In Silico Prediction of Blood Brain Barrier Permeability: An Artificial Neural Network Model

Prabha Garg* and Jitender Verma

National Institute of Pharmaceutical Education & Research, Sector 67, S.A.S. Nagar, Punjab -160 062, India

Received July 26, 2005

This paper has two objectives: first to develop an in silico model for the prediction of blood brain barrier permeability of new chemical entities and second to find the role of active transport specific to the P-glycoprotein (P-gp) substrate probability in blood brain barrier permeability. An Artificial Neural Network (ANN) model has been developed to predict the ratios of the steady-state concentrations of drugs in the brain to those in the blood (logBB) from their molecular structural parameters. Seven descriptors including P-gp substrate probability have been used for model development. The developed model is able to capture a relationship between P-gp and logBB. The predictive ability of the ANN model has also been compared with earlier computational models.

INTRODUCTION

The Blood Brain Barrier (BBB) is a physical barrier in the circulatory system that stops many substances from traveling into the Central Nervous System (CNS). The drugs which are intended to interact with their molecular targets in the CNS must cross the BBB in order to be used as therapeutic agents. The peripherily acting agents should not cross the BBB so as to avoid the CNS related side effects. In both cases the BBB permeability of the molecules must be known. BBB permeability is expressed as $\log(C_{brain}/C_{blood})$ i.e., the ratio of the steady-state concentrations of the drug molecule in the brain and in the blood. Recently it has also been quantified as logPS i.e., permeability-surface area product.[1]

The experimental BBB permeability can be determined either by in vivo or in vitro methods. In vivo methods involves the measurement of the fraction of drug transported either by direct assay in brain tissues or in blood samples in order to study the disappearance of the drug.[2] The various methods include brain blood partitioning, brain perfusion, the indicator dilution technique, brain uptake index, the capillary depletion technique, and intracerebral microdialysis.[3] Out of the various in vivo methods, the commonly used are determinations of brain-plasma ratio (logBB) and measurement of the permeability−surface area product (PS or logPS), from which permeability P can be derived provided the vessel surface area (S) can be estimated. All of these in vivo methods are laborious and low throughput in nature. Therefore usually high throughput in vitro methods are used.[4] In these methods, the brain microvessels are isolated from the brain and seeded in the culture medium. This results in the growth of the endothelial cells to form monolayers suitable for experimental examination.[5] These models may be cell-based e.g. MDCK cell lines or noncell-based e.g. IAM, PAMPA.[6]

In silico prediction methods are gaining popularity in drug discovery processes as they are inexpensive and less time-consuming. Various computational methods that have been commonly used by earlier researchers for BBB model generation includes Multiple Linear Regression (MLR), Partial Least Square (PLS), and grid based approaches such as Volsurf.[7] Recently BBB models have also been generated using Genetic Algorithms[8] and Artificial Neural Networks.[9]

Modeling the BBB permeability of drugs is a challenge in drug design both because of the quality and quantity of in vivo BBB data available and the difficult task of establishing a useful relationship between the molecular structure and the measured blood brain partitioning.[10] The data that are available on BBB permeability is not only limited but also often uncertain, contradictory, and generated from different experimental protocols.[11] Furthermore there are various other factors that can influence the BBB penetration e.g. plasma protein binding, active efflux from the CNS by transporters such as P-glycoprotein, and metabolism. CNS active drugs can cross the BBB by different mechanisms. CNS inactive drugs show an even more complex situation, some simply do not penetrate at all, whereas others are rapidly metabolized or expelled by active efflux processes.[7] All the BBB models that have been built so far are based on the assumption that the majority of the drugs are transported across the BBB by passive diffusion. Absence of data in the public domain regarding active transport or P-gp efflux as well as the limited understanding and complexity of the behavior of drugs as substrates or inhibitors of these transporters[12] have restricted the people from generating BBB models based on active transport. All of these factors limit the development of highly predictive models of BBB permeation.

Despite all the limiting factors, a number of attempts have been made in the last 17 years to generate improved computational models for predicting the BBB permeability of the molecules. In 1988, Young and co-workers investigated the important physicochemical properties (Delta logP) for brain penetration using the linear regression method on

* Corresponding author phone: +91172-2214682; e-mail: prabhagarg@niper.ac.in, gargprabha@yahoo.com.

10.1021/ci050303i CCC: $33.50
Published on Web 11/09/2005

centrally acting histamine H_2 antagonists.[13] Van de Waterbeemd and Kansy used Young's data set to further investigate the importance of hydrogen bonding to logBB by the Multiple Linear Regression (MLR) method and obtained a good relationship with logBB using the Polar Surface Area (PSA) and molecular volume.[14] Calder and Ganellin also investigated Young's data set using experimental delta logP values as well as theoretically computed parameters such as PSA and molecular volume by the MLR method.[15] Abraham and workers extended Young's data set with an additional 35 compounds, to form Abraham's data set of 57 structures.[16] They developed MLR models using descriptors such as $\log P_{oct}$, excess molar refraction, dipolarity/polarizability, overall H-bonding acidity and basicity, and McGowan volume. They have studied the factors influencing the distribution of solutes between blood and brain, determination of blood brain distribution using octanol−water partition coefficients,[17] and the blood brain barrier partitioning of ampholytes.[18] Lombardo and co-workers analyzed Abraham's data set using MLR on free solvation energy (in water) parameter.[19] Basak and co-workers used discriminant function analysis to determine the efficacy of using nonempirical parameters (logP, H-bonding parameter and topological indices) in the estimation of blood brain transport, inferred from central nervous system (CNS) activity, for a set of 28 compounds.[20] Kaliszan and Markuszewski used the 20 compound Young's data set to relate BBB permeability with the partition coefficient (logP) and the molecular weight by the MLR method.[21] Salminen and colleagues presented a new data set of 29 diverse molecules and related their BBB penetration with immobilized artificial membrane chromatographic retention, molecular volume, and an ionization indicator using MLR.[22]

Norinder and co-workers analyzed Abraham's data set using a quantum chemistry-based approach with descriptors related to lipophilicity, polarizability, H-bonding, Lewis acid−base strength, polarity, charge-transfer interactions, and Partial Least Square (PLS) multivariate data analysis.[23] Clark used the data set of Lombardo and obtained a good MLR model using PSA and ClogP.[24] Kelder and workers also performed MLR analysis and obtained a good correlation with logBB using the dynamic polar surface area.[25] Luco related the BBB permeation with topological and constitutional descriptors using PLS statistics.[26] Osterberg and Norinder used a PLS multivariate approach to relate the polar surface area with logP and the number of H-bond acceptors and H-bond donors.[27] Ertl and co-workers used MLR with topological PSA and achieved good statistics for the data sets of Kelder and Clark.[28] Feher and colleagues used the data set of Luco and related the BBB permeability with PSA, logP, and the number of H-bond acceptors in an aqueous medium using the MLR method.[29] Crivori and co-workers used the Volsurf approach to relate BBB penetration with descriptors derived from 3D molecular fields such as molecular volume, rugosity, globularity, integy moment, amphiphilic moment, critical packing, molecular weight, polarizability, hydrophilic−lipophilic balances, etc. using the Principal Component Analysis (PCA) and discriminant PLS methods.[7] Keseru and Molnar used a thermodynamic approach similar to Lombardo and related the solvation free energies with BBB permeation using a linear regression method.[30] Kaznessis and workers used Monte Carlo simulations to calculate physically significant descriptors such as solvent accessible surface area, solute dipole, number of H-bond acceptors and donors, and molecular weight and related them with BBB permeability using the MLR method.[31] Liu and co-workers introduced a new molecular lipoaffinity descriptor to account for the effect of molecular hydrophobicity on blood brain barrier penetration along with molecular weight, using the MLR method.[32] Platts and workers generated linear free energy relation models of the equilibrium distribution of molecules between blood and brain, as logBB values, using MLR analysis with descriptors related to H-bonding capacity, polarity/polarizability, and molecular size.[33]

