Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Pharmaceutics 304 (2005) 115–123
international journal of pharmaceutics
www.elsevier.com/locate/ijpharm

Predicting permeability coefficient in ADMET evaluation by using different membranes-interaction QSAR

Jianzhong Liu*, Yi Li, Dahua Pan, Anton J. Hopfinger

Laboratory of Molecular Modeling and Design (M/C 781), College of Pharmacy, The University of Illinois at Chicago, 833 South Wood Street, Chicago, IL 60612-7231, USA

Received 8 March 2005; received in revised form 7 June 2005; accepted 1 August 2005
Available online 22 September 2005

Abstract

Membrane-interaction quantitative structure activity relationship (MI-QSAR) analysis was applied to a data set with 18 compounds in 18 different membranes. MI-QSAR was used to estimate the ADMET properties including the transport of organic solutes through biological membranes. The most important descriptors are the aqueous solvation free energy, FH_2O, and diffusion coefficient for all membranes. The correlation coefficient, r^2, and cross-validation correlation coefficient, q^2, for DMPG membrane is 0.850 and 0.770, respectively. The relationship between FH_2O and permeability is nonlinear. But the detail effect of aqueous solvation free energy and diffusion coefficient to the permeability depends on the type of membrane. The final models also support the solution–diffusion mechanism of transport is important in membrane.

Keywords: Molecular dynamic simulation; Permeation coefficient; Diffusion coefficient; QSAR

1. Introduction

Drug discovery and development is an extremely time-consuming and costly process. For every drug that reaches the market, there are more than 10,000 compounds synthesized, characterized, and tested for biological effects. Hundreds of millions of dollars are invested in basic research and clinical studies which lead to the FDA approval and subsequent marketing of a new drug. Traditionally, drugs have been "discovered" predominantly through random, or targeted, screening efforts, followed by small structural changes in the lead molecule to optimize the properties responsible for the desired activity. It takes between 7 and 10 years from initiation of the preclinical drug discovery program to initial marketing of resulting drug (excluding regulatory delays) (Smith, 2002).

The study conducted to determine the expense of bringing a drug to the pharmaceutical market took place in 1979, at which time the cost was estimated to be US$ 54 million (Outlook, 2003). The most recent

* Corresponding author. Present address: Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA. Tel.: +1 302 831 3522; fax: +1 302 831 6335.
E-mail address: zhong@udel.edu (J. Liu).

0378-5173/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2005.08.003

data indicates that this figure has ballooned to US$ 897 million in 2003 (Kaitin, 2003). Moreover, despite the tremendous development of many new technologies, and a 200-fold increase in initial screening capacity over the past decade, the number of drug candidates entering development remains relatively unchanged over the past 20 years. The fundamental problem seems to be the lack of quality in the lead compounds and the corresponding low success rate in the attempted development of these compounds into drug candidate. It is clear that the pharmaceutical industry needs new approaches to lead discovery that will generate lead compounds that will be translate into a higher rate of success in subsequent development efforts.

Computer-aided molecular modeling is still a relatively new field, with exciting new methods and applications being reported at a staggering rate. New methods of performing pharmacophore searching, docking, structure-based design (SBD), quantitative structure–activity relationship (QSAR), quantitative structure–property relationship (QSPR), and molecular similarity (MS) comparison studies are continually being proposed. Some of the approaches are ligand-based (Liu et al., 2003), others are receptor-dependent (Pan et al., 2004). Many statistical methods are also used in computer-aided drug design (CADD), such as multiple linear regression (Ecker et al., 1996), partial least-square (PLS) regression (Glen et al., 1989), genetic algorithms (Fleischer et al., 1997), cluster analyses (Cronin, 1996) and artificial neural networks (Winkler and Burden, 2000; Burden et al., 2000; Jalali-Heravi and Parastar, 2000).

It has been suggested that computational models for reliable prediction of ADMET properties are promising as early screening tools for drug candidates and for designing more successful combinatorial libraries (Rose and Stevens, 2003; Winkler and Burden, 2002). In our preceding study (Iyer et al., 2002), we successfully applied membrane-interaction quantitative structure activity relationship (MI-QSAR) analysis to develop predictive models of blood–brain barrier partitioning of organic compounds by simulating the interaction of an organic compound with phospholipids-rich regions of cellular membranes with resulting significant MI-QSAR models ($r^2 = 0.845$, $q^2 = 0.795$). These computational models have been applied to high throughput screening and were demonstrated to be highly predictive. MI-QSAR analysis also has been successfully applied to construct robust models for both eye and skin irritation for a structurally diverse training set (Kulkarni et al., 2001), and to predict the Caco-2 cell permeability of a diverse set of drugs (Kulkarni et al., 2002). It has subsequently occurred to us that MI-QSAR is ideally suited for construction QSAR models for membrane uptake and transport measures so that it can be used to evaluate the absorption and distribution properties in the early stage of drug development.

