

Full-length article

MI-QSAR models for prediction of corneal permeability of organic compounds¹

Cheng CHEN, Jie YANG²*State Key Laboratory of Pharmaceutical Biotechnology, College of Life Science, Nanjing University, Nanjing 210093, China*

Key words

cornea; permeability; quantitative structure-activity relationship; theoretical models

¹ Project supported by the National Natural Science Foundation of China (No 30171094 and No 30271497).² Correspondence to Dr Jie YANG.

Phn 86-25-8359-4060.

Fax 86-25-8332-4605.

E-mail luckyjyj@sina.com.cn

Received 2005-03-21

Accepted 2005-08-18

doi: 10.1111/j.1745-7254.2006.00241.x

Abstract

Aim: To derive a theoretical model for the prediction of corneal permeability of miscellaneous organic compounds in drug design. **Methods:** A training set of 28 structurally diverse compounds was used to build up the membrane-interaction quantitative structure-activity relationship (MI-QSAR) models. Intermolecular and intramolecular solute descriptors were computed using molecular mechanics, molecular dynamics simulations and quantum chemistry. The QSAR models were optimized using multidimensional linear regression fitting and a stepwise method. A test set of 8 compounds was evaluated using the models as part of a validation process. **Results:** Significant MI-QSAR models ($R=0.976$, $S=0.1301$, $F=70.957$) of corneal permeability of organic compounds were constructed. Corneal permeability was found to depend upon the sum of net atomic charges of hydrogen atoms attached to the heteroatoms (N, O), the sum of the absolute values of the net atomic charges of oxygen and nitrogen atoms, the principal moment of inertia (X), the Connolly accessible area and the conformational flexibility of the solute-membrane complex. **Conclusion:** The MI-QSAR models indicated that the corneal permeability of organic molecules was not only influenced by the organic solutes themselves, but also related to the properties of the solute-membrane complex, that is, the interactions of the molecule with the phospholipid-rich regions of cellular membranes.

Introduction

Most ocular diseases are treated by drugs that are applied to the eye to achieve a localized action on the surface or in the interior of the eye^[1]. However, typically less than 5% of the applied drug reaches the intraocular tissues. There are 2 main ways to improve the ocular absorption of ophthalmic drugs: increasing corneal permeability and prolonging the contact time on the ocular surface. Because a major part (50%–100%) of the instilled dose is absorbed systemically and may cause systemic side effects, it seems that a more effective way to improve the bioavailability of ophthalmic drugs is to increase their corneal permeability.

The main route for intraocular drug absorption is across the cornea^[2]. The corneal barrier we discussed in the present article is composed of epithelium, stroma, and endothelium^[3]. Usually, the corneal epithelium is the main barrier for drug

penetration into the eye^[4]. Intercellular tight junctions (zonula occludens), which completely surround the most superficial cells, serve as a selective barrier for small molecules and fully prevent the diffusion of macromolecules via the paracellular route. Compared with epithelium, which is mainly composed of lipids, stroma is a highly hydrophilic tissue containing mostly water. The major cellular components of the corneal stroma are the corneal fibroblasts (keratocytes) that occupy 2%–3% of the total volume of corneal stroma. Because of the relatively open structure, drugs with a molecular size of up to 500 000 can diffuse in normal stroma^[5]. Only for the most lipophilic drugs does the hydrophilic stroma represent the rate-limiting barrier to corneal penetration. This is because of the slow partitioning of lipophilic compounds from the lipophilic epithelium to the hydrophilic stroma^[6]. The corneal endothelium is a single layer of hexagonal cells that maintains normal corneal hydration.

It has been estimated that drugs with molecular sizes up to approximately 20 nm can permeate normal endothelium^[7].

As mentioned earlier, the corneal permeability of ophthalmic drugs is mainly determined by the epithelium. There are 2 ways for drugs to penetrate across epithelium: for lipophilic drugs, penetration primarily occurs via the transcellular route, and for hydrophilic drugs, it occurs via the paracellular route that involves passive or altered diffusion through intercellular spaces^[8]. Physicochemical drug properties, such as lipophilicity^[9], molecular size^[10], charge^[11], solubility^[12], degree of ionization^[12,13], and pH^[14], affect the route and rate of corneal permeability. For highly hydrophilic drugs, the rate-limiting barrier is the lipoidal corneal epithelium, whereas for highly lipophilic drugs, partitioning from the epithelium to the hydrophilic stroma is rate-limiting and, in most cases, determines the corneal permeability^[15].

To research corneal permeability, we need to determine the permeability coefficient. For measuring permeability coefficients *in vitro*, an isolated rabbit cornea mounted in a perfusion chamber is usually used, which maintains its viability and integrity during the experiment^[16]. This method is very time consuming and expensive. Therefore it seems necessary to construct theoretical models of the corneal permeability of drug candidates on the basis of their physicochemical properties and other structural parameters. Since the 1970s, several congeneric quantitative structure-activity relationships (QSAR) models have been reported for, amongst others, β -blocking agents^[10] ($r=0.9756$) and steroids^[17]. These models are functions of the partition coefficient ($\log P$) or the distribution coefficient ($\log D$, for an ionizable compound) of the drug^[9,17]. However, Grass and Robinson pointed out that some steroids did not fit the parabolic curve established for β -blockers^[18]. Yoshida and Topliss tried to develop a non-congeneric model on the basis of former research results^[19]; they used the difference between the octanol-water partition coefficient and the alkane-water partition coefficient ($\Delta \log P$) and the distribution coefficient ($\log D$) as predictors, and obtained a preferable noncongeneric model ($R=0.950$, $S=0.221$) that unified β -blockers and steroids. Furthermore, they then introduced a further 9 structurally different compounds, including methanol, into the model, using the same method, and also obtained a good result ($R=0.917$, $S=0.282$). However, $\Delta \log P$ values are usually difficult to obtain. Therefore, Fu and Liang attempted to predict the corneal permeabilities of miscellaneous compounds from simpler parameters^[20]. They primarily considered charge parameters and solid parameters and constructed a better model ($R=0.921$, $S=0.233$), which included β -blockers, steroids and other structurally different compounds. Moving on from tradi-

tional QSAR, Iyer *et al* developed a methodology called membrane-interaction (MI)-QSAR analysis, which combines a structure-based design methodology with classic intramolecular QSAR analysis to model the interactions of different compounds with cellular membranes^[21]. They successfully constructed MI-QSAR models predicting the blood-brain barrier penetration of organic compounds.

