## **ORIGINAL ARTICLES**

Department of Pharmacy, Zhejiang University City College, Hangzhou<sup>1</sup>, College of Medicine, Zhejiang University, Hangzhou<sup>2</sup>, Chou Zhou Middle School, Yiwu<sup>3</sup>, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou<sup>4</sup>, P.R. China

# A simple predictive model for blood-brain barrier penetration

X. C. Fu<sup>1</sup>, Z. F. Song <sup>2</sup>, C. Y. Fu<sup>3</sup>, W. Q. LIANG<sup>4</sup>

Received May 24, 2004; accepted June 6, 2004

Prof. Xuchun Fu, Department of Pharmacy, Zhejiang University City College, Hangzhou, 310015, P.R. China Fuxc@zucc.edu.cn

Pharmazie 60: 354–358 (2005)

A simple two-descriptor model to predict blood-brain barrier penetration is derived from a training set of 79 compounds:  $\log BB = -13.31V^2 + 9.601V - 2.231PSA - 0.5290$  (n = 79, r<sup>2</sup> = 0.83) where log BB is the logarithm of the ratio of the steady-state concentration of the compound in the brain to in the blood, V (nm<sup>3</sup>) is the molecular volume, PSA (nm<sup>2</sup>) is the polar surface area which is defined as the sum of the van der Waals surface areas of oxygen atoms, nitrogen atoms, and attached hydrogen atoms in a molecule, n is the number of compounds, and r is the correlation coefficient. The model is validated by a leave-one-out procedure and an external test set (25 compounds). The results indicate that the model developed is statistically sound and is sufficiently reliable and robust for predictive use. The descriptors in the model can be easily computed and it is suitable for the rapid prediction of the blood-brain barrier penetration for a wide range of drug candidates.

#### 1. Introduction

In drug design it is important to determine whether a candidate molecule is capable of penetrating the bloodbrain barrier (BBB). Drugs that act in the central nervous system (CNS) need to cross the BBB to reach their molecular target. By contrast, for drugs with a peripheral target, little or no BBB penetration might be required in order to avoid or minimize CNS side effects. A common measure of the degree of BBB penetration is the ratio of the steady-state concentration of the drug molecule in the brain to in the blood, usually expressed as log(C<sub>brain/blood</sub>) or log BB. The experimental determination of log BB is a time-consuming, expensive, and difficult technique, requiring animal experiments and the synthesis of the test compounds, usually in radiolabeled form (Pardridrge and Mietus 1979; Young et al. 1988; Chikhale et al. 1994; Sveigaard and Dalgaard 2000). It is of considerable value to predict log BB values of compounds from their physicochemical parameters or, ideally, from their molecular structures.

Young et al. (1988) showed that log BB values of 20 H<sub>2</sub> receptor histamine antagonists were correlated with  $\Delta \log P$  (octanol-cyclohexane). van de Waterbeemd and Kansy (1992) examined the same series of 20 compounds and found a significant correlation between log BB and the cyclohexane-water partition coefficient when the molecular volume was included in the parameterization. They also found that log BB was correlated with polar surface area (PSA, defined as the sum of the van der Waals surface areas of oxygen atoms, nitrogen atoms, and attached hydrogen atoms in a molecule), but the model showed it to be poorly predictive when tested with compounds outside its training set (Calder and Ganellin 1994), suggesting that the structural diversity

of the 20 H<sub>2</sub> receptor histamine antagonists might be insufficient to develop a generally applicable model for predicting log BB. Thus Abraham et al. (1994) constructed a larger training set of 65 compounds and derived a correlation between log BB and solvato-chromatic parameters for 57 compounds (8 compounds were excluded as outliers). With a set of 57 compounds drawn from the Abraham training set mentioned above, Lombardo (1996), Norinder (1998), Clark (1999), and their co-workers developed the models for log BB prediction using calculated molecular structural parameters such as free energy of solvation in water,  $\Delta G_w^0$  (Lombardo et al. 1996), Molsurf parameters (Norinder et al. 1998), PSA, and calculated octanol-water partition coefficient, C log P or M log P (Clark 1999), respectively. More recently, a variety of models to predict BBB penetration for larger dataset has been developed (Luco 1999; Feher et al. 2000; Crivori et al. 2000; Kaznessis et al. 2001; Rose and Hall 2002; Ooms et al. 2002) using different descriptors such as the three-dimensional molecular field descriptors, electropological state indices, and so on. In summary, the BBB penetration of a compound is thought to be dependent on its hydrogen-bonding potential, lipophilicity and size. Weak hydrogen-bonding potential, high lipophilicity, and small size are favorable to BBB penetration.

