Characterization of Lipophilicity Scales Using Vectors from Solvation Energy Descriptors

YASUSHI ISHIHAMA* AND NAOKI ASAKAWA

Contribution from *Analytical Research Laboratories, Eisai Co., Ltd., 5-1-3 Tokodai, Tsukuba, Ibaraki 300-2635, Japan.*

Received April 12, 1999. Final revised manuscript received September 29, 1999. Accepted for publication October 5, 1999.

Abstract □ Lipophilicity scales were characterized by an approach using vectors provided from solvation energy descriptors (SED) of solutes such as an excess molar refraction, the dipolarity/polarizability, the hydrogen-bond acidity, the basicity, and the McGowan characteristic volume. The five components of the SED vector were obtained from the coefficients of the five SED terms of the linear solvation energy relationship (LSER) equation for the lipophilicity scales. The analogy between two lipophilicity scales was expressed as the angle between the two SED vectors, while the difference in the contribution of the five independent SEDs to these two lipophilicity scales was quantified by the difference of the unit vectors of the SED vectors. These approaches were applied to several lipophilicity scales measured using microemulsions, micelles, an immobilized artificial membrane column, and an octanol–water system. As a result, the quantitative classification of these scales was successfully carried out, and the difference in the scales was well characterized. In addition, this vector approach was extended to the estimation of the contribution of each constituent of the microemulsions to the lipophilicity scale. Furthermore, some biological parameters such as skin permeability and the distribution between blood and brain could be predicted by the summation of the SED vectors obtained from the chromatographic systems. These results suggest that complex biological systems can be expressed quantitatively by simple chemical models with their SED vectors.

Introduction

In the passage of drugs throughout the body, the permeability of cell membranes is quite important for the prediction of their in vivo activities from their in vitro results. Therefore, the lipophilicity of drugs such as the logarithm of the partition coefficients between 1-octanol and water (log P_{ow}) has been used as a parameter for the structure optimization of the drug candidates. Because recent developments in combinatorial chemistry allowed the synthesis of a large number of compounds as drug candidates, the demand for high-throughput measurement of biologically appropriate lipophilicity is steadily increasing.

In our previous study, the lipophilicity scale by electrokinetic chromatography (EKC) with the microemulsion of sodium dodecyl sulfate (SDS), 1-butanol, *n*-heptane, and a buffer provided the excellent correlation with log P_{ow} for neutral compounds with various hydrogen-bonding abilities.[1] In quantitative structure−activity relationship studies (QSAR) for some bioactivities of the drugs, their lipophilicity from the microemulsion provided a better correlation with their bioactivities than other lipophilicity scales.[2,3] In addition, it was suggested that the lipophilicity scales from the microemulsions could be designed by selecting the constituents and their concentration.[4]

On the other hand, other lipophilicity scales such as the logarithm of the capacity factors (log k') in various chromatographic systems including HPLC[5−7] and micellar EKC (MEKC)[8−15] have been developed, and they provided diverse and unique properties as lipophilicity scales. In optimizing the structure of drug candidates, these lipophilicity scales often provided different results, although they have viability for high-throughput analysis. Therefore, selecting the lipophilicity scales suitable for predicting the bioactivities of drug candidates is quite important for optimizing their structures. For this purpose, the characterization and classification of these scales were required.

The correlation coefficients of the linear relationship between two scales have been often used for the comparison between these two scales. It is, however, well-known that the correlation coefficient strongly depends on the test set of solutes. Alternatively, the linear solvation energy relationship (LSER) analysis has been used for the characterization of the retention behaviors of solutes in many chromatographic media, and the quantitative prediction of the retention times of the solutes from their structure was performed.[16−18] Recently, this approach was applied to MEKC to classify the separation selectivity of the micelles[19−22] and was also used for evaluating the correlation between log P_{ow} and migration index (MI) measured by microemulsion EKC (MEEKC).[4]

The general equation of LSER based on the solvation energy descriptors (SED) of solutes is as follows:

$$\log \mathrm{SP} = c + rR_2 + s\pi_2^{\mathrm{H}} + a\sum\alpha_2^{\mathrm{H}} + b\sum\beta_2 + vV_x \quad (1)$$

where log SP is the dependent variable, i.e., the lipophilicity scales (LS) such as log k' and MI in this case, and the independent variables are solute descriptors as follows: R_2 is an excess molar refraction, π_2^{H} is the solute dipolarity/polarizability, $\sum\alpha_2^{\mathrm{H}}$ and $\sum\beta_2$ are the solute hydrogen-bond acidity and basicity, and Vx is the McGowan characteristic volume in units of $cm^3\ mol^{-1}/100$.[17] The obtained coefficients of eq 1 were used for characterization of the lipophilicity scales as well as the prediction of the separation selectivity in the chromatographic media. To classify these scales, the ratios of the coefficients such as r/v, a/v, b/v, and s/v were calculated and compared with those from another scale. This approach was quite useful to judge the analogy between two scales. Unfortunately, however, it was difficult to simultaneously compare the set of the coefficients or coefficient ratios between the plural scales. For example, Abraham et al. reported on the following scales:[4]

$(r/v, s/v, a/v, b/v) =$
$(0.12, -0.23, 0.01, -0.94)$ for log P_{ow}

$(r/v, s/v, a/v, b/v) =$
$(0.18, -0.24, 0.00, -0.87)$ for log P (pentanol−water)

$(r/v, s/v, a/v, b/v) =$
$(0.09, -0.23, -0.02, -0.92)$ for log k' (microemulsion)

It was impossible to judge which was similar to octanol−water, pentanol−water, or the microemulsion. Therefore, an approach for analyzing these scales simultaneously was required.

* Corresponding author. Tel: (+81-298) 47-5763, fax: (+81-298) 47-5771, e-mail: y-ishihama@hhc.eisai.co.jp.

10.1021/js990119+ CCC: $18.00
Published on Web 11/04/1999

Table 1—System for Measurement of Lipophilicity Scales

system	scale	constituents
ME(SDS)	MI_{SDS}	SDS (1.44%)/1-butanol (6.49%)/ heptane (0.82%)/buffer[a]
ME(CTAC)	MI_{CTAC}	CTAC (1.44%)/1-butanol (6.49%)/ heptane (0.82%)/buffer[a]
ME(DTAC)	MI_{DTAC}	DTAC (1.44%)/1-butanol (6.49%)/ heptane (0.82%)/buffer[a]
MC(SDS)	$\log k'_{SDS}$	50 mM SDS in buffer[a]
MC(DTAC)	$\log k'_{DTAC}$	50 mM DTAC in buffer[a]
MC(S/B)	$\log k'_{S/B}$	25 mM SDS and 25 mM Brij 35 in buffer[a]
OW	$\log P_{OW}$	1-octanol and water
IAM	$\log k'_{IAM}$	IAM column with buffer[a] as the mobile phase

[a] buffer: 50 mM sodium phosphate and 100 mM sodium borate (pH 7.0).

