Et	13.8448	19.0274	28.6292	−5.6	−6.0	−5.6
9	H	Pr	Pr	Pr	19.5991	26.8107	34.3306	2.4	3.3	3.0
10	Me	Me	Me	Me	10.0869	24.2182	8.0869	−6.4	−6.0	−6.5
11	Me	Me	Me	Et	12.0662	25.2548	20.5771	−3.9	−4.7	−4.1
12	Me	Me	Me	Pr	13.9627	27.5838	22.2227	−0.6	−1.7	−0.4
13 ^b	Me	Me	Me	Bu	15.8500	28.6771	23.6324	2.4	0.9	1.7
14	Me	Et	Et	Et	16.7614	47.5082	86.7694	0.7	1.9	0.7
15	Et	Et	Et	Et	17.9155	30.0055	54.2513	2.7	0.2	2.6

^a Notes: Me, methyl; Et, ethyl; Pr, propyl; Bu, butyl. ^b Member of test set.

Table 2. Regression Analysis Results of Table 1

indices	R	S	F	n
$A_{\times 1}$	0.9420	2.6306	102.46	15
$A_{\times 2}$	0.8867	3.6249	47.80	15
$A_{\times 3}$	0.6988	5.6073	12.41	15
$A_{\times 1}, A_{\times 2}$	0.9648	2.1453	80.79	15
$A_{\times 1}, A_{\times 3}$	0.9431	2.7141	48.23	15
$A_{\times 2}, A_{\times 3}$	0.8980	3.5902	24.99	15
$A_{\times 1}, A_{\times 2}, A_{\times 3}$	0.9851	1.4663	120.19	15

equation. The full details of BP training have been published previously.⁷

2. The Quasi-Newton (BFGS) Training Algorithm. The BFGS (Broyden–Fletcher–Goldfarb–Shanno)^{8–12} quasi-Newton optimization method is an alternative way to minimize the sum-squared-error of eq 5. The advantages of using the BFGS method over the BP method are that one need not specify a learning rate or momentum, and training progresses much more rapidly.

The basis of all quasi-Newton methods is that in cycle $k + 1$ of the optimization, the error E and gradient \mathbf{g}_{k+1} are assumed to be expressible as truncated Taylor series in the parameters \mathbf{x}

$$E(\mathbf{x}_{k+1}) \approx E(\mathbf{x}_k) + \mathbf{g}_k^T \Delta \mathbf{x}_k + \Delta \mathbf{x}_k^T \mathbf{H}_k \Delta \mathbf{x}_k \quad (6)$$

$$\mathbf{g}_{k+1} \approx \mathbf{g}_k + \mathbf{H}_k \Delta \mathbf{x}_k \quad (7)$$

where $\Delta \mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ is the change in the parameters from cycle k to cycle $k + 1$, and \mathbf{H}_k is the Hessian matrix of cycle k . The Hessian is defined as the matrix of the second derivatives of the error function with respect to the parameters. For a neural network, the parameters are just the weights and biases.

If \mathbf{x}_{k+1} is to correspond to a minimum of the error functional, then $\mathbf{g}_{k+1} = 0$. This leads to the “Newton” step formula

$$\Delta \mathbf{x}_k = -\mathbf{H}_k^{-1} \mathbf{g}_k \quad (8)$$

In quasi-Newton methods, the inverse Hessian matrix \mathbf{H}_k^{-1} is never computed directly. Instead, it is iteratively estimated and updated as the optimization proceeds. With an estimate \mathbf{G}_k of the inverse Hessian matrix, the steps in a BFGS cycle are as follows:

- (1) Choose search direction \mathbf{d}_k according to $\mathbf{d}_k = \mathbf{G}_k \mathbf{g}_k$
- (2) Determine the scalar α_k to minimize $E(\mathbf{x}_k + \alpha \mathbf{d}_k)$
- (3) Let $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{d}_k$
- (4) Compute the gradient \mathbf{g}_{k+1} corresponding to the parameters \mathbf{x}_{k+1}
- (5) Update the inverse Hessian matrix \mathbf{G} by the BFGS method
- (6) Iterate

