articles

molecular pharmaceutics

Characterization of Skin Penetration Processes of Organic Molecules Using Molecular Similarity and QSAR Analysis

Osvaldo A. Santos-Filho,[†] A. J. Hopfinger,[†] and Tao Zheng*[,‡]

Laboratory of Molecular Modeling and Design (M/C 781), College of Pharmacy, The University of Illinois at Chicago, 833 South Wood Street, Chicago, Illinois 60612-7231, and New Technology Department, Global Research & Development, Avon Products, Inc., 1 Avon Place, Suffern, New York 10901-5605

Received July 19, 2004

Abstract: Molecular similarity and QSAR analyses have been used to develop compact, robust, and definitive models for skin penetration by organic compounds. The QSAR models have been sought to provide an interpretation and characterization of plausible molecular mechanisms of skin penetration. A training set of 40 structurally diverse compounds were selected to be representative of a parent set of 152 compounds in terms of both structural diversity and range in measured skin penetration. The subset of 40 compounds was used in a series of QSAR analyses in the search for the most significant, compact, and straightforward skin penetration QSAR models. Molecular dynamics simulations were employed to determine a set of MI (membrane-interaction) descriptors for each test compound (solute) interacting with a model DMPC monolayer membrane model. The MI-QSAR models may capture features of cellular membrane lateral transverse transport involved in the overall skin penetration process by organic compounds. An additional set of intramolecular solute descriptors, the non-MI-QSAR descriptors, were computed and added to the trial pool of descriptors for building QSAR models. All QSAR models were constructed using multidimensional linear regression fitting and a genetic algorithm optimization function. QSAR models were constructed using only non-MI-QSAR descriptors and using a combination of both these descriptor sets. It was found that a combination of non-MI-QSAR and MI-QSAR descriptors yielded the optimum models, not only with respect to the statistical measures of fit but also regarding model predictivity.

Keywords: QSAR; skin penetration; lipophilicity; polar surface area; molecular shape

Introduction

Transdermal drug delivery is an attractive route for administering systemically active drugs. The selective uptake of compounds into the outermost layer of skin, the stratum corneum, for beauty and skin care applications is also a popular research goal. However, the stratum corneum provides a protective barrier that prevents the loss of physiologically significant molecular entities, and guards against the entry of toxic agents from the external environment. Data for structure−skin penetration relationships have been, and are continuing to be, gathered, but a definitive and quantitative way of predicting percutaneous penetration as a function of chemical structure remains elusive. Specification of the corresponding mechanism of action for penetrating the stratum corneum remains largely unknown. Elucidation of the mechanism of skin penetration, and its implementation for making quantitative predictions on test compounds, would be a major benefit to both drug delivery and beauty and skin care research.

* To whom correspondence should be addressed: New Technology Department, Global Research & Development, Avon Products, Inc., 1 Avon Place, Suffern, NY 10901-5605. Tel: (845)369-2690. Fax: (845)369-2402. E-mail: tao.zheng@avon.com.

[†] The University of Illinois at Chicago.

[‡] Avon Products, Inc.

10.1021/mp049924+ CCC: $27.50
Published on Web 09/30/2004

Several predictive skin penetration models have been reported in which different descriptor sets have been used to build the models.[1−4] Moreover, a variety of optimization and fitting procedures, including ANN (artificial neural network analysis), MLR (multiple linear regression), and PCA (principal component analysis) have been used in model building.[5,6] The recent paper by Geinoz and co-workers[7] provided a general guide for the developments and reported models in this area. However, most of these models are based on small data sets and consider only a few descriptors, such as log *P* and molecular weight. In order to provide a general and accurate prediction model that can be used in virtual screening, a comprehensive skin penetration prediction tool, based on larger datasets and more descriptors, is needed. In addition, not only the physicochemical properties of organic compounds but also the interactions between organic compounds and lipid layers in skin stratum corneum should be considered as potential descriptors in constructing a general and reliable prediction model. Interpretation of such a comprehensive model would provide insight on key factors of skin penetration and help formulators to choose appropriate skin delivery systems.

We have developed a methodology called *membrane interaction QSAR (MI-QSAR)* analysis where structure-based design methodology is combined with classic intramolecular QSAR analysis to model chemically and structurally diverse compounds interacting with cellular membranes.[8−10] In MI-QSAR analysis the assumption is made that the phospholipid regions of a cellular membrane constitute the "receptor" required in structure-based design that permits incorporation of structural and chemical diversity into a training set. A set of *membrane−solute intermolecular properties* are determined and added to a set of comprehensive intramolecular solute QSAR descriptors to enlarge the trial QSAR descriptor pool and, ostensibly, to provide the information needed to incorporate chemical and structural diversity into the QSAR analysis. The MI-QSAR descriptor terms have proven to be significant in generating models for several ADMET properties.[8−11] A major part of the study reported here was to determine if MI-QSAR descriptors are essential to building good QSAR models for skin penetration.

Methods

A. Skin Penetration Coefficients. The dependent variable used in the QSAR analyses is the logarithm of the skin penetration coefficient, log k_p, also known as skin permeability. In vitro skin penetration coefficients for 152 organic compounds through human skin were used as the parent data set in units of centimeters/hour (Table 1). These data are a compilation from several literature sources.[6,12−14] Hence, it may be assumed that there is some intrinsic variability among the methods of measurement which gets incorporated into the resultant dataset (Table 1).

B. Building Solute Molecules and a DMPC Monolayer. All the solute molecules were built using the Chemlab-II molecular modeling package.[15] A single dimyristoylphosphatidylcholine (DMPC) molecule was built using HyperChem program[16] from available crystal structure data.[17] The AM1 Hamiltonian in Mopac 6.0 was used for the estimation of partial atomic charges on all molecules.[18]

DMPC was selected as the model phospholipid in this study. The structure of a DMPC molecule is shown in Figure

(1) Kasting, G. B.; Smith, R. L.; Cooper, E. R. Effect of Lipid Solubility and Molecular Size on Percutaneous Absorption. In *Pharmacology and the Skin. Vol. 1. Skin Pharmacokinetics*; Shroot, B., Schaefer, H., Eds.; Kanger: Basel, 1987; pp 138−153.

(2) Flynn, G. L. Physicochemical Determinants of Skin Absorption. In *Principles of Route-to-Route Extrapolation for Risk Assessment*; Gerrity, T. R., Henry, C. J., Eds.; Elsevier: New York, 1990; pp 93−127.

(3) Patel, H.; ten Berge, W. F.; Cronin, M. T. D. Quantitative Structure−Activity Relationships (QSARs) for the Prediction of Skin Permeation of Exogenous Chemicals. *Chemosphere* **2002**, *48*, 603−613.

(4) Fitzpatrick, D.; Corish, J. Modelling Skin Permeability in Risk Assessment: The Future. *Chemosphere* **2004**, *55*, 1309−1314.