It is known that solute transport through a polymer membrane is either via the pore or partition mechanism. In the pore mechanism, the solute diffuses through the water filled pores and in the partition mechanism the solute transport is presumed to occur by a process involving the dissolution of the solute within the polymer followed by the diffusion through the membrane. Recent experimental studies of membrane lipids have produced increasingly detailed pictures of the structure of fluid phase lipid at the level of mean positions and conformations of constituent molecular groups. These precise experiments yielded structure details that arise from the complex set of intermolecular interactions. Computer simulation is the only theoretical approach capable of probing the nature of the interactions between atoms on lipid and water molecules, and the mechanism by which these interactions lead to the observed structures. Studies of Caco-2 cell permeability in dimyristoylphosphatidylcholine (DMPC) membrane have examined the atomic level structure of this system. By doing simulations that are consistent with available experimental data, we have already gained an atomic level structural and dynamical picture of the system. Extension of the database of predictions from simulations to a wider class of biologically relevant lipid systems is the aim of the work presented here.

Atomic level simulations of lipid of dipalmitoylphosphatidylcholine (DPPC) and even more complex systems, including lipid–cholesterol and lipid protein systems have been carried out by several groups over the past several years (Essmann and Berkowitz, 1999; Katragadda et al., 1999; Petrache et al., 2002). Our studies in DPPC, dimyristoylphosphatidylcholine (DMPC) and dipalmitoylphosphatidic acid (DPPA) cumulated plenty of experience in membrane dynamic simulation (Klein et al., 1999) and provided the impetus for us to go further. The goal of the study reported here is to demonstrate the applicability of MI-QSAR

analysis in different phospholipids and to predict the drug ADMET properties. We explored structurally diverse drugs and their interactions with variety of phospholipids through molecular dynamics simulation. The ability of a molecule to permeate cell membranes by passive diffusion is primarily dependent on its partitioning into the membrane layer. The most frequently used physicochemical property to represent this partitioning, and the prediction of cellular permeability, is the log of the (1-octanol/water) partition coefficient, log *P*. Other descriptors and modeling/QSAR work has been done to better understand cell permeability. The hydrogen bonding capacity and molecular surface properties of the solute have been used to construct correlation models for cell permeability. Based on our experience in previous studies, our focus here is to construct a model using as few as possible descriptors.

2. Materials and methods

2.1. Permeation coefficients

The dependent variable used in MI-QSAR analysis is the Caco-2 cell permeability coefficient. Yazdanian et al. (1998) performed permeability experiments on a data set of 38 structurally and chemically diverse drugs ranging in molecular weight from 60 to 515 amu and varying in net charge at pH 7.4. We used only 18 of them as a training set, which proportionally distributed between the maximum and minimum value of the permeability coefficient and fully represents the whole dataset. Table 1 contains the permeability coefficient values for 18 structurally diverse drugs used as the training set of compounds.

Table 1
Drug name, structure, and permeability coefficient for the training set

Compound	Permeability (10^{-6} cm/s)
Acyclovir	0.25
Alprenolol	25.3
Atenolol	0.53
Bremazocine	8.02
Caffeine	30.8
Clonidine	21.8
Dexamethasone	12.2
Diazepam	33.4
Dopamine	9.33
Ganciclovir	0.38
Metoprolol	23.7
Nadolol	3.88
Phenytoin	26.7
Pindolol	16.7
Salicylic acid	22.00
Sulfasalazine	0.30
Terbutaline	0.47
Timolol	12.8

2.2. MI-QSAR analysis applied to the training set

2.2.1. Step 1: Building solute molecules and phospholipids monolayer

All the solute molecules of the training and test sets, see Table 1, were built using HyperChem 6.03 software. Partial atomic charges were computed using the AM1 semi-empirical method, which was implemented in the HyperChem program. Each structure was energy minimized using the quantum mechanical method, also implemented in HyperChem, without any geometric constraint. The energy-minimized structures were used as the initial structures in conformational sampling.

