Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2004) 499–505
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Modelling blood–brain barrier partitioning using Bayesian neural nets

David A. Winkler [a,b,*], Frank. R. Burden [b,c]

[a] CSIRO Molecular Science, Private Bag 10, Clayton South MDC 3169, Australia
[b] School of Chemistry, Monash University, 3800 Monash, Australia
[c] Scimetrics, 23 Harrow Street, Blackburn, Australia

Accepted 4 March 2004

Available online 27 April 2004

Abstract

We have employed three families of molecular molecular descriptors, together with Bayesian regularized neural nets, to model the partitioning of a diverse range of drugs and other small molecules across the blood–brain barrier (BBB). The relative efficacy of each descriptors class is compared, and the advantages of flexible, parsimonious, model free mapping methods, like Bayesian neural nets, illustrated. The relative importance of the molecular descriptors for the most predictive BBB model were determined by use of automatic relevance determination (ARD), and compared with the important descriptors from other literature models of BBB partitioning.

Keywords: QSAR; Blood–brain barrier; ARD; Neural net; ADMET; Bayesian methods

1. Introduction

With the accelerating cost of drug discovery and development, it has become increasingly important to understand the reasons why candidates fail late in the development pathway. In many cases, unfavorable absorption, distribution, metabolism or excretion (ADME) properties are the reason for these failures [1]. An understanding of ADME properties, the ability to create useful mathematical models of them, and the use of these models earlier in the drug design and development process has the potential to increase the success rate of drug candidates, and impact substantially on both the cost and time that drugs take to get to market [2].

Blood–brain barrier (BBB) permeability is an important ADME property. Since the surface area of the human BBB is very large, the BBB is considered to be the main region controlling the uptake of drugs into the brain and is the target for delivering drugs to the brain. The BBB is formed by the complex tight junctions between the endothelial cells of the brain capillaries and their low endocytic activity. This results in a capillary wall that behaves as a continuous lipid bilayer and prevents the passage of polar and lipid insoluble substances. Scherrmann [3], and Rubin and Staddon [4] have published accessible, brief reviews of the processes influencing the permeability of the BBB. An understanding of the relationship between BBB partition or permeability and molecular structure is essential for central nervous system (CNS) drugs to be rationally targeted to the brain. Conversely, it is also important to block entry to the brain for many other types of drugs, to reduce or eliminate CNS side effects or toxicities. Scherrman [3], Abbott and Romero [5], and Habgood et al. [6] have reviewed the relationships between physicochemical properties of drugs and their BB permeability, and have outlined strategies, including manipulation of physicochemical properties to mimic those of endogenous compounds, to modulate BBB permeability.

Blood–brain barrier partitioning is a difficult property to model because of the complexity of the underlying biological processes, and the paucity of data. Most of the published studies in which experimentally determined BBB partitions are reported involve a small number of compounds. Recent studies use datasets of the order of 100 compounds with good molecular diversity. Feher et al. [7] published a simple, linear three descriptor QSAR model in which the BBB was dependent on the lipophilicity, polar surface area, and hydrogen bond acceptor properties of the drugs. Labute [8] used a BB partitioning dataset to showcase the properties of his general molecular descriptors. Crivori et al. [9] developed

* Corresponding author. Tel.: +61-3-9545-2477; fax: +61-3-9545-2446.
E-mail addresses: dave.winkler@csiro.au (D.A. Winkler), frank.burden@scimetrics.com (F.R. Burden).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.03.010

a BBB partition model from a small set of 40 compounds using descriptors from 3D molecular fields. Rose et al. [10] compiled data from eleven experimental studies to obtain a comprehensive dataset of 106 compounds.

The development of robust QSAR models (i.e. those which are stable, reproducible, difficult to overtrain or overfit, and relatively insensitive to poor or missing data) is not a trivial undertaking. Devising robust models for BBB partitioning depends on the correct choice of molecular descriptors, and the application of structure-activity mapping methods that can extract information most efficiently from small datasets with substantial experimental uncertainty. Development of improved molecular descriptors is an area of intense activity in our group. We have also applied Bayesian regularized neural networks to QSAR modelling with excellent results [11–15]. We report the use of such Bayesian neural nets in the modelling of BBB partition data. This study, and others that we have recently carried out, suggest Bayesian neural networks can generate very good, robust QSAR models for a range of ADME properties. Given the complexity of the processes modulating partitioning of drugs across the BBB, ranging from passive diffusion, active transport via amino acid and organic acid transporters, and vesicular transport [5], a flexible, model free mapping system is highly desirable to account for a number of simultaneous, underlying processes. We have shown previously that Bayesian neural networks are able to model several parallel mechanisms simultaneously to produce a robust QSAR model [14].

