Study on the Correlation and Prediction of Octanol/Water Partition Coefficients by Quantum Chemical Calculations

Wolfgang Eisfeld[†] **and Gerd Maurer***

Lehrstuhl für Technische Thermodynamik und Fachbereich Chemie, Universität Kaiserslautern, D-67653 Kaiserslautern, Federal Republic of Germany

The influence and applicability of several quantum chemically obtained descriptors on the correlation/prediction of octanol/water partition coefficients (log P_{ow}) at 298 K is thoroughly studied. The log P_{ow} values are expressed as a linear equation in terms of several contributions to the chemical potential. The chosen parameters reflect the molecular interactions between the solute and the solvent molecules and account for cavity effects, polarity and polarizability, and the distribution of the electrostatic potential. The description of hydrogen-bonding by the sums of minima and maxima of the electrostatic potential is studied in detail for a wide variety of molecules. All descriptors can be easily computed by standard ab initio methods and result in good correlations with experimental log P_{ow} values. The prediction power of the method is tested for a set of solutes, possessing multiple functional groups, intramolecular hydrogen-bonding, and zwitterionic structures.

I. Introduction

Partition coefficients of organic solutes between two liquid phases are of considerable interest, e.g., in liquid−liquid extraction in industrial processes and for pharmacological and environmental research. Especially octanol/water partition coefficients (log P_{ow}) are widely used in quantitative structure activity relationships (QSAR).[1] Thus, there is a strong interest in predicting log P_{ow} values by theoretical methods in order to be independent from expensive and time consuming measurements. An extensive overview and discussion of different approaches to calculate log P_{ow} values was given by Leo in 1993.[2] Several methods were developed on the basis of group contributions like the well-known method of Hansch and Leo,[3] CLOGP,[4] or the application of UNIFAC.[5] Another successful approach is the LSER method developed by Kamlet, Taft, et al.,[6] which later was extended to other organic solvents.[7,8]

Considerable efforts were made to calculate log P_{ow} values by application of computational chemistry methods like molecular mechanics, semiempirical, and ab initio MO theoretical approximations. Most of the earlier methods are based on atomic charges and a "surface area".[9−11] Other groups suggested to calculate descriptors that are in some way related to the parameters of the LSER method.[12−15] Especially the use of the electrostatic potential instead of atomic charges seems to be promising for the modeling of electrostatic solvent−solute interactions.

In order to find a mathematical description of log P_{ow} a basic question must be answered: which physical and chemical effects are responsible for the distribution of a solute between two liquid phases? From the thermodynamic point of view the condition of phase equilibrium must be fulfilled.

$$\mu_i' = \mu_i'' \tag{1}$$

The chemical potential μ_i is split into the potential of a reference state, $\mu_{i,\text{ref}}$, and a deviation term containing the activity of component i, $x_i\gamma_i$.

$$\mu_i = \mu_{i,\text{ref}} + RT \ln x_i\gamma_i \tag{2}$$

The activity coefficient is a measure for the difference in the interactions of component i with its environment in comparison to the reference state. On a molecular scale this means that the sum of all interactions of the solute molecule with the solvation shell in its environment has to be known. The driving force for a solute to favor one of the phases must be the different strength of these interactions.

In general, three basic types of contributions to these forces can be defined: A volume term, electrostatic interactions, and hydrogen-bonding. The importance of the volume term is generally accepted.[4] Kamlet et al.[6] suggested that this term is a measure for the energy, necessary to form a cavity for the solute in the solvent. Other authors suggested that the volume term also contains the dispersion interaction.[16] However, a closer look on the given equations reveals that the polarizability is not treated explicitly. Since the polarizability scales somehow with the molecular volume, the volume term seems to represent the cavity effect together with the dispersion, which is physically not entirely correct.

Also, the electrostatic interactions are physically well-defined. If only uncharged molecules are considered, the interaction of multipoles and induced multipoles results in attractive forces between the molecules. For the simplest contribution, the interaction between two dipoles with dipole moments μ_1 and μ_2 at distance r, the energy $\bar{U}(r)$ is

$$\bar{U}_{\text{dipole}}(r) = -\frac{2\mu_1^2\mu_2^2}{3(4\pi\epsilon\epsilon_0)^2 kTr^6} \tag{3}$$

While for molecules possessing a dipole moment the dipole−dipole interaction is the major electrostatic force, in absence of a dipole moment the inductive interaction becomes important.

[†] Present address: Department of Chemistry, Emory University, Atlanta, GA. E-mail: eisfeld@chemie.uni-kl.de or eisfeld@euch4e.chem.emory.edu.

* Author to whom correspondence should be sent. E-mail: gmaurer@rhrk.uni-kl.de.

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5717**

The dipole moment μ_1 of a permanent dipole causes an induced dipole moment in a second molecule which is dependent on the polarizability α_2 of that particle. The dipole induction energy between a permanent dipole and a nonpolar molecule is

$$\bar{U}_{\text{ind}}(r) = -\frac{2\mu_1{}^2\alpha_2}{(4\pi\epsilon\epsilon_0)^2 r^6} \qquad (4)$$

In many cases, hydrogen-bonding is a very important interaction because the acting forces are orders of magnitude larger than the purely electrostatic contributions. Nevertheless, several authors claim that by means of electrostatic effects they can give a qualitative or even quantitative description of hydrogen-bonding[17−19] while others disagree in this question.[20] But, regarding the short bonding distances of approximately 2 Å, it becomes clear that considerable quantum mechanical effects must be taken into account to avoid the violation of Pauli's principle.[21] However, the group of Politzer successfully used the electrostic potential to model hydrogen-bonding acidity and basicity of organic molecules.[22−25]

In addition to these categories of interaction, several organic molecules possess the ability to form weak complexes (e.g., ethylene[26] or benzene[27]), which usually is not considered as hydrogen-bonding because the interaction energy is much lower. The binding forces of these complexes can be attributed to higher order multipole moments or dispersion (e.g., benzene− Ar complex), but orbital effects like those active in classical hydrogen-bonding are likely to be important as well. This latter contribution to the intermolecular forces was not explicitly taken into account by any of the cited methods to predict partition coefficients.

The aim of the present paper is to investigate the influence of the different interaction contributions on the correlation of octanol/water partition coefficients. Following the LSER concept, the partition coefficient ($P_{iw}^{(S)} = x_i/x_w$) for solute S between the organic solvent i and water at constant temperature (298 K) is expressed in terms of a linear equation.

$$\log P_{iw}^{(S)} = \sum_j \alpha_{ij}\beta_j^{(S)} \qquad (5)$$

We developed methods to compute the solute specific parameters $\beta_j^{(S)}$ from standard quantum chemical calculations. Since the parameters are derived from the electronic wave function of the solute, which gives a physically straightforward description of its properties, these descriptors should be capable of accounting for many problematic effects like intramolecular hydrogen-bonding, resonance or structure-related polarity changes. These parameters are used to correlate experimental partition coefficients for the system octanol/water (ow) at 298 K. The system specific coefficients α_{oj} of the linear equations are obtained by multiple linear regression analysis using experimental $\log P_{ow}$ values. The obtained equations are used to predict the partition coefficients for a set of test molecules that were not used to determine the coefficients.

II. Methodology

In the following we outline the methods for obtaining the parameters—also called descriptors—we use in the linear equations to correlate/predict the partition coefficients. The descriptors can be divided into three groups: (a) parameters $\beta_j^{(S)}$, which are directly obtained by quantum chemistry (dipole moment, polarizability), (b) the molecular volume, which is computed from the electron density, and (c) several parameters related to and obtained from the electrostatic potential.

A. Dipole Moment and Polarizability. These molecular properties are directly used as descriptors and are computed by use of standard quantum chemistry methods. They are clearly defined physical observables. They are based on the electron density distribution and thus are highly dependent on the theoretical method and the basis sets applied to calculate the wave function.[28,29] Throughout this study the Hartree−Fock approximation is used, which is known to overestimate the dipole moments. In contrast, the polarizabilities are frequently underestimated and are strongly dependent on the basis set, especially its diffuseness.[28] But, because these errors are more or less systematic ones, a good correlation with the electrostatic interaction forces can be expected.

B. Molecular Volume. The *molecular* volume, i.e., the volume of a single, discrete molecule, is not a physical observable and thus cannot be uniquely defined. From basic quantum mechanics it is obvious that a molecule should be infinitely expanded, but from everyday experience we know that a well-defined macroscopic volume contains a well-defined number of atoms or molecules. The reason for this seeming contradiction is that the electron density of a molecule or atom decreases exponentially with the distance at the long range tail of the wave function. Hence, at long distances the Pauli principle is no longer violated, if the tails of the particles penetrate each other.

Taking this into account, an envelope around the atom or molecule can be defined by a distinct low electron density value.[30] The volume included by this envelope can be considered as the molecular volume. Of course, then the absolute value of the molecular volume is dependent on the choice of the electron density threshold (see below). As a reasonable guideline for this choice, the integral of the electron density over the molecular volume should come close to the total number of electrons contained in the treated system.

An easy method to calculate the molecular volume is to compute the electron density of the particle and write it out in the shape of a three-dimensional rectangular grid. Then all cells are counted that possess an electron density higher than the chosen threshold, and this number finally is multiplied by the volume of the unit cell. By the same density grid the vectors that define the molecular surface can be obtained which will be used to compute properties like the surface electrostatic potential. Furthermore, the number of grid points on the molecular surface can be used to improve the volume calculation, but in the present work these corrections are of negligible magnitude because the unit cells are sufficiently small.

C. Parameters Derived from Electrostatic Potential. Several descriptors were already obtained from the electrostatic potential by other groups.[14,15,22−25] Especially the study of Haeberlein and Brinck[15] on the prediction of $\log P_{ow}$ values, which uses two such parameters, showed promising results. One of these parameters was interpreted as a polarity term and was named $\overline{V_-^2}$, while the other, denoted by ΣV_{min}, was connected to the hydrogen-bonding basicity. In order to avoid confusions with the volume, we will use U instead of V for the potential derived descriptors in the following.

While, in general, the electrostatic potential is a three-dimensional property, it was suggested that the potential values at the surface of a molecule are most influential for the interaction forces. Hence, a surface electrostatic potential U_S can be defined, if the molecular surface S is known (in fact the surface must be defined as well).

a)

b)
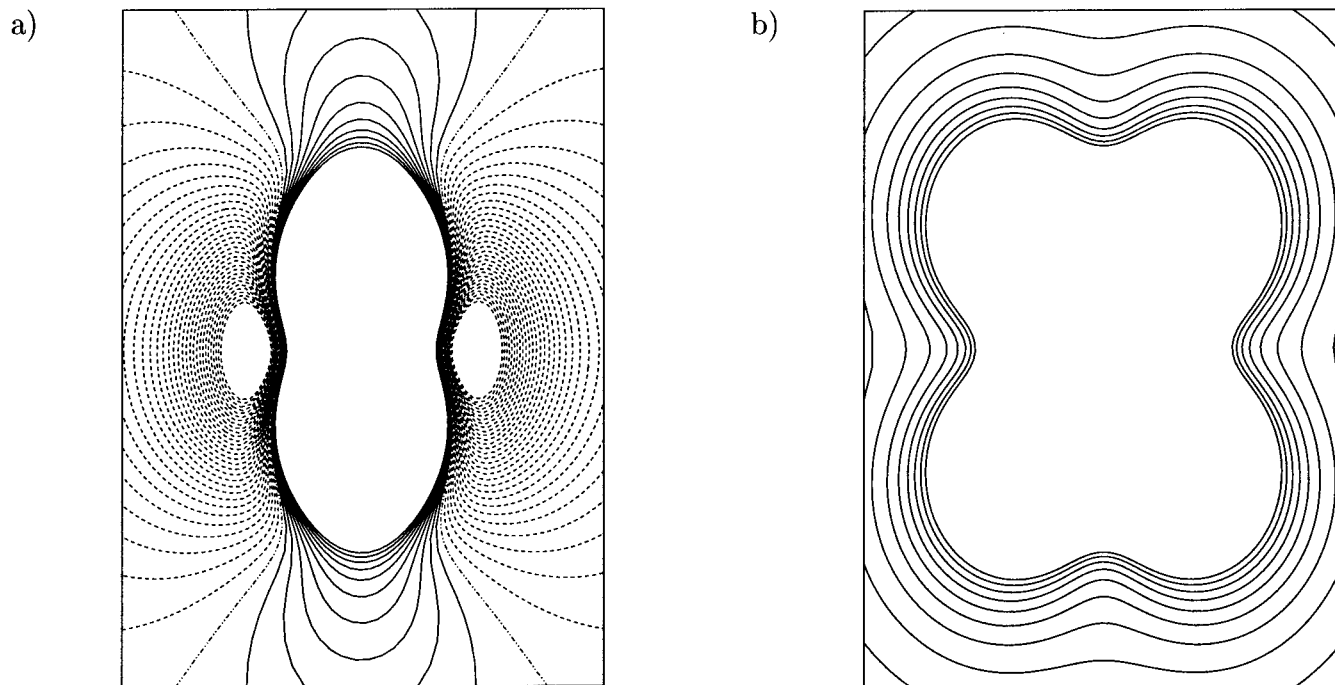
Figure 1. Contour plots of the electrostatic potential of ethylene. Solid lines indicate values above zero; dashed lines correspond to values below zero. (a) Potential cut through C−C bond axis, perpendicular to molecular plane. (b) Potential cut through C−C bond axis, in molecular plane.

