In Silico Prediction of Blood−Brain Barrier Permeation Using the Calculated Molecular Cross-Sectional Area as Main Parameter

Grégori Gerebtzoff and Anna Seelig*

Biophysical Chemistry, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Received March 13, 2006

The cross-sectional area, A_D, of a compound oriented in an amphiphilic gradient such as the air−water or lipid−water interface has previously been shown to be crucial for membrane partitioning and permeation, respectively. Here, we developed an algorithm that determines the molecular axis of amphiphilicity and the cross-sectional area, A_{Dcalc}, perpendicular to this axis. Starting from the conformational ensemble of each molecule, the three-dimensional conformation selected as the membrane-binding conformation was the one with the smallest cross-sectional area, A_{DcalcM}, and the strongest amphiphilicity. The calculated, A_{DcalcM}, and the measured, A_D, cross-sectional areas correlated linearly ($n = 55$, slope, $m = 1.04$, determination coefficient, $r^2 = 0.95$). The calculated cross-sectional areas, A_{DcalcM}, were then used together with the calculated octanol−water distribution coefficients, log $D_{7.4}$, of the 55 compounds (with a known ability to permeate the blood−brain barrier) to establish a calibration diagram for the prediction of blood−brain barrier permeation. It yielded a limiting cross-sectional area ($A_{DcalcM} = 70$ Å^2) and an optimal range of octanol−water distribution coefficients ($-1.4 \leq \log D_{7.4} < 7.0$). The calibration diagram was validated with an independent set of 43 compounds with the known ability to permeate the blood−brain barrier, yielding a prediction accuracy of 86%. The incorrectly predicted compounds exhibited log $D_{7.4}$ values comprised between −0.6 and −1.4, suggesting that the limitation for log $D_{7.4}$ is less rigorous than the limitation for A_D. An accuracy of 83% has been obtained for a second validation set of 42 compounds which were previously shown to be difficult to predict. The calculated parameters, A_{DcalcM} and log $D_{7.4}$, thus allow for a fast and accurate prediction of blood−brain barrier permeation. Analogous calibration diagrams can be established for other membrane barriers.

INTRODUCTION

Brain-targeted drugs have to permeate the blood−brain barrier (BBB) to be therapeutically active, whereas drugs targeted to peripheral sites should ideally not reach the central nervous system (CNS) not to induce central side effects. Finding simple and unambiguous criteria for membrane permeation that can be used for structural optimization of drug candidates in the early stage of drug discovery would therefore be desirable.

To assess BBB permeation, in vivo models have been used for decades, and more recently, many in vitro assays based on confluent cell monolayers have been established (e.g., ref 1). The parameters obtained from such assays are complex because they reflect the resultant of different active and passive transport processes and can therefore not be easily translated into simple molecular parameters. Moreover, these assays are rather time-consuming and are thus not applicable to the early screening of a large number of compounds.

The most commonly used physical−chemical approaches to predict membrane permeation are octanol−water[2] or hexadecane−water[3] partition coefficient measurements, high-performance liquid chromatography (HPLC)-related techniques (ref 4 and references therein), and parallel artificial membrane permeation (PAMPA) assays.[5] Although predictions of membrane permeation on the basis of these techniques are generally successful for small molecules, they often fail for larger molecules.[4] This is due to the fact that the membrane mimicking systems are either fully isotropic (e.g., octanol and hexadecane) or exhibit a low anisotropy (e.g., HPLC or PAMPA systems). As shown by Lipinski et al.,[7] the prediction of membrane permeation can be improved by combining octanol−water partition coefficients with additional parameters such as the molecular weight and the number of H-bond acceptors and donors.

Lipid bilayers exhibit a fluidity which is comparable to that of olive oil. However, in contrast to the random, isotropic arrangement of the molecules in olive oil, the molecules in a lipid bilayer exhibit an average structural order relative to each other along their molecular axis and are thus highly anisotropic liquids or liquid crystals. Using solid-state deuterium−nuclear magnetic resonance, the packing of the fatty acyl chains was described in terms of statistical order profiles. For a given membrane, the average order of the fatty acyl chains is comparatively high, close to the head-group region, and decreases toward the center of the membrane (for a review, see refs 6 and 8).

The lateral packing density of the bilayer membrane, π_M, can also be assessed in comparison to that of a lipid monolayer.[9] It depends on the lipid composition and can vary considerably (for a review, see ref 10). Insertion into a lipid bilayer requires energy, ΔW, because a cavity has to be formed in the well-ordered membrane which is large enough to accommodate the drug. This energy, ΔW, is proportional

* Corresponding author phone: +41-61-267-22-06; fax: +41-61-267-21-89; e-mail: Anna.Seelig@unibas.ch.

10.1021/ci0600814 CCC: $33.50
Published on Web 10/04/2006

to the lateral packing density, π_M, of the membrane and the cross-sectional area, A_D, of the inserting molecule ($\Delta W = \pi_M \times A_D$).[11] For molecules with small cross-sectional areas, the energy is low; however, for molecules with large cross-sectional areas, it can become prohibitively high. It should be noted that the cross-sectional area, A_D, of the molecule is not necessarily proportional to the molecular weight but depends on the conformation and on the orientation of the molecule.[12] The relevance of the molecular cross-sectional area (rather than the molecular weight) was also demonstrated by measuring the passive diffusion of linear and branched molecules.[13]

The lipid−water partition coefficient, K_{lw}, of a molecule decreases exponentially with increasing energy of cavity formation.[12,14] To give a numerical example, doubling the cross-sectional area, A_D, of the molecule from 50 Å^2 to 100 Å^2 reduces the lipid−water partition coefficient for a membrane with a lateral packing density, $\pi_M = 35$ mN/m (e.g., BBB) by a factor of 67. This is in contrast to the octanol−water partition coefficient, *P*, which increases with increasing size of the molecule,[15] because the energy for cavity formation is relatively small.

Most membrane-permeating drugs are amphiphilic and, if brought into contact with the air−water or the lipid−water interface, they organize themselves in an anisotropic manner comparable to that of lipid molecules. Because the dielectric constant of air ($\epsilon = 1$) and that of the lipid core region ($\epsilon = 2$) are similar and much lower than that of water ($\epsilon = 80$), the amphiphilic orientation of the molecule is identical at the two interfaces. Measurement of the Gibbs adsorption isotherm, that is, the surface pressure of the drug in a buffer solution as a function of the concentration, yields the air−water partition coefficient, K_{aw}, the critical micelle concentration, CMC, and the surface area requirement of the compound, A_S, in its amphiphilic orientation. If measurements are performed under conditions of minimal charge repulsion, the surface area requirement, A_S, corresponds to the cross-sectional area, A_D, of the molecule perpendicular to its axis of amphiphilicity.[12]

The three parameters, K_{aw}, CMC, and A_D, have been used to establish 3D calibration diagrams for membrane permeation with high predictive values.[12] For BBB permeation, the calibration diagram was established with 53 drugs of known ability to cross that BBB. It revealed that permeation is only possible if a compound exhibits a cross-sectional area $A_D < 80$ Å^2 [the limit of 80 Å^2 corresponds to a rounded value from the experimental cross-sectional area of spiradoline (73 ± 5 Å^2)]; an intermediate air−water partition coefficient, K_{aw}; and an ionization constant p$K_a < 10$ for bases and p$K_a > 4$ for acids. An analogous analysis was also performed for the intestinal barrier.[16]

For fast screening of preclinical drug candidates, several in silico models have been developed (for a review, see refs 17−19). The parameters used most frequently are the molecular weight; the calculated octanol−water partition coefficient, *P*, of the neutral form of the compound (expressed as log *P*) or the partition coefficient of the salt form at pH 7.4; and the so-called distribution coefficient, *D* (expressed as log $D_{7.4}$). Further parameters are the number of nitrogen and oxygen atoms, the number of heteroatoms, and the number of hydrogen-bond donors and acceptors. These parameters can all be calculated on the basis of the structural formula. A second group of parameters also used frequently, such as the molecular volume, the molecular surface, the solvent-accessible surface area, the polar surface area, the molecular shape, and flexibility, requires the three-dimensional structure of the compound for calculation.

Because the cross-sectional area, A_D, measured at the air−water interface under conditions of minimal electrostatic repulsion[12] has been shown to be highly predictive, the aim was to implement this parameter in the in silico prediction of membrane permeation. So far, the cross-sectional area, A_D, of the molecule oriented at an amphiphilic interface has not been used as a parameter for the in silico prediction of membrane permeation. Although a related algorithm has been described by Rohrbaugh and Jurs[20] to assess the shape of molecules from their 3D structure, using the projected surface of the atoms toward three opposite spatial directions ("shadow areas"), and has been implemented in a QSAR model,[21] it does not consider the conformational ensemble or the amphiphilic orientation of the molecule. Therefore, (i) we developed an algorithm to orient a molecule in an amphiphilic gradient such as the air−water or lipid−water interface and to calculate the cross-sectional area, A_{Dcalc}, of the molecule perpendicular to the axis of amphiphilicity, taking into account the conformational ensemble of a molecule. (ii) A calibration of the calculated cross-sectional area, A_{Dcalc}, with the cross-sectional areas, A_D, determined by surface activity measurements yielded the cross-sectional areas relevant for membrane permeation, A_{DcalcM}. (iii) The predictive value of A_{DcalcM} for BBB permeation was then tested using three data sets.

METHODS

Amphiphilicity Axis. Molecules in an amphiphilic gradient such as the air−water or lipid−water interface orient such that the hydrophilic part remains immersed in the aqueous phase and the hydrophobic part reaches into the air or the lipid phase, respectively. The amphiphilicity of a compound has been determined previously by two different approaches: (i) Fischer et al. described it as the sum of the hydrophilic/hydrophobic contribution of an atom/fragment as described by Meylan and Howard[22] multiplied by the distance between these atoms/fragments and the charged part of the molecule;[23] (ii) Cruciani et al. determined it by assessing the center of the hydrophilic and the hydrophobic domains. A closer description of how the individual contributions were calculated is lacking.[24] Here, we used the principle of the latter approach.

To find the orientation of the molecule in such an amphiphilic gradient, the axis of amphiphilicity, defined as the line crossing the hydrophobic (Figure 1A) and the hydrophilic center of the molecule (Figure 1B, purple line) was determined.

