High-Throughput Prediction of Blood−Brain Partitioning: A Thermodynamic Approach

György M. Keserü* and László Molnár

Computer Assisted Drug Discovery, Gedeon Richter Ltd., P.O. Box 27, H-1475 Budapest, Hungary

Received June 8, 2000

A high-throughput in silico screening tool for potentially CNS active compounds was developed on the basis of the correlation of solvation free energies and blood−brain partitioning ($\log(c_{brain}/c_{blood}) = \log BB$) data available from experimental sources. Utilizing a thermodynamic approach, solvation free energies were calculated by the fast and efficient generalized Born/surface area continuum solvation model, which enabled us to evaluate more than 10 compounds/min. Our training set involved a structurally diverse set of 55 compounds and yielded a function of $\log BB = 0.035G_{solv} + 0.2592$ ($r = 0.85$, standard error 0.37). Calculation of solvation free energies for 8700 CNS active compounds (CIPSLINE database) revealed that G_{solv} is higher than −50 kJ/mol for the 96% of these compounds which can be used as suitable criteria for the identification of compounds preferable for CNS penetration.

1. INTRODUCTION

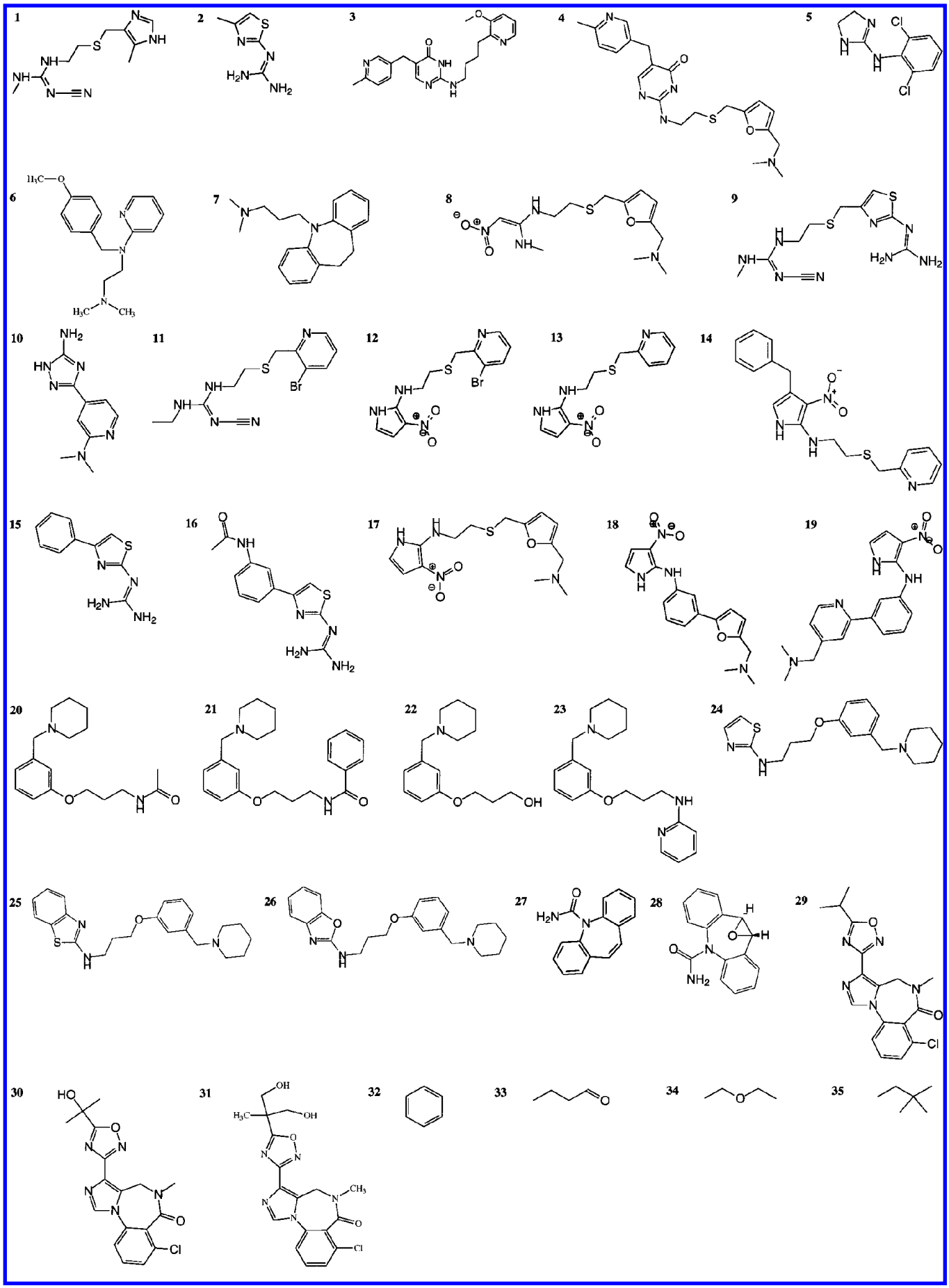
Analyzing the R&D activity of seven British pharma companies, Prentis et al. concluded that 66% of the new chemical entities (NCEs) were withdrawn between 1964 and 1985.[1] Dropped investigational new drug (IND) applications in the U.S. between 1964 and 1989 follow a similar trend; only one in six NCEs nominated to IND status became a marketed drug.[2] In addition to unproved efficiency, toxicity, and adverse reactions detected in humans, Lin and Lu demonstrated[3] that inadequate pharmacokinetic properties accounted for nearly 40% of the withdrawals of drugs from development. Penetration over the brain−blood barrier is one of the most critical pharmacokinetic issues in the design of CNS active drugs. Experimental methodology used for the measurement of log BB (=$\log(c_{brain}/c_{blood})$, however, is rather complicated and time-consuming, which prevents its application on the high-throughput scale. There is a clear trend in computer-aided drug design (CADD) that well-established, fast, and efficient CADD tools should be on-desk available for bench medicinal chemists. On the other hand, combinatorial technologies also require high-throughput CADD techniques. Identified needs at our company active on the CNS field prompt us to develop fast and effective in silico screens for CNS active compounds.[4]

