Journal of Computer-Aided Molecular Design, **12:** 451–470, 1998.
KLUWER/ESCOM

Heuristic lipophilicity potential for computer-aided rational drug design: Optimizations of screening functions and parameters

Qishi Du & Paul G. Mezey*
Department of Chemistry, University of Saskatchewan, 110 Science Place, Saskatoon, Saskatchewan, Canada S7N 5C9

Received 20 November 1997; Accepted 20 March 1998

Key words: combinatorial chemistry, drug design, electrostatic potential, lipophilicity potential, molecular modeling, molecular shape

Summary

In this research we test and compare three possible atom-based screening functions used in the heuristic molecular lipophilicity potential (HMLP). Screening function 1 is a power distance-dependent function, $b_i / \| \mathbf{R}_i - \mathbf{r} \|^{\gamma}$, screening function 2 is an exponential distance-dependent function, $b_i \exp(-\| \mathbf{R}_i - \mathbf{r} \| / d_0)$, and screening function 3 is a weighted distance-dependent function, $\mathrm{sign}(b_i) \exp[-\xi(\| \mathbf{R}_i - \mathbf{r} \| / | b_i |)]$. For every screening function, the parameters (γ, d_0, and ξ) are optimized using 41 common organic molecules of 4 types of compounds: aliphatic alcohols, aliphatic carboxylic acids, aliphatic amines, and aliphatic alkanes. The results of calculations show that screening function 3 cannot give chemically reasonable results, however, both the power screening function and the exponential screening function give chemically satisfactory results. There are two notable differences between screening functions 1 and 2. First, the exponential screening function has larger values in the short distance than the power screening function, therefore more influence from the nearest neighbors is involved using screening function 2 than screening function 1. Second, the power screening function has larger values in the long distance than the exponential screening function, therefore screening function 1 is effected by atoms at long distance more than screening function 2. For screening function 1, the suitable range of parameter γ is $1.0 < \gamma < 3.0$, $\gamma = 2.3$ is recommended, and $\gamma = 2.0$ is the nearest integral value. For screening function 2, the suitable range of parameter d_0 is $1.5 < d_0 < 3.0$, and $d_0 = 2.0$ is recommended. HMLP developed in this research provides a potential tool for computer-aided three-dimensional drug design.

Introduction

In our first papers [1] we suggested a model of heuristic molecular lipophilicity potential (HMLP), and presented some examples and simple applications of this model. HMLP is a fully structural, three-dimensional, unified lipophilicity and hydrophilicity potential requiring no empirical lipophilicity indices. However, because of the complexity of this task, the formulas and equations used in this model are not all derived from first principles, therefore we say it is a heuristic model.

Heuristic molecular lipophilicity potential takes the form of the following equation,

$$L_{\mathbf{r}\in\mathbf{S}_\alpha}(\mathbf{e}) \;=\; V(\mathbf{r}) \sum_{i\neq\alpha} M_i(\mathbf{r}; \mathbf{R}_i, b_i) \qquad (1)$$

where $V(\mathbf{r})$ is molecular electrostatic potential (MEP) at the point $\mathbf{r}$, and $\mathbf{r}$ is on the 'surface' S_α of atom α. In summation, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ is an atom-based screening function of atom *i* at position $\mathbf{r}$ and summation is over all constituent atoms, except atom α, on which point $\mathbf{r}$ sits. In the atom-based screening function, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$, $\mathbf{R}_i$ is the nuclear position of atom *i*, and b_i is the atomic surface-MEP descriptor of atom *i* [2],

*To whom correspondence should be addressed.

$$b_i = \sum_{k \in S_i} V(r_k) \Delta S_k \quad (2)$$

where Δs_k's are the area elements on the surface of atom *i*. In HMLP the positive values of L(**r**) represent lipophilicity, and the negative values of L(**r**) are for hydrophilicity. This convention has a theoretical background. The logarithm of partition coefficient between water and octanol, $\log P_{ow}$, has been used as an overall measure of molecular lipophilicity. Partition coefficient is connected with the transfer free energy, and assumed to be the sum of the contributions from all constituent fragments or atoms [3–14],

$$\log P_{ow} = \log \frac{C_{organic}}{C_{water}} = \frac{-\Delta G^{\circ}_{tr}}{2.303RT} = \sum_i n_i f_i. \quad (3)$$

It is obvious that, if a compound is a 'water-lover', there is a higher concentration in the water phase than in organic phase. Otherwise, if a compound is an 'oil-lover', there is a higher concentration in the organic phase than in water phase. This means that the hydrophilic group has a negative contribution ($f_i < 0$) to $\log P_{ow}$. On the other hand, the lipophilic group has a positive contribution ($f_i > 0$) to $\log P_{ow}$. The sign of HMLP can be understood in another way: lipophilicity is the repulsive interaction between solute and water molecules and the potential energy is positive, while hydrophilicity is the attractive interaction between solute and water molecules and the potential energy is negative. These conventions are consistent with the Chinese traditions of *YIN* and *YANG*. Lipophilicity, meaning dry, corresponds to *YANG*, and is positive, while hydrophilicity, meaning wet, corresponds to *YIN*, and is negative.

HMLP is a unified lipophilicity and hydrophilicity potential. Here hydrophilicity includes the interactions of dipole moments, hydrogen bonds, and charged atoms of solute molecules with water molecules. The basic ideas of heuristic molecular lipophilicity potential defined by Equation (1) is that the interactions between organic molecules and water molecules at point **r** on the molecular surface are decided not only by the atom to which point **r** belongs, but by a large neighborhood. Atom-based screening function $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ conveys this idea, which represents the influence on point **r** from atom *i*. If the influence on point **r** from all surrounding atoms, $\Sigma_{i \neq \alpha} M_i(\mathbf{r}; \mathbf{R}_i, b_i)$, has the same sign as electrostatic potential V(**r**), the interaction between solute and water molecules is repulsive, the lipophilicity potential L(**r**) is positive, and at point **r,** the molecule is lipophilic. Otherwise, if the influence on point **r** from all surrounding atoms, $\Sigma_{i \neq \alpha} M_i(\mathbf{r}; \mathbf{R}_i, b_i)$, has the opposite sign to electrostatic potential V(**r**), the interaction between solute and water molecules is attractive, the lipophilicity potential L(**r**) is negative, and at point **r**, the molecule is hydrophilic. The model of HMLP does not include the structure of the hydration shell explicitly, but considers the effects of water molecules implicitly through the screening function.

As shown by Equation (1), HMLP, L(**r**), is a modified electrostatic potential, V(**r**), combining the averaged influences from the molecular environments. HMLP includes the information of electrostatic potential, therefore it is also a unified electrostatic and lipophilic potential. Molecular electrostatic potential (MEP) V(**r**) is computed at each point **r**, by integrating over the *ab initio* electron density in the usual fashion [16] (in atomic units),

$$V(r) = \sum_A \frac{Z_A}{| R_A - r |} - \int \frac{\rho(r')dr'}{| r' - r |}, \quad (4)$$

where $\{Z_A\}$ and $\{\mathbf{R}_A\}$ are the sets of nuclear charges and position vectors, respectively. The physical meaning of molecular electrostatic potential (MEP) can be illustrated simply as follows: V(**r**) *is the interaction energy between a molecule and a unit point probe charge at position* **r**. MEP has a rigorous theoretical definition and a solid experimental background. MEP can be tested experimentally by electron diffraction. Politzer and his research group have successfully used MEP in the studies of hydrogen-bondings and various molecular interactions [17, 18], as well as in the studies of nucleophilic and electrophilic attacks [19]. In recent years MEP has been more and more often used to illustrate the chemical and physical properties of macromolecules.

Imagining that a water molecule consists of four tetrahedrally distributed point charges with two hydrogen bonding donors and two acceptors, it is easy to understand why electrostatic potential is the best physical property for describing molecular lipophilicity. However, it has not yet been successfully used in describing lipophilicity by MEP, as in the studies of hydrogen-bonds and nucleophilic and electrophilic attacks, though a number of promising efforts have been made by some authors [20–25]. The reason for this

is that molecular lipophilicity is an entropy-dominated phenomenon, dealing with the interactions of a huge number of water molecules, unlike hydrogen-bonding, in which case the maximum and minimum of MEP on a few atoms are sufficient for the qualitative description. However, the description of lipophilicity needs a large neighborhood on the molecular surface. Lipophilic effect is mainly dominated by negative entropic changes of water molecules. Usually, people think that interactions of dipole moments and multipole moments, charged atoms, hydrogen-bonding, dispersion, and polarization, are electrostatic interactions, and they assume that lipophilic interaction has a non-electrostatic origin and is dominated entropically [15]. In a macroscopic point of view, this is true, however, in a microscopic point of view, lipophilic interaction is of electrostatic origin, too. Molecular lipophilicity is a property that may also manifest itself in the electrostatic interaction between a solute and a collection of water molecules.

The atom-based screening function, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$, plays an important role in HMPL. In this study, we focus on the improvement of the screening function in HMLP. We will compare three possible screening functions, and optimize the parameters based on the calculation results of 41 common organic compounds of several families.

Screening functions

It is common knowledge that a lipophilic surface region is a nonpolar area, while a hydrophilic region is the polar area on the molecular surface. This means that the designation of lipophilic or hydrophilic surface area is decided by the molecular environment. In the study of molecular lipophilicity, the effects from surrounding atoms have drawn the attention of many authors. Israelachvili [15] believes that the hydrophilic and hydrophobic interactions are interdependent, unlike various electrostatic and dispersion interactions, which are independent interactions. He presents an example where the hydrophobic energy per CH_2 group of an alkane chain is greatly changed when a hydrophilic head-group, such as OH, is attached to the end of the chain [15]. Ghose and Crippen [3] have developed a data set of empirical atomic lipophilicity indices for carbon, oxygen, nitrogen, sulfur, and halogens. They have classified these elements into 110 'atomic types'. The factors taken into consideration for the classification are: (1) the electron distribution around the atom, (2) the approachability of the solvent, (3) the nature of the nearest atoms attached to the atom concerned, and (4) the influences from the next nearest neighbors. Carbon atoms may have as many as 4 directly connected neighbors, therefore the situations are extremely complex. For carbon alone, there are as many as 51 'atomic types', each assigned different values. Hansch and Leo [26], Rekker [27], Eisenberg and McLachlan [28] have conducted similar studies.

Náray-Szabó has studied this problem using a different method [20]. He emphasizes that besides a geometric fit, electrostatic attraction and the matching of nonpolar regions are also necessary to ensure optimal binding of a ligand to a biological acceptor. The electrostatic attraction accounts for ion-pair interactions and hydrogen bonding, while the matching of the nonpolar region represents the hydrophobic interactions. He uses MEP to describe the electrostatic interactions, and uses MEF (molecular electrostatic field) [21, 22, 29, 30] to describe the hydrophobic interactions. MEF is the gradient of MEP,

$$E(\mathbf{r}) = -\nabla V(\mathbf{r}). \tag{5}$$

MEF's are vectors. Actually, MEF represents the changes in the magnitude and direction of MEP in a molecular space. The results of calculation of MEF indicate that in the surroundings of the polar part of a molecular surface, MEF's show big changes in both direction and magnitude. On the other hand, in the surroundings the nonpolar lipophilic part, there are no direction changes, and only small changes in magnitude. Platt and Silverman [31] developed a new technique for the expansion of multipolar decomposition of electrostatic potential (MDEP). Multipolar expansion of MEP provides an immediate characterization of the MEP distributions on the molecular space. Both MEF and MDEP are modifications of MEP, and the motivations for introducing MEF and MDEP are to attempt to describe the distributions of MEP on the molecular space, which is essential for the description of molecular lipophilicity.

In our heuristic molecular lipophilicity potential, the atom-based screening function $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ represents the influence at point $\mathbf{r}$ from atom *i*, and plays a key role in the HMLP. It is a function of position $\mathbf{r}$. Nuclear coordinates $\mathbf{R}_i$ and atomic surface-MEP descriptors b_i are parameters in the screening function. There may be more parameters needed for the description of some special considerations, however, at the first stage, we hope to keep the mathematical form of the screening function as simple as possible. A

complete theoretical derivation of the screening function from first principles, explicitly including a huge number of water molecules, is not easy. On the other hand, in some cases a heuristic lipophilicity potential is enough for the purpose of molecular modeling.

