Prediction of Receptor Properties and Binding Affinity of Ligands to Benzodiazepine/$GABA_A$ Receptors Using Artificial Neural Networks

Desmond J. Maddalena* and Graham A. R. Johnston

The Department of Pharmacology, The University of Sydney, NSW 2006, Australia

Received February 8, 1994[⊗]

To date the use of artificial neural networks (ANNs) in quantitative structure–activity relationship (QSAR) studies has been primarily concerned in comparing the predictive accuracy of the technique using known data sets where the data set parameters had been preselected and optimized for use with other statistical methods. Little effort has been directed at optimizing the input parameters for use with ANNs or exploring other potential strengths of ANNs. In this study, back-propagation ANNs and multilinear regression (MLR) were used to examine the QSAR between substituent constants and random noise at six positions on 57 1,4-benzodiazepin-2-ones (1,4-BZs) and their binding affinities (log IC_{50}) for benzodiazepine $GABA_A$ receptor preparations. By using selective pruning and cross-validation techniques, it was found possible to use ANNs to indicate an optimum set of 10 input parameters from a choice of 48 which were then used to train back-propagation ANNs that best predicted the receptor binding affinity with a high correlation between known and predicted data sets. Using the optimum set of input parameters, three-layer ANNs performed no better than the two-layer ANNs which gave marginally better results than MLR. Using the trained ANNs to examine the individual parameters showed that increases in the lipophilicity and $\mathcal{F}$polar value at position 7, $\mathcal{F}$polar value at position 2′, and dipole at position 1 on the molecule all enhanced receptor binding affinity of 1,4-BZ ligands. Increases in molar refractivity and resonance parameters at position 1, molar refractivity at positions 6′ and 2′, Hammet *meta* constant at position 3′, and Hammet *para* constant at position 8 on the molecule all caused decreases in receptor binding affinity. By considering the optimal ANNs as pharmacophore models representing the internal physicochemical structure of the receptor site, it was found that they could be used to critically examine the properties of the receptor site.

Introduction

Over the last two decades, benzodiazepines (BZs) have been widely used therapeutically for their ability to reduce anxiety and act as tranquilizers and for their anticonvulsant effects in epilepsy.[1] More recently, via positron emission tomography (PET), radionuclide-labeled BZs have been found useful in diagnosis of a variety of neurological disorders.[2] In both the therapeutic and diagnostic cases, the usefulness of BZs has been found to be related to their benzodiazepine/$GABA_A$ receptor affinity. For therapy, a high *in vitro* receptor affinity is usually synonymous with high biological activity,[3] and in clinical diagnosis, a high *in vitro* receptor affinity is necessary for adequate *in vivo* imaging specificity.[2] While more than 20 quantitative structure–activity relationship (QSAR) studies have been carried out on BZs, only a few of these were wholly concerned with molecular structure–receptor affinity[4–9] and the relationships found had variable predictive capabilities and gave little information about the receptor itself.

Recently artificial neural networks (ANNs) have found uses in several biological chemistry-related areas[10–12] including QSAR studies.[13] ANNs have been used to relate the molecular physicochemical parameters of carboqinones and benzodiazepines;[14] dihydrofolate reductase inhibitors[15,16] and 4-(*R*)-phenylthiopropyl heterocycles to their biological activities;[17] and benzodiazepines to their receptor affinities.[9]

In four of the studies, the performance of the ANNs at predicting results (i.e., generalizing) from input parameters such as Hansch type substituent constants[18] was compared with similar studies carried out using statistical techniques such as multilinear regression (MLR). In three of the four studies, the ANNs were found to give statistically significantly better results at prediction than the MLR methods. In the fourth study,[9] the predictive performance of ANNs and the MLR techniques was similar.

Classically QSAR studies are used for both selection of principal parameters involved in structure–activity relationships and the derivation of such relationships so that they may be used for predictive purposes in drug design. The ANN QSAR studies to date have been concerned primarily in demonstrating the utility of the technique at providing more accurate predictions from known data sets where the data set parameters were already selected and optimized for the MLR studies. No efforts were directed at optimizing the input parameters for the ANNs.

Since the ANNs characteristically have a high resistance to noise, they might be expected to give a strong internal weighting to those input parameters that have a high information content and a low internal weighting to those parameters with little or none. Consequently, it should be possible to use the ANNs to both select the important parameters and derive useful predictive relationships between them.

In the current study, two- and three-layer back-propagation ANNs and MLR were used to explore the relationships between the seven substituent constants:

* Author for correspondence. Tel: 612 521 1120. Fax: 612 545 4174.
[⊗] Abstract published in *Advance ACS Abstracts*, December 15, 1994.

0022-2623/95/1838-0715$09.00/0

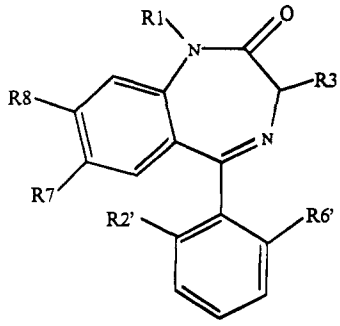

Figure 1. 1,4-Benzodiazepin-2-ones.

lipophilicity (π), molar refractivity (MR), polar constant ($\mathcal{F}$), resonance constant ($\mathcal{R}$), Hammet *meta* constant (σ_m), Hammet *para* constant (σ_p), and aromatic group dipole (μ) for the functional groups R7, R1, R3, R8, R2', and R6' (see Figure 1) of 57 1,4-benzodiazepin-2-ones (1,4-BZs) and their binding affinities measured as the log IC_{50} of the compounds.

Methods

Neural Network Model. A back-propagation ANN program called Bioactivnet[19] was used in the studies. The program was run in Microsoft Windows 3.1 on an IBM clone 486/33 MHz personal microcomputer in conjunction with Microsoft Excel 4.0, a spreadsheet program, with which it interacts. The back-propagation algorithm used was similar to the classic back-propagation algorithm[20] with an added momentum term. The momentum term was to keep the network from oscillating at high learning rates. The theory has been discussed by several others[10,11,14-16,20] so will not be further discussed herein.

The program was constructed to allow variable numbers of neurons in the input and hidden layers with a single output neuron corresponding to the predicted log IC_{50}. Both two- and three-layer ANNs were used during the studies. All inputs were scaled to a range of 0.1−0.9 on the basis of the minimum and maximum of the input source range. Inverse scaling was used for the outputs. Sigmoidal transfer functions were used in all layers.

The networks were trained for 8000 learning cycles. The first 3000 cycles were run at a learning rate of 0.6 and a momentum of 0.4; thereafter cycles were run at a learning rate of 0.2 and a momentum of 0.1. The program saved the ANN parameters that gave the lowest error during training. This typically occurred between 5000 and 8000 learning cycles when the standard deviation of the error between the net output and the training data for a randomly submitted epoch of patterns was less than 10% of the maximum of the training data.

Optimization of ANN Input Parameters. The optimum input parameters used in prediction of receptor affinity (log IC_{50}) were determined using a method where the inputs not contributing significantly to the prediction of the result were systematically eliminated (pruned) to improve generalization[21] as follows.

A two-layer back-propagation ANN using 48 neurons in the input layer and one in the output layer was used. The 48 input parameters consisted of seven substituent constants and a randomly generated data point for each of the six positions around the molecule. The random data points taken from a uniformly distributed random data set were to give the network a choice between real data and random data. If the random data had greater weights than the real data, then it could be concluded that the real data, was no better than the random data. The single output represented the predicted log IC_{50}.

