

# Prediction of Atomic Ionization Potentials I-III Using an Artificial Neural Network

Michael E. Sigman\* and Stephen S. Rives†

Chemical and Analytical Sciences Division, Oak Ridge National Laboratory,  
P.O. Box 2008, MS 6100, Oak Ridge, Tennessee 37831-6100

Received September 27, 1993\*

A simple three-layer back-propagation neural network is shown to learn the complex relationship between the electronic structure and the first three ionization potentials ( $\log\{\text{IP}(\text{eV})\}$ ) of 222 atoms and ions for which spectroscopic data have been determined. The neural network was trained to a root mean square error of 0.043 log units and was subsequently used to predict ionization potentials which were not included in the training set and values which have not been experimentally determined. The neural network predictions are in very good agreement with experimental values not included in the training data set. In addition, the neural network predicts IP for elements 73, 104, and 105 which are also in good agreement with those values predicted by highly sophisticated multiconfiguration Dirac-Fock (MCDF) calculations.

## INTRODUCTION

The application of neural networks to modeling problems in chemistry is a relatively new area of research. Some applications reported to date include quantitative structure-activity and structure-property relationships,<sup>4-8</sup> spectroscopic modeling,<sup>9-21</sup> and quantum mechanical<sup>22-26</sup> modeling as well as others.<sup>27,28</sup> Neural networks are useful when a mathematical relationship is not available to describe a phenomenon to be modeled. If the property in question can be modeled by very complex and highly demanding computational techniques (ca. *ab initio* quantum mechanical calculations), neural networks may provide an alternative approach to obtaining accurate numerical values in a computationally less intensive fashion. The trade-off is that a clear physical interpretation of the factors which influence the quantity of interest is not generally available from neural network models at this time.

One problem that is of interest to chemists, and one which is amenable to a neural network model, is the prediction of physical properties for naturally occurring and man-made elements. Some properties, such as ionization potential (IP), cannot be directly measured for short-lived elements. Computational work directed toward prediction of the first few ionization potentials of heavy elements has recently utilized MCDF calculations; however, these calculations are so computationally intensive as to be impractical for some elements.<sup>2,3</sup> We demonstrate in this paper that a simple three-layer neural network can learn the complex relationship between IP and a single electronic configuration for the atom or ion in question. The neural network model does not offer the physical insight which the quantum mechanical approach does; however, the advantages of the neural net model include the rapid prediction of a large number of previously unmeasured IP, a better estimate of the error associated with the predictions due to a larger training set and a much less computationally intensive method.

## RESULTS AND DISCUSSION

The neural network employed for this modeling was a fully-connected three layer network trained by back-propagation

of error as outlined by Eberhart and Dobbins.<sup>29</sup> The sigmoidal transfer function used in the network is given by eq 1. In eq

$$O_j = \frac{1}{[1 + \exp(-\sum_i W_{ji} O_i)]} \quad (1)$$

1,  $O_j$  is the output of neurode  $j$  and  $W_{ji}$  is the weight connecting  $O_j$  with the output of neurode  $i$ ,  $O_i$ . Feed-forward calculations and error back-propagation were done by standard techniques previously described elsewhere with learning rates and momentum factors both set at 0.04.<sup>29,30</sup> The software used in this study was written in-house and bench-marked for speed and accuracy against commercial software available from Eberhart and Dobbins.<sup>29</sup> All calculations were done on an 80486 personal computer running at 25 MHz. One thousand training epochs required approximately 2.5 min to compute. A single epoch involved 222 feed-forward calculations and the associated error back-propagation.

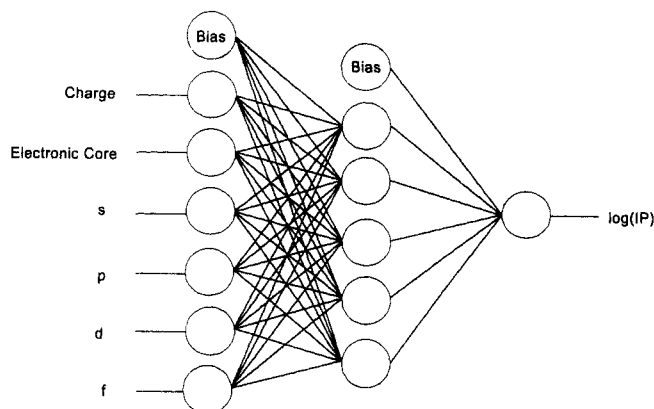
Large data sets are required for training neural networks; otherwise the network can memorize the data and may not perform well when generalizing to new examples. A minimum value of three has recently been suggested for the ratio of training examples to weights.<sup>29</sup> The data set for these calculations included 232 experimental values which have been reported for the first three IP of elements from H (element 1) to Am (element 95).<sup>1</sup> Of these values, 222 were used to train the network and 10 were used as a cross-validation set to ensure that the net was generalizing. The cross-validation set consisted of the following elements (charges): Se (0), Ac (0), Al (1), Sr (1), Pd (1), Yb (1), B (2), Mg (2), Ce (2), and Th (2).