Totally 18 different phospholipids were selected as model phospholipids in this study, which are listed in Table 2. For those phospholipids with available atomic coordinates, their structures were constructed based on the atomic coordinates. Take dimyristoylphosphatidylglycerol (DMPG, Fig. 1) as example: its structure was

Table 2
Phospholipids used to construct the membrane monolayer

DMPE	Dimyristoylphosphatidylethanolamine
DMPG	Dimyristoylphosphatidylglycerol
DMPI	Dimyristoylphosphatidylinositol
DOPC	Dioleoylphosphatidylcholine
DOPE	Dioleoylphosphatidylethanolamine
DOPS	Dioleoylphosphatidylserine
DPPC	Dipalmitoylphosphatidylcholine
DPPE	Dipalmitoylphosphatidylethanolamine
DPPG	Dipalmitoylphosphatidylglycerol
DPPI	Dipalmitoylphosphatidylinositol
DPPS	Dipalmitoylphosphatidylserine
DSPC	Distearoylphosphatidylcholine
DSPE	Distearoylphosphatidylethanolamine
DSPG	Distearoylphosphatidylglycerol
PDHS	*N*-Palmitoyldihydrosphingomyelin
PSPM	*N*-Palmitoylsphingomyelin
SPM240	*N*-Tetracosanoicsphingomyelin
SPM241	*N*-Tetracosenoicsphingomyelin

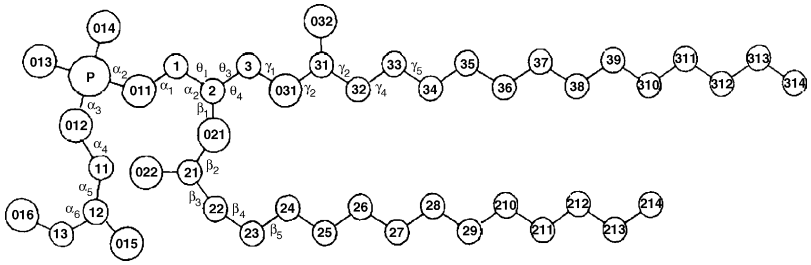

Fig. 1. Numbering of atoms and torsion angles for the DMPG molecule according to Pashcher. Atoms other than P, O, and N are carbon atoms with 1, 2, or 3 hydrogen atoms attached.

constructed in HyperChem 6.03 based on the crystal structure atomic coordinates from Pascher et al. (1987). Calculations at the semi-empirical AM1 level were carried out, and DMPG was energy minimized using the quantum mechanical method. An assembly of 25 DMPG molecules ($5 \times 5 \times 1$) in (x,y,z) directions, respectively, was used as the model membrane monolayer. The size of the monolayer simulation system was constructed according to the symmetry of the crystal structure. The cell parameters for an individual phospholipids molecule were $10.5\,\text{Å} \times 8.5\,\text{Å} \times 45\,\text{Å}$, $\gamma = 95.2$. These parameters result in an average surface area per phospholipid of $89.25\,\text{Å}^2$, which is close to the reported value of about $88\,\text{Å}^2$ for the fully hydrated fluid lamellar phase of DMPG. It was found that the estimated order parameters for these two model bilayers agree with one another suggesting that smaller assembly is adequate for modeling short-range properties.

Phospholipids for which atomic coordinates are not available, their structure were constructed based on the atomic coordinates of their analogues. For example, DSPC was built from the crystal structure atomic coordinates of DMPC by extending the four -CH_2- groups in a *trans* conformation to each of two chains.

When surface areas are known for selected phospholipids the data is used to assembly the monolayer. If surface areas are not available, the head group volume must be the same for conditions where headgroups are chemically identical. The Eq. (1) (Sugar, 1979) was used to obtain relating surface area from its analogues and then to build the structure.

$$S = \frac{n\Sigma}{\cos\varphi} \quad (1)$$

where S is the molecular surface area, n the number of hydrocarbon chains per lipid molecule, Σ the cross-section of the hydrocarbon chains perpendicular to the chain axis, and φ is the angle of the tilts between the hydrocarbon chain axis and the layer normal.