(5) Degim, T.; Hadgraft, J.; Ilbasmis, S.; Ozkan, Y. Predicition of Skin Penetration Using Artificial Neural Network (ANN) Modeling. *J. Pharm. Sci.* **2003**, *92*, 656−664.

(6) Barratt, M. D. Quantitative Structure-Activity Relationships for Skin Permeability. *Toxicol. in Vitro* **1995**, *9*, 27−37.

(7) Geinoz, S.; Guy, R. H.; Testa, B.; Carrupt, P. A. Quantitative Structure-Permeation Relationships to Predict Skin Permeation: A Critical Evaluation. *Pharm. Res.* **2004**, *21*, 83−92.

(8) Kulkarni, A. S.; Hopfinger, A. J.; Osborne, R.; Bruner, L. H.; Thompson, E. D. Prediction of Eye Irritation from Organic Chemicals Using Membrane-Interaction QSAR Analysis. *Toxicol. Sci*. **2001**, *59*, 335−345.

(9) Kulkarni, A. S.; Hopfinger, A. J. Membrane-Interaction QSAR Analysis: Application to the Estimation of Eye Irritation by Organic Compounds. *Pharm. Res.* **1999**, *16*, 1244−1252.

(10) Kulkarni, A. S.; Han, Y.; Hopfinger, A. J. Predicting Caco-2 Cell Permeation Coefficients of Organic Molecules Using Membrane-Interaction QSAR Analysis. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 331−342.

(11) Iyer, M.; Mishra, R.; Han, Y.; Hopfinger, A. J. Predicting Blood-Brain Barrier Partitioning of Organic Molecules Using Membrane-Interaction QSAR Analysis. *Pharm. Res.* **2002**, *19*, 1611−1621.

(12) Ursin, C.; Hansen, C. M.; Van Dyk, J. W.; Jensen, P. O.; Christensen, I. J.; Ebbehoej, J. Permeability of Commercial Solvents Through Living Human Skin. *Am. Ind. Hyg. Assoc. J.* **1995**, *56*, 651−660.

(13) Johnson, M. E.; Blankschtein, D.; Langer, R. Evaluation of Solute Permeation Through the Stratum Corneum: Lateral Bilayer Diffusion as the Primary Transport Mechanism. *J. Pharm. Sci.* **1997**, *86*, 1162−1172.

(14) Wilschut, A.; ten Berge, W. F.; Robinson, P. J.; McKone, T. E. Estimating Skin Permeation: The Validation of Five Mathematical Skin Permeation Models. *Chemosphere* **1995**, *30*, 1275−1296.

(15) Pearlstein, R. A. *CHEMLAB-II Users Guide*; CHEMLAB Inc: Chicago, 1988.

(16) HyperChem. *HyperChem Release 4.5 for MS Windows*; Hypercube Inc.: Waterloo, Ontario, 1998.

(17) Hauser, H.; Pascher, I.; Pearson, R. H.; Sundell, S. Preferred Conformation and Molecular Packing of Phosphatidylethanolamine and Phosphatidylcholine. *Biochim. Biophys. Acta* **1981**, *650*, 21−51.

(18) Mopac. *Mopac 6.0 Release Notes*; Frank J. Seiler Research Laboratory: United States Air Force Academy, 1990.

Table 1. The Parent Data Set of Skin Penetration Compounds

compound	log k_p (cm/h)	compound	log k_p (cm/h)	compound	log k_p (cm/h)
1,1,1-trichloroethane	−2.34	corticosterone	−4.22	isoquinoline	−1.78
1,3-dichloropropene	−2.00	cortisone	−5.00	lidocaine	−1.78
17-hydroxyprogesterone	−3.22	cyclohexanone	−2.74	meperidine	−2.43
2,3-butanediol	−4.40	dexamethasone	−4.19	methanol	−3.30
2,4,6-trichlorophenol	−1.23	diclofenac	−1.74	methyl 4-OH-benzoate	−2.04
2,4-dichlorophenol	−1.22	diethyl ether	−1.80	methyl acrylate	−2.68
2-butoxyethanol	−2.85	diethylamine	−2.75	methyl acrylic acid	−2.58
2-chlorophenol	−1.48	diethylcarbamazine	−3.89	methyl Cellosolve	−3.73
2-cresol	−1.80	digitoxin	−4.89	methylene chloride	−2.74
2-ethoxyethanol	−3.60	dimethyl acetamide	−2.80	monomethylhydrazine	−3.75
2-heptanone	−2.00	dimethyl formamide	−1.98	morphine	−5.03
2-hexanone	−2.35	dimethyl sulfoxide	−1.80	morpholine	−3.86
2-naphthol	−1.55	ephedrine	−2.22	*N,N*-dimethylaniline	−1.70
2-pentanone	−2.60	epichlorohydrin	−3.43	*n*-butanol	−2.60
2-phenylethanol	−1.88	estradiol	−2.38	*n*-decanol	−1.10
2-toluidine	−1.44	estriol	−4.40	*n*-heptanol	−1.50
3,4-xylenol	−1.44	estrone	−2.44	*n*-hexanol	−1.89
3-cresol	−1.82	ethanol	−3.10	*n*-methyl-2-pyrrolidone	−1.80
3-nitrophenol	−2.25	ethyl acrylate	−2.39	*n*-nonanol	−1.22
3-xylene	−1.10	ethyl formate	−3.01	*n*-octanol	−1.28
4-bromophenol	−1.44	ethylbenzene	0.08	*n*-pentanol	−2.22
4-chloro-3,5-xylenol	−1.28	ethylene dichloride	−2.00	*n*-propanol	−2.85
4-chloro-*o*-cresol	−1.26	ethylene glycol	−4.07	naproxen	−3.40
4-chlorophenol	−1.44	ethylhexyl phthalate	−1.52	ndela	−5.22
4-cresol	−1.75	etorphine	−2.44	nicotine	−1.71
4-ethylphenol	−1.46	fentanyl	−2.25	nitroglycerine	−1.96
4-methyl-2-pentanol	−2.33	fluocinonide	−2.77	octanoic acid	−1.60
4-nitrophenol	−2.25	formaldehyde	−2.65	ouabain	−6.11
acetic acid	−3.21	γ-butyrolactone	−4.00	pentanoic acid	−2.70
acrylic acid	−3.05	heptanoic acid	−1.70	phenobarbital	−3.34
acrylonitrile	−2.87	hexachloroethane	−1.40	phenol	−2.09
aldosterone	−5.52	hexanoic acid	−1.85	phenylglycidyl ether	−2.84
allyl alcohol	−2.95	hydrocortisone 21-(6-hydroxy)hexanoate	−3.04	pregnenolone	−2.82
amobarbital	−2.64	hydrocortisone 21-(*N,N*-dimethyl)succinamate	−4.17	progesterone	−2.82
aniline	−2.65	hydrocortisone 21-dimethylsuccinamate	−4.17	propionic acid	−2.94
anisole	−1.13	hydrocortisone 21-hemipimelate	−2.75	propylene carbonate	−4.22
atropine	−5.07	hydrocortisone 21-hemisuccinate	−3.20	propylene oxide	−3.05
barbital	−3.95	hydrocortisone 21-hexanoate	−1.75	resorcinol	−3.62
benzaldehyde	−1.21	hydrocortisone 21-methylpimelate	−2.27	salicyclic acid	−2.20
benzene	−0.86	hydrocortisone 21-methylsuccinate	−3.68	scopolamine	−4.30
benzyl alcohol	−2.22	hydrocortisone 21-octanoate	−1.21	styrene	−0.19
butanoic acid	−3.00	hydrocortisone 21-pimelamate	−3.05	sucrose	−5.28
butanone	−2.34	hydrocortisone 21-propionate	−2.47	sufentanyl	−1.92
butobarbital	−3.71	hydrocortisone 21-succinamate	−4.59	sulfolane	−4.70
butyl acrylate	−2.00	hydrocortisone	−5.52	testosterone	−3.40
caffeine	−4.09	hydromorphone	−4.82	thymol	−1.28
catechol	−2.77	hydroxypregnenolone	−3.22	toluene	0.00
chlorpheniramine	−2.66	indometacin	−1.83	triethylamine	−2.31
codeine	−4.31	isoamyl alcohol	−2.00	vinyl acetate	−2.73
cortexolone	−4.13	isobutanol	−2.65	water	−3.30
cortexone	−3.35	isopropylamine	−2.90		