The electronic configuration of an atom or ion and the associated atomic charge are natural choices for parametrizing a model of IP because of the fundamental role which they play in controlling this and other atomic properties. These values were used as input parameters for training the network after the appropriate normalization (*vide infra*). Electronic configurations which have previously been reported were mapped to the first IP for all uncharged atoms in the data set.<sup>31,32</sup> The electronic configurations for ions, which were mapped to the second and third IP, were derived from the configurations used for the uncharged species. These new

\* To whom correspondence should be addressed.

† Oak Ridge Semester Engineering Research Student, visiting from Southwest Baptist University, Bolivar, MO.

• Abstract published in *Advance ACS Abstracts*, March 1, 1994.



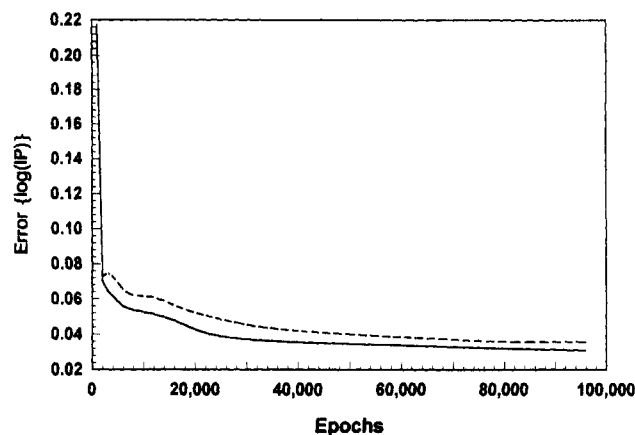
**Figure 1.** Schematic of the neural network used for predicting the  $\log\{IP(eV)\}$  values in Table 1. Six inputs included the atomic charge, the electronic core (see text), and the number of s, p, d, and f valence electrons. The single hidden layer has five nodes, and there is a single output. The input and hidden layers each have a bias.

configurations were generated by removal of highest lying electrons, following the principles outlined by Huheey.<sup>31</sup>

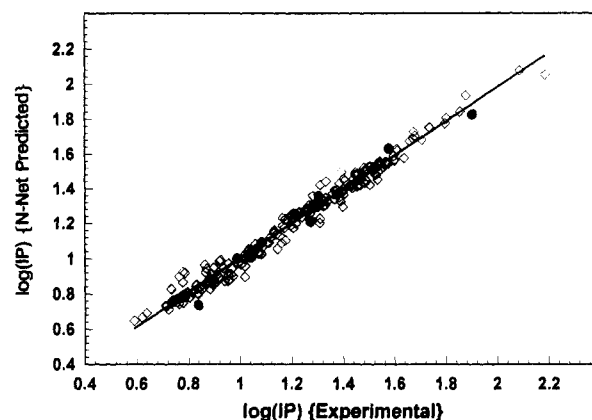
The total electronic configuration used as input to the network consisted of specifying both a core and a valence contribution. The core configuration for each atom or ion was assigned a numerical value of 0 (for species with fewer electrons than He), 1 (He), 2 (Ne), 3 (Ar), 4 (Kr), 5 (Xe), or 6 (Rn). The valence configuration was designated by giving the number of s, p, d, and f electrons occupying the respective orbitals. No designation of period was made (e.g. no differentiation was made between 3p and 4p electrons). This information is encoded in the core electronic configuration specified for each atom or ion. Although d and f electrons are never removed in the ionization processes under consideration here, they are included in the model because of the role they play in controlling IP through nuclear screening effects. The input data for each atom or ion consisted of a list of six numbers representing the atomic charge, core configuration, s, p, d, and f electron counts. For example, the input for the first IP of Pu (element 94), which has an electron configuration of  $[Rn]5f^67s^2$ , was 0, 6, 2, 0, 0, 6. The electron configurations and ionization potentials used in the training set are available as supplemental material.

