A Study of the Relationship between Cornea Permeability and Eye Irritation Using Membrane-Interaction QSAR Analysis

Yi Li, Jianzhong Liu, Dahua Pan, and A. J. Hopfinger

Laboratory of Molecular Modeling and Design (MC 781), College of Pharmacy, University of Illinois at Chicago, 833 South Wood Street, Chicago, Illinois 60612-7231

Received July 7, 2005; accepted September 6, 2005

A methodology termed membrane-interaction QSAR (MI-QSAR) analysis has been used to develop QSAR models to predict drug permeability coefficients across cornea and its component layers (epithelium, stroma, and endothelium). From a training set of 25 structurally diverse drugs, significant QSAR models are constructed and compared for the permeability of the cornea, epithelium, and stroma plus endothelium. Cornea permeability is found to depend on the measured distribution coefficient of the drug, the cohesive energy of the drug, the total potential energy of the drug-membrane “complex,” and three other energy refinement descriptor terms. The endothelium may be a more important barrier in cornea permeation than the stroma. Moreover, an investigation of the correlation between cornea permeation and eye irritation is presented as an example of a cross study on different ADME properties using MI-QSAR analysis. Thirteen structurally diverse drugs, whose molar-adjusted eye irritation scores (MES) have been measured using the Draize rabbit-eye test, were chosen as an eye irritation comparison set. A poor correlation ($R^2 = 0.0232$) between the MES measures and the predicted cornea permeability coefficients for the drugs in the eye irritation set suggests there is no significant relationship between eye irritation potency and the cornea permeability.

Key Words: cornea permeability; eye irritation; membrane-interaction quantitative structure-activity relationships (MI-QSAR).

A challenge in drug delivery is the local administration of drugs to the eye (Lang, 1995; Tasman, 1995). To be effective, most drugs must pass through the eye’s tissue barriers to reach therapeutic targets within the globe. The cornea is often the dominant barrier to drug transport (Tasman, 1995). Thus, the ability to predict drug transport across the cornea would be useful in the development of new topical drugs for ophthalmic diseases. Generating data to characterize corneal permeability is time-consuming and uses animals. As an alternative, a few models have been developed to predict cornea permeability of drug candidates. Most models of transcorneal transport can be divided into two classes based either on classic transport thermodynamics and kinetics (Edward and Prausnitz, 2001; Grass et al., 1988; Yoshida and Topliss, 1996), or on statistical analysis (Worth and Cronin, 2000).

Classic transport models require assumptions regarding corneal structure, components and transport pathways, and a number of physical properties, such as LogP and molecular radius (Edward and Prausnitz, 1998, 2001). Classic transport models can partially explain cornea permeation. Cellular membrane permeation has a nonspecific component with regard to the chemical structure of the solute, largely captured by its LogP. However, there is also a structure-specific component to cornea permeability, reflected by the limitation of simply using LogP in cornea permeability models. There is improvement in model quality when structure-specific features are included, e.g., hydrogen-bonding capacity and molecular shape (Worth and Cronin, 2000).

Construction of a model from statistical analysis, such as a QSAR model, requires few assumptions. However, the number of calculated trial descriptors of the compounds of the training set is made as large as possible. A statistical method is then employed to fit the data in constructing QSAR models. Moreover, modeling chemically and structurally diverse solute data sets only involves computation of intramolecular solute properties. The underlying idea is that if enough solute features are included, the key intramolecular solute properties of, in this case, the cornea permeation mechanism will be captured into the QSAR model. But two questions can be raised: Firstly, is the trial set of the intramolecular solute descriptors chosen adequate to capture the requisite mechanistic information? Secondly, can intramolecular descriptors capture the intramolecular interactions between solutes and cornea cell membranes to adequately explain cornea permeability? Some type of structure-based design QSAR approach may be necessary to build a significant QSAR model.

A methodology termed membrane-interaction, MI, QSAR analysis, where structure-based design is combined with classic intramolecular QSAR analysis to model diverse compounds...
interacting with cellular membranes, has been designed to build ADMET QSAR models (Kulkarni et al., 2001, 2002; Kulkarni and Hopfinger, 1999). The composite phospholipid-rich regions of a membrane bilayer of the cell is assumed to constitute the “receptor” required in structure-based design. Membrane-solute interaction properties determined from modeling the solute-membrane complex are added to the intramolecular physicochemical properties of the solute to enlarge the trial descriptor pool, and supply information needed to incorporate chemical and structural diversity into the QSAR analysis.

Three tissues form the cornea: the epithelium, stroma, and endothelium. The epithelium is composed of seven or so layers of keratocyte on the external surface of the cornea. The stroma is primarily composed of large collagen fibers embedded in a proteoglycan matrix forming the bulk of the cornea. The corneal endothelium is a monolayer of hexagonal packed cells at the internal base of the cornea. The rate-limiting tissue barrier for cornea permeation is the epithelium (Edward and Prausnitz, 2001; Prausnitz and Noonan, 1998). Since the epithelium is composed of a multi-layer packing of cells, MI-QSAR analysis is ideally suited to construct QSAR models that can capture the descriptors indicative of the cornea permeability of a drug. In this article we describe a set of MI-QSAR analyses that model the permeability of cornea, as well as the permeability of its components: the epithelium and stroma plus endothelium. The role each tissue component plays in net cornea permeability is discussed based on the MI-QSAR models.

In a previous study, MI-QSAR analysis has been successfully applied to build a QSAR model from an eye irritation data set in which the training set compounds have molar-adjusted eye irritation scores (MES) measured using the Draize rabbit-eye test (Kulkarni et al., 2001). In this current study, based on the observed MES values and the predicted values of cornea permeability for 13 compounds, an exploration of the possible relationship between the cornea permeability and eye irritation potency of a solute is presented.

MATERIALS AND METHODS

Permeabilities of the cornea, endothelium plus stroma, and epithelium. The dependent variables used in MI-QSAR analysis are the log values of the permeabilities across the cornea and its components, the endothelium plus stroma, and the epithelium. The reason for using log values of the permeabilities is to obtain an approximately uniform distribution of this dependent variable over its range.