1. An assembly of 25 DMPC molecules (5 × 5 × 1) in (x, y, z) directions, respectively, was used as the model membrane monolayer. Additional information regarding construction of the model monolayer used in this MI-QSAR analysis can be found in refs 3−5.

C. Compound Selection. The 4D molecular similarity analysis software package[19] was used to estimate the molecular similarity of the set of 40 compounds chosen as the training set. The training set of 40 compounds is given in Table 2, and a histogram plot of the distribution of molecular

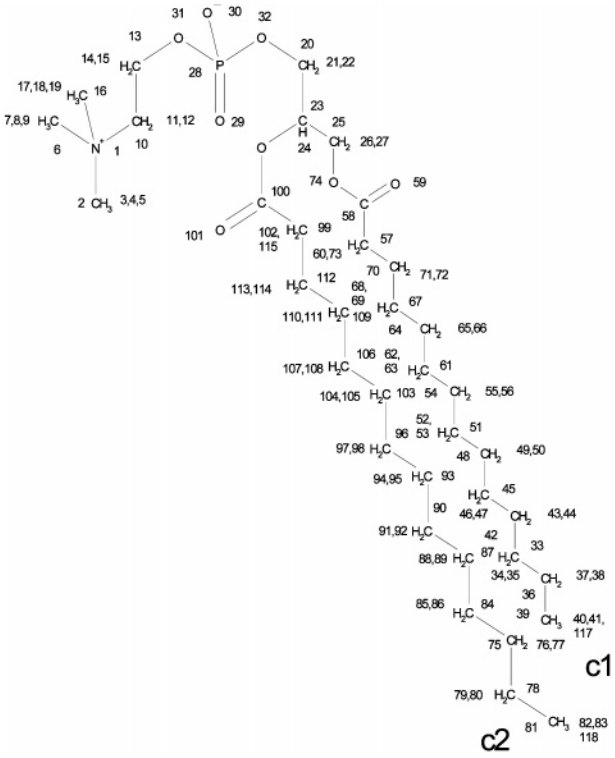

Figure 1. The chemical structure of a DMPC phospholipid molecule with an arbitrary atom numbering assignment. c1 and c2 denote the two aliphatic chains of a DMPC molecule.

Table 2. The 40 Compounds Forming the Training Set

compound	logk_p (cm/h)	compound	logk_p (cm/h)
2-chlorophenol	−1.48	hexanol	−1.89
2,4,6-trichlorophenol	−1.23	hydrocortisone	−5.52
2,4-dichlorophenol	−1.22	hydromorphone	−4.82
4-bromophenol	−1.44	methanol	−3.30
4-chlorophenol	−1.44	methylene chloride	−2.74
4-chloro-*o*-cresol	−1.26	morphine	−5.03
aldosterone	−5.52	ndela	−5.22
atropine	−5.07	ouabain	−6.11
benzaldehyde	−1.21	pentanol	−2.22
benzene	−0.86	phenol	−2.09
butanone	−2.34	propanol	−2.85
butyric acid	−3.00	propylene carbonate	−4.22
caffeine	−4.09	resorcinol	−3.62
corticosterone	−4.22	salicilic acid	−2.20
decanol	−1.10	scopolamine	−4.30
diethyl ether	−1.80	styrene	−0.19
estradiol	−2.38	sucrose	−5.28
estriol	−4.40	sulfolane	−4.70
ethanol	−3.10	testosterone	−3.40
ethyl benzene	0.08	thymol	−1.28

similarity of this training set is given in Figure 2. The distributions of skin penetration coefficients, log k_p, for the parent and training sets are given in Figure 3, parts a and b,

(19) Chem21. *4D-MS [Molecular Similarity] Users Manual*; The Chem21 Group, Inc.: 1780 Wilson Drive, Lake Forest, IL 60045, 2002.

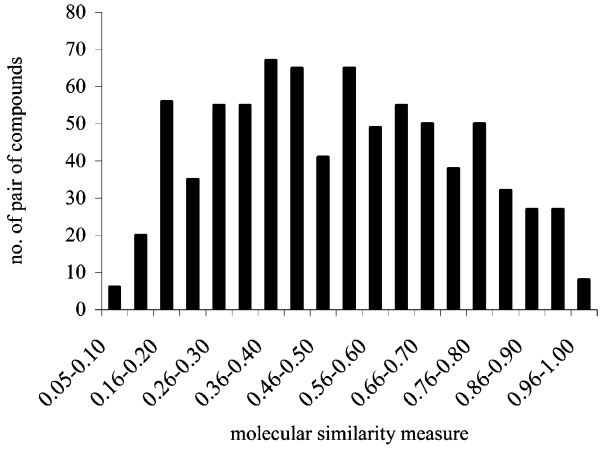

Figure 2. Distribution of 4D molecular similarity measures for the training set.

