ISSN 0253-9756 Acta Pharmacol Sin 中国药理学报 2001 Jul; 22 (7): 663-668
http://www.chinaphar.com

Predicting blood-brain barrier penetration of drugs by polar molecular surface area and molecular volume

FU Xu-Chun[1], CHEN Chun-Xiao[2], LIANG Wen-Quan, YU Qing-Sen[3] (*College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310031*; [2]*First Affiliated Hospital of Zhejiang University, Hangzhou 310003*; [3]*Department of Chemistry, College of Science, Zhejiang University, Hangzhou 310027, China*)

KEY WORDS blood-brain barrier; molecular structure; statistical models; regression analysis

ABSTRACT

AIM: To predict the blood-brain barrier penetration by polar molecular surface area and molecular volume. **METHODS**: Polar molecular surface area and molecular volume are calculated by Monte Carlo method from the lowest energy conformation obtained using the semiempirical self-consistent field molecular orbital calculation AM1 method. The stepwise multiple regression analysis is used to derive the correlation equations between the ratios of the steady-state concentrations of the training compounds in the brain to in the blood (logBB) and their structural parameters. **RESULTS**: For a training set of 56 compounds, logBB values are well correlated with the sums of surface areas of oxygen and nitrogen atoms ($S_{O,N}$, A^2, excluding the nitrogen atoms in nitrogen molecule or in nitro) and molecular volumes (V, A^3). The regression equation is $\log BB = -1.331 \times 10^{-5} V^2 + 9.228 \times 10^{-3} V - 0.02439 S_{O,N} - 0.4318$ ($n = 56$, $r = 0.9043$). The calculated logBB values of a test set of 10 compounds from the model agree well with their experimental logBB values. **CONCLUSION**: The model is simple and effective. It can be used to predict the logBB values of candidate molecule in drug design.

INTRODUCTION

It is important to determine whether a candidate molecule is capable of penetrating the blood-brain barrier (BBB) in drug design. High penetration is needed for drugs targeted at the central nervous system (CNS), while low penetration may be desirable in order to minimize CNS-related side effects for drugs aimed at other sites of action. 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 $\log(C_{brain/blood})$ or logBB. The experimental determination of logBB is a time-consuming, expensive, and difficult technique, requiring the animal experiments and the synthesis of the test compounds (often in radiolabeled form)[1-4]. It is of very considerable value to predict logBB of drugs from their physicochemical parameters or, ideally, from their molecular structures. Young *et al*[2] showed that logBB values of 20 H_2 receptor histamine antagonists were correlated with $\Delta\log P$(octanol - cyclohexane). van de Waterbeemd *et al*[5] examined the same series of 20 compounds and found a significant correlation between logBB and the cyclohexane-water partition coefficient when the molecular volume was included in the parameterization. They also found that logBB was correlated with polar molecular surface area (PSA, defined as the area of its van de Waals surface that arises from oxygen or nitrogen atoms or hydrogen atoms attached to oxygen or nitrogen atoms)[5], but the model showed it to be poorly predictive when tested with compounds outside its training set[6], suggesting that the structural diversity of the 20 H_2 receptor histamine antagonists might be insufficient to develop a generally applicable model for predicting logBB. Thus Abraham *et al*[7] constructed a larger training set of 65 compounds and derived a correlation between logBB and solvatochromatic parameters for 57 compounds (8 compounds were excluded as outliers). With a set of 57 compounds drawn from the Abraham training set[7] mentioned above, Lombardo[8], Norinder[9], Clark[10], and their coworkers developed the models for logBB prediction using calculated molecular structural parameters such as free energy of solvation in water (ΔG_W^0)[8], Molsurf parameters[9], PSA, and octanol-water partition coefficient (Clog$P$ or Mlog$P$)[10] respectively. Recently, other two models for the prediction of blood-brain barrier pene-

[1] Correspondence to Prof FU Xu-Chun. Phn 86-571-721-7211. Fax 86-571-721-7412. E-mail fuxuchun@mail.hz.zj.cn
Received 2000-09-04 Accepted 2001-03-20

ISSN 0253-9756 Acta Pharmacol Sin 中国药理学报 2001 Jul; 22 (7)
E-mail aps@mail.shcnc.ac.cn Phn/Fax 86-21-6474-2029

tration from three-dimensional molecular structure[11] and from the calculated octanol-water partition coefficient, the number of hydrogen-bond acceptors, and the polar surface area[12] were reported. In this paper, we derive a new and simple model for the prediction of logBB.