We defined the hydrophobic center of the molecule as the center of mass of all hydrophobic atoms (carbon, fluorine, chlorine, bromine, and iodine); these atoms were weighted according to their contribution to log *P* (logarithm of the octanol−water partition coefficient) as described in the atom-additive method proposed by Wang et al.[25] They calculated log *P* as the sum of the specific contributions from the individual atoms. The contribution to log *P* further depends on the hybridization state, the number of attached hydrogen

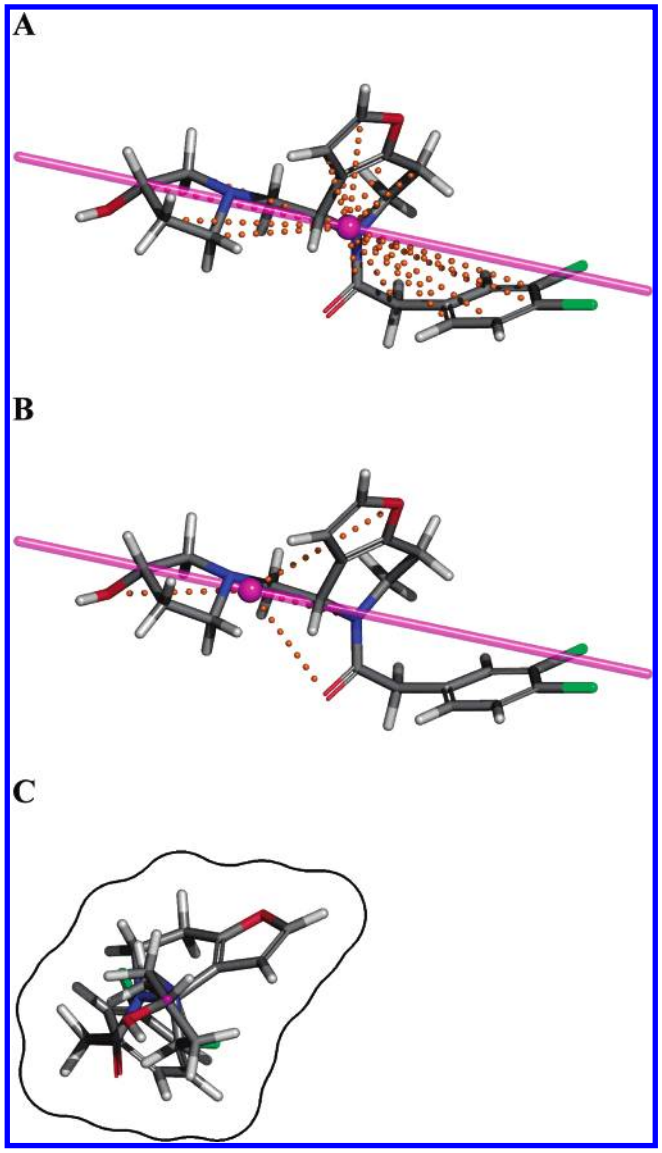

Figure 1. Calculation procedure of the cross-sectional area of GR 91272; the purple line represents the axis of amphiphilicity of the molecule crossing the hydrophobic (A) and the hydrophilic (B) centers of mass. (A) Determination of the hydrophobic center of the molecule, (B) determination of the hydrophilic center of the molecule, and (C) top view of the molecule (perpendicular to the amphiphilicity axis) and projected surface area smoothed by a water probe of 1.4 Å radius. The structures were rendered with Pymol.[40]

atoms, the nature of neighboring atoms, and the adjacency to π systems. To find these contributions, the covalent neighborhood of each hydrophobic atom (carbon, fluorine, chlorine, bromine, and iodine) was screened and its electronic state was determined. The aromatic systems were defined as cycles or groups of cycles (found using the algorithm of Balducci and Pearlman[26]) obeying Hückel's $(4n + 2)$ aromaticity rule.[27, 28]

In a similar manner, the hydrophilic center of the molecule was defined as the center of mass of all polar atoms of the molecule. To this purpose, nitrogen, oxygen, phosphorus, and sulfur atoms were weighted according to Ertl et al.[29] (Figure 1B). However, we modified this approach by taking into account the net charge, z, of the ionizable group, because charged atoms do not penetrate into the membrane and therefore remain in the aqueous environment. To a fully uncharged atom, we attributed the weight, C_O, proposed by Ertl et al.,[29] and to a fully charged atom, we attributed an arbitrary weight of 100. The contribution of a given atom corrected for charge, C_C, was calculated assuming a linear relationship from C_O to 100 for $z = 0-1$

$$C_C = 100z + C_O(1 - z) \quad (1)$$

Figure 1B shows as an example the influence of charge on the hydrophilic center of GR91272 (**V15**) with an ionization constant, pK_a 7.58, which means that 60.42% of the species is protonated at pH 7.4. The ionizable nitrogen thus strongly contributes to the hydrophilic center of mass and dominates the contributions of the uncharged nitrogen and the oxygen atoms.

To determine the ionizable atoms and their ionization state as a function of the pH, we embedded in our algorithm a pK_a determination module, based on the approach of Sayle.[30] The algorithm finds the ionizable atoms of a compound and attributes to them a user-given (if available) or a calculated pK_a value.

2D and 3D Structures. The positions of the hydrophobic and hydrophilic centers of mass, and thus the calculated cross-sectional area (see below), depend on the three-dimensional structure of the compound; multiple 3D conformations for each molecule were therefore required. The 2D structures were either imported from SciFinder Scholar (American Chemical Society) or drawn using ChemDraw (Cambridge Software). The 3D conformations were then calculated with MOE 2005.06 (Chemical Computing Group) using two different algorithms: (i) the systematic search which covers the whole conformational space by systematically rotating all rotatable bonds by discrete increments, without energy minimization, and (ii) the stochastic search, which samples local minima of the conformational space using a random approach and an energy minimization with the MMFF94X force field and chiral inversion in order to cover the whole conformational space. The hydrogen atoms were added to the structure (if required) and the partial charges were calculated prior to energy minimization. The 3D structures generated by the stochastic search algorithm were finally exported as MDL SDF files.

Calculation of the Molecular Cross-Sectional Area. To determine the cross-sectional area, A_{Dcalc}, of the molecule, the amphiphilicity axis determined for each 3D structure was oriented perpendicular to the plane of the lipid bilayer surface (assumed to be parallel to the plane of the paper) as seen in Figure 1C, and the atoms of the molecule were then projected onto this plane. The van der Waals radii of the atoms[31] were taken into account, and the contour of the projection was smoothened using a water probe of 1.4 Å radius (Figure 1C) following the definition of the molecular surface, according to Richards.[32] The area inside this smoothed contour was then defined as the calculated cross-sectional area, A_{Dcalc}. The current procedure was applied to each individual 3D structure generated with MOE, which leads to a range of A_{Dcalc} values.

To assess the difference in the calculated cross-sectional areas, A_{Dcalc}, obtained with the two different algorithms within MOE, we used verapamil (**55**) with 13 rotatable bonds as an example. The systematic search generated 257 647 conformations with calculated cross-sectional areas, A_{Dcalc}, varying from 63.43 to 131.61 Å^2 (Figure 2A, black bars). The stochastic search generated 79 conformations if the

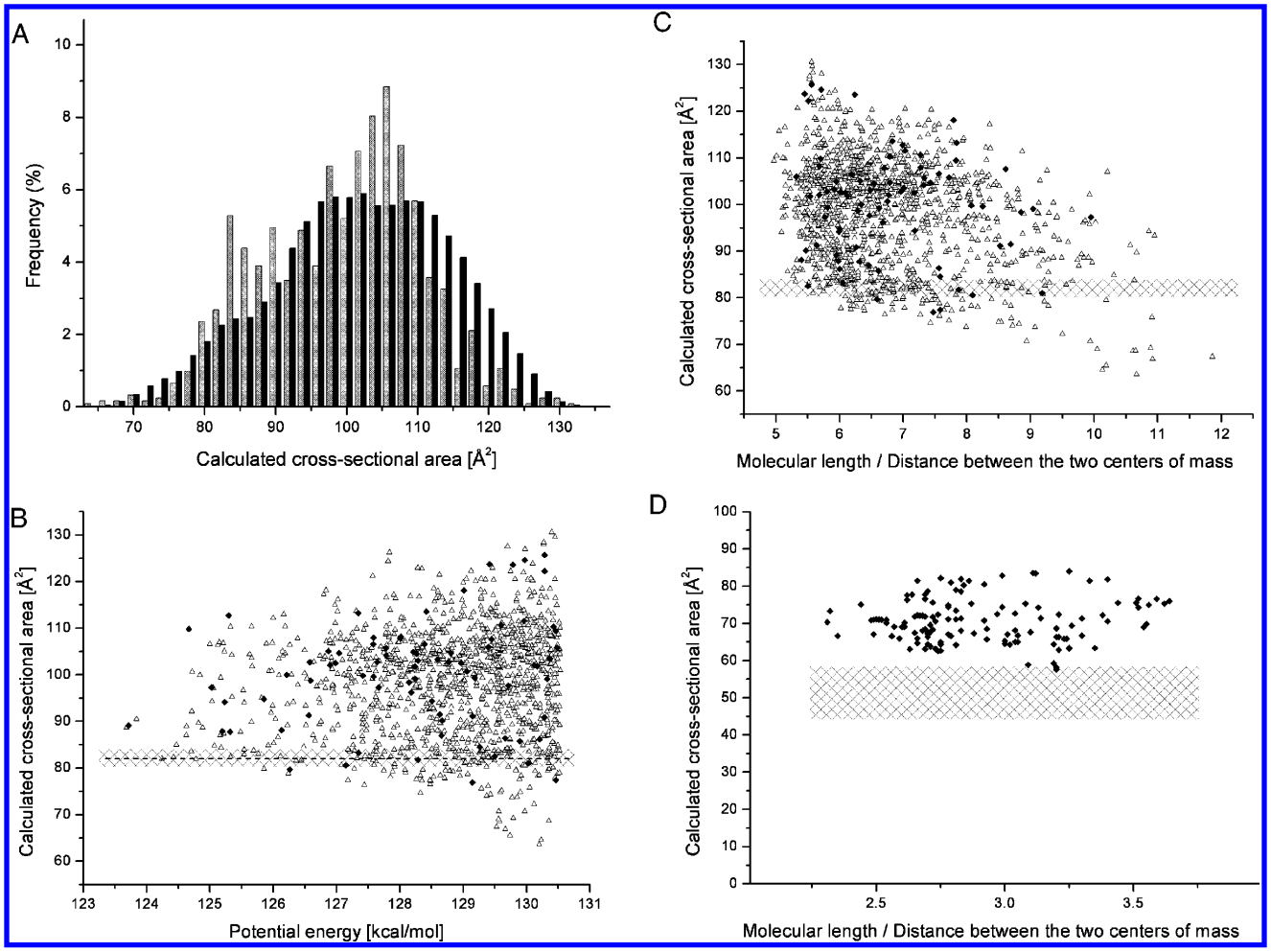

Figure 2. Calculated cross-sectional areas of verapamil (A−C) and fluoxetine (D). (A) The 257 647 conformations generated with the systematic search (black bars), and 1309 energy-minimized conformations generated with the stochastic search (grey bars); three of the latter conformations are shown in Figure 3 as an example. (B) Calculated cross-sectional area versus the potential energy of conformations generated with the stochastic search, with the failure limit set to 20 (77 conformations, black lozenges) and the failure limit set to 50 (1232 conformations, open triangles); the cross-hatched rectangles represent the experimental cross-sectional area with the error range. (C) Calculated cross-sectional area versus the quotient of molecular length (oriented toward the amphiphilicity axis) and the distance between the two centers of mass. Symbols are identical to those in B. (D) Calculated cross-sectional area versus the ratio molecular length/distance between the two centers of mass. Symbols are identical to those in B.