In the present article, we focused on constructing predictive models of corneal permeability for miscellaneous organic compounds on the basis of MI-QSAR analysis.

Materials and methods

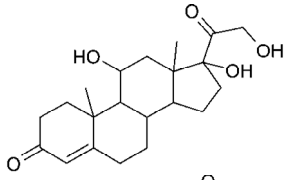
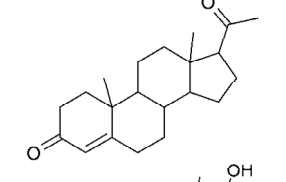
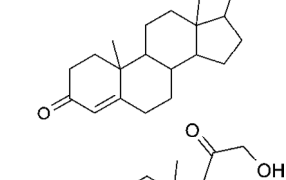
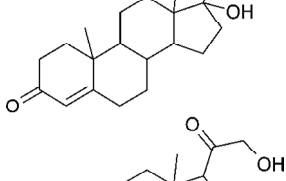
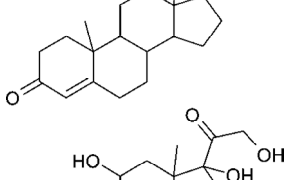
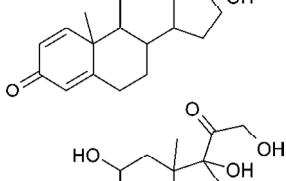
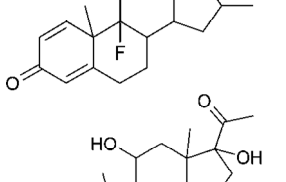
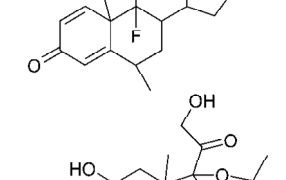
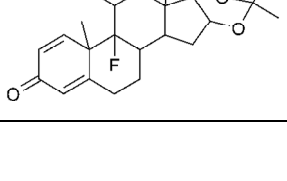
Corneal permeability coefficient The dependent variable used in this theoretical model is the logarithm of the corneal permeability coefficient (P), $\log P$. *In vitro* P (per cm) is calculated using the equation below:

$$P = \frac{7\Delta C}{\Delta t A C_0} \quad (18)$$

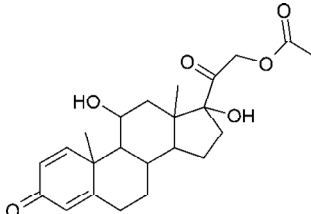
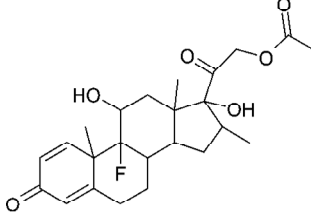
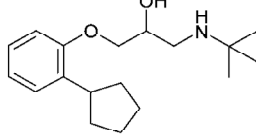
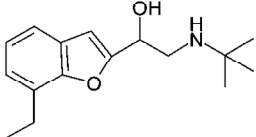
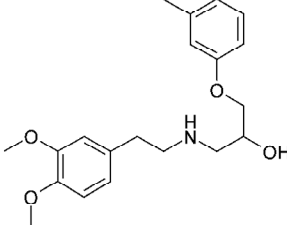
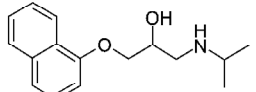
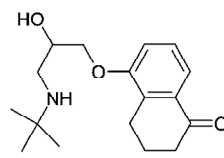
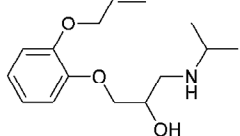
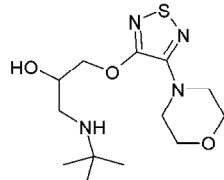
where $\Delta C/\Delta t$ is the change in concentration in dpm/mL of sample per unit time, and 7 is the volume of the sampling chamber in mL. Therefore, the quantity $7\Delta C/\Delta t$ is the steady-state flux in dpm/h. A is the surface area of the cornea (taken to be 1.087 cm²), and C_0 is the initial concentration of the diffusing drug.

Building solute molecules and a DMPC monolayer A training set of 28 organic compounds^[19] (Table 1) and a test set of 8 organic compounds^[19] (Table 2) were chosen. These compounds had a range of molecular weights, from 18.02 to 468.37. Experimental values of $\log P$ published to date cover a range between -6.17 and -3.82. All these compounds were built on a personal computer using the Build module of the commercial software package Hyperchem 7.5 (release 7.5 for Windows; Hypercube, Waterloo, Ontario, Canada). First, the geometry of these compounds was optimized by using the Amber 94 force field in gas state. Second, they were placed in a periodic solvent box whose volume was $X=12 \text{ \AA}$, $Y=10 \text{ \AA}$, $Z=24 \text{ \AA}$, which included 96 water molecules. The temperature was 298 K and the pressure was 101.325 kPa. Then, the compounds in water were minimized by using the method described earlier. The compounds in water were simulated by using the Monte Carlo method at 300 K, and minimized by using the method described earlier. The Monte Carlo method, which samples from the random conformations generated by the Boltzmann distribution under a certain temperature, simulates the molecular movement and kinetic properties of the equilibrium state. It uses a logarithm to calculate a conformation based on the previous one, and

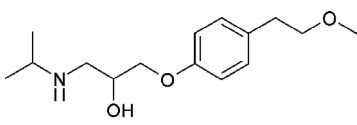
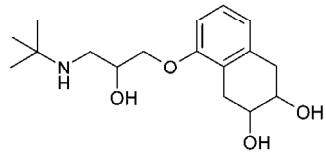
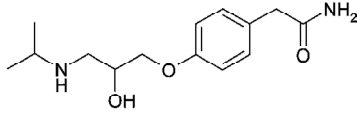


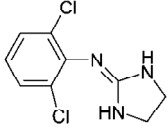
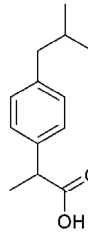
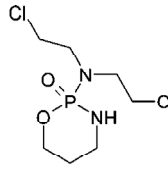
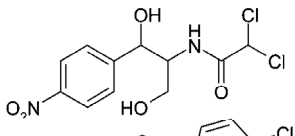
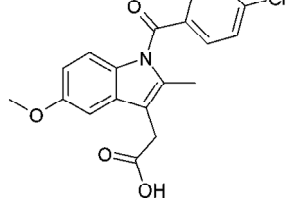
Table 1. The structure, molecular weight, formula, and log *P* of 28 organic compounds in the training set^[19].