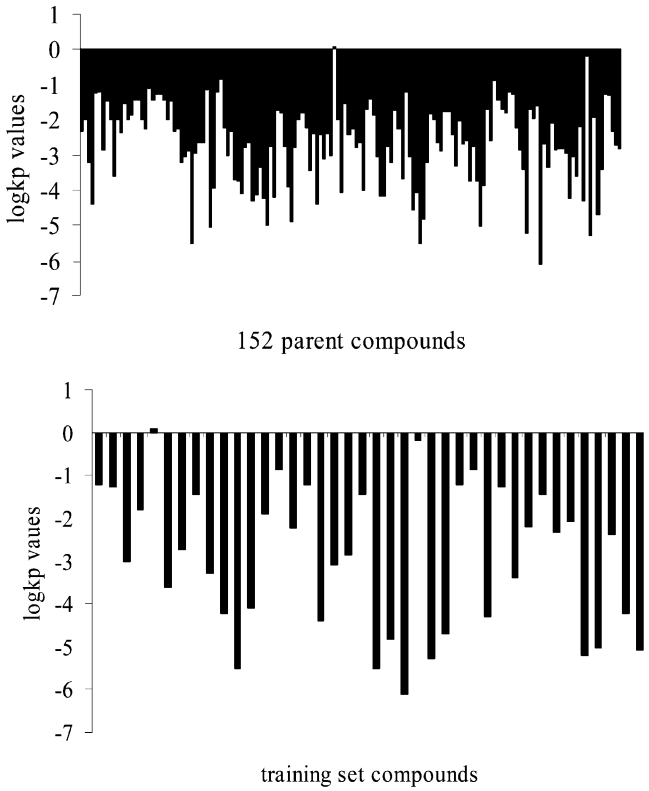

Figure 3. Distribution of log k_p values over the (a) parent data set and (b) training set. The *X*-axes record compound numbers which are arbitrary.

respectively. It is clear from Figures 2 and 3 that the training set of 40 compounds is a "miniature" version of the parent data set.

D. Molecular Dynamics Simulations (MDS). The conditions set for the MDS were established in previous MI-QSAR analyses[8−11] and are only summarized here. An initial MDS on the model membrane, without a solute molecule present, was carried out to allow for structural relaxation and distribution of the kinetic energy over the monolayer. In order 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 equilibrated monolayer and a test solute molecule inserted in the space created by the missing DMPC molecule. Each of the test solute molecules of the 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 models were generated for each solute molecule with regard to the trial positions of the solute molecule in the monolayer. The three trial positions were (1) solute molecule in the headgroup region; (2) solute molecule between the headgroup region and the aliphatic chains; and (3) solute molecule in the tail region of the aliphatic chains.

The lowest energy geometry of the solute molecule in the monolayer was sought using each of the three trial solute positions. The three different initial MDS positions of ethanol are shown in Figure 4a to illustrate this modeling procedure. The energetically most favorable geometry of this solute molecule in the model DMPC monolayer is shown in Figure 4b.

MDS were carried out using the Molsim package with an extended MM2 force field.[20] The simulation temperature of 311 K was selected since it is the normal body temperature. Temperature was held constant in the MDS by coupling the system to an external fixed temperature bath.[21] The trajectory step size was 0.001 ps over a total simulation time of 20 ps for each test compound. Two-dimensional periodic boundary conditions, corresponding to the "surface plane" of the monolayer, were employed ($a = 50$ Å^2, $b = 50$ Å^2, $c = 80$ Å^2, and $\gamma = 90°$) for the DMPC molecules of the monolayer model, but not the test solute molecule. The angle γ 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. Before the MDS of the solute−membrane complexes, each of the solute molecules was placed at each of the three different positions in the monolayer, as described above, with the most polar portion of the solute "facing" toward the headgroup region.

E. Calculation of Descriptors. Both *intramolecular* physicochemical properties and features of the solute molecules and *intermolecular* solute−membrane interaction properties were calculated. "Properties" and "features" will both be referred to as descriptors from this point forward as they constitute the trial pool of independent variables used to build the QSAR models. The intramolecular solute descriptors used in building QSAR models were calculated using Cerius^2.[22] The members of this set of descriptors are listed and defined in Table 3.

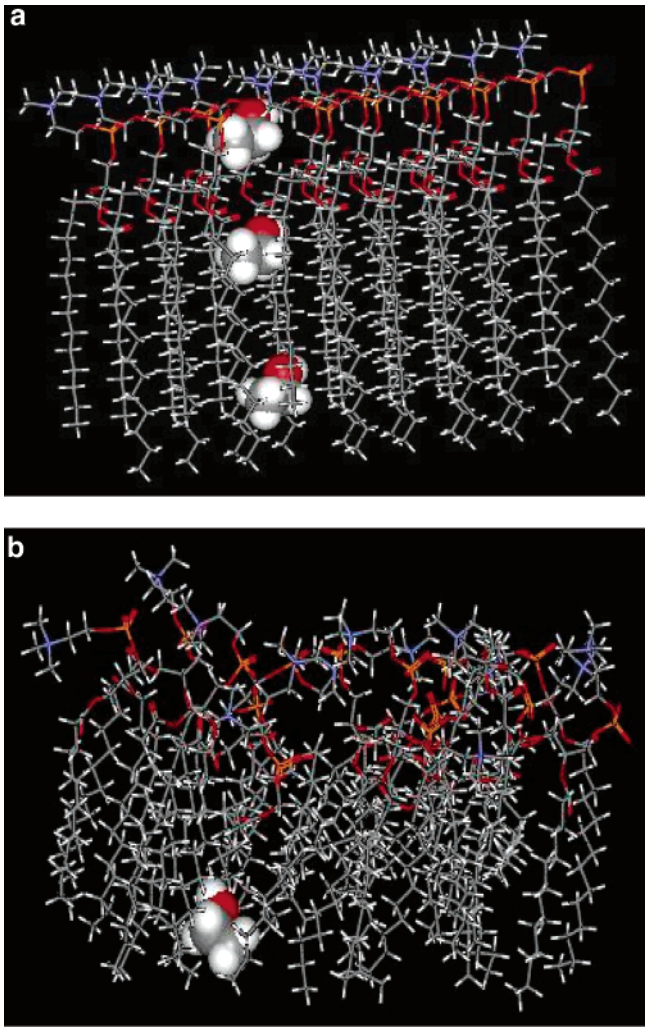

Figure 4. (a) A "side" view of an ethanol molecule inserted at three different positions in the DMPC model monolayer prior to the start of each of the three corresponding MDS used in the MI-QSAR modeling. (b) The lowest energy geometry of a DMPC−ethanol complex in the MDS.

The *intermolecular solute−membrane interaction descriptors* were extracted directly from the MDS trajectories. 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 4a) for each of the solutes. The intermolecular membrane−solute descriptors are given in Table 4.

