In silico prediction of
blood-brain barrier permeation
and P-glycoprotein activity

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Grégori Gerebtzoff

aus

Liège, Belgium

Basel, 2006

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

Auf Antrag von

Prof. Dr. Joachim Seelig

Prof. Dr. Anna Seelig

Prof. Dr. Torsten Schwede

Basel, den 04.07.06

<div align="right">

Dekan:

Prof. Dr. Hans-Jakob Wirz

</div>

Computers are incredibly fast, accurate and stupid.
Human beings are incredibly slow, inaccurate and brilliant.
Together they are powerful beyond imagination.

A. Einstein

I tried to do my best...

Contents

List of Figures

List of Tables

Introduction

P-glycoprotein is an ATP-dependent efflux transport protein which is highly expressed in many human tissues such as the intestinal epithelium and the blood-brain barrier, and is over-expressed in many cancer cells.[1] This transporter carries a wide variety of chemically unrelated compounds. It binds them within the cell lipid membrane, and flips them to the outer leaflet or exports them to the extracellular medium.[2] Since P-glycoprotein affects the distribution of many drugs, assessing the interactions between drugs and P-glycoprotein at an early stage of drug development is important.

It has been shown that the binding of a drug to the transporter occurs in a two-step process.[3–5] (i) The drug partitions from the extracellular environment to the lipid membrane, and after diffusion to the inner cytosolic leaflet of the bilayer (Figure 1, arrow **1**), (ii) it binds to P-glycoprotein most likely *via* hydrogen bond formation (Figure 1, arrow **2**).
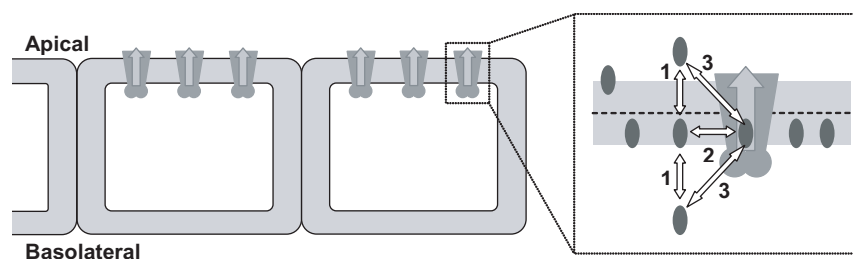
Figure 1: Two-step transport mechanism of P-glycoprotein. The drug has first to partition from extracellular or intracellular aqueous environment into the inner leaflet of the lipid bilayer (**1**), and then from the lipid to the transporter (**2**) (taken from ref.[6]).

Different methods have been used to assess the lipid-water partition coefficient, such as isothermal titration calorimetry, and lipid monolayer insertion measurements. However, the lipid-water partition coefficient depends on the lipid used, and in turn on the lateral packing density of the lipid layer. Therefore an approach based on surface activity measurements was developed, which allows the prediction of the lipid-water partition coefficient for membranes of different lateral packing densities.[7] Measurements of the surface pressure of

the drug in buffer solution as a function of concentration (Gibbs adsorption isotherm) yields the air-water partition coefficient (K_{aw}), the critical micellar concentration (CMC), and the cross-sectional area of the compound (A_D), provided experiments are performed under conditions of minimal electrostatic repulsion. Since air has a dielectric constant close to that of the lipid core region of a membrane, there is a direct relationship between the partition of a drug into the air-water interface, and the partition into the lipid-water interface.[8] The cross-sectional area, as well as the lipid-water partition coefficient (and by extension the air-water partition coefficient), are thus crucial parameters to assess the binding and diffusion of a drug into a lipid bilayer.

In a first part of the thesis, I focused on the membrane binding step. Since the cross-sectional area of a compound is a crucial parameter for drug partitioning into the lipid bilayer, the quality of the data obtained by mean of surface activity measurements are most important. For this purpose, in a first step, I improved the calibration of the experimental settings, by assessing several factors like the evaporation or the solvent effect. In a second step, I developed computer routines for unbiased evaluation of these measurements. In a third step, I developed an algorithm to calculate the cross-sectional area of a compound oriented at a hydrophilic-hydrophobic interface; this algorithm has been calibrated on a set of measured data, in order to find from a conformational ensemble the conformation of the membrane-bound drug.

In a second part of the thesis, I focused on the binding of a drug to P-glycoprotein. P-glycoprotein is monitored essentially by three types of assays, (i) the measurement of ATP hydrolysis activity of the transporter, (ii) a competition assay against calcein-AM, and (iii) a transcellular transport assay through polarized P-glycoprotein over-expressing cell monolayer. Based on a modular binding approach to assess the two-step binding of a drug to P-glycoprotein (Figure 1),[5] I developed several rules to predict the outcome of these experimental assays. Each rule, predicting one particular assay, has been tested on experimental datasets.

In a third part of the thesis, I developed a working interface to handle multiple structures of compounds, to calculate the new descriptors involved in the two-step binding of drugs to P-glycoprotein (membrane partitioning, and binding to the transporter), and to calculate the outcome of the prediction rules. Moreover the working interface has been designed in a way the user can easily define new rules, or even introduce a new multidrug transporter (e.g. the multidrug transporter MRP1).

Starting from well characterized physical-chemical parameters, I developed a coherent ensemble of descriptors to assess by a rule-based approach the thermodynamics and kinetics of P-glycoprotein activation. This ensemble has been embedded in a customizable working interface, allowing easy evaluation of the *in silico* predictions.

Chapter 1

Assessment of the cross-sectional area

1.1 Introduction

1.1.1 Experimental part

For accurate use of the two monolayer Langmuir troughs (20 mL and 3 mL) a calibration is required. It has been shown[9] that several parameters have to be taken into account, like the evaporation of the liquid, the influence of the injected volume and the influence of the solvent in which the tested compound is dissolved. These factors have an important impact on the measured surface pressure. By making measurements over a broad concentration range, or successive measurements with different concentrations containing overlapping regions, we noted that the corrections applied to the data obtained in the 3 mL trough were not optimal. Indeed, as depicted in Figure 1.5, we could not see an overlap in the surface pressure for two identical concentrations when two different stock solutions with a low and a high concentration were used, especially for large injected volumes.

In order to enhance the accuracy of the data evaluation, we improved the correction factors with respect to the following points:

- temperature dependant evaporation of the buffer,

- influence of the total volume (buoyancy) on the measured surface pressure, and

- influence of the solvent (water, methanol, DMSO) on the measured surface pressure.

1.1.2   Computational part

The need of a software for a more effective evaluation of the monolayer data has become important with the multiplication of the measurements made in both 3 mL and 20 mL troughs. For this purpose, we developed a way to automatize the selection of the quasi-linear part of the $\pi = f(log(C))$ plot (Gibbs adsorption isotherm), in order to remove the bias inherent to the manual selection of these points.

We then finalized a method to determine the *critical micellar concentration* (CMC) in an automatical way.

Finally, we improved the determination of the averages of the measured surface pressures and the injection times of the raw data given by the automatic recorder coupled to the 3 mL trough.

1.2   Experimental part: calibration of the monolayer troughs

1.2.1   Correction factor for the water evaporation

Fischer[9] showed that in the 20 mL trough, the injected amount exactly compensates the evaporation, as long as the injections are made every 30 minutes. For the 3 mL trough, we could see that the evaporation and the injected volume did not compensate each other. In order to estimate the evaporation in the 3 mL trough, we placed the trough covered with a plastic hood and filled with 3 mL of buffer solution (50 mM TRIS, 114 mM NaCl, pH 7.4) on an analytical balance at room temperature (21.5 °C), with a standard filter paper plunged into the buffer to mimic the usual conditions of experiment. Figure 1.1(a)

(a)   (b)
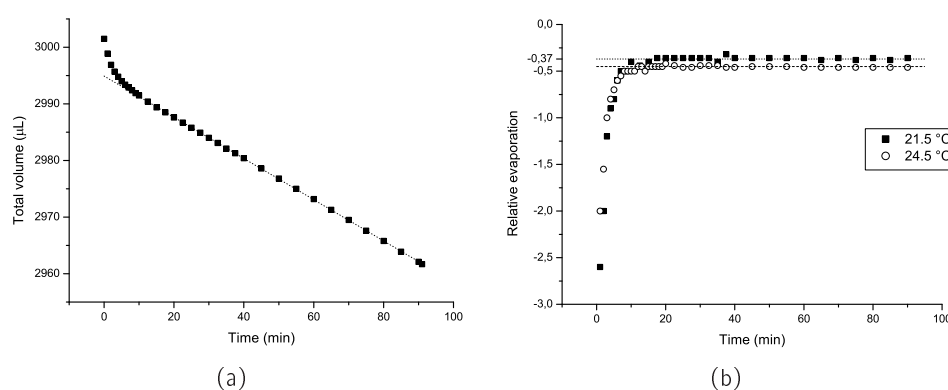
Figure 1.1: Evaporation in the 3 mL monolayer trough: (a) shows the absolute volume decrease with the time at 21.5 °C; (b) shows the relative evaporation of the water.

shows the linear decrease of the total volume with time (a measurement was made every minute during 20 minutes and every five minutes during additional 70 minutes). By calculating the relative evaporation per minute:

$$\text{relative evaporation} \ = \frac{1}{2} \cdot \left(\frac{V_{i+1} - V_i}{t_{i+1} - t_i} + \frac{V_i - V_{i-1}}{t_i - t_{i-1}} \right), \qquad (1.1)$$

where V_i is the measured volume at time t_i, V_{i+1} the volume at time t_{i+1}, and V_{i-1} the volume at time t_{i-1}, it appears that the volume decrease in the trough is constant after 15 minutes, as shown in Figure 1.1(b). The average evaporation was estimated as -0.37 μL/min at 21.5 °C (dotted line) and -0.45 μL/min at 24.5 °C (dashed line) over a period of 90 minutes (average duration of an experiment).

The influence of the presence of a monolayer of drugs on the surface of the buffer on the evaporation was also measured. As an example for a surface active compound, we used the local anesthetic dibucaine and applied it at a concentration of 1 mM (by adding 100 μL of a stock solution of 31 mM dibucaine prepared in pure water in 3 mL of buffer solution), which elicits around half maximum surface pressure. The negative control was 100 μL of pure water in 3 mL of buffer solution. The results of this experiment, shown in
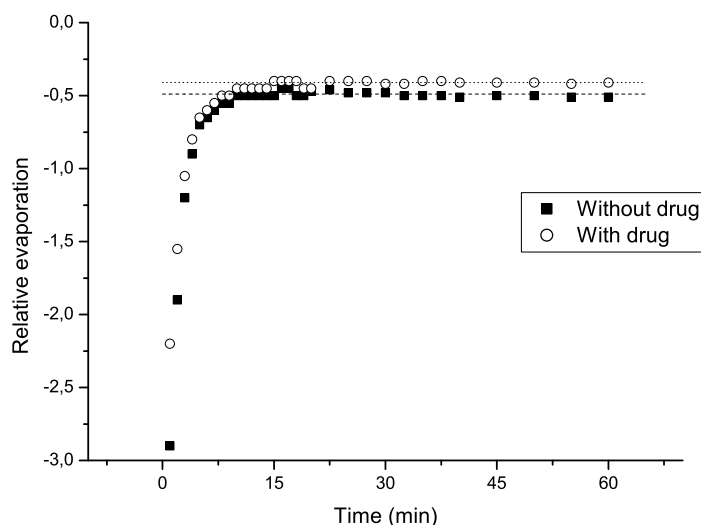
Figure 1.2: Relative evaporation in the 3 mL monolayer trough without (■) and with (○) 1 mM of dibucaine

Figure 1.2, demonstrate that there is a perceptible influence of the presence of a drug monolayer at the air-water interface on the water evaporation. It is, however, small enough to be neglected.

The influence of stirring the buffer on evaporation was determined by using the same setup as before but applying in addition the stirring device. To

avoid disturbances by vibrations, the experiment was performed outside of the analytical balance because of its high sensitivity. The weight of the trough was measured every 5 minutes during 65 minutes. The measurement was performed twice at 23.5 °C, with 100 μL of pure water (control) and with 100 μL of dibucaine (final concentration: 1 mM). The results are shown in Figure 1.3. The average evaporation without drug (-0.304) is very close to
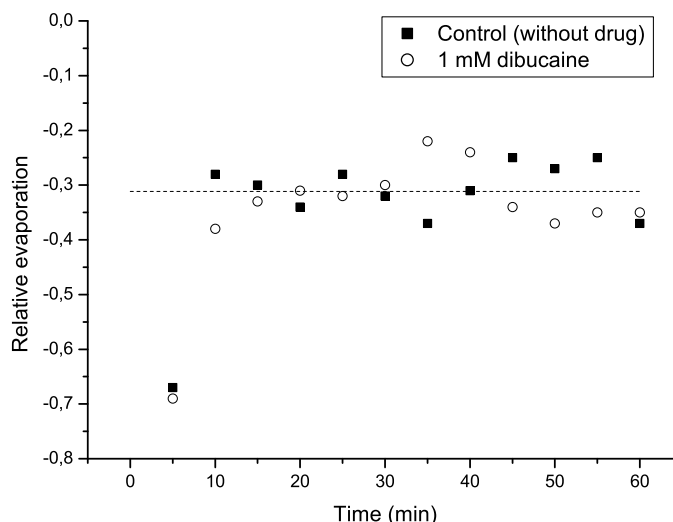
Figure 1.3: Relative evaporation in the 3 mL monolayer trough under stirring condition, without (■) and with (○) 1 mM of dibucaine. The dashed line corresponds to the average of all values.

the one with drug (-0,319), confirming the negligible impact of a monolayer of surface-active compound at the air-water interface of the 3 mL trough. Interestingly, the stirring seems to decrease the evaporation by 30%.

As a consequence of these observations, the only parameters taken into account for the evaporation are:

temperature of the local environment (room temperature), and

elapsed time between each injection.

These two parameters are important for the accuracy of our measurements, but the total volume and the solvent in which the compounds are dissolved are even more crucial as shown below.

1.2.2 Correction factor for the liquid level

We then tested the influence of the injected amount of a stock solution on the surface pressure, π, particularly in the 3 mL trough. An increase of the volume

will lead to a change in the buoyancy force exerted on the filter paper. The impact of the variation of total volume on the measured surface pressure was determined by adding increasing amounts of water to 3 mL of buffer (50 mM TRIS, 114 mM NaCl, pH 7.4) every 5 minutes. Figure 1.4 shows that the
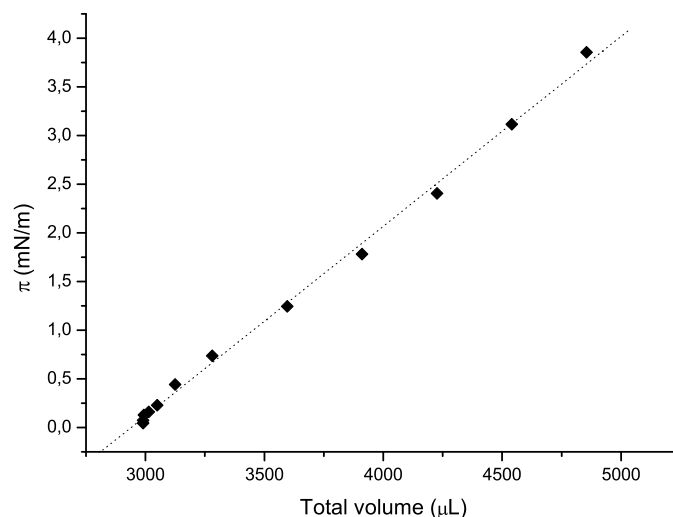
Figure 1.4: Effect of addition of water to 3 mL of buffer on the measured surface pressure.

increase of the total volume has a linear correlation with the increase of the measured surface pressure in the 3 mL trough. We went up to an addition of 1920 μL, a volume which is never reached experimentally. The linear fit of these data points leads to equation 1.2:

$$\pi = 1.976 \cdot 10^{-3} \cdot V - 5.928. \tag{1.2}$$

The previously defined evaporation parameter was applied to the measured surface pressure. This equation will be used as an additional correction factor for the measured surface pressure.

These correction factors for evaporation and liquid level lead to a better evaluation of the surface pressure measurements, as shown in Figure 1.5: the data evaluated with the new corrections (open symbols) are much closer to the linear fit (dotted line) of the surface pressure measurements than the data evaluated previously.[9]

1.2.3 Correction factor for the solvent

Many compounds are not well soluble in water, and therefore stock solutions have to be prepared in an organic solvent, usually methanol or DMSO. We measured the effect of methanol in the 20 mL trough (T. Alt and X. Li-Blatter, unpublished results), as well as in the 3 mL through. The troughs
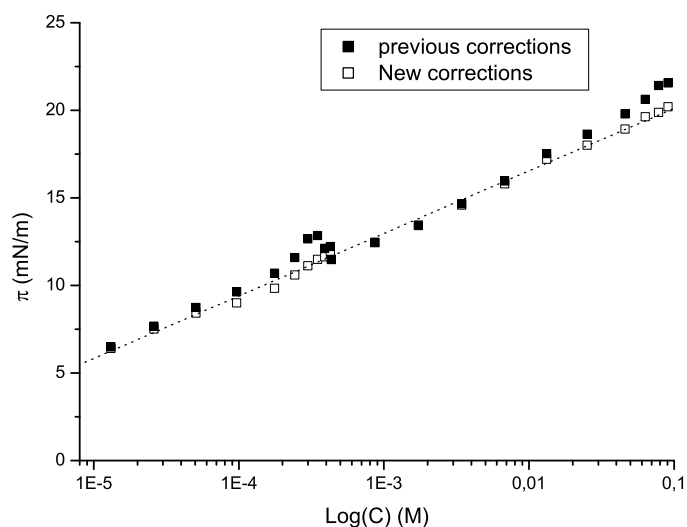
Figure 1.5: Comparison between previous[9] corrections (■) and the new defined corrections (□) for the $\pi = f(log(C))$ plot of methyl-cyclodextrin in the 3 mL trough; the dotted line represents the linear fit of the new corrected data.

were filled with buffer (50 mM TRIS, 114 mM NaCl, pH 7.4). The results are shown in Figure 1.6. Both sets of data could be fitted with the extended Langmuir model:

$$\pi = \left(\frac{a \cdot b \cdot V^{1-c}}{1 + b \cdot V^{1-c}} \right), \tag{1.3}$$

where π is the surface pressure in mN/m, V, the total injected volume in μL and a, b and c are variable parameters. The fit parameters obtained were $a = 45.92$, $b = 1.92 \cdot 10^{-3}$, $c = 6.91 \cdot 10^{-2}$ and $a = 11.66$, $b = 2 \cdot 10^{-5}$, $c = -5.81 \cdot 10^{-1}$ for the 3 mL (Figure 1.6(a)) and the 20 mL trough (Figure 1.6(b)), respectively. To check the relevancy of the correction factor, we compared previous results obtained with the same compound (amitriptyline, chlorpromazine) dissolved either in water or in methanol (data not shown). We obtained identical $\pi = f(log(C))$ plots by subtracting the correction factor from the methanol data set.

For the DMSO, the same approach was made, with the same buffer. Data for the 20 mL trough measurement were taken from a previous experiment (T. Alt and X. Li-Blatter, unpublished results). The two results are shown in Figure 1.7(a) for the 3 mL trough and in Figure 1.7(b) for the 20 mL trough. The DMSO effect on the 3 mL trough was also fitted with equation 1.3. However, the effect of the DMSO on the 20 mL trough could not be fitted with the extended Langmuir model; the only satisfying fitting model was the
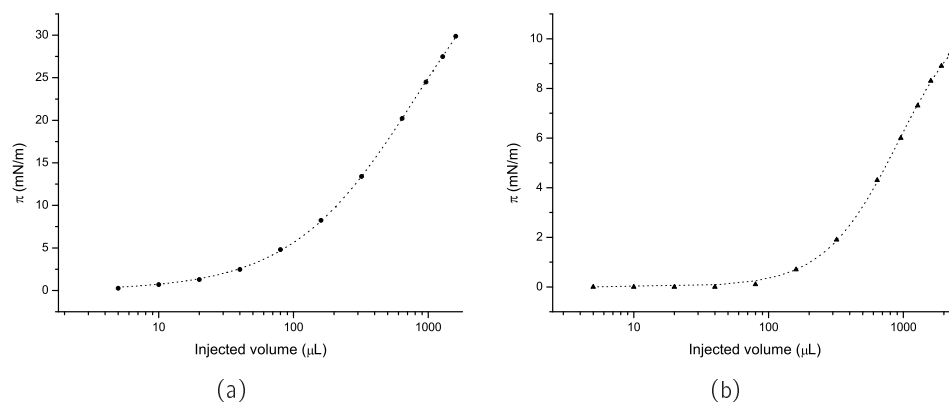
Figure 1.6: Effect of methanol on the measured surface pressure: (a) shows the effect on the 3 mL trough; (b) shows the effect on the 20 mL trough; the dotted curves represent the extended Langmuir model fit.
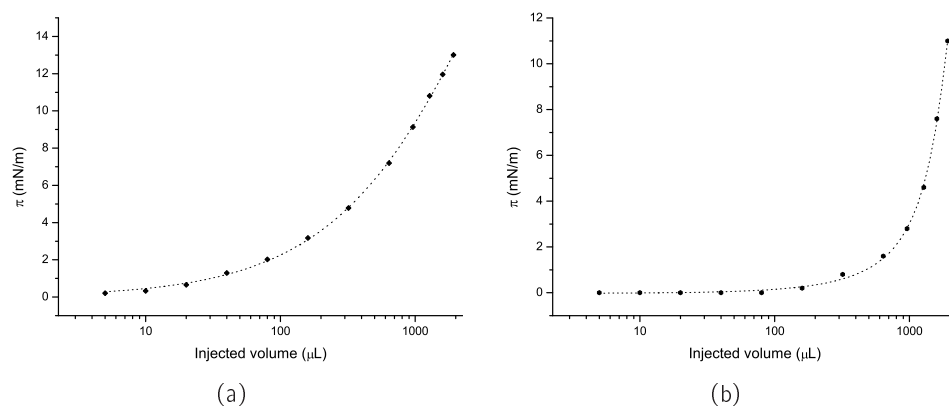
Figure 1.7: Effect of DMSO on the measured surface pressure: (a) shows the effect on the 3 mL trough; (b) shows the effect on the 20 mL trough; the dotted curves represent the extended Langmuir model fit (a), and the Lorenz model fit (b).

Lorenz model:

$$\pi = \pi_0 + \frac{2A}{\Pi}\frac{w}{4(V - V_c)^2 + w^2}\,, \tag{1.4}$$

where π is the surface pressure in mN/m, V the total injected volume in μL and π_0, A, w and V_c four variable parameters. The fit parameters obtained were $\pi_0 = -2.36$, $A = 5.14 \cdot 10^4$, $w = 1.94 \cdot 10^3$ and $V_c = 2.42 \cdot 10^3$.

1.3   Computational part

1.3.1   Automatic evaluation of π-Log(C) plots

After having establish consistent correction factors for the evaluation of the measured surface pressures, the next step was to develop a software to facilitate the analysis of the $\pi = f(log(C))$ plots. The objective was also to remove the bias induced by the manual selection of the quasi-linear part of these plots. Indeed, it can be very difficult to choose by eye the best quasi-linear part, and this method will not be reproducible and user-independent.

The most logical approach was to develop an automatized system which would select the best quasi-linear part of the $\pi = f(log(C))$ plot in a rational way. Thus, a Microsoft® Excel macro was programmed to fulfil this task.
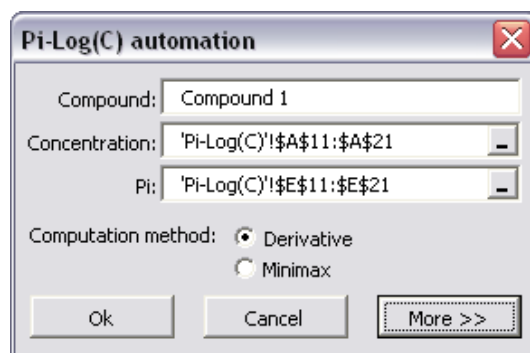
Figure 1.8: User interface of the *Pi-Log(C) automation* Excel macro

The user interface shown in Figure 1.8 allows a choice of the range of concentrations and surface pressure (π) measurements, and two different methods of calculation for the automatic evaluation, which will be described further below. The "More" button opens a second panel, in which one can manually choose the quasi-linear part of the plot and the surface pressure where the CMC is reached, in case the program was unable to determine these parameters.

The automatic selection of the quasi-linear part will be discussed in the first section, and some other features will be described in the second section.

Automatic determination of the linear part

Here we will describe briefly how the Excel macro determines the "best" quasi-linear part of the plot; two methods for evaluation, "derivative" (A) and "minimax" (B), were developed. "n" represents the total number of data points (i.e. the number of distinct concentrations for which a surface pressure was measured). A datapoint is given by the log of a concentration and its corresponding measured surface pressure, and is abbreviated as "point". The method described below is the "derivative" method (A).

1. Calculate the derivative for each point from 2 to $n-1$

2. Set the derivative of the first and last point to 0

3. Calculate the sum of the derivatives for each group of k points ($3 \le k \le n$) as defined below:

$$3 \text{ points:} \sum_{i=1}^{3} \partial_i \, ; \sum_{i=2}^{4} \partial_i \, ; \ldots ; \sum_{i=k}^{k+2} \partial_i \, ; \ldots ; \sum_{i=n-2}^{n} \partial_i$$

$$4 \text{ points:} \sum_{i=1}^{4} \partial_i \, ; \ldots ; \sum_{i=k}^{k+3} \partial_i \, ; \ldots ; \sum_{i=n-3}^{n} \partial_i \tag{1.5}$$

$$\vdots$$

$$\text{n points:} \sum_{i=1}^{n} \partial_i$$

4. Select for each group of k points ($3 \le k \le n$) the one which has the highest sum (i.e. the steepest part of the plot)

5. Calculate for the selected groups of 3 to n points their r^2 (correlation coefficient)

6. Adjust all calculated r^2 to a base 100 where 0 is the lowest r^2 and 100 the highest r^2

7. Choose the group which has the highest k and for which r^2 is superior or equal to 90

This group of k points represents the steepest part of the $\pi = f(log(C))$ plot and shows a good ratio between the correlation coefficient and the number of points.

A second possibility of evaluation is the "minimax" method (B) which works on the same basis as the one described above (A), but the weighting of

all p points ($1 \leq p \leq n$) is not defined as their derivative, but as follows:

$$
\begin{aligned}
&\texttt{If } \pi_p > (\pi_{min} + \pi_{max}) \ / \ 2 \texttt{ Then} \\
&\qquad \texttt{weight}(\pi_p) = \pi_{max} - \pi_p \\
&\texttt{Else} \\
&\qquad \texttt{weight}(\pi_p) = \pi_p - \pi_{min} \\
&\texttt{End If}
\end{aligned}
\tag{1.6}
$$

with π_{min} and π_{max} the lowest and the highest measured value, respectively.

This method can be useful for example when the shape of the $\pi = f(log(C))$ plot is sigmoidal; in this case, the "derivative" method is unable to determine a relevant group of point, whereas the "minimax" method can lead to useful data.

Other features

Besides the automatic determination of the linear part of the $\pi = f(log(C))$ plot, the Excel macro contains other features which improve the evaluation of data as outlined below:
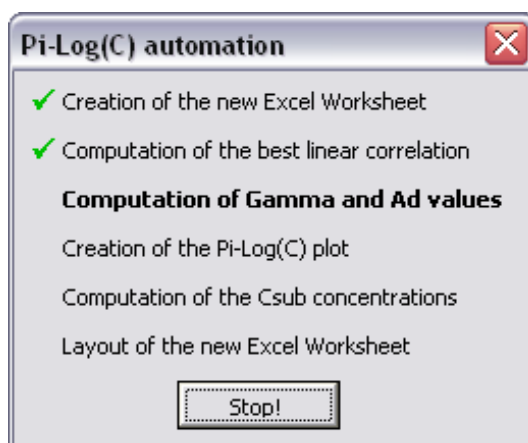
Figure 1.9: Progress window of the *Pi-Log(C) automation* Excel macro

Progress window: The user is informed of the calculation process in real time during the computation and can stop it at any time, as shown on the screen capture above (Figure 1.9).

Cross-sectional area: Using the slope of the linear part of the plot chosen with the method described above, the macro calculates the cross-sectional area (A_D) of the tested compound. From A_D, it can directly calculate the air-water partition coefficient ($K_{aw_{calc}}$) and the lipid-water partition coefficient ($K_{lw_{calc}}$).

CMC: The macro is able to determine automatically if the concentration range used for the measurement was large enough to reach the *critical micellar concentration* and it specifies also the CMC value, if required. In this case, the whole macro is automatically run again by giving a weight of zero to the datapoint which is closest to the CMC. Indeed, it often appears that the measured surface pressure slightly "jumps" just before reaching the CMC, leading to a non-optimal determination of the quasi-linear part of the $\pi = f(log(C))$ plot. By weighting the last datapoint before CMC to zero, this point will have less chance to be included in the quasi-linear part of the $\pi = f(log(C))$ plot, especially in the case of a surface pressure jump. Consequently, different K_{aw} and A_D ($K'_{aw_{calc}}$ and $A'_{D_{calc}}$) will be calculated. The average of the two values (resp. $(K_{aw_{calc}} + K'_{aw_{calc}})/2$ and $(A_{D_{calc}} + A'_{D_{calc}})/2$) and the corresponding standard deviations will be displayed on the Excel sheet.
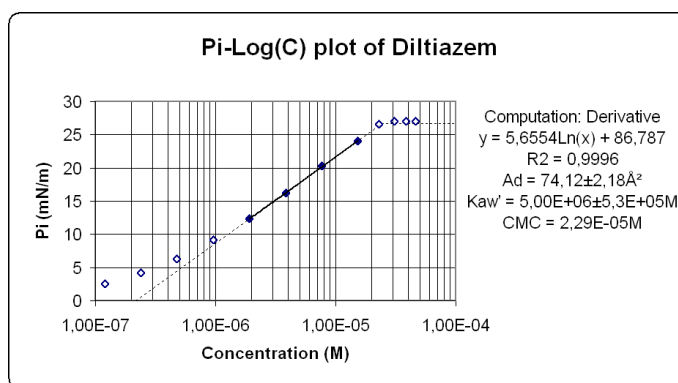
Figure 1.10: Screen capture of the graphic displayed by the *Pi-Log(C) automation* Excel macro

Graphic: The macro draws an explicit $\pi = f(log(C))$ graphic ($\diamond$) showing the best linear part chosen by the macro ($\blacklozenge$) and the CMC if required (see Figure 1.10). The graphic also shows the numerical values for A_D and CMC.

Chemical structure: If the chemical structure of the tested compound is present in a predefined folder, it is added to the output worksheet.

Future enhancements: The macro is structured in a way that it is easy to add some new modules (functionalities), like the calculation of the surface concentration (C_{sub}) of the compound.

This Excel macro strongly simplifies the evaluation of the surface pressure measurements, without the inconvenience of many automated systems: it has not become a black box where the user only sees the output, as he always

keeps the control on the crucial part of the automation which is the selection of the quasi-linear part of the $\pi = f\left(log(C)\right)$ plot.

1.3.2   Automatic data extraction for the 3 mL trough

For the monolayer measurements, we either use a 20 mL trough which is coupled to a mechanical recorder, or a 3 mL trough which is connected to a computer for the recording of the surface pressure. Both troughs are covered with a hood to lower the evaporation effect. The electronic data acquisition for the 3 mL trough gives an output file containing a large data table with the surface pressure measured in real-time during the whole experiment. The interval between two data acquisitions can be configured and is usually set to 3 seconds; as a consequence, a measurement of an average duration (i.e. 1h30) will generate 1800 data points.

As the injection is not yet automatic, the elapsed time between each injection is not recorded, therefore the user has to find out the injection steps in this large data table. He also has to calculate an average value of the equilibrated part of the surface pressure for each concentration, to take into account the background noise generated by interferences with the magnetic stirrer.

This task is of course very time consuming, therefore it was important to develop an automated method to extract the injection times (to take into account the evaporation of the buffer, see chapter 1.2.1 on page 4) and the average surface pressures for each concentration.

Extraction of injection times and surface pressures

The first user interface shown in Figure 1.11(a) allows the selection of the whole data table. The user can also select the sensitivity for the extraction of the injection times and surface pressures, to handle the importance of the background noise. From one measurement to the other, the background noise can vary, and if the sensitivity isn't high enough, some injections will not be found by the *PiXtract* macro. During the calculation, the progress window (Figure 1.11(b)) is displayed to inform the user on the task progress in real time.

The data are first smoothed to lower the background noise, and then derived to reveal the injection peaks. The smoothing works by averaging the 14 surrounding points weighted according to a gaussian distribution, in order to get rid of the periodic noise due to the stirring device.

Then, the injection times are extracted (from one peak of injection to the other). Afterwards, the surface pressure of each injection is defined as an average of measured surface pressures, from the injection peak towards the previous injection peak until the measured surface pressure is lower or higher than a threshold defined from global background noise.

(a)   (b)
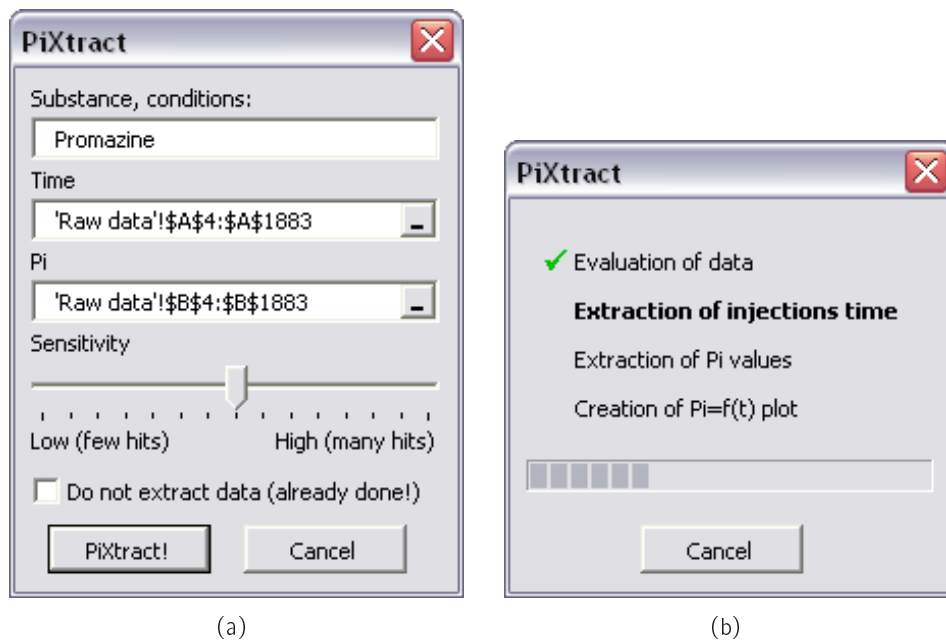
Figure 1.11: User interface (a) and progress window (b) of the *PiXtract* Excel
macro

Refinement of the extracted data

The second user interface (Figure 1.12) is displayed when the calculation is
over, simultaneously with the graphic shown in Figure 1.13. This graphic
presents the injection peaks found by the Excel macro, and the user can either
add or remove some injection steps intuitively; the selected steps are displayed
in real time on the graphic. If the global threshold was too high (too many
peaks found) or too low (not enough peaks found), the Excel macro can be
run again with a different sensitivity. When the correct peaks are selected,
the user can rerun the Excel macro with a higher sensitivity to improve the
calculated average of the surface pressure.

Later on, the user can set the duration between the beginning of the exper-
iment and the start of the recording (to take into account the real evaporation)
and also deduct the baseline from all data. Finally, the data (surface pressures
and time between each injection) can be pasted into the worksheet described
in section 1.3.1 on page 10. The two Excel macros are intimately coupled for
better efficiency in the data evaluation. Moreover, the exact determination of
the injection times allows a more accurate estimation of the evaporation of
the water.
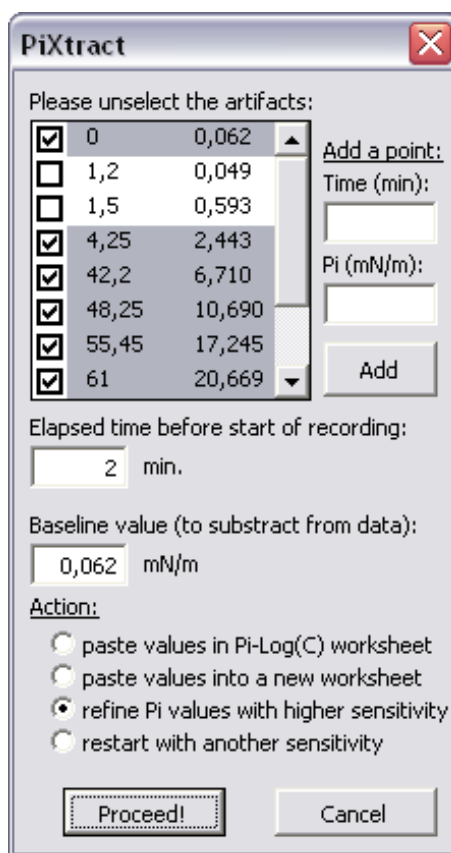
Figure 1.12: Second user interface of the *PiXtract* Excel macro
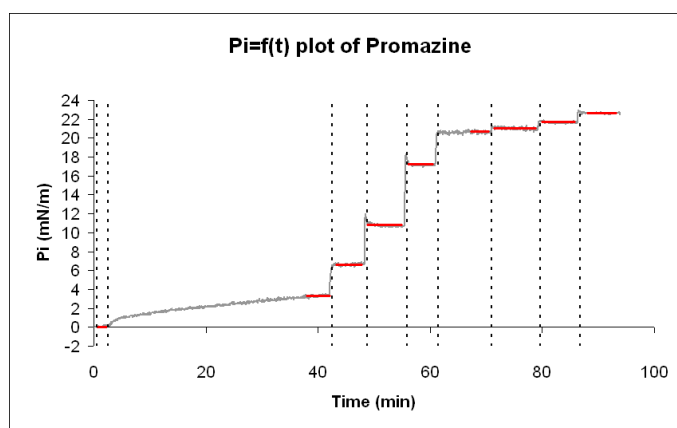
Figure 1.13: Screen capture of the graphic displayed by the *PiXtract* Excel macro

1.3.3 Calculation of the cross-sectional area

This section has been thoroughly described in ref.[10] (cf. Appendix).

1.4 Conclusion

In a first step, we assessed new correction factors for the solvent evaporation, the liquid level, and the influence of the solvent on the measured surface pressure. In a second step, the *PiXtract* Excel macro was developed, in order to precisely evaluates the raw data generated by the automatic recorder of the 3 mL trough (injection times, surface pressure reached at the equilibrium before the next injection). Finally, the *Pi-Log(C) automation* Excel macro was programmed to provide a more accurate and unbiased evaluation of the $\pi = f(log(C))$ plot. This may lead to a better reproducibility of our experiments and a decrease of the margin of error.

Chapter 2

Prediction of P-glycoprotein interaction

2.1 Introduction

During the development of a new pharmaceutical compound, the pharmacokinetics of the drug has to be assessed precisely;[11] this will be directly related to the plasma, as well as the concentration of the drug at the target site. Achievement of required concentration might be hindered at different steps of the disposition of the drug within the organism:[12]

Absorption   The first step of the disposition of the drug is the absorption, which will critically influence its bioavailability. Several parameters have to be fulfilled in order to get the compound in a sufficient concentration into the bloodstream: the compound has to be *soluble*, *chemically stable* in the stomach if the compound is orally taken, and it has to be able to cross the *intestinal barrier*.

Distribution   Once the drug is present in the bloodstream, it has to reach its effector site (target) and distribute into tissues and organs. To this purpose the compound might have to cross several barriers, such as the *blood-brain barrier*.

Metabolization   The organism has several ways to eliminate any exogenous material, as a drug. One of the best investigated ways is the metabolism by enzymes present in the liver, the cytochromes P450 family. The drugs' half-time in the organism is thus a crucial parameter to assess its bioavailability.

Elimination   The last step of the disposition of the drug is its elimination: the drug, metabolized or not, will be removed from the organism by excretion,

usually through the kidneys or in the feces. Since multidrug transporters are expressed in the kidneys, they might play an important role by increasing the clearance of the drug.[13, 14]   Moreover since metabolites of the drug might also have a pharmacological activity or toxicity, their elimination is also an important factor to assess.

2.1.1   Multidrug transporters

During *absorption* and *distribution*, the drug might thus have to cross several barriers, such as the *intestinal barrier* (IB) or the *blood-brain barrier* (BBB). Several factors might prevent the diffusion of the drug through these barriers. First of all the compound must achieve passive diffusion, thus if it is highly charged or too large it will not be able to diffuse or diffuse only slowly.[15, 16] Secondly, these barriers have a high expression level of a wide variety of multidrug transporters,[17–19] some of them are localized in the apical side of the epithelial cells, others on the basolateral side (Figure 2.1(a)).

These multidrug transporters have a high sequence homology (Figure 2.1(b)) and share similar properties: they have two transmembrane domains constituted of 6 membrane-spanning α-helices providing specificity for the substrate (Figure 2.2, orange and pink), and two nucleotide-binding domains, which bind ATP and uses the energy of ATP hydrolysis to drive the transport of a wide panel of structurally unrelated molecules across the cell membrane (Figure 2.2, green and limon).[20, 21]   The best characterized member of the ATP-binding cassette (ABC) multidrug transporters is the P-glycoprotein, which has a molecular weight of 170kDa and is encoded by the MDR1 gene, is mainly expressed in the intestinal epithelium and the blood-brain barrier where it might reduce the diffusion of the drugs into the organism, and the kidneys and the liver where it plays an important role in the clearance of drugs. It is also highly expressed in many cancer cells, to which it confers multidrug resistance.[22]   However the three-dimensional structure of P-glycoprotein has not been resolved yet.

2.1.2   P-glycoprotein

P-glycoprotein recognizes a wide spectrum of chemically diverse molecules, that are hydrophobic and may carry a positive charge at physiological pH; they are e.g. anticancer drugs (vinca alkaloids, anthracyclines, epipodophyllotoxins, taxoids. . .), therapeutic agents like HIV-protease inhibitors, or cyclic peptides.[24–26] A drug that will be recognized and translocated by P-glycoprotein (or any other multidrug transporter) will not be able to reach its target if the targeted organ is protected by a high expression level of the transporter, or if the drug is in a decreased concentration.[27, 28]   On the other hand, a compound having a high affinity for the transporter may change the bioavailability
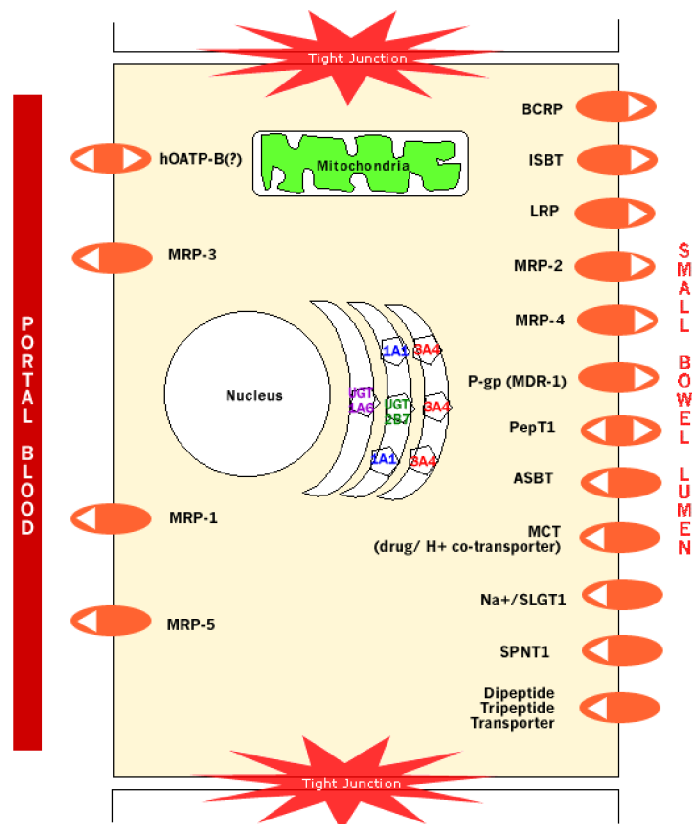
(a)
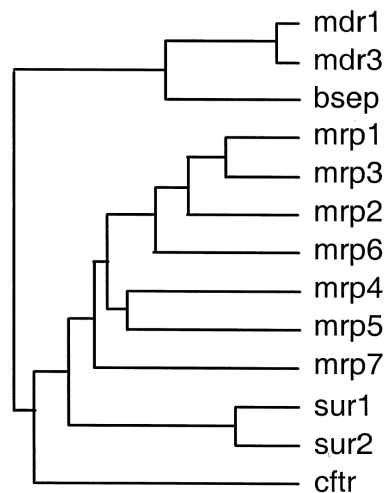
(b)

Figure 2.1: Multidrug transporters: (a) schematic representation of transporters involved in intestinal transport through enterocytes (adapted from http://bigfoot.med.unc.edu/watkinsLab); (b) Phylogenetic tree of some selected transporters.