Fu and co-workers predicted the blood brain barrier penetration using the stepwise MLR method with the polar molecular surface area and the molecular volume.[34] Rose and workers modeled the BBB permeability using MLR analysis with electrotopological state descriptors such as the E-State index for hydrogen bond donors, the hydrogen E-State index for aromatic CHs, and the valence molecular connectivity index.[10] Iyer and co-workers used the Membrane-interaction QSAR approach to develop predictive models of BBB partitioning by simulating the interactions of the organic compounds with the phospholipid-rich regions of cellular membranes and other intramolecular solute descriptors, using methods such as MLR and Genetic Algorithm.[8] Hutter generated a model using the MLR and PCA methods for BBB prediction using descriptors derived from quantum chemical information like Mean Electrostatic Potential (MESP), product of total variance of ESP and balance parameter to ESP, number of H-bond donors, covalent H-bond basicity, sum of halogens, number of nitro groups, sulfur atoms, aromatic six-membered rings, ionization potential, molecular geometry, and number of rotatable sigma bonds.[35] Hou and Xu developed a QSPR model for BBB prediction using the MLR method with logP, high-charged polar surface areas and excessive molecular weight larger than 360.[36] Subramanian and Kitchen modeled the BBB permeability (as logBB) of a structurally diverse set of 281 compounds using linear regression and a multivariate genetic partial least squares (G/PLS) approach. They used logP, polar surface area, and a variety of electrotopological indices for generating the models for predicting BBB permeability of the molecules.[37] Abraham studied the factors influencing permeation across the blood brain barrier and related the same set of descriptors used earlier by him with the product of permeability and surface area i.e., logPS using the PLS method.[38] Fu and co-workers developed an artificial neural network model to predict the BBB permeability of the molecules from their structural parameters such as molecular volume, sum of net atomic charges of O- and N-atoms which are H-bond acceptors, and the sum of net atomic charges of H-atoms attached to N- and O-atoms.[39] Liu and workers generated a MLR based model to relate logD, van der Waals surface area, and polar surface area of the molecules with the permeability-surface area product.[1] Winkler and Burden related the BBB permeability of the molecules with logP, PSA, number of rotatable bonds, and H-bond donors and acceptors using Bayesian neural nets.[40] Cabrera[41] and co-workers predicted the BBB permeability of the molecules using a topological substructural molecular design approach (TOPS-MODE) with MLR and descriptors such as hydro-

phobicity, PSA, etc. Sun[42] developed the predictive models for partition coefficient, aqueous solubility, blood brain barrier permeability, and human intestinal absorption using the Partial Least Squares (PLS) approach. Burns[43] and Weaver developed a MLR based BBB permeability model using 14 theoretically derived biophysical descriptors based on topological and hydrogen bonding properties of the molecules. Yap[9] and Chen generated Quantitative Structure-Pharmacokinetic Relationships (QSPkR) for BBB permeability, human serum albumin binding, and milk-plasma distribution by using general regression neural network, multilayer feedforward neural network, and MLR. Ma[44] and co-workers developed a predictive MLR based model of blood brain barrier penetration of organic compounds using intermolecular and intramolecular solute descriptors such as polar surface area, octanol/water partition coefficient, Balaban Index, the strength of a small molecule to combine with the membrane−water complex, and the changeability of the structure of a solute−membrane−water complex. Recently Narayanan[45] and Gunturi had generated prediction models for blood brain barrier permeation using a systematic variable selection method along with MLR and descriptors such as the Kappa shape index of order 1, Atomic type E-state index (SsssN), Atomic level based AI topological descriptor (AIssssC), and AlogP98.

The models developed above are based on the assumption of passive diffusion. The role of P-glycoprotein in extruding a variety of structurally unrelated compounds and preventing their accumulation within the brain had already been established experimentally.[46] There is no simple way of dealing with this with the existing statistical models as the relationship is nonlinear. In this work, active transport of the molecules has been taken into consideration in the form of their probability of becoming a substrate to P-glycoprotein so as to generate accurate and reliable BBB models that can more closely mimic the in vivo situation. The results show that the P-gp plays a role in BBB permeability for some of the molecules.

MATERIALS AND METHODS

Descriptor Selection. The selection of the descriptor is very crucial for a reliable and accurate prediction model. The descriptors used by other researchers to predict BBB permeability includes LogP, Molecular Weight, Number of H-Bond Donors, Number of H-Bond Acceptors, Polar Surface Area (PSA), Molecular Volume, Molar Refraction, Polarizability, Overall H-Bonding Acidity, Overall H-Bonding Basicity, McGowan Volume, Free Energy of Solvation, Non-Polar Surface Area, Dynamic Polar Surface Area, Static Polar Surface Area, Number of O-Atoms, Number of N-Atoms, Number of Hydrogen bound to O- and N-Atoms, and various topological descriptors. The first four descriptors (from Lipinski's "Rule of five") have been used more frequently.

Initially 18 descriptors were selected to develop the NN model including LogP, Molecular Weight, Number of H-Bond Donors, Number of H-Bond Acceptors, Number of Rotatable Bonds, and P-gp Substrate Probability. The developed model was not showing a good correlation between the observed vs predicted logBB values. Strongly correlated descriptors were not selected for the model development. Some of the descriptors were also removed based on their significance in the first model. Finally, seven descriptors including P-gp Substrate Probability were used for model building. Other descriptors include Molecular Weight, Topological Polar Surface Area, ClogP, No. of H-Bond Acceptors and Donors, and No. of Rotatable Bonds.

P-gp Substrate Probability. The ability of a molecule to act as a substrate or an inhibitor of P-gp can be determined using in vitro assay systems which generally involve the measurement of bidirectional transport i.e., transcellular (from apical to basolateral chamber) and carrier-mediated (from basolateral to apical chamber) efflux transport.[47] The chambers referred to the monolayers of differentiated epithelial cells joined by intercellular tight junctions. The in vitro assay systems can be classified into three main categories:[48]

1. Transport across the polarized cell monolayers expressing P-gp on the apical membrane.
2. Drug uptake in to cells overexpressing P-gp.
3. Direct binding to P-gp using inside-out membrane vesicles or the reconstituted P-gp.

