Tetrahedron Vol. 48, No. 17, pp. 3463-3472, 1992
Printed in Great Britain

0040-4020/92 $5.00+.00

NEURAL NETWORK STUDIES 2.[1] USE OF A NEURAL NET TO ESTIMATE OXIDATION ENERGIES FOR SUBSTITUTED DIHYDROPYRIDINES AND RELATED HETEROCYCLES

Marcus E. Brewster,[a,b] Ming-Ju Huang,[b] Alan Harget[b,c] and Nicholas Bodor[a,b,2]

[a] Pharmatec, Inc., PO Box 730, Alachua, FL 32615, [b] Center for Drug Discovery, University of Florida, Gainesville, FL 32610, and [c] University of Aston, Birmingham, England

(*Received in USA* 26 *February* 1992)

Abstract: A perceptron-type neural net was applied to the prediction of oxidation energies for dihydropyridines and related heterocycles. A set of 71 energies corresponding to heats of formation differences for reduced and oxidized derivatives were used to train the net using a back propagation algorithm. A system using four hidden units produced good duplication of the input values (SD = 0.298) and was useful in predicting unknown values (SD = 1.80). Finally, the output matrix was helpful in determining substituent effects.

Introduction

The redox characteristics of dihydropyridines and related compounds have been studied extensively as these structures represent the active electron donating components of such co-enzymes as NADH and NADPH[3]. Given this, cellular respiration is in large part based on this biochemically mediated interconversion. Understanding the kinetics and thermodynamics of dihydropyridine oxidation has therefore been an important and fruitful area of endeavor. In particular, various experimental and theoretical approaches have been applied to defining those structural parameters which add to or detract from compound stability. In this vein, semiempirical molecular orbital approximations such as the AM1 method have been useful[4-6]. In the current investigation, a model was sought to give greater insight into the relative importance of various molecular manipulations to the structure of dihydropyridines and related heterocycles. To this end, we here report on the use of a neural net as a method of both predicting reaction characteristics such as the energies associated with dihydropyridine oxidation, as well as to estimate the relative importance of various substitutions in these determinations.

Neural Networks

A neural network is a parallel processing system in which the processing components or neurons are interconnected in a plexus (Fig. 1)[7]. In simple terms, this network is designed to mimic the multimodular architecture of biological neural systems in which nerve cells interconnect with one another by way of axonal synapses. Neural nets have found particular application in scenarios when the relationships between cause and effect cannot be precisely defined, i.e., circumstances in which complex nonlinear correlations exist between input and output, as in text-to-speech conversions[8], language processing[9] and various pattern recognitions[10-15]. Neural nets have only recently been applied to chemical problems and, as such, relatively few examples of such applications are available[7,16-18]. Various researchers have used artificial neural systems to predict protein structure from amino acid sequence information[19,20] while others have applied neural nets to spectral identification programs[17], nucleic acid sequence analysis[18] and QSAR problems[10-15]. Bodor et al., have recently developed a program for estimating the aqueous solubility of organic compounds using this processing technology[21]. Importantly, dissection of neural net methodologies has shown that they operate as a nonlinear multiregressional technique, although linear operators can be introduced[13].

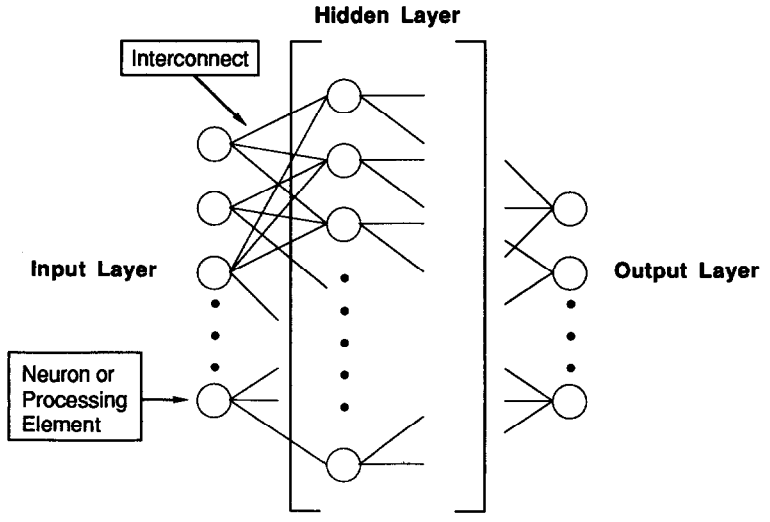


Figure 1. Schematic of a Neural Net

The network is composed of at least three distinct parts: an input layer, an output layer and any number of intermediate, interconnecting strata termed hidden layers. In the superstructure, each processor in the net is affected by those neurons to which it is attached, with the degree of influence being determined by a weighing factor. The nature of the system can therefore be modified by modulating the weights through a learning routine such as the delta or Widrow-Hoff rules. In the application presented here, the specific algorithm used is the back propagation system[12,22].