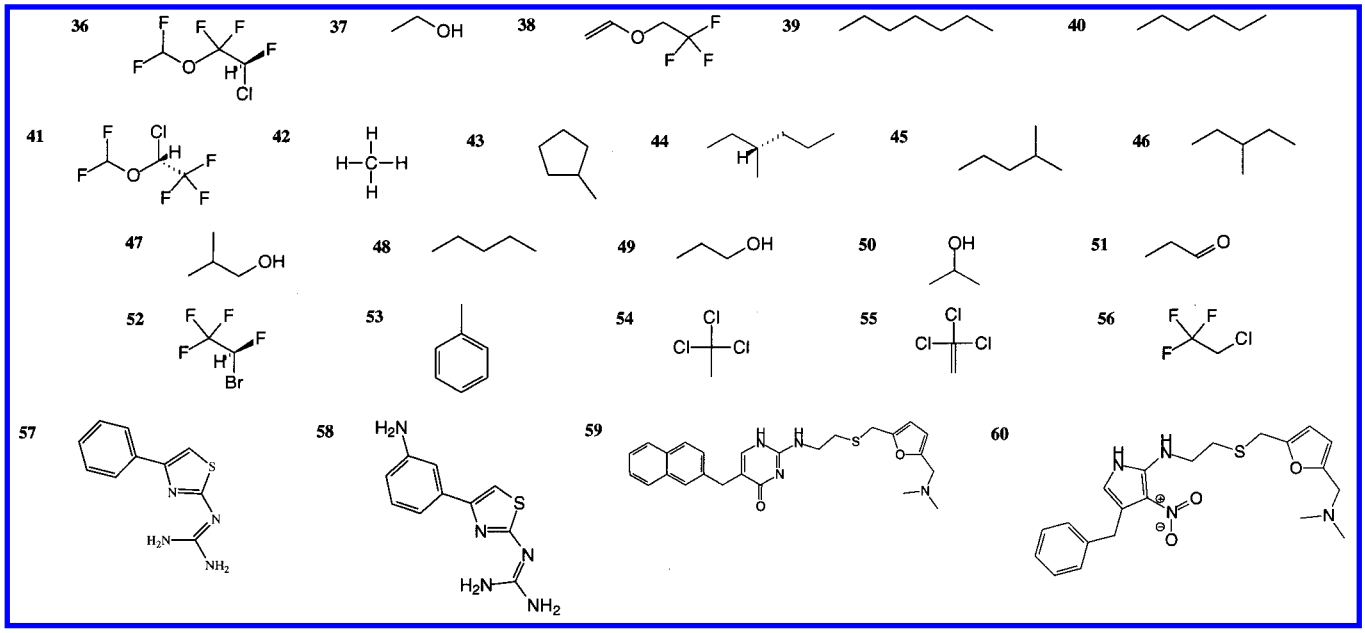
Early predictions of log BB involved classical QSAR approaches using various physicochemical parameters such as the octanol−water partition coefficient ($\log P_{oct}$),[5] molecular size descriptors,[6,7] and solvation parameters.[8,9] A combination of molecular descriptors was used by Norinder et al., predicting log BB using lipophilicity, polarity, polarizability, and hydrogen-bonding parameters and partial least-squares (PLS) statistics.[10] Recently Luco demonstrated that topological and constitutional descriptors used in PLS correlate well with experimental log BB data.[11] Although the quantitative structure−property relationship (QSPR) methodology was found to be useful in the accurate prediction of log BB data for a relatively small set of compounds, calculation of molecular descriptors is often the speed-limiting step. On the other hand, standard errors associated with the calculation of descriptors are usually accumulated when the derived function predicts log BB. Calculation of the polar surface area (PSA) defined as the surface area (Å^2) occupied by nitrogen and oxygen atoms and polar hydrogens connected to these heteroatoms represents a novel approach. One of the first applications of this descriptor in QSPR was published by Kansy et al., combining PSA with molecular volumes to yield a function to calculate log BB in an acceptable accuracy ($r = 0.84$, $s = 0.45$).[12] On the basis of the good experiment in the prediction of human intestinal absorption published by Palm et al.,[13,14] PSA was declared as a dominating determinant for oral absorption and brain penetration. Calculating the dynamic polar surface area (dPSA) from all low-energy conformations of a particular compound, Kelder et al. gave an excellent correlation with experimental log BB data ($r = 0.91$, $F = 229$).[15] The protocol, however, includes an extensive conformational analysis for each molecule, which clearly prevents its application in virtual high-throughput screening. This drawback of the dPSA approach was overruled by Clark,[16] demonstrating that static PSA in combination with ClogP[17] is a good estimate of log BB ($r = 0.88$, $s = 0.35$, $F = 96$). Correlation was significantly weaker when only PSA was considered ($n = 55$, $r = 0.84$, $s = 0.41$, $F = 128.4$), and we could not obtain any correlation using ClogP alone ($r = 0.38$). In our hands, there was a significant set of compounds where ClogP could only be calculated by large errors or the calculation was practically impossible. Although ClogP parameters are available for many fragments, breaking molecules into smaller fragments requires considerable care, the additivity of fragmental contributions may be difficult to judge, and finally stereochemical and conformational issues are usually not considered in fragmental approaches. Recent situations can be summarized as we have a single descriptor, dynamic PSA, which performs well but slowly. On the other hand, there is a rapid method utilizing static

* To whom correspondence should be addressed. Phone: +36-1-4314605. Fax: +36-1-2623008. E-mail: gy.keseru@richter.hu.

10.1021/ci000043z CCC: $20.00
Published on Web 11/05/2000

Chart 1


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
28
29
30
31
32
33
34
35



PSA and ClogP with some limitations. All PSA-based approaches, however, suffer from a phenomenological problem: in spite of intuitive and speculative reasons, it is not clear why PSA seems to work. Looking for a single, rapidly available, and physically meaningful descriptor, we concluded that the thermodynamics of membrane penetration should be considered.

2. MATERIALS AND METHODS

2.1. Biological Data. Experimental log BB data for the 60 compounds used in the training set were taken from studies (Chart 1) reported by Young et al.[18] (**1**−**25**), VanBelle et al. (**26** and **27**),[19] Lin et al. (**28**−**30**),[20] and Abraham et al.[8] (**31**−**60** in Table 1). The test set including 8700 CNS active compounds was extracted from the CIPSLINE database.[21]

2.2. Calculation of Solvation Free Energies. For analysis of available implicit solvation models regarding their speed and accuracy,[22] we chose the generalized Born/surface area (GB/SA) continuum solvation model as a high-throughput-scale solution for solvation free energy calculation. In a comparative work, Edinger et al.[23] demonstrated that among continuum solvation models the GB/SA model displays qualitatively superior performance on reproducing accurate solution of the Poisson−Boltzmann equation.