The experiments to determine the permeability of a rabbit cornea for various compounds have been conducted by researchers from different labs, the details of which can be found in the cited literatures in a study of library construction for literature permeability data (Prausnitz and Noonan, 1998). For example, the measurement of cornea permeability of hydroxypropyl beta cyclodextrin has been described in detail in a previous study (Tirucherai and Mitra, 2003). Briefly, it is performed in an incubator at 34°C (temperature of the rabbit cornea). Each compound of interest is placed in isotonic phosphate buffer solution (IPBS) of a particular pH, and then added to the epithelial side of the cornea (donor chamber) at concentrations of saturation solubility. In the other half chamber (receiver chamber), pure IPBS at the same pH value is added. The cumulative amount of the compound in the receiver chamber is plotted as a function of time to determine the cornea permeability based on Equation 1:

\[
\text{Permeability} = \frac{dM}{dt} \div A \times C_d
\]

In Equation 1, \(dM/dt\) is the rate of appearance of solute in the receiver chamber, \(C_d\) is the initial concentration of the solute in the donor chamber, and \(A\) is the cross-sectional area of the barrier available for diffusion.

Prausnitz and coworkers have collected literature permeability data for almost 150 compounds for transport across the cornea, sclera, and conjunctiva, as well as the epithelium, stroma, and endothelium (Prausnitz and Noonan, 1998). They subsequently developed theoretical models, based on 25 compounds from the original data set which have permeabilities measures for both the cornea, and endothelium plus stroma at pH 7.6 – 7.65 (Edward and Prausnitz, 2001). These 25 structurally and chemically diverse compounds, which also vary in net charge at the common pH, were used as the training set in the study represented here. Since the cornea is composed of epithelium, stroma and endothelium layers, the permeability of the epithelium can be expressed as:

\[
\frac{1}{P_{ep}} = \frac{1}{P_{cornea}} - \frac{1}{P_{endo+stroma}}
\]

In Equation 2, \(P_{ep}\) is the estimated permeability of epithelium, and \(P_{cornea}\) and \(P_{endo+stroma}\) are the measured permeabilities of the cornea and endothelium plus stroma, respectively. Log values of the measured and calculated permeabilities of the 25 compounds in the training set are listed in Table 1.

Seven structurally diverse compounds from the original literature permeability data set were used as a test set in this study, and are shown in Table 1. The cornea permeabilities of the test set compounds were measured at the same pH as the compounds of the training set.

Thirteen compounds from different chemical classes were used as the eye irritation data set and are shown in Table 2. These compounds were selected from a previous MI-QSAR study using the European Center for Ecotoxicology and Toxicology of Chemicals (ECETOC) data set (Ubels et al., 2000). The measured molar-adjusted eye scores (MES) of these thirteen compounds are close to the predicted scores from the resulting MI-QSAR model suggesting the descriptors of the MI-QSAR model for these compounds are significant.

Solute molecules. The compounds of the training set, test set and eye irritation data set were constructed using HyperChem software (Hyperchem, 2000). The AM1 partial atomic charges were then added to these molecules. Structural optimization of each of these training set compounds was carried out using the Chemlab-II software (Perlstein, 1988).

Construction of the DMPC monolayer. A single dimyristoylphosphatidylcholine (DMPC) molecule was built using HyperChem from available crystal structure data (Hauser et al., 1981). The aliphatic chains of the DMPC molecule were assigned the trans-planar, local minimum energy conformation. The AM1 Hamiltonian in Mopac 6.0 was used for the estimation of partial atomic charges on all molecules (Mopac, 1990). DMPC is the phospholipid used in the model membrane in this study. The structure of a DMPC molecule is shown in Figure 1. An assembly of 25 DMPC molecules (5 × 5 × 1) in (x,y,z) directions, respectively, was used as the model membrane monolayer. The size of the monolayer simulation system was selected based on the work done by van der Ploeg and Berendsen (1982). Additional information regarding construction of the model DMPC monolayer used in this MI-QSAR analysis is given in previous references (Kulkarni et al., 2001, 2002; Kulkarni and Hopfinger, 1999).

To prevent unfavorable van der Waals interactions between a solute molecule and the membrane DMPC molecules, the “center” DMPC molecule, located at position (x,y) = (3,3) of the 5 × 5 DMPC monolayer model, was removed from the monolayer, and a test solute molecule inserted in the space created by the missing DMPC molecule. Each of the test solute molecules of
the cornea permeation data set was inserted at three different positions (depths) in the DMPC monolayer with the most polar group of the solute molecule "facing" toward the headgroup region of the monolayer. Three corresponding MDS trajectories were generated for each solute molecule with regard to the trial positions of the solute molecule in the monolayer. The three trial positions are (1) the solute molecule in the headgroup region, (2) the solute molecule between the headgroup region and the aliphatic chains, and (3) the solute molecule in the tail region of the aliphatic chains.

The lowest energy geometry of the solute molecule in the monolayer was sought from the MDS trajectories of each of the three trial solute positions. The three different initial MDS positions of phenylephrine, one of the training set solute molecules, are shown in Figure 2a to illustrate this modeling and simulation procedure. The energetically most favorable geometry of this solute molecule in the model DMPC monolayer is shown in Figure 2b.

**Molecular dynamic simulation.** MDS were carried out using the Molsim package with an extended MM2 force field (Doherty, 2000). The selection of the simulation temperature was based on the phase transition temperature for DMPC, which is 297 K (Bloom et al., 1991). A simulation temperature of 311 K was selected since it is body temperature, and it is also above the primary DMPC phase transition temperature. Temperature was held constant in the MDS by coupling the system to an external fixed temperature bath (Berendsen et al., 1984). The trajectory step size was 0.001 ps over a total simulation time of 10 ps for each solute of the training set, the test set, and the eye irritation data set. Every membrane-solute system, for each of the solutes of the training and test sets of this study, reached equilibrium by about 1500 trajectory steps, that is 1.7 ps. The MDS trajectory for phenylephrine is shown in Figure 3, and the lowest energy equilibrium step on the trajectory, near 5300 steps, corresponds to the phenylephrine-membrane complex geometry shown in Figure 2b. Two-dimensional periodic boundary conditions, PBC, corresponding to the "surface plane" of the monolayer, were adhered to \((a = 38 \text{ Å}^2, b = 38 \text{ Å}^2, c = 80 \text{ Å}^2, \gamma = 96.0^\circ)\) for the DMPC molecules of the monolayer model but not the test solute molecule. By using PBC, it is possible to simulate an infinite system. The angle \(\gamma\) is the angle an extended DMPC molecule makes with the "planar surface" of the monolayer. Only a single solute molecule was explicitly considered in each MDS. Additional details of the membrane-solute MDS can be found in previous references (Kulkarni et al., 2001, 2002; Kulkarni and Hopfinger, 1999).