Figure 2.2: Homology model of the transporter P-glycoprotein.[23] The two transmembrane domains are shown in orange and pink, and the two nucleotide binding domains in green and limon.

of another co-administered drug that was usually effluxed by P-glycoprotein, leading to unexpected side-effects.[29]

It is therefore crucial, in an early stage of the development of a new drug, to find out if it is a potential substrate or inhibitor of P-glycoprotein, to optimize the bioavailability of the drug, and to avoid any unwanted toxicological effect that may occur with co-administration of other substances.

It has been proposed that P-glycoprotein has two binding regions, one activating binding region occupied at low substrate concentrations, and one inhibitory binding region, occupied at high substrate concentrations.[1, 30, 31] Binding of drug to the activating region will enhance P-glycoprotein activity and thus ATP hydrolysis, whereas binding to the inhibitory region will reduce P-glycoprotein activity.[32–34]

The putative transmembrane helices of P-glycoprotein have a high number of residues with hydrogen bond donor side chains arranged in an amphipathic manner.[25] It has been suggested that the binding of the drug to the transporter occurs *via* hydrogen bonds arranged in a particular spatial distance, called *type I* and *type II* units, between hydrogen bond acceptors on the drug and the hydrogen bond donor side chains of the transmembrane α helices of P-glycoprotein.[35]

2.1.3 Assays to study the influence of P-glycoprotein

Some *transgenic animal models* have been established (like double knock-out mice, since mice have two genes coding for P-glycoprotein, mdr1a and mdr1b) to study the effect of the absence of P-glycoprotein on the bioavailability of drugs.[27] Moreover, several *in vitro* methods based on transgenic cells over-expressing P-glycoprotein have been established.[36] The following section will enumerate and describe the most commonly used techniques.

ATPase activity assay As stated previously, P-glycoprotein needs ATP to perform the transport of the drug out of the cell. Thus a process coupled to drug efflux will be hydrolysis of ATP, leading to apparition of ADP and inorganic phosphate, P_i. However the hydrolyzing step will occur inside the cell, which is rather difficult to monitor. The solution is to work with inside-out vesicles: the cells are disrupted, then the membrane fragments containing the P-glycoprotein are isolated with several centrifugation steps. Afterward, these fragments will reassemble predominantly as inside-out (the nucleotide-binding sites toward the exterior) vesicles. In those vesicles, the inorganic phosphate will appear in the medium upon P-glycoprotein activation; a colorimetric reaction is used to measure afterward the extravesicular P_i concentration.[37, 38]

Competition assay The competition assay uses a "prefluorochrome", the calcein-AM, that is a known substrate for P-glycoprotein. Once in the cell,

calcein-AM is hydrolyzed by endogenous esterases into calcein, which is a highly negatively charged fluorescent compound (Figure 2.3). Calcein will thus not be able to leave the cell once hydrolyzed. In P-glycoprotein over-expressing cells, calcein-AM will only poorly penetrate the cell due to its efflux by the multidrug transporter, leading to a slow apparition of fluorescence. However upon co-administration of another P-glycoprotein substrate, this drug will compete with calcein-AM, leading to a higher diffusion of the prefluorochrome into the cell, and thus a faster apparition of fluorescence. This assay is carried out with intact cells; a control might be performed with wild-type cells, leading to a fast apparition of fluorescence both without and with co-administration of a second drug.[39]
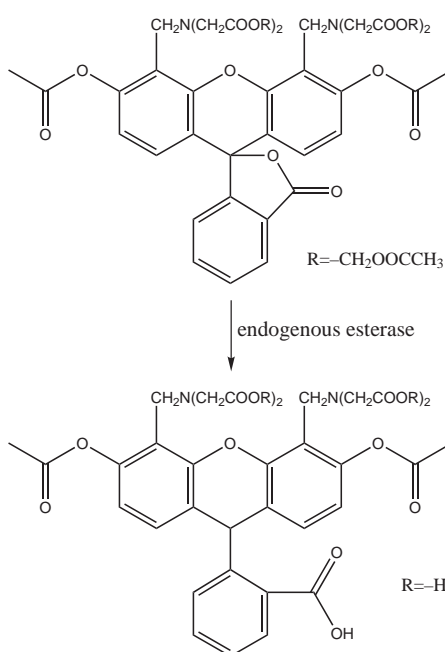
Figure 2.3: Non-florescent calcein-AM hydrolyzed by esterases into fluorescent and negatively charged calcein.

Transcellular transport assay   The last kind of commonly used *in vitro* assays is the transcellular transport assay.[40] It is performed using either wild-type cells (e.g. porcine brain endothelial cells, PBEC), or transgenic cells over-expressing P-glycoprotein (e.g. PBEC or Caco-2 cells). The cells are grown on a polyethylene membrane filter (with pore size of 3μm) coated with collagen; this allows the cells to be polarized, and to express the multidrug transporter on one side only, namely the apical side (corresponding to the lumen of the stomach, or the blood vessel at the blood-brain barrier, depicted as *donor compartiment* in the model) (Figure 2.4). The cells will grow until con-
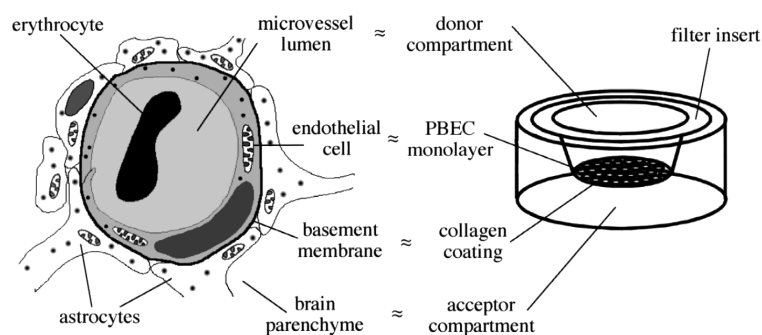
Figure 2.4: Schematic view of the paracellular transport assay, compared to the blood-brain barrier.[36] PBEC stands for porcine brain endothelial cells.

fluence is reached, and the endothelial cells will build tight junctions between them. The tested compound is then first applied in the *donor compartment* and the apparition of drug is monitored (usually by HPLC and mass spectrometry) in the *acceptor compartment* (corresponding to the brain parenchyme at the blood-brain barrier, or the blood at the intestinal barrier); this experiment is called B→A. In a second experiment, the drug is applied in the acceptor compartment, and the apparition of drug is monitored in the donor compartment; this experiment is called A→B.

If a drug is not a substrate for P-glycoprotein, the apparition of the drug in both compartments will be identical, thus the ratio $\frac{B \to A}{A \to B}$ will be ~1. Since the transporter is expressed only in the apical side (towards the donor compartment), a substrate applied in the acceptor compartment the drug will be effluxed by the transporter, leading to no or a slow apparition of the drug in the donor compartment. When applied to the donor compartment it will accumulate more rapidly in the acceptor compartment. Thus the ratio $\frac{B \to A}{A \to B}$ will be larger than 1. A compound is generally defined as substrate if this ratio is larger than 1.5.[36]

Cyctosensor microphysiometer In our group we use a new approach to study the substrates of P-glycoprotein. It has been described previously that the P-glycoprotein uses the energy of ATP hydrolysis to transport drugs out of the cell.[32] The use of ATP is directly correlated to the production of lactate, and as a coupled phenomenon, an extrusion of a proton out of the cell.[33] Thus upon P-glycoprotein activation, the environmental medium will be *acidified*; this acidification can be assessed using a cytosensor microphysiometer, which is a micro pH-meter. Results obtained with this method are comparable to those obtained with the ATPase activity assay, however it has the advantage to handle living cells instead of reconstituted inside-out membrane vesicles.[33, 34]

2.1.4   Kinetic parameters of these models

The four assays described above allow the measurement of different kinetic parameters. A kinetic model was proposed by Litman et al. in 1997;[30, 31] it states that P-glycoprotein has two binding regions, one activating binding region and one inhibitory binding region. At low drug concentration, the activating binding region will be occupied, leading to an increase of ATP hydrolysis. Drugs enhancing ATP hydrolysis by P-glycoprotein are called *intrinsic substrates*.[6] At high drug concentration, the inhibitory binding region is likely to be occupied, leading to a decrease in the ATP hydrolysis. Drugs causing a decrease in the ATP hydrolysis already at low concentrations are called *inhibitors*.[4] Moreover, as described for the competition assay, if two different intrinsic substrates (like calcein-AM and another substrate of P-glycoprotein) are applied simultaneously, the substrate with the higher binding affinity to the transporter will influence the bioavailability of the second drug;[29] these compounds having a high affinity to P-glycoprotein are called *modulators*.[4, 41] Modulators generally enhance P-glycoprotein activity, and thus ATP hydrolysis.

ATPase activity assays can identify unambiguously intrinsic substrates, modulators and inhibitors. Since this method monitors the ATP hydrolysis resulting from P-glycoprotein activity, any drug that will interact with P-glycoprotein influences the ATP consumption and thus the apparition of inorganic phosphate. Moreover, since P-glycoprotein has a basal activity in the absence of externally applied drug due to transport of endogenous substrate or uncoupled cycling,[42–44] even the effect of an inhibitor of P-glycoprotein will be measurable, since inhibitors decrease the transporter activity.

Competition assays, using calcein-AM (or rhodamine 123), can be used to detect *modulators* (compounds that will modulate the diffusion of calcein-AM into the cell, and thus lead to a faster apparition of fluorescence of calcein), or inhibitors of P-glycoprotein. This method only detects the modulators with higher affinity to P-glycoprotein than calcein-AM.

One should emphasize the difference between intrinsic substrates and apparent substrates. Apparent substrates are intrinsic substrates (leading to an increase in the ATP consumption of P-glycoprotein), which give a value larger than 1.5 in the $\frac{B \to A}{A \to B}$ ratio in *transcellular transport assays*.[36] To be detected by this method, the substrates of P-glycoprotein should be large enough (i.e. cross-sectional area $\gtrsim 70$ Å^2 for blood-brain barrier permeation), or highly charged, in order to diffuse slowly through the cell membrane.[4, 7] Indeed intrinsic substrates having a small cross-sectional area might diffuse too fast to be effectively accumulated in the apical side (donor compartment), leading thus to a ratio $\frac{B \to A}{A \to B} \sim 1$.[6]

Available data on P-glycoprotein interaction with drugs have to be considered carefully: considering the fundamental differences in the setup and thus the information collected with these assays, one might obtain various outcome for one single drug, that has not been fully understood for a long time. This is

mainly due to the difference between intrinsic substrate and apparent substrate explained above, the presence of two binding regions on P-glycoprotein leading to inhibition at high substrate concentration, but also due to the difference in the lipid composition and the lateral packing density of membranes from different cell lines used in these assays.[6]

2.2 Literature review of in silico prediction of P-glycoprotein interaction

In the above section we described the influence that P-glycoprotein might have on the absorption and distribution of a drug in the organism, and in a more general manner on its bioavailability. Several *in vivo*, as well as *in vitro* models and methods have been developed to assess the interaction of drugs with P-glycoprotein; these methods are of particular interest especially during the development of new pharmaceutical compounds, like anticancer or antiviral drugs where numerous cases of resistance to the treatment due to P-glycoprotein interaction have been reported.

Since these methods are poorly adapted to screening of a large number of compounds, it has been a great challenge to develop relevant computational tools to anticipate whether drug candidates are P-glycoprotein substrates, modulators or inhibitors. Several tools based on different approaches have been published, and the following section will shortly review these methods of prediction of P-glycoprotein interaction.

2.2.1 Datasets used in the literature for models training

Building a new model to assess the interaction between P-glycoprotein and drugs requires experimental data. It is therefore of importance to first carefully investigate by which technique the data used as training sets were obtained. Some models are essentially based on the results of transcellular transport assay on Caco-2 cells or MDR1-MDCK cells (dog kidney epithelial cells),[45–47] some others on competition assays[48–51] or ATPase assays[52,53] and some models combine *in vitro* results from different experiments.[54,55] Some authors do not describe the source of their experimental data, but define P-glycoprotein substrates as compounds being transported or inducing overexpression of the multidrug transporter, whereas non-substrates are compounds specifically described as not being transported by P-glycoprotein.[56–58] One can assume that these authors use the results of transcellular transport assays and protein expression level analysis, however, they combine the results of two processes which cannot be *a priori* mixed. Compounds inducing overexpression of P-glycoprotein are likely to be intercalating agents or molecules interacting with DNA, but they are not necessarily apparent substrates of P-glycoprotein.[59] Since intercalating agents often contain one or several phenyl groups, which

have been shows to be a frequent recognition element by the multidrug transporter,[35] the confusion between *substrates* and *inducers* is regularly observed.

2.2.2 Recent models for P-glycoprotein substrates prediction

In the last years, many models have been published, using different approaches to assess the affinity of drugs to P-glycoprotein. Most of them are using the results of transcellular transport assays, however the results of ATPase activity assays or competition assays are also employed. Three main approaches are used for the prediction of substrates of P-glycoprotein.

Pharmacophoric model A pharmacophore corresponds to the essential features (hydrophobic area, aromatic ring, hydrogen bond acceptors, hydrogen bond donors, anions, cations etc.) of one or several molecules having the same biological activity. The published pharmacophore models have several (usually 4) features, each feature having a specific spatial localization.[45, 48–51, 53, 54, 60] Ekins *et al.*[50, 51] collected data from different experiments and proposed one pharmacophore for each type of experiment: one for inhibition of digoxin transport by P-glycoprotein through Caco-2 cells (transcellular transport assay), one for the inhibition of vinblastine binding (competition assay), one for the accumulation of vinblastine in P-glycoprotein over-expressing cells (competition assay), and a fourth one for the calcein accumulation in the same cells (competition assay). Garrigues *et al.*[53] proposed two pharmacophores for P-glycoprotein binding; both binding regions are bound together, one region will bind one kind of small drugs (like bromocriptine), the second another kind of small drugs (like tentoxin), and large drugs will bind using both binding regions (like verapamil or vinblastine). Penzotti *et al.*[52] generated 100 pharmacophores of either 2 (8 models), 3 (39 models) or 4 (53 models) features; 88% of the pharmacophores contained at least one hydrogen bond acceptor feature, 97% a hydrogen bond donor feature, and most of them a hydrophobic group, and an aromatic ring. Moreover, 18% of the models contained either type I or type II units,[35] as described in section 3.2.9.

These models have all been established on relatively small datasets (no more than 30 compounds, and 50 compounds for the training set in Penzotti *et al.*,[52] mainly collected from Seelig[35]), thus they might fail in predicting substrates of P-glycoprotein of chemical families not represented in the original training set. It might be surprising that the models proposed by Ekins *et al.*[50, 51] for the three competition assays are quite different, but this can be explained by the small size of the experimental datasets; it highlights the weaknesses of the pharmacophore approach. This is also visible in the external validation (screening of an external database) of the model proposed by Ecker *et al.*,[60] providing homologous compounds to the ones included in the training set. It is worth noting that the model described by Garrigues *et al.*[53] might

be relevant to assess binding of long molecules, shown to inhibit the multidrug transporter.[4] Finally, all authors used only one single three-dimensional structure of each compound, which will obviously have a huge impact on the generated 3D pharmacophore. This aspect was highlighted by Cianchetta *et al.*[54] who noted the similarity in the features of different published models, but noticed the divergence in their geometrical and spatial repartition within the pharmacophores.

P-glycoprotein transports a wide variety of structurally unrelated compounds, and has at least two binding regions (activating and inhibitory binding regions).[1,30,31] The traditional pharmacophoric approach based on a three dimensional "lock and key" model, which is the method applied in traditional Structure Activity Relationship (SAR) studies where a protein will have one single binding site and will bind one kind of substrate, might thus not be the method giving the highest predictivity for the interaction between drugs and P-glycoprotein.

Discriminant linear model The other common approach for predicting substrates of P-glycoprotein is the discriminant linear model, a multivariate generalized regression method.[46,55] Several methods are derived from this model, like the support vector machine approach (SVM),[56] the principal component analysis (PCA),[56] partial least squares discriminant analysis (PLSD),[45,48,54] or machine learning approach (neural network).[45] These methods start with numerous descriptors and try to find a linear correlation between experimental data and the descriptors. These descriptors are similar as those listed above, however the spatial property is lost. The geometry of the molecules is thus generally assessed by other descriptors like the number of rotatable bonds, and the ability of membrane partitioning with LogP or similar descriptors; it is also possible to transform the 3D information (from a single 3D structure of the drugs) into 2D descriptors, for instance the length of the molecule, the distance between particular regions, the globularity, etc.[45,48,54,56] The model established by Gombar *et al.*[46] is based on 254 descriptors, from which 27 were significant, including hydrogen bond donor and acceptor, calculated molar refraction, and molecular E-state. Cianchetta *et al.*[54] used the 94 Volsurf descriptors, and the 940 Almond descriptors, in which the most significant descriptors were molecular weight, rugosity, polarisability, molecular volume, hydrogen bonding ability, and hydrophobicity. Crivori *et al.*[45] used the same descriptor sets, but in this case the relevant descriptors were size, shape, flexibility, molecular surface, globularity, elongation, hydrophobic regions, and hydrogen bonding ability. Cabrera *et al.*[55] used only 9 descriptors, from which 6 were significant: atomic charge, molar refraction, polar surface area, polarisability, partition coefficient, and van der Waals radii. Li *et al.*[48] also used a small number of descriptors: steric field, electrostatic field, hydrogen bonding ability, hydrophobic field, and hydrophobicity. Finally, Xue *et al.*[56] selected

159 significant descriptors from the 1000 obtained with Dragon.

The datasets used to establish the models were from various sources, and vary also in number of compounds. Li *et al.* used the data of 20 steroids which are substrates and/or inhibitors of P-glycoprotein, measured on SW600 Ad300 cells by drug accumulation and efflux studies, without including any non-substrate or non-inhibitor. For substrates, their training set contained 10 compounds, whereas the validation set 3 compounds; for inhibitors, 18 compounds were included in the training set, and only 2 compounds in the validation set. They obtained for both models a cross-validated q^2 of 0.720 and 0.520, respectively. Crivori *et al.* included 53 drugs (22 substrates, 31 non-substates) in their training set, measured in Caco-2 cells (transcellular transport assay) and against calcein-AM (competition assay); the test set was composed of 272 in-house compounds (115 substrates, 157 non-substrates). The PLSD model on the training set gave a sensitivity of 86%, and a specificity of 90%. The model applied on the validation set gave a sensitivity of 61% and a specificity of 81%. Cianchetta *et al.* used the same assays, on 129 compounds (100 in-house substances), all of them being P-glycoprotein substrates, dispatched into a training set of 109 compounds, and a test set of 20 compounds. They obtained a coefficient of correlation r^2 of 0.72 and 0.83 using the Volsurf and the Almond descriptors, respectively. Xue *et al.* claimed that they used 116 substrates (described as being transported or inducing over-expression) and 85 non-substrates from the literature, without citing their sources. The training set contained 142 compounds (74 substrates, 68 non-substrates), the optimization set (used to calibrate the model) 34 drugs (22/12), whereas the validation set had 25 compounds (19/6). The described model gave a sensitivity of 84% and a specificity of 67%, and a 5 time cross-validation a sensitivity of 81% and a specificity of 79%. Gombar *et al.* used 95 compounds described by Seelig *et al.*,[35] containing 32 substrates and 63 non-substrates. The validation set contained 58 in-house compounds (35/23), giving a sensitivity of 94% and a specificity of 78%. Finally, Cabrera *et al.* used 163 compounds (91/72) for the training set, originating from various sources (transcellular transport, ATPase assay, competition assay, drugs inducing over-expression), and 40 (22/18) compounds in the validation set. The authors described a sensitivity of 81% and a specificity of 72% on their validation set.

Like for pharmacophore models, the discriminant linear models have some drawbacks inherent to the method applied. Most problems of the discriminant linear model approach might come from the intrinsic mechanism of drug binding to P-glycoprotein. A drug, which has to be amphiphilic, will first have to partition into the lipid membrane, be flipped to the inner leaflet, and finally bind to the transporter. This two-step process might require a separate structure-activity relationship analysis for membrane binding, as suggested by Seelig *et al.*[3] (for review, see Stouch and Gudmundsson[61]). Even for the

ATPase assay, this binding is relevant since drugs will also have to partition into the outer membrane leaflet of the inside-out vesicles, in order to reach P-glycoprotein.

Rule-based approach A ruled-based model has been proposed by Didži-apetris *et al.*[47] They used 11 descriptors (hydrogen acidity and basicity, hydrogen bonding ability, pKa, number of rotatable bonds, polar surface area, number of aromatic rings, molecular weight, hydrophobicity, and characteristic volume), from which only five were significant for P-glycoprotein substrates prediction: hydrogen basicity, hydrogen bonding ability (acceptors and donors), molecular weight, and pKa. Using these descriptors they defined simple rules to discriminate between substrates and non-substrates of P-glycoprotein: compounds with (N+O) $\geq$ 8, MW > 400 and acid Pk_a > 4 are likely to be P-glycoprotein substrates, whereas compounds with (N+O) $\leq$ 4, MW < 400 and basic pK_a < 8 are likely to be non-substrates. This approach resembles the "rule of five" for drug-like molecules, as described by Lipinski.[62,63]

The calibration set contained 220 compounds for which P-glycoprotein affinity was determined by transcellular transport experiment, with 101 substrates and 119 non-substrates. On this dataset, the model was unable to classify 89 compounds (40%). On the remaining 131 compounds (65/66), the model lead to a sensitivity of 86%, and a specificity of 85%. The validation set contained the 220 compounds of the calibration set, as well as 780 compounds collected from around 600 original publications and review papers. The model could not classify 453 compounds (45%). On the 547 (166/381) remaining compounds, the model gave a sensitivity of 89%, and a specificity of 90%.

Contrary to pharmacophore models, Didžiapetris *et al.* did not find a particular importance of aromaticity for P-glycoprotein substrates. This can be explained by the much larger size of their training set comparing to the sets used to develop the pharmacophore models. The authors also highlight the importance of amphiphilicity for P-glycoprotein substrates, and note that 3D pharmacophores might not be adequate to assess this parameter due to the variety of the conformational space for each compound. They describe drug P-glycoprotein interaction as "fuzzy specificity" with no clearly expressed binding sites, as previously suggested by Sharom (one single binding site that can accommodate several ligand molecules),[26] and Litman *et al.*[30,31] However the fact that experimental data were "binarized" (1 for substrates, 0 for non-substrates) and the rigidity of the rules might lead to misprediction, especially for compounds having borderline properties (like a molecular weight around 400).

2.3   Development of a new model for drug binding to P-glycoprotein based on a two-step mechanism

As described previously, a substrate of P-glycoprotein will first have to partition into the outer leaflet of the lipid membrane of the cell, then to diffuse to the inner leaflet, where it will bind to the multidrug transporter.[5] Thus the binding constant of the substrate from the aqueous environment (w) to the transporter (t), K_{tw}, can be considered as the product of its partition coefficient from the aqueous environment to the lipid membrane (l), K_{lw}, and its binding constant from the lipid environment to the transporter, K_{tl}.[4] Consequently, the free energy of substrate binding from the aqueous environment to the transporter, ΔG_{tw}, will be the sum of the free energy of its binding from the aqueous environment to the lipid membrane, ΔG_{lw}, and its binding from the lipid environment to the transporter, ΔG_{tl}.

This two-step process, membrane partition and protein binding, has so far not been taken into account by other published *in silico* P-glycoprotein models, as presented in the above section. Moreover, due to the confusion still reigning between substrates, modulators or inhibitors of P-glycoprotein, the available models might not always be relevant to predict *intrinsic* or *apparent* substrates since they have been elaborated on experimental data coming from one type of *in vitro* assay, or even mixing the results of several assays describing different mechanisms. One should also keep in mind that numerous models try to find different features for substrates and inhibitors, which is irrelevant since most inhibitors will be transported depending on the applied concentration.[4]

Intrinsic substrates   are compounds recognized by P-glycoprotein, thus leading to a change in the basal ATP hydrolysis activity of the transporter. Those drugs will be unambiguously detected by the ATPase activity assay. However they might not be detected by the transcellular transport assay or the competition assay if they diffuse too rapidly through the cell membrane.

Apparent substrates   are intrinsic substrates large or charged enough to diffuse slowly through the cell membrane and accumulate in the inner leaflet of the lipid bilayer, so that the transporter will be able to efficiently efflux them out of the cell since the local membrane concentration of the drug will be high.

Considering the particular binding sequence of a P-glycoprotein substrate, in a first step from the aqueous environment to the lipid membrane, and in a second step from the lipid environment to the transporter, and taking into account the differences of experimental settings between the different *in vitro* assays to assess P-glycoprotein substrates and inhibitors, we were able to develop several models, named *prediction rules*, to describe the results of the different experimental assays. As mentioned previously, the binding constant

of the substrate from the aqueous environment to the transporter corresponds to the product of its partition coefficient from the aqueous environment to the lipid membrane, and its binding constant from the lipid environment to the transporter. It has been shown that the lipid-water partition coefficient can be approximated from the substrate's cross-sectional area, A_D,[10] and its air-water partition coefficient, K_{aw}, by the following equation:

$$K_{lw} = K_{aw} \exp^{-\pi_M A_D / kT},\qquad(2.1)$$

where π_M is the lateral packing density of the membrane.[7] Additionally, the free energy of substrate binding from the lipid membrane to the transporter was shown to correlate with the sum of the free energies of hydrogen-bond formation between hydrogen-bond acceptor groups on the substrate,[35] to hydrogen-bond donor groups in the transmembrane helices of the transporter,[25]

$$\Delta G_{tl} = \sum \Delta G Hi,\qquad(2.2)$$

given the fact that these hydrogen bonds are arranged in a particular spatial distance, called *type I* and *type II* units (see section 3.2.8 on page 67).[4]

Using the calculated cross-sectional area and the distribution coefficient, LogD (or the octanol-water partition coefficient, LogP, and the pKa), of the molecules, which correlate to the air-water partition coefficient, K_{aw}, the membrane-binding step could be assessed. Using the hydrogen-bonding ability to P-glycoprotein (*via* type I and type II units), the transporter-binding step could be assessed. Finally the prediction rules could be developed, combining these descriptors into a modular binding approach, to simulate the results obtained with the different experimental *in vitro* assays.

Thus we developed one prediction rule for each *in vitro* assay, named the ATPase activity assay, the competition assay, and the transcellular transport assay. The prediction rules are based on reliable *in vitro* data, collected in the publications of Schwab *et al.*,[36] Litman *et al.*,[30] and Mahar Doan *et al.*[28] The first paper presents results of all three *in vitro* assays, whereas the second one does not have results of the ATPase activity assay.

2.3.1 Prediction of the ATPase activity assay

As stated previously (on page 26), ATPase activity assay detects all *intrinsic substrates*, as well as inhibitors of P-glycoprotein, since they will modify the basal ATP hydrolysis of the transporter. The prerequisite is to make a concentration-dependant measure of the P-glycoprotein activity (and not a single concentration measurement) to get all kinetic parameters, since different drugs will have different concentrations of half-maximum P-glycoprotein activation, K_1. This explains inconsistent results observed in the publication of Schwab *et al.*,[36] where some drugs (terfenadine and miconazole) gave a

positive result in the competition assay, whereas they did not enhance the P-glycoprotein ATP hydrolysis activity.

The prediction rule for the ATPase activity assay has been established in a straightforward way: since any drug interacting with P-glycoprotein will change the basal ATP hydrolysis activity of the transporter,[30,31] only the ability of the drug to bind to the transporter, and the membrane concentration of the drug, thus the membrane binding step from the aqueous environment to the lipid bilayer, will have some influence on the outcome of the *in vitro* assay. Thus the prediction rule has only to assess the binding of the drug to the transporter *via* hydrogen bond formation, given the fact that the drug will partition into the lipid membrane. The prediction rule is the following:

```
If EU.TOTAL >= 1 And EU.TOTAL < 8 And (LogD(7.4) >= -0.5
    And LogD(7.4) < 4.5)
    Then Display("activation", 1)
ElseIf (EU.TOTAL >= 8 And LogD(7.4) >= 0) Or (EU.TOTAL >= 1
    And EU.TOTAL <= 4.5 And LogD(7.4) >= 6.1)
    Then Display("inhibition", 2)
Else Display("no activity", 3)
```

The distribution coefficient, LogD, calculated at pH 7.4, assesses the hydrophobicity of the molecule (highly hydrophilic drugs will not partition into the lipid bilayer, thus they will not be able to reach the transporter), whereas the binding to the transporter is assessed with the type I and type II units as described in the previous paragraph, whose total is implemented in the `EU.TOTAL` variable (total of energy units, corresponding to the free energy of drug binding from the lipid to the transporter, ΔG_{tl}) (see section 3.2.8 on page 67).[35] Thus compounds having a ΔG_{tl} of 0 will not change the activity of ATP hydrolysis of P-glycoprotein, whereas inhibitors will be hydrophobic compounds with a high ΔG_{tl}, and intrinsic substrates hydrophobic compounds with a low ΔG_{tl}.

Table 2.1 presents the experimental and predicted results of the ATPase activity assay on a set of 28 compounds collected from the publication of Schwab *et al.*[36] The limit between active and inactive compound has been set by the authors at 1.2 nmol P_i/mg protein/min, which corresponds to the 98% confidence level of mean background activity.

The prediction rule was able to correctly predict 21 out of 28 drugs in this dataset, which corresponds to a prediction accuracy of 75%. However several mispredicted compounds can be easily explained. Since the measurements of P-glycoprotein ATPase activity were made at one single concentration (20 μM), it is difficult to resolve intrinsic substrates from inhibitors; moreover, some drugs might give no change in the basal ATP hydrolysis if the applied concentration is too low or too high. Ivermectin has been shown to be a potent P-glycoprotein inhibitor by preventing its ATPase activation by verapamil, at a

Table 2.1: Comparison of experimental and predicted results of ATPase activity assay. EU_H corresponds to the variable *EU.TOTAL*.

Compound	LogD$_{7.4}$	EU$_H$	In vitro result[a] [nmol/mg/min]	Prediction[§]
[D-Pen2,5]Enkephalin	-2.28	3.5	4.0	*no activity*
Astemizole	4.79	1	18	activation
Cimetidine	0.32	1	2.8	activation
Clotrimazole	6.24	2	7.5	inhibition
Colchicine	1.3	5	0.1	*activation*
Cyclosporin A	2.92	11.5	9.8	inhibition
Dexamethasone	1.83	3.5	4.0	activation
Digoxin	1.26	9.5	0.3	*inhibition*
Erythromycin	2.25	10	4.0	inhibition
Etoposide	0.6	11	-0.2	*inhibition*
Hydrocortisone	1.61	2.25	0.7	*activation*
Itraconazole	5.62	5.5	18	*no activation*
Ivermectin[b]	4.04	11	-3.4	*inhibition*
Ketoconazole	4.29	5	27	activation
Mibefradil	3.43	3	8.9	activation
Miconazole[b]	6.18	1.5	1.0	no activation
Midazolam	4.32	2.5	6.0	activation
Morphine	-0.1	2.5	3.7	activation
Nelfinavir	5.16	3	6.2	activation
Nicardipine	3.64	3.75	18	activation
Pimozide	4.28	2.5	11	activation
Quinidine	1.55	1.5	20	activation
Ranitidine	-0.77	0	-2.4	no activation
Ritonavir	4.24	5.75	19	activation
Saquinavir	3.75	4	30	activation
Terfenadine[b]	5.45	1	0.4	no activation
Verapamil	2.66	3.75	21	activation
Vinblastine	3.26	9	33	inhibition

[a] The limit between no activation and activation has been set at 1.2 nmol/mg/min by the authors.

[b] These compounds have been tested within a range of 0.1-20 μM.

[§] Italic values correspond to mispredicted compounds.

concentration as low as 4 μM.[64] Colchicine has also been shown to be bound to P-glycoprotein and modify verapamil efflux by P-glycoprotein.[31,65] In the case of etoposide, it has been shown to bind to P-glycoprotein and inhibit verapamil accumulation, however the maximal inhibition occurs at 60 μM,[66] whereas the applied concentration in the publication of Schwab *et al.* was only 20 μM;[36] at this concentration no inhibition was found by Muller *et al.*[66] In a similar manner, Matsunaga *et al.* showed that 20 μM of digoxin significantly increased the ATPase activity of P-glycoprotein.[67] [D-Pen2,5]enkephalin showed an ATPase activity of 4.0 nmol/mg/min, whereas it has been predicted as non-activating. However this compound contains a carboxy group with a calculated pK_a of 3.08 (calculated with the Advanced Chemistry Development (ACD/Labs) Software V8.14 for Solaris). P-glycoprotein has been shown not to transport negatively charged compounds, thus the result of this experiment is disputable.

Litman *et al.*[30] measured the ATPase activity of 30 compounds, over a broad range of concentrations (0.2 μM-400 μM). Using a modified form of the Michaelis-Menten equation taking into account both the ascent and the decline of the ATPase activity curve (corresponding to an activating binding site, occupied at low drug concentration, and an inhibitory binding site, occupied at high drug concentration), the authors fitted the ATP hydrolysis activity of P-glycoprotein. From the kinetic parameters, we calculated the ATPase activity for each drug at 20 μM, and used the prediction rule displayed below to predict the activation-inhibition.

```
If pKa(acidic) > 5 And EU.TOTAL >= 1 And EU.TOTAL < 7 And
   LogP > -0.5 And LogP < 4.5
   Then Display("activation", 0)
ElseIf pKa(acidic) > 5 And ((EU.TOTAL >= 7 And LogP > 0) Or
   (EU.TOTAL >= 1 And EU.TOTAL <= 4.5 And LogP > 6))
   Then Display("inhibition", 0)
Else Display("no activity", 0)
```

Table 2.2 presents the experimental and the predicted values for the ATPase activity assay for 30 compounds, collected from the publication of Litman *et al.*,[30] at a drug concentration of 20 μM. Colchicine has been excluded from the dataset due to solubility problem.

The prediction rule presented above was able to correctly predict the ATPase activity of P-glycoprotein for 29 out of 30 drugs (96.7%), with 13 compounds leading to inhibition, two compounds showing no activity, and 15 compounds leading to ATPase activation. The mispredicted compound, valinomycin, showed a strong activation at low concentration ($V_1 = 2.3$, $K_1 = 0.03$ μM) and a strong inhibition at high drug concentration ($V_2 = 0$, $K_2 = 16.1$ μM). Thus the K_2 of the ATPase activity at 20 μM is close to 1, leading to the experimental misprediction if 1 μM is taken as concentration.

Table 2.2: Comparison of experimental and predicted results of ATPase activity assay. EU_H corresponds to the variable *EU.TOTAL*.

Compound	LogP	EU_H	$pK_a^†$	ATPase[‡]	Prediction[§]
Amiodarone	7.8	1.66	8.45 / 43	0.83	inhibition
Amitriptyline	4.92	1	9.24 / 43	1.19	activation
Chlorpromazine	5.41	1.75	9.43 / 43	1.38	activation
Cyclosporin A	2.92	12.25	4.74 / 15.5	0.64	inhibition
Daunorubicin	1.83	9.25	8.64 / 7.15	0.83	inhibition
Diltiazem	2.79	2.75	8.91 / 43	1.48	activation
Dipyridamole	3.07	3.9	6.37 /13.54	1.83	activation
Epirubicin	1.27	9.7	8.64 / 7.12	0.93	inhibition
Fluphenazine	4.36	2.7	7.9 / 15.5	1.39	activation
Fusidine	4.97	1.55	– / 4.8	1.06	no activity
Gramicidin S	8	8.17	10.5 / 43	0.22	inhibition
Mefloquine	2.87	1.75	10.13/13.13	0.94	activation
Methotrexate	-1.85	1.7	5.23 / 4.7	1	no activity
Pimozide	6.3	2.5	9.42 /12.11	0.7	inhibition
Progesterone	3.87	1.5	– / 50	1.8	activation
Promethazine	4.81	1.33	8.98 / 43	1.36	activation
Propafenone	3.37	1.7	9.31 /13.82	1.78	activation
Propranolol	3.48	1.25	9.14 /13.84	1.11	activation
PSC-833	1.64	13	4.74 / 50	0.31	inhibition
Quinidine	3.44	1.5	8.56 /13.05	1.46	activation
Reserpine	3.32	7	6.6 / 43	0.94	inhibition
Spironolactone	2.78	4.8	– / 43	1.7	activation
Tamoxifen	6.3	1.1	8.69 / 43	0.73	inhibition
Terfenadine	7.62	1	9.57 /13.31	0.75	inhibition
Trifluoperazine	5.03	2.66	7.82 / 43	1.31	activation
Triflupromazine	5.54	2.25	9.43 / 43	1.44	activation
Valinomycin	4.49	14	-0.77 /11.32	1.03	*inhibition*
Verapamil	3.79	3.65	8.92 / 43	1.77	activation
Vinblastine	3.7	9	7.64 /11.36	0.96	inhibition
Vincristine	2.82	7.75	5 / 43	0.91	inhibition

[†] The first value corresponds to the basic pK_a, the second value to the acidic pK_a.

[‡] These compounds have been tested within a range of 0.2-400 μM; the value corresponds to the ATPase activity at 20 μM.

[§] Italic values correspond to mispredicted compounds.

2.3.2   Prediction of the competition assay

As discussed in section 2.1.3, the competition assay is based on the displacement of the P-glycoprotein substrate calcein-AM, which becomes fluorescent and negatively charged (thus preventing its diffusion out of the cell) once inside the cytoplasm. A modulator is an *intrinsic* P-glycoprotein substrate, having a higher binding affinity toward the transporter than calcein-AM. To predict the competition of a molecule against calcein-AM, we thus established a prediction rule taking into account (i) the number of type I or type II units present on the molecule (to be able to bind to the transporter), and the lipophilicity of the molecule, to assess whether the molecule will enter the cell membrane. Indeed, a molecule which is too hydrophilic will not be able to partition into the lipid bilayer. The size of the molecule, i.e. its cross-sectional area, has not to be taken into account, since the competition assay will detect all *intrinsic* substrates.

We collected experimental data from the publication of Schwab *et al.*,[36] in which 28 compounds were tested in a concentration range of 0.05 to 50 μM, against a concentration of 0.25 μM of calcein-AM. The positive control was the P-glycoprotein substrate verapamil applied at a concentration of 50 μM. The prediction rule developed for this dataset of 28 compounds was the following:

```
If LogP > 3 And EU.TOTAL >= 0.5 And pKa(acidic) > 5
    Then Display("competition", 0)
Else Display("no competition", 0)
```

Indeed, the molecule has to be hydrophobic enough (large octanol-water partition coefficient), and contain at least one type I or type II unit, in order to bind to P-glycoprotein. Moreover, since P-glycoprotein does not transport negatively charged compounds, we set a limit for acidic pK_a to rule-out compounds having for instance a carboxy group which will be negatively charged at physiological pH. Table 2.3 presents the experimental and the predicted values for the calcein-AM competition assay, for the dataset of 28 drugs tested by Schwab *et al.*[36]

The prediction rule was able to correctly predict 25 out of 28 drugs (89.3%), with one mispredicted inhibitor (saquinavir), and two mispredicted non-inhibitors (erythromycin and midazolam). Interestingly, saquinavir was a border-line inhibitor, since it showed an IC_{50} higher than 50 μM on cell lines overexpressing either mouse mdr1a or mouse mdr1b P-glycoprotein. On the other hand, midazolam showed an IC_{50} lower than 50 μM (which corresponds for the authors to the limit between inhibitor and non-inhibitors) on cell lines overexpressing mdr1b P-glycoprotein ($IC_{50} = 33$ μM). In a similar manner, two border-lines compounds, erythromycin and vinblastine, were classified as inhibitors: the former showed an IC_{50} of 43 μM on porcine P-glycoprotein overexpressing cell

Table 2.3: Comparison of experimental and predicted results of calcein-AM competition assay. EU_H corresponds to the variable *EU.TOTAL*.

Compound	LogP	EU_H	calcein-AM[a]	Prediction[§]
[d-pen]enkephalin	2.01	3.5	>50	no competition
astemizole	6.43	2	1.3	competition
cimetidine	0.57	2	>50	no competition
clotrimazole	6.26	2	6.7	competition
colchicine	1.3	7	>50	no competition
cyclosporin a	4.12	13	0.8	competition
dexamethasone	1.83	3.25	>50	no competition
digoxin	0.5	14	>50	no competition
erythromycin	3.06	14.5	>50	*competition*
etoposide	0.6	13	>50	no competition
hydrocortisone	1.61	3.5	>50	no competition
itraconazole	5.66	7	2.1	competition
ivermectin	4.11	15.5	0.1	competition
ketoconazole	4.35	7	4.8	competition
mibefradil	6.27	4.5	1.8	competition
miconazole	6.25	2.5	3.5	competition
midazolam	4.33	2.75	>50	*competition*
morphine	0.89	3	>50	no competition
nelfinavir	5.53	4	3.4	competition
nicardipine	3.82	7	2.3	competition
pimozide	6.3	2.5	2.9	competition
quinidine	3.44	2	5.6	competition
ranitidine	0.27	1.5	>50	no competition
ritonavir	6.27	6.5	12	competition
saquinavir	2.55	4	12	*no competition*
terfenadine	7.62	1	1.4	competition
verapamil	3.79	6	6.3	competition
vinblastine	3.7	9	>50	competition

[a] Calcein-AM corresponds to the concentration, in μM, of the drug inhibiting 50% of calcein-AM, compared to the inhibition of 50 μM verapamil.

[§] Italic values correspond to mispredicted compounds.

lines, and the latter showed an IC_{50} higher than 50 μM in mdr1a mouse P-glycoprotein and human P-glycoprotein overexpressing cell lines. However they were as well close the the separation between inhibitors and non-inhibitors.

We collected a second set of *in vitro* data from the publication of Mahar Doan *et al.*[28] They measured the calcein-AM inhibition of 93 drugs, applied at a concentration of 100 μM. Calcein-AM concentration was 0.25 μM, whereas the positive control was the P-glycoprotein substrate GF120918 (elacridar) applied at a concentration of 1 μM. The discrimination between inhibitors and non-inhibitors has been set by the authors at a limit of 10% of maximum inhibition (inhibition observed with 1 μM GF120918 against calcein-AM). Since the experimental settings were different from those used by Schwab *et al.*,[36] we had to define a new prediction rule, to discriminate inhibitors from non-inhibitors. The prediction rule was the following:

```
If (EU.TOTAL >= 0.5 And LogP >= 3.5) Or (EU.TOTAL >= 2.5
    And LogP > 2.5) And pKa(acidic) > 5
    Then Display("inhibition", 0)
Else Display("non inhibition", 0)
```

The prediction rule is however very similar to the rule established for the dataset of Schwab *et al.*[36] The dataset of Mahar Doan *et al.*[28] contained four highly negatively charged compounds which are not transported by P-glycoprotein (acrivastine, cetirizine, indomethacin, and sulfasalazine). Table 2.4 presents the experimental and the predicted values for the dataset of 93 compounds.

Table 2.4: Comparison of experimental and predicted results of calcein-AM competition assay.

Compound	LogP	EU_H[1]	calcein-AM[2]	Prediction[3]
Acrivastine	4.33	2	3.57	no inhibition
Alprenolol	3.1	2	15.4	inhibition
Amantadine	2.44	0	2.1	no inhibition
Amitriptyline	4.92	1	25.5	inhibition
Amprenavir	3.1	6	14.7	inhibition
Antipyrine	0.38	1.5	1.02	no inhibition
Astemizole	7.88	2	71.5	inhibition
Atenolol	0.16	2	0.07	no inhibition
Biperiden	4.25	0	4.97	no inhibition

[1]The value of EU_H corresponds to the maximal number of type I and type II units found in all conformations for each compound.

[2]The values correspond to the percentage of inhibitory potential of the drug at 100 μM, compared to the inhibitory activity of 1 μM GF120918, a potent and specific P-glycoprotein inhibitor. The limit between non-inhibition and inhibition has been set at 10% by the authors.

[3]Italic values correspond to mispredicted compounds.