Recently, Valko et al. reported on the characterization of the various HPLC columns using the gradient retention parameter named chromatographic hydrophobicity indices (CHI) and their SED coefficients.[23] They described that the principal component analysis and the nonlinear mapping technique provide an appropriate tool for comparison of various HPLC partition systems.

In this paper, we demonstrate that the lipophilicity scales could be characterized using the coefficients of the five SED terms of the LSER equation as vector components. The analogy of the lipophilicity scales was evaluated using the scalar product of the vectors, and the difference in the five independent factors which affect the lipophilicity scales was quantified by these unit vectors. This vector approach was also employed to obtain the actual structural information of the microemulsions from the vectors of their individual constituent system. Furthermore, it was applied to predict some biological systems according to the summation of plural vectors from simple chemical systems.

Experimental Section

Capillary electrophoresis was performed using P/ACE 2100 (Beckman, Fullerton, CA). For three microemulsions (ME) and three micelles (MC), EKC was used for the determination of the lipophilicity scales from these systems. The experimental details were described in previous papers.[1−3,8] In all cases, 50 mM phosphate−100 mM borate solution (pH 7.0) was used as the buffer. The ME and MC solutions employed are listed in Table 1. Uncoated fused silica capillary with 50-μm i.d. and 27-cm length (GL Sciences, Tokyo, Japan) was employed. The capillary was thermostated at 25 °C. The applied voltage was 7.5 kV, and the detection wavelength was 214 nm. The injection was performed by pressure (0.5 psi, 2 s). In the cases of the micellar systems, the values of log k' were used as lipophilicity scales, while in the cases of the microemulsions, the migration indexes (MI) were calculated from the log k' of test solutes and references.[1−3]

A Shimadzu LC-10A system (Kyoto, Japan) equipped with an SPD-10A UV detector (Shimadzu) was used for the measurement of the lipophilicity scale from immobilized artificial membrane (IAM) column (4.6-mm i.d., 100-mm length, Resis, Morten Grove, IL). In this case, log k' was used as a lipophilicity scale using a phosphate buffer at pH 7.0 (ionic strength: 0.05) as a mobile phase and UV 220 nm as a detection wavelength. All tested samples listed in Table 2 were purchased from Aldrich (Milwaukee, WI), Sigma (St. Louis, MO), Wako (Osaka, Japan), and Tokyo Kasei Kogyo (Tokyo, Japan).

The measured log P_{ow} values were obtained from the database of Mac-logP ver. 1.0.3 (BioByte Corp., Claremont, CA).

Methodology

To obtain the set of the five coefficients (r, s, a, b, v) of eq 1, a regression analysis is performed using the measured lipophilicity values and the SED values of solutes listed in Table 2. The SED coefficient vector of lipophilicity scale i (LS_i), $\vec{\omega}$, is defined as follows:

$$\vec{\omega}_i = (r_i, s_i, a_i, b_i, v_i) \tag{2}$$

The analogy between LS_i and LS_j is expressed as cos θ_{ij} between $\vec{\omega}_i$ and $\vec{\omega}_j$ as follows:

$$\cos\theta_{ij} = \frac{\vec{\omega}_i.\vec{\omega}_j}{|\vec{\omega}_i||\vec{\omega}_j|} = \frac{r_i r_j + s_i s_j + a_i a_j + b_i b_j + v_i v_j}{\sqrt{r_i^2 + s_i^2 + a_i^2 + b_i^2 + v_i^2}\sqrt{r_j^2 + s_j^2 + a_j^2 + b_j^2 + v_j^2}} \tag{3}$$

Table 2—Solutes Employed

sample name	solvation energy descriptor				
	R_2	π_2^H	$\Sigma\alpha_2^H$	$\Sigma\beta_2$	V_x
pyrimidine	0.606	1.00	0	0.65	0.634
pyrazine	0.629	0.95	0	0.61	0.634
4-methylpyrimidine	0.595	1.00	0	0.63	0.775
methylpyrazine	0.629	0.90	0	0.65	0.775
4,6-dimethylpyrimidine	0.580	1.00	0	0.65	0.916
ethylpyrazine	0.616	0.90	0	0.66	0.916
pyrrole	0.613	0.73	0.41	0.29	0.577
resorcinol	0.980	1.00	1.10	0.58	0.834
n-methylbenzamide	0.950	1.49	0.40	0.71	1.114
methyl 2-furoate	0.560	1.00	0	0.50	0.893
benzyl alcohol	0.803	0.87	0.39	0.56	0.916
1-methylpyrrole	0.559	0.79	0	0.31	0.718
acetanilide	0.870	1.40	0.50	0.67	1.113
p-methoxyphenol	0.900	1.17	0.57	0.48	0.975
furan	0.369	0.53	0	0.13	0.536
p-nitroaniline	1.220	1.91	0.42	0.38	0.990
phenol	0.805	0.89	0.60	0.30	0.775
2,5-dimethylpyrrole	0.639	0.70	0.35	0.44	0.859
benzaldehyde	0.820	1.00	0	0.39	0.873
quinoxaline	1.300	1.22	0	0.59	1.003
ethyl 2-furoate	0.560	1.00	0	0.50	1.033
benzonitrile	0.742	1.11	0	0.33	0.871
acetophenone	0.818	1.01	0	0.48	1.014
thiophene	0.687	0.57	0	0.15	0.641
2-methylfuran	0.372	0.50	0	0.14	0.677
nitrobenzene	0.871	1.11	0	0.28	0.891
p-cresol	0.820	0.87	0.57	0.31	0.916
o-cresol	0.840	0.86	0.52	0.30	0.916
m-cresol	0.822	0.88	0.57	0.34	0.916
p-nitroanisole	0.970	1.29	0	0.40	1.090
anisole	0.708	0.75	0	0.29	0.916
methyl benzoate	0.733	0.85	0	0.46	1.073
benzene	0.610	0.52	0	0.14	0.716
indole	1.200	1.12	0.44	0.31	0.946
propiophenone	0.804	0.95	0	0.51	1.155
p-nitrotoluene	0.870	1.11	0	0.28	1.032
p-chlorophenol	0.915	1.08	0.67	0.20	0.898
2-ethylfuran	0.361	0.50	0	0.14	0.818
p-ethylphenol	0.800	0.90	0.55	0.36	1.057
2-methylindole	1.200	1.05	0.44	0.37	1.087
3-methylindole	1.200	1.06	0.44	0.35	1.087
1-methylindole	1.206	1.03	0	0.37	1.087
butyrophenone	0.797	0.95	0	0.51	1.296
benzofuran	0.888	0.83	0	0.15	0.905
toluene	0.601	0.52	0	0.14	0.857
2-naphthol	1.520	1.08	0.61	0.40	1.144
chlorobenzene	0.718	0.65	0	0.07	0.839
p-propylphenol	0.793	0.88	0.55	0.37	1.198
ethylbenzene	0.613	0.51	0	0.15	0.998
naphthalene	1.340	0.92	0	0.20	1.085
propylbenzene	0.604	0.50	0	0.15	1.139
butylbenzene	0.600	0.51	0	0.15	1.280
anthrathene	2.290	1.34	0	0.28	1.454