2.2.2. Step 2: Docking

To prevent unfavorable van der Waals interactions between a solute molecule and the membrane molecules, the center phospholipid molecule, located at position (x,y) = (3,3) of the 5×5 monolayer, and a test solute molecule was docked into the space created by the missing phospholipid molecule. Each of the test solute molecules of the permeation data set was docked at three different positions in each phospholipid monolayer with the most polar group of the solute molecule facing toward the headgroup region of the monolayer. Three corresponding MDS models were generated for each solute molecule with regard to the trial positions of the solute molecule in the monolayer, which were: (1) solute molecule in the headgroup region, (2) solute molecule between the headgroup and the aliphatic chains, and (3) solute molecule in the tail region of the aliphatic chains. The three different initial MDS positions of acyclovir, one of the training set solute molecules in DMPG, are shown in Fig. 2a to illustrate this dock.

2.2.3. Step 3: Energy minimization

The energy minimization and following molecular dynamic simulation were performed using a Molsim package with an extended MM2 force field. To remove unfavorable high-energy van der Waals interactions between solute and phospholipid molecules, the energy of the system was minimized by a series of steepest

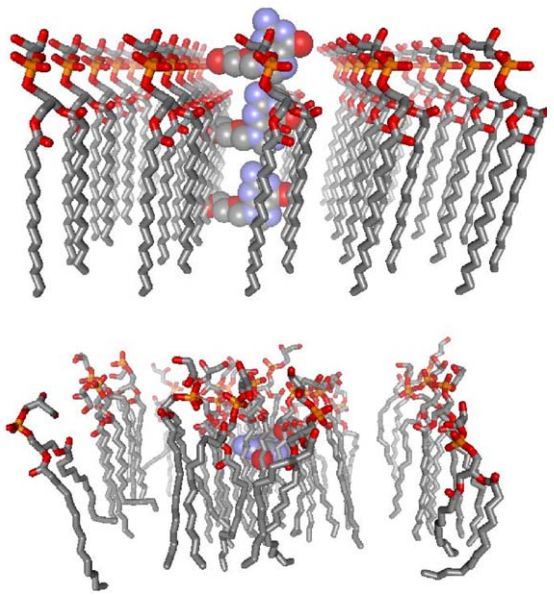

Fig. 2. (a) The side view of an acyclovir molecule docked into DMPG monolayer at three different positions before the energy minimization with hydrogen atoms not shown. (b) The lowest energy geometry of a DMPG–acyclovir complex during MDS with hydrogen atoms not shown.

descent and conjugate gradient minimization steps. The energy convergence criterion was a gradient of less than 0.5 kcal/(Å mol). Convergence was generally achieved within less than 10 ps.

2.2.4. Step 4: Molecular dynamics simulation

The model monolayer was first heated to 20 K and then to 50 K and from that point in increments of 50 K to a final temperature of 311 K. At each temperature increment, 2 ps of MDS was carried out to allow for structural relaxation and distribution of kinetic energy throughout the system. After 311 K temperature was achieved, the whole system was held at this temperature and production trajectories were recorded every 0.1 ps over a 50 ps range for each solute of the dataset. Concurrently, the potential energy of the whole system and the interaction of each phospholipid and the solute were also recorded. Only a single solute molecule in one kind of membrane was explicitly considered in each MDS. Each of the solute molecules was placed at each of the three different positions in the monolayer as described above. The lowest energy geometry of the solute molecule in the monolayer was sought

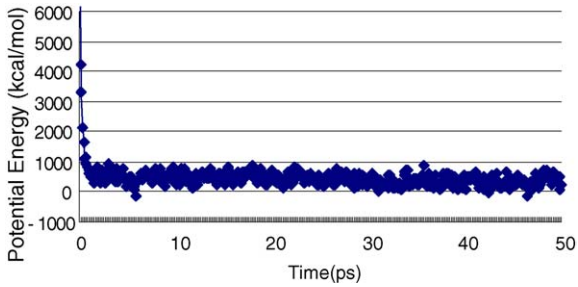

Fig. 3. Plot of total potential energy vs. time for acyclovir embedded in the model of DMPG monolayer.

using each of the three trial solute positions. The energetically most favorable geometry of acyclovir in the model DMPG monolayer is shown in Fig. 2b. And the MDS trajectory of acyclovir in the model DMPG monolayer is also shown in Fig. 3. It is obvious from the plot that this complex system deceases its potential energy and that equilibrium is achieved gradually.

