laboratory is directed toward enhancing the delivery of the (chloroethyl)diazonium ion to the tumor tissue, without the undesirable side reactions. The methyl group in the 3-position will be replaced by groups which favor the dissociation to the (2-chloroethyl)diazonium and which do not produce toxic intermediates. The benzyl group appears to have these qualities. Likewise, the experience with the apparent activation of CMM by metabolism of the *N*-methylcarbamoyl moiety suggests that an appropriately designed acyl group, which can be cleaved by a tumor-specific enzyme, may impart much greater selectivity and greatly decreased toxicity.

## Experimental Section

**Synthesis.** The compounds in this study were prepared by previously published methods. Thus, DMA, DMP, and DMC were prepared by the acylation of the anion of 1,3-dimethyltriazene (DMA, CMP, DMC) or, in the case of DMM, by the direct reaction of 1,3-dimethyltriazene with methyl isocyanate.[12] The compounds were isolated and characterized as described previously and were >99% pure. The (2-chloroethyl)triazenes CMA, CMC, and CMM were prepared by a multistep synthesis, also described previously.[13] These compounds were also analytically pure.

**Kinetics.** Rates of triazene decomposition were determined spectrophotometrically, as described previously,[12] on a Hewlett-Packard Model 8450A diode-array spectrophotometer. The thermostated (±0.1 °C) 1-cm cuvettes were charged with 1.341 mL of 0.1 M lysine buffer at the appropriate pH. The reaction was initiated by addition of 9 μL of a $3 \times 10^{-3}$ M solution of the triazene in acetonitrile. The reference cuvette contained the same buffer and 9 μL of acetonitrile. The reactions were followed for at least 3.5 half-lives and at least 100 points were used to evaluate each rate constant. The calculations employed the Guggenheim approximation to determine the infinity absorbance and the rate constants were evaluated by a least-squares method. The calculations were carried out by utilizing a program written in our laboratory. Each kinetic run was carried out in duplicate and, when deviations were >3%, three or more runs were used to obtain a more accurate value.

**MTT-Microculture Tetrazolium Assay.** Cellular growth in the presence or absence of experimental agents was determined by using the previously described MTT assay.[17] Briefly, cells were harvested and inoculated into 96-well microtiter plates at 1000 cells/well. After 24 h, drugs were applied and cultures were incubated an additional 6 days at 37 °C. MTT was added, the formazan product was solubilized, and the absorbancy was measured at 540 nm with a Bio-Tek Model EL 312 microplate reader.

**Acknowledgment.** Research was sponsored by the National Cancer Institute, DHHS, under Contract Nos. N01-C0-74101 with ABL and N01-C0-74102 with PRI. R.H.S. was supported in part by NSF Grant No. CHE-8521385. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We are grateful to Dr. Kenneth Paull, Division of Cancer Treatment, NCI, for facilitating the in vivo assays.

(17) Scudiero, D. A.; Shoemaker, R. H.; Paull, K. D.; Monks, A.; Tierney, S.; Nofziger, T. H.; Currens, M. J.; Seniff, D.; Boyd, M. R. *Cancer Res.* 1988, *48*, 4827.

# Neural Networks Applied to Quantitative Structure–Activity Relationship Analysis[1]

Tomoo Aoyama,[†] Yuji Suzuki,[‡] and Hiroshi Ichikawa*[,‡]

*Hitachi Computer Engineering Co., Ltd., Horiyamashita, Hadano, Kanagawa 259-13, Japan, and Hoshi College of Pharmacy, Shinagawa, Tokyo 142, Japan.* *Received December 27, 1989*

An application of the neural network to quantitative structure–activity relationship (QSAR) analysis has been studied. The new method was compared with the linear multiregression analysis in various ways. It was found that the neural network can be a potential tool in the routine work of QSAR analysis. The mathematical relationship of operation between the neural network and the multiregression analysis was described. It was shown that the neural network can exceed the level of the linear multiregression analysis.

## Introduction

The first quantitative structure–activity relationship (QSAR) method is the model proposed by Hansch and co-workers.[2-4] It was the seminal contribution to this field. The success of this method has prompted many workers to reexamine the derivation of the Hansch equation by using the principles of theoretical pharmacology[5,6] or pharmacokinetics.[7-10] This model, the free energy model,[11] and its elaborations[12] have been by far the most widely used. This may be due to its direct conceptual linkage to established physical organic chemical principles. However, the method is totally dependent on the multiregression analysis. This causes the problems of orthogonality of the variables as well as the size of population.

QSAR is also regarded as the problem of pattern recognition. From this view point, techniques of pattern recognition have been applied to QSAR study, examining structural features and/or chemical properties underlying

[†] Hitachi Computer Engineering Co., Ltd.
[‡] Hoshi College.

(1) Neural Networks Applied to Pharmaceutical Problems. III. For part II, see ref 24.
(2) Hansch, C.; Muir, R. M.; Fujita, T.; Maloney, P.; Geiger, E.; Streich, M. *J. Am. Chem. Soc.* 1963, *86*, 2817.
(3) Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* 1964, *86*, 1616.
(4) Hansch, C. *J. Med. Chem.* 1976, *19*, 1.
(5) Austel, V.; Kutter, E. In *Drug Design*; Ariens, E. J., Ed. Academic Press: New York, 1980; Vol 10, Medicinal Chemistry, p 1.
(6) Goldstein, A.; Aronow, L.; Kalman, N. S. *Principles of Drug Action—Basis of Pharmacology*, 2nd ed.; Wiley: New York, 1974.
(7) van Rossum, J. M. *Proc. Int. Pharmacol. Meet., 3rd, Sao Paulo, 1966*, 1968, *7*, 237.
(8) van Rossum, J. M. In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1971; Vol 1, Medicinal Chemistry, p 469.
(9) Seydel, J. K. In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1971; Vol. 1, Medicinal Chemistry, p 343.

0022-2623/90/1833-2583$02.50/0 

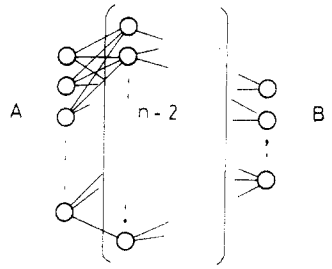



**Figure 1.** *n*-Layer neural network.

patterns that are associated with differing biological effects. A number of pattern recognition systems have been developed: the earliest used in QSAR work were that of Kirschner and Kowalski, called ARTHUR,[13] and that of Stuper and Jurs, named ADAPT.[14,15] Some difficulties of the earliest methods prompted attempts in different approaches to QSAR. One of the fruitful outputs is the SIMCA system of Wold et al.[16-19] This method makes use of principal components analysis to provide a structure and limits to the classification groups so that not only group membership but also the level of activity within each group can be determined. The most successful approach may be the method called adaptive least squares (ALS) proposed by Moriguchi et al.[20,21] which is related to discriminant analysis. However, the resolution and prediction abilities of the ALS as well as other pattern-recognition methods are still far from satisfactory.

Recently, the neural network has been the center of attention in the field of pattern recognition. The neural network is one of typical parallel-distributed processing methods and is a computer-based system derived from simplified concept of the brain in which a number of nodes, called processing elements or neurons, are interconnected in a netlike structure.[22] Since the characteristics of the neural network have been found to be suitable for the processing of data in which the relationship between the cause and its results cannot be exactly defined, such pattern recognitions as those of handwriting letters and human's voice are most expected to be targets of application. We considered that the effective application of such neural networks may bring forth a breakthrough in the current state of QSAR analysis.

As our preceding reports show, the neural networks were successfully applied to decision making[23] and to the study of structure–activity relationship.[24] In the latter application, we could show that the resolution ability of the neural network exceeded that of the ALS method. These are examples of the application of the classification ability which demonstrates that the neural network would be a valuable tool in clinical media as well as in developing new drugs.

It has been said that one cannot give rationalization to the results by the neural network. And this was considered to be the fatal defect of this method as a theoretical tool. In order to remove this obstacle, we studied the reasoning and found the fact that the operation of the neural networks is, indeed, one of the nonlinear multiregression analyses.

## Theory

**A. The Standard Operation of a Neural Network with Back-Propagation Algorithm.** Shown in Figure 1 is the perceptron-type neural network: the circles are neurons which are actually variables taking a value ranging from 0 to 1. The number of the layer is arbitrary and generally consists of $n$ layers. The data are input to A and are output from B. The value of a neuron ($O_j$) at the $n$th layer can be expressed by eq 1 where $x_i$ is one of the values

$$O_j = 1/[1 + \exp(-\alpha y_j)] \equiv f(y_j) \qquad y_j = (\sum W_{ij}x_i) - \theta_j \quad (1)$$

of the neurons at the $n - 1$ layer; $W_{ij}$, an element of the weight matrix, expresses the weight value between neurons i and j and can take either a positive or negative value; $\theta_j$ is a threshold value for neuron j, $\alpha$ is a parameter which expresses the nonlinearity of the neuron's operation. On feeding the input data, the value of every neuron expressed by eq 1 is synchronously renewed.