**Calculation of descriptors.** The QSAR descriptors are the various properties and features of molecules. A descriptor can be intramolecular—an inherent property of a molecule obtained solely from its chemical structure. The other class of descriptors is the set of intermolecular properties and features that are considered in each MDS. Additional details of the membrane-solute MDS can be found in previous references (Kulkarni et al., 2001, 2002; Kulkarni and Hopfinger, 1999).
depend, in part, on the environment in which a molecule is located. These properties and features are normally computed from the interaction between two, or more, molecules. Moreover, the descriptors used in the MI-QSAR analysis can also be divided into (a) general intramolecular solute descriptors, (b) solute aqueous dissolution and solvation descriptors, and (c) solute-membrane interaction descriptors. A trial pool composed of 65 descriptors was used in this study.

The general intramolecular solute descriptors included as part of the trial descriptor pool are defined in Part A of Table 3. The term general is used because solute descriptors in this class may be useful in building models for a range of endpoint measures of a solute.

The intermolecular solvation and dissolution descriptors are reported as Part B of Table 3. $E_{coh}$, the cohesive energy, is a measure of the energy required to remove a molecule from being surrounded by other molecules identical to itself. $T_{cr}$, the hypothetical crystal-melt temperature of a solute, measures the crystal packing strength of a molecule. $T_{g}$, the glass transition temperature of a solute, measures the amorphous packing strength of a molecule. $E_{coh}$, $T_{cr}$, and $T_{g}$ can be considered the properties reflecting the dissolution behavior of a solute. $F_{H_{2}O}$, $FOCT$, and LogP, are the aqueous and 1-octanol solvation free energies of the solutes, and the corresponding 1-octanol/water partition coefficient, respectively. LogDM, the octanol-to-water distribution coefficient of a solute, is the product of the partition coefficient of the un-ionized solute and the un-ionized fraction of the solute at a selected reference pH. LogDM is included in the descriptor pool of this study because most of the solutes in the training and test sets are ionizable molecules, and it is assumed that only the un-ionized form can partition into membrane. $F_{H_{2}O}$, $FOCT$, LogP, and LogDM can be used to describe the aqueous solvation and lipophilic properties of a solute. It should be noted that although these seven descriptors are called intermolecular descriptors, they are all normally computed using intramolecular computational methods.

The intermolecular solute-membrane interaction descriptors can be extracted directly from the MDS trajectories and are listed in Part C of Table 3. These particular intermolecular descriptors were calculated using the most stable (lowest total potential energy) solute-membrane geometry realized from MDS sampling of the three initial positions, see Figure 2a, for each of the solutes. Figure 3 shows a plot of the total potential energy versus simulation time from which the most energetically favorable position of phenylephrine (one of the solutes of the training set) in DMPC is identified which is shown in Figure 2b. Details regarding the methods and algorithms used to compute these solute-membrane descriptors can be found in previous references (Kulkarni et al., 2001, 2002; Kulkarni and Hopfinger, 1999).

Construction and testing of QSAR models. MI-QSAR models were fit using multidimensional linear regression (MLR) and optimized using the genetic function approximation, GFA. GFA is a multidimensional optimization method based on the genetic algorithm paradigm (Rogers, 1994; Rogers and Hopfinger, 1994). Statistical significance in the optimization of a QSAR model using GFA is based on Friedman’s lack of fit (LOF) measure (Friedman, 1988).
The LOF measure is designed to resist overfitting, which is a problem often encountered in constructing statistical models. Since the number of descriptors available in a MI-QSAR analysis normally exceeds the number of observations (training set compounds), the ability to prevent overfitting using GFA is critical to the successful construction of a statistically significant MI-QSAR model.

Smoothing factors of 0.5–3.0, which control model size, and 100,000 GFA crossover operations were used to optimize the MI-QSAR models having different numbers of descriptor terms using the WOLF software (Rogers, 1994). For a given smoothing factor, optimization of a QSAR model was considered to be realized when descriptor usage became constant and independent of an increasing number of crossover operations. A crossover operation is the “birth” of a child model from its parent models.

All 64 intramolecular and intermolecular descriptors in the MI-QSAR trial descriptor pool were used as linear terms during the evolution of genetic function approximation to generate MI-QSAR models. In addition, non-linear terms consisting of spline functions and quadratic term representations of the descriptors were permitted to be randomly created during the GFA optimization process and, therefore, used to generate trial MI-QSAR models.

The quality of a QSAR model can be described in two ways, goodness-of-fit and predictive power. The goodness-of-fit of a MI-QSAR model can be evaluated using the correlation coefficient of fit, $r^2$. Higher values of $r^2$ indicate better fitting of the model to the training set data. The predictive power can be evaluated in two ways. First, the higher the value of the leave-one-out cross validation coefficient, $xv-r^2$, the higher the predictive capability of the MI-QSAR model. Second, the $LogP_{cornea}$ values of the test set compounds can be predicted using the MI-QSAR model constructed from the training set. Small differences (residuals) between corresponding observed and predicted $LogP_{cornea}$ values of the test set compounds indicate high predictivity of the MI-QSAR model. Moreover, the extent of chance effects/correlations can be evaluated using “scrambling” experiments, in which the measured dependent variables, the $LogP_{cornea}$ values of the training set, are randomly “scrambled” with respect to the training set compounds. If meaningful correlations (QSARs) are found among the scrambled data sets, the significance of a QSAR model is suspect (Liu et al., 2003; Waterbeemd, 1995). The absence of any significant correlation for each of the scrambled data sets is taken as evidence of the significance of the MI-QSAR models with respect to the original, non-scrambled data set.