Some computational approaches have also been made to predict the potential of a compound to interact with P-gp. Some of the earliest models indicate P-gp substrates to be large amphipathic molecules with a basic amine and two or more aromatic rings.[49,50] But exceptions to these criteria have already been identified. Some researchers have now found that the compounds with H-bond acceptors, separated by 4.6 Å or 3-acceptors placed 2.5 Å from each other, can interact with P-gp.[51,52] Others have also suggested a relationship between H-bond acceptor ability and the potential to interact with P-gp.[53,54] 3D-QSAR studies of P-gp substrates and inhibitors have also been performed by some workers.[55] Some have developed a computational ensemble pharmacophore model for identifying P-gp substrates,[56] while others have used the classification SAR (C-SAR) for predicting P-gp substrate specificity.[57] P-gp interacting agents have also been tried to be predicted using Molsurf parametrization and PLS statistics.[58]

Data Collection. A large data set of experimental logBB have been collected from the following published papers and online sources:

1. The work of James A. Platts and associates.[33]
2. The work of Kimberly Rose and associates.[10]
3. ChemSilico Blood Brain Barrier (CSBBB) training set compounds.[59]
4. ChemSilico Blood Brain Barrier (CSBBB) external validation set compounds.[60]

The structures of the molecules were obtained from the ADME Boxes 2.0[61] or drawn in ChemOffice.[62] The model was built using seven descriptors. Five of these descriptors i.e., molecular weight (Mol Wt), topological polar surface area (TPSA), number of H-bond acceptors (HBA), number of H-bond donors (HBD), and number of rotatable bonds (NRB) were calculated using the software DS ViewerPro Property Calculator.[63] ClogP was calculated using ChemOffice,[62] and P-glycoprotein substrate probability was calculated using ADME Boxes 2.0.[61]

Model Development. The data set ($n = 191$) was divided into training ($n = 141$) and test sets ($n = 50$). A 4-layered 7−5−2−1 neural network architecture was used with log-sigmoid transfer function in the hidden layers and linear

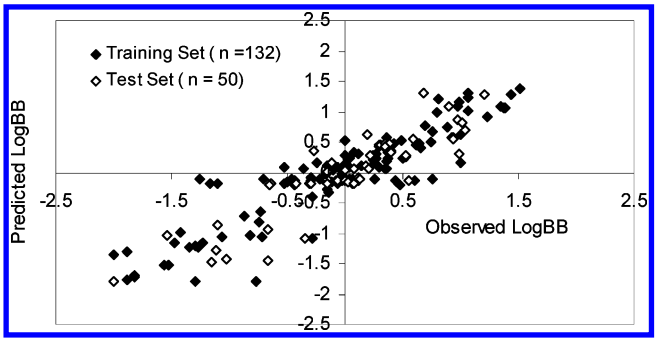

Figure 1. Observed vs predicted values of logBB.

Table 1. Regression Statistics

parameter	training set	test set	working set
correlation coefficient (*r*)	0.90	0.89	0.90
coefficient of determination	0.82	0.80	0.81
adjusted R square	0.81	0.80	0.81
standard error	0.30	0.32	0.30
observations	132	50	182

transfer function in the output layer. MatLab 6.5.1[64] was used for training the model. Feed-forward scaled conjugate gradient back-propagation learning algorithm was used for weight adjustment.

RESULTS AND DISCUSSION

Model Validation. Training of the neural network model started with 141 molecules. But 9 molecules were found to be the outliers (discussed in the later part of the section). After removing these outliers, the remaining set of 132 molecules was used to train the neural network model. The trained model was validated with a test set of 50 molecules. Observed vs predicted values for the working data set are shown in Figure 1.

The regression statistics are shown in Table 1. Results show a good correlation between observed vs predicted values. Hence the model is validated. The model must be validated on a larger data set, but due to limitations of availability of experimental data in the public domain, the model was tested with 50 molecules only. The test set size used in this study is one of the largest data sets used in logBB prediction models.

Outliers. During training, 9 molecules were found to be the outliers. They include arduan (pipecurium), pavulon (pancuronium), nuromax (doxacurium), mivacron (mivacurium), norcuron (vecuronium), zemuron (rocuronium), tracrium (atracurium), raplon (rapacuronium), and brezal (choline-L-alfoscerate). Interestingly except brezal all these molecules belong to the class of neuromuscular blockers, and the experimental logBB values of all of them are extremely low. There are many possible reasons for such low logBB values. For example, if the analytical method is radiochemical detection, any biological degradation will lead to much smaller experimental logBB values than predicted ones.[33] Similarly efflux mechanisms such as P-glycoprotein will also result in more negative experimental logBB values than the predicted ones.[65] To identify the reason for impermeability of nuromuscular drugs, descriptor values of these molecules were analyzed. Table 2 shows the descriptors values of outliers. As can be seen from their descriptor values, the most probable reason for these molecules to act as outliers is their high molecular weight. Except Brezal, all other molecules have a molecular weight greater than Lipinski's cut off i.e., 500. Norcuron, pavulon, raplon, and zemuron are having the P-gp Sub. Prob. above 0.9. However in the case of brezal, it is not the molecular weight or P-gp Sub. prob. but the extremely low ClogP that might be responsible for its low BBB permeability.

Comparison with Other Models. Based on the results obtained from the neural network model, it was of interest to compare and validate this model with other computational methods. The logBB values of 182 molecules used in the training and test sets were also predicted using the online ADMET prediction tools such as PreADME[66] (from Research Institute of Bioinformatics and Molecular Design, Korea, based on the ANN method), CSBBB[67] (from Chem-Silico LLC, Tewksbury, based on the ANN method), and a commonly used commercial software Cerius^2 (from Accelrys Software Inc., U.S.A., based on the MLR method).[68] The predicted values from these software were then compared with those predicted using this ANN model developed, and the comparison is shown in Table 3.

LogBB of 24 molecules could not be predicted from Cerius^2 software as shown −100 in Table 3. Cerius^2 calculates logBB only for molecules whose polar surface area and AlogP98 values lie within 99% confidence limit ellipse.[69] So in order to compare the Cerius^2 predictions with those of this ANN model, these values (24 molecules) were removed, and regression statistics was calculated from the remaining predicted values (n = 158). Table 4 shows their regression statistics.

Clearly the ANN model developed shows better results in comparison to the other software (Table 4). The model has shown very good logBB predictions (R^2 = 0.81) followed by PreADME (R^2 = 0.58), CSBBB (R^2 = 0.57), and then Cerius^2 (R^2 = 0.44).