METHODS

The molecular geometries of the compounds studied in this paper are optimized using the semiempirical self-consistent field molecular orbital calculation AM1 method[13]. Polar molecular surface areas and molecular volumes are calculated by Monte Carlo method[14] and the atomic radii used to calculate polar molecular surface areas and molecular volumes are same with those Clark[10] used. The stepwise multiple regression analysis is used to obtain the correlation equations between the logBB values of training compounds and their structural parameters.

RESULTS

The 57 compounds, previously studied by Lombardo *et al*[8] and other researchers[9,10] are illustrated in Fig 1 and listed in Tab 1 along with experimental logBB values taken from reference 8.

The predicted logBB values for a test set are listed in Tab 2. Compounds 58 – 63 were also predicted as an external test set by other groups[8-10] and compounds 64 – 67 were reported recently[4].

Using PSA and V as regression variables, the following regression equation is obtained from the stepwise multiple regression analysis (including square and product terms) for the 57 compounds.

$$\log BB = -1.037 \times 10^{-5} V^2 + 6.735 \times 10^{-3} V - 0.01808PSA - 0.1581$$

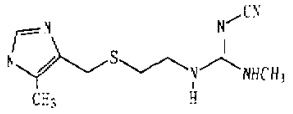
$$n = 56,\ r = 0.8732,\ s = 0.3733,\ F = 55.62 \qquad (1)$$

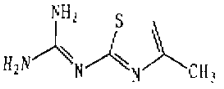
where n is the number of compounds, r is the correlation

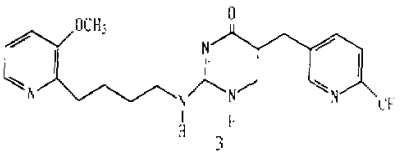
Tab 1. logBB values of 57 compouds and their structurl parameters

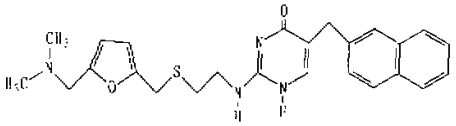
Compound	V	$S_{O,N}$	logBB Exp[8]	logBB Calc[1]	Compound	V	$S_{O,N}$	logBB Exp[8]	logBB Calc[1]
1	309.3	84.8	−1.42	−0.92	Butanone	116.5	20.2	−0.08	−0.03
2	173.6	63.3	−0.04	−0.78	Benzene	114.5	0.0	0.37	0.45
3	472.5	89.6	−2.00	−1.23	3-Methylpentane	159.7	0.0	1.01	0.70
4	547.5	70.1	−1.30	−1.08	3-Methylhexane	182.7	0.0	0.90	0.81
5	508.0	78.5	−1.06	−1.09	2-Propanol	99.3	18.7	−0.15	−0.10
6	242.2	33.8	0.11	0.20	2-Methylpropanol	122.1	17.9	−0.17	0.06
7	385.3	31.6	0.49	0.38	2-Methylpentane	161.2	0.0	0.97	0.71
8	383.8	8.3	0.83	0.95	2,2-Dimethylbutane	159.0	0.0	1.04	0.70
9	388.1	73.8	−1.23	−0.65	CF_3CH_2Cl	100.9	0.0	0.08	0.36
10	350.5	118.3	−0.82	−1.72	CH_3CCl_3	123.8	0.0	0.40	0.51
11	240.3	77.6	−1.17	−0.88	Diethyl ether	127.6	11.0	0.00	0.26
12	357.9	71.4	−2.15	–	CHF_2OCF_2CHFCl	144.9	9.9	0.24	0.39
13	342.6	74.3	−0.67	−0.64	Ethanol	76.0	19.7	−0.16	−0.29
14	316.8	74.6	−0.66	−0.66	$CF_3CH_2OCH=CH_2$	131.1	10.8	0.13	0.29
15	433.8	70.8	−0.12	−0.66	$CF_3CHClBr$	127.2	0.0	0.35	0.53
16	241.9	61.8	−0.18	−0.48	Heptane	185.8	0.0	0.81	0.82
17	251.6	81.1	−1.15	−0.93	Hexane	162.4	0.0	0.80	0.72
18	301.5	88.9	−1.57	−1.03	$CHF_2OCHClCF_3$	144.2	9.9	0.42	0.38
19	342.0	115.1	−1.54	−1.64	Methane	44.2	0.0	0.04	−0.05
20	387.6	77.5	−1.12	−0.74	Methylcyclopentane	145.9	0.0	0.93	0.63
21	502.5	72.1	−0.73	−0.91	Nitrogen	39.4	0.0	0.03	−0.09
22	382.6	71.6	−0.27	−0.60	Pentane	138.7	0.0	0.36	0.59
23	398.8	76.8	−0.28	−0.74	Propanol	99.4	19.5	−0.16	−0.12
24	397.5	46.1	−0.46	0.01	Propanone	93.1	22.1	−0.15	−0.23
25	466.4	40.6	−0.24	−0.01	CF_3CHFBr	114.1	0.0	0.27	0.45
26	343.3	34.3	−0.02	0.33	Toluene	137.5	0.0	0.37	0.59
27	437.3	32.6	0.69	0.26	Trichloroethane	113.5	0.0	0.34	0.44
28	425.7	36.2	0.44	0.20					
29	479.3	37.4	0.14	0.02					
30	467.3	46.1	0.22	−0.15					

