Predicting Blood-**Brain Barrier Permeation from Three-Dimensional Molecular Structure**

Patrizia Crivori,[†] Gabriele Cruciani,*^{,‡} Pierre-Alain Carrupt,[†] and Bernard Testa[†]

Institute of Medicinal Chemistry, BEP, University of Lausanne, CH-1015 Lausanne-Dorigny, Switzerland, and Laboratory for Chemometrics, University of Perugia, Via Elce di Sotto 10, I-06123 Perugia, Italy

Received December 15, 1999

Predicting blood-brain barrier (BBB) permeation remains a challenge in drug design. Since it is impossible to determine experimentally the BBB partitioning of large numbers of preclinical candidates, alternative evaluation methods based on computerized models are desirable. The present study was conducted to demonstrate the value of descriptors derived from 3D molecular fields in estimating the BBB permeation of a large set of compounds and to produce a simple mathematical model suitable for external prediction. The method used (VolSurf) transforms 3D fields into descriptors and correlates them to the experimental permeation by a discriminant partial least squares procedure. The model obtained here correctly predicts more than 90% of the BBB permeation data. By quantifying the favorable and unfavorable contributions of physicochemical and structural properties, it also offers valuable insights for drug design, pharmacological profiling, and screening. The computational procedure is fully automated and quite fast. The method thus appears as a valuable new tool in virtual screening where selection or prioritization of candidates is required from large collections of compounds.

Introduction

To be effective as therapeutic agents, centrally acting drugs must cross the blood-brain barrier (BBB). Conversely, to be devoid of unwanted central nervous system (CNS) effects, peripherally acting drugs must show limited ability to cross the BBB. In both cases, the BBB permeability of drug candidates must be known. However, the experimental determination of brain-blood partitioning is difficult, time-consuming, and expensive and not suitable to screen large collections of chemicals.1 A broadly applicable method for predicting the BBB permeation of candidates at an early stage of discovery would have a great impact in drug research and development.

Entry into the brain is a complex phenomenon which depends on multiple factors. It is known that relatively lipophilic drugs can cross the BBB by passive diffusion as influenced by their H-bonding capacity. Polar molecules normally do not cross the BBB, but sometimes a process of active transport facilitates their penetration. Local hydrophobicity, ionization profile, molecular size, lipophilicity, and flexibility are other important parameters which play a role in BBB permeation.^{1,2} Not only is the number of accepted or donated H-bonds important but also their 3D distribution, due to the anisotropic nature of all biological membranes.3

Furthermore, plasma protein binding, active efflux from the CNS, and metabolism can also influence BBB penetration. Thus, unambiguous and reliable data on BBB permeation are difficult to find, since literature values are often uncertain and contradictory. CNSactive drugs (CNS+) can cross the BBB by different

mechanisms. For CNS-inactive drugs (CNS-) the situation is even more complex; some simply do not penetrate, whereas others are rapidly metabolized or expelled by active efflux processes.

Various authors^{1-2,4,5} have attempted to predict BBB transport using lipophilicity (log *P*), solvatochromic parameters, topological indexes, or combinations of these. To the best of our knowledge, no attempt has been to use descriptors derived from 3D molecular interaction fields. We have been involved in the development and validation of VolSurf, a method able to convert 3D fields into new descriptors well-suited for structure-pharmacokinetic relationships and have proven its efficacy, simplicity of use, and chemical interpretability.⁶ Here, we use this methodology to develop a quantitative model for BBB permeation. Such a model should be easy to use and to interpret and well-suited for external predictions.

Computational Methods

Database. The starting compounds analyzed in this study were the 40 agents presented by Giardina et al.,7 to which four compounds taken from the work of Shaw et al.8 were added (Table 1). The ability of the drugs to enter the brain was obtained by comparing their antinociceptive potency in the mouse abdominal constriction test (MAC),⁷ following subcutaneous and intracerebroventricular administration. A basic assumption in our study was passive permeation. The present dataset contains a number of related, but chemically diverse, compounds which are either brain-penetrating (BBB+), have a moderate permeation (BBB \pm), or have little if any ability to cross the blood-brain barrier (BBB-). All the reported compounds contain one or two chiral centers, and all except eight had been measured as racemates or mixtures of racemates.7 Because it is impossible to model mixtures of stereoisomers in 3D, the reported mixtures were modeled separately for their respective stereoisomers, assuming negligible stereoselectivity in permeation, a reasonable assumption for passive permeation. As a result, the number of compounds in the set increased from 44 to 110 (Table 1).

^{*} Correspondence author. Fax: +39 075 45646. E-mail: gabri@ chemiome.chm.unipg.it.

[†] University of Lausanne.

[‡] University of Perugia.

The compounds used as the test set for external prediction were also obtained from the literature. Instead of inspecting large directories of chemicals which require a subjective

automatable classification and where CNS activity is often reported but not BBB permeation, we used reliable data on experimental BBB permeation. Our literature search led to