energy cutoff was set to 7 kcal/mol, the failure limit was set to 20, and the RMS tolerance was set to 0.1 Å (Figure 2B,C, black lozenges), and 1232 conformations were generated with the failure limit set to 50 (Figure 2B,C, open triangles) [The energy cutoff means all conformations with an energy greater than the global minimum (the minimum energy value of those conformations generated) plus the value specified in this field will be discarded. The failure limit specifies how many contiguous attempts, *k*, should be made to generate a new conformation prior to terminating the search. If *k* contiguous attempts at generating a new conformation all result in conformations already generated, the search will terminate. RMS tolerance means if two conformations have a heavy-atom RMSD less than the specified value, then they are considered duplicates. Molecular symmetry is taken into account in the superposition (definitions taken from the MOE manual)]. The calculated cross-sectional areas for the conformational ensemble obtained with the stochastic search varied from A_{Dcalc} = 63.57 to 130.56 Å^2 (Figure 2A, gray bars).

The dashed rectangles in Figure 2B and C represent the experimental cross-sectional area, A_D (with error range), of verapamil determined by means of surface-activity measurements at 24 °C and pH 7.4 (50 mM Tris, 114 mM NaCl) (Li and Seelig, unpublished result; for experimental details, see ref 12). Figure 2A highlights the fact that the repartition of calculated cross-sectional areas for conformations obtained with both systematic search and stochastic search is similar, varying from A_{Dcalc} = 63 to 132 Å^2. Moreover, Figure 2B and C show that the stochastic search with the failure limit set to 20 is sufficient to cover the whole conformational space, leading to much fewer conformations (93.75% less in the case of verapamil) and, thus, a much faster processing of the calculations. Figure 2B further shows that neither the smallest calculated cross-sectional area (A_{Dcalc} = 63.57 Å^2, Figure 3A) nor the one having the smallest potential energy in vacuo (which has a calculated cross-sectional area, A_{Dcalc} = 88.65 Å^2, Figure 3B) corresponds to the experimental value. For membrane insertion, two conditions have to be fulfilled: (i) the molecule has to be amphiphilic, whereby the hydrophilic or charged portion remains in the headgroup region of the lipid bilayer,[6] and (ii) the molecule has to be as slim as possible for energetic reasons.[11,12] We therefore searched for the most amphiphilic conformations

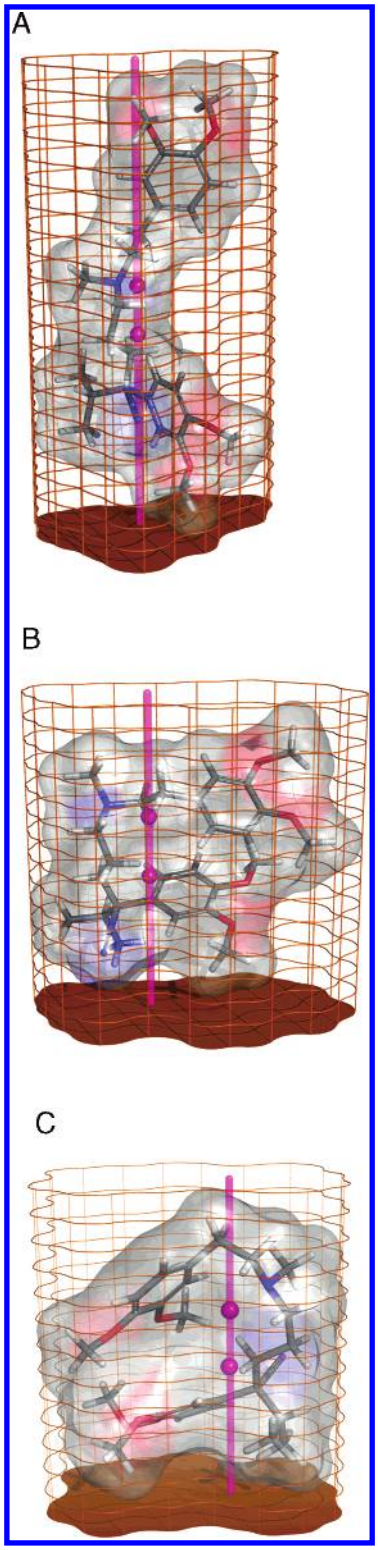

Figure 3. Three-dimensional conformations of verapamil. (A) Conformation yielding the smallest cross-sectional area. (B) Conformation with the lowest potential energy. (C) Smallest amphiphilic conformation, corresponding to the membrane-bound conformation (A_{DcalcM}). All figures were rendered with Pymol.[40]

with the smallest cross-sectional areas. This was achieved by plotting the calculated cross-sectional area versus the quotient of the length of the molecule (oriented in an amphiphilic gradient), L_M, and the distance between the hydrophilic and hydrophobic centers of mass, D_{hh} (Figure 2C). The quotient, L_M/D_{hh}, reveals the tendency of a molecule to regroup its hydrophilic fragments on one side and the hydrophobic fragments on the other side, creating a long distance between the two centers of mass, relative to the molecular length, while keeping the cross-sectional area as small as possible by eventually folding the structure (as, e.g., for verapamil, **55**, Figure 3C). Hence, the relevant conformations are those with the smallest value for the calculated cross-sectional area and a quotient $L_M/D_{hh} < 6$; for verapamil, this led to a calculated cross-sectional area, $A_{DcalcM} = 82.52$ Å^2, which is close to the experimental value ($A_D = 82 \pm 2$ Å^2) (Figure 3C).

For the evaluation of the cross-sectional area, A_{Dcalc}, we therefore always generated multiple conformations using the stochastic search algorithm with the failure limit set to 20. The number of conformations varied from one conformation (e.g., tranylcypromine, **34**) to several thousand conformations (e.g., 5105 conformations for amiodarone, **36**, which has 11 rotatable bonds). We then chose the smallest calculated cross-sectional area of each data set having a ratio $L_M/D_{hh} < 6$, A_{DcalcM}. Comparing the calculated cross-sectional areas versus the quotient L_M/D_{hh} led to the observation that nonfoldable molecules, like fluoxetine (**9**, Figure 2D), show a relatively small variation in both parameters, whereas foldable molecules, like verapamil (**55**, Figure 2C), show a much broader variation of the quotient.

Correction for Missing Hydrogen Atoms. Three-dimensional structure-generation software often omits hydrogen atoms in the output file. To be able to also use the output files lacking hydrogen atoms, we assessed a correction factor for the missing hydrogen atoms. To this purpose, the cross-sectional area with, $A_{Dcalc+H}$, and without, $A_{Dcalc-H}$, hydrogen atoms was calculated for 46 compounds (most of the drugs of Table 1). The plot of $A_{Dcalc+H}$ versus $A_{Dcalc-H}$ yielded the following linear correlation with a coefficient of determination $r^2 = 0.995$ (Figure 4)

$$A_{Dcalc+H} = A_{Dcalc-H} \times 1.13 \qquad (2)$$

Because we used the MMFF94X force field to generate multiple conformations of each compound, the hydrogen atoms were always present in the structures used for the calculation of the cross-sectional area. However, because hydrogen atoms might be absent when using other modeling software, eq 2 was implemented in the algorithm to correct for missing hydrogen atoms in case no hydrogen atoms were found in the processed structure.

Octanol−Water Partition Coefficients. The octanol−water partition coefficients log P were obtained with Kowwin v.1.67 (U.S. Environmental Protection Agency).

Distribution coefficient log D was calculated using the following equations:

$$\log D = \log P - \log[1 + 10^{pH - pK_a(acidic)}] \qquad (3)$$

for acids,

$$\log D = \log P - \log[1 + 10^{pK_a(basic) - pH}] \qquad (4)$$

for bases, and

$$\log D = \log P - \log[1 + 10^{pH - pK_a(acidic)}] - \log[1 + 10^{pK_a(basic) - pH}] \qquad (5)$$

for zwitterions.