Table 2.4: (continued)

Compound	LogP	EU_H	calcein-AM	Prediction
Bromocriptine	3.82	6.5	61.6	inhibition
Bufuralol	3.5	2	14.3	inhibition
Buspirone	2.63	4	24.1	inhibition
Carbamazepine	2.45	2.5	1.22	no inhibition
Cetirizine	4.66	2.5	3.32	no inhibition
Chlorpheniramine	3.38	2	2.36	*inhibition*
Chlorpromazine	5.41	2	46.9	inhibition
Chlorprothixene	5.18	1	57.5	inhibition
Cimetidine	2.05	2	0.16	no inhibition
Clemastine	5.95	2.5	28.4	inhibition
Clomipramine	5.19	2	48.2	inhibition
Clonidine	3.98	0	0.63	no inhibition
Cyclobenzaprine	4.78	1	23.1	inhibition
Desipramine	4.9	1	12.8	inhibition
Diltiazem	2.7	4.5	37.3	inhibition
Diphenhydramine	3.27	2.5	11.1	inhibition
Domperidone	3.9	3	10.4	inhibition
Doxapram	5.75	4	11.7	inhibition
Doxepin	4.03	2.5	15.3	inhibition
Doxylamine	3.27	2.5	5.37	*inhibition*
Eletriptan	5.08	3	2.6	*inhibition*
Famciclovir	2.46	6	3.12	no inhibition
Flumazenil	1	4.5	3.89	no inhibition
Fluoxetine	3.82	1.5	23.5	inhibition
Flurazepam	6.58	3.25	22.6	inhibition
Fluvoxamine	4.03	2.5	20.8	inhibition
Guanabenz	3.94	0	8.6	no inhibition
Guanfacine	2.2	0	5.05	no inhibition
Haloperidol	4.3	2.5	37.2	inhibition
Imipramine	4.8	2	5.37	*inhibition*
Indinavir	2.92	4.5	1.81	*inhibition*
Indomethacin	4.27	2.5	4.63	no inhibition
Ketamine	2.18	0	1.23	no inhibition
Labetalol	3.09	3	7.11	*inhibition*
Levomeprazine	4.68	3.5	25.6	inhibition
Lidocaine	2.44	0	1.7	no inhibition
Loperamide	6.29	4	76.3	inhibition
Lorcainide	4.85	2.5	16.4	inhibition
Mannitol	-2.2	0	0.35	no inhibition
Maprotiline	4.49	1	36.7	inhibition

Table 2.4: (continued)

Compound	LogP	EU_H	calcein-AM	Prediction
Mephentermine	2.78	0	1.43	no inhibition
Meprobamate	0.7	4	2.24	no inhibition
Mequitazine	5.68	2	33.5	inhibition
Metergoline	4.79	1.5	84.4	inhibition
Methysergide	2.45	2	3.83	no inhibition
Metoprolol	1.88	2	5.08	no inhibition
Mexilitene	2.15	0	0.37	no inhibition
Midazolam	6.72	2.75	74.9	inhibition
Nalbuphine	1.31	2	0.07	no inhibition
Naloxone	2.09	3	0.73	no inhibition
Naltrexone	1.92	3	2.03	no inhibition
Nelfinavir	5.91	4	30.6	inhibition
Neostigmine	1.77	1.5	1.91	no inhibition
Nitrazepam	2.46	1	3.66	no inhibition
Nordiazepam	2.93	1	5.17	no inhibition
Nortriptylene	4.32	1	13.7	inhibition
Noscapine	1.94	6.5	35.5	*no inhibition*
Oxprenolol	2.1	2	5.53	no inhibition
Perphenazine	4.2	2	65.4	inhibition
Pheniramine	2.89	2	1.49	no inhibition
Pirenzepine	2.4	3.5	2.16	no inhibition
Procyclidine	4.35	0	16.2	*no inhibition*
Progabide	3.06	3	14.1	inhibition
Promazine	4.55	2	16.3	inhibition
Promethazine	4.81	2	16.3	inhibition
Propranolol	3.48	2	6.22	*inhibition*
Protriptylene	4.09	1	15	inhibition
Risperidone	4.11	3.75	52.6	inhibition
Roxindole	5.02	1	32.7	inhibition
Saquinavir	3.58	4	35.5	inhibition
Scopolamine	0.62	4	0.54	no inhibition
Selegiline	2.9	0	3.02	no inhibition
Sulfasalazine	2.88	4	3.45	no inhibition
Sumatriptan	0.93	3	0.76	no inhibition
Tacrine	2.71	0	4.76	no inhibition
Terfenadine	7.08	1	100	inhibition
Trazodone	3.58	3.5	21.1	inhibition
Trimethoprim	0.91	5	1.35	no inhibition
Trimipramine	4.97	2	22.8	inhibition
Verapamil	3.79	6	56.4	inhibition

Table 2.4: (continued)

Compound	LogP	EU_H	calcein-AM	Prediction
Warfarin	2.6	4	1.43	no inhibition
Zimeldine	3.99	2	7.4	*inhibition*
Zolmitriptan	2.22	2	1.01	no inhibition
Zolpidem	3.4	1.5	17.3	inhibition

From the 93 compounds, with 45 central nervous system targeted drugs and 48 peripherically targeted drugs, the prediction rule was able to correctly discriminate 83 compounds (89.2%). From the 47 compounds showing an inhibition lower than 10%, 6 were mispredicted (82.2% of correct prediction), and from the 46 compounds showing an inhibition higher than 10%, 4 compounds were mispredicted (91.3% of correct prediction). Only one inhibitor was mispredicted due to its absence of type I or type II units, procyclidine (Figure 2.5). Indeed, this compound shows no usual type I unit, however it
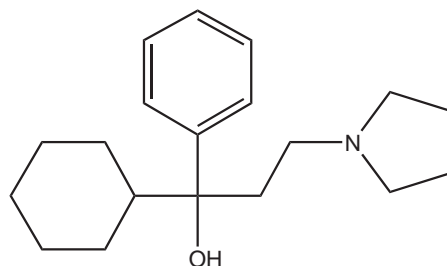
Figure 2.5: Two-dimensional structure of procyclidine.

has been shown that phenyl groups and tertiary nitrogen can be involved in type I or type II units,[35] but the fact that a phenyl ring can be involved in a type I or type II unit together with a tertiary amine has not been described so far. However this is in accordance to a type II (1, 5) pattern as described by Seelig.[35] Another explanation might be that two molecules of procyclidine bind together in the transmembrane channel of P-glycoprotein. If the presence of a type I or type II unit between a phenyl group and a tertiary nitrogen, or the presence of two molecules in the transporter at the same time is confirmed, this would imply that procyclidine would be correctly predicted, since its LogP is 4.35.

A closer look at the mispredicted compounds shows that generally all of them are close to the limit defined in the prediction rule, except procyclidine (as explain above). Several mispredicted non-inhibitors showed a positive result in the transcellular transport assay (indinavir, labetalol, eletriptan, diphen-

hydramine). These incoherences between transcellular transport assay and competition assay indicates that these drugs are *apparent* substrates of P-glycoprotein, but most likely bind to P-glycoprotein more weakly than calcein-AM, and are thus not able to efficiently compete against it.

We also modelled the substrate of P-glycoprotein used by the authors as positive control, GF120918; with a LogP of 4.98 and a maximal number of type I and type II units of 5, this compound was correctly predicted as modulator of calcein-AM towards P-glycoprotein (see page 23). As a comparison, the modelling of calcein-AM showed a maximal number of type I and type II units of 11, and a LogP of 3.49. This compound is thus also correctly predicted as a P-glycoprotein substrate.

2.3.3   Prediction of the transcellular transport assay

The results of the transcellular transport assay might be the most difficult to predict. As stated previously, an *intrinsic* substrate, meaning a compound that will be transported by P-glycoprotein and thus give a change in the ATP hydrolysis by the transporter, might not be detected by the transcellular transport assay. Indeed, this assay will only give positive results for *apparent* substrates, which are intrinsic substrates diffusing slowly enough to be efficiently effluxed by P-glycoprotein, and thus accumulating in the apical region of the cells. For the prediction of the transcellular transport assay, one has to consider several parameters:

- The hydrophobicity of the molecule: if a compound is too hydrophilic, it will not be able to partition into the lipid bilayer of the cells, and thus will not reach the transporter. This parameter is assessed by means of the octanol-water partition coefficient, LogP.

- The charge of the molecule: only the uncharged fraction of a drug will permeate through the lipid membrane of the cells; the ionization constant of the molecules plays thus an important role in the diffusion through the cell membrane.

- The size of the molecule: if a compound diffuses too rapidly through the membrane, P-glycoprotein will not be able to efflux it in an efficient manner, thus the compound will not accumulate in the apical part of the cells, giving a negative result in the transcellular transport assay. In contrast, a compound which is "big" enough to diffuse slowly, and is efficiently effluxed, will give a positive result for the transport assay. However, if a molecule is too large, it will not be able to partition into the lipid bilayer. In a first step, we assessed the size of the molecule with its molecular weight. Once the algorithm for the calculation of the cross-sectional area was achieved, the size of the molecule was assessed

in a second step with the calculated cross-sectional area, as described elsewhere.[10]

- The interaction with P-glycoprotein: as described previously, the binding affinity of a molecule to the transporter is assessed using the number of hydrogen-bond formation between them, using type I and type II units.[35]

We collected experimental data for transcellular transport assay for 13 drugs, in the publication of Schwab *et al.*[36] These drugs were also tested for the calcein-AM competition assay, as well as ATPase assay. The authors used wild type pig kidney cells (LLC-PK1), as well as transfected either with human (MDR1) or mouse (mdr1a) P-glycoprotein gene, at a density of $1.4 \cdot 10^5$ cells/well leading to a confluent cells monolayer. The tightness of the monolayer was assessed with inulin, a paracellular flux marker (as it cannot diffuse through the cell membrane). The specific activity of the test compounds was 3 μCi/mL, corresponding to compound concentrations of less than 350 nM. The tested drug was applied in the apical part of the cells, and samples were collected from the basolateral part of the cells; in a second experiment, the drug was applied in the basolateral part, and samples collected from the apical part. The first prediction rule developed for this dataset of 13 compounds was the following:

```
If EU.TOTAL >= 0.5 and pKa(acidic) > 5 and (pKa(basic) >= 9 Or
   MW >= 500 Or (LogP < 2 And LogP > 0.5))
   Then Display("transport", 1)
Else Display("no transport", 3)
```

Table 2.5 presents the experimental and predicted results of the transcellular transport assay on the set of 13 compounds collected from the publication of Schwab *et al.*[36] The limit between transported and not transported compounds has been set by the authors at 1.5 for the ratio $\frac{B \to A}{A \to B}$. The first prediction rule to estimate the transport of drug through P-glycoprotein overexpressing cells, using the molecular weight as a parameter to assess the size of the molecules, was able to correctly predict the transcellular transport of 11 out of 12 drugs (91.7%). The only mispredicted compound, erythromycin, was predicted as being accumulated on the apical side of the cells, whereas the *in vitro* transcellular transport assay gave a negative result. However this compound showed an activation of the ATPase activity of P-glycoprotein, proving that it is an *intrinsic* substrate. Moreover, Schuetz *et al.* showed that the rate of $A \to B$ flux of erythromycin was diminished in P-glycoprotein overexpressing cells;[68] their setup was identical as the one used by Schwab *et al.*,[36] plating $2 \cdot 10^6$ cells/well, and applying the drug at a final concentration of 2 or 5 μM (0.25 μCi/mL). The $\frac{B \to A}{A \to B}$ ratios are not given, however, they can be estimated from the transcellular transport measurements. For wild type, the transport ratio was $\frac{7.6}{4.7} = 1.62$ after 4 hours of experiment. For

Table 2.5: Comparison of experimental and predicted results of the transcellular transport assay. EU_H corresponds to the variable *EU.TOTAL*.

Compound	LogP	EU_H	MW	pK_a[†]	$\frac{B \to A}{A \to B}$	Prediction[§]
[d-pen]enkephalin	2.01	3.5	645.79	7.9 / 3.08	0.8	no transport
benzylpenicillin	1.83	3	334.39	-1.32/ 2.62	1	no transport
cimetidine	0.57	2	252.34	6.73 / 16.5	0.5	no transport
colchicine	1.3	7	399.44	– /14.87	2.2	transport
cyclosporin a	4.12	13	1202.61	4.74 / 15.5	7.7	transport
dexamethasone	1.83	3.25	392.46	– /12.14	2.8	transport
digoxin	0.5	14	780.94	– / 13.5	4.4	transport
erythromycin	3.06	14.5	733.93	8.14 /13.08	1.2	*transport*
mibefradil	6.27	4.5	495.63	9.3 /12.55	2.1	transport
morphine	0.89	3	285.34	8.26 / 9.5	1.5	transport
saquinavir	2.55	4	670.84	7.35 / 15.5	3.4	transport
verapamil	3.79	6	454.6	8.5 / 43	5.7	transport
vinblastine	3.7	9	810.97	7.64 /11.36	1.7	transport

[†] The first value corresponds to the basic pK_a, the second value to the acidic pK_a.
[§] Italic values correspond to mispredicted compounds.
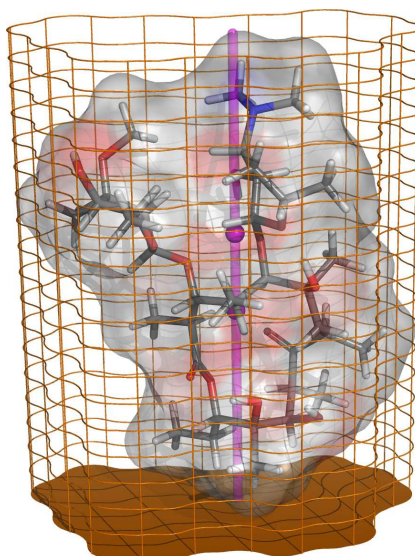
Figure 2.6: Three-dimensional structure of erythromycin, oriented at an hydrophilic-hydrophobic interface (see Table 2.6).

the P-glycoprotein overexpressing pig kidney cells, the transport ratio was $\frac{10.35}{1.55} = 6.68$ for human MDR1 gene, and $\frac{7.6}{1.55} = 4.9$ for mouse mdr1a gene. The corrected value (corresponding to the experimental value in Table 2.5) was thus 4.12 for the human P-glycoprotein. As as consequence, this compound was detected as transported by Schuetz *et al.*

Once the algorithm for the calculation of the cross-sectional area has been validated,[10] we established a second prediction rule for estimation of the results of the transcellular transport assay, which consider the cross-sectional area instead of the molecular weight to assess the size of the molecules. The prediction rule was the following:

```
If pKa(acidic) >= 5 And Ad[7.4] < 170 And EU.TOTAL >= 0.5 And
   (Ad[7.5] > 75 Or pKa(basic) >= 9)
   Then Display("transport", 1)
Else Display("no transport", 3)
```

Table 2.6 presents the results of the experimental transcellular transport assay, as well as the results of the prediction using the second prediction rule. The calculated cross-sectional area has been also included in the table.

Table 2.6: Comparison of experimental and predicted results of the transcellular transport assay, using the second prediction rule. EU_H corresponds to the variable *EU.TOTAL*, A_D to the calculated cross-sectional area.

Compound	LogP	EU_H	A_D [Å²]	pK_a[†]	$\frac{B \to A}{A \to B}$	Prediction[§]
[d-pen]enkephalin	2.01	3.5	100.72	7.9 / 3.08	0.8	no transport
benzylpenicillin	1.83	3	42.52	-1.32/ 2.62	1	no transport
cimetidine	0.57	2	42.01	6.73 / 16.5	0.5	no transport
colchicine	1.3	7	90.71	– /14.87	2.2	transport
cyclosporin a	4.12	13	155.44	4.74 / 15.5	7.7	transport
dexamethasone	1.83	3.25	76.28	– /12.14	2.8	transport
digoxin	0.5	14	86.56	– / 13.5	4.4	transport
erythromycin	3.06	14.5	95.44	8.14 /13.08	1.2	*transport*
mibefradil	6.27	4.5	79.43	9.3 /12.55	2.1	transport
morphine	0.89	3	48.01	8.26 / 9.5	1.5	*no transport*
saquinavir	2.55	4	140.22	7.35 / 15.5	3.4	transport
verapamil	3.79	6	81.9	8.5 / 43	5.7	transport
vinblastine	3.7	9	129.18	7.64 /11.36	1.7	transport

[†] The first value corresponds to the basic pK_a, the second to the acidic pK_a.
[§] Italic values correspond to mispredicted compounds.

The second prediction rule was able to correctly predict 11 out of 13 compounds (84.6%); one of the mispredicted compound was erythromycin (as with the first prediction rule), whereas the second mispredicted compound was

morphine. This compound showed no modulation in the calcein-AM assay,[36] and a $\frac{B \rightarrow A}{A \rightarrow B}$ ratio of 1.5 in the same publication (which corresponds to the limit between transported and non-transported compounds), and a ratio lower than 1.5 ($\frac{B \rightarrow A}{A \rightarrow B} = 1.47$ at 10 μM morphine, and $\frac{B \rightarrow A}{A \rightarrow B} = 1.37$ at 150 μM morphine) in the publication of Crowe (on Caco-2 cells).[69] Morphine has a relatively small cross-sectional area, thus this compound will diffuse rapidly through the cell membrane. It contains a total of 3 type I or type II units, it will thus interact with P-glycoprotein. Morphine is a *intrinsic* substrate (giving a change in the ATP hydrolysis by the transporter), and a weak *apparent* substrate due to its small cross-sectional area and pretty low basic pK_a. However the calculated cross-sectional area represents a better descriptor for the passive diffusion of the molecules through the lipid bilayer, and to assess the accumulation of the drugs in the inner leaflet of the membrane, which is an important parameter for efficient efflux of the drugs out of the cells by P-glycoprotein.

To assess the results of the *in vitro* data of a second dataset of transcellular transport experiments, we adapted the second prediction rule to the experimental settings of Mahar Doan *et al.*[28] They measured the bidirectional efflux of the same dataset of 93 drugs presented previously (on page 40), using canine kidney cells (MDCK) overexpressing human MDR1 P-glycoprotein gene. The cells were seeded at a density of $3 \cdot 10^5$ cells/well and the drugs applied at a concentration of 10 μM. Paracellular diffusion was assessed using mannitol, a highly hydrophilic compound (LogP$_{exp}$ = -2.47), and the positive control for P-glycoprotein transport was amprenavir. A drug was defined as transported when the $\frac{B \rightarrow A}{A \rightarrow B}$ ratio was higher than 1.5. All transported drugs were then tested in the presence of 2 μM of GF120918 (cf. page 40), and the transport ratio decreased in all cases to $\simeq 1$. The prediction rule was the following:

```
If LogD(7.4) < 0.49 And EU.TOTAL >= 0.5 And Ad >= 42
   Then Display("if entry", "0, 0, 0, 118, 156, 234")
ElseIf ((EU.TOTAL >= 0.75 And Ad > 75) Or (EU.TOTAL >= 2 And Ad > 68)
   Or (EU.TOTAL >= 3 And Ad > 43)) And Ad < 142 And LogD(7.4) >= 0.5
   Then Display("transport", 3)
Else Display("no transport", 1)
```

Table 2.7 contains the results of the experimental data of the transport assay (corresponding to the quotient between the $\frac{B \rightarrow A}{A \rightarrow B}$ ratio in the presence of the tested drug, and the $\frac{B \rightarrow A}{A \rightarrow B}$ ratio in the presence of the tested drug and 2 μM GF120918, a potent inhibitor of P-glycoprotein), and the prediction of the *in vitro* assay using the third prediction rule defined above. The limit between transport and no transport has been set by the authors at 1.5 for the ratio $\frac{B \rightarrow A}{A \rightarrow B}$.

Table 2.7: Comparison of experimental and predicted results of the transcellular transport assay.

Compound	LogP	EU_H	A_D [Å^2]	pK$_a$[1]	$\frac{B \to A}{A \to B}$	Prediction[2]
Acrivastine	4.33	2	59.41	9.01/1.99	4.26	if entry
Alprenolol	3.1	2	39.98	9.17/13.88	1.06	no transport
Amantadine	2.44	0	38	10.75/-	1.25	no transport
Amitriptyline	4.92	1	53.75	9.24/-	0.93	no transport
Amprenavir	3.1	6	72.75	1.76/11.54	22.66	transport
Antipyrine	0.38	1.5	36.58	0.7/-	1.09	no transport
Astemizole	7.88	2	77.13	9.03/-	1.66	transport
Atenolol	0.16	2	75.82	9.17/13.88	1	*if entry*
Biperiden	4.25	0	59.14	9.8/-	1.07	no transport
Bromocriptine	3.82	6.5	106	6.72/9.6	1.21	*transport*
Bufuralol	3.5	2	48.28	8.97/13.67	0.83	no transport
Buspirone	2.63	4	36.14	6.43/-	1.04	no transport
Carbamazepine	2.45	2.5	62.15	13.94/-	0.87	*if entry*
Cetirizine	4.66	2.5	61.84	6.43/3.27	7.52	if entry
Chlorpheniramine	3.38	2	53.67	9.33/-	0.99	no transport
Chlorpromazine	5.41	2	48.44	9.43/-	1.03	no transport
Chlorprothixene	5.18	1	54.5	9.15/-	0.83	no transport
Cimetidine	2.05	2	41.99	6.73/-	3.82	if entry
Clemastine	5.95	2.5	57.48	10.23/-	1.39	no transport
Clomipramine	5.19	2	52.92	9.43/-	1.29	no transport
Clonidine	3.98	0	37.7	9.16/-	0.94	no transport
Cyclobenzaprine	4.78	1	53.76	9.21/-	0.88	no transport
Desipramine	4.9	1	47.71	10.4/-	1.1	no transport
Diltiazem	2.7	4.5	68.34	8.94/15	1.74	transport
Diphenhydramine	3.27	2.5	54.59	1.14/43	1.11	no transport
Domperidone	3.9	3	79.84	8.76/-	32.5	transport
Doxapram	5.75	4	67.53	9.01/11.11	1.41	*transport*
Doxepin	4.03	2.5	50.7	8.47/-	0.88	no transport
Doxylamine	3.27	2.5	51.39	9.19/15	0.88	no transport
Eletriptan	5.08	3	51.88	8.69/-	41.01	transport
Famciclovir	2.46	6	43.59	10.35/-	2.33	transport
Flumazenil	1	4.5	40.73	4.58/43	1.05	no transport
Fluoxetine	3.82	1.5	57.55	10.05/43	1.27	no transport
Flurazepam	6.58	3.25	55.28	9.79/15	1.06	*transport*

[1]The first value corresponds to the basic pK$_a$, the second to the acidic pK$_a$.

[2]"transport" corresponds to a transported compound; "no transport" corresponds to a compound which is not transported or whose passive diffusion is too fast to allow apical accumulation; "if entry" corresponds to a compound which might be efficiently transported if it manages to enter the cell membrane. Italic values correspond to mispredicted compounds.

Table 2.7: (continued)

Compound	LogP	EU_H	A_D [Å^2]	pK_a	$\frac{B\rightarrow A}{A\rightarrow B}$	Prediction
Fluvoxamine	4.03	2.5	44.77	9.39/43	1.02	no transport
Guanabenz	3.94	0	31.62	9.66/-	1.11	no transport
Guanfacine	2.2	0	37.7	7.34/11.07	1.02	no transport
Haloperidol	4.3	2.5	56.7	8.25/13.85	1.11	no transport
Imipramine	4.8	2	47.86	9.49/-	1.01	no transport
Indinavir	2.92	4.5	106.25	5.19/14.4	9.76	transport
Indomethacin	4.27	2.5	44.25	-9.2/3.96	0.88	no transport
Ketamine	2.18	0	49.8	6.59/-	1	no transport
Labetalol	3.09	3	45.07	9.2/8.21	8.43	transport
Levomeprazine	4.68	3.5	50.03	9.34/43	1.22	*transport*
Lidocaine	2.44	0	42.63	8.53/-	0.98	no transport
Loperamide	6.29	4	96.2	8.05/13.85	9.2	transport
Lorcainide	4.85	2.5	57.45	9.54/-	1.45	no transport
Mannitol	-2.2	0	41.24	-/13.14	0.75	no transport
Maprotiline	4.49	1	48.13	10.63/-	0.9	no transport
Mephentermine	2.78	0	33.96	10.38/-	0.76	no transport
Meprobamate	0.7	4	39.38	-1.09/13.09	1.02	no transport
Mequitazine	5.68	2	62.89	10.3/-	2.2	transport
Metergoline	4.79	1.5	74.45	9.54/12.45	1.02	no transport
Methysergide	2.45	2	54.73	7.48/14.46	4.46	*no transport*
Metoprolol	1.88	2	40.85	9.17/13.89	1.26	no transport
Mexilitene	2.15	0	34.55	8.58/43	0.74	no transport
Midazolam	6.72	2.75	65.79	5.65/-	0.96	*transport*
Nalbuphine	1.31	2	60.45	7.35/9.39	2.15	transport
Naloxone	2.09	3	54.36	6.73/9.16	1.48	*transport*
Naltrexone	1.92	3	57.09	7.53/9.16	1.01	*transport*
Nelfinavir	5.91	4	115.52	7.53/9.58	13.94	transport
Neostigmine	1.77	1.5	60.61	-3.12/-	2.3	if entry
Nitrazepam	2.46	1	56.72	3.19/11.35	1.23	*if entry*
Nordiazepam	2.93	1	55.22	3.4/11.72	0.92	no transport
Nortriptylene	4.32	1	52	10.08/-	1.32	no transport
Noscapine	0.56	6.5	78.21	6.32/43	1.06	*transport*
Oxprenolol	2.1	2	43.28	9.17/13.89	1.56	if entry
Perphenazine	4.2	2	52.54	6.85/14.96	1.08	no transport
Pheniramine	2.89	2	53.28	9.35/-	1.09	no transport
Pirenzepine	2.4	3.5	57.79	7.72/11.29	2.16	transport
Procyclidine	4.35	0	57.62	10.48/14.2	0.93	no transport
Progabide	3.06	3	50.28	10.31/16.48	0.93	*if entry*
Promazine	4.55	2	44.49	9.43/-	0.98	no transport
Promethazine	4.81	2	50.76	9.25/-	0.83	no transport

Table 2.7: (continued)

Compound	LogP	EU_H	A_D [Å^2]	pK_a	$\frac{B \rightarrow A}{A \rightarrow B}$	Prediction
Propranolol	3.48	2	36.85	9.14/13.84	1.21	no transport
Protriptylene	4.09	1	51.92	10.61/-	2.35	*no transport*
Risperidone	4.11	3.75	54.37	7.89/43	1.66	transport
Roxindole	5.02	1	62.07	8.64/10.15	1.15	no transport
Saquinavir	3.58	4	140.63	7.61/11.05	53.16	transport
Scopolamine	0.62	4	38.69	8.01/14.11	1.09	no transport
Selegiline	2.9	0	32.53	7.53/-	0.62	no transport
Sulfasalazine	2.88	4	43.45	1.86/2.88	0.76	*if entry*
Sumatriptan	0.93	3	54	9.49/11.31	1.12	*if entry*
Tacrine	2.71	0	36.87	9.64/-	1.18	no transport
Terfenadine	7.08	1	95.12	9.57/13.32	2.38	transport
Trazodone	3.58	3.5	40.48	6.73/-	1.12	no transport
Trimethoprim	0.91	5	48.19	7.34/-	1.76	if entry
Trimipramine	4.97	2	56.12	9.37/-	0.79	no transport
Verapamil	3.79	6	81.95	8.92/-	1.75	transport
Warfarin	2.6	4	70.22	-/9.08	0.9	*transport*
Zimeldine	3.99	2	50.38	8.07/-	0.97	no transport
Zolmitriptan	2.22	2	51.56	9.52/12.57	2	if entry
Zolpidem	3.4	1.5	51.07	6.91/-	1.37	no transport

The prediction rule was able to correctly predict the result of the transcellular transport assay of 82 out of 93 drugs (88.2%), with two false negatives, and nine false positives. Having a closer look to the mispredicted compounds led to the observation that from the 9 false positives, two had a $\frac{B \rightarrow A}{A \rightarrow B}$ ratio between 1.4 and 1.5 (naloxone and doxapram), and five had a positive result for the calcein-AM competition assay (midazolam, flurazepam, noscapine, bromocriptine, levomeprazine). One of the two false negatives, methysergide, had a negative result for the calcein-AM competition assay, suggesting that the transport assay was a false positive observation.

2.3.4 Prediction of potential substrates from the National Cancer Institute

The Developmental Therapeutics Program (DTP) of the National Cancer Institute (NCI/NIH) has screened tens of thousands of compounds against a panel of 60 human tumor cell lines representing nine tissue sites. The compounds are generally tested at different concentrations (usually at five 10-fold dilutions, up to 100 μM) for the ability to inhibit the growth of these cell

lines. The dose-response data is used to calculate several concentration parameters, in particular the GI_{50}, which corresponds to the concentration of the test drug causing 50% growth inhibition. Furthermore thousands of molecular targets including protein or gene expression levels have been measured in the NCI panel of human tumor cell lines. We downloaded the expression profile of P-glycoprotein obtained with different techniques (RT-PCR, Affymetrix microarrays) and normalized the data: the average expression level of all cell lines was subtracted from the expression level of each individual cell line. Figure 2.7(a) depicts the P-glycoprotein gene expression level obtained by RT-PCR. The analysis of the expression level obtained with other techniques (PCR, Affymetrix microarrays) all lead to comparable data. It appears that only five cell lines have a high P-glycoprotein expression level: HCT15 and SW-620 (colon), CAKI-1 and UO-31 (renal), and NCI/ADR-RESD (breast). In a second step, we compared the activity profile of 44'652 compounds taken from the DTP, and selected the 6'547 drugs having a global growth inhibitory power on the different cancer cell lines. This allowed the reduction of the initial dataset of 85.5% to a subset of only active compounds. In a third step, the activity profile of the 6'547 selected drugs, each drug being previously tested against each cancer cell line of the panel, was compared to the P-glycoprotein expression level on the same cell lines. As expected, some drugs showed a *reduced* activity on the five cell lines over-expressing P-glycoprotein (Figure 2.7(b), whereas other drugs showed no particular reduction of activity on these five cell lines (Figure 2.7(c)). We selected the first 37 compounds showing the highest decrease of efficiency on the cell lines over-expressing P-glycoprotein (highest correlation between P-glycoprotein expression profile and activity profile), called *putative substrates*, and the last 43 drugs showing no decrease of efficiency on these cell lines (lowest correlation between P-glycoprotein expression profile and activity profile), called *putative non-substrates.*

The cross-sectional areas of the 80 compounds were calculated as described elsewhere,[10] and the total number of type I and type II units was obtained as described above. Octanol-water distribution coefficients at pH 7.4 (LogP and pK_a) were obtained with Marvin Sketch. As described above, to be an *apparent* substrate, a drug has (i) to have at least one type I or type II unit to interact with P-glycoprotein, (ii) be hydrophobic enough to partition into the lipid bilayer, (iii) to be cationic or neutral, and (iv) be large enough to diffuse slowly through the membrane, and be thus efficiently effluxed by the multidrug transporter. Hence we plotted in a three-dimensional diagram the distribution coefficient (LogD), the number of type I and type UU units (EU_H), as well as the calculated cross-sectional area (A_D). We expected to obtain a discrimination between putative substrates (at least one type I or type II unit, hydrophobic and large molecules), and putative non-substrates (eventually small, no type I or type II units, and negatively charged, thus low $LogD_{7.4}$). Figure 2.8 shows the three-dimensional plot of LogD, EU_H, and A_D; it demon-

(a)

(b)

(c)

Figure 2.7: (a) normalized P-glycoprotein expression level on 50 cell lines investigated in the Developmental Therapeutics Program; (b) activity profile of NSC 80469, a putative substrate of P-glycoprotein; (c) activity profile of NSC 600223, a putative non-substrate of P-glycoprotein.

Table 2.8: 37 putative substrates of P-glycoprotein.

Compound	LogD	pK$_a$[†]	EU$_H$	A$_D$ [Å^2]	Prediction
NSC 125973	4.2	-2.44/12.53	12	135.08	substrate
NSC 139105	4.42	8.21 / –	1.99	53.5	non substrate
NSC 289922	5.99	– / –	1	73.07	non substrate
NSC 328426	2.8	– /12.66	15.83	97.83	substrate
NSC 337766	3.74	6.49 / –	0	65.52	non substrate
NSC 342443	2.67	– / 12.6	11.33	91.72	substrate
NSC 352670	0.73	– / –	3.39	77.28	non substrate
NSC 368891	4.35	6.66 / –	0.64	53.07	non substrate
NSC 600223	-2.75	-2.52/ 3.99	13.73	129.91	non substrate
NSC 646946	1.09	– / –	1	143.74	substrate
NSC 656177	4.07	-2.44/11.88	12.5	143.67	substrate
NSC 661748	10.92	– / 8.48	1.33	169.26	substrate
NSC 664401	3.95	-2.44/11.93	11.75	125.63	substrate
NSC 664403	4.26	-3.7 /12.07	12.5	133.36	substrate
NSC 664404	3.82	-3.69/12.07	12.67	140.37	substrate
NSC 666606	2.28	-4.43/11.07	13.08	124.62	substrate
NSC 667645	2.74	7.72 /17.93	0	61.63	non substrate
NSC 668404	3.38	-3.37/11.75	14.5	156.1	substrate
NSC 671864	3.23	-4.34/ 11.9	13	139.84	substrate
NSC 671865	4.82	-4.34/ 11.9	14.75	145.33	substrate
NSC 671866	4.46	4.69 /12.53	12.5	145.34	substrate
NSC 671867	3.47	-4.9 / 9.62	12.83	168.45	substrate
NSC 671871	2.27	-2.81/11.91	12.75	118.82	substrate
NSC 671872	4.15	-2.44/11.92	11.12	115.07	substrate
NSC 673185	3.15	-4.41/10.75	13.5	170.5	substrate
NSC 673186	3.38	-2.44/11.92	14	133.82	substrate
NSC 673187	3.15	-2.44/11.92	12.5	121.87	substrate
NSC 673188	4.34	-2.44/11.92	14	132.94	substrate
NSC 673191	5.08	-2.44/11.92	13	142.7	substrate
NSC 673192	3.41	3.24 /11.95	12.75	115.5	substrate
NSC 673193	4.46	3.48 /11.95	10.5	138.03	substrate
NSC 675252	2.42	8.49 /15.67	0.53	85.22	non substrate
NSC 675256	4.46	8.21 /11.34	1	38.13	non substrate
NSC 701744	2.38	3.5 /10.88	4.5	150.11	substrate
NSC 80465	-2.11	– / –	1	60.33	non substrate
NSC 80467	-1.65	– / –	2.37	58.11	non substrate
NSC 80469	-0.31	– / –	1.23	44.83	non substrate

[†] The first value corresponds to the basic pK$_a$, the second value to the acidic pK$_a$.

Table 2.9: 43 putative non-substrates of P-glycoprotein.

Compound	LogD	pK$_a$†	EU$_H$	A$_D$ [Å²]	Prediction
NSC 45383	-4.16	4.03 / 0.95	6.5	81.65	non substrate
NSC 52947	-0.83	9.17 / 9.94	4.5	72.76	non substrate
NSC 94739	2.44	7.65 / 9.27	1.5	57.57	non substrate
NSC 123976	2.88	- /12.75	5.25	105.02	non substrate
NSC 126728	1.84	- /12.69	6	88.3	non substrate
NSC 138780	0.85	- /13.42	7	69.23	non substrate
NSC 163062	0.27	- /12.94	4	48.01	non substrate
NSC 177499	1.7	-0.13/ 9.48	3.5	46.66	non substrate
NSC 200737	2.02	- /12.68	7.5	83.93	non substrate
NSC 251698	3.5	- /13.49	4.75	111.34	non substrate
NSC 295505	-0.57	8.63 /12.75	0	54.51	non substrate
NSC 326408	2.39	0.3 /11.91	4.75	97.95	non substrate
NSC 352890	-1.15	4.54 /12.83	1	48.56	non substrate
NSC 407808	0.28	- /12.65	7	110.26	non substrate
NSC 603724	-0.09	- /11.59	8	82.85	non substrate
NSC 626168	1.36	-5.62/ 9.74	6.5	99.73	non substrate
NSC 626169	1.36	-5.62/ 9.74	7.25	106.26	non substrate
NSC 626170	1.64	-5.62/ 9.74	6	114.29	non substrate
NSC 626369	0.85	- /12.87	1.5	58.1	non substrate
NSC 645804	-0.94	8.98 /12.75	0.25	54.2	non substrate
NSC 648772	3.11	0.58 /12.15	3.5	84.18	non substrate
NSC 648910	1.42	- /10.07	6	55.08	non substrate
NSC 650393	3.2	8.14 /12.83	2.25	78.65	non substrate
NSC 650394	3.36	6.35 /12.83	3.5	78.66	non substrate
NSC 650395	2.65	7.74 / 9.6	4.25	72.03	non substrate
NSC 650396	3.51	6.56 / 9.5	4.5	75.49	non substrate
NSC 668421	5.41	5.58 /13.32	1.25	79.71	non substrate
NSC 673805	0.25	9.18 /12.74	0.5	45.87	non substrate
NSC 676836	2.55	- /11.15	7.5	74.38	non substrate
NSC 677588	2.12	-0.69/ 9.17	4	50.9	non substrate
NSC 681237	2.3	8.68 /12.59	0	61.28	non substrate
NSC 681239	1.53	-0.7 /11.25	2	59.01	non substrate
NSC 683094	4.27	-2.78/12.11	0.5	68.49	non substrate
NSC 683098	2.98	1.1 /12.13	1	61.1	non substrate
NSC 687939	2.73	- / 11.9	5.5	92.55	non substrate
NSC 690267	1.34	-3.2 /12.22	3.75	86.96	non substrate
NSC 693541	3.13	-1.4 /12.19	1.5	83.35	non substrate
NSC 693542	2.6	3.7 /12.19	1.5	53.59	non substrate
NSC 696128	2.63	3.19 /11.45	3.75	58.57	non substrate
NSC 698103	3.11	-0.78/ 8.76	4	58.84	non substrate
NSC 698104	4.11	-1.36/22.65	1.5	67.48	non substrate
NSC 699482	1.8	8.42 /12.17	3	49.02	non substrate
NSC 700011	1.78	8.45 /11.35	2.25	61.86	non substrate

† The first value corresponds to the basic pK$_a$, the second value to the acidic pK$_a$.
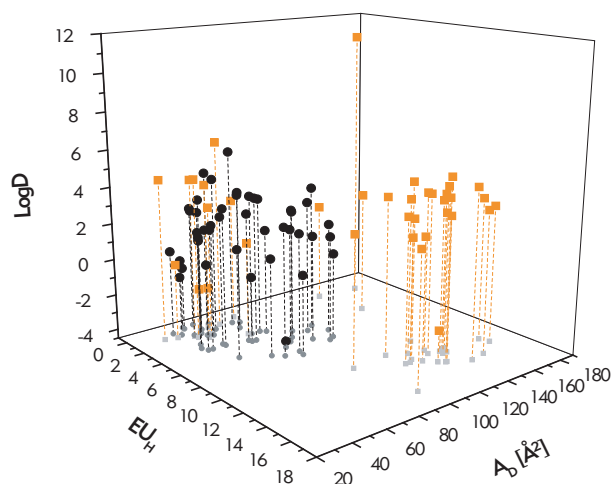
Figure 2.8: Three-dimensional diagram for the discrimination between putative substrates (■) and putative non-substrates (●).

strates the difference in the activity profile of the two populations of drugs (the one having a decreased activity on the cell lines over-expressing P-glycoprotein, and the second one showing no particular decrease in the activity), can be seen on the 3D diagram. Indeed, as expected, putative substrates have at least one type I or type II unit, and are large, whereas the putative non-substrates are localized in the portion of the diagram with low cross-sectional area and low number of type I or type II units. They also show slightly lower LogD; some of these drugs are highly negatively charged. However 10 putative substrates are in the same range as the putative non-substrates; we suppose that these compounds might be effluxed by other multidrug transporters, like the MRP1. For instance NSC 600223 contains a carboxy group, and thus cannot be transported by the P-glycoprotein.

2.4   Conclusion

Several methods have been developed to predict substrates and non-substrates of the multidrug transporter P-glycoprotein. The binding of a drug to the transporter occurs (i) with a partition from the aqueous phase to the lipid membrane, and (ii) with the binding from the lipid membrane to the transporter, we developed specific prediction rules to take into account this two-step binding mechanism. The two steps have to be predicted individually, this is possible if lipid-water, K_{lw}, or the air-water partition coefficient, K_{aw}, is known.[4] The lipid-water partition coefficient, K_{lw}, of a specific membrane can either be measured or it can be derived from the air-water partition coefficient, K_{aw}, according to Eq. (2.1), Taking into account the cross-sectional area, A_D, of

the molecule and the packing density, Π_m, of the membrane.[4] The binding constant from the lipid membrane to the transporter, K_{tl}, can be estimated by counting the number of hydrogen bond acceptor patterns (see Eq. (2.2)). Since each *in vitro* assay monitors a specific phenomenon (where the difference between *intrinsic* and *apparent* substrate can be assessed), one prediction rule has been established to predict the result of each assay.

The prediction rules have been tested on several datasets, leading to high prediction accuracies. Moreover, the understanding of the binding mechanism allowed to resolve the apparent inconsistencies observed by several authors between different *in vitro* assays, and can be used to develop new prediction rules to assess the results of new experimental settings.

Chapter 3

Computational tool for the predictions

3.1 Introduction

The effect of P-glycoprotein on drug's efflux out of targeted cells has gained more and more importance, especially for the development of a new drug. Several *in silico* models have been established with the aim to predict if a compound will be carried out by P-glycoprotein.

Since we developed several models to assess substrates of P-glycoprotein (one model for each *in vitro* assay), and since these models are based on new descriptors, we had to develop some algorithms to handle them; we decided to this purpose to establish a whole working interface to handle compounds, multiple conformations, some usual descriptors and our original ones, as well as our rule-based models for the prediction of the passive diffusion and the active efflux of drugs by multidrug transporters, in particular P-glycoprotein (Figure 3.1).

3.2 Presentation of the software

3.2.1 Aim and roots of the software

The software is designed to handle within a graphical interface, several files, compounds, multiple conformations, and to be able to calculate for each compound in a first step all parameters required for the prediction of the drug's passive diffusion and active efflux, and in a second step to calculate the prediction itself.

The software has been programmed using *Microsoft Visual Studio .NET 2002*; thus it works under Microsoft Windows. All the libraries implemented into the software have been written during the PhD, except:
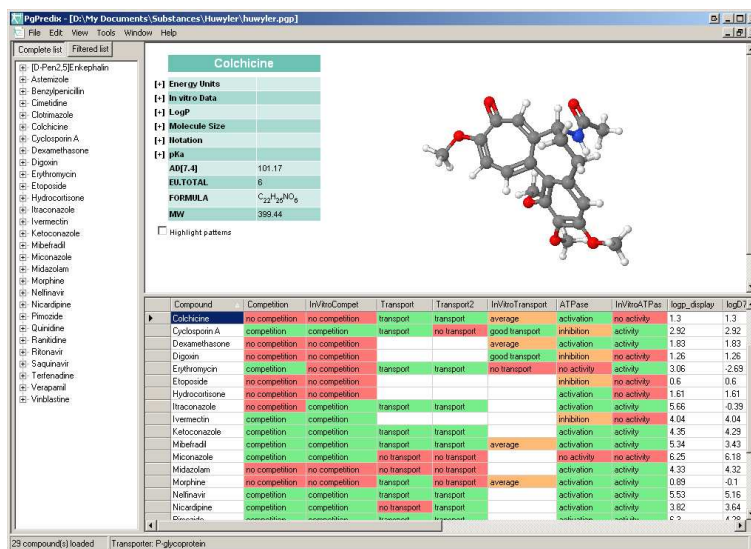
Figure 3.1: Main user interface of the software. The left panel is the *compound panel*, the top panel is the *visualization panel*, and the bottom panel is the *results panel*.

- Jmol applet v10.2 (jmol.sourceforge.net), used for the three-dimensional visualization of the molecule in the *visualization panel* of the software.

- OpenBabel library (openbabel.sourceforge.net), used to transform MOL files into SMILES strings.

- ModEncryption library (www.freevbcode.com/ShowCode.asp?ID=2919), used to encrypt/decrypt informations and prediction rules that are not intended to be readable by the user.

Moreover the software uses some external plugins to calculate the LogP using different algorithms, however the XlogP algorithm has been also implemented within the software.[70]

3.2.2   Input

For an optimal compatibility with other softwares and to avoid unnecessary conversion steps, we decided to use for the input of structures the Structure Data Format (SDF); this format is the most widely used public standard to handle structural and data information on chemicals. SDF files are simple ASCII text files containing any number of records, each record consisting in two blocks, one structural block (which contains a list of atoms and a connection table for atoms binding) and one information block (consisting of zero or more data fields). The SDF format was originally developed and published by Molecular Design Limited (MDL).