In this method, the output, O_j, of a neuron, j, at the nth layer is a sigmoidal function of the total input of the unit as given by the equation:

$$O_j = (1 + \exp(-y_j))^{-1} \qquad [1]$$

where

$$y_i = \sum_i O_i W_{ji} + \Theta_j \qquad [2]$$

in which O_j is the output of the neuron, i, in the previous layer, W_{ji} is a weight factor which modulates the connectivity between units j and i and is therefore a weight matrix element and Θ_j is termed the bias or the threshold for neuron, j.

In teaching the net, each training set pattern is introduced at the input neurons and iteratively minimized until a desired error is obtained between the net output, O_k, and the target pattern, t_k. The basic relationship in training the net is

$$\Delta W_{kj} = \eta \delta_k O_j \qquad [3]$$

where η is a learning factor and δ_k is defined as

$$\delta_k = (t_k - O_k) f'(y_k) \qquad [4]$$

with

$$f'(y_k) = df(y_k) = O_k (1-O_k) \qquad [5]$$

Weight correcting begins with equation [3], i.e., the output unit, and is back propagated to the input. The δ_k term for the hidden layers (for which there is no predetermined value) is obtained recursively in terms of the δ terms of the neurons to which it directly interdigitates and weights for those interdigitations. This is defined by

$$\delta_j = f'(y_j) \sum_k \delta_k W_{kj} = O_j(1-O_j) \sum_k \delta_k W_{kj}.$$

The set is then trained either until there is convergence between actual outputs and target outputs or until a predetermined number of iterations is reached. If the trained network

demonstrates an acceptable error between actual and target outputs, the system can determine input patterns of the training set to the desired level of accuracy.

Methods

In the calculations described, a MicroVax II computer was employed. The neural network was implemented in Pascal and configured so as to allow for an arbitrary number of hidden layers and an arbitrary number of units within each layer. In training exercises, iterations were varied until the desired error was reached. Heats of formation differences ($\Delta\Delta H_f$) for various dihydropyridines and their corresponding quaternary salts were obtained from the literature[23] or estimated using the AM1 method[24,25]. Input values for the neural net were adjusted to between 0 and 1 by applying the following equation:

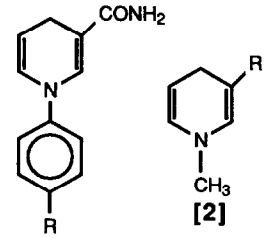
$$\text{input} = \frac{\Delta\Delta H_f - \Delta\Delta H_f(\text{min})}{\Delta\Delta H_f(\text{max}) - \Delta\Delta H_f(\text{min})}$$

where $\Delta\Delta H_f(\text{min})$ and $\Delta\Delta H_f(\text{max})$ refer to the minimum and maximum energy values in the training set, i.e., $\Delta\Delta H_f(\text{min})$ = 146.6 kcal/mol for the oxidation of 1-methyl-3-dimethylamino-1,4-dihydropyridine and $\Delta\Delta H_f(\text{max})$ = 173.0 kcal/mol for the oxidation of 1-methyl-3-trifluoroacetyl-1,4-dihydroquinoline. The program has both training and prognostic capabilities.

Results and Discussion

The study described applied a neural net to the study of the oxidation of various dihydro systems including 1-(4-substituted phenyl)-1,4-dihydronicotinamides, 1-methyl-3-substituted-1,4-dihydropyridines, 1-methyl-3-substituted-1,4-dihydroquinolines, 1-methyl-3-substituted-1,2-dihydroquinolines and 2-methyl-4-substituted-1,2-dihydroisoquinolines to the corresponding aromatic salts[23]. A group of 71 compounds was evaluated and the energy associated with oxidation obtained using AM1 semiempirical molecular orbital approximation as the energy difference between the dihydro derivative and the corresponding quaternary salt. A group of 22 parameters were assigned which included the five parent structures as well as 17 substituents. The training set matrix is given in Table I with associated $\Delta\Delta H_f$ values. The $\Delta\Delta H_f$ values were adjusted to between 0 and 1 as described in the Methods section and these input values then iteratively processed with the output minimized against the actual data. In all simulations, one hidden layer was used and the number of hidden units varied as summarized in Table II. One output neuron was used in the net. Three to ten hidden units gave standard deviations between 0.045 and 0.32. Importantly, while high accuracy can be obtained, overtraining the system often

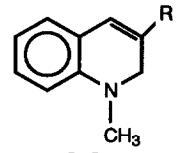
Table I. Heats of Formation Differences (Energies of Oxidation, $\Delta\Delta H_f$) for a Series of Dihydropyridines and Related Heterocycles as well as Assigned Parameter Designations [in Brackets].