The mean absolute weights of the connections between the input and output layers were calculated, and the inputs with the lowest weights for each of the six positions around the molecule were pruned. Where two parameters had weights similar in magnitude, but the first was highly correlated with a third parameter with a greater weight and the second independent of all other parameters, the highly correlated parameter was pruned to force a selection of independent parameters. The process was repeated with 42, 36, 30, 24, 18, 12, and 6 inputs. The bias connection weights were ignored as they had no effect on the selection process.

The ability of the ANNs to generalize was examined by using a leave-one-out ($N - 1$) cross-validation technique as follows. For each given set of input parameters using 57 patterns, 57 ANNs were trained each using 56 patterns. At the end of training, each trained ANN was used to predict the log IC_{50} of the missing pattern. In every study the correlation, standard error, and leave-one-out cross-validation correlation between the predicted and known data sets were calculated using 57 trained ANNs. Once an optimum training set was obtained, the ability of the ANNs to generalize was further tested by dividing the total set of 57 patterns into 20 random training sets of 38 and test sets of 19.

Using three-layer ANNs and the optimal parameter set from the two-layer study, leave-one-out cross-validation studies were used to determine the optimum number of neurons in the hidden layer. Successive series of ANNs where the number of inputs was kept constant and the numbers in the hidden layer increased between one and eight were cross-validated. The ANNs with the highest correlation between cross-validation predicted and known data were considered optimum.

Calculation of Optimal Function Groups. To calculate the optimal function groups, a score was determined for each functional group by summing the proportional contributions of the normalized optimal parameters for each position. The set of functional groups was then sorted in descending order of scores, the functional groups with the highest scores being the best functional groups for use at that position. For example, at position 7, the sum of the π and $\mathcal{F}$ parameters produced the optimal result. The score was determined by first scaling each of the elements of the two function group parameter sets to lie in the range from 0 to 1, multiplying each element by its connection weight proportion, and then summing the pairs of modified parameter values. The connection weight proportions were determined by dividing the connection weight of each parameter by the sum of the connection weights of the two parameters acting at position 7.

Statistical Studies. Since two-layer ANNs have been considered to be functionally equivalent to MLR,[14,22] the prediction abilities of the ANNs were compared to those of MLR using the leave-one-out and leave-19-out cross-validation techniques. In all studies the optimal parameter set from the parameter optimization studies and matched data sets were used. The statistical analyses and random number generation were carried out using the statistical analysis tools available in the Microsoft Excel 4 program.

The Data Set. The data set (see Table 1) of 57 1,4-benzodiazepin-2-ones with groups substituted in six positions around the molecule (see Figure 1) was obtained from the literature.[3] The data set was selected as it was a homogeneous set from a single well-established laboratory. Not all compounds available were used in the study. The only criterion for nonuse was that the substituent constants were not readily available from the source literature. Where compounds were substituted in position 3, the results were for the more active (*S*) enantiomer or, where this was not available, the racemic mixture as indicated in the source literature.[3] The substituent constants were obtained from three sources.[18,23,24] The dipole substituent constants of three compounds marked with an asterisk were calculated using MLR (see Table 2).

Results

Optimization of Input Parameters. The results of pruning the inputs are shown in Table 3, which shows the training and cross-validation correlations (R_t and R_{cv}) and the training standard errors (SE_t) between the ANN-predicted and known receptor affinities for nets

Table 1. Data Set[a]

name	R7	R1	R2′	R6′	R3	R8	known	pred	diff
Ro 05-3061	F	H	H	H	H	H	1.602	1.372	0.230
Ro 05-4865	F	Me	H	H	H	H	1.230	1.638	−0.408
Ro 05-6820	F	H	F	H	H	H	0.869	0.637	0.232
Ro 05-6822	F	Me	F	H	H	H	0.708	0.861	−0.153
nordazepam	Cl	H	H	H	H	H	0.973	1.090	−0.116
diazepam	Cl	Me	H	H	H	H	0.908	1.352	−0.444
Ro 05-3367	Cl	H	F	H	H	H	0.301	0.432	−0.131
delorazepam	Cl	H	Cl	H	H	H	0.255	0.274	−0.019
Ro 07-9957	I	Me	F	H	H	H	0.462	0.469	−0.006
Ro 05-2904	CF_3	H	H	H	H	H	1.114	1.039	0.075
Ro 14-3074	N_3	H	F	H	H	H	0.724	0.638	0.086
nitrazepam	NO_2	H	H	H	H	H	1.000	1.277	−0.277
Ro 05−4435	NO_2	H	F	H	H	H	0.176	0.565	−0.389
flunitrazepam	NO_2	Me	F	H	H	H	0.580	0.778	−0.198
clonazepam	NO_2	H	Cl	H	H	H	0.255	0.384	−0.128
Ro 05-4082	NO_2	Me	Cl	H	H	H	0.342	0.564	−0.222
Ro 05-3590	NO_2	H	CF_3	H	H	H	0.544	0.455	0.089
Ro 20-7736	NHOH	Me	F	H	H	H	1.982	2.117	−0.135
Ro 05-3072	NH_2	H	H	H	H	H	2.587	2.500	0.087
Ro 05-3418	NH_2	Me	H	H	H	H	2.663	2.646	0.017
Ro 20-1815	NH_2	Me	F	H	H	H	1.813	2.112	−0.299
Ro 05-4619	NH_2	H	Cl	H	H	H	1.875	1.615	0.260
Ro 05-4528	CN	Me	H	H	H	H	2.580	1.909	0.671
Ro 20-2541	CN	Me	F	H	H	H	1.477	0.123	0.354
Ro 20-2533	Et	H	H	H	H	H	1.556	1.541	0.015
Ro 20-5747	$CH{=}CH_2$	H	H	H	H	H	1.380	1.488	−0.108
Ro 20-5397	CHO	H	H	H	H	H	1.633	1.954	−0.321
Ro 20-3053	COMe	H	F	H	H	H	1.255	1.104	0.151
Ro 05-2921	H	H	H	H	H	H	2.544	2.012	0.532
Ro 05-4336	H	H	F	H	H	H	1.322	1.234	0.088
Ro 05-4520	H	Me	F	H	H	H	1.146	1.501	−0.355
Ro 05-4608	H	Me	Cl	H	H	H	0.580	1.233	−0.653
halazepam	Cl	CH_2CF_3	H	H	H	H	1.964	2.404	−0.440
Ro 06-9098	NO_2	CH_2OCH_3	H	H	H	H	2.633	2.478	0.156
Ro 20-1310	Cl	$C(CH_3)_3$	H	H	H	H	2.792	2.544	0.249
Ro 07-2750	Cl	$(CH_2)_2OH$	F	H	H	H	1.389	1.027	0.362
Ro 22-4683	NO_2	$C(CH_3)_3$	F	H	H	H	2.477	2.111	0.366
Ro 07-4419	H	H	F	F	H	H	1.279	1.312	−0.034
Ro 07-3953	Cl	H	F	F	H	H	0.204	0.485	−0.281
Ro 07-4065	Cl	Me	F	F	H	H	0.613	0.685	−0.072
Ro 07-5193	Cl	H	Cl	F	H	H	0.477	0.317	0.160
Ro 22-3294	Cl	H	Cl	Cl	H	H	0.845	0.609	0.236
Ro 07-5220	Cl	Me	Cl	Cl	H	H	0.740	0.828	−0.088
Ro 13-3780	Br	Me	F	F	H	H	0.380	0.589	−0.208
Ro 11-4878	Cl	H	F	H	Me	H	0.544	0.405	0.139
meclonazepam	NO_2	H	Cl	H	Me	H	0.079	0.357	−0.278
Ro 11-6896	NO_2	Me	F	H	Me	H	0.845	0.741	0.104
Ro 06-7263	Cl	Cl	H	H	Me	H	1.690	1.574	0.117
oxazepam	Cl	H	H	H	OH	H	1.255	1.161	0.094
temazepam	Cl	Me	H	H	OH	H	1.204	1.427	−0.222
lorazepam	Cl	H	Cl	H	OH	H	0.544	0.315	0.229
Ro 20-7078	Cl	H	F	H	Cl	H	0.724	0.593	0.131
Ro 07-6198	H	H	F	F	H	Cl	1.447	1.478	−0.031
Ro 20-8895	H	H	F	H	H	Me	1.279	1.114	0.164
Ro 22-6762	Cl	Me	H	H	H	Cl	1.602	1.518	0.084
Ro 20-8065	Cl	H	F	H	H	Cl	0.556	0.546	0.010
Ro 20-8552	Me	H	F	H	H	Cl	1.146	1.147	−0.001
no. groups[b]	51	27	38	8	8	5			