2.2.5. Step 5: Calculation of descriptors

Many experiments (Patel and Manley, 1995; Stern et al., 1993) suggest that the permeability coefficient (Pm) can be estimated as the product of the solubility and diffusion coefficients, which has a relationship as Eq. (2),

$$\mathrm{Pm} = KD_{\mathrm{m}} \tag{2}$$

where D_{m} is the membrane diffusion coefficient, K is the membrane-donor partition coefficient.

We have previously demonstrated (Kulkarni et al., 2002.) that the permeation coefficient of solutes correlates with the aqueous solvation free energy, F(H_2O) ($r^2 = 0.75$, $q^2 = 0.71$). Our purpose in this study are to correlate the relationship between permeation coefficient and diffusion coefficient by using dynamic parameters based on other membrane phospholipids also to correlate this relationship with as few as possible descriptors. Table 3 lists the calculated descriptors in

Table 3
The descriptors used in the MI-QSAR descriptor pool

FH_2O	The aqueous solvation free energy
FOCT	The 1-octanol solvation free energy
log P	The 1-octanol/water partition coefficient
Ecoh	The cohesive packing energy of the solute molecules
D	Diffusion coefficient of a solute in the membrane

this study. It should be noted that FH_2O, FOCT, and log P are computed using intra-molecular computational methods. This is also true for Ecoh, the cohesive energy, which is measurement of the energy required to remove a molecule from being surrounded by other molecules identical to itself.

The mean square displacement method was used to calculate diffusion coefficients. In a molecular system, a molecule moves in three dimensions and its motion of an individual molecule does not follow a simple path. If the path is examined in close detail, it will be seen to be a good approximation to a random walk. Mathematically, Einstein showed that mean square displacement grows linearly with time, which is shown in Eq. (3) (Einstein, 1926):

$$< r^2 >= 6Dt + C \tag{3}$$

where $<r^2>$ is the mean square distance and t is time. D and C are constants. The constant D is the most important of these and defines the diffusion rate. It is called the diffusion coefficient. Also, since we have many atoms to consider we can calculate a square displacement for all of them. This is what makes the mean square displacement (or MSD for short) scientifically significant. Through its relation to diffusion it is a measurable quantity, one which relates directly to the underlying motion of the molecules. For each trajectory frame in the MDS trajectory, the average movement Δd_i of a solute molecule, relative to the previous frame, was determined.

2.2.6. Step 6: Construction of a model for the membrane

MI-QSAR models for each membrane were built and optimized using multidimensional linear regression fitting and the genetic function approximation (GFA) (Rogers and Hopfinger, 1994), which is a multidimensional optimization method based on the genetic algorithm paradigm. Both linear and quadratic representations of each of the descriptor values were included in the trial descriptor pool, and MI-QSAR models were built as a function of the number of descriptor terms in the model. Statistical significance in the optimization of an MI-QSAR model was based on Friedman's lack of fit (LOF) measure. The LOF measure is designed to resist overfitting which is a problem often encountered in construction of statistical models. Detailed descriptions and methods about the GFA were published in previous papers (Rogers and Hopfinger, 1994; Rogers, 1991). The best model was judged jointly by the correlation coefficient of fit, r^2, and the leave-one-out cross-validation correlation coefficient, q^2.

3. Results

The aqueous solvation free energy, FH_2O has been shown to correlate to aqueous solubility as would be expected. For all phospholipids membranes, the one-term models are same, which contain only FH_2O, and are described by:

$$P = 55.212 - 31.634 \times FH_2O;$$
$$n = 18,\ r^2 = 0.686,\ q^2 = 0.635 \tag{4}$$

Here, n is the number of the trials, r^2 the correlation coefficient, and q^2 is the leave-one-out cross-validation correlation coefficient. Because FH_2O is an intra-molecular solute descriptor (i.e. it is independent of membrane or interaction of the solute-membrane complex system, for all membrane monolayers, the one-term model remains constant.

By taking into account both first order and second order terms, the following model produces higher r^2 and q^2 value for all membranes than model (4) has:

$$P = 40.156 - 14.653 \times FH_2O^2;$$
$$n = 18,\ r^2 = 0.764,\ q^2 = 0.704 \tag{5}$$

This model suggests a nonlinear relationship between FH_2O and the permeability coefficient. Eq. (5) is mathematically better than Eq. (4) based on r^2 value, and it is also more statistically significant based on q^2 value.