ID	Structure	<i>M_r</i>	Formula	log <i>P</i>
1. Hydrocortisone		362.46	C ₂₁ H ₃₀ O ₅	-5.07
2. Progesterone		314.47	C ₂₁ H ₃₀ O ₂	-4.71
3. Testosterone		288.50	C ₁₉ H ₂₈ O ₂	-4.37
4. Cortisolone		346.46	C ₂₁ H ₃₀ O ₄	-4.52
5. Deoxycorticosterone		330.47	C ₂₁ H ₃₀ O ₃	-4.40
6. Prednisolone		360.44	C ₂₁ H ₂₈ O ₅	-5.43
7. Dexamethasone		392.46	C ₂₂ H ₂₉ FO ₅	-5.30
8. Fluorometholone		376.46	C ₂₂ H ₂₉ FO ₄	-4.78
9. Triamcinolone acetonide		434.50	C ₂₄ H ₃₁ FO ₆	-4.80

Continue

ID	Structure	M_r	Formula	$\log P$
10. Prednisolone acetate		402.49	C ₂₃ H ₃₀ O ₆	-4.48
11. Dexamethasone acetate		434.50	C ₂₄ H ₃₁ FO ₆	-4.43
12. Penbutolol		291.43	C ₁₈ H ₂₉ NO ₂	-4.35
13. Bufuralol		261.36	C ₁₆ H ₂₃ NO ₂	-4.14
14. Bevantolol		345.44	C ₂₀ H ₂₇ NO ₄	-4.24
15. Propranolol		259.35	C ₁₆ H ₂₁ NO ₂	-4.32
16. Levobunolol		291.39	C ₁₇ H ₂₅ NO ₃	-4.79
17. Oxprenalol		265.35	C ₁₅ H ₂₃ NO ₃	-4.60
18. Timolol		316.42	C ₁₃ H ₂₄ N ₄ O ₃ S	-4.91

Continue

ID	Structure	M_r	Formula	$\log P$
19. Metoprolol		267.37	C ₁₅ H ₂₅ NO ₃	-4.66
20. Nadolol		309.41	C ₁₇ H ₂₇ NO ₄	-6.00
21. Atenolol		266.34	C ₁₄ H ₂₂ N ₂ O ₃	-6.17
22. Methanol		32.042	CH ₄ O	-4.04
23. Butanol		74.122	C ₄ H ₁₀ O	-4.12
24. Clonidine		230.1	C ₉ H ₉ Cl ₂ N ₃	-4.36
25. Ibuprofen		206.28	C ₁₃ H ₁₈ O ₂	-4.65
26. Cyclophosphamide		261.09	C ₇ H ₁₅ C ₁₂ N ₂ O ₂ P	-4.95
27. Chloramphenicol		323.13	C ₁₁ H ₁₂ Cl ₂ N ₂ O ₅	-5.17
28. Indomethacin		357.79	C ₁₉ H ₁₆ ClNO ₄	-4.16

obtains a continuous conformation sequence that forms a trace in image space. Dynamic simulation uses dynamic equations to generate new conformations, whereas the Monte Carlo method uses a statistical sampling technique to pro-

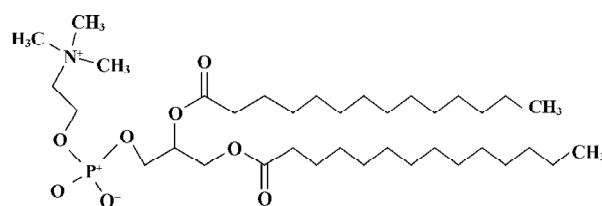
duce the trace of the image space. That is to say, dynamic simulation calculates average value according to time, whereas the Monte Carlo method computes the statistical mean value from averaging each conformation. If the pri-

Table 2. The structure, molecular weight, formula, and log *P* of 8 organic compounds in the test set^[19].

ID	Structure	M_r	Formula	log <i>P</i>
1. Acebutolol		336.43	C ₁₈ H ₂₈ N ₂ O ₄	-6.07
2. Phenylephrine		167.21	C ₉ H ₁₃ NO ₂	-6.03
3. Pilocarpine		208.26	C ₁₁ H ₁₆ N ₂ O ₂	-4.77
4. Water		18.02	H ₂ O	-3.82
5. Cromolyn		468.37	C ₂₃ H ₁₆ O ₁₁	-5.97
6. Alprenolol		249.35	C ₁₅ H ₂₃ NO ₂	-4.54
7. Betaxolol		307.43	C ₁₈ H ₂₉ NO ₃	-4.57
8. Pindolol		248.32	C ₁₄ H ₂₀ N ₂ O ₂	-5.00

mary parameter is appropriate, both the Monte Carlo and the dynamic method can reach equilibrium. However, the Monte Carlo method is usually faster. Here, we used a Monte Carlo simulation to gain the steady-state conformation of the solute in the solvent^[22].

The phospholipid dimyristoylphosphatidylcholine (DMPC) was selected as the model phospholipid in the present study. A single DMPC (Figure 1) molecule was built using Hyperchem 7.5 from the available crystal structure data^[23]. A model of the DMPC membrane monolayer was constructed by using the software Material Studio (version 2.2.1; Accelrys, San Diego, CA, USA). According to the work done by van der Ploeg and Berendsen, the DMPC monolayer is composed of 25 DMPC molecules (5×5×1)^[24]. The unit cell parameters used for building the DMPC monolayer were a=8 Å, b=8 Å, c=32 Å, and γ=96.0°, which yield an average surface area per

**Figure 1.** Structure of a single dimyristoylphosphatidylcholine molecule.

phospholipid of 64 Å², similar to the value found by Stouch experimentally^[25]. Therefore, we can consider the DMPC membrane monolayer model to be reasonable.