2. Methods

2.1. Blood–brain barrier partition data

We used the 106 compound BBB partition dataset reported by Rose et al. [10], which was compiled from studies by Young [16], Abraham et al. [17], Salminem et al. [18], Clark [19], Luco [20], Yazdanian and Glynn [21], Grieg et al. [22], Lin et al. [23], Lombardo et al. [24], Van Belle et al. [25] and Calder and Ganellin [26]. These were in vivo measurements in rats of the partition coefficient of the compound between the brain and in blood. Although the experimental protocols for the eleven studies were not identical, they were sufficiently similar to justify them being combined into a single dataset. However, this difference in protocol, together with the difficulty of experimentally measuring BBB partitions, suggests that there is a significant degree of uncertainty in the experimental data. The log of this partition coefficient was used as the dependent variable.

$$\log \mathrm{BB} = \log\left(\frac{\mathrm{B_{brain}}}{\mathrm{C_{blood}}}\right)$$

The diversity of the dataset is very high, with structures ranging from low molecular weight simple gases such as nitrogen and nitrous oxide, to complex drug molecules such as codeine and imipramine. Table 1 lists the names and log BB values for the 106 compounds. Structures for these compounds are listed in the paper of Rose et al. [10]. The dataset is available from the authors on request.

2.2. Molecular descriptors

We investigated three families of descriptors:

1. feature based (number of hydrogen bond donors, acceptors, rotatable bonds, hydrophobes, log *P*, molecular weight, polar surface area);
2. topological indices (Randic, electrotopological, atomistic, and functional group) [27];
3. those based on eigenvalues of modified adjacency matrices (CIMI) [28] and atomic charges binned into fingerprints.

The charge binning was accomplished by defining a set of 11 bins with charge values between −0.5 and 0.5 eu in 0.1 eu increments. Each bin contained the number of atoms of a given elemental type in the molecule with a charge in the given range. log *P* values were calculated using the HINT! package [29], using explicit hydrogens, polar proximity effects through bonds, and e^{-r} distance function. Polar surface areas were calculated from structures in lowest energy conformations using Gasteiger–Huckel charges as implemented in the molecular spreadsheet algorithms in Sybyl 6.8 [30] on an SGI R5200 O2 workstation.

Within each family, descriptors were not used if they were not represented at least three times in the dataset. The families were not combined because we wanted to evaluate different descriptor types. There are often high correlations among chemical descriptors. In addition, the use of a large number of descriptors can lead to overfitting of the model unless care is taken [31]. Selecting descriptors from a large pool can also lead to chance correlations.

2.3. Bayesian neural nets and automatic relevance determination (ARD)

We have reported theory of Bayesian regularized neural networks applied to QSAR modelling, and a number of applications of this robust SAR mapping method, in several prior publications [11–15]. Consequently we will present only a brief account here.

Conventional backpropagation neural net training methods are usually variations of maximum likelihood algorithms that aim to find a single set of network weights that maximize the fit to training data. Bayesian regularization considers a probability distribution of weights with the probabilities being varied during training in response to how well a particular set of weights models the data. When predicting net output for new data Bayesian regularized nets provide a predictive output distribution for the new case, whereas conventional neural nets give a single value.