Only some of the drugs are transported through active transport, and there is no linear relationship between P-gp substrate probability and BBB permeability. Hence to use the predictive model in early drug discovery phases, P-gp

Table 2. Descriptors of Outliers

name	mol wt	TPSA	HBA	HBD	NRB	P-gp	ClogP	logBB
arduan	602.909	59.080	6	0	6	0.553	0.628	−25.320
brezal	258.237	106.030	6	3	8	0.242	−5.643	−10.990
mivacron	1029.292	144.900	14	0	30	0.461	2.636	−21.620
norcuron	557.844	55.840	5	0	6	0.955	4.334	−9.820
nuromax	1005.226	154.130	15	0	28	0.568	1.400	−22.250
pavulon	572.879	52.600	4	0	6	1.000	1.206	−24.500
raplon	597.909	55.840	5	0	9	0.980	5.637	−9.300
tracrium	929.173	126.440	12	0	26	0.366	3.497	−24.330
zemuron	529.790	59.000	5	1	6	0.917	2.433	−9.620

Table 3. Comparison of This Model with Other Available Software

s.no.	compound	exp. logBB	Cerius[2]	CSBBB	Pre ADME	ANN	data set
1	1,1,1-trichloroethane	0.400	0.474	−0.700	0.300	0.361	test
2	1,1,1-trifluoro-2-chloroethane	0.080	0.419	0.250	0.245	0.053	test
3	1,2,3,4-tetrahydroquinoline	0.650	0.314	0.170	0.202	0.392	training
4	1-hydroxymidazolam	−0.070	0.313	−0.660	0.205	0.097	training
5	1-propanol	−0.160	−0.324	−0.030	0.051	−0.189	training
6	2,2-dimethylbutane	1.040	0.667	0.600	0.582	0.702	test
7	2-methylpentane	0.970	0.744	0.690	0.922	0.854	test
8	2-methylpropanol	−0.170	−0.225	0.040	0.084	−0.169	training
9	2-propanol	−0.150	−0.369	0.000	0.037	−0.189	test
10	3-methylhexane	0.900	0.885	0.710	1.045	1.093	test
11	3-methylpentane	1.010	0.744	0.700	0.922	0.821	test
12	4-hydroxyalprazolam	−1.480	−0.055	−0.410	−0.249	−1.164	training
13	4-hydroxymidazolam	−0.300	0.313	−0.460	0.001	−0.073	training
14	9-hydroxy risperidone	−0.670	−0.637	−0.750	1.344	−0.933	test
15	acetaminophen	−0.310	−0.741	−0.580	−0.007	−0.190	training
16	acetylsalicylic acid	−0.500	−0.792	−0.680	−0.463	−0.162	training
17	albuterol	−1.030	−1.053	−1.040	−1.648	−1.437	test
18	alprazolam	0.044	0.312	0.210	−0.112	0.236	training
19	aminopyrine	0.000	−0.161	0.270	0.057	0.279	training
20	amitriptyline	0.886	1.268	0.810	1.049	0.739	training
21	amobarbital	0.040	−0.670	−0.330	−0.246	−0.121	training
22	amphetamine	0.930	−0.068	0.270	−0.077	0.574	training
23	antipyrine	−0.097	−0.037	0.200	0.130	0.031	training
24	argon	0.030	−100.00	−0.050	0.015	0.074	training
25	atenolol	−0.870	−1.314	−0.750	−1.054	−0.714	training
26	atropine	−0.060	−0.419	0.160	1.240	−0.146	test
27	BBcpd10	−1.170	−1.228	−0.480	−0.322	−0.193	training
28	BBcpd12 (cimetidine derivative)	−0.670	−0.597	−0.830	−0.953	−1.449	test
29	BBcpd13 (cimetidine derivative)	−0.660	−0.828	−0.430	−0.667	−0.214	test
30	BBcpd14 (cimetidine derivative)	−0.120	−0.218	−0.210	−1.584	−0.155	training
31	BBcpd15 (guanidinothiazole der)	−0.180	−0.754	−0.190	−0.053	−0.114	training
32	BBcpd16 (guanidinothiazole der)	−1.570	−1.502	−0.850	−0.979	−1.536	training
33	BBcpd17 (ranitidine analog)	−1.120	−0.818	−0.520	−0.220	−1.278	test
34	BBcpd18 (ranitidine analog)	−0.270	−0.382	−0.780	0.063	0.348	test
35	BBcpd19 (ranitidine analog)	−0.280	−0.787	−0.620	−1.125	−1.074	training
36	BBcpd21 (ranitidine analog)	−0.240	0.326	0.230	0.666	0.169	training
37	BBcpd22 (ranitidine analog)	−0.020	0.042	0.040	1.383	−0.048	training
38	BBcpd23 (ranitidine analog)	0.690	0.419	0.320	1.138	0.759	training
39	BBcpd24 (ranitidine analog)	0.440	0.357	0.380	1.124	−0.139	training
40	BBcpd26 (ranitidine analog)	0.220	0.430	0.030	0.269	0.291	test
41	BBcpd57 (guanidinothiazole der)	−1.150	−1.404	−0.550	−0.899	−1.465	test
42	BBcpd58 (guanidinothiazole der)	−1.540	−100.00	−1.500	−1.405	−1.037	test
43	BBcpd60 (ranitidine analog)	−0.730	−0.208	−0.460	0.761	−0.653	training
44	BCNU	−0.520	−0.432	−0.560	−0.666	−0.191	training
45	benzene	0.370	0.412	0.420	0.175	0.225	training
46	bishydroxy L-663,581 metabolite	−1.820	−100.00	−1.540	−2.069	−1.696	training
47	bromocriptine	−1.100	−100.00	−1.220	0.163	−0.864	test
48	bromperidol	1.380	0.319	−0.170	1.416	1.049	training
49	buspirone	0.480	−0.430	0.410	1.191	−0.216	training
50	butanone	−0.080	−0.253	0.230	0.055	−0.136	test
51	caffeine	−0.055	−0.939	0.000	−0.043	−0.127	test
52	carbamazepine	−0.140	−0.072	0.020	0.211	0.015	training
53	carbamazepine-10,11-epoxide	−0.350	−0.437	0.070	0.053	−1.096	test
54	carbon disulfide (CS_2)	0.600	−0.357	0.170	0.193	−0.138	training
55	cefotetan	−1.890	−100.00	−1.900	−1.005	−1.777	training
56	CF_3CH_2Cl	0.080	0.419	0.250	0.245	0.053	training
57	$CF_3CH_2OCH{=}CH_2$	0.130	0.058	0.240	0.173	−0.117	test
58	chlorpromazine	1.060	1.205	0.740	0.956	1.224	training
59	cimetidine	−1.420	−1.484	−0.700	−1.332	−0.979	training
60	clobazam	0.350	0.039	0.100	0.332	0.050	training
61	clonidine	0.110	−0.011	−0.630	0.160	−0.076	test
62	codeine	0.550	−0.312	0.030	0.991	−0.122	test
63	cyclohexane	0.920	0.692	0.530	0.628	0.571	training
64	cyclopropane	0.000	−100.00	0.450	0.138	0.127	training
65	desflurane	0.110	0.396	0.510	0.560	0.297	training
66	desipramine	1.200	0.781	0.530	0.979	1.270	test
67	desmethylclobazam	0.360	−0.174	−0.260	0.166	0.055	training
68	desmethyldesipramine	1.060	0.430	0.500	0.924	1.008	training
69	desmethyldiazepam	0.500	0.107	−0.090	0.223	0.231	training
70	desmonomethylpromazine	0.590	0.685	0.700	0.980	0.543	test
71	diazepam	0.520	0.321	0.260	0.450	0.273	test
72	dichloromethane	−0.110	−100.00	−0.080	0.097	0.152	test
73	didanosine	−1.300	−100.00	−0.730	−1.430	−1.796	training
74	diethyl ether	0.000	−100.00	0.450	0.098	−0.130	test