Table 1. Molecular Weights, log *P* (calculated with Kowwin 1.67), pK_a and Calculated Cross-Sectional Area of the Membrane-Bound Conformation (A_{DcalcM}), and Experimental Cross-Sectional Area (A_{Dexp}) at pH 7.4 and pH 8.0 of a Selection of 55 Compounds (Fischer et al. [12]) Used for the Validation of the A_{Dcalc} Algorithm

							A_{DcalcM} [Å^2]	A_{Dexp} [Å^2]	
activity	no	name	MW	log *P*	p$K_{a(base)}$	p$K_{a(acid)}$	pH 7.4	pH 7.4	pH 8
BBB+	1	amitriptyline	277.4	4.95	9.4[41]		53.75	56 ± 3	52 ± 4
	2	apomorphine	267.32	2.78	8.92[42]	9.43[a]	50.61		
	3	chlorpromazine	318.86	5.2	9.3[43]		48.44	42 ± 3	
	4	chlordiazepoxide	297.74	2.42	4.8[42]		61.48	67 ± 4	
	5	clomipramine	314.85	5.65	9.5[44]		52.92	64 ± 8	50 ± 4
	6	clozapine	326.82	2.84	7.6[42]		51.29		
	7	diazepam	284.74	2.7	3.4[42]		60.17	48 ± 2	
	8	flunitrazepam	313.28	1.91	1.8[45]		64.57		
	9	fluoxetine	309.33	4.65	9.62[46]		57.55	51 ± 7	51 ± 1
	10	*cis*-flupenthixol	434.52	4.07	7.8[47]	14.96[a]	59.43	63 ± 1	60 ± 2
	11	haloperidol	375.86	4.2	8.66[43]	13.85[a]	56.67		
	12	hydroxyzine	374.9	2.36	7.1[45]	14.4[a]	60.94	71 ± 7	62 ± 1
	13	perphenazine	403.97	3.82	7.94[41]	14.96[a]	52.13	55 ± 4	55 ± 4
	14	promethazine	284.42	4.49	9.1[41]		57.75	59 ± 3	63 ± 1
	15	roxindole	346.47	5.71	8.64[a]	10.15[a]	62.07		
	16	spiradoline	425.39	4.78	9.85[a]		69.08	73 ± 5	
	17	thiopental	242.34	2.87	7.55[41]		42.66		
BBB$^+$ (sometimes described as BBB− because of low-dose administration)	18	thioridazine	370.57	6.45	9.5[42]		60.56	56 ± 8	56 ± 8
	19	clonidine	230.09	1.85	8.05[48]		37.64	34 ± 6	34 ± 1
	20	mCPP	196.68	2.19	8.72[a]		29.99	30 ± 1	30 ± 1
	21	desipramine	266.38	4.8	10.4[41]		47.71	36 ± 3	
	22	doxylamine	270.37	2.37	9.3[49]		51.34		
	23	imipramine	280.41	5.01	9.4[41]		47.86	39 ± 1	
	24	lidocaine	234.34	1.66	8.01[41]		42.83	49 ± 4	45 ± 1
	25	mequitazine	322.47	5.66	10.43[a]		62.89	45 ± 3	45 ± 1
	26	metoprolol	267.36	1.69	9.56[50]	13.89[a]	40.83		
	27	naltrexone	341.4	1.41	8.4[51]	10.3[51]	62.76		
	28	noxiptilin	294.39	4.29	9.07[a]		55.32	63 ± 2	52 ± 2
	29	piracetam	142.16	−1.4	−0.62[a]	15.67[a]	30.49		
	30	promazine	284.42	4.56	9.36[41]		55.78	57 ± 2	57 ± 4
	31	salbutamol	239.31	0.64	9.3[52]	10.3[52]	57.23	44 ± 5	27 ± 1
	32	sumatriptan	295.4	1.05	9.49[a]	11.31[a]	54.00		26 ± 2
	33	tamitinol	226.34	1.18	9.06[a]	8.11[b]	52.40	55 ± 4	
	34	tranylcypromine	133.19	1.57	8.78[a]		27.71	38 ± 1	45 ± 1
	35	zimelidine	317.22	3.18	7.91[a]		50.22		44 ± 6
BBB$^-$	36	amiodarone	645.31	8.81	6.56[53]		58.08		
	37	asimadoline	414.54	3.52	8.1[b]	14.95[b]	76.89	76 ± 5	81 ± 1
	38	astemizole	458.57	6.43	6.71[54]		92.97	94 ± 5	94 ± 3
	39	domperidone	425.91	3.35	7.9[43]		79.82		
	40	ebastine	469.66	7.64	8.78[55]		89.51		
	41	loperamide	477.04	5.15	8.66[56]	13.85[a]	100.33	102 ± 10	147 ± 9
	42	terfenadine	471.67	7.62	9.53[57]	13.32[a]	94.04	110 ± 7	92 ± 11
	54	methyl-β-cyclodextrin	1303.3	−3.87	n.d.		243.61	243	
	55	verapamil	454.61	4.8	8.92[45]		82.52	82 ± 2	90 ± 5
BBB$^-$ (not surface-active or strongly charged)	43	captopril	217.29	0.84	9.8[53]	3.7[53]	47.07		
	44	carmoxirole	374.48	6.07	8.66[a]	4.47[a]	72.91		
	45	D-mannitol	182.17	−3.01	13.5[42]		41.23		
	46	furosemide	330.74	2.32		3.65[58]	61.25		
	47	pirenzepine	351.4	1.68	8.2[59]	11.29[a]	57.17		
	48	acrivastine	348.44	2.08	8.88[a]	1.99[a]	59.41		
	49	ampicillin	349.4	1.45	7.06[54]	2.6[54]	64.10	67 ± 5	
	50	carebastine	499.64	2.83	8.48[a]	4.16[a]	89.41	141 ± 9	78 ± 5
	51	cetirizine	388.89	−0.61	8.27[60]	1.52[60]	61.45	57 ± 2	
	52	ICI204448	465.37	4.05	8.42[b]	2.89[b]	73.38	73 ± 5	
	53	penicillin G	334.39	1.85		2.74[42]	48.58		

[a] pK_a calculated using Advanced Chemistry Development (ACD/Labs) software v.8.14 for Solaris. [b] pK_a calculated using ChemAxon software MarvinSketch v.4.0.1.

Calculation Speed. The algorithm for the calculation of the cross-sectional areas, A_{Dcalc}, can process hundreds of compounds within a few seconds on a normal desktop PC. Moreover the speed of computing can be enhanced by decreasing the contour resolution, which has only negligible effects on the final result. The computational speed can be further enhanced by omitting the hydrogen atoms of the structures, which strongly decreases the number of atoms to be processed, and then correcting the obtained cross-sectional area without hydrogen atoms, $A_{Dcalc-H}$, with an average correction factor, $f = 1.13$ (eq 2), to obtain the cross-sectional area, A_{Dcalc}, with hydrogen atoms.

RESULTS

Comparison of Calculated and Measured Cross-Sectional Areas. To validate our algorithm, we compared

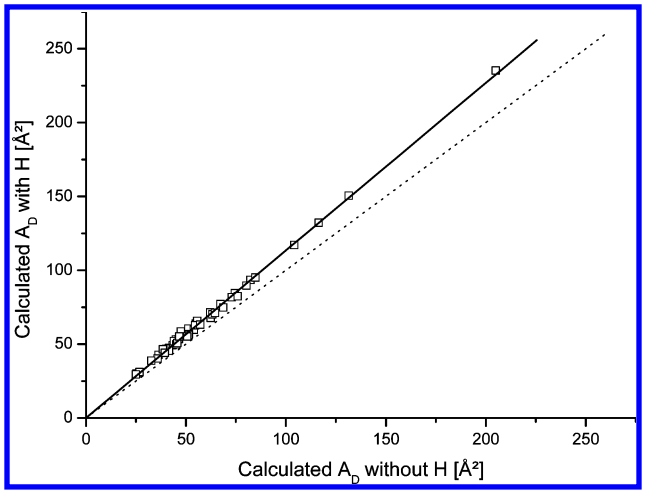

Figure 4. Comparison between the cross-sectional areas calculated with, $A_{Dcalc+H}$, and without, $A_{Dcalc-H}$, hydrogen atoms for 46 compounds. The straight line represents the linear regression of the data, and the dotted line, with the slope of 1, was drawn to guide the eye.

the calculated, A_{DcalcM}, with the measured, A_D, cross-sectional areas for 55 compounds (Table 1), covering a broad range of molecular weights (130−1300 Da).[12] The experimentally determined cross-sectional areas, A_D, were obtained from measurements of the Gibbs adsorption isotherm, that is, the surface pressure, π, as a function of the concentration of the drug in a buffer solution (for experimental details, see ref 12). Measurements were performed at pH 7.4 (Figure 5A) and pH 8.0 (for cations) or pH 6.8 (for anions) (Figure 5B). The latter two pH values were chosen to minimize electrostatic repulsion between the compounds in the air−water interface. The linear regression of A_D (pH 7.4) versus A_{DcalcM} yielded a coefficient of determination $r^2 = 0.903$, a standard deviation SD = 12.22, and a slope close to 1 ($A_D = 1.027A_{DcalcM}$). As expected, the correlation under conditions of minimal electrostatic repulsion (Figure 5B) is slightly better than that at pH 7.4: $r^2 = 0.948$, SD = 8.67, and $A_D = 0.962A_{DcalcM}$ (Figure 5B). It has to be noted that the pK_a values of cationic (anionic) compounds in the air−water interface are distinctly lower (higher) than those in a dilute solution.[14] Therefore, charge repulsion minima are generally reached in the range of pH 7.4−8 for cations and around pH 6.8 for anions. (No good agreement was achieved between experimental data by Fischer et al.,[12] and experimental data by Suomalainen et al.[33] This is due to the fact that drug stock solutions in the latter case were prepared in a buffer, not in pure water,[12] which can lead to molecular association in the solution and hence generally larger apparent cross-sectional areas.)

Packing Density of Molecules at the Air−Water Interface. For compounds such as the phenothiazine mequitazine (**25**), the calculated cross-sectional area was larger than the experimental cross-sectional area ($A_{DcalcM} < A_D$). The measured value was 50 ± 5 Å² in the range of pH 7.4−8.0 (Li and Seelig, unpublished data), whereas A_{DcalcM} is 62.9 Å² at pH 7.4 for one single molecule (Table 1). As shown in Figure 6, packing several mequitazine molecules in an antiparallel manner can lead to distinctly smaller cross-sectional areas. As calculated from an ensemble of 12 molecules of mequitazine, the cross-sectional area can be as small as $A_{Dcalc} = 56.2$ Å² (Figure 6).

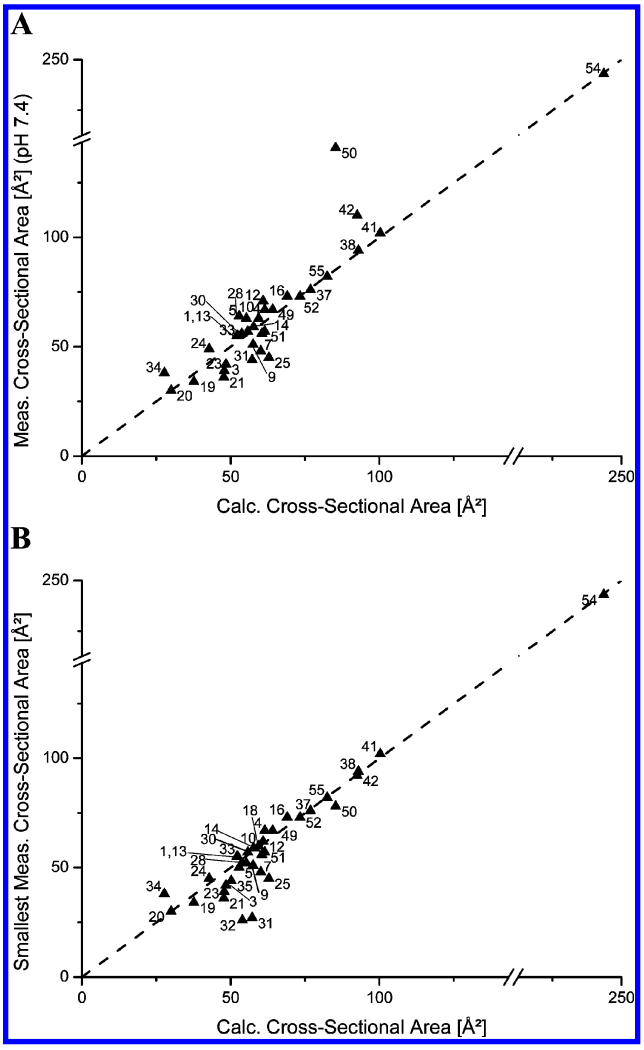

Figure 5. Plot of the measured versus the calculated cross-sectional area of the membrane-bound conformation. The labels refer to the compounds in Table 1. The dashed lines correspond to a slope of 1. Measurements were performed at pH 7.4 (A) and pH 8.0 (bases) or pH 6.8 (acids) (B).