$$U_S = \int U \, dS \approx \sum_{i=1}^{n} U_S(i) \, \Delta S \approx \frac{S}{n} \sum_{i=1}^{n} U_S(i) \qquad (6)$$

If U_S is split into a negative and a positive part ($U_{S,-}$ and $U_{S,+}$), some information about the surface polarity can be obtained. The parameter used by Haeberlein and Brinck[15] is defined as the mean of the squares of $U_{S,-}$.

$$\overline{U_-^2} = \frac{1}{n} \sum_{i=1}^{n} [U_{S,-}(i)]^2 \qquad (7)$$

We added a descriptor $U_{S,\%}$, which is the percental rate of $U_{S,-}$ in relation to U_S.

$$U_{S,\%} = \frac{\dfrac{1}{n} \sum_{i=1}^{n} |U_{S,-}|}{\dfrac{1}{n} \sum_{i=1}^{n} |U_{S,-}| + \dfrac{1}{m} \sum_{i=1}^{m} U_{S,+}} \qquad (8)$$

One further descriptor, $\sum U_{max}$ can be calculated from the surface potential U_S, which correlates with the hydrogen-bonding acidity (donor capability).[31] In this case the restriction to the maxima of the *surface* potential originates from the problem that otherwise the maxima would be found at the extremely high electron density regions of the core orbitals. Under this restraint and the use of an energy threshold, the located maxima are always attributed to hydrogens usually attached to oxygen or nitrogen atoms.

In contrast, the *spatial minima* of the electrostatic potential correlate with the hydrogen-bonding basicity (acceptor capability) and the sum of the single contributions, $\sum U_{min}$, was shown to be a useful descriptor for the partition coefficients.[15] For the use as a descriptor, several additional boundary conditions must be fulfilled by the single minima. E.g., these are minimal distances between the located extrema and energy thresholds (see below).

To account for the well-known influence of carbon−carbon multiple bonds,[3] we introduce a new parameter I_-, which is derived from the spatial electrostatic potential. This simply is the integral of the negative part of the potential, U_-, over the molecular volume divided by the integral of the total electrostatic potential over the molecular volume.

$$I_- = \frac{\int U_- \, dV}{\int U \, dV} \approx \frac{\displaystyle\sum_{i=1}^{n} U_-(i) \, \Delta V}{\displaystyle\sum_{j=1}^{m} U(j) \, \Delta V} \qquad (9)$$

This parameter is connected to the high electron density of the π bonds off the bonding axis, as becomes clear from the contour plots of the electrostatic potential of ethylene, shown in Figure 1.

III. Computational Details

All quantum chemical calculations were performed using GAUSSIAN 94.[32] The geometries of all solute molecules were optimized at Hartree−Fock level. The 3-21G* basis set was used,[33] which was supplemented by polarization functions of the d-type for N and O with exponents of 0.80 and 1.00, respectively (the original 3-21G* contains no d-functions for second row atoms). Only the lowest energy conformers were used to compute the electron density and the electrostatic potential over a three-dimensional rectangular grid. The grid was automatically generated by GAUSSIAN, resulting in a typical number of 515 000 grid points on average. Depending on the shape of the molecule, the points per dimension of three-dimensional space vary between approximately 60 and 100. Numerical integration of the electron density over the complete grid resulted in at least 99.98% of the number of valence electrons. First calculations of the electron density and electro-

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5719**
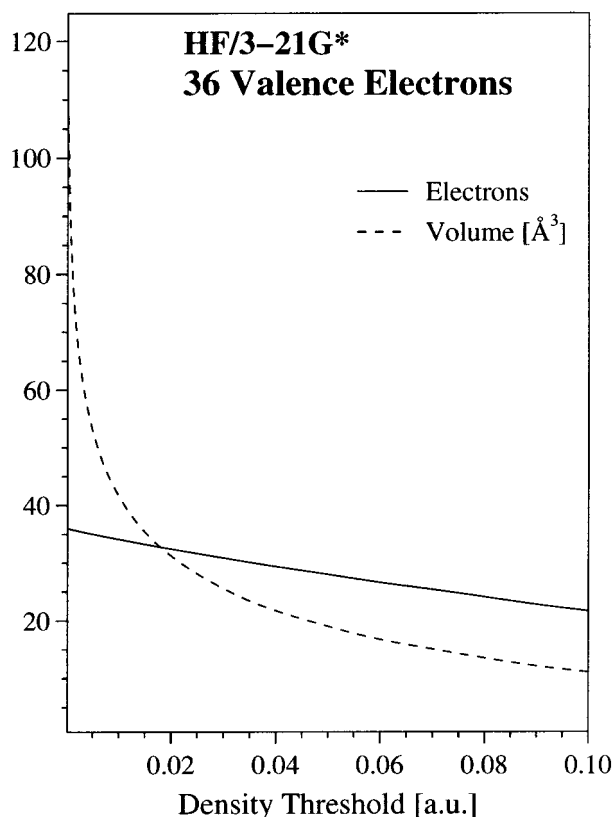
Figure 2. Dependency of the molecular volume and the integral of the electron density over the molecular volume (number of electrons) on the electron density threshold.

TABLE 1: Deviations Δ of Correlated log P_{ow} Values for Alkanes and Cycloalkanes[a]

solute	exp[b]	Δ[c]	Δ[d]	Δ[e]
butane	2.890	−0.084	−0.054	−0.061
cyclohexane	3.440	0.029	0.061	0.115
cyclopentane	3.000	−0.067	−0.047	−0.047
cyclopropane	1.720	0.186	−0.006	−0.008
2,2-dimethylbutane	3.820	0.017	0.025	0.011
2,3-dimethylbutane	3.850	0.008	0.023	0.002
ethane	1.810	−0.158	−0.118	−0.103
ethylcyclohexane	4.750	−0.168	−0.154	−0.113
heptane	4.500	0.031	0.021	0.006
hexane	4.000	−0.046	−0.042	−0.055
isobutane	2.760	0.033	0.059	0.056
methylcyclohexane	3.880	0.149	0.161	0.139
methylcyclopentane	3.370	0.130	0.141	0.134
neopentane	3.110	0.205	0.217	0.214
nonane	5.650	0.025	−0.013	−0.031
octane	5.150	−0.039	−0.063	−0.079
pentane	3.450	−0.073	−0.057	−0.066
propane	2.360	−0.127	−0.089	−0.089
tert-butylcyclohexane	5.620	−0.052	−0.065	−0.024
std dev		0.106	0.093	0.089

[a] $\Delta = [\log P_{ow}]_{pred} - [\log P_{ow}]_{exp}$. [b] Experimental log P_{ow} values from refs 3 and 34. [c] Equation 10 with parameters given in Table 7. [d] Equation 11 with parameters given in Table 7. [e] Equation 12 with parameters given in Table 7.

TABLE 2: Influence of the Density Threshold on the Correlation of log P_{ow} of the Solutes given in Table 1

density threshold[a]	0.0001	0.0005	0.0010	0.0015	0.0020
D_{max}	0.204	0.205	0.210	0.213	0.211
$\bar{D}$	0.088	0.086	0.088	0.097	0.103
SD	0.109	0.106	0.110	0.115	0.119

[a] In au. D_{max}: maximal absolute error. $\bar{D}$: mean absolute error. SD: standard deviation.

static potential were performed using the original as well as our supplemented 3-21G* basis set. To improve the accuracy of the calculated electron and charge distribution, the 6-31G* basis set was applied for the final computations. Nevertheless, the previous tests showed that the supplemented 3-21G* basis set gives results close to that of the larger 6-31G* basis set while the original 3-21G* is clearly inferior. The quantum chemical descriptors were calculated for a set of 78 non-hydrogen-bonding and 145 hydrogen-bonding molecules. A total of 202 solutes was used to determine the coefficients of the linear equations. The remaining 21 molecules form the test set for predicting log P_{ow} values for solutes not included in the fitting base. The numerical values of the parameters for all solutes are given in Tables A1−A3 in the Appendix.

IV. Results and Discussion

In order to evaluate the importance and the size dependency of the single descriptors, we investigated different sets of solute molecules. For alkanes and cycloalkanes hydrogen-bonding can be completely ruled out and other influences besides the molecular volume are of minor importance. Thus, from this group of molecules the volume dependency can be studied in detail. The influence of polarity and polarizability can be assessed by investigating hydrocarbons containing multiple bonds and halogenated hydrocarbons. Finally, the description of hydrogen-bonding effects can be examined best for mono-functional compounds containing N and O atoms.

A. Correlations for Nonpolar Solutes. The linear dependency of log P_{ow} of alkanes on the molecular volume is already well established,[2] but the way to obtain the volume is not unique. It was suggested that a volume calculated from the electron

density and included by a 0.002 or 0.001 au density contour should correlate well with the van der Waals volume.[30] But the exponential decay of the density at the long range tail of a molecule leads to a very strong dependency of the calculated volume on the chosen density threshold (cf. Figure 2).

Therfore, the volume V included by density envelopes between 1×10^{-4} and 2×10^{-3} au was computed for a test set of 19 molecules (cf. Table 1) and correlated to the log $P_{ow}^{(alk)}$ values according to eq 10.

$$\log P_{ow}^{(alk)} = aV + b \tag{10}$$

The best result (standard deviation SD = 0.106) was obtained for a density threshold of 5×10^{-4} au (cf. Table 2), which is less than the suggestions of other authors.[15,29] It also becomes obvious that the molecular volume correlates very well with log $P_{ow}^{(alk)}$. The effects of dipole moment and polarizability proved to be negligible because dipole moments are small or zero and the polarizability is proportional to the molecular volume for alkanes. In contrast, the inclusion of the parameters $\overline{U_-^2}$ or I_- into the correlation slightly improves the result (eq 11, SD = 0.093; eq 12, SD = 0.089). While $\overline{U_-^2}$ represents a local polarity, I_- accounts for the attractive forces between the highly negative charge density of C−C π bonds with H-donors.