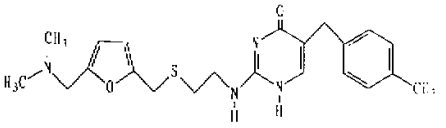
1) From equation (3)

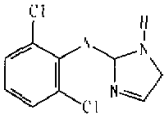

1

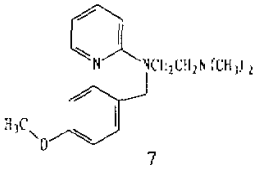

2

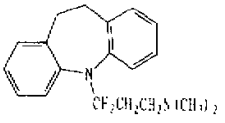

3

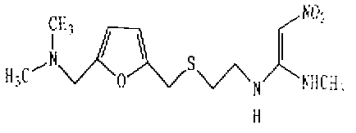

4

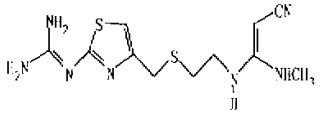

5

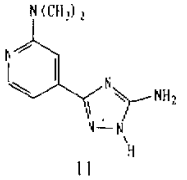

6

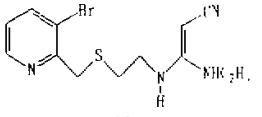

7

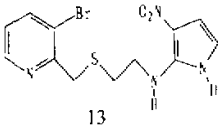

8

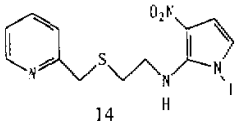

9

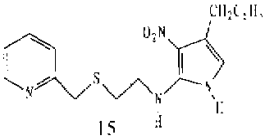

10

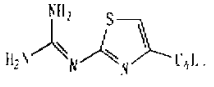

11

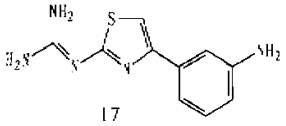

12

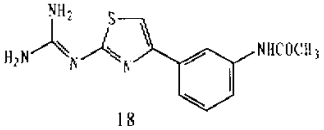