Calibration Diagram for BBB Prediction Using the Calculated Cross-Sectional Area, A_{DcalcM}, and Calculated Octanol−Water Distribution Coefficient, log $D_{7.4}$. To establish a calibration diagram for the prediction of BBB permeation, we plotted the calculated cross-sectional area, A_{DcalcM}, and calculated octanol−water partition coefficient, log $D_{7.4}$, for a data set of 55 compounds (Table 1), comprising 35 drugs able to cross the blood−brain barrier (BBB^+) and 20 peripherally acting drugs which are not able to reach the brain (BBB^-) (for details, see ref 12).

The plot of log $D_{7.4}$ versus A_{DcalcM} (Figure 7) shows an excellent discrimination between BBB^+ and BBB^- drugs, suggesting a limiting calculated cross-sectional area of $A_{DcalcM} = 70$ Å² (A_{DcalcM} of spiradoline, **16**) and a limiting range of octanol−water distribution coefficients of $-1.4 \leq \log D_{7.4} < 7.0$ for BBB permeation. When these limits are used, the only misinterpreted compound was pirenzepine (**47**), which is not surface-active and was therefore predicted correctly on the basis of surface-activity measurements.[12] The prediction accuracy based on log $D_{7.4}$ and A_{DcalcM} was thus 100% for BBB^+ drugs and 95% for BBB^- drugs, leading to an overall prediction accuracy of 98%.

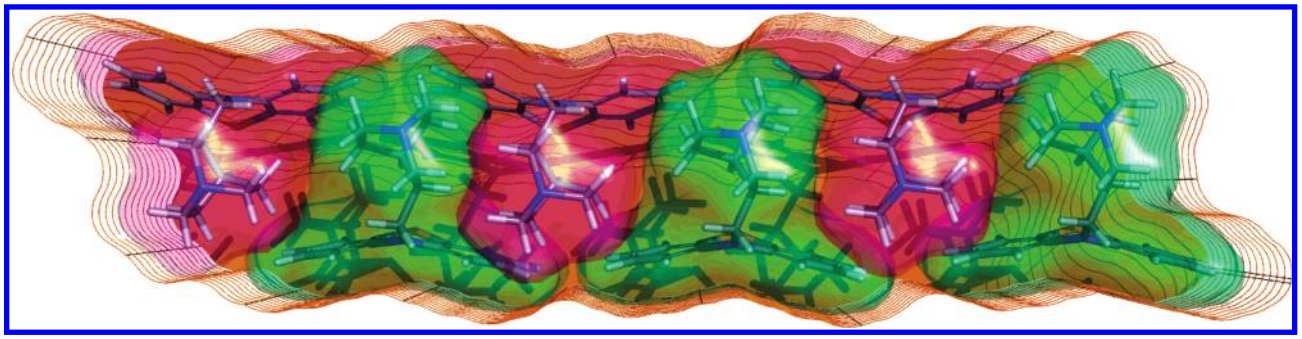

Figure 6. Possible arrangement of six molecules of mequitazine at a lipid−water or air−water interface. The figure was rendered with Pymol.[40]

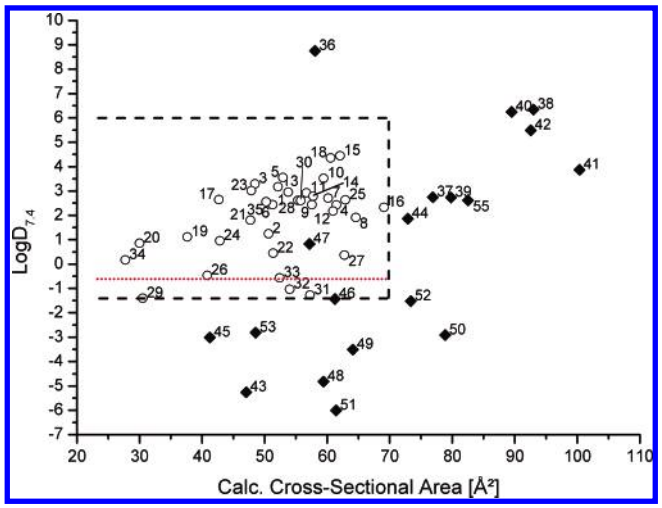

Figure 7. Log *D* calculated at pH 7.4 versus the calculated cross-sectional area of the membrane-bound conformation calibration diagram. Open circles are BBB^+ compounds; closed symbols are BBB^- compounds. Labels refer to compounds in Table 1. The dashed lines correspond to the limits set between BBB^+ and BBB^- compounds (calibration set). The dotted line is the limit defined after evaluation of the validation set.

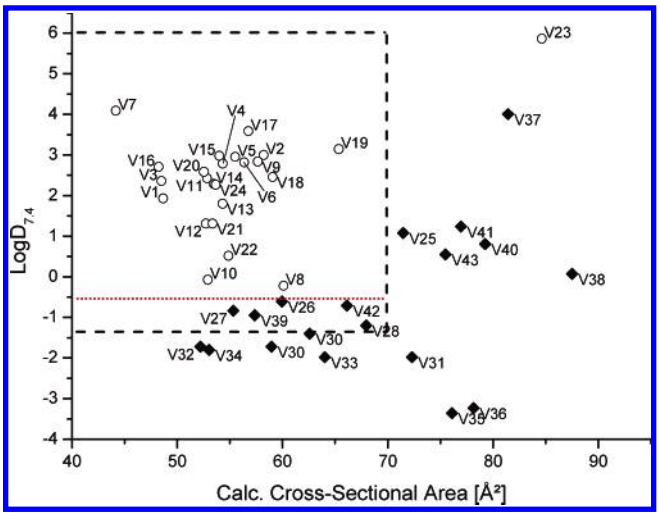

Figure 8. Log *D* calculated at pH 7.4 versus the calculated cross-sectional area of the membrane-bound conformation validation diagram. The dashed lines correspond to the limits set between BBB^+ and BBB^- compounds. Refer to Figure 7 for the symbol legend.

Validation of the BBB Permeation Diagram on the Basis of Calculated Cross-Sectional Areas. To validate the approach using the calculated cross-sectional area, A_{DcalcM}, together with the calculated log $D_{7.4}$, for the prediction of BBB permeation, we used a second data set of 43 compounds[34] (Table 2), comprising 24 drugs able to cross the BBB (BBB^+) and 19 drugs which are not able to cross the BBB at the concentrations applied and have a peripheral target (BBB^-).

Figure 8 shows a plot of log $D_{7.4}$ versus A_{DcalcM} where the dashed lines represent the limits defined above ($A_{DcalcM} < 70$ Å^2 and $-1.4 \leq \log D_{7.4} < 7$). Again, a good discrimination between BBB^+ and BBB^- drugs was observed. The two parameters, A_{DcalcM} and log $D_{7.4}$ were able to correctly predict 96% of the brain-targeted compounds and 74% of the peripherally targeted drugs, leading to an overall prediction accuracy of 86%. However, all of the BBB^- compounds which were incorrectly predicted have an octanol−water distribution coefficient log $D_{7.4} \leq -0.6$. By increasing the limit between brain-targeted drugs and peripherally acting drugs from −1.4 to −0.6 for log $D_{7.4}$ (Figures 7 and 8, red dots), we increased the overall prediction accuracy of the validation set to 97.7%, whereas the prediction accuracy for the calibration set decreased from 98% to 93%.

Three other published models used the same data set of 43 compounds (Table 2) either as a training set (Deconinck et al.[35] and Crivori et al.[36]) or as a validation set (Narayanan and Gunturi[37]). The prediction accuracies obtained with our model were comparable to those obtained by these three other models, with overall accuracies of 97.7%, 81%, and 95.3% for the models of Deconinck et al. (1639 descriptors, 150 classification trees of size 4), Narayanan and Gunturi (324 descriptors, three- and four-descriptor models), and Crivori et al. (72 descriptors, three significant principal components), respectively (Table 3).

To further test our approach, we selected all of the compounds that were mispredicted in at least one of these publications. This led to a third subset of 42 compounds (15 BBB^+ drugs and 27 BBB^- drugs) from which three compounds were missing in the data set of Deconinck et al.,[35] nine missing in the data set of Narayanan and Gunturi,[37] and covered in totality in the data sets of Crivori et al.[36] (Table 4). It is worthy to note that, in the publication of Deconinck et al., all compounds were present in the training set, and we considered the results of the best discrimination tree of size 4 obtained with the Gini index as a split criterion.