Given $N$ neurons at the first layer. A set of the input data can be expressed by a vector with $N$ elements for $N$ neurons which is, here, called and "input pattern". Likewise, the output data can also be regarded as a vector and be called an "output pattern". The vector which is compared with an output pattern to obtain the fixed $W_{ij}$ is called a "training pattern" ($t_j$). The training of the network is based on the following equations.

$$\delta W_{ij} = -d_j x_i \epsilon \quad (2)$$

$$d_j = (O_j - t_j)f'(y_j) \quad (3a)$$

$$d_j = (W'_{jl}d'_l)f'(y_j) \quad (3b)$$

Here, $\epsilon$ is a parameter which determines the shift for correction in recursive cycles. Equation 3a is used only for the correction of the last (output) layer and 3b for other layers where $W'_{jl}$ and $d'_l$ at the $n$th layer are $W_{ij}$ and $d_j$ at the $n + 1$ layer, respectively. The function f' in eq 3 is

$$f'(y_j) = f(y_j)[1 - f(y_j)]\alpha \quad (4)$$

where both $\epsilon$ and $\alpha$ can be set to be independent of the layer.

The training is carried out according to the above back-propagation algorithm[22] until

$$E = \sum (O_j - t_j)^2 \quad (5)$$

becomes small enough. Even in case that $M$ sets of the input and training patterns are given, all of output patterns can be made close enough to the training patterns by the iteration of eqs 1 and 2. If the convergence is attained,

(10) Seydel, J. K.; Schaper, K.-J. *Chemische Struktur* und *Biologische Activitat von Wirkstoffen*; Verlag Chemie: Weinheim, 1979.
(11) Leffler, J. E.; Grunwald, E. *Rates and Equilibria of Organic Reactions*; Wiley: New York, 1963. See also: ref 14, Chapter 1.
(12) For a review, see, for example, *Quantitative Structure–Activity Relationships of Drugs*; Topliss, J. G., Ed.; Academic Press: New York, 1983.
(13) Kirschner, G. L.; Kowalski, B. R. In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1979; Vol 8, Medicinal Chemistry, p 73.
(14) Stuper, A. J.; Jurs, P. C. *J. Am. Chem. Soc.* **1975**, *97*, 182.
(15) Stuper, A. J.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1976**, *16*, 99, 105.
(16) Dunn, W. J., III; Wold, S.; Martin, Y. C. *J. Med. Chem.* **1978**, *21*, 922.
(17) Dunn, W. J., III; Wold, S. *J. Med. Chem.* **1980**, *23*, 595.
(18) Dunn, W. J., III; Wold, S. *Bioorg. Chem.* **1980**, *9*, 505.
(19) Wold, S. *Experientia, Suppl. No.* 23, **1976**, 87.
(20) Moriguichi, I.; Komatsu, K.; Matsushita, Y. *J. Med. Chem.* **1979**, *23*, 20.
(21) Moriguichi, I.; Komatsu, K. *Eur. J. Med. Chem.-Chim. Ther.* **1981**, *16*, 19.
(22) *Parallel Distributed Processing Exploration in Microstructure of Cognition*; Rumelhart, D. E., McClelland, N. L., Eds.; MIT Press: Cambridge, MA, 1986; Vols. 1 and 2.
(23) Aoyama, T.; Suzuki, Y.; Ichikawa, H. *Chem. Pharm. Bull.* **1989**, *37*, 2558.
(24) Aoyama, T.; Suzuki, Y. Ichikawa, H. *J. Med. Chem.* **1990**, *33*, 905.

**Table I.** Structures and Parameters of Neural Networks

| A | | | | | | B | | | | | | C | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ly[a] | nnr[b] | | $\alpha$ | $\epsilon$ | $\theta$ | ly[a] | nnr[b] | | $\alpha$ | $\epsilon$ | $\theta$ | ly[a] | nnr[b] | | $\alpha$ | $\epsilon$ | $\theta$ |
| 1 | 7 | | | | | 1 | 13 | | | | | 1 | 14 | | | | |
| 2 | 12 | eq 1 | 2 | 0.10 | 0 | 2 | 26 | eq 1 | 5 | 0.05 | 0 | 2 | 28 | eq 1 | 5 | 0.05 | 0 |
| 3 | 1 | eq 6 | - | 0.05 | - | 3 | 1 | eq 6 | - | 0.05 | - | 3 | 1 | eq 6 | - | 0.05 | - |

[a] Layer. [b] Number of neurons.

then the neural network has an ability to classify the input patterns into $M$ groups.

Those procedures can easily be programmed in BASIC or FORTRAN languages. The length of the program is ca. 200 steps and one can practically perform the operation of the neural network on a small personal computer.

**B. The Relationship between the Operation of Neural Network and the Multiregression Analysis.** Here, we describe the relationship between the operation of the neural network and the multiregression analysis. For simplification, let us consider a three-layer network. Since the operation expressed by eq 1 results in vector elements that are too close to 0 or 1, eq 1 is not very suitable when it is applied to the problems where the values between 0 and 1 are important. Therefore, we considered a new operation equation. Without losing generality, one can omit $\theta_j$ in eq 1, giving

$$y_i = \sum W_{ij} x_i \tag{6}$$

Namely,

$$y = Wx \tag{7}$$

where $W$ and $x$ are the weight matrix and the input vector, respectively. Thus, if all neurons of each layer are governed by eq 6, i.e.

$$y = W_1 x \qquad z = W_2 y \tag{8}$$

then, the output pattern, $z$, becomes

$$z = (W_1 W_2) x = Wx \tag{9}$$

where $W_1$ and $W_2$ are the matrices which express the weights between the layers 1 and 2 and those between the layers 2 and 3, respectively.

The method of the multiregression analysis seeks the optimal coefficients of the linear equation

$$z_i = a_i + \sum b_i x_i \tag{10}$$

where $z$ and $x$ are, respectively, the expectation vector and input data. Equation 10 is equivalently rewritten as

$$z = B'(1 + x) \tag{11}$$

Equation 9, a special case of the neural network's operation, shows that the operation is equivalent to that of the 2-layer network and to that of a generalized multiregression analysis if the variables are so set as $x$ to be the observed values plus the constant 1. It should be emphasized here that addition of the constant 1 to the input data means that the optimization of $\theta_j$ in eq 1 is carried out through the weight matrix, $W_{ij}$.

**C. An Improvement of the Operation of the Neural Network.** The neural network with eq 9 performs the linear operation equivalent to that of multiregression analysis. In order to exceed this level, it is necessary to introduce a nonlinear operation in the network. This is possible by incorporating the hidden layers. Thus we used a three-layer network, letting $O_j = y_j$ and using eq 6 for the last year. However, the larger number of the neurons in the hidden layers must be adapted, rather than the input layer, to avoid loss of the information that the input pattern has.[25]

## Results and Discussion

We considered that, in order to show the usefulness of a new method, it may be most appealing to use the data well-studied by the conventional methods to compare the results.

**A. QSAR in Carboquinones.** Carboquinones were synthesized by Nakao et al.[26,27] and other groups[28-30] and were developed to an anticarcinogenic drug for the clinical media. A detailed QSAR study based on the Hansch method has been carried out by Yoshimoto et al.[31] We first used those data to compare the results of the neural network with those of conventional QSAR techniques.

For comparison, we tried two sets of the structure and parameters of the network which are shown in Table I, parts A and B. The input data, physicochemical parameters, are the molecular refractivity constants (MR), hydrophobicity constant ($\pi$), substituent constants (F and R), as well as, $MR_{1,2}$ and $\pi_{1,2}$ to estimate the steric effects of $R^1$ and $R^2$ and the total hydrophobicity. Biological data are minimum effective dose (MED) and optimal dose (OD) on a chronic treatment schedule and those in single injection. MED is the dose giving a 40% increase in lifespan compared to the controls, and OD is the dose giving maximum increase of lifespan. The input data are shown in Table II.