[a] The known values were those from the literature; the predicted (pred) values were those found by a trained ANN. The difference (diff) is the known − predicted. [b] No. groups indicates the number of compounds with a non-H substituent group present.

as the number of input parameters were pruned from 48 down to 6. It was found that as each group of six input parameters (one representing each position around the molecule) with lowest weights was successively pruned, the training correlation and standard error stayed constant but the cross-validation correlation improved until the number of inputs reached 24. As further input parameters were pruned, the training correlation and standard errors deteriorated slowly but the cross-validation correlation continued to improve until there were only 12 input parameters left, two for each position around the molecule. Further pruning to six input parameters caused a substantial loss of both training and cross-validation correlation and an increase in standard error, suggesting that at least two parameters for some of the positions around the molecule were required. The cross-validation correlation peak at 12 inputs suggested that the optimum number of inputs lies between 6 and 18.

In further studies, groups of nets were trained with combinations from 10 to 18 inputs where individual parameters were pruned one at a time. This procedure showed that eight of the input parameters including the random numbers representing positions 3 and 8 added no information. The remaining 10 input parameters, MR1, π7, $\mathscr{F}$7, $\mathscr{R}$1, $\mathscr{F}$2, MR2, MR6, σ_p8, σ_m3 and μ1, were

Table 2. Substituent Values[a]

substituent	μ	π	MR	$\mathcal{F}$	$\mathcal{R}$	σ_m	σ_p
Br	−1.57	0.86	8.88	0.44	−0.17	0.39	0.23
$C(CH_3)_3$	0.52	1.98	19.62	−0.07	−0.13	−0.10	−0.20
CF_3	−2.61	0.88	5.02	0.38	0.19	0.43	0.54
$CH{=}CH_2$	0.20	0.82	10.99	0.07	−0.08	0.05	−0.02
CH_2CF_3	−2.07*	1.34	9.64	0.34	0.09	0.40	0.50
CH_2CH_2OH	−0.60*	−0.31	12.10	0.01	−0.29	−0.05	−0.23
CH_2OCH_3	−1.01*	−0.78	12.07	0.01	0.02	0.02	0.03
CHO	−3.02	−0.65	6.88	0.31	0.13	0.35	0.42
Cl	−1.59	0.71	6.03	0.41	−0.15	0.37	0.23
CN	−4.08	−0.57	6.33	0.51	0.19	0.56	0.66
COMe	−2.90	−0.55	11.18	0.32	0.20	0.38	0.50
Et	0.39	1.02	10.30	−0.05	−0.10	−0.07	−0.15
F	−1.43	0.14	0.92	0.43	−0.34	0.34	0.06
H	0.00	0.00	1.03	0.00	0.00	0.00	0.00
I	−1.36	1.12	13.94	0.40	−0.19	0.35	0.18
Me	0.36	0.56	5.65	−0.04	−0.13	−0.07	−0.17
N_3	−1.56	0.46	10.20	0.30	−0.13	0.27	0.15
NH_2	1.53	−1.23	5.42	0.02	−0.68	−0.16	−0.66
NHCOMe	−3.65	−0.97	14.93	0.28	−0.26	0.21	0.00
NHOH	−0.14	−1.34	7.22	0.06	−0.40	−0.04	−0.34
NO_2	−4.13	−0.28	7.36	0.67	0.16	0.71	0.78
OH	−1.59	−0.67	2.85	0.29	−0.64	0.12	−0.37

[a] Substituent dipole values for three compounds marked with an asterisk were calculated using a multiple linear regression created for the purpose using $N = 55$ samples as follows. $\mu = 0.56\mu a - 2.63\mathcal{R} - 2.69\mathcal{F} + 0.17\pi - 0.06$, where $R = 0.929$, SE = 0.689, and μa = aliphatic dipole from ref 24.

the important parameters of the substituent groups which contributed to the receptor affinity. This set had a correlation on training (R_t) of 0.938 ± 0.003, standard error on training (SE_t) of 0.254 ± 0.006, and an optimum correlation on cross-validation (R_{cv}) of 0.896. This set of input parameters was termed the optimal set and was used for all other studies. The receptor affinities predicted by the network using $N - 1$ cross-validation for the optimal data set are shown in Table 1.

All parameters of the optimum input set were found to be independent by cross-correlation (see Table 4). The highest squared correlation between any two of the parameters was between MR1 and $\mathcal{R}1$ (0.41). However, removal of either caused a significant reduction in both training and cross-validation correlation suggesting that both had important contributions to make.

An examination of the contribution of the various substituent parameters to the ANN weights between the inputs and the output layer for the optimal input set showed that the parameters representing positions 1 (39%), 7 (30%), and 2′ (15%) around the molecules contributed 84% of the total weights (see Figure 2) with the parameters representing positions 6′ (8%), 3 (4%), and 8 (4%) making much smaller but still significant contributions. These results suggest that in this study with the current data set the substituent parameters at positions 1, 7, and 2′ would be the main parameters determining the receptor affinity; however, since positions 6′, 3, and 8 are represented by only a few examples in the current data set, it is possible that they could be shown to play more important roles if a more uniform set of samples was available.

Comparison of Prediction Abilities of Two-Layer ANNs and MLR. When two-layer ANNs and MLR were used to examine the same data sets by two different cross-validation methods, the ANNs gave marginally better results than the MLR (see Table 5). In the $N - 1$ cross-validation study, the ANNs gave slightly higher but not significant correlations on cross-validation (0.896 vs 0.886) and lower standard errors (0.321 vs 0.339) than the MLR method. In the $N - 19$ cross-validation study, 20 random sets of 38 compounds were used to train the ANNs or used in the MLR program and the 20 residual sets of 19 compounds were used as test series. The ANNs were found to give both a significantly higher ($p < 0.05$) cross-validation correlation than the MLR method (0.910 vs 0.865) and a significantly lower ($p < 0.05$) cross-validation standard error (0.308 vs 0.384).

Effects of the Hidden Layers. The effect of hidden layers on the ANNs with the optimum input set from the pruning studies is shown in Table 6. The results showed no significant improvement in training correlation (R_t) or standard error (SE_t). The cross-validation correlation (R_{cv}) and standard errors (SEcv) were both found to reach an optimum with a hidden layer of three neurons, but this peak was only marginally greater than the results of the two-layer ANNs, suggesting no improvement could be gained by using a hidden layer in this study.