F. Construction QSAR Models. QSAR models derived from relatively large data sets (>100 compounds) usually contain a combination of *major* and *minor* descriptors irrespective of the manner in which the data are fit and the fitting is optimized. Major descriptors are those that are likely to reflect the global mechanism of action associated with the data set. Minor descriptors are present in the QSAR to better provide an overall fit, and/or incorporate specific attributes of some compounds. These minor descriptors are usually not reflective of the global mechanism of action.

(20) Doherty, D. C. *Molsim Version 3.0 User's Guide*; Chicago, 1994.
(21) Berendsen, H. J. C.; Postman, J. P. M.; Gunsteren, W. F. V.; Nola, A. D.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys*. **1984**, *81*, 3684−3690.
(22) MSI. *Cerius^2 Molecular Simulations Users Guide Ver. 3.0*; Molecular Simulation Inc.: San Diego, 1997.

Table 3. Definitions of the Intramolecular Solute Descriptors

symbol	description of the intramolecular solute descriptors
Chi$Z$	Kier and Hall molecular connectivity index, $Z = 0, 1, 2, ...$
J_X	Balaban index: characterizes the shape of a molecule, which can take account of the covalent radii
CHI-V-3_P	Kier and Hall valence-modified connectivity index CHI-3_P means third-order CHI index with three paths (bonds) connected; V means that electron configuration of the atom (single or multiple bonds) is considered
Ecoh	cohesive packing energy of the solute with itself
FH2O	desolvation free energy for water
Jurs-FNSA-1	fractional charged partial surface areas: total charge weighted negative surface area divided by the total molecular solvent-accessible surface area
Jurs-PNSA-1	partial negative surface area: sum of the solvent-accessible surface areas of all negatively charged atoms
Jurs-PPSA-1	partial positive surface area: sum of the solvent-accessible surface areas of all positively charged atoms
Jurs-PPSA-3	atomic charge weighted positive surface area: sum of the product of solvent-accessible surface area X partial charge for all positively charged atoms
Jurs-RPCS	relative positive charge surface area: solvent-accessible surface area of most positive atom divided by descriptor
Jurs-RPCG	relative positive charge: charge of most positive atom divided by the total positive charge
Jurs-RNCG	relative negative charge: charge of most negative atom divided by the total negative charge
Jurs-FPSA-1	fractional charged partial surface areas: partial positive surface area divided by the total molecular solvent-accessible surface area
Jurs-FPSA-3	fractional charged partial surface areas: total charge weighted positive surface area divided by the total molecular solvent-accessible surface area
Jurs-RNCS	relative negative charge surface area: solvent-accessible surface area of most negative atom divided by relative negative charge
Jurs-DPSA-1	difference in charged partial surface areas: partial positive solvent-accessible surface area minus partial negative solvent-accessible surface area
Kappa6	Kier's shape indices; Kappa6 is the sixth-order index, compares the molecule graph with "minimal" and "maximal" graphs
log P	octanol/water partition coefficient calculated by Advanced Chemistry Development (ACD, www.acdlabs.com); log P is related to the hydrophobic character of the molecule
Area	molecular surface area: a 3D spatial descriptor that describes the van der Waals area of a molecule
Shadow-Zlength	length of molecule in the Z dimension

Table 4. The Explicit Intermolecular Membrane−Solute Interaction Descriptors Forming the Trial MI-QSAR Descriptor Pool

symbol	description of the membrane−solute descriptors
$\langle F(\text{total})\rangle$	average total free energy of interaction of the solute and membrane
$\langle E(\text{total})\rangle$	average total interaction energy of the solute and membrane
$E_{\text{inter}}(\text{total})$	interaction energy between the solute and the membrane at the total intermolecular system minimum potential energy
$E_{XY}(Z)$	Z = 1,4-nonbonded, general van der Waals, electrostatic, hydrogen-bonding, torsion, and combinations thereof energies at the total intermolecular system minimum potential energy. X and Y can be the solute, S, and/or membrane, M
$\Delta E_{XY}(Z)$	change in the Z = 1,4-nonbonded, general van der Waals, electrostatic, hydrogen-bonding, torsion, and combinations thereof energies due to the uptake of the solute to the total intermolecular system minimum potential energy. X and Y can be the solute, S, and/or membrane, M
$E_{TT}(Z)$	Z = 1,4-nonbonded, general van der Waals, electrostatic, hydrogen-bonding, torsion, and combinations thereof energies of the total (solute and membrane model) intermolecular minimum potential energy
$\Delta E_{TT}(Z)$	change in the Z = 1,4-nonbonded, general van der Waals, electrostatic, hydrogen-bonding, and combinations thereof of the total (solute and membrane model) intermolecular minimum potential energy
ΔS	change in entropy of the membrane due to the uptake of the solute
S	absolute entropy of the solute−membrane system
$\Delta\rho$	change in density of the model membrane due to the permeating solute
$\langle d\rangle$	average depth of the solute molecule from the membrane surface
D	diffusion coefficient of the solute in the membrane model

Table 5. The Five Test Set Compounds Used To Evaluate Each Best Model from Each Descriptor Model Class[a]

compound		residuals of fit: 3-term model	4-term model	5-term model	6-term model
		Linear Non-MI-QSAR			
2-pentanone	[−2.60]	7.27	1.20	0.04	−5.86
3-cresol	[−1.82]	0.11	2.80	2.42	−13.05
acetic acid	[−3.21]	−41.21	−18.85	−17.40	−166.70
cortisone	[−5.00]	−18.33	−7.81	−4.99	−18.18
γ-butyrolactone	[−4.00]	−16.05	−13.73	−13.04	−24.41
		Quadratic Non-MI-QSAR			
2-pentanone		0.89	0.30	−8.58	−4.27
3-cresol		−1.71	−1.99	−5.39	−3.06
acetic acid		41.59	−31.65	−9.22	7.24
cortisone		0.67	−5.77	1.40	−3.11
γ-butyrolactone		16.01	−10.54	−17.53	−3.80
		Linear Composite			
2-pentanone		0.01	−0.13	−0.24	−0.14
3-cresol		−0.04	−0.19	−0.46	0.15
acetic acid		0.01	−0.62	−0.16	−0.21
cortisone		0.22	0.02	−0.16	−0.41
γ-butyrolactone		0.00	−1.12	−1.22	−0.62
		Quadratic Composite			
2-pentanone		0.29	−0.11	−0.51	−0.77
3-cresol		0.30	−0.53	−0.07	−0.24
acetic acid		−0.59	−1.15	−0.50	−0.64
cortisone		0.62	0.36	−0.22	−0.25
γ-butyrolactone		−1.01	−1.45	−1.13	−1.43

[a] The observed log k_p values are given in brackets in the first compound listing, and the corresponding residuals of fit to the observed log k_p values are given for each model in columns 2−5.