There is no experimental evidence to show in what mathematical form the screening function should be. An experiment for the study of hydrophilic force law conducted by Israelachvili and Pashley [32] shows that the hydrophobic force decays exponentially in the range 0–10 nm. Marcelja et al. [33] propose an equation for the calculation of hydrophobic force, which is an exponentially distance-dependent function. Hydrophobic force law describes the interaction between two organic molecules in an aqueous solution. However, screening function is not designed for the hydrophobic force law between two solute molecules, but for the influence of other atoms on the interaction ability with water of one atom in the solute molecule. Experimental results can provide some hints, however, they cannot give direct help for the selection of screening functions.

Empirical molecular lipophilicity potential (EMLP) takes a number of forms [5, 34–37]. Audry et al. [35, 36] present an equation for the 3-dimensional representation of EMLP,

$$\mathrm{MLP} = \sum_{\mathrm{i}} \frac{\mathrm{f_i}}{1 + \mathrm{d_i}} \tag{6}$$

where $\mathrm{d_i}$ is the distance (in Å) between a given point outside the molecular surface and the atom *i*, and $\mathrm{f_i}$'s are empirical atomic lipophilicity indices. Fauchère et al. [38] propose another form for EMLP. They have defined an exponential distance dependence for a fragmental contribution: EMLP $\sim \exp(-\mathrm{d})$. Heiden et al. [5] extend the selection of EMLP functions widely. They point out that there is no physical reason for the use of one or another distance dependent functions in the empirical lipophilicity potential. They suggest a general form for EMLP [5],

$$\mathrm{MLP} = \frac{\sum_{\mathrm{i}} \mathrm{g(d_i)f_i}}{\mathrm{N}} = \frac{\sum_{\mathrm{i}} \mathrm{g(d_i)f_i}}{\sum_{\mathrm{i}} \mathrm{g(d_i)}}, \tag{7}$$

where $\mathrm{g(d_i)}$ is a distance dependent function, and $\mathrm{N} = \Sigma_{\mathrm{i}}\ \mathrm{g(d_i)}$ is the normalization factor. Heiden al. [5] have discussed the conditions fulfilled by function $\mathrm{g(d_i)}$. They use the Fermi function as $\mathrm{g(d_i)}$,

$$\mathrm{g(d_i)} = \frac{1}{\exp\left[\mathrm{a(d_i - d_{cut-off}}\right] + 1}, \tag{8}$$

where $\mathrm{d_{cut-off}}$ is an assigned cut-off value called the proximity distance.

All of the above research results of EMLP are valuable for the selection of a screening function for HMLP. Unlike empirical MLP, in which it is difficult to tell the physical meaning of function $\mathrm{g(d_i)}$ because of the ambiguity of empirical parameters $\mathrm{f_i}$'s, the screening function $\mathrm{M_i}(\mathbf{r}; \mathbf{R_i}, \mathrm{b_i})$ of heuristic MLP may have a certain physical meaning. Molecular electrostatic potential $\mathrm{V}(\mathbf{r})$ is the measure of the interaction ability of a solute molecule with a unit test charge at point $\mathbf{r}$. Suppose a water molecule is a dipole consisting of two opposite point charges ($\mathrm{q_w^+}$ and $\mathrm{q_w^-}$), then $\mathrm{V}(\mathbf{r}_A)$ and $\mathrm{V}(\mathbf{r}_\mathrm{B})$ are the measures of the interactive abilities of atom A and atom B in a solute molecule with water molecules at point $\mathbf{r}_\mathrm{A}$ and $\mathbf{r}_\mathrm{B}$, respectively. If the MEP at atom A is positive, $\mathrm{V^+}(\mathbf{r}_\mathrm{A}) > 0$, then the interaction between atom A and negative charge $\mathrm{q_w^-}$ of a water molecule is attractive,

$$\mathrm{E_{A,W}} = \mathrm{V^+}(\mathbf{r}_\mathrm{A})\mathrm{q_w^-} < 0. \tag{9}$$

If the MEP at the neighboring atom B of atom A is negative, $\mathrm{V^-}(\mathbf{r}_\mathrm{B}) < 0$, there is an attractive interaction between atom B and positive charge $\mathrm{q_w^+}$ of a water molecule,

$$\mathrm{E_{B,W}} = \mathrm{V^-}(\mathbf{r}_\mathrm{B})\mathrm{q_w^+} < 0. \tag{10}$$

The interaction between two water molecules interacting with atom A and B is attractive, too, $\mathrm{q_w^+}\ \mathrm{q_w^-}\ /\mathbf{r}_\mathrm{AB} < 0$. Therefore, both atom A and B are hydrophilic. If the MEP at atom B is positive, $\mathrm{V^+}(\mathbf{r}_\mathrm{B})$, there is an attractive interaction between atom B and the negative charge $\mathrm{q_w^-}$ of a water molecule,

$$\mathrm{E_{B,W}} = \mathrm{V^+}(\mathbf{r}_\mathrm{B})\mathrm{q_w^-}. \tag{11}$$

However, the interaction between two water molecules binding on atoms A and B is repulsive, $\mathrm{q_w^-}\ \mathrm{q_w^-}\ /\mathbf{r}_\mathrm{AB} > 0$ and $\mathrm{q_w^+}\ \mathrm{q_w^+}\ /\mathbf{r}_\mathrm{AB} > 0$, and is much stronger than the attractive interactions with atoms A and B. In this case, both atom A and B are lipophilic.

The influence of atom B on the atom A is determined by MEP $\mathrm{V}(\mathbf{r}_\mathrm{B})$ of atom B and the distance between atom A and B. If we use the ST2 water model [39], which contains two hydrogen bonding donors and two hydrogen bonding acceptors located along four tetrahedral arms radiating out from the center of the O atom, the interaction between two water molecules is assumed to involve 16 Coulombic terms representing the interactions between four point charges on one molecule with four on the other.

There are an additional 16 Coulombic terms between two atoms (q_A and q_B) and 8 point charges of two water molecules. Because of the complexity of this task, we do not initially pursue the physical meaning of the screening function, but instead think of the screening function as a mathematical function, and select its mathematical form and optimize its parameters based on the calculation results. From the chemical and physical facts, an atom-based screening function should satisfy the following four conditions:

1) If at the point **r** a molecule is lipophilic, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ has the same sign as $V(\mathbf{r})$; otherwise, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ has the opposite sign as $V(\mathbf{r})$,
2) If the absolute value of atomic surface-MEP descriptor b_i is higher, the absolute value of $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ is higher too; otherwise, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ is smaller,
3) $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ decays with the distance $\| \mathbf{R}_i - \mathbf{r} \|$;
4) $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ is a dimensionless function.

Condition 1 ensures that EMLP is chemically reasonable. Condition 2 conveys the idea that the atomic surface-MEP descriptor b_i plays an important role in the screening function. Condition 3 is based on physical fact. Condition 4 makes HMLP have the same unit as MEP. The above four conditions are assumptions for the screening function, yet there may be other conditions. For example, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ may decay with the distance in an oscillatory fashion, and may contain geometric and topological parameters which represent the tendency of an atom to interfere with the hydrogen bonding network of water molecules. The reasonableness of conditions should be examined by calculation results based on the chemical and physical facts. The atom-based screening function $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ can take a number of forms. Three possible screening functions, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$, are suggested as follows:

$$M_i(\mathbf{r}; \mathbf{R}_i, b_i) = \frac{r_0^\gamma}{b_0} \frac{b_i}{\| \mathbf{R}_i - \mathbf{r} \|^\gamma} = \varsigma \frac{b_i}{\| \mathbf{R}_i - \mathbf{r} \|^\gamma}, \quad (12)$$

$$M_i(\mathbf{r}; \mathbf{R}_i, b_i) = \frac{b_i}{b_0} e^{-\frac{\|\mathbf{R}_i - \mathbf{r}\|}{d_0}}, \quad (13)$$

$$M_i(\mathbf{r}; \mathbf{R}_i, b_i) = \frac{b_i}{| b_i |} e^{-\frac{b_0}{\lambda_0} \frac{\|\mathbf{R}_i - \mathbf{r}\|}{|b_i|}} = \mathrm{sign}(b_i) e^{-\zeta \frac{\|\mathbf{R}_i - \mathbf{r}\|}{|b_i|}}. \quad (14)$$

In the above three functions (r_0, b_0, γ), (b_0, d_0) and (b_0, λ_0) are parameters. The unit of b_0 is the same as b_i (energy · area) and r_0 , d_0, and λ_0 have a units of length. Therefore, $M_i(\mathbf{r}; \mathbf{R}_i, b_i)$ is a dimensionless function in all of the above three equations. In the screening function 1, Equation (12), $\zeta = (r_0)^\gamma / b_0$ is a simple scaling factor. In our calculations, we take $\zeta = 1$. Exponent γ is the parameter that decides how strong the influence is and how rapidly the influence decays with distance. It will be optimized in this study. In screening function 2, Equation (13), b_0 is a simple scaling factor, and is assigned value $b_0 = 1$ in this study. Parameter d_0 in Equation (13) plays the same role as γ in Equation (12), and will be optimized later. In screening function 3, Equation (14), $\xi = b_0/\lambda_0$ makes the exponent a dimensionless quantity, and affects the behavior of the screening function, like γ and d_0 in Equations (12) and (13). It will be optimized, too. The positions of atomic MEP-surface descriptor b_i in the three functions are different. In Equations (12) and (13), b_i is a factor of distance dependent functions, $1/ | \mathbf{R}_i - \mathbf{r} |^\gamma$ and $\exp(- | \mathbf{R}_i - \mathbf{r} | / d_0)$. In Equation (14), the sign of b_i is a factor, however, the value of b_i is put in the exponent, like a weighting function. In this research, we will test and compare all three screening functions and optimize parameters in the three functions based on the chemical and physical facts and experimental partition coefficient data $\log P_{ow}$.

Optimizations of screening functions and parameters

In this study, the *ab initio* quantum chemical program package Gaussian 92 is used to calculate electron density $\rho(\mathbf{r})$ and MEP's on the grid of molecular surfaces at the 6–31G* level. Molecular geometries are optimized using Gaussian 92 at the level STO-3G. Fused sphere van der Waals surfaces are used, and atomic radii are optimized based on MEP criteria [40]. Molecular surfaces are generated by program MS [41–43] using point-density 25 points/Å^2. Figure 1 shows the optimizations of parameter γ in the screening function 1, Equation (12), using (a) ethanol, (b) propionic acid, (c) ethylamine, and (d) propane. In Figure 1, atomic lipophilicity indices l_a's, molecular lipophilic indices L_M's, and molecular hydrophobic indices H_M's are