Table 1. The 110 Compounds Used To Obtain the Training Set Model

Number	Name & BBB	Structure	Number	Name & BBB	Structure
1(S)	BRL 52537 $\ddot{}$		$20\;(\mathrm{R})$ 21(S)	GR 88377 $\ddot{}$	
$2\;(\mathrm{S})$	BRL 52580 $\ddot{}$		$22\;(\mathrm{R})$ $23(S)$	GR 89696 $\ddot{}$	
3(S)	BRL 52656 $\ddot{}$		24 (R) $25(S)$	GR 89696_et $\ddot{}$	
4(S)	BRL 52871 $\ddot{}$		26(R) $27 \, (\mathrm{S})$	GR 89696_pr	
5(R) 6(S)	BRL 52974 ä,		28 (RR) 29 (RS) 30 (SR) 31 (SS)	GR 91272 $\ddot{}$	
$7(R)$ 8(S)	BRL 53080 $\ddot{}$		32 (RR) 33 (RS) 34 ${\rm (SR)}$ 35 (SS)	GR 94839_A	
9(R) 10(S)	BRL 53087 $^{+}$		36 (RR) $37 \; (\mathrm{RS})$ 38 $\rm (SR)$ 39 (SS)	GR 94839_B	
11 $\left(\textrm{RRR}\right)$ 12 (RRS) 13 (SSR) 14 ${\rm (SSS)}$	Cyclazocine $\ddot{}$		40 (RR) 41 (RS) 42 (SR) 43 (SS)	GR 94839_C	HO
15 (SS)	EMD 60400	NH ₂ ö	44 (RR) 45 $\left(\text{RS}\right)$ 46 $\rm (SR)$ 47 (SS)	GR 94839_D	
$16(R)$ $17(S)$	GR 45809 $\ddot{}$	ö	48 (RR) 49 (RS) 50 $\left(\text{SR} \right)$ 51 $\rm (SS)$	GR 94839_E $\qquad \qquad \blacksquare$ \sim	
$18(R)$ $19(S)$	GR 85571 $\ddot{}$		$52\;(\mathrm{RR})$ 53 $\left(\text{RS}\right)$ 54 $\left(\text{SR} \right)$ $55~\mathrm{(SS)}$	GR 94839_F	

108 drugs with well-characterized BBB behavior. Here also, racemic drugs were modeled as separate enantiomers, resulting in a final test set of 120 compounds (Tables 2 and 3).

Computational Approach. The overall procedure contained the following four major steps:

(1) The three-dimensional structure of the compounds was built.

(2) The compounds were submitted to multivariate characterization based on their interaction energy with chemical probes. Molecular mechanics force fields or semiempirical methods as well as ab initio methods can be used for this purpose. Here, we used the GRID9 program to calculate the 3D molecular interaction fields.

(3) Molecular descriptors were calculated using the VolSurf program.6

(4) Chemometric tools (PCA, discriminant PLS) were used to correlate the data and build a BBB permeation model.

It is important to note that steps 2-4 were performed automatically by the VolSurf program.

Modeling of 3D Molecular Structures. It has been demonstrated that most VolSurf descriptors are only marginally influenced by conformational sampling.^{6,10} However, these results may depend on the set of compounds and their property space. To study the impact of conformation on the BBB model obtained, two different protocols were used to model the

compounds, each protocol being based on a different conformation for each compound.

In a first protocol, 2D-to-3D structure conversions were carried out by the program $Concord₁₁$ and the resulting conformations were refined by energy minimization with the Merck force field as implemented in Sybyl 6.4.¹²

In a second protocol the 3D structures were imported in the Amsol program¹³ and fully minimized using a AM1 semiem-

^a Descriptors 1-34 are generated with the water probe, descriptors 35-54 with the hydrophobic probe, and the remaining descriptors with the oxygen carbonyl probe.

pirical method¹³ including solvation effects. In the presence of multiple low-energy conformations, the most diverse molecular structures were selected in order to obtain the maximum variability in VolSurf descriptors.

All chemicals were modeled in their neutral form. The results of both protocols were analyzed as described below.

Multivariate Characterization of the Compounds by Their Molecular Interaction Fields. A molecular interaction field may be viewed as a 3D matrix, the elements of which (called grid nodes) are the attractive and repulsive forces between an interacting partner (the probe) and a target (the molecule or macromolecule). Most properties related to molecular interactions can be represented in a 3D molecular field. Well-known examples of a 3D molecular field are the molecular $electrostatic potential (MEP),¹⁴ the molecular lipophilicity$ potential (MLP),¹⁵ or GRID⁹ fields, one of the most widely used computational tools to map molecular surfaces of drugs and macromolecules.

The interaction of molecules with biological membranes is mediated by surface properties such as shape, electrostatic forces, H-bonds, and hydrophobicity. Therefore, the GRID force field, which uses a potential based on the total energy of interaction (the sum of Lennard-Jones, H-bonding and electrostatic terms) between a target molecule and a probe, was used to characterize putative polar and hydrophobic interaction sites around target molecules. The water probe was used to simulate solvation-desolvation processes, while the hydrophobic probe (called DRY in the GRID program) and the carbonyl probe (O) were used to simulate drug-membrane interactions. The DRY probe is a specific probe to compute the hydrophobic energy;¹⁶ the overall energy of the hydrophobic probe is computed at each grid point as *^E*entropy ⁺ *^E*LJ - *^E*HB, where $E_{entropy}$ is the ideal entropic component of the hydrophobic effect in an aqueous environment, *E*LJ measures the induction and dispersion interactions occurring between any pair of molecules, and E_{HB} measures the H-bonding interactions between water molecules and polar groups on the target surface.

Calculation of Molecular Descriptors. 3D molecular interaction fields can be automatically converted into simpler molecular descriptors using a procedure called VolSurf.6 The method is simple to apply and is specifically designed to produce descriptors *related to pharmacokinetic properties*, starting from 3D molecular field maps. In the standard procedure, GRID interaction fields are calculated around the target molecules.

The basic concept of VolSurf is to transform the information present in 3D molecular field maps into a limited number of quantitative numerical descriptors which are easy to understand and to interpret. In computational molecular graphics the use of molecular surfaces always involves partitioning the surface into small portions of tesserae or polyhedra in order to allow rendering, back-illumination, and other graphic effects. In such cases, a single feature and information, the molecular surface, is spread out into many small contiguous pieces of information. The VolSurf methodology proceeds in an exactly opposite manner. From many tesserae containing the same information, VolSurf builds a single framework (a volume and/or a surface) related to specific molecular properties.

The literature contains various algorithms to compute molecular descriptors derived from molecular surfaces and volumes.17 The originality of VolSurf lies in the fact that surfaces, volumes, and other related descriptors can be obtained *directly* from 3D molecular interaction fields without complex algorithms of trigonometric projections, recursive generations, and tessellation. Molecular recognition is achieved using software as in image analysis, but the image information is extracted by adding external chemical knowledge. VolSurf does so by selecting the most appropriate descriptors and parametrization according to the type of 3D maps under study.

