Brief Articles

Robust QSAR Models Using Bayesian Regularized Neural Networks

Frank R. Burden*

Chemistry Department, Monash University, Clayton, Victoria 3168, Australia

David A. Winkler

CSIRO Division of Molecular Science, Private Bag 10, Clayton South MDC, Clayton, Victoria 3169, Australia

Received December 11, 1998

We describe the use of Bayesian regularized artificial neural networks (BRANNs) in the development of QSAR models. These networks have the potential to solve a number of problems which arise in QSAR modeling such as: choice of model; robustness of model; choice of validation set; size of validation effort; and optimization of network architecture. The application of the methods to QSAR of compounds active at the benzodiazepine and muscarinic receptors is illustrated.

Introduction

Quantitative structure−activity relationship (QSAR) methods were developed by Hansch and Fujita,[1] and they have been successfully applied to many drug and agrochemical design and optimization problems. As well as speed and simplicity, QSAR has advantages of being capable of accounting for some transport and metabolic processes which occur once the compound is administered. Hence, the method is often applicable to the analysis of in vivo data. As useful as 'traditional' QSAR methods have been, they still exhibit a number of difficulties and shortcomings which relate to the molecular representations used and the methods by which SAR models are developed and validated.

Finding SARs is essentially a regression or pattern recognition process. Historically, linear regression methods such as MLR (multiple linear regression) and PLS (partial least squares) have been used to develop QSAR models. Regression is an "ill-posed" problem in statistics, which sometimes results in QSAR models exhibiting instability when trained with noisy data. In addition traditional regression techniques often require subjective decisions to be made on the part of the investigator as to the likely functional (e.g. nonlinear) relationships between structure and activity. It is important that QSAR methods be quick, give unambiguous models, not rely on any subjective decisions about the functional relationships between structure and activity, and be easy to validate. Recently, regression methods based on neural networks have been shown to overcome some of these problems as they can account for nonlinear SARs and can deal with linear dependencies which sometimes appear in real SAR problems. Neural network training can be regularized, a mathematical process which converts the regression into a well-behaved, "well-posed" problem. The mathematics of 'well-posedness' and regularization can be found in the papers by Hadamard and Tikhonov.[2] Feed-forward, back-propagation neural nets still present some problems, principal of which are overtraining, overfitting, network architecture optimization, and selection of the best QSAR model. Overtraining results from running the neural network training for too long and results in a loss of ability of the trained net to generalize. Overtraining can be avoided by used of a validation set. However, the effort to cross-validate QSAR models scales as $O(N^2P^2)^3$, where N is the number of data points and P is the number of input parameters. Where the training data set is large and diverse, as may occur in combinatorial discovery, this can result in prohibitively large validation times. Validation procedures also produce a family of similar QSAR models, and it is not clear which of these models is preferred or how they may be combined to give the 'best' model. It is also not obvious which neural net architecture (e.g. number of hidden layers; number of nodes per hidden layer; fully connected or not) gives the best QSAR model necessitating an additional architecture optimization step. Overfitting results from the use of too many adjustable parameters to fit the training data and is avoided by use of test sets of data, not used in the training and validation steps.

The primary purpose of this paper is to show how to produce a robust QSAR model using a special type of neural network, a Bayesian regularized artificial neural network (BRANN). These types of neural networks have found applicability in numerous other areas of data modeling but, with a very recent exception,[4] have not been used for QSAR analysis. BRANNs offer substantial advantages in QSAR analysis compared with other methods such as conventional back-propagation neural networks and regression techniques.

Methods

Molecular Indices. We employed a combination of three easily computed molecular indices in this work: the well-

* To receive all correspondence.

10.1021/jm980697n CCC: $18.00
Published on Web 07/24/1999

studied Randic[5] index (R); the valence modification to the Randic index by Kier and Hall (K);[6] an atomistic (A)[7] index developed by Burden. The R and K indices are produced from the path length and valence electron counts in the molecule. The atomistic index (A) counts the numbers of each type of atom present in the molecule.