In the subset of 42 compounds (Table 4), 10 compounds were present in our calibration set and seven compounds in our validation set. The 25 remaining drugs were modeled using the same procedure as described before; the log P_{ow} values were calculated using Kowwin v.1.67 and the pK_a using ACD/Labs v.8.14. We also used the limit of −0.6 for log $D_{7.4}$ and 70 Å^2 for the calculated cross-sectional area as

Table 2. Molecular Weights, log *P* (calculated with Kowwin 1.67), pK_a and Calculated Cross-Sectional Area of the Membrane-Bound Conformation (A_{DcalcM}) of 43 Compounds Used as a Validation Set

activity	no	name	MW	log *P*	p$K_{a(base)}$	p$K_{a(acid)}$	A_{DcalcM} [Å^2]
BBB$^+$	V1	BRL52537	355.3	4.37	9.84[a]		48.66
	V2	BRL52580	403.34	5.15	9.55[a]		58.22
	V3	BRL52656	354.41	4.04	9.07[b]		48.50
	V4	BRL52871	409.37	4.97	9.58[a]		54.32
	V5	BRL53080	409.37	4.97	9.41[a]		55.52
	V6	BRL53087	383.36	5.24	9.82[a]		56.36
	V7	cyclazocine	257.37	4.09	9.15[a]	10.21[a]	44.18
	V8	EMD60400	353.46	0.97	8.56[a]		60.11
	V9	GR45809	413.34	4.52	9.07[b]		57.64
	V10	GR85571	398.33	2.07	9.54[a]		52.92
	V11	GR88377	393.31	4.52	9.49[a]		52.84
	V12	GR89696	414.33	2.92	9.00[b]		52.73
	V13	GR89696et	428.35	3.41	9.00[b]		54.30
	V14	GR89696pr	442.38	3.9	9.00[b]		53.48
	V15	GR91272	409.31	2.98	7.58[b]	14.95[b]	53.99
	V16	ICI197067	357.32	4.41	9.09[b]		48.25
	V17	ICI199441	391.33	4.71	8.49[b]		56.77
	V18	nalorphine	311.37	2.46	7.39[a]	9.49[a]	59.06
	V19	RP60180	395.56	5.55	9.80[a]		65.35
	V20	sankyo	370.32	2.59	5.00[a]	12.07[a]	52.54
	V21	SB201708	393.31	3.21	9.29[a]		53.35
	V22	SB204484	394.3	2.39	9.26[a]		54.87
	V23	tifluadom	415.5	5.87	4.50[a]	14.13[a]	84.65
	V24	U50488	369.33	4.78	9.91[a]		53.64
BBB$^-$	V25	BRL52974	393.31	3.21	9.53[a]	13.79[a]	71.46
	V26	GR94839	414.33	0.53	8.51[a]	14.95[b]	59.95
	V27	GR94839A	413.43	0.2	8.40[b]	14.95[b]	55.33
	V28	GR94839B	391.53	−0.16	8.40[b]	14.95[b]	67.95
	V30	GR94839C	381.42	−0.36	8.40[b]	14.95[b]	62.57
	V30	GR94839D	375.46	−0.68	8.40[b]	14.95[b]	58.96
	V31	GR94839E	390.43	−0.94	8.40[b]	14.95[b]	72.30
	V32	GR94839F	375.46	−0.68	8.40[b]	14.95[b]	52.19
	V33	GR94839G	390.43	−0.94	8.40[b]	14.95[b]	64.01
	V34	GR94839H	345.44	−0.76	8.40[b]	14.95[b]	53.02
	V35	GR94839I	409.54	−2.32	8.40[b]	14.95[b]	76.11
	V36	GR94839L	423.53	−2.19	8.40[b]	14.95[b]	78.16
	V37	ICI204448	465.37	4.05	2.89[b]	8.42[b]	73.38
	V38	ICI205640	493.42	5.03	9.56[a]	4.60[a]	87.52
	V39	SB204454	439.34	1.38	9.72[a]	9.44[a]	57.37
	V40	SB204457	439.34	2.28	8.86[b]		79.25
	V41	SB204459	409.31	1.68	7.65[b]	14.95[a]	76.96
	V42	SB205563	416.34	0.48	8.56[a]		66.12
	V43	SB205605	409.31	1.68	8.50[a]	13.79[a]	75.48

[a] pK_a calculated using Advanced Chemistry Development (ACD/Labs) software v.8.14 for Solaris. [b] pK_a calculated using ChemAxon software MarvinSketch v.4.0.1.

Table 3. Comparison of Different in Silico Methods to Assess Blood−Brain Barrier Permeation

source	set	descriptors	number of compounds (BBB$^+$/BBB$^-$)	compounds from training set (BBB$^+$/BBB$^-$)	correct prediction (BBB$^+$/BBB$^-$)	prediction accuracy (%)
Table 1	calibration	1[a]	54 (35/19)		45 (35/10)	83.3%
	calibration	2[b]	54 (35/19)		50 (32/18)	92.6%
Table 2	validation	1[a]	43 (24/19)	0	32 (23/9)	74.4%
	validation	2[b]	43 (24/19)	0	42 (23/19)	97.7%
	training[28]	4[c]	43 (24/19)		42 (24/18)	97.7%
	validation[29]	4[d]	42 (23/19)	0	34 (16/18)	81.0%
	training[27]	3[e]	43 (24/19)		41 (23/18)	95.3%
Table 4	validation	1[a]	42 (16/26)	10 (3/7)	30 (15/15)	71.4%
	validation	2[b]	42 (16/26)	10 (3/7)	35 (15/20)	83.3%
	training[28]	4[c]	39 (14/25)		29 (11/18)	74.4%
	validation[29]	4[d]	33 (15/18)	6 (3/3)	10 (3/7)	30.3%
	validation[27]	3[e]	41 (15/26)	5 (4/1)	14 (11/3)	34.1%

[a] Cross-sectional area A_{DcalcM}. [b] Cross-sectional area A_{DcalcM} and log $D_{7.4}$. [c] Starting from 1639 descriptors, 150 trees of size 4 were built. [d] Starting from 324 descriptors, three four-descriptor models were developed. [e] Starting from 72 descriptors, a model based on three significant principal components was obtained.

defined previously to discriminate between BBB$^+$ and BBB$^-$ compounds.

Our model was able to correctly predict the blood−brain barrier permeation of 35 compounds out of 42 (83.3%), with

Table 4. Calculated Cross-Sectional Areas of the Membrane-Bound Conformation (A_{DcalcM}) and log $D_{7.4}$ (Kowwin Used for log *P* and ACD/Labs for pK_a) of 42 Compounds Mispredicted in Some Other Works

name	A_{DcalcM} [Å^2]	log $D_{7.4}$	exptl. BBB	this work	Deconinck	Narayanan	Crivori
alprazolam	64.19	2.12	+	+		−[a]	+
astemizole	84.44	6.34	−	−[a]	−[a]	−	+
BRL52974	71.46	1.08	−	−	+[a]	+	+[a]
carboxymefloquine	37.59	−1.76	−	−	+[a]		+
carebastine	78.87	−2.91	−	−[a]	−[a]	+	−
carmoxirol	72.91	1.86	−	−[a]	+[a]	+	+
cimetidine	64.31	0.49	−	+	−[a]	−[a]	+
clonidine	37.64	1.11	+	+[a]	−[a]	+[a]	+
corticosterone	78.86	1.94	−	−	+[a]		−
cp102	55.96	−0.66	−	−	−[a]		+
cp107	46.66	−0.12	−	+	−[a]		+
cp41	55.73	−0.92	−	−	−[a]		+
desloratadine	58.35	1.76	−	+	+[a]	+	+
difloxacin	75.91	−0.96	−	−	−[a]	−	+
dopamine	29.88	−2.34	−	−	−[a]	−	+
ebastine	90.06	6.24	−	−[a]	−[a]	+	+
EMD 61753	77.23	2.34	−	−			+
fexofenadine	92.5	1.08	−	−	−[a]	+	−
GR85571	52.92	−0.07	+	+	+[a]	−	+[a]
GR89696et	54.3	1.8	+	+	+[a]	+	−[a]
GR91272	53.99	2.98	+	+	+[a]	−	+
L364,718	79.87	3.06	−	−	+[a]		+
L663,581	44.41	2.1	+	+	−[a]	−	−
loperamide	96.95	3.87	−	−[a]	−[a]	+	+
loratidine	72.35	2.6	−	−	+[a]	+	+
mefloquine	58.36	1.21	−	+	−[a]	+	+
mequitazine	62.63	2.63	−	+[a]	−[a]	+	+
morphine	48.01	−0.2	+	+	+[a]	−	+
nalorphine	59.06	2.46	+	+	+[a]	−	+
naltrexone	54.84	0.37	+	+[a]	+[a]	−	−
nordazepam	55.22	2.87	+	+	+[a]	−	+
oxacepam	56.13	2.24	+	+		−[a]	+
ranitidine	42.47	−0.77	−	−	−[a]	−[a]	+
SB201708	53.35	1.31	+	+	+[a]	−	+[a]
SB204484	54.87	0.52	+	+	+[a]	−	+[a]
skb-i	77.19	6.02	−	−	−[a]		+
tamitinol	46.52	−0.57	+	+[a]	−[a]		+
temelastine	78.24	4.96	−	−	−[a]	−	+
terfenadine	92.57	5.49	−	−[a]	−[a]	+	+
thiopental	42.69	2.64	+	+	+[a]	±	−
tifluadom	84.65	5.87	+	−	+[a]	−	?[a,b]
tiotidine derivative (BBCPD 16)	33.32	−0.58	−	+	−[a]	−[a]	+
total				42	39	33	41
correct prediction				35 (83.3%)	29 (74.4%)	10 (30.3%)	13 (31.7%)

[a] The compound belongs to the training set in the original publication. [b] The prediction for tifluadom could not be found in the original publication.

a prediction accuracy of 93.3% for BBB^+ compounds and 77.8% for BBB^- compounds. On the same subset of compounds, Deconinck et al. obtained an overall prediction of 74.4% by correctly predicting 29 out of 39 drugs, Narayanan and Gunturi an accuracy of 30.3% (10 out of 33 drugs), and Crivori et al. an accuracy of 31.7% (13 out of 41; Table 3). We should mention that several mispredicted compounds, like mequitazine or mefloquine, showed an alteration of vigilance and sleepiness; these symptoms observed in double-blind crossover placebo-controlled trials suggest that these drugs reach the central nervous system if applied at higher concentrations. For example, mequitazine is applied at 10 mg/day, whereas the structurally related CNS-targeted phenothiazines are applied at 100−200 mg/day (see ref 12).

DISCUSSION

To determine the optimal membrane partitioning conformation of a drug, we first generated the conformational ensemble for a given compound using MOE, combined with a stochastic approach to reduce the number of conformations and thus the computational time. We then assessed the axis of amphiphilicity and calculated the cross-sectional area of the compound in its amphiphilic orientation. To obtain the optimal conformation for membrane partitioning, we then calibrated the ensemble of the calculated cross-sectional areas, A_{Dcalc}, with the cross-sectional area, A_D, measured under conditions of minimal charge repulsion. The calibration revealed that the measured cross-sectional area corresponds well to the smallest cross-sectional area in the case of nonfoldable compounds and to the smallest cross-sectional area of an amphiphilic conformation ($L_M/D_{HH} < 6$) in the case of foldable compounds. The amphiphilic conformation with the smallest calculated cross-sectional areas (from various conformers), A_{DcalcM}, led to an excellent linear correlation with the measured cross-sectional area, A_D (coefficient of determination, $r^2 = 0.94$, and a slope close to 1). Only a small variation of A_D between the different

stereoisomers of racemic compounds was observed (data not shown); thus, the stochastic conformation search within MOE was configured in order to cover the whole conformational space, including racemates. The quality of the calculated cross-sectional areas, A_{DcalcM}, as a descriptor for the prediction of blood−brain barrier permeation was then tested on several data sets.