In this paper, we derive a simple model for the prediction of log BB from a dataset of 79 compounds.

#### 2. Investigations and results

The dataset of 111 compounds and their corresponding log BB values is taken from the literature (Young et al. 1988; Abraham et al. 1994; Salminen et al. 1997; Greig et al. 1995; Abraham et al. 1995; Calder and Ganellin 1994; Kelder et al. 1999; Lombardo et al. 1996; von Spre-

Table 1:	Experimental and calculated log BB values for the
	training set compounds and their computed descrip-
	tors

## Table 1 (continued)

Con	npound	$V_{(nm^3)}$	PSA (nm <sup>2</sup> )	log BB			
		(	(1111)	Exp. <sup>a</sup>	Calc. <sup>b</sup>	Pred. <sup>c</sup>	
1		0.3097	0.9784	-1.42	-1.02	-0.99	
2	H <sub>2</sub> N-S-N-	0.1735	0.7807	-0.04	-	_	
3	-N_S_NH_N_O	0.5088	0.8774	-1.06	-1.05	-1.05	
4		0.3812	0.3011	0.49	0.52	0.53	
5		0.3828	0.0540	0.83	1.08	1.10	
6	NH2 2 S NH NH	0.3488	1.4402	-0.82	-	-	
7		0.3424	0.8425	-0.67	-0.68	-0.68	
8	Option of the second se	0.3169	0.8517	-0.66	-0.72	-0.73	
9		0.4313	0.8171	-0.12	-0.69	-0.71	
10	H2 S-	0.2418	0.7636	-0.18	-0.69	-0.71	
11	NH2 S-	0.2516	1.0403	-1.15	-1.28	-1.29	
12		0.3016	1.0698	-1.57	-1.23	-1.20	
13		0.3420	1.3859	-1.54	-1.89	-1.95	
14		0.3902	0.9170	-0.27	-0.86	-0.88	
15		0.3897	0.9412	-0.28	-0.91	-0.93	
16	C. C. NIL	0.3941	0.4831	-0.46	0.11	0.13	
17	C. C. Martin	0.4633	0.4442	-0.24	0.07	0.09	
18		0.3383	0.3815	-0.02	0.34	0.36	
19		0.4327	0.3664	0.69	0.32	0.30	
20	CN CO NH S	0.4219	0.3753	0.44	0.32	0.31	