Combining the FH_2O with other descriptors by using nonlinear expression, all of phospholipids monolayers produce statistically significant prediction model than only using FH_2O square as descriptor. Take DMPG as example, the model is:

$$P = 29.613 - 52.320 \times (FH_2O - 0.826)^2 + 0.455 \times D;\quad n = 18,\ r^2 = 0.850,\ q^2 = 0.770 \tag{6}$$

This shows that the model has significant improvement both on correlation coefficient, r^2, and

Table 4
The equation of permeability coefficient relating to FH_2O and diffusion coefficient in different membranes

Membrane	β_1	a	β_2	r^2	q^2
DMPE	−51.984	0.817	−1.591	0.853	0.777
DMPG	−52.320	0.826	0.455	0.850	0.770
DMPI	−51.679	0.821	−1.452	0.852	0.755
DOPC	−51.864	0.822	0.556	0.850	0.754
DOPE	−52.304	0.826	0.041	0.850	0.753
DOPS	−50.043	0.817	−3.290	0.858	0.757
DPPC	−57.153	0.856	4.681	0.866	0.798
DPPE	−49.456	0.800	7.263	0.885	0.824
DPPG	−51.477	0.818	2.865	0.859	0.787
DPPI	−52.895	0.832	−1.213	0.851	0.770
DPPS	−54.714	0.848	−3.255	0.859	0.783
DSPC	−50.273	0.814	3.627	0.865	0.802
DSPE	−50.165	0.815	−1.574	0.852	0.783
DSPG	−47.820	0.797	6.617	0.876	0.811
PDHS	−51.842	0.824	−1.003	0.851	0.769
PSPM	−51.633	0.820	−2.247	0.854	0.768
SPM240	−52.867	0.825	1.895	0.853	0.770
SPM241	−52.634	0.829	0.394	0.851	0.764

The general expression is $P = \alpha + \beta_1 \times (FH_2O - a)^2 + \beta_2 \times D$.

cross-validation correlation coefficient, q^2. The addition of diffusion coefficient term enhances the model reliability and predictability. And there is no obvious improvement with further increase of terms. The best values for two-terms FH_2O and D in each membrane are shown in Table 4.

4. Discussion

Our previous finding (Kulkarni et al., 2002.) demonstrated that the aqueous solvation free energy, $F(H_2O)$, has been shown to correlate to aqueous solubility as would be expected. Increasingly negative $F(H_2O)$ values corresponds to increasing aqueous solubility of a solute. This is also in agreement with our study. In our data, the larger the FH_2O value, the higher the aqueous solubility. And this descriptor ranges from 0.56 to 1.64. But the quadratic term about this descriptor suggests that the permeability coefficient decreases with the increase of the aqueous solubility from a value nearly at 0.80, while the permeability coefficient decreases with the decrease of the aqueous solubility below 0.80. This suggests that a drug with suitable aqueous solubility is important for this compound to have good oral absorption ability. Our observations are in agreement with Hilgers' report that cell permeability is not a linear relationship with $\log P$ (Hilgers et al., 1990). An increase in $\log P$ reflects an increase in lipophilicity.

Although the one-term model suggests that the permeability coefficient is linear to FH_2O and the two-term model shows nonlinear relationship, there is no contradiction. GFA algorithm generates multiple QSAR models when it is applied to optimize MI-QSAR models. These different models are multiple interpretations of the data. The "best" model can be selected among the multiple good models, rather than being forced to accept a single arbitrarily chosen "best" model. Similarities and differences among these models can be used to consider alternative mechanisms, which explain the data, or may lead to new experiments to better understand the system being studied.

In this study, the two-term model suggests diffusion coefficient is the second significant term to predict permeability coefficient. The influence of the diffusion coefficient depends on membrane lipids. For example, the regression coefficient of term D is 0.455 in DMPG model. For DPPG and DSPG, which have longer acyl chains, the regression coefficient becomes 2.865 and 6.617, respectively. This result shows that longer acyl chains will enhance the impact of diffusion coefficient on permeability. But this tendency is not obvious for non-glycerol head phospholipids membranes.