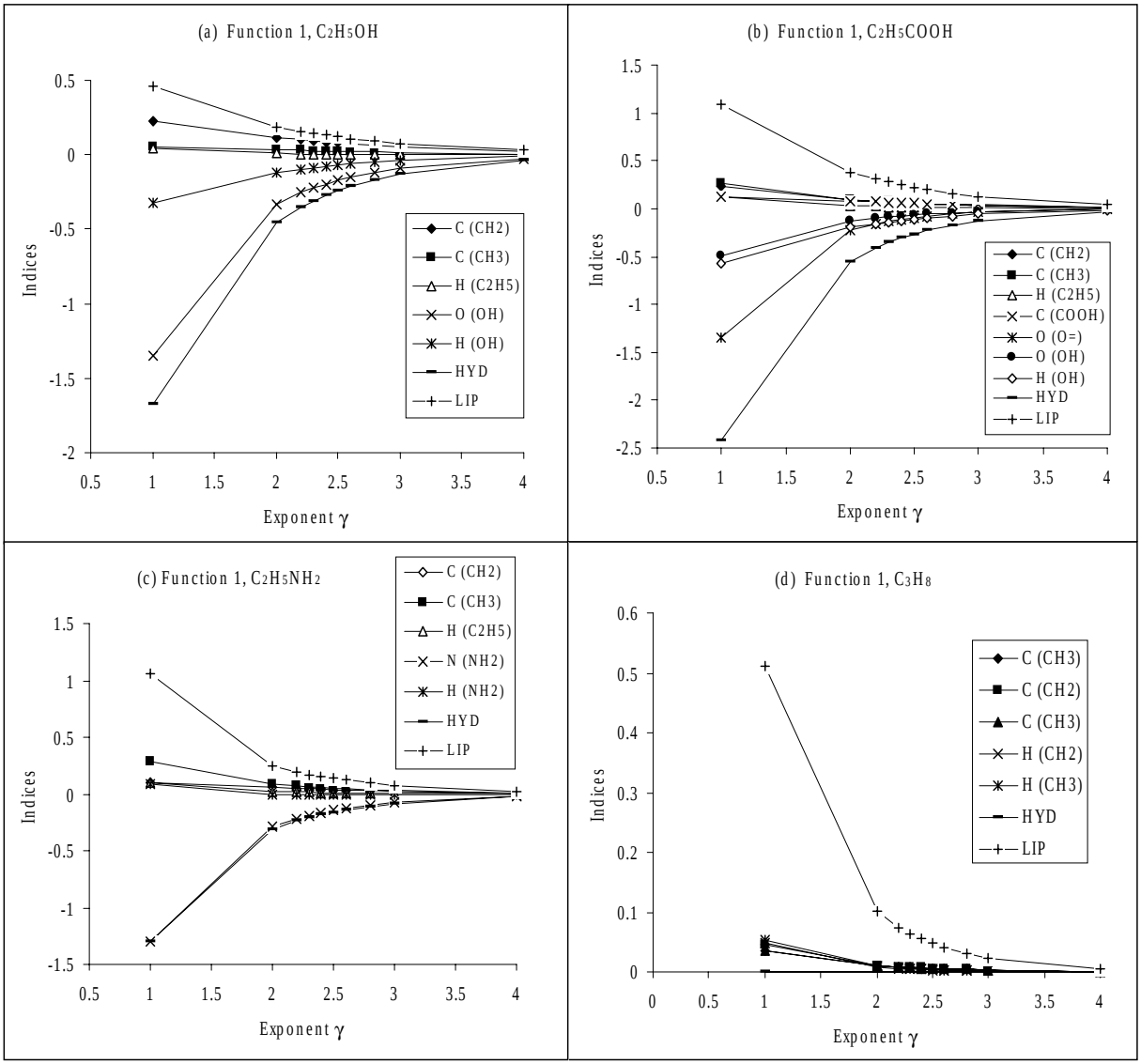

Figure 1. Optimization of γ parameter in screening function 1, Equation (12), using (a) ethanol, (b) propionic acid, (c) ethylamine, and (d) propane. Molecular lipophilic indices, L_M, molecular hydrophilic indices, H_M, and atomic lipophilicity indices, l_a, of functional groups, carbons, and several of the hydrogen atoms are shown.

shown for each molecule as defined by Equations (15), (16), and (17),

$$l_a = \sum_{i \in S_a} L(\mathbf{r}_i)\Delta s_i, \tag{15}$$

where the summation is over all the exposed area S_a of atom *a*. If $l_a > 0$, then atom *a* is lipophilic, whereas, if $l_a < 0$, then atom *a* is hydrophilic. Molecular lipophilic index (L_M) and hydrophilic index (H_M) is the sum of the corresponding values for all lipophilic atoms and hydrophilic atoms, respectively,

$$L_M = \sum_{\substack{a \\ (l_a>0)}} l_a, \tag{16}$$

$$H_M = \sum_{\substack{a \\ (l_a>0)}} l_a. \tag{17}$$

Figure 1 shows the optimizations of parameter γ in the screening function 1, using (a) ethanol, (b) propionic acid, (c) ethylamine, and (d) propane. Various

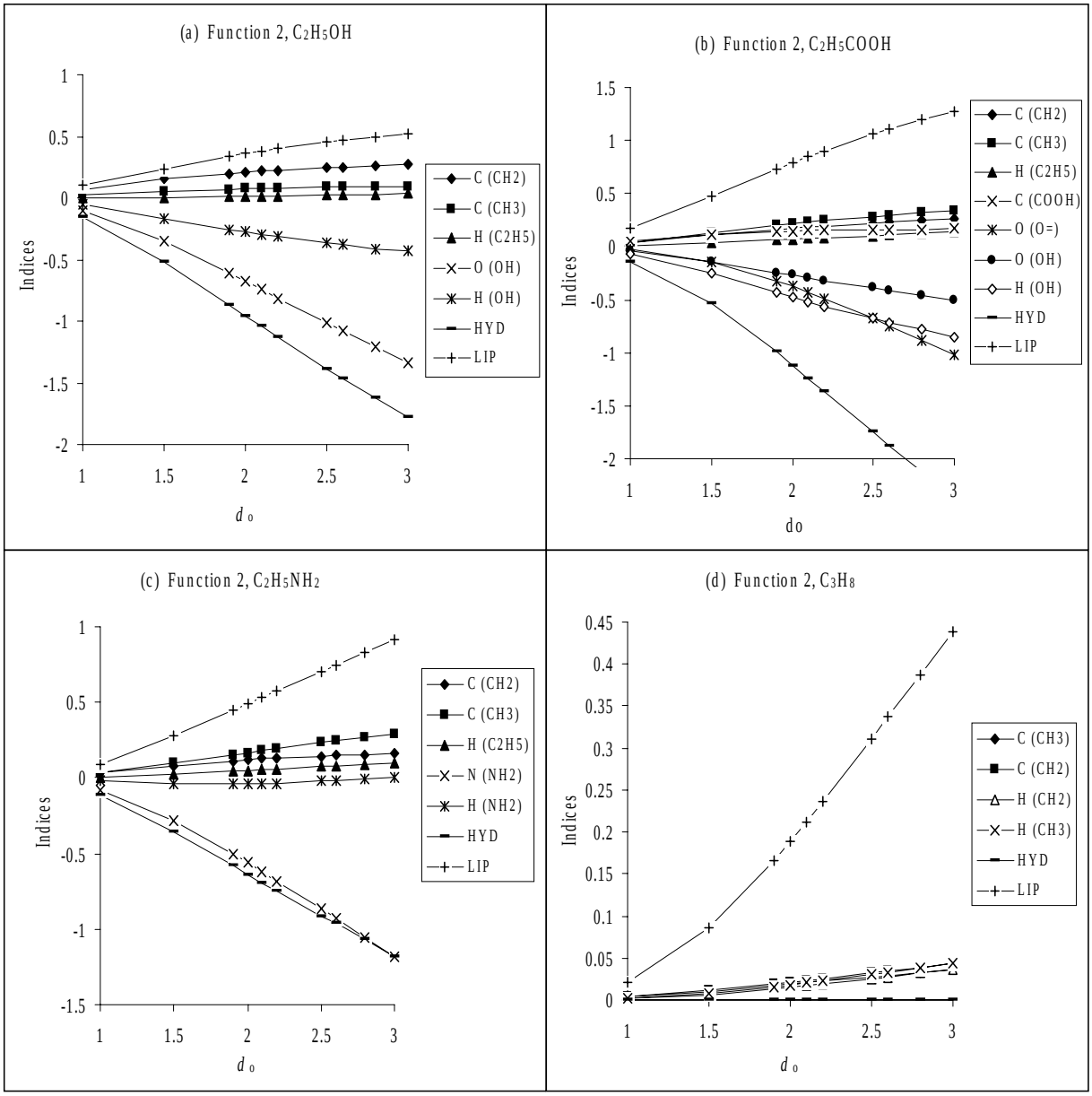

Figure 2. Optimizations of parameters d_o in screening function 2, Equations (4–9). (a) ethanol, (b) propionic acid, (c) ethylamine, and (d) propane. Molecular lipophilic indices, L_M, molecular hydrophilic indices, H_M, and atomic lipophilicity indices, l_a, of functional groups, carbons, and several hydrogen atoms are shown.

indices of ethanol are shown in Figure 1 (a). Atomic lipophilicity indices l_O and l_H in hydroxyl group are negative. This means that O and H in the hydroxyl group are hydrophilic atoms. All carbons and hydrogens in the hydrocarbon chain are lipophilic, having positive lipophilicity indices. All values of γ give qualitatively reasonable results based on the chemical facts. The absolute values of all indices decrease with increasing the value of γ, however, in the range $\gamma = 2.0 - 2.5$, decreases in the indices are getting smaller. In Figure 1 (b), propionic acid shows behavior very similar to that of ethanol. A detailed examination shows that there is an order of magnitude change in the atomic lipophilicity indices ($l_{=O}$, l_O and l_H) in the carboxyl group with an increase of γ. When $\gamma \leq 2.0$, the order is $l_{=O} < l_H < l_O$; then, when $\gamma > 2.0$, the order

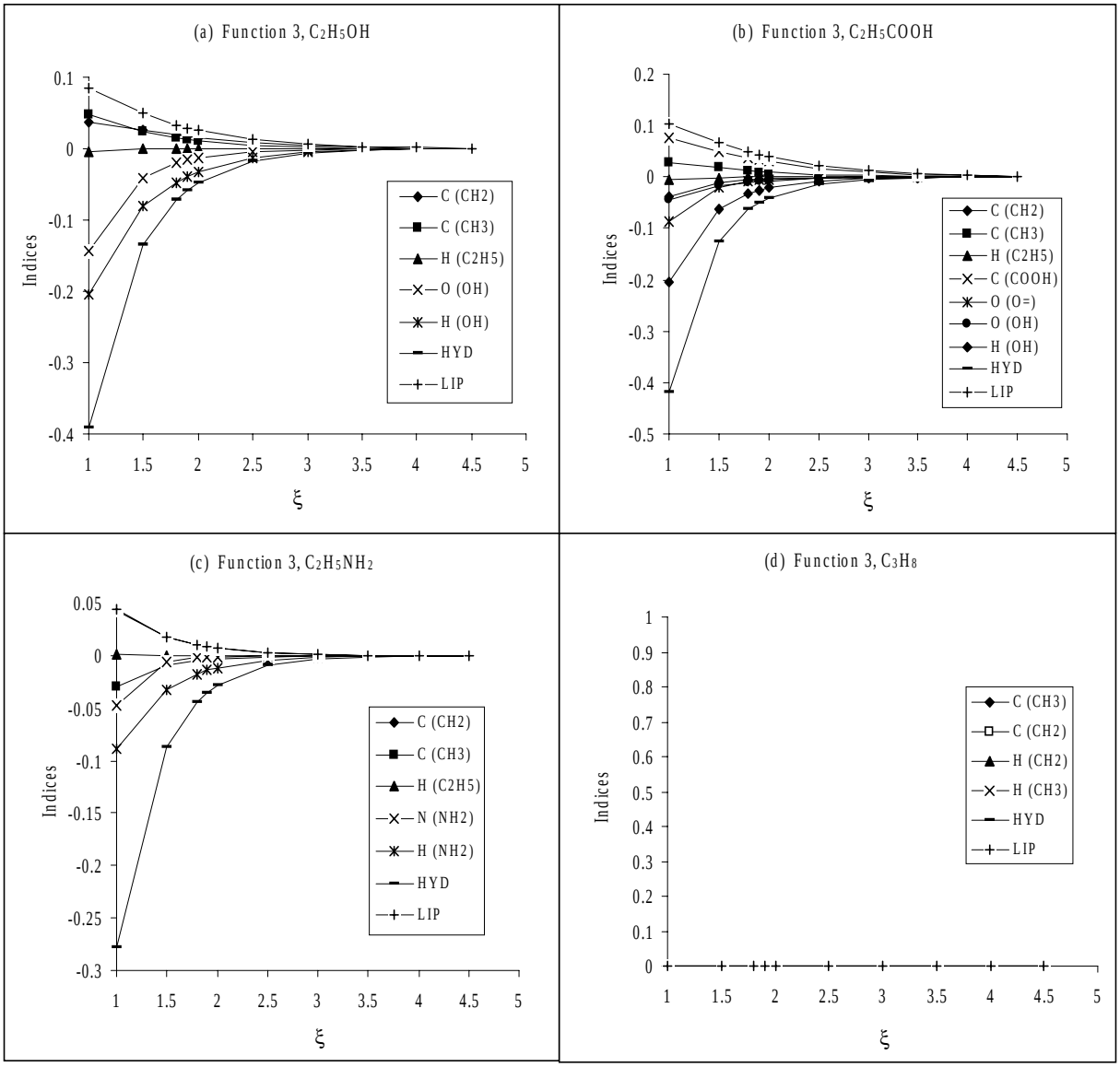

Figure 3. Optimizations of ξ parameter in screening function 3, Equation (14), (a) ethanol, (b) propionic acid, (c) ethylamine, and (d) propane. Molecular lipophilic indices, L_M, molecular hydrophilic indices, H_M, and atomic lipophilicity indices, l_a, of functional groups, carbons, and part of hydrogens are shown.

becomes $l_H < l_{=O} < l_O$; finally, when γ reaches the value $\gamma = 4.0$, the order is $l_H < l_O < l_{=O}$. The details of these observations can be found in Figure 4 (a). Ethylamine has something special, as shown in Figure 1 (c). Two hydrogens in the amino group -NH_2 have positive lipophilicity indices, l_H, in the range $\gamma \leq 1.0$. This is unreasonable from a chemical point of view. However, when $\gamma \geq 2.0$, the indices, l_H, of the two hydrogens turn to negative, and at the value $\gamma = 2.3$, the l_H's reach their minimum ($l_H = -0.0113$). The details of these observations can be found in Figure 4 (b). For propane in Figure 1 (d), there are no negative atomic lipophilicity indices, and the molecular hydrophilic index H_M is 0. This means that all atoms (carbons and hydrogens) are lipophilic. This is reasonable based on the chemical facts.