The overall operational steps and strategy of an MI-QSAR analysis has been described in detail in a previous reference (Kulkarni et al., 2001).

**Correlation between cornea permeability and eye irritation potency.** The $LogP_{cornea}$ for each of the 13 compounds of the eye irritation data set can be predicted using the MI-QSAR cornea permeation model. The correlation coefficient of the observed molarity-adjusted eye irritation scores, MES, versus the predicted $LogP_{cornea}$ values of the compounds in the eye irritation set is then used to explore the relationship between cornea permeability and eye irritation potency of a solute molecule.
The combination of linear terms, e.g., 65 general intramolecular solute, intermolecular solvation/dissolution, and intermolecular membrane-solute descriptors, and non-linear terms, e.g., the spline functions and quadratic functions of those descriptors, were used in the construction of the MI-QSAR models. Interestingly, no non-linear terms are found in any of the optimized MI-QSAR models.

Based on the genetic function approximation (GFA) optimization, the best MI-QSAR models, with different numbers of descriptor terms, for cornea permeability are:

**1-term model**

\[
\text{LogP}_{\text{cornea}} = -5.390 + 0.254\text{LogDM} \\
\]

\[n = 25, r^2 = 0.365, xv - r^2 = 0.279\]

**2-term model**

\[
\text{LogP}_{\text{cornea}} = -2.460 + 0.363\text{LogDM} - 0.087\text{Ecoh} \\
\]

\[n = 25, r^2 = 0.683, xv - r^2 = 0.562\]

**3-term model**

\[
\text{LogP}_{\text{cornea}} = -9.440 + 0.329\text{LogDM} \\
- 0.070\text{Ecoh} + 0.004\text{ETOT} \\
\]

\[n = 25, r^2 = 0.804, xv - r^2 = 0.699\]
4-term model

\[
\log P_{\text{cornea}} = -7.080 + 0.337 \log DM - 0.068 \text{Ecoh} + 0.004 \text{ETOT} - 0.013 \text{ETOR}
\]

\[n = 25, \quad r^2 = 0.872, \quad \text{xv} - r^2 = 0.777\]

5-term model

\[
\log P_{\text{cornea}} = -5.870 + 0.347 \log DM - 0.076 \text{Ecoh} + 0.003 \text{ETOT} - 0.015 \text{ETOR} - 0.006 \text{DEVWD}
\]

\[n = 25, \quad r^2 = 0.914, \quad \text{xv} - r^2 = 0.830\]

6-term model

\[
\log P_{\text{cornea}} = -5.310 + 0.238 \log DM - 0.054 \text{Ecoh} + 0.003 \text{ETOT} - 0.018 \text{ETOR} - 0.009 \text{DEVWD} - 0.006 \text{INTC}
\]

\[n = 25, \quad r^2 = 0.933, \quad \text{xv} - r^2 = 0.886\]

7-term model

\[
\log P_{\text{cornea}} = -13.800 + 0.225 \log DM - 0.053 \text{Ecoh} + 0.004 \text{ETOT} - 0.017 \text{ETOR} - 0.010 \text{DEVWD} - 0.006 \text{INTC} - 0.043 \text{Kappa5}
\]

\[n = 25, \quad r^2 = 0.939, \quad \text{xv} - r^2 = 0.892\]

The number of compounds used to construct the correlation equation is given by \(n\), \(r^2\) is the coefficient of determination, and \(\text{xv} - r^2\) is the leave-one-out-cross-validation coefficient of determination.

\(\log DM\), defined as the observed octanol/water distribution coefficient, is the product of un-ionized octanol/water partition coefficient and the fraction of un-ionized solute at pH of 7.6. \(\text{Ecoh}\), the solute cohesive energy, is the average energy needed to remove a solute molecule from its nearest neighbors. \(\log DM\) and \(\text{Ecoh}\) are intermolecular solvation and dissolution descriptors.

\(\text{ETOT}\), is the total energy of the entire solute-membrane system and \(\text{ETOR}\) is the torsional energy component of the entire solute-membrane system. \(\text{DEVWD}\) is the difference between the van der Waals interaction energy of the entire solute-membrane system and the van der Waals interaction energy of solute alone plus that of membrane alone. \(\text{INTC}\) is the intermolecular electrostatic energy between solute and membrane. It is re-iterated that \(\text{ETOT}\), \(\text{ETOR}\), \(\text{DEVWD}\), and \(\text{INTC}\) are the energies corresponding to the lowest total potential energy state of the model system found in the three molecular dynamics simulations (MDS) for the three initial solute positions in the membrane.

\(\text{Kappa5}\), is a molecular connectivity index, which measures size and shape features of a molecule (Kier, 1980). It is a general intramolecular solute descriptor.

The computed values of \(\log DM\), \(\text{Ecoh}\), \(\text{ETOT}\), \(\text{ETOR}\), \(\text{DEVWD}\), and \(\text{INTC}\) found in the 1–6 term MI-QSAR models for each compound in the training and test sets are listed in Table 4. The energy term descriptors are expressed in kcal/mol. The observed and predicted values of cornea permeability of the training and test set compounds, using the 4–7 term MI-QSAR models, are given in Table 5. None of the compounds in