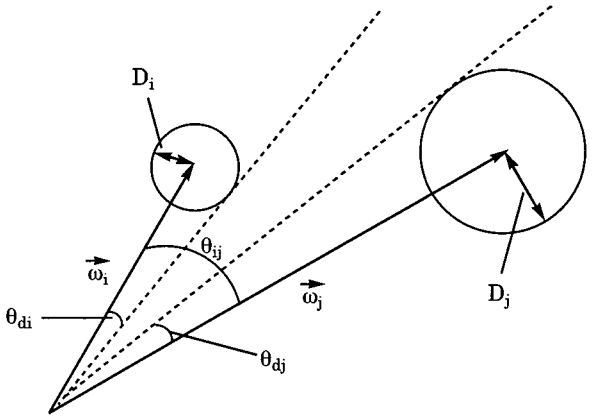

Figure 1—Two-dimensional model space of SED vectors.

Thus, as the correlation is higher, the value of cos θ_{ij} becomes closer to 1. When the analogy of LS_j ($j = 1, 2, ...$) to LS_i is examined, the analogy ranking of LS_j ($j = 1, 2,...$) to LS_i is established according to cos θ_{ij}. However, to judge the analogy between LS_i and LS_j, cos θ_{ij} is insufficient and the deviation of the vector should be also considered. In this study, D, which is a 95% confidence level of the coefficients of SED, is used as the deviation of each vector as follows:

$$D = TINV(0.05, N) \times \mathrm{SE} \quad (4)$$

where TINV is the inverse of the Student's t-distribution for the specified degrees of freedom, N, and SE is the average of the standard errors of the coefficients of SED.

Therefore, the analogy between two systems can be evaluated by the following equation:

$$J = \cos\theta_{ij} - \cos(\theta_{di} + \theta_{dj}) \quad (5)$$

where θ_{di} and θ_{dj} are the angles of the deviations of $\vec{\omega}_i$ and $\vec{\omega}_j$, respectively, as shown in Figure 1. The value of $\cos(\theta_{di} + \theta_{dj})$ can be calculated as follows:

$$\cos(\theta_{di} + \theta_{dj}) = \sqrt{\left(1 - \frac{D_i^2}{|\vec{\omega}_i|^2}\right)\left(1 - \frac{D_j^2}{|\vec{\omega}_j|^2}\right)} - \frac{D_i D_j}{|\vec{\omega}_i||\vec{\omega}_j|} \quad (6)$$

In eq 5, when J is greater than zero, these two systems are found to be analogue systems, and in the opposite case, these systems should be distinguished. This vector algorithm is based on the commercial UV spectra database searching program in Shimadzu CLASS−VP Chromatography Data System (Kyoto, Japan).

Concerning the difference in each SED factor of the lipophilicity, the unit vector, $\vec{\omega}_u$, is employed:

$$\vec{\omega}_u = \frac{\vec{\omega}_u}{|\vec{\omega}_u|} = (r_u, s_u, a_u, b_u, v_u) \quad (7)$$

To quantify the difference in each SED factor, the difference in the components of the unit vectors between LS_i and LS_j is evaluated as follows:

$$\Delta\vec{\omega}_u = (\Delta r_u, \Delta s_u, \Delta a_u, \Delta b_u, \Delta v_u) = (r_{ui} - r_{uj}, s_{ui} - s_{uj}, a_{ui} - a_{uj}, b_{ui} - b_{uj}, v_{ui} - v_{uj}) \quad (8)$$

Using the vector, the contribution of each SED factor to the lipophilicity from one system can be compared with that from other systems.

Results and Discussion

Characterization of Lipophilicity Scales by Vector Approaches—lipophilicity scales measured in this study are summarized in Table 1. As for the microemulsion (ME) systems, an anionic ME using SDS as a surfactant and two cationic MEs using CTAC and DTAC, which have different hydrocarbon chain length with the same hydrophilic group, were employed to measure the lipophilicity of 53 compounds listed in Table 2. For micelles (MCs), three different surfactants, anionic SDS, cationic DTAC, and neutral Brij 35, were used. In these MC systems, the capacity factors of several hydrophobic compounds could not be measured because they coeluted with the MC tracers (AO-10-dodecyl bromide).[1,24] Therefore, only 49 compounds were used in this study. Concerning the IAM column, which has the zwitterionic phosphatidyl choline moiety as the stationary phase, 53 compounds were used although analysis was quite time-consuming because no organic modifier was used. The results of LSER regression analyses are listed in Table 3. The correlation coefficients of the analysis were quite high, except for IAM and MC(S/B). The results of MC(SDS) and IAM by others[18] are also listed in Table 3 to compare with our results. For these two MC(SDS)s, little difference was observed, whereas obvious difference was observed between two IAM systems. This might be caused by the difference in the mobile phase, i.e., 10% acetonitrile was employed in ref 18, whereas no organic modifier was used in this study. This was supported by another LSER result from the fast-gradient IAM system recently reported.[23] In this study, considering the reasonability, the results from the IAM system without organic modifier were used in further study.