13

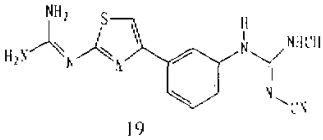

14

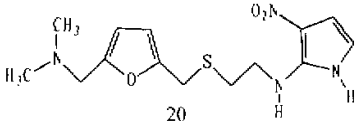

15

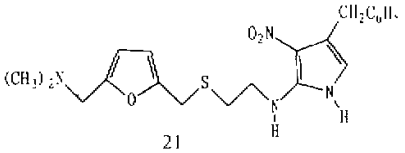

16

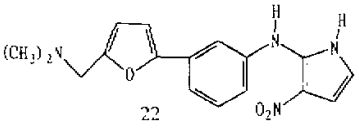

17

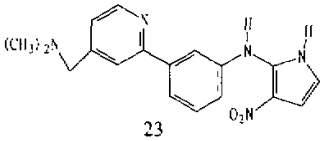

18

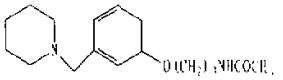

19

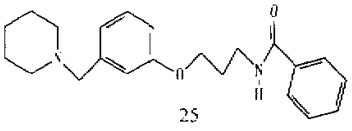

20

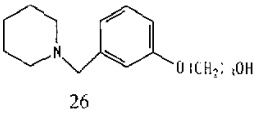

21

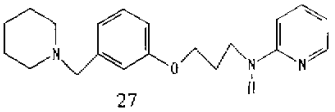

22


23


24


25


26


27

ISSN 0253-9756 Acta Pharmacol Sin 中国药理学报 2001 Jul; 22 (7)
E-mail aps@mail.shcnc.ac.cn Phn/Fax 86-21-6474-2629

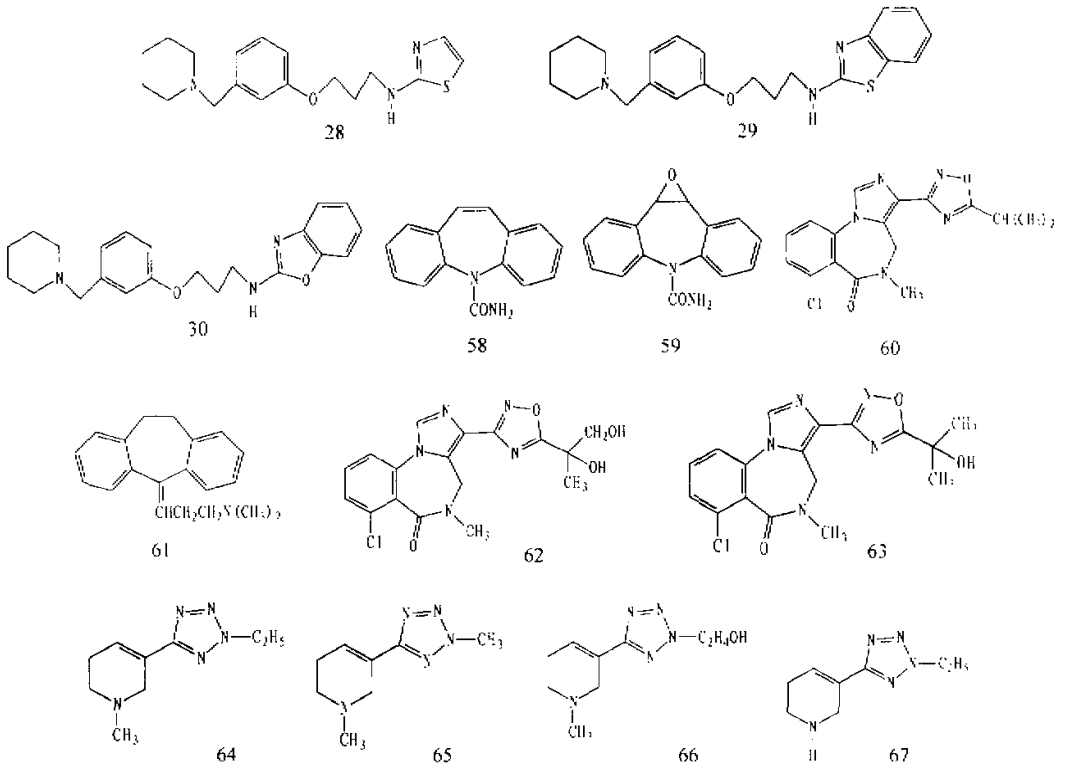


Fig 1. Compounds 1 – 30 and 58 – 67.