Separating and assigning major and minor descriptors after QSAR model construction is difficult and ill-defined. Thus, it can be considered advantageous in the search for the major descriptors of a training set to define a representative subset of the parent data set with respect to both chemical diversity and distribution of the endpoint measures (dependent variables). The subset can be used as a representative training set of the parent data set and lead to a small model (in terms of number of independent variables) where the independent variables have a high chance of being major descriptors. The minor descriptors can be thought of as "noise" in the fitting, which is reduced by reducing the size of the system. That is the strategy employed here in reducing the size of the data set from 152 to 40 in the manner described above.

All QSAR models were built and optimized using multidimensional linear regression fitting and genetic function approximation, GFA, which is a multidimensional optimization method based on the genetic algorithm paradigm.[23,24] Both linear and quadratic representations of each of the descriptor values were included in the trial descriptor pool, and QSAR models were built as a function of the number of descriptor terms in a model. Statistical significance in the optimization of a QSAR model was judged jointly by the correlation coefficient of fit, r^2, and the leave-one-out cross validation correlation coefficient, q^2. The leave-one-out q^2 value is determined by systematically leaving out each compound of the training set, and then predicting its value based on a model derived from the remainder of the training set. The average of the corresponding r^2 values, based on the predictions from each of these leave-one-out experiments, is essentially the q^2 value for the training set. In addition, GFA uses the Friedman's lack of fit (LOF) measure to resist overfitting, which is a problem often encountered in constructing statistical models.[25]

A test set of five compounds of diverse structures that span the training set range in log k_p values was used to further explore the robustness and predictivity of each QSAR model. The five test compounds are listed in Table 5 along with the observed log k_p values. The residual log k_p values for the test set compounds, using each constructed QSAR model, are also reported in Table 5. The "observed log k_p" of two of the test set compounds, 2-pentanone and acetic acid, are actually computed values from previously reported QSAR

(23) Rogers, D.; Hopfinger, A. J. Applications of Genetic Function Approximation to Quantitative Structure−Activity Relationships and Quantitative Structure−Property Relationships. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854−866.

(24) Rogers, D. WOLF 6.2 GFA Program; Molecular Simulation Inc.: San Diego, 1994.

(25) Friedman, J. *Multivariate Adaptive Regression Splines*; Stanford University: Stanford, 1988.

models.[26] The inclusion of these two compounds in the test set has been done to explore the consistency between the models developed here to previously reported models.

G. Classes and Properties of the QSAR Models. Two classes of QSAR models were constructed on the basis of the set of descriptors employed in the trial descriptor pool:

(1) *Non-MI-QSAR models* derived from a descriptor pool that contains only the intramolecular solute descriptors listed in Table 3. These models are constructed in an equivalent manner to the very large majority of skin penetration QSAR models reported in the literature.

(2) *Combined (MI-QSAR) models* derived from a descriptor pool that contains both the intramolecular and intermolecular membrane−solute descriptors.

These two classes of QSAR models were sought with respect to both the number of descriptors (3- to 6-descriptor-term models) and linear and quadratic representation of the descriptor terms. Outlier refinement analyses were carried out for models that have reasonable statistical fits using all of the compounds of the training set.

Results

A. The QSAR Models. The optimized 3- to 6-term QSAR models, for the model-building constraints listed in Methods, are given below. The definitions of the descriptors used in the models are given in Tables 3 and 4.

(1) Non-MI-QSAR

(a) Linear Models

$$\log k_p = 496.76 - 599.33\ \text{Jurs-FPSA-1} + 0.52\ \text{Area} - 1.84\ \text{Jurs-PNSA-1} \quad (1)$$

$$(N = 40;\quad r^2 = 0.67;\quad q^2 = -0.56)$$

$$\log k_p = 31.34 - 1908.61\ \text{Jurs-FPSA-3} + 4.80\ \text{Jurs-PPSA-3} + 478.84\ \text{Jurs-FNSA-1} - 1.57\ \text{Jurs-PNSA-1} \quad (2)$$

$$(N = 40;\quad r^2 = 0.88;\quad q^2 = -0.31)$$

$$\log k_p = 35.27 - 1.56\ \text{Jurs-PNSA-1} - 1.72\ \text{CHI-V-3\_P} - 2037.39\ \text{Jurs-FPSA-3} + 5.19\ \text{Jurs-PPSA-3} + 474.97\ \text{Jurs-FNSA-1} \quad (3)$$

$$(N = 40;\quad r^2 = 0.78;\quad q^2 = -0.30)$$

$$\log k_p = 120.67 - 216.58\ \text{Jurs-RNCG} - 37.50\ J_X - 7.17\ \text{CHI-V-3\_P} + 4.37\ \text{Jurs-RPCS} + 109.70\ \text{Jurs-FNSA-1} + 4.02\ \text{Kappa6} \quad (4)$$

$$(N = 40;\quad r^2 = 0.80;\quad q^2 = -0.26)$$

(b) Quadratic Models

$$\log k_p = -4.85 - 0.10\ (\text{Jurs-RNCS} - 5.35)^2 - 687.92\ (\text{Jurs-RNCG} - 0.21)^2 + 1.34\ (\text{Jurs-RPCS} - 2.91)^2 \quad (5)$$

$$(N = 40;\quad r^2 = 0.96;\quad q^2 = -0.63)$$

$$\log k_p = -3.20 - 0.002\ (\text{Jurs-DPSA-1} - 277.20)^2 - 0.002\ (\text{Jurs-PNSA-1} - 40.09)^2 + 0.001\ (\text{Jurs-PPSA-1} - 352.39)^2 + 998.66\ (\text{Jurs-FNSA-1} - 0.11)^2 \quad (6)$$

$$(N = 40;\quad r^2 = 0.94;\quad q^2 = -0.18)$$

$$\log k_p = -0.49 + 0.69\ (\text{Jurs-RPCS} - 3.00)^2 + 0.62\ (\text{Jurs-RPCS} - 2.47)^2 - 702.03\ (\text{Jurs-RNCG} - 0.21)^2 - 24.91\ \text{Jurs-RPCG} + 0.12\ (\text{Jurs-RNCS} - 5.10)^2 \quad (7)$$

$$(N = 40;\quad r^2 = 0.97;\quad q^2 = -0.59)$$

$$\log k_p = -3.26 + 0.77\ (\text{Jurs-RPCS} - 3.47)^2 + 0.10\ (\text{Jurs-RNCS} - 3.20)^2 - 747.21\ (\text{Jurs-RNCG} - 0.23)^2 + 0.62\ (\text{Jurs-RPCS} - 2.32)^2 + 0.94\ \text{Shadow-Zlength} - 38.88\ \text{Jurs-RPCG} \quad (8)$$

$$(N = 40;\quad r^2 = 0.97;\quad q^2 = -0.59)$$

(2) Combined (MI-QSAR) Models

(a) Linear Models

$$\log k_p = -2.97 + 0.22 \log P - 0.14\ E_{SS}(\text{tor}) - 0.05\ E_{\text{inter}}(\text{vdW}) \quad (9)$$