The line minimization parameter α_k can be estimated using a parabolic fit of the error along \mathbf{d}_k

$$E(\mathbf{x}_k + \alpha \mathbf{d}_k) \approx E(\mathbf{x}_k) + a\alpha + b\alpha^2 \quad (9)$$

Two pieces of data are needed to find the constants a and b . Satisfactory results are obtained by using (a) the slope $\partial E / \partial \alpha$ at \mathbf{x}_k , which is given by $\mathbf{d}_k^T \mathbf{g}_k$ and (b) the $E(\mathbf{x}_k + s \mathbf{d}_k)$, where s is an appropriately chosen step size. The minimum of this curve occurs when $\alpha = -(2b)$.

When using any iterative training procedure, a criterion must be available for deciding when to stop the iterations. Three approaches were used in this study. (1) The weights were adjusted for each observation until the sum-squared-error reached an acceptable value for the entire training set. (2) The number of training cycles was limited, and training was stopped after a fixed number of training cycles had been reached. (3) Training was stopped when the minimum test set error was achieved.

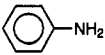
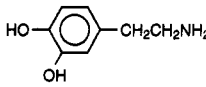
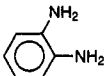
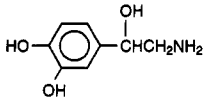
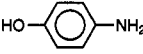
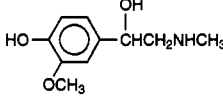
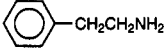
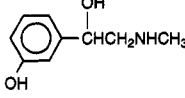
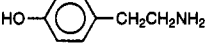
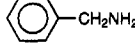
RESULTS AND DISCUSSION

The kinetic and thermodynamic parameters in studies of liquid–liquid interface electrochemistry for many systems are not easy to obtain because of the limits of potential windows. The measurement of Gibbs energy of phase transfer is usually performed through the determination of potential difference between ion interfaces for a set of given solvents. In this area, the liquid interface of water/nitrobenzene is regularly preferred to simulate biomembranes.

The Gibbs energy of phase transfer and kinetic as well as thermodynamic properties of drug molecules will play an important role for the studies of correlation and kinetic process of drugs.

1. Gibbs Energies of Phase Transfer of Alkylammonium Ions. The standard Gibbs energies of phase transfer

Table 3. Structures and Gibbs Energy (kJ/mol) of Transfer (Water to Nitrobenzene) of 10 Phenylamines

compd no.	structure	ΔG_{tr}°	compd no.	structure	ΔG_{tr}°
1		17.3	6		22.6
2		18.0	7		28.7
3		23.6	8		15.5
4		9.65	9		19.0
5		17.6	10		12.0

(water to nitrobenzene) of 15 mono- to tetraalkylammonium ions (Table 1) have been measured using electrochemical methods,¹³ such as cyclic voltammetry and differential stripping voltammetry. The objective of this paper is to use the indices A_{x1} – A_{x3} to model the standard Gibbs energies of phase transfer of these compounds shown in Table 1.

(1) Results with Multiple Regression Analysis. Table 2 shows a summary of the correlation coefficients, R , standard deviations, s , and F -test values by various combinations of A_{x1} – A_{x3} . It can be seen from Table 2 that the best result has been achieved by using the combination of A_{x1} – A_{x3} simultaneously. Therefore, A_{x1} – A_{x3} will be taken as the variables in the subsequent sections.