For choline head phospholipids membranes, the addition of double bonds into the hydrocarbon chains lowers the value of regression coefficient of term D. There is a cis double bond in each hydrocarbon chain in DOPC, compared to DSPC. However, the regression coefficients for DSPC and DOPC are 3.627 and 0.556, respectively. This tendency is also shown in SPM240. The regression coefficient of the term D is 1.895. It decreases to 0.394 while adding one double bond into the longer hydrocarbon chain (SPM241). This suggests the addition of the double bond will decrease the impact of the diffusion coefficient to the permeability. PDHS has a choline headgroup. It becomes PSPM while adding the double bond into the longer chain. The regression coefficient of term D changed from −1.003 to −2.247.

This phenomenon is not shown in ethanolamine head phospholipids, such as DSPE. The addition of the double bonds changes the regression coefficient of the term D from positive to negative. These models, shown in Table 4, suggest that different membrane monolayers

induce different membrane phospholipids–solute interaction mechanism. Diffusion coefficient is not simply positive or negative proportional to permeability for all membranes. The bigger diffusion coefficient only suggests the solute move quicker within the membrane–solute complex system. It does not mean that the solute passes through membrane quicker from one side to another side because the movement is in the three dimensions. Sometime, the solute moves back and forth within the membranes. This depends on the electrostatic/steric properties of the phospholipids and the interaction between solute and headgroup or acyl chain. But different phospholipids have different inherent electrostatic/steric properties. Also hydration condition of membrane system will produce influence on solute permeability. Our final models support the solution–diffusion mechanism of transport is important in membrane.

It is worthy noting that our models do not show the linear relationship between the permeability coefficient and the product of the solubility and diffusion coefficient as we expected. Also no statistically significant relationship exists between them. The possible reason is that our solubility and diffusion coefficient are empirical calculation. There are errors among these descriptors. The product of these two terms will magnify these errors. Therefore, more sophisticated membrane systems, such as lipid/cholesterol system, mixed lipids/cholesterol system, will be helpful to describe the interaction among solutes and lipid molecules (Falck et al., 2004). Further investigation about solute–lipid complex simulation and more accurate prediction of solubility will be highly significant to resolve this problem. Part of these work are already under the way in our group.

Acknowledgments

We gratefully acknowledge support from the Laboratory of Molecular Modeling & Design at UIC and from The Chem21 Group, Incorporated.

References

Burden, F.R., Ford, M.G., Whitley, D.C., Winkler, D.A., 2000. Use of automatic relevance determination in QSAR studies using Bayesian neural networks. J. Chem. Inf. Comput. Sci. 40, 1423–1430.

Cronin, M.T., 1996. The use of cluster significance analysis to identify asymmetric QSAR data sets in toxicology. An example with eye irritation data. SAR QSAR Environ. Res. 5, 167–175.

Ecker, G., Chiba, P., Hitzler, M., Schmid, D., Visser, K., Cordes, H.P., Csolleri, J., Seydel, J.K., Klaus-Jürgen Schaper, K.J., 1996. Structure–activity relationship studies on benzofuran analogs of propafenone-type modulators of tumor cell multidrug resistance. J. Med. Chem. 39, 4767–4774.

Einstein, A., 1926. Investigations On the Theory of Brownian Movement. Methuen & Co. Ltd., London.

Essmann, U., Berkowitz, M.L., 1999. Dynamical properties of phospholipids bilayers from computer simulation. Biophys. J. 76, 2081–2089.

Falck, E., Patra, M., Karttunen, M., Hyvonen, M.T., Vattulainen, I., 2004. Impact of cholesterol on voids in phospholipid membranes. J. Chem. Phys. 121, 12676–12689.

Fleischer, R., Wiese, M., Troschutz, R., Zink, M., 1997. 3D-QSAR analysis and molecular modeling investigations of piritrexim and analogues. J. Mol. Mod. 3, 338–346.

Glen, W.G., Dunn, W.J., Scott, D.R., 1989. Principal components analysis and partial least squares. Tetrahedron Comput. Methods 2, 349–354.

Hilgers, A.R., Conradi, R.A., Burton, P.S., 1990. Caco-2 cell monolayers as a model for drug transport across the intestinal mucosa. Pharm. Res. 7, 902–910.

HyperChem Program Release 6.03 for windows, Hypercube, Inc., 1999.

Iyer, M., Mishra, R., Han, Y., Hopfinger, A.J., 2002. Predicting blood–brain barrier partitioning of organic molecules using membrane-interaction QSAR analysis. Pharm. Res. 19, 1611–1621.