The structural block can exist on its own in individual files, having the *.MOL* extension. Each MOL file contains only one structure and cannot contain any additional information (data fields). The software can handle MOL files as well, and these missing data fields can be added afterward from the software itself. A parser designed to read SDF files as well as MOL files has been programmed and implemented into the software.

3.2.3 Project

The software handles SDF files, MOL files, data fields, as well as calculations and prediction results in *project files*. These files carry the *.PGP* extension. A project file keeps track to zero or more input files (SDF or MOL), contains the description of the transporters, the variables, the prediction rules, it also contains all results of the calculations that have been carried out for each compound, and some additional display information (for instance the size of the display window, or the way data fields are displayed and regrouped).

Each new project (empty project, or new project based on a SDF file or a MOL file) will be initially fed with the information contained in a *template file*. This file is by default localized in the same folder as the running application, and is called default.pgc (PgPredix Configuration File). Each modification in the transporters definition, the prediction rules, the variables and the way data fields are organized in the visualization panel can be saved in the template file; these settings will be used for each new project.

The configuration of an opened project can be saved in a *.PGC* file; this can be useful for the portability of the software, or to save the current configuration (as a backup) before making some modifications on prediction rules or transporters definition. Inversely, the configuration of an opened project can be overwritten by importing the configuration of another PGC file. The whole configuration or only parts of the configuration (for instance only parameters of the variables, or the one of the transporters) might be imported.

Since the PGP project file does not embed the structures, but only maintains a link to the various SDF or MOL files that have been added to the project, these structure files should always be kept together with the PGP project file. However at any time SDF files or MOL files can be added to the opened project, and structure files already attached to a project might be removed from the project. This will automatically remove from the compound list all compounds belonging to the structure file (SDF or MOL file) that has been removed from the project. Moreover single records (compounds) might be excluded from the project; this might be useful to perform some calculations on only a subset of conformations, for instance. The excluded compounds might be recovered at any time.

3.2.4   Output

Since the software represents a complete working interface and has the possibility to handle and display structures, perform simple mathematical operations, or sort the results, one can use it without any other additional softwares. The PGP project file will keep all the information needed for proper operation. However one might want to use the generated results in other softwares.

- The results can be exported in a *tab-delimited file*, which has the *.XLS* extension. This file will be directly compatible with Excel, or any other spreadsheet analysis software.

- The results can also be exported directly in the source SDF files: the information block of the input files will be updated with the new data fields corresponding to the prediction rules name or variable name and the related information. All input files can also be merged into a single SDF file; this can be useful to have to handle only two files at a time (one SDF file and one PGP project file), instead of a whole list of, for instance, MOL files coupled to one project file. If this merging process is executed, the newly generated SDF file will be included into the current PGP project file, and all other structure files (SDF of MOL files) will be automatically removed from the project.

- Each individual structure can also be exported as MOL file, thus containing only the structure information (atoms and connections between atoms) and no additional information. The structures can be exported one at a time or all of them. Moreover, the amphiphilicity axis, meaning the axis representing the way the molecule will be oriented at an hydrophilic-hydrophobic gradient (Figure 3.2), can be embedded in the MOL file. Since the MOL files can only handle "real" atoms and no additional information, the amphiphilicity had to be embedded as two "real" atoms connected together. We chose the represent the hydrophilic center of mass with a chloride atom (Cl), and the hydrophobic center of mass with a magnesium atom (Mg). The reason of choosing these two atoms is that (i) they have nearly identical van der Waals radii (1.75 Å for Cl, 1.73 Å for Mg), which is useful for representations as seen on Figure 2.6, and (ii) the Cl-Mg covalent bond will never occur in reality, this allows to unequivocally find the amphiphilicity axis embedded in the MOL file.

3.2.5   Data fields

As described previously, each record in a SDF file contains one structural block, and one information block. The information might contain one or more *data fields*, each data field consisting in one *field name* and one *value*. When a
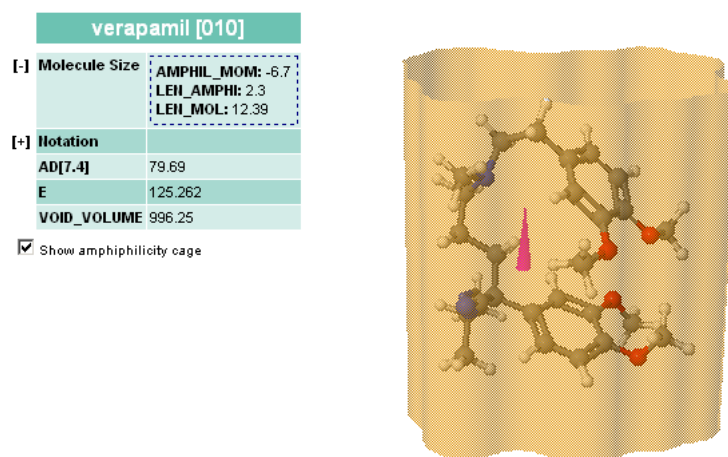
Figure 3.2: Visualization panel showing a folded conformation of verapamil, with the amphiphilicity axis (purple) and the cross-sectional area of the molecule (orange cage)

SDF file is included in a project, the software will read for each record the information block and collect all field names and their attached values. These information will be displayed in the software, in the tree panel as well as in the display panel.
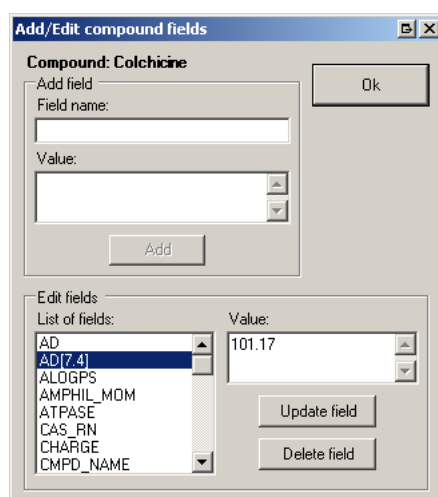

Figure 3.3: User interface to manage the data fields of the compounds.

The data fields of each compound can be edited within the software (Figure 3.3). The user can add a new data field (consisting of one field name and one value), edit or delete an existing field. The user can also add a new data field to all compounds loaded in a project, and edit or delete an existing data

field only if all records (compounds) contain the same value for the selected field.

One should emphasize the fact that these modifications on data fields (new, modified or deleted fields) will *not* be saved in the project, since these informations are not maintained in the PGP project file, unless the user exports these informations in the source SDF file (as described in the section 3.2.4).

In the *display panel*, the data fields can be grouped into *groups*. Each group will have a name and a list of attached data fields, which are called *items* (Figure 3.4). The groups will be displayed like folders and items like subfolders. All items that are attached to a group will not appear at the root level on the display panel. However the user might create a group named *root*, which will contain items that are present in a group but will also appear at the root level.
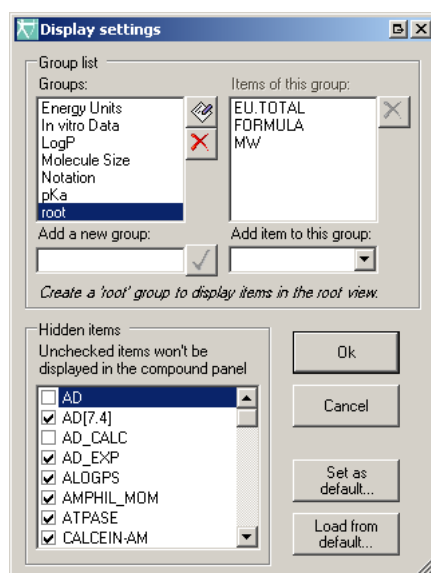
Figure 3.4: User interface to manage the display groups.

All data fields are displayed in the tree panel. However the user might want to have only some useful information displayed in the display panel. Thus one can hide some items; these items will still be available for calculation but they simply will not appear on the display panel. All items that are not included in any group and that are not hidden will be shown at the root level on the display panel.

3.2.6   Variables

The previously described data fields are not available directly in the calculations (prediction rules). One have to link them to a *variable*. Each variable might

be referring to one or several data fields, and the data fields will be classified by order of priority. The aim of variables is the following: one might have several data fields describing the same parameter, several of these fields might be present for one single record (compound), or several records in one project might contain data fields having different field names but describing the same parameter. Instead of having to define complicated prediction rules handling these different field names, or having to define several prediction rules to handle each field name, the user will simply have to use the *variable* in the prediction rule, and the software will search for the availability of data within the listed data fields, following the priority order (Figure 3.5).
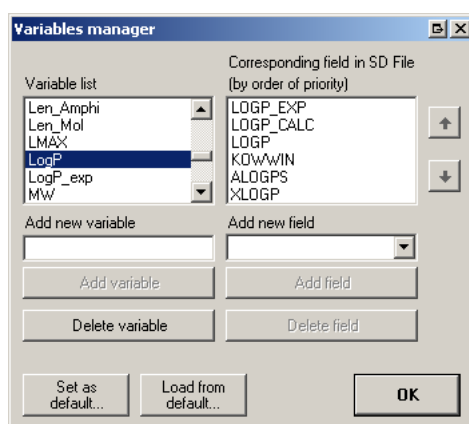

Figure 3.5: User interface to manage the variables and their dependencies.

Let us consider a simple example. The octanol-water partition coefficient, LogP, might be obtained from several sources, experimental or computational ones. Thus one might define a variable named *LogP*, which will refer to several data fields, by order of priority, `LOGP_EXP` (for experimental value), `CLOGP`, `KOWWIN`, `PUBCHEM_CACTVS_XLOGP` (which is the field for LogP from National Institutes of Health's PubChem substances). Thus when calling the variable *LogP*, the software will first search in the record data fields for `LOGP_EXP`, then if no experimental data is available it will search for the `CLOGP` data field and so on. If no listed data field is available, the variable will return an error (and as a consequence, the prediction rule referring to this variable will also return an error).

An important variable is *Cmpd_Name*. When reading a new SDF file, the software will search in the information block of each record (compound) for the field names listed under *Cmpd_Name*. If one of these data fields is found, the corresponding value will be used as compound name, which is the unique identifier of each record. If no compound name is found, the software will attribute an arbitrary unique identifier, which will be *Compound [#]*, where # is an incrementing number. If a compound name is found but this name

was already attributed to another record as unique identifier, this record will have as unique identifier the compound name followed by *[#]*, where # is an incrementing number.

One might define as data fields for the variable *Cmpd_ Name* the following names: `CAS_RN` (American Chemical Society identifying number), `NSC` (Cancer Chemotherapy National Service Center identifying number), `PUBCHEM_COMPOUND_CID` (PubChem identifying number).

3.2.7   Constants

The *variables* described previously are designed to handle the data fields of each individual compounds. However one might want to have global variables which can be modified, but which have the same value for *all* compounds included in a project. This particular kind of variables are called *constants*.

Five constants are embedded in the software, the Avogadro's number (Na), the Boltzmann constant (kB), the Pi value (Pi), and two constants that can be modified in the options window within the software: the temperature (configuration.Temperature) and the pH (configuration.pH). Other constants can be created/updated/deleted and saved in the default template file. The constants, like variables, are available in the prediction rules.

3.2.8   Calculations

Several algorithms and plugins have been implemented in the software. They allow the calculation of several parameters that might be used afterward in prediction rules. The following section will list each calculation item, and the corresponding result(s). Indeed one calculation item might generate one or several data fields. Moreover, each generated data field will also be automatically added to a *variable* having the same name as the field name; this allows the user to directly apply the software's generated data in prediction rules.

MW+Formula will calculate, for each record (compound), its exact molecular weight (by eventually adding missing hydrogens), and the empirical formula. The generated fields will be `MW` for the molecular weight and `FORMULA` for the empirical formula.

SMILES will "calculate" the SMILES (Simplified Molecular Input Line Entry System) compound formula, using the openBabel plugin. The generated field will be `SMILES`.

Maximal length will calculate the distance between each pair of atoms of a structure, an the output will be the largest distance found. The generated field will be `LMAX`.

pKa is used to search for ionizable atom(s), and attribute them pK_a values. It will also search for carboxy groups, as well as quaternary ammonium ions. The compound charge is also returned (based on source file information and not on the pH and pK_a). The user might feed the software with pK_a values, and these values must have the following syntax: the field names `PKA1_TYPE` and `PKA2_TYPE` must have the values *acidic* or *basic*, and the field names `PKA1_VALUE` and `PKA2_VALUE` must have as values numerical entries. The software will search for basic and acidic ionizable atoms, and attribute them the fed pK_a values. These fed values might be present in the information block of records in the source SDF files, or added afterward as described previously.

The generated fields will be `CHARGE` for the global charge of the molecule, `COOH` for the number of carboxy groups, `PKA1.VALUE` for the value of the first pK_a, `PKA1.TYPE` for the type of the first pK_a (either acidic or basic), `PKA2.VALUE` for the value of the second pK_a, `PKA2.TYPE` for the type of the second pK_a (either acidic or basic).

The pKa module has been developed based on the method described by Sayle.[71]

LogP with XLogP will calculate the LogP as described by Wang et al.[70] The XLogP algorithm has been implemented in the software. The generated field will be `XLOGP`.

LogP with KowWin will calculate the LogP using the Syracuse Research Corporation (SRC) online KowWin plugin. The generated fields will be `KOWWIN` for the calculated LogP, and `LOGP_EXP` for the experimental LogP, if the value was available in the SRC online database.

LogP with ALogPS will calculate the LogP using the VCC Lab ALogPS plugin. The generated field will be `ALOGPS`.

Polar Surface Area will calculate the topological polar surface area, according to Ertl et al.[72] The described algorithm has been implemented in the software. The generated field will be `PSA`.

EU(H) will calculate the hydrogen-bonding ability of the molecule to the selected transporter (see below). The generated fields will be `EU.TOTAL` for the total number of hydrogen-bond formation, `EU.TYPE0` for the number of single atoms able to form hydrogen-bonds, `EU.TYPE1` for the number of type I units, and `EU.TYPE2` for the number of type II units.

Cross-sectional area will calculate the cross-sectional area of the molecule, oriented at an hydrophilic-hydrophobic gradient. This algorithm will first run the *pKa* calculation since the charges and their localization have to be known. The generated fields will be `AMPHIL_MOM` which has been a

tryout for the prediction of the amphiphilic moment of the molecule, `AD` or `AD[#]` which is the calculated cross-sectional area, in Å^2, where # in the current defined pH, `LEN_AMPHI` is the distance between the hydrophilic and the hydrophobic centers of mass, in Å, `LEN_MOL` is the total length of the molecule oriented at the amphiphilic interface, in Å, and `VOID_VOLUME` is the void volume generated by a molecule inserted into a lipid membrane, in Å^3.

The algorithm developed and implemented in the software has been described elsewhere.[10]

3.2.9 Transporters

We presented in section 3.2.8 the possibility to calculate the hydrogen-bonding ability of the molecules to a transporter. By default, the software comes with one already configured transporter, P-glycoprotein. However one might want to define a new transporter; the following section will describe the available settings, this should also help in understanding how the software assesses the interaction of a molecule to the transporter, and the meaning of the previously cited fields `EU.TOTAL`, `EU.TYPE0`, `EU.TYPE1` and `EU.TYPE2`.
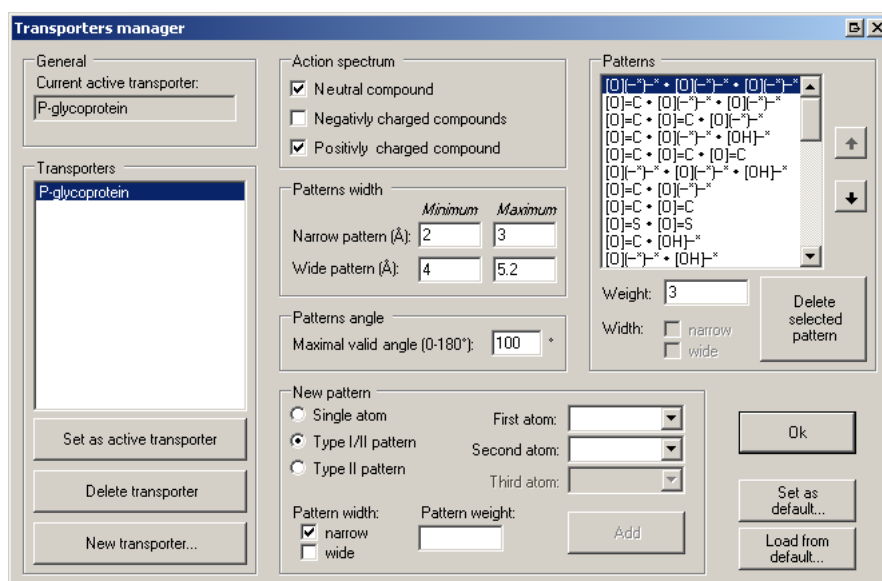

Figure 3.6: User interface for the configuration of the transporters.

One single PGP project file can handle several transporters; the *template file* as well. However for the calculations of *EU(H)*, one can calculate the interactions of the molecules to one transporter at a time since the user has to choose the current "active transporter".

It has been suggested that a substrate of P-glycoprotein will interact to the transporter via hydrogen bonds formation; these bonds will form only if the molecule shares some specific hydrogen bond acceptor patterns, which can be of two types:

Type I units are two hydrogen bond acceptors, spatially distant of 2.5 ± 0.5 Å, and

Type II units are either two hydrogen bond acceptor groups, spatially distant of 4.6 ± 0.6 Å, or three hydrogen bon acceptors groups with a distance of 4.6 ± 0.6 Å between the first and the third group, with an intermediate hydrogen bond acceptor group.

These hydrogen bond acceptor groups must be electronegative atoms (O, N, S, F, Cl) with an unshared electron pair, or unsaturated systems with a π-electron orbital (like a phenyl ring). Oxygens involved in a specific pattern might be a carbonyl group ($>C{=}O$), an alkoxy group ($-OR$) or one (but not two) hydroxy group ($-OH$). Nitrogens involved in a specific pattern might be an imine ($-N{=}$), a pyrrole (N1CCCC1), a pyridine (n1ccccc1), or a tertiary amine ($-NR_2$).

It has also been shown that the free orbitals of the two or three hydrogen bond acceptors must be oriented in the same direction, within a range of 0 to 123°.

A new transporter can be parametered as follows (Figure 3.6):

Transporter name is the name attributed to the transporter.

Action spectrum can be any of these three parameters, *neutral compounds*, *negatively charged compounds* and *positively charged compounds*. This setting will be available in the prediction rules (see further).

Patterns width represents the width of narrow (type I) patterns and wide (type II) patterns. For P-glycoprotein, the type I unit has a width of 2.5 ± 0.5 Å, whereas the type II unit has a width of 4.6 ± 0.6 Å. When searching for interaction patterns between the molecule and the transporter, the software will exclude all groups of two or three hydrogen bond acceptor groups that are outside of these ranges.

Patterns angle corresponds to the maximal angle between the free orbitals of the two hydrogen bond acceptors of one pattern. The acceptable range is 0 to 180°; the latter value corresponds to antiparallel free orbitals.

Type I pattern: as described previously, type I units are formed by two hydrogen bond acceptor groups, separated by a specific spatial distance and a specific angle between free orbitals. The user might select two hydrogen bond acceptor groups to define a new type I unit; moreover,

each pattern will have a specific *weight*, corresponding to the contribution to the total free energy of binding of the drug to the transporter (ΔG_{tI}). This weight can be any numerical value. See below for the list of available hydrogen bond acceptor groups. The sum of these weights of all type I units carried by a molecule gives the EU.TYPE1 field in the *EU(H)* calculation.

Type II pattern: we also described that the type II units are formed by two or three hydrogen bond acceptor groups. The user has to specify if the type II unit is formed by two or three groups, and select the corresponding groups in the list of available hydrogen bond acceptor groups (see below for the complete list). As for type I units, type II units have a specific weight for the ΔG_{tI}. The sum of these weights of all type II units carried by a molecule gives the EU.TYPE2 field in the *EU(H)* calculation.

Single atom (aka Type 0 pattern): the software has the possibility to count single atoms able to form hydrogen bonds with the transporter. For single atoms, the user has to choose one hydrogen bond acceptor group (see below for the complete list), and specify one weight for the free energy of binding to the transporter (ΔG_{tI}). The sum of the weights of all single atoms forming hydrogen bonds of one molecule gives the EU.TYPE0 field in the *EU(H)* calculation.

The last calculation output of the *EU(H)* section is the field EU.TOTAL. This field is the sum of all free energy contributions of *Type I Units*, *Type II Units* and *Single atoms*.

The available hydrogen bond acceptor groups are the following (the $\star$ symbol represents any non-hydrogen atom, and the considered atom is depicted in bold):

- **O**$=$C

- **O**$=$S

- **O**$=$N

- **O**$=$P

- $\star$—**O**—$\star$

- $\star$—**O**H

- N—**O**H

- P—**O**H

- $\star$—**N**$<^{\star}_{\star}$

- $\star\!\!=\!\!N$

- $\star\!\!-\!\!N\!\!=\!\!O$

- $\star\!\!-\!\!N\!<^O_O$

- $\star\!\!-\!\!NH\!\!-\!\!\star$

- $\star\!\!=\!\!NH$

- $\star\!\!-\!\!NH_2$

- $\star\!\!-\!\!S\!\!-\!\!\star$

- $\star\!\!-\!\!SH$

- Aromatic cycle

- $\star\!\!-\!\!C\!\!=\!\!C\!\!-\!\!\star$

- $\star\!\!-\!\!F$

- $O\!\!=\!\!C\!\!-\!\!OH$

- $HO\!\!-\!\!C\!\!=\!\!O$

3.2.10 Prediction rules

The *prediction rules* represent the core of the software. This is were the user can really interact with the software. A prediction rule is a "command" that will collect information in *variables* and *constants*, do simple mathematical operations, compare variables or different results together, and finally either display a result (for instance the result of a comparison), or export a structure to a specific folder. The results of the prediction rules will be displayed in the *results panel* of the user interface. The prediction rules are evaluated using a parser implemented in the software, which will convert the rules into *Reverse Polish Notation*, to allow a faster processing of the rules on large datasets. The following section will describe the *objects* the user can work with, then the *values* and the *operators*, and finally the *statements* and the *methods* (Figure 3.7).

Objects

The *objects* of the prediction rules are the elements the user can work with. They belong to two groups, the *variables* and *constants*, and the *rules*. The variables and constants have been described in section 3.2.6 and section 3.2.7, they give access to any information stored in the input SDF files, or information fed afterward from the software itself. As stated previously, several *field*
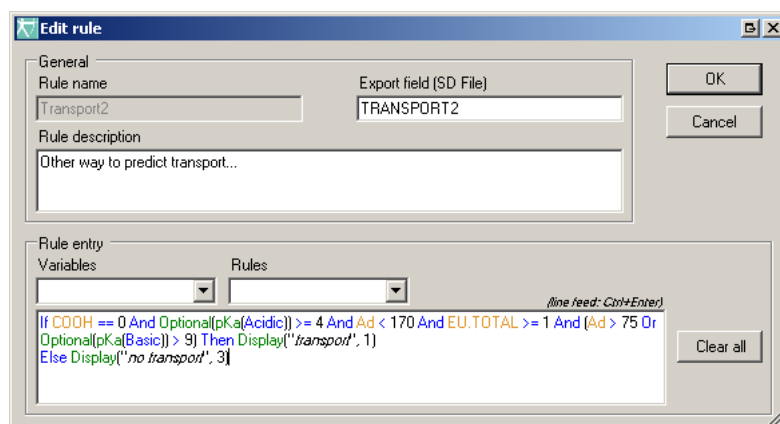
Figure 3.7: User interface for the edition of the prediction rules.

names can be stored within one single variable, for an easier handling of several *data fields* describing the same parameter.

The rules are the results generated by other (and previously evaluated) prediction rules. The aim is to avoid long and complicated rules, but the user has to be warned that a prediction rule referring to the result of another prediction rule will generate an error if the latter has not been evaluated yet. Thus the prediction rules should be *ordered* to avoid any unexpected halt in the evaluation process.

Values and Operators

Prediction rules can deal with three types of values: *numerical values*, *logical values*, and *character strings*.

Numerical values can be any number, integer or not; scientific notations constitute valid numerical values as well.

Logical values should be recognized in prediction rules, however in case of problem the user should try to use string-embedded logical values (that means, instead of using `True` or `False`, the user should use `"True"` or `"False"`).

Character strings are text enclosed in double quotation marks.

The operators belong to three groups: *arithmetic operators*, *logical operators*, and *comparison operators*.

Arithmetic operators take numerical values (either literals or variables) as their operands and return a single numerical value. The standard arithmetic operators are addition ($+$), substraction ($-$), multiplication ($*$), and division ($/$). These operators work in the standard way. One should

note that the usual arithmetic rules apply, hence multiplication and division will have priority over addition and substraction; one can use parenthesis to change the priority order. One fifth arithmetic operator is the string concatenation operator (&). It is used to join two character strings.

Logical operators: three logical operators are available: the bitwise `And` operator returns `True` if both operands are identicals; the bitwise `Or` operator returns `True` if at least one of the two operands is True. The bitwise `Xor` operator returns `True` if one but not both operands are True.

Comparison operators: six comparison operators are available. The *equal* operator (==) returns `True` if the operands are equal. The *not equal* operator (<>) returns `True` if both operands are different. The *greater than* operator (>) returns `True` if the first operand is greater than the second one, and the *greater than or equal to* operator (>=) returns `True` il the first operand is greater or equal to the second operand. The *smaller than* operator (<) returns `True` if the first operand is smaller than the second one, and the *smaller than or equal to* operator (<=) returns `True` if the first operand is smaller or equal to the second one.

Statements and methods

Only one type of *conditional statements* is available. These statements are:

```
If...Then...
Elseif...Then...
Else...
```

Only the first two statements are required, whereas `Elseif...Then` and `Else` statements are optional. One can use as many `Elseif...Then` statements are needed, but one can use only one `Else` statement.

There are many *methods* available in the software. They belong to two groups: the methods of the first group allow manipulation of variables or rules values, whereas the methods of the second group are designed to *output* the results. The two output methods can be used only alone or after a `Then` or a `Else` statement. The manipulation methods cannot be used after a `Then` or a `Else` statement. The following section will enumerate alphabetically all available methods, starting with the two output methods.

Output methods:

Display(x, y): displays the result (x) without background (y=0) or with a "Good" (y=1), "Medium" (y=2), or "Bad" (y=3) background. The

symbol y can also be a string containing six comma-delimited number corresponding to the RGB value of the text and the RGB value of the background of the cell. Ex: `Display("test", "0, 0, 0, 118, 156, 234")` will display the text test with a black font on a blue background.

Export_Mol(x, y): exports the current molecule into the specified folder (x) by adding (y=1) or not (y=0) the amphiphilicity axis (see section 3.2.4 for the definition of the amphiphilicity axis).

Manipulation methods:

Abs(x): absolute value of x.

Angle(x, y): gives the angle between the orbitals of atom x and atom y.

Average(x): gives the average value of a variable (x) within a whole dataset.[1]

Distance(x, y): gives the distance between atom #x and atom #y (atoms number can be displayed by right-clicking on a structure and selecting Text → Atom number), in ångström.

Exp(x): exponential value of x.

Ln(x): neperian logarithm of x.

Log(x): base 10 Log of x.

LogD(x): log of the distribution coefficient at pH x. This function uses the *LogP* variable.

Export_Mol(x, y): exports a structure in the x folder with (y=1) or without (y=0) amphiphilicity vector embedded in the mol file.

Max(x): gives the maximal value of a variable (x) within a whole dataset.[1] Example of usage:

```
If Ad[7.4] == Max(Ad[7.4]) Then
Display("Largest structure")
```

Min(x): gives the minimal value of a variable (x) within a whole dataset.[1]

Optional(x): it will skip the variable x if the corresponding field(s) is/are absent from the database.

---

[1]These three methods (*Average*, *Max*, and *Min*) do not work on the whole dataset, but on each group of compounds, that means compounds sharing the same compound name (with the incrementing number in brackets, as unique identifier) will be evaluated together.

pKa(x): gives the acidic or basic dissociation constant (pK_a) of the compound, with x="acidic" or x="basic". This function requires the result of the `pKa` calculation.

Power(x, y): raises the number x at the power y.

Required(x): it will skip the whole section (from `If` or `ElseIf` to the next `ElseIf` or `Else` or the end of the rule) if the variable x is absent from the database.

Round(x, y): rounds the number x with y decimals.

SpeciesNeg(x): amount of negatively charged species at a given pH (x). This method requires the result of the `pKa` calculation. The returned value is comprised between 0 (no negatively charged species) and 1 (100% of negatively charged species).

SpeciesPos(x): amount of positively charged species at a given pH (x). This method requires the result of the `pKa` calculation. The returned value is comprised between 0 (no positively charged species) and 1 (100% of positively charged species).

Sqrt(x): square root of x. Same as `Power(x, 0.5)`.

Zwitterion(x): calculates the amount of zwitterionic species at pH x. The output is a number comprised between 0 and 1. Since this function requires two pKas, one can use it with the `Optional( )` argument, skipping the function if only one pKa is provided.

3.2.11 Tuning the calculation of the cross-sectional area

Several parameters of the software can be modified by the user. Some of them concern the interface, some others will influence the calculation of the cross-sectional area (Figure 3.8). Here we will discuss only the settings having a consequence on the calculation results.

Missing hydrogens. As described elsewhere,[10] there is a linear relationship between cross-sectional area calculated by considering the hydrogens, or by removing them from the structure. If no hydrogen are found in the structure, the software will automatically correct the calculated cross-sectional area to take into account the missing hydrogens. However, one can increase the processing speed by removing the hydrogens from the structure for the calculation of the cross-sectional area (in fact the hydrogens are not removed but they are not considered), and applying the correction for the missing atoms. This feature might be useful when calculating the cross-sectional area of many structures.
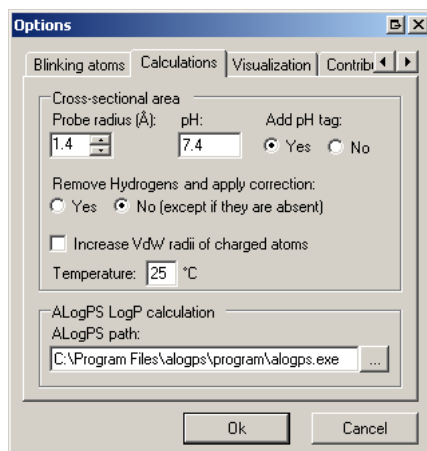
Figure 3.8: Configuration window for the tuning of the cross-sectional area
        calculation

Water probe.   The surface of the projection of all atoms into the plan is
smoothed by a water probe which has by default a radius of 1.4 Å. The user
might want to change the size of the water probe; a radius of 0 Å will give the
exact surface of the contour of the projection, whereas a larger radius might
be more relevant to simulate *in vitro* data.

pH.   Since the charge has an strong influence on the orientation of the
molecule at a hydrophilic-hydrophobic interface, the software calculates the
repartition of charges within the molecule and attributes the weights for the
calculation of the hydrophilic center of mass according to these charges. Thus
the pH will dramatically influence the charge repartition, and the orientation
of the molecule. Hence the user can change the pH at which the charges
are calculated; this will affect the orientation of the molecule, and thus the
calculated cross-sectional area.

Hydrophilic contributions.   To take into account the effect of the charge on
the orientation of the molecules, an arbitrary weight of 100 has been attributed
to a fully charged tertiary amine or a carboxy group. However the user might
want to change these weights, for instance to give a higher or a lower influence
of the charge on the orientation of the molecule, and thus on the calculated
cross-sectional area.

3.2.12   Scripting language

It is often pretty useful to have a three-dimensional representation of the
structures the user is working with, and to see directly all the information and

the results on the user interface. However when working on large databases or on many files, it might be useful to use the scripting language of the software, since the graphical interface is *much* slower than pure processing. Hence the software embed a scripting language that gives the user access to all functions he might use in the software.

Description of the scripting language

The *script files* must have the *.PGS* extension (which stands for PgPredix Scripting File). Scripts are plain ascii texts, and each command must be on a new line. Empty lines are allowed, and comment lines must start with the // tag. Scripting commands are not case sensitives, but *variables* and *rules* are case sensitive. Character strings can be in simple or double quotes, but these are not required. The following section will describe each scripting command alphabetically.

Scripting commands

add_field(string compound; string field_name; string field_value) Adds a new data field to the specified compound, or to all compounds (using the wildcard symbol, ∗) (data fields were described in section 3.2.5).

add_mol(string file_name) Adds a substance (MOL file) or a group of substances (SDF file) to an opened project.

add_rule(string rule_name; string rule_definition; {string rule_description}) Adds a new *prediction rule* or edits an existing prediction rule (prediction rules were described in section 3.2.10). The `rule_description` field is optional.

add_varfield(string variable_name; string field_name) Adds a new *field name* into the specified variable; the variable is created if it does not yet exists. If the variable exists, the field name is added at the bottom of the list (lowest priority in the processing). The variables have been discussed thoroughly in section 3.2.6.

calc(string function; string compound) Calculates the chosen function for the specified compound or all compounds (using the wildcard symbol, ∗). The calculations have been described in section 3.2.8; the *variables* and *data fields* generated are those specified in this section.

List of available functions:

- "ad ": calculates the cross-sectional area. The settings for the calcula-
  tion (as described in section 3.2.11) can be parameterized using the `set`
  command.

- "alogps ": calculates the LogP using the ALogPS plugin.

- "eu ": finds the hydrogen bond donnors/acceptors patterns.

- "kowwin ": gets the experimental LogP (if available) and the calculated
  LogP using the Kowwin algorithm; this requires an Internet connection.

- "lmax ": calculates the maximal length of the compound.

- "mw ": calculates the molecular weight and the summation formula of
  the compound.

- "smiles ": gets the SMILES string of the compound.

- "species ": calculates the acidic and basic pKa's.

- "tpsa ": calculates the polar surface area, using the TPSA algorithm
  implemented in the software.

- "xlogp ": calculates the LogP, using the XLogP algorithm implemented
  in the software.

close_project()   Closes the current project.

del_rule(string rule_name)   Deletes the specified *prediction rule*.

del_varfield(string variable_name; string field_name)   Deletes the speci-
fied *data field* of the selected *variable*.

del_var(string variable_name)   Deletes the *variable* and all its *data fields*.

export_mol(string compound; string file_name; integer flag)   Exports the
specified compound (using it unique identifier, as described in section 3.2.6) in
a MOL file (file_name); if flag=0, the amphiphilicity axis will not be embedded,
if flag=1 the axis will be embedded. If flag=2, the axis will be embedded and
the files cage.py and vector.py will be created in the c:\temp\ folder (for use
with PyMol). The definition of the amphiphilicity axis has been described in
section 3.2.4.

export_sdf(string compound; string file_name; integer include_rules) Exports structure(s) of one compound or all compounds (use the wildcard symbol, ∗) to one SDF file (file_name), including (include_rules=1) or not (include_rules=0) the results of the *prediction rules*.

export_xls(string file_name) Exports all the results (obtained with the run_rule function) to a tab-delimited file (which name is specified in file_name).

new_project(string file_name) Opens a PGP project file, or creates a new project based on a MOL or a SDF file.

open_project(string file_name) Opens an existing PGP project file.

run_rule(string rule_name; string compound) Runs the chosen rule or all rules (using the wildcard symbol, ∗) for the specified compound or all compounds (using the wildcard symbol, ∗).

save_project(string file_name) Saves the project in the specified PGP project file.

save_settings() Saves the settings changed with the set function. If settings are not saved, they won't be available during the next session.

set(string setting, string value) Sets the setting to the specified value. List of customizable settings:

- "ph": sets the pH for the calculation of the cross-sectional area.

- "ph_tag": defines if a tag is added to the *field name* generated by the calculation of the cross-sectional area; if *ph_tag* is set to true, the field name will be AD[#] with # the current pH, otherwise it will be AD.

- "digits": sets the minimal number of digits for compound duplicates.

- "hydrogen": value must be "true" to remove the hydrogens and apply a correction in the calculation of the cross-sectional area, otherwise "false".

- "increase_vdw": set to "true" to increase the size of the Van der Waals radii of charged atoms, for the calculation of the cross-sectional area; set to "false" otherwise. The increase will be by a factor of $|charge|+1$.

- "probe": sets the radius of the water probe, for the calculation of the cross-sectional area (in Å).

- "rule_def": if rule_def="true", the rules definitions will be included in the tab-delimited file (obtained with the "export_xls" function), otherwise the rules definitions will not be included.

- "temperature": set the temperature.

swap_rule(string rule1, string rule2)   Exchanges the rule1 with the rule2; this will permute the columns in the output file (obtained with the export_xls function). This might be also useful for rules referring to the result of another rule; in this case this rule has to be placed *after* the other one.

Scripting example

The following section will show an example of a *script file*. It will open a source SDF file containing multiple conformations of the same compound (benzylpenicillin), add several *prediction rules*, feed the source with experimental pKas, calculate the cross-sectional area at pH 7.4 and export the structure of the smallest calculated cross-sectional area, calculate the cross-sectional area at pH 8 and exporting all results in a tab-delimited file.

```
//
//Benzylpenicillin
//
new_project("D:\Substances\Benzylpenicillin_all.sdf")
add_rule("Ad_Display74"; "Display(Ad[7.4], 0)")
add_rule("Ad_Display8"; "Display(Ad[8], 0)")
add_rule("Energy_display"; "Display(Energy, 0)")
swap_rule("Ad_Display74"; "Energy_display")
add_rule("Export_Smallest"; "If AD[7.4] == Min(Ad[7.4]) Then
Export_Mol("D:\Substances\smallest\", 0)")
add_field(*; "PKA1_VALUE"; "2.62")
add_field(*; "PKA1_TYPE"; "acidic")
add_field(*; "PKA2_VALUE"; "-1.32")
add_field(*; "PKA2_TYPE";"basic")
set("ph"; "7.4")
set("ph_tag"; "true")
calc("ad"; *)
add_rule("EU_Avg"; "Display(Average(EU.TOTAL), 0)")
calc("eu"; *)
set("ph"; 8)
calc("ad"; *)
run_rule(*; *)
del_rule("Export_Smallest")
export_xls("D:\Substances\Benzylpenicillin_all.xls")
```

```
close_project()
```

3.3 Conclusion

The software represents a versatile and useful tool to handle many compound structures, calculate usual descriptors as well as the original descriptors developed in order to assess the particularity of two-step binding of drugs to the multidrug transporter P-glycoprotein, first the partitioning into the lipid bilayer (where the cross-sectional area represents a highly predictive parameter), and then the binding to the transporter *via* specific hydrogen bond patterns (type I and type II units). The working interface offers a convenient way to handle structures and related information, and gives the user the opportunity to visualize the three dimensional structure of each compound, as well as important parameters, like the type I and type II units, or the way the compound will be oriented at a hydrophilic-hydrophobic interface. Moreover, the scripting language offers the possibility to realize larger calculations, and to automatize some repetitive tasks on several structure files. Finally, the high flexibility of the software, on input, variables, transporter definition, and prediction rules, allows a wide variety of applications, for instance the elaboration of a whole set of hydrogen bonding definition and prediction rules to assess the binding of a drug to another transporter, like the multidrug resistance transporter MRP1.

Bibliography

[1] S. V. Ambudkar, S. Dey, C. A. Hrycyna, M. Ramachandra, I. Pastan, and M. M. Gottesman. Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annual Review of Pharmacology and Toxicology*, 39:361–98, 1999.

[2] T. Litman, T. E. Druley, W. D. Stein, and S. E. Bates. From mdr to mxr: new understanding of multidrug resistance systems, their properties and clinical significance. *Cellular and Molecular Life Sciences*, 58(7):931–59, 2001.

[3] A. Seelig, E. Landwojtowicz, H. Fischer, and X. L. Blatter. Towards p-glycoprotein structure-activity relationships. In van der Waterbeemd, Lennernas, and Arthurson, editors, *Drug bioavailability/Estimation of solubility, permeability and absorption*, Methods and principles in medicinal chemistry. Wiley-VCH Verlag GmbH, Weinheim, 2003.

[4] A. Seelig and E. Gatlik-Landwojtowicz. Inhibitors of multidrug efflux transporters: their membrane and protein interactions. *Mini Reviews in Medicinal Chemistry*, 5(2):135–51, 2005.

[5] E. Gatlik-Landwojtowicz, P. Aanismaa, and A. Seelig. Quantification and characterization of p-glycoprotein-substrate interactions. *Biochemistry*, 45(9):3020–32, 2006.

[6] A. Seelig and G. Gerebtzoff. Enhancement of drug absorption by non-charged detergents through membrane and p-glycoprotein binding. *Expert Opin Drug Metab Toxicol*, 2(5):733–52, 2006.

[7] H. Fischer, R. Gottschlich, and A. Seelig. Blood-brain barrier permeation: molecular parameters governing passive diffusion. *Journal of Membrane Biology*, 165(3):201–11, 1998.

[8] G. Gerebtzoff, X. Li-Blatter, H. Fischer, A. Frentzel, and A. Seelig. Halogenation of drugs enhances membrane binding and permeation. *Chembiochem*, 5(5):676–84, 2004.

[9] H. Fischer. *Passive diffusion and active transport through biological membranes - Binding of drugs to transmembrane receptors*. PhD thesis, Universität Basel, 1998.

[10] G. Gerebtzoff and A. Seelig. In silico prediction of blood-brain barrier permeation using the calculated molecular cross-sectional area as main parameter. *Journal of Chemical Information and Modeling*, 46(6):2638–50, 2006.

[11] J. Lin, D. C. Sahakian, S. M. de Morais, J. J. Xu, R. J. Polzer, and S. M. Winter. The role of absorption, distribution, metabolism, excretion and toxicity in drug discovery. *Current Topics in Medicinal Chemistry*, 3(10):1125–54, 2003.

[12] J. H. Lin and M. Yamazaki. Role of p-glycoprotein in pharmacokinetics: clinical implications. *Clinical Pharmacokinetics*, 42(1):59–98, 2003.

[13] D. Keppler, J. Konig, and M. Buchler. The canalicular multidrug resistance protein, cmrp/mrp2, a novel conjugate export pump expressed in the apical membrane of hepatocytes. *Advances in Enzyme Regulation*, 37:321–33, 1997.

[14] M. F. Fromm. Importance of p-glycoprotein for drug disposition in humans. *European Journal of Clinical Investigation*, 33 Suppl 2:6–9, 2003.

[15] A. Avdeef. Physicochemical profiling (solubility, permeability and charge state). *Current Topics in Medicinal Chemistry*, 1(4):277–351, 2001.

[16] A. Seelig. The role of size and charge for blood-brain barrier permeation of drugs and fatty acids. *Journal of Molecular Neuroscience*, 33(1):32–41, 2007.

[17] M. F. Fromm. Importance of p-glycoprotein at blood-tissue barriers. *Trends in Pharmacological Sciences*, 25(8):423–9, 2004.

[18] P. Borst, R. Evers, M. Kool, and J. Wijnholds. The multidrug resistance protein family. *Biochimica et Biophysica Acta*, 1461(2):347–57, 1999.

[19] M. Dean, A. Rzhetsky, and R. Allikmets. The human atp-binding cassette (abc) transporter superfamily. *Genome Research*, 11(7):1156–66, 2001.

[20] G. Chang. Multidrug resistance abc transporters. *FEBS Letters*, 555(1):102–5, 2003.

[21] A. H. Schinkel and J. W. Jonker. Mammalian drug efflux transporters of the atp binding cassette (abc) family: an overview. *Advanced drug delivery reviews*, 55(1):3–29, 2003.

[22] S. P. Cole, G. Bhardwaj, J. H. Gerlach, J. E. Mackie, C. E. Grant, K. C. Almquist, A. J. Stewart, E. U. Kurz, A. M. Duncan, and R. G. Deeley. Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science*, 258(5088):1650–4, 1992.

[23] H. Omote and M. K. Al-Shawi. Interaction of transported drugs with the lipid bilayer and p-glycoprotein through a solvation exchange mechanism. *Biophysical Journal*, 90(11):4046–59, 2006.

[24] A. H. Schinkel. P-glycoprotein, a gatekeeper in the blood-brain barrier. *Advanced drug delivery reviews*, 36(2-3):179–194, 1999.

[25] A. Seelig, X. L. Blatter, and F. Wohnsland. Substrate recognition by p-glycoprotein and the multidrug resistance-associated protein mrp1: a comparison. *International Journal of Clinical Pharmacology and Therapeutics*, 38(3):111–21, 2000.

[26] F. J. Sharom. The p-glycoprotein efflux pump: how does it transport drugs? *Journal of Membrane Biology*, 160(3):161–75, 1997.

[27] A. H. Schinkel, E. Wagenaar, C. A. Mol, and L. van Deemter. P-glycoprotein in the blood-brain barrier of mice influences the brain penetration and pharmacological activity of many drugs. *Journal of Clinical Investigation*, 97(11):2517–24, 1996.