VolSurf has the advantage of producing descriptors using the 3D information embedded in any map. Not all the information can be transferred from 3D to these new molecular descriptors, but practical examples do exist $6,10,18$ to show that relevant information is indeed extracted. Moreover, the Vol-Surf transformation is fast and the results are usually easy to interpret, as shown below. Also, the descriptors have a clear chemical meaning and are lattice-independent, and some of them can be projected back into the original 3D grid map to help interpretation.

The molecular descriptors obtained are reported in Table 4. They refer to molecular size and shape, to size and shape of hydrophilic and hydrophobic regions, and to the balance between them. Hydrogen bonding, amphiphilic moments, and critical packing parameters are other useful descriptors. The VolSurf descriptors have been presented and explained in detail elsewhere.⁶ However, a more detailed description of the *nonstandard* descriptor terms is reported in Table 5. It is important to note that VolSurf descriptors can be obtained for small, medium, and large molecules, as well as for biopolymers such as DNA sequences, peptides, and proteins.

Statistical Analysis. Principal component analysis (PCA)19 and partial least squares discriminant analysis $(PLS)^{20}$ are chemometric tools for extracting and rationalizing the information from any multivariate description of a biological system. Complexity reduction and data simplification are two of the most important features of such tools. PCA and PLS condense the overall information into two smaller matrixes, namely the score plot (which shows the pattern of compounds) and the loading plot (which shows the pattern of descriptors).²¹ Because the chemical interpretation of score and loading plots is simple and straightforward, PCA and PLS are usually preferred to other nonlinear methods, especially when the noise is relatively high.

Score and loading plots are interconnected so that any descriptor change in the loading plot is reflected by changes in the position of compounds in the score plot. Pairwise comparison can be made directly with interactive plots²² as developed in the VolSurf program,⁶ and the relative contributions to the property under study are shown in the related descriptors space.

PCA is a least-squares method and for this reason its results depend on data scaling. The initial variance of a column variable partly determines its importance in the model. To avoid this problem, column variables were scaled to unit variance before analysis.²³ The column average was then subtracted from each variable. From a statistical point of view, this corresponds to moving the multivariate system to the center of the data, which becomes the starting point of the mathematical analysis. The same autoscaling and centering procedures were applied to the PLS discriminant analysis.

Once the PCA model was developed, PCA predictions for new compounds or external test compounds were made by projecting the compound descriptors into the PCA model. This was made by calculating the score vector **T** of descriptors *X* and average *^x*j for the new compounds, using the loading *^P* of the PCA model, according to the following eq 1:

$$
\mathbf{T} = (X - \bar{x})P(PP)^{-1} \tag{1}
$$

For the PLS discrimination, external predictions were made using the following equation²⁴ (2) :

$$
Y = \bar{y} - \bar{x}P(PP)B\mathbf{Q} + XP(PP)^{-1}B\mathbf{Q}
$$
 (2)

where \bar{y} is the *Y* column average and **Q** is the loading vector for the *y* space and *B* the coefficient between the *X* and *Y* spaces.

Software Packages. Three-dimensional molecular structures were obtained from $Concord¹¹$ and minimized using Sybyl¹² and Amsol.¹³ Semiempirical calculations were made

with Amsol AM1aq.13 Molecular interaction fields were produced using the GRID v 17 program^{9,16} graphically interfaced with VolSurf 2.0. GRID is a free software for nonprofit institutions (contact peter@goodford.daemon.co.uk). Molecular descriptors, PCA, and discriminant PLS were produced with the VolSurf 2.0 program.6,25 VolSurf is a free program for nonprofit institutions, available at http://www.cgate.chm. unipg.it. [The dataset, the 3D molecular structures, the interaction fields, the Volsurf descriptors, and the BBB model are available at the VolSurf Web page.]

The GRID-VolSurf procedure is completely automated and, in contrast to previous versions of the interface, does not require any user assistance. For this reason, all the work can be handled and submitted in batch queue. The computation of molecular interaction fields and the subsequent computation of molecular descriptors using the GRID-VolSurf procedure is rather fast. For example, processing the overall database of 100 compounds used for external prediction, starting from 3D molecular structures, takes about 2 min at low resolution, and about 20 min at high resolution with a R10000 Silicon Graphics 270 MHz CPU.

Results and Discussion

First BBB Permeation Model. In a first investigation, we searched for a relation between the 3D structure and the BBB permeation of the dataset consisting in 110 compounds (Table 1) and 72 descriptors (Table 4). No biological input was given to the model. Three significant principal components (PCs) were found by a cross-validation technique. These components explained about 65% of the total variance of the matrix. The score plot for the first two PCs is reported in Figure 1, where the compounds are color-coded by their ability to cross the BBB.

Whereas the second and third PCs describe the chemical variability and spatial geometry, the first PC is able to discriminate between the BBB- (open circles), the BBB \pm (gray circles), and the BBB $+$ compounds (filled circles). Such a result appears of high interest given that no classification of compounds nor any training information was given to the PCA model, in contrast to training procedures such as neural networks.26,27 In other words, the first PC in the model was able, without any external information or training, to qualitatively and correctly predict the BBB permeation of the compounds.

Figure 1. PCA score plot for the compounds reported in Table 1. Filled circles represent the BBB+ compounds, gray circles the $BBB\pm$ compounds, and white circles the BBB- compounds. Numbers are those in Table 1. The first PC clusters the compounds according to their ability to cross the BBB.

A more detailed inspection of the score plot in Figure 1 indicates the misclassification of the two enantiomers of BRL52974 (compounds **5** and **6**). This compound has been reported as CNS-, its presence in the brain after single systemic administration being undetectable. In Figure 1, the two enantiomers appear near the borderline between BBB+ and BBB \pm compounds. The reason for BRL52974 to be an outlier may be due to a limitation of the statistical model or to a peculiar biological behavior (e.g., fast efflux or metabolism).