Jalali-Heravi, M., Parastar, F., 2000. Use of artificial neural networks in a QSAR study of anti-HIV activity for a large group of HEPT derivatives. J. Chem. Inf. Comput. Sci. 40, 147–154.

Kaitin, K.I., 2003. Post-approval R&D raises total drug development costs to $897 million. Tufts Center for the Study of Drug Development Impact Report.

Katragadda, A.K., Singh, M., Betageri, G.V., 1999. Encapsulation, stability, and in vitro release characteristics of liposomal formulations of stavudine (d4t). Drug Deliv. 6, 31–37.

Klein, C.D.P., Klingmueller, M., Schellinski, C., Landmann, S., Hauschild, S., Heber, D., Mohr, K., Hopfinger, A.J., 1999. Synthesis, pharmacological and biophysical characterization, and membrane-interaction QSAR analysis of cationic amphiphilic model compounds. J. Med. Chem. 19, 3874–3888.

Kulkarni, A., Han, Y., Hopfinger, A.J., 2002. Predicting Caco-2 cell permeation coefficients of organic molecules using membrane-interaction QSAR analysis. J. Chem. Inf. Comput. Sci. 42, 331–342.

Kulkarni, A., Hopfinger, A.J., Osborne, R., Bruner, L.H., Thompson, E.D., 2001. Prediction of eye irritation from organic chemicals using membrane-interaction QSAR analysis. Toxicol. Sci. 59, 335–345.

Liu, J., Pan, D., Tseng, Y., Hopfinger, A.J., 2003. 4D-QSAR analysis of a series of antifungal P450 inhibitors and 3D-pharmacophore

comparisons as a function of alignment. J. Chem. Inf. Comput. Sci. 43, 2170–2179.

Outlook, 2003. Tufts Center for the Study of Drug Development, Boston.

Pan, D., Liu, J., Senese, C., Hopfinger, A.J., Tseng, Y., 2004. Characterization of a ligand–receptor binding event using receptor–dependent four-dimensional quantitative structure–activity relationship analysis. J. Med. Chem. 47, 3075–3088.

Pascher, I., Sundell, S., Harlos, K., Eibl, H., 1987. Conformation and packing properties of membrane lipids: the crystal structure of sodium dimyristoylphosphatidylglycerol. Biochim. Biophys. Acta Biomembr. 896, 77–88.

Patel, K., Manley, R., 1995. Carbon dioxide sorption and transport in miscible cellulose/poly(vinyl alcohol) blends. Macromolecules 28, 5793–5798.

Petrache, H.I., Zuckerman, D.M., Sachs, J.N., Killian, J.A., Koeppe, R.E., Woolf, T.B., 2002. Hydrophobic matching mechanism investigated by molecular dynamics simulations. Langmuir 18, 1340–1351.

Rogers, D., 1991. G/SPLINES, A hybrid of Friedman's multivariate adaptive regression splines (MARS) algorithm with Holland's genetic algorithm.

Rogers, D., Hopfinger, A.J., 1994. Application of genetic function approximation to quantitative structure–activity relationships and quantitative structure–property relationships. J. Chem. Inf. Comput. Sci. 34, 854–866.

Rose, S., Stevens, A., 2003. Computational design strategies for combinatorial libraries. Curr. Opin. Chem. Bio. 7, 331–339.

Smith, P., 2002. Chemical development. In: Medicinal Chemistry, Principles and Practice, second ed.

Stern, S.A., Liu, Y., Feld, W.A., 1993. Structure/permeability relationships of polyimides with branched or extended diamine moieties. J. Poly. Sci. Part B Polym. Phys. 31, 939–951.

Sugar, I.P., 1979. A theory of the electric field-induce phase transition of phospholipids bilayers. Biochim. Biophys. Acta. Biomembr. 556, 72–85.

Winkler, D.A., Burden, F.R., 2000. Robust QSAR models from novel descriptors and Bayesian regularized neural networks. Mol. Simulat. 24, 4–6.

Winkler, D.A., Burden, F.R., 2002. Application of neural networks to large dataset QSAR, virtual screening, and library design. Methods Mol. Bio. 201, 325–367.

Yazdanian, M., Glynn, S.L., Wright, J.L., Hawi, A., 1998. Correlating partitioning and Caco-2 cell permeability of structurally diverse small molecular weight compounds. Pharm. Res. 15, 1490–1494.