Tab 2. Predicted logBB values for a test set.

Compound	V	$S_{O,N}$	logBB Exp[8,4]	This work	Lombardo[8]	Norinder[9]	Clark-1[10]	Clark-2[10]
58	276.8	39.3	0.00	0.14	−0.14	−0.58	−0.25	−0.01
59	284.7	54.7	−0.34	−0.22	−0.28	−1.11	−0.75	−0.37
60	398.8	80.5	−0.30	−0.83	−0.46	−0.75	−0.70	−0.38
61	404.5	97.6	−1.34	−1.26	−0.64	−0.99	−1.26	−0.83
62	413.1	114.1	−1.82	−1.67	−0.82	−1.35	−1.77	−1.28
63	381.3	5.3	0.76 – 0.98	1.02	0.28	1.03	0.76	0.80
64	258.0	53.3	−0.51	−0.24				
65	234.0	54.8	−0.41	−0.34				
66	265.9	73.1	−0.48	−0.70				
67	231.7	59.6	−0.37	−0.46				
MAE[1]				0.20	0.41	0.56	0.24	0.23

1) The mean absolute errors over compounds 58 – 62.

coefficient, s is the standard error, F is the Fisher value. As other groups[7-10] have found, compound 12 is an outlier which is omitted from the original set of 57.

Two more significant regression equations are obtained when PSA in equation (1) is replaced with other parameters relevant to the polar molecular surface areas.

$$\log BB = -1.390 \times 10^{-5} V^2 + 9.181 \times 10^{-3} V - 0.01996SA - 0.4274$$
$$n = 56,\ r = 0.8951,\ s = 0.3415,\ F = 69.85 \quad (2)$$

$$\log BB = -1.331 \times 10^{-5} V^2 + 9.228 \times 10^{-3} V - 0.02439S_{O,N} - 0.4318$$
$$n = 56,\ r = 0.9043,\ s = 0.3269,\ F = 77.78 \quad (3)$$

here SA is PSA minus the surface areas of the nitrogen atoms in nitrogen molecule or in nitro and $S_{O,N}$ is the sum of surface areas of oxygen atoms and the nitrogen atoms except those in nitrogen molecule or in nitro.

DISCUSION

The BBB penetration of a compound is generally believed to be dependent on its hydrogen bond potential, lipophilicity and size[7-10,12,15]. Weak hydrogen bond potential, high lipophilicity, and small size are favorable to BBB penetration. Because the nitrogen atoms in nitrogen molecule or in nitro are difficult to form hydrogen bonds with hydrogen bond donors, maybe SA is a more appropriate parameter indicating the capacity of a molecule to form hydrogen bonds than PSA and better correlated with logBB.

Among the atoms involved in SA, hydrogen atom is hydrogen bond donor whereas oxygen and nitrogen atoms are hydrogen bond acceptors. When we divide SA into $S_{O,N}$ and S_H which is the sum of the surface areas of hydrogen atoms attached to oxygen or nitrogen atoms and relate them and V with logBB, we obtain the equation (3) in which there is no S_H term. The fact shows that oxygen and nitrogen atoms are more important in logBB prediction than the hydrogen atoms.

Equations (1) – (3) indicate that there is a parabolic correlation between logBB and V. This comes from the dual function of molecular size on BBB penetration. Increasing V decreases molecular diffusion through a lipid membrane and therefore decreases logBB value. On the other hand, bigger molecular volume also means higher lipophilicity which facilitates BBB penetration when polar surface areas remain unchanged. The results in Tab 2 shows that calculated logBB values of the compounds from equation (3) agree well with their experimental logBB values. Equation (3) performs as well as other models[8-10] or even better than some of them. Furthermore, the parameters used in equation (3) can be obtained very easily. It is a simple and effective model for logBB prediction. However, it can't completely take the place of animal experiments and the logBB values obtained from a predictive model usually need the verification of animal experiments.