Table 3. (Continued)

s.no.	compound	exp. logBB	Cerius[2]	CSBBB	Pre ADME	ANN	data set
75	divinyl ether	0.110	−100.00	0.130	0.000	−0.140	training
76	enflurane	0.240	0.529	0.330	0.864	0.218	training
77	ethanol	−0.160	−100.00	−0.080	0.011	−0.193	training
78	ethylbenzene	0.197	0.703	0.440	0.610	0.621	test
79	etoposide	−2.000	−100.00	−2.000	−2.065	−1.797	test
80	flunitrazepam	0.060	−0.531	−0.570	0.266	−0.103	training
81	fluphenazine	1.510	0.729	0.050	1.309	1.373	training
82	gentisic acid	0.080	−1.114	−1.310	−0.961	−0.193	test
83	haloperidol	1.340	0.392	−0.010	1.338	1.086	training
84	halothane	0.350	0.622	0.070	0.015	0.486	test
85	heptane	0.810	0.948	0.670	0.800	1.213	training
86	hexane	0.800	0.807	0.670	0.967	0.979	training
87	hexobarbital	0.100	−0.547	−0.240	0.166	−0.141	training
88	hydroxyzine	0.390	0.341	0.140	0.664	0.476	test
89	ibuprofen	−0.180	0.358	−0.010	0.265	−0.076	training
90	ICI 17148	−0.040	−1.268	−0.280	−0.207	−0.096	training
91	icotidine	−2.000	−0.900	−1.330	−1.979	−1.360	training
92	imipramine	1.060	−100.00	0.890	0.967	1.303	training
93	indinavir	−0.745	0.711	−0.340	−1.542	−0.819	training
94	indomethacin	−1.260	0.053	−1.010	−0.621	−0.110	training
95	isoflurane	0.420	0.488	0.330	0.719	0.485	training
96	krypton	−0.160	−100.00	−0.050	0.015	0.106	training
97	L-663,581	−0.300	−0.489	−0.400	−0.847	−0.131	training
98	levodopa	−0.770	−1.715	−1.640	−0.185	−1.797	training
99	levorphanol	0.000	0.530	0.290	1.505	0.535	training
100	lin_test_35	−1.820	−100.00	−1.650	−1.730	−1.718	training
101	lin_train_45	0.350	0.533	0.070	0.473	0.423	test
102	lin_train_55	0.270	0.441	0.310	0.347	0.242	training
103	lupitidine	−1.060	−0.871	−0.660	−1.167	−1.053	training
104	mepyramine	0.490	0.375	0.550	0.733	0.517	training
105	mesoridazine	−0.360	0.842	1.110	0.965	0.063	training
106	methamphetamine	0.990	0.282	0.320	1.244	0.297	test
107	methane	0.040	−100.00	0.910	0.139	0.080	training
108	methohexital	−0.060	0.293	−0.610	−0.095	−0.126	training
109	methotrexate	−1.520	−100.00	−2.000	−0.910	−1.530	training
110	methoxyflurane	0.250	0.367	0.120	0.514	0.162	training
111	methylcyclopentane	0.930	0.629	0.550	0.517	0.551	test
112	mianserin	0.990	0.886	0.900	1.210	1.148	training
113	midazolam	0.360	0.680	0.170	0.090	0.561	training
114	mirtazapine	0.530	0.519	0.730	1.178	0.257	training
115	monohydroxy L-663,581 metabolite	−1.340	−1.129	−0.870	−0.951	−1.230	training
116	morphine	−0.160	−0.570	−0.380	0.067	−0.285	training
117	*m*-xylene	0.295	0.713	0.420	0.565	0.460	training
118	N-desmethylclobazam	0.000	−0.174	−0.260	0.166	0.055	training
119	neon	0.200	−100.00	−0.050	0.015	0.071	test
120	nevirapine	0.000	−0.331	−0.110	0.021	0.032	test
121	nitrogen	0.030	−0.796	0.040	0.032	−0.133	training
122	nitrous oxide	0.030	−100.00	0.030	−0.241	−0.152	test
123	nor-1-chlorpromazine	1.370	0.890	0.430	0.973	1.053	training
124	nor-2-chlorpromazine	0.970	0.539	0.390	1.081	1.075	training
125	northioridazine	0.750	1.144	0.830	0.949	−0.116	training
126	norverapamil	−0.640	0.261	−0.860	−1.624	−0.185	test
127	oxazepam	0.610	−0.259	−0.690	−0.013	0.458	training
128	*o*-xylene	0.366	0.713	0.420	0.565	0.457	training
129	paraxanthine	0.060	−1.152	−0.340	−0.077	−0.193	training
130	pentane	0.760	0.666	0.670	0.785	0.662	training
131	pentobarbital	0.120	−0.670	−0.430	−0.246	−0.121	training
132	pergolide	0.300	0.826	0.520	1.453	0.455	training
133	phencyclidine	0.680	1.195	0.820	1.360	1.306	test
134	phenserine	1.000	0.336	0.410	1.475	0.168	training
135	phenylbutazone	−0.520	0.415	0.550	0.040	0.079	training
136	phenytoin	−0.040	−0.396	−0.110	−0.299	−0.141	training
137	physostigmine	0.079	−0.151	0.160	1.379	0.333	training
138	primidone	−0.070	−0.819	−0.410	−0.217	−0.193	training
139	promazine	1.230	1.000	1.250	0.974	0.909	training
140	propan-1-ol	−0.160	−0.324	−0.020	0.051	−0.189	training
141	propan-2-ol	−0.150	−0.369	0.010	0.037	−0.189	training
142	propanone	−0.150	−100.00	0.420	−0.003	−0.114	test
143	propranolol	0.640	−0.042	−0.010	1.168	0.481	training
144	*p*-xylene	0.314	0.713	0.680	0.565	0.460	test
145	quinidine	−0.460	−0.011	−0.060	1.378	−0.118	training
146	ranitidine	−1.230	−1.167	−0.760	0.152	−1.149	training
147	risperidone	−0.020	−0.064	−0.390	0.979	−0.128	training
148	salicylic acid	−1.100	−0.710	−0.780	−0.648	−0.193	training