$$(N = 39;\quad r^2 = 0.80;\quad q^2 = 0.77)$$

$$\log k_p = -2.43 - 0.10\ E_{SS}(\text{tor}) + 0.44 \log P - 0.04\ E_{SS}(\text{hb}) - 0.04\ \text{Ecoh} \quad (10)$$

$$(N = 40;\quad r^2 = 0.81;\quad q^2 = 0.76)$$

$$\log k_p = -1.01 - 0.004\ E_{SS}(\text{chg}) - 0.10\ E_{SS}(\text{tor}) + 0.37 \log P - 0.01\ E_{TT}(\text{tor}) - 0.03\ E_{SS}(\text{hb}) \quad (11)$$

$$(N = 38;\quad r^2 = 0.84;\quad q^2 = 0.77)$$

$$\log k_p = -2.57 - 0.05\ E_{SS}(\text{hb}) + 0.55 \log P - 0.02\ \text{Ecoh} - 0.003\ E_{SS}(\text{tot}) - 0.34\ \text{Chi10} - 0.00003\ E_{TT}(\text{tot}) \quad (12)$$

$$(N = 38;\quad r^2 = 0.85;\quad q^2 = 0.79)$$

(b) Quadratic Models

$$\log k_p = -7.20 + 0.04\ (\log P + 3.28)^2 + 0.01\ (\text{Ecoh} - 45.14)^2 + 0.002\ (E_{SS}(\text{tor}) - 34.45)^2 \quad (13)$$

$$(N = 40;\quad r^2 = 0.85;\quad q^2 = 0.81)$$

(26) Kirchner, L. A.; Moody, R. P.; Doyle, E.; Bose, R.; Jeffery, J.; Chu, I. The Prediction of Skin Permeability by Using Physicochemical Data. *ATLA* **1997**, *25*, 359−370.

$$\log k_p = -8.21 + 0.05\,(\log P + 3.28)^2 + 0.003\,(E_{SS}(\text{tor}) - 34.45)^2 + 0.001\,(\text{Ecoh} - 45.14)^2 + 0.02\,E_{TT}(\text{bend}) \quad (14)$$

$$(N = 40; \quad r^2 = 0.88; \quad q^2 = 0.82)$$

$$\log k_p = -8.38 + 0.002\,(E_{SS}(\text{tor}) - 34.45)^2 + 0.44 \log P + 0.0005(\text{FH2O} + 83.58)^2 + 0.00006\,(E_{TT}(\text{hb}) + 11.39)^2 + 2.64\,\text{Chi12}^2 \quad (15)$$

$$(N = 40; \quad r^2 = 0.95; \quad q^2 = 0.81)$$

$$\log k_p = -8.41 + 0.003\,(E_{SS}(\text{tor}) - 21.85)^2 + 0.45 \log P + 0.0005\,(\text{FH2O} + 83.58)^2 + 0.00006\,(E_{TT}(\text{hb}) + 11.39)^2 + 4.05\,\text{Chi12}^2 + 0.06\,(\text{Chi8} - 4.93)^2 \quad (16)$$

$$(N = 40; \quad r^2 = 0.91; \quad q^2 = 0.81)$$

B. Analysis and Comparison of the QSAR Models. Comparing the QSAR models constructed from the two classes of descriptors clearly indicates that models composed of a combination of both non-MI-QSAR and MI-QSAR descriptors yield, by far, the best models. All of the non-MI-QSAR models show negative values of q^2 and, consequently, are rather unstable models in terms of predictivity. Thus, the MI-QSAR descriptors are providing information about the skin penetration process that is not captured by the intramolecular (non-MI-QSAR) solute descriptors.

In the outlier refinement analysis, analogues of the training set were considered outliers when the residuals of fit exceed 2 standard deviations, SD, from the mean of the residual of fit of the whole training set. Using this criterion outlier compounds were found in the composite linear models. The outliers are scopolamine and benzaldehyde (for the 6-term model), estriol and benzaldehyde (for the 5-term model), and costicosterone (for the 3-term model). The quadratic composite models do not have any outliers, suggesting that slightly nonlinear dependencies exist between the descriptors and log k_p. Outlier refinement analyses were not carried out for non-MI-QSAR models since each of these models have negative values of q^2 and, therefore, lack predictive stability.

The observed and residual log k_p values for a test set of five compounds are given as part of Table 5. An inspection of the residual values of prediction for the test set suggests that the models from the combined descriptor sets are, indeed, most predictive as suggested by the q^2 values of the models. The two test set compounds, 2-pentanone and acetic acid, whose "observed" log k_p values were actually computed,[26] are predicted with fidelity about equal to that of the other three compounds. This suggests that there may be a general consistency among the skin penetration QSAR models, with the differences among the QSAR models mainly arising from the extent of structural diversity that can be handled by a model and, perhaps, the differences in predictive resolution among the QSAR models.

Potts and Guy[27] described a skin permeability prediction model based on molecular weight (MW) and the octanol−water partition coefficient (log P) using a skin permeability data set collected by Flynn.[2] This optimum skin permeation QSAR model is

$$\log k_p = 0.72 \log P - 0.0059\,\text{MW} - 2.80 \quad (17)$$

$$(N = 93; \quad r^2 = 0.67; \quad q^2 = 0.65)$$

This QSAR model indicates that percutaneous absorption is mediated by the hydrophobicity and the molecular weight of the penetrant. The Potts and Guy model is regarded as one of the better QSAR models for skin penetration prediction. However, the linear non-MI-QSAR descriptor models, eqs 1−4, constructed in this study are far different in descriptor terms from the Potts and Guy model, eq 17. Thus, in order to explore if the training set used in this work would provide a better QSAR model using the same descriptors of the Potts and Guy model, a "new" Potts and Guy model was constructed and optimized using the GFA. The resulting model is

$$\log k_p = 0.28 \log P - 0.007\,\text{MW} - 2.00 \quad (18)$$

$$(N = 40; \quad r^2 = 0.56; \quad q^2 = 0.51)$$

The use of only log P and MW as descriptors leads to a QSAR model having lower values of both r^2 and q^2 for the training set of this study as compared to the r^2 and q^2 of the Potts and Guy QSAR model, eq 17. In order to test the predictive power of eq 18, it was applied to the same test set of five compounds used to evaluate the other QSAR models developed in this study. The predicted skin penetration values for the test set compounds using eq 18 are given in Table 6. The results of applying the Potts and Guy QSAR model, eq 17, directly to the test set of five compounds is also given in Table 6. It is clear from the residuals of fit for the test set predictions that the linear composite QSAR models (Table 5) perform better than both of the Potts and Guy models. Of course, more descriptor terms are present in the composite QSAR models, but these additional descriptor terms are seemingly needed to be able to make good predictions.