The input data, $MR_{1,2}$, $\pi_{1,2}$, $\pi_2$, $MR_1$, F, and R, are rescaled to have the values between ca. 0.1 and 1 by the following equation

$$\tilde{x}_i = (x_i - x_{min} + 0.1)/(x_{max} - x_{min} + 0.1) \tag{12}$$

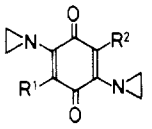
and fed to the network together with the constant 1. As the training pattern, $\log(1/c)$ of the observed values were rescaled to have the values between 0.0 and 1 and given to the third layer.

In order to make best use of the information embodied in the data, we also examined the effect of incorporating squares, $MR_{1,2}^2$, $\pi_{1,2}^2$, $\pi_2^2$, $MR_1^2$, $F^2$, and $R^2$, in the input data, although the original authors did not consider them. The parameter set of the network was shown in Table I, part B, where $\alpha$ was set to be 5 in order to increase the nonlinearity of the neuron's operation. Generally, as the nonlinear operation is increased, the convergence is not easily attained.

The results are shown in Table III, where columns for clc 1, 2, 3, and 4 (columns 3, 7, 11, and 15, respectively) are the calculated results in the literature[31] and those for

(25) An extensive program, PSDD (Perceptron Simulator for Drug Design) ca. 1500 steps in FORTRAN, will be submitted to QCPE (Quantum Chemistry Program Exchange, Indiana University, IN).
(26) Nakao, H.; Arakawa, M; Nakamura, T.; Fukushima, M. *Chem. Pharm. Bull.* **1972**, *20*, 1962.
(27) Nakao, H.; Arakawa, M. *Chem. Pharm. Bull.* **1972**, *20*, 1968.
(28) Driscoll, J. S.; Hazard, G. F., Jr.; Wood, H. B., Jr.; Goldin, A. *Cancer Chemother. Rep. Part 2* **1976**, *4*, 1.
(29) Khan, A. H.; Driscoll, J. S. *J. Med. Chem.* **1976**, *19*, 313.
(30) Chou, F.; Khan, A. H.; Driscoll, J. S. *J. Med. Chem.* **1976**, *19*, 1302.
(31) Yoshimoto, M.; Miyazawa, H.; Nakao, H.; Shinkai, K.; Arakawa, M. *J. Med. Chem.* **1979**, *22*, 491.

Table II. Input Data for Carboquinones[a]



| no. | $R_1$ | $R_2$ | $MR_{1,2}$ | $\pi_{1,2}$ | $\pi_2$ | $MR_1$ | F | R |
|---|---|---|---|---|---|---|---|---|
| 1 | $CH_3$ | $COCH_3$ | 1.69 | −0.05 | −0.55 | 0.57 | 0.28 | 0.07 |
| 2 | $C_6H_5$ | $C_6H_5$ | 5.08 | 3.92 | 1.96 | 2.54 | 0.16 | −0.16 |
| 3 | $CH_3$ | $(CH_2)_3C_6H_5$ | 4.50 | 3.66 | 3.16 | 0.57 | −0.08 | −0.26 |
| 4 | $C_5H_{11}$ | $C_5H_{11}$ | 4.86 | 5.00 | 2.50 | 2.43 | −0.08 | −0.26 |
| 5 | $CH(CH_3)_2$ | $CH(CH_3)_2$ | 3.00 | 2.60 | 1.30 | 1.50 | −0.08 | −0.26 |
| 6 | $CH_3$ | $CH_2C_6H_5$ | 3.57 | 2.51 | 2.01 | 0.57 | −0.12 | −0.14 |
| 7 | $C_3H_7$ | $C_3H_7$ | 3.00 | 3.00 | 1.50 | 1.50 | −0.08 | −0.26 |
| 8 | $CH_3$ | $CH_2OC_6H_5$ | 3.79 | 2.16 | 1.66 | 0.57 | −0.04 | −0.13 |
| 9 | $R^1 = R^2 = CH_2CH_2OCON(CH_3)_2$ | | 6.14 | 0.72 | 0.36 | 3.07 | −0.08 | −0.26 |
| 10 | $C_2H_5$ | $C_2H_5$ | 2.06 | 2.00 | 1.00 | 1.03 | −0.08 | −0.26 |
| 11 | $CH_3$ | $CH_2CH_2OCH_3$ | 2.28 | 1.03 | 0.53 | 0.57 | −0.08 | −0.26 |
| 12 | $OCH_3$ | $OCH_3$ | 1.58 | −0.04 | −0.02 | 0.79 | 0.52 | −1.02 |
| 13 | $CH_3$ | $CH(CH_3)_2$ | 2.07 | 1.80 | 1.30 | 0.57 | −0.08 | −0.26 |
| 14 | $C_3H_7$ | $CH(OCH_3)CH_2OCONH_2$ | 4.24 | 0.98 | −0.52 | 1.50 | −0.04 | −0.13 |
| 15 | $CH_3$ | $CH_2CH_2OCON(CH_3)_2$ | 3.64 | 0.86 | 0.36 | 0.57 | −0.08 | −0.26 |
| 16 | $CH_3$ | $CH_3$ | 1.14 | 1.00 | 0.50 | 0.57 | −0.08 | −0.26 |
| 17 | H | $CH(CH_3)_2$ | 1.60 | 1.30 | 1.30 | 0.10 | −0.04 | −0.13 |
| 18 | $CH_3$ | $CH(OCH_3)C_2H_5$ | 2.75 | 1.53 | 1.03 | 0.57 | −0.04 | −0.13 |
| 19 | $C_3H_7$ | $CH_2CH_2OCONH_2$ | 3.56 | 1.45 | −0.05 | 1.50 | −0.08 | −0.26 |
| 20 | $R^1 = R^2 = CH_2CH_2OCH_3$ | | 3.42 | 1.03 | 0.53 | 1.71 | −0.08 | −0.26 |
| 21 | $C_2H_5$ | $CH(OC_2H_5)CH_2OCONH_2$ | 4.23 | 0.98 | −0.02 | 1.03 | −0.04 | −0.13 |
| 22 | $CH_3$ | $CH_2CH_2OCOCH_3$ | 2.78 | 1.23 | 0.73 | 0.57 | −0.08 | −0.26 |
| 23 | $CH_3$ | $(CH_2)_3$-dimer | 1.96 | 2.00 | 1.50 | 0.57 | −0.08 | −0.26 |
| 24 | $CH_3$ | $C_2H_5$ | 1.60 | 1.50 | 1.00 | 0.57 | −0.08 | −0.26 |
| 25 | $CH_3$ | $CH(OCH_2CH_2OCH_3)CH_2OCONH_2$ | 4.45 | 0.01 | −0.49 | 0.57 | −0.04 | −0.13 |
| 26 | $CH_3$ | $CH_2CH(CH_3)OCONH_2$ | 3.09 | 0.75 | 0.25 | 0.57 | −0.08 | −0.26 |
| 27 | $C_2H_5$ | $CH(OCH_3)CH_2OCONH_2$ | 3.77 | 0.48 | −0.52 | 1.03 | −0.04 | −0.13 |
| 28 | $CH_3$ | $CH(C_2H_5)CH_2OCONH_2$ | 3.55 | 1.25 | 0.75 | 0.57 | −0.08 | −0.26 |
| 29 | $CH_3$ | $CH(OC_2H_5)CH_2OCONH_2$ | 3.77 | 0.48 | −0.02 | 0.57 | −0.04 | −0.13 |
| 30 | $CH_3$ | $(CH_2)_3OCONH_2$ | 3.09 | 0.95 | 0.45 | 0.57 | −0.08 | −0.26 |
| 31 | $CH_3$ | $(CH_2)_2OCONH_2$ | 2.63 | 0.45 | −0.05 | 0.57 | −0.08 | −0.26 |
| 32 | $C_2H_5$ | $(CH_2)_2OCONH_2$ | 3.09 | 0.95 | −0.05 | 1.03 | −0.08 | −0.26 |
| 33 | $CH_3$ | $CH_2CH_2OH$ | 1.78 | 0.34 | −0.16 | 0.57 | −0.08 | −0.26 |
| 34 | $CH_3$ | $CH(CH_3)CH_2OCONH_2$ | 3.09 | 0.75 | 0.25 | 0.57 | −0.08 | −0.26 |
| 35 | $CH_3$ | $CH(OCH_3)CH_2OCONH_2$ | 3.31 | −0.02 | −0.52 | 0.57 | −0.04 | −0.13 |
| 36 | H | $N(CH_2)_2$ | 1.66 | 0.18 | 0.18 | 0.10 | 0.10 | −0.92 |
| 37 | $R^1 = R^2 = CH_2CH_2OH$ | | 2.42 | −0.32 | −0.16 | 1.21 | −0.08 | −0.26 |
| 38 | $CH_3$ | $N(CH_2)_2$ | 2.13 | 0.68 | 0.18 | 0.57 | 0.06 | −1.05 |
| 39 | $CH_3$ | $CH(OCH_3)CH_2OH$ | 2.47 | −0.13 | −0.63 | 0.57 | −0.04 | −0.13 |

[a] The data were taken from the literature (ref 31).

set A are the results for the parameter set A while set B are those for the parameter set B. The numbers in the rows with +, ±, and − show cases of superior (i.e., much closer to the observed values), equivalent, and inferior to those by the multiregression analysis, respectively. Apparently, the neural networks give better results than the multiregression analysis does. It should be noted here that incorporation of the second-order contributions ($MR_{12}^2$, etc.) considerably improves the ability. Besides, increase of input parameters does not cause as much trouble in the neural network as it does in the multiregression analysis. Therefore, it may be always recommended to incorporate the squares of the input physicochemical parameters.