The three types of indices, R, K, and A, are complementary, and we have shown in previous studies[8] that their combination yields better QSAR models than the individual indices. Principal component analysis (PCA) is used to reduce redundant information and minimize overfitting. When PCA (a linear transformation) is used prior to a nonlinear regression, the usual criterion of ignoring those components with small variance is inappropriate. The number of principal components that gives the lowest standard error of prediction is the proper measure, and the use of a test set allows selection of models with best predictivity.

BRANNs. The Bayesian method is summarized in the papers by Mackay[9] and Buntine and Weigend,[10] and only a brief summary will be provided here. Bayesian methods are optimal methods for solving learning problems. Any other method not approximating them should not perform as well on average.[10] They are very useful for comparison of data models (something orthodox statistics cannot do well) as they automatically and quantitatively embody "Occam's Razor". Complex models are automatically self-penalizing under Bayes' Rule.[9] Bayesian methods are complementary to neural networks as they overcome the tendency of an overflexible network to discover nonexistent, or overly complex, data models.

Unlike a standard back-propagation neural network training method where a single set of parameters (weights, biases, etc.) are used, the Bayesian approach to neural network modeling considers all possible values of network parameters weighted by the probability of each set of weights. Bayesian inference is used to determine the posterior probability distribution of weights and related properties from a prior probability distribution according to updates provided by the training set D using the BRANN model, H_i. Where orthodox statistics provide several models with several different criteria for deciding which model is best, Bayesian statistics only offer one answer to a well-posed problem.

$$P(\mathbf{w}|D,H_i) = \frac{P(D|\mathbf{w},H_i)P(\mathbf{w}|H_i)}{P(D|H_i)}$$

Bayesian methods can simultaneously optimize the regularization constants in neural nets, a process which is very laborious using cross-validation. There is no better method for reliably finding and identifying better models using only the training set.[9]

Characteristics of the Neural Network. We used the Bayesian regularized neural network package in the MATLAB[11] Neural Network Toolbox for the work described made here. This module is incorporated in a comprehensive chemometrics package written in MATLAB language by one of the authors (F. R. Burden). Our Bayesian neural networks were three-layer fully connected feed-forward networks with a variable number of neurodes in the hidden layer, each neurode in the middle and output layer using a sigmoidal transfer function. The basic method used in the network training is derived from the Levenberg−Marquardt algorithm,[12] and the MATLAB implementation of the algorithm uses an automatic adjustment of the learning rate and the Marquardt μ parameter. The Bayesian regularization takes place within the Levenberg−Marquardt algorithm and uses back-propagation to minimize the linear combination of squared errors and weights. The training is stopped if: the maximum number of epochs is reached; the performance has been minimized to a suitable small goal; the performance gradient falls below a suitably target; the Marquardt μ parameter exceeds a suitable maximum. Each of these targets and goals were set at the default values set by the MATLAB implementation. The training was carried out many times and the final model chosen with reference to the test set to assess robustness.

Small Benzodiazepine Data Set. This was a set of 57 1,4-benzodiazepin-2-ones used in a study by Maddelena and Johnston.[13] Benzodiazepines have been used therapeutically as anxiolytics, as tranquilizers, and as anticonvulsants in epilepsy. They act via the benzodiazepine site (BzR) on the γ-aminobutyric acid receptor ($GABA_A$) family and have been subject to extensive research, with over 20 QSAR studies having been carried out (e.g. ref 14). The IC_{50} values reported represent the concentration of compound causing 50% inhibition of binding of [^{3}H]diazepam at the BzR. The set was reduced from 57 to 55 compounds to remove those with a single atomistic entry and hence be untrainable. The activities of the compounds in Maddelena and Johnston's data set have been very carefully measured[15] with errors estimated at less than 10% at the 95% confidence level. A test set of 11 (20% of data set) was used.

Large Benzodiazepine Data Set. This is a set of 245 compounds that act on the BzR and was culled from the literature.[16−23]. They do not have a common substructure so that the nature of the molecular indices used in forming the model becomes more important. The data, which falls into several subsets, is likely to be less accurate than those in the smaller data set. The larger size of the set also means that the test set of 50 compounds will be more representative of the set as a whole. A test set of 20% of the compounds was used.