Data normalization is often observed to be a critical factor in determining how well a neural network learns the associated input and output patterns.<sup>29</sup> A simple normalization of each independent input parameter proved to be sufficient in this example. The charge associated with each species was normalized independently from the number of s electrons for each and so on. For example, the charge on each atom or ion before the ionization in question ranged from 0 to 2; therefore, the parameter of charge was normalized over this range. Similarly, the number of s electrons can vary from 0 to a maximum of 2, and this parameter (the count of valence s electrons) was normalized over the range of 0 to 2. Analogous normalizations were used for the number of p, d, and f electrons and the core electron configuration using the range of 0 to 6, as previously discussed. The  $\log\{IP(eV)\}$  values were also normalized independently over a range from 0.5 to 2.5. Alternative normalization procedures were not explored.

Several network architectures were investigated, and the simplest one which gave an error for the cross-validation set which approximated that of the training set throughout the training process is depicted in Figure 1. The network had six inputs (*vide supra*) and five nodes in a single hidden layer. Biases were added to both the input and hidden layers. The

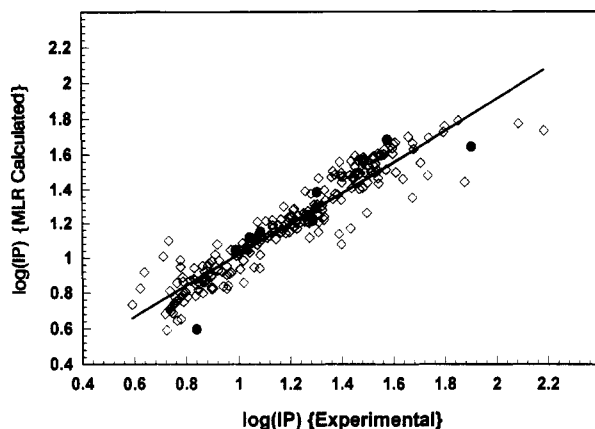


**Figure 2.** Error in the predicted  $\log\{IP(eV)\}$  values for the training (solid line) and cross-validation (dashed line) data sets over the 96 000 epochs required to train the network.



**Figure 3.** Experimental versus neural network predicted values of  $\log\{IP(eV)\}$  for the training set (open squares) and the cross-validation set (filled circles).

network had 41 adjustable weights, giving a 5.4:1 ratio of training examples to weights which exceeds the recommended ratio of 3.<sup>29</sup> The net was trained to a root mean square (RMS) error of 0.043 log units which is slightly larger than 2% of the total range of  $\log\{IP(eV)\}$  in the training set (0.5–2.5 *vide supra*). A plot of the error for both the training and cross-validation sets through the 96 000 epochs required to train the network is shown in Figure 2. The RMS error for the cross-validation set, 0.052 log units, was slightly higher than the error for the training set. Attempts to train the network with larger values for the momentum factor and learning rate led to minimization of the error for the cross-validation after a much smaller number of epochs and larger RMS error for the resulting model. The use of small values for these parameters, 0.04 (*vide supra*), led to longer training times but resulted in a smaller RMS error. The error for the cross-validation set increased if training was continued beyond the 96 000 epochs shown in Figure 2. A plot of the experimental  $\log\{IP(eV)\}$  values versus the neural network calculated values for both the training and cross-validation sets is shown in Figure 3. The best fit line through the data, slope = 0.9764 and intercept = 0.0276, has a correlation coefficient of 0.9898 ( $n = 222$ ). It is evident from the plot that the quality of the model is quite good and that the neural network has successfully generalized to calculate the IP values for the test set. It is also notable that the network is associating the measured IP with a single electronic configuration as opposed to multiple configurations used in MCDF calculations. In addition, there are no systematic errors in the predicted values with increasing  $\log\{IP(eV)\}$ , indicating that the values



**Figure 4.** Experimental versus MLR predicted values of  $\log\{IP(eV)\}$  for the training set (open squares) and the cross-validation set (filled circles).

for heavy elements are predicted with no less accuracy than those for light atoms. This is an impressive result given the simple nature of the network architecture. Multiple linear regression, MLR, with the same parameter set, yields a much poorer model of the data as shown in Figure 4. The correlation coefficient for the plot of experimental versus calculated  $\log\{IP(eV)\}$  from the MLR model is only 0.9417 ( $n = 222$ ). No attempt was made to find an optimal parameter set for the MLR model.