The comparisons of the mean deviations, variances, and standard deviations by the neural network with those by the multiregression analysis are shown in Table IV. As seen from the table, the results by the neural networks were found superior in all cases. The ratios of variance ($F$) of the multiregression analysis over the neural network were $1.14 \leq F \leq 1.5$ for the parameter set A and $1.26 \leq F \leq 1.97$ for the set B. Since the number of the regression coefficients was 6, the ratios can be regarded as valid ones.

We have tested the results of the neural network by the following method. The leave-$n$-out method ($n$ = 1, 2, 5, and 10) was applied to the network with the parameter set A. Namely, the $37 - n$ data which are formed by randomly removing $n$ number of data from the total 37 data, were fed to the network for training. Then, the removed data were input to the trained network to calculate the mean deviations. This operation was repeated 37 times and the deviations ($\sigma^2$) were averaged. The results are shown in Table V, where for comparison the deviation of the data used for training ($\sigma_L^2$) are also recorded.

Although the deviations of the untrained data were found to be sightly larger than those of trained data, it may be said that the neural network well reproduces the observed values. The small variances also indicate that the relevant data have a good linear relationship.

**B. QSAR in Benzodiazepines.** Randall et al. first introduced chlorodiazepoxide, a derivative of benzodiazepine, as a minor tranquilizer.[32] Since then, the biological activities of 1,4-benzodiazepines have been extensively studied[33] and played the major role in the field of minor tranquilizers. The QSAR study on this series of compounds has been carried out by Kubota et al.[34] We

(32) Randall, L. O.; Schallek, W.; Heise, G. A.; Eith, E. F.; Bagdon, R. E. *J. Pharmacol. Exp. Ther.* **1960**, *129*, 163.
(33) Sternbach, L. H.; Randall, L. O.; Banziger, R.; Lehr, H. In *Medicinal Research Series 2. Drugs Affecting the Central Nervous System*; Burger, A., Ed.; Mercell Dekker, Inc.: New York, 1968; p 237.

**Table III.** Comparison of Results by Neural Networks with Those by Multiregression Analysis

| | chronic injection, log (1/C) | | | | | | | | singular injection, (1/C) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MED | | | | OD | | | | MED | | | | OD | | | |
| no. | obsd[a] | clc 1[b] | set A[c] | set B[d] | obsd[a] | clc 2[b] | set A[c] | set B[d] | obsd[a] | clc 3[b] | set A[c] | set B[d] | obsd[a] | clc 4[b] | set A[d] | set B[d] |
| 1 | | | | | | | | | 3.94 | 4.12 | 4.260 | 4.387 | 3.48 | 3.72 | 3.757 | 3.582 |
| 2 | 4.33 | 4.05 | 4.260 | 4.362 | 4.14 | 3.81 | 4.062 | 4.155 | | | | | 3.53 | 2.97 | 3.301 | 3.426 |
| 3 | 4.47 | 4.61 | 4.629 | 4.549 | 4.21 | 4.57 | 4.561 | 4.452 | 3.93 | 4.23 | 4.268 | 4.178 | 3.60 | 3.99 | 3.923 | 3.717 |
| 4 | 4.63 | 4.31 | 4.327 | 4.297 | 4.52 | 4.27 | 4.211 | 4.167 | 4.07 | 3.58 | 3.825 | 3.840 | 3.62 | 3.50 | 3.615 | 3.491 |
| 5 | 4.77 | 5.26 | 5.097 | 5.142 | 4.59 | 4.96 | 4.899 | 4.920 | 4.36 | 4.74 | 4.658 | 4.624 | 4.14 | 4.38 | 4.292 | 4.195 |
| 6 | 4.85 | 5.18 | 5.109 | 5.087 | 4.69 | 4.92 | 4.944 | 4.874 | 4.74 | 4.77 | 4.746 | 4.672 | 4.26 | 4.37 | 4.320 | 4.169 |
| 7 | 4.92 | 5.15 | 4.976 | 5.006 | 4.44 | 4.89 | 4.802 | 4.804 | 4.32 | 4.55 | 4.530 | 4.488 | 4.14 | 4.23 | 4.171 | 4.047 |
| 8 | 5.15 | 5.21 | 5.132 | 5.206 | 4.71 | 4.87 | 4.893 | 4.916 | 4.68 | 4.61 | 4.611 | 4.692 | 3.89 | 4.23 | 4.167 | 4.123 |
| 9 | 5.16 | 5.21 | 5.193 | 5.048 | 4.85 | 4.83 | 4.862 | 4.815 | | | | | 4.62 | 5.07 | 4.647 | 4.537 |
| 10 | 5.46 | 5.57 | 5.463 | 5.525 | 5.09 | 5.20 | 5.227 | 5.253 | 4.94 | 5.01 | 4.997 | 4.992 | 4.79 | 4.58 | 4.575 | 4.489 |
| 11 | 5.57 | 5.98 | 5.977 | 6.021 | 5.42 | 5.50 | 5.621 | 5.657 | 5.19 | 5.51 | 5.524 | 5.579 | 5.12 | 4.95 | 5.019 | 4.998 |
| 12 | 5.59 | 5.74 | 5.707 | 5.729 | 5.17 | 5.27 | 5.294 | 5.264 | 4.81 | 4.79 | 4.559 | 4.629 | 4.32 | 4.45 | 4.144 | 4.196 |
| 13 | 5.60 | 5.58 | 5.536 | 5.579 | 5.21 | 5.23 | 5.325 | 5.339 | 4.96 | 5.13 | 5.100 | 5.108 | 4.69 | 4.67 | 4.657 | 4.616 |
| 14 | 5.63 | 6.03 | 5.797 | 5.808 | 5.07 | 5.37 | 5.229 | 5.239 | 5.01 | 5.18 | 5.191 | 5.117 | 4.64 | 4.66 | 4.645 | 4.446 |
| 15 | | | | | | | | | 5.09 | 5.59 | 5.626 | 5.684 | 4.84 | 5.01 | 5.068 | 5.023 |
| 16 | 5.66 | 5.99 | 6.067 | 6.088 | 5.36 | 5.51 | 5.742 | 5.730 | 5.36 | 5.52 | 5.539 | 5.505 | 4.79 | 4.96 | 5.067 | 4.971 |
| 17 | 5.68 | 5.56 | 5.657 | 5.728 | 5.37 | 5.13 | 5.393 | 5.388 | 5.16 | 5.02 | 5.024 | 5.032 | 4.59 | 4.54 | 4.548 | 4.564 |
| 18 | 5.68 | 5.54 | 5.498 | 5.617 | 5.33 | 5.09 | 5.186 | 5.236 | 5.26 | 4.91 | 4.907 | 5.014 | 4.84 | 4.46 | 4.430 | 4.429 |
| 19 | 5.68 | 5.96 | 5.720 | 5.736 | 5.23 | 5.43 | 5.291 | 5.329 | 4.90 | 5.30 | 5.286 | 5.214 | 4.42 | 4.80 | 4.792 | 4.650 |
| 20 | 5.69 | 5.59 | 5.564 | 5.611 | 5.31 | 5.17 | 5.247 | 5.298 | 5.18 | 5.51 | 5.157 | 5.089 | 4.71 | 4.95 | 4.799 | 4.786 |
| 21 | 5.76 | 5.93 | 5.763 | 5.829 | 5.24 | 5.33 | 5.253 | 5.294 | 5.40 | 5.18 | 5.195 | 5.241 | 4.64 | 4.66 | 4.641 | 4.515 |
| 22 | 5.78 | 5.87 | 5.824 | 5.874 | 5.78 | 5.43 | 5.492 | 5.539 | | | | | | | | |
| 23 | 5.82 | 5.47 | 5.434 | 5.468 | 5.39 | 5.16 | 5.257 | 5.256 | | | | | | | | |
| 24 | 5.86 | 5.73 | 5.732 | 5.776 | 5.37 | 5.33 | 5.485 | 5.495 | 5.16 | 5.28 | 5.259 | 5.255 | 4.52 | 4.78 | 4.807 | 4.755 |
| 25 | 6.03 | 6.33 | 6.326 | 6.272 | 5.39 | 5.62 | 5.649 | 5.613 | 5.45 | 5.65 | 5.751 | 5.739 | 4.96 | 5.01 | 5.096 | 4.891 |
| 26 | 6.14 | 6.12 | 6.104 | 6.125 | 5.79 | 5.60 | 5.667 | 5.706 | 5.86 | 5.64 | 5.690 | 5.759 | 5.18 | 5.05 | 5.140 | 5.106 |
| 27 | 6.16 | 6.19 | 6.100 | 6.115 | 5.22 | 5.50 | 5.488 | 5.494 | 5.62 | 5.42 | 5.471 | 5.487 | 4.92 | 4.84 | 4.885 | 4.712 |
| 28 | 6.18 | 5.86 | 5.767 | 5.805 | 5.66 | 5.42 | 5.417 | 5.467 | 6.03 | 5.40 | 5.402 | 5.462 | 5.20 | 4.87 | 4.877 | 4.850 |
| 29 | 6.18 | 6.09 | 6.061 | 6.113 | 5.22 | 5.46 | 5.512 | 5.541 | 5.53 | 5.42 | 5.472 | 5.564 | 4.62 | 4.84 | 4.879 | 4.805 |
| 30 | 6.18 | 6.02 | 5.977 | 6.013 | 5.93 | 5.53 | 5.581 | 5.628 | 5.55 | 5.55 | 5.573 | 5.644 | 5.48 | 4.98 | 5.037 | 5.017 |
| 31 | 6.21 | 6.28 | 6.333 | 6.323 | 5.75 | 5.50 | 5.844 | 5.857 | 5.83 | 5.79 | 5.871 | 5.921 | 5.46 | 5.16 | 5.313 | 5.249 |
| 32 | 6.25 | 6.12 | 6.016 | 6.038 | 5.48 | 5.57 | 5.557 | 5.596 | 5.98 | 5.55 | 5.571 | 5.591 | 4.88 | 4.98 | 5.043 | 4.940 |
| 33 | 6.39 | 6.34 | 6.466 | 6.433 | 5.79 | 5.74 | 5.980 | 5.956 | 5.89 | 5.84 | 5.938 | 5.935 | 5.25 | 5.20 | 5.399 | 5.287 |
| 34 | 6.41 | 6.12 | 6.104 | 6.125 | 5.71 | 5.60 | 5.667 | 5.706 | 5.93 | 5.64 | 5.690 | 5.759 | 5.31 | 5.05 | 5.140 | 5.106 |
| 35 | 6.41 | 6.35 | 6.424 | 6.404 | 5.66 | 5.64 | 5.770 | 5.744 | 5.81 | 5.67 | 5.767 | 5.813 | 5.03 | 5.03 | 5.145 | 5.001 |
| 36 | 6.45 | 6.54 | 6.623 | 6.581 | 6.19 | 6.16 | 6.300 | 6.319 | 6.02 | 6.19 | 6.291 | 6.287 | 5.74 | 5.60 | 5.750 | 5.663 |
| 37 | 6.54 | 6.12 | 6.376 | 6.333 | 6.05 | 5.56 | 5.907 | 5.875 | 5.93 | 6.16 | 5.894 | 5.852 | 5.60 | 5.45 | 5.494 | 5.429 |
| 38 | 6.77 | 6.56 | 6.468 | 6.426 | 6.21 | 6.25 | 6.210 | 6.194 | 6.54 | 6.30 | 6.376 | 6.310 | 5.69 | 5.72 | 5.878 | 5.638 |
| 39 | 6.90 | 6.40 | 6.559 | 6.522 | 5.75 | 5.67 | 5.903 | 5.848 | 6.05 | 5.72 | 5.833 | 5.851 | 5.27 | 5.07 | 5.228 | 5.060 |
| +[e] | | | 17 | 19 | | | 21 | 22 | | | 17 | 24 | | | 22 | 24 |
| ±[e] | | | 3 | 0 | | | 1 | 0 | | | 6 | 1 | | | 1 | 1 |
| −[e] | | | 17 | 18 | | | 15 | 15 | | | 12 | 10 | | | 14 | 12 |