Effect of Substituent Groups on Predicted Receptor Affinity. It is possible to use a trained ANN, where the connection weights are kept constant, to examine the effects of an individual parameter on the output of the ANN. By fixing $n - 1$ of the input parameters at some predetermined value and varying the nth parameter from the minimum to the maximum of its known range, a curve plotted from the ANN output indicates the effect of the nth parameter.[15,16] Consequently, when 9 of the optimum 10 parameters were kept constant at the minima of their respective ranges, it was possible to examine the effect of the 10th upon the predicted log IC_{50} (see Figure 3). Increases in four of the substituent parameters, $\mathcal{F}7$, $\pi7$, $\mu1$ and $\mathcal{F}2$, from their range minima to maxima were found to improve the predicted receptor affinity, while the remaining six, σ_m3, σ_p8, MR1, $\mathcal{R}1$, MR2, and MR6, were found to decrease it.

Curves showing the effects of increases in the magnitude of the substituent parameters at positions 7, 3, and 8 around the molecule on the predicted log IC_{50} are shown in Figure 3a. Increases in lipophilicity ($\pi7$) and polar effects ($\mathcal{F}7$) of substituent groups at position 7 were found to strongly favor an increase in the predicted receptor affinity (i.e., a lower log IC_{50}). An increase in the electronic parameters, at position 3 (σ_m3) and position 8 (σ_p8), was found to have a detrimental effect on predicted receptor affinity. In Figure 3b, the independent effects of the three important parameters at position 1 are shown. Group size reflected as molar refractivity (MR1) and resonance effects ($\mathcal{R}1$) both showed a detrimental effect on predicted receptor binding affinity, while increases in dipole ($\mu1$), enhanced predicted receptor affinity.

In Figure 3c, an increase in the polar effect ($\mathcal{F}2$) of groups at position 2′ on the phenyl ring was found to enhance predicted receptor affinity. However, this was reduced if bulky groups indicated by the molar refractivity were placed at positions 2′ or 6′ on the molecule.

While Figure 3 gave an indication as to how a particular substituent parameter behaved in isolation, it did not show how those parameters performed under more realistic conditions where several groups might be exerting positive and negative effects on the predicted receptor affinity. Since it has been shown that the

Table 3. Optimal Parameter Set Indicated by Pruning[a]

inputs	7	1	2′	6′	3	8	R_t	SD	SE_t	SD	R_{cv}	ϱ
48	μ	π	X	μ	π	σ_m	0.947	0.004	0.236	0.010	0.826	1.1
42	σ_p	σ_m	σ_p	$\mathcal{F}$	MR	μ	0.948	0.003	0.235	0.010	0.848	1.3
36	X	$\mathcal{F}$	μ	X	$\mathcal{R}$	$\mathcal{R}$	0.948	0.004	0.239	0.015	0.865	1.5
30	MR	σ_p	π	σ_p	μ	$\mathcal{F}$	0.947	0.003	0.237	0.012	0.880	1.8
24	$\mathcal{R}$	X	σ_m	σ_m	σ_p	MR	0.947	0.003	0.237	0.008	0.888	2.2
18	σ_m	μ	$\mathcal{R}$	$\mathcal{R}$	$\mathcal{F}$	π	0.943	0.003	0.245	0.008	0.909	2.9
12	$\mathcal{F}$	$\mathcal{R}$	**MR**	π	X	X	0.936	0.004	0.260	0.013	0.891	4.1
6	$\boldsymbol{\pi}$	**MR**	$\mathcal{F}$	**MR**	$\boldsymbol{\sigma_m}$	$\boldsymbol{\sigma_p}$	0.826	0.007	0.411	0.008	0.776	7.1
10[b]							0.938	0.003	0.254	0.006	0.896	4.8

[a] The table shows the results of seven cycles of pruning. In each cycle, six parameters were pruned from the initial data set of 48. The top row lists the parameters pruned and the results for the first cycle, the second row the second cylcle, etc. R_t = mean correlation or SE_t = standard error between known and ANN-predicted log IC_{50} during training. R_{cv} = correlation between known and ANN-predicted log IC_{50} based on $(N-1)$ cross-validation study. ϱ = ratio of input data to net connections. X = random data. SD = standard deviation. [b] Results of study with optimal 10 parameters (indicated by bold typeface). N = 57 ANNs for all studies.

Table 4. Squared Cross-Correlation Table of Optimal Parameter Set

	π7	MR1	MR2	MR6	$\mathcal{F}$7	$\mathcal{F}$2	$\mathcal{R}$1	σ_m3	σ_p8	μ1
π7	1									
MR1	0.00	1								
MR2	0.00	0.04	1							
MR6	0.04	0.00	0.15	1						
$\mathcal{F}$7	0.02	0.06	0.03	0.00	1					
$\mathcal{F}$2	0.00	0.01	0.24	0.04	0.00	1				
$\mathcal{R}$1	0.01	0.41	0.01	0.00	0.01	0.00	1			
σ_m3	0.03	0.01	0.00	0.00	0.00	0.00	0.01	1		
σ_p8	0.02	0.00	0.01	0.00	0.01	0.00	0.00	0.00	1	
μ1	0.05	0.02	0.01	0.01	0.00	0.02	0.08	0.01	0.00	1

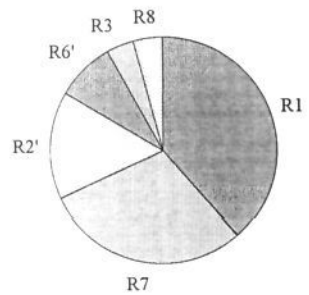

Figure 2. Contribution of groups to total connection weights.

Table 5. Comparison of Two-Layer ANN and Multilinear Regression Using Cross-Validation Studies[a]

	N	R_t	SD	SE_t	SD	R_{cv}	SD	SE_{cv}	SD
ANN-1	56−1	0.938	0.003	0.254	0.006	0.896		0.321	
MLR-1	56−1	0.936	0.003	0.254	0.005	0.886		0.339	
ANN-2	38−19	0.941	0.010	0.248	0.021	0.910*	0.027	0.308*	0.036
MLR-2	38−19	0.940	0.011	0.244	0.020	0.865*	0.076	0.384*	0.140

[a] Study 1—no. in training set = 56, test set = 1, means and standard deviations of 57 studies. Study 2—no. in training set = 38, test set = 19, means and standard deviations of 20 studies. R_t = correlation or SE_t = standard error between known and ANN- or MLR-predicted data during training. R_{cv} = correlation or SE_{cv} = standard error between known and ANN- or MLR-predicted data for test sets. All studies used optimal 10 input parameters. ANN = artificial neural networks. MLR = multilinear regressions. *$p < 0.05$.

receptor binding affinity is strongly related to the lipophilicity of the position 7 substituent,[5,8,9] the graphs were replotted in Figure 4 as a change in lipophilicity at position 7 (π7) versus predicted log IC_{50} with selected parameters set at values greater than their range minima.