Table 3—Results of LSER Analyses

systems	SED coefficients					
	r	*s*	*a*	*b*	*v*	R^2
ME(SDS)	0.699	–1.721	–0.129	–6.932	7.474	0.988
S.E.	0.153	0.183	0.119	0.230	0.214	
ME(CTAC)	0.909	–1.533	0.688	–7.509	7.301	0.994
S.E.	0.295	0.194	0.202	0.250	0.273	
ME(DTAC)	1.262	–1.507	0.700	–7.858	7.572	0.995
S.E.	0.276	0.182	0.189	0.234	0.256	
OW	0.537	–0.926	0.019	–3.537	3.794	0.996
S.E.	0.049	0.058	0.038	0.073	0.068	
IAM	0.280	–0.225	0.517	–2.306	2.657	0.939
S.E.	0.123	0.148	0.096	0.186	0.173	
MC(SDS)	0.497	–0.399	–0.254	–1.669	2.765	0.994
S.E.	0.075	0.057	0.073	0.160	0.167	
MC(DTAC)	0.749	–0.430	0.871	–2.667	2.823	0.976
S.E.	0.100	0.098	0.066	0.127	0.121	
MC(S/B)	–0.094	–0.032	0.615	–2.695	2.519	0.947
S.E.	0.162	0.093	0.123	0.257	0.266	
IAM from ref 18	0.81	–0.42	0.69	–2.00	1.87	
S.E.						
MC(SDS) from ref 22[a]	0.46	–0.48	–0.16	–1.71	2.81	0.982
S.E.	0.06	0.07	0.04	0.08	0.09	
MC(SC) from ref 22[a]	0.56	–0.74	0.15	–2.49	2.65	0.970
S.E.	0.08	0.1	0.06	0.11	0.12	

[a] The original data were cited from ref 19.

Regarding cos θ values between SED vectors, it would be necessary to indicate what value of cos θ could be regarded as "good" analogy because this parameter was not familiar. Therefore, for the eight scales described in Table 1, cos θ values were compared with the corresponding correlation coefficients (r). As shown in Figure 2, $r = 0.90$ corresponds to about cos $\theta = 0.96$, while this relationship was only a yardstick and some deviation was observed for this linear relationship.

Using the LSER coefficients and their D values, analyses of the SED vectors were performed. In Table 4, analogy ranking between one system and the other systems are performed using cos θ. The values of cos θ between three ME systems were quite close to 1, while the values of cos

Table 4—Analogy Ranking between Lipophilicity Scales *i* and *j*

	analogy ranking	LS_i								
		ME (SDS)	ME (CTAC)	ME (DTAC)	OW	IAM	MC (SDS)	MC (DTAC)	MC (S/B)	MC (SC)[a]
	1	OW	ME (DTAC)	ME (CTAC)	ME (SDS)	ME (CTAC)	OW	IAM	IAM	OW
		0.9992	0.9996	0.9996	0.9992	0.9900	0.9736	0.9898	0.9877	0.9979
	2	ME (CTAC)	OW	OW	MC (SC)	MC (DTAC)	ME (SDS)	ME (DTAC)	ME (CTAC)	ME (DTAC)
		0.9953	0.9964	0.9960	0.9979	0.9898	0.9734	0.9842	0.9796	0.9960
	3	MC (SC)	ME (SDS)	MC (SC)	ME (CTAC)	ME (DTAC)	MC (SC)	ME (CTAC)	ME (DTAC)	ME (CTAC)
		0.9947	0.9953	0.9960	0.9964	0.9896	0.9692	0.9820	0.9770	0.9952
	4	ME (DTAC)	MC (SC)	ME (SDS)	ME (DTAC)	MC (S/B)	IAM	MC (SC)	MC (DTAC)	ME (SDS)
		0.9941	0.9952	0.9941	0.9960	0.9877	0.9570	0.9800	0.9699	0.9947
LS_j ($\cos\theta_{ij}$)	5	IAM	IAM	IAM	IAM	OW	ME (DTAC)	OW	ME (SDS)	IAM
		0.9818	0.9900	0.9896	0.9829	0.9829	0.9569	0.9718	0.9652	0.9822
	6	MC (SDS)	MC (DTAC)	MC (DTAC)	MC (SDS)	MC (SC)	ME (CTAC)	MC (S/B)	OW	MC (DTAC)
		0.9734	0.9820	0.9842	0.9736	0.9822	0.9561	0.9699	0.9630	0.9800
	7	MC (DTAC)	MC (S/B)	MC (S/B)	MC (DTAC)	ME (SDS)	MC (DTAC)	ME (SDS)	MC (SC)	MC (SDS)
		0.9653	0.9796	0.9770	0.9718	0.9818	0.9355	0.9653	0.9572	0.9692
	8	MC (S/B)	MC (SDS)	MC (SDS)	MC (S/B)	MC (SDS)	MC (S/B)	MC (SDS)	MC (SDS)	MC (S/B)
		0.9652	0.9561	0.9569	0.9630	0.9570	0.9126	0.9355	0.9126	0.9572

[a] Data were cited from ref 22; original data were reported by ref 19.

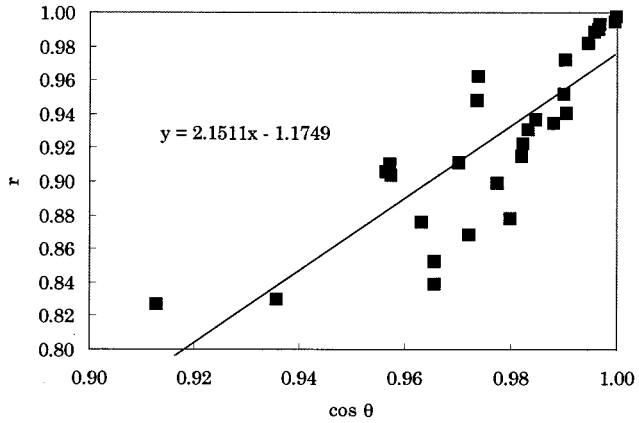

Figure 2—Relationship between cos θ and r for eight lipophilicity scales described in Table 1.

Table 5—Analogy Judgment by *J* Values[a]

system	ME (SDS)	ME (CTAC)	ME (DTAC)	OW	IAM	MC (SDS)	MC (DTAC)	MC (S/B)	MC (SC)[b]
ME(SDS)	○								
ME(CTAC)	×	○							
ME(DTAC)	×	○	○						
OW	○	×	×	○					
IAM	×	×	×	×	○				
MC(SDS)	×	×	×	×	×	○			
MC(DTAC)	×	×	×	×	×	×	○		
MC(S/B)	×	×	×	×	△	×	×	○	
MC(SC)[a]	×	△	△	○	×	×	×	×	○

[a] $J \geq 0$: ○, $J < 0$: ×, Out of Judgment: △. When either the analogy of LS_i to LS_j or the analogy of LS_j to LS_i was out of judgment, △ was indicated. [b] Data were cited from ref 22; original data were reported by ref 19.