Muscarinic Data Set. This is a set of 162 compounds that act on the M_1 muscarinic receptor and was culled from the literature.[24,−29] Muscarinic compounds are used in the treatment of memory-related problems such as Alzheimer's disease. The compounds in this data set do not have a common substructure but do fall into small subsets with common structures. IC_{50} values were measured as the concentration necessary to displace 50% of [^{3}H]quinuclidinyl benzilate (QNB) from the M_1 muscarinic receptor. A test set of 20% of the compounds was used.

Procedure Used in Forming the Model. For each data set the following steps were taken. The data set was divided into a training set and a test set chosen by a K-means clustering algorithm clustering on $X|Y$ values. Clustering on $X|Y$ generally gives slightly poorer training statistics than Y clustering, but superior predictive statistics. Clustering on $X|Y$ data is our preferred method in that it clusters the compounds according to all of the given information in a manner akin to PLS. This may lead to different test sets for different groups of indices but is appropriate when searching for the best model of a data set.

The training set data was mean-centered, and this mean was subtracted from the test set data. Several training sessions were carried out with different neural net architectures using different numbers of principal components (PCs) derived from the X data. Since the modeling procedure using neural networks is nonlinear, the number of PCs used was determined by the standard error of prediction of the test set rather than by the minimum variance described by the PCs. The number of effective parameters N_{Par} can be calculated by the BRANN. With optimal number of PCs and architecture, the BRANN was trained independently 30 times to eliminate spurious effects caused by the random set of initial weights, and the one that gave the lowest standard error of prediction, SEP, is the one reported.

We also compared the models produced by BRANNs with those generated by PLS. We used the same descriptors as BRANNs with the PLS Toolbox developed by Wise.[30]

Network Architecture. To test whether the QSAR models were essentially independent of network architecture beyond a minimum, we developed QSAR models for the small Bz data set using the Bayesian neural net in which the number of nodes comprising the hidden layers was varied between 1 and 6. Bayesian neural nets used to generate QSAR models for the two large data sets employed one hidden layer with four neurodes.

Table 1. 55-Benzodiazepine Data Set with a Common Substructure[a]

method (descriptors)	N_I	no. of PCs or LVs	training *SEE*, R^2	validation *SEV*, R^2_{cv}	test *SEP*, Q^2	N_{Par}	ρ_{eff}
BRANN (*R,K,A*)	21	12	0.076, 0.91	na[b]	0.14, 0.82	25	1.8
PLS (*R,K,A*)	21	5	0.146, 0.69	na	0.183, 0.50	21	na
previous paper (atomistic)	11	11	na	0.12, 0.91	na	na	0.70[c]
Maddalena & Johnston (substituent)	10	10	na	0.12, 0.88	na	na	na
So & Karplus (substituent)	10	10	na	na, 0.94	na	na	na

[a] Clustering on *X|Y* data and unshuffled. Data scaled 0 to 1. [b] na, not available/not applicable. [c] Four hidden modes.

Table 2. Variation of the Effective Number of Parameters and Model Quality with the Number of Hidden Neurodes for the Muscarinic Data Set and the *R,K,A* Indices[a]

hidden nodes	*SEE*	R^2	*SEP*	Q^2	N_{Weight}	N_{Par}	ρ_{eff}
1	0.14	0.38	0.16	0.40	17	14	9.3
2	0.12	0.57	0.13	0.57	33	25	5.2
3	0.12	0.63	0.12	0.64	49	32	4.1
4	0.11	0.65	0.13	0.59	65	36	3.6
5	0.11	0.67	0.13	0.58	81	37	3.5
6	0.11	0.66	0.14	0.54	97	38	3.4

[a] Using four hidden nodes. 162 in training set, 30 in test set. Data scaled 0 to 1.

Results and Discussion

Robustness of the Models. All 30 training sessions gave results that were very similar to the one reported in Tables 1−3. This contrasts with the common experience with neural networks used in QSAR studies where the use of initial randomized weights often leads to different models, with different weights, though often leading to similar *SEP*s. The use of BRANNs overcomes this shortcoming where models giving similar *SEP*s have similar weights.