$$\log P_{ow}^{(alk)} = aV + b\overline{U_-^2} + c \tag{11}$$

$$\log P_{ow}^{(alk)} = aV + bI_- + c \tag{12}$$

TABLE 3: Deviations Δ of Correlated log P_{ow} Values of Non-Hydrogen-Bonding Solutes[a]

solute	exp[b]	Δ^c	Δ^d	solute	exp[b]	Δ^c	Δ^d
allene	1.450	0.024	0.056	1-hexene	3.400	0.028	0.038
anthracene	4.500	0.159	0.152	1,5-hexadiene	2.800	0.164	0.176
azulene	3.220	0.101	0.123	hexane	4.000	−0.097	−0.097
benzene	2.130	−0.065	−0.054	hexyne	2.730	−0.010	0.017
biphenyl	3.980	0.044	0.040	indane	3.330	−0.005	−0.003
bromobenzene	2.990	−0.265	−0.185	indene	2.920	0.167	0.176
1-butene	2.400	−0.174	−0.158	isobutane	2.760	−0.001	0.008
2-butyne	1.460	0.255	0.300	isobutene	2.350	−0.191	−0.177
butadiene	1.990	−0.125	−0.118	isopropylbenzene	3.660	−0.063	−0.060
butane	2.890	−0.128	−0.121	methylcyclohexane	3.880	0.138	0.139
cis-butene	2.330	−0.124	−0.119	methylcyclopentane	3.370	0.028	0.028
trans-butene	2.310	−0.080	−0.076	naphthalene	3.350	0.117	0.123
chlorobenzene	2.840	−0.311	−0.229	neopentane	3.110	0.136	0.140
chloroform	1.970	0.403	0.274	nonane	5.650	−0.004	−0.017
1,3-cyclohexadiene	2.470	−0.007	0.003	1-nonene	5.150	0.004	−0.005
1,4-cyclohexadiene	2.300	0.226	0.227	octane	5.150	−0.083	−0.093
cyclohexane	3.440	0.086	0.085	1-octene	4.570	0.013	0.009
cyclohexene	2.860	0.081	0.080	1-pentyne	1.980	0.113	0.150
cyclopentane	3.000	−0.175	−0.173	1,4-pentadiene	2.480	0.003	0.018
cyclopropane	1.720	−0.063	−0.047	pentane	3.450	−0.129	−0.124
dibromethane	1.880	−0.009	0.046	phenanthrene	4.520	0.030	0.021
1,2-dichlorobenzene	3.280	−0.270	−0.188	phenylacetylene	2.400	−0.086	−0.211
1,3-dichlorobenzene	3.480	0.248	0.036	phenylallene	3.230	0.069	0.070
1,4-dichlorobenzene	3.440	−0.331	−0.300	phenylcyclopropane	3.270	−0.049	−0.038
dichloromethane	1.250	0.107	0.046	phenylpropene	3.350	−0.050	−0.037
1,2-dichloropropane	1.990	−0.035	−0.051	propane	2.360	−0.172	−0.161
difluoromethane	0.200	0.271	0.324	propene	1.770	−0.200	−0.185
1,2-dimethylbenzene	3.120	0.041	0.045	propyne	0.940	0.121	0.180
1,3-dimethylbenzene	3.200	−0.141	−0.130	styrene	3.050	−0.171	−0.161
1,4-dimethylbenzene	3.150	0.044	0.042	*tert*-butylbenzene	4.110	0.002	−0.003
2,2-dimethylbutane	3.820	−0.049	−0.048	*tert*-butylcyclohexane	5.620	0.031	0.013
2,3-dimethylbutane	3.850	−0.026	−0.024	tetrachloromethane	2.830	0.291	0.297
ethane	1.810	−0.219	−0.211	tetrafluoromethane	1.180	0.035	0.050
ethene	1.130	−0.072	−0.044	toluene	2.730	−0.098	−0.088
ethyne	0.370	−0.026	−0.231	1,1,1-trichloroethane	2.490	0.150	0.208
ethylbenzene	3.150	−0.021	−0.015	trichloroethene	2.350	0.327	0.224
ethylcyclohexane	4.750	−0.089	−0.097	trichlorofluoromethane	2.530	0.087	0.101
fluorobenzene	2.270	0.019	0.058	trifluoromethane	0.640	0.127	0.029
heptane	4.500	−0.011	−0.014	std dev		0.144	0.135
1-heptene	3.990	−0.066	−0.061				

[a] $\Delta = [\log P_{ow}]_{pred} - [\log P_{ow}]_{exp}$. [b] Experimental log P_{ow} values from refs 3 and 34. [c] Equation 13 with parameters given in Table 7. [d] Equation 14 with parameters given in Table 7.

B. Correlations for Polar Solutes without Strong Hydrogen-Bonding Capability. For this class of compounds the polarity must be taken into account, which can be expressed in terms of dipole moment (μ), polarizability (α), the percentage of the negative electrostatic potential on the molecular surface ($U_{S,\%}$), and the mean square of the negative electrostatic potential on the molecular surface ($\overline{U_-^2}$). To model the effect of π bonds, the integral of the spatial electrostatic potential over the molecular volume (I_-) can be included. $U_{S,\%}$ and $\overline{U_-^2}$ are dependent on the choice of the density threshold by which the molecular surface is defined. On the other hand α and I_- depend on the molecular size. Therefore, the influence of the surface density threshold and the molecular volume on the calculated log P_{ow} values was investigated.

The descriptors were determined for a set of 78 molecules (cf. Table 3) applying surface density thresholds between 5×10^{-4} and 2×10^{-3} au. Multiple linear regression of V, μ, α, $U_{S,\%}$, I_-, and $\overline{U_-^2}$ showed best results for a threshold value of 1×10^{-3} au. The corresponding standard deviation of the calculated log P_{ow} values is 0.155, and the largest absolute deviation is 0.449. The correlation can be improved by scaling the single contributions by different powers of the volume. The terms μV^{-2} and αV^{-2} would be proportional to the interaction energies of eqs 3 and 4. However, upon optimization of the powers of V with respect to the standard deviation, a term μV is found to give better results. The scaling of I_- and $\overline{U_-^2}$ is purely empirical.

$$\log P_{ow} = aV + b\mu V + c\frac{\alpha}{V^2} + dU_{S,\%} + e\sqrt{V}I_- + fV^2\overline{U_-^2} + g \quad (13)$$

The standard deviation and the maximum absolute deviation of the correlation by eq 13 are 0.144 and 0.403, respectively. A closer look at the properties of the electrostatic potential shows that some of the molecules obviously have hydrogen-bonding donor capabilities, even though no strong hydrogen bonds would be expected. Maxima of the surface electrostatic potential possessing values larger than 150 kJ/mol were located for chloroform, 1,3-dichlorobenzene, dichloromethane, 1,2-dichloropropane, phenylacetylene, trichloroethylene, and trifluoromethane (at a surface density threshold of 1×10^{-3} au). These extrema of the potential are always located at hydrogen atoms. For 1,3-dichlorobenzene two maxima are found for the hydrogens in the 4- and 6-positions[35] and the maximum for 1,2-dichloropropane is located at the hydrogen in the 2-position. If the sum of the potential maxima, ΣU_{max}, is used as an additional parameter, the correlation is further improved and results in a

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5721**

TABLE 4: Deviations Δ of Correlated log P_{ow} Values for Hydrogen-Bonding and Non-Hydrogen-Bonding Solutes[a]

solute	exp[b]	Δ[c]	Δ[d]	Δ[e]	solute	exp[b]	Δ[c]	Δ[d]	Δ[e]
acetaldehyde	0.450	−0.985	−0.760	−0.781	methyl phenyl sulfone	0.500	−0.177	0.218	0.420
acetamide	−1.260	0.428	0.097	−0.033	methyl *tert*-butyl ether	0.940	0.856	0.620	0.587
acetic acid	−0.170	−0.026	−0.073	−0.048	methyl acetate	0.180	−0.268	0.083	0.069
acetone	−0.240	−0.245	0.030	0.045	methylamine	−0.570	0.107	−0.283	−0.259
acetonitrile	−0.340	−0.318	−0.589	−0.599	methyl benzoate	2.200	0.181	−0.075	−0.118
acetophenone	1.630	0.091	0.163	0.189	methyl formate	0.030	−0.114	−0.116	−0.147
acroleine	−0.010	−0.213	0.146	0.137	4-methylpyridine	1.220	0.403	−0.084	−0.072
acrylonitrile	0.250	−0.082	−0.442	−0.454	*N*-methylpyrrole	1.210	0.699	0.603	0.427
acrylophenone	1.880	0.453	0.332	0.326	2-naphtol	2.700	0.274	−0.423	−0.421
adenine	−0.090	2.264	−0.676	−0.269	neopentanol	1.310	0.049	0.079	0.100
aminocyclohexane	1.490	0.204	0.081	0.088	*p*-nitroaniline	1.390	−0.420	−0.182	−0.209
1-aminonaphthalene	2.250	0.296	−0.043	−0.046	nitrobenzene	1.850	−0.687	−0.258	−0.483
2-aminonaphthalene	2.400	0.326	−0.354	−0.320	nitrobutane	1.470	−0.720	−0.121	−0.373
m-aminophenol	0.210	1.004	−0.379	−0.396	nitroethane	0.180	0.085	0.150	−0.128
o-aminophenol	0.620	0.188	0.470	0.217	nitromethane	−0.330	0.150	0.340	0.048
p-aminophenyl methyl sulfone	−0.120	1.731	−0.155	0.064	nitropropane	0.870	−0.521	−0.019	−0.285
3-amino-1-propene	0.030	0.433	−0.274	−0.244	*p*-nitrotoluene	2.420	−0.570	−0.165	−0.386
2-aminopyridine	0.320	0.592	0.240	0.272	octanol	3.070	0.355	0.062	0.075
aniline	0.900	0.757	−0.061	−0.027	pentachlorphenol	5.120	−0.700	−0.295	−0.423
anisole	2.110	0.213	−0.283	−0.347	2-pentanone	0.910	−0.866	−0.103	−0.072
benzaldehyde	1.480	−0.100	−0.039	−0.031	phenol	1.500	0.265	−0.431	−0.437
benzamide	0.640	0.307	0.488	0.330	phenyl trifluoromethyl sulfone	2.680	−2.610	−0.731	−0.600
benzoic acid	1.870	−0.173	−0.421	−0.398	piperidine	0.840	0.036	0.381	0.374
benzonitrile	1.560	−0.395	−0.392	−0.364	propanal	0.590	−0.657	−0.396	−0.414
benzophenone	3.180	−0.020	0.331	0.346	1-propanol	0.250	−0.239	−0.101	−0.065
butanal	0.880	−0.525	−0.164	−0.177	2-propanol	0.050	0.287	0.198	0.224
1-butanol	0.840	0.126	−0.069	−0.041	propanethiol	1.810	−0.124	0.227	0.287
tert-butanol	0.350	0.390	0.210	0.273	2-propen-1-ol	0.170	0.136	−0.264	−0.235
butanone	0.290	−0.411	0.071	0.087	propionic acid	0.330	0.175	−0.012	0.002
butanethiol	2.280	−0.391	0.343	0.403	propionitrile	0.160	0.116	−0.363	−0.369
3-butenenitrile	0.400	0.323	−0.113	−0.115	2-propyn-1-ol	−0.380	0.518	−0.228	−0.200
butylamine	0.860	0.607	0.119	0.123	pyridine	0.650	0.083	−0.131	−0.127
butyric acid	0.790	−0.269	−0.005	0.014	pyridine *N*-oxide	−1.200	−2.040	−0.252	−0.310
butyronitrile	0.600	0.061	−0.276	−0.267	pyrimidine	−0.400	0.809	−0.130	0.006
cyclohexanol	1.230	−0.407	−0.040	0.037	pyrrole	0.750	0.590	0.292	0.159
cyclohexanone	0.810	−0.608	0.240	0.271	pyrrolidine	0.460	0.333	0.120	0.128
cyclohexenone	0.610	−0.133	0.293	0.330	quinoline	2.030	0.241	−0.108	−0.113
cyclopropylamine	0.070	0.657	0.197	0.187	quinuclidine	1.380	0.345	0.368	0.368
dibutyl ether	3.210	0.861	0.329	0.269	THF	0.460	−0.435	0.047	0.050
3,5-dichloropyridine	2.560	0.035	−0.076	−0.186	thiophenol	2.520	0.264	0.143	0.206
diethyl ether	0.890	0.499	0.260	0.228	toluidene	1.390	1.440	0.939	0.821
diisopropyl ether	1.520	1.020	0.817	0.772	trifluoroethylamine	0.240	0.672	−0.295	−0.242
1,2-dimethoxybenzene	1.750	0.014	−0.036	−0.064	2,2,2-trifluoroethanol	0.410	0.505	−0.402	−0.372
dimethyl ether	0.100	−0.123	−0.182	−0.198	*p*-trifluoromethylaniline	2.390	0.714	0.193	0.173
dimethylacetamide	−0.770	0.281	1.004	0.815	trimethylamine	0.160	0.601	0.441	0.428
dimethylamine	−0.380	0.645	0.338	0.342	urea	−2.110	0.767	0.267	0.184
N,N-dimethylaniline	2.310	0.421	−0.116	−0.174	vinyl acetate	0.730	0.845	0.312	0.265
dimethylformamide	−1.010	0.595	0.806	0.607	allene	1.450	−0.166	0.330	0.349
dimethyl sulfone	−1.410	−0.146	0.117	0.274	anthracene	4.500	−0.258	0.288	0.267
dimethyl sulfoxide	−1.350	0.829	−0.193	−0.050	azulene	3.220	0.053	0.097	0.156
2,4-dinitrophenole	1.670	0.454	0.638	0.492	benzene	2.130	−0.130	−0.130	−0.086
dioxane	−0.270	0.150	0.032	0.122	benzothiophene	3.120	0.002	−0.099	−0.052
diphenyl sulfide	4.450	0.065	0.167	0.225	biphenyl	3.980	0.230	0.015	0.021
dipropyl ether	2.030	0.852	0.346	0.294	brombenzene	2.990	−0.083	−0.194	−0.149
ethanol	−0.300	0.162	−0.099	−0.069	butadiene	1.990	−0.173	−0.141	−0.096
ethyl vinyl ether	1.040	0.544	−0.069	−0.116	butane	2.890	−0.521	−0.154	−0.141
ethylamine	−0.130	0.455	−0.010	0.001	1-butene	2.400	−0.352	−0.194	−0.160
ethylene oxide	−0.300	−0.669	−0.261	−0.270	*cis*-butene	2.330	−0.299	−0.197	−0.157
ethyl propionate	1.210	−0.257	0.153	0.134	*trans*-butene	2.310	−0.210	−0.148	−0.111
formamide	−1.510	0.603	0.237	0.071	2-butyne	1.460	0.086	0.431	0.446
formanilide	1.150	0.408	0.344	0.157	chlorbenzene	2.840	−0.128	−0.255	−0.196
formic acid	−0.540	−0.224	−0.184	−0.135	chloroform	1.970	0.229	0.187	0.275
furan	1.340	−0.314	−0.341	−0.435	1,3-cyclohexadiene	2.470	−0.198	−0.099	−0.057
1-hexanol	2.030	−0.273	−0.111	−0.081	1,4-cyclohexadiene	2.300	0.067	0.115	0.151
2-hexanol	1.760	0.601	0.260	0.277	cyclohexane	3.440	−0.590	0.027	0.025
3-hexanol	1.650	0.540	0.417	0.426	cyclohexene	2.860	−0.292	−0.029	0.000
1-hexyn-5-one	0.580	0.981	0.834	0.817	cyclopentane	3.000	−1.119	−0.223	−0.206
indole	2.140	0.534	0.246	0.117	cyclopropane	1.720	−0.361	0.068	0.100
methacrylic acid	0.930	−0.080	−0.251	−0.244	dibromomethane	1.880	−0.272	0.027	0.083
methanol	−0.740	−0.045	−0.258	−0.224	1,2-dichlorobenzene	3.280	−0.921	−0.273	−0.187
2-methylfuran	1.850	−0.240	0.076	−0.111	1,3-dichlorobenzene	3.480	−0.410	0.134	0.186
1-methylimidazole	−0.060	0.719	−0.022	0.045	1,4-dichlorobenzene	3.440	−0.708	−0.283	−0.285
2-methyl-2-nitropropane	1.170	−0.331	0.055	−0.194	dichloromethane	1.250	0.000	0.149	0.213
2-methylphenole	1.980	0.125	−0.309	−0.324	dichloropropane	1.990	0.162	−0.012	0.086