[28] K. M. Mahar Doan, J. E. Humphreys, L. O. Webster, S. A. Wring, L. J. Shampine, C. J. Serabjit-Singh, K. K. Adkison, and J. W. Polli. Passive permeability and p-glycoprotein-mediated efflux differentiate central nervous system (cns) and non-cns marketed drugs. *Journal of Pharmacology and Experimental Therapeutics*, 303(3):1029–37, 2002.

[29] J. Van Asperen, O. Van Tellingen, and J. H. Beijnen. The pharmacological role of p-glycoprotein in the intestinal epithelium. *Pharmacological Research*, 37(6):429–35, 1998.

[30] T. Litman, T. Zeuthen, T. Skovsgaard, and W. D. Stein. Structure-activity relationships of p-glycoprotein interacting drugs: kinetic characterization of their effects on atpase activity. *Biochimica et Biophysica Acta*, 1361(2):159–68, 1997.

[31] T. Litman, T. Zeuthen, T. Skovsgaard, and W. D. Stein. Competitive, non-competitive and cooperative interactions between substrates of p-glycoprotein as measured by its atpase activity. *Biochimica et Biophysica Acta*, 1361(2):169–76, 1997.

[32] Z. E. Sauna and S. V. Ambudkar. Characterization of the catalytic cycle of atp hydrolysis by human p-glycoprotein. the two atp hydrolysis events in

a single catalytic cycle are kinetically similar but affect different functional outcomes. *Journal of Biological Chemistry*, 276(15):11653–61, 2001.

[33] E. Landwojtowicz, P. Nervi, and A. Seelig. Real-time monitoring of p-glycoprotein activation in living cells. *Biochemistry*, 41(25):8050–7, 2002.

[34] E. Gatlik-Landwojtowicz, P. Aanismaa, and A. Seelig. The rate of p-glycoprotein activation depends on the metabolic state of the cell. *Biochemistry*, 43(46):14840–51, 2004.

[35] A. Seelig. A general pattern for substrate recognition by p-glycoprotein. *European Journal of Biochemistry*, 251(1-2):252–61, 1998.

[36] D. Schwab, H. Fischer, A. Tabatabaei, S. Poli, and J. Huwyler. Comparison of in vitro p-glycoprotein screening assays: recommendations for their use in drug discovery. *Journal of Medicinal Chemistry*, 46(9):1716–25, 2003.

[37] C. A. Doige, X. Yu, and F. J. Sharom. Atpase activity of partially purified p-glycoprotein from multidrug-resistant chinese hamster ovary cells. *Biochimica et Biophysica Acta*, 1109(2):149–60, 1992.

[38] S. V. Ambudkar, I. H. Lelong, J. Zhang, C. O. Cardarelli, M. M. Gottesman, and I. Pastan. Partial purification and reconstitution of the human multidrug-resistance pump: characterization of the drug-stimulatable atp hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 89(18):8472–6, 1992.

[39] Z. Hollo, L. Homolya, C. W. Davis, and B. Sarkadi. Calcein accumulation as a fluorometric functional assay of the multidrug transporter. *Biochimica et Biophysica Acta*, 1191(2):384–8, 1994.

[40] M. Yamazaki, W. E. Neway, T. Ohe, I. Chen, J. F. Rowe, J. H. Hochman, M. Chiba, and J. H. Lin. In vitro substrate identification studies for p-glycoprotein-mediated transport: species difference and predictability of in vivo results. *Journal of Pharmacology and Experimental Therapeutics*, 296(3):723–35, 2001.

[41] A. Seelig and E. Landwojtowicz. Structure-activity relationship of p-glycoprotein substrates and modifiers. *European Journal of Pharmaceutical Sciences*, 12(1):31–40, 2000.

[42] M. K. Al-Shawi, M. K. Polar, H. Omote, and R. A. Figler. Transition state analysis of the coupling of drug transport to atp hydrolysis by p-glycoprotein. *Journal of Biological Chemistry*, 278(52):52629–40, 2003.

[43] R. J. Raggers, T. Pomorski, J. C. Holthuis, N. Kalin, and G. van Meer. Lipid traffic: the abc of transbilayer movement. *Traffic*, 1(3):226–34, 2000.

[44] Y. Romsicki and F. J. Sharom. Phospholipid flippase activity of the reconstituted p-glycoprotein multidrug transporter. *Biochemistry*, 40(23):6937–47, 2001.

[45] P. Crivori, B. Reinach, D. Pezzetta, and I. Poggesi. Computational models for identifying potential p-glycoprotein substrates and inhibitors. *Mol Pharm*, 3(1):33–44, 2006.

[46] V. K. Gombar, J. W. Polli, J. E. Humphreys, S. A. Wring, and C. S. Serabjit-Singh. Predicting p-glycoprotein substrates by a quantitative structure-activity relationship model. *Journal of Pharmaceutical Sciences*, 93(4):957–68, 2004.

[47] R. Didziapetris, P. Japertas, A. Avdeef, and A. Petrauskas. Classification analysis of p-glycoprotein substrate specificity. *Journal of Drug Targeting*, 11(7):391–406, 2003.

[48] Y. Li, Y.-H. Wang, L. Yang, S.-W. Zhang, C.-H. Liu, and S.-L. Yang. Comparison of steroid substrates and inhibitors of p-glycoprotein by 3d-qsar analysis. *Journal of Molecular Structure*, 733(1-3):111–118, 2005.

[49] I. K. Pajeva and M. Wiese. Pharmacophore model of drugs involved in p-glycoprotein multidrug resistance: explanation of structural variety (hypothesis). *Journal of Medicinal Chemistry*, 45(26):5671–86, 2002.

[50] S. Ekins, R. B. Kim, B. F. Leake, A. H. Dantzig, E. G. Schuetz, L. B. Lan, K. Yasuda, R. L. Shepard, M. A. Winter, J. D. Schuetz, J. H. Wikel, and S. A. Wrighton. Three-dimensional quantitative structure-activity relationships of inhibitors of p-glycoprotein. *Molecular Pharmacology*, 61(5):964–73, 2002.

[51] S. Ekins, R. B. Kim, B. F. Leake, A. H. Dantzig, E. G. Schuetz, L. B. Lan, K. Yasuda, R. L. Shepard, M. A. Winter, J. D. Schuetz, J. H. Wikel, and S. A. Wrighton. Application of three-dimensional quantitative structure-activity relationships of p-glycoprotein inhibitors and substrates. *Molecular Pharmacology*, 61(5):974–81, 2002.

[52] J. E. Penzotti, M. L. Lamb, E. Evensen, and P. D. Grootenhuis. A computational ensemble pharmacophore model for identifying substrates of p-glycoprotein. *Journal of Medicinal Chemistry*, 45(9):1737–40, 2002.

[53] A. Garrigues, N. Loiseau, M. Delaforge, J. Ferte, M. Garrigos, F. Andre, and S. Orlowski. Characterization of two pharmacophores on the multidrug transporter p-glycoprotein. *Molecular Pharmacology*, 62(6):1288–98, 2002.

[54] G. Cianchetta, R. W. Singleton, M. Zhang, M. Wildgoose, D. Giesing, A. Fravolini, G. Cruciani, and R. J. Vaz. A pharmacophore hypothesis for p-glycoprotein substrate recognition using grind-based 3d-qsar. *Journal of Medicinal Chemistry*, 48(8):2927–2935, 2005.

[55] M. A. Cabrera, I. Gonzlez, C. Fernndez, C. Navarro, and M. Bermejo. A topological substructural approach for the prediction of p-glycoprotein substrates. *Journal of Pharmaceutical Sciences*, 95(3):589–606, 2006.

[56] Y. Xue, C. W. Yap, L. Z. Sun, Z. W. Cao, J. F. Wang, and Y. Z. Chen. Prediction of p-glycoprotein substrates by a support vector machine approach. *Journal of Chemical Information and Computer Sciences*, 44(4):1497–505, 2004.

[57] Y. H. Wang, Y. Li, S. L. Yang, and L. Yang. Classification of substrates and inhibitors of p-glycoprotein using unsupervised machine learning approach. *Journal of Chemical Information and Modeling*, 45(3):750–7, 2005.

[58] T. Langer, M. Eder, R. D. Hoffmann, P. Chiba, and G. F. Ecker. Lead identification for modulators of multidrug resistance based on in silico screening with a pharmacophoric feature model. *Archiv der Pharmazie*, 337(6):317–27, 2004.

[59] M. Gniazdowski, W. A. Denny, S. M. Nelson, and M. Czyz. Transcription factors as targets for dna-interacting drugs. *Current Medicinal Chemistry*, 10(11):909–24, 2003.

[60] G. F. Ecker and C. R. Noe. In silico prediction models for blood-brain barrier permeation. *Current Medicinal Chemistry*, 11(12):1617–28, 2004.

[61] T. R. Stouch and O. Gudmundsson. Progress in understanding the structure-activity relationships of p-glycoprotein. *Advanced drug delivery reviews*, 54(3):315–28, 2002.

[62] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 23(1-3):3–25, 1997.

[63] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability

in drug discovery and development settings. *Advanced drug delivery reviews*, 46(1-3):3–26, 2001.

[64] J. Karwatsky, M. C. Lincoln, and E. Georges. A mechanism for p-glycoprotein-mediated apoptosis as revealed by verapamil hypersensitivity. *Biochemistry*, 42(42):12163–73, 2003.

[65] D. F. Tang-Wai, A. Brossi, L. D. Arnold, and P. Gros. The nitrogen of the acetamido group of colchicine modulates p-glycoprotein-mediated multidrug resistance. *Biochemistry*, 32(25):6470–6, 1993.

[66] M. Muller, E. Bakos, E. Welker, A. Varadi, U. A. Germann, M. M. Gottesman, B. S. Morse, I. B. Roninson, and B. Sarkadi. Altered drug-stimulated atpase activity in mutants of the human multidrug resistance protein. *Journal of Biological Chemistry*, 271(4):1877–83, 1996.

[67] T. Matsunaga, E. Kose, S. Yasuda, H. Ise, U. Ikeda, and S. Ohmori. Determination of p-glycoprotein atpase activity using luciferase. *Biological and Pharmaceutical Bulletin*, 29(3):560–4, 2006.

[68] E. G. Schuetz, K. Yasuda, K. Arimori, and J. D. Schuetz. Human mdr1 and mouse mdr1a p-glycoprotein alter the cellular retention and disposition of erythromycin, but not of retinoic acid or benzo(a)pyrene. *Archives of Biochemistry and Biophysics*, 350(2):340–7, 1998.

[69] A. Crowe. The influence of p-glycoprotein on morphine transport in caco-2 cells. comparison with paclitaxel. *European Journal of Pharmacology*, 440(1):7–16, 2002.

[70] R. X. Wang, Y. Gao, and L. H. Lai. Calculating partition coefficient by atom-additive method. *Perspectives in Drug Discovery and Design*, 19(1):47–66, 2000.

[71] R. Sayle. Physiological ionization and pka prediction, 2000.

[72] P. Ertl, B. Rohde, and P. Selzer. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, 43(20):3714–7, 2000.

Halogenation of Drugs Enhances Membrane Binding and Permeation

Grégori Gerebtzoff,[a] Xiaochun Li-Blatter,[a] Holger Fischer,[b] Adrian Frentzel,[c] and Anna Seelig*[a]

Halogenation of drugs is commonly used to enhance membrane binding and permeation. We quantify the effect of replacing a hydrogen residue by a chlorine or a trifluoromethyl residue in position C-2 of promazine, perazine, and perphenazine analogues. Moreover, we investigate the influence of the position (C-6 and C-7) of residue CF_3 in benzopyranols. The twelve drugs are characterized by surface activity measurements, which yield the cross-sectional area, the air–water partition coefficient, and the critical micelle concentration. By using the first two parameters (A_D and K_{aw}) and the appropriate membrane packing density, the lipid–water partition coefficients, are calculated in excellent agreement with the lipid–water partition coefficients measured by means of isothermal titration calorimetry for small unilamellar vesicles of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine. Replacement of a hydrogen residue by a chlorine and a trifluoromethyl residue enhances the free energy of partitioning into the lipid membrane, on average by $\Delta G_{lw} \approx -1.3$ or -4.5 kJ mol^{-1}, respectively, and the permeability coefficient by a factor of ~ 2 or ~ 9, respectively. Despite exhibiting practically identical hydrophobicities, the two benzopyranol analogues differ in their permeability coefficients by almost an order of magnitude; this is due to their different cross-sectional areas at the air–water and lipid–water interfaces.

Introduction

Most drugs permeate biological membranes by passive diffusion. The extent of permeation depends, on one hand, on the properties of the membrane and, on the other, on those of the diffusing molecule. Membranes are highly organized, anisotropic systems that are nevertheless fluid enough to allow considerable translational, rotational, and flexing movements of the constituent lipid and protein molecules. Under physiological conditions, the lipid-bilayer membrane is in a liquid crystalline state and behaves like optically uniaxial crystals with the optical axis perpendicular to the surface of the membrane.[1] This is in contrast to membrane-mimicking systems such as octanol or hexadecane, which are isotropic organic solvents. By using solid-state NMR techniques, a quantitative analysis of the molecular ordering and dynamics of a lipid bilayer has become possible with a segment-to-segment resolution. The packing density of the hydrocarbon chains is well described in terms of a statistical order profile. Membrane packing and ordering increase with the cholesterol content and decrease with increasing temperature or increasing fatty acyl chain unsaturation (for a review see ref. [2]).

The lateral packing density of a lipid bilayer can, alternatively, be assessed in comparison to the packing density of a lipid monolayer (for a review see ref. [3]). The packing density of planar bilayers consisting of the most abundant natural lipid, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine, POPC, was determined as $\pi_M = 32$ mN m^{-1} at ambient temperature.[4] That of unilamellar lipid vesicles is somewhat lower and varies between $\pi_M = 25$–32 mN m^{-1} depending on the size of the vesicle,[5] while that of cholesterol-containing membranes such as erythrocyte membranes is higher ($\pi_M = 32$–35 mN m^{-1}).[6] The membrane packing density influences binding[7,8] and permeation[9] of drugs in an exponential manner.

The properties of the drug with the strongest impact on membrane binding are hydrophobicity (which is suitably reflected by the air–water partition coefficient) and the cross-sectional area.[9,10] Whereas partitioning into an isotropic organic solvent increases with the molecular volume,[11] partitioning into an anisotropic lipid bilayer decreases exponentially with increasing cross-sectional area of the molecule.[7,8,9] Drugs are generally weak bases or weak acids that are present in a charged as well as an uncharged form under physiological conditions. Charged and uncharged molecules can insert into the lipid–water interface; however, only a fraction of uncharged drugs can permeate a lipid membrane.

Strategies to enhance passive diffusion often consist of replacing a hydrogen residue by a chlorine or a trifluoromethyl residue. Despite the relatively high numbers of halogenated drugs on the market, little information is available on the quantitative effects of halogenation on drug binding to membranes and drug diffusion through membranes. We therefore quantify the influence of such replacements on the lipid–water partition coefficient, K_{lw}, and the permeability coefficient, P. We chose chlorinated and fluorinated phenothiazine analogues

[a] G. Gerebtzoff, X. Li-Blatter, Dr. A. Seelig
Biophysical Chemistry, Biozentrum, University of Basel
Klingelbergstrasse 70, 4056 Basel (Switzerland)
Fax: (+41) 61-267-21-89
E-mail: anna.seelig@unibas.ch

[b] Dr. H. Fischer
Present address: F. Hoffmann–La Roche Ltd.
Pharma Research, Molecular Properties
Grenzacherstrasse, 4070 Basel (Switzerland)

[c] Dr. A. Frentzel
Present Address: Mepha Pharma
Dornacherstrasse 14, 4147 Aesch (Switzerland)

(analogues of promazine, perazine, and perphenazine) to investigate the influence of the type of halogenation, and 6- and 7-trifluoromethyl benzopyranol to investigate the influence of the position of a $-CF_3$ group.[12]

We characterized the drugs in terms of their cross-sectional area, A_D, their air–water partition coefficient, K_{aw}, and their critical micelle concentration, CMC_D, which are all obtained by means of surface-activity measurements (SAMs).[9,13] Using the first two parameters, A_D and K_{aw}, we calculated K_{lw} for membranes of different packing densities, π_M. This approach is validated by comparing the lipid–water partition coefficients predicted on the basis of surface-activity measurements with those determined by means of isothermal titration calorimetry, ITC.[10]

The permeability coefficient, P, is then calculated by taking into account the lipid–water partition coefficient derived from surface-activity measurements, the ionization constant, and the pK_a of the compound as outlined previously.[14]

Experimental Section

Compounds: Promazine·HCl, triflupromazine·HCl, trifluoperazine·2HCl, and fluphenazine·HCl were obtained from Sigma–Aldrich, Steinheim, Germany. *cis*-Flupenthixol·2HCl, chlorpromazine·HCl, chlorperphenazine·HCl, 6-trifluoromethyl benzopyranol (6-trifluoromethyl-3,4-dihydro-4-(1,6-dihydro-1-methyl-6-oxo-3-pyridazinyl-oxy)-2,2-dimethyl-2H-1-benzopyran-3-ol) and 7-trifluoromethyl benzopyranol (7-trifluoromethyl-3,4-dihydro-4-(1,6-dihydro-1-methyl-6-oxo-3-pyridazinyloxy)-2,2-dimethyl-2H-1-benzopyran-3-ol) were kind gifts from Merck, Darmstadt, Germany. Perazine·dimaleate and chlorperazine·HCl were kindly provided by Jacek Wójcikowski, Institute of Pharmacology, Polish Academy of Science, Kraków, Poland, and by F. Hoffmann–La Roche Ltd., Basel, Switzerland, respectively.

Buffers: For SAMs and ITC, Tris/HCl buffer (50 mM), containing NaCl (114 mM) was used. SAMs were performed at ambient temperature ($T = 24 \pm 1$°C), and ITC was measured at 37°C. Buffers were adjusted to pH 7.4 at the temperatures used for the respective measurements. For SAMs, stock solutions of drugs were prepared at concentrations of 10^{-4}–10^{-2} M either in nanopure water with a resistance of 17–18 MΩm^{-1} or in methanol. For ITC drugs were dissolved in buffer solution.

For SAMs we either used a Teflon trough designed by Fromherz (Mayer Feintechnik, Göttingen Germany)[15] with a filling volume of one compartment of 20 mL or a home-built Teflon trough (filling volume 3 mL). To maintain a constant humidity, the troughs were covered by a Plexiglas hood. The surface pressure, $\pi = \gamma_0 - \gamma$, where γ_0 is the surface tension of the pure buffer and γ the surface tension of the drug solution, was monitored with filter paper (Whatman No. 1) connected to a Wilhelmy balance. For drugs dissolved in methanol, the measured surface pressure was corrected for the intrinsic surface pressure of methanol.[13] The total methanol concentration in the final drug solution was 10% (v/v). For the home-built trough, in which evaporation is not compensated for by an added volume as in the Fromherz trough, the surface pressure was corrected for the effects of evaporation and buoyancy.

The thermodynamics of drug adsorption at the air–water interface is described by the Gibbs adsorption isotherm:

$$d\pi = RT\Gamma d\ln C \tag{1}$$

Here C is the concentration of the amphiphile in the Teflon trough, RT is the thermal energy per mole, and Γ is the surface excess concentration defined as the inverse of the product of the Avogadro number, N_A, and the area requirement of the surface active molecule at the interface, A_S:

$$\Gamma = (N_A A_S)^{-1} \tag{2}$$

The surface excess concentration, Γ, increases with C up to a limiting value Γ_∞. As long as Γ is constant, a plot of π versus $\log C$ yields a straight line. The area requirement of the compound, A_S, was evaluated from the quasilinear part, $d\pi/d\ln C$, of the Gibbs adsorption isotherm:

$$\Gamma_\infty = (1/RT)\, d\pi/d\ln C \tag{3}$$

For data analysis, we developed a program that selects the quasilinear part of the $\pi/\log C$ plot in an automatic and reproducible manner.

To evaluate the air–water partition coefficient, the Szyszkowski equation was used:

$$\pi = RT\Gamma_\infty \ln(K_{aw}C + 1) \tag{4}$$

This is an integral version of Equation (3) combined with a Langmuir adsorption isotherm. Combining Equations (2) and (4) allows evaluation of K_{aw} by calculating the slope of the linear regression line through data points corresponding to the quasilinear part of the π versus $\log C$ plot by using A_S, determined as described above:

$$K_{aw}C = e^{\pi A_S N_A/RT} - 1 \tag{5}$$

C is the equilibrium concentration of the drug, C_{eq}, in bulk solution, which is defined as the total concentration, C_{tot}, minus the concentration of the drug adsorbed to the air–water interface, C_b:

$$C_{eq} = C_{tot} - C_b \tag{6}$$

C_b is the product of Γ and the surface area of the solution, A, per total volume, V ($C_b = \Gamma A V^{-1}$). C_b is negligibly small as long as K_{aw} is small ($< 10^6$ M^{-1}) and is therefore generally neglected ($C_{eq} \sim C_{tot}$).[9] In the present evaluation procedure, however, we correct for C_b; this slightly influences the parameters of hydrophobic compounds, such as *trans*-flupenthixol.

Lipid–water partition coefficients determined from surface-activity measurements: Knowledge of the K_{aw} and the A_D of a drug allows estimation of K_{lw} according to Equation (7):[9]

$$K_{lw} = K_{aw}e^{-\pi_M A_D/kT} \tag{7}$$

Here kT is the thermal energy, and π_M is the lipid packing density of the membrane. K_{lw} [M^{-1}] is defined as the quotient of mole fraction of drugs bound to the membrane, X_b, and the concentration of the drug in aqueous solution, C_{eq} [mol L^{-1}]:

$$K_{lw} = X_b/C_{eq} \tag{8}$$

K_{lw} can be transformed to the dimensionless partition coefficient, γ_{lw}, defined as the quotient of the drug concentration in the membrane and the drug concentration in the aqueous phase, both

given in [mol L^{-1}]:

$$\gamma_{lw} = C_m/C_{eq} = C_l.K_{lw} \tag{9}$$

C_l is the molar concentration of lipid ($C_l = 1.05$ mol L^{-1}) if the density of lipids is assumed to be $\rho = 0.8$ kg L^{-1} and the molecular weight is $M_W = 760.1$ Da (POPC).

Free energies: The free energy of self-association or micelle formation, ΔG_{mic}, the free energy of partitioning into the air–water interface, ΔG_{aw}, and the free energy of partitioning into the lipid–water interface, ΔG_{lw}, are obtained as follows:

$$\Delta G_{mic} = RT \ln (CMC_D/C_w) \tag{10}$$

$$\Delta G_{aw} = -RT \ln (C_w K_{aw}) \tag{11}$$

$$\Delta G_{lw} = -RT \ln (C_w K_{lw}) \tag{12}$$

where C_w is the molar concentration of water ($C_w = 55.5$ mol L^{-1} at $24 \pm 1\,°C$ and $C_w = 55.3$ mol L^{-1} at $37\,°C$).

Permeability coefficient: As outlined previously,[14, 16] P can be estimated on the basis of surface activity measurements. For a nonelectrolyte P is proportional to the product of γ_{lw} and the diffusion coefficient, D:

$$P = \gamma_{lw}D/\Delta x \tag{13}$$

where Δx is the thickness of the membrane. The diffusion coefficient is defined as:

$$D = kT/(6\pi\eta r) \tag{14}$$

where kT is the thermal energy, η is the membrane viscosity, and r is the molecular radius, which is derived from A_D. For the following calculations, the membrane viscosity is assumed to be $\eta = 1$ poise and the membrane thickness, $\Delta x = 50$ Å.

The fraction of the nonionized form of a basic drug, f_A, that is present at a particular pH can be calculated as:

$$f_A = \frac{[A]}{[AH^+] + [A]} = (1 + 10^{pK_a - pH})^{-1} \tag{15}$$

where [A] and [HA$^+$] are the concentration of the uncharged and the ionized form of the drug, respectively. For the calculation of the permeability coefficients, we use the standard pK_a values corrected for temperature. The partition coefficient for the permeating species of a basic drug, γ_{lw}^*, is thus:

$$\gamma_{lw}^* = \gamma.f_A \tag{16}$$

and the permeability coefficient is:

$$P = \gamma_{lw}^* D/\Delta x \tag{17}$$

Isothermal titration calorimetry: Drug partitioning into small unilamellar vesicles (SUVs) was measured by means of high-sensitivity ITC with a Microcal VP-ITC calorimeter (Microcal, Northampton, MA). A suspension of small unilamellar vesicles formed from 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) ($C_l = 30–35$ μM) was injected in 3–10 μL aliquots into the drug solution in the calorimeter cell ($V_{cell} = 1.4037$ mL) by using a Hamilton syringe

coupled with a stepping motor. For all drugs, injections gave rise to exothermic heats of reaction, produced by the partitioning of the drug into the membrane (for further details see refs. [10] and [17]).

For uncharged drugs, binding to the membrane is best described by a simple partition equilibrium:

$$K_{lw} = X_b/C_{eq} \tag{18}$$

For the charged cationic drugs used in the present investigation, this simple approach [Eq. (18)] is not adequate and yields concentration-dependent partition coefficients. This is due to the fact that partitioning of the drug into the electrically neutral lipid–water interface leads to a positive surface charge density, σ, (defined as total surface charge, Q_T per total membrane surface area, A_T) and, in turn, to a positive surface potential, Ψ. As a consequence, the lipid–water partition coefficient decreases with increasing concentration.[18] A concentration-independent binding constant, K_{lw_0}, can either be obtained by plotting X_b/C_{eq} against C_{eq} and extrapolating to $C_{eq} = 0$, as shown previously for the binding of cationic peptides to POPC vesicles, or by applying the Gouy–Chapman theory.[18]
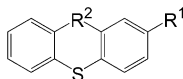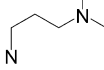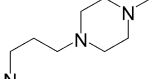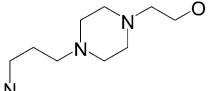
Results
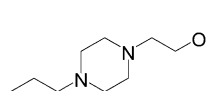
The compounds investigated are displayed in Table 1 A and B. The phenothiazine analogues, series A–C, as well as the benzopyranol analogues, series D, carry an uncharged or hydrophobic residue, R^1, and a cationic or hydrophilic residue, R^2, and are thus amphiphilic compounds. Series A–C represent promazine, perazine, and perphenazine analogues carrying a hydrogen atom, a chlorine atom, or a trifluoromethyl group as residue R^1. Compounds in series D exhibit identical sum formulas but differ with respect to the position of the nonpolar –CF$_3$ group (R^{1a} and R^{1b}).

Surface activity measurements, SAMs

Injection of an amphiphilic drug into a monolayer trough filled with buffer is followed by partitioning of the drug between the aqueous phase and the air–water interface. Molecules in the air–water interface orient such that the hydrophilic residue, –R^2, remains immersed in the aqueous phase and the hydrophobic residue, –R^1, reaches into the air. The surfa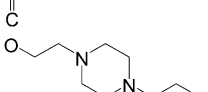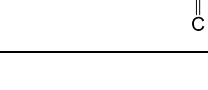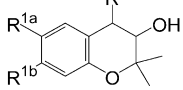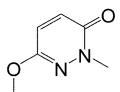ce activity as a function of concentration (π/log C plot or Gibbs adsorption isotherm) was measured for all compounds at pH 7.4 and pH 8.0. The lower pH corresponds to the condition in aqueous bulk solution, and the higher pH reflects that close to an electrically neutral lipid membrane to which cationic drugs are bound and which therefore exhibits a positive surface potential, Ψ.[5, 19, 20]

Cross-sectional areas, air–water partition coefficients, and critical micelle concentrations

Figure 1 shows the π/log C plots of the promazine and perazine analogues, series A and B, respectively. The slope of the quasilinear part of the Gibbs adsorption isotherm (solid line) yields the surface excess concentration, Γ_∞ [Eq. (3), below], and

Table 1. Analogues investigated.

Series	Name	R¹	R²
A	promazine	H	
	chlorpromazine	Cl	
	triflupromazine	CF₃	
B	perazine	H	
	chlorperazine	Cl	
	trifluoperazine	CF₃	
C	chlorperphenazine	Cl	
	fluphenazine	CF₃	
	cis-flupenthixol	CF₃	
	trans-flupenthixol	CF₃	

Series	Name	R¹ᵃ	R¹ᵇ	R²
D	6-(trifluoromethyl)benzopyranol	CF₃	H	
	7-(trifluoromethyl)benzopyranol	H	CF₃	

Figure 1. $\pi/\log C$ plots of promazine (series A) and perazine (series B) analogues. Solid symbols indicate the quasilinear part of the Gibbs adsorption isotherm. The corresponding slope is shown as a solid line. A) Promazine (squares), chlorpromazine (circles), and triflupromazine (triangles). The $\pi/\log C$ plots of promazine and triflupromazine consist of two independent measurements. B) Perazine (lozenges), chlorperazine (pentagons), and trifluoperazine (triangles). Measurements were performed at pH 8.0 (50 mM Tris/HCl containing 114 mM NaCl).

in turn the surface area requirement, A_s, of the drug molecule at the air–water interface [Eq. (2), below]. The intersection of the linear slope and the solid line drawn through the points of constant surface pressure at high concentrations is defined as the critical micelle concentration of the drug, CMC$_D$. Due to a comparatively low amphiphilicity (see Discussion) and a concomitant high tendency to aggregate at higher pH values, chlorpromazine could only be measured up to pH 7.8.

Figure 2 A shows the $\pi/\log C$ plots of triflupromazine as a function of the pH. At high pH the compounds are only partially charged and therefore partition into the air–water interface even at low concentrations ($C < 10^{-6}$ M). The lower the pH of the solution, the lower is the air–water partition coefficient, K_{aw}, the flatter the slope of the $\pi/\log C$ plot and the larger the area requirement of the molecule at the air–water interface, A_s. The pH dependence of the surface area requirement, A_s, (■) and the critical micelle concentration, CMC$_D$ (△) for triflupromazine are summarized in Figure 2 B. At low pH values where the drug is fully protonated, the area requirement at the interface, A_s, is large due to charge-repulsion effects,[9] but decreases with increasing pH. For promazine (not shown) and triflupromazine (Figure 2 B), the minimum area would, in principle,
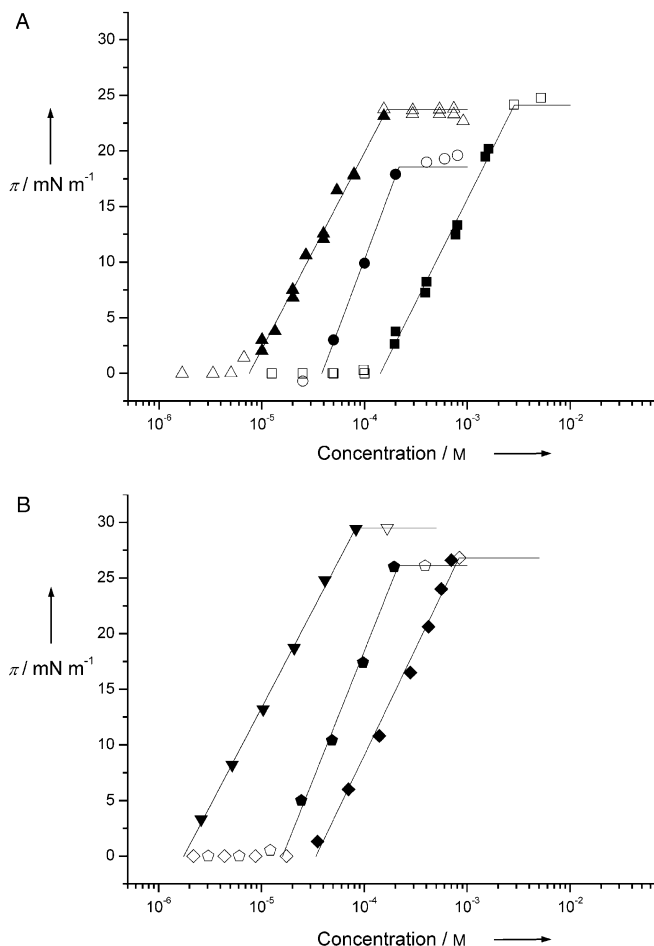
be reached close to pH 8.5. However, these drugs tend to form small micelles or aggregates, even below the apparent CMC$_D$ at pH > 7.5. This can lead to a decrease in the slope of the Gibbs adsorption isotherm and, in turn, to a small *apparent* increase in area as seen in Figure 2 B. Therefore, for the highly charged compounds promazine and triflupromazine, we used the extrapolated minimal values at pH 8.5 for the following calculations.

For most cationic drugs, however, the surface area requirement, A_s, measured at pH 8.0 corresponds well to the minimum area and thus reflects the cross-sectional area, A_D, of a drug as shown previously.[9] Table 2 summarizes A_s, assumed to correspond to A_D, the corresponding K_{aw}, and CMC$_D$.

Figure 2 C shows the free energy of self association, ΔG_{mic} (◇) [Eq. (12)] in comparison to the free energy of partitioning into the air–water interface, ΔG_{aw} (■) [Equation (10)] for triflupromazine. The difference between ΔG_{aw} and ΔG_{mic} was defined as the amphiphilicity, $\Delta\Delta G_{am}$, of the compound:[9]
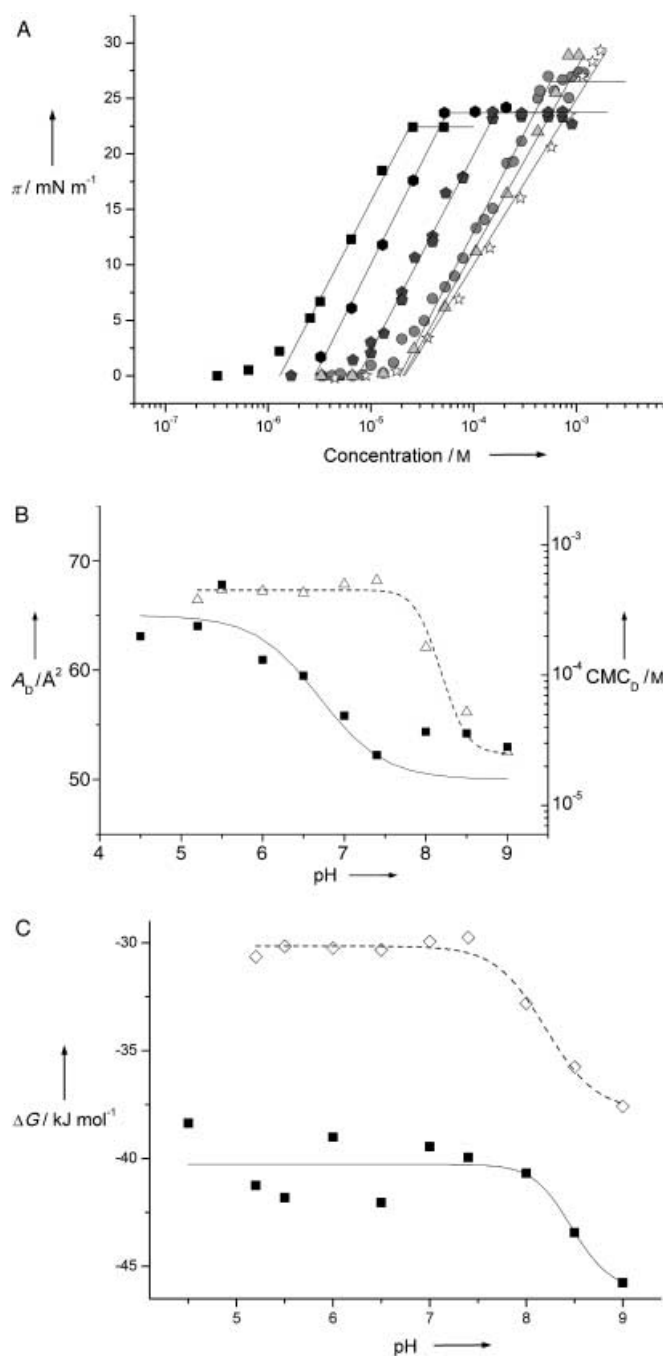
Figure 2. Surface-activity measurements of triflupromazine as a function of pH. A) $\pi/\log C$ plots at pH 4.5 (stars), pH 6.0 (triangles), pH 7.4 (circles), pH 8.0 (pentagons), pH 8.5 (hexagons), and pH 9.0 (squares). B) Cross-sectional area (squares), and critical micelle concentration (triangles). C) Free energy of partitioning into the air–water interface, ΔG_{aw} (squares), and free energy of micelle formation, ΔG_{mic} (diamonds). The solid and dashed lines in B and C are sigmoidal fits to the data.

$$\Delta G_{aw} - \Delta G_{mic} = \Delta\Delta G_{am} \qquad (19)$$

As seen in Figure 2C, the amphiphilicity of triflupromazine is largest at low pH.

Figure 3A displays the $\pi/\log C$ plots of 6- and 7-trifluoromethyl benzopyranol at pH 8.0. The two compounds differ distinctly with respect to the slopes of the quasilinear part of the

Gibbs adsorption isotherms and thus with respect to the cross-sectional areas, A_D, as seen in Table 2. Moreover, they differ with respect to their amphiphilicity, 7-trifluoromethyl benzopyranol ($\Delta\Delta G_{am} = -7.73$ kJ mol^{-1}) is more amphiphilic than 6-trifluoromethyl benzopyranol ($\Delta\Delta G_{am} = -6.94$ kJ mol^{-1}). This can be rationalized by calculating the sum of the vectors of amphiphilicity. These were calculated for a set of multiple conformers from which the conformer with the highest amphiphilic moment was selected.[21] Figure 3B displays the calculated vectors of amphiphilicity,[21] which are $\Delta\Delta G_{am} = -7.41$ and -5.54 kJ mol^{-1}, respectively; this is in reasonable agreement with the measured amphiphilicities.[9] Here the amphiphilicities are given for $T = 24$ °C.

Lipid–water partition coefficient

Knowledge of K_{aw} and A_D allows calculation of K_{lw} for membranes with a specific π_M according to Equation (7). For the present calculations we used $\pi_M = 27$ and 35 mN m^{-1}, for SUVs formed from POPC at 37 °C[5, 10] or cholesterol-containing natural membranes, respectively.[6] Calculated lipid–water partition coefficients are summarized in Table 2. For comparison, K_{lw}'s of the phenothiazine analogues were also measured by means of ITC. As an example, the titration of a fluphenazine solution by SUVs formed from POPC is displayed in Figure 4A. The titration of all phenothiazines gave rise to exothermic titration patterns. With increasing lipid concentration, the free-drug concentration in the measuring cell and the heat flow decreased concomitantly. Figure 4B shows the heats of reaction, h_i, obtained by integration of the heat-flow peaks. The molar binding enthalpy ΔH^0_{exp} was determined directly from the cumulative heat release.[10, 17]

Permeability coefficients

The permeability coefficients, P, were calculated according to Equation (17) for membranes with lateral packing densities of $\pi = 27$ and 35 mN m^{-1} by using the pK_a values of the drugs at 37 °C.[23] The data are summarized in Table 2.

Discussion

In the following, we discuss the effects of drug halogenation on a drug's ability to permeate membranes. We place special emphasis on analyzing the difference between replacing a hydrogen residue by a chlorine or a trifluoromethyl residue in phenothiazine analogues (Table 1 series A–C). Moreover, we analyze the influence of the position of a $-$CF$_3$ residue in benzopyranol analogues (series D). To this purpose, we characterized twelve different drugs by means of SAMs. Measurements of the surface pressure as a function of concentration (Gibbs adsorption isotherms) yield i) the cross-sectional area, A_D, of the molecule when it is oriented in the amphiphilic gradient of the air–water interface or the lipid–water interface, ii) K_{aw}, which mainly reflects the hydrophobicity of the compound, and iii) CMC$_D$, which reflects the tendency of the compound to
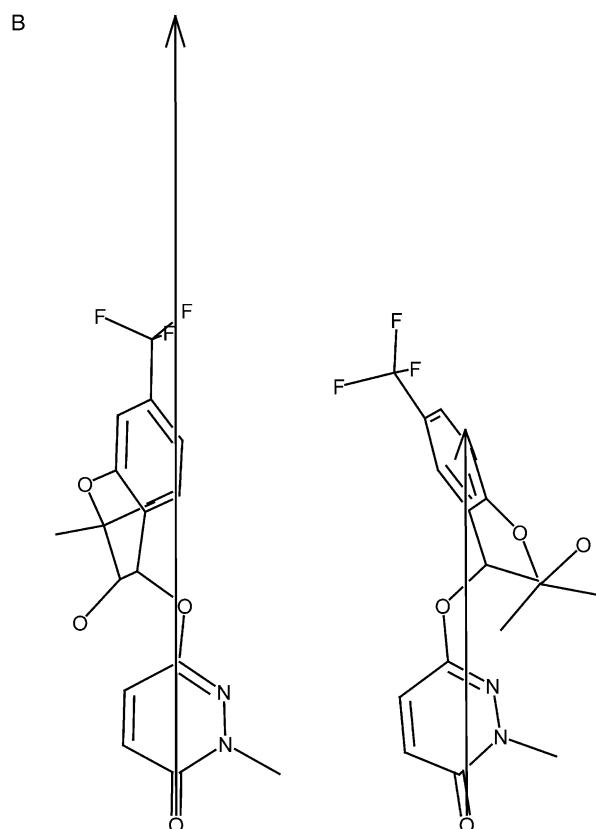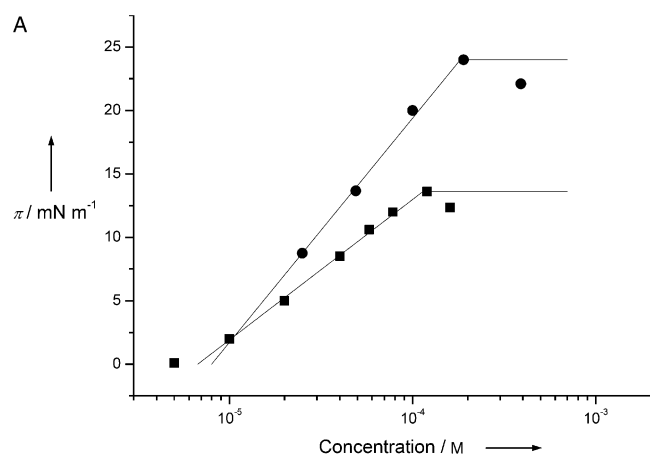
A

B

Figure 3. A) Gibbs adsorption isotherms (π/logC plots) of 6-trifluoromethyl ben-zopyranol (squares) and 7-trifluoromethyl benzopyranol (circles) measured at pH 8.0 (50 mm Tris/HCl containing 114 mm NaCl). B) Calculation of the vector of amphiphilicity for 6- (right) and 7-trifluoromethyl benzopyranol (left) with the program CAFCA.[21] The vector addition starts from the most hydrophilic residue (the oxygen atom of the pyridazin-3-one moiety was taken as an initial point) and points towards the most hydrophobic region of the molecule. The direction of the vector indicates the most probable orientation of the molecule in the amphiphilic gradient of the air–water interface. Conformer selection was performed according to the procedure described previously.[21] Briefly, vectors of amphiphilicity were calculated for a set of multiple conformers from which the conformer with the highest amphiphilic moment was selected.

self associate in solution. K_{lw} and P are calculated on the basis of the data obtained from SAMs as outlined previously.[10, 14]

The effect of halogenation is discussed on three levels: the first level deals with drugs in solution and at the air–water interface, which are characterized by the ionization constant,
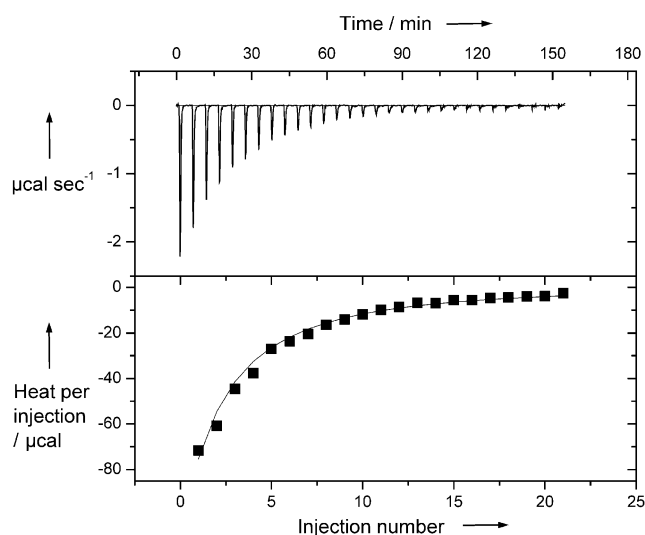

Figure 4. Isothermal titration calorimetry of fluphenazine. The fluphenazine sol-ution (101.88 μM) was contained in the measuring cell of a calorimeter and SUVs (10.1 mM) were injected (4 μL at each injection). Measurements were per-formed in buffer solution (50 mM Tris, 114 mM NaCl) at pH 7.4 and a tempera-ture of 37°C. Top: Titration curve. Bottom: Heats of reaction, h_i. The solid line is the theoretical binding isotherm calculated according to the Gouy–Chapman theory.

pK_a, the critical micelle concentration, and the air–water parti-tion coefficient. The second level deals with the thermodynam-ics of interfacial membrane partitioning, and the third level with the kinetics of membrane permeation.