Molecular dynamic simulation To prevent unfavorable van der Waals interactions between a solute molecule and the membrane DMPC molecules, one of the “center” DMPC

molecules was removed from the DMPC model described and an organic compound (solute) was inserted in the space created by the missing DMPC molecule to form a solute-membrane complex. The solute was inserted at 3 different positions in the DMPC model and 3 corresponding molecular dynamic simulation (MDS) models were generated for each compound. The 3 different positions (depths) were: (1) solute molecule in the head group region; (2) solute molecule between the head-group region and the aliphatic chains; and (3) solute molecule in the tail region of the aliphatic chains. MDS of the complex was performed by using the Discover module of Material Studio, using a compass force field. A simulation temperature of 311 K was selected and 2-dimensional periodic boundary conditions corresponding to the "surface plane" of the monolayer were used ($a=32 \text{ \AA}$, $b=32 \text{ \AA}$, $c=80 \text{ \AA}$, and $\gamma=96.0^\circ$) for the DMPC model. In order to equilibrate the solute-membrane complex gradually and fully, the MDS course was carried out with 3 phases: (1) simulation at 120 K for 1500 steps (1.5 ps); (2) simulation at 220 K for 1500 steps (1.5 ps); and (3) simulation at 311 K for 10 000 steps (10 ps).

Calculation of descriptors Most of the intramolecular solute descriptors were calculated by using the commercial software package CS Chem3D Ultra 7.0 (Chemoffice 2002; CambridgeSoft, Cambridge, MA, USA), which included molecular mechanism (MM) parameters (such as bending energy, torsion energy and van der Waals energy), quantum chemistry parameters (such as electronic energy, HOMO, and LUMO energy), hydrophobic parameters (such as Clog P), and stereo parameters (such as E_s and Balaban index). The data of $Q_{O,N}$ and Q_H comes from Fu and Liang's study^[20].

The intermolecular solute-membrane interaction descriptors were extracted directly from the MDS trajectories in which the solute-membrane complex had the lowest energy geometry. These descriptors were mainly energy parameters. The total energy of a system can be expressed as follows^[21]:

$$E_{\text{total}} = E_{\text{valence}} + E_{\text{crossterm}} + E_{\text{nonbond}}$$

Construction and testing of MI-QSAR models MI-QSAR models of permeation of the cornea by organic compounds were constructed by using the partial sum of squares for regression using the SPSS software package (Chicago, USA). A training set of 28 structurally diverse compounds whose corneal permeability coefficients have been measured *in vivo* was used to construct the MI-QSAR models. Molecular dynamics simulations were used to determine the explicit interaction of each test compound with the DMPC model. An additional set of intramolecular solute descriptors were computed and considered in the trial pool of descriptors for building MI-QSAR models. The MI-QSAR models were op-

timized by using multidimensional linear regression fitting and a stepwise method. A test set of 8 compounds was evaluated by using the MI-QSAR models as part of a validation process. A principal components analysis (PCA) was performed by using SPSS to identify the principal components of the constructed models.

Application package for models A forecasting application package of the constructed models was built using the MFC module of the commercial software packages Microsoft Visual C++ (version 6.0; Microsoft Corporation, USA).

Results

Construction of solute molecules and a DMPC monolayer Figure 2 shows the dominant conformation of compound No 1 labeled by atom-type in water. The box denotes the water solvent box defined in the Monte Carlo simulation.

Molecular dynamic simulation The energy of a solute inserted in the middle position of the DMPC model was lower than that of a solute inserted in the other 2 positions. Figure 3 shows the dominant conformation of a solute-membrane complex in the MDS. The DMPC molecules appear as white sticks. The molecule depicted using spheres represents an organic compound. The white box indicates the border of the volume. Figure 4 is a "side" view of the molecule in Figure 3.

Construction and testing of MI-QSAR models MI-QSAR analysis was used to develop predictive models of corneal permeability of some organic compounds and simulate the interaction of a solute with the phospholipid-rich regions of cellular membranes. Molecular descriptors of 28 compounds in the training set and 8 compounds in the test set are listed in Table 3 and Table 4, respectively. Six MI-QSAR equations were constructed based on the information in Table 3:

$$\log P = -4.201 - 2.585 Q_H^2$$

$n=28 \quad R=0.860 \quad S=0.2750 \quad R^2=0.730 \quad F=74.071 \quad (1)$

$$\log P = -3.972 - 2.388 Q_H^2 - 0.207 Q_{O,N}$$

$n=28 \quad R=0.897 \quad S=0.2433 \quad R^2=0.789 \quad F=51.407 \quad (2)$

$$\log P = -4.488 - 5.230 Q_H^2 - 0.236 Q_{O,N} + 2.768 Q_H$$

$n=28 \quad R=0.939 \quad S=0.1931 \quad R^2=0.867 \quad F=59.628 \quad (3)$

$$\log P = -4.354 - 4.925 Q_H^2 - 0.216 Q_{O,N} + 2.533 Q_H + 1.576 \times 10^{-3} dT$$

$n=28 \quad R=0.950 \quad S=0.1785 \quad R^2=0.886 \quad F=53.656 \quad (4)$

$$\log P = -4.324 - 4.396 Q_H^2 - 0.288 Q_{O,N} + 2.106 Q_H + 2.163 \times 10^{-3} dT + 1.795 \times 10^{-4} PMIX$$

$n=28 \quad R=0.962 \quad S=0.1605 \quad R^2=0.908 \quad F=54.355 \quad (5)$

$$\log P = -3.885 - 4.290 Q_H^2 - 0.304 Q_{O,N} + 2.171 Q_H + 2.366 \times 10^{-3} dT + 3.453 \times 10^{-4} PMIX - 1.22 \times 10^{-3} SAS$$

$n=28 \quad R=0.976 \quad S=0.1301 \quad R^2=0.940 \quad F=70.957 \quad (6)$

where Q_H is the sum of net atomic charges of hydrogen

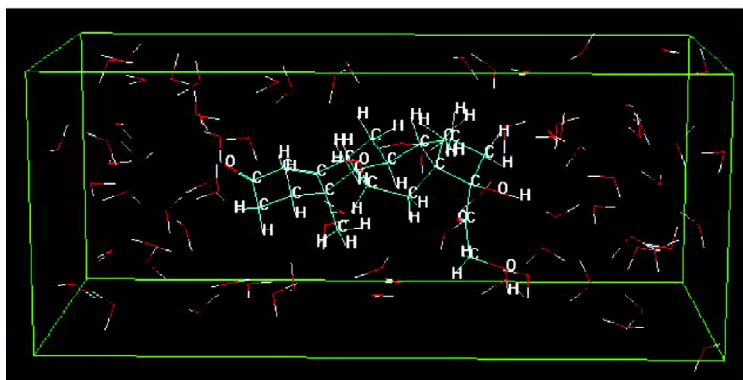


Figure 2. The dominant conformation of compound No 1 in water. The compound is labeled by atom-type. The box denotes the water solvent box defined in the Monte Carlo simulation.