TABLE 4 (Continued)

solute	$\exp^b$	Δ^c	Δ^d	Δ^e	solute	$\exp^b$	Δ^c	Δ^d	Δ^e
diethyl sulfide	1.950	0.264	0.535	0.598	nonane	5.650	−0.269	−0.280	−0.311
difluoromethane	0.200	0.284	0.391	0.462	1-nonene	5.150	−0.275	−0.193	−0.201
1,2-dimethylbenzene	3.120	−0.125	−0.057	−0.021	octane	5.150	−0.378	−0.319	−0.343
1,3-dimethylbenzene	3.200	−0.037	−0.547	−0.511	1-octene	4.570	−0.250	−0.157	−0.158
1,4-dimethylbenzene	3.150	0.029	−0.064	−0.037	1,4-pentadiene	2.480	−0.039	−0.036	−0.011
2,2-dimethylbutane	3.820	−0.682	−0.194	−0.194	pentane	3.450	−0.471	−0.218	−0.212
2,3-dimethylbutane	3.850	−0.624	−0.148	−0.154	1-pentyne	1.980	0.239	0.068	0.114
ethane	1.810	−0.728	−0.060	−0.026	phenanthrene	4.520	−0.859	0.193	0.176
ethene	1.130	−0.292	0.200	0.235	phenylacetylene	2.400	−0.424	−0.055	0.025
ethyne	0.370	0.097	0.232	0.275	phenylallene	3.230	0.057	−0.008	0.015
ethylbenzene	3.150	0.011	−0.140	−0.108	phenylcyclopropane	3.270	0.020	−0.129	−0.096
ethylcyclohexane	4.750	−0.843	−0.264	−0.282	phenylpropene	3.350	0.168	−0.417	−0.387
fluorobenzene	2.270	−0.128	0.015	0.073	propane	2.360	−0.592	−0.111	−0.089
heptane	4.500	−0.335	−0.210	−0.223	propene	1.770	−0.328	−0.182	−0.132
1-heptene	3.990	−0.271	−0.216	−0.205	propyne	0.940	0.211	0.316	0.352
1,5-hexadiene	2.800	0.200	0.038	0.065	styrene	3.050	−0.098	−0.252	−0.222
hexane	4.000	−0.426	−0.243	−0.248	*tert*-butylbenzene	4.110	−0.070	−0.107	−0.091
1-hexene	3.400	−0.160	−0.088	−0.073	*tert*-butylcyclohexane	5.620	−1.071	−0.225	−0.256
hexyne	2.730	0.072	−0.075	−0.034	tetrachloromethane	2.830	−0.169	0.165	0.178
indan	3.330	−0.158	−0.079	−0.044	tetrafluoromethane	1.180	−0.180	0.099	0.144
indene	2.920	0.026	0.094	0.136	thiophen	1.810	0.100	0.266	0.302
isobutane	2.760	−0.610	−0.021	−0.008	toluene	2.730	−0.135	−0.195	−0.155
isubutene	2.350	−0.361	−0.254	−0.209	1,1,1-trichloroethane	2.490	0.094	0.017	0.089
isopropylbenzene	3.660	−0.057	−0.175	−0.153	trichloroethene	2.350	0.025	0.119	0.195
methylcyclohexane	3.880	−0.506	0.020	0.012	trichlorofluoromethane	2.530	−0.270	−0.002	0.025
methylcyclopentane	3.370	−0.221	−0.095	−0.085	trifluoromethane	0.640	0.158	0.093	0.190
naphthalene	3.350	−0.346	0.160	0.163	std dev		0.540	0.287	0.274
neopentane	3.110	−0.499	0.045	0.052	max. absolute error		2.610	1.004	0.821

[a] $\Delta = [\log P_{ow}]_{pred} - [\log P_{ow}]_{exp}$. [b] Experimental $\log P_{ow}$ values from refs 3 and 34. [c] Equation 7 of ref 15. [d] Equation 15 with parameters given in Table 7. [e] Equation 16 with parameters given in Table 7.

standard deviation of 0.135 and a largest absolute deviation of 0.324.

$$\log P_{ow} = aV + b\mu V + c\frac{\alpha}{V^2} + dU_{S,\%} + e\sqrt{V}I_- +$$

$$fV^2\overline{U_-^2} + g\sum U_{max} + h \quad (14)$$

It turns out that the descriptor I_- is of significant importance to model the influence of π bonds and that the use of $\overline{U_-^2}$ alone is not sufficient. For comparison, eq 7 of ref 15 was also used to correlate the $\log P_{ow}$ for the solutes listed in Table 3. The descriptors were calculated by following the procedure and applying the thresholds given by Haeberlein and Brinck. Surprisingly, the result was much worse than could be expected. A standard deviation of 0.238 and a largest absolute deviation of 0.769 was obtained. Deviations larger than 0.5 were found for phenanthrene, 1,2-dichlorobenzene, cyclopentane, 2-butyne, and chloroform. A closer look at the data revealed that in the present work a considerably different molecular surface area (A_d) is obtained for phenanthrene while for all other molecules contained in ref 15 the A_d values could be reproduced within numerical errors.

C. Correlations for Solutes with Hydrogen-Bonding Capability. The correct description of hydrogen-bonding is extremely important for the prediction of partition coefficients for the major part of the molecules of potential interest. However, that effect is very difficult to assess for a number of reasons. The direct quantum chemical calculation of the respective interaction energies would be a research project for every single molecule and is unapplicable for large molecules and too time consuming for a large set of molecules. On the other hand, the approximation of hydrogen-bonding by purely electrostatic properties neglects the considerable effects of orbital interactions. However, Haeberlein and Brinck[15] published

encouraging results about accounting for hydrogen-bonding by the descriptor $\sum U_{min}$. Therefore, we used $\sum U_{min}$ and the related parameter $\sum U_{max}$ for our study of solutes with hydrogen-bonding capability.

Surprisingly, we were faced with substantial problems in correlating the $\log P_{ow}$ values of solutes containing more than one nitrogen or oxygen atom. Therefore, the set of molecules studied in ref 15 was reinvestigated. We were not able to reproduce the results of that study and a closer look at the data revealed that for 14 of the 74 treated molecules considerably different values for the $\sum U_{min}$ parameter were obtained. In general, for acids, esters, nitro compounds, and sulfones we obviously located one more interaction site than Haeberlein and Brinck. For urea, *p*-aminophenyl methyl sulfone, nitroaniline, and *p*-toluidine the differences are even more complicated or cannot be explained. Additionally, it must be mentioned that for several molecules multiple interaction sites were located at one and the same atom. In these cases only the strongest interaction was counted and summed up for the descriptor in order to obtain results comparable to those by Haeberlein and Brinck.

The localization of the electrostatic potential extrema depends on two boundary conditions that are more or less arbitrarily chosen. These are an energy threshold and a minimal distance between any pair of minima or maxima. Thus, we tried to adjust the threshold values in order to obtain data that can be correlated with the partition coefficients. First, it was found that at a minimal distance of 6 au the multiple minima vanish, which are found at the majority of the oxygen atoms when smaller distance thresholds are applied. Enhancing this value to 8 au reduces the interaction sites of nitro groups to a single oxygen atom. A closer look at the obtained minima energies revealed that for acids and esters the second interaction at the −OR oxygen always contributes less than −165 kJ/mol. Finally, it turns out that for hydrogens of potential hydrogen-bonding

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5723**

TABLE 5: Deviations Δ of Correlated log P_{ow} Values of Solutes Containing Only Nitrogen Heteroatoms[a]

solute	exp[b]	Δ[c]
acetonitrile	−0.340	0.097
acrylonitrile	0.250	−0.036
benzonitrile	1.560	−0.054
3-butenenitrile	0.400	0.033
butyronitrile	0.600	−0.124
propionitrile	0.160	−0.128
aminocyclohexane	1.490	0.151
1-aminonaphthalene	2.250	0.099
2-aminonaphthalene	2.400	−0.172
3-amino-1-propene	0.030	−0.250
aniline	0.900	−0.229
butylamine	0.860	−0.155
cyclopropylamine	0.070	0.037
dimethylamine	−0.380	0.115
N,N-dimethylaniline	2.310	−0.235
ethylamine	−0.130	−0.138
methylamine	−0.570	−0.063
toluidene	1.390	0.420
trimethylamine	0.160	−0.036
N-methylpyrrole	1.210	0.131
quinoline	2.030	0.000
quinuclidine	1.380	0.119
piperidine	0.840	0.298
indole	2.140	−0.079
1-methylimidazole	−0.060	0.210
pyridine	0.650	−0.013
pyrimidine	−0.400	0.291
pyrrole	0.750	−0.123
pyrrolidine	0.460	−0.041
3,5-dichloropyridine	2.560	−0.222
4-methylpyridine	1.220	−0.111
2-aminopyridine	0.320	0.271
adenine	−0.090	−0.061
std dev		0.168

[a] Δ = [log P_{ow}]$_{pred}$ − [log P_{ow}]$_{exp}$. [b] Experimental log P_{ow} values from refs 3 and 34. [c] Equation 15 with parameters given in Table 7.

activity, maximal potential values of at least 160 kJ/mol can be expected. Applying the following thresholds improved the correlation results considerably: minimal distance between minima of 8 au; minimal distance between maxima of 6 au; maximum energy of −165 kJ/mol for $\sum U_{min}$; minimal energy of 160 kJ/mol for $\sum U_{max}$; surface density threshold of 1×10^{-3} au; volume density threshold of 5×10^{-4}. For a set of 202 molecules experimental partition coefficients were correlated by

$$\log P_{ow} = aV + b\mu V + c\frac{\alpha}{V^2} + dU_{S,\%} +$$
$$e\sqrt{VI_-} - fV^2\overline{U_-^2} + g\sum U_{max} + h\sum U_{min} + i \quad (15)$$

The results are given in Table 4.