Drugs in solution and at the air–water interface

Ionization constants of the drugs in aqueous solution (standard values) are given in Table 2. For the promazine analogues, the pK_a values decrease with increasing electronegativity of residue R^1 in the order H > Cl > CF₃; this is in agreement with previous observations.[22] For the perazine and perphenazine analogues this effect is less pronounced.

Apparent ionization constants depend on many factors such as temperature,[23] the dielectric constant, ε, of the environ-ment, and the association state of the drug.[24] The pK_a values of a drug at the air–water interface and in a drug micelle differ from that of a drug in solution since neighboring charged groups influence each other.[24] The ionization constant of a drug at the air–water interface under conditions in which only half of the air–water interface is occupied ($\Gamma = \Gamma_\infty/2$) is ob-tained from the pH dependence of ΔG_{aw} (Figure 2C). The ioni-zation constant under conditions in which the air–water inter-face is fully occupied ($\Gamma = \Gamma_\infty$) is obtained from the pH depen-dence of surface area requirement, A_S (Figure 2B). For triflupro-mazine the respective values are $pK_{a(\Gamma_\infty/2)} = 8.5$ and $pK_{a(\Gamma_\infty)} = 6.7$. From these two values it is possible to estimate the pK_a value of the monomer in solution as outlined previously.[24] An inter-mediate value is obtained for the pH dependence of critical micelle concentration, CMC_D (Figure 2B and C) ($pK_{a(CMC)} = 8.2$). This is consistent with small, highly curved vesicles for which the splay of the head groups is larger than in a tightly packed

Table 2. *Data obtained from surface-activity measurements and isothermal titration calorimetry.*

Series	Compound	M_W (base) [g mol^{-1}]	pK_a	pK_a 37°C	A_D[a] [Å^2]	K_{aw} [mM^{-1}]	CMC_D [mM]	$\Delta\Delta G_{am}$ [kJ mol^{-1}]	$K_{lw\,calc}$ [mM^{-1}] (27 mN m^{-1})	$K_{lw\,ITC}$ [mM^{-1}]	$D \times 10^8$ [cm^2 s^{-1}]	P [cm s^{-1}] (27 mN m^{-1})	P [cm s^{-1}] (35 mN m^{-1})
A	Promazine	284.42	9.42[27]	9.1	42.0	15	2.86	−9.69	1.20	1.55[10]	5.38	11	5
	Chlorpromazine	318.86	9.2[28]	8.9	40.0	20	0.22	−3.78	1.61	2.34[10]	6.36	26	12
	Triflupromazine	352.42	9.07[29]	8.7	50.0	129	0.16	−7.88	5.52	5.10[10]	5.34	102	40
B	Perazine	339.5	8.01[30]	7.7	42.0	30	0.83	−8.32	2.14	2.20	5.53	184	84
	Chlorperazine	373.94	8.1[31]	7.8	41.1	61	0.21	−6.60	4.53	4.53	6.28	367	170
	Trifluoperazine	407.5	8.08[32]	7.8	57.4	609	0.08	−10.11	15.71	17.60[10]	5.29	1088	368
C	Chlorperphenazine	403.97	7.9[29]	7.6	50.3	105	0.11	−6.35	4.42	4.50	5.67	377	147
	Fluphenazine	437.52	8.1[31]	7.8	55.4	520	0.10	−10.22	15.78	14.00	5.41	1038	390
	cis-flupenthixol	434.52	7.8[33]	7.5	63.0	1081	0.05	−10.47	20.31	20.00[10]	5.07	1633	503
	trans-flupenthixol	434.52	7.8[33]	7.5	66.0	4504	0.032	−12.8	70.02	24.00	4.95	5503	1602
D	6-trifluoromethyl benzopyranol	370.11	2.61	2.4	85.4	138	0.12	−7.24	0.63		4.35	58	12
	7-trifluoromethyl benzopyranol	370.11	2.61	2.4	53.9	121	0.19	−8.07	4.04		5.48	465	1670

[a] For the highly charged molecules promazine and triflupromazine the minimum area, A_D, was extrapolated to pH 8.5. Data given represent average values from several measurements. Maximum error range is ±5%. A_D is the cross-sectional area, CMC_D, the critical micelle concentration, and K_{aw} the air–water partition coefficient [Equation (5)]. K_{lw} is the lipid–water partition coefficient which was either calculated according to Equation (7), $K_{lw\,calc}$, or measured by means of isothermal titration calorimetry, $K_{lw\,ITC}$. D is the diffusion coefficient [Equation (14)], and P, the permeability coefficient [Equation (17)]. SAM and ITC were performed at 24 ± 1°C and 37°C, respectively. For simplicity all ΔG values were calculated at 37°C.

planar drug layer at the air–water interface but smaller than that at a half-occupied air–water interface ($\Gamma_\infty/2$).

Shifts of ionization constants to lower values are also observed if drugs insert into the lipid–water interface.

The cross-sectional areas of the phenothiazine analogues in series A–C vary between $A_D = 40$ and 58 Å^2. Only the flupenthixols exhibit slightly larger cross-sectional areas; this might be due to the rigidity of the double bond in residue R^2. Replacement of a H atom by a Cl atom does not, on average, lead to a significant increase in A_D, in contrast to a replacement by a −CF$_3$ group, which leads to a measurable increase. The A_D ratios for the different analogues vary between $r = 1.0$ and $r = 1.4$ and are given in Table 3.

Measurement of the two analogues of series D, 6- and 7-trifluoromethyl benzopyranol, revealed an A_D ratio $r = 1.6$ (Table 3). The change in the cross-sectional area upon changing the position of the CF$_3$ group is thus due to a change in the orientation of the vector of amphiphilicity drawn from the most hydrophilic residue R^2, immersed in aqueous solution ($\varepsilon \sim 80$), to the most distant hydrophobic residue R^1, reaching into the air ($\varepsilon \sim 1$), as illustrated in Figure 3B.[21] Since the dielectric constant of air is similar to that of lipids ($\varepsilon \sim 2$), it can be assumed that the molecular orientation at the two interfaces is identical.

The air–water partition coefficient increases for compounds in series A–C in the order: H < Cl < CF$_3$ and reflects the increase in hydrophobicity of residues R^1. Replacement of H by Cl or CF$_3$ leads, on average, to an increase in K_{aw} of a factor of approximately $r = 2$ or $r = 14$, respectively, or to an increase in the negative free energy of partitioning into the air–water interface of −1.5 or −6.6 kJ mol^{-1}, respectively. The effect of hal-

Table 3. *Ratios, r, of cross-sectional areas, air–water partition coefficients, critical micelle concentrations, lipid–water partition coefficients (calculated from SAMs), and permeability coefficients for analogues with and without halogen residues.*

Residue R^1	Series	rA_D	rK_{aw}	rK_{lw}	rP (27 mN m^{-1})	rP (35 mN m^{-1})
Cl/H	A	1	1.34	1.3	2.40	2.50
	B	1	2.0	2.13	1.99	2.02
CF$_3$/H	A	1.25	8.62	4.48	9.56	8.24
	B	1.38	20.13	7.08	5.91	4.38
CF$_3$/Cl	A	1.25	6.45	2.85	3.98	3.30
	B	1.4	10.04	2.53	2.97	2.17
	C	1.1	4.93	3.25	2.92	2.65
CF$_3$ *trans*/*cis*	C	1.05	4.17	3.45	3.37	3.19
6-CF$_3$/7-CF$_3$	D	1.59	1.15	0.45	0.12	0.07

ogenation is somewhat more pronounced in perazine than in promazine analogues. This may be due to the fact that the effect of charge still dominates in the latter analogues. The air–water partition coefficients of the two benzopyranol analogues in series D are practically identical.

The free energy of self-association or micelle formation, ΔG_{mic}, is significantly enhanced by the replacement of a H by a Cl or a CF$_3$ residue. Surprisingly, at first, the difference between the two halogen residues is small.

The amphiphilicity, $\Delta\Delta G_{am}$, which is the difference between the free energy of partitioning into the air–water interface, ΔG_{aw} and the free energy of micelle formation, ΔG_{mic},[9] increases with increasing charge of the compound at constant hydrophobicity or with increasing hydrophobicity at constant charge (Table 4). For the drugs under study $\Delta\Delta G_{am}$ increases in

Table 4. *Difference in free energies of air–water partitioning, micelle formation or self-association, and membrane partitioning of analogues with and without halogen residues calculated from SAMs.*

Residue R¹	Series	ΔG_{aw} [kJ mol⁻¹]	ΔG_{mic} [kJ mol⁻¹]	ΔG_{lw} [kJ mol⁻¹]
Cl/H	A	−0.75	−6.66	−0.75
	B	−1.79	−3.51	−1.94
CF₃/H	A	−5.55	−7.36	−3.93
	B	−7.74	−5.95	−5.14
CF₃/Cl	A	−4.81	−0.70	−3.18
	B	−5.95	−2.44	−3.20
	C	−4.11	−0.24	−3.28
CF₃ trans/cis	C	−3.68	−1.35	−3.19
7-CF₃/6-CF₃	D	−0.35	−1.18	−4.79

the order: $Cl < H < CF_3$. The comparatively low critical micelle concentrations and amphiphilicities of the chlorinated analogues in series A, B and C is most probably due to the relatively strong reduction in pK_a values combined with a negligibly small increase in hydrophobicity upon chlorination, and explains the comparatively high tendency of these analogues to aggregate in solution.

Interfacial membrane partitioning

The lipid–water partition coefficients were calculated according to Equation (7) by using the K_{aw} and A_D of the compound measured under conditions of minimal electrostatic repulsion for membranes with a packing density, $\pi_M = 27$ and 35 mN m⁻¹, corresponding to that of small unilamellar vesicles formed from POPC at physiological temperature and that of cholesterol-containing membranes, respectively. As seen in Table 2, the lipid–water partition coefficients calculated for a membrane packing density of $\pi_M = 27$ mN m⁻¹ are in excellent agreement with those measured for small unilamellar POPC vesicles by means of ITC.[10] The lipid–water partition coefficients for the promazine analogues have also been measured by means of spectrophotometric techniques and are also in good agreement (if transformed to the same units).[26]

The lipid–water partition coefficients of the compounds in series A–C increase in the order of residue R¹: $H > Cl > CF_3$. The exchange of a H to Cl or CF₃ leads on average to an increase in the negative free energy of membrane partitioning of $\Delta G_{lw} = -1.5$ or -4.5 kJ mol⁻¹, respectively. The dominant factor is the increase in hydrophobicity. For compounds in series D that exhibit similar hydrophobicities, the lipid–water partition coefficients are dominated by the different cross-sectional areas, A_D (Table 3).

Membrane permeation

On a third level, the kinetics of passive diffusion through the lipid membrane are calculated on the basis of simple Stokesian diffusion by using the parameters obtained from surface-activity measurements and taking into account the pK_a value of the drugs. For a replacement of H by a Cl or a CF₃ residue, the permeability coefficient, P, increases on average by a factor of approximately $r = 2$ or $r = 9$, respectively. The increase is again somewhat larger for promazine than for perazine analogues. Despite the differences between promazines and perazines, the increase in P upon replacement of Cl by CF₃ is rather constant for all three types of analogues (promazines, perazines, and perphenazines) and amounts to about $r = 3.5$ (Table 3). It is interesting to note that the two isomers, *cis*- and *trans*-flupenthixol, differ distinctly in their K_{aw} and K_{lw} as well as in P. Due to the relatively small cross-sectional areas of the above analogues, the packing density dependence of P is relatively small for the phenothiazine analogues. This is different for 7- and 6-trifluoromethyl benzopyranol, for which the ratios in permeability coefficients are $r = 0.12$ at the lower packing density investigated and only $r = 0.07$ at the higher.

Conclusion

A characterization of drugs by SAMs allows a detailed analysis of the effect of halogenation. The air–water and the lipid–water partition coefficients of promazine, perazine, and perphenazine analogues increase in the order, R¹: $H < Cl < CF_3$ due to the increase in hydrophobicity, despite a small increase in cross-sectional area. The permeability coefficient increases in the same order due to the increase in the lipid–water partition coefficient and the decrease in the pK_a values. For the small phenothiazine analogues, the packing density dependence of the permeability coefficient is rather small. The amphiphilicity of the halogenated analogues increases in the order R¹: $Cl < H < CH_3$; this explains the higher tendency of the chlorinated analogues to aggregate. As shown for benzopyranols, the position of a trifluoromethyl residue can change the amphiphilicity, the cross-sectional area, and, as a consequence, the permeability coefficient of a molecule.

Biological situation

The pH close to the surface of a living cell is generally acidic despite the fact that the extracellular lipid leaflet is electrically neutral. If one takes into account an acidic pH and the correspondingly low air–water partition coefficient (cf. Figure 2A) permeability coefficients can be orders of magnitude lower than those given in Table 2.

Even compounds with low permeability coefficients can in principle cross a membrane, provided the time to reach equilibrium is given. In natural environments however, equilibration time is limited by metabolic processes—for example, the action of cytochrome P450 and ATP-driven efflux transporters, such as P-glycoprotein, which bind molecules within the lipid membrane and export them out of the cell. If the export rate

of a drug is faster than the rate of passive diffusion into the cell (influx), the drug will barely reach the cytosol. However, if passive influx is distinctly faster than active efflux, the drug will reach the cytosol even if it is a substrate for an efflux transporter.[14] The simple permeability predictions on the basis of SAMs provide an estimate of rates of passive influx of drugs and allow for comparison with the rates of efflux processes.[14] The present approach opens new possibilities for a detailed understanding of membrane permeation in biological systems.

[1] A. Seelig, J. Seelig in *Encyclopedia of Physical Science and Technology*, 3rd ed. (Ed.: R. A. Meyers), Academic Press, N.Y., **2002**.

[2] J. Seelig, A. Seelig, *Q. Rev. Biophys.* **1980**, *13*, 19.

[3] D. Marsh, *Biochim. Biophys. Acta* **1996**, *1286*, 183.

[4] A. Seelig, *Biochim. Biophys. Acta* **1987**, *899*, 196.

[5] A. Seelig, *Biochemistry.* **1992**, *31*, 2897.

[6] R. A. Demel, W. S. Geurts van Kessel, R. F. Zwaal, B. Roelofsen, L. L. van Deenen, *Biochim. Biophys. Acta* **1975**, *406*, 97.

[7] V. Boguslavsky, M. Rebecchi, A. J. Morris, D. Y. Jhon, S. G. Rhee, S. Mclaughlin, *Biochemistry.* **1994**, *33*, 3032.

[8] F. Hanakam, G. Gerisch, S. Lotz, T. Alt, A. Seelig, *Biochemistry.* **1996**, *35*, 11 036.

[9] H. Fischer, R. Gottschlich, A. Seelig, *J. Membr. Biol.* **1998**, *165*, 201.

[10] X. Li-Blatter, E. Gatlik-Landwojtowicz, A. Seelig, unpublished results.

[11] F. A. Gobas, J. M. Lahittete, G. Garofalo, W. Y. Shiu, D. Mackay, *J. Pharm. Sci.* **1988**, *77*, 265.

[12] R. Bergmann, V. Eiermann, R. Gericke, *J. Med. Chem.* **1990**, *33*, 2759.

[13] A. Seelig, R. Gottschlich, R. M. Devant, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 68.

[14] A. Seelig, E. Gatlik-Landwojtowicz, *Mini-Rev. Med. Chem.* **2004**, in press.

[15] P. Fromherz, *Rev. Sci. Instrum.* **1975**, *46*, 1380.

[16] E. Gatlik-Landwojtowicz, X. Li-Blatter, A. Seelig, unpublished results.

[17] M. R. Wenk, J. Seelig, *Biophys. J.* **1997**, *73*, 2565.

[18] J. Seelig, S. Nebel, P. Ganz, C. Bruns, *Biochemistry.* **1993**, *32*, 9714.

[19] S. Mclaughlin, *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 113.

[20] S. Mclaughlin, *Curr. Top. Membr. Transp.* **1977**, *9*, 71.

[21] H. Fischer, M. Kansy, D. Bur, *Chimia* **2000**, *54*, 640.

[22] J. E. True, T. D. Thomas, R. W. Winter, G. L. Gard, *Inorg. Chem.* **2003**, *42*, 4437.

[23] R. F. Cookson, *Chem. Rev.* **1974**, *74*, 5.

[24] A. Seelig, *Biochim. Biophys. Acta* **1990**, *1030*, 111.

[25] K. Kitamura, S. Takegami, T. Kobayashi, K. Makihara, C. Kotani, T. Kitade, M. Moriguchi, Y. Inoue, T. Hashimoto, M. Takeuchi, *Biochim. Biophys. Acta* **2004**, in press.

[26] S. Takegami, K. Kitamura, T. Kitade, A. Kitagawa, K. Kawamura, *Chem. Pharm. Bull. (Tokyo).* **2003**, *51*, 1056.

[27] P. Seiler, *Eur. J. Med. Chem.* **1974**, *9*, 473.

[28] F. H. Clarke, *J. Pharm. Sci.* **1984**, *73*, 226.

[29] U. Franke, A. Munk, M. Wiese, *J. Pharm. Sci.* **1999**, *88*, 89.

[30] R. Mannhold, K. P. Dross, R. F. Rekker, *Quant. Struct-Act. Relat.* **1990**, *9*, 21.

[31] D. W. Newton, R. B. Kluza, *Drug Intell. Clin. Pharm.* **1978**, *12*, 546.

[32] F. H. Clarke, N. M. Cahoon, *J. Pharm. Sci.* **1987**, *76*, 611.

[33] J. P. Tollenaere, H. Moereels, M. H. J. Koch, *Eur. J. Med. Chem.* **1977**, *12*, 199.

Expert Opinion

Enhancement of drug absorption by noncharged detergents through membrane and P-glycoprotein binding

Anna Seelig[†] & Grégori Gerebtzoff

[†]*Biozentrum, Biophysical Chemistry, University of Basel, Klingelbergstrasse 70, CH-4057 Basel, Switzerland*

Noncharged detergents are used as excipients in drug formulations. Until recently, they were considered as inert compounds, enhancing drug absorption essentially by improving drug solubility. However, many detergents insert into lipid membranes, although to different extents, and change the lateral packing density of membranes at high concentrations. Moreover, they bind to the efflux transporter P-glycoprotein (P-gp) and most likely to related transporters and metabolising enzymes with overlapping substrate specificities. If their affinity to P-gp is higher than that of the coadministered drug they act as modulators or inhibitors of P-gp and enhance drug absorption. Inhibition of P-gp and related proteins can, however, cause severe side effects. This paper first reviews the membrane binding propensity of different noncharged detergents (including poloxamers) and discusses their ability to bind to P-gp. Second, literature data on drug uptake enhancement by noncharged detergents, obtained *in vivo* and *in vitro*, are analysed at the molecular level. The present analysis provides the tools for an approximate and simple prior estimate of the membrane and P-gp binding ability of noncharged detergents based on a modular binding approach.

1. Introduction

The first observation of the chemosensitising effect of detergents dates probably back to 1972 when Riehm and Biedler [1] noted that Tween 80 enhanced the effect of actinomycin D and daunomycin in drug-resistant cells. In the early 1990s several groups [2-9] reported that nonionic detergents, commonly used as excipients in drug formulations, modulated the activity of P-glycoprotein (P-gp, multi-drug resistance protein MDR1, ABCB1), an ATP-binding cassette (ABC) efflux transporter of broad substrate specificity (for review see [10,11]). P-gp is highly expressed in different tissues with protective functions such as the blood–brain barrier (BBB) [12], liver, pancreas, kidney, colon and the intestinal barrier (IB) [13]. This tissue distribution indicates that P-gp prevents the accumulation of xenobiotics and metabolites in the brain and plays a significant role in their excretion into urine, bile and the intestinal lumen (for review see [14]). High levels of P-gp are also observed in certain cancer cells. In a given tissue the expression level of efflux transporters can vary depending on genetic predisposition, age, diet and medication [15] (and references therein).

P-gp is the best-investigated representative of a large family of 49 ABC transporters identified so far [16], which all play a significant role in ADME [17]. Many ABC transporters show overlapping substrate specificities. A detailed analysis has been
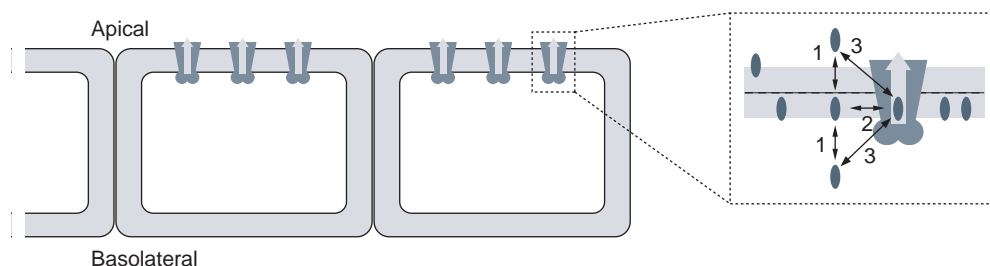
informa
healthcare

Figure 1. Cartoon showing the apical location of P-glycoprotein in the intestinal barrier, where it pumps its substrates into the intestinal lumen removing them from the body, or in the blood–brain barrier, where it pumps its substrates into the blood, removing them from the brain. Substrate binding to P-glycoprotein takes place in the cytosolic membrane leaflet. Substrate affinity to the transporter from the aqueous phase (expressed as free energy of transporter–water binding, ΔG^0_{tw}; arrow 3) is the sum of the substrate affinity to the lipid membrane (expressed as free energy of lipid–water partitioning, ΔG^0_{lw}; arrow 1) and the drug affinity to the transporter in the lipid membrane (expressed as free energy of transporter–lipid binding, ΔG^0_{tl}; arrow 2).

performed for P-gp and the multi-drug resistance-associated protein, MRP1 [18], but much remains to be clarified in this field. P-gp transports cationic or electrically neutral compounds, whereas MRP1 transports electrically neutral or negatively charged compounds. Substrates common to both transporters are electrically neutral compounds that carry specific hydrogen bond acceptor patterns. As shown in this review, many neutral detergents have such structural elements and, therefore, react most likely with both transporters and possibly even with other ABC transporters. Overlapping substrate specificity has also been observed for P-gp and the metabolising enzyme CYP [19].

P-gp binds its substrates in the cytosolic lipid leaflet of the plasma membrane and flips them to the outer leaflet or exports them to the extracellular medium at the expense of ATP hydrolysis [20]. To reach the P-gp binding region exogenous substrates first partition into the membrane and then cross the lipid bilayer, whereas endogenous substrates directly insert into the cytosolic membrane leaflet as illustrated in **Figure 1**. Substrate binding to P-gp from the aqueous phase can be divided in two consecutive steps: a lipid–water partitioning step and a transporter–lipid binding step. The free energy of transporter–water binding, ΔG^0_{tw} (reflecting the binding affinity of the substrate to the transporter from the aqueous phase) can, therefore, be expressed as sum of the free energy of lipid–water partitioning, ΔG^0_{lw} (reflecting the binding affinity of the substrate for the lipid membrane), and the free energy of transporter–lipid binding, ΔG^0_{tl} (reflecting the affinity of the substrate to the transporter after insertion into the lipid membrane) [21].
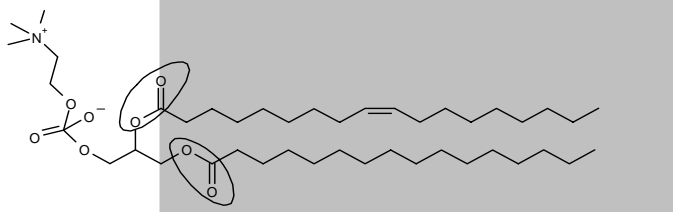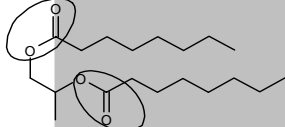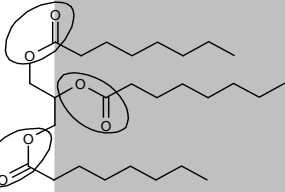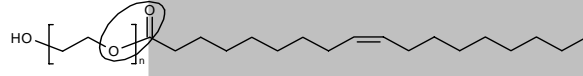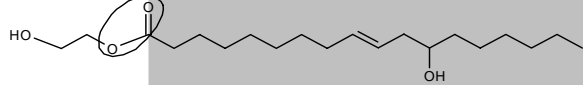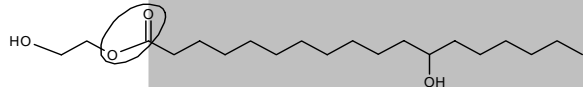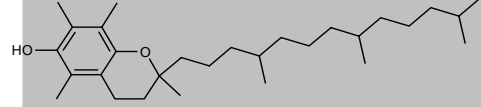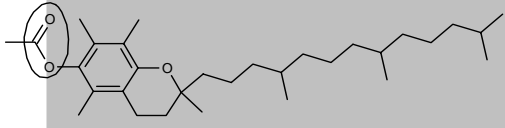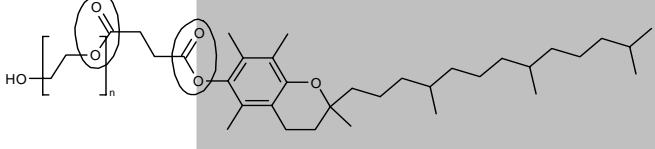
P-gp inhibition by detergents has been widely discussed, and was assumed to result either from a direct detergent–P-gp interaction [2,4,7,22] or from membrane fluidisation leading to an indirect protein destabilisation [5,6,23]. It should be noted that not only detergents (see Section 3.2), but also many drugs that are substrates for P-gp, such as verapamil [24] or dibucaine [21,25], disorder membranes. To fully understand the effect of

membrane fluidisation on P-gp, further experiments are required. Rege *et al.* have shown that the detergents Tween 80 and Cremophor EL, which fluidise lipid bilayers, as well as Vitamin E D-α-tocopheryl poly(ethylene glycol) 1000 succinate (TPGS), which rigidifies lipid bilayers, inhibit multiple efflux transporters, and concluded that membrane fluidisation may not be a general mechanism to reduce transporter activity [26]. Recently, it was observed that detergents exhibiting the characteristics of P-gp substrates or modulators [15] activate P-gp in a concentration-dependent manner in contrast to detergents such as CHAPS, which are nonsubstrates [27]. This further supports a direct interaction with P-gp.

Numerous *in vitro* and *in vivo* investigations on the enhancement of drug absorption by noncharged detergents have been published (see references cited in **Table 3**). The experiments were performed with different combinations of drugs and detergents either *in vivo* or *in vitro* using different cellular systems. At first sight, the resulting data seem complex; however, a detailed analysis shows a consistent and rational picture of detergent–membrane and detergent–P-gp interactions.
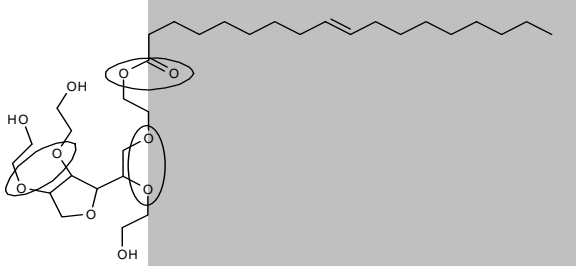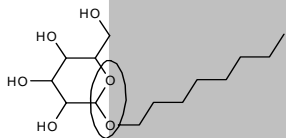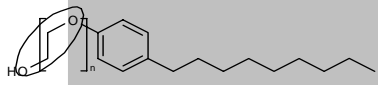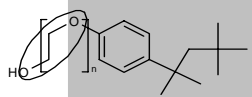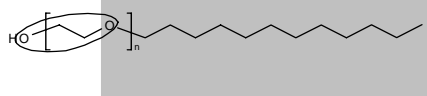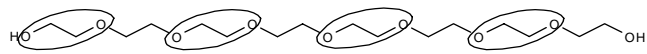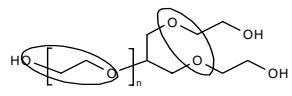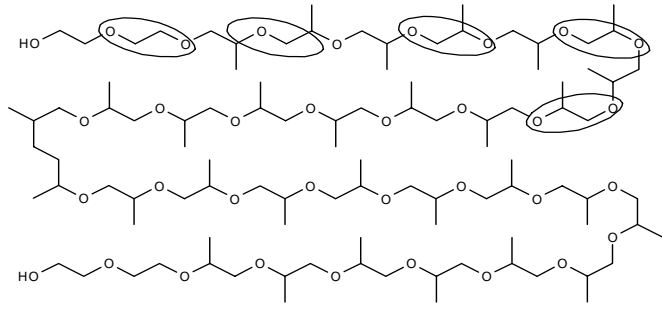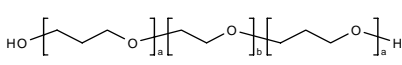
This review analyses the available data on absorption enhancement of drugs by noncharged detergents to unmask the relevant molecular processes. For this purpose: i) the non-charged detergents, comprising membrane-anchoring detergents and poloxamers; ii) the membranes used as model systems; and iii) the transporter P-gp are characterised. This review also addresses: i) whether the selected detergents insert into the lipid membrane and ii) whether they subsequently bind to P-gp. The analysis reveals that most of the selected membrane-anchoring detergents (**Table 1**) are modulators of P-gp and are able to enhance the absorption of the coadministered drug, provided their affinity to P-gp is higher than that of the coadministered drug. Under these conditions the detergents inhibit (or modulate) the P-gp efflux pump, thus allowing the drugs to cross the membrane by passive diffusion. Specific poloxamers, which do not insert into biological membranes, may enhance absorption by mere solubilisation

Table 1. Noncharged detergents as potential P-glycoprotein substrates

No.	Structure	Name
1		POPC (26853-31-6)
2		1,2-Dicaprylin (7384-98-7)
3		Tricaprylin Miglyol 808 (538-23-8)
4		PEG oleate (9004-96-0)
5		Cremophor EL (9004-97-1) (9006-38-6)
6		Solutol HS-15 (61909-81-7)
7		Vitamin E (59-02-9)
8		Vitamin E acetate (7695-91-2)
9		Vitamin E TPGS (9002-96-4)

Ellipses indicate type I units, grey background indicates nonpolar and white background polar environment, respectively.
POPC: 1-Palmitoyl-2-oleoyl-*sn*-3-glycerophosphocholine; TPGS: D-α-Tocopheryl poly(ethylene glycol) 1000 succinate.

Table 1. Noncharged detergents as potential P-glycoprotein substrates (continued)

No.	Structure	Name
10		Tween 80 (9005-65-6)
11		Octyl-β-D-glucoside (29836-26-8)
12		(26027-38-3)
13		n = 1: Triton X-100 (9002-93-1) n = 2: Nonidet P-40 (9036-19-5)
14		n = 1: 9002-92-0 n = 3: 2315-62-0 n = 8: 35056-97-4
15		Pentaethylene glycol (25322-68-3)
16		Polyoxyethylene glycerol ether (31694-55-0)
17		Pluronic L61
18		CRL-1605 (a = 40, b = 20)

Ellipses indicate type I units, grey background indicates nonpolar and white background polar environment, respectively.

POPC: 1-Palmitoyl-2-oleoyl-sn-3-glycerophosphocholine; TPGS: D-α-Tocopheryl poly(ethylene glycol) 1000 succinate.

of the compounds or, in combination with highly cationic drugs, by endocytosis. The analysis further shows that the severe side effects that have been observed with several detergents at high concentrations can also be traced back to a direct interaction of detergents with P-gp, related ABC transporters and CYP.

2. Noncharged detergents

Table 1 displays the most abundant lipid in mammalian cell membranes, 1-palmitoyl-2-oleoyl-*sn*-3-glycerophosphocholine (POPC; compound **1**), together with some of the most frequently used noncharged detergents. The detergents are divided in two sets: a first one, comprising compounds **2 – 14**; and a second one, comprising compounds **15 – 18**. The compounds in the first set exhibit a hydrophobic tail, which is responsible for membrane anchoring (shown on grey background), and a hydrophilic headgroup (shown on white background). The latter generally consist either of ester (lactone) or ether groups or combinations of the two. The numerous oxygen atoms can act as hydrogen bond acceptors and are probably responsible for hydrogen bonding interactions with P-gp, as in Section 4.

The second set of compounds comprises polyethylene glycols (PEGs) and polyoxyethylene–polyoxypropylene block copolymers, the so-called poloxamers or Pluronics®, which lack a typical membrane anchor. Poloxamers are formed from alternating sequences of polyoxyethylene $(EO)_a$ and polyoxypropylene $(PO)_b$, in which the length of the different blocks (indicated by the subscripts a and b) can vary considerably. They usually consist of a central hydrophobic polyoxypropylene $(PO)_a$ block and two flanking hydrophilic polyoxyethylene $(EO)_b$ blocks. Typical examples are the Pluronics L61 $(EO_2\text{-}PO_{30}\text{-}EO_2)$, L81 $(EO_3\text{-}PO_{43}\text{-}EO_3)$, P85 $(EO_{26}\text{-}PO_{40}\text{-}EO_{26})$ (or poloxamer 235) and F68 $(EO_{76}\text{-}PO_{30}\text{-}EO_{76})$ (or poloxamer 188). Poloxamers with a hydrophilic centre and hydrophobic flanking regions such as CRL-1605 $(PO_{40}\text{-}EO_{20}\text{-}PO_{40})$ are also available (for further examples of Pluronics see [28]).

3. The biological membranes and membrane model systems

Lipid molecules, when brought into contact with water, spontaneously organise themselves into a bilayer leaflet, with the polar lipid headgroups remaining in the aqueous environment and the fatty acid tails forming the inner hydrophobic core. The fluid (liquid crystalline) bilayer is the main organisational element in biological membranes and behaves like optically uniaxial crystals with the optical axis perpendicular to the surface of the membrane. The anisotropic character of membranes is not reflected by isotropic organic solvents such as octanol, which are often used as membrane model systems.

Using solid-state NMR techniques a quantitative analysis of the molecular ordering and dynamics of a lipid bilayer is possible with a segment-to-segment resolution. Well-defined conformations are observed for the glycerol backbone and, to some extent, also for the polar headgroups. The packing of the fatty acyl chains can be 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 headgroup region and decreases towards the centre of the membrane. The shape of the order profiles are similar for different membranes; however, the absolute values differ and depend on the lipid composition. An increase in membrane order is observed on insertion of cholesterol or replacement of phosphatidylcholine headgroups by the smaller phosphatidylethanolamine headgroups. A disordering is observed on incorporation of double bonds into the fatty acyl chains or an increase in temperature. However, the incorporation of transmembrane proteins barely influences the order profiles of lipid bilayers [29,30].

A lipid monolayer exhibits the same anisotropic organisation as a lipid bilayer, provided the lipids are compressed to the same lateral packing density. The monolayer–bilayer equivalence pressure, reflecting the lateral packing density of a particular lipid bilayer, π_M, was assessed for several biological membranes and membrane model systems (see **Table 2**). The highest lateral packing densities were observed for mixed bilayers containing cholesterol and high levels of *Escherichia coli* lipids, which are rich in phosphatidylethanolamines (π_M = 40.5 – 46.5 mN/m at 37°C). Plasma membranes of mammalian cell lines, such as Chinese hamster ovary cells, CR1R12, murine embryo fibroblasts, NIH-MDR1-G185, and porcine kidney cells, LLC-MDR1, cover an intermediate range of lateral packing densities, π_M = 30 – 35 mN/m at 37°C [21]. Because MDR1 transfection seems not to alter the lipid composition [31], it can be assumed that the corresponding wild-type cells (NIH3T3 and LLC) exhibit similar lateral packing densities.

For the human BBB, the lateral packing density was estimated as $\pi_M \approx 35$ mN/m. Using Equation 1 (see Section 1) and taking into account the limiting molecular cross-sectional area, $A_D = 80$ Å² for the BBB [32] and $A_D = 100$ Å² for the IB [33], the lateral packing density of the human IB was calculated as $\pi_M \approx 28$ mN/m.

The lateral packing density of planar lipid membranes formed from POPC was determined as $\pi_M = 32 \pm 1$ mN/m at ambient temperature [34], and that of large (100 nm), and small (30 nm), unilamellar vesicles formed from POPC at 37°C was determined as $\pi_M = 28 \pm 1$ mN/m and $\pi_M = 25 \pm 1$ mN/m, respectively (Blatter and Seelig, unpublished results). The decrease in packing density observed with increasing vesicle curvature is due to the increasing splay of the lipid headgroups in the outer membrane leaflet. The lateral packing density of black lipid membranes depends on the content of short chain alkanes and is generally significantly lower than that of pure lipid bilayers ($\pi_M < 25$ mN/m).

In summary, biological membranes are highly organised, anisotropic systems exhibiting different lateral packing densities

Table 2. The packing density of biological membranes and lipid bilayers.

Lipid bilayer membrane	Temperature T [°C]	Packing density π_M [mN/m]	Ref.
Bilayers containing *Escherichia coli* lipids	37	46 ± 5	[21]
CR1R12 (Chinese hamster ovary cells)	37	~ 35	[21]
Blood–brain barrier	37	~ 35	[32]
Erythrocyte membrane	Ambient	31 – 35	[94]
LLC-MDR1 (porcine kidney cells)	37	~ 33	[21]
POPC (planar lipid bilayers)	Ambient	32 ± 1	[34]
NIH-MDR1-G185 ≈ NIH3T3 (murine embryo fibroblasts)	37	~ 30	[21]
DMPC bilayers	Ambient	~ 30	[95]
Intestinal barrier	37	~ 28	[33]
POPC (large unilamellar vesicles)	37	28 ± 1	[96]
POPC (small unilamellar vesicles)	37	25 ± 1	
Black lipid membranes/PAMPA	Ambient	< 25	

DMPC: Dimyristoylphosphatidylcholine; PAMPA: Parallel Artificial Membrane Permeation Assay; POPC: 1-Palmitoyl-2-oleoyl-*sn*-3-glycerophosphocholine.

of the lipid constituents. The lateral packing densities of biological membranes are often higher than those of the commonly used membrane model systems. Despite their anisotropic organisation biological membranes are fluid enough to allow considerable translational, rotational and flexing movements of the constituent lipid and protein molecules.

3.1 Determination of lipid–water partition coefficients

To illustrate the influence of the lateral membrane packing density for membrane partitioning, a method is discussed based on surface activity measurements. Measurements of the surface pressure, π, as a function of concentration (Gibbs adsorption isotherm) allows to quantitatively measure the tendency of a drug to move to the air–water interface. Because air has a similar dielectric constant as the hydrocarbon region of the lipid membrane, the air–water interface provides an ideal model for drug orientation at the lipid–water interface. The analysis of the Gibbs adsorption isotherm yields three characteristic parameters: i) the air–water partition coefficient, K_{aw}; ii) the surface area of the molecule at the air–water interface, A_D; and iii) the critical micelle concentration, *CMC*.

Whereas partitioning into the air–water interface at low concentrations is essentially determined by the amphiphilicity of the molecule, penetration of a drug between the hydrocarbon chains in a lipid bilayer requires additional energy, which is small for small molecules, but can become prohibitively high for molecules with large cross sections. To form a hole in a monolayer or bilayer, a penetrating substance with a cross-sectional area, A_D, has to perform work, ΔW, against the lateral packing density, π_M, of the lipids [35]:

$$\Delta W = A_D \pi_M \tag{1}$$

The lipid–water partition coefficient, K_{lw}, is thus given by:

$$K_{lw} = K_0 \cdot e^{-\pi_M A_D / kT} \tag{2}$$

where K_0 is a proportionality constant. Because the surface pressure of the drug monolayer at the concentration $1/K_{aw}$ is negligibly small, it was assumed that K_0 can be replaced by K_{aw} (note: approximately constant lateral packing density as assumed in Equation 1 is valid for low detergent concentrations only). This assumption was confirmed by measuring both K_{lw} and K_{aw} for a series of compounds [32,36]. With the parameters, K_{aw} and A_D, obtained from one single measurement of the Gibbs adsorption isotherm, the lipid–water partition coefficients, K_{lw}, for membranes of different lateral packing densities, π_M, can be calculated.

The influence of the membrane packing density, π_M, on the lipid–water partition coefficient, K_{lw}, is illustrated in the following examples. Decreasing the packing density from $\pi_M \approx 35$ mN/m (BBB) to $\pi_M \approx 28$ mN/m (IB) for a compound with a cross-sectional area, $A_D = 100$ Å^2 ($A_D = 50$ Å^2), leads to an increase in the lipid–water partition coefficient, K_{lw}, by a

factor of 5.4 (3.3). Decreasing the cross-sectional area from $A_D = 100$ Å² to $A_D = 50$ Å² for a membrane with a lateral packing density, $\pi_M = 35$ mN/m ($\pi_M = 28$ mN/m), leads to an increase of the lipid–water partition coefficient by a factor of ~ 70 (~ 20).

In contrast to the partitioning into lipid membranes, partitioning into isotropic solvents such as hexane or octanol has been shown to increase with the molecular volume of the compound [37]. Predicting the lipid–water partition coefficient without taking into account the different lateral packing densities of membranes can be misleading for molecules with large cross-sectional areas. Unfortunately, the air–water partition coefficient, K_{aw}, is not available for all detergents shown in **Table 1**. The lipid–water partition coefficients, K_{lw}, and the free energies of lipid–water partitioning, ΔG^0_{lw}, given in **Table 3** were, therefore, estimated on the basis of the *CMC* [38].

A detergent molecule in an aqueous phase containing lipid vesicles has two possibilities, either to associate with other monomers or to partition into the lipid membrane. The free energy of the two processes, that is, the free energy of micelle formation, ΔG^0_{CMC}, and the free energy of lipid–water partitioning, ΔG^0_{lw}, are similar. The free energy of lipid–water partitioning, ΔG^0_{lw}, can thus be estimated as follows:

$$(3)$$

$$\Delta G^0_{lw} \approx RT \ln(CMC/C_w)$$

where $C_W = 55.5$ mol/l is the molar concentration of water. The free energy of lipid–water partitioning, ΔG^0_{lw}, derived from *CMC*s according to Equation 3 is included in **Table 3**. For comparison, the calculated octanol–water partition coefficient LogP values (KowWin) are also given. For small compounds and 'soft' membranes ($\pi_M = 25 – 28$ mN/m) a reasonable linear correlation is observed between LogP and ΔG^0_{lw} values derived from *CMC* and the ΔG^0_{lw} values obtained from air–water partition coefficients, K_{aw} (not shown). However, for compounds with larger cross-sectional areas and densely packed membranes, the ΔG^0_{lw} values obtained from air–water partition coefficients, K_{aw}, are generally less negative.

3.1.1 Lipid–water partitioning of membrane-anchoring detergents

The lipid–water partition coefficients of several detergents have been measured for POPC vesicles (diameter of 100 nm, large unilamellar vesicles, $\pi_M \approx 28$ mN/m) using isothermal titration calorimetry. Depending on the nature of the polar and hydrophobic groups the lipid–water partition coefficients for POPC vesicles were shown to vary considerably from $K_{lw} = 10^5$/M for the oligo(ethylene oxide) alkyl ether, C_{12} EO$_3$, to only $K_{lw} = 25$/M for octylmaltoside [38]. For the present

investigation the lipid–water partition coefficients estimated on the basis of CMC values are summarised in **Table 3**.

3.1.2 Lipid–water partitioning of polyethyleneglycols and polyoxyethylene–polyoxypropylene block copolymers

PEGs exhibit low LogP values and are generally used to keep compounds in solution or to prevent, rather than enhance, cell adhesion or membrane partitioning. Poloxamers, which are closely related, are also rather hydrophilic compounds, exhibiting low LogP and high *CMC* values, as seen in **Table 3**.

Experimentally, membrane insertion of poloxamers was investigated by modelling the outer leaflet of a lipid membrane with a lipid monolayer. To study the role of the lateral lipid packing density for insertion, lipid monolayers were spread on a buffer surface at a given lateral packing density (or surface pressure) and the poloxamers were injected underneath the lipid monolayer. Wu *et al.* [39] observed insertion of poloxamer P188 into dipalmitoylphosphatidylcholine (DPPC) or dipalmitoylphosphatidyglycerol (DPPG) monolayers at very low surface pressures of $\pi < 5$ mN/m. On compression of the monolayers to surface pressures of $\pi > 35$ mN/m poloxamers were squeezed out. Results for the two phospholipids were similar, thus indicating that P188 insertion is not influenced by headgroup electrostatics.