Independence of Network Architecture. QSAR models with differing numbers of nodes within the hidden layer are illustrated in Table 2 for the *R,K,A* indices and the muscarinic data set. It is clear that the training and test set statistics converge to a constant set for neural networks containing a hidden layer with more than four nodes. The number of effective parameters also converges when the number of hidden layer nodes is increased beyond a minimum value as Table 2 shows. The number of effective parameters increases from 14 when there is 1 node and settles at near 37 when the number of nodes is 4 or greater. This allows the use of four nodes with some confidence knowing that the BRANN will have produced a set of weights with good generalization properties.

A 'rule of thumb' has been reported by those employing neural networks for QSAR to allow network architectures to be defined approximately. It has been accepted that the ratio, ρ, of the number of input samples to the number of weights[31] should be greater than 2.0, though cross-validation allows for the use of smaller values.

Performance of Test Set. Tables 1−3 show that in most cases the test set statistics of the standard error of prediction and corresponding squared correlation coefficient (*SEP* and Q^2) were inferior to the corresponding values of the training set (*SEE* and R^2). There is one case where the Q^2 is slightly larger than R^2, which is possible with some choices of test set.

Table 1 compares the QSAR models for the small Bz data derived using BRANNS with a PLS analysis using the same *R,K,A* indices and results from previous QSAR studies of this data set. The BRANN models are clearly superior to the PLS models, suggesting that significant nonlinearity exists in the SAR model which PLS is not addressing. Comparison of BRANN models with other QSAR models for this data set is problematic. Many QSAR studies rely on leave-one-out (LOO) or similar cross-validation methods for estimating the predictivity of the model. We have adopted a more stringent measure of predictivity in using a test test which is never involved in training the model, unlike cross-validation. It is therefore not correct to compare statistics from LOO validation sets with those from test sets. This is why we distinguish between training (R^2, *SEE*), validation (R^2_{cv}, *SEV*), and test (Q^2, *SEP*) statistics in Table 1. The results show that the care taken in the present method leads to a very good model with a *SEP* value of 0.14 and a Q^2 of 082. Previous studies by Maddalena and Johnston[13] used physicochemical parameters rather than the structural indices used here, further confounding comparisons of models. Maddalena and Johnston's method is not sufficiently general for the analysis of large SAR data sets or screening of large data sets where there is no common substructure. Similar arguments apply to So and Karplus' neural network analysis of this data set.[32]

Table 3 compares the QSAR models for the large Bz data set and the muscarinic data set derived using BRANNs with a PLS analysis using the same *R,K,A* indices. The overall statistics for the larger Bz data set model are not as good as those for the small benzodiazepine data set. As this smaller set has a common

Table 3. Benzodiazepine and Muscarinic Data Sets without a Common Substructure Using the *R,K,A* Indices[a]

data set	*N*	N_I	method	no. of PCs or LVs	*SEE*	R^2	*SEP*	Q^2	N_{Par}	ρ_{eff}
benzodiazepine	245	25	BRANNs	14	0.14	0.63	0.14	0.69	50	4.1
	245	25	PLS	13	0.20	0.38	0.21	0.28	25	na[b]
muscarinic	162	22	BRANNs	14	0.10	0.69	0.12	0.63	41	3.2
	162	22	PLS	17	0.14	0.44	0.17	0.21	22	na

[a] Data scaled 0 to 1. [b] na, not available/not applicable.

structural motif, it is likely the quality of the models is a reflection of the suitability of the molecular indices to cope with molecular diversity and the quality of the measured biological response data (the large data sets were an amalgam of values measured by several laboratories at different times). The results for both large data sets show that the *R,K,A* indices give a good predictive Q^2 and *SEP* compared with the PLS analysis of these data also shown for comparison, again suggesting nonlinearity in the SAR model. We suggest that the *R,K,A* set of indices, which are simple to calculate, are likely to produce a useful model for the screening of very large virtual data sets.