In Figure 4a, when most parameters were kept at their range minima but the two parameters (MR2 and $\mathcal{F}$2) at position 2′ were simultaneously increased from 0% to 50% and then 100% of their range, three curves were generated which showed that, for any given

Table 6. Effect of PEs in the Hidden Layer on Two- and Three-Layer ANN Predictive Performance[a]

PEs	R_t	SD	SE_t	SD	R_{cv}	SE_{cv}
0	0.938	0.003	0.254	0.006	0.896	0.321
1	0.931	0.006	0.267	0.014	0.871	0.355
2	0.933	0.005	0.264	0.012	0.878	0.345
3	0.935	0.003	0.262	0.009	0.899	0.316
4	0.935	0.004	0.261	0.007	0.895	0.323
6	0.932	0.004	0.266	0.009	0.886	0.335
8	0.931	0.004	0.268	0.008	0.879	0.344

[a] PEs = no. of processing elements in hidden layer. R_t = mean correlation between known and predicted log IC_{50} at training. SE_t = mean standard error between known and predicted log IC_{50} at training. R_{cv} = correlation between known and predicted IC_{50} for $N-1$ cross-validated results. SD = standard deviation. N= 57 studies.

lipophilicity of a substituent group at position 7, an increase in the magnitude of the parameter pair at position 2′ strongly increased predicted receptor binding affinity. In this case the beneficial $\mathcal{F}$2 parameter was much stronger than the detrimental MR2 parameter. The addition of the $\mathcal{F}$7 parameter at its range maxima (see Figure 4a) together with the MR2 and $\mathcal{F}$2 parameters at their range maxima showed that this parameter strongly enhanced predicted receptor affinity.

In Figure 4b,c, the effects of the remaining parameters were examined when MR2, $\mathcal{F}$2, and $\mathcal{F}$7 were set at a control level of 50% of their range values. In Figure 4b, the effects of changing the substituent parameters at positions 3, 8, and 6′ to their range maxima are shown. All three substituent parameters were found to reduce the predicted receptor affinity, the largest effect being caused by the MR6 parameter, while the effects of the σ_m3 and σ_p8 parameters were approximately the same but about one-half that of the MR6. In Figure 4c, the effects of changing the substituent parameters at position 1 to their range maxima were examined. The greatest effect was found to be due to the molar refractivity (MR1) which caused the largest reduction in predicted receptor affinity seen and tended to act like a buffer, retarding improvement in receptor affinity with increasing lipophilicity. The resonance parameter ($\mathcal{R}$1) also had a detrimental effect on predicted receptor affinity but of much less magnitude than the effect caused by MR1. The dipole (μ1) at position 1 was found to improve the predicted receptor binding affinity, but its positive influence was much less than the negative influences of the MR1 and $\mathcal{R}$1 parameters.

Prediction of High-Affinity Compounds. The graphs in the preceding section showed that the main parameters affecting high predicted receptor binding

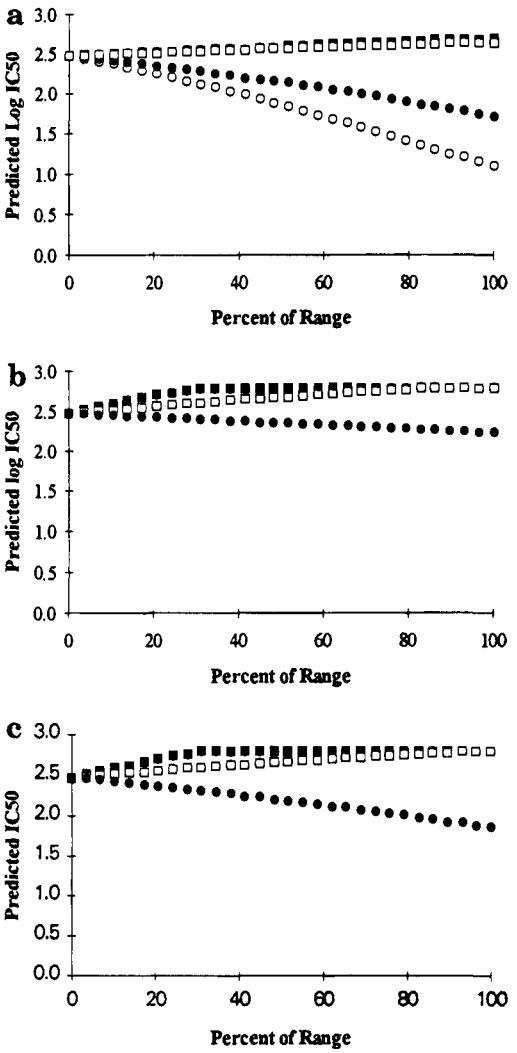

Figure 3. (a) Effect of substituent parameters on predicted log IC_{50} at positions 3, 8, and 7. (b) Effect of substituent parameters on predicted log IC_{50} at position 1. (c) Effect of substituent parameters on predicted log IC_{50} at positions 2′ and 6′. All parameters were set at their minima, and the selected parameters were increased in turn to their maxima. Panel a shows that increases in the magnitude of lipophilicity (π7, ○) and polar effects ($\mathcal{F}$7, ●) of substituent groups at position 7 from their range minima to maxima strongly favor a decrease in the log IC_{50} (increase in the receptor affinity). An increase in the σ_m3, (■) and σ_p8, (□) parameters reduces predicted receptor affinity. In panel b the molar refractivity (MR1, ■) and resonance effects ($\mathcal{R}$1, □) both showed a detrimental effect on the predicted receptor affinity, while increases in dipole (μ1, ●), enhanced receptor affinity. In panel c increases in the polar effect ($\mathcal{F}$2, ●) of groups at position 2′ on the phenyl ring was found to enhance predicted receptor affinity; however, this was reduced if bulky groups, indicated by the molar refractivity (MR2, □; MR6, ■), were placed at either position 2′ or 6′ on the molecule.

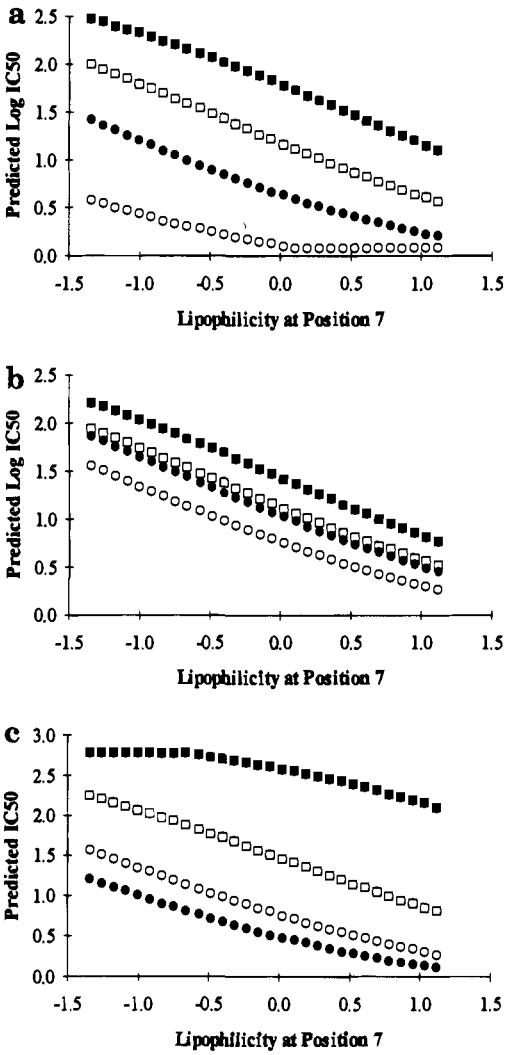

Figure 4. (a) Effect of lipophilicity of position 7 with substitution at position 2 on predicted log IC_{50}. (b) Effect of lipophilicity of position 7 with substitution at positions 6′, 3, and 8 on predicted log IC_{50}. (c) Effect of lipophilicity of position 7 with substitution at position 1 on predicted log IC_{50}. In panel a, when MR2 and $\mathcal{F}$2 were set at the minima, (■), 50%, (□), and maxima (●) of their ranges and other parameters at range minima, log IC_{50} decreases with lipophilicity; hence, increases in the magnitude of the parameters at position 2′ strongly increase predicted receptor binding affinity. When MR2, $\mathcal{F}$2, and $\mathcal{F}$7 were set at their range maxima (○) and other parameters at their range minima, then the maximum predicted receptor binding affinity of log IC_{50} = 0.08 was found. In panel b, as a control, (○), MR2, $\mathcal{F}$2, and $\mathcal{F}$7 were set at 50% and other substituents at their range minima. When the control was changed by setting MR6, (■), σ_m3, (□), or σ_p8, (●) to range maxima, all three substituent parameters were found to reduce the predicted receptor affinity, the largest effect being caused by the MR6 parameter. In panel c, when the control set (○) from panel b was changed by setting MR1 (■), $\mathcal{R}$1, (□), or μ1 (●) to their range maxima, both MR1 and $\mathcal{R}$1 were found to substantially reduce predicted receptor affinity while increases in μ1 were found to have a slightly beneficial effect on it.