In Figure 2 (a) the absolute values of various indices of ethanol increase almost linearly with parameter d_0 of screening function 2. As in Figure 1 (a), O and H of hydroxyl group are hydrophilic, having

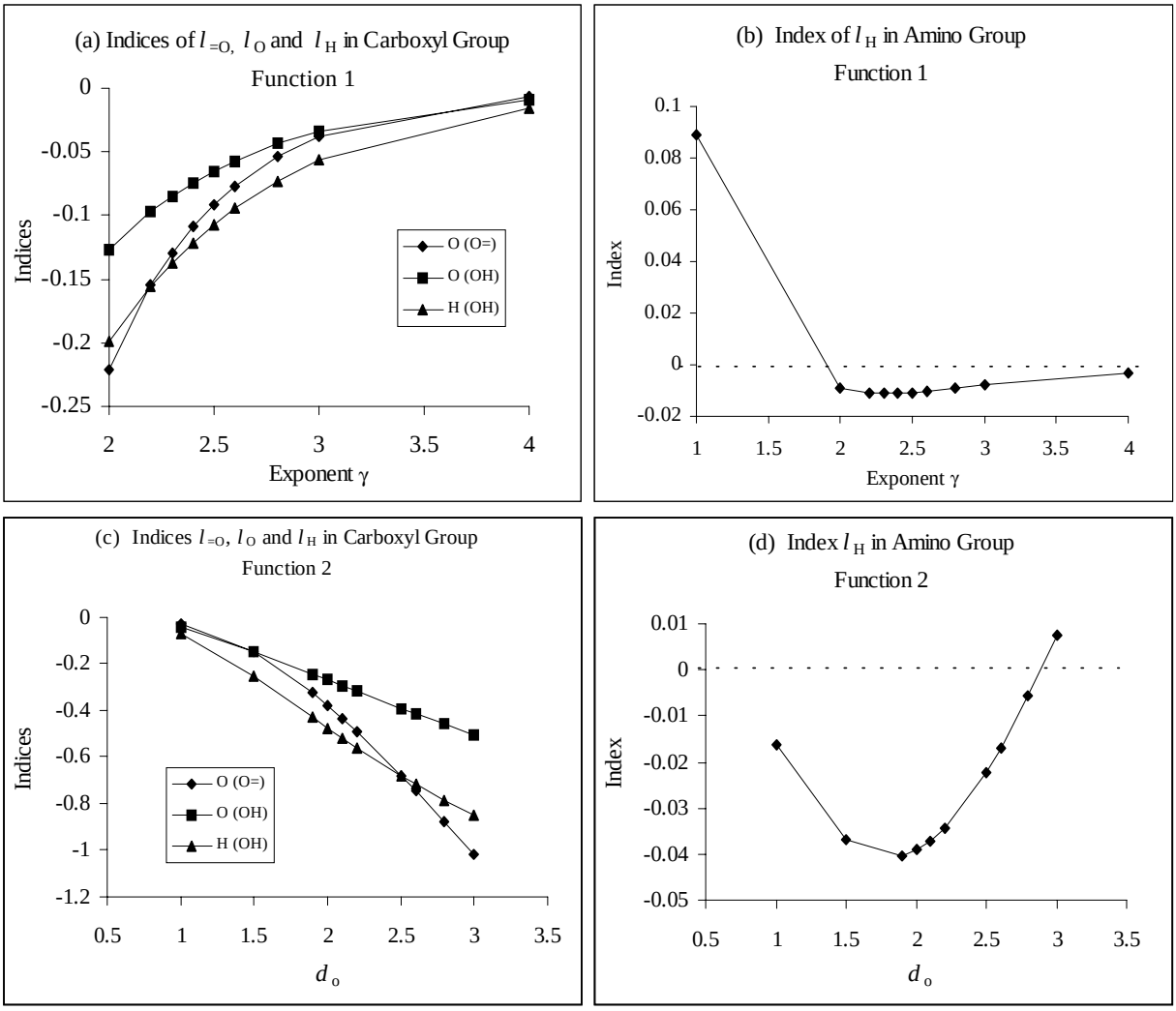

Figure 4. Changes in the orders of indices of $l_{=O}$, l_O, and l_H in the carboxyl group, COOH, and the minimum of l_H in amino group, NH_2, using screening functions 1 and 2.

negative atomic lipophilicity indices l_O and l_H. All carbons and hydrogens in the hydrocarbon chain are lipophilic, having positive lipophilicity indices. The behaviors of propionic acid, see Figure 2 (b), are very similar to that of ethanol. Once again, an order of magnitude change is found for the atomic lipophilicity indices of the carboxyl group. When $d_o \leq 1.0$, the order is $l_H < l_O < l_{=O}$; then, for $d_o > 1.5$, the order becomes $l_H < l_{=O} < l_O$; finally, when d_o reaches the value $d_o > 2.6$, the order is $l_{=O} < l_H < l_O$. Details can be found in Figure 4(c). As in Figure 1 (c) of ethylamine, the atomic lipophilicity indices l_H's of two hydrogens in the amino group NH_2 have a minimum ($l_H = -0.0411$) at $d_o = 1.9$. The details can be found in Figure 4 (d). For propane in Figure 2 (d), there are no negative atomic lipophilicity indices, and the molecular hydrophilic index H_M is 0, as in Figure 1 (d).

At first glance, the general tendencies of various indices in Figure 3 are much similar to those in Figure 1; however, more careful examinations shows that there are some differences between the two figures. The absolute values of various indices in Figure 3 are much smaller than those in Figure 1. For hydrocarbon propane, Figure 3 (d), almost all indices are zero. In Figure 3 (a) of ethanol, atomic lipophilicity indices of l_O and l_H of the hydroxyl group are negative, and atomic lipophilicity indices of two carbons in the hy-

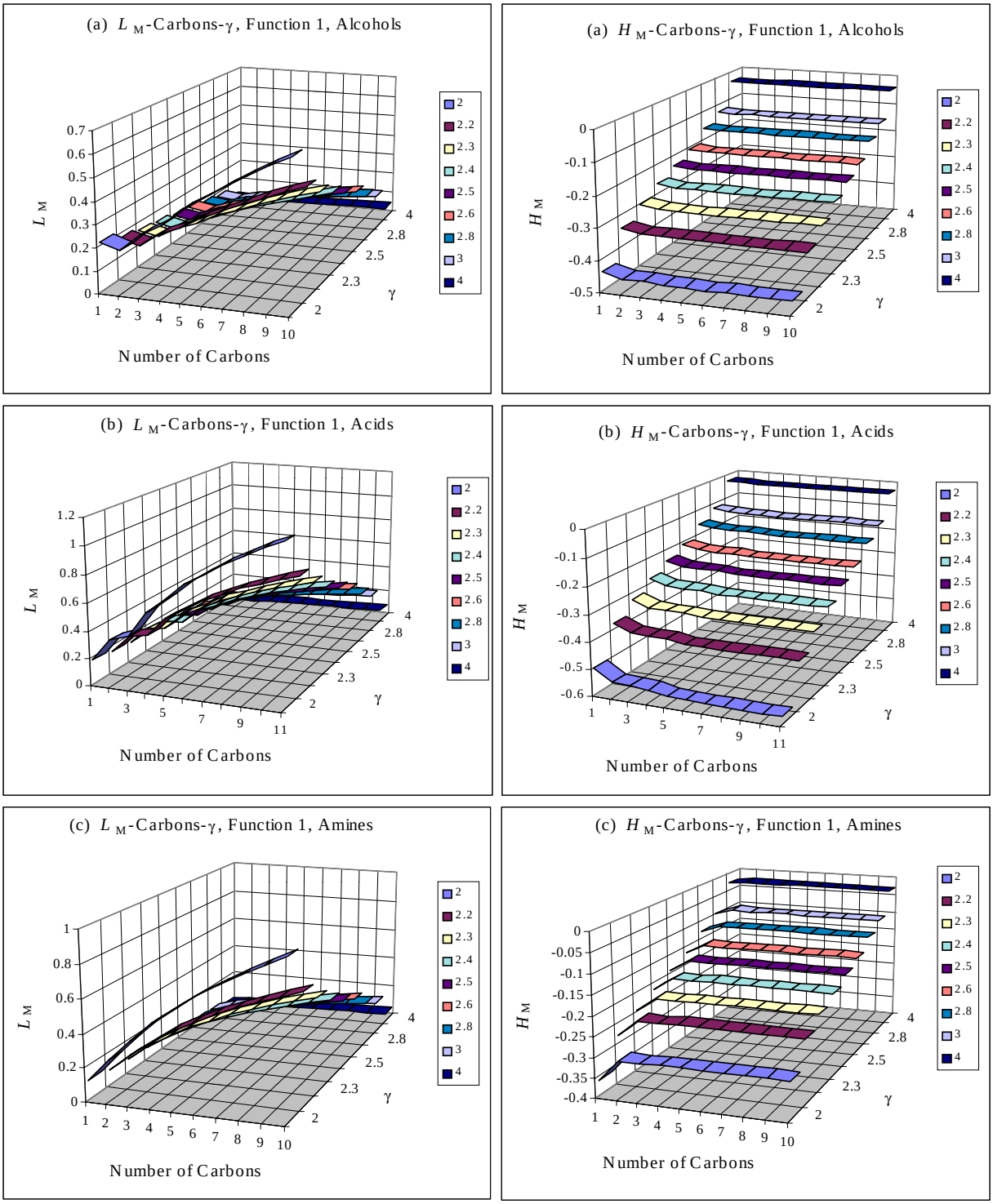

Figure 5. Optimization of parameter γ in the screening function 1, Equation (12), using (a) aliphatic alcohols, (b) aliphatic carboxylic acids, and (c) aliphatic amines. Molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, are shown as functions of number of carbon atoms and parameter γ.

drocarbon chain have positive values, as in Figure 1 (a). However, the hydrogen indices in the hydrocarbon chain have very small negative values. In Figure 3 (b), in the hydrocarbon chain of propionic acid, the hydrogen and the carbon atoms, connected directly with the carboxyl group, have very small negative atomic lipophilicity indices. In Figure 3 (c) of ethylamine, a negative atomic lipophilicity index is found for the carbon atom in the methyl of the hydrocarbon chain. All these results are unreasonable based on the chemical facts.

In the following part, we show the optimizations of three screening functions and parameters, using 4 series of compounds: aliphatic alcohols, aliphatic carboxylic acids, aliphatic amines, and linear hydrocarbons, which contain carbon atoms from 1 to 10. We will show the calculated results of molecular lipophilic indices, L_M, and hydrophilic indices, H_M, as functions of the number of carbon atoms and the parameters γ, d_o, and ξ in 3 types of screening functions.