Related experiments were performed by Maskarinec *et al.* [40]. They injected poloxamer P188 into the subphase of densely packed DPPC or DPPG monolayers ($\pi = 30$ mN/m) and slowly released the surface pressure until insertion was observed that occurred at surface pressures of $\pi \leq 22$ mN/m. Both experiments were performed at 30°C. Because the lateral packing density of a monolayer, equivalent to that of a biologically relevant bilayer, is generally $\pi_M \geq 28$ mN/m (**Table 2**), the lack of insertion above a lateral packing density of $\pi = 22$ mN/m suggests that poloxamer P188 does not insert into intact biological membranes. Poloxamer P188 was, however, reported to selectively adsorb to damaged portions of electroporated membranes and to promote healing of these membranes [41], which is consistent with the above observation. In a further investigation the 'squeeze-out pressure' in DPPC monolayers was determined as $\pi = 25$ mN/m for Pluronic F68 (P188) and $\pi = 29$ mN/m for Pluronic P85 [42].

Krylova and Pohl [43] reported insertion of Pluronic L61 (EO$_2$-PO$_{30}$-EO$_2$) into planar black lipid membranes and large unilamellar vesicles formed from diphytanoylphosphatidyl-choline (DPhPC). At high concentrations L61 acted as mobile carrier and formed ion channels. Black lipid membranes generally exhibit rather low lateral packing densities ($\pi_M < 25$ mN/m) and vesicles formed from DPhPC, which carry branched fatty acyl chains, may also exhibit relatively low packing densities of $\pi_M < 28$ mN/m.

Demina *et al.* [44] investigated the influence of different polymers on the doxorubicin flip-flop rate in sonified egg lecithin vesicles, which exhibit lipid packing densities of $\pi_M \leq 25$ mN/m. They observed that increasing the length of the hydrophobic

Table 3. Examples of P-glycoprotein inhibition by membrane-anchoring detergents and poloxamers

No.	System	Detergent	LogP KowWin	Detergent [M]	CMC [M]	A_D [Å²]	ΔG^0_{tw} (detergent) [kJ/mol]	Drug	ΔG^0_{tw} (drug) [kJ/mol]	Absorption enhancement	Ref.
1	Rats	Solutol HS	6.13	-	5.8×10^{-4} [97]	-	-33.7	Colchicine	-27.7 [64]	Yes	[98]
2	Rats	TPGS	11.31	1.74×10^{-1}	4.3×10^{-6} [96]	71	-50.9	Colchicine	-27.7 [64]	Yes	[98]
3	NIH-MDR1-G185 cells	TPGS	11.31	-	4.3×10^{-6} [96]	71	-50.9	Rhodamine 123	-33.6 [66]	Yes	[99]
4	NIH-MDR1-G185 cells	TPGS	11.31	-	4.3×10^{-6} [96]	71	-50.9	Vinblastine	-43.9 [64]	Yes	[99]
5	Caco-2 cells	TPGS	11.31	2.5×10^{-5}	4.3×10^{-6} [96]	71	-50.9	Rhodamine 123	-33.6 [66]	Yes	[26]
6	Rat jejunal tissue	TPGS	11.31	1.74×10^{-4}	4.3×10^{-6} [96]	71	-50.9	Digoxin	-21.9 [100]	Yes	[90]
7	Caco-2 cells	Tween 80	3.72	-	1×10^{-5} [96]	69	-53.8	Epirubicin	-48.6 [64]	Yes	[101]
8	Primary rat hepatocytes (MDR1)	Triton X-100	4.86	8×10^{-6}	2.3×10^{-4} [96]	54	-36.0	Rhodamine 123	-33.6 [66]	Yes	[102]
9	Drug-resistant cells	Triton X-100	4.86	8×10^{-6}	2.3×10^{-4} [96]	54	-36.0	[3H]azidopine	-	Yes	[7]
10	Caco-2 cells	TPGS	11.31	6.96×10^{-4}	4.3×10^{-6} [96]	71	-50.9	Talinolol	-	Yes	[78]
11	Healthy male volunteers	TPGS	11.31	-	4.3×10^{-6} [96]	71	-50.9	Talinolol	-	Yes	[78]
12	Caco-2 cells	Tween 80	3.72	8.27×10^{-4}	1×10^{-5} [96]	69	-53.8	Taxol	-	Yes	[51]
13	Caco-2 cells	Cremophor EL	5.91	2.9×10^{-3}	3×10^{-5} [69]	-	-41.1	Taxol	-	Yes	[51]
14	MDR1-MDCK cells	Cremophor EL	5.91	2.9×10^{-3}	3×10^{-5} [69]	-	-41.1	Taxol	-	No (weak)	[51]
15	MDR1-MDCK cells	Tween 80	3.72	8.27×10^{-4}	1×10^{-5} [96]	69	-53.8	Taxol	-	No (weak)	[51]
16	KB-8-5-11 cells (MDR1)	Octyl-β-D-glucoside	0.93	3.4×10^{-5}	2.34×10^{-2} [48]	51	-24.5	Rhodamine 123	-33.6 [66]	No	[75]
17	MDR1-MDCK cells	PEG-300	-2.57	7.08×10^{-1}	-	-	-15.0	Taxol	-	Yes	[51]
18	Rat jejunal tissue	PEG-400	-2.34	-	-	-	-	Digoxin	-21.9 [100]	Yes	[90]
19	Rat jejunal tissue	Poloxamer 235 (Pluronic P85)	-	2.17×10^{-4}	6.5×10^{-5} [28]	-	-59.11	Digoxin	-21.9 [100]	Yes	[90]
20	Caco-2 cells	Poloxamer 188 (Pluronic F68)	-	9.52×10^{-4}	4.8×10^{-4} [28]	-	-54.2	Talinolol	-	No	[78]

A_D: Cross-sectional area of the molecule in its amphiphilic orientation; CMC: Critical micelle concentration; LogP (KowWin): Calculated octanol–water partition coefficient; MDR: Multi-drug resistance protein; TPGS: D-α-Tocopheryl poly(ethylene glycol) 1000 succinate.

ΔG^0_{tw} (detergent) is the free energy of the detergent binding from the aqueous phase to the transporter (or free energy of transporter–water binding) estimated as the sum of ΔG^0_{tw}, (derived from CMC values) and ΔG^0_{tl} (derived from the number of hydrogen bond acceptor groups in type I units); ΔG^0_{tw} (drug) is derived from the concentrations of half-maximum P-glycoprotein activation [21].

Table 3. Examples of P-glycoprotein inhibition by membrane-anchoring detergents and poloxamers (continued)

No.	System	Detergent	LogP KowWin	Detergent [M]	CMC [M]	A_D [Å²]	ΔG^0_{tw} (detergent) [kJ/mol]	Drug	ΔG^0_{tw} (drug) [kJ/mol]	Absorption enhancement	Ref.
21	Healthy male volunteers	Poloxamer 188 (Pluronic F68)	-	-	4.8×10^{-4} [28]	-	-54.2	Talinolol	-	No	[78]
22	Mice	CRL-1605	-	132 mg/kg	-	-	-	Amikacin	-	Yes	[79]
23	Mice	CRL-1605	-	25 mg/kg	-	-	-	Tobramycin	-	Yes	[80]

A_D: Cross-sectional area of the molecule in its amphiphilic orientation; CMC: Critical micelle concentration; LogP (KowWin): Calculated octanol–water partition coefficient; MDCK: Madin–Darby Canine Kidney; MDR: Multi-drug resistance protein; TPGS: D-α-Tocopheryl poly(ethylene glycol) 1000 succinate.

ΔG^0_{tw} (detergent) is the free energy of the detergent binding from the aqueous phase to the transporter (or free energy of transporter–water binding) estimated as the sum of ΔG^0_{lw}, (derived from CMC values) and ΔG^0_{tl} (derived from the number of hydrogen bond acceptor groups in type I units); ΔG^0_{tw} (drug) is derived from the concentrations of half-maximum P-glycoprotein activation [21].

Figure 2. Homology model of P-glycoprotein [53]. It was built using as templates the crystal structure coordinates of the *Vibrio cholera* lipid A transporter, MsbA, for the transmembrane domains and those of the *Salmonella typhimurium* histidine permease, HisP, and the human TAP1 for the nucleotide-binding domains. Helices are shown in turquoise, and β-sheets in magenta. Hydrogen bond donor groups of P-glycoprotein in the transmembrane sequences and in the extramembranous regions are indicated in blue and red, respectively. The membrane is schematically indicated in yellow. POPC (**1**) and the membrane-anchoring detergents (numbers correspond to those in Table 1), 1,2-dicaprylin (**2**), Cremophor EL (**5**), Vitamin E TPGS (**9**), Tween 80 (**10**) and Triton X-100 (**13**), are inserted in the cytosolic membrane leaflet from where they can diffuse to the P-glycoprotein-binding region, whereas Pluronic L61 (**17**) is shown on the extracellular side of the membrane as membrane insertion is unlikely. The detergents and lipids are represented as sticks, the van der Waals surface are indicated as shadow and the hydrogen acceptors (oxygens) are in red.

propylene oxide (PO) blocks increased the flip-flop activity, whereas polymers with simple fatty acyl chain anchors such as the Brij series had no effect. This suggests that flip-flop increases with the defect induced in the loosely packed outer bilayer leaflet by partially inserting polymers. Membrane disordering by poloxamers was also observed in bovine brain microvessel endothelial cells (BBMECs), whereby the highest effect was obtained with Pluronic P85 [44,45].

In summary, detergents with hydrophobic anchors insert into membranes exhibiting a wide range of lipid–water partition coefficients. However, poloxamers are unlikely to enter intact living cells by passive diffusion if the lateral packing density of the plasma membrane is $\pi_M \geq 29$ mN/m. Lipid–water partition coefficients, K_{lw}, of poloxamers are, therefore, most likely negligibly small for tightly packed membranes. The IB, with an estimated lateral packing density

of $\pi_M \approx 28$ mN/m, may be in a range where certain poloxamers (e.g., P85) start to penetrate.

3.2 The influence of detergents on the lateral packing density of membranes

Short-chain, lipid-anchoring detergents such as oligo(ethylene oxide) alkyl ether (e.g., $C_{12}EO_8$) [46], n-alkanols [47], octyl-β-D-glucopyranoside [48] and Triton X-100 [49] generally reduce the order parameter, S, of the lipid chains, the bending modulus, κ, or the lipid lateral packing density, π_M, at high concentrations and physiological temperature. PEG-300 led to a disordering of the membrane headgroup region in both Caco-2 cells and Madin–Darby canine kidney cells transfected with the human *MDR1* gene (Madin–Darby Canine Kidney [MDCK]-MDR1 cells). Vitamin E incorporation (20 mol%) into DPPC membranes lead in

Figure 3. Hydrogen bond acceptor patterns that are present in all P-glycoprotein substrates. Patterns formed from two hydrogen bond acceptor pairs with a spatial 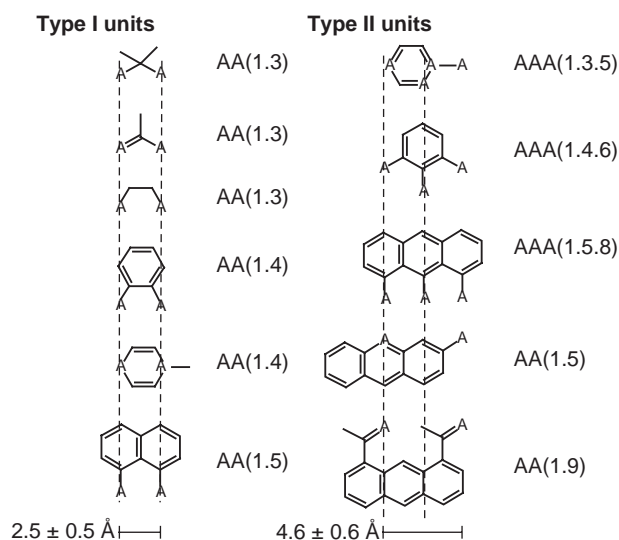separation of 2.5 ± 0.5 Å are called type I units and hydrogen bond acceptor pairs with a spatial separation of 4.6 ± 0.5 Å are called type II. The latter consists of two or three hydrogen bond acceptor [55,57].

contrast to an increase in the acyl chain ordering by 17% [50]. Membrane ordering/disordering effects are generally observed at much higher concentrations than used to inhibit P-gp [51].

The micromechanical properties of poloxamer/egg phosphatidylcholine vesicles have been investigated using atomic force microscopy. To obtain membrane-spanning poloxamers, the lipid and the poloxamers were cosolubilised in organic solvent. Water was added only after solvent evaporation. The mixed vesicles were much more rigid than the pure lipid vesicles and showed a several-fold higher bending modulus, κ [52]. This is in contrast to the membrane-disordering effects observed by Demina *et al.* [44], which are most likely due to partial insertion of poloxamers in loosely packed bilayers.

In summary, the detergent concentrations required for absorption enhancement are generally much lower than the concentrations used for measurements of membrane-ordering/disordering effects. It can, therefore, be concluded that the membrane-ordering/disordering effects of the detergents required for P-gp inhibition are relatively small.

4. P-glycoprotein

Figure 2 displays a recent homology model of P-gp by Omote and Al-Shawi [53]. The model reveals the two nucleotide-binding domains located at the cytosolic side of the membrane and the two transmembrane domains each comprising a bundle of six transmembrane helices, which together form one functional entity. An analysis of the amino acids in the transmembrane helices revealed a high density of residues with hydrogen bond donor side chains arranged in an amphipathic manner [18]. Because P-gp and related ABC transporters of hydrophobic compounds bind their substrates not from the aqueous phase, but from the cytosolic membrane leaflet, the lipid bilayer plays a synergistic role in substrate binding. It accumulates the substrates due to their high lipid–water partition coefficient and, moreover, orients them such that the polar parts reside in the lipid headgroup region and the more hydrophobic parts in the hydrophobic core region. This is illustrated in **Figure 2**, in which the membrane-anchoring detergents are shown inserted into the cytosolic side of the membrane. Pluronic L61, which does not insert into densely packed membranes, is shown on the extracellular side. Due to its low dielectric constant, ε, the lipid environment prioritises electrostatic and hydrogen bonding interactions [21,54].

Screening the three-dimensional structures of a large number of drugs revealed that the minimal common binding element consisting of two or three hydrogen bond acceptor groups in a specific spatial distance (so-called type I or II units, shown in **Figure 3**) [54-58]. Because the transmembrane sequences of P-gp are rich in hydrogen bond donor groups, they may recognise the hydrogen bond acceptor groups of the substrates through hydrogen bond formation in the lipid membrane environment [53,59].

Examples for type I units include ester groups (**Figure 3**, column 1, line 2), which are abundant in detergents as seen in **Table 1**. It has recently been shown that polyoxyethylene sequences, which exhibit no preformed type I units, are also recognised by P-gp [27], thus suggesting that type I units can be induced by rotation of carbon bonds following binding to the transporter (**Figure 3**, column 1, line 3). Most detergents exhibit preformed, rigid and/or inducible type I units as indicated in **Table 1**. It is interesting to note that Vitamin E, which is not an intrinsic P-gp substrate in its natural form, is transformed to a substrate by pegylation. POPC is not a
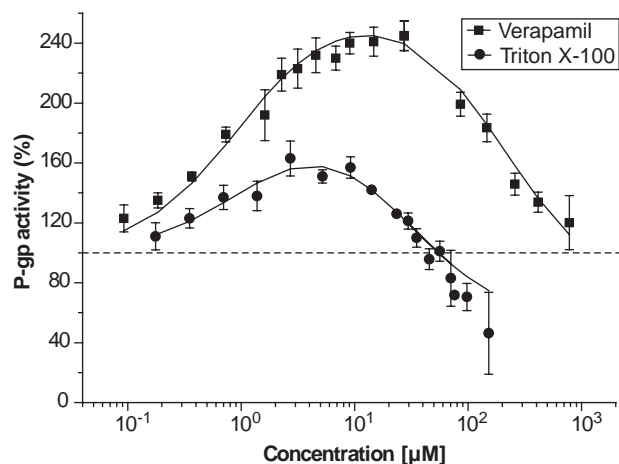
Figure 4. Triton X-100 activates P-gp. P-gp ATPase activation as a function of the concentration of Triton X-100 (Li-Blatter and Seelig, unpublished results) and verapamil [24] measured in inside-out plasma membrane vesicle from mouse embryo fibroblasts, NIH-MDR1-G185, measured at pH 7.4 at 37°C. The data are fitted to a modified Michaelis–Menten equation (Equation 4) [21,64].
P-gp: P-glycoprotein.

substrate in its zwitterionic form; however, it may be a substrate in its protonated, cationic form [60].

In summary, substrates have first to partition into the lipid membrane in order to bind to P-gp. For the direct interaction between substrates in the lipid environment and P-gp hydrogen bonding interactions seem to be important. The broad substrate specificity of P-gp is compatible with a modular binding concept, in which the binding modules are type I and II units. The membrane-mediated binding of substrates to P-gp thus differs distinctly from the water-mediated interactions.

4.1 Identifying substrates, modulators and inhibitors of P-glycoprotein

P-gp has a basal activity in the absence of externally added intrinsic substrates, which is either due to uncoupled cycling [61] or to the transport of endogenous substrates [62,63]. Addition of intrinsic substrates generally induces a change (increase or decrease) in the P-gp activity. **Figure 4** displays the P-gp activity as a function of Triton X-100 concentration (Log scale) (Blatter and Seelig, unpublished results). For comparison, the P-gp activity as a function of verapamil concentration is also displayed. Verapamil is a known P-gp substrate [24]. The activation profiles show the characteristic bell-shaped dependence observed previously for different drugs in living MDR1-transfected cells [21], in plasma membrane vesicles of transfected cells [64], and in reconstituted proteoliposomes [61]. The same type of activation profiles as shown in **Figure 4** was also obtained for Triton X-100 and other detergents such as Tween 80 and $C_{12}EO_8$ in living MDR1-transfected cells [27].

The solid lines in **Figure 4** are fits to the modified Michaelis–Menten equation (Equation 4) proposed by Litman *et al.* [64]. The model assumes an activating binding region (occupied at low substrate concentrations) and an inhibitory binding region (occupied at high substrate concentrations) [21]:

$$(4)$$

$$V_{sw} = \frac{K_1 K_2 V_{bas} + K_2 K_1 C_{sw} + V_2 C_{sw}^2}{K_1 K_2 + K_2 C_{sw} + C_{sw}^2}$$

where V_{sw} is the rate of phosphate release as a function of substrate concentration in aqueous solution, C_{sw}, and V_{bas} is the basal activity in the absence of drug; V_1, is the maximum transporter activity and V_2 is the minimum activity at infinite substrate concentration. At the concentration $C_{sw} = K_1$, half-maximum binding at the activating binding region and at the concentration $C_{sw} = K_2$, half-maximum binding at the inhibitory binding region, is reached.

The intrinsic drug transport rate is directly proportional to the turnover number of P-gp. For example, the verapamil (10 µM)-induced turnover was determined as 4.7 ± 0.8 ATP hydrolysed/P-gp/s [65,66]. Whether the transport cycle of P-gp requires one [67] or two [68] ATP molecules per drug transported is still a matter of debate. Intrinsic substrates are thus defined as compounds that are transported at the rate of ATP hydrolysis (or half the rate of ATP hydrolysis). It has been shown that the rate of ATP hydrolysis and concomitantly the rate of transport decreases exponentially with the binding affinity of the drug to the transporter, whereby the binding affinity to the transporter (in the lipid membrane) seems to be essentially due to hydrogen bond formation [15,53].

At high concentrations each substrate can in principle act as a modulator or inhibitor of P-gp, as seen in **Figure 4**. However, toxicity often prohibits the use of high concentrations. If two different intrinsic substrates are added simultaneously to P-gp-containing membranes, the substrate with the higher binding affinity to P-gp (or higher free energy of transporter–water binding, ΔG^0_{tw}) acts as a modulator [15,57]. As shown below, detergents with a high free energy of drug binding from water to the transporter (or free energy of transporter–water binding ΔG^0_{tw}) are modulators or inhibitors and enhance drug absorption.

In summary, intrinsic substrates for P-gp can be identified *in silico* as compounds that: i) are cationic or noncharged; ii) are able to reach the inner cytosolic leaflet of the cytoplasmic membrane; and iii) carry hydrogen bond acceptor groups arranged in type I and/or II units. Experimentally, intrinsic substrates, modulators and inhibitors can be identified unambiguously by ATPase activation measurements. Intrinsic substrates, modulators and, in most cases, inhibitors are

transported. The rate of the intrinsic transport (which is not the rate of apparent transport as shown in Section 4.2) is proportional to the rate of ATP hydrolysis. The rate of intrinsic transport decreases with increasing affinity of the compound to the transporter.

4.2 Identifying apparent substrates

Experimentally, apparent substrates are generally determined by means of transport measurements across confluent cell monolayers that express high levels of P-gp [69]. If P-gp is located in the apical membrane (**Figure 1**) the apparent transport across the confluent cell layer is composed of the active transport by P-gp into the apical medium in combination with passive influx into the cell and the passive bidirectional diffusion across the basolateral membrane. Under the assumption that the passive diffusion across the basolateral membrane is identical in both directions, the apparent efflux can be estimated using Equation 5 [15]. It shows that the net flux, J, of compounds depends on the sum of the active efflux, $-V$, by P-gp and the passive influx, ϕ (both given in molecules transported $s^{-1}cell^{-1}$):

$$(5)$$

$$J = \phi - V$$

Passive influx, ϕ, is proportional to the lipid–water partition coefficient, K_{lw}, which in turn strongly depends on the cross-sectional area, A_D, of the molecule and the lateral packing density, π_M, of the lipid bilayer (see Equation 2). Apparent transport can thus be estimated as the sum of active efflux (determined by ATPase measurements) and passive influx (calculated on the basis of data obtained from surface activity measurements) according to Equation 5 [15].

Intrinsic substrates with a large cross-sectional area and/or a high charge that diffuse slowly are detected as apparent substrates and do not reach the cytosol, whereas intrinsic substrates that diffuse rapidly appear as 'nonsubstrates'. Whether a compound is an apparent substrate depends, however, not exclusively on the characteristics of the compound, but also on those of the lipid membrane. Typical intrinsic substrates (e.g., verapamil with $A_D \approx 80$ Å²) can, therefore, be apparent substrates in cell lines with a high lateral packing density and apparent 'nonsubstrates' in cell lines with a lower lateral packing density, π_M.

Most detergents have cross-sectional areas, $A_D < 80$ Å², and, therefore, are rather intrinsic than apparent substrates [15]. Because detergents can change the lateral packing density of the lipid membrane at high concentrations they can also modulate the passive influx, ϕ, of drugs. However, as

discussed in Section 3.2, the effect of detergents on the membrane packing density is generally relatively low at the concentrations used for P-gp inhibition.

In summary, apparent substrates are intrinsic substrates with a low rate of passive influx due either to a large cross-sectional area and/or a high charge. Because most detergents diffuse rapidly they are generally intrinsic, but not apparent substrates. At high concentrations they may enhance passive influx of a coadministered drug by loosening the membrane packing density and enhancing its influx rate (and by thus transforming an apparent substrate to an apparent 'nonsubstrate'). At the concentrations required for P-gp inhibition these effects are, however, relatively small.

4.3 Estimation of the substrate binding affinity to P-glycoprotein

As illustrated in **Figure 1**, the intramembrane binding site of P-gp implies that the free energy of drug binding from water to the transporter (or the free energy of transporter–water binding, ΔG^0_{tw}) of the substrate is the sum of the free energy of lipid–water partitioning (ΔG^0_{lw}), and the free energy of binding from the lipid membrane to the transporter (or the free energy of transporter–lipid binding ΔG^0_{tl}):

$$(6)$$

$$\Delta G^0_{tw} = \Delta G^0_{lw} + \Delta G^0_{tl}$$

The free energy of transporter–water binding can be expressed as:

$$(7)$$

$$\Delta G^0_{tw} \cong -RT(\ln[C_w K_{lw}] + \ln K_{tl})$$

where RT is the thermal energy per mole, K_{lw} is the lipid–water partition coefficient, K_{tl} is the transporter–lipid binding constant and $C_w = 55.5$ mol/l is the molar concentration of water. The free energy of lipid–water partitioning, ΔG^0_{lw}, of 15 structurally different drugs has previously been determined from surface activity measurements, using the membrane packing density of 30 mN/m, for MDR1-transfected mouse embryo fibroblasts (see **Table 2**), as $\Delta G^0_{lw} = -23$ to -34 kJ/mol. The free energy of transporter–water binding of the drugs was derived from the concentration of half-maximum P-gp activation as $\Delta G^0_{tw} = -30$ to -54 kJ/mol. The free energy of lipid–transporter binding, ΔG^0_{tl}, which is experimentally not directly accessible, was determined as the difference of the latter

two free energies as ΔG^0_{tl} = -6 to -27 kJ/mol [21]. The free energy of lipid–water partitioning, ΔG^0_{lw}, is thus generally more negative (reflecting a higher affinity to the transporter) than the free energy of binding of the drug from the lipid membrane to the transporter, ΔG^0_{tl}. At half-maximum P-gp activation, the concentration of drugs bound to the membrane, C_b [mmol/l lipid], is generally one to three orders of magnitude higher than the corresponding aqueous concentration. The relatively weak substrate–transporter interactions are thus compensated by the high membrane concentrations of the drugs. For a given drug (e.g., verapamil) the membrane concentration for half-maximum activation (mmol/l lipid) is a constant, although the aqueous concentration for half-maximum activation (mmol/l water) varies by up to two orders of magnitude, increasing with increasing membrane packing density [21].

The free energy of transporter–lipid binding of the substrate, ΔG^0_{tl}, was proposed to be the sum of the free energies, ΔG^0_{Hi}, of the individual potential hydrogen bonds formed between substrate and transporter [15,55]:

$$(8)$$

$$\Delta G^0_{tl} \approx \sum_{i=1}^{n} \Delta G^0_{Hi}$$

This hypothesis was tested by dividing the experimentally determined ΔG^0_{tl} values by the number of hydrogen bond acceptor groups per substrate yielding the apparent free energy per hydrogen bond, ΔG^0_{Hi}. For a compound with only one type I unit ΔG^0_{Hi} = -7.8 kJ/mol. With increasing number of patterns the free energy of hydrogen bond formation decreased and reached a limiting value of $\Delta G^0_{Hi} \sim$ -2.5 kJ/mol for compounds with more than four hydrogen bond acceptor patterns (eight hydrogen bonds) [21].

4.4 Membrane-anchoring detergents as modulators of P-glycoprotein: estimation of binding affinity

For the present investigation the free energy of transporter–water binding, ΔG^0_{tw}, was estimated according to Equation 6 as the sum of the free energy of lipid–water partitioning, ΔG^0_{lw}, and the free energy of transporter–lipid binding, ΔG^0_{tl}, where the former was estimated from the CMC according to Equation 3, and the latter was calculated according to Equation 8 as the sum of the free energies per hydrogen bond formation using the limiting value of ΔG^0_{Hi} = -2.5 kJ/mol [21]. For the estimation of ΔG^0_{tl} all hydrogen bond acceptors in stable as well as in inducible type I units were counted, except for poloxamers. For the latter, the value of hydrogen bond acceptors was limited to 10, as a sequence of 10 PO units would span a lipid bilayer. Compounds with similarly high numbers of hydrogen bond acceptor groups are ciclosporin A and PSC-833. For the membrane-anchored detergents investigated the estimated free energy of transporter–water binding, of lipid–water partitioning and of transporter–lipid binding was ΔG^0_{tw} = -24 to -54 kJ/mol, ΔG^0_{lw} = -20 to -39 kJ/mol and ΔG^0_{tl} = -5 to -25 kJ/mol, respectively, which is similar to the corresponding values obtained for drugs (see Section 4.3). Because the free energy of lipid–water partitioning of detergents was estimated on the basis of CMC values [38], it may be somewhat overestimated for cells with high lateral packing densities. As seen in **Table 3**, the present estimation of the free energy of water–transporter binding, ΔG^0_{tw} (or binding affinity of detergents dissolved in water to P-gp) increases in the order: octyl-β-D-glucoside < Triton X-100 < Vitamin E TPGS < Tween 80 < Pluronic L61. As discussed in Section 3.1.2, the largest molecules insert only into loosely packed membranes ($\pi_M \leq$ 28 mN/m). The cross-sectional area, A_D, of the molecules increases in the order: octyl-β-D-glucoside < Triton X-100 < Vitamin E TPGS < Tween 80 < Pluronic L61. If the compounds exhibit large cross-sectional areas the order of binding constants given above may, therefore, change for membranes with high lateral packing densities, π_M.

In summary, the binding affinity of detergents from water to the transporter can be estimated as the sum of their binding affinity to the lipid membrane (derived from CMC values) and their binding affinity to the transporter in the lipid membrane (estimated as the sum of the potential free energies of hydrogen bond formation between detergent and transporter).

5. Analysis of data

The individual experiments summarised in **Table 3** are numbered from 1 to 23 (column 1). **Table 3** lists the *in vivo* or *in vitro* system (column 2), the detergent used (column 3), the calculated LogP (KowWin) (column 4), the detergent concentration (column 5), the CMC (column 6), the cross-sectional area, A_D (measured or calculated) (column 7) and the detergent affinity to the transporter, ΔG^0_{tw} (detergent) (column 8). For comparison the drug applied (column 9) and its affinity to the transporter, ΔG^0_{tw} (drug) (column 10) is given. The latter was derived from the concentration of half-maximum P-gp activation [21]. The binding affinity of the detergents has been estimated using the approach described in Section 4.4. Column 11 finally shows whether absorption of the given drug is enhanced by the presence of the detergent.

It should be noted that enhancement of drug absorption observed in *in vitro* experiments with P-gp overexpressing cells mainly reflects the direct interaction between the detergent and P-gp. The enhancement of drug absorption in *in vivo* experiments reflects the interaction of the detergent with P-gp, with other ABC transporters, with the metabolising enzyme, CYP, and possibly with other metabolising enzymes. In the latter case enhancement of drug absorption is generally coupled with a decreased plasma clearance and an increased excretion of the nonmetabolised drug [70-73].

5.1 Absorption enhancement by membrane-anchoring detergents

5.1.1 Absorption enhancement due to P-gp inhibition

An enhancement of drug absorption by detergents due to P-gp inhibition is expected if the binding affinity of the detergent to P-gp is higher than that of the drug (i.e., if the free energy of binding, ΔG^0_{tw}, of the detergent to P-gp is more negative than that of the coadministered drug). This applies to experiments 1 – 8 and most likely also to experiments 9 – 13, although the concentrations of half-maximum P-gp activation and thus the ΔG^0_{tw} values of the drugs were not available in the latter experiments.

5.1.2 Cells with plasma membranes exhibiting relatively high lateral packing densities

Hugger *et al.* [51] have shown that Cremophor EL and Tween 80 are less active in inhibiting P-gp in MDCK-MDR1 cells than in Caco-2 cells. The packing density of the plasma membrane of canine kidney cells (which may be similar to that of LLC-MDR1 cells) is most likely higher than that of Caco-2 cells resembling the IB (see **Table 2**), and, therefore, insertion of Tween 80, which exhibits a relatively large cross-sectional area (A_D = 69 Å^2), is more difficult in MDCK-MDR1. The cross-sectional area of Cremophor EL has not yet been determined; however, due to the hydroxyl group in the hydrophobic extension it may insert in a bent conformation, which would lead to a relatively large cross-sectional area, A_D, and thus also to a reduced ability to insert into densely packed membranes.

Experiments by Wang *et al.* [74] (not included in **Table 3**) showed that PEG, with fatty acyl chain extensions (small cross-sectional areas) was more effective in the human colon cancer cells, HCT-8, than Vitamin E TPGS (large cross-sectional area), whereas Vitamin E TPGS was more effective in mouse embryo fibroblasts, NIH3T3, exhibiting a lower packing density than HCT-8 cells. This example may, thus, also be traced back to differences in the cellular membrane packing density.

5.1.3 No P-glycoprotein inhibition and no absorption enhancement

If the detergent exhibits a lower binding affinity to the transporter than the drug (i.e., less negative free energy of binding to P-gp) no absorption enhancement is expected. This behaviour is observed for the detergent, octyl-β-D-glucoside in the presence of rhodamine 123 [75]. The low inhibitory power of octyl-β-D-glucoside can be traced back to the low free energy of lipid–water partitioning, ΔG^0_{lw} (see Section 3.1.1).

5.2 Absorption enhancement by polyethyleneglycols and poloxamers

5.2.1 Absorption enhancement by means of drug solubilisation

Experiments 17 and 18 reveal an absorption enhancement of taxol and digoxin by PEG. Neither PEGs nor poloxamers show a strong tendency to partition into lipid membranes, as shown in Section 3.1.2. Whether pre-emptive binding and transport, as proposed for verapamil [76], could play a role for highly water soluble compounds such as PEGs has to be tested. Because taxol and digoxin are compounds with a tendency to self-associate the absorption enhancement by PEG-300 and -400 may be simply due to a solubilising effect. The same may hold true for experiment 19, in which poloxamer is used as a detergent.

5.2.2 P-glycoprotein inhibition by poloxamers

Provided PEG or poloxamer can insert into the membrane (e.g., at low membrane packing density, π_M > 28 mN/m, see Section 3.1.2), they are likely to be substrates for P-gp. A systematic analysis of the P-gp inhibitory potential of a large number of Pluronic block copolymers with varying length of EO and PO segments was performed in BBMEC cells [28]. The highest efficacy was obtained with copolymers exhibiting intermediate length of ~ 30 to ~ 60 PO segments and a hydrophile–lipophile balance, HLB < 20 (e.g., P85 with ~ 40 PO segments and an HLB of 16) [77].

No membrane insertion and no P-gp interaction in inside-out plasma membrane vesicles of NIH-MDR1-G185 cells (Egli and Seelig, unpublished results) were observed for Pluronic L61 with ~ 30 PO segments and an HLB of 3. Pluronic F68 (poloxamer P188) with ~ 29 PO segments and an HLB of 29 was inactive in Caco-2 cells and healthy male volunteers [78]. Pluronic P188 is a component of many laxatives [201]. Because talinolol used as a drug is highly soluble, the solubilising potential of poloxamers played no role, as seen in experiments 20 and 21.

5.2.3 Endocytosis

The positive effect of CRL-1605 on amikacin [79] and tobramycin [80] uptake is unlikely to be due to P-gp inhibition. The two compounds are highly cationic and very hydrophilic, exhibiting large negative LogP values (XLogP: -9.048 for amikacin and XLogP: -6.899 for tobramycin) and are thus unable to cross a membrane barrier by passive diffusion, even if P-gp is inhibited. Under certain conditions poloxamers may form cationic complexes with highly charged drugs, which enter the cell by means of endocytosis [81].

In summary, it has to be differentiated between detergents that can insert into the lipid membrane and detergents that cannot, whereby the potential to insert not only depends on the detergent, but also on the the membrane. Compounds that insert into the membrane can bind to P-gp, provided they carry hydrogen bond acceptor patterns (**Figure 3**). If the binding affinity of the detergent to P-gp is higher than that of the coadministered drug, the latter will act as a modulator by keeping P-gp busy and allowing the drug to cross the membrane by passive diffusion. The absorption-enhancing effect of compounds that cannot insert into densely packed membranes, such as certain Pluronics or PEGs, may be due to mere solubilisation in the presence of self-associating drugs or to endocytosis in the presence of strongly cationic compounds.

6. Side effects and toxicity due to detergent interactions with transporters and metabolising enzymes

In vivo, detergents commonly used as excipients can exert a range of biological effects, some of which have important clinical implications. Their use has been associated with severe anaphylactoid hypersensitivity reactions, hyperlipidaemia, abnormal lipoprotein patterns, aggregation of erythrocytes, peripheral neuropathy and modification of the toxicity profile of certain anticancer agents given concomitantly [82].

Anaphylactoid hypersensitivity reactions, which have been observed with several ubiquitously used detergents such as polysorbate 80, Tween 80 and Cremophor EL (polyethoxylated castor oil) [83-85], may be traced back to P-gp modulation. P-gp has been linked to the immune response and may play a role in antigen presentation [86,87].

Massive hyperlipidaemia and artherosclerosis in rodents was observed following treatment with poloxamer 407 [88]. Detergents including Pluronic L81, P85 and F68 and Cremophor EL inhibited intestinal secretion of lipoproteins, including triglyceride-rich chylomicrons in Caco-2 cells [89]. A strong correlation was observed between detergent-mediated inhibition of lipoprotein secretion and inhibition of P-gp efflux implying a link between the two processes. The fact that triglycerides display the characteristics of P-gp substrates (see **Table 1**) suggested a direct involvement of P-gp or related transporters in these processes [89]. Because many ABC transporters play a role in the maintenance of the lipid bilayer and in the transport of fatty acids and sterols [16] transporters other than P-gp may also be inhibited by detergents.

Modification of the toxicity profile due to interaction with CYP3A has been observed. P-gp and CYP3A also show overlapping substrate specificities and share a preference for compounds with many hydrogen acceptor patterns (type I patterns) such as PSC-833 [15]. Hydrogen bond acceptor-rich compounds, therefore, generally have a significant impact on drug metabolism and efflux, as shown, for example, for PEG-400 and P85, which are used as solubilising agents during *in vitro* permeability assessment in excised rat intestine [90].

An overlap between P-gp and MRPs has been observed for the Pluronics P85, L81 and F108 in bovine brain microvessel endothelial, BBMEC, cells [91,92]. In contrast, none of the detergents, TPGS, Pluronic PE8100, Cremophor EL, Pluronic PE6100 and Tween 80, showed a significant inhibition of MRP2-mediated efflux in MDCK/MRP2 cells [93].

7. Conclusion

This review provides the tools for a systematic analysis of the different factors that influence absorption enhancement of drugs by different detergents. Many detergents fulfill the requirements for intrinsic P-gp substrates, that is, they: i) are noncharged; ii) reach the cytosolic membrane leaflet, which is the site of interaction with P-gp; and iii) carry

hydrogen bond acceptor groups arranged in type I units, which are required for an interaction with P-gp. The binding affinity of a substrate from water to P-gp can be considered as the sum of the binding affinity of the substrate to the lipid membrane and the binding affinity of the substrate to P-gp in the lipid membrane. The binding affinity to the lipid membrane was derived from the free energy of micelle formation. The fact that different membranes exhibit different lateral packing densities, which in turn affects membrane partitioning, was also addressed. The binding affinity to the transporter in the lipid membrane increases with the number of hydrogen bond acceptor groups arranged in type I units (or hydrogen bond acceptor patterns) per substrate molecule. It was, therefore, estimated by attributing a free energy of $\Delta G^0_{Hi} = -2.5$ kJ/mol to each hydrogen bond acceptor group. If the binding affinity of the detergent from water to P-gp (sum of the two binding contributions) is higher than that of the coadministered drug, the detergent acts as modulator of P-gp and enhances passive diffusion of the drug into the cells. Examples for detergents with high binding affinities to P-gp that work as modulators in the presence of many different drugs include Tween 80 and Vitamin E TPGS. The latter is, however, not highly water soluble. Provided poloxamers enter the cell, they exhibit a large binding affinity to P-gp due to the high number of oxygen molecules that can be arranged in type I units. However, their effect in inhibiting P-gp is limited to membranes with low lateral packing densities. In combination with highly cationic compounds, poloxamers may induce endocytosis. The modular binding approach predicts the absorption enhancement of drugs by various noncharged detergents in good agreement with published *in vivo* and *in vitro* data.

8. Expert opinion

Noncharged detergents that carry hydrogen bond acceptor patterns directly bind to P-gp and modulate or inhibit the interaction of drugs with P-gp, provided they can reach the cytosolic membrane leaflet. As a result drugs can cross the membrane by passive diffusion without being exported by P-gp, which leads to an enhancement of drug absorption. Experiments made with confluent cell monolayers of MDR1-transfected cells reveal essentially this P-gp inhibitory effect. In *in vivo* experiments the enhancement of drug absorption due to P-gp inhibition by detergents is accompanied by additional effects, such as decreased plasma clearance and increased excretion of the nonmetabolised drug, thus suggesting an interaction of detergents with other ABC transporters and with the metabolising enzyme, CYP. Many ABC transporters are charge sensitive; however, they most likely share noncharged substrates with P-gp and CYP. Noncharged detergents and drugs may thus interact with a larger number of transporters than charged compounds. This may explain their considerable effect on the ADME properties and their

enhancement of the toxicity of the coadministered drug. With this in mind, dosing of noncharged detergents used as excipients in drug formulations becomes a difficult task, especially if one considers in addition the variability of the expression level of the different transporters and metabolising enzymes in different individuals. Alternative ways to enhance drug absorption may, therefore, also be considered. As shown above, decreasing the cross-sectional area of a drug molecule in its amphiphilic orientation significantly enhances passive influx of the drug and concomitantly reduces active efflux by P-gp. Because detergents are omnipresent, not only as cleaning agents and excipients in drug formulations, but also as food and cosmetic additives, it is of importance to identify those

with strong inhibitory effects on ABC transporters and metabolising enzymes and, moreover, to quantify their modulating and inhibitory interactions. The present analysis provides the tools for an approximate and simple *a priori* estimate of the membrane and P-gp binding ability of noncharged detergents based on a modular binding approach. However, much remains to be clarified in the field of the interaction of detergents with other ABC transporters.

Acknowledgement

Bibliography

1. RIEHM H, BIEDLER JL: Potentiation of drug effect by Tween 80 in Chinese hamster cells resistant to actinomycin D and daunomycin. *Cancer Res.* (1972) **32**(6):1195-1200.

2. FRICHE E, JENSEN PB, SEHESTED M, DEMANT EJ, NISSEN NN: The solvents Cremophor EL and Tween 80 modulate daunorubicin resistance in the multidrug resistant Ehrlich ascites tumor. *Cancer Commun.* (1990) **2**(9):297-303.

3. COON JS, KNUDSON W, CLODFELTER K, LU B, WEINSTEIN RS: Solutol HS 15, nontoxic polyoxyethylene esters of 12-hydroxystearic acid, reverses multidrug resistance. *Cancer Res.* (1991) **51**(3):897-902.

4. SPOELSTRA EC, DEKKER H, SCHUURHUIS GJ, BROXTERMAN HJ, LANKELMA J: P-glycoprotein drug efflux pump involved in the mechanisms of intrinsic drug resistance in various colon cancer cell lines. Evidence for a saturation of active daunorubicin transport. *Biochem. Pharmacol.* (1991) **41**(3):349-359.

5. WOODCOCK DM, LINSENMEYER ME, CHOJNOWSKI G *et al.*: Reversal of multidrug resistance by surfactants. *Br. J. Cancer* (1992) **66**(1):62-68.

6. LOE DW, SHAROM FJ: Interaction of multidrug-resistant Chinese hamster ovary cells with amphiphiles. *Br. J. Cancer* (1993) **68**(2):342-351.

7. ZORDAN-NUDO T, LING V, LIU Z, GEORGES E: Effects of nonionic detergents on P-glycoprotein drug binding and reversal of multidrug resistance. *Cancer Res.* (1993) **53**(24):5994-6000.

8. DRORI S, EYTAN GD, ASSARAF YG: Potentiation of anticancer-drug cytotoxicity by multidrug-resistance chemosensitizers involves alterations in membrane fluidity leading to increased membrane permeability. *Eur. J. Biochem.* (1995) **228**(3):1020-1029.

9. BORREL MN, FIALLO M, VERESS I, GARNIER-SUILLEROT A: The effect of crown ethers, tetraalkylammonium salts, and polyoxyethylene amphiphiles on pirarubicin incorporation in K562 resistant cells. *Biochem. Pharmacol.* (1995) **50**(12):2069-2076.

10. LITMAN T, DRULEY TE, STEIN WD, BATES SE: From MDR to MXR: new understanding of multidrug resistance systems, their properties and clinical significance. *Cell. Mol. Life Sci.* (2001) **58**(7):931-959.

11. STOUCH TR, GUDMUNDSSON O: Progress in understanding the structure–activity relationships of P-glycoprotein. *Adv. Drug Deliv. Rev.* (2002) **54**(3):315-328.

12. CORDON-CARDO C, O'BRIEN JP, CASALS D *et al.*: Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood–brain barrier sites. *Proc. Natl. Acad. Sci. USA* (1989) **86**(2):695-698.

13. THIEBAUT F, TSURUO T, HAMADA H, GOTTESMAN MM, PASTAN I, WILLINGHAM MC: Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. *Proc. Natl. Acad. Sci. USA* (1987) **84**(21):7735-7738.