In the GB/SA model the solvation free energy (G_{solv}) can be approximated by the following equation:

$$G_{solv} = G_{cav} + G_{vdW} + G_{pol} \quad (1)$$

where G_{cav} is the solvent−solvent cavity term, G_{vdW} is the solute−solvent van der Waals term, and G_{pol} is the solute−solvent electrostatic polarization term.

The G_{pol} term can be approximated by the generalized Born equation:

$$G_{pol} = -166.0\left(1 - \frac{1}{\epsilon}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{q_i q_j}{(r_{ij}^2 + \alpha_{ij}^2 e^{-Dij})^{0.5}} \quad (2)$$

where $\alpha_{ij} = (\alpha_i\alpha_j)^{0.5}$, $D_{ij} = r_{ij}^2/(2\alpha_{ij})^2$, r_{ij} is the distance between atoms i and j, α_i is the so-called Born radius of atom i, q_i and q_j are the charges on atoms i and j, respectively, and ϵ is the dielectric constant of the surrounding medium. For the calculation of Born radii, we used the method evaluated by Still and co-workers.[24] This analytical method uses a pairwise atom evaluation approach. The polarization energy for atom i, where all atoms of the molecule are electrically neutral except for i, which has a unit charge, and where the surrounding medium has a high dielectric constant ($1 - 1/\epsilon \approx 1.0$), is

$$G'_{pol,i} = \frac{G_{pol,i}}{1 - 1/\epsilon} = \frac{-166.0}{R_{vdW,i} + \phi + P_1} + \sum^{stretch}\frac{P_2 V_j}{r_{ij}^4} + \sum^{bend}\frac{P_3 V_j}{r_{ij}^4} + \sum^{nonbonded}\frac{P_4 V_j (CCF)}{r_{ij}^4} \quad (3)$$

where $G_{pol,i}$ = polarization energy of atom i, ϕ = dielectric offset,[22] r_{ij} = distance between atoms i and j, V_j = volume of atom j, $R_{vdW,i}$ = van der Waals radius of atom i, P_1 = single-atom scaling factor, $P_2 = 1$, 2 scaling factor, $P_3 = 1$, 3 scaling factor, $P_4 = 1$, ≥4 scaling factor, P_5 = soft cutoff parameter, and CCF is the close contact function for the 1, ≥4 interaction, where

$$CCF = 1.0 \quad \text{if} \quad \left(\frac{r_{ij}}{R_{vdW,i} + R_{vdW,j}}\right)^2 > \frac{1}{P_5}$$

otherwise

$$CCF = \left\{0.5\left[1.0 - \cos\left\{\left(\frac{r_{ij}}{R_{vdW,i} + R_{vdW,j}}\right)^2 P_5 \pi\right\}\right]\right\}^2$$

Assuming the same conditions concerning the charge on atom i and the dielectric constant of the medium, we can compute the approximate Born radius from the simplified Born equation:

$$\alpha_i = -(-166.0/G'_{pol,i}) \quad (5)$$

We used P_1, P_2, P_3, P_4, and P_5 parameters and OPLS[25] radii optimized previously.[26] For the dielectric offset param-

Table 1. Computed G_{solv} and Calculated vs Experimental log BB for Compounds 1−60

compd		G_{solv} (kJ/mol)	calcd log BB	exptl log BB
1		−26.75	−0.73	−1.42
2		−4.95	0.12	−0.04
3		−24.00	−0.63	−2.00
4		−27.54	−0.76	−1.06
5		−15.78	−0.31	0.11
6		6.92	0.58	0.49
7		12.51	0.80	0.83
8		−28.66	−0.81	−1.23
9		−33.03	−0.98	−0.82
10		−22.73	−0.58	−1.17
11		−26.58	−0.73	−2.15
12		−38.02	−1.17	−0.67
13		−34.05	−1.02	−0.66
14		−25.79	−0.70	−0.12
15		−2.46	0.21	−0.18
16		−36.76	−1.12	−1.57
17		−31.09	−0.90	−1.12
18		−20.55	−0.49	−0.27
19		−26.39	−0.72	−0.28
20		−17.43	−0.37	−0.46
21		−11.95	−0.16	−0.24
22		5.69	0.53	−0.02
23		3.55	0.45	0.69
24		5.81	0.54	0.44
25		9.05	0.66	0.14
26		10.51	0.72	0.22
27		−21.63	−0.53	0.00
28		−25.90	−0.70	−0.34
29		−32.29	−0.95	−0.30
30		−37.95	−1.17	−1.34
31		−44.50	−1.43	−1.82
32	benzene	6.91	0.58	0.37
33	butanone	−15.19	−0.28	−0.08
34	diethyl ether	7.86	0.62	0.00
35	2,2-dimethylbutane	10.67	0.73	1.04
36	enflurane	−13.54	−0.22	0.24
37	ethanol	−1.71	0.24	−0.16
38	fluroxene	0.86	0.34	0.13
39	heptane	11.08	0.74	0.81
40	hexane	10.22	0.71	0.80
41	isoflurane	−6.08	0.07	0.42
42	methane	4.11	0.47	0.04
43	methylcyclopentane	8.38	0.64	0.93
44	3-methylhexane	11.31	0.75	0.90
45	2-methylpentane	10.59	0.72	0.97
46	3-methylpentane	10.34	0.71	1.01
47	2-methylpropanol	1.68	0.38	−0.17
48	pentane	9.37	0.67	0.76
49	propanol	−0.90	0.27	−0.16
50	2-propanol	0.68	0.34	−0.15
51	propanone	−16.24	−0.32	−0.15
52	teflurane	−16.83	−0.35	0.27
53	toluene	8.43	0.64	0.37
54	1,1,1-trichloroethane	2.28	0.40	0.40
55	trichloroethene	−10.58	−0.10	0.34
56	1,1,1-trifluoro-2-chloroethane	−11.66	−0.15	0.08
57		−7.37	0.02	−1.15
58		−27.24	−0.75	−1.54
59		−16.57	−0.34	−1.30
60		−25.67	−0.69	−0.73

eter (ϕ), we used −0.09 as described in ref 24 for OPLS parameters and water as solvent.