Figure 5 tells us that the molecular lipophilic indices, L_M, increase with the number of carbon atoms and decrease with the increasing magnitude of exponent γ in screening function 1 for all three types of compounds, respectively. The increases and decreases are approximately linear, however, more careful examination shows that the increases in L_M are not exactly linear, but become smaller with the increasing number of carbons. Molecular hydrophilic indices, H_M, remain basically constant with the increasing number of carbon atoms, and increase with the increasing magnitude of exponent γ.

Figure 6 shows that the molecular lipophilic indices, L_M, increase with the number of carbon atoms for all three types of compounds, and the increases are approximately linear, as in Figure 5. However, L_M's increase with the parameter d_0 of screening function 2, unlike γ of screening function 1, where L_M's were observed to decrease. Careful examination shows that the increases of L_M become smaller as the number of carbons increases, as in Figure 5. Molecular hydrophilic indices, H_M, remain basically constant as the number of carbon atoms increase, and decrease with the increasing magnitude of d_0.

Figure 7 is completely different from Figure 5 and 6. Molecular lipophilic indices, L_M, increase slightly with the number of carbon atoms for all three types of compounds. However, when ξ is smaller (e.g., $\xi = 1.0$), the increases are greater. For the first three carbons, there is a big fluctuation in the increases. The L_M's decrease with increasing ξ. The molecular hydrophilic indices H_M's are basically constant with the increasing number of carbon atoms, however, there is a fluctuation in the first three carbons. The H_M's increase with an increase in the parameter ξ.

Figure 8 shows the molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, as functions of the number of carbon atoms for aliphatic alcohols, aliphatic carboxylic acids, and aliphatic amines, using screening function 1. In Figure 8, parameter γ in Equation (12) takes several values: (a) $\gamma = 1.0$, (b) $\gamma = 2.0$, (c) $\gamma = 2.5$, and (d) $\gamma = 3.0$. From Figure 8, we find that when $\gamma = 1.0$, the order of the L_M's of the three families of compounds is amines>acids>alcohols, and the order of the H_M's of the three compounds is acids<alcohols<amines. However, when γ becomes larger, e.g., $\gamma = 2.0$, 2.5, 3.0, the order of L_M's of the three compounds changes to acids>amines>alcohols. The order of H_M's remains the same, however, when $\gamma = 3.0$, values of H_M's for the alcohols and acids almost overlap each other, ref. Figure 8 (d). For all 4 values of γ, the molecular hydrophilic indices, H_M, remain basically constant with the increasing number of carbon atoms. The molecular lipophilic indices, L_M, increase with the number of carbon atoms. However, the increases are not exactly linear; if γ is smaller then the linearity is better.

Molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, are shown in Figure 9 as functions of the number of carbon atoms for aliphatic alcohols, aliphatic carboxylic acids, and aliphatic amines using screening function 2. In Figure 9, parameter d_o in Equation (13) takes several values: (a) d_0=1.0, (b) $d_0 = 1.5$, (c) $d_0 = 2.0$, and (d) $d_0 = 3.0$. From Figure 9, we find that when $d_0 = 1.0$, the order of L_M for the three compounds is acids>amines>alcohols, and the order of H_M for the three compounds is alcohols<acids<amines. However, when d_0 becomes lager, e.g., $d_0 = 1.5$, 2.0, and 3.0, one order is changed: the order of H_M of the three compounds becomes acids<alcohols<amines while the order for L_M remains the same. For the 4 different values of d_0, the molecular hydrophilic indices, H_M, remain basically constant with an increase in the number of carbon atoms. Molecular lipophilic indices, L_M, increase with the number of carbon atoms. However, the increase is not exactly linear; the d_0 is larger and the linearity is better. Comparing Figure 8 and Figure 9, we find that the orders of L_M and H_M of the three types of compounds are the same when $2.0 \leq \gamma \leq 3.0$ in function 1 and $1.5 \leq d_0 \leq 3.0$ in function 2.

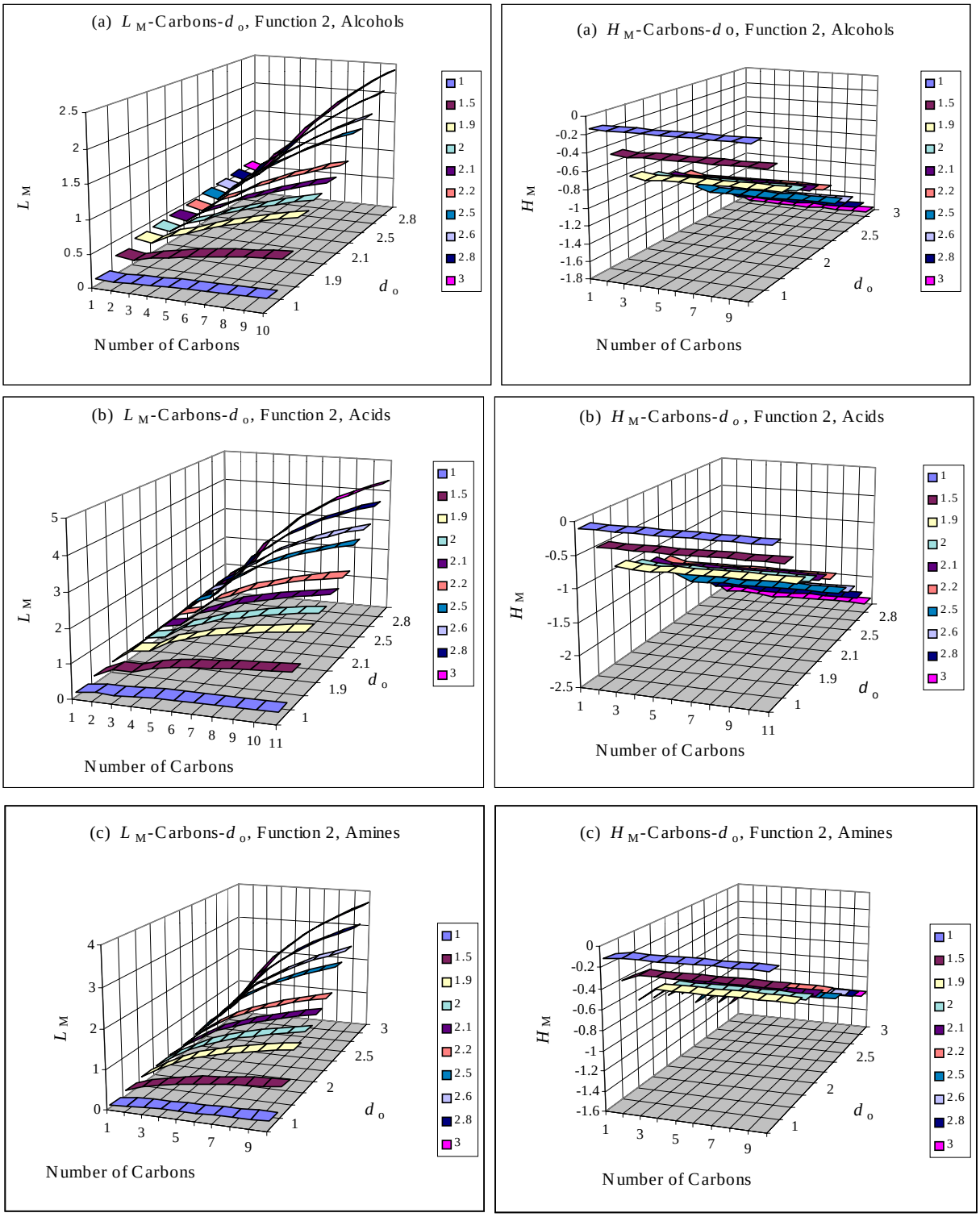

Figure 6. Optimization of parameters d_0 in the screening function 2, Equation (13), using (a) aliphatic alcohols, (b) aliphatic carboxylic acids, and (c) aliphatic amines. Molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, are shown as functions of the number of carbon atoms and parameter d_0.

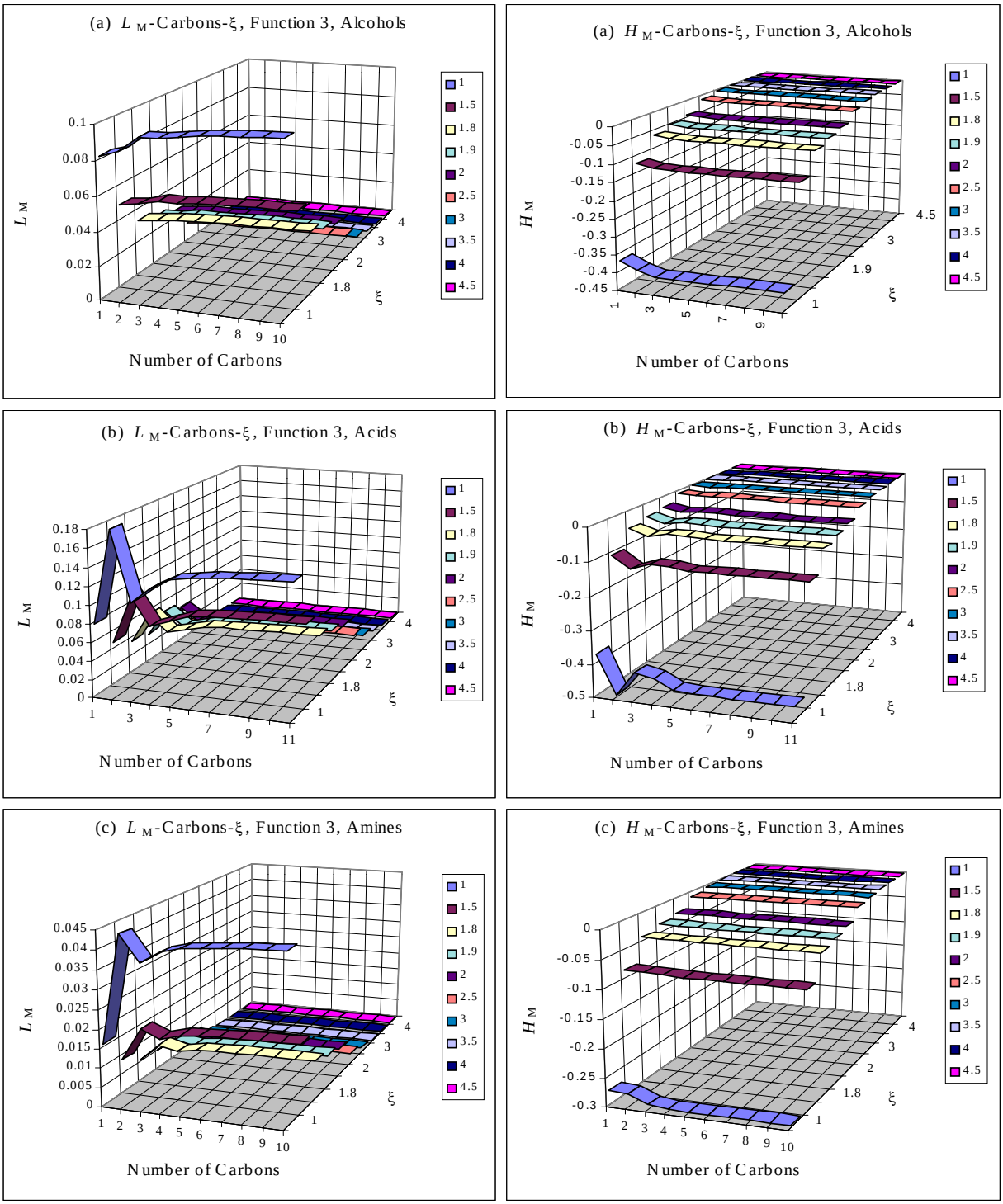

Figure 7. Optimization of parameters ξ in the screening function 3, Equation (14), using (a) aliphatic alcohols, (b) aliphatic carboxylic acids, and (c) aliphatic amines. Molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, are shown as functions of the number of carbon atoms and parameter ξ.