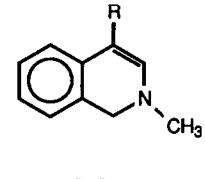
[1]

[2]

[3]



[4]



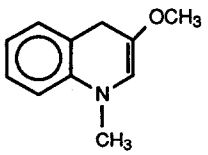
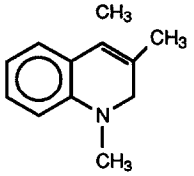
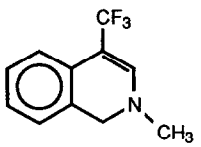
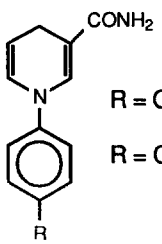
[5]

	R	$\Delta\Delta H_f$ (Kcal/mol)				
[6]	$N(CH_3)_2$	152.6	146.6	153.9	151.4	148.7
[7]	OCH_3	157.0	153.7		156.2	157.3
[8]	CH_3	158.0	152.8	157.6		167.7
[9]	F	162.0	161.3	165.0	161.5	159.2
[10]	Cl	161.7	159.4	162.9	158.7	158.2
[11]	CF_3	166.3	168.2	170.1	164.4	
[12]	$CONH_2$		159.3	164.3	157.7	159.9
[13]	$COOCH_3$		162.5	165.0	158.8	162.7
[14]	$COCH_3$		160.7	164.0	157.3	161.3
[15]	$C(=NOH)CH_3$		154.5	158.5	153.2	162.5
[16]	COC_6H_5		158.1	161.5	155.7	157.8
[17]	$COCF_3$		171.4	173.0	165.6	170.7
[18]	$C(=NOH)CF_3$		161.4	164.3	158.6	168.4
[19]	CH_3S		155.3	158.8	154.8	158.5
[20]	CH_3SO		165.7	165.3	157.1	164.7
[21]	CN		165.7	168.2	162.7	165.6
[22]	$(CH_3)_2NO$		159.0	161.9	156.2	160.4

Table II. Effect of Various Hidden Layers on Neural Net Training and Prediction.

Hidden Unit	Standard Deviation Training Set	Standard Deviation Test Set	Number of Iterations
3	0.372	2.98	50,000
4	0.298	1.80	10,000
5	0.233	2.52	2,400
6	0.171	2.00	2,400
7	0.045	2.79	6,530
8	0.318	1.44	3,900
9	0.156	1.87	810
10	0.141	2.24	1,150

Table III. Calculated (AM1) and Predicted Heats of Formation Differences ($\Delta\Delta H_f$) for Various Dihydropyridines and Related Heterocycles. The Trained Neural Net Incorporated Four Hidden Units.

Compound	$\Delta\Delta H_f$ (Kcal/Mol) AM1 Calculated	$\Delta\Delta H_f$ (Kcal/Mol) Neural Net Predicted
	159.62	156.84
	153.26	151.67
	167.84	169.16
 R = $CONH_2$	163.11	163.11
R = $COOCH_3$	163.76	162.80

yields a net with poor predictive power as the input matrix tends to be "memorized" when large numbers of hidden units are used.[11] The best balance, with this in mind, was a system containing four hidden units. The configuration selected (Figure 2) gave a standard deviation of 0.298 for the training set. The ability of the neural net to predict unknown values was then tested using 5 previously unexamined compounds including 1-methyl-3-methoxy-1,4-dihydroquinoline [3-7], 1-methyl-3-methyl-1,2-dihydroquinoline [4-8], 2-methyl-4-trifluoromethyl-1,2-dihydroisoquinoline [5-11], 1-(4-carbamoylphenyl)-1,4-dihydronicotinamide [1-12] and 1-(4-acetylphenyl)-1,4-dihydronicotinamide [1-14]. The testing set was selected randomly based on missing elements in the training set matrix (See Table I). As indicated in Table III, the system accurately predicted the $\Delta\Delta H_f$ values for the five compounds with a standard deviation of 1.8 kcal/mol.