Table 1
Compounds and BBB measured partition values

Compound	log BBB	Compound	log BBB
Indinavir	−0.745	Didanosine	−1.301
Verapamil	−0.700	Guanadinothiazole analogue 57	−1.150
SK&F 93319	−1.300	Pentobarbital	−0.120
Lupitidine	−1.060	BBcpd10	−1.170
Ranitidine Analogue 60	−0.730	Guanadinothiazole analogue 15	−0.180
SB-222200	0.300	Ibuprofen	−0.180
Icotidine	−2.000	Clonidine	0.110
bis-Hydroxylated L-663,581	−1.820	Y-G19	−0.430
Trifluoroperazine	1.440	Caffeine	−0.055
Temelastine	−1.880	Antipyrine	−0.097
Zolantidine	0.140	Theophyline	−0.290
ranitidine analogue 26	0.220	Acetylsalicylic acid	−0.500
Cimetidine analogue 14	−0.120	Y-G20	0.250
Ranitidine analogue 21	−0.240	BCNU	−0.520
Mono-hydroxylated L-663,581	−1.340	*p*-Acetamidophenol	−0.310
Hydroxyzine	0.390	ICI17148	−0.040
Ranitidine analogue 19	−0.280	Enflurane	0.240
L-663,581	−0.300	Isoflurane	0.420
Indomethacin	−1.260	Y-G14	−0.420
Thioridazine	0.240	Valproic acid	−0.220
Phenserine	1.000	Salicylic acid	−1.100
Ranitidine analogue 18	−0.270	Fluroxene	0.130
Ranitidine analogue 23	0.690	Heptane	0.810
Ranitidine analogue 24	0.440	3-Methylhexane	0.900
Midazolam	0.360	Teflurane	0.270
Tertbutylchlorambucil	1.000	Toluene	0.370
Ranitidine analogue17	−1.120	Halothane	0.350
Guanidinothiazol analogue 17	−1.540	Sulfur hexafluoride	0.360
Codeine	0.55	Benzene	0.370
Alprazolam	0.044	2,2-Dimethylbutane	1.040
Mepyramine	0.490	Hexane	0.800
Imipramine	1.070	Methylcyclopentane	0.930
Ranitidine	−1.230	2-Methylpentane	0.970
Ranitidine analogue 20	−0.460	3-Methylpentane	1.010
Amitryptalline	0.886	1,1,1-Trifluoro-2-chloroethane	0.080
Chlorpromazine	1.060	Butanone	−0.080
Tiotidine	−0.820	Diethyl ether	0.000
Cimetidine analogue 12	−0.670	2-Methylpropanol	−0.170
SKF89124	−0.060	Pentane	0.760
Oxazepam	0.610	1,1,1-Trichloroethane	0.400
Desipramine	1.200	Trichloroethene	0.340
Promazine	1.230	1-Propanol	−0.160
Physostigmine	0.079	2-Propanol	−0.150
Nevirapine	0.000	Propanone	−0.150
Thioperamide	−0.160	Ethanol	−0.160
Cimetidine analogue 13	−0.660	Nitrous oxide	0.030
Guanidinothiazole analogue 16	−1.570	Carbon disulfide	0.600
CBZ-EPO	−0.350	Nitrogen	0.030
SKF101468	−0.300	Methane	0.040
Zidovudine	−0.720	Y-G15	−1.300
Ranitidine analogue 22	−0.020	Cimetidine analogue 11	−2.150
Carbamazepine	−0.140	Y-G16	−1.400
Cimetidine	−1.420	Chlorambucil	−1.700

Thus the Bayesian neural net can provide a measure of uncertainty or quality of the prediction.

ARD [32,33] model uses multiple regularization constants, one associated with each input. By applying Bayes theorem, the regularization constants for noninformative inputs are automatically inferred to be large, preventing those inputs from causing overfitting.

2.4. QSAR modelling

Descriptors were mean autoscaled (mean centered and divided by the standard deviation), and biological data scaled to between zero and one. We employed a Bayesian neural net implemented in MATLAB code on a PC. As Bayesian nets produce models which are essentially independent of

network architecture, and they are very resistant to overtraining and overfitting, we used a single hidden layer network architecture, with a variable number of nodes in the hidden layer, depending on the number descriptors chosen. Cross-validation is unnecessary for Bayesian net models [34] but we used a test set of 20% of the training set to assess how predictive the models were. There has been extensive discussion in the literature over whether there is a best way to choose a test set and this issue has not been resolved. In our QSAR studies, we have elected to use *k*-means cluster analysis on both independent (X) and independent (Y) data taken together. Such a test set selection method is clearly defined, reproducible, robust, and should be a more rigorous test of predictive ability of models than leave-one-out cross-validation.

We trained the Bayesian net using a gradient descent method until the log of the evidence was a maximum. Typically we repeat the training several times but models are essentially identical, due to the robustness introduced into the training process by the Bayesian regularization.

3. Results and discussion

3.1. QSAR modelling

The Bayesian neural net models were robust with respect to training; repeated training with random initial weights generally yielded essentially identical models. The statistics of the QSAR models derived from the three families of descriptors using Bayesian regularized neural nets are presented in Table 2. It is clear that the property-based descriptors give the best training set statistics, with the topological indices and CIMI/binned charges descriptors producing models with inferior training statistics.