Compound		V (nm <sup>3</sup> )	PSA	log BB			
		()	(IIII )	Exp. <sup>a</sup>	Calc. <sup>b</sup>	Pred.c	
21		0.4773	0.3608	0.14	0.22	0.22	
22		0.4654	0.5428	0.22	-0.15	-0.18	
23		0.4736	0.9747	-2.00	-1.14	-1.08	
24		0.5482	0.7260	-1.30	-0.89	-0.77	
25		0.2404	0.4206	0.11	0.07	0.07	
26	A A A A A A A A A A A A A A A A A A A	0.3875	0.8629	-1.12	-0.73	-0.72	
27		0.5010	0.8539	-0.73	-0.97	-0.99	
28	N Nel2	0.2415	0.9040	-1.17	-1.00	-0.99	
29		0.3882	0.8955	-1.23	-0.81	-0.79	
30	Br N	0.3562	0.7315	-2.15	-	-	
31	Butanone	0.1164	0.1998	-0.08	-0.04	-0.04	
32	Benzene	0.1147	0.0000	0.37	0.40	0.40	
33	3-Methylpentane	0.1597	0.0000	1.01	0.67	0.65	
34	3-Methylhexane	0.1828	0.0000	0.90	0.78	0.78	
35	2-Propanol	0.0989	0.2311	-0.15	-0.23	-0.23	
36	2-Methylpropanol	0.1223	0.2201	-0.17	-0.05	-0.04	
31	2-Methylpentane	0.1587	0.0000	1.04	0.67	0.65	
39	1.1.1-Trifluoro-2-chloroethane	0.1009	0.0000	0.08	0.30	0.05	
40	1,1,1-Trichloroethane	0.1237	0.0000	0.40	0.46	0.46	
41	Diethyl ether	0.1272	0.1052	0.00	0.24	0.25	
42	Enflurane	0.1446	0.0918	0.24	0.38	0.38	
43	Ethanol	0.0760	0.2421	-0.16	-0.42	-0.45	
44	Fluroxene	0.1311	0.1104	0.13	0.25	0.26	
45	Halothane	0.1273	0.0000	0.35	0.48	0.48	
40	Неране	0.1630	0.0000	0.81	0.80	0.79	
48	Isoflurane	0.1444	0.1003	0.42	0.36	0.35	
49	Methylcyclopentane	0.1460	0.0000	0.93	0.59	0.58	
50	Pentane	0.1388	0.0000	0.76	0.55	0.54	
51	Propanol	0.0995	0.2417	-0.16	-0.24	-0.25	
52	Propanone	0.0932	0.2201	-0.15	-0.24	-0.25	
53 	Teflurane	0.1141	0.0000	0.27	0.39	0.40	
54 55	Trichloroethene	0.1389	0.0000	0.37	0.55	0.55	
55 56	Acetylsalicylic acid	0.2048	0.6940	-0.50	-0.67	-0.68	
57	Valproic acid	0.2155	0.4233	-0.22	-0.02	-0.02	
58	Salicylic acid	0.1522	0.6312	-1.10	-0.78	-0.77	
59	p-Acetamidophenol	0.1817	0.5959	-0.31	-0.55	-0.56	
60	Chlorambucil	0.3575	0.4884	-1.70	-	-	

## **ORIGINAL ARTICLES**

#### Table 1 (continued)

Compound		V	PSA	log BB			
		(nm²)	(nm²)	Exp. <sup>a</sup>	Pred. <sup>c</sup>		
61	C C N Nh	0.2477	0.4004	-1.30	_	-	
62	, Ma	0.2051	0.4765	-1.40	-	-	
63		0.3696	0.6736	-0.43	-0.30	-0.30	
64		0.3624	0.4342	0.25	0.23	0.23	
65	NH_N	0.1936	0.2813	-0.30	0.20	0.22	
66		0.2164	0.1880	-0.06	0.51	0.52	
67	NH2 S	0.1560	0.4216	-0.42	-0.30	-0.29	
68	BUT NH	0.3755	0.4031	-0.16	0.30	0.32	
69		0.2763	0.4667	0.00	0.07	0.07	
70	O NH2	0.2858	0.6592	-0.34	-0.34	-0.34	
71		0.3981	0.7959	-0.30	-0.59	-0.60	
72	CI O	0.4053	1.0088	-1.34	-1.07	-1.06	
73		0.4124	1.2201	-1.82	-1.56	-1.53	
74	<i>7</i> 0	0.3774	0.0560	0.89	1.07	1.09	
75		0.3425	0.0839	0.99	1.01	1.01	
76		0.3619	0.3054	0.82	0.52	0.51	
77		0.3435	0.3384	1.03	0.44	0.42	

Table 1	(continued)
Iubic I	(continueu)

Compoun	d	V (7773)	PSA	log BB			
		(nm <sup>*</sup> )	(nm)	Exp. <sup>a</sup>	Calc. <sup>b</sup>	Pred.c	
78		0.2698	0.2965	1.64	_	-	
79		0.3373	0.4139	0.52	0.27	0.26	
80		0.3184	0.4533	0.39	0.17	0.16	
81		0.3379	0.2052	0.53	0.74	0.75	
82		0.4110	0.4138	0.40	0.25	0.24	
83		0.4774	0.8300	-0.78	-0.83	-0.83	
84		0.3254	0.5289	0.00	0.01	0.01	
85		0.4932	0.6306	-0.02	-0.44	-0.47	
86		0.5010	0.8453	-0.67	-0.95	-0.98	