affinity were at position 7 (π7, $\mathcal{F}$7), position 2′ ($\mathcal{F}$2, MR2), and position 1 (MR1, $\mathcal{R}$1, μ1) with much lesser effect from the parameters at positions 6′, 3, and 8. These results can be used to predict the optimal functional groups to be used for making compounds with high receptor affinity.

For functional groups at position 7, the sum of the scaled values of π7 and $\mathcal{F}$7 (see the Methods section) should be high since both contribute to high predicted receptor binding affinity. The optimal 10 functional groups for placing at position 7, in order from best to worst, would be CH_2CF_3 > I > Br > CF_3 > Cl > $C(CH_3)_3$ > NO_2 > F > N_3 > $CH{=}CH_2$. A comparison of three compounds, Ro 07-9957, flunitrazepam, and Ro 05-6822 from Table 1, which differ only in that they have an I, NO_2, and F at position 7 showed that they had binding affinities (log IC_{50}) of 0.462, 0.580, and 0.708 which was the predicted order.

At position 2′, the value of the $\mathcal{F}$2 less the value of the MR2 should be high since $\mathcal{F}$2 enhanced predicted

receptor binding affinity and MR2 reduced it. The five functional groups predicted to be optimal for placing at position 2′, in order from best to worst, would be NO_2, F, CN, Cl, and CF_3. Only F, Cl, and CF_3 were used in our test data set. A comparison of three compounds, Ro 05-4435, clonazepam, and Ro 05-3590 from Table 1, which differed only by placement of F, Cl, and CF_3 at position 2′ showed that they had binding affinities (log IC_{50}) of 0.176, 0.255, and 0.544 which was the predicted order.

At position 1, the value of the substituent's dipole reduced by the sum of the values of the other two parameters values, i.e., $\mu 1 - (MR1 + \mathcal{R}1)$, should be high since the dipole enhanced the predicted receptor binding affinity while the other two parameters had a detrimental effect. The optimal 10 functional groups for locating at position 1, in order from best to worst, would be OH > F > NH_2 > H > NHOH > Me > Cl > CF_3 > Br > Et. None of our test compounds had OH, F, or NH_2 at position 1; most had either H or Me. An examination of 10 pairs of compounds from Table 1 which differed only by having a H or Me at position 1 showed that in four cases the compound with the H at position 1 had the best binding affinity while in the other six cases the reverse was true, suggesting that H or Me at position 1 is roughly equivalent. The substituents at positions 6′, 3, and 8 should be H or have a value less than that of H.

Discussion

Neural Networks and QSAR. In previous studies using ANNs for QSAR, it was shown that there were advantages in using at least one hidden layer to give a nonlinear response and hence produce a good relationship between input and output data sets. However, those studies also restrained the number of inputs to a number based upon earlier statistical regression QSAR studies. This meant that the ANNs used were forced to create the best relationship that was possible between limited input parameter sets and the output sets. In those cases the best relationships were generated with three-layer networks.

Many types of physicochemical parameters have been used in QSAR studies on 1,4-BZs at the molecular, substituent group and atomic levels. The molecular parameters include dipole,[6,8,26] lipophilicity,[5] and HOMO and LUMO energies;[6,8,26] the substituent group parameters include lipophilicity,[8,9] $\mathcal{F}$polar,[8,9] dipole,[6,9] molar refractivity,[9] B1,[8] indicator variables,[5,6,9] and Free–Wilson variables.[4] At the atomic level, the parameters include CNDO- and AM1-calculated charges on various atoms,[6,7,26] lipophilicity, and molar refractivity.[7] Consequently the choice of input parameters was wide.

In the current study, by selectively pruning a large input set, we allowed the ANNs to indicate the input parameter set that produced the optimal results. We at first used three-layer networks but found that when the ANNs were given a wide choice of inputs the results were only marginally improved by varying the number of hidden units. Indeed the optimum number of hidden units was often one or two, suggesting that only a two-layer network was required. A similar result was found in an earlier study.[9] On re-examining the data using two-layer networks, the results were found to be more consistent than with the three-layer network suggesting that it was not necessary to use a higher order ANN for this particular study.

The use of a two-layer network also had some significant advantages over a three-layer network in that there was only one layer of connection weights (between the input and output layers), and consequently the choice of which input parameters to prune was simplified. As all inputs were scaled to an input range of 0.1–0.9, the magnitude of the connection weights after training gave a direct indication of the contribution of each input parameter to the total weight. Thus the input parameters with the smallest connection weights contributed least to the total and hence could be pruned. Pruning of connections has also been found by others to be useful for both reducing the complexity of multilayer networks and improving their ability to generalize.[21]

There has recently been some concern as to the most appropriate number of training samples to use for a given ANN size and complexity. It was found that ratio of input data set size to number of ANN connections (ϱ) had a bearing on network performance in QSAR studies on DHFR inhibitors.[15] The authors found that the optimum range for ϱ was between 1.8 and 2.2. At ϱ values lower than 1.8, the ANN memorized the data, while at values higher than 2.2, generalization was poor. In other studies[25] it was found that the probability of chance correlation between the input and output data sets for simple three-layer ANNs was unacceptably high when the ϱ value was less than 3. In the current studies the ϱ value was approximately 1 before pruning but rose to a value greater than 4 after pruning the input set to its optimum size of 10 (see Table 3).

Two methods were used in the current studies to ensure the ANNs performed as reliably as possible. The first was that the networks were optimized for the best generalization using the cross-validation technique rather than a high correlation or low standard error on training, and the second was the use of random data input sets. The results showed quite unequivocally that the optimal input data set on cross-validation was not the input set with the highest correlation or lowest standard error on training.

The addition of random data inputs might at first seem unusual; however, others have found that the addition of random noise to the inputs improved the ability of ANNs to tolerate noisy inputs and gave better generalization.[21] In the current study the ANNs were found to be able to discriminate between the real data and random data input sets without a great deal of difficulty. The random data sets were useful as they gave an internal indication as to the lower bounds of the connection weights of input data that would provide meaningful information. For example, the ANNs preferred the random numbers in position 8 on the molecule to all other inputs except the $\sigma_p 8$ parameter indicating that this was the only useful input (see Table 3). This input also had the lowest connection weight of all the inputs, suggesting that it was just above the lower bound of useful information. Subsequent studies (not shown) where other parameters representing position 8 were substituted for $\sigma_p 8$ or where a second parameter in addition to $\sigma_p 8$ was used gave no improvement in correlation or standard error on training or correlation on cross-validation.