To calculate the solvation free energy from eq 1, one has to consider the following approximation:

$$G_{cav} + G_{vdW} = \sum \sigma_k(SA_k) \tag{6}$$

where SA_k is the total solvent-accessible surface area of atoms of type k and σ_k is an empirical atomic solvation parameter. This approach can be applied because of the evidence that the solvation free energy of saturated hydrocarbons in water is linearly related to the solvent-accessible surface area. We used 7.2 as the value of σ_k for all atom types.[24] For the calculation of SA_k, we used the algorithm of Hasel et al.[27] to calculate one atom's surface with no overlap:

$$S_i = 4\pi(r_i + r_s)^2 \tag{7}$$

where S_i = surface area for atom i, r_i = radius of atom i, and r_s = radius of a solvent molecule.

For one overlapping atom the excluded volume can be written as

$$b_{ij} = \pi(r_i + r_s)(r_i + r_j + 2r_s - r_{ij})\left(1 + \frac{r_j - r_i}{r_{ij}}\right) \tag{8}$$

where r_{ij} is the distance between the two atoms. For multiple overlaps the atomic surface with excluded volumes (A_i) can be written as

$$A_i = S_i \prod_{j=1}^{N}\left(1 - \frac{p_i p_{ij} b_{ij}}{S_i}\right) \tag{9}$$

where p_i = parameter depending on the type of atom i, p_{ij} = parameter depending on the connectivity of atom j to atom i, and N = number of atoms overlapping atom i. The total solvent-accessible surface can be evaluated by the following equation:

$$SA = \sum_{i=1}^{N} A_i \tag{10}$$

Our program written in C programming language can calculate the GB/SA solvation free energy on the basis of the equations above. The source code was compiled by MIPSpro Compilers, version 7.30, on an SGI machine and by egcs-1.1.2 release on a PC under Linux.

2.3. Protocol for High-Throughput Computation of log BB. To calculate the solvation free energy and predicted log BB of molecules, we used the following protocol. We started from a 2D MACCS file containing all molecules of interest. First we used CONCORD v4.0 to produce a 3D MACCS file. This file was imported into a SYBYL mdb database, which was minimized by the built-in MAXIMIN2 according to the following settings:

```
SETVAR TAILOR!MAXIMIN2!TERMINATION_OPTION GRADIENT
SETVAR TAILOR!MAXIMIN2!MINIMIZATION_METHOD CONJUGATE_GRADIENT
SETVAR TAILOR!MAXIMIN2!MAXIMUM_ITERATIONS 400
SETVAR TAILOR!MAXIMIN2!SIMPLEX_ITERATIONS 20
SETVAR TAILOR!MAXIMIN2!NON_BONDED_RESET 10
SETVAR TAILOR!MAXIMIN2!RMS_GRADIENT 0.050000
SETVAR TAILOR!FORCE_FIELD!FF_CHOICE TRIPOS
SETVAR TAILOR!FORCE_FIELD!PARAMETER_SET Tripos
SETVAR TAILOR!FORCE_FIELD!NON_BONDED_CUTOFF 8.000000
SETVAR TAILOR!FORCE_FIELD!DIELECTRIC_CONSTANT 1.000000
SETVAR TAILOR!FORCE_FIELD!DIELECTRIC_FUNCTION distance
CHARGE COMPUTE GAST_HUCK
```

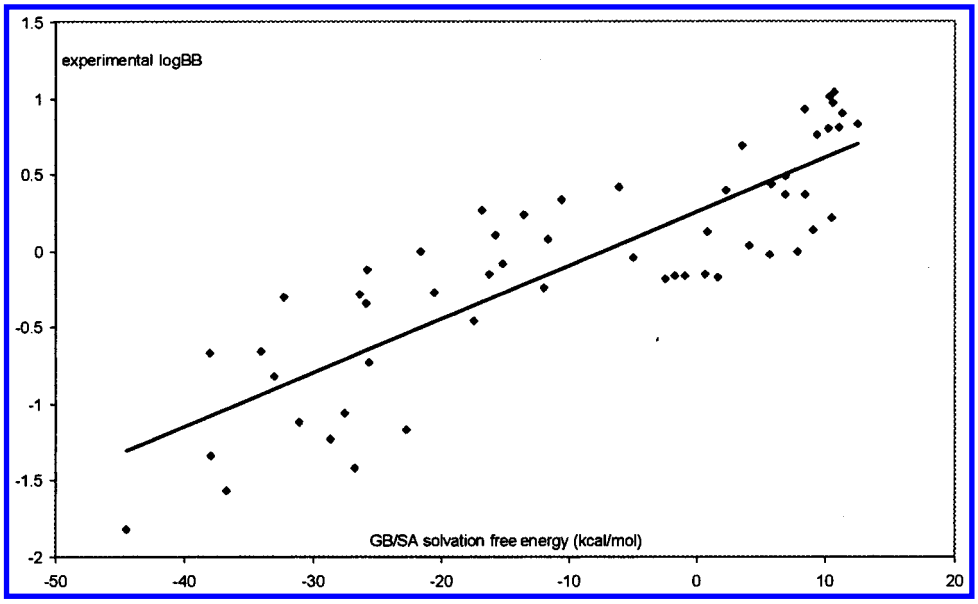


Figure 1. Relationship between the calculated GB/SA solvation free energy (G_{solv}) and the experimental log BB data obtained for the 55 compounds used in the training set.

Table 2. Comparison of Mean Absolute Errors of Prediction of Different log BB Prediction Methods Using Three Test Sets Published by Picket et al.[35]

		mean absolute errors of prediction[a]					
test set	no. of compds	Abraham[9]	Lombardo[29]	Norinder[10]	Luco[11]	RPR[16]	Richter/GSOLV
1[9]	7	0.30	NA	NA	0.42	0.37	0.16
2[29]	5	NA	0.41	0.52	0.29	0.24	0.14
3[11]	25	NA	NA	NA	0.43	0.50	0.37

[a] NA indicates that no results were presented for this set.