Figure 1 also shows that moderately penetrating compounds (BBB \pm) tend to cluster together and at the left of poorly penetrating compounds (BBB-).

The present model was obtained with the Amsol minimization protocol, but similar results were also obtained using the minimization procedure in Sybyl. This demonstrates that a conformational search is of modest relevance here, and that the faster method for 3D structure generation is also valid.

Before embarking on chemical interpretation and refinement, it was necessary to test the predictive capacity of the model using a large test set of compounds with well-documented BBB behavior.

External Prediction of the First Model. Tables 2 and 3 report the results of a literature search for drugs with a well-documented BBB behavior profile. Different chemical and pharmacological classes are included in these tables, e.g. antipsychotics, dopaminergics, antihistamines of first and second generation, anxiolytics, hypnotics, opioids, *â*-blockers, CCK antagonists, acetylcholinesterase inhibitors, quinolone antibacterials, NSAIDs, antiparkinsonians, iron chelators, and hormones. A total of 49 compounds are BBB+ and the remaining 71 BBB-. The overall 120 drugs were modeled like the training set.

The results for the BBB+ and BBB- compounds in the test set are shown in Figures 2 and 3 which are PC1

versus PC2 plots such as Figure 1. It clearly appears that the prediction of BBB permeation is quite satisfactory, with a correct classification of 90% of the BBB+ compounds (40 out of 44), and about 65% accuracy for the BBB- compounds (46 out of 71).

A detailed investigation of the results of the external predictions shows interesting findings. Figure 2 reports the BBB+ compounds projected in the PCA model made with the training set. The only clear outlier is thiopental, a very well-known anaesthetic, which is a very active CNS+ compound. According to our model, thiopental is to be considered a moderate BBB penetrating compound. However, experimental evidence shows²⁸ that thiopental can lead to changes in the permeability of the blood-brain barrier, thus crossing it by a different mechanism. Our model, being based on passive diffusion, appears to resolve the contradiction .

Caffeine is correctly located in the BBB+ region and it is close to the borderline of the moderately penetrating compounds. Caffeine enters the brain by a combination of transport mechanisms, namely passive diffusion and carrier-mediated transport.²⁹ The correct prediction suggests a predominantly passive diffusion for caffeine. The diffusion model also correctly predicts the permeation of a number of compounds. For example, the selective BBB permeation properties of the major steroid hormones are well-calculated and are coherent with the experimental data.30

Figure 3 reports the BBB- compounds projected on the same PCA model. It is clear that the prediction of BBB- compounds appears more difficult than that of BBB+ compounds. This is in agreement with the fact that CNS- drugs can be stopped by the BBB, metabolized before producing their effects, and/or readily expelled back by active efflux processes. Since our model is able to predict only passive diffusion, it is not surprising that predictions for BBB- compounds are

Figure 2. PCA predictions for the BBB+ compounds reported in Tables 2 and 3. Some of the apparently mispredicted compounds are labeled in the plot. The plot can be superimposed on Figure 1.

Figure 3. PCA predictions for the BBB- compounds reported in Tables 2 and 3. Some apparently mispredicted compounds are labeled in the plot. The plot can be superimposed to Figure 1.

apparently less reliable than for BBB+ compounds. Some of the mispredicted compounds are discussed below.

Mequitazine is a first-generation antihistaminic drug reported4 as BBB-. However, its great structural similarity with BBB+ antihistamines casts doubt on this classification. This drug was also an outlier in another study,⁴ thus reinforcing the doubts on its BBB profile.

Loperamide (compound 87 in Figure 3) is a welldocumented CNS- compound, but our model predicts a BBB+ behavior. Experimental data³¹ suggest that the lack of CNS activity of loperamide is due to its active removal from the brain by P-glycoprotein at the BBB. Such a result thus explains why our model failed to predict the BBB profile of loperamide. The same is true of asimadoline EMD6175332 (compound **54**).

Terfenadine (**108** and **109** in Figure 3) is a secondgeneration antihistamine with very low CNS activity.33 However, the model predicts terfenadine to be able to cross the blood-brain barrier. Assuming that the model is correct, how can terfenadine be explained to lack CNS effects? Experimental evidence shows³³ a fast metabolism of terfenadine to fexofenadine. Fexofenadine (**79** and **80**) is predicted by the model to be hardly able to cross the BBB (see Figure 3). This could therefore be the reason terfenadine shows very few CNS effects, even if by itself it could cross the blood-brain barrier. The same applies to ebastine (**77** in Figure 3), which is rapidly metabolized to carebastine (**64**), a zwitterionic

Figure 4. Discriminant PLS t1-t2 score plot for the global model. The model offers a good discrimination between the BBB+ and BBB- compounds, since it assigned a correct BBB profile to > 90% of the compounds. A confidence interval is built in the t1-t2 space, where BBB prediction can be borderline and doubtful. \bullet , BBB+ compounds; \circ , BBB- compounds.

compound well predicted by the model as being a BBBcompound.34

Although the model fails to correctly predict the BBB profile of a few compounds, the results of the external predictions can be considered as satisfactory. Indeed, there is other experimental evidence for the majority of mispredicted compounds, indicating that the predictions may in fact be correct. Moreover, our model is also able to predict the penetration trend of some hitherto poorly predicted compounds.

For example, the three benzodiazepines L663581, L365260, and L364718, the cholecystokinin (CCK) antagonists,35 and two hydroxylated metabolites of L663581 were compared with the BBB permeation of diazepam (a well-known BBB+ drug).³⁶ The sequence of experimental BBB permeation was diazepam > $L663581 \geq L364718 = L365260 \gg$ mono-OH-L663581 > di-OH-L663581,35 with the last two compounds as BBB-. The model correctly predicts this BBB permeation trend. Interestingly, this trend also confirms that lipophilicity is not the only or main property controling BBB penetration. Alprazolam and clobazam, two other benzodiazepines, have a lipophilicity similar to that of mono-OH- and di-OH-L668581.35 However, alprazolam and clobazam readily enter the brain, whereas the two metabolites are unable to do so. 35 As clearly demonstrated here, other factors than lipophilicity play an important role in the BBB transfer.