Conclusions

The advantages of Bayesian methods are that they produce models that are robust, are well-matched to the data, and make optimal predictions. Multiple training of a given data set/index/architecture combination results in models which are very similar, suggesting that the method is robust. No test or validation sets are strictly necessary so that, in principle, all available training data can be devoted to the model and the potentially lengthy validation process discussed above is avoided.[9] The Bayesian objective function is not noisy, in contrast to the cross-validation measure. Although there is no need for a test set, since the application of the Bayesian statistics provides a network that has maximum generalization,[9] it is still considered prudent to use a test set. The Bayesian neural net has the potential to give models which are relatively independent of neural network architecture, above a minimum, at least for the cases reported here, and the Bayesian regularization method estimates the number of effective parameters. The number of effective parameters used in the model is less than the number of weights, as some weights do not contribute to the models. This minimizes the likelihood of overfitting. The concerns about overfitting and overtraining are also removed by this method so that the production of a definitive and reproducible model is attained.

The results indicate that BRANNs possess several properties useful for the analysis of structure−activity data. Our studies on the application of BRANN to the development of QSAR models suggest that the method has the potential to become a universal robust method applicable to a wide range of problems, and we feel that the method merits consideration by others developing QSAR models in the drug and agrochemical research areas.

We are now investigating using the automatic relevance detection (ARD) for input variables method to eliminate the need for PCA variable reduction prior to training the BRANN. This technique, reviewed by Mackay,[9] allows all input parameters to be used in the neural net, with Bayesian inference eliminating those which contain no or redundant information. Successful application of ARD will further simplify, and increase the robustness of, QSAR models developed by BRANN.

Our research focus is also on developing information-rich, computationally cheap molecular descriptors and robust SAR mapping methods for use in QSAR, database mining, and simulation of combinatorial discovery. We are using BRANN methods to carry out virtual screening and identification of novel lead structures in chemical space in conjunction with novel molecular descriptors to be described in further publications.

References

(1) Hansch, C.; Fujita, T. $\rho-\sigma-\pi$ Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.* **1964**, *86*, 1616.

(2) (a) Hadamard, J. Sur les problemes aux derivees parielies et leur signification physique. *Bull. Univ. Princeton* **1902**, *49*−*52*. (b) Tikhonov, A.; Arsenin, V. *Solution of Ill-posed Problems*; Winston: Washington, DC, 1977.

(3) Goutte, C. Statistical Learning and Regularization for Regression. Ph.D. Thesis, University of Paris, 1997.

(4) Ajay; Walters, W. P.; Murcko, M. Can we Learn to Distinguish between "Drug-like" and "Nondrug-like" Molecules? *J. Med. Chem.* **1998**, *41*, 3314−3324.

(5) Randic, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.

(6) Kier, L. B.; Hall, L. H. The Molecular Connectivity Chi Indexes and kappa Shape Indexes in Structure−Property Modelling. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1995; Vol. 2, pp 367−422 and references therein.

(7) Burden, F. R. Using Artificial Neural Networks to Predict Biological Activity from Simple Molecular Strctural considerations. *Quant. Struct.-Act. Relat.* **1996**, *15*, 7−11.

(8) Winkler, D. A.; Burden, F. R.; Watkins, A. J. R. Atomistic Topological Indices Applied to Benzodiazepines using Various Regression Methods. *Quant. Struct.-Act. Relat.* **1998**, *17*, 14−19.

(9) MacKay, D. J. C. A Practical Bayesian Framework for Backprop Networks. *Neural Comput.* **1992**, *4*, 415−447. (b) Mackay, D. J. C. Probable Networks and Plausible Predictions − a Review of Practical Bayesian Methods for Supervised Neural Networks, Comput. *Neural Sys.* **1995**, *6*, 469−505. (c) Mackay, D. J. C. Bayesian Interpolation. *Neural Comput.* **1992**, *4*, 415−447.

(10) Buntine, W. L.; Weigend, A. S. Bayesian Back-Propagation. *Complex Sys.* **1991**, *5*, 603−643.