Table 3. (Continued)

s.no.	compound	exp. logBB	Cerius^2	CSBBB	Pre ADME	ANN	data set
149	salicyluric acid	−0.440	−1.477	−1.440	−0.466	−0.202	training
150	SB-222200	0.300	1.090	1.000	0.168	0.420	training
151	SK&F 93319	−1.300	−0.144	−0.230	−1.574	−1.217	training
152	SKF101468	−0.300	0.232	0.300	1.406	−0.077	training
153	SKF89124	−0.060	−0.173	−0.530	0.739	−0.124	test
154	sulforidazine	0.180	0.598	1.250	0.962	0.119	test
155	sulfur hexafluoride (SF6)	0.360	−100.00	0.570	0.112	0.172	training
156	tacrine	−0.130	0.111	−0.040	0.249	0.088	training
157	teflurane	0.270	0.530	0.540	0.369	0.301	training
158	temelastine	−1.880	−0.372	−1.370	−2.063	−1.297	training
159	tertbutylchlorambucil	1.000	0.941	0.880	0.876	0.632	training
160	theobromine	−0.280	−1.254	−0.540	−0.077	−0.391	training
161	theophylline	−0.290	−1.156	−0.560	−0.282	−0.193	test
162	thiopental	−0.140	−0.096	−0.250	0.368	−0.324	training
163	thioperamide	−0.160	−0.010	−0.110	−0.053	0.023	test
164	thioridazine	0.240	1.460	1.330	0.917	0.289	training
165	tibolone	0.400	0.585	0.410	1.258	0.330	test
166	tiotidine	−0.820	−100.00	−1.220	−1.305	−1.038	training
167	toluene	0.370	0.562	0.680	0.331	0.337	training
168	triazolam	0.740	0.517	0.370	0.100	0.499	training
169	trichloroethene	0.340	0.385	0.560	0.215	0.398	training
170	trichloromethane	0.290	0.346	0.510	0.186	0.094	training
171	trifluoperazine	1.440	1.225	0.510	0.947	1.269	training
172	valproic acid	−0.220	0.093	−0.060	−0.203	−0.120	training
173	verapamil	−0.700	0.576	−0.420	−1.349	−0.114	training
174	xenon	0.030	−100.00	−0.050	0.015	0.208	training
175	Y-G14	−0.420	−0.324	0.120	0.782	−0.192	test
176	Y-G15	−0.060	−0.009	0.450	1.090	−0.127	training
177	Y-G16	−0.420	−0.832	0.070	−0.795	−0.171	training
178	Y-G19	−0.430	−0.230	0.280	−0.258	−0.137	training
179	Y-G20	0.250	−0.712	−0.130	−0.732	−0.114	training
180	zidovudine	−0.720	−100.00	−1.000	−0.700	−1.052	training
181	zolantidine	0.140	0.903	0.580	0.105	0.103	training
182	α-hydroxyalprazolam	−1.270	−0.225	−0.460	−0.820	−1.231	training

Table 4. Regression Statistics

parameter	Cerius^2	CSBBB	PreADME	ANN
correlation coefficient (r)	0.66	0.76	0.76	0.90
coefficient of determination (R^2)	0.44	0.57	0.58	0.81
adjusted R square	0.43	0.57	0.58	0.81
standard error	0.52	0.43	0.54	0.30
observations	158	182	182	182
training set size	>120	103	88	132
test set size		74	42	50

substrate probability has been used as one of the input variables in predicting the logBB. The developed model is able to capture the effect of P-gp in BBB permeability (Table 5). The descriptor values of some of the molecules in the data set were carefully analyzed, and it indicates that P-gp substrate probability plays a role in determining the BBB permeability of the molecules. For example, the logBB value of icotidine is very low (−2.0) despite its acceptable molecular weight (379.47) and ClogP (1.99) values. This may be due to the reason that its P-gp substrate probability is sufficiently high (0.78). Similarly a high P-gp substrate probability (0.81) of quinidine is responsible for its low logBB value (−0.46) despite having molecular weight (324.43) and ClogP (2.79) values within the acceptable range.

Table 5. Role of P-gp Substrate Probability in Blood Brain Barrier Permeability

name	mol wt	TPSA	HBA	HBD	NRB	P-gp	ClogP	exp logBB	pred logBB
9-hydroxy risperidone	426.50	82.17	8	1	4	0.51	0.95	−0.67	−0.93
albuterol	239.32	72.72	4	4	5	0.93	0.06	−1.03	−1.44
amitriptyline	277.41	3.24	1	0	3	0.58	4.85	0.89	0.74
atenolol	266.34	84.58	5	3	8	0.52	−0.11	−0.87	−0.71
BBcpd14 (cimetidine derivative)	368.46	114.84	5	2	8	0.78	3.98	−0.12	−0.16
bishydroxy L-663,581 metabolite	403.83	116.99	8	2	4	0.76	−0.40	−1.82	−1.70
bromocriptine	654.61	118.21	8	3	6	0.85	6.58	−1.10	−0.86
codeine	299.37	41.93	4	1	1	0.68	0.98	0.55	−0.12
icotidine	379.47	88.50	5	2	9	0.78	1.99	−2.00	−1.36
indinavir	613.81	118.03	7	4	14	0.99	3.68	−0.75	−0.82
lupitidine	413.55	108.06	6	2	10	0.51	1.63	−1.06	−1.05
morphine	285.35	52.93	4	2	0	0.65	0.57	−0.16	−0.29
northioridazine	356.56	65.87	4	1	4	0.75	6.17	0.75	−0.12
norverapamil	440.59	72.74	6	1	14	0.64	3.93	−0.64	−0.19
quinidine	324.43	45.59	4	1	4	0.81	2.79	−0.46	−0.12
SB-222200	380.49	41.99	3	1	6	0.65	6.26	0.30	0.42
temelastine	442.36	79.27	7	2	8	0.51	2.96	−1.88	−1.30
zolantidine	381.55	65.63	4	1	8	0.51	5.88	0.14	0.10

LIMITATIONS

The list of outliers shows that the ANN model developed in this study is not suitable for neuromuscular blockers. However rule based classification has been implemented in the software which identifies impermeable molecules having very low logBB (<−8.0). If the molecules are not impermeable, the logBB value is predicted using the ANN model developed in this study.

CONCLUSIONS

Software based on an Artificial Neural Network model has been developed with seven molecular structural descriptors for predicting the Blood Brain Barrier permeability of the molecules. Significance of each descriptor was checked by setting its value to zero. Out of the various descriptors that have been used, molecular weight appeared to be the most significant parameter for predicting the BBB permeability, followed by topological polar surface area, ClogP, number of H-bond acceptors, number of H-bond donors, and P-gp substrate probability. The number of rotatable bonds was found to be the least contributing parameter towards the BBB permeability of the molecules. Unlike all earlier models which were based on the assumption of passive diffusion, active transport phenomenon had been tried to be modeled in this work in the form of P-gp substrate probability. Out of 182 molecules, only 18 molecules were having the value of P-gp substrate probability >= 0.5 as shown in Table 5. The coefficient of determination between observed and predicted values of logBB is 0.81 for 18 molecules also, having P-gp >=0.5, hence this model appears to be a robust and accurate BBB permeability predictor for most of the druglike molecules. Based on the results, it can be concluded that this model can be used as a high-throughput virtual screening tool in drug discovery particularly in the CNS area.