Another example of the utility of our model is seen in the ranking of antibacterial quinolones. Although opinions differ on the BBB transport of quinolones,³⁷ it is well-known that they exhibit a lower concentration in brain than in most other tissues, 38 implying a low BBB permeability. Indeed, our model classifies all quinolones as BBB-. However, compounds a60616, a56726, and difloxacin are close to the borderline of moderate BBB+ compounds, whereas the others are

predicted as increasingly stronger BBB- in the sequence rufloxacin, ciprofloxacin, pefloxacin, sparfloxacin, lomefloxacin, fleroxacin, enoxacin, norfloxacin, and ofloxacin. The predicted trend is in agreement with experimental results.38 It is interesting to note that the recently discovered 6-aminoquinolone antibacterials³⁹ are predicted to be less BBB penetrating than the correspondent 6-fluoroquinolones. In fact, the 6-amino derivative of ciprofloxacin³⁹ is predicted a much poorer BBB permeator than ciprofloxacin itself.

A Refined BBB Permeation Model. The two datasets containing the training set^{7,8} and the test set (except the major outlier mequitazine) were combined in a unique matrix containing 229 compounds and 72 descriptors. PLS discriminant analysis was carried out, assigning to the BBB+ compounds a categorical (not thermodynamic) score 1, and to the BBB- and BBB \pm compounds a score -1 . This procedure is equivalent to those used by other researchers^{26,27} when training neural nets. However, in contrast to procedures in neural networks, we checked the existence of relevant information on BBB permeation before building and training the model. It is suggested here that simple linear models such as PCA should always be used before training models, especially when nonlinear methods are used in combination with binary descriptors such as structural keys, fingerprints, or topological indices.

Two significant latent variables emerged from the PLS model and cross- validation. The PLS t1-t2 score plot of the resulting model is shown in Figure 4. The model distinguishes well between the BBB+ and BBBcompounds, better than the unbiased PCA model. Because this result was obtained by adding information to the model, it can be trusted only because a strong signal existed before training.

BBB evaluation is quite straightforward with the discriminant PLS. A BBB+ behavior was assigned to a

Figure 5. PLS coefficient plot for the global model (training and test sets combined) for the correlation of VolSurf descriptors with blood-brain barrier permeation. Shading refers to the different energy levels used (see Table 4).

compound having a score >0.0 , and a BBB- behavior to a compound with a score ≤ 0.0 . The model correctly assigned a BBB profile to more than 90% of the compounds. However, since the prediction error (SDEP)⁴⁰ of the discriminant PLS was 0.6 unit, a confidence interval was built in the $t1-t2$ space between the BBB+ and BBB- regions, as shown in Figure 4. In this interval, BBB prediction can be borderline and doubtful.

The coefficient plot of the model so obtained (Figure 5) reports the contribution of all VolSurf descriptors in Table 4. The vertical bars represent the contribution of each single descriptor, with a short bar being an unimportant descriptor and a long bar an important descriptor. The last bar on the right represents the biological response, here the BBB permeation. The conclusions to emerge from this plot are as follows:

Descriptors of polarity such as hydrophilic regions (W1-W8), capacity factors (C1-C8), and H-bonding (H1-H8) are inversely correlated with BBB permeability. W1-W8 descriptors refer to polar water-accessible surface areas (PWASA), indicating that BBB permeation decreases when the polar surface increases. This means that, besides H-bonding potential, other factors influence BBB permeation, e.g. charge distribution and electron lone pairs. Capacity refers to polar interactions per surface unit. While diffuse polar regions are tolerable for BBB permeation, dense and localized polar regions are markedly detrimental.

An increase in H-bonding capacity is known to be detrimental for permeation. In addition, the contribution of the integy moment (descriptors Iw1-Iw8) demonstrates that, besides the number of H-bonds, their 3D distribution also influences BBB permeation.

The descriptors of hydrophobic interactions $(D1-D8)$ are directly correlated with BBB permeation, but their role appears smaller than that of the polar descriptors.

The size and shape descriptors have no marked impact on BBB permeation. In contrast, critical packing (CP) and the hydrophilic-lipophilic balance (HL) are important descriptors.

Globally, it is the balance of all descriptors, in other words of molecular properties, which is seen to control BBB permeation.

As an example of the interpretative value of VolSurf, Figure 6 reports a visual comparison of the GRID 3D molecular fields of the BBB- compound SB204459 and of the BBB+ compound GR88377 (see Table 1), chosen because of their similar 2D structure. The gray zones around the molecules represent the hydrophilic regions. The arrows represent the vectors of the integy moments. The hydrophilic regions of SB204459 are larger than those of GR88377 and distributed in different regions of the 3D space. From Figure 5 we can deduce that the size of these regions is inversely correlated with BBB permeability, thus explaining why SB204459 is a poorly penetrating compound. The pattern of integy moment vectors for SB204459 is noticeably different from that of GR88377, demonstrating the importance of the location of the polar regions.

Conclusion

The good results of the PCA prediction and PLS discriminant models demonstrate that it is possible to predict BBB permeation from the 3D molecular structure of drug candidates.

The model interpretation is in good agreement with the known molecular factors influencing BBB permeation. In addition, and this outlines the originality of the method, VolSurf allows the relevant 3D molecular properties to be quantified. As such, VolSurf affords much structural information of use in designing BBB+ or BBB- candidates, and in defining an ideal property profile in similarity searches.