The results in Table 2 illustrate one of the important advantages of the Bayesian neural nets. The models obtained were relatively independent of the architecture of the nets, provided that a minimum number of hidden layer nodes were used. The statistics for models with varying number of hidden layer nodes were very similar, as the Table shows. With the property-based descriptors (seven descriptors), it was necessary to use at least three hidden layer nodes to obtain good models. Increasing the number of hidden layer nodes beyond this minimum resulted in very similar models. This robustness was even more apparent with the other two descriptors classes that had 37 and 21 descriptors compared with seven for the property-based descriptor class. The topological and CIMI/bc descriptor families produced essentially identical models with as few as one hidden layer node. This is suggestive of a fairly linear model when these descriptors are employed.

Another aspect of robustness is also manifest in Table 2. The final two entries under each descriptor family (shown in bold) represent numbers of hidden layer neurodes in the neural network for which a standard backpropagation neural

Table 2
Statistics of QSAR models formed from different descriptor families

n_{hidden}	SEE	r^2	SEP	q^2	n_w	n_{eff}	ρ_{eff}
Property-based descriptors [7]							
1	0.50	0.63	0.54	0.59	10	8	10.8
2	0.43	0.69	0.54	0.61	19	12	9.6
3	0.37	0.79	0.17	0.59	28	21	4.5
4	0.37	0.81	0.54	0.65	37	26	3.3
5	0.37	0.80	0.54	0.65	46	28	3.3
6	0.37	0.80	0.54	0.64	55	26	3.3
7	0.37	0.79	0.54	0.65	64	25	3.4
8	0.37	0.78	0.54	0.65	73	26	3.4
9	**0.39**	**0.76**	**0.54**	**0.65**	**82**	**24**	**3.6**
10	**0.39**	**0.74**	**0.54**	**0.65**	**91**	**22**	**3.7**
Topological indices [37]							
1	0.50	0.62	0.57	0.65	41	21	4.1
2	0.50	0.62	0.57	0.65	80	22	3.7
3	**0.50**	**0.62**	**0.57**	**0.65**	**119**	**24**	**3.6**
4	**0.50**	**0.61**	**0.57**	**0.65**	**157**	**24**	**3.6**
CIMI/bc descriptors [21]							
1	0.54	0.55	0.54	0.65	24	15	5.7
2	0.54	0.55	0.54	0.65	47	16	5.1
3	0.54	0.55	0.54	0.65	70	17	5.1
4	**0.54**	**0.54**	**0.54**	**0.64**	**93**	**18**	**4.8**
5	**0.54**	**0.54**	**0.54**	**0.64**	**116**	**18**	**4.8**

Statistics in bold represent models for which a standard backpropagation neural net would be overfitted. Where r^2 is the squared correlation coefficient for the fit (training set); SEE the standard error of estimation (training set); q^2 the squared correlation coefficient for the predictions (test set); SEP the standard error of prediction (test set); n_w the number of weights in the neural network; n_{eff} the number of effective weights in the network (after Bayesian regularization); and ρ_{eff} is the ratio of the number of compounds in the training set to the number of effective weights. The numbers appeared in square brackets following the table subheadings are not references. They refer to the number of descriptors in the model.

network would be formally overfitted (number of weights is greater than the number of training examples). This situation is analogous to multiple regression case where the number of fitted variables is greater than the number of training examples. As the Bayesian neural nets choose the most parsimonious models, which utilize the number of effective parameters (shown in the second last column of Table 2) not the total number of weights (third last column), the Bayesian neural net models are not overfitted in this regime. The 'overfitted' models are essentially the same as the models with fewer numbers of hidden layer neurodes. Clearly, this is due to the convergence of the number of effective parameters as the number of hidden layer neurodes is increased, as can be seen in Table 2. The ρ_{eff} numbers in the table are the ratio of number of training examples to the number of effective parameters. As can be seen this converges (unlike a standard backpropagation neural net) and their values are always greater than three, even when a standard net would be overtrained ($\rho < 1$). In all models, it is clear that the number of effective parameters in the Bayesian neural net models (15–26) are substantially less than the number of compounds in the training set ($n = 85$) as Table 2 shows.