<sup>a</sup> From references (Young et al. 1988; Abraham et al. 1994; Salminen et al. 1997; Greig et al. 1995; Abraham et al. 1995; Calder and Ganellin 1994; Kelder et al. 1999; Lombardo et al. 1996)

<sup>b</sup> Calculated from eq. (1)

<sup>c</sup> Predicted using the leave-one-out cross validation procedure

cher, et al. 1998; Yazdanian and Glynn 1998). These compounds are divided into a training set (86 compounds) and a test set (25 compounds). Molecular volumes and polar surface areas are selected as the structural descriptors to develop a predictive model for BBB penetration. These structural descriptors are obtained from the molecular conformations optimized using the semiempirical self-consistent field molecular orbital calculation AM1 method (Dewar et al. 1985) and the atomic radii used by Clark (1999). The model to predict blood-brain barrier penetration is derived on the training set using the stepwise multiple regression analysis and then cross-validated using leave-one-out procedure (Wold 1978) in which one compound is left out from the training set and predicted from the model based on the remaining data and tested on the external prediction.

The 86 compounds of the training set are listed in Table 1 along with their experimental log BB values.



Fig. 1: Relationship between experimental and calculated log BB values for the training set

Using PSA and V as regression variables, the following regression equation is obtained from the stepwise multiple regression analysis (including quadratic terms) for the 86 compounds,

 $\log BB = -13.31V^2 + 9.601V - 2.231PSA - 0.5290 \quad (1)$ n = 79  $r^2 = 0.83$   $q^2 = 0.82$ s = 0.31 F = 126

where n is the number of compounds, r is the correlation coefficient, q is the cross validation coefficient, s is the standard deviation, F is the Fisher F-statistic. Compounds 2, 6, 30, 60, 61, 62 and 78 are removed from above equation as outliers. The calculated log BB values for the training set are presented in Table 1 and the experimental and calculated log BB values are plotted in Fig. 1.

## 3. Discussion

### 3.1 The predictive model of blood-brain barrier penetration including only molecular volume and polar surface area

Eq. (1) displays good statistical significance. As shown in Table 1 and Fig. 1, the calculated log BB values are in good agreement with the respective experimental ones. The log BB value of a compound is correlated with its molecular size parabolically and its polar surface area inversely.

Because the polar surface area is a descriptor of hydrogen-bonding potential (Stenberg et al. 2001), eq. (1) indicates that the log BB of a compound is inversely correlated with its hydrogen-bonding capacity.

Eq. (1) shows the parabolic relation between log BB and molecular volume. The explicit descriptor for lipophilicity is absent from eq. (1) and the molecular volume terms in the equation represent a combination of the impacts of molecular size and lipophilicity on BBB penetration. Increasing molecular volume decreases molecular diffusion through a lipid membrane and therefore decreases log BB value. On the other hand, a bigger molecular volume also means higher lipophilicity which facilitates BBB penetration.

Table 2:	Experimer	tal and	l calculated	logBB	values for	the	test set	compounds	and	their of	computed	descri	ptors
----------	-----------	---------	--------------	-------	------------	-----	----------	-----------	-----	----------	----------	--------	-------