(11) *MATLAB*; The MathWorks, Inc.: Natick.

(12) Hagen, M. T.; Menhaj, M. Training Feedforward Networks With The Marquardt Algorithm. *IEEE Trans. Neural Networks* **1994**, *5*, 989−993.

(13) Maddalena, D.; Johnston, G. A. R. Prediction of Receptor Properties and Binding Affinities of Ligands to Benzodiazepine/$GABA_A$ Receptors Using Neural Networks. *J. Med. Chem.* **1995**, *38*, 715−724.

(14) (a) Blair, T.; Webb, G. A. Electronic Factors in the Structure−Activity Relationship of Some 1,4-Benzodiazepin-2-ones. *J. Med. Chem.* **1977**, *20*, 1206−1210. (b) Greco, G.; Novellino, E.; Silipo, C.; Vittoria, A. Study of Benzodiazepines Receptor Sites Using Combined QSAR−CoMFA Approach. *Quant. Struct.-Act. Relat.* **1992**, *11*, 461−477. (c) Gupta, S. P.; Paleti, A. Quantitative Structure−Activity relationship Studies on Benzodiazepine Receptor Binding of Some Nonbenzodiazepine Series of Ligands. *Quant. Struct.-Act. Relat.* **1996**, *15*, 12−16. (d) Gupta, S. P. QSAR Studies on Drugs Acting at the Central Nervous System. *Chem. Rev.* **1989**, *89*, 1765−1800.

(15) Haefely, W.; Kyburz, E.; Gerecke, M.; Möhler, H. Recent Advances in the Molecular Pharmacology of Benzodiazepine Receptors and the Structure−Activity Relationships of their Agonists and Antagonists. *Adv. Drug Res.* **1985**, *14*, 165−322.

(16) Harrison, P. W.; Barlin, G. B.; Davies, L. P.; Ireland, S. J.; Matyus, P.; Wong, M. G. Syntheses, pharmacological evaluation and molecular modelling of substituted 6-alkoxyimidazo[1,2-b]-pyridazines as new ligands for the benzodiazepine receptor. *Eur. J. Med. Chem.* **1996**, *31*, 651−662.

(17) Davies, L. P.; Barlin, G. B.; Ireland, S. J.; Ngu, M. M. L. Substituted imidazo[1,2-β]pyridazines. New compounds with activity at central and peripheral benzodiazepine receptors. *Biochem. Pharmacol.* **1992**, *44*, 1555−1561.

(18) Barlin, G. B.; Davies, L. P.; Davis, R. A.; Harrison, P. W. Imidazo-[1,2-β]pyridazines. XVII* Synthesis and central nervous system activity of some 6-(alkylthio and chloro)-3-(methoxy, unsubstituted and benzamidomethyl)-2-aryl-imidazo[1,2-β]pyridazines containing methoxy, methylenedioxy and methyl substituents. *Aust. J. Chem.* **1994**, *47*, 2001−2012.

(19) Fryer, R. I.; Zhang, P.; Rios, R.; Gu, Z.-Q.; Basile, A. S.; Skolnick, P. Structure−activity relationship studies at the benzodiazepine receptor (BzR): A comparison of the substituent effects of pyrazoloquinoline analogues. *J. Med. Chem.* **1993**, *36*, 1669−1673.

(20) Wang, C.-G.; Langer, T.; Kamath, P. G.; Gu, Z.-Q.; Skolnick, P.; Fryer, R. I. Computer-aided molecular modelling, synthesis and biological evaluation of 8-(benzyloxy)-2-phenylpyrazolo[4,3-c]-quinoline as a novel benzodiazepine receptor agonist ligand. *J. Med. Chem.* **1995**, *38*, 950−957.