Prediction of Receptor Affinity. While there were several attempts at relating physicochemical properties of 1,4-BZs with several different types of biological activity,[26,27] the first to attempt to correlate receptor affinity to physicochemical properties was Borea et al. (1983) who studied 1,4-BZs with substitutions in the 1, 3, 7, and 2′ positions. They found a highly significant parabolic relationship between the molecular lipophilicity of the BZs with receptor affinity measured as -log K_i. By using an indicator variable to account for the difference between compounds substituted and unsubstituted in the 2′ position, they were able to get an excellent correlation (0.973) between predicted and observed data sets for 22 compounds.

Borea (1983) used a Free−Wilson model which calculated electronic variables to examine the substituent group contributions to receptor affinity of 39 1,4-BZs substituted in the 1−4, 7, and 2′ positions. A high correlation (0.968) was found between the predicted and observed receptor affinity values, but the electronic variables were difficult to relate to molecular properties.

Ghose et al. (1990) used a three-dimensional receptor cavity model which was developed using 28 compounds from 14 different groups of BZ receptor binding ligands. Their data fitted the model with very good correlation (r = 0.980) between the predicted and observed data sets, allowing them to predict five members of a highly heterogeneous test set with good accuracy.

Greco et al. (1992) used conventional multiple regression with cross-validation and comparative molecular field analysis on a homogeneous set of 30 1,4-BZs substituted at the 7, 1, 2′, 3, and 6′ positions and a heterogeneous set of 48 ligands. They found good correlations on training (0.932) and cross-validation (0.906) for the homogeneous training set which related the receptor affinity and the lipophilicity at position 7, the LUMO energy, and an indicator variable indicating the difference between compounds substituted and unsubstituted at position 2′. They also found reasonable correlation on training (0.867) and cross-validation (0.834) between the predicted and observed data sets for the heterogeneous set.

In the current studies using a two-layer ANN model and 10 input parameters, the ANNs gave a correlation between the predicted and known data sets of 0.941 on training (R_t) and 0.910 on cross-validation (R_{cv}) when the data set was divided into 38 samples for training and 19 for a test series. These results are very similar to those of the earlier studies especially those of Greco et al. (1992).

The finding that the two-layer ANNs had significantly better correlations ($p < 0.05$) on cross-validation than the MLR and the standard deviations of the R_{cv} and SE_{cv} were much lower in the ANN study than in the equivalent MLR study suggests that, while the two methods may be considered theoretically equivalent, the ANNs through the use of the nonlinear sigmoidal transfer functions had greater internal flexibility to accommodate the variability in the input data than the MLR method.

The use of 10 input parameters may seem large by comparison to Borea (1983) who used two or Greco et al. (1992) who used three but is less than the number used by Ghose et al. (1990) who used 17. It should be noted, however, that each of the 10 input parameters held information about the physicochemical interactions at six positions around the molecule, which will be considered in the following section.

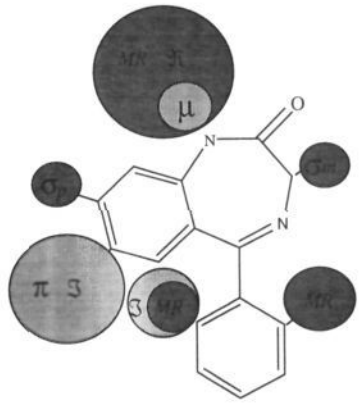

Figure 5. Physicochemical regions indicated by ANN as a receptor analog. Dark and light areas indicate negative and positive effects on receptor binding affinity. An area of circle indicates the relative size of the effects based upon ANN connection weights.

The ANN as a Receptor Site Mapper. Since each of the six positions on the 1,4-BZ molecules examined are not just numbers but actual optimal positions in three-dimensional space located within a receptor site at the time of binding, assuming a similar binding mode, then an ANN trained and optimized using physicochemical parameters can be used to derive the internal physicochemical structure of the receptor site. The physicochemical parameters selected by the ANN (provided that adequate choice was given and correlation between inputs and receptor affinity is high) should then represent the main effects occurring between the ligand and the receptor, and the weights allocated to the parameters should give an indication as to the relative importance of the parameters at those locations. If this is true, then the ANN can be used as a pharmacophore model and consequently used to predict some of the properties of the receptor binding site (see Figure 5).

Position 7. Position 7 was the most important location on the molecules for increasing the receptor affinity. Increases in the lipophilicity and electronic charge as indicated by the $\mathcal{F}$polar value were found to be directly related to increases in receptor affinity. This position also accounted for 30% of the total connection weight.

These results compare well with those of Greco et al. (1992) who found, using conventional multiple regression analysis, that lipophilicity of groups at position 7 was one of the main parameters in determination of receptor affinity. Borea (1983) suggested that lipophilicity at position 7 might be important and later suggested that the lipophilicity of the whole molecule was the prime determinant of receptor affinity.[5]

Loew et al. (1984) found that substituents in position 7 create a large negative potential with well-defined directionality indicating a direct interaction with a cationic receptor subsite. They also found an inverse relationship between dipole at position 7 and receptor affinity. We have found[9] that dipole at position 7 could be used in lieu of $\mathcal{F}$.

The results support the hypothesis of Greco et al. (1992) that groups substituted at position 7 interact

with a complementary hydrophobic pocket in the receptor but are in disagreement with those of Ghose et al. (1990) who found, using a three-dimensional receptor cavity model, that position 7 favored hydrophilic groups with a partial positive charge.

Position 2′. In the current study the ANNs found that substitution at position 2′ was the second most important position on the molecules for influencing the receptor affinity. Increases in the $\mathscr{F}$polar value were found to be directly related to increases in receptor affinity, although this effect was reduced if the groups were bulky as indicated by MR. While these results were quite significant, care needs to be taken in their interpretation as they were based on a limited variety of substituent types (F, Cl, CF_3, and H). Substitution at this position accounted for 15% of the total connection weight.

All authors are in agreement that the presence of an electrophilic group at position 2′ leads to a strong increase in receptor affinity. However, the cause is more equivocal. Boria (1983) felt that the increase in binding affinity was related to increase in electron-withdrawing properties and steric hindrance which is consistent with an increase in $\mathscr{F}$and a decrease in MR found in the current studies. Loew et al. (1984) suggested that the groups at position 2′ would increase the negative charge on the N4 nitrogen, supporting the hypothesis of a cationic receptor interaction at that site. Greco et al. (1992) agreed that substituent groups at position 2′ caused a large increase in receptor affinity but were undecided about the cause. Ghose et al. (1990) found that their receptor cavity model favored hydrophillic and dispersive groups, i.e., increases in MR, but not electrostatic interactions which disagree with the current model. Thus there is reasonable agreement that molecular bulk (measured by MR) is important at position 2′ but less agreement as to other physicochemical factors.

Position 1. Three parameters, MR, $\mathscr{R}$, and μ, were needed to describe the effects of substitution at position 1. The most important was MR, which caused a large reduction in receptor affinity as MR increased. An increase in $\mathscr{R}$ also caused a reduction in receptor affinity, but an increase in μ was beneficial, causing a small but significant increase in receptor affinity. Substitution at this position accounted for 39% of the total connection weight.

These results are consistent with those of others in that all appear to agree that molecular bulk as indicated by molar reactivity is important, but there is less agreement on other factors. Ghose et al. (1990) found that MR and π were both important descriptors of the binding at position 1. Their model suggested that small hydrophillic groups at this position improved receptor affinity. Haefely et al. (1985) noted that large groups at position 1 caused a reduction in receptor affinity; however, they stated that this was not mirrored in reduced biological activity owing to the fact that groups substituted at this position were readily removed *in vivo* creating the more active metabolites substituted with an H at position 1. Boria (1983) also noted that while an increase in size of the 1 substituent did not appear to affect the biological activity of the compounds it did appear to cause increasing stearic hindrance which caused a reduced binding affinity.