The 15 compounds were randomly divided into two groups: a training set containing 12 compounds and a test set containing three (compounds labeled footnote a in Table 1, i.e., compounds 3, 7, 12, and 13). The mathematical model was obtained using multiple regression analysis

$$\Delta G_{tr}^\circ = -26.6651 + 1.1975A_{x1} + 0.3950A_{x2} - 0.1182A_{x3} \quad (10)$$

$$R = 0.9862 \quad F = 94.4461 \quad s = 1.4806 \quad n = 12$$

where R is the correlation coefficient, F denotes the F -test value, s is standard deviation, and n represents the number of samples. Pairwise correlations of A_{x1} – A_{x3} were observed. The correlation coefficients between A_{x1} and A_{x2} and A_{x1} and A_{x3} as well as A_{x2} and A_{x3} are 0.863, 0.747, and 0.868, respectively.

The Gibbs energies calculated by eq 1 for training set and test set are also listed in Table 1, column $\Delta G_{tr}^\circ(\text{reg})$. The standard deviation to test set is 1.9548.

(2) Results with Neural Network. The architecture of the neural network in this paper is as follows.

(a) The number of layers is three, i.e., layers input, hidden, and output.

(b) The input neurons are A_{x1} – A_{x3} . Usually, the value of each neuron is defined between 0 and 1, thus the input data should be scaled within the defined region. Note that if the value of a neuron in input layer is zero, the connections from this neuron are always zero, i.e., the information from that neuron cannot be propagated to the following layers. To avoid this situation, the values were set between 0.05 and 0.95 in our research.

(c) As an usual rule of thumb, the total number of weights including the biases is to limit less than the number of the observations. Finding a compromise between the performance and the training time of the network, the optimum number of neurons in hidden layer for this application was found to be two.

(d) The number of output neuron is one. There are two ways to express the output neurons. Typically, single output neuron networks are used usually to predict continuous values. In this manner, the neural network is shown to be analogous to nonlinear regression analysis. When multiple neurons are used in output layer, the neural networks are employed for pattern recognition. Therefore, in our research, single output neural networks were used.

The initial weights are usually important for the training of a network. The results may be quite different with different sets of initial weights. The initial weights were generated randomly. To obtain the best model, the set of random starting weights needs to be determined. This is a very difficult procedure and requires training the network several times using different sets of starting weights.

It is noteworthy that owing to the tremendous number of adjustable parameters, extreme caution must be taken to avoid overtraining. In this study we utilized a test set to monitor the training process. In this way, the performance of the neural network can be enhanced markedly.¹⁴

The 15 compounds were divided into the same two sets as in the regression analysis above. The calculated results by the best model obtained using network are also listed in Table 1, column $\Delta G_{tr}^\circ(\text{neural})$. The standard deviations for training set and test set are 0.5997 and 0.6531, respectively, which all are less than the deviations obtained by using regression analysis.

2. Gibbs Energies of Phase Transfer of Protonated Amines Having Phenyl Rings. It is well-known that an amine molecule in organic phases may act as a proton acceptor and thereby facilitate proton transfer across the water/organic solvent interface.

Recently, the study on partition of a series of protonated amines with aromatic rings, between water and nitrobenzene, was reported by Homka et al.¹⁵ using an electrochemical

Table 4. Logarithm Gibbs Energy (kJ/mol) and Indices A_{x1} – A_{x3} of the Diamines in Table 3

compound no. ^a	A_{x1}	A_{x2}	A_{x3}	$\log \Delta G_{tr}^{\circ}$ (exp.)	$\log \Delta G_{tr}^{\circ}$ (reg)	$\log \Delta G_{tr}^{\circ}$ (neural)
1 ^b	18.26	27.07	26.76	1.238	1.077	1.235
2	20.11	30.77	38.62	1.255	1.221	1.353
3	20.51	33.39	57.63	1.373	1.449	1.381
4	22.05	31.61	35.62	0.984	1.041	0.991
5 ^b	23.81	34.94	42.83	1.243	1.156	1.199
6	25.57	38.69	52.22	1.354	1.317	1.355
7	27.33	41.46	61.34	1.458	1.371	1.363
8	30.86	45.23	67.89	1.190	1.287	1.227
9	27.42	40.29	53.02	1.279	1.232	1.233
10	20.18	29.39	33.42	1.086	1.062	1.106

^a The number corresponds to Table 3. ^b Member of test set.