Analysis of the test set q^2 values shows that all three descriptor families generate models with essentially the same predictive power. The standard errors of prediction (SEP) are also essentially identical for all three descriptor types. The SEE values for training are less than or equal to the SEP values in all cases. All three model families explain 65% of the variance. Given the paucity of data, the differences in experimental protocol, and the difficulty of measuring BB partition coefficients, it is likely that the residual error is largely due to these factors.

These results suggest that the property-based descriptors are the most efficient, as they predict new molecules as well or better than the other two families, and involve fewer descriptors than the topological descriptor family. However, the CIMI/bc descriptors are relatively simplistic as they do not recognize differences between some important atom types (e.g. halogens). Consequently, we are investigating extensions of this promising descriptor class.

We compared the statistics of our model with those of the BBB QSAR model reported by Rose et al. based on electrotopological indices and multiple linear regression using the RSQUARE selection method in SAS. Their training r^2 and cross-validated (leave-one-out) q^2 values are very similar to those of our model employing topological (including electrotopological) indices. Our best training model using feature-based descriptors has a r^2 value of 0.81 and a standard error of estimation (SEE rms) of 0.37. In addition, our model was based on all 106 compounds whereas Rose et al. omitted four outliers that were poorly predicted by their model. The corresponding statistics for their best model using the resultant 102 compounds are r^2 of 0.66 and SEE (rms) of 0.45.

The q^2 values for cross-validation, jack-knifing, or prediction of a separate test set, provide measures of varying degrees of rigor for assessing the predictive ability of models. Rose et al. reported a leave-one-out cross-validated q^2 value of 0.62 and an s_{press} of 0.48 for their best model with four outliers omitted. They also reported an rms error of 0.38 for prediction of a test set partitioned as 20% of the dataset (predictive r^2 or q^2 values not reported). Our models using any descriptor families yielded q^2 values of 0.65 with SEP values of 0.54. Our models appear to be more general than those reported by Rose et al. as they were based on all compounds in the model (no outliers omitted). These authors note that models of the quality we report approach the best that can be derived given the likely experimental error in the BBB partition data. The overall similarity of prediction of Rose et al.'s model and ours suggests that the response surface being modelled is close to linear, at least for some families of descriptors. Abraham [17] has pointed out that, with an appropriate choice of descriptors, the BBB partition can be modelled by a linear relationship. The effect of Bayesian regularization of neural networks is to produce the most parsimonious models in which the optimum compromise is found between bias (model too simple to describe the underlying relationship) and variance (model so complex it fits the noise). Modelling an almost linear relationship is a test of how well this parsimony is implemented, as it represents one extreme in the continuum of response surface nonlinearity. Our results mimic those of Abraham's group, suggesting that the neural net has minimal complexity. Further description of the properties of Bayesian regularized neural nets may be found in our original paper on this technique and the references therein [13], and the seminal papers by Mackay [34].

3.2. Model interpretation

We used the ARD method, which also uses Bayesian statistics, to determine which of the descriptors were the most relevant. The ARD analyses used the entire dataset ($n = 106$) in the training set. Fig. 1 summarized the relative importance of the feature-based descriptors on a log scale. Clearly the octanol–water partition coefficient ($\log P$), the molecular flexibility (number of rotatable bonds), the polar surface area (PSA), and the number of hydrogen bond donors were the most important descriptors in the model. None of the descriptors were significantly correlated with each other except $\log P$ and PSA for which the correlation was not high ($\sim$0.5). The molecular weight was the least significant descriptor in the model. The relevance of these property-based descriptors is consistent with the three electrotopological descriptors from Rose et al.'s model. They found that the HS^T (HBd), HS^T (arom) and $d^2\chi^2$ descriptors were the most important. HS^T (HBd) is their representation of hydrogen bond donor ability (analogous to our donor descriptor), HS^T (arom) represents non-polar effects (analogous to our $\log P$ and hydrophobe descriptors) and $d^2\chi^2$ is related to structural skeletal (branching) variation (which