θ between four different MC systems were not more than 0.98. As for IAM system, higher analogy to systems using cationic surfactants such as ME(CTAC), MC(DTAC), and ME(DTAC) was observed. This would be caused by the cationic choline moiety of the packing material of the IAM column. In Table 5, the results of analogy judgment using the J values are indicated. Note that varied D values may provide the ranking/judgment reversal, which is the phenomenon that the high-ranked vector with larger cos θ and smaller D value provides $J < 0$ and low-ranked vector with smaller cos θ and larger D value provides $J > 0$ when the deviations of the LSER regression analyses for these LSs are varied. Therefore, the accuracy of the analogy judgment

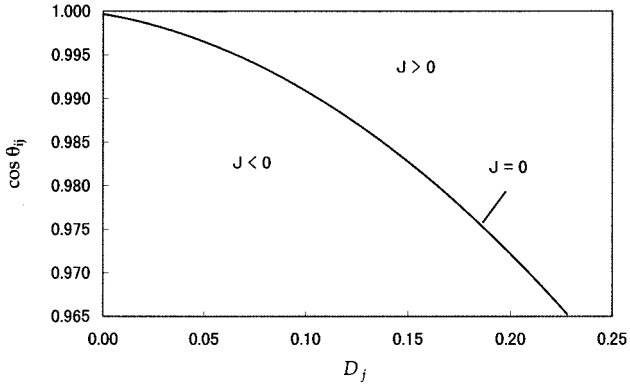

Figure 3—Relationship between cos θ_{ij}, D_j, and J values using the SED unit vectors of ME(SDS) ($\vec{\omega}_i$) and simulated unit vectors ($\vec{\omega}_j$). $\vec{\omega}_i$ = (0.015, 0.018, 0.012, 0.022, 0.021), D_i = 0.035.

depended on D values of SED vectors employed. The relationship between cos θ, D, and J values was simulated using the unit vector of ME(SDS)(LS_i) and the unit vectors of the other scales(LS_j). As shown in Figure 3, the ranking/judgment reversal would occur when D_j of lower-ranked LS_j is large enough. Therefore, the restriction to judge the analogy between two LSs was required, considering the purpose of the study. In this case, to prevent the ranking/judgment reversal, the analogy judgment was restricted by the rule that if the J value of one of the ranked vectors is less than zero, the judgment of the lower-ranked vectors with $J > 0$ should not be performed, i.e., "Out of Judgment" should be indicated for the lower-ranked vectors with $J > 0$. In Table 5, this rule was applied and the ranking/judgment reversal between some SED vectors was prevented.

Concerning the analogy to OW, so far as we know, the ME(SDS)[1,4] and sodium cholate micelle (MC(SC))[12,25] systems provided the best correlation. In this study, the LSER data from MC(SC) was cited from the report by Poole et al.[22] Using the data, the vector analyses were performed. As a result, both ME(SDS) and MC(SC) were each correlated with OW, whereas the J value between ME(SDS) and MC(SC) indicated that the analogy between them was not found despite a large cos θ value (0.9947). From the J values, two cationic ME systems were different from the anionic ME(SDS) although the cos θ values between these three systems were quite close to 1. The relationship between three MEs, OW, and MC(SC) on the basis of their J values is roughly illustrated in Figure 4. ME(SDS) should

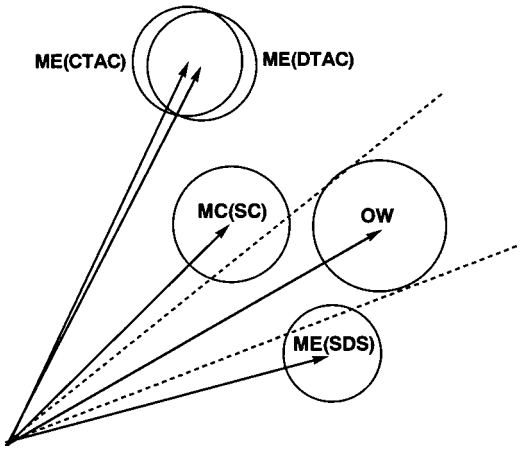

Figure 4—Relationship between ME(SDS), OW, MC(SC), ME(DTAC), and ME(CTAC) in two-dimensional model space of SED vectors. The angles and the length of these vectors are not accurate, because the actual space of these vectors are five-dimensional.

be, in the strict sense, distinguished from MC(SC) and cationic ME systems, while Poole et al. classified ME(SDS) and MC(SC) as the same group.[22]

Next, the unit vectors of the SED vectors were calculated to evaluate the difference in each independent descriptor of these LSs. Using the vector approach, it was quite easy to analyze plural scales simultaneously, and the contribution of each SED between all scales could be compared. The results of the unit vector analysis are shown in Figure 5. Note that this approach was based on the assumption that these five descriptors were almost equivalent to each other. This assumption would be valid because LSER descriptors of solutes had been originally adjusted to be almost the same order, and no artifact caused by the LSER descriptors has been reported so far.[16,17,26,27] Additionally, because the results were used only for the comparison in the same descriptor between the different scales, the assumption would not cause inappropriate evaluation. As shown in Figure 5a, among the components of the unit vector, the a_u component, i.e., the contribution of the hydrogen-bond basicity of the systems or hydrogen-bond acidity of the solutes to their lipophilicity scale was the most diverse between the employed scales. The scale with the most positive contribution of the a_u component was MC(CTAC), and the most negative one was MC(SDS). On the other hand, the b_u component was almost independent of the scales. These three MEs, OW, and MC(SC) systems provided almost the same values in all components, whereas IAM, MC(SDS, DTAC, and S/B) provided different values. The difference in each component of the unit vectors between OW and other scales is shown in Figure 5b. As expected from Figure 4, the a_u and r_u components of MC(SC) were larger than those of OW, while the a_u and r_u components of ME(SDS) were smaller than those of OW. The b_u, v_u and s_u components of MC(SC) were smaller than those of OW, while the b_u, v_u and s_u components of ME(SDS) were larger than those of OW.

As a consequence, the analogy between these employed scales and the difference in the contribution of each descriptor to the lipophilicity scales were well quantified by the SED vectors and their unit vectors, respectively.