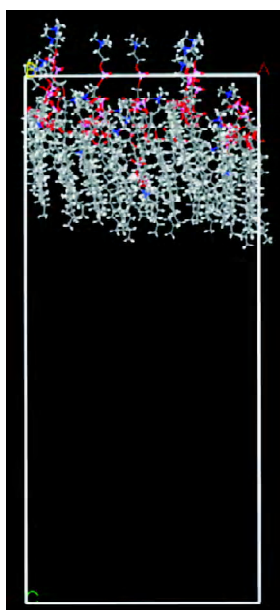


Figure 3. The dominant conformation of a solute-membrane complex in the MDS. The DMPC molecules appear as white sticks. The molecule depicted using spheres represents an organic compound. The white box represents the border of the volume.

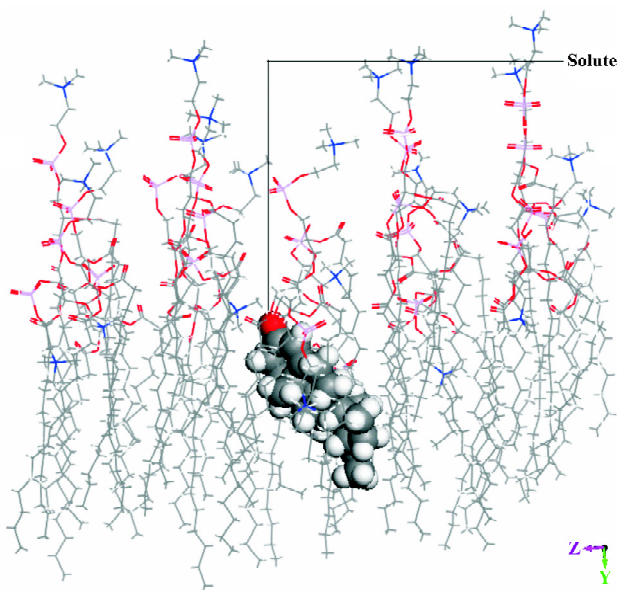


Figure 4. A "side" view of the solute-membrane complex in Figure 3.

atoms attached to the heteroatoms (N, O), $Q_{O,N}$ is the sum of the absolute values of the net atomic charges of oxygen and nitrogen atoms, dT is the conformational flexibility of the solute-membrane complex, $PMIX$ is the principal moment of inertia (X), and SAS is the Connolly accessible area. $PMIX$ and SAS are intramolecular solute descriptors that came from the CS calculation. dT is related to interactions between a solute and the DMPC model. It represents the change in the dihedral torsion energy of the solute-membrane complex, compared with that of the DMPC model, that is, $\Delta E_{\text{torsion}}$. Here, n is the number of organic compounds, R is the correlation coefficient, S is the standard deviation and F is the F -statistic.

Figure 5 is a diagnostic plot of the MI-QSAR models: R is the correlation coefficient and S is the standard deviation of the best x -term model, where x is plotted on the X-axis for the 28 compounds of the training set. From this plot we can see that with the increase of the variable from 1 to 6, the relativity of MI-QSAR equations is also improved, and the predictive ability of the models is enhanced. The observed and predicted $\log P$ values of the training set are listed in Table 5.

A test set of 8 organic compounds was constructed as one way to attempt to validate the MI-QSAR models given by the 6 equations mentioned. The compounds of the test set were selected from different groups of ophthalmic drugs. The observed and predicted $\log P$ values for this test set are given in Table 6. Figure 6 gives the linear relationship between the experimental $\log P$ values (shown as the abscissa Predict) and the corresponding predicted $\log P$ as predicted by the 6-term MI-QSAR model (shown as the ordinate Observed) for all the molecules in the training set and the test set.

According to the result of the PCA, which has a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy value of 0.543 and a Bartlett value of 94.955 ($P < 0.01$), there are 3 principal components in the models. The cumulative variance that they explain is 84.550%. Table 7 shows the rotated component matrix. From it we can see that $Q_{O,N}$, $PMIX$ and

Table 3. Molecular descriptors of the 28 organic compounds in the training set.

ID	Q_H^2	$Q_{O,N}$	Q_H	dT (kcal·mol ⁻¹)	$PMIX$ (g·mol ⁻¹ ·Å ⁻²)	SAS (Å ²)
1	0.3900	1.5441	0.6245	-107.04	1221.87	524.58
2	0.0000	0.5957	0.0000	-128.32	809.42	515.02
3	0.0393	0.6164	0.1982	-98.97	615.15	478.90
4	0.1788	1.2137	0.4228	-103.81	991.81	524.69
5	0.0481	0.9107	0.2194	-91.81	783.781	528.28
6	0.3921	1.5442	0.6262	-133.96	1122.19	531.39
7	0.4040	1.5344	0.6356	-134.11	1257.19	541.23
8	0.1704	1.2247	0.4128	-141.39	1608.80	534.99
9	0.0426	2.0059	0.2063	-168.04	1588.30	591.12
10	0.1705	1.8494	0.4129	-144.75	2261.20	598.05
11	0.1782	1.8362	0.4221	-133.93	2219.61	575.73
12	0.1481	0.8562	0.3849	-92.56	950.79	559.49
13	0.1472	0.7283	0.3837	-23.11	692.23	510.32
14	0.1486	1.2833	0.3855	-16.61	1841.94	668.07
15	0.1490	0.8562	0.3860	41.45	511.32	513.05
16	0.1369	1.1553	0.3700	-73.34	606.90	552.88
17	0.1537	1.0307	0.3921	-62.95	958.28	557.32
18	0.1424	1.9042	0.3774	-73.18	1008.90	543.64
19	0.1501	1.1426	0.3874	-66.73	558.74	563.90
20	0.6386	1.4927	0.7991	-99.70	843.22	564.42
21	0.6996	1.6587	0.8364	-82.71	708.75	526.27
22	0.0382	0.326	0.1954	-57.47	3.94	149.21
23	0.0389	0.3292	0.1972	-44.57	64.81	242.63
24	0.1796	0.7789	0.4238	-84.52	678.11	385.39
25	0.2033	0.7605	0.4509	-56.20	360.24	428.80
26	0.0605	3.6249	0.2460	-53.74	1135.71	430.07
27	0.4260	2.0257	0.6527	-60.66	763.56	479.77
28	0.0591	1.3898	0.2431	45.72	1683.40	556.67

Note: Q_H is the sum of net atomic charges of hydrogen atoms attached to the heteroatoms (N, O), and $Q_{O,N}$ is the sum of the absolute values of the net atomic charges of oxygen and nitrogen atoms. This information comes from Fu and Liang^[20]. dT is the conformational flexibility of the solute-membrane complex which related to interactions between a solute and the DMPC model. It represents the change in the dihedral torsion energy of the solute-membrane complex compared with that of the DMPC model. $PMIX$ is the principal moment of inertia (X), and SAS is the Connolly accessible area. These values come from the CS calculation. Å: Angstrome.