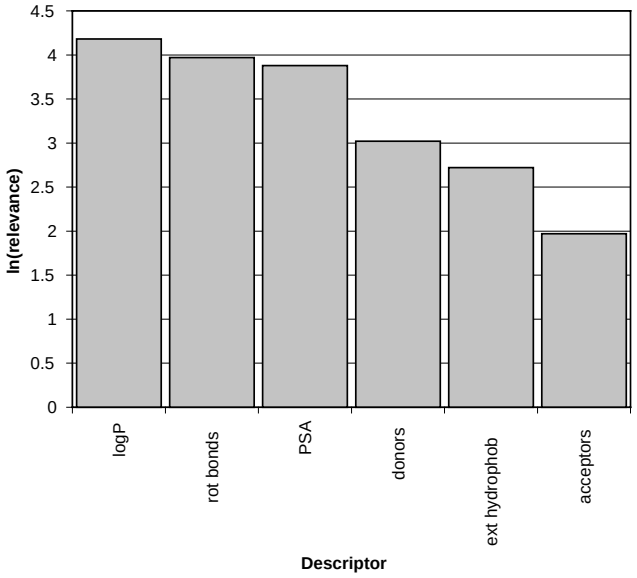

Fig. 1. Automatic relevance determination (ARD) bar plot showing relevances of property-based descriptors in the blood–brain barrier model. Four hidden layer neurodes were used in the neural net.

may correlate positively or negatively with molecular flexibility).

4. Conclusions

Bayesian neural nets have formed robust models of blood–brain barrier partitioning using three families of molecular descriptors. The models were of similar predictive power and similar to those reported by Rose et al. The models showed that the most relevant molecular properties were related to hydrophobic, hydrogen bond donor, and structural flexibility of the very diverse range of compounds in the training set. The robustness of the models, and their quality are remarkable given the difficulty of measuring this ADME property, and the large diversity of the training set. Such successful modelling using relatively few, interpretable descriptors augers well for extrapolating to larger datasets. Such general models would be valuable for assessing the likely success of CNS targeting, or likely CNS side effects of new drugs and helping prioritize lead areas for further development. In addition, such models would be useful for biasing appropriate BBB behavior into discovery screening sets or libraries. Clearly the major issue holding back the building better BBB models is the lack of large, high quality, diverse datasets.

References

[1] E.K. Wilson, Picking the winners, Chem. Eng. News 80 (2002) 35–39.
[2] M.B. Brennan, Drug discovery: Filtering out failures early in the game, Chem. Eng. News 78 (2000) 63–73.
[3] J.-M. Scherrmann, Drug delivery to brain via the blood–brain barrier, Vasc. Pharmacol. 38 (2002) 349–354.
[4] L.L. Rubin, J.M. Staddon, The cell biology of the blood–brain barrier, Annu. Rev. Neurosci. 22 (1999) 11–28.
[5] M.J. Abbott, I.J. Romero, Transporting therapeutics across the blood–brain barrier. Mol. Med. Today March 1996, pp. 106–113.
[6] M.D. Habgood, D.J. Begley, N.J. Abbott, Determinants of passive drug entry into the central nervous system, Cell. Mol. Neurobiol. 20 (2000) 231–253.
[7] M. Feher, E. Sourial, J.M. Schmidt, A simple model for the prediction of blood–brain partitioning, Int. J. Pharmaceut. 201 (2000) 239–247.
[8] P. Labute, A widely applicable set of descriptors, J. Mol. Graph. Mod. 18 (2000) 464–477.
[9] P. Crivori, G. Cruciani, P. Carrupt, B. Testa, Predicting blood–brain barrier permeation from three-dimensional molecular structure, J. Med. Chem. 43 (2000) 2204–2216.
[10] K. Rose, L. Hall, L.B. Kier, Modelling Blood–brain Barrier Partitioning Using the Electrotopological State, J. Chem. Inf. Comput. Sci. 42 (2002) 651–666.
[11] F.R. Burden, D.A. Winkler, New SAR methods applied to structure-activity mapping and combinatorial chemistry, J. Chem. Inf. Comput. Sci. 39 (1999) 236–242.
[12] F.R. Burden, D.A. Winkler, The computer simulation of high throughput screening of bioactive molecules, in: K. Gundertofte, F.S. Jorgensen (Eds.), Molecular Modelling and Prediction of Bioactivity, Plenum Press, 1998.
[13] F.R. Burden, D.A. Winkler, Robust QSAR models using Bayesian regularized artificial neural networks, J. Med. Chem. 42 (1999) 3183–3187.
[14] F.R. Burden, D.A. Winkler, A QSAR model for the acute toxicity of substituted benzenes towards *Tetrahymena pyriformis* using Bayesian regularized neural networks, Chem. Res. Toxicol. 13 (2000) 436–440.
[15] D.A. Winkler, F.R. Burden, Robust QSAR models from novel descriptors and Bayesian regularized neural networks, Mol. Simul. 24 (2000) 243–258.
[16] R.C. Young, Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H2 receptor histamine antagonists, J. Med. Chem. 31 (1988) 656–671.
[17] M.H. Abraham, H.S. Chadha, R.C. Mitchell, Hydrogen bonding. Part 33. Factors that influence the distribution of solutes between blood and brain, J. Pharm. Sci. 83 (1994) 1257–1268;
(b) H.S. Chadha, M.H. Abraham, R.C. Mitchell, Correlation and prediction of blood–brain distribution, J. Mol. Graph. 11 (1993) 281–282;
(c) J.A. Platts, M.H. Abraham, Y.H. Zhao, A. Hersey, L. Ijaz, D. Butina, Correlation and prediction of a large blood–brain distribution dataset—an LFER study, Eur. J. Med. Chem. 36 (2001) 719–730.
[18] T. Salminem, A. Pulli, J. Taskinen, Relationship between immobilized artificial membrane chromatographic retention and the brain penetration of structurally diverse drugs, J. Pharm. Biomed. Anal. 15 (1997) 469–677.
[19] D.E. Clark, Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood–brain barrier penetration, J Pharm. Sci. 88 (1999) 815–821.
[20] J.M. Luco, Prediction of brain-blood distribution of a large set of drugs from structurally derived descriptors using partial least squares (PLS) modelling, J. Chem. Inf. Comput. Sci. 39 (1999) 396–404.
[21] M. Yazdanian, S.L. Glynn, In vitro blood–brain barrier permeability of nevirapine compared to other HIV antiretroviral agents, J. Pharm. Sci. 87 (1998) 306–310.
[22] N.H. Grieg, A. Brossi, P. Xue-Feng, D.K. Ingram, T. Soncrant, in: J. Greenwood, et al. (Eds.), New Concepts of a Blood–brain Barrier, Plenum: New York, 1995; pp 251–264.
[23] J.H. Lin, I. Chen, T. Lin, Blood–brain barrier permeability and in vivo activity of partial agonists of benzodiazepine receptor: a study of L-663, 581 and its metabolites in rats, J. Pharmacol. Exptl. Therapeut. 271 (1994) 1197–1202.