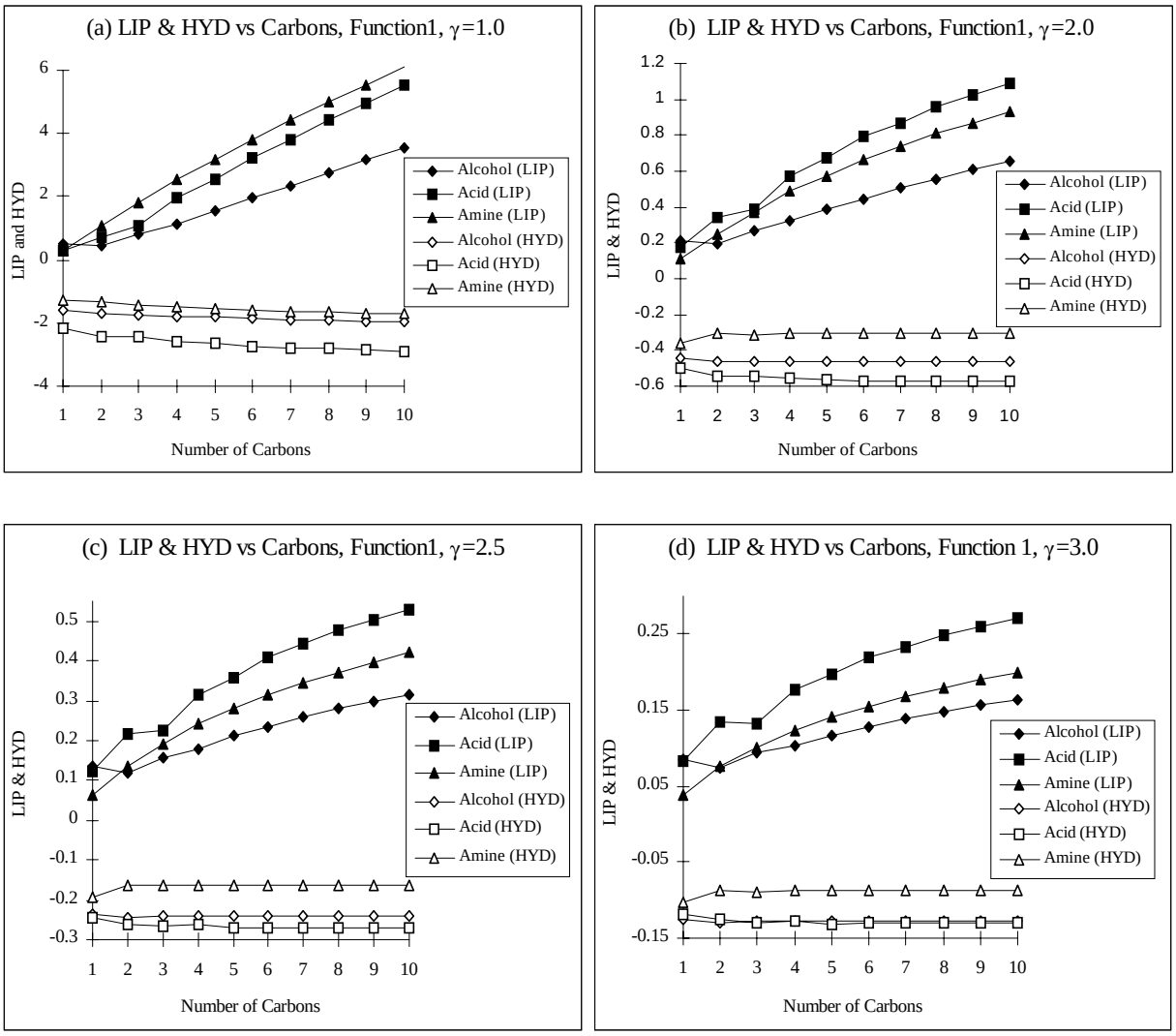

Figure 8. Molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, are shown as functions of the number of carbons for aliphatic alcohols, aliphatic carboxylic acids, and aliphatic amines. Parameter γ in the screening function 1 takes several values: (a) $\gamma = 1.0$, (b) $\gamma = 2.0$, (c) $\gamma = 2.5$, and (d) $\gamma = 3.0$.

Figure 10 shows the experimental $\log P_{ow}$ data [44] as a function of the number of carbon atoms, which are nearly linear for the three types of compounds. However, unlike Figure 8 and 9, we cannot find any difference in the slopes for the three series of compounds from Figure 10. Three lines show similar dependence, and the alcohol and acid lines almost overlap. Generally speaking, the $\log P_{ow}$ values for the amines are smaller than acids and alcohols, and the $\log P_{ow}$ of the acids and alcohols are almost the same based on Figure 10. If one assumes that the relationship between $\log P_{ow}$ and indices L_M and H_M is linear,

$$\log P_{ow} = C_0 + C_1 L_M + C_2 H_M \tag{18}$$

then a fair correlation coefficient and standard deviation are obtained for 23 molecules of the three types of compounds, $r = 0.833$, $\sigma = 0.698$, using screening function 1 and $\gamma = 2.0$. Similar results are obtained using screening function 2 and $d_0 = 2.3$. We are not satisfied with these results. The results of correlation calculations of $\log P_{ow}$ for each family of the three types of compounds are much better than for miscellaneous compounds (for aliphatic alcohols $r = 0.997$ and $\sigma = 0.152$, for aliphatic carboxylic acids $r = 0.978$ and $\sigma = 0.305$, and for aliphatic amines r

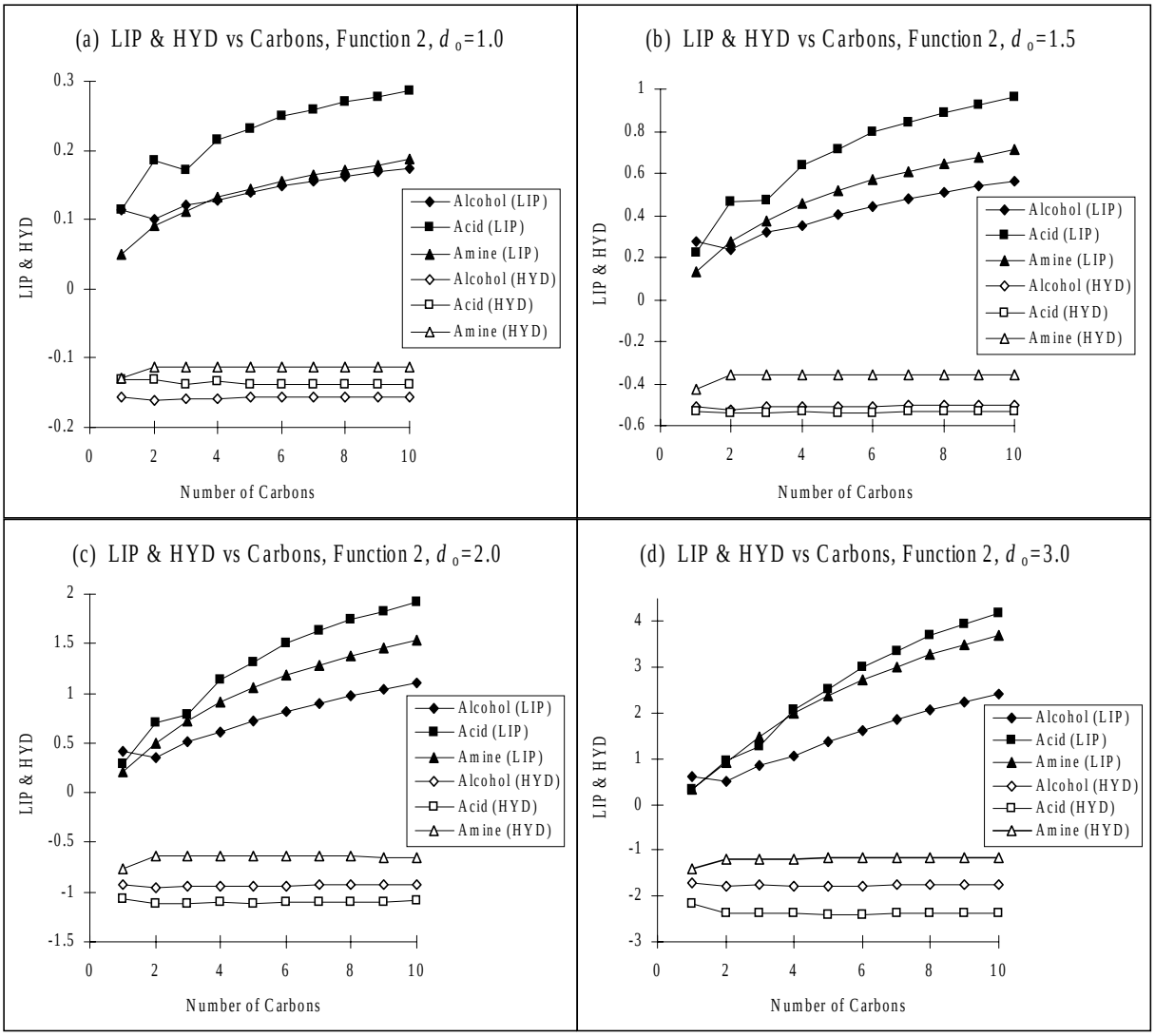

Figure 9. Molecular lipophilic indices, L_M, and molecular hydrophilic indices, H_M, are shown as functions of the number of carbons for aliphatic alcohols, aliphatic carboxylic acids, and aliphatic amines. Parameter d_0 in screening function 2 takes several values: (a) $d_0 = 1.0$, (b) $d_0 = 1.5$, (c) $d_0 = 2.0$, and (d) $d_0 = 3.0$.

$= 0.980$ and $\sigma = 0.282$). In the next section, we will discuss this observation.

Discussion and conclusions

Figure 11 shows the curves for the two types of distance-dependent functions used in screening functions 1 and 2. One is a power function used in screening function 1, $1/\parallel \mathbf{R}_i - \mathbf{r} \parallel^{\gamma}$, the other is an exponential function used in screening function 2, $\exp(-\parallel \mathbf{R}_i - \mathbf{r} \parallel / d_0)$. Actually, the weight distance-dependent function in screening function 3, $\exp[-\xi(\parallel \mathbf{R}_i - \mathbf{r} \parallel / \mid b_i \mid)]$, is the same as in screening function 2, if one takes $b_i = 1$, and the $\xi = 1/d_0$. In the defining equation of HMLP, Equation (1), the summation does not include the atom α on which point $\mathbf{r}$ is located. Therefore, there is an exclusive region around point $\mathbf{r}$. Typically, in atomic units, the radius of the exclusive region is $2 \sim 3\ a_0$ ($1a_0 = 0.509$ Å). The largest difference between exponential and power distance-functions is around zero ($\mathbf{r} \sim 0$), where the exclusive region originates. Outside this exclusive region, the two types of distance-dependent functions are very similar. However, there still are two notable

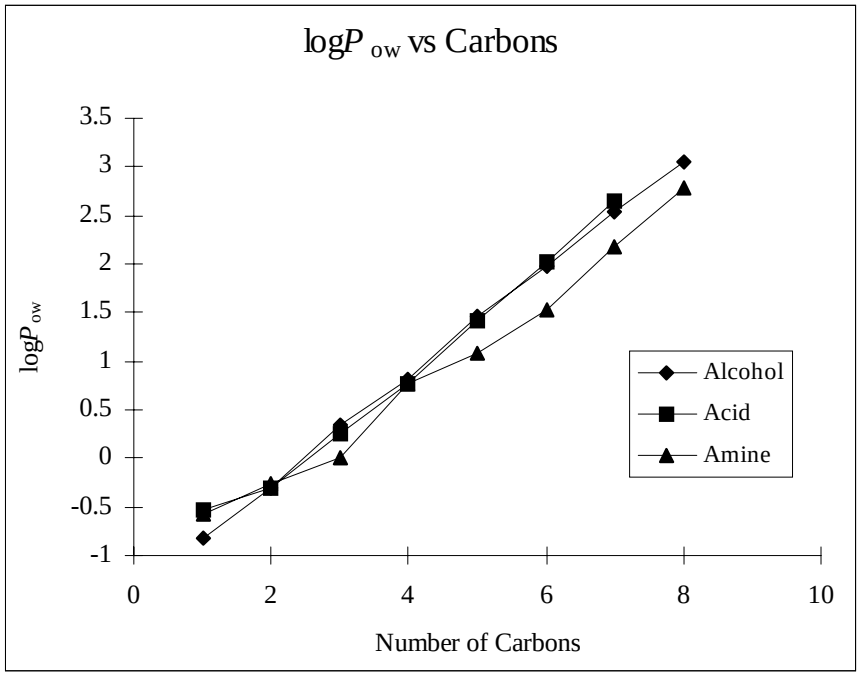

Figure 10. Experimental partition coefficients $\log P_{ow}$ of aliphatic alcohols, aliphatic carboxylic acids, and aliphatic amines as functions of the number of carbons [40].

differences. First, the exponential distance-function has larger values at short distances than the power distance-function. This means that the exponential distance-function is more influenced by its the nearest neighbors than the power distance-function. Second, the power distance-function has larger values at long distances than the exponential distance-function. This means that the power distance-function is more affected by atoms at longer distances.