SAS can be replaced by the first component, Q_H^2 and Q_H can be replaced by the second component, and dT can be replaced by the third component.

Discussion

We have constructed a theoretical model of corneal permeability of organic compounds. With the increase in the number of variables, the relativity of the MI-QSAR equation

Table 4. Molecular descriptors of the 8 organic compounds in the training set.

ID	Q_H^2	$Q_{O,N}$	Q_H	dT (kcal·mol ⁻¹)	$PMIX$ (g·mol ⁻¹ ·Å ⁻²)	SAS (Å ²)
1	0.4051	1.8220	0.6365	-69.60	1044.55	652.66
2	0.3315	0.8887	0.5758	-84.99	216.24	354.85
3	0.0000	0.8258	0.0000	-62.47	413.89	386.25
4	0.1465	0.3828	0.3828	-29.50	0.66	97.07
5	0.4853	2.7024	0.6966	-178.79	1351.57	644.18
6	0.1487	0.8518	0.3856	-56.07	709.48	511.98
7	0.1481	1.1241	0.3848	-62.94	1292.37	610.83
8	0.4023	1.0675	0.6343	58.32	459.70	481.06

Note: see the footnote of Table 3 for explanations of Q_H , $Q_{O,N}$, Q_H^2 , dT , $PMIX$, SAS , and Å.

Table 5. Experimental value of log P and the predictive value of log P for the 28 organic compounds in the training set.

ID	Experi- mental value of log P	Predictive value of log P					
		Term 1	Term 2	Term 3	Term 4	Term 5	Term 6
1	-5.07	-5.21	-5.22	-5.16	-5.20	-5.18	-5.14
2	-4.71	-4.20	-4.09	-4.63	-4.68	-4.63	-4.72
3	-4.37	-4.30	-4.19	-4.29	-4.33	-4.36	-4.42
4	-4.52	-4.66	-4.65	-4.54	-4.59	-4.62	-4.65
5	-4.40	-4.33	-4.28	-4.35	-4.38	-4.39	-4.48
6	-5.43	-5.21	-5.23	-5.17	-5.24	-5.26	-5.25
7	-5.30	-5.25	-5.25	-5.20	-5.28	-5.27	-5.25
8	-4.78	-4.64	-4.63	-4.52	-4.63	-4.57	-4.52
9	-4.80	-4.31	-4.49	-4.61	-4.74	-4.73	-4.80
10	-4.48	-4.64	-4.76	-4.67	-4.78	-4.64	-4.57
11	-4.43	-4.66	-4.78	-4.68	-4.77	-4.64	-4.54
12	-4.35	-4.58	-4.50	-4.40	-4.44	-4.44	-4.52
13	-4.14	-4.58	-4.47	-4.37	-4.30	-4.30	-4.34
14	-4.24	-4.59	-4.59	-4.50	-4.41	-4.24	-4.29
15	-4.32	-4.59	-4.51	-4.40	-4.23	-4.23	-4.30
16	-4.79	-4.55	-4.54	-4.45	-4.46	-4.53	-4.66
17	-4.60	-4.60	-4.55	-4.45	-4.44	-4.44	-4.50
18	-4.91	-4.57	-4.71	-4.64	-4.63	-4.68	-4.74
19	-4.66	-4.59	-4.57	-4.47	-4.46	-4.54	-4.69
20	-6.00	-5.85	-5.81	-5.97	-5.95	-5.94	-5.98
21	-6.17	-6.01	-5.99	-6.22	-6.17	-6.17	-6.17
22	-4.04	-4.30	-4.13	-4.22	-4.21	-4.30	-4.04
23	-4.12	-4.30	-4.13	-4.22	-4.19	-4.26	-4.10
24	-4.36	-4.67	-4.56	-4.44	-4.47	-4.51	-4.41
25	-4.65	-4.72	-4.61	-4.48	-4.47	-4.54	-4.54
26	-4.95	-4.36	-4.87	-4.98	-4.90	-5.03	-4.97
27	-5.17	-5.30	-5.41	-5.39	-5.33	-5.40	-5.37
28	-4.16	-4.35	-4.40	-4.45	-4.26	-4.07	-4.02

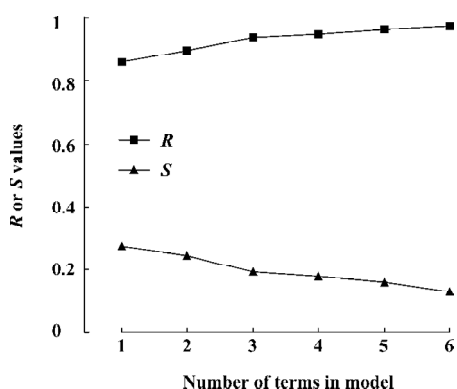


Figure 5. A diagnostic plot of the MI-QSAR models. R is the correlation coefficient and S is the standard deviation of the best x -term model, where x is plotted on the X-axis for the 28 compounds of the training set. With an increase in the number of variables from 1 to 6, the relativity of the MI-QSAR equation is also improved, and the predictive ability of the models is enhanced.

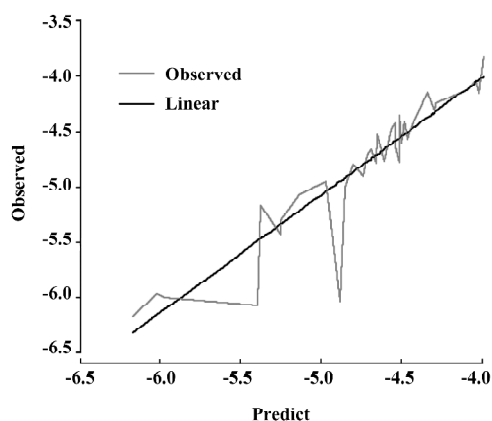


Figure 6. The linear relationship between the experimental $\log P$ values (shown as the abscissa Predict) and the corresponding predicted $\log P$ as predicted by the 6-term MI-QSAR model (shown as the ordinate Observed) for all the molecules in the training set and the test set.

is also improved, and the predictive ability of the model is enhanced. Equation 6 is most significant. Moreover, the models have been validated by using the compounds of the test set, and the 5–6 term MI-QSAR models could be used to predict $\log P$ for other compounds during drug design.