14. TANIGAWARA Y: Role of P-glycoprotein in drug disposition. *Ther. Drug Monit.* (2000) **22**(1):137-140.

15. SEELIG A, GATLIK-LANDWOJTOWICZ E: Inhibitors of multidrug efflux transporters: their membrane and protein interactions. *Mini Rev. Med. Chem.* (2005) **5**(2):135-151.

16. DEAN M, HAMON Y, CHIMINI G: The human ATP-binding cassette (ABC) transporter superfamily. *J. Lipid Res.* (2001) **42**(7):1007-1017.

17. ZAMEK-GLISZCZYNSKI MJ, HOFFMASTER KA, NEZASA K, TALLMAN MN, BROUWER KL: Integration of hepatic drug transporters and phase II metabolizing enzymes: mechanisms of hepatic excretion of sulfate, glucuronide, and glutathione metabolites. *Eur. J. Pharm. Sci.* (2006) **27**(5):447-486.

18. SEELIG A, BLATTER XL, WOHNSLAND F: Substrate recognition by P-glycoprotein and the multidrug resistance-associated protein MRP1: a comparison. *Int. J. Clin. Pharmacol. Ther.* (2000) **38**(3):111-121.

19. ZHANG Y, BENET LZ: The gut as a barrier to drug absorption: combined role of cytochrome P450 3A and P-glycoprotein. *Clin. Pharmacokinet.* (2001) **40**(3):159-168.

20. HIGGINS CF, GOTTESMAN MM: Is the multidrug transporter a flippase? *Trends Biochem. Sci.* (1992) **17**(1):18-21.

21. GATLIK-LANDWOJTOWICZ E, AANISMAA P, SEELIG A: Quantification and characterization of P-glycoprotein-substrate interactions. *Biochemistry* (2006) **45**(9):3020-3032.

22. ORLOWSKI S, SELOSSE MA, BOUDON C *et al.*: Effects of detergents on P-glycoprotein atpase activity: differences in perturbations of basal and verapamil-dependent activities. *Cancer Biochem. Biophys.* (1998) **16**(1-2):85-110.

23. REGEV R, ASSARAF YG, EYTAN GD: Membrane fluidization by ether, other anesthetics, and certain agents abolishes P-glycoprotein ATPase activity and modulates efflux from multidrug-resistant cells. *Eur. J. Biochem.* (1999) **259**(1-2):18-24.

24. MEIER M, LI-BLATTER X, SEELIG A, SEELIG J: Interaction of verapamil with lipid membranes and P-glycoprotein. Connecting thermodynamics and membrane structure with functional activity. *Biophysical J.* (In Press).

25. WAKITA M, KURODA Y, NAKAGAWA T: Interactions between local anesthetic dibucaine and pig erythrocyte membranes as studied by proton and phosphorus-31 nuclear magnetic resonance spectroscopy. *Chem. Pharm. Bull. (Tokyo)* (1992) **40**(6):1361-1365.

26. REGE BD, KAO JP, POLLI JE: Effects of nonionic surfactants on membrane transporters in Caco-2 cell monolayers. *Eur. J. Pharm. Sci.* (2002) **16**(4-5):237-246.

27. NERVI P: Peptide and drug interactions in murine models *in vitro* and *in vivo* studies with microphysiometry, confocal microscopy and functional magnetic resonance imaging. PhD Thesis, University of Basel, Switzerland (2004).

28. KABANOV AV, BATRAKOVA EV, MILLER DW: Pluronic block copolymers as modulators of drug efflux transporter activity in the blood–brain barrier. *Adv. Drug Deliv. Rev.* (2003) **55**(1):151-164.

29. SEELIG J, SEELIG A: Lipid conformation in model membranes and biological membranes. *Q. Rev. Biophys.* (1980) **13**(1):19-61.

30. SEELIG A, SEELIG J: Membrane Structure. In: *Encyclopedia of Physical Science and Technology, 3rd edn.* Academic Press, NY, USA (2002):355-367.

31. LITMAN T, NIELSEN D, SKOVSGAARD T, BUKHAVE K: Lipid composition of sensitive and multidrug resistant Ehrlich ascites tumour cells. *Cell. Pharmacol.* (1995) **2**:9-14.

32. FISCHER H, GOTTSCHLICH R, SEELIG A: Blood–brain barrier permeation: molecular parameters governing passive diffusion. *J. Membr. Biol.* (1998) **165**(3):201-211.

33. FISCHER H, SEELIG A, CHOU RC, VAN DE WATERBEEMD J: The difference between the diffusion through the blood–brain barrier and the gastro-intestinal membrane. *4th International Conference on Drug Absorption.* Edinburgh, UK (1997).

34. SEELIG A: Local anesthetics and pressure: a comparison of dibucaine binding to lipid monolayers and bilayers. *Biochim. Biophys. Acta* (1987) **899**(2):196-204.

35. BOGUSLAVSKY V, REBECCHI M, MORRIS AJ, JHON DY, RHEE SG, MCLAUGHLIN S: Effect of monolayer surface pressure on the activities of phosphoinoside-specific phospholipase C-β1, -γ1, and -δ1. *Biochemistry* (1994) **33**(10):3032-3037.

36. GEREBTZOFF G, LI-BLATTER X, FISCHER H, FRENTZEL A, SEELIG A: Halogenation of drugs enhances membrane binding and permeation. *Chembiochem* (2004) **5**(5):676-684.

37. GOBAS FA, LAHITTETE JM, GAROFALO G, SHIU WY, 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):265-272.

38. HEERKLOTZ H, SEELIG J: Correlation of membrane/water partition coefficients of detergents with the critical micelle concentration. *Biophys. J.* (2000) **78**(5):2435-2440.

39. WU G, MAJEWSKI J, EGE C, KJAER K, WEYGAND MJ, LEE KY: Lipid corralling and poloxamer squeeze-out in membranes. *Phys. Rev. Lett.* (2004) **93**(2):028101.

40. MASKARINEC SA, HANNIG J, LEE RC, LEE KY: Direct observation of poloxamer 188 insertion into lipid monolayers. *Biophys. J.* (2002) **82**(3):1453-1459.

41. WU G, MAJEWSKI J, EGE C, KJAER K, WEYGAND MJ, LEE KY: Interaction between lipid monolayers and poloxamer 188: an X-ray reflectivity and diffraction study. *Biophys. J.* (2005) **89**(5):3159-3173.

42. CHANG LC, LIN CY, KUO MW, GAU CS: Interactions of Pluronics with phospholipid monolayers at the air–water interface. *J. Colloid Interface Sci.* (2005) **285**(2):640-652.

43. KRYLOVA OO, POHL P: Ionophoric activity of pluronic block copolymers. *Biochemistry* (2004) **43**(12):3696-3703.

44. DEMINA T, GROZDOVA I, KRYLOVA O *et al.*: Relationship between the structure of amphiphilic copolymers and their ability to disturb lipid bilayers. *Biochemistry* (2005) **44**(10):4042-4054.

45. BATRAKOVA EV, MILLER DW, LI S, ALAKHOV VY, KABANOV AV, ELMQUIST WF: Pluronic P85 enhances the delivery of digoxin to the brain: *in vitro* and *in vivo* studies. *J. Pharmacol. Exp. Ther.* (2001) **296**(2):551-557.

46. OTTEN D, BROWN MF, BEYER K: Softening of membrane bilayers by detergents elucidated by deuterium NMR spectroscopy. *J. Phys. Chem. B* (2000) **104**:12119-12129.

47. THEWALT JL, TULLOCH AP, CUSHLEY RJ: A deuterium NMR study of labelled n-alkanol anesthetics in a model membrane. *Chem. Phys. Lipids* (1986) **39**(1-2):93-107.

48. WENK MR, ALT T, SEELIG A, SEELIG J: Octyl-β-D-glucopyranoside partitioning into lipid bilayers: thermodynamics of binding and structural changes of the bilayer. *Biophys. J.* (1997) **72**(4):1719-1731.

49. HEERKLOTZ H: Triton promotes domain formation in lipid raft mixtures. *Biophys. J.* (2002) **83**(5):2693-2701.

50. WASSALL SR, THEWALT JL, WONG L, GORRISSEN H, CUSHLEY RJ: Deuterium NMR study of the interaction of α-tocopherol with a phospholipid model membrane. *Biochemistry* (1986) **25**(2):319-326.

51. HUGGER ED, NOVAK BL, BURTON PS, AUDUS KL, BORCHARDT RT: A comparison of commonly used polyethoxylated pharmaceutical excipients on their ability to inhibit P-glycoprotein activity *in vitro*. *J. Pharm. Sci.* (2002) **91**(9):1991-2002.

52. LIANG X, MAO G, NG KY: Effect of chain lengths of PEO-PPO-PEO on small unilamellar liposome morphology and stability: an AFM investigation. *J. Colloid Interface Sci.* (2005) **285**(1):360-372.

53. OMOTE H, AL-SHAWI MK: Interaction of transported drugs with the lipid bilayer and P-glycoprotein through a solvation exchange mechanism. *Biophys. J.* (2006) **90**(11):4046-4059.

54. SEELIG A, LANDWOJTOWICZ E, FISCHER H, BLATTER XL: Towards P-glycoprotein structure–activity relationships. In: *Drug bioavailability/Estimation of solubility, permeability, absorption and bioavailability.* Van der Waterbeemd H, Lennernas H,

Arthurson P (Eds), Wiley-VCH Verlag GmbH, Weinheim, Germany (2003).

55. SEELIG A: A general pattern for substrate recognition by P-glycoprotein. *Eur. J. Biochem.* (1998) **251**(1-2):252-261.

56. ECKER G, HUBER M, SCHMID D, CHIBA P: The importance of a nitrogen atom in modulators of multidrug resistance. *Mol. Pharmacol.* (1999) **56**(4):791-796.

57. SEELIG A, LANDWOJTOWICZ E: Structure–activity relationship of P-glycoprotein substrates and modifiers. *Eur. J. Pharm. Sci.* (2000) **12**(1):31-40.

58. CHIBA P, ERKER T, GALANSKI M, HITZLER M, ECKER GF: Synthesis and multidrug-resistance modulating activity of a series of thienothiazines. *Arch. Pharm. (Weinheim)* (2002) **335**(5):223-228.

59. SEELIG A: Unraveling membrane-mediated substrate–transporter interactions. *Biophys. J.* (2006) **90**(11):3825-3826.

60. LANDWOJTOWICZ E, NERVI P, SEELIG A: Real-time monitoring of P-glycoprotein activation in living cells. *Biochemistry* (2002) **41**(25):8050-8057.

61. AL-SHAWI MK, POLAR MK, OMOTE H, FIGLER RA: Transition state analysis of the coupling of drug transport to ATP hydrolysis by P-glycoprotein. *J. Biol. Chem.* (2003) **278**(52):52629-52640.

62. RAGGERS RJ, POMORSKI T, HOLTHUIS JC, KALIN N, VAN MEER G: Lipid traffic: the ABC of transbilayer movement. *Traffic* (2000) **1**(3):226-234.

63. ROMSICKI Y, SHAROM FJ: Phospholipid flippase activity of the reconstituted P-glycoprotein multidrug transporter. *Biochemistry* (2001) **40**(23):6937-6947.

64. LITMAN T, ZEUTHEN T, SKOVSGAARD T, STEIN WD: Structure-activity relationships of P-glycoprotein interacting drugs: kinetic characterization of their effects on ATPase activity. *Biochim. Biophys. Acta* (1997) **1361**(2):159-168.

65. GATLIK-LANDWOJTOWICZ E, AANISMAA P, SEELIG A: The rate of P-glycoprotein activation depends on the metabolic state of the cell. *Biochemistry* (2004) **43**(46):14840-14851.

66. URBATSCH IL, AL-SHAWI MK, SENIOR AE: Characterization of the ATPase activity of purified Chinese hamster P-glycoprotein. *Biochemistry* (1994) **33**(23):7069-7076.

67. SHAPIRO AB, LING V: Stoichiometry of coupling of rhodamine 123 transport to ATP hydrolysis by P-glycoprotein. *Eur. J. Biochem.* (1998) **254**(1):189-193.

68. SAUNA ZE, AMBUDKAR SV: Characterization of the catalytic cycle of ATP hydrolysis by human P-glycoprotein. The two ATP hydrolysis events in a single catalytic cycle are kinetically similar but affect different functional outcomes. *J. Biol. Chem.* (2001) **276**(15):11653-11661.

69. NERURKAR MM, HO NF, BURTON PS, VIDMAR TJ, BORCHARDT RT: Mechanistic roles of neutral surfactants on concurrent polarized and passive membrane transport of a model peptide in Caco-2 cells. *J. Pharm. Sci.* (1997) **86**(7):813-821.

70. BRAVO GONZALEZ RC, HUWYLER J, BOESS F, WALTER I, BITTNER B: *In vitro* investigation on the impact of the surface-active excipients Cremophor EL, Tween 80 and Solutol HS-15 on the metabolism of midazolam. *Biopharm. Drug Dispos.* (2004) **25**(1):37-49.

71. BITTNER B, GONZALEZ RC, WALTER I, KAPPS M, HUWYLER J: Impact of Solutol HS 15 on the pharmacokinetic behaviour of colchicine upon intravenous administration to male Wistar rats. *Biopharm. Drug Dispos.* (2003) **24**(4):173-181.

72. BRAVO GONZALEZ RC, HUWYLER J, WALTER I, MOUNTFIELD R, BITTNER B: Improved oral bioavailability of cyclosporin A in male Wistar rats. Comparison of a Solutol HS 15 containing self-dispersing formulation and a microsuspension. *Int. J. Pharm.* (2002) **245**(1-2):143-151.

73. BITTNER B, GONZALEZ RCB, WALTER I, HUWYLER J: Impact of oral administration of the surface-active excipient solutol HS-15 on the pharmacokinetics of intravenously administered colchicine. *Lett. Drug Des. Disc.* (2005) **2**(3):193-195.

74. WANG SW, MONAGLE J, MCNULTY C, PUTNAM D, CHEN H: Determination of P-glycoprotein inhibition by excipients and their combinations using an integrated high-throughput process. *J. Pharm. Sci.* (2004) **93**(11):2755-2767.

75. DUDEJA PK, ANDERSON KM, HARRIS JS, BUCKINGHAM L, COON JS: Reversal of multidrug resistance phenotype by surfactants: relationship to membrane lipid fluidity. *Arch. Biochem. Biophys.* (1995) **319**(1):309-315.

76. LITMAN T, SKOVSGAARD T, STEIN WD: Pumping of drugs by P-glycoprotein: a two-step process? *J. Pharmacol. Exp. Ther.* (2003) **307**(3):846-853.

77. ATTWOOD D, FLORENCE AT: *Surfactant systems: their chemistry, pharmacy and biology.* Chapman and Hall, London (UK) and New York (USA) (1983).

78. BOGMAN K, ZYSSET Y, DEGEN L *et al.*: P-glycoprotein and surfactants: effect on intestinal talinolol absorption. *Clin. Pharmacol. Ther.* (2005) **77**(1):24-32.

79. JAGANNATH C, WELLS A, MSHVILDADZE M *et al.*: Significantly improved oral uptake of amikacin in FVB mice in the presence of CRL-1605 copolymer. *Life Sci.* (1999) **64**(19):1733-1738.

80. BANERJEE SK, JAGANNATH C, HUNTER RL, DASGUPTA A: Bioavailability of tobramycin after oral delivery in FVB mice using CRL-1605 copolymer, an inhibitor of P-glycoprotein. *Life Sci.* (2000) **67**(16):2011-2016.

81. GAUCHER G, DUFRESNE MH, SANT VP, KANG N, MAYSINGER D, LEROUX JC: Block copolymer micelles: preparation, characterization and application in drug delivery. *J. Control. Release* (2005) **109**(1-3):169-188.

82. GELDERBLOM H, VERWEIJ J, NOOTER K, SPARREBOOM A: Cremophor EL: the drawbacks and advantages of vehicle selection for drug formulation. *Eur. J. Cancer* (2001) **37**(13):1590-1598.

83. COORS EA, SEYBOLD H, MERK HF, MAHLER V: Polysorbate 80 in medical products and nonimmunologic anaphylactoid reactions. *Ann. Allergy Asthma Immunol.* (2005) **95**(6):593-599.

84. HAVEL M, MULLER M, GRANINGER W, KURZ R, LINDEMAYR H: Tolerability of a new vitamin K1 preparation for parenteral administration to adults: one case of anaphylactoid reaction. *Clin. Ther.* (1987) **9**(4):373-379.

85. THEIS JG, LIAU-CHU M, CHAN HS, DOYLE J, GREENBERG ML, KOREN G: Anaphylactoid reactions in children receiving high-dose intravenous cyclosporine for reversal of tumor resistance: the causative role of improper dissolution of Cremophor EL. *J. Clin. Oncol.* (1995) **13**(10):2508-2516.

86. PENDSE S, SAYEGH MH, FRANK MH: P-glycoprotein-a novel therapeutic target for immunomodulation in clinical transplantation and autoimmunity? *Curr. Drug Targets* (2003) **4**(6):469-476.

87. BEGLEY GS, HORVATH AR, TAYLOR JC, HIGGINS CF: Cytoplasmic domains of the transporter associated with antigen processing and P-glycoprotein interact with subunits of the proteasome. *Mol. Immunol.* (2005) **42**(1):137-141.

88. COGGER VC, HILMER SN, SULLIVAN D, MULLER M, FRASER R, LE COUTEUR DG: Hyperlipidemia and surfactants: the liver sieve is a link. *Atherosclerosis* (2006) (In Press).

89. SEEBALLUCK F, ASHFORD MB, O'DRISCOLL CM: The effects of pluronics block copolymers and Cremophor EL on intestinal lipoprotein processing and the potential link with P-glycoprotein in Caco-2 cells. *Pharm. Res.* (2003) **20**(7):1085-1092.

90. JOHNSON BM, CHARMAN WN, PORTER CJ: An *in vitro* examination of the impact of polyethylene glycol 400, Pluronic P85, and vitamin E D-α-tocopheryl polyethylene glycol 1000 succinate on P-glycoprotein efflux and enterocyte-based metabolism in excised rat intestine. *AAPS PharmSci* (2002) **4**(4):40.

91. BATRAKOVA EV, LI S, MILLER DW, KABANOV AV: Pluronic P85 increases permeability of a broad spectrum of drugs in polarized BBMEC and Caco-2 cell monolayers. *Pharm. Res.* (1999) **16**(9):1366-1372.

92. MILLER DW, BATRAKOVA EV, KABANOV AV: Inhibition of multidrug resistance-associated protein (MRP) functional activity with pluronic block copolymers. *Pharm. Res.* (1999) **16**(3):396-401.

93. BOGMAN K, ERNE-BRAND F, ALSENZ J, DREWE J: The role of surfactants in the reversal of active transport mediated by multidrug resistance proteins. *J. Pharm. Sci.* (2003) **92**(6):1250-1261.

94. DEMEL RA, GEURTS VAN KESSEL WS, ZWAAL RF, ROELOFSEN B, VAN DEENEN LL: Relation between various phospholipase actions on human red cell membranes and the interfacial phospholipid pressure in monolayers. *Biochim. Biophys. Acta* (1975) **406**(1):97-107.

95. BLUME A: A comparative study of the phase transitions of phospholipid bilayers and monolayers. *Biochim. Biophys. Acta* (1979) **557**(1):32-44.

96. NERVI P, LI BLATTER X, SEELIG A: Surfactants as P-glycoprotein substrates and inhibitors. (In preparation).

97. BRAVO GONZALEZ RC, BOESS F, DURR E, SCHAUB N, BITTNER B: *In vitro* investigation on the impact of Solutol HS 15 on the uptake of colchicine into rat hepatocytes. *Int. J. Pharm.* (2004) **279**(1-2):27-31.

98. BITTNER B, GUENZI A, FULLHARDT P, ZUERCHER G, GONZALEZ RC, MOUNTFIELD RJ: Improvement of the bioavailability of colchicine in rats by co-administration of D-α-tocopherol polyethylene glycol 1000 succinate and a polyethoxylated derivative of 12-hydroxy-stearic acid. *Arzneimittelforschung* (2002) **52**(9):684-688.

99. DINTAMAN JM, SILVERMAN JA: Inhibition of P-glycoprotein by D-α-tocopheryl polyethylene glycol 1000 succinate (TPGS). *Pharm. Res.* (1999) **16**(10):1550-1556.

100. REBBEOR JF, SENIOR AE: Effects of cardiovascular drugs on ATPase activity of P-glycoprotein in plasma membranes and in purified reconstituted form. *Biochim. Biophys. Acta* (1998) **1369**(1):85-93.

101. LO YL: Relationships between the hydrophilic-lipophilic balance values of pharmaceutical excipients and their multidrug resistance modulating effect in Caco-2 cells and rat intestines. *J. Control. Release* (2003) **90**(1):37-48.

102. HIRSCH-ERNST KI, ZIEMANN C, RUSTENBECK I, KAHL GF: Inhibitors of mdr1-dependent transport activity delay accumulation of the mdr1 substrate rhodamine 123 in primary rat hepatocyte cultures. *Toxicology* (2001) **167**(1):47-57.

Website

201. http://www.drugs.com/cons/Poloxamer_188.html
Poloxamer 188 Drug Information. Drugsite Trust, NZ (2006).

Affiliation

Anna Seelig[†1] & Grégori Gerebtzoff[2]
[†]Author for correspondence
[1]Biozentrum, Biophysical Chemistry, University of Basel, Klingelbergstrasse 70, CH-4057 Basel, Switzerland
Tel: +41 61 267 22 06; Fax: +41 61 267 21 89; E-mail: Anna.Seelig@unibas.ch
[2]F. Hofmann–La Roche Ltd, Grenzacherstrasse 124, CH-4070 Basel, Switzerland

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

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.

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 $pK_a < 10$ for bases and $pK_a > 4$ 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
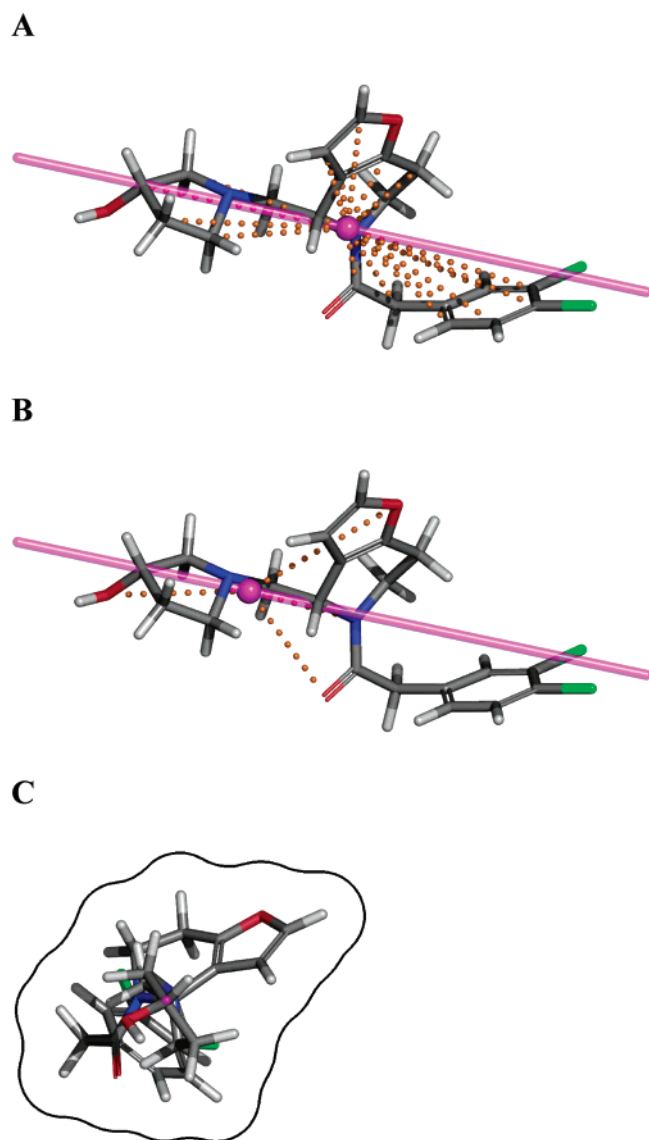
A

B

C

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 $z = 0-1$

$$C_C = 100z + C_O(1 - z) \tag{1}$$

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
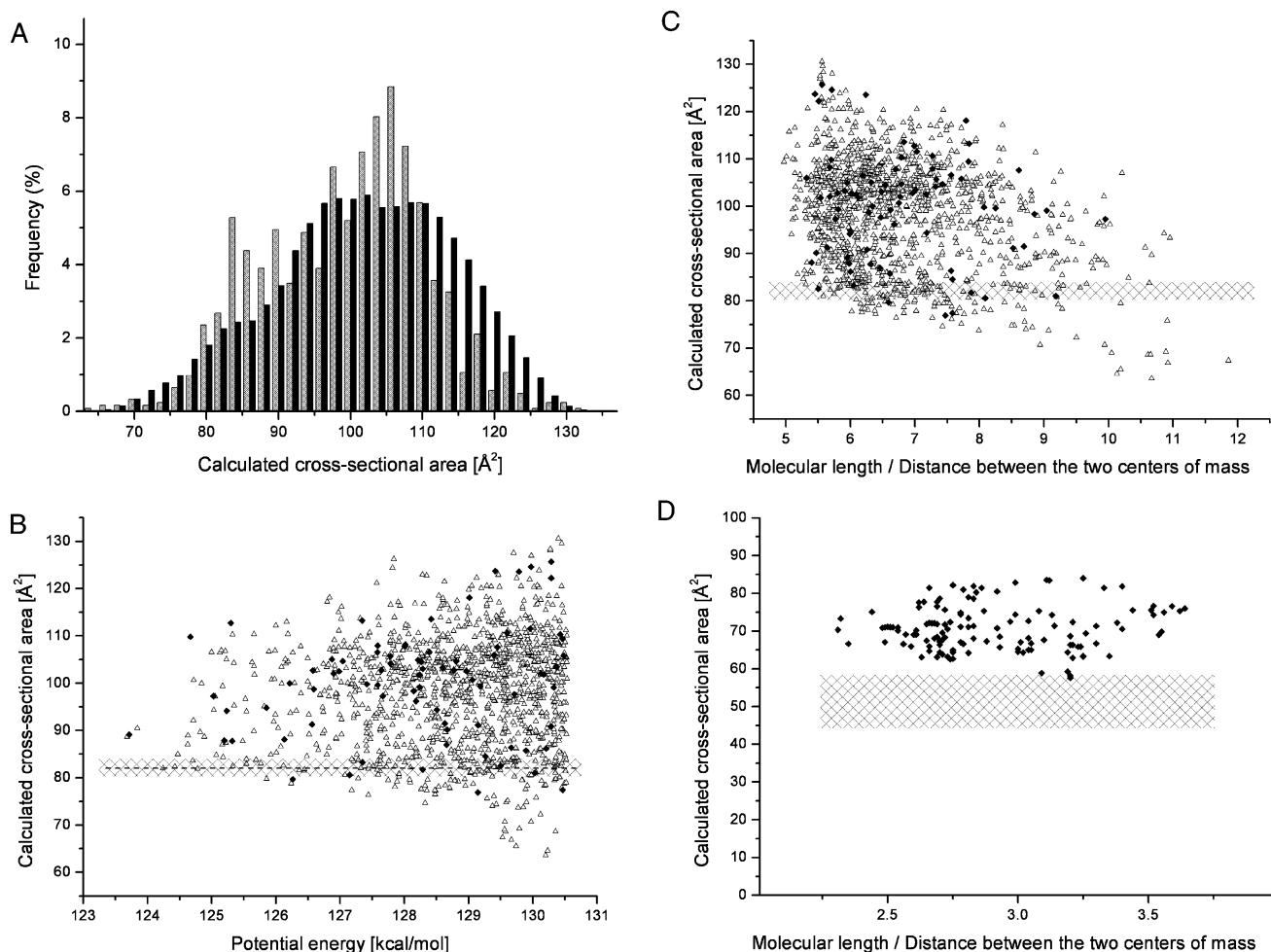
Figure 2. Calculated cross-sectional areas of verapamil (A—C) and fluoxetine (D). (A) The 257 647 conformations generated with the systematic search (black bars), and 1309 energy-minimized conformations generated with the stochastic search (grey bars); three of the latter conformations are shown in Figure 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
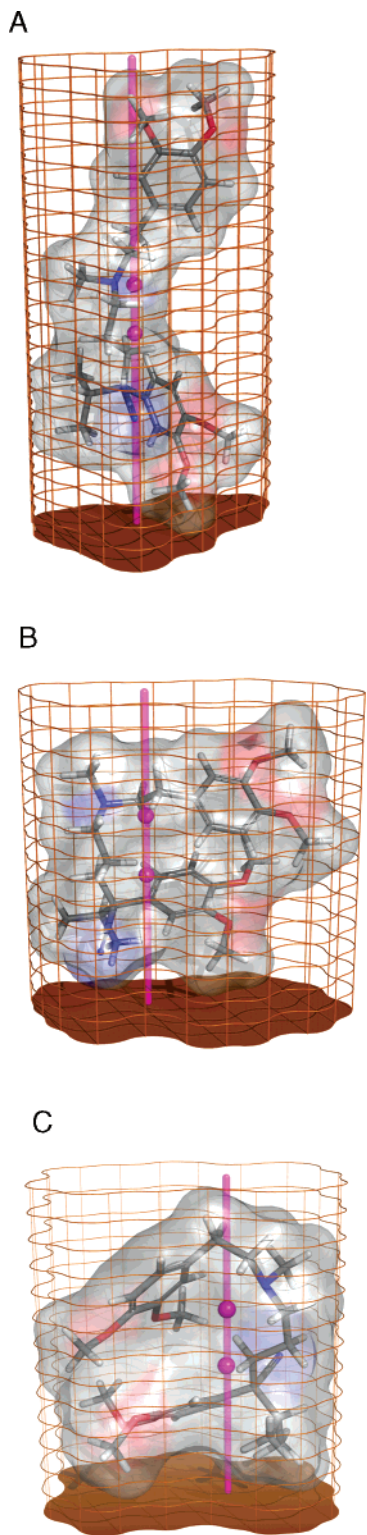
Figure 3. Three-dimensional conformations of verapamil. (A) Conformation yielding the smallest cross-sectional area. (B) Conformation with the lowest potential energy. (C) Smallest amphiphilic conformation, corresponding to the membrane-bound conformation (A_{DcalcM}). All figures were rendered with 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

activity	no	name	MW	log P	$pK_{a(base)}$	$pK_{a(acid)}$	A_{DcalcM} [Å²] pH 7.4	A_{Dexp} [Å²] 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
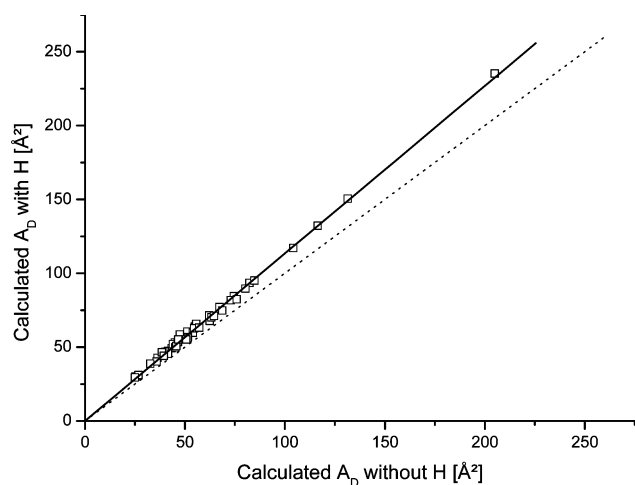
Figure 4. Comparison between the cross-sectional areas calculated with, $A_{Dcalc+H}$, and without, $A_{Dcalc-H}$, hydrogen atoms for 46 compounds. The straight line represents the linear regression of the data, and the dotted line, with the slope of 1, was drawn to guide the eye.

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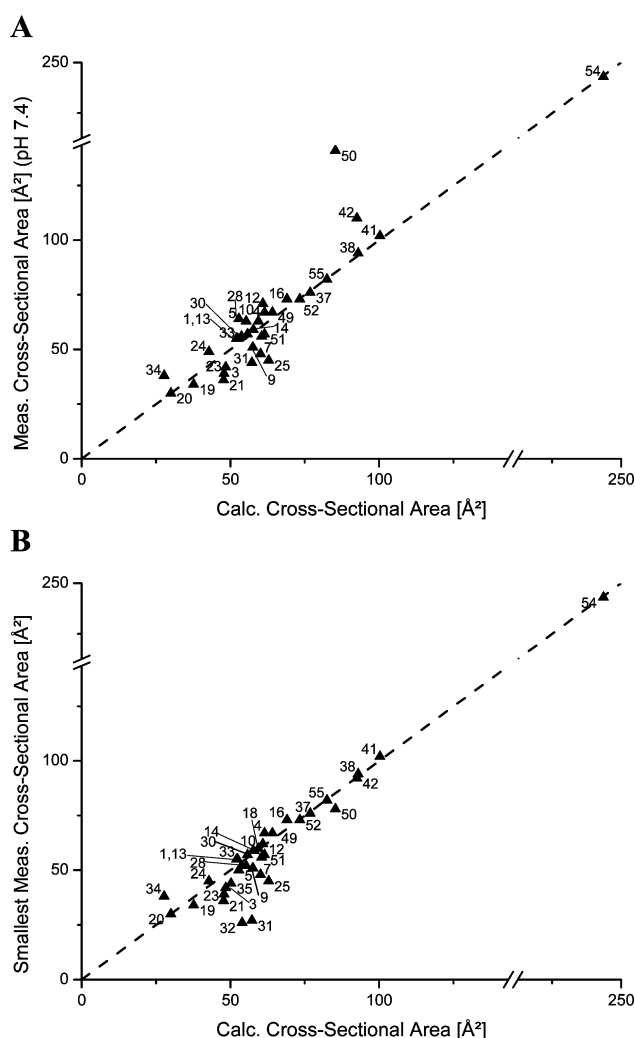
Figure 5. Plot of the measured versus the calculated cross-sectional area of the membrane-bound conformation. The labels refer to the compounds in Table 1. The dashed lines correspond to a slope of 1. Measurements were performed at pH 7.4 (A) and pH 8.0 (bases) or pH 6.8 (acids) (B).

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]
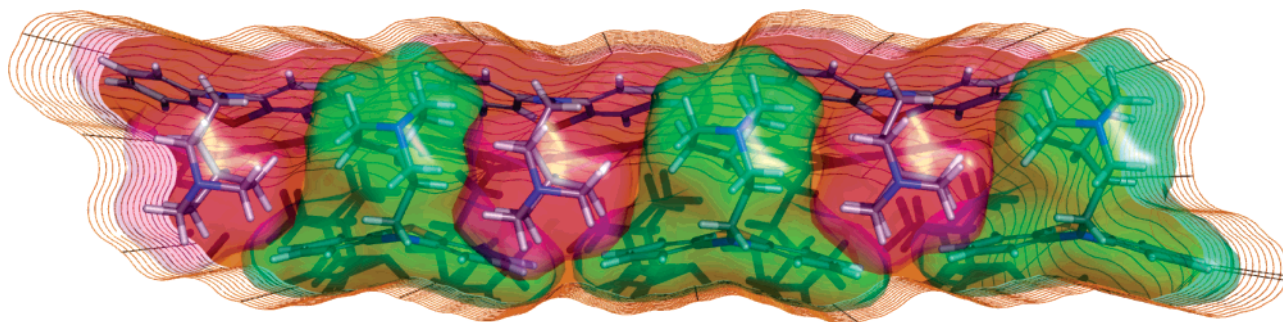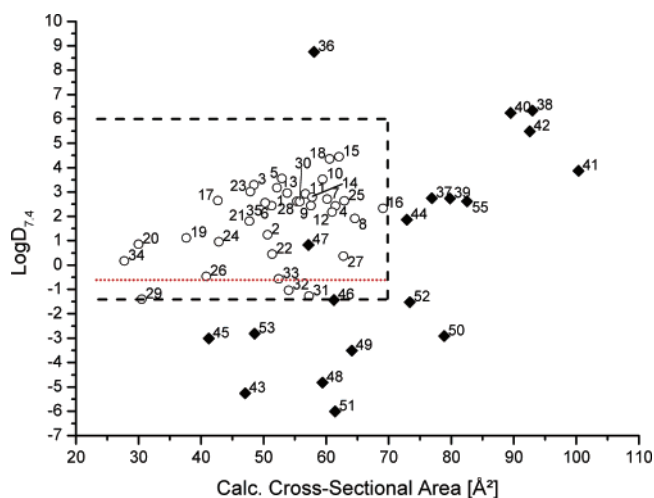

Figure 7. Log D calculated at pH 7.4 versus the calculated cross-sectional area of the membrane-bound conformation calibration diagram. Open circles are BBB⁺ compounds; closed symbols are BBB⁻ compounds. Labels refer to compounds in Table 1. The dashed lines correspond to the limits set between BBB⁺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

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 ($A_{DcalcM} < 70$ Å² 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
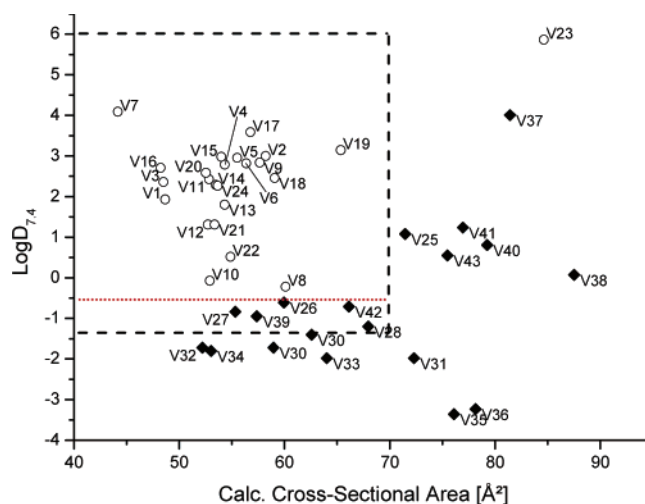
Figure 8. Log D calculated at pH 7.4 versus the calculated cross-sectional area of the membrane-bound conformation validation diagram. The dashed lines correspond to the limits set between BBB⁺ and BBB⁻ compounds. Refer to Figure 7 for the symbol legend.

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 Å² 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} [Å²]
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

PREDICTION OF BLOOD−BRAIN BARRIER PERMEATION

J. Chem. Inf. Model., Vol. 46, No. 6, 2006 **2647**

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} [Å²]	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

PREDICTION OF BLOOD−BRAIN BARRIER PERMEATION

J. Chem. Inf. Model., Vol. 46, No. 6, 2006 **2649**

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.

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 Through-put 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 Non-charged 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 Perme-ation: 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 Coef-ficients of Hydrophobic Organic Chemicals: Comparison with 1-Oc-tanol−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 Prob-lem. I. The Electron Configuration of Benzene and Related Com-pounds. *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. *Biochem-istry* **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.

Grégori GEREBTZOFF
F. Hoffmann-La Roche Ltd.
Bldg. 92/1.56C   CH-4070 Basel
Tel.: +41 (0)61 68**8 93 69**
E-mail: gregori.gerebtzoff@roche.com

Date of birth: January 6th, 1977
Place of birth: Liège (Belgium)
Nationality: Belgian

CURRENT POSITION

Postdoctoral fellow in the *Molecular Properties – Structure-properties correlation* group of
Dr. Manfred Kansy, F. Hoffmann-La Roche Ltd. (Roche), Basel

PUBLICATIONS & AWARDS

Deleuze-Masquefa, C., Gerebtzoff, G., et al. Design and synthesis of novel imidazo[1,2-a]
quinoxalines as PDE4 inhibitors. *Bioorg Med Chem* 2004 12(5)

Gerebtzoff, G., Li Blatter, X., et al. Halogenation of drugs enhances membrane binding and
permeation. *ChemBioChem* 2004 5(5)

Gerebtzoff, G, and Seelig, A. In silico prediction of blood-brain barrier permeation using the
calculated molecular cross-sectional area as main parameter. *J Chem Inf Model* 2006 46(6)

Seelig, A., and Gerebtzoff, G. Enhancement of drug absorption by noncharged detergents through
membrane and P-glycoprotein binding. *Expert Opin Drug Metab Toxicol* 2006 2(5)

Gerebtzoff, G, and Seelig, A. A high quality descriptor for the prediction of drug's passive diffusion
through biological membrane. Winner of the Basel Award for the Best Poster at MipTec 2005

EDUCATION

2002-2006:
Biozentrum, University of Basel
Ph.D.: Division of Biophysical Chemistry, laboratory of PD Dr. A. Seelig
In silico prediction of blood-brain barrier permeation and P-gylcoprotein activity

2001-2002:
University Montpellier II, Montpellier *(France)*
Fifth year postgraduate certificate: "Interface between Biology and Chemistry: molecular
systems with therapeutic aiming", passed with distinction

1999-2001:
University François Rabelais, Tours *(France)*
Degree and Master's degree in Biochemistry, passed with distinction

1997-1999:
University Louis Pasteur, Strasbourg *(France)*
1st and 2nd year of Biophysical Chemistry (University diploma awarded after two years of graduate
study)

1996:
Lycee Jean Mermoz, Saint-Louis *(France)*
General Scientific Baccalaureate (equivalent to 'A' level), with specialization in mathematics,
physics, biology and chemistry

EXPERIENCE

January 14th - June 28th 2002:
Centre of Structural Biochemistry and Organic Chemistry Laboratory, Montpellier *(France)*
Determination of potential inhibitors of phosphodiesterase: purification, activity measurements,
establishment of structure-activity relationship, molecular docking (FlexX, working under Unix
on a SGI station)

June 12th - September 29th 2000, June 11th - September 14th 2001 and
July 29th - August 30th 2002:
F. Hoffmann-La Roche Ltd. (Roche), Basel
Automation of screening tests on a robotized pipeting station, adapted for use with a
spectrophotometer, implementing transgenic cells overexpressing the human P-glycoprotein
(Schwab, D. *et al.* J Med Chem, 2003. 46(9)). Programming of software to convert data from ASCII
into Excel format (Visual Basic); laboratory notebook converted to HTML form

July 2Q[th] - September 30th 1999:
Wunschheim & Partner GmbH, Basel
Realization of a website (HTML, ASP); data-processing maintenance (on Microsoft Windows NT4 Workstation) for the Hoffmann-La Roche Ltd. Company

October 9th 1996 - July 31st 1997:
Novartis, Basel
Part-time employee in the electronic mail department of the Information Services division. Programming of automation systems for the monitoring and the maintenance, with Microsoft Visual BASIC and Microsoft Visual C++, under Microsoft Windows NT

September 9th - October 4th 1996:
Ciba-Geigy, Basel
Employee within the Information Services division, department of data-processing support (using Microsoft Word, Excel, Access). Duties included programming of Excel and Word macros in VBA

July 1st - 6th and August 8th - September 1st 1994:
Ciba-Geigy, Groot-Bijgaarden *(Belgium)*
Employee within the pharma division, department of maintenance and data-processing support (Microsoft Word, Excel, Access)

August 1993:
Baden-Wurtemberg *(Germany)*
Agricultural operator on a pig and cereal farm

August 1992:
Canton of Bern
Agricultural operator in the agricultural federal service

GENERAL AND BIOCHEMICAL SKILLS
- General biochemical techniques
- Pharmacia FPLC
- Cell Culture
- Scheduling
- Teamwork
- Flexibility
- Inquisitiveness
- Problem Solving

COMPUTER SCIENCE SKILLS
- Programming Languages: VB .NET, VBA, HTML; occasional use of Perl, C++, PHP, ASP, MySQL, SVL, C#
- Software and OS: Microsoft Office, Windows XP Pro, Unix …

LANGUAGE SKILLS
- German: Good oral and written, several jobs in German-speaking Switzerland
- English: Good oral and written, linguistic course in Reading (UK), 1995
- French: Mother tongue

EXTRACURRICULAR ACTIVITIES
- Paragliding: confirmed pilot's licence, federal two-seater qualification
- Deep-sea diving: level 2; climbing; badminton (Badminton Club Aesch)
- Former gymnast member, in charge of the young people commission in 1996-97
- Webmaster of the website for the former students of Biochemistry in Tours (www.biochimie.fr.fm) made with PHP and MySQL

REFERENCES

Dr. Marie-Paule Strub	Prof. Dr. Jörg Huwyler	Prof. Dr. Anna Seelig
LBC, National Institute of Health	Institute of Pharma Technology	Biophysikalische Chemie
50 Center Drive	University of Applied Sciences	Biozentrum, Universität Basel
Bethesda, MD 20892-8013 (US)	4132 Muttenz (Switzerland)	4056 Basel (Switzerland)
+1 301 451 4071	+41 (0)61 467 46 89	+41(0)61 267 22 06
strubm@nhlbi.nih.gov	joerg.huwyler@fhnw.ch	anna.seelig@unibas.ch