Equation 7 of ref 15 was used for comparison. The parameters for that equation were computed using the thresholds suggested by Haeberlein and Brinck. Equation 15 gives a standard deviation of 0.287 and a largest absolute deviation of 1.004. This result can at least compare with other theoretical methods reported so far, especially if the broad variety of solutes treated in the present study is taken into account. The correlation using eq 7 of ref 15 results in a standard deviation of 0.540 and a largest absolute deviation of 2.610. The deviations should be discussed also with respect to the sometimes substantial scattering of experimental partition coefficients, which often is larger than 0.5 in log P_{ow}. Only 7% of the log P_{ow} values calculated by eq 15 show deviations larger than 0.5. A similar

TABLE 6: Deviations Δ of Correlated log P_{ow} Values of Solutes Containing Only Oxygen Heteroatoms[a]

solute	exp[b]	Δ[c]
acetone	−0.240	0.179
acetophenone	1.630	−0.019
acrylophenone	1.880	0.209
benzophenone	3.180	0.161
butanone	0.290	0.042
cyclohexanone	0.810	0.132
cyclohexenone	0.610	0.209
pentan-2-one	0.910	−0.205
acetaldehyde	0.450	−0.502
acrolein	−0.010	0.281
benzaldehyde	1.480	−0.190
butanal	0.880	−0.235
propanal	0.590	−0.351
acetic acid	−0.170	0.118
benzoic acid	1.870	−0.377
butyric acid	0.790	−0.017
formic acid	−0.540	0.312
methacrylic acid	0.930	−0.180
propionic acid	0.330	0.044
ethyl propionate	1.210	0.000
methyl acetate	0.180	0.046
methyl benzoate	2.200	−0.185
methyl formate	0.030	−0.025
vinyl acetate	0.730	0.211
1-butanol	0.840	−0.173
tert-butanol	0.350	0.243
cyclohexanol	1.230	0.020
ethanol	−0.300	0.092
1-hexanol	2.030	−0.261
2-hexanol	1.760	0.101
3-hexanol	1.650	0.233
methanol	−0.740	0.216
2-methylphenol	1.980	−0.219
2-naphthol	2.700	−0.230
neopentanol	1.310	−0.077
octanol	3.070	−0.104
phenol	1.500	−0.276
2-propan-l-ol	0.170	−0.088
2-propyn-l-ol	−0.380	0.197
1-propanol	0.250	−0.053
2-propanol	0.050	0.192
anisole	2.110	−0.292
dibutyl ether	3.210	0.085
diethyl ether	0.890	0.022
diisopropyl ether	1.520	0.498
1,2-dimethoxybenzene	1.750	0.128
dimethyl ether	0.100	−0.205
dipropyl ether	2.030	0.075
ethyl vinyl ether	1.040	−0.138
methyl *tert*-butyl ether	0.940	0.374
dioxane	−0.270	−0.131
ethylene oxide	−0.300	0.137
furan	1.340	−0.236
2-methylfuran	1.850	0.249
THF	0.460	−0.035
std dev		0.209

[a] Δ = [log P_{ow}]pred − [log P_{ow}]$_{exp}$. [b] Experimental log P_{ow} values from refs 3 and 34. [c] Equation 15 with parameters given in Table 7.

comparison with the results of eq 7 of ref 15 reveals that 29% of the calulated log P_{ow} values exceed this arbitrary limit.

Nevertheless, substantial deviations are encountered for molecules like acetaldehyde, acetonitrile, diisopropyl ether, *N,N*-dimethylacetamide, and *p*-toluidine. This finding is somewhat surprising because these molecules do not possess multiple functional groups appropriate for hydrogen-bonding and hence should be described properly and without arbitrariness by our method. To gain better insight into the effects, all molecules containing only nitrogen as a heteroatom were selected from Table 4 and a new regression analysis was performed. The

TABLE 7: Coefficients for Quantum Chemical Descriptors in Equations 10−16

eqn	V	μV	α/V^2	$U_{S,\%}$	I_-	$\sum U_{min}$	$\sum U_{max}$	$\overline{U_-^2}V^2$	N_{ind}	O_{ind}	const
10	3.435										−0.034
11	3.427							−0.500			0.039
12	3.341				−8.056						0.147
13	3.586	0.481	0.203	−0.563	−6.695			−7.664			−0.164
14	3.549	0.619	0.175	−0.595	−6.192		−0.863	−7.916			−0.107
15	3.460	0.381	0.758	−0.942	−5.688	3.987	−0.788	−0.036			−0.371
16	3.393	0.595	0.739	−0.876	−5.769	3.393	−0.586	−0.034	−0.259	−0.241	−0.305
15 (N)[a]	5.101	0.067	3.707	−1.973	−1.994	2.878	−1.070	−0.020			−4.051
15 (O)[b]	4.083	−0.679	2.071	−1.094	−1.592	5.141	−0.297	−0.025			−1.800

[a] Fit of compounds in Table 5, containing only N heteroatoms. [b] Fit of compounds in Table 6, containing only O heteroatoms.

results are given in Table 5. It turns out that the octanol/water partition coefficients of most of the selected compounds can be represented excellently by eq 15, as indicated by a standard deviation of 0.168. Especially, the log P_{ow} values of acetonitrile, acrylonitrile, benzonitrile, *N*-methylpyrrole, and adenine, which counted among the worst represented solutes of Table 4, are now calculated with absolute errors of less than 0.14. Only for *p*-toluidine does the obtained partition coefficient show a considerable deviation of 0.42, which may indicate that eq 15 neglects an effect that is important for this compound.

For another regression analysis all molecules were selected that contain only oxygen as heteroatoms and the multiple linear regression was repeated (cf. Table 6). For this correlation a standard deviation SD = 0.209 was obtained, which is noticeably better than the correlation of eq 15 given in Table 4, but it does not reach the accuracy of the correlated values for the solutes listed in Table 5. The largest deviations (0.502 and 0.498) are found for acetaldehyde and diisopropyl ether, respectively, indicating some kind of systematic error. If the results for diisopropyl and di-*n*-propyl ether are compared, it becomes obvious that eq 15 results in very similar partition coefficients for both isomers while the experimental values are markedly different. In fact, all the theoretical descriptors are quite similar for both compounds. But while for di-*n*-propyl ether these parameters result in a very good description of the log P_{ow} value, they fail to represent the correct value for the isomer. A similar behavior is found for 1-propanol and 2-propanol. While the partition coefficient of the primary alcohol is described correctly, that of the secondary one shows a considerable deviation. Even the tendency of the log P_{ow} values is not matched, as the experimental value for the secondary alcohol is lower than that of the isomer. The same phenomenon is observed for the three hexanol isomers. From basic organic chemistry this behavior could be explained by the well-known electron pushing effect of alkyl groups. However, the effects on the electron distribution should be covered by the method. Hence, it is assumed that the neglect of the influence of orbital interactions is responsible for that deficiency.

Finally, one regression analysis was made for a set of solutes containing all molecules of Tables 3, 5, and 6. A standard deviation of 0.252 is obtained, which is larger than every single standard deviation for the three subsets, but still less than the one obtained for correlating all 202 molecules treated in Table 4. The largest absolute deviation of 0.982, found for *p*-toluidine, is considerably larger than the respective values for the subsets. This indicates that in principle the main contributions to the partition coefficients are covered by eq 15 but that at least one factor is neglected, which could balance the results for the different subsets. It is tempting to suppose that by including some kind of orbital effect the correlation could be further improved, but no suitable approximation has been found yet. However, a rather simple way to adjust the correlation properties

of the different kinds of molecules is the introduction of indicator variables for the presence of nitrogen (N_{ind}) and oxygen (O_{ind}) atoms in the system. These variables are set to unity if the molecule contains the respective atom and otherwise they are zero.

$$\log P_{ow} = aV + b\mu V + c\frac{\alpha}{V^2} + dU_{S,\%} + e\sqrt{V}I_- + fV^2\overline{U_-^2} +$$
$$g\sum U_{max} + h\sum U_{min} + iN_{ind} + jO_{ind} + k \quad (16)$$

The results for the correlation of eq 16 are also given in Table 4. In fact, the use of these indicators leads to a further improvement of the results. For the set of the 202 solutes of Table 4 a standard deviation of 0.274 was obtained, which is slightly better than the correlation without the indicators. The largest absolute deviation is reduced to 0.821 and the number of solutes for which deviations larger than 0.5 are obtained decreases to 5%. Of course, from the theoretical point of view, unlike the calculated parameters, the empirical indicators lack a deeper physical meaning and are thus not really satisfactory. Nevertheless, from practical aspects they seem to be quite useful until a better description of the physical behavior is found.

D. Significance of the Fitting Coefficients. Since the quantum chemical descriptors are developed to carry transparent physical meaning, it is very tempting to analyze them in connection with the respective linear coefficients to see the contributions of a particular parameter to the log P_{ow} value. However, one has to keep in mind that our parameter set contains several descriptors for closely related phenomena while other effects seem to be missing, which prevents an easy interpretation. Also, the present theoretical description needs to include all interactions separately while in the experimentally based LSER each measured parameter accounts for the effect of a whole group of interactions. That means that most parameters of the present TLSER method cannot be directly compared to the descriptors of conventional LSER.

In principle, it can be distinguished between effects that increase or decrease the log P_{ow} value. The only increasing effect should be the volume, and as Table 7 shows, the volume in fact gives a large positive contribution. However, smaller positive effects are also found for the dipole moment and the polarizability, which is not to be expected. Since the polarizability scales approximately with the size of the solute molecule and the induced dipole interaction should be very small anyway, this indicates that the α parameter picks up some of the significance of the volume. The positive influence of the dipole moment is even less expected because this should give a negative contribution that should be at least effective for solutes with significant polarity but without hydrogen-bonding capability. However, the parameters $U_{S,\%}$ and I_- also contain informa-

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5725**

TABLE 8: Experimental and Predicted log P_{ow} Values and Their Deviations Δ for Solutes Not Used for Fitting[a]

solute	exp[b]	pred[c]	Δ
1,2-ethanediol	−1.36	−1.126	0.234
1,3-propanediol	−1.04	−0.693	0.347
ethyl acetate	0.73	0.793	0.063
butyrolactone	−0.36	−0.091	0.269
glycine[d]	−3.21	−3.099	0.111
alanine[d]	−2.96	−2.556	0.404
1,2-dibromobenzene	3.64	3.362	−0.278
1,3-dibromobenzene	3.75	3.736	−0.014
1,4-dibromobenzene	3.79	3.777	−0.013
2-chloroaniline	1.90	1.445	−0.455
3-chloroaniline	1.88	1.417	−0.463
4-chloroaniline	1.83	1.374	−0.456
2-chlorobenzoic acid	2.05	1.934	−0.116
3-chlorobenzoic acid	2.68	2.442	−0.238
4-chlorobenzoic acid	2.65	2.553	−0.097
2-hydroxybenzoic acid	1.50	1.230	−0.270
3-hydroxybenzoic acid	1.50	1.045	−0.455
4-hydroxybenzoic acid	1.58	1.443	−0.137
2-nitrophenol	1.79	1.818	0.028
3-nitrophenol	2.00	1.569	−0.431
4-nitrophenol	1.91	1.479	−0.431
std dev			0.300

[a] $\Delta = [\log P_{ow}]_{pred} - [\log P_{ow}]_{exp}$. [b] Experimental log P_{ow} values from ref 3. [c] Equation 16 with coefficients given in Table 7. [d] Molecular geometry optimized at the HF/6-31G* level.

tion about the polarity of the solute and it is reasonable to assume that the positive μ contribution is a result of balancing these effects. It is also possible that this parameter acts in a way to correct the overestimation of the influence of hydrogen-bonds by the remaining descriptors.