---

<table>
<thead>
<tr>
<th>Name</th>
<th>(\log DM)</th>
<th>(\text{Ecoh})</th>
<th>(\text{ETOT})</th>
<th>(\text{ETOR})</th>
<th>(\text{DEVWD})</th>
<th>(\text{INTC})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetobutol</td>
<td>-0.07</td>
<td>39.57</td>
<td>1826.22</td>
<td>217.97</td>
<td>-7.84</td>
<td>57.24</td>
</tr>
<tr>
<td>Acetazolamide</td>
<td>-1.31</td>
<td>32.81</td>
<td>1703.35</td>
<td>199.04</td>
<td>43.30</td>
<td>54.21</td>
</tr>
<tr>
<td>Acetazolamide der. 1</td>
<td>-0.38</td>
<td>39.29</td>
<td>1611.11</td>
<td>221.80</td>
<td>-13.40</td>
<td>65.38</td>
</tr>
<tr>
<td>Acetazolamide der. 2</td>
<td>-0.53</td>
<td>31.13</td>
<td>1715.65</td>
<td>239.36</td>
<td>-1.88</td>
<td>52.84</td>
</tr>
<tr>
<td>Alpha_Yohimbine</td>
<td>2.15</td>
<td>36.04</td>
<td>1842.40</td>
<td>195.29</td>
<td>30.81</td>
<td>10.15</td>
</tr>
<tr>
<td>Atenolol</td>
<td>-2.11</td>
<td>33.01</td>
<td>1803.19</td>
<td>205.86</td>
<td>-21.50</td>
<td>125.13</td>
</tr>
<tr>
<td>Benzolamide</td>
<td>-1.38</td>
<td>39.34</td>
<td>1743.09</td>
<td>200.48</td>
<td>23.10</td>
<td>79.76</td>
</tr>
<tr>
<td>Bromacetolazolamide</td>
<td>-2.02</td>
<td>36.55</td>
<td>1715.99</td>
<td>183.64</td>
<td>23.61</td>
<td>96.53</td>
</tr>
<tr>
<td>Chlorozolamide</td>
<td>1.49</td>
<td>33.47</td>
<td>1749.03</td>
<td>191.35</td>
<td>-12.00</td>
<td>46.37</td>
</tr>
<tr>
<td>Clonidine</td>
<td>-2.16</td>
<td>20.62</td>
<td>1841.75</td>
<td>187.20</td>
<td>21.79</td>
<td>21.75</td>
</tr>
<tr>
<td>Corynathine</td>
<td>2.24</td>
<td>36.77</td>
<td>1749.75</td>
<td>193.49</td>
<td>7.37</td>
<td>42.71</td>
</tr>
<tr>
<td>Levobunolol</td>
<td>0.56</td>
<td>31.39</td>
<td>1910.03</td>
<td>231.37</td>
<td>-3.47</td>
<td>18.07</td>
</tr>
<tr>
<td>Methazolamide</td>
<td>-0.38</td>
<td>32.20</td>
<td>1802.06</td>
<td>203.94</td>
<td>6.43</td>
<td>11.43</td>
</tr>
<tr>
<td>Methazolamide der. 1</td>
<td>-3.60</td>
<td>26.05</td>
<td>1774.71</td>
<td>201.36</td>
<td>12.39</td>
<td>132.17</td>
</tr>
<tr>
<td>Metoprolol</td>
<td>-0.80</td>
<td>27.09</td>
<td>1826.72</td>
<td>187.92</td>
<td>-7.65</td>
<td>10.47</td>
</tr>
<tr>
<td>Nadolol</td>
<td>-1.77</td>
<td>39.86</td>
<td>1850.61</td>
<td>190.66</td>
<td>-42.20</td>
<td>125.40</td>
</tr>
<tr>
<td>Oxprenolol</td>
<td>-0.01</td>
<td>26.85</td>
<td>1939.57</td>
<td>200.47</td>
<td>16.44</td>
<td>56.34</td>
</tr>
<tr>
<td>Phenylephrine</td>
<td>-2.72</td>
<td>25.11</td>
<td>1759.84</td>
<td>202.97</td>
<td>38.2</td>
<td>43.86</td>
</tr>
<tr>
<td>Rauwolffine</td>
<td>0.70</td>
<td>34.65</td>
<td>1875.58</td>
<td>220.23</td>
<td>-3.22</td>
<td>28.49</td>
</tr>
<tr>
<td>Sotalol</td>
<td>-2.30</td>
<td>30.04</td>
<td>1844.41</td>
<td>210.92</td>
<td>-47.82</td>
<td>72.15</td>
</tr>
<tr>
<td>Timolo</td>
<td>0.09</td>
<td>36.97</td>
<td>1884.3</td>
<td>203.26</td>
<td>-21.70</td>
<td>58.54</td>
</tr>
<tr>
<td>Trichlormethazolamide</td>
<td>2.50</td>
<td>37.56</td>
<td>1825.33</td>
<td>216.35</td>
<td>48.70</td>
<td>29.98</td>
</tr>
<tr>
<td>Trifluoromethazolamide</td>
<td>1.28</td>
<td>33.07</td>
<td>1787.50</td>
<td>197.90</td>
<td>39.96</td>
<td>84.52</td>
</tr>
<tr>
<td>Vidarabine</td>
<td>-1.46</td>
<td>37.90</td>
<td>1826.56</td>
<td>188.97</td>
<td>47.70</td>
<td>49.19</td>
</tr>
<tr>
<td>Yohimbine</td>
<td>2.10</td>
<td>36.77</td>
<td>1730.95</td>
<td>181.80</td>
<td>10.39</td>
<td>43.59</td>
</tr>
</tbody>
</table>

**Note.** The descriptors are defined in Table 3. All energies are in kcal/mol.
either the training or test sets is an outlier for any of the 4–7 term MI-QSAR models.

A QSAR model is usually considered significant if it has a coefficient of determination \( r^2 \) greater than 0.7. The 1-term MI-QSAR model, Equation 3, has \( r^2 \) of 0.365, which suggests that no descriptor, by itself, can significantly explain the variance of the dependent variable, \( \log P_{\text{cornea}} \). The 2-term and 3-term models represent a distinctly large statistical improvement over the 1-term model since the coefficient of determination, \( r^2 \), increases from 0.365 to 0.804 as the number of terms increases from 1 to 3. The 3-term MI-QSAR model may capture the essential features of the mechanism responsible for cornea permeability as represented by the \( \log P_{\text{cornea}} \) values, see the Discussion section. However, increasing the number of descriptor terms in the MI-QSAR models from 3 to 7, increases \( r^2 \) from 0.804 to 0.939, which may indicate that the 4–7-term MI-QSAR models successively refine, as opposed to overfit, the 3-term model for the training set. The possible significance of the descriptors added in the 4–7 term MI-QSAR models is that they begin to further reveal the essence of the mechanism of cornea permeation that may only be further ascertained by consideration of an expanded training set.