Compound		$V_{(nm^3)}$	PSA 1 $(nm^2)$		log BB				
		()	(IIII)	Exp. <sup>a</sup>	Pred. <sup>b</sup>	Pred. <sup>c</sup>	Pred. <sup>d</sup>		
87	Theophylline	0.1993	0.7688	-0.29	-0.86	-1.43	-0.512		
88	Caffeine	0.2253	0.6075	-0.06	-0.40	-1.03	-0.219		
89	Antipyrine	0.2357	0.2728	-0.10	0.39	-0.03	0.474		
90	Ibuprofen	0.2816	0.4133	-0.18	0.20	-0.09	-0.555		
91	Codeine	0.3596	0.4836	0.55	0.12	-0.75	0.271		
92	Pentobarbital	0.2822	0.8646	0.12	-0.81	-0.77	-0.191		
93	Alprazolam	0.3467	0.4675	0.04	0.16	-0.58	0.332		
94	Indomethacin	0.3988	0.7630	-1.26	-0.52	-1.07	-1.032		
95	Oxazepam	0.3072	0.6951	0.61	-0.39	-0.70	-0.476		
96	Hydroxyzine	0.4674	0.4264	0.39	0.10	-0.20	0.128		
97	Desipramine	0.3769	0.0932	1.20	0.99	0.77	0.426		
98	Midazolam	0.3677	0.3206	0.36	0.49	-0.02	0.400		
99	Verapamil	0.5994	0.6787	-0.70	-1.07	-1.32	-1.111		
100	Promazine	0.3607	0.0834	1.23	1.02	0.78	0.832		
101	Chlorpromazine	0.3788	0.0831	1.06	1.01	0.86	0.710		
102	Trifluoroperazine	0.3944	0.0948	1.44	0.98	0.70	0.459		
103	Thioridazine	0.4579	0.0698	0.24	0.92	0.89	1.062		
104	BCNU	0.2258	0.6703	-0.52	-0.54	-0.56	-0.570		
105	Phenserine	0.4191	0.4825	1.00	0.08	-0.23	0.230		
106	Physostigmine	0.3514	0.5167	0.08	0.05	-0.50	0.007		
107	Terbutylchlorambucil	0.4528	0.2624	1.00	0.50	0.28	-0.227		
108	Didanosine	0.2625	1.0139	-1.30	-1.19	-1.95	-0.816		
109	Zidovudine	0.2941	1.3735	-0.72	-1.92	-2.37	-1.024		
110	Nevirapine	0.3132	0.5732	0.00	-0.11	-0.95	0.285		
111	SB-222200	0.4817	0.4306	0.30	0.05	0.19	0.426		

<sup>a</sup> From references (Salminen et al. 1997; Greig et al. 1995; von Sprecher, et al. 1998; Yazdanian and Glynn 1998)

<sup>b</sup> Predicted from eq. (1) <sup>c</sup> Predicted from the model developed by Feher et al. (2000)

<sup>d</sup> Predicted from the model developed by Luco (1999)



Fig. 2: Relationship between experimental and predicted log BB values for the test set

#### 3.2 Model validation using the leave-one-out procedure

The predictive model, eq. (1), is validated using leave-oneout procedure. Its cross validation coefficient ( $q^2 = 0.82$ ) is almost the same as its correlation coefficient ( $r^2 = 0.83$ ). The predicted values using the leave-one-out cross validation procedure (shown in Table 1) are also very close to the respective calculated values from eq. (1). The predictive model appears to be reliable and robust.

#### 3.3 Model validation using test set outside the training set

In order to further assess the predictive power of eq. (1), a test set of log BB values are predicted. The experimental and predicted log BB values are listed in Table 2 and plotted in Fig. 2.

As may be seen from Table 2 and Fig. 2, the predicted log BB values from eq. (1) are in good agreement with the respective experimental ones and only four compounds (92, 95, 105, and 109) are predicted above or near three standard deviations. The RMSE value calculated on the 25 validation compounds is 0.53. Considering the experimental difficulties and the varied experimental conditions under which the log BB values have been obtained, the predictive model for BBB penetration containing only molecular volume and polar surface area performs reasonably well.

As shown in Table 2, these prediction results are superior to the one obtained by the model reported by Feher et al. (2000) (RMSE = 0.79) and as good as the three-component model based on 25 descriptors using the multivariate partial least-squares procedure (Luco 1999) (RMSE = 0.54). However, our model is much simpler than the three-component model (Luco 1999), and thus more suitable for the rapid prediction of the BBB penetration for a wide range of drug candidates.

#### 3.4. Conclusion

The model derived in this paper for the prediction of BBB penetration shows a good predictive power. It contains only

two descriptors, namely molecular volume and polar surface area which are easy to interpret and compute. The model appears to be very simple but robust and effective for predictive use, so it is suitable for the rapid prediction of the BBB penetration for a wide range of drug candidates.