The hydrogen-bond basicity is modeled by $\sum U_{min}$ and $\overline{U_-^2}$, and from Table 7 it becomes obvious that only in the absence of $\sum U_{min}$ does the second descriptor dominate this contribution. The influence of hydrogen-bond acidity is comparably small, which is in agreement with the findings of the experimental LSER.

The fitting coefficients for the test sets of Table 5 and 6 show some significant deviations between the N and the O set. This indicates that there must be effects specific to the respective atoms that are not sufficiently described by the present method. It seems that the purely electrostatic modeling of the partition coefficients cannot overcome this problem, which means that effects like orbital interactions may play a significant role. The inclusion of this type of effect will be the subject of further work.

E. Prediction of log P_{ow} Values. To test the prediction power of the developed method, a test set of solutes was chosen that were not included in the fitting base. The selected molecules possess combinations of functional groups and structural features that were not used in the multiple linear regressions. Special attention was payed to intramolecular hydrogen-bonding, structure effects, and the zwitterionic form of amino acids. For the solutes given in Table 8 the solute specific parameters were determined and then used to predict the log P_{ow} values by eq 16. The experimental and predicted values together with the respective deviations are also given in Table 8.

The obtained standard deviation of 0.300 is perceivably larger then for the correlation given in Table 4 which, of course, is not surprising. However, the deviations are not unreasonably large, especially if the fact is taken into account that the test set contains many problematic solutes. The standard deviation of the predictions is comparable to the estimated reliability of the experimental determination of partition coefficients.[34] Of

the two diols only 1,2-ethanediol contains an intramolecular hydrogen-bond. For both compounds log P_{ow} values slightly larger than the experimental ones are predicted. This shows that the intramolecular hydrogen-bond is recognized by the method because otherwise the calculation should lead to a much lower log P_{ow} value for 1,2-ethanediol. The results for ethyl acetate and butyrolactone reflect the strong structural effect that causes much lower partition coefficients for cyclic esters in comparison to their open chain analogues. In principle, this effect is described correctly by eq 16, though the deviation for the lactone is noticeably larger.

The amino acids are a remarkable challenge for the method since they possess intramolecular hydrogen-bonding along with an extremely high polarity, caused by their zwitterionic structure.[36] Nevertheless, the predicted log P_{ow} values compare fairly well to very well with the measurements. It is also worth noting that the partition coefficients for the amino acids are considerably lower than any value used in the fitting process. That means that the validity range of the method is not limited by the range of fitted data points.

The set of dibromobenzenes addresses a simple form of isomeric effect that results in slightly different predictions for the log P_{ow} values. As can be expected, the lowest partition coefficient is found for the 1,2-isomer, which possesses the highest polarity. The value for the 1,3-isomer is significantly larger than that of the 1,2-isomer and slightly smaller than the value found for the 1,4-conformation. While the 1,3- and 1,4-isomers show excellent agreement with the measurements, the deviation for the 1,2-structure is slightly larger. The set of chloroanilines demonstrates that the isomer effect can be matched rather well, although the absolute deviation from the measurements is somewhat larger. The three chlorobenzoic acids show also the expected trends, partly in contrast to the measurements. But in this case a concealed superposition of different effects may lead to the unexpected ordering of the experimental values.

In the set of hydroxybenzoic acids again intramolecular hydrogen-bonding is encountered in the case of salicylic acid. This is reflected by the prediction of a larger partition coefficient of the 2-isomer with respect to the 3-structure. Surprisingly, the experimental values for both solutes are equal, which results in a noticeably larger deviation for the 3-isomer while for the 2- and 4-structures the errors are fairly small.

A comparable type of intramolecular hydrogen-bonding is present in the nitrophenols. Here, the effect of the hydrogen-bond present in the 2-isomer leads to a remarkable agreement with the measurements while for the 3- and 4-isomers a considerably larger deviation is found. However, these deviations are of the same magnitude and direction as the ones found for aromatic nitro compounds in the multiple linear regressions, indicating that additional effects may be active in this class of solutes.

V. Conclusions

For a wide variety of organic compounds, including molecules with several functional groups and hetero atoms, we correlated experimental octanol/water partition coefficients (log P_{ow}) at 298 K by linear equations using theoretical solute descriptors following the concept of the linear solvation energy relationship (LSER). The descriptors are obtained from quantum chemical ab initio calculations and account for cavity effects, electrostatic interactions, and hydrogen-bonding. The majority of solutes is represented very well by our method (standard deviation in log P_{ow} for all correlated 202 molecules of 0.274). Only for very

few systems are larger deviations from experimental values obtained. Predictions for partition coefficients that are not included in the fitting base show fairly good to very good agreement with experiments (standard deviation in log P_{ow} for 21 solutes of 0.300). Especially intramolecular hydrogen-bonds are described properly by the method. However, our results indicate that not all effects that determine the partition coefficients are covered by the suggested parameters. We presume that the description of hydrogen-bonding by purely electrostatic properties derived from the electrostatic potential cannot sufficiently account for the true interactions. Hence, we suggest

that orbital effects must be included in the model in order to further improve the correlation. This will be the subject of future work.

Appendix

Tables A1−A3 contain the numerical values of the parameters for all the solutes.

TABLE A1: Quantum Chemical Descriptors To Correlate log P_{ow} Values of Non-Hydrogen-Bonding Solutes

solute	V^a	μV^b	$\alpha/V^{2\ c}$	$U_{S,\%}$	I_-	$\sum U_{max}$	$\overline{U_-^2}V^{2\ d}$	solute	V^a	μV^b	$\alpha/V^{2\ c}$	$U_{S,\%}$	I_-	$\sum U_{max}$	$\overline{U_-^2}V^{2\ d}$
allene	0.535	0.000	1.057	0.151	0.059	0.000	0.002	1-heptene	1.263	0.043	0.450	0.334	0.034	0.000	0.018
anthracene	1.651	0.000	0.520	0.350	0.077	0.000	0.064	1-hexene	1.099	0.035	0.511	0.302	0.031	0.000	0.012
azulene	1.269	0.176	0.643	0.643	0.079	0.000	0.051	1,5-hexadiene	1.040	0.024	0.559	0.481	0.051	0.000	0.015
benzene	0.845	0.000	0.730	0.654	0.064	0.000	0.020	hexane	1.158	0.000	0.469	0.178	0.011	0.000	0.001
biphenyl	1.523	0.000	0.477	0.622	0.086	0.000	0.058	hexyne	1.040	0.078	0.527	0.438	0.071	0.000	0.035
brombenzene	1.039	0.220	0.643	0.239	0.122	0.000	0.016	indan	1.223	0.056	0.546	0.585	0.055	0.000	0.044
1-butene	0.766	0.021	0.700	0.379	0.035	0.000	0.008	indene	1.167	0.068	0.596	0.629	0.065	0.000	0.039
2-butyne	0.700	0.000	0.763	0.016	0.084	0.000	0.028	isobutane	0.821	0.007	0.624	0.162	0.009	0.000	0.000
butadiene	0.709	0.000	0.851	0.597	0.042	0.000	0.009	isubutene	0.764	0.034	0.701	0.484	0.035	0.000	0.009
butane	0.825	0.000	0.624	0.177	0.009	0.000	0.000	isopropylbenzene	1.329	0.029	0.476	0.603	0.069	0.000	0.041
cis-butene	0.761	0.009	0.713	0.499	0.026	0.000	0.007	methylcyclohexane	1.180	0.009	0.493	0.130	0.011	0.000	0.001
trans-butene	0.768	0.000	0.708	0.492	0.026	0.000	0.007	methylcyclopentane	1.026	0.007	0.558	0.284	0.010	0.000	0.001
chlorobenzene	0.993	0.247	0.639	0.285	0.120	0.000	0.020	naphthalene	1.243	0.000	0.598	0.370	0.067	0.000	0.038
chloroform	0.819	0.121	0.598	0.333	0.054	0.198	0.004	neopentane	0.973	0.000	0.546	0.206	0.010	0.000	0.001
1,3-cyclohexadiene	0.907	0.029	0.671	0.576	0.048	0.000	0.017	nonane	1.657	0.010	0.342	0.185	0.013	0.000	0.002
1,4-cyclohexadiene	0.907	0.000	0.667	0.585	0.038	0.000	0.015	1-nonene	1.600	0.054	0.361	0.299	0.025	0.000	0.024
cyclohexane	1.017	0.000	0.567	0.115	0.000	0.000	0.001	octane	1.493	0.000	0.375	0.185	0.012	0.000	0.002
cyclohexene	0.962	0.017	0.617	0.449	0.020	0.000	0.012	1-octene	1.430	0.046	0.402	0.291	0.024	0.000	0.021
cyclopentane	0.861	0.000	0.657	0.292	0.009	0.000	0.001	1-pentyne	0.874	0.060	0.610	0.463	0.084	0.000	0.027
cyclopropane	0.564	0.000	0.910	0.303	0.030	0.000	0.002	1,4-pentadiene	0.872	0.004	0.665	0.401	0.047	0.000	0.010
dibromomethane	0.752	0.132	0.741	0.413	0.087	0.000	0.008	pentane	0.990	0.006	0.537	0.182	0.010	0.000	0.001
1,2-dichlorobenzene	1.135	0.362	0.562	0.410	0.107	0.000	0.031	phenanthrene	1.644	0.003	0.496	0.384	0.077	0.000	0.072
1,3-dichlorobenzene	1.234	0.252	0.480	0.322	0.067	0.305	0.016	phenylacetylene	1.050	0.072	0.654	0.456	0.092	0.170	0.076
1,4-dichlorobenzene	1.166	0.000	0.540	0.261	0.097	0.000	0.029	phenylallene	1.226	0.010	0.595	0.620	0.066	0.000	0.035
dichloromethane	0.657	0.147	0.699	0.409	0.113	0.158	0.008	phenylcyclopropane	1.239	0.048	0.535	0.615	0.078	0.000	0.042
1,2-dichloropropane	0.973	0.362	0.531	0.434	0.168	0.154	0.037	phenylpropene	1.284	0.000	0.000	0.605	0.068	0.000	0.045
difluoromethane	0.352	0.071	0.969	0.456	0.083	0.000	0.006	propane	0.657	0.004	0.748	0.185	0.008	0.000	0.000
1,2-dimethylbenzene	1.164	0.050	0.547	0.589	0.054	0.000	0.038	propene	0.600	0.017	0.849	0.517	0.039	0.000	0.006
1,3-dimethylbenzene	1.175	0.000	0.000	0.601	0.054	0.000	0.038	propyne	0.533	0.034	0.900	0.224	0.102	0.000	0.010
1,4-dimethylbenzene	1.178	0.008	0.541	0.595	0.054	0.000	0.037	styrene	1.114	0.013	0.608	0.620	0.074	0.000	0.031
2,2-dimethylbutane	1.125	0.004	0.485	0.210	0.011	0.000	0.001	*tert*-butylbenzene	1.471	0.035	0.434	0.583	0.061	0.000	0.048
2,3-dimethylbutane	1.130	0.100	0.482	0.144	0.011	0.000	0.001	*tert*-butylcyclohexane	1.627	0.010	0.369	0.150	0.000	0.000	0.002
ethane	0.489	0.000	0.935	0.252	0.007	0.000	0.000	tetrachloromethane	0.972	0.000	0.535	0.299	0.020	0.000	0.001
ethene	0.427	0.000	1.086	0.282	0.052	0.000	0.003	tetrafluoromethane	0.397	0.000	0.769	0.286	0.006	0.000	0.000
ethyne	0.367	0.000	1.058	0.382	0.115	0.305	0.005	toluene	1.009	0.026	0.623	0.613	0.060	0.000	0.028
ethylbenzene	1.175	0.027	0.536	0.617	0.065	0.000	0.034	1,1,1-trichloroethane	0.969	0.223	0.541	0.488	0.079	0.000	0.011
ethylcyclohexane	1.341	0.005	0.438	0.120	0.000	0.000	0.001	trichloroethene	0.919	0.118	0.588	0.364	0.058	0.164	0.005
fluorobenzene	0.854	0.139	0.714	0.428	0.065	0.000	0.019	trichlorofluoromethane	0.821	0.029	0.568	0.293	0.018	0.000	0.001
heptane	1.326	0.008	0.416	0.183	0.012	0.000	0.001	trifluoromethane	0.373	0.064	0.863	0.408	0.055	0.162	0.002

^{a} In $Å^3 \times 10^{-2}$. ^{b} In $D \times Å^3 \times 10^{-2}$. ^{c} In 6.755×10^2 bohr^{-3}. ^{d} In $Å^6 \cdot kJ^2 \cdot mol^{-2} \times 10^{-5}$.