These MI-QSAR models indicate that corneal permeability depends on 5 parameters: Q_H , $Q_{O,N}$, dT , $PMIX$, and SAS . Q_H seems to be the dominant descriptor in these MI-QSAR models, which is a parameter that is closely related to the hydrophilic groups such as $-\text{COOH}$, $-\text{NH}_2$, $-\text{OH}$, and $-\text{NH}$. A compound with greater corneal permeability usually has adequate hydrophilic groups. Relative to the stroma, the bar-

Table 6. Experimental value of $\log P$ and the predictive value of $\log P$ for the 8 organic compounds in the test set.

ID	Experimental value of $\log P$	Predictive value of $\log P$					
		Term 1	Term 2	Term 3	Term 4	Term 5	Term 6
1	-6.07	-5.25	-5.32	-5.28	-5.24	-5.25	-5.39
2	-6.03	-5.06	-4.95	-4.84	-4.85	-4.97	-4.89
3	-4.77	-4.20	-4.14	-4.68	-4.63	-4.62	-4.61
4	-3.82	-4.58	-4.40	-4.29	-4.24	-4.34	-3.99
5	-5.97	-5.46	-5.69	-5.74	-5.84	-5.91	-6.02
6	-4.54	-4.59	-4.50	-4.40	-4.38	-4.40	-4.46
7	-4.57	-4.58	-4.56	-4.46	-4.45	-4.39	-4.47
8	-5.00	-5.24	-5.15	-5.09	-4.87	-4.85	-4.85

Table 7. Rotated component matrix of principal components analysis.

	Component		
	1	2	3
Q_H^2	0.08293	0.983	-0.08513
$Q_{O,N}$	0.705	0.218	-0.04657
Q_H	0.153	0.974	-0.03045
dT	-0.152	-0.08532	0.982
$PMIX$	0.898	-0.08819	-0.231
SAS	0.831	0.144	-0.02631

rier effect of epithelium is more prominent, so when a compound has a Q_H value that is too high, its corneal permeability will decrease markedly. This is why $\log P$ is directly proportional to Q_H , but inversely proportional to Q_H^2 . The Q_H value is also relevant to the capacity of a compound to form hydrogen bonds, which is the same as the descriptor $Q_{O,N}$. When $Q_{O,N}$ lessens, the value of $\log P$ will increase, which indicates that weak hydrogen bond potential is favorable for corneal penetration. This is a similar situation to the transport of compounds through other biological membranes such as skin^[26,27], small intestine^[28], Caco-2 cell monolayers^[29,30], and the blood-brain barrier^[31]. $PMIX$ is the moment of inertia (X) when the Cartesian coordinate axes are the principal axes of the molecule. The inertia of a molecule is determined by its 3-dimensional structure, and describes the molecule shape. The MI-QSAR models reveal that with the accretion of inertia (X), the solute compound becomes more easily to penetrate through the cornea. It can be inferred that quadrature molecules are more likely to permeate the cornea. SAS represents the area of the solute that contacts the solvent, and

can be considered as an index of the solute's hydrophobic properties. When *SAS* increases, the value of $\log P$ decreases. Thus, it can be concluded that hydrophobic (lipophilic) molecules can cross the corneal barrier more easily.

The other descriptor, dT , reflects the interaction of the solute with the membrane and the behavior of the entire membrane-solute complex. Here, the greater dT is, the more the value of $\log P$ increases. This indicates that for a small molecule-membrane complex in the combined state, the $\log P$ value is low when its E_{torsion} is stable, which means that the more tightly a small molecule combines with the membrane, the more difficultly it has in penetrating through the corneal barrier. This suggests that as the solute becomes more flexible within the membrane, its $\log P$ value would decrease. This may be due to the complex amphiphilic structure of the corneal barrier.

On the basis of the results of the PCA, the capability of an organic compound to permeate the cornea is mainly related to 3 principal components, which can be related to the molecular structure and shape, the hydrophilicity of a solute molecule, and the strength of the combination of a small molecule with the membrane, respectively. The molecular structure and shape are described by 3 parameters: $Q_{O,N}$, *PMIX* and *SAS*. $Q_{O,N}$ and *SAS* represent the strength of hydrogen bonds in a compound and *PMIX* represents the shape of a solute molecule. Generally, a quadrate molecule with weak hydrogen bond potential has good potential to penetrate the cornea. The hydrophilicity of a solute molecule is described by 2 descriptors, namely Q_H^2 and Q_H . Generally, organic compounds with proper hydrophilic groups, in which lipophilicity is greater than hydrophilicity, are easily able to penetrate through the cornea. The strength of the combination of a small molecule with the membrane is described by one parameter, namely $\Delta E_{\text{torsion}}$. The more tightly a small molecule combines with the membrane, the more difficulty it has in penetrating through the corneal barrier.

For QSAR models, the less comparable the studied molecules are, the more universally significant the equations are. On the other hand, the precision of a MI-QSAR simulation may be greatly increased when a series of organic compounds with similar structures comprise the training set. Although our MI-QSAR models are noncongeneric models based on several types of organic compounds, there are still many types of compounds that cannot be included in these models. For further studies, we should pay more attention to the experimental component, in order to expand the research range of molecules and search for new models that can unify more types of compounds, to improve the practical value of the constructed models.

In conclusion, we have developed an extension of the traditional QSAR approach by combining it with a solute-membrane complex that simulates the corneal environment. MI-QSAR analysis is a structure-based design methodology combined with classic intramolecular QSAR analysis to model the interactions of different compounds with cellular membranes. Although still applying the structural information in a 2-dimensional, "structure-function relationship" manner, this method also takes into account the powerful 3-dimensional behavior displayed by membrane structures, and thus improves on past QSAR methods. With the help of the model, the membrane penetration process can be reliably described for structurally diverse molecules whose interactions with the phospholipid-rich regions of cellular membranes are explicitly considered. The MI-QSAR models indicate that the corneal permeability of organic molecules is not only influenced by organic solutes themselves, but is also related to the properties of the solute-membrane complex, that is, interactions between the molecule and the phospholipid-rich regions of cellular membranes. Compared with the work of Iyer and coworkers, which paid more attention to the energy parameters of the solute-membrane system^[21], our models achieve a better correlation coefficient and smaller standard deviation by taking a wider range of descriptors into account, and thus have a better predictive ability. Moreover, we use PCA to identify the descriptors involved in the models, which makes the data analysis more clear and reliable. In addition, by building an application for the constructed models, we recognize the human-machine dialog in the prediction process, and make the work easier.