Positions 3, 6′, and 8. Increases in σ_m at position 3, MR at 6′, and σ_p at 8 were all found to be consistant with a reduction in predicted receptor affinity. The results suggest that electrostatic influences are important at positions 3 and 8 and molecular bulk at position 6′. However, they should be treated with caution as only a small number of compounds (eight or less) with a limited variety of substituents (four or less) were used in the study and substituents at these positions only contributed a small (<6%) amount to the total connection weight.

Borea (1983) and Haefely et al. (1985) pointed out that substitution of position 3 was difficult to interpret owing to the formation of two enantiomeric forms with dramatically different three-dimensional structures. Ghose et al. (1990) found that hydrophobic, dispersive, and positively charged atoms were favored in this position. Few studies have commented on the effects of substitution at positions 6′ and 8. Haefely et al. (1985) noted that substitution at position 8 produced compounds with less affinity than at position 7 which agrees with the current results.

Conclusions

In the current study we have shown that by using selective pruning and cross-validation techniques, ANNs can be used to force a selection of an optimal data input parameter set from a much larger set that included random data. Two-layered back-propagation networks can be trained to predict the receptor binding affinity of 1,4-BZ ligands with a good cross-validated correlation (0.910) between predicted and observed data sets and less variability than an equivalent MLR method. ANNs can be used to select an optimum set of physicochemical parameters that describe important aspects of the ligand−receptor interaction and might be considered a pharmacophore representing the internal physicochemical structure of the receptor site. The results suggest that ANNs might be considered to be useful for a wider role in QSAR studies.

Acknowledgment. This work was carried out while one of the authors (D.J.M.) was a recipient of a University of Sydney Consolidated Medical Research Fund Scholarship.

References

(1) Tallman, J. F.; Paul, S. M.; Skolnick, P.; Gallager, D. W. Receptors for the age of anxiety: pharmacology of the benzodiazepines. *Science* **1980**, *207*, 274−281.
(2) Pike, V. W.; Halldin, C.; Crouzel, C.; Barré, L.; Nutt, D. J.; Osman, S.; Shah, F.; Turton, D. R.; Waters, S. L. Radioligands for PET studies of Central Benzodiazepine receptors and PK (Peripheral Benzodiazepine) Binding Sites - Current Status. *Nucl. Med. Biol.* **1993**, *20*, 503−525.
(3) Haefely, W.; Kyburz, E.; Gerecke, M.; Möhler, H. Recent advances in the molecular pharmacology of benzodiazepine receptors and in the structure-activity relationships of their agonists and antagonists. *Adv. Drug Res.* **1985**, *14*, 165−322.
(4) Borea, P. A. De Novo analysis of receptor binding affinity data of benzodiazepines. *Arzneim.-Forsch.* **1983**, *33* (II), 1086−1088.
(5) Borea, P. A.; Bonora, A. Brain receptor binding and lipophilic character of benzodiazepines. *Biochem. Pharmacol.* **1983**, *32*, 603−607.
(6) Loew, G. H.; Nienow, J. R.; Poulsen, M. Theoretical structure-activity studies of benzodiazepine analogues. Requirements for receptor affinity and activity. *Mol. Pharmacol.* **1984**, *26*, 19−34.
(7) Ghose, A. K.; Crippen, G. M. Modeling the benzodiazepine receptor binding site by the general three-dimensional structure-directed quantitative structure-activity relationship method REMOTEDISC. *Mol. Pharmacol.* **1990**, *37*, 725−734.

(8) Greco, G.; Novellino, E.; Silipo, C.; Vittoria, A. Study of benzodiazepines receptor sites using a combined QSAR-CoMFA approach. *Quant. Struct.-Act. Relat.* **1992**, *11*, 461–477.
(9) Maddalena, D. J.; Johnston, G. A. R. Application of neural networks to quantitative structure-activity relationships of benzodiazepine / $GABA_A$ receptor binding compounds. *Proc. 4th Aust. Conf. Neural Networks*, Melbourne, Australia; Sydney University Electrical Engineering Press: Sydney, Australia, 1993; pp 228–231.
(10) Holley, L. H.; Karplus, M. Neural networks for protein structure prediction. *Methods Enzymol.* **1991**, *202*, 204–224.
(11) 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.
(12) Maddalena, D. J. Applications of neural networks in chemistry. *Chem. Aust.* **1993**, *60*, 218–221.
(13) Salt, D. W.; Yildiz, N.; Livingstone, D. J.; Tinsley, C. J. The use of artificial neural networks in QSAR. *Pestic. Sci.* **1992**, *36*, 161–170.
(14) Aoyama, T.; Suzuki, Y.; Ichikawa, H. Neural networks applied to quantitative structure-activity relationship analysis. *J. Med. Chem.* **1990**, *33*, 2583–2590.
(15) Andrea, T. A.; Kalayeh, H. Applications of neural networks in quantitative structure-activity relationships of dihydrofolate reductase inhibitors. *J. Med. Chem.* **1991**, *34*, 2824–2836.
(16) So, S.-S.; Richards, W. G. Application of neural networks: Quantitative structure-activity relationships of derivatives of 2,4-diamino-5-(substituted-benzyl) pyrimidines as DHFR inhibitors. *J. Med. Chem.* **1992**, *35*, 3201–3207.
(17) Wiese, M. Application of neural networks in the analysis of percent effect biological data. *Quant. Struct.-Act. Relat.* **1991**, *10*, 369–371.
(18) Hansch, C.; Leo, A. *Substituent constants for correlation analysis in chemistry and biology*; John Wiley & Sons: Brisbane, 1979.
(19) Bioactivnet is an artificial neural network generation program based upon the algorithms from Wards Systems Group NeuroWindows Dynamic Link Library. The program is available from Answers From Computers, Sydney, Australia.
(20) Rumelhart, D. E.; McClelland, J. L. *Parallel distributed processing: Explorations in the microstructure of cognition;* MIT Press: Cambridge, 1986; Vol. I, II.
(21) Sietsma, J.; Dow, R. J. F. Creating artificial neural networks that generalise. *Neural Networks* **1991**, *4*, 67–79.
(22) Ajay. A unifled framework for using neural networks to build QSARs. *J. Med. Chem.* **1993**, *36*, 3565–3571.
(23) Lien, E. J.; Guo, Z.-R.; Li, R.-L.; Su, C.-T. Use of dipole moment as a parameter in drug-receptor interaction and quantitative structure-activity relationship studies. *J. Pharm. Sci.* **1982**, *71*, 641–655.
(24) Li, W.-Y.; Guo, Z.-R.; Lien, E. J. Examination of the interrelationship between aliphatic group dipole moment and polar substituent constants. *J. Pharm. Sci.* **1984**, *73*, 553–558.
(25) Livingstone, D. J.; Manallack, D. T. Statistics using neural networks: Chance effects. *J. Med. Chem.* **1993**, *36*, 1295–1297.
(26) Blair, T.; Webb, G. A. Electronic factors in the structure-activity relationships of some 1,4-benzodiazepin-2-ones. *J. Med. Chem.* **1977**, *20*, 1206–1210.
(27) Sternbach, L. H.; Randall, L. O.; Bazinger, R.; Lehr, H. Structure-activity relationships in the 1,4-benzodiazepine series. In *Drugs Affecting the Central Nervous System*; Burger, A., Ed.; Edward Arnold: London, 1968; pp 237–264.

JM940090G