TABLE A2: Quantum Chemical Descriptors To Correlate log P_{ow} Values of Hydrogen-Bonding Solutes

solute	V^a	μV^b	$\alpha/V^{2\ c}$	$U_{S,\%}$	I_-	$\sum U_{min}$	$\sum U_{max}$	$\overline{U_-^2}V^{2\ d}$	N_{ind}	O_{ind}
acetaldehyde	0.497	0.149	0.907	0.396	0.134	−0.226	0.000	10.10	0	1
acetamide	0.596	0.250	0.774	0.480	0.185	−0.287	0.415	15.46	1	1
acetic acid	0.553	0.104	0.805	0.447	0.104	−0.226	0.272	8.47	0	1
acetone	0.672	0.226	0.735	0.408	0.148	−0.253	0.000	15.77	0	1
acetonitrile	0.491	0.199	0.954	0.425	0.189	−0.236	0.000	17.64	1	0
acetophenone	1.186	0.425	0.552	0.487	0.120	−0.257	0.000	9.75	0	1
acrolein	0.604	0.213	0.876	0.417	0.117	−0.230	0.000	9.70	0	1
acrylonitrile	0.598	0.254	0.891	0.450	0.155	−0.236	0.000	11.46	1	0
acrylophenone	1.274	0.367	0.530	0.558	0.113	−0.229	0.000	7.85	0	
adenine	1.111	0.757	0.603	0.501	0.211	−0.565	0.683	14.38	1	0
aminocyclohexane	1.124	0.170	0.513	0.320	0.064	−0.352	0.000	9.14	1	0
1-aminonaphthalene	1.356	0.197	0.543	0.637	0.093	−0.270	0.166	7.27	1	0
2-aminonaphthalene	1.367	0.219	0.544	0.621	0.105	−0.267	0.334	8.21	1	0
m-aminophenol	1.011	0.259	0.619	0.572	0.131	−0.454	0.608	8.00	1	1
o-aminophenol	1.002	0.068	0.630	0.631	0.110	−0.175	0.447	6.59	1	1
p-aminophenyl methyl sulfone	1.441	1.182	0.469	0.499	0.252	−0.524	0.497	36.16	1	1

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5727**

TABLE A2 (Continued)

solute	V^a	μV^b	$\alpha/V^{2\ c}$	$U_{S,\%}$	I_-	ΣU_{min}	ΣU_{max}	$\overline{U^2_-}V^{2\ d}$	N_{ind}	O_{ind}
3-amino-l-propene	0.710	0.121	0.734	0.549	0.118	−0.353	0.000	9.27	1	0
2-aminopyridine	0.922	0.191	0.685	0.598	0.125	−0.292	0.184	7.37	1	0
aniline	0.965	0.149	0.662	0.623	0.098	−0.254	0.341	7.21	1	0
anisole	1.097	0.149	0.575	0.611	0.094	−0.192	0.000	5.91	0	1
benzaldehyde	1.012	0.363	0.639	0.476	0.111	−0.240	0.000	7.59	0	1
benzamide	1.109	0.407	0.570	0.518	0.137	−0.252	0.378	9.81	1	1
benzoic acid	1.064	0.244	0.595	0.550	0.103	−0.218	0.260	6.25	0	1
benzonitrile	0.998	0.484	0.668	0.399	0.120	−0.241	0.000	16.15	1	0
benzophenone	1.680	0.548	0.428	0.552	0.117	−0.236	0.000	9.30	0	1
butanal	0.826	0.230	0.617	0.354	0.109	−0.228	0.000	12.79	0	1
1-butanol	0.885	0.155	0.578	0.451	0.085	−0.260	0.204	8.58	0	1
tert-butanol	0.877	0.150	0.580	0.421	0.094	−0.261	0.389	8.87	0	1
butanone	0.825	0.257	0.614	0.406	0.118	−0.243	0.000	18.28	0	1
butanthiol	1.047	0.208	0.521	0.304	0.102	0.000	0.000	6.51	0	0
3-butenenitrile	0.763	0.315	0.703	0.403	0.148	−0.244	0.000	12.14	1	0
butylamine	0.932	0.113	0.557	0.422	0.058	−0.344	0.000	6.63	1	0
butyric acid	0.882	0.147	0.570	0.416	0.085	−0.227	0.248	11.24	0	1
butyronitrile	0.822	0.350	0.629	0.349	0.154	−0.248	0.000	15.59	1	0
cyclohexanol	1.075	0.207	0.528	0.303	0.073	−0.258	0.423	15.86	0	1
cyclohexanone	1.022	0.347	0.562	0.368	0.101	−0.248	0.000	21.04	0	1
cyclohexenone	0.966	0.392	0.623	0.415	0.128	−0.256	0.000	15.16	0	1
cyclopropylamine	0.672	0.079	0.777	0.520	0.074	−0.314	0.000	3.93	1	0
dibutyl ether	1.571	0.195	0.352	0.409	0.050	−0.237	0.000	6.94	0	1
3,5-dichloropyridine	1.129	0.033	0.552	0.334	0.032	−0.224	0.000	2.55	1	0
diethyl ether	0.906	0.123	0.560	0.492	0.076	−0.245	0.000	5.85	0	1
diisopropyl ether	1.224	0.162	0.433	0.458	0.055	−0.239	0.000	6.04	0	1
1,2-dimethoxybenzene	1.316	0.097	0.473	0.659	0.126	−0.267	0.000	12.74	0	1
dimethyl ether	0.566	0.092	0.784	0.501	0.098	−0.251	0.000	7.39	0	1
dimethylacetamide	0.924	0.379	0.566	0.451	0.154	−0.297	0.000	18.77	1	1
dimethylamine	0.608	0.062	0.769	0.486	0.062	−0.326	0.000	7.46	1	0
N,N-dimethylaniline	1.282	0.115	0.492	0.620	0.079	−0.255	0.000	6.53	1	0
dimethylformamide	0.770	0.327	0.655	0.417	0.158	−0.290	0.000	18.49	1	1
dimethyl sulfone	0.783	0.395	0.635	0.428	0.230	−0.417	0.000	24.50	0	1
dimethyl sulfoxide	0.762	0.402	0.730	0.435	0.253	−0.362	0.000	33.71	0	1
2,4-dinitrophenol	1.280	0.484	0.539	0.239	0.090	−0.179	0.698	9.43	1	1
dioxane	0.835	0.000	0.639	0.360	0.082	−0.474	0.000	15.04	0	1
diphenyl sulfide	1.726	0.345	0.396	0.510	0.131	0.000	0.000	5.23	0	0
dipropyl ether	1.237	0.142	0.432	0.433	0.056	−0.236	0.000	6.81	0	1
ethanol	0.552	0.099	0.828	0.508	0.119	−0.259	0.210	6.87	0	1
ethyl vinyl ether	0.845	0.146	0.627	0.524	0.101	−0.207	0.000	6.04	0	1
ethylamine	0.599	0.074	0.785	0.470	0.077	−0.344	0.000	5.86	1	0
ethylene oxide	0.474	0.125	0.993	0.459	0.138	−0.241	0.000	12.47	0	1
ethyl propionate	1.059	0.191	0.487	0.464	0.093	−0.239	0.000	12.48	0	1
formamide	0.432	0.169	0.971	0.476	0.177	−0.260	0.415	10.41	1	1
formanilide	1.122	0.437	0.574	0.424	0.117	−0.258	0.209	9.94	1	1
formic acid	0.388	0.061	1.010	0.416	0.112	−0.225	0.449	5.63	0	1
furan	0.653	0.062	0.832	0.323	0.073	−0.170	0.000	4.03	0	1
1-hexanol	1.217	0.235	0.441	0.298	0.066	−0.258	0.223	13.36	0	1
2-hexanol	1.214	0.210	0.441	0.315	0.066	−0.253	0.224	10.11	0	1
3-hexanol	1.208	0.199	0.445	0.388	0.066	−0.246	0.222	7.11	0	1
1-hexyn-5-one	1.040	0.294	0.521	0.417	0.122	−0.224	0.000	9.38	0	1
indole	1.112	0.227	0.623	0.651	0.105	0.000	0.235	6.98	1	0
methacrylic acid	0.816	0.142	0.649	0.478	0.099	−0.214	0.249	7.08	0	1
methanol	0.383	0.074	1.025	0.500	0.142	−0.261	0.222	7.23	0	1
methyl acetate	0.731	0.146	0.649	0.470	0.111	−0.239	0.000	11.46	0	1
methylamine	0.432	0.056	0.965	0.508	0.099	−0.349	0.000	8.16	1	0
methyl benzoate	1.243	0.210	0.466	0.564	0.100	−0.219	0.000	7.30	0	1
methyl *tert*-butyl ether	1.039	0.144	0.503	0.480	0.071	−0.242	0.000	7.69	0	1
methyl formate	0.563	0.111	0.780	0.441	0.120	−0.224	0.000	8.41	0	1
2-methylfuran	0.829	0.057	0.679	0.588	0.073	0.000	0.000	3.82	0	1
1-methylimidazole	0.827	0.345	0.675	0.497	0.164	−0.347	0.000	11.87	1	0
2-methyl-2-nitropropane	0.985	0.407	0.539	0.432	0.119	−0.209	0.000	12.63	1	1
2-methylphenol	1.064	0.133	0.590	0.584	0.083	−0.187	0.284	4.00	0	1
methyl phenyl sulfone	1.295	0.724	0.486	0.491	0.193	−0.423	0.000	21.77	0	1
4-methylpyridine	0.966	0.260	0.634	0.477	0.088	−0.310	0.000	6.30	1	0
N-methylpyrrole	0.871	0.183	0.662	0.593	0.103	0.000	0.000	7.08	1	0
2-naphthol	1.318	0.228	0.558	0.614	0.103	−0.201	0.282	6.46	0	1
neopentanol	1.023	0.180	0.514	0.376	0.071	−0.251	0.231	8.19	0	1
p-nitroaniline	1.135	0.900	0.613	0.477	0.149	−0.245	0.517	13.08	1	1
nitrobenzene	1.032	0.551	0.634	0.419	0.112	−0.211	0.000	11.73	1	1
nitrobutane	1.010	0.431	0.526	0.325	0.121	−0.205	0.000	14.50	1	1
nitroethane	0.677	0.274	0.730	0.371	0.140	−0.202	0.000	9.58	1	1
nitromethane	0.515	0.204	0.895	0.411	0.129	−0.195	0.000	7.12	1	1