(21) Hollinshead, S. P.; Trudell, M. L.; Skolnick, P.; Cook, J. M. Structural requirements for agonist actions at the benzodiazepine receptor: studies with analogues of 6-(benzyloxy)-4-(methoxymethyl)-β-carboline-3-carboxylic acid ethyl ester. *J. Med. Chem.* **1990**, *33*, 1062−1069.
(22) Allen, M. S.; Hagen, T. J.; Trudell, M. L.; Codding, P. W.; Skolnick, P.; Cook, J. M. Synthesis of novel 3-substituted b-carbolines as benzodiazepine receptor ligands: Probing the benzodiazepine pharmacophore. *J. Med. Chem.* **1988**, *31*, 1854−1861.
(23) Yokoyama, N.; Ritter, B.; Neubert, A. D. 2-Arylpyrazolo[4,3-c]-quinolin-3-ones: Novel agonist, partial agonist and antagonist benzodiazepines. *J. Med. Chem.* **1982**, *25*, 337−339.
(24) Orlek, B. S.; Blaney, F. E.; Brown, F.; Clark, M. S. G.; Hadley, M. S.; Hatcher, J.; Riley, G. J.; Rosenberg, H. E.; Wadsworth, H. J.; Wyman, P. Comparison of Azabicyclic Esters and Oxadiazoles as Ligands for the Muscarinic Receptor. *J. Med. Chem.* **1991**, *34*, 2726−2735.
(25) Wadsworth, H. J.; Jenkins, S. M.; Orlek, B. S.; Cassidy, F.; Clark, M. S. G.; Brown, F.; Riley, G. J.; Graves, D.; Hawkins, J.; Naylor, C. Synthesis and Muscarinic Activity of Quinuclidin-3-yltriazole and -tetrazole Derivatives. *J. Med. Chem.*, **1992**, *35*, 1280−1290.
(26) Ward, J. S.; Merritt, L.; Klimkowski, V. J.; Lamb, M. L.; Mitch, C. H.; Bymaster, F. P.; Sawyer, B.; Shannon, H. E.; Olesen, P. H.; Honoré, T.; Sheardown, M. J.; Sauerberg, P. Novel functional M1 selective muscarinic agonists. 2. Synthesis and structure−activity relationships of 3-pyrazinyl-1,2,5,6-tetrahydro-1-methylpyridines. Construction of a molecular model for the M1 pharmacophore. *J. Med. Chem.* **1992**, *35*, 4011−4019.
(27) Sauerberg, P.; Olesen, P. H.; Nielsen, S.; Treppendahl, S.; Sheardown, M. J.; Honoré, T.; Mitch, C. H.; Ward, J. S.; Pike, A. J.; Bymaster, F. P.; Sawyer, B. D.; Shannon, H. E. Novel functional M1 selective muscarinic agonists. Synthesis and structure−activity relationships of 3-(1,2,5-thiadiazolyl)-1,2,5,6-tetrahydro-1-methylpyridines. *J. Med. Chem.* **1992**, *35*, 2274−2283.
(28) Jenkins, S. M.; Wadsworth, H. J.; Bromidge, S.; Orlek, B. S.; Wyman, P. A.; Wiley, G. J.; Hawkins, J. Substituent Variation in Azabicyclic Triazole and Tetrazole-Based Muscarinic Receptor Ligands. *J. Med. Chem.* **1991**, *35*, 2392−2406.
(29) Sauerberg, P.; Kindtlet, J. W.; Nielsen, L.; Sheardown, M. J.; Honoré, T. Muscarinic Cholinergic Agonists and Antagonists of the 3-(3-Alkyl-1,2,4-oxadiazol-5-yl-1,2,5,6-Tetrahydropyridine Type. Synthesis and Structure−Activity Relationships. *J. Med. Chem.* **1991**, *34*, 687−692.
(30) PLS Toolbox version 1.5, Eigenvalue Research Inc.
(31) Manallack, D. T.; Livingstone, D. J. Artificial Neural Networks: Application and Chance Effects for QSAR Data Analysis. *Med. Chem. Res.* **1992**, *2*, 181−190.
(32) So, S.-S.; Karplus, M. Genetic Neural Networks for Quantitative Structure−Activity Relationships: Improvements and Application of Benzodiazepine Affinity for Benzodiazepine/$GABA_A$ Receptors. *J. Med. Chem.* **1996**, *39*, 5246−5256.

JM980697N