The predictive power of a QSAR model can be explored using the cross-validation coefficient of determination, \( xv-r^2 \), for the training set, and the differences between observed and predicted values of the dependent variable for the test set. Figure 4 contains a plot of \( xv-r^2 \) for the training set as a function of the number of descriptor terms in the model. The 6-term and 7-term models have the highest values of \( xv-r^2 \), 0.886 and 0.892, which may indicate that the 6-term and 7-term models have the most predictive capability. Figure 5 shows the observed versus predicted \( \log P_{\text{cornea}} \) values for the test set using the 4–7 term MI-QSAR models. The smallest average difference between the observed and the measured cornea permeability of the test set is obtained using the 6-term model. This finding suggests that the 7-term model may somewhat overfit the training set data, and that the 6-term model has the highest predictive power.

Scrambling experiments to ascertain the validity of the 6-term model lead to scrambled models with an average \( xv-r^2 \) value of 0.477 as shown in Table 6. These low \( xv-r^2 \) values for the scrambling experiments suggest that the 6-term model is a significant MI-QSAR model, and not the result of chance correlation.

Still, the 4-term model, Equation 6, seemingly captures the essential features of cornea permeation. This model has an \( r^2 \) of 0.872 and a \( xv-r^2 \) of 0.777, indicating it is a significant model. Therefore, descriptors in the 4-term model might most reflect the biological mechanism of cornea permeation. The significance of other descriptor terms in the 5-term and 6-term models need to be further explored by expanding the size of the training set.

![FIG. 4. The leave-one-out-cross-validation coefficients of determination, \( xv-r^2 \), for 1–7 term MI-QSAR models versus the numbers of terms in the corresponding models.](https://academic.oup.com/toxsci/article-abstract/88/2/434/1691204)
MI-QSAR Models for Permeability of the Cornea Components

The best MI-QSAR models with different numbers of descriptor terms for epithelium permeability of the training set compounds listed in Table 1 are:

1-term model

$$\log_{\text{Pepi}} = -5.186 + 0.299 \log_{\text{DM}}$$

$$n = 25, r^2 = 0.318, \text{svr} = 0.228$$

2-term model

$$\log_{\text{Pepi}} = -1.343 + 0.442 \log_{\text{DM}} - 0.114 \text{Ecoh}$$

$$n = 25, r^2 = 0.659, \text{svr} = 0.537$$

3-term model

$$\log_{\text{Pepi}} = -10.126 + 0.400 \log_{\text{DM}} - 0.093 \text{Ecoh} + 0.004 \text{ETOT}$$

$$n = 25, r^2 = 0.780, \text{svr} = 0.675$$

4-term model

$$\log_{\text{Pepi}} = -7.098 + 0.405 \log_{\text{DM}} - 0.091 \text{Ecoh} + 0.005 \text{ETOT} - 0.017 \text{ETOR}$$

$$n = 25, r^2 = 0.851, \text{svr} = 0.761$$

5-term model

$$\log_{\text{Pepi}} = -5.470 + 0.423 \log_{\text{DM}} - 0.102 \text{Ecoh} + 0.004 \text{ETOT} - 0.019 \text{ETOR} - 0.009 \text{DEVDW}$$

$$n = 25, r^2 = 0.904, \text{svr} = 0.814$$

6-term model

$$\log_{\text{Pepi}} = -4.710 + 0.275 \log_{\text{DM}} - 0.072 \text{Ecoh} + 0.004 \text{ETOT} - 0.023 \text{ETOR} - 0.013 \text{DEVDW} - 0.009 \text{INTC}$$

$$n = 25, r^2 = 0.936, \text{svr} = 0.891$$

The descriptors found in Equations 10–15 are exactly the same as those found in Equations 3–8. The values of the coefficient of determination, $r^2$, and the cross-validation coefficient of determination, svr, of Equations 10–15, are similar to those corresponding values of Equations 3–8, indicating that the MI-QSAR models for epithelium permeability, Equations 3–8, may have a similar goodness-of-fit and predictive power to the MI-QSAR models for cornea permeability, Equations 10–15.

FIG. 5. The observed versus predicted log values of cornea permeability, $\log_{\text{Pcornea}}$, for the compounds of the test set using the 4–7 term MI-QSAR models.

TABLE 6

The Leave-One-Out Cross Validation Coefficients, svr, of the MI-QSAR Models from the Ten Scrambled Data Sets

<table>
<thead>
<tr>
<th>Set</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>svr</td>
<td>0.587</td>
<td>0.798</td>
<td>0.459</td>
<td>0.541</td>
<td>0.341</td>
<td>0.477</td>
<td>0.343</td>
<td>0.387</td>
<td>0.543</td>
<td>0.292</td>
</tr>
</tbody>
</table>

Average $0.477 < 0.886^a$

svr

$^a0.886$ is the value of svr of the 6-term MI-QSAR model, Equation 8, from the original training set.
The 4-term model for epithelium permeability, Equation 13, can be considered as the best QSAR model to describe the basic features of epithelium permeation behavior of a compound relative to the size of the training set.

Finally, the best MI-QSAR models with different numbers of descriptor terms for endothelium plus stroma permeability for the training set compounds listed in Table 1 are:

1-term model

\[
\log P_{\text{endo+stroma}} = -4.676 + 0.063 \log DM
\]

\[n = 25, \ r^2 = 0.223, \ xv - r^2 = 0.107\]

2-term model

\[
\log P_{\text{endo+stroma}} = -3.643 + 0.102 \log DM - 0.031 \text{Ecoh}
\]

\[n = 25, \ r^2 = 0.609, \ xv - r^2 = 0.480\]

3-term model

\[
\log P_{\text{endo+stroma}} = -5.691 + 0.092 \log DM - 0.026 \text{Ecoh} + 0.001 \text{ETOT}
\]

\[n = 25, \ r^2 = 0.711, \ xv - r^2 = 0.569\]

4-term model

\[
\log P_{\text{endo+stroma}} = -4.624 + 0.094 \log DM - 0.025 \text{Ecoh} + 0.001 \text{ETOT} - 0.006 \text{ETOR}
\]

\[n = 25, \ r^2 = 0.837, \ xv - r^2 = 0.754\]

5-term model

\[
\log P_{\text{endo+stroma}} = -4.814 + 0.095 \log DM - 0.024 \text{Ecoh} + 0.001 \text{ETOT} - 0.004 \text{ETOR} - 0.007 \text{LETOR}
\]

\[n = 25, \ r^2 = 0.863, \ xv - r^2 = 0.774\]