It is also of interest to note that VolSurf descriptors are independent of the alignment of molecules and relatively independent of conformational sampling and averaging. They are also fast to compute and easy to interpret. Since the determination of molecular descrip-

Figure 6. GRID 3D molecular fields of SB204459 (upper part) and GR88377 (lower part) calculated with a water probe. The zones shown are the hydrophilic regions contoured at -3 kcal/ mol. The arrows represent the integy moment's pattern calculated at eight energy levels as reported in Table 4.

tors and the generation of the model have been designed for easy transport, the method is accessible to all researchers and should help them in designing and selecting candidates.

Acknowledgment. We are grateful to Dr. M. Grugni (SmithKline Beecham Farmaceutici, Milano, Italy) for information on some of the compounds used in the test set. We also thank Prof. S. Clementi, Dr. M. Pastor, Dr. P. Benedetti, and Prof. R. Mannhold for valuable discussions and advice. B.T. and P.A.C. are indebted to the Swiss National Science Foundation for support.

References

- (1) ombardo, 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.
- (2) Abraham, M. H.; Chada, S.; Mitchell, R. Hydrogen Bonding. 33. Factors that Influence the Distribution of Solutes between Blood and Brain. *J. Pharm. Sci*. **¹⁹⁹⁴**, *⁸³*, 1257-1268.
- Mouritsen O. G.; Jorgensen K. A New Look at Lipid-Membrane Structure in Relation to Drug Research. *Pharm. Res.* **1998**, *15*,
- 1507–1519.
(4) Basak, S. C.; Gute, B. D.; Drewes, L. R. Predicting Blood–Brain
Transport of Drugs: A Computational Approach, *Pharm, Res* Transport of Drugs: A Computational Approach. *Pharm. Res.* **¹⁹⁹⁶**, *¹³*, 775-778.

Predicting Blood-*Brain Barrier Permeation Journal of Medicinal Chemistry, 2000, Vol. 43, No. 11* **²²¹⁵**