[a] Data by Yoshimoto et al.[30] [b] The results calculated by eqs 8, 10, 12, and 14 in ref 30. [c] Results by the neural network with the parameter set A (Table I, part A). [d] Results by the neural network with the parameter set B (Table I, part B). [e] +, ±, and − show the numbers of cases in which the neural network is superior, equivalent, and inferior to the multiregression analysis, respectively.

**Table IV.** Comparison of Mean Deviation, Variance, and Standard Deviation

| | chronic injection | | | | | | single injection | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MED | | | OD | | | MED | | | OD | | |
| | set A[a] | set B[b] | MR[c] | set A[a] | set B[b] | MR[c] | set A[a] | set B[b] | MR[c] | set A[a] | set B[b] | MR[c] |
| MD | 0.17 | 0.16 | 0.20 | 0.15 | 0.14 | 0.19 | 0.20 | 0.19 | 0.23 | 0.16 | 0.14 | 0.20 |
| variance | 0.044 | 0.044 | 0.059 | 0.036 | 0.032 | 0.054 | 0.064 | 0.058 | 0.073 | 0.040 | 0.030 | 0.059 |
| SD | 0.21 | 0.21 | 0.24 | 0.19 | 0.18 | 0.23 | 0.25 | 0.24 | 0.27 | 0.20 | 0.17 | 0.24 |

[a] Results by the parameter set A (Table I, part A). [b] Results by the parameter set B (Table I, part B). [c] Results by the multiregression analysis.

quote those data and compare them with the results of the neural network.

The same biological data and structural parameters were used as those in the literature.[33,34] The input parameters are MR-3, $\pi$-3, MR-7, $\sigma_m$-3, F-4, R-4, I-1 and the squares of MR-3, $\pi$-3, MR-7, $\rho_m$-3, F-4, and R-4 and the constant 1 to make best use of the information of the structural parameters (where the number (-3, etc.) indicates the position of the structure shown in Table VI). As a rule, the input data were rescaled to have the values between ca. 0.1 and 1. The number of neurons of each layer and the parameters are shown in Table I, part C.

The input data and the results together with those in the literature are recorded in Table VI. If one compares

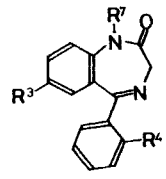
**Table V.** Variances of Leave-*n*-Out Results in Carboquinones[a]

| leave-*n*-out | $\sigma^2$ | $\sigma_L^2$ |
|---|---|---|
| 1[b] | 0.069 | 0.043 |
| 2 | 0.016 | 0.043 |
| 5 | 0.044 | 0.041 |
| 10 | 0.059 | 0.040 |

[a] Applied to MED of chronic injection. [b] Simple mean value of differences (observed value − calculated value).

(34) Kubota, T.; Yamakawa, M.; Terada, H.; Yoshimoto, M. In Structure-Activity Relationships—Quantative Approaches; the QSAR Research Group, Ed. *Kagaku no Ryoichi Supl. Ed., No. 122*, 1979.