Supporting Information Available: Smiles notations for all the compounds used in the training and test sets and indication of the training/test set of investigated compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Liu, X.; Tu, M.; Kelly, R. S.; Chen, C.; Smith, B. J. Development of a computational approach to predict blood-brain barrier permeability. *Drug Metab. Dispos.* **2004**, *32*, 132−9.

(2) Smith, Q. R. A review of blood-brain barrier transport techniques. *Methods Mol. Med.* **2003**, *89*, 193−208.

(3) Jolliet-Riant, P.; Tillement, J. P. Drug transfer across the blood-brain barrier and improvement of brain delivery. *Fundam. Clin. Pharmacol.* **1999**, *13*, 16−26.

(4) Gumbleton, M.; Audus, K. L. Progress and limitations in the use of in vitro cell cultures to serve as a permeability screen for the blood-brain barrier. *J. Pharm. Sci.* **2001**, *90*, 1681−98.

(5) Reichel, A.; Begley, D. J.; Abbott, N. J. An overview of in vitro techniques for blood-brain barrier studies. *Methods Mol. Med.* **2003**, *89*, 307−24.

(6) Abbott Joan, N. Prediction of blood-brain barrier permeation in drug discovery from in vivo, in vitro and in silico models. *Drug Discovery Today: Technol.* **2004**, *1*, 407−416.

(7) Crivori, P.; Cruciani, G.; Carrupt, P. A.; Testa, B. Predicting blood-brain barrier permeation from three-dimensional molecular structure. *J. Med. Chem.* **2000**, *43*, 2204−16.

(8) Iyer, M.; Mishru, R.; Han, Y.; Hopfinger, A. J. Predicting blood-brain barrier partitioning of organic molecules using membrane-interaction QSAR analysis. *Pharm. Res.* **2002**, *19*, 1611−21.

(9) Yap, C. W.; Chen, Y. Z. Quantitative Structure-Pharmacokinetic Relationships for drug distribution properties by using general regression neural network. *J. Pharm. Sci.* **2005**, *94*, 153−68.

(10) Rose, K.; Hall, L. H.; Kier, L. B. Modeling blood-brain barrier partitioning using the electrotopological state. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 651−666.

(11) Van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: towards prediction paradise? *Nat. Rev. Drug Discov.* **2003**, *2*, 192−204.

(12) Scala, S.; Akhmed, N.; Rao, U. S.; Paull, K.; Lan, L. B.; Dickstein, B.; Lee, J. S.; Elgemeie, G. H.; Stein, W. D.; Bates, S. E. P-glycoprotein substrates and antagonists cluster into two distinct groups. *Mol. Pharmacol.* **1997**, *51*, 1024−33.

(13) Young, R. C.; Mitchell, R. C.; Brown, T. H.; Ganellin, C. R.; Griffiths, R.; Jones, M.; Rana, K. K.; Saunders: D.; Smith, I. R.; Sore, N. E.; et al. 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.* **1988**, *31*, 656−71.

(14) Van de Waterbeemd, H.; Kansy, M. Hydrogen-bonding capacity and brain penetration. *Chimia* **1992**, *46*, 299−303.

(15) Calder, J. A.; Ganellin, C. R. Predicting the brain-penetrating capability of histaminergic compounds. *Drug Des. Discovery* **1994**, *11*, 259−68.

(16) Abraham, M. H.; Chadha, H. S.; Mitchell, R. C. Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain. *J. Pharm. Sci.* **1994**, *83*, 1257−68.

(17) Abraham, M. H.; Chadha, H. S.; Mitchell, R. C. Hydrogen-bonding. Part 36. Determination of blood brain distribution using octanol−water partition coefficients. *Drug Des. Discovery* **1995**, *13*, 123−31.

(18) Abraham, M. H.; Takacs-Novak, K.; Mitchell, R. C. On the partition of ampholytes: application to blood-brain distribution. *J. Pharm. Sci.* **1997**, *86*, 310−5.

(19) Lombardo, F.; Blake, J. F.; Curatolo, W. J. Computation of brain-blood partitioning of organic solutes via free energy calculations. *J. Med. Chem.* **1996**, *39*, 4750−5.

(20) Basak, S. C.; Gute, B. D.; Drewes, L. R. Predicting blood-brain transport of drugs: a computational approach. *Pharm. Res.* **1996**, *13*, 775−8.

(21) Kaliszan, R.; Markuszewski, M. Brain/blood distribution described by a combination of partition coefficient and molecular mass. *Int. J. Pharm.* **1996**, *145*, 9−16.

(22) Salminen, T.; Pulli, A.; Taskinen, J. Relationship between immobilised artificial membrane chromatographic retention and the brain penetration of structurally diverse drugs. *J. Pharm. Biomed. Anal.* **1997**, *15*, 469−77.

(23) Norinder, U.; Sjoberg, P.; Osterberg, T. Theoretical calculation and prediction of brain-blood partitioning of organic solutes using MolSurf parametrization and PLS statistics. *J. Pharm. Sci.* **1998**, *87*, 952−959.

(24) Clark, D. E. 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.* **1999**, *88*, 815−21.

(25) Kelder, J.; Grootenhuis, P. D.; Bayada, D. M.; Delbressine, L. P.; Ploemen, J. P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* **1999**, *16*, 1514−9.

(26) Luco, J. M. Prediction of the brain-blood distribution of a large set of drugs from structurally derived descriptors using partial least-squares (PLS) modeling. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 396−404.

(27) Osterberg, T.; Norinder, U. Prediction of polar surface area and drug transport processes using simple parameters and PLS statistics. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1408−11.

(28) Ertl, P.; Rohde, B.; Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* **2000**, *43*, 3714−7.

(29) Feher, M.; Sourial, E.; Schmidt, J. M. A simple model for the prediction of blood-brain partitioning. *Int. J. Pharm.* **2000**, *201*, 239−247.

(30) Keseru, G. M.; Molnar, L. High-throughput prediction of blood-brain partitioning: a thermodynamic approach. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 120−8.

(31) Kaznessis, Y. N.; Snow, M. E.; Blankley, C. J. Prediction of blood-brain partitioning using Monte Carlo simulations of molecules in water. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 697−708.

(32) Liu, R.; Sun, H.; So, S. S. Development of quantitative structure−property relationship models for early ADME evaluation in drug discovery. 2. Blood-brain barrier penetration. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1623−1632.

(33) Platts, J. A.; Abraham, M. H.; Zhao, Y. H.; Hersey, A.; Ijaz, L.; Butina, D. Correlation and prediction of a large blood-brain distribution data set−an LFER study. *Eur. J. Med. Chem.* **2001**, *36*, 719−30.

(34) Fu, X. C.; Chen, C. X.; Liang, W. Q.; Yu, Q. S. Predicting blood-brain barrier penetration of drugs by polar molecular surface area and molecular volume. *Acta Pharmacol. Sin.* **2001**, *22*, 663−8.