The resulting structures served as input into our GSOLV program. The final results were placed into a tab delimited text file which can be easily imported to a SYBYL spreadsheet to summarize the data.

3. RESULTS AND DISCUSSION

The brain−blood barrier, an important element in the regulation of the internal environment of the brain, is formed at the level of endothelial cells of the cerebral capillaries. Tight junctions between these cells as well as the lack of an aqueous pathway between cells restricts the movement of polar molecules across the cerebral endothelium.[28] In addition to specific transport mechanisms (P-glycoprotein- or receptor-mediated transport, peptide transporters, and other transport systems such as GLUT-1, system L1, and system ASC) passive diffusion is one of the most important ways to penetrate across the barrier. Analyzing the brain uptake of different drugs, it was concluded that activation energy is required to pluck a molecule out of the aqueous phase and into the lipid phase of the membrane;[28] i.e., molecules should be desoluted before entering the brain. Desolution can therefore be identified as the thermodynamic condition of the penetration across the brain−blood barrier. Since the hydrogen-bonding potential of a particular solute is in correlation with the activation energy, application of hydrogen-bonding descriptors in QSPR approaches is rationalized. PSA is proportional to the hydrogen-bonding capacity, and therefore, it can be used as a measure of the activation energy required for membrane penetration. Although these approaches have some physical rationale, calculation of the free energy of desolution is the only exact way from a thermodynamic point of view. The free energy of desolution (the negate of the solvation free energy, G_{solv}) can be calculated by several methods. Lombardo et al. calculated solvation free energies in water for 55 molecules by AMSOL 5.0 using the the AM1-SM2 solvation model.[29] Although calculated G_{solv} values correlate well with experimental log BB data ($r = 0.82$, $s = 0.42$), semiempirical calculations required for each molecule made this approach very time-consuming. Truhlar and co-workers recently published a faster model called OMNISOL[30] which allows the prediction of solvation free energies about 2 orders of magnitudes faster than AMSOL, but still slow enough to apply in a virtual HTS protocol. Our previous studies on solvated proteins demonstrated the effectiveness of the GB/SA continuum solvation model.[31,32] This good experience and several other successful applications[24,33] of the GB/SA model prompted us to use this approach for the rapid calculation of solvation free energies.

In the framework of the GB/SA model G_{solv} is calculated by eq 1, where the electrostatic term is approximated by $G_{GB} = G_{pol}$, while $G_{cav} + G_{vdw}$ is calculated by the surface area (SA) term (eq 6). Since PSA can be formulated as

$$\mathrm{PSA} = \sum_{i=\mathrm{O,N,OH,NH}} A_i \qquad (11)$$

where A_i stands for the atomic surface area of atom i with excluded volumes, neither an electrostatic contribution nor nonpolar interactions are considered in this approach. Furthermore, the polar surface area calculated in PSA-based

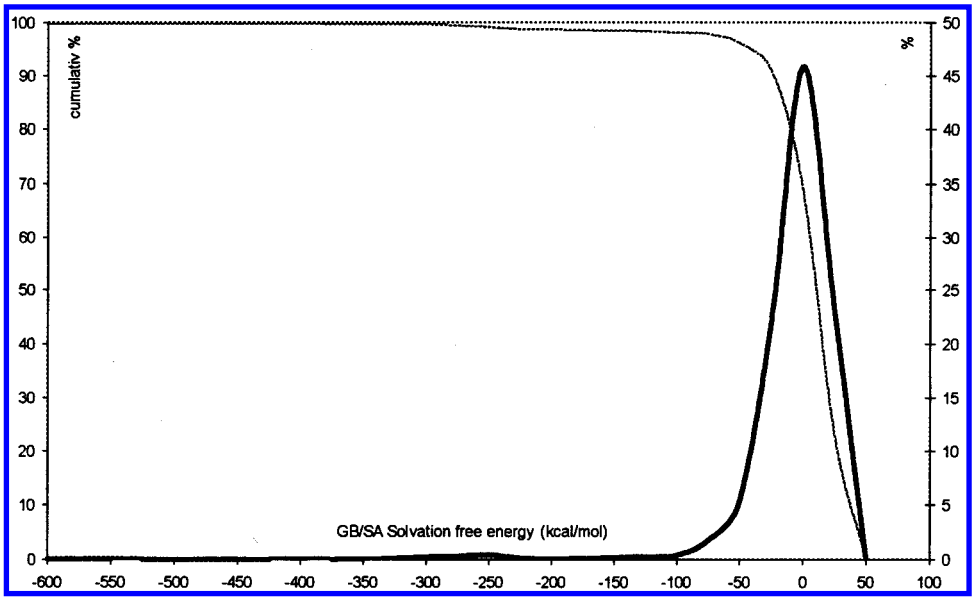


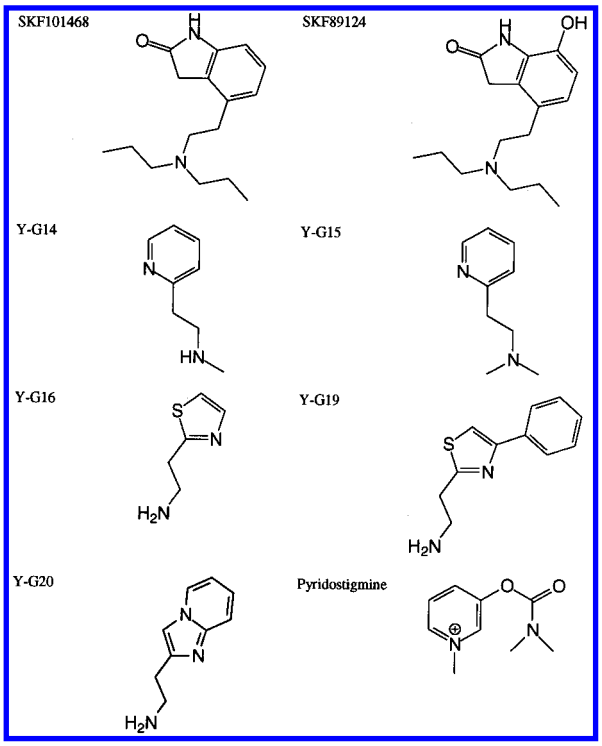
Figure 2. Distribution of 8700 CNS active compounds (CIPSLINE) concerning the GB/SA solvation free energy.