The trained network was then used to predict the missing values of the first three ionization potentials not included in the training set for elements up through Unq (element 106), (see Table 1). Those values missing from Table 1 correspond to IP that were included in the training set. Some calculated  $\log\{IP(eV)\}$  values reported in Table 1 have also been experimentally determined (reported in parentheses) and serve as a further test of the network model.<sup>33</sup> The calculated and experimental values in Table 1 are seen to be in fairly good agreement and have an RMS error of 0.039 log units which is comparable to the RMS error of 0.43 log units for the training set. Table 1 also shows a comparison between the neural network predictions and the values calculated by MCDF for elements 73, 104, and 105. The neural network predictions for the  $\log\{IP(eV)\}$  values are seen to compare very well with the values predicted by MCDF calculations,<sup>2,3</sup> giving additional confidence in the neural net model.

The  $\log\{IP(eV)\}$  values predicted for a given ionization potential (I, II, or III) in Table 1 show small variations in some cases. This is due to these atoms or ions falling within the lanthanides or actinides (atomic numbers 58–71 and 90–103, respectively) and the subtle changes in IP values that result from addition of  $f$  electrons while progressing through a given period. This periodic trend is observed for the experimentally measured ionization potentials of the lanthanides.<sup>1,33</sup> The trend is also observed for the predicted values in Table 1. More importantly, the model does a good job of predicting the change in successive IP values for a given element. In this regard, the total range in predicted  $\log\{IP(eV)\}$  values in Table 1 (0.92 log units) spans approximately half of the 2 log unit range in the training set, and the predicted values do not exceed the upper or lower limits of the training set. The predicted values in Table 1 are therefore consistent with observed periodic trends in measured IP values, and the associated error is small relative to the total range of predicted values in Table 1.

**Table 1.** Neural Network Predicted Values of  $\log\{IP(eV)\}$  for Ionization Potentials I–III<sup>a</sup>

symbol	atomic no.	$\log\{IP(eV)\}$		
		I	II	III
Cs	55			1.55
Ba	56			1.49
Nd	60			1.43 (1.34 <sup>b</sup> )
Pm	61			1.42 (1.35 <sup>b</sup> )
Sm	62			1.41 (1.37 <sup>b</sup> )
Eu	63			1.40 (1.40 <sup>b</sup> )
Gd	64			1.35 (1.31 <sup>b</sup> )
Tb	65			1.39 (1.34 <sup>b</sup> )
Dy	66			1.38 (1.36 <sup>b</sup> )
Ho	67			1.38 (1.36 <sup>b</sup> )
Er	68			1.37 (1.36 <sup>b</sup> )
Lu	71			1.36 (1.32 <sup>b</sup> )
Ta	73		1.20 (1.19 <sup>d</sup> )	1.43 (1.38 <sup>d</sup> )
W	74		1.22	1.46
Re	75		1.23	1.49
Os	76		1.24	1.51
Ir	77		1.25	1.52
Pt	78			1.53
Au	79			1.54
Po	84		1.27	1.47
At	85	1.02	1.30	1.49
Rn	86		1.32	1.52
Fr	87	0.63	1.20	1.54
Ra	88			1.41
Ac	89			1.31
Th	90	0.75 (0.78 <sup>b</sup> )		
Pa	91	0.75 (0.77 <sup>b</sup> )	1.04	1.30
U	92	0.75 (0.78 <sup>b</sup> )	1.05	1.30
Np	93	0.75 (0.79 <sup>b</sup> )	1.06	1.30
Pu	94		1.02	1.35
Am	95		1.02	1.34
Cm	96	0.77 (0.78 <sup>b</sup> )	1.08	1.31
Bk	97	0.76 (0.79 <sup>b</sup> )	1.04	1.33
Cf	98	0.76 (0.80 <sup>b</sup> )	1.05	1.33
Es	99	0.77 (0.81 <sup>b</sup> )	1.05	1.32
Fm	100	0.77 (0.81 <sup>b</sup> )	1.06	1.32
Md	101	0.78 (0.82 <sup>b</sup> )	1.07	1.32
No	102	0.78 (0.82 <sup>b</sup> )	1.07	1.32
Lr	103	0.80	1.12	1.33
Unq	104	0.82 (0.81 <sup>c</sup> )	1.16 (1.17 <sup>c</sup> )	1.37 (1.38 <sup>c</sup> )
Unp	105	0.83 (0.87 <sup>d</sup> )	1.18 (1.21 <sup>d</sup> )	1.41 (1.39 <sup>d</sup> )
Unh	106	0.83	1.20	1.45

<sup>a</sup> Missing values were included in the training and test data sets.