From the results of previous calculations, we can say that both screening functions 1 (Equation (12)) and 2 (Equation (13)) give chemically reasonable results. However, the results of screening function 3 (Equation (9)) are not good. The distance-dependent functions in screening functions 2 and 3 are the same, however, the mathematical roles of atomic surface-MEP descriptors b_i's in these two screening functions are different. In screening function 2, b_i is a factor of the distance-dependent function, while in screening function 3, b_i is part of the exponent in the exponentially distance-dependent function, $\exp[-\xi(\| \mathbf{R}_i - \mathbf{r} \| / | b_i |)]$. In some cases, the b_i's are too small to provide non negligible values. For example, in propane, both carbons and hydrogen atoms have small atomic surface-MEP descriptors, b_i, therefore, their atomic lipophilicity indices are almost zero, (ref. Figure 3 (d)). For the hydrophilic groups, e.g., the hydroxyl group OH, the carboxyl group COOH, and the amino group NH_2, the results of screening function 3 are not too bad because of the large values of atomic surface-MEP descriptors in the functional groups. However, for the hydrocarbon chains of the three compounds (ethanol, propionic acid, and ethylamine), screening function 3 does not give reasonable results because of very small atomic surface-MEP descriptors b_i's. Molecular lipophilic indices, L_M, of the three types of compounds (alcohols, acids, and amines) do not increase much with the increasing number of carbon atoms in the hydrocarbon chains, using Equation (14) as the screening function, ref. Figure (7). This is a major fault with screening function 3. However, when parameter ξ becomes smaller, the results are better.

In both screening functions 1 (Equation (12)) and 2 (Equation (13)), atomic surface-MEP descriptors b_i's are factors of the distance-dependent functions. The difference is the distance-dependent functions. Screening function 1 uses a power decay function, $1/ \| \mathbf{R}_i - \mathbf{r} \|^{\gamma}$, and screening function 2 uses an exponential decay function, $\exp(- \| \mathbf{R}_i - \mathbf{r} \| / d_0)$. In Figures 1 and 2, the atoms in the hydrophilic functional groups, OH, COOH, and NH_2, have negative

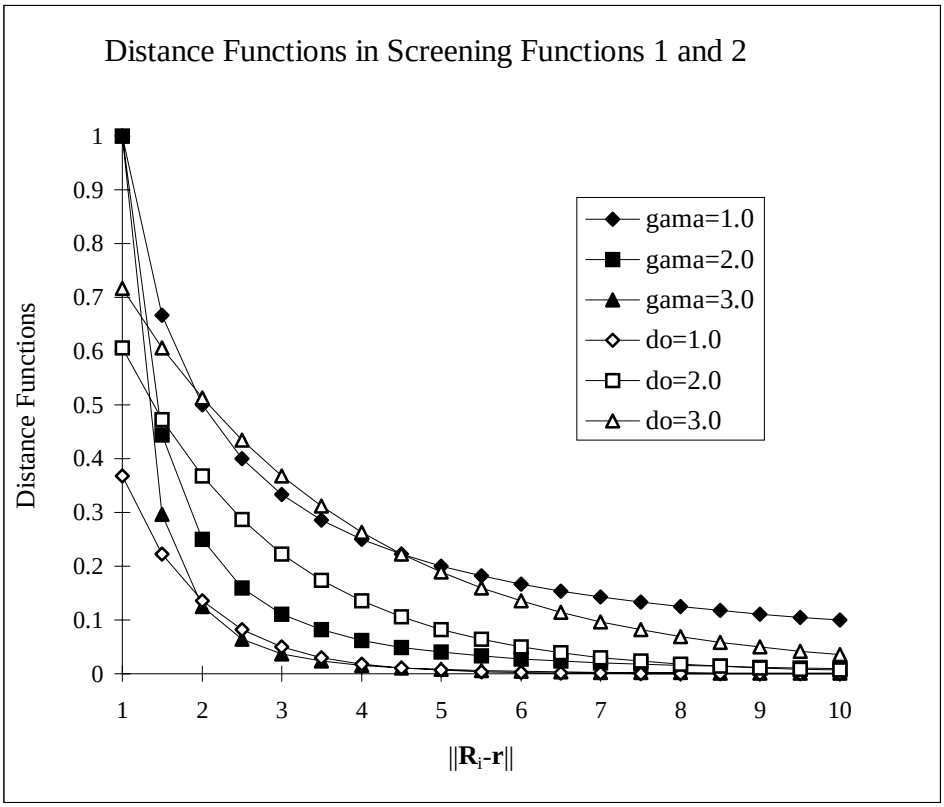

Figure 11. The curves of the power and exponential distance functions in screening functions 1 and 2 using different values of parameters γ and d_0. The black symbols are for the power distance function $1/\parallel \mathbf{R}_i - \mathbf{r} \parallel^{\gamma}$, and the white symbols are for exponential distance function $\exp(-\parallel \mathbf{R}_i - \mathbf{r} \parallel /d_0)$. The curves of the distance function in screening function 3 for $\xi = 1$, 1/2 and 1/3 are the same as those of the exponential distance function in screening function 2 for $d_0 = 1$, 2 and 3.

atomic lipophilicity indices, and the atoms in the lipophilic hydrocarbon chains have positive atomic lipophilicity indices. Based on Figures 1 and 2, the parameters γ and d_0 in Equations (12) and (13) affect the values of various indices in two different ways: the values of the indices decrease with the increasing γ in screening function 1, and increase with the increasing d_0 in screening function 2. In the carboxyl functional group COOH, there are 3 strongly charged atoms, therefore their atomic lipophilicity indices are very sensitive to the values of γ and d_0. We find an order of magnitude change of the indices l_H, l_O, and $l_{=O}$, caused by the changes in value of parameters γ and d_0, cf. Figure 4 (a) and (c). In the range $2.0 < \gamma < 3.5$ for Equation (12), the order of the indices is $l_H < l_{=O} < l_O$. In the range $1.5 < d_0 < 2.6$ for Equation (3), the order of the indices is the same as in Equation (13). In a carboxyl group, H tends to become more like the hydronium ion H_3O^+. It should be the most hydrophilic atom among the three atoms. Carbonyl oxygen, =O, has two lone electron pairs and a widely exposed surface area, therefore it is more hydrophilic than the oxygen in the hydroxyl group OH. The order $l_H < l_{=O} < l_O$ is reasonable, and it gives us a range for the optimization of parameters γ and d_0. In the amino group NH_2, there are three heavily charged atoms, and two hydrogens have the same sign of MEP. Their atomic lipophilicity idices are sensitive to the parameters γ and d_o, too. An interesting observation is that for both γ and d_0, there is a minimum in the atomic lipophilicity indices l_H. In the optimization of γ in screening function 1, the minimum of index l_H is -0.0113, at $\gamma = 2.3$. In the optimization of d_0, a minimum, -0.0411, in the index l_H is found at $d_0 = 1.9$, cf. Figure 4 (b) and (d). The minimum in l_H means that the effects from neighboring atoms are maximum. This finding provides a useful suggestion for the selection of parameters γ and d_o. However, it is not the only condition for the optimizations of parameters γ and d_o.

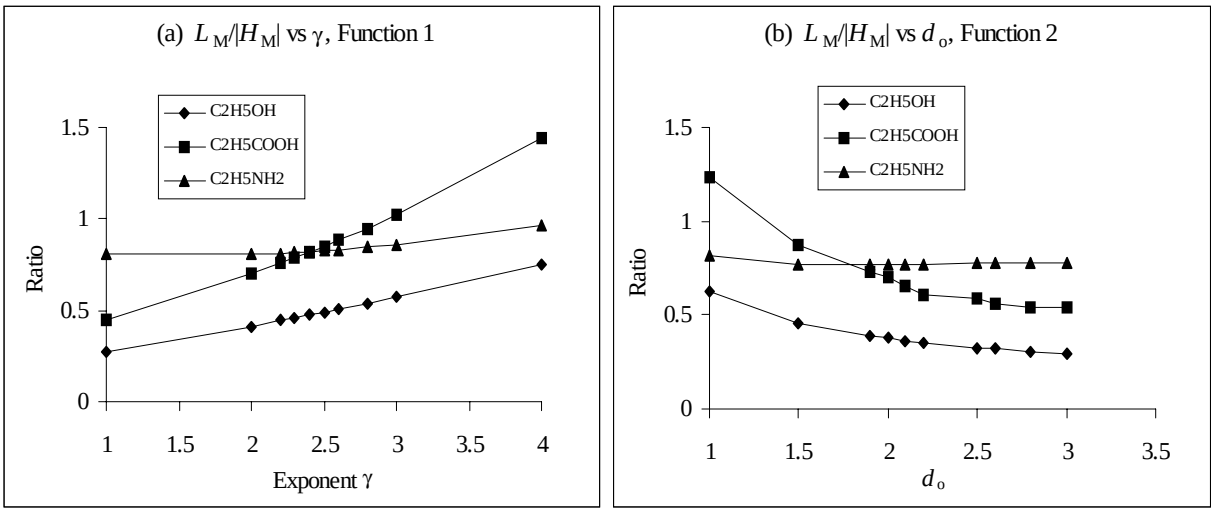

Figure 12. Ratio of indices $L_M/|H_M|$ of ethanol, propionic acid, and ethylamine as functions of parameters γ and d_o in screening functions 1 (a) and 2 (b).

For a long time, an additivity scheme of fragmental or atomic contributions to molecular lipophilicity has been used in the calculation of molecular partition coefficients. It is the basic principle of empirical MLP, which is the summation of atomic lipophilicity indices; cf. Equation (3). Israelachvili [15] said 'The hydrophilic and hydrophobic interactions, unlike electrostatic and dispersion interactions, are interdependent and therefore not additive'. It is obvious that simple additivity is not accurate enough for the calculation of molecular lipophilicity. However, much experimental evidence shows that the whole molecular lipophilicity is roughly the sum of all constituent fragments or atoms [3]. Here we do not wish to discuss and check the additivity of molecular lipophilicity, rather we focus on the validity of using partition coefficients as a criterion for HMLP, and discuss the meaning of the two indices: molecular lipophilic index L_M and molecular hydrophilic index H_M.

In the defining equations (15) and (16), L_M is the sum of atomic lipophilicity indices of all lipophilic atoms ($l_a > 0$), and H_M is the sum of atomic lipophilicity indices of all hydrophilic atoms ($l_a < 0$). MLP is a three dimensional representation of molecular lipophilicity. The indices L_M and H_M are the simplified and approximate representation of MLP. A rough picture is that the atomic lipophilicity indices l_a tell us the local lipophilicity on a molecule, while the molecular lipophilic index L_M is the overall measure of the lipophilic part of a molecule, and the hydrophilic index H_M is the overall measure of the hydrophilic part of a molecule, respectively. Reducing the three dimensional molecular lipophilicity potential into two simple scalar descriptors, L_M and H_M, is a drastic approximation, therefore the two indices are rather rough measures of molecular lipophilicity and hydrophilicity. However, these two indices are very convenient and useful for the approximate comparisons of molecular lipophilicity.