Table 3. Computed G_{solv} and Calculated vs Experimental log BB for the Test Set of Clark[16]

compd	G_{solv} (kJ/mol)	calcd log BB	exptl log BB
Y-G14	−12.83	−0.19	−0.30
Y-G15	−15.97	−0.30	−0.06
Y-G16	−23.12	−0.55	−0.42
Y-G19	−17.98	−0.37	−1.30
Y-G20	−12.83	−0.19	−1.40
SKF89124	−17.40	−0.35	−0.43
SKF101468	0.88	0.29	0.25
pyridostigmine	−298.48	−10.2	nd[a]

[a] Not detected in the brain.

Chart 2



predictions is the van der Waals surface of polar atoms and attached hydrogens, while in GB/SA we calculate the solvent accessible surface instead. In conclusion, PSA contains surface fragments which might have no impact on desolution, and therefore calculation of the SA term is highly preferred from a thermodynamic point of view. Although the van der Waals surface of nonpolar atoms is omitted in PSA, Stenberg et al. demonstrated that consideration of the nonpolar part of the surface made the correlation with experimental membrane permeability significantly better,[34] indicating this term to be significant with respect to the penetration. In contrast to the original PSA approach, the solvent accessible surface of nonpolar atoms is considered in the SA term of the GB/SA model. Most of the false positives identified by the PSA approach originate from the lack of the electrostatic term. Since G_{pol} is an important term in G_{solv}, the calculation of PSA is clearly not a thermodynamic approach.

On the basis of original publications of Still et al.,[22,23] we developed a C code (GSOLV) for the rapid calculation of G_{solv}. Calculations start from a 2D MACCS file which is converted to 3D MACCS using Concord. Resulting 3D structures are collected to a Sybyl database (MDB) and are minimized by MAXIMIN to yield an input for GSOLV. Application of GSOLV is limited by the atomic parameters currently available.[24] The method can be used for C-, N-, O-, S-, H-, P-, F-, Cl-, Br-, and I-containing organic molecules excluding organometallics and complexes. GSOLV produces a text output file which is inserted into the corresponding Sybyl spreadsheet (MSS) containing 3D structures.

Calculated solvation free energies for the 60 compounds used in this study are presented in Table 1. Five outliers (**3**, **11**, **57**−**59**) were removed to yield a training set of 55 compounds. It is interesting to note that **3** and **11** were also omitted by Abraham et al.,[9] Lombardo et al.[29] and Clark.[16] Compound **59** was identified as an outlier by Abraham et al.[8,9] Compounds **2** and **9**, structural analogues of **57** and **58**, were outliers in the analysis of Luco et al.[11] Since we predicted log BB for **2** and **9** correctly but overestimated experimental values of **57** and **58**, this suggests the influence of measurement uncertainties, active transporters, or metabolic factors. Other authors also rationalized these outliers