<sup>b</sup> Spectroscopic data.<sup>33</sup> <sup>c</sup> MCDF calculated value.<sup>2</sup> <sup>d</sup> MCDF calculated value.<sup>3</sup>

## CONCLUSIONS

We have shown that a simple artificial neural network can learn the relationship between ionization potentials and electronic structure for atoms and ions. The neural network permits rapid prediction of a large number of previously uncalculated and unmeasured ionization potentials. Estimation of error in the calculated IP is based on the quality of the model for 222 training examples and substantiated by the similar error for 10 cross-validation examples and 23  $\log\{IP(eV)\}$  values predicted in Table 1. The network training process is not computationally demanding, and the results are in very good agreement with experimental values not included in the training set and with values previously calculated by much more sophisticated quantum mechanical methods.

## ACKNOWLEDGMENT

The authors would like to thank Elijah Johnson, Chemical and Analytical Sciences Division, Oak Ridge National Laboratory (ORNL), for informative discussions concerning MCDF calculations. ORNL is managed by Martin Marietta Energy Systems under Contract DE-AC05-84OR21400 with U. S. Department of Energy.

**Supplementary Material Available:** A table listing training and cross-validation sets (6 pages). Ordering information is given on any current masthead page.

## REFERENCES AND NOTES

- (1) *CRC Handbook of Chemistry and Physics*, 63rd ed.; Weast, R. C., Astle, M. J., Eds.; CRC Press: Boca Raton, FL, 1982; pp E64-E65.
- (2) Johnson, E.; Fricke, B.; Keller, O. L., Jr.; Nestor, C. W., Jr.; Tucker, T. C. Ionization Potentials and Radii of Atoms and Ions of Element 104 (Unnilquadium) and of Hafnium (2+) Derived from Multiconfiguration Dirac-Fock Calculations. *J. Chem. Phys.* **1990**, *93*, 8041-8050.
- (3) Fricke, B.; Johnson, E.; Rivera, G. M. Ionization Potentials and Radii of Atoms and Ions of Element 105 (Unnilpentium) and Ions of Tantalum Derived from Multiconfiguration Dirac-Fock Calculations. *Radiochim. Acta* **1993**, *62*, 17-25.
- (4) Egolf, M.; Jurs, P. C. Prediction of Boiling Points of Organic Heterocyclic Compounds Using Regression and Neural Network Techniques. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 616-625.
- (5) Aoyama, T.; Ichikawa, H. Neural Networks as Nonlinear Structure-Activity Relationship Analyzers. Useful Functions of the Partial Derivative Method in Multilayer Neural Networks. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 492-500.
- (6) Bodor, N.; Hargert, A.; Huang, M.-J. Neural Network Studies. 1. Estimation of the Aqueous Solubility of Organic Compounds. *J. Am. Chem. Soc.* **1991**, *113*, 9480-9483.
- (7) Andrea, T. A.; Kalayeh, H. Applications of Neural Networks by Quantitative Structure-Activity Relationships of Dihydrofolate Reductase Inhibitors. *J. Med. Chem.* **1991**, *34*, 2824-2836.
- (8) Aoyama, T.; Suzuki, Y.; Ichikawa, H. Neural Networks Applied to Quantitative Structure-Activity Relationship Analysis. *J. Med. Chem.* **1990**, *33*, 2583-2590.
- (9) Anker, L. S.; Jurs, P. C. Prediction of Carbon-13 Nuclear Magnetic Resonance Chemical Shifts by Artificial Neural Networks. *Anal. Chem.* **1992**, *64*, 1157-1164.
- (10) Long, J. R.; Mayfield, H. T.; Henley, M. V.; Kromann, P. R. Pattern Recognition of Jet Fuel Chromatographic Data by Artificial Neural Networks with Back-Propagation of Error. *Anal. Chem.* **1991**, *63*, 1256-1261.
- (11) West, G. M. J. Predicting Phosphorous NMR Shifts Using Neural Networks. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 577-589.
- (12) Meyer, B.; Hansen, T.; Nute, B.; Albersheim, P.; Darvill, A.; York, W.; Sellers, J. Identification of the <sup>1</sup>H NMR Spectra of Complex Oligosaccharides with Artificial Neural Networks. *Science* **1991**, *251*, 542-544.