Limiting Values of the Cross-Sectional Area and log $D_{7.4}$ for BBB Permeation. The calibration and validation procedures yielded an upper limiting cross-sectional area for BBB permeation $A_{DcalcM} = 70$ Å^2. This is in good agreement with the previous calibration results obtained from surface activity measurements which yielded a limiting cross-sectional area for BBB permeation $A_D = 80$ Å^2.[12]

The cutoff for passive diffusion across the BBB is due to the rather high packing density of the lipid bilayer, as discussed above. Because of the high expression level of efflux transporters in the BBB, which efficiently export the slowly diffusing (large or charged) molecules, the cutoff is even more rigorous than that in a simple lipid bilayer. If the rate of passive drug influx drops below the rate of active drug efflux by multidrug resistance transporters, the drugs no longer reach the cytosol or cross the BBB.[38]

For the octanol−water distribution coefficient, a lower and upper value of $-0.6 \leq \log D_{7.4} < 7$ was assessed. This agrees well with the intermediate range of air−water partition coefficients.[12] If compounds are small and soluble, the lower limit can be reduced down to −1.4 as shown in Figure 7 because partitioning into the lipid bilayer can be easily enhanced by increasing the concentration of drugs. The upper limit is more difficult to overcome because increasing the concentration of hydrophobic compounds may lead to aggregation. As mentioned above, hydrophobic compounds with large cross-sectional areas tend to be substrates for multidrug resistance transporters. Provided the drugs will not aggregate in solution, increasing the concentration may help to modulate or inhibit the multidrug resistance transporters and to enhance passive diffusion.[39]

Prediction Accuracy. The quality of the calculated cross-sectional area A_{DcalcM} and the calculated log $D_{7.4}$ as parameters for the prediction of the blood−brain barrier permeation have then been evaluated with a first data set of 55 compounds and have been validated with a second independent data set of 43 compounds and a third data set of 42 compounds. It is interesting to note that the calculated cross-sectional area alone can correctly predict 83% of blood−brain barrier permeation for the first data set, 74.4% for the second data set, and 71.4% for the third data set (Table 3). Together with the calculated log $D_{7.4}$, the prediction accuracy for the present set of 55 + 43 + 42 compounds (total of 122 different drugs) increased to 85.2% with the lower log $D_{7.4}$ limit set to −1.4 and to 90.2% with this limit set to −0.6 (87.3% prediction accuracy for BBB^- compounds and 92.4% prediction accuracy for BBB^+ compounds). However, the limits between BBB^+ and BBB^- compounds should not be set in a too restrictive manner because, depending on the concentration applied or the formulation[10] used, the experimental outcome may vary.

Comparison of Different Prediction Methods. We compared our method to predict BBB permeation with three methods described previously.[35−37] All three methods are heuristic approaches and start with a large number of descriptors (72−1639) and then filter out the most relevant descriptors. The latter are related to (i) the size of the molecule (κ shape index used by Narayanan and Gunturi[37] and the sum of topological distances between N and O atoms used by Deconinck et al.[35]), (ii) the charge of the molecule [an electrotopological descriptor (SsssN) used by Narayanan and Gunturi[37] and a descriptor yielding information on the atomic charges and atomic polarizability (BEHm4) used by Deconinck et al.[35]], and (iii) the hydrophilicity/lipophilicity of the molecule (log P used by Narayanan and Gunturi[37] and a hydrophilicity index used by Deconinck et al.[35]). In the third approach,[36] the most relevant descriptors are not revealed.

The present approach is based on clear physical−chemical principles developed on the basis of structural and dynamic investigations of the lipid bilayer membrane and drug partitioning measurements performed with lipid monolayers and bilayers (cf. Introduction). They revealed that partitioning into an amphiphilic lipid leaflet is not a random process but requires conformations in which the polar and nonpolar groups of a drug are arranged in an amphiphilic manner. Because cavity formation in an anisotropic membrane requires energy, the partitioning of molecules with small cross-sectional areas is favored, and because the hydrophobic core of the lipid bilayer exhibits a low dielectric constant, molecules can only diffuse in their uncharged form. For the prediction of membrane permeation, we therefore chose (i) the cross-sectional area, A_D; (ii) the ionization constant, pK_a; and (iii) the octanol−water partition coefficient, P, as parameters, whereby the latter two were combined and expressed as log $D_{7.4}$. As expected, the parameters found to be most relevant in the three heuristic approaches are related to the parameters based on physical−chemical principles used in the present approach. The main difference is, however, that we investigated the full conformational space of the molecule, defined the cross-sectional area in the amphiphilic direction for each conformation, A_{Dcalc}, and chose the most amphiphilic and at the same time smallest cross-sectional area, A_{DcalcM}. The last step, that is, the determination of the cross-sectional area, A_{DcalcM}, of the membrane-binding conformation of the drug, was possible by a calibration with measured cross-sectional areas, A_D.

CONCLUSIONS

The full conformational space of a molecule was investigated, and the cross-sectional areas of all of the conformations were calculated perpendicular to their axis of amphiphilicity, yielding a large number of A_{Dcalc} values. For calibration, the cross-sectional area, A_D, measured at the air−water interface was used and yielded the conformations relevant for membrane partitioning, A_{DcalcM}. This calibration revealed that the most amphiphilic conformation with the smallest cross-sectional area had to be chosen from the ensemble of conformations, which is in accordance with the physical−chemical principles for membrane partitioning. For several compounds, including verapamil, the conformation which meets the above demands is folded. A plot of the measured cross-sectional, A_D, versus the calculated minimal and most amphiphilic cross-sectional area, A_{DcalcM}, yielded an excellent linear correlation. The present data reveal that partitioning into a membrane requires a specific conformation and

orientation. Membrane partitioning thus clearly differs from partitioning into an isotropic solvent and resembles to some extent receptor docking, which also requires a specific molecular conformation and orientation. The predictive value of the parameter A_{DcalcM}, for the BBB permeation, was then tested with different data sets, revealing that the calculated cross-sectional area alone (when excluding highly charged compounds) correctly predicted 83.3% of blood−brain barrier permeation on the first data set, 74.4% for the second data set, and 71.4% for the third data set. Together with the calculated log $D_{7.4}$, the prediction accuracy for the present set of 55 + 43 + 42 compounds (total of 122 different drugs) increased to 90.2%. This high prediction accuracy obtained with a model based on two descriptors only (calculated cross-sectional area, A_{DcalcM}, and distribution coefficient at pH 7.4, log $D_{7.4}$) can be explained by the fact that these descriptors fulfill the conditions required for membrane partitioning.

The simple and unambiguous criteria for membrane partitioning should help for the structural optimization of drug candidates in the early stages of drug discovery.

ACKNOWLEDGMENT

This work was supported by the Swiss National Science Foundation (grant number: 3100-107793) and F. Hoffmann La-Roche AG.

REFERENCES AND NOTES

(1) Cecchelli, R.; Dehouck, B.; Descamps, L.; Fenart, L.; Buee-Scherrer, V. V.; Duhem, C.; Lundquist, S.; Rentfel, M.; Torpier, G.; Dehouck, M. P. In Vitro Model for Evaluating Drug Transport across the Blood−Brain Barrier. *Adv. Drug Delivery Rev.* **1999**, *36* (2−3), 165−178.

(2) van De Waterbeemd, H. Physico-chemical Approaches to Drug Absorption. In *Drug Bioavailability − Estimation of Solubility, Permeability, Absorption and Bioavailability*; van De Waterbeemd, H., Lennernas, H., Artursson, P., Eds.; Wiley-VCH: New York, 2003.

(3) Wohnsland, F.; Faller, B. High-Throughput Permeability pH Profile and High-Throughput Alkane/Water log P with Artificial Membranes. *J. Med. Chem.* **2001**, *44* (6), 923−30.

(4) Seelig, A.; Gottschlich, R.; Devant, R. M. A Method to Determine the Ability of Drugs to Diffuse through the Blood−Brain Barrier. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91* (1), 68−72.

(5) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J. Med. Chem.* **1998**, *41* (7), 1007−10.

(6) Seelig, A.; Seelig, J. Membrane Structure. In *Encyclopedia of Physical Science and Technology*, 3rd ed.; Meyers, R. A., Ed.; Academic Press: New York, 2002; Vol. 9, pp 355−367.

(7) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **1997**, *23* (1−3), 3−25.

(8) Seelig, J.; Seelig, A. Lipid Conformation in Model Membranes and Biological Membranes. *Q. Rev. Biophys.* **1980**, *13* (1), 19−61.

(9) Seelig, A. Local Anesthetics and Pressure: A Comparison of Dibucaine Binding to Lipid Monolayers and Bilayers. *Biochim. Biophys. Acta* **1987**, *899*, 196−204.

(10) Seelig, A.; Gerebtzoff, G. Enhancement of Drug Absorption by Noncharged Detergents through Membrane and P-glycoprotein Binding. *Expert Opin. Drug Metab. Tox.* **2006** in press.

(11) Boguslavsky, V.; Rebecchi, M.; Morris, A. J.; Jhon, D. Y.; Rhee, S. G.; McLaughlin, S. Effect of Monolayer Surface Pressure on the Activities of Phosphoinositide-Specific Phospholipase C-beta 1, -gamma 1, and -delta 1. *Biochemistry* **1994**, *33* (10), 3032−7.

(12) Fischer, H.; Gottschlich, R.; Seelig, A. Blood−Brain Barrier Permeation: Molecular Parameters Governing Passive Diffusion. *J. Membr. Biol.* **1998**, *165* (3), 201−11.

(13) Finkelstein, A. *Water Movement through Lipid Bilayers, Pores, and Plasma Membranes. Theory and Reality*; Wiley-Interscience: New York, 1987; Vol. 4, Chapter 6.

(14) Gerebtzoff, G.; Li-Blatter, X.; Fischer, H.; Frentzel, A.; Seelig, A. Halogenation of Drugs Enhances Membrane Binding and Permeation. *ChemBioChem* **2004**, *5* (5), 676−84.