Chart 3

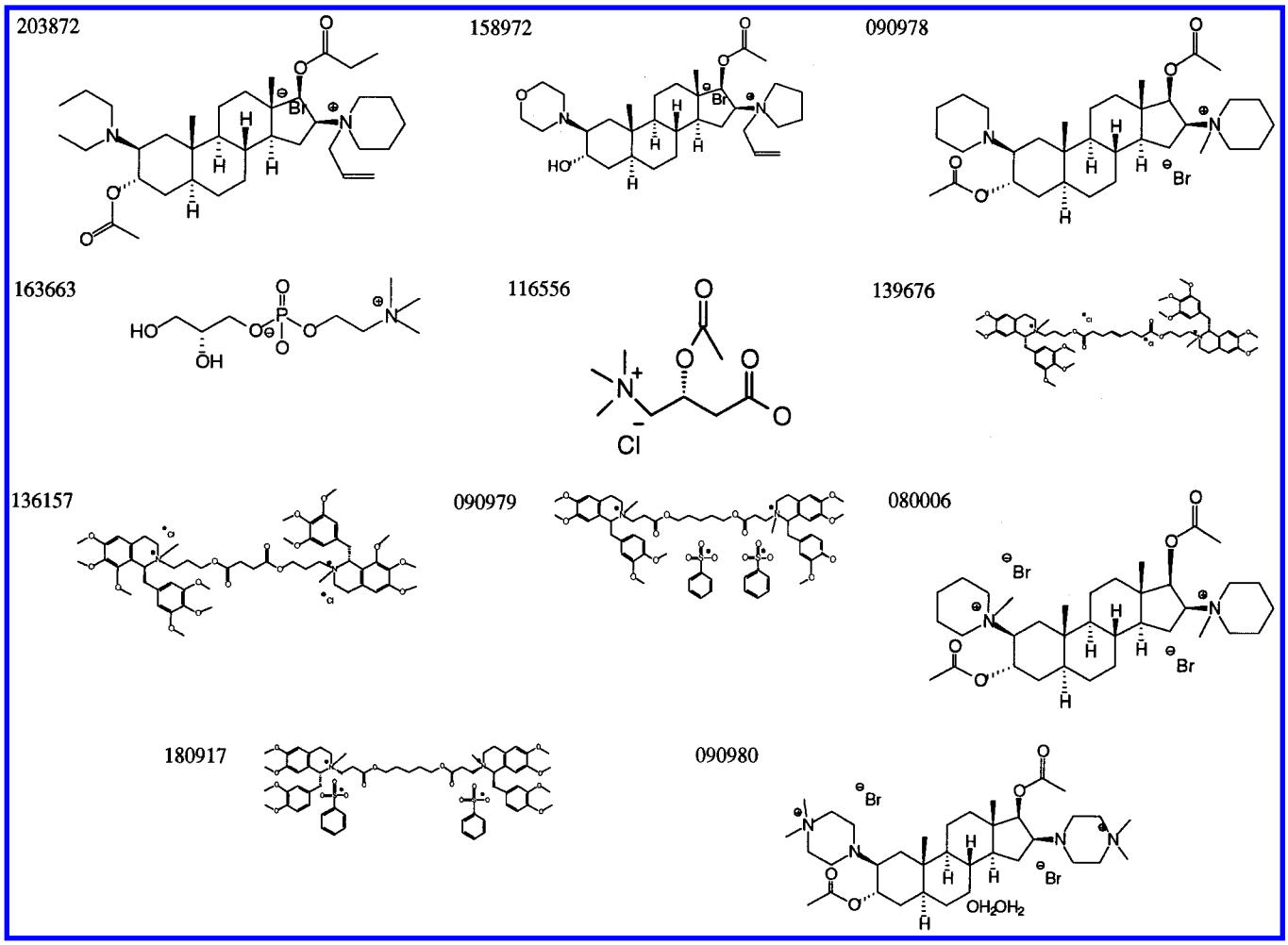


Table 4. Compounds with Extremely Low log BB in the CIPSLINE Database

Prous code	trade name	company	pharmacological activity	CAS RN	log BB
203872	Raplon	Organon	nondepolarizing neuromuscular blocker	156137-99-4	−9.30
158972	Zemuron	Organon	neuromuscular blocker	119302-91-9	−9.62
090978	Norcuron	Organon	nondepolarizing muscle relaxant	050700-72-6	−9.82
163663	Brezal	Novartis	cholinomimetic, treatment for age-associated cerebral involutive syndrome	028319-77-9	−10.99
116556	Carnitor	Glaxo Wellcome	coadjuvant in the treatment of involutive syndrome in cerebral vasculopathies	005080-50-2	−11.64
139676	Mivacron	Glaxo Wellcome	short-acting and nondepolarizing neuromuscular blocker	106861-44-3	−21.62
136157	Nuromax	Glaxo Wellcome	neuromuscular blocker, a long-acting, nondepolarizing relaxant	106791-39-3	−22.25
090979	Tracrium	Glaxo Wellcome	neuromuscular blocker	064228-81-5	−24.33
080006	Pavulon	Organon	nondepolarizing muscle relaxant	015500-66-0	−24.50
180917	Nimbex	Glaxo Wellcome	neuromuscular blocker	096946-42-8	−24.60
090980	Arduan	Gedeon Richter	long-acting, nondepolarizing neuromuscular blocker	068399-57-5	−25.32

due to nonaccounted metabolic effects and also to the fact that standard deviations of experimental log BB data might be as large as the mean value (see, e.g., ref 34).

Correlating calculated G_{solv} with experimental log BB, we obtained the following equation:

$$\log BB = 0.035G_{solv} + 0.2592 \qquad (12)$$

($r = 0.85$, standard error 0.37). The relationship between experimental log BB values and those computed by eq 12 is depicted in Figure 1. The prediction power of eq 12 was compared to those of other methods published in the literature (Table 2). Comparison of mean absolute errors of predictions of different log BB prediction methods using three test sets[35] demonstrated that our GSOLV approach gave the best prediction on these sets. Furthermore, it should be noted that GSOLV is as fast as the RPR code[16] and considers polarization notoriously neglected by others. Equation 12 was finally evaluated on a test set used by Clark[16] (Table 3). In agreement with Abraham et al.[9] and Clark,[16] we considered Y-G19 and Y-G20 (Chart 2) as outliers and calculated the mean absolute error to be 0.12. Reasonable mean errors obtained for all test sets suggest the predictive power of eq 12 to be sufficient for virtual HTS purposes. The most remarkable compound collected to the last test set (Table 3) is pyridostigmine, a reversible inhibitor of acetylcholinest-

erase used during the Persian Gulf War as prophylaxis against exposure to nerve gas (e.g., Soman). PSA calculations suggest the ability of pyridostigmine to penetrate across the blood−brain barrier (PSA ≈ 34 Å^2, log BB_{Clark} ≈ −0.25), while our thermodynamic approach indicated that its brain penetration is practically impossible (log BB = −10.2). Since it was suggested that emotional stress, as during the Gulf War, can make the blood−brain barrier to be permeable for this compound, it has been subjected to in-depth investigations. On the basis of the isotope studies of Lallement et al.,[36] no entry of pyridostigmine was detected into the CNS.

On the basis of experimental log BB values, it was concluded that compounds with log BB > 0.3 cross the barrier readily, while compounds with log BB < −1.0 are only poorly distributed to the brain. The linear relationship between brain penetration and the solvation free energy of 55 compounds (Figure 1) suggests that considering only passive diffusion orally administered CNS active drugs should have a solvation free energy higher than −40 kJ/mol. This conclusion was confirmed by calculating G_{solv} for 8700 CNS active compounds available from the CIPSLINE database[20] (Prous identification codes and/or CAS registry numbers are available upon request). Figure 2 shows that 96% of the CNS active molecules have G_{solv} higher than −50 kJ/mol, which can be used for the rapid identification of molecules that might penetrate to the brain via passive transport. Analysis of launched compounds revealed that there are only 11 compounds with extremely low log BB (Chart 3). Table 4 summarizes their pharmacological data and computed log BB values. It is interesting to note that most of these compounds are peripheral neuromuscular blockers having at least side effects on the CNS. Considering the excellent hit rate of this high-throughput virtual test, we concluded that its application might be beneficial in the early-phase CNS screening of large compound libraries.

4. CONCLUSION

On the basis of the thermodynamic condition of the passive transport across membranes, we demonstrated that brain penetration of molecules can be described as a function of solvation free energy. The equation derived from 55 molecules showed a good predictive ability. Rapid calculation of G_{solv} by the GB/SA continuum solvation model enabled us to develop a virtual screening tool on the high-throughput scale. Comparing the predictive power of Richter`s GSOLV to literature methods, we concluded that our approach outperforms other log BB prediction tools. We calculate log BB faster than 6 s/molecule; to the best of our knowledge, this is the fastest approach for the prediction of brain penetration.

REFERENCES AND NOTES

(1) Prentis, R. A.; Lis, Y.; Walker, S. R. Pharmaceutical innovation by seven UK-owned pharmaceutical companies (1964−1985). *Br. J. Clin. Pharmacol.* **1988**, *25*, 387−396.

(2) Caldwell, G. W. Compound optimization in early- and late-phase drug discovery: Acceptable pharmacokinetic properties utilizing combined physicochemical, in vitro and in vivo screens. *Curr. Opin. Drug Discuss.* **2000**, *3*, 30−41.

(3) Lin, J. H.; Lu, A. Y. Role of pharmacokinetics and metabolism in drug discovery and development. *Pharmacol. Rev.* **1997**, *49*, 403−449.