Application of These Vector Approaches—The approach using the SED vectors and their unit vectors from various chemical two-phases partition models would be useful to express other scales using plural and diverse vectors because the independent vectors are suitable to describe another vector by addition or subtraction of these vectors with a certain ratio. Two typical examples are shown below:

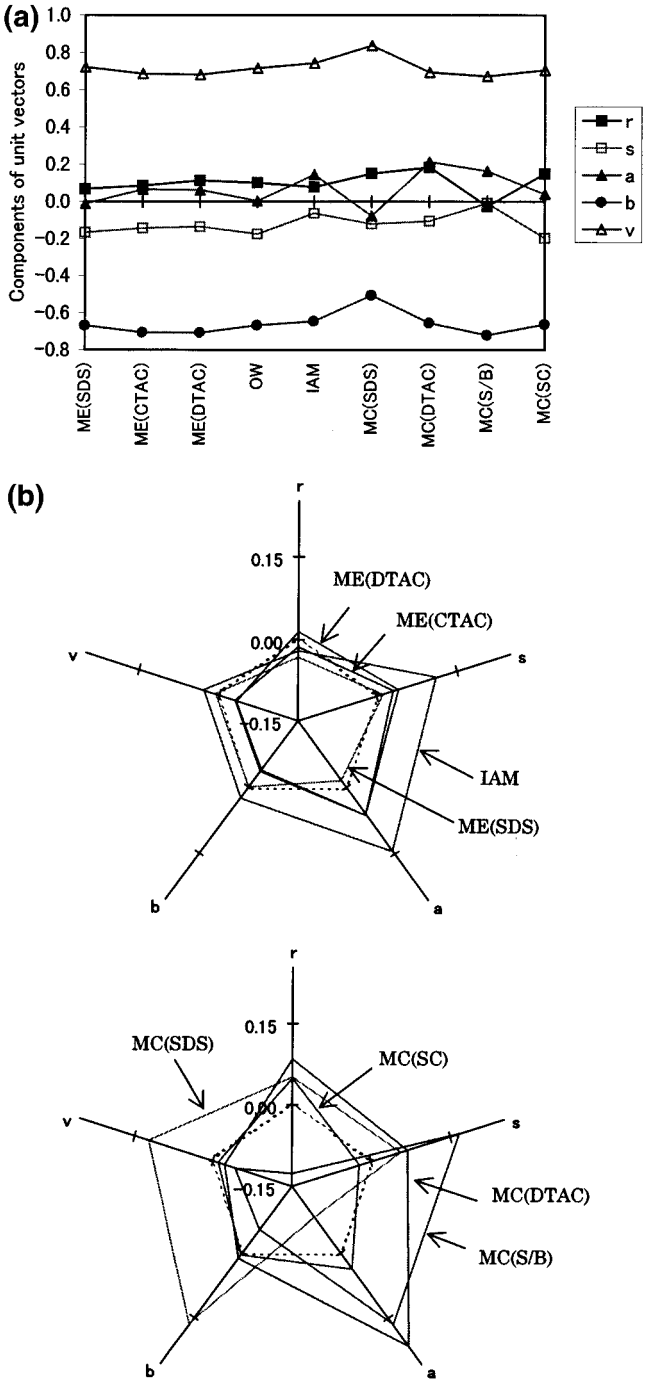

Figure 5—(a) Component of unit vectors. (b) Difference in the unit vector of OW from those of other systems. The dot lines indicate OW.

(1) Structural Information of Microemulsion from Each Constituent Vector—It was previously reported that the partition behaviors of solutes in ME(SDS) could be expressed by the behavior in each constituent system such as water−surfactant, water−alcohol, and water−alkane systems.[4] In this study, the same results were obtained using this vector approach for not only ME(SDS) but also ME(DTAC) as shown in Tables 6 and 7. In this approach, reproducible MI could be directly used as the lipophilicity scale, while MI had to convert to irreproducible log k' in a previous study because the lipophilicity scales such as MI and log P were not equivalent (e.g., log P = 0.518MI − 0.854 for ME(SDS)).[1] As a result, ME(DTAC) as well as ME(SDS) could be expressed by the summation of the SED vectors of the constituents with actual mixing molar ratio. In addition, the vector approach allowed the

Table 6—Estimation of Lipophilicity Scales from ME(SDS) by Its Constituents

system	SED coefficients					ME constituent ratio[a]	
	r	*s*	*a*	*b*	*v*	actual	regressed
MC(SDS)	0.497	–0.399	–0.254	–1.669	2.765	0.050	0.000
water–alkane	0.65	–1.66	–3.52	–4.82	4.28	0.081	0.030
water–pentanol	0.58	–0.79	0.02	–2.84	3.25	0.869	0.970
ME (observed)	0.699	–1.721	–0.129	–6.932	7.474		
ME (calcd with actual ratio)	0.582	–0.842	–0.283	–2.941	3.310		
ME (calcd with regressed ratio)	0.582	–0.816	–0.088	–2.900	3.281		

analogy	cos θ
ME (obs) and ME (calcd with actual ratio)	0.9965
ME (obs) and ME (calcd with regressed ratio)	0.9976

[a] These ratios were expressed as molar ratio. The regressed ratio was calculated from the regression analysis described in the text.

Table 7—Estimation of Lipophilicity Scales from ME(DTAC) by Its Constituents

system	SED coefficients					ME constituent ratio[a]	
	r	*s*	*a*	*b*	*v*	actual	regressed
MC(DTAC)	0.749	–0.430	0.871	–2.667	2.823	0.050	0.671
water–alkane	0.65	–1.66	–3.52	–4.82	4.28	0.081	0.103
water–pentanol	0.58	–0.79	0.02	–2.84	3.25	0.869	0.226
ME (observed)	0.909	–1.533	0.688	–7.509	7.301		
ME (calcd with actual ratio)	0.594	–0.842	–0.224	–2.992	3.312		
ME (calcd with regressed ratio)	0.701	–0.638	0.227	–2.928	3.069		

analogy	cos θ
ME (obs) and ME (calcd with actual ratio)	0.9897
ME (obs) and ME (calcd with regressed ratio)	0.9963

[a] These ratios were expressed as molar ratio. The regressed ratio was calculated from the regression analysis described in the text.

regression analyses using the mixing ratio of the constituents as a variable parameter to obtain the minimum values of $(1 - \cos\theta)$. The comparison of the obtained regressed mixing ratio with the actual ones would provide the structural information of the microemulsion. As shown in Table 6, the contribution of SDS was more suppressed than that expected by the actual mixing ratio. Interestingly, however, in the case of ME(DTAC), the contribution of DTAC in the mixing ratio was more increased than the actual one. This would be caused by the influence of the bulkiness of the hydrophilic portion of the surfactants on the surface-shielding effect of 1-butanol.[1,2] A similar phenomenon was observed in the case of micelles.[8] Although ME(SDS) and ME(DTAC) provided the large cos θ value, the analogy was not found according to the *J* value, as previously shown. The results of the regressed mixing ratio in ME(SDS) and ME(DTAC) supported the reasonability of the *J* value analysis, and this approach would be useful to investigate complex partition systems such as microemulsions.