REFERENCES

1 Pardridrge WM, Mietus LJ. Transport of steroid hormones through the rat blood-brain barrier. J Clin Invest 1979; 64: 145-54.
2 Young RC, Mitchell RC, Brown TH, *et al*. Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H_2 receptor histamine antagonists. J Med Chem 1988; 31: 656-71.
3 Chikhale EG, Ng KY, Burton PS, Borchardt RT. Hydrogen bonding potential as a determinant of the *in vitro* and *in situ* blood-brain barrier permeability of peptides. Pharm Res 1994; 11: 412-9.
4 Sveigaard HH, Dalgaard L. Evaluation of blood-brain barrier passage of a muscarine M1 agonist and a series of analogous tetrahydropyridines measured by *in vivo* microdialysis. Pharm Res 2000; 17: 70-6.
5 van de Waterbeemd H, Kansy M. Hydrogen bonding capacity and brain penetration. Chimia 1992; 46: 299-303.
6 Calder JAD, Ganellin CR. Predicting the brain-penetraing capability of histaminergic compounds. Drug Des Discov 1994; 11: 259-68.
7 Abraham MH, Chadha HS, Mitchell RC. Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain. J Pharm Sci 1994; 83: 1257-68.
8 Lombardo F, Blake JF, Curatolo WJ. Computation of brain-blood partitioning of organic solutes via free energy calculation. J Med Chem 1996; 39: 4750-5.
9 Norinder U, Sjoberg P, Osterberg T. Theoretical calculation and prediction of brain-blood partitioning of organic solutes using Molsurf parameterization and PLS statistics. J Pharm Sci 1998; 87: 952-9.
10 Clark DE. 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.
11 Crivori P, Cruciani G, Carrupt PA, Testa B. Predicting blood-brain barrier permeation from three-dimensional molecular structure. J Med Chem 2000; 43: 2204-16.
12 Feher M, Sourial E, Schmidt JM. A simple model for the prediction of blood-brain partitioning. Int J Pharm 2000; 201: 239-47.
13 Dewar MJS, Zoebisch GE, Healy EF, Stewart JJP. AM1: A new general purpose quamtum mechanical molecular model. J Am Chem Soc 1985; 107: 3902-9.
14 Shang ZC, Yu QS, Lin RS. Monte Carlo simulation calculation of volume and surface area of molecule. Acta Phys Chim Sin 1997; 13: 1097-100.
15 Shah MV, Audus KL, Borchardt RT. The application of bovine brain microvessel endothelial-cell monolayers grown onto polycarbonate membranes *in vitro* to estimate the potential permeability of solutes through the blood-brain barrier. Pharm Res 1989; 6: 624-7.

用极性表面积和分子体积预测药物对血脑屏障的穿透性

傅旭春[1], 陈春晓[2], 梁文权, 俞庆森[3]

ISSN 0253-9756 Acta Pharmacol Sin 中国药理学报 2001 Jul; 22 (7)
E-mail aps@mail.shene.ac.cn Phn/Fax 86-21-6474-2629

(浙江大学湖滨校区药学院，杭州 310031；[2]浙江大学附属第一医院，杭州 310003；[3]浙江大学玉泉校区理学院化学系，杭州 310027，中国)

关键词 血脑屏障；分子结构；统计学模型；回归分析

目的：用极性表面积和分子体积预测药物对血脑屏障的穿透性. **方法**：用 Monte Carlo 法从半经验自洽场分子轨道 AM1 法得到最低能构型计算极性表面积和分子体积，用逐步多元回归法导出药物分子分别在脑组织和血液中的稳态浓度之比(logBB)和分子结构参数之间的回归方程式. **结果**：对于 56 个化合物组成的训练样本，logBB 与氧原子和氮原子(不包括氮分子和硝基中的氮原子)的表面积之和($S_{O,N}$, A^2)以及分子体积(V, A^3)具有较好的相关性，回归方程式为：$logBB = -1.331 \times 10^{-5} V^2 + 9.228 \times 10^{-3} V - 0.02439 S_{O,N} - 0.4318$ ($n = 56$, $r = 0.9043$). 对于 10 个化合物组成的预测样本，预测值与实验值相当符合. **结论**：本模型简单有效，在药物设计中可以用来预测候选药物的 logBB 值.

(责任编辑 吕 静)