(15) Gobas, F. A.; Lahittete, J. M.; Garofalo, G.; Shiu, W. Y.; Mackay, D. A Novel Method for Measuring Membrane−Water Partition Coefficients of Hydrophobic Organic Chemicals: Comparison with 1-Octanol−Water Partitioning. *J. Pharm. Sci.* **1988**, *77* (3), 265−72.

(16) Fischer, H.; Seelig, A.; Chou, R. C.; van De Waterbeemd, H. The Difference between the Diffusion through the Blood−Brain Barrier and the Gastro-Intestinal Membrane, 4th International Conference on Drug Absorption, Edinburgh, Scotland, June 13−15, 1997.

(17) Norinder, U.; Haeberlein, M. Computational Approaches to the Prediction of the Blood−Brain Distribution. *Adv. Drug Delivery Rev.* **2002**, *54* (3), 291−313.

(18) Clark, D. E. In Silico Prediction of Blood−Brain Barrier Permeation. *Drug Discovery Today* **2003**, *8* (20), 927−33.

(19) Ecker, G. F.; Noe, C. R. In Silico Prediction Models for Blood−Brain Barrier Permeation. *Curr. Med. Chem.* **2004**, *11* (12), 1617−28.

(20) Rohrbaugh, R. H.; Jurs, P. C. Molecular Shape and the Prediction of High-Performance Liquid Chromatographic Retention Indexes of Polycyclic Aromatic Hydrocarbons. *Anal. Chem.* **1987**, *59* (7), 1048−54.

(21) Zhang, Q.-Y.; Luo, C.-C.; Qi, Y.-H.; Dong, L.; Wang, J.; Xu, L. Peripheries of Molecular Projection in Three-Dimensional Space and Studies of Quantitative Structure−Activity Relationships. *Chin. J. Chem.* **2004**, *22* (6), 605−610.

(22) Meylan, W. M.; Howard, P. H. Atom/Fragment Contribution Method for Estimating Octanol−Water Partition Coefficients. *J. Pharm. Sci.* **1995**, *84* (1), 83−92.

(23) Fischer, H.; Kansy, M.; Bur, D. CAFCA: A Novel Tool for the Calculation of Amphiphilic Properties of Charged Drug Molecules. *Chimia* **2000**, *54* (11), 640−45.

(24) Cruciani, C.; Crivori, P.; Carrupt, P. A.; Testa, B. Molecular Fields in Quantitative Structure−Permeation Relationships: The VolSurf approach. *THEOCHEM* **2000**, *503* (1−2), 17−30.

(25) Wang, R. X.; Gao, Y.; Lai, L. H. Calculating Partition Coefficient by atom-Additive Method. *Perspect. Drug Discovery Des.* **2000**, *19* (1), 47−66.

(26) Balducci, R.; Pearlman, R. S. Efficient Exact Solution of the Ring Perception Problem. *J. Chem. Inf. Comput. Sci.* **1994**, *34* (4), 822−831.

(27) Huckel, E. Quantum-Theoretical Contributions to the Benzene Problem. I. The Electron Configuration of Benzene and Related Compounds. *Z. Phys.* **1931**, *70*, 204−86.

(28) Huckel, E. Quantum Theory Treatment of the Benzene Problem. II. Quantum Theory of Induced Polarity. *Z. Phys.* **1931**, *72*, 310−37.

(29) 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* (20), 3714−7.

(30) Sayle, R. Physiological Ionization and pK_a Prediction. http://www.daylight.com/meetings/emug00/Sayle/pkapredict.html (accessed Aug 2005).

(31) Bondi, A. Van Der Waals Volumes + Radii. *J. Phys. Chem.* **1964**, *68* (3), 441−51.

(32) Richards, F. M. Areas, Volumes, Packing, and Protein Structure. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−76.

(33) Suomalainen, P.; Johans, C.; Soderlund, T.; Kinnunen, P. K. Surface Activity Profiling of Drugs Applied to the Prediction of Blood−Brain Barrier Permeability. *J. Med. Chem.* **2004**, *47* (7), 1783−8.

(34) Giardina, G.; Clarke, G. D.; Dondio, G.; Petrone, G.; Sbacchi, M.; Vecchietti, V. Selective kappa-Opioid Agonists: Synthesis and Structure−Activity Relationships of Piperidines Incorporating on oxo-Containing Acyl Group. *J. Med. Chem.* **1994**, *37*, 7 (21), 3482−91.

(35) Deconinck, E.; Zhang, M. H.; Coomans, D.; Vander Heyden, Y. Classification Tree Models for the Prediction of Blood−Brain Barrier Passage of Drugs. *J. Chem. Inf. Model.* **2006**, *46* (3), 1410−1419.

(36) Crivori, P.; Cruciani, G.; Carrupt, P. A.; Testa, B. Predicting Blood−Brain Barrier Permeation from Three-Dimensional Molecular Structure. *J. Med. Chem.* **2000**, *43* (11), 2204−16.

(37) 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* (8), 3017−28.

(38) Seelig, A.; Gatlik-Landwojtowicz, E. Inhibitors of Multidrug Efflux Transporters: Their Membrane and Protein Interactions. *Mini-Rev. Med. Chem.* **2005**, *5* (2), 135−51.

(39) Gatlik-Landwojtowicz, E.; Aanismaa, P.; Seelig, A. Quantification and Characterization of P-Glycoprotein−Substrate Interactions. *Biochemistry* **2006**, *45* (19), 3020−32.

(40) Delano, W. L. The PyMOL Molecular Graphics System. http://www.pymol.org (accessed Jan 2006).

(41) Sangster, J. Phase-Diagrams and Thermodynamic Properties of Binary Organic-Systems Based on 1,2-Diaminobenzene, 1,3-Diaminobenzene, 1,4-Diaminobenzene or Benzidine. *J. Phys. Chem. Ref. Data* **1994**, *23* (2), 295−338.
(42) Budavari, S. *The Merck Index an Encyclopedia of Chemicals, Drugs, and Biologicals*, 12th ed.; Merck: Whitehouse Station, NJ, 1996.
(43) Howard, P. H.; Meylan, W. M. *Handbook of Physical Properties of Organic Chemicals*; CRC Press: Boca Raton, FL, 1996.
(44) Peters, M. D., II; Davis, S. K.; Austin, L. S. Clomipramine: An Antiobsessional Tricyclic Antidepressant. *Clin. Pharm.* **1990**, *9* (3), 165−78.
(45) Hansch, C.; Sammes, P. G.; Taylor, J. B. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study, and the Therapeutic Applications of Chemical Compounds*; Pergamon Press: New York, 1990.
(46) Balon, K.; Riebesehl, B. U.; Muller, B. W. Drug Liposome Partitioning as a Tool for the Prediction of Human Passive Intestinal Absorption. *Pharm. Res.* **1999**, *16* (6), 882−8.
(47) Tollenaere, J. P.; Moereels, H.; Koch, M. H. J. Conformation of Neuroleptic Drugs in 3 Aggregation States and Their Conformational Resemblance to Dopamine. *Eur. J. Med. Chem.* **1977**, *12* (3), 199−211.
(48) Kontturi, K.; Murtomaki, L. Electrochemical Determination of Partition Coefficients of Drugs. *J. Pharm. Sci.* **1992**, *81* (10), 970−5.
(49) Drugs.com Doxylamine Succinate Advanced Consumer Drug Information. http://www.drugs.com/MMX/Doxylamine_Succinate.html (accessed Jan 2006).
(50) Ruell, J. A.; Tsinman, K. L.; Avdeef, A. PAMPA−A Drug Absorption in Vitro Model. 5. Unstirred Water Layer in iso-pH Mapping Assays and pK_a(flux)−Optimized Design (pOD-PAMPA). *Eur. J. Pharm. Sci.* **2003**, *20* (4−5), 393−402.
(51) Palermo, P. J.; Colucci, R. D.; Kaiko, R. F. Method of Preventing Abuse of Opioid Dosage Forms. U.S. Patent 6,228,863, 1998.
(52) de Boer, T.; Bijma, R.; Ensing, K. Modelling of Conditions for the Enantiomeric Separation of β2-Adrenergic Sympathicomimetics by Capillary Electrophoresis Using Cyclodextrins as Chiral Selectors in a Polyethylene Glycol Gel. *J. Pharm Biomed. Anal.* **1999**, *19* (3−4), 529−37.
(53) O'Neill, M. *The Merck Index: An Encyclopedia of Chemicals, Drugs and Biologicals*, 13th ed.; John Wiley & Sons: Whitehouse Station, NJ, 2001.
(54) Wan, H.; Holmen, A. G.; Wang, Y.; Lindberg, W.; Englund, M.; Nagard, M. B.; Thompson, R. A. High-Throughput Screening of pK_a Values of Pharmaceuticals by Pressure-Assisted Capillary Electrophoresis and Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, *17* (23), 2639−48.
(55) Yamaguchi, T.; Hashizume, T.; Matsuda, M.; Sakashita, M.; Fujii, T.; Sekine, Y.; Nakashima, M.; Uematsu, T. Pharmacokinetics of the H1-Receptor Antagonist Ebastine and its Active Metabolite Carebastine in Healthy Subjects. *Arzneimittelforschung* **1994**, *44* (1), 59−64.
(56) Crowe, A.; Wong, P. pH Dependent Uptake of Loperamide across the Gastrointestinal Tract: An in Vitro Study. *Drug Dev. Ind. Pharm.* **2004**, *30* (5), 449−59.
(57) Nielsen, P. E.; Du, C.; Tsinman, K. L.; Voloboy, D.; Avdeef, A. A New Technique for High-Throughput Solubility Assay. AAPS Annual Meeting, Boston, MA, 1997.
(58) Tam, K. Y.; Takacs-Novak, K. Multi-wavelength Spectrophotometric Determination of Acid Dissociation Constants: A Validation Study. *Anal. Chim. Acta* **2001**, *434* (1), 157−167.
(59) Barlow, R. B.; Chan, M. The Effects of pH on the Affinity of Pirenzepine for Muscarinic Receptors in the Guinea-Pig Ileum and Rat Fundus Strip. *Br. J. Pharmacol.* **1982**, *77* (3), 559−63.
(60) Pagliara, A.; Testa, B.; Carrupt, P. A.; Jolliet, P.; Morin, C.; Morin, D.; Urien, S.; Tillement, J. P.; Rihoux, J. P. Molecular Properties and Pharmacokinetic Behavior of Cetirizine, a Zwitterionic H1-Receptor Antagonist. *J. Med. Chem.* **1998**, *41* (6), 853−63.

CI0600814