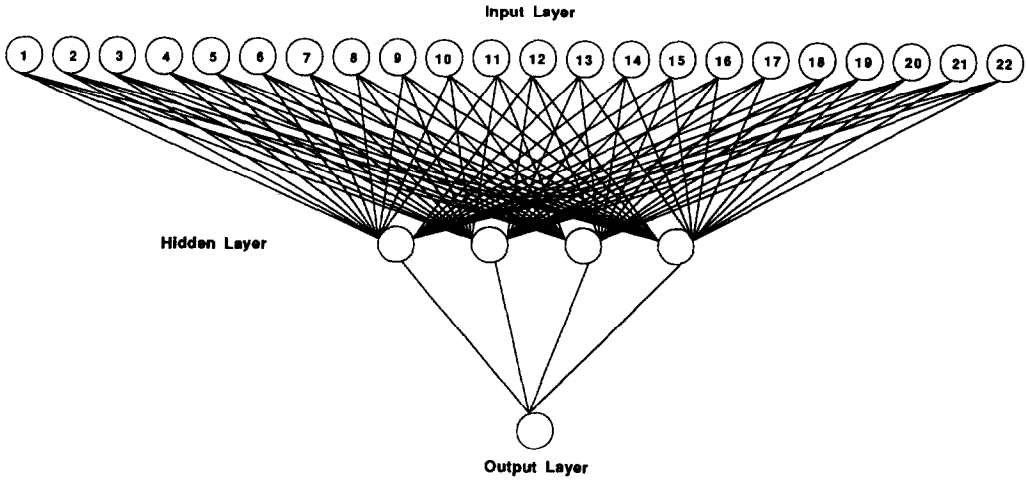


Fig. 2. Neural Net configuration used in the present study. The parameter numbers are included in the input matrix and both input layer, hidden layer, and output layer are identified. Biases are not included in the representation.

The first goal of the neural net was therefore met, i.e., a system for predicting energies associated with dihydropyridine oxidation. The second aspect of this study was to examine the relative importance of various substituents using the calculated output matrix. In this assessment, the relative significance of each parameter as to how it contributes to $\Delta\Delta H_f$ predictions was considered by including or excluding each parameter in the neural net output.

Table IV. Selected Output Matrix Values Reported as a Ratio Between Substituted and Unsubstituted Components for Various Neural Net Parameters

Parameter	Substituent	Output Ratio
[6]	$N(CH_3)_2$	0.235
[7]	OCH_3	0.829
[8]	CH_3	0.881
[9]	F	1.152
[10]	Cl	1.054
[11]	CF_3	1.640
[12]	$CONH_2$	1.104
[13]	$COOCH_3$	1.170
[14]	$COCH_3$	1.094
[15]	$C(=NOH)CH_3$	0.854
[16]	COC_6H_5	0.946
[17]	$COCF_3$	2.076
[18]	$C(=NOH)CF_3$	1.261
[19]	CH_3S	0.875
[20]	CH_3SO	1.237
[21]	CN	1.425
[22]	$(CH_3)_2NO$	1.023

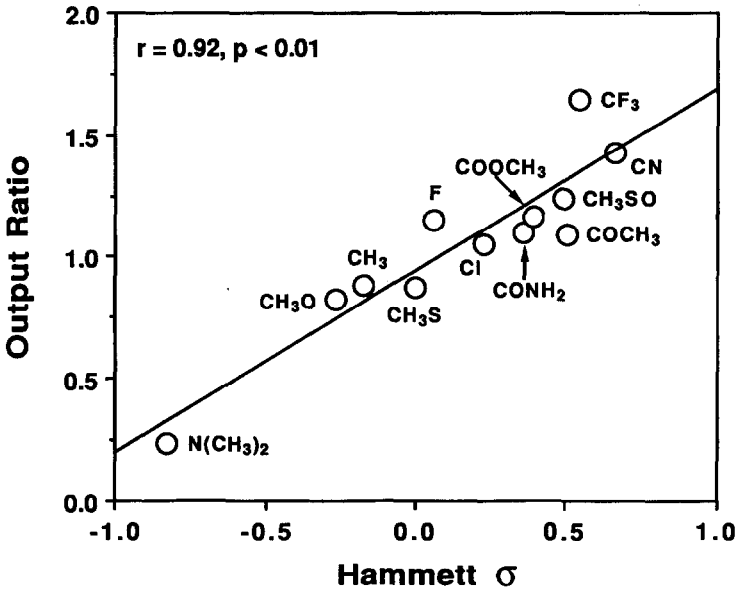


Fig. 3. Correlation between Hammett sigma (σ) values for para substitution and weight matrix output ratios of the substituted and unsubstituted conditions.