- (5) Young, 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 H2 Receptor Histamine Antagonists. *J. Med.*
- *Chem.* **¹⁹⁹⁸**, *³¹*, 656-671. (6) Cruciani, G.; Crivori, P.; Carrupt, P. A.; Testa, B. Molecular Fields in Quantitative Structure-Permeation Relationships: The VolSurf Approach. *Theochem* In press.
- (7) Gardina, G.; Clarke, G. D.; Grugni, M.; Sbacchi, M.; Vecchietti, V.; Central and Pheripheral Analgesic Agents: Chemical Strategies for Limiting Brain Penetration in Kappa-Opioid Agonists Belonging to Different Chemical Classes. *Farmaco* **1995**, *50*, ⁴⁰⁵-418. (8) Shaw, J. S.; Carrol, J. A.; Alcock, P.; Main, B. G. ICI204448: a
- k-Opioid Agonist with Limited Access to the CNS. *Br. J. Pharmacol.* **¹⁹⁸⁹**, *⁹⁶*, 986-992.
- (9) Goodford, P. J. Computational Procedure for Determining Energetically Favourable Binding Sites on Biologically Impor-
tant Macromolecules. J. Med. Chem. 1985, 28, 849–857. (b) tant Macromolecules. *J. Med. Chem.* **¹⁹⁸⁵**, *²⁸*, 849-857. (b) Bobbyer, D. N. A.; Goodford, P. J.; McWhinnie, P. M. New Hydrogen-Bond Potential for Use in Determining Energetically Favourable Binding Sites of Molecules of Known Structure. *J. Med. Chem.* **¹⁹⁸⁹**, *³²*, 1083-1094.
- (10) Guba, W.; Cruciani, G. A Novel Approach for the Multivariate Modeling of Pharmacological Data Using Molecular Field Derived Descriptors. In *Molecular Modelling and Prediction of Bioreactivity*; Guberrtofte, K., Jorgensen, F. S., Eds.; Plenum: New York, 2000; pp 89-95.
- (11) Pearlman, R. S. 3D Molecular Structure: Generation and Use in 3D Searching. In *3D QSAR in Drug Design*-*Theory Methods and Applications*; Kubyni, H. Ed.; Escom Science Publishers: Leiden; 1993, p 41-79.
- (12) SYBYL, Tripos Associates, St. Louis, MO.
- (13) Cramer, C. J.; Hawkins, G. D.; Lynch, G. C.; Giesen, D. J.; Rossi, I.; Storer, W. J.; Thrular, D. G.; Liotard, D. A. AMSOL: An SCF Program Incorporating Free Energy in Aqueous Solution and Semiempirical Charge Models. *QCPE Bull.* **1995**, *15*, 41.
- (14) Chirlian, L. E.; Francl, M. M. Atomic Charge Derived from Electrostatic Potentials: A Detailed Study. *J. Comput. Chem.* **¹⁹⁸⁷**, *⁸*, 894-905. (15) Carrupt, P. A.; Gaillard, P.; Billois, F.; Weber, P.; Testa, B.;
- Meyer, C.; Perez, S. The Molecular Lipophilicity Potential (MLP): A New Tool for logP Calculations and Docking, and in Comparative Molecular Field Analysis (CoMFA). In *Lipophilicity in Drug Action and Toxicology*; Pliska, V., Testa, B., van de Waterbeemd, H., Eds.; VCH: Weinheim, **¹⁹⁹⁵**, pp 195-215.
- (16) GRID v.17, Molecular Discovery Ltd. Oxford, GB, 1999.
- (17) Connolly, M. L. *http:// www.netsci.org/ Science/Comp-chem/ feature14.html.*
- (18) Mannhold, R.; Cruciani, G.; Weber, H.; Lemoine , H.; Derix, A.; Weichel, C.; Clementi, M. 6-Substituted Benzopyrans as Potassium Channel Activators: Synthesis, Vasodilator Properties, and Multivariate Analysis. J. Med. Chem, 1999, 42, 981-991. Multivariate Analysis. *J. Med. Chem*, **¹⁹⁹⁹**, *⁴²*, 981-991. (19) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis.
- *Chem. Intell. Lab. Syst.* **¹⁹⁸⁷**, *²*, 37-52.
- (20) Dunn, W. J.; Wold, S. Pattern Recognition Techniques in Drug Design. In *Comprehensive Medicinal Chemistry*; Vol. 4, Hansch, C., Sammes, P. G., Taylor, J. B., Eds.; Pergamon Press: Oxford, **¹⁹⁹⁰**; p 691-714.
- (21) Cruciani, G.; Clementi, S. GOLPE: Philosophy and Applications in 3D-QSAR. In *Advanced Computer-Assisted Techniques in Drug Discovery*. van de Waterbeemd, H. Ed.; VCH: Weinheim, **¹⁹⁹⁴**, p 61-88.
- (22) Cruciani, G.; Clementi, S.; Baroni, M.; Pastor, M. Recent Development in 3D-QSAR Methodologies. In *Rational Molecular Design in Drug Research*; Alfred Benzon Symposium 42, Liljefors, T., Jorgensen, F. S., and Krogsgaard-Larsen, P., Eds.; Munskgaard: Copenhagen, 1998, p 87-97.
- (23) Wold, S.; Albano, C.; Dunn, W. J., III; Edlund, U.; Esbensen, K.; Geladi, P.; Helberg, S.; Johansson, E.; Lindberg, W.; Sjostrom, M. Multivariate Data Analysis in Chemistry. In *Chemometrics Mathematics and Statistics in Chemistry*. Kowalsky, B. R. Ed.; Dordrecht: Holland, 1983; p 17-96. (24) Clementi, S.; Cruciani, G.; Curti, G.; Skagerberg, B.; PLS
- Response Surface Optimisation: The CARSO Procedure. *J. Chemom.* **¹⁹⁸⁹**, *³*, 499-509. (25) VolSurf v.2.0, Multivariate Infometrics Analysis, Perugia, Italy
- 1999.
- (26) Sadowski, J.; Kubinyi, H. A Scoring Scheme for Discriminating Between Drugs and Nondrugs. *J. Med. Chem.* **¹⁹⁹⁸**, *⁴¹*, 3325- 3329.
- (27) Ajay, W.; Walters, P.; Murckp, M. A. Can We Learn to Distinguish Between Drug Like and Nondrug Like Molecules? *J. Med. Chem*. **¹⁹⁹⁸**, *⁴¹*, 3314-3324.
- (28) Fischer, S.; Renz, D.; Schaper, W.; Karliczek, G. F. In Vitro Effects of Fentanyl, Methohexital, and Thiopental on Brain Endothelial Permeability. *Anesthesiology* **¹⁹⁹⁵**, *⁸²*, 451-458.
- (29) McCall, A. L.; Mellington, W. R.; Wurtman, R. J. Blood-Brain Barrier Transport of Caffeine: Dose-Related Restriction of Adenine Transport. *Life Sci.* **¹⁹⁸²**, *³¹*, 2709-2715. (30) Pardrigde W. M. Transport of Protein-Bound Hormones into
- Tissues in Vivo. *Endocr. Rev.* **¹⁹⁸¹**, *²*, 103-123. (31) Schinkel, A. H.; Wagenaar, E.; Carla A. A. M.; van Deemter, L.
- P-Glycoprotein in the Blood-Brain Barrier of Mice Influences the Brain Penetration and Pharmacological Activity of Many Drugs. *J. Clin. Invest.* **¹⁹⁹⁶**, *⁹⁷*, 2517-2524.
- (32) Jonker, J. W.; Wagenaar, E.; van Deemter, L.; Gottschlich, R.; Bender, H. M.; Dasenbrock, J.; Schinkel, A. H. Role of Blood-Brain Barrier P-glycoprotein in Limiting Brain Accumulation and Sedative Side-Effects of Asimadoline, a Peripherally Acting Analgesic Drug. *Br. J. Pharmacol.* **¹⁹⁹⁹**, *¹²⁷*, 43-50.