[24] F. Lombardo, J.F. Blake, W. Curatolo, Computation of brain-blood partitioning of organic solutes via free energy calculations, J. Med. Chem. 39 (1996) 4750–4755.
[25] K. Van Belle, S. Sarre, G. Ebinger, Y. Michotte, Brain, liver, and blood distribution kinetics of carbamazepine and its metabolic interaction with clomipramine in rats: a quantitative microdialysis study, J. Pharmacol. Exptl. Therapeut. 272 (1995) 1217–1222.
[26] J.A.D. Calder, C.R. Ganellin, Predicting the brain-penetrating capability of histaminergic compounds, Drug Des. Disov. 11 (1994) 259–268.
[27] L.B. Kier, L.H. Hall, Molecular Structure Description: The Electrotopological State, Academic Press, San Diego, 1999.
[28] F.R. Burden, A chemically intuitive molecular index based on the eigenvalues of a modified adjacency matrix, Quant. Struct.-Act. Relat. 16 (1997) 309–314.
[29] G.E. Kellogg, S.F. Semus, D.J. Abraham, HINT—A new method of empirical hydrophobic field calculation for CoMFA, J. Comput.-Aided Mol. Design 5 (1991) 545–552.
[30] Tripos Associates, St. Louis (http://www.tripos.com).
[31] P.H. Westfall, S.S. Young, D.K.J. Lin, Forward selection error control in the analysis of supersaturated designs, Statistica Sinica 8 (1998) 101–117.
[32] F.R. Burden, M. Ford, D. Whitley, D.A. Winkler, The use of automatic relevance determination in QSAR studies using Bayesian neural nets, J. Chem. Inf. Comput. Sci. 40 (2000) 1423–1430.
[33] R.M. Neal, Bayesian Learning for Neural Networks, Springer-Verlag, New York, 1996.
[34] D.J.C. MacKay, A practical Bayesian framework for backprop networks, Neural Comput. 4 (1992) 415–447;
(b) D.J.C. Mackay, Probable networks and plausible predictions—a review of practical Bayesian methods for supervised neural networks, Comput. Neural Sys. 6 (1995) 469–505;
(c) D.J.C. Mackay, Bayesian interpolation, Neural Comput. 4 (1992) 415–447.