Figures 1, 2 and 3 show atomic lipophilic indices l_a, L_M and H_M as the functions of parameters (γ, d_0, and ξ) for ethanol, propionic acid, and ethylamine. Figure 12 shows the ratio of $L_M/|H_M|$ for these three compounds as functions of parameters γ and d_o. From these figures, we find that parameters (γ and d_0) affect both the values and ratio of L_M and H_M. Figure 12 (a) tells us that the absolute values of L_M and H_M decrease with increasing γ, but the ratio $L_M/|H_M|$ increases with increasing γ in screening function 1 for ethanol and propionic acid. The ratio $L_M/|H_M|$ of ethylamine has a minimum at $\gamma = 2.3$. In Figure 12 (b), for screening function 2, the absolute values of L_M and H_M increase with the increasing d_o, but the ratio $L_M/|H_M|$ of ethanol and propionic acid decreases with d_o, and the ratio $L_M/|H_M|$ of ethylamine has a minimum at $d_o = 1.9$. The order of the ratio $L_M/|H_M|$ for the three molecules changes with the parameters γ and d_0.

Figures 8 and 9 compare values of the L_M and H_M indices of the three series of compounds for various values of γ and d_o. The molecular hydrophilic index H_M is a rough measure of the hydrophilic

strength of the hydrophilic functional groups in a molecule. Figures 8 and 9 show that in the range $1.0 < \gamma < 3.0$ and $1.5 < d_0 < 3.0$ the orders of H_M, using screening functions 1 and 2, is the same: acids<alcohols<amines. This may be a reasonable order for the three functional groups (COOH, OH, and NH_2). This result provides another criterion for the selection of parameters γ and d_0. As mentioned before, the molecular lipophilic index L_M is a rough measure of the lipophilic strength of the lipophilic part of a molecule. As shown by Figures 8 and 9, the increase of L_M with the number of carbon atoms is approximately linear for the three types of compounds, albeit with different slopes. The order of the slopes of L_M for the three compounds is alcohols< amines< acids when $1.5 < \gamma < 3.0$ using screening function 1. The same order is found in Figure 9 in the range $1.5 < d_o < 3.0$ using screening function 2. This means that the hydrocarbon chains of aliphatic acids are the most lipophilic and the hydrocarbon chains of aliphatic alcohols are the least lipophilic among the three types of compounds. In Figure 12, we find the ratio $L_M/|H_M|$ of ethanol is the smallest for the three types of compounds when using two screening functions.

Partition coefficients $\log P_{ow}$ have been used as the overall measure of molecular lipophilicity for a long time. As an equilibrium constant for a two-phase system, $\log P_{ow}$ is determined by the difference of the solvation free energies ΔG° of the solute in each phase. This difference is represented by the molar standard free energy of transfer , from the aqueous phase to the organic phase,

$$\Delta\overline{G}^\circ_{tr} = \Delta\overline{H}^\circ_{tr} - T\Delta\overline{S}^\circ_{tr} \tag{19}$$

Common chemical knowledge is the following: a larger positive $\log P_{ow}$ means the molecule is more lipophilic; otherwise, a larger negative $\log P_{ow}$ means the molecule is more hydrophilic. Transfer free energy has two components, enthalpy and entropy. As pointed out by Israelachvili [15], lipophilic interaction is an entropy-controlled phenomenon. Similarly, we can say that hydrophilic interaction is an enthalpy-controlled phenomenon. A good guess is that there is a close relationship between the molecular lipophilic index L_M and transfer entropy $\Delta\overline{S}^\circ_{tr}$, and a close relationship exists between the molecular hydrophilic index H_M and transfer enthalpy $\Delta\overline{H}^\circ_{tr}$. Many authors [45, 46] point out that values of partition coefficients are not very sensitive to the changes of molecular structures. The reason is that sometimes the change of the two components of transfer free energy, enthalpy and entropy, caused by changes in chemical structure, cancel each other. Partition coefficients are also affected by the intramolecular hydrogen bonding. Therefore, partition coefficients cannot be used as a good criterion for HMLP in the comparison of miscellaneous compounds, however, in a family of compounds, partition coefficients are good criterion for HMLP.

Table 1. Summary of optimizations of parameters γ and d_0 in screening functions 1 and 2

Criteria	γ in Function 1	d_0 in Function 2
$l_H < l_{=O} < l_O$ in COOH	$2.0 < \gamma < 3.5$	$1.5 < d_0 < 2.6$
Minimum of l_H in NH_2	$\gamma = 2.3$	$d_0 = 1.9$
Minimum of $L_M/\|H_M\|$	$\gamma = 2.3$	$d_0 = 1.9$
Order of H_M's		
	$1.0 < \gamma < 3.0$	$1.5 < d_0 < 2.6$
Acids<alcohols<amines		
Best value	$\gamma = 2.3$	$d_0 = 2.0$

HMLP gives a good description of molecular lipophilicity distributions in a molecular surface or space, and atomic lipophilic indices l_a's are the representations of local lipophilicity. The experimental method that is best suited to provide the answers to the calculation results of HMLP is nuclear magnetic resonance (NMR) [47], which gives information about water structure near certain atomic groups [48–51]. In this study, we have not compared the calculation results with the experimental data of NMR, because the experimental data are not available. Besides NMR, there are several other experimental methods that may be used to get information for the improvements of HMLP. The methods most often used for the description of molecular motions in solution are infrared absorption, Raman and Rayleigh scattering, coherent and incoherent neutron scattering, dielectric and Kerr relaxation, and fluorescence depolarization. Measurements of heat capacities may also provide useful information for the development of a quantitative HMLP. More work should be done to check the calculation results of HMLP with various experimental data.

Generally speaking, both screening functions 1 and 2 give good results, and there is no large difference between the two screening functions. For screening function 1, the suitable range of parameter γ is $2.0 < \gamma < 3.0$, and $\gamma = 2.3$ is recommended. For screening function 2, the suitable range of parameter d_0 is

$1.5 < d_0 < 3.0$, and $d_0 = 2.0$ is recommended. Table 1 gives a summary of the above conclusions.

References

1. Du, Q., Arteca, G.A. and Mezey, P.G., J. Comput.-Aided Mol. Design, 11 (1997) 503.
2. Du, Q. and Arteca, G.A., J. Comput.-Aided Mol. Design, 10 (1996) 133.
3. Ghose, A.K. and Crippen, G.M., J. Comput. Chem., 7 (1986) 565.
4. Furet, P., Sele, A. and Cohen, N.C., J. Mol. Graph., 6 (1988) 182.
5. Heiden, W., Moeckel, G. and Brickmann, J., J. Comput.-Aided Mol. Design, 7 (1993) 503.
6. Viswanadhan, V.N., Ghose, A.K., Revankar, G.R. and Robins, R.K., J. Chem. Inf. Comput. Sci., 29 (1989) 163.
7. Kantola, A., Villar, H.O. and Loew, G. H., J. Comput. Chem., 12 (1991) 681.
8. Alkorta, I. and Villar H.O., Int. J. Quantum Chem., 44 (1992) 203.
9. Eisenberg, D., Weiss, R.M. and Terwilliger, T.C., Nature, 299 (1982) 371.
10. von Heijne, G. and Blomberg, C., Eur. J. Biochem., 97 (1979) 175
11. Ooi, T., Oobatake, M., Némethy, G. and Scheraga, H.A., Proc. Natl. Acad. Sci. USA, 84 (1987) 3086.
12. Cabani, S., Conti, G. and Lepori, L., Trans. Faraday Soc., 67 (1971) 1933.
13. Pleiss, M.A. and Grunewald, G.L., J. Med. Chem., 26 (1983) 1760.
14. Klopman, G., Namboodiri, K. and Schochet, M., J. Comput. Chem., 6 (1985) 28.
15. Israelachvili, J.N., Intermolecular and Surface Forces, Academic Press, London, 1992.
16. Politzer, P., In: Politzer, P. and Truhlar, D.G. (Eds.), Chemical Applications of Atomic and Molecular Electrostatic Potentials, Use of the Electrostatic Potential as a Guide to Understanding Molecular Properties, Plenum Press, New York, NY, 1981.
17. Murray, J.S., Ranganathan, S. and Politzer, P., J. Org. Chem., 56 (1991) 3734.
18. Murray, J.S., Brinck, T., Lane, P., Paulsen, K. and Politzer, P., J. Mol. Struct. (Theochem), 307 (1994) 55.
19. Tomasi, J., In: Politzer, P. and Truhlar, D.G. (Eds.), Chemical Applications of Atomic and Molecular Electrostatic Potentials, Use of the Electrostatic Potential as a Guide to Understanding Molecular Properties, Plenum Press, New York, NY, 1981.
20. Náray-Szabó, G., J. Mol. Struct. (Theochem), 138 (1986) 197.
21. Náray-Szabó, G., J. Mol. Graph., 7 (1989) 76.
22. Náray-Szabó, G., Int. J. Quantum. Chem.: Quant. Biol. Symp., 16 (1989) 87.
23. Náray-Szabó, G. and Nagy, P., Int. J. Quantum. Chem., 35 (1989) 215.
24. Murray, J.S., Brinck, T., Lane, P., Paulsen, K. and Politzer, P., J. Mol. Struct. (Theochem), 307 (1993) 55.
25. Murray, J.S., Brinck, T. and Politzer, P., J. Phys. Chem., 97 (1993) 13807.
26. Hansch, C. and Leo, A., Substituent Constants for Correlation Analysis in Chemistry and Biology, Wiley, New York, NY (1979).
27. Rekker, R.F., The Hydrophobic Fragmental Constants, Elsevier, New York, NY (1977).
28. Eisenberg, D. and McLachlan, A.D., Nature, 319 (1986) 199.
29. Dughan, L., Burt, C. and Richards, W.G., J. Mol. Struct. (Theochem), 481 (1991) 235.
30. Mishra, P.C. and Kumar, A., In: Sen, K. (Ed.), 'Topics in Current Chemistry (174), Molecular similarity II', Springer-Verlag, New York, NY.
31. Platt, D.E. and Silverman, B.D., J. Comput. Chem., 17 (1996) 358.
32. Israelachvili, J.N. and Pashley, R.M., Nature, 300 (1982) 341.
33. Marcelja, S., Mitchell, D.J., Ninham, B.W. and Sculley, M.J., J. Chem. Soc. Faraday Trans, 73 (1977) 630.
34. Audry, E., Dubost, J.P., Colleter, J.C. and Dallet, P., Eur. J. Med. Chem. Chim. Thér., 21 (1986) 71.
35. Audry, E., Dubost, J.P., Dallet, P., Langlois, M.H. and Colleter, J.C., Eur. J. Med. Chem. Chim. Thér., 24 (1989) 155.
36. Audry, E., Dubost, J.P., Dallet, P., Langlois, M.H. and Colleter, J.C., Eur. J. Med. Chem., 24 (1989) 155.
37. Croizet, F., Langlois, M.H., Dubost, J.P., Braquet, P., Audry, E., Dallet, Ph. and Colleter, J.C., J. Mol. Graph., 8 (1990) 153.
38. Fauchére, J.L., Quarendon, P. and Kaetterer, L., J. Mol. Graph., 6 (1988) 203.
39. Stillinger, F.H. and Rahman, A., J. Chem. Phys., 60 (1974) 1545.
40. Du, Q. and Arteca, G.A., J. Comput. Chem., 17 (1996) 1258.
41. Connolly, M.L., Science, 221 (1983) 709.
42. Connolly, M.L., J. Appl. Crystallog., 16 (1983) 548.
43. Connolly, M.L., J. Am. Chem. Soc., 107 (1985) 1118.
44. Leo, A., Hansch, C. and Elkins, D., Chem. Rev., 71 (1971) 525.
45. Cabani, S. and Gianni, P., J. Chem. Soc., Faraday Trans. I, 75 (1979) 1184.
46. Cabani, S., Gianni, P., Mollica, V. and Lepori, L., J. Sol. Chem., 10 (1981) 563.
47. Lee, S.H. and Rossky, P.J., J. Chem. Phys., 100 (1994) 3334.
48. Piculell, L., J. Chem. Soc., Faraday Trans., 82 (1986) 387.
49. Piculell, L. and Halle, B., J. Chem. Soc., Faraday Trans., 82 (1986) 401.
50. Halle, B. and Piculell, L., J. Chem. Soc., Faraday Trans., 82 (1986) 415.
51. Halle, B., Andersson, T., Forsén, S. and Lindman B., J. Am. Chem. Soc., 103 (1981) 500.