- (33) Brown, A.; Griffiths, R.; Harvey, C. A.; Owen, D. A. A. Pharma-cological Studies with SK&F93944 (Temelastine), a Novel Histamine H₁-receptor Antagonist with Negligible Ability to Penetrate the Central Nervous System. *J. Pharmacol.* **1986**, *87*,
- ⁵⁶⁹-578. (34) Yamaguchi, T.; Hashizume, T.; Matzuda, M.; Sakashita, M.; Fujii, T.; Sekine, Y.; Nakashima, M.; Uematsu, T. Pharmacokinetics of the H1-Receptor Antagonist Ebastine andd its Active Metabolite Carebastine in Healthy Subjects. *Arzneim.-Forsch.*
- *(Drug Res.)* **¹⁹⁹⁴**, *⁴⁴*, 59-64. (35) Lin, H. J.; Chen, I.-Wu.; Lin, T.-H. Blood-Brain Barrier Permeability and in Vivo Activity of Partial Agonists of Benzazepines Receptor: A Study of L-663, 581 and Its Metabolites in Rats. *J. Pharmacol. Exp. Ther.* **¹⁹⁹⁴**, *²⁷¹*, 1197-1202.
- (36) Jones, D. R.; Hall, S. D.; Jackson, E. K.; Branch, R. A.; Wilkinson, G. R. Brain Uptake of Benzodiazepines: Effects of Lipophilicity and Plasma Protein Binding. *J. Pharmacol. Exp. Ther.* **1988**,
- *²⁴⁵*, 816-822. (37) Ooie, T.; Terasaki, T.; Suzuki, H.; Sugiyama, Y. Kinetic Evidence for Active Efflux Transport Across the Blood-Brain Barrier of Quinolone Antibiotics. *J. Pharmacol. Exp. Ther.* **¹⁹⁹⁷**, *²⁸³*, 293- 304.
(38) Jaehde, U.; Goto, T.; de Boer, A. G.; Breimer, D. D. Blood–Brain
- (38) Jaehde, U.; Goto, T.; de Boer, A. G.; Breimer, D. D. Blood-Brain Barrier Transport Rate of Quinolone Antibacterials Evaluated in Cerebrovascular Endothelial Cell Cultures. *Eur. J. Pharm.*
- *Sci.* **¹⁹⁹³**, *¹*, 49-59. (39) Cecchetti, V.; Filipponi, E.; Fravolini, A.; Tabarrini, O.; Bonelli, D.; Clementi, M.; Cruciani, G.; Clementi, S. Chemometric Methodologies in a Quantitative Structure-Activity relationship Study: The Antibacterial Activity of 6-Aminoquinolones. *J. Med.*
- *Chem.* **¹⁹⁹⁷**, *⁴⁰*, 1698-1706. (40) Cruciani, G.; Baroni, M.; Clementi, S.; Costantino, G.; Riganelli, D.; Skagerberg, B. Predictive Ability of Regression Models. Part I: Standard Deviation of Prediction Errors (SDEP). *J. Chemom.*
- **¹⁹⁹²**, *⁶*, 335-346. (41) Lin, T. H.; Lin, J. H. Effect of Protein Binding and Experimental Disease States on Brain Uptake of Benzazepines in Rats. *J. Pharmacol. Exp. Ther.* **¹⁹⁹⁰**, *²⁵³*, 45-50.
- (42) Aasmundstad, T. A.; Morland, J.; Paulsen, R. R. Distribution of Morphine 6-Glucuronide and Morphine Across the Blood-Brain Barrier in Awake, Freely Moving Rats Investigated by in Vivo Microdialysi Sampling. *J. Pharmacol. Exp. Ther.* **1995**, *275*, ⁴³⁵-441.
- (43) Habgood, M. D.; Liu, Z. D.; Dehkordi, L. S.; Khodr, H. H.; Abbott, J.; Hider, R. C. Investigation Into the Correlation Between the Structure of Hydroxypyridinones and Blood-Brain Barrier Permeability. *Biochem. Pharmacol.* **¹⁹⁹⁹**, *⁵⁷*, 1305-1310.
- (44) Tse, F. L. S.; Laplanche, R. Absorption, Metabolism, and Disposition of [14C]SDZ ENA 713, an Acetylcholinesterase Inhibitor, in Minipigs Following Oral, Intravenous, and Dermal Administration. *Pharm. Res.* **¹⁹⁹⁸**, *¹⁵*, 1614-1620.
- (45) Calcutt, C. R.; Ganellin, C. R.; Griffiths, B. K.; Maguire, L. J. P.; Mitchell, R. C.; Mylek, M. E.; Parsons, M. E.; Smith, I. R.; Young, R. C. Zolantidine (SK&F95282) Is a Potent Selective Brain-Penetrating Histamine H2-Receptor Antagonist. *Br. J. Pharmacol.* **¹⁹⁸⁸**, *⁹³*, 69-78.
- (46) Jolliet, P.; Simon, N.; Brée, F.; Urien, S.; Pagliara, A.; Carrupt, P. A.; Testa, B.; Tillement, J.-P. Blood-to-Brain Transfer of Various Oxicams: Effects of Plasma Binding on Their Brain Delivery. *Pharm. Res.* **¹⁹⁹⁷**, *¹⁴*, 650-656.
- (47) Raveglia, L. F.; Giardina, G. A. M.; Grugni, M.; Hay, D. W.; Farina, C.; Graziani, D.; Luttman, M. A.; Potts, W. M.; Salter, C. J.; Discovery of Potent and Selective, Non Brain Penetrant NK-3 Receptor Antagonists. Abstract of the First Italian-Swiss Meeting on Medicinal Chemistry, Torino, Italy, Abstract, 1997, p 49.
- (48) Babe, S. K., Jr.; Serafin, W. E. Histamine, Bradykinin, and their Antagonists. In *Goodman and Gilmans. The Pharmacological Basis of Therapeutics*, 9th ed.; Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Gilman A. G., Eds.; New York, 1996; p 581-600.
- (49) Pagliara, A.; Testa, B.; Carrupt, P.-A.; Jolliet, P.; Morin, C.; Morin, D.; Urien, S.; Tillement, J.-P.; Rihoux, J.-P. Molecular Properties and Phamacokinetic Behavior of Cetirizine, a Zwitterionic H1-Receptor Antagonist. *J. Med. Chem.* **¹⁹⁹⁸**, *⁴¹*, 853- 863.
- (50) Gonzàlez, M. A.; Estes, K. S. Pharmacokinetic Overview of Oral Second-Generation H1 Antihistamines. *Int. J. Clin. Pharmacol.*
- *Ther.* **¹⁹⁹⁸**, *³⁶*, 292-300. (51) Hoffman, B. B.; Lefkowitz, R. J. Catecholamines Sympathomimetic Drugs, and Adrenergic receptor Antagonists. In *Goodman and Gilmans. The Pharmacological Basis of Therapeutics,* 9th ed.; Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Gilman A. G., Eds.; New York, **¹⁹⁹⁶**; p 199-248.
- (52) Standaert, D. G.; Young, A. B. Treatment of Central Nervous System Degenerative Disorders. In *Goodman and Gilmans. The Pharmacological Basis of Therapeutics*, 9th ed. Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Gilman A. G., Eds. New York, **¹⁹⁹⁶**, pp 503-519.
- (53) Cecchetti, V.; Fravolini, A. personal comunication.
- (54) Baudry, S.; Pham, Y. T.; Baune, B.; Vildrequin, S.; Crevoisier, C. H.; Gimenez, F; Farinotti, R. Stereoselective Passage of Mefloquine Through the Blood-Brain Barrier in the Rat. *J. Pharm. Pharmacol.* **¹⁹⁹⁷**, *⁴⁹*, 1086-1090.

JM990968+