Table 5. Gibbs Energy (kJ/mol) of Transfer (Water to Nitrobenzene) and Indices A_{x1} – A_{x3} of Seven Protonated Diamines

compound	A_{x1}	A_{x2}	A_{x3}	ΔG_{tr}° (exp.)	ΔG_{tr}° (reg)
ethylenediamine	7.57	8.28	8.80	20.90	19.41
N-methylethylenediamine	9.60	11.37	11.93	18.50	17.63
N-ethylethylenediamine	11.60	14.03	14.55	14.60	15.88
N,N-dimethylethylenediamine	11.70	17.30	16.68	16.10	15.79
N-propylethylenediamine	13.57	16.60	18.50	12.30	14.16
2,2'-dipyridyl	31.93	45.14	57.62	-1.70	-1.91
phenanthroline	38.04	57.67	84.31	-7.00	-7.27

approach. Table 2 shows these Gibbs energies and A_{x1} – A_{x3} of the compounds. Similar to the previous section, the 10 compounds shown in Table 3 were randomly divided into two sets, a training set containing eight compounds and a test set containing two compounds (see Table 4, labeled footnote a, *i.e.*, compounds 1 and 5). Statistical analysis has been performed by using various combinations of A_{x1} – A_{x3} . The better result can be obtained using a combination of A_{x1} and A_{x2} .

$$\log \Delta G_{tr}^{\circ} = 0.6958 - 0.1395A_{x1} + 0.1082A_{x2} \quad (11)$$

$$R = 0.9040 \quad F = 11.1796 \quad s = 0.07912 \quad n = 8$$

The calculated $\log \Delta G_{tr}^{\circ}$ values for a training set and a test set are listed in Table 4. The standard deviation to the test set is 0.1815. If we use A_{x1} – A_{x3} simultaneously, the regression model will be

$$\log \Delta G_{tr}^{\circ} = 0.2007 - 0.2579A_{x1} + 0.2281A_{x2} + 0.01980A_{x3} \quad (12)$$

$$R = 0.9692 \quad F = 20.6582 \quad s = 0.05101 \quad n = 8$$

Although the regression analysis can be improved by introduction of A_{x3} , the weight of A_{x3} is small for eq 12. The standard deviation to test set is 0.1197. Similarly, pairwise correlations of A_{x1} – A_{x3} for this situation were also observed, and it was found that the correlation coefficients between A_{x1} and A_{x2} and A_{x1} and A_{x3} as well as A_{x2} and A_{x3} are 0.981, 0.823, and 0.911, respectively. Obviously, A_{x1} to A_{x2} and A_{x2} to A_{x3} possess slightly high correlation coefficients.

The architecture of neural network for this group is 2:2:1, *i.e.*, the number of layers is three and the numbers of neurons in input, hidden, and output layers are 2, 2, and 1, respectively.

The predicted $\log \Delta G_{tr}^{\circ}$ values calculated by the model achieved using neural network are shown in Table 4. The standard deviations for the training set and the test set are 0.067 60 and 0.044 46, respectively, which also all are less

than the result from regression analysis, *i.e.*, more accurate results can be obtained using neural network.

3. Gibbs Energies of Phase Transfer of Diamines. Studies on thermodynamic properties of phase transfer of a series of nitrogen-containing complexes to drive protons to transfer across the water/nitrobenzene interface have been performed by Baruzzi et al.¹⁶ using electrochemical techniques. Table 5 shows the Gibbs energies and the indices A_{x1} – A_{x3} of these protonated diamines. In this group, the structural characteristics among the compounds are quite different, *e.g.*, compounds 6 and 7 are nitrogen-containing heterocycles, whereas the others are aliphatic compounds. Even so, satisfactory results can be obtained from regression analysis using only a single index; therefore, no combination of A_{x1} – A_{x3} and no neural network were used. The best regression analysis is

$$\Delta G_{tr}^{\circ} = 26.0385 - 0.8754A_{x1} \quad (13)$$

$$r = -0.9939 \quad F = 405.0800 \quad s = 1.2868 \quad n = 7$$

The correlation coefficients using A_{x2} and A_{x3} are -0.9892 and -0.9796, respectively.