(35) Hutter, M. C. Prediction of blood-brain barrier permeation using quantum chemically derived information. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 415−33.
(36) Hou, T. J.; Xu, X. J. ADME Evaluation in Drug Discovery. 3. Modeling Blood-Brain Barrier Partitioning Using Simple Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* **2003**, *44*, 766−70.
(37) Subramanian, G.; Kitchen, D. B. Computational models to predict blood-brain barrier permeation and CNS activity. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 643−64.
(38) Abraham, M. H. The factors that influence permeation across the blood-brain barrier. *Eur. J. Med. Chem.* **2004**, *39*, 235−40.
(39) Fu, X. C.; Wang, G. P.; Liang, W. Q.; Yu, Q. S. Predicting blood-brain barrier penetration of drugs using an artificial neural network. *Pharmazie* **2004**, *59*, 126−30.
(40) Winkler, D. A.; Burden, F. R. Modelling blood-brain barrier partitioning using Bayesian neural nets. *J. Mol. Graph. Model.* **2004**, *22*, 499−505.
(41) Cabrera, M. A.; Bermejo, M.; Perez, M.; Ramos, R. TOPS-MODE approach for the prediction of blood-brain barrier permeation. *J. Pharm. Sci.* **2004**, *93*, 1701−17.
(42) Sun, H. A Universal Molecular Descriptor System for Prediction of LogP, LogS, LogBB, and Absorption. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 748−57.
(43) Burns, J.; Weaver, D. F. A mathematical model for prediction of drug molecule diffusion across the blood-brain barrier. *Can. J. Neurol. Sci.* **2004**, *31*, 520−7.
(44) Ma, X. L.; Chen, C.; Yang, J. Predictive model of blood-brain barrier penetration of organic compounds. *Acta Pharmacol. Sin.* **2005**, *26*, 500−12.
(45) Narayanan, R.; Gunturi, S. B. In silico ADME modelling: prediction models for blood-brain barrier permeation using a systematic variable selection method. *Bioorg. Med. Chem.* **2005**, *13*, 3017−28.
(46) Schinkel, A. H. P-Glycoprotein, a gatekeeper in the blood-brain barrier. *Adv. Drug Deliv. Rev.* **1999**, *36*, 179−194.
(47) Yu, S.; Michael, S. H.; Graham, T.; Caco-2 Bi-Directional Transport Assay Using Beckman Coulter's Biomek Automated Platforms; http://www.beckman.com/resourcecenter/literature/BioLit/BioPdf.asp?OrderNumber=A-1985A.
(48) Hochman, J. H.; Yamazaki, M.; Ohe, T.; Lin, J. H. Evaluation of drug interactions with P-glycoprotein in drug discovery: in vitro assessment of the potential for drug-drug interactions with P-glycoprotein. *Curr. Drug Metab.* **2002**, *3*, 257−73.
(49) Zamora, J. M.; Pearce, H. L.; Beck, W. T. Physical-chemical properties shared by compounds that modulate multidrug resistance in human leukemic cells. *Mol. Pharmacol.* **1988**, *33*, 454−62.
(50) Pearce, H. L.; Safa, A. R.; Bach, N. J.; Winter, M. A.; Cirtain, M. C.; Beck, W. T. Essential features of the P-glycoprotein pharmacophore as defined by a series of reserpine analogues that modulate multidrug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5128−32.
(51) Seelig, A. A general pattern for substrate recognition by P-glycoprotein. *Eur. J. Biochem.* **1998**, *251*, 252−61.
(52) Seelig, A.; Landwojtowicz, E. Structure−activity relationship of P-glycoprotein substrates and modifiers. *Eur. J. Pharm. Sci.* **2000**, *12*, 31−40.
(53) Chiba, P.; Holzer, W.; Landau, M.; Bechmann, G.; Lorenz, K.; Plagens, B.; Hitzler, M.; Richter, E.; Ecker, G. Substituted 4-acylpyrazoles and 4-acylpyrazolones: synthesis and multidrug resistance-modulating activity. *J. Med. Chem.* **1998**, *41*, 4001−11.
(54) Ecker, G.; Huber, M.; Schmid, D.; Chiba, P. The importance of a nitrogen atom in modulators of multidrug resistance. *Mol. Pharmacol.* **1999**, *56*, 791−6.
(55) Ekins, S.; Kim, R. B.; Leake, B. F.; Dantzig, A. H.; Schuetz, E. G.; Lan, L. B.; Yasuda, K.; Shepard, R. L.; Winter, M. A.; Schuetz, J. D.; Wikel, J. H.; Wrighton, S. A. Application of three-dimensional quantitative structure−activity relationships of P-glycoprotein inhibitors and substrates. *Mol. Pharmacol.* **2002**, *61*, 974−81.
(56) Penzotti, J. E.; Lamb, M. L.; Evensen, E.; Grootenhuis, P. D. A computational ensemble pharmacophore model for identifying substrates of P-glycoprotein. *J. Med. Chem.* **2002**, *45*, 1737−40.
(57) Didziapetris, R.; Japertas, P.; Petrauskas, A.; Riauba, L. Classification SAR in the Prediction of P-glycoprotein Substrate Specificity Presentation at EuroQSAR 2002, Bournemouth, U.K.
(58) Osterberg, T.; Norinder, U. Theoretical calculation and prediction of P-glycoprotein-interacting drugs using MolSurf parametrization and PLS statistics. *Eur. J. Pharm. Sci.* **2000**, *10*, 295−303.
(59) ChemSilico; CSBBB Training − Set Compounds; http://www.chemsilico.com/CS_prBBB/BBBcomp.html.
(60) ChemSilico; CSBBB External Validation Set Compounds; http://www.chemsilico.com/CS_prBBB/BBBExValcomp.html.
(61) ADME Boxes, ver 2.0; Pharma Algorithms, 591 Indian Road, Toronto, Ontario, M6P 2C4, Canada.
(62) ChemOffice, ver 6.0.1; CambridgeSoft.Com, 100 CambridgePark Drive, Cambridge, MA 02140, U.S.A.
(63) DS ViewerPro, ver 5.0; Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121, U.S.A.
(64) MATLAB, ver 6.5.1; The MathWorks Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, U.S.A.
(65) Demeule, M.; Regina, A.; Jodoin, J.; Laplante, A.; Dagenais, C.; Berthelet, F.; Moghrabi, A.; Beliveau, R. Drug transport to the brain: key roles for the efflux pump P-glycoprotein in the blood-brain barrier. *Vasc. Pharmacol.* **2002**, *38*, 339−48.
(66) PreADME, ver 1.0; B138A, Yonsei Engineering Research Complex, Yonsei University 134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea.
(67) ChemSilico; CSBBB − A new Log BB Predictor; http://www.chemsilico.com/CS_prBBB/BBBhome.html.
(68) Cerius^2, ver 4.8.1; Accelrys, Inc., 10188 Telesis Court, Suite 100, San Diego, CA 92121, U.S.A.
(69) Accelrys, Working with Descriptors. In *Cerius2 4.8.1 QSAR*; Accelrys Inc. 9685 Scranton Rd., San Diego, CA 92121-3752, 2003; p 144.

CI050303I