#### References

- Abraham MH, Chadha HS, Michell RC (1994) Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain. J Pharm Sci 83: 1257–1268.
- Abraham MH, Chadha HS, Michell RC (1995) Hydrogen bonding. Part 36. Determination of blood brain distribution using octanol-water partition coefficients. Drug Des Discov 13: 123–131.
- Calder JAD, Ganellin ČR (1994) Predicting the brain-penetrating capability of histaminergic compounds. Drug Des Discov 11: 259–268.
- Chikhale EG, Ng KY, Burton PS, Borchardt RT (1994) Hydrogen bonding potential as a determinant of the in vitro and in situ blood-brain barrier permeability of peptides. Pharm Res 11: 412–419.
- Clark DE (1999) Rapid calculation of polar surface area and its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration. J Pharm Sci 88: 815–821.
- Crivori P, Cruciani G, Carrupt P, Testa B (2000) Predicting blood-brain barrier permeation from three-dimensional molecular structure. J Med Chem 43: 2204–2216.
- Dewar MJS, Zoebisch GE, Healy EF, Stewart JJP (1985) AM1: A new general purpose quantum mechanical molecular model. J Am Chem Soc 107: 3902–3909.
- Feher M, Sourial E, Schmidt JM (2000) A simple model for the prediction of blood-brain partitioning. Int J Pharm 201: 239–247.
- Greig NH, Brossi A, Pei XF, Ingram DK, Soncrant TT (1995). In: Greenwood, J. et al. (Eds.), New concepts of a blood-brain barrier. Plenum, New York, pp. 251–264.
- Kaznessis YN, Snow ME, Blankley CJ (2001) Prediction of blood-brain partioning using Monte Carlo simulations of molecules in water. J Comput Aid Mol Des 15: 697–708.
- Kelder J, Grootenhuis PDJ, Bayada DM, Delbressine LPC, Ploemen JP (1999) Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. Pharm Res 16: 1514–1519.
- Lombardo F, Blake JF, Curatolo WJ (1996) Computation of brain-blood partitioning of organic solutes via free energy calculation. J Med Chem 39: 4750–4755.
- Luco JM (1999) 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 Comp Sci 39: 396–404.
- Norinder U, Sjoberg P, Osterberg T (1998) Theoretical calculation and prediction of brain-blood partitioning of organic solutes using Molsurf parameterization and PLS statistics. J Pharm Sci 87: 952–959.
- Ooms F, Weber T, Carrupt PA, Testa B (2002) A simple model to predict blood-brain barrier permeation from 3D molecular fields. Biochim Biophys Acta 1587: 118–125.
- Pardridrge WM, Mietus LJ (1979) Transport of steroid hormones through the rat blood-brain barrier. J Clin Invest 64: 145–154.
- Rose K, Hall LH (2002) Modeling blood-brain barrier partitioning using the electropological state. J Chem Inf Comput Sci 42: 651–666.
- Salminen T, Pulli A, Taskinen J (1997) Relationship between immobilized artificial membrane chromatographic retention and the brain penetration of structurally diverse drug. J Pharm Biomed Anal 15: 469–477.
- Stenberg P, Norinder U, Luthman K, Artursson P (2001) Experimental and computational screening models for the prediction of intestinal drug absorption. J Med Chem 44: 1927–1937.
- Sveigaard HH, Dalgaard L (2000) Evaluation of blood-brain barrier passage of a muscarine M1 agonist and a series of analogous tetrahydropyridines measured by in vivo microdialysis. Pharm Res 17: 70–76.
- van de Waterbeemd H, Kansy M (1992) Hydrogen bonding capacity and brain penetration. Chimia 46: 299–303.
- von Sprecher A, Gerpacher M, Anderson GP (1998) Neurokinin antagonists as potential therapies for inflammation and rheumatoid arthritis. Idrugs 1: 73–91.
- Wold S (1978) Cross-validatory estimation of the number of components in factor and principal component models. Technometrics 20: 397–406.
- Yazdanian M, Glynn SL (1998) In vitro blood-brain barrier permeability of nevirapine compared to other HIV antiretroviral agents. J Pharm Sci 87: 306–310.
- Young RC, Mitchell RC, Brown TH, Ganellin CR, Griffiths R, Jones M, Rana KK, Saunders D, Smith IR, Sore NE, Wilks TJ (1988) Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H<sub>2</sub> receptor histamine antagonists. J Med Chem 31: 656–671.