TABLE A2 (Continued)

solute	V^a	μV^b	$\alpha/V^{2\ c}$	$U_{S,\%}$	I_-	$\sum U_{min}$	$\sum U_{max}$	$\overline{U_-^2}V^{2\ d}$	N_{ind}	O_{ind}
nitropropane	0.844	0.352	0.611	0.340	0.129	−0.204	0.000	11.81	1	1
p-nitrotoluene	1.193	0.650	0.547	0.377	0.109	−0.206	0.000	10.14	1	1
octanol	1.550	0.268	0.359	0.402	0.062	−0.254	0.223	8.68	0	1
pentachlorphenole	1.673	0.345	0.398	0.346	0.065	0.000	0.315	2.26	0	1
pentan-2-one	0.992	0.320	0.529	0.382	0.110	−0.251	0.000	21.82	0	1
phenole	0.905	0.133	0.680	0.593	0.095	−0.192	0.273	4.91	0	1
phenyl trifluoromethyl sulfone	1.385	0.907	0.453	0.425	0.153	−0.229	0.000	26.96	0	1
piperidine	0.963	0.112	0.591	0.359	0.049	−0.333	0.000	7.87	1	0
propanal	0.659	0.189	0.737	0.396	0.122	−0.227	0.000	10.31	1	1
1-propanol	0.717	0.136	0.677	0.364	0.093	−0.256	0.225	12.56	0	1
2-propanol	0.717	0.126	0.678	0.445	0.093	−0.253	0.208	8.35	0	1
propanethiol	0.879	0.171	0.605	0.325	0.113	0.000	0.000	5.73	0	0
2-propen-l-ol	0.659	0.121	0.768	0.553	0.122	−0.250	0.222	6.73	0	1
2-propyn-l-ol	0.602	0.125	0.792	0.577	0.178	−0.238	0.241	7.53	0	1
propionic acid	0.714	0.114	0.669	0.438	0.093	−0.221	0.255	8.50	0	1
propionitrile	0.655	0.271	0.751	0.385	0.170	−0.246	0.000	12.69	1	0
pyridine	0.792	0.182	0.768	0.502	0.089	−0.300	0.000	9.00	1	0
pyridine N-oxide	0.859	0.520	0.741	0.432	0.195	−0.360	0.000	51.34	1	1
pyrimidine	0.754	0.179	0.772	0.451	0.113	−0.538	0.000	5.74	1	0
pyrrole	0.696	0.131	0.805	0.646	0.117	0.000	0.232	5.47	1	0
pyrrolidine	0.817	0.098	0.674	0.483	0.054	−0.333	0.000	9.20	1	0
quinoline	1.212	0.262	0.601	0.616	0.088	−0.292	0.000	5.80	1	0
quinuclidine	1.161	0.132	0.520	0.572	0.032	−0.327	0.000	8.75	1	0
THF	0.772	0.150	0.695	0.361	0.070	−0.264	0.000	16.17	0	1
thiophenol	1.064	0.204	0.607	0.517	0.103	0.000	0.000	3.10	0	0
toluidene	1.117	0.155	0.574	0.639	0.095	0.000	0.377	6.06	1	0
2,2,2-trifluoroethanol	0.607	0.225	0.667	0.469	0.117	−0.198	0.281	5.34	0	1
trifluoroethylamine	0.650	0.214	0.641	0.451	0.105	−0.257	0.340	5.11	1	0
p-trifluoromethylaniline	1.167	0.601	0.521	0.520	0.119	0.000	0.458	4.95	1	0
trimethylamine	0.774	0.067	0.638	0.525	0.044	−0.306	0.000	6.91	1	0
urea	0.532	0.251	0.797	0.545	0.233	−0.302	0.598	13.75	1	1
vinyl acetate	0.839	0.149	0.627	0.461	0.092	−0.209	0.000	6.36	0	1

^{a} In $\text{Å}^3 \times 10^{-2}$. ^{b} In $\text{D} \times \text{Å}^3 \times 10^{-2}$. ^{c} In 6.755×10^2 bohr^{-3}. ^{d} In $\text{Å}^6 \cdot kJ^2 \cdot \text{mol}^{-2} \times 10^{-5}$.

TABLE A3: Quantum Chemical Descriptors To Predict log P_{ow} Values of Solutes Not Included in the Fitting Base

solute	V^a	μV^b	$\alpha/V^{2\ c}$	$U_{S,\%}$	I_-	$\sum U_{min}$	$\sum U_{max}$	$\overline{U_-^2}V^{2\ d}$	N_{ind}	O_{ind}
1,2-ethanediol	0.604	0.166	0.705	0.476	0.132	−0.476	0.251	9.091	0	1
1,3-propanediol	0.773	0.207	0.621	0.366	0.141	−0.492	0.250	11.861	0	1
ethyl acetate	0.851	0.138	0.578	0.486	0.083	−0.235	0.000	10.507	0	1
butyrolactone	0.779	0.389	0.677	0.422	0.177	−0.261	0.000	18.906	0	1
glycine6 ionic	0.637	0.642	0.752	0.466	0.367	−0.391	0.355	39.078	1	1
alanine ionic	0.800	0.794	0.627	0.458	0.349	−0.395	0.352	42.235	1	1
1,2-dibromobenzene	1.237	0.366	0.562	0.415	0.122	0.000	0.000	2.935	0	0
1,3-dibromobenzene	1.258	0.245	0.553	0.324	0.079	0.000	0.000	1.372	0	0
1,4-dibromobenzene	1.246	0.000	0.569	0.222	0.056	0.000	0.000	1.546	0	0
2-chloroaniline	1.105	0.254	0.572	0.555	0.126	−0.242	0.185	5.140	1	0
3-chloroaniline	1.113	0.334	0.568	0.545	0.148	−0.240	0.187	4.780	1	0
4-chloroaniline	1.114	0.372	0.569	0.530	0.158	−0.242	0.190	5.254	1	0
2-hydroxybenzoic acid	1.111	0.242	0.527	0.525	0.084	−0.365	0.306	4.962	0	1
3-hydroxybenzoic acid	1.123	0.111	0.560	0.510	0.085	−0.369	0.541	5.792	0	1
4-hydroxybenzoic acid	1.123	0.174	0.566	0.521	0.106	−0.214	0.574	6.370	0	1
2-nitrophenol	1.080	0.343	0.549	0.400	0.083	−0.177	0.000	6.603	1	1
3-nitrophenol	1.088	0.386	0.588	0.380	0.094	−0.191	0.323	8.001	1	1
4-nitrophenol	1.088	0.568	0.595	0.394	0.104	−0.206	0.349	9.990	1	1

^{a} In $\text{Å}^3 \times 10^{-2}$. ^{b} In $\text{D} \times \text{Å}^3 \times 10^{-2}$. ^{c} In 6.755×10^2 bohr^{-3}. ^{d} In $\text{Å}^6 \cdot kJ^2 \cdot \text{mol}^{-2} \times 10^{-5}$.

References and Notes

(1) Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.

(2) Leo, A. J. *Chem. Rev.* **1993**, *93*, 1281.

(3) (a) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley: New York, 1979. (b) Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR - Hydrophobic, Electronic, and Steric Constants*; ACS Professional Reference Book; American Chemical Society: Washington, DC, 1995.

(4) (a) Leo, A. In *Methods in Enzymology*; Langone, J., Ed.; Academic Press: San Diego, 1991; Vol. 202, p 544. (b) Leo, A. In *Comprehensive Medicinal Chemistry*; Hansch, C., Ed.; Pergamon Press: Oxford, U.K., 1990; Vol. 4, p 315.

(5) Wienke, G.; Gmehling, J. *Tox. Environ. Chem.* **1998**, *65*, 57.

(6) (a) Kamlet, M. J.; Abboud, J. M.; Abraham, M. H.; Taft, R. W. *J. Org. Chem.* **1983**, *48*, 2877. (b) Taft, R. W.; Abboud, J. M.; Kamlet, M. J.; Abraham. M. H. *J. Solution Chem.* **1985**, *3*, 153. (c) Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Marcus, Y.; Taft, R. W. *J. Phys. Chem.* **1988**, *92*, 5244.

(7) Meyer, P.; Maurer, G. *Ind. Eng. Chem. Res.* **1993**, *32*, 2105.

(8) Meyer, P.; Maurer, G. *Ind. Eng. Chem. Res.* **1995**, *34*, 373.

(9) Klopman, G.; Iroff, D. *J. Comput. Chem.* **1981**, *2*, 157.

(10) (a) Bodor, N.; Babanyi, Z.; Wong, C. *J. Am. Chem. Soc.* **1989**, *111*, 3783. (b) Bodor, N.; Huang, M. *J. Pharm. Sci.* **1992**, *81*, 272.

(11) Kantola, A.; Villar, H. O.; Loew, G. H. *J. Comput. Chem.* **1991**, *12*, 681.

(12) Sasaki, Y.; Kubodera, H.; Matuszaki, T.; Umeyama, H. *J. Pharmacobio-Dyn.* **1991**, *14*, 207.

Octanol/Water Partition Coefficients

J. Phys. Chem. B, Vol. 103, No. 27, 1999 **5729**

(13) Famini, G. R.; Penski, C. A.; Wilson, L. Y. *J. Phys. Org. Chem.* **1992**, *5*, 395.

(14) Brinck, T.; Murray, J. S.; Politzer, P. *J. Org. Chem.* **1993**, *58*, 7070.

(15) Haeberlein, M.; Brinck, T. *J. Chem. Soc., Perkin Trans.* **1997**, 289.

(16) (a) Abraham, M. H. *Chem. Soc. Rev.* **1993**, 73. (b) Abraham, M. H.; Andonian-Haftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1777.

(17) Buckingham, A. D.; Fowler, P. W. *Can. J. Chem.* **1985**, *63*, 2018.

(18) Lin, S.; Dykstra, C. E. *Chem. Phys.* **1986**, *107*, 343.

(19) Legon, A. C.; Millen, D. J. *Acc. Chem. Res.* **1987**, *20*, 39 and literature cited therein.

(20) Baiocchi, F. A.; Reiher, W.; Klemperer, W. *J. Chem. Phys.* **1983**, *79*, 6428.

(21) van der Avoird, A.; Wormer, P. E. S.; Mulder, F.; Berns, R. M. *Top. Curr. Chem.* **1980**, *93*, 1.

(22) Murray, J. S.; Ranganathan, S.; Politzer, P *J. Org. Chem.* **1991**, *56*, 3734.

(23) Murray, J. S.; Politzer, P. *J. Org. Chem.* **1991**, *56*, 6715.

(24) Murray, J. S.; Politzer, P. *J. Chem. Res.* **1992**, 1100.

(25) Hagelin, H.; Murray, J. S.; Brinck, T.; Berthelot, M.; Politzer, P. *Can. J. Chem.* **1995**, *73*, 483.

(26) (a) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202. (b) Chan, M. C.; Block, P. A.; Miller, R. E. *J. Chem. Phys.* **1995**, *102*, 3993.

(27) (a) Sun, S.; Bernstein, E. R. *J. Phys. Chem.* **1996**, *100*, 13348. (b) Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 11217. (c) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790. (d) Hobza, P.; Špirko, V.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem. A* **1998**, *102*, 2501. (e) Neusser, H. J.; Krause, H. *Chem. Rev.* **1994**, *94*, 1829.

(28) Werner, H.-J.; Meyer, W. *Mol. Phys.* **1976**, *31*, 855.

(29) (a) Diercksen, G. H. F.; Sadlej, A. J. *J. Chem. Phys.* **1981**, *75*, 1253. (b) Zeiss, G. D.; Scott, W. R.; Suzuki, N.; Chong, D. P.; Langhoff, S. R. *Mol. Phys.* **1979**, *37*, 1543.

(30) (a) Bader, R. F. W.; Carroll, M. T.; Cheeseman, J. R.; Chang, C. *J. Am. Chem. Soc.* **1987**, *109*, 7968. (b) Bader, R. F. W.; Henneker, W. H.; Cade, P. E. *J. Chem. Phys.* **1967**, *46*, 3341.

(31) Murray, J. S.; Brinck, T.; Grice, M. E.; Politzer, P. *THEOCHEM* **1992**, *256*, 29.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 94, Revision C.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(33) (a) Binkley, J. S.; Pople J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939. (b) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1983**, *104*, 2797. (c) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; DeFrees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039.

(34) Sangster, J. *J. Phys. Chem. Ref. Data* **1989**, *18*, 1111.

(35) As a further restriction only maxima are counted that have a minimal distance of 6 au to any other minimum.

(36) For the zwitterionic structure of the amino acids no stationary point could be obtained at Hartree−Fock with the supplemented 3-21G* basis set. Thus, these geometries were optimized at HF/6-31G*.