The Gibbs energies calculated by eq 13 are listed in Table 5.

CONCLUSION

Topological indices A_{x1} – A_{x3} recently introduced by us have been successfully used to model the physicochemical properties of alkanes and alcohols, such as the boiling point, and to predict the activities of barbiturates as well as to predict the Gibbs energies of phase transfer of a variety of protonated amines with satisfactory results. The experiments reveal that the results obtained using a new quasi-Newton neural network is more accurate than that achieved using regression analysis. The study demonstrates convincingly again that A_{x1} – A_{x3} are useful topological indices.

ACKNOWLEDGMENT

We thank Prof. P. C. Jurs (Department of Chemistry, The Penn State University, U.S.A.) very much for providing us

with program neural network. The financial support of the National Natural Science Foundation of China is gratefully acknowledged.

REFERENCES AND NOTES

- (1) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
- (2) Hosoya, H. Topological Index, A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332–2339.
- (3) Randic, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
- (4) Balaban, A. T. Topological Indices Based on Topological Distances in Molecular Graphs. *Pure Appl. Chem.* **1983**, *55*, 199–206.
- (5) Xu, L.; Wang, H. Y.; Su, Q. A Newly Proposed Molecular Topological Index for the Discrimination of cis/trans Isomers and for the Studies of QSAR/QSPR. *Comput. Chem.* **1992**, *16*, 187–194.
- (6) Yao, Y. Y.; Xu, L.; Yang, Y. Q.; Yuan, X. S. Study Structure–Activity Relationships of Organic Compounds: Three New Topological Indices and Their Applications. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 590–594.
- (7) Janson, P. A. Neural Network: An Overview. *Anal. Chem.* **1991**, *63*, 357A–362A.
- (8) Broyden, C. G. The Convergence of a Double-Rank Minimization Algorithms. Part 1 and 2. *J. Inst. Maths. Appl.* **1970**, *6*, 76–90; **1970**, *6*, 222–231.
- (9) Fletcher, R. A New Approach to Variable Metric Algorithms. *Comp. Jour.* **1970**, *13*, 317–322.
- (10) Goldfarb, D. A Family of Variable-Metric Methods Derived by Variational Means. *Math. Comp.* **1970**, *24*, 23–26.
- (11) Shanno, D. F. Conditioning of Quasi-Newton Methods for Function Minimization. *Math. Comp.* **1970**, *24*, 647–656.
- (12) Fletcher, R. *Practical Methods of Optimization*; Wiley: New York, 1980; Vol. 1.
- (13) Vanijsek, P. *Electrochemistry on Liquid/Liquid Interfaces*. In *Lecture Notes in Chemistry*; Springer-Verlag: Berlin, 1985.
- (14) Xu, L.; Ball, J. W.; Dixon, S. L.; Jurs, P. C. Quantitative Structure–Activity Relationships for Toxicity of Phenols Using Regression Analysis and Computational Neural Network. *Envir. Tox. Chem.* **1994**, *13*, 841–851.
- (15) Homlka, D.; Marecek, V.; Samec, Z. The Partition of Amines between Water and Organic Solvent Phase. *J. Electroanal. Chem.* **1984**, *163*, 159–170.
- (16) Baruzzi, A. M.; Wendt, H. Interfacial Phase Transfer of Amino-Complexed Ni(II) Cations. A Model for Solvent Extraction of Transition Metals. *J. Electroanal. Chem.* **1990**, *279*, 19–30.

CI940044P