In this context of descriptor terms, a comparison was done between the Potts and Guy model (eq 17) and the three-term linear composite QSAR model (eq 9),

$$\log k_p = 0.22 \log P - 0.14\,E_{SS}(\text{tor}) - 0.05\,E_{\text{inter}}(\text{vdW}) - 2.97$$

$$(N = 39; \quad r^2 = 0.80; \quad q^2 = 0.77)$$

Both of these QSAR models have similar regression constants (−2.80 and −2.97) and similar dependencies on log P (regression coefficients of 0.72 and 0.22). The molecular weight term of the Potts and Guy model is replaced

(27) Potts, R. O.; Guy, R. H. Predicting Skin Permeability. *Pharm. Res.* **1992**, *9*, 663−669.

Table 6. Residuals of Fit for the Test Set Predictions Using the Optimized Potts and Guy 2-Term (log *P* and MW) QSAR Model for the Training Set of This Study, Eq 18, and the Original Potts and Guy QSAR Model, Eq 17

test compound	log k_p prediction, eq 18	log k_p prediction, eq 17
2-pentanone	−0.22	0.13
3-cresol	0.27	−0.10
acetic acid	−0.67	0.24
cortisone	−1.21	−1.96
γ-butyrolactone	−1.23	−0.27

by two descriptor terms in the composite descriptor QSAR model. The second descriptor term of eq 9, $-0.14\ E_{SS}$(tor), indicates that skin penetration decreases as the torsion energy of a penetrant solute increases. Such an increase in torsion energy can arise as the penetrant solute becomes larger, that is, as molecular weight increases, and/or as the penetrant becomes more rigid in conformation. The other descriptor term in the composite descriptor QSAR model is $-0.05\ E_{inter}$-(vdW), which indicates that skin penetration of the penetrant decreases as the van der Waals penetrant−membrane interaction energy favorably increases. This term is dependent on molecular weight but can both increase and decrease with molecular weight. This term is, most likely, a compensating term to the log *P* term. Better van der Waals interactions, which dominate the origin of hydrophobicity, are possible when the solute is in the membrane than when in an aqueous medium. The E_{inter}(vdW) descriptor captures this thermodynamic feature.

Thus, the three-term combined descriptor QSAR can be thought to "dissect" the molecular weight term of the Potts and Guy model and replace it with two terms that incorporate more molecular detail. The higher r^2 value in the linear composite descriptor QSAR model, as compared to both eqs 17 and 18, indicates a better data fitting for the regression model. The higher q^2 value of the combined descriptor QSAR model indicates that it is more predictive and captures more structural information than does the Potts and Guy model.

Discussion

Three major conclusions can be made from the findings of this study which relate to both understanding the skin penetration process, and how best to model and correspondingly extract information from data generated in skin penetration studies. The first major conclusion is that the composite descriptor QSAR models in this study are far better than those constructed using only non-MI-QSAR trial descriptor sets. The composite descriptor QSAR models are very "robust" statistically. Moreover, on the basis of the high values of the q^2 and of the test set prediction, these models should have considerable predictive power skin penetration studies, which seems borne out for the test set predictions. Overall, the MI-QSAR descriptors do, in fact, win out over the other descriptors in building optimum QSAR models.

A corollary to this conclusion is that solute−membrane interactions, as represented by the MI-QSAR descriptors, provide information regarding the skin penetration process that cannot be realized from properties (descriptors) derived solely from the solute. Thus, QSAR models for skin penetration employing only solute-derived descriptors fail to capture significant features of the skin penetration process, and they can be expected to be limited both in predictive power and in the insight they provide regarding this biochemical process.

The second major conclusion is that partitioning between aqueous (polar) and lipid (nonpolar) media, which is best represented by log *P*, is the most significant factor in the skin penetration mechanism. log *P* appears in all composite QSAR models of this study and nearly every non-MI-QSAR model reported in the literature. Interestingly, however, log *P* is not found in any of the non-MI-QSAR models, eqs 1−4, in this study. Certainly the q^2 of eqs 1−4 are poor, but the criterion for building these models was r^2 which are significant for these models. It would seem that the partitioning process between aqueous and lipid media can be represented by descriptors other than log *P* and/or such descriptors can augment log *P* in describing such partitioning. The presence and importance of log *P* are seemingly dependent upon the number, structure, and size of compounds in a training set, and on the set of descriptors available for building a model. For medium size data sets and/or for descriptors like those used in this study, log *P* can be absent because specific hydrogen bonding, hydrophobic and aqueous solvation, and related descriptors are better representations of compound partitioning into a membrane than log *P*. However, as a training set becomes very large and/or very structurally diverse, log *P* becomes the best single descriptor to capture the average behavior of membrane partitioning. But capturing the average behavior means that the precision of prediction of the corresponding model will drop off.

Last, there are four major factors which seemingly govern skin penetration of a compound and reveal features of the mechanism of action:

(1) *Polar/nonpolar partitioning* of the compound as generally represented by log *P*. This descriptor is present in all of the composite QSAR models and indicates that skin penetration increases with increasing log *P*.

(2) *Polar, electrostatic, hydrogen-bonding, and/or aqueous solvation interactions* involving the compound as represented by FH2O, E_{SS}(hb) and, E_{SS}(chg). The presence of FH2O only as quadratic terms in the models suggests that there is an optimum aqueous solvation free energy (aqueous solubility) for a compound that, in turn, optimizes skin penetration. It is this factor, relative aqueous solubility, which is being captured by this set of descriptors.

(3) *Size and shape* of the compound as represented by E_{SS}(vdW), Chi8, Chi10, and Chi12. Presumably this factor accounts for the steric component to diffusing through the skin media. In this particular study, the use of size- and shape-related descriptors, which substitute for a molecular size only descriptor found in many non-MI-QSAR models,

obviates the need to divide the training set into structural/chemical classes of compounds for which individual QSAR models are built for each class in other studies of skin penetration.[2,6]

(4) *Conformational flexibility* of the compound as represented by E_{SS}(tor) and E_{TT}(tor). This pseudo entropic factor can also be related to the *size* and *rigidity* of a penetrant solute.

Ongoing work is focusing upon extending and refining the present QSAR models by using clustering and discriminate analyses on the parent data set of compounds. This approach is designed to build a manifold set of skin penetration QSAR models which, in composite, will provide accurate estimates of skin penetration for a wide range of chemistries as well as to obtain more information about the overall skin penetration mechanism.

Acknowledgment. The UIC authors are pleased to acknowledge financial support from Avon. Resources of the Laboratory of Molecular Modeling and Design at UIC and of The Chem21 Group, Inc., were used in performing this work.

MP049924+