**Table VI.** Input Data[a] and Results for Benzodiazepines




| | | | | | | | | | anti-pent effect[b] | | | anti-fighting behavior | | | clined screen test | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| no. | substituent | MR-3 | $\pi$-3 | MR-7 | $\sigma_m$-3 | F-4 | R-4 | I-1 | obsd[a] | clc 1[c] | NN[d] | obsd[a] | clc 2[c] | NN[d] | obsd[a] | clc 3[c] | NN[d] |
| 1 | 3-Cl-7-*i*-$C_5H_{11}$ | 0.60 | 0.71 | 2.42 | 0.37 | 0.0 | 0.0 | 0.0 | 4.99 | 4.76 | 4.876 | 3.53 | 3.78 | 3.791 | 2.83 | 3.15 | 2.856 |
| 2 | 3,4-$F_2$-7-$CH_3$ | 0.09 | 0.14 | 0.57 | 0.34 | 0.43 | −0.34 | 0.0 | 3.33 | 6.40 | 3.447 | | | | 2.85 | 4.05 | 2.658 |
| 3 | 3-$SC_2H_5$ | 1.84 | 1.07 | 0.10 | 0.15 | 0.0 | 0.0 | 1.0 | 3.57 | 4.00 | 3.874 | | | | | | |
| 4 | 3-Cl-7-$CH_2CONHCH_3$ | 0.60 | 0.71 | 1.92 | 0.37 | 0.0 | 0.0 | 0.0 | 5.83 | 4.95 | 5.012 | 3.63 | 3.95 | 3.917 | 2.88 | 3.36 | 2.932 |
| 5 | 3-$SC_4H_9$ | 2.77 | 2.07 | 0.10 | 0.15 | 0.0 | 0.0 | 1.0 | 3.79 | 3.50 | 3.859 | | | | | | |
| 6 | 3-$NO_2$-7-*i*-$C_5H_{11}$ | 0.74 | −0.28 | 2.42 | 0.71 | 0.0 | 0.0 | 0.0 | 4.80 | 5.02 | 5.027 | 3.77 | 4.28 | 3.989 | 2.94 | 3.42 | 2.691 |
| 7 | 3-$N(CH_3)_2$ | 1.56 | 0.18 | 0.10 | −0.15 | 0.0 | 0.0 | 1.0 | 3.84 | 3.85 | 3.994 | | | | | | |
| 8 | 3-Cl-4-$OCH_3$ | 0.60 | 0.71 | 0.10 | 0.37 | 0.26 | −0.51 | 1.0 | 4.60 | 5.30 | 4.638 | 3.48 | 4.83 | 3.448 | 3.00 | 3.38 | 2.610 |
| 9 | 3-Cl-7-$(CH_2)_3N(CH_3)_2$ | 0.60 | 0.71 | 2.95 | 0.37 | 0.0 | 0.0 | 0.0 | 3.38 | 4.56 | 4.712 | | | | 3.01 | 2.93 | 2.901 |
| 10 | 3-Cl-7-$(CH_2)_3OH$ | 0.60 | 0.71 | 1.65 | 0.37 | 0.0 | 0.0 | 0.0 | 4.34 | 5.05 | 5.081 | | | | 3.04 | 3.48 | 3.047 |
| 11 | 3-$NO_2$-7-$CH_2CONHCH_3$ | 0.74 | −0.28 | 1.92 | 0.71 | 0.0 | 0.0 | 0.0 | 5.29 | 5.21 | 5.269 | 4.55 | 4.45 | 4.419 | 3.07 | 3.63 | 2.859 |
| 12 | 3-Cl-7-$(CH_2)_2N(C_2H_5)_2$ | 0.60 | 0.71 | 3.41 | 0.37 | 0.0 | 0.0 | 0.0 | 5.06 | 4.39 | 4.545 | 3.57 | 3.45 | 3.724 | 3.09 | 2.73 | 3.009 |
| 13 | 3-Cl-7-$CH_2CON(CH_3)_2$ | 0.60 | 0.71 | 2.39 | 0.37 | 0.0 | 0.0 | 0.0 | 5.35 | 4.77 | 4.884 | 4.25 | 3.79 | 3.796 | 3.11 | 3.16 | 2.856 |
| 14 | 4-F-7-$CH_3$ | 0.10 | 0.0 | 0.57 | 0.0 | 0.43 | −0.34 | 0.0 | 4.53 | 6.06 | 4.359 | 3.83 | 4.26 | 3.922 | 3.13 | 3.56 | 2.681 |
| 15 | 3-Cl-7-$CH_2C_6H_5$ | 0.60 | 0.71 | 3.00 | 0.37 | 0.0 | 0.0 | 0.0 | 4.64 | 4.54 | 4.696 | 3.56 | 3.59 | 3.732 | | | |
| 16 | 3-Cl-7-$(CH_2)_2N(CH_3)_2$ | 0.60 | 0.71 | 2.48 | 0.37 | 0.0 | 0.0 | 0.0 | 4.73 | 4.74 | 4.858 | 3.53 | 3.76 | 3.781 | 3.14 | 3.13 | 2.856 |
| 17 | 3-Cl-4-F-7-$(CH_2)_3N(CH_3)_2$ | 0.60 | 0.71 | 2.95 | 0.37 | 0.43 | −0.34 | 0.0 | 4.76 | 4.27 | 5.193 | 3.97 | 4.18 | 4.244 | 3.17 | 3.19 | 3.167 |
| 18 | H | 0.10 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 1.0 | | | | 3.37 | 3.44 | 3.301 | 3.20 | 3.08 | 2.967 |
| 19 | 3-$CF_3$-7-$CH_2CONHCH_3$ | 0.50 | 0.88 | 1.92 | 0.43 | 0.00 | 0.00 | 0.0 | 5.06 | 5.06 | 5.088 | 3.69 | 4.07 | 4.108 | 3.20 | 3.48 | 3.080 |
| 20 | 3-$SCH_3$ | 1.38 | 0.61 | 0.10 | 0.15 | 0.00 | 0.00 | 1.0 | 4.15 | 5.24 | 3.953 | 4.15 | 3.76 | 4.109 | 3.21 | 3.41 | 3.231 |
| 21 | 3-Cl-7-$(CH_2)_3OH$ | 0.60 | 0.71 | 1.65 | 0.37 | 0.00 | 0.00 | 0.0 | | | | 3.61 | 4.04 | 4.024 | | | |
| 22 | 3-Cl-7-$CH_2CONH_2$ | 0.60 | 0.71 | 1.44 | 0.37 | 0.00 | 0.00 | 0.0 | 5.07 | 5.13 | 5.132 | 4.21 | 4.11 | 4.124 | 3.21 | 3.57 | 3.182 |
| 23 | 3-F | 0.09 | 0.14 | 0.10 | 0.34 | 0.00 | 0.00 | 1.0 | | | | 3.40 | 4.05 | 3.490 | 3.23 | 3.57 | 3.259 |
| 24 | 3-$SOCH_3$ | 1.37 | −1.58 | 0.10 | 0.52 | 0.00 | 0.00 | 1.0 | 4.08 | 4.61 | 4.338 | | | | 3.23 | 3.48 | 2.823 |
| 25 | 3-Cl-4-$CH_3$ | 0.60 | 0.71 | 0.10 | 0.37 | −0.04 | −0.13 | 1.0 | 4.57 | 4.81 | 5.115 | 3.45 | 4.28 | 3.727 | 3.28 | 3.36 | 3.583 |
| 26 | 3-$N(CH_3)_2$-7-$CH_3$ | 1.56 | 0.18 | 0.57 | −0.15 | 0.00 | 0.00 | 0.0 | 4.69 | 4.43 | 4.561 | 3.86 | 3.44 | 3.849 | 3.29 | 3.13 | 2.952 |
| 27 | 3-Cl-4-F-7-$(CH_2)_2N(C_2H_5)_2$ | 0.60 | 0.71 | 3.41 | 0.37 | 0.43 | −0.34 | 0.00 | 5.38 | 5.09 | 5.187 | 4.29 | 4.02 | 4.215 | 3.29 | 3.00 | 3.322 |
| 28 | 3-$NO_2$-7-$(CH_2)_3N(CH_3)_2$ | 1.56 | 0.18 | 0.57 | 0.37 | 0.41 | −0.15 | 0.0 | 5.82 | 5.10 | 6.100 | 3.61 | 3.80 | 3.608 | 3.34 | 3.01 | 2.792 |
| 29 | 3-$NO_2$-7-$(CH_2)_3N(CH_3)_2$ | 0.74 | −0.28 | 2.95 | −0.15 | 0.00 | 0.00 | 0.0 | 4.28 | 4.82 | 4.573 | 3.96 | 4.10 | 3.885 | 3.35 | 3.20 | 3.218 |