The weight matrix derived output values are given in Table IV. The magnitude of these values, which represent the ratio of the substituted and unsubstituted conditions, were found to be useful in examining experimental behavior. For example, in the 1-(4-substituted phenyl)-1,4-dihydronicotinamide series, the rank order of kinetic stability is well described by the parameter contributions of various substituents with a plot of the log of the second order rate constant for oxidation and output matrix ratios giving a positive linear relationship ($r = 0.97$)[26]. Furthermore, the output ratios significantly correlated with Hammett sigma values with $r > 0.9$ (Figure 3). It has been reported that weight data are a highly complicated function of neural net structure and, therefore, may not be useful in simple interpretation. Such observations have been made in protein structure predictions. In the chemical substitution examples investigated herein, the weights obtained appear to give good linear correlations with electronic parameters and reactivity trends.

These methods may prove to be helpful in a number of circumstances. In the present application, calculation of unknown energy values could be obtained theoretically. In examining experimentally derived values such estimations may not be possible. In addition, neural net analysis also may allow one to suggest which parameters in a data set significantly affect the process examined and which do not. In the current evaluation, a neural net was applied to estimation of oxidation energies. A training set was obtained and the resulting neural net useful in predicting unknown values.

Acknowledgement

The authors would like to acknowledge the expert editorial assistance of Kathe Burkhead.

References

1. This article is also contribution No. 10 in the series "Reactivity of Biologically Important Reduced Pyridines." Part 1 of this series: Bodor, N., Harget, A., Huang, M., J. Am. Chem. Soc., 1991, 113, 9480.
2. To whom correspondence should be addressed.
3. Powell, M.F., Bruice, T.C., Prog. Clin. Biol. Res. 1988, 274, 369.
4. Brewster, M., Kaminski, J., Gabanyi, Z., Czako, K., Simay, A., Bodor, N. Tetrahedron 1989, 45, 4395.
5. Brewster, M., Kaminski, J., Gabanyi, Z., Czako, K., Simay, A., Bodor, N., Tetrahedron 1990, 46, 319.
6. Bodor, N., Brewster, M., Kaminski, J., J. Molec. Struct. (Theochem) 1990, 206, 315.
7. Lacy, M. Tetrahedron Comput. Method. 1990, 3, 119.
8. Sejnowski, T., Rosensberg, C., Complex Systems, 1987, 1, 75.
9. Allen, R., Proc. IEEE First Internat. Conf. Neural Networks 1987, 2, 335.
10. Aoyama, T., Suzuki, Y., Ichikawa, H., J. Med. Chem. 1990, 33, 2583.
11. Andrew, T., Kalayeh, H., J. Med. Chem., 1991, 34, 2824.
12. Aoyama, T., Ichikawa, H., Chem. Pharm. Bull., 1991, 39, 358.

13. Aoyama, T., Ichikawa, H., Chem. Pharm. Bull., 1991, 39, 372.
14. Aoyama, T., Ichikawa, H., Chem. Pharm. Bull., 1991, 39, 1222.
15. Aoyama, T., Ichikawa, H., Chem. Pharm. Bull., 1991, 37, 2558.
16. Bryngelson, J., Hopfield, J., Southard, S., Tetrahedron Comput. Method. 1990, 3, 129.
17. Curry, B., Rumelhart, D., Tetrahedron Comput. Method. 1990, 3, 213.
18. Lukashin, A., Anshelevich, V., Amirikyan, B., Gragerov, A., Frank-Kamenetskii,M., J. Biomol. Struct. Dyn. 1989, 6, 1123.
19. Qian, N., Sejnowski, T., J. Mol. Biol. 1988, 202, 865.
20. Holley, L., Karplus, M., Proc. Natl. Acad. Sci. USA, 1989, 86, 152.
21. Bodor, N., Harget, A., Huang, M., J. Am. Chem. Soc., 1991, 113, 9480.
22. Rumelhart, D., McClelland, J., Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Bradford Books, MIT Press, Cambridge, MS (1986).
23. Brewster, M., Kaminski, J., Huang, M., Bodor, N., J. Org. Chem. 1990, 55, 2361.
24. Dewar, M., Zoebisch, E., Healy, E., Stewart, J., J. Am. Chem. Soc. 1985, 107, 3902.
25. Dewar, M., Storch, D., J. Am. Chem. Soc. 1985, 107, 3898.
26. Brewster, M., Simay, A., Czako, K., Winwood, D., Farag, H., Bodor, N., J. Org. Chem. 1989, 54, 3721.