(4) Keserû, G. M.; Molnár, L.; Greiner, I. A neural network based high throughput screening test for CNS active compounds. *Comb. Chem. High Throughput Screening*, submitted for publication.

(5) Hansch, C.; Bjorkroth, J. P.; Leo, A. Hydrophobicity and central nervous system agents: On the principle of minimal hydrophobicity in drug design. *J. Pharm. Sci.* **1987**, *76*, 663−687.

(6) van Bree, J. B. M. M.; De Boer, A. G.; Danhof, M.; Ginsel, L. A.; Breimer, D. D. Characterization of an "in vitro" blood-brain barrier: Effects of molecular size and lipophilicity on cerebrovascular endothelial transport rates of drugs. *J. Pharmacol. Exp. Ther.* **1988**, *247*, 1233−1239.

(7) Kaliszan, R.; Markuszewski, M. Brain/blood distribution described by a combination of partition coefficients and molecular mass. *Int. J. Pharm.* **1996**, *45*, 9−16.

(8) Abraham, M. H.; Chadha, H. S.; Mitchel, R. C. Hydrogen bonding. 33. Factors that influence the distribution of solutes between blood and brain. *J. Pharm. Sci.* **1994**, *83*, 1257−1268.

(9) Abraham, M. H.; Chadha, H. S.; Mitchel, R. C. Hydrogen bonding. 36. Determination of blood brain distribution using octanol-water partition coefficients. *Drug Des. Discuss.* **1995**, *13*, 123−131.

(10) Norinder, U.; Sjoberg, P.; Ostenberg, 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.

(11) 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.

(12) Kansy, M.; van de Waterbemd, H. Hydrogen bonding capacity and brain penetration. *Chimia* **1992**, *46*, 299−303.

(13) Palm, K.; Stenberg, P.; Luthman, K.; Artursson, P. Polar molecular surface properties predict the intestinal absorption of drugs in human. *Pharm. Res.* **1997**, *14*, 568−571.

(14) Palm, K.; Luthman, K.; Ungell, A. L.; Strandlund, G.; Beigi, F.; Lundahl, P.; Artursson, P. Evaluation of dynamic polar molecular surface area as predictor of drug absorption: Comparison with other computational and experimental predictors. *J. Med. Chem.* **1998**, *41*, 5382−5392.

(15) Kelder, J.; Grootenhuis, P. D. J.; Bayada, D. M.; Delbressine, L. P. C.; Ploemen, J. P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* **1999**, *16*, 1514−1519.

(16) Clark, D. E. Rapid calculation of polar molecular surface area ands its application to the prediction of transport phenomena. 2. Prediction of blood-brain barrier penetration. *J. Pharm. Sci.* **1999**, *88*, 815−821.

(17) ClogP. Daylight Chemical Information Software, Daylight Chemical Information Inc., 27401 Los Altos, Mission Viejo, CA 92691.

(18) 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.; Wilks, T. J. Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H_2 receptor histamine antagonist. *J. Med. Chem.* **1988**, *31*, 656−671.

(19) VanBelle, K.; Sarre, S.; Ebinger, G.; Michotte, Y. Brain, liver and blood distribution kinetics of carbazepine and its metabolic interaction with clomipramine in rats: A quantitative microdialysis study. *J. Pharmacol. Exp. Ther.* **1995**, *272*, 1217−1222.

(20) Lin, J. H.; Chen, I. W.; Lin, T. H. Blood brain barrier permeability and in vivo activity of partial benzodiazepine receptor: A study of L-663,581 and its metabolites in rats. *J. Pharmacol. Exp. Ther.* **1994**, *271*, 1197−1202.

(21) CIPSLINE database (March 2000 release), Prous Science, Barcelona, Spain.

(22) Cramer, C. J.; Truhlar, D. G. Implicit solvation models: Equilibra, structure, spectra, and dynamics. *Chem. Rev.* **1999**, *99*, 2161−2200.

(23) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. Solvation free energies of peptides: Comparison of approximate continuum solvation models with accurate solution of the Poisson-Boltzmann equation. *J. Phys. Chem. B* **1997**, *101*, 1190−1197.

(24) Qui, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.

(25) Still, W. C.; Tempczyk, A.; Hawley, R.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(26) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(27) Hasel, W.; Hendrickson, T. F.; Still, W. C. A rapid approximation to the solvent accessible surface areas of atoms. *Tetrahedron Comput. Methodol.* **1988**, *1*, 103−116.

(28) Begley, D. J. The blood-brain barrier: Principles for targeting peptides and drugs to the central nervous system. *J. Pharm. Pharmacol.* **1996**, *48*, 136−140.

(29) 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*, 4750−4755.
(30) Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. Omnisol: Fast prediction of free energies of solvation and partition coefficients. *J. Org. Chem.* **1998**, *63*, 4305−4313.
(31) Keserü, G. M.; Kolossváry, I.; Székely, I. Inhibitors of cytochrome P450 catalyzed insecticide metabolism: A rational approach. *Int. J. Quantum Chem.* **1999**, *73*, 123−135.
(32) Keserü, G. M.; Menyhárd, D. K. Role of proximal His93 in nitric oxide binding to metmyoglobin. Application of continuum solvation in Monte Carlo protein simulations. *Biochemistry* **1999**, *38*, 6614−6622.
(33) Reddy, R. M.; Erion, M. D.; Agarwal, A.; Viswanadhan, Vellarkad, N.; McDonald, D. Q.; Still, W. C. Solvation free energies calculated using the GB/SA model: sensitivity of results on charge sets, protocols, and force fields. *J. Comput. Chem.* **1998**, *19*, 769−780.
(34) Stenberg, P.; Luthman, K.; Artursson, P. Prediction of membrane permeability to peptides from calculated dynamic surface properties. *Pharm. Res.* **1999**, *16*, 205−212.
(35) Pickett, S. D.; McLay, I. M.; Clark, D. E. Enhancing the Hit-to-Lead Properties of Lead Optimization Libraries. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 263−272.
(36) Lallement, G.; Foquin, A.; Baubichon, D.; Burckhart, M. F.; Carpentier, P.; Canini, F. Heat stress, even extreme, does not induce penetration of pyridostigmine into the brain of guinea pigs. *Neurotoxicology* **1998**, *19*, 759−766.

CI000043Z