| 30 | 3-$NO_2$-7-$(CH_2)_2N(CH_3)_2$ | 0.74 | −0.28 | 2.48 | 0.71 | 0.00 | 0.00 | 0.0 | 4.70 | 5.00 | 5.002 | 4.25 | 4.26 | 3.938 | 3.37 | 3.40 | 2.680 |
| 31 | 3-Cl-4-Cl | 0.60 | 0.71 | 0.10 | 0.37 | 0.41 | −0.15 | 1.0 | 5.85 | 5.54 | 6.216 | 5.18 | 4.52 | 4.960 | 3.48 | 3.61 | 3.654 |
| 32 | 3-Cl-7-$CH_2$-cyc-$C_3H_5$ | 0.60 | 0.71 | 1.82 | 0.37 | 0.00 | 0.00 | 0.0 | 4.90 | 4.99 | 5.038 | 4.51 | 3.99 | 3.953 | 3.51 | 3.41 | 2.967 |
| 33 | 3-CN | 0.63 | −0.57 | 0.10 | 0.56 | 0.00 | 0.00 | 1.0 | 5.30 | 5.04 | 5.458 | 3.81 | 4.37 | 4.004 | 3.54 | 3.73 | 3.358 |
| 34 | 3-$NO_2$-4-$CF_3$ | 0.74 | −0.28 | 0.10 | 0.71 | 0.38 | 0.19 | 1.0 | 5.70 | 5.76 | 5.835 | 4.54 | 4.64 | 4.371 | 3.54 | 3.93 | 3.229 |
| 35 | 3-Cl | 0.60 | 0.71 | 0.10 | 0.37 | 0.00 | 0.00 | 1.0 | 4.65 | 4.87 | 5.118 | 4.13 | 4.16 | 4.168 | 3.56 | 3.73 | 3.924 |
| 36 | 3-CN-4-F | 0.63 | −0.57 | 0.10 | 0.56 | 0.43 | −0.34 | 1.0 | 5.63 | 5.75 | 5.819 | 4.75 | 4.94 | 4.663 | 3.57 | 4.00 | 3.136 |
| 37 | 3-Cl-7-$C_2H_5$ | 0.60 | 0.71 | 1.03 | 0.37 | 0.00 | 0.00 | 0.0 | 4.90 | 5.28 | 5.226 | 4.17 | 4.25 | 4.350 | 3.60 | 3.74 | 3.549 |
| 38 | 3-$SCH_3$-7-$CH_3$ | 1.38 | 0.61 | 0.57 | 0.15 | 0.00 | 0.00 | 0.0 | 3.60 | 3.64 | 4.267 | 3.87 | 4.34 | 3.915 | 3.77 | 4.25 | 3.190 |
| 39 | 3-Cl-7-$CH_2COCH_3$ | 0.60 | 0.71 | 1.51 | 0.37 | 0.00 | 0.00 | 0.0 | 5.28 | 5.10 | 5.115 | 4.21 | 4.09 | 4.089 | 3.82 | 3.54 | 3.133 |
| 40 | 3-Cl-4-Br | 0.60 | 0.71 | 0.10 | 0.37 | 0.44 | −0.17 | 1.0 | 5.77 | 5.59 | 6.298 | 4.85 | 4.56 | 5.003 | 3.85 | 3.42 | 3.648 |
| 41 | 3-Cl-4-F | 0.60 | 0.71 | 0.10 | 0.37 | 0.43 | −0.34 | 1.0 | 6.46 | 5.58 | 6.112 | 4.76 | 4.73 | 4.846 | 3.86 | 3.99 | 3.483 |
| 42 | 3-$NO_2$-4-$CF_3$-7-$CH_3$ | 0.74 | −0.28 | 0.57 | 0.71 | 0.38 | 0.19 | 0.0 | 5.71 | 5.58 | 5.542 | 4.86 | 4.89 | 4.907 | 3.86 | 4.15 | 3.021 |
| 43 | 3-$CF_3$-7-$(CH_2)_2N(CH_3)_2$ | 0.50 | 0.88 | 2.48 | 0.43 | 0.00 | 0.00 | 0.0 | 4.76 | 4.85 | 4.939 | 4.27 | 3.89 | 3.897 | 3.88 | 3.24 | 2.902 |
| 44 | 3-Cl-7-$CH_2CH{=}CH_2$ | 0.60 | 0.71 | 1.45 | 0.37 | 0.00 | 0.00 | 0.0 | 5.35 | 5.13 | 5.130 | 4.49 | 4.11 | 4.119 | 3.89 | 3.56 | 3.175 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 45 | 3,4-$Cl_2$-7-$CH_3$ | 0.60 | 0.71 | 0.57 | 0.37 | 0.41 | −0.15 | 0.0 | 6.03 | 6.13 | 5.929 | 4.50 | 4.76 | 4.400 | 3.90 | 3.82 | 3.297 |
| 46 | 3-Cl-7-$CH_3$ | 0.60 | 0.71 | 0.57 | 0.37 | 0.00 | 0.00 | 0.0 | 5.31 | 5.46 | 5.313 | 4.45 | 4.41 | 4.606 | 3.98 | 3.94 | 4.043 |
| 47 | 3-$NO_2$-4-Cl-7-$CH_3$ | 0.74 | −0.28 | 0.57 | 0.71 | 0.41 | −0.15 | 0.0 | 6.92 | 6.39 | 6.732 | 5.12 | 5.26 | 5.210 | 4.04 | 4.09 | 3.618 |
| 48 | 3-$CF_3$-4-$CF_3$ | 0.50 | 0.88 | 0.10 | 0.43 | 0.38 | 0.19 | 1.0 | 4.97 | 5.61 | 5.239 | 4.27 | 4.27 | 4.400 | 4.09 | 3.78 | 3.992 |
| 49 | 3-Br | 0.89 | 0.86 | 0.10 | 0.39 | 0.00 | 0.00 | 1.0 | 5.20 | 4.74 | 4.825 | 4.20 | 4.21 | 4.408 | 4.10 | 3.78 | 4.063 |
| 50 | 3-CN-7-$CH_3$ | 0.63 | −0.57 | 0.57 | 0.56 | 0.00 | 0.00 | 0.0 | 5.44 | 5.63 | 5.524 | 4.74 | 4.61 | 4.811 | 4.14 | 3.95 | 3.380 |
| 51 | 3-$NO_2$ | 0.74 | −0.28 | 0.10 | 0.71 | 0.00 | 0.00 | 1.0 | 5.60 | 5.13 | 5.652 | 4.75 | 4.66 | 4.568 | 4.27 | 4.00 | 4.152 |
| 52 | 3-$NO_2$-7-$CH_3$ | 0.74 | −0.28 | 0.57 | 0.71 | 0.00 | 0.00 | 0.0 | 5.62 | 5.72 | 5.986 | 5.07 | 4.90 | 5.130 | 4.47 | 4.21 | 4.286 |
| 53 | 3-$CF_3$ | 0.50 | 0.88 | 0.10 | 0.43 | 0.00 | 0.00 | 1.0 | 5.48 | 4.99 | 5.375 | 4.48 | 4.28 | 4.378 | 4.48 | 3.84 | 4.185 |
| 54 | 3-Cl-4-F-7-$CH_3$ | 0.60 | 0.71 | 0.57 | 0.37 | 0.00 | −0.34 | 0.0 | 5.88 | 5.46 | 5.806 | 4.78 | 4.77 | 4.823 | 4.48 | 3.94 | 4.071 |
| 55 | 3-$CF_3$-7-$CH_2CH{=}CH_2$ | 0.50 | 0.88 | 1.45 | 0.43 | 0.00 | 0.00 | 0.0 | 5.03 | 5.24 | 5.211 | 4.54 | 4.23 | 4.382 | 4.54 | 3.68 | 3.469 |
| 56 | 3-$NO_2$-4-Cl | 0.74 | −0.28 | 0.10 | 0.71 | 0.41 | −0.15 | 1.0 | 6.30 | 5.80 | 6.355 | 4.20 | 5.02 | 4.537 | | | |
| 57 | 3-$NO_2$-7-$NH_2$ | 0.74 | −0.28 | 0.54 | 0.71 | 0.00 | 0.00 | 0.0 | 5.77 | 5.73 | 5.998 | 5.17 | 4.91 | 5.140 | 4.77 | 4.22 | 4.326 |
| 58 | 3-$NO_2$-4-$NO_2$ | 0.74 | −0.28 | 0.10 | 0.71 | 0.67 | 0.16 | 1.0 | 5.97 | 6.23 | 6.011 | 3.91 | 4.81 | 3.951 | | | |
| 59 | 3-$NO_2$-4-F | 0.74 | −0.28 | 0.10 | 0.71 | 0.43 | −0.34 | 1.0 | 6.00 | 5.84 | 6.013 | 5.17 | 5.23 | 5.140 | 4.88 | 4.26 | 4.502 |
| 60 | 3-$NO_2$-4-F-7-$CH_3$ | 0.74 | −0.28 | 0.57 | 0.71 | 0.43 | −0.34 | 0.0 | 6.50 | 6.42 | 6.554 | 5.80 | 5.47 | 5.626 | 5.50 | 4.47 | 4.826 |