6-term model

\[
\log P_{\text{endo+stroma}} = -5.370 + 0.095 \log DM - 0.020 \text{Ecoh} + 0.001 \text{ETOT} - 0.004 \text{ETOR} - 0.010 \text{LETOR} - 0.025 \text{Kappa3}
\]

\[n = 25, \ r^2 = 0.904, \ xv - r^2 = 0.832\]

The four terms in Equations 16–19 are the same as the corresponding terms in Equations 3–6. However, two terms, LETOR and Kappa3, in Equations 20 and 21, are different from two terms, DEVDW and INTC, in Equations 7 and 8. LETOR is the intramolecular torsional energy of the solute molecule in the solute-membrane complex having the lowest total potential energy state. Although LETOR describes an intramolecular property of the solute, it is classified as an intermolecular solute-membrane interaction descriptor because it is calculated using data explicitly determined for the membrane-solute complex. Kappa3, a topological index, is a general intramolecular descriptor measuring the size and shape of the solute.

The values of \(r^2\) and \(xv-r^2\) of Equations 16–21 are less than those of Equations 3–8, respectively. Thus, the MI-QSAR models for endothelium plus stroma, Equations 16–21, have less significant fit to the training set, and less predictive power, than the MI-QSAR models for cornea permeability, Equations 3–8.

The 4-term model for endothelium plus stroma permeability, Equation 19, can be considered as the best QSAR model to describe the basic features of endothelium plus stroma permeation behavior of a compound relative to the limited size of the training set.

**Correlation between Eye Irritation Potency and Cornea Permeability**

The 6-term MI-QSAR model, Equation 8, is used to predict the \(\log P_{\text{cornea}}\) values of the thirteen compounds in the eye irritation set listed in Table 2. The \(\log P_{\text{cornea}}\) values in Table 2 are the molar-adjusted eye irritation scores for the thirteen compounds. The larger the MES value, the greater the eye irritation potency. Moreover, Figure 6 contains a plot of the predicted values of \(\log P_{\text{cornea}}\) versus the MES values of the eye irritation set of compounds. The corresponding correlation coefficient squared, \(R^2\), is 0.023, which indicates there is no correlation between eye irritation potency and the cornea permeability of a solute. That is, for the data sets used in this study, eye irritation is not correlated to cornea permeability.

![FIG. 6. The predicted log values of cornea permeability versus the observed molar-adjusted eye irritation score (MES) for the compounds of the eye irritation set.](https://academic.oup.com/toxsci/article-abstract/88/2/434/1691204)
DISCUSSION

**MI-QSAR Models for Cornea Permeability**

The octanol/water distribution coefficient, LogDM, of a molecule is the product of the neutral octanol/water partition coefficient, LogP, and the fraction of the un-ionized solute molecule. In the MI-QSAR model of highest prediction for cornea permeability, Equation 8, LogDM is positively correlated to LogP\textsubscript{cornea} indicating that increasing solute LogP, and/or increasing the population of the solute in the un-ionized form, can increase cornea permeation. This finding is similar to reported observations where logDM has been shown to have a direct positive relationship to cornea permeability (Edward and Prausnitz, 2001; Prausnitz and Noonan, 1998). An increase in LogP, reflecting an increase in lipophilicity, always corresponds to an increase in cornea permeability. Moreover, it is thought that only the un-ionized form of a solute permeates the membrane. Hence, less ionized solute molecules permeate the cornea in higher concentrations of un-ionized form, which corresponds to higher cornea permeabilities. It is important to note that although the descriptor pool used for MI-QSAR model construction includes both LogP and LogDM, only LogDM is found in the optimized MI-QSAR models. This finding strongly suggests that the ionization behavior of a solute in water is, in turn, a factor in its cornea permeation behavior.

Ecoh, the cohesive energy, which is the interaction energy of a molecule with its like neighbors, describes the dissolution behavior of a molecule. The greater the value of Ecoh, the more difficult it is to dissolve a solute molecule in a solvent. Ecoh has a negative regression coefficient in Equation 8, indicating an increase in Ecoh corresponds to a decrease in cornea permeability. This relationship suggests that as dissolution of a solute becomes more difficult, the lower will be cornea permeability of the solute molecule.

ETOT, the total potential energy of the solute-membrane complex, is positively correlated to cornea permeability in Equation 8. This relationship indicates that a higher total potential energy of the complex corresponds to higher cornea permeability of the solute. The value of ETOT represents the strength of solute binding to the membrane and the stability of solute-membrane complex. A higher ETOT corresponds to less binding of the solute to the membrane structure and, hence, greater permeation across the cornea membrane.

ETOR, the total torsion energy of the solute-membrane complex, is always positive in energy value and measures the extent of torsional deformation energy that occurs in the membrane-solute system. The greater the value of ETOR, the greater torsional deformation energy needed for the solute to navigate through the membrane (Kulkarni et al., 2002). The regression coefficient for this descriptor is negative in Equation 8, and LogP\textsubscript{cornea} is predicted, as expected, to decrease as ETOR increases.

DEVDW is the difference between the van der Waals interaction energy of the entire solute-membrane system, and the van der Waals interaction energy of solute alone plus that of membrane alone. This term is a measure of the steric deformation in the solute-membrane complex that occurs with uptake of the solute. The negative regression coefficient of DEVDW in Equation 8 indicates, as expected, that an increase in steric deformation energy interaction of the membrane, due to solute uptake, corresponds to a decrease in solute cornea permeability.

INTC, the electrostatic interaction energy between the solute and membrane, is negatively correlated to LogP\textsubscript{cornea} in Equation 8, indicating that cornea permeability decreases as the value of INTC increases. This relationship between LogP\textsubscript{cornea} and INTC is counter-intuitive. As INTC increases, there is less electrostatic binding of the solute to the membrane, and, consequently, LogP\textsubscript{cornea} should increase. Perhaps INTC is a "correction term" to the more significant descriptors involving membrane-solute interactions found in Equations 3–5.