References

- 1 Davies NM. Biopharmaceutical considerations in topical ocular drug delivery. *Clin Exp Pharmacol Physiol* 2000; 27: 558–62.
- 2 Ahmed I, Patton TF. Disposition of timolol and insulin in the rabbit eye following corneal versus noncorneal absorption. *Int J Pharm* 1987; 38: 9–21.
- 3 Kumar MT, Pandit JK, Balasubramaniam J. Novel therapeutic approaches for uveitis and retinitis. *J Pharm Pharm Sci* 2001; 4: 248–54.
- 4 Toropainen E, Ranta VP, Talvitie A, Suhonen P, Urtti A. Culture model of human corneal epithelium for prediction of ocular drug absorption. *Invest Ophthalmol Vis Sci* 2001; 42: 2942–8.
- 5 Maurice DM, Mishima S. Ocular pharmacokinetics. In: Sears MC, editor. vol 69. *Handbook of experimental pharmacology. Pharmacology of the eye*. Berlin: Springer-Verlag; 1984. p19–116.
- 6 Schornwald RD. Ocular drug delivery. Pharmacokinetic considerations. *Clin Pharmacokinet* 1990; 18: 255–69.
- 7 Liaw J, Robinson JR. Ocular penetration enhancers. In: Mitra AK, editor. *Ophthalmic drug delivery system*. New York: Marcel Dekker; 1993. p 369–81.

- 8 Borchardt RT. Assessment of transport barriers using cell and tissue culture systems. *Drug Dev Ind Pharm* 1990; 16: 2595–612.
- 9 Schoenwald RD, Huang HS. Corneal penetration behavior of beta-blocking agents I: physicochemical factors. *J Pharm Sci* 1983; 72: 1266–72.
- 10 Liaw J, Robinson JR. The effect of polyethylene glycol molecular weight on corneal transport and the related influence of penetration enhancers. *Int J Pharm* 1992; 88: 125–40.
- 11 Liaw J, Rojanasakul Y, Robinson JR. The effect of drug charge type and charge density on corneal transport. *Int J Pharm* 1992; 88: 111–24.
- 12 Sieg JW, Robinson JR. Vehicle effects on ocular drug bioavailability II: evaluation of pilocarpine. *J Pharm Sci* 1977; 66: 1222–8.
- 13 Maren TH, Jankowska L. Ocular pharmacology of sulfonamides: the cornea as barrier and depot. *Curr Eye Res* 1985; 4: 399–408.
- 14 Wei G, Xu H, Ma Y, Li SM, Zheng JM. Effect of pH on the permeability of timolol maleate across isolated rabbit cornea. *Acta Pharm Sin* 2001; 36: 707–10.
- 15 Shih RL, Lee VHL. Rate limiting barrier to the penetration of ocular hypotensive beta blockers across the corneal epithelium in the pigmented rabbit. *J Ocul Pharmacol* 1990; 6: 329–36.
- 16 Loftsson T, Jarvinen T. Cyclodextrins in ophthalmic drug delivery. *Adv Drug Deliv Rev* 1999; 36: 59–79.
- 17 Schoenwald RD, Ward RL. Relationship between steroid permeability across excised rabbit cornea and octanol-water partition coefficients. *J Pharm Sci* 1978; 67: 786–8.
- 18 Grass GM, Robinson JR. Mechanisms of corneal drug penetration: *in vivo* and *in vitro* kinetics. *J Pharm Sci* 1988; 77: 3–17.
- 19 Yoshida F, Topliss JG. Unified model for the corneal permeability of related and diverse compounds with respect to their physicochemical properties. *J Pharm Sci* 1996; 85: 819–23.
- 20 Fu XC, Liang WQ. A simple model for the prediction of corneal permeability. *Int J Pharm* 2002; 232: 193–7.
- 21 Iyer M, Mishra R, Han Y, Hopfinger AJ. Predicting blood-brain barrier partitioning of organic molecules using membrane-interaction QSAR analysis. *Pharm Res* 2002; 19: 1611–21.
- 22 Ma XL, Chen C, Yang J. Predictive model of blood-brain barrier penetration of organic compounds. *Acta Pharmacol Sin* 2005; 26: 500–12.
- 23 Hauser H, Pascher L, Pearson RH, Sundell S. Preferred conformation and molecular packing of phosphatidylethanolamine and phosphatidylcholine. *Biochem Biophys Acta* 1981; 650: 21–51.
- 24 van der Ploeg P, Berendsen HJC. Molecular dynamics simulation of a bilayer membrane. *J Chem Phys* 1982; 76: 3271–6.
- 25 Stouch TR. Lipid membrane structure and dynamics studied by all atom molecular dynamics simulations of hydrated phosphatidylcholine vesicles. *Mol Simulation* 1993; 1: 335–62.
- 26 Abraham MH, Martins F, Mitchell RC. Algorithms for skin permeability using hydrogen bond descriptors: the problem of steroids. *J Pharm Pharmacol* 1997; 49: 858–65.
- 27 Fu XC, Yu QS, Liang WQ. A modified mathematical model for percutaneous absorption of drugs. *Chin Pharm J* 2000; 35: 276–7.
- 28 Clark DE. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J Pharm Sci* 1999; 88: 807–14.
- 29 van de Waterbeemd H, Camenish G, Folkers G, Raevsky OA. Estimation of Caco-2 cell permeability using calculated molecular descriptors. *Quant Struct Act Relat* 1996; 15: 480–90.
- 30 Fu XC, Liang WQ, Yu QS. Correlation of drug absorption with molecular charge distribution. *Pharmazie* 2001; 56: 267–8.
- 31 Norinder U, Sjoberg P, Osterberg T. Theoretical calculation and prediction of brain-blood partitioning of organic solutes using Molsurf parameterization and PLS statistics. *J Pharm Sci* 1998; 87: 952–9.