[a] The data were taken from the literature (ref 34). [b] Anti-pentylenetetrazole effect. [c] Results by the multiregression analysis.[34] [d] Results by the neural network.

**Table VII.** Comparison of Mean Deviation, Variation, and Standard Deviation

| | anti-pent[a] | | anti-fight[b] | | clined scr[c] | |
|---|---|---|---|---|---|---|
| | NN[d] | MR[e] | NN[d] | MR[f] | NN[d] | MR[g] |
| MD[h] | 0.24 | 0.39 | 0.16 | 0.30 | 0.25 | 0.33 |
| variance | 0.11 | 0.37 | 0.04 | 0.15 | 0.09 | 0.17 |
| SD[i] | 0.33 | 0.61 | 0.21 | 0.39 | 0.31 | 0.41 |

[a] Anti-pentylenetetrazole effect. [b] Anti-fighting behavior. [c] Clined screen test. [d] Results by the neural network. [e-g] Results calculated by eqs 33, 34, and 35 in ref 34, respectively. [h] Mean deviation. [i] Standard deviation.

**Table VIII.** Variance of Leave-$n$-Out Results in Benzodiazepines[a]

| leave-$n$-out | $\sigma^2$ | $\sigma_L{}^2$ |
|---|---|---|
| 1[b] | 0.280 | 0.081 |
| 2 | 0.460 | 0.084 |
| 5 | 0.539 | 0.076 |
| 10 | 0.913 | 0.077 |

[a] Applied to anti-pentylenetetrazole effect. [b] Simple mean value of differences (observed value − calculated value).

the present results with those of multiregression analysis, it is found that the neural network gives better results in 96 cases, worse results in 62 cases, and the comparable results in 5 cases.

In order to compare the reliability of the calculated results in Table VI, we obtained the average deviations, variances, and standard deviations for the results by both the neural network and the multiregression analysis. The results are shown in Table VII. The number of samples are about 54. The table shows that the ratios of variance ($F$: the multiregression analysis over the neural network) are $1.89 \leq F \leq 3.75$. Therefore, it may be definitely said that the neural network reproduces the observed values better than the multiregression analysis.

The leave-$n$-out experiments have been applied to the data of anti-pentylenetetrazole effect in Table VI. The $57 - n$ data which are formed by haphazardly removing $n$ number of data from the 57 data, were input to the network in training phase. Then the removed data were calculated to give the variances. This procedure was repeated 57 times and the averaged variances ($\sigma^2$) and those for the data used for training ($\sigma_L{}^2$) were shown in Table VIII.

The neural network performs a nonlinearity fitting for the input data with nonlinear relationships. The degree of such a fitting is determined by the characteristics (or quality in terms of the sense of linearity) of the given data. It is easily expected that the larger the degree of the nonlinearity fitting is, the larger the deviation of expectation is. Unlike the case of carboquinones, nonlinearity fitting in the case of benzodiazepines appears intensive resulting in one-digit larger $\sigma^2$ than those of carboquinones. Such large variances totally stem form the input data of poor linearity.

## Concluding Remarks

It may be useful to note the differences of operation between the linear multiregression analysis and the neural network. In the linear multiregresson analysis, the relationship between the biological activities and the structural parameters is expressed by a linear combination of the contributing terms. The coefficients of contribution are determined by the least-squares method. Here, it is necessary to effectively select the contributing terms considering the rationale of each term. Therefore, the quality and the number of the terms are greatly dependent on the experience and the knowledge of the analyzer.

The neural network, on the contrary, does not require such a comprehensive term selection. This is a merit as well as a shortcoming of this method. Namely, one can analyze the given data without knowing special techniques. However, even if the appropriate results are obtained, the definitive reason may not be given.

The neural network studied here performs two processes at the same time: the process to convert the input data to the effective form and the process to classify the converted data referring to the characteristics. The former process is carried out by the first and second layers and the latter, by the second and third layers. Since these two processes are optimized to the training patterns, it is very probable for such a neural network to exceed the level of the multiregressional analysis as shown in this paper. Especially, the neural network becomes superior in such cases that the analysis includes a large number of the structural parameters or expansion terms compared to the number of the obtained biological data. However, one may wonder why it should be possible to determine, for example, 96 weights in model A or 420 in model C in Table I when a considerably small number of experimental data is used.

The operation of the neural network is very different from that of the usual multiregression analysis. The information in the given data is accumulated in the weight matrices as the number of the input data is increased. The decision by the network is very much like that of the brain of human: the number of given data seems to be how much a man experienced the situation. Namely, the larger the number of data and the better the quality of the data, the better the network gives the decision. Unlike the multiregression analysis, however, the reliability of the decision cannot be treated statistically at present.

Finally, it should be mentioned that one of the unfamiliar situations in the network is the uncertainty of weight matrices. Namely, the operation can be exactly defined by the mathematical expressions (eqs 1–5). However, the matrices do not always take the same definitive elements even if they given the definitive decision. For example, consider the case that the weight between the first and second layers which coagulates into neuron $j_0$ (of the second layer) can be dispersed into $j_1$ and $j_2$ (by simply increasing the number of neurons in the second layer). Then,

$$W_{ij0} = \lambda_1 W_{ij1} + \lambda_2 W_{ij2} \quad (13)$$

where $\lambda_1$ and $\lambda_2$ are coefficients. Noticeably, $W_{ij1}$ and $W_{ij2}$ are indefinite although they are controlled by the $\lambda$ values. Therefore, the weight matrices do not always take the fixed elements even if they give the same results.

# Cyclization-Activated Prodrugs. Basic Esters of 5-Bromo-2′-deoxyuridine

Walfred S. Saari,* John E. Schwering, Paulette A. Lyle, Steven J. Smith, and Edward L. Engelhardt

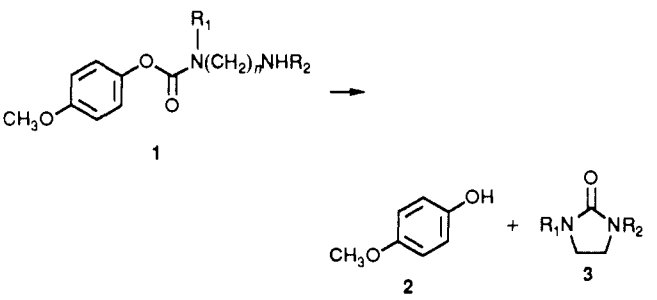
*Merck Sharp & Dohme Research Laboratories, West Point, Pennsylvania 19486.* *Received February 20, 1990*

Some 3′- and 5′-[[(alkylamino)ethyl]glycyl] esters of 5-bromo-2′-deoxyuridine were prepared and evaluated in vitro as progenitors of the parent alcohol. The esters proved to be relatively stable at low pH but released 5-bromo-2′-deoxyuridine cleanly at rates which were pH and structure dependent. These basic esters are examples of cyclization-activated prodrugs in which generation of active drug is not linked to enzymatic cleavage but rather results from an intramolecular cyclization–elimination reaction.

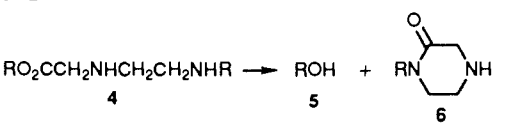
Ester prodrugs of alcohols are frequently utilized to circumvent adverse physicochemical limitations or to extend the duration of action of the parent drug.[1-3] Generally, ester prodrugs have depended upon chemical or enzymatic hydrolysis of the ester bond for conversion of prodrug to drug. However this strategy can only be successful in those cases where the alcohol is generated from the ester at a practical rate under physiological conditions. When this requirement is not attainable, this approach will fail or be of limited value. In addition, generation of drug by enzymatic mechanisms may be subject to much variability between species or even among individual members of a particular species.

A previous report[4] described some basic carbamate prodrugs (1) of the melanocytoxic agent 4-hydroxyanisole which generated the parent phenol 2 by a cyclization–activated mechanism under physiological conditions (Scheme I). In this approach, prodrug is converted to active drug by an intramolecular cyclization–elimination reaction and not through mechanisms involving intermolecular hydrolysis of the ester bond. By this method, ideally, drug formation is not dependent upon the host environment but instead solely upon the rate of the intramolecular cyclization reaction.

Scheme I




Scheme II




Although basic carbamates of phenols are sufficiently activated to generate phenol at useful rates under physiological conditions, the corresponding carbamates of al-

(1) Stella, V. J.; Charman, W. N. A.; Naringrekar, V. H. *Drugs* **1985**, *29*, 455.
(2) Stevenson, L-A.; Tunek, A. *Drug Metab. Rev.* **1988**, *19*, 165.
(3) Bundgaard, H.; Falch, E.; Jensen, E. *J. Med. Chem.* **1989**, *32*, 2507.
(4) Saari, W. S.; Schwering, J. E.; Lyle, P. A.; Smith, S. J.; Engelhardt, E. L. *J. Med. Chem.* **1990**, *33*, 97.

0022-2623/90/1833-2590$02.50/0 