- (13) Peterson, K. Counter-Propagation Neural Networks in the Modeling and Prediction of Kovats Indices for Substituted Phenols. *Anal. Chem.* **1992**, *64*, 379-386.
- (14) Tanabe, K.; Tamura, T.; Uesaka, H. Neural Network System for the Identification of Infrared Spectra. *Appl. Spectrosc.* **1992**, *46*, 807-810.
- (15) Bell, S. E.; Mead, W. C.; Jones, R. D.; Eiceman, G. A.; Ewing, R. G. Connectionist Hyperprism Neural Network for the Analysis of Ion Mobility Spectra: An Empirical Evaluation. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 609-615.
- (16) Doucet, J. P.; Panaye, A.; Feuillebois, E.; Ladd, P. Neural Networks and <sup>13</sup>C NMR Shift Prediction. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 320-324.
- (17) Ricard, D.; Cachet, C.; Cabrol-Bass, D.; Forrest, T. P. Neural Network Approach to Structural Feature Recognition from Infrared Spectra. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 202-210.
- (18) Weigel, U.-M.; Herges, R. Automatic Interpretation of Infrared Spectra: Recognition of Aromatic Substitution Patterns Using Neural Networks. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 723-731.
- (19) Kvanicka, V.; Sklenak, S.; Pospichal, J. Application of Recurrent Neural Networks in Chemistry. Prediction and Classification of <sup>13</sup>C NMR Chemical Shifts in a Series of Monosubstituted Benzenes. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 742-747.
- (20) McAvoy, T. J.; Su, H. T.; Wang, N. S.; He, M.; Horvath, J.; Semerjian, H. A Comparison of Neural Networks and Partial Least Squares for Deconvoluting Fluorescence Spectra. *Biotechnol. Bioeng.* **1992**, *40*, 53-62.
- (21) Allanic, A. L.; Jexquel, J. Y.; Andre, J. C. Application of Neural Networks Theory to Identify Two-Dimensional Fluorescence Spectra. *Anal. Chem.* **1992**, *64*, 2618-2622.
- (22) Noid, D. W.; Varma-Nair, M.; Wunderlich, B.; Darsey, J. A. Neural Network Inversion of the Tarasov Function Used for the Computation of Polymer Heat Capacities. *J. Therm. Anal.* **1991**, *37*, 2295-2300.
- (23) Sumpter, B. G.; Noid, D. W. Potential Energy Surfaces for Macromolecules. A Neural Network Technique. *Chem. Phys. Lett.* **1992**, *192*, 455-461.
- (24) Sumpter, B. G.; Getino, C.; Noid, D. W. Neural Network Predictions of Energy Transfer in Macromolecules. *J. Phys. Chem.* **1992**, *96*, 2761-2767.
- (25) Sumpter, B. G.; Getino, C.; Noid, D. W. A Neural Network Approach to the Study of Internal Energy Flow in Molecular Systems. *J. Chem. Phys.* **1992**, *97*, 293-306.
- (26) Noid, D. W.; Darsey, J. A. Neural Net Simulation of Polymer Dynamics. *Comput. Polym. Sci.* **1991**, *1*, 157-160.
- (27) Zupan, J.; Gasteiger, J. Neural Networks: A New Method for Solving Chemical Problems or Just a Passing Phase? *Anal. Chim. Acta* **1991**, *248*, 1-30.
- (28) Maggiora, G. M.; Elrod, D. W.; Trenary, R. G. Computational Neural Networks as Model-Free Mapping Devices. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 732-741.
- (29) Eberhart, R. C.; Dobbins, R. W. *Neural Network PC Tools, A Practical Guide*; Academic Press: New York, 1990.
- (30) Wasserman, P. D. *Neural Computing: Theory and Practice*; Van Nostrand Reinhold: New York, 1989. Blum, A. *Neural Networks in C++: An Object-Oriented Framework for Building Connectionist Systems*; Wiley: New York, 1992.
- (31) Huheey, J. E. *Inorganic Chemistry, Principles of Structure and Reactivity*, 2nd ed.; Harper & Row: New York, 1978; p 30.
- (32) Table of Periodic Properties of the Elements; Sargent-Welch Scientific Co.: Skokie, IL, 1979.
- (33) Martin, W. C.; Hagan, L.; Reader, J.; Sugar, J. Ground Levels and Ionization Potentials for Lanthanide and Actinide Atoms and Ions. *J. Phys. Chem. Ref. Data* **1974**, *3*, 771-779.