Based on the analysis of the terms in Equation 8, MI-QSAR models for cornea permeability can be mechanistically interpreted as consisting of the following contributing factors:

- **Lipophilicity**—An increase in lipophilicity of the solute, as measured by an increasing value of LogDM, increases cornea permeability.
- **Fraction of un-ionized solute**—A higher fraction of un-ionized solute corresponds to a higher cornea permeability.
- **Dissolution**—A solute that is hard to completely dissolve in an aqueous solvent, as measured by a high value of Ecoh, has low cornea permeability.
- **Distortion of the membrane-solute complex**—The more the structural distortion of the membrane-solute complex, the lower the cornea permeability.
- **Solute-membrane binding**—The greater the binding of the solute to the membrane, the less is the cornea permeability.

**MI-QSAR Models for Permeability of the Cornea Components**

The MI-QSAR models for epithelium permeability, Equations 10–15, have the same respective descriptor terms as the MI-QSAR models for cornea permeability, Equations 3–8. That is, epithelium permeation appears to be nearly identical, in terms of mechanistic action, to permeation across the entire cornea. Moreover, the epithelium is the main barrier in passing through the cornea for 20 of the 25 compounds of the training set. This behavior is indicated by the lower permeability values of epithelium of the 20 compounds than the corresponding permeation values for the stroma plus endothelium given in
Table 1. Thus, permeation of the epithelium largely governs permeation behavior across the entire cornea.

The endothelium and stroma are, in composite, the main barrier of cornea permeation for the other five compounds in the training set, which are alpha-yohimbine, clonidine, levobunolol, metoprolol, and oxprenolol. Based on a comparison of the descriptor terms, and the corresponding regression coefficients in the 4-term models, Equations 13 and 19, permeation in the epithelium is faster than permeation in the endothelium and stroma if a solute has high values of LogDM, ETOT, and/or low values of Ecoh, ETOR. As shown in Table 4, alpha-yohimbine and levobunolol have high values of LogDM, while clonidine, metoprolol and oxprenolol have low values of Ecoh. Moreover, all five compounds have high values of ETOT and low values of ETOR. The values of $r^2$ and $xv-r^2$ of the 4-term MI-QSAR model for permeability of the stroma plus endothelium, Equation 19, are 0.837 and 0.754, respectively. This information may give some hint as to which component, stroma or endothelium, plays a more important role in the permeation through the cornea in the absence of the epithelium. MI-QSAR analysis is designed for investigating ADMET properties where the cell membranes play a role. Therefore, MI-QSAR analysis can only yield significant QSAR models that involve membrane-solute interaction descriptors for a training set in which the main barriers to cornea permeation involve membranes, that is, the epithelium or the endothelium, but not the stroma. The MI-QSAR model for stroma plus endothelium, Equation 19, is significant since it has a value of $r^2$ greater than 0.7, and half the descriptors are explicit membrane-solute descriptors. These QSAR features suggest that the endothelium cellular component is the main-barrier to solute permeation across a cornea without an epithelial layer. Since both Equations 13 and 19 are QSAR models for permeation dominated by the endothelium and the epithelium cellular components, respectively, it is not surprising that these two models have common descriptor terms. The comparison between the MI-QSAR models for the cornea and its components leads to revealing information with respect to the main barriers in solute permeation.

**Correlation between Eye Irritation Potency and Cornea Permeability**

A major finding from this study is that there is no meaningful relationship between eye irritation potency and cornea permeability based on the data sets used in the analysis. This non-relationship can be gleaned from the low value of the square of the correlation coefficient, $R^2$, between the predicted values of LogP$_{cornea}$ versus the observed MES values for the thirteen compounds in the eye irritation set.

Employing a similar descriptor pool to that used in this current study, a significant QSAR model for eye irritation has been constructed using MI-QSAR analysis in a previous study (Kulkarni et al., 2001):

\[
\text{MES} = -0.660 - 0.010E(\text{chg} + \text{vdw}) - 0.480F(H_2O) + 0.390\text{LUMO}
\]

\[n = 36, r^2 = 0.710, xv - r^2 = 0.650\]

Where, MES is the molar-adjusted eye score, $E(\text{chg} + \text{vdw})$ is the electrostatic plus van de Waals interaction energy between the solute and the membrane at the total system minimum potential energy, $F(H_2O)$ is the aqueous solvation energy computed using a hydration shell model (Hopfinger, 1973), and LUMO is the lowest unoccupied molecular orbital energy. This QSAR model suggests that the eye irritation potency of a solute is partially related to its chemical reactivity potency, as represented by the LUMO descriptor. However, in the QSAR model for cornea permeability, Equation 6, no descriptor related to reactivity potencies present. This difference in the two QSAR models may account, at least in part, for the lack of a relationship between eye irritation potency and cornea permeability. In addition, for Equation 22 as $E(\text{chg} + \text{vdw})$ becomes more negative in value, the binding of the solute to the membrane increases, and the eye-irritation potency correspondingly increases. However, in the QSAR model for cornea permeability, Equation 6, as ETOT becomes more negative in value, the binding of the solute to the membrane increases, but cornea permeability decreases. Thus, solute-membrane binding leads to something of an inverse relationship between eye irritation potency and cornea permeability according to the QSAR models. The lack of a relationship between cornea permeability and eye irritation potency, based on the MI-QSAR analysis, suggests that a change in the value of any term (molecular property) in the MI-QSAR model for cornea permeability, Equation 6, will change cornea permeability, but have a minimal effect on eye irritation potency, and vice versa.

In general, MI-QSAR analysis can be employed to do cross comparison studies on pairs of ADMET properties provided significant MI-QSAR models can be constructed for each property of the pair. If there is no apparent relationship between a pair of ADMET properties, then it is likely that a change in the value of any descriptor (property) of the MI-QSAR model for one of the pair, will not produce a change in the other ADMET property. If a relationship is established for a pair of ADMET properties, then a change in the value of any term (property) in the MI-QSAR model of one ADMET property of the pair, can be used to estimate the change in the other ADMET property. This “design” capability should be useful to optimize the influence of multiple, and coupled, ADMET properties that may be both good and bad features, but which are governed by very similar molecular mechanisms and interactions.
ACKNOWLEDGMENTS

Resources of the Laboratory of Molecular Modeling and Design at UIC, and of The Chem21 Group Inc. were used in performing this work. We are also grateful for support from Avon and The Procter & Gamble Company in performing this study.

REFERENCES