(2) Prediction of Biological Systems from Chemical Systems—it is important to predict the scales from complex biological systems whose lipophilicity scales are difficult or time-consuming to measure with high reproducibility. Using the vector approach, the SED vectors from the biological systems would be promptly expressed by the vectors from some chemical systems, although the partition behavior of the minimum test set of solutes in the biological systems had to be measured to obtain the SED coefficients. In this study, two biological systems, skin and blood–brain barrier (BBB), were investigated.

(2.1) Skin—Water/skin partition coefficients (K_m) of some alcohols and steroids (22 compounds) were measured using excised human skin, and the LSER equation was calculated by Abraham et al.[26] as follows:

$$\log K_m = -(0.03 \pm 0.14) - (0.37 \pm 0.11)\pi_2^H + (0.33 \pm 0.15)\sum\alpha_2^H - (1.67 \pm 0.16)\sum\beta_2 + (1.87 \pm 0.17)V_x \quad (9)$$

$$n = 22,\ r^2 = 0.943,\ \text{sd} = 0.166,\ F = 70$$

In this case, *R* was removed because each descriptor of the compounds employed was not independent. Using the four SED coefficients, the analogy ranking of various chemical systems to the skin system was evaluated (Table 8). Interestingly, the skin system showed the high-ranked analogy to cationic IAM, cationic ME(CTAC), and ME(DTAC). In this case, the analogy judgment was not performed because the *D* value of the skin SED unit vector was quite large (0.06). Next, the analysis of their unit vectors was performed (Figure 6). As a result, the a_u, b_u, and v_u components of the skin system were between IAM and others (cationic MEs). Thus, the skin system was examined to be expressed by IAM and ME(CTAC). To obtain the ratio of IAM and ME(CTAC), a regression analysis for the unit vector of the skin system using the ratio as a variable parameter was performed, and the results are shown as follows:

$$\vec{\omega}_u(\text{skin}) = 0.746\vec{\omega}_u(\text{ME(CTAC)}) + 0.252\vec{\omega}_u(\text{IAM}) \quad (10)$$

$$\cos\theta = 0.9960$$

The value of cos θ was improved in comparison with that by each single vector. Thus, the skin permeation of the drug candidates would be efficiently predicted by the two chromatographic systems such as ME(CTAC) and IAM with higher reproducibility, although the preliminary LSER regression analysis was not accurate enough to judge the analogy between the observed skin vector and the regressed skin vector.

Table 8—Analogy Ranking to Skin Permeability

	ME(SDS)	ME(CTAC)	ME(DTAC)	OW	IAM	MC(SDS)	MC(DTAC)	MC(S/B)
analogy ranking	4	1	2	5	3	8	7	6
cos θ	0.9904	0.9954	0.9941	0.9903	0.9920	0.9795	0.9841	0.9872

Table 9—Analogy Ranking to BBB

	ME(SDS)	ME(CTAC)	ME(DTAC)	OW	IAM	MC(SDS)	MC(DTAC)	MC(S/B)	MC[a] (LiPFOS)	AW[b]
analogy ranking	3	5	6	4	7	2	9	10	8	1
cos θ	0.8372	0.7890	0.7879	0.8345	0.7268	0.8580	0.7002	0.6701	0.7145	0.9544
analogy ranking (reg)[c]	6	4	2	8	2	1	5	8	6	–
cos θ (reg)[c]	0.9545	0.9551	0.9552	0.9544	0.9552	0.9558	0.9547	0.9544	0.9545	–

[a] The data were cited from ref 22. LiPFOS: lithium perfluorooctanesulfonate. [b] The data were cited from ref 27. AW: alkane–water. [c] cos θ (reg): the angle between BBB vector and regressed BBB vector which was calculated by the summation of alkane–water and the indicated system with the regressed ratio.

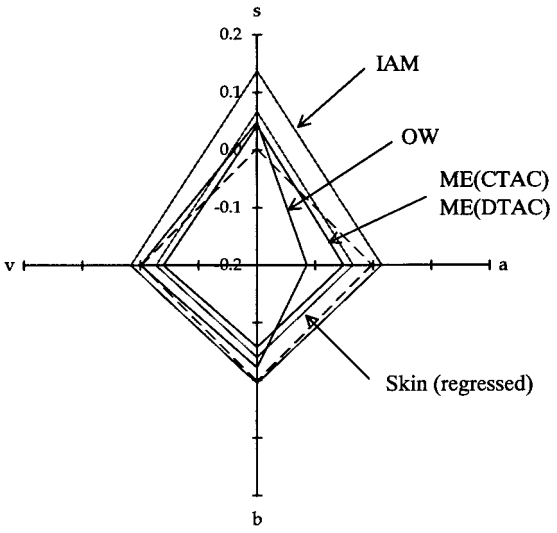

Figure 6—Difference in the unit vector of skin from those of other chemical systems including the regressed skin system.

(2.2) Blood−Brain Barrier—Young et al.[28] and Abraham et al.[27] determined the distribution coefficients between blood and brain (K_{BB}) for 57 compounds, and the following LSER equation was found:

$$\log K_{BB} = -(0.04 \pm 0.06) + (0.20 \pm 0.10)R_2 - (0.69 \pm 0.12)\pi_2^H - (0.71 \pm 0.33)\sum\alpha_2^H - (0.70 \pm 0.11)\sum\beta_2 + (1.00 \pm 0.10)V_x \quad (11)$$

$$n = 57,\ r^2 = 0.907,\ \text{sd} = 0.197,\ \text{F} = 99$$

Using these SED coefficients, the analogy ranking of various chemical systems was investigated (Table 9). In this case, the values of cos θ was not so close to 1 except the case of AW. In addition, the analogy judgment was not also performed because the D value of the BBB SED unit vector was 0.10. Next, the combination of AW with another system was employed to express the BBB system. The ratio of the two vectors was calculated by regression analysis using the ratio as a variable parameter to obtain the minimum value of (1 − cos θ) between the regressed and actual vectors from the BBB. The obtained values of cos θ are also listed in Table 9. As a result, the best combination for the BBB system was MC(SDS) and AW, and the cos θ was slightly improved as follows:

$$\vec{\omega}_u(\text{BBB}) = 0.107\vec{\omega}_u(\text{MC(SDS)}) + 0.861\vec{\omega}_u(\text{AW}) \quad (12)$$

$$\cos\theta = 0.9558$$

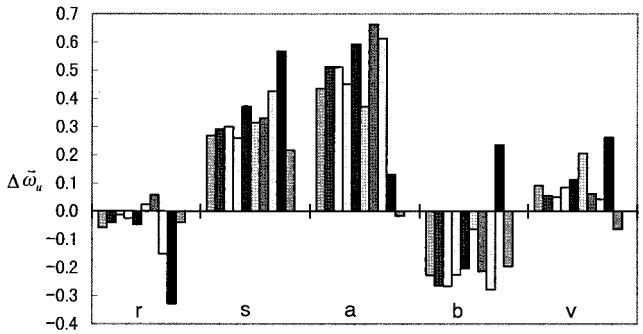

Figure 7—Difference in the unit vector of BBB from those of other chemical systems. Columns: (left) ME(SDS), ME(CTAC), ME(DTAC), OW, IAM, MC(SDS), MC(DTAC), MC(S/B), MC(LiPFOS), and AW.

This equation was not sufficient to express the complex BBB system by the simple chemical systems because the s_u and a_u components of the unit vector of BBB were smaller than those of other 10 unit vectors employed in this study (Figure 7). It would be necessary to search other simple chemical scales with smaller s_u, a_u, and v_u and larger r_u and b_u components as well as to improve the accuracy of the LSER regression analysis for log k_{BB}.

Conclusions

We developed the vector approaches to treat LSER analysis data more efficiently and quantitatively. This allows us to characterize various lipophilicity scales simultaneously. Using the vector approach, not only a complex chemical system such as a microemulsion, but also biological systems such as the skin and BBB, could be expressed by some simple chemical systems, although some improvement on the accuracy of the biological SED vectors would be required. These approaches would facilitate the selection of the chemical systems suitable for the prediction of the hydrophobic interaction of drugs in the body.

References and Notes

1. Ishihama, Y.; Oda, Y.; Uchikawa, K.; Asakawa, N. Evaluation of Solute Hydrophobicity by Microemulsion Electrokinetic Chromatography. *Anal. Chem.* **1995**, *67*, 1588−1595.
2. Ishihama, Y.; Oda, Y.; Asakawa, N. A Hydrophobicity Scale Based on the Migration Index from Microemulsion Electrokinitic Chromatography of Anionic Solutes. *Anal. Chem.* **1996**, *68*, 1028−1032.
3. Ishihama, Y.; Oda, Y.; Asakawa, N. Hydrophobicity of Cationic Solutes Measured by Electrokinetic Chromatography with Cationic Microemulsions. *Anal. Chem.* **1996**, *68*, 4281−4284.

4. Abraham, M. H.; Treiner, C.; Roses, M.; Rafols, Y.; Ishihama, Y. Linear Free Energy Relationship Analysis of Microemulsion Electrokinetic Chromatographic Determination of Lipophilicity. *J. Chromatogr. A* **1996**, *752*, 243−249.
5. Kaliszan, R. Quantitative Structure-Retention Relationships. *Anal. Chem*. **1992**, *64*, 619A−631A.
6. Ong, S.; Liu, H.; Pidgeon, C. Immobilized-Artificial-Membrane Chromatography: Measurements of Membrane Partition Coefficient and Prediction Drug Membrane Permeability. *J. Chromatogr. A* **1996**, *728*, 113−128.
7. Valko, K.; Bevan, C.; Reynolds, D. Chromatographic Hydrophobicity Index by Fast-Gradient RP-HPLC: A High-Throughput Alternative to log P/log D. *Anal. Chem*. **1997**, *69*, 2022−2029.
8. Ishihama, Y.; Oda, Y.; Uchikawa, K.; Asakawa, N. Correlation of Octanol−Water Partition Coefficients with Capacity Factors Measured by Micellar Electrokinetic Chromatography. *Chem. Pharm. Bull*. **1994**, *42*, 1525−1527.
9. Chen, N.; Zhang, Y.; Terabe, S.; Nakagawa, T. Effect of Physicochemical Properties and Molecular Structure on the Micelle-Water Partition Coefficient in Micellar Electrokinetic Chromatography. *J. Chromatogr. A* **1994**, *678*, 327−332.
10. Smith, J.T.; Vinjamoori, D. V. Rapid Determination of Logarithmic Partition Coefficients between *n*-Octanol and Water Using Micellar Electrokinetic Capillary Chromatography. *J. Chromatogr. B* **1995**, *669*, 59−66.
11. Herbert, B. J.; Dorsey, J. G. *n*-Octanol−Water Partition Coefficient Estimation by Micellar Electrokinetic Capillary Chromatography. *Anal. Chem*. **1995**, *67*, 744−749.
12. Yang, S.; Bumgarner, J. G.; Kruk, L. F. R.; Khaledi, M. G. Quantitative Structure−Activity Relationships Studies with Micellar Electrokinetic Chromatography Influence of Surfactant Type and Mixed Micelles on Estimation of Hydrophobicity and Bioavailability. *J. Chromatogr. A* **1996**, *721*, 323−335.
13. Garcia, M. A.; Diez-Masa, J. C.; Marina, M. L. Correlation between the Logarithm of Capacity Factors for Aromatic Compounds in Micellar Electrokinetic Chromatography and their Octanol−Water Partition Coefficients. *J. Chromatogr. A* **1996**, *742*, 251−256.
14. Hanna, M.; Biasi, V.; Bond, B.; Salter, C.; Hutt, A. J.; Camilleri, P. Estimation of the Partition Characteristics of Drugs: A Comparison of a Large and Diverse Drug Series Utilizing Chromatographic and Electrophoretic Methodology. *Anal. Chem*. **1998**, *70*, 2092−2099.
15. Masucci, J. A.; Caldwell, G. W.; Foley, J. P. Comparison of the Retention Behavior of β-blockers Using Immobilized Artificial Membrane Chromatography and Lysophospholipid Micellar Electrokinetic Chromatography. *J. Chromatogr. A* **1998**, *810*, 95−103.
16. Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Taft, R. W. Solubility Properties in Biological Media. 12. Regarding the Mechanism of Nonspecific Toxicity or Narcosis by Organic Nonelectrolytes. *Quant. Struct.-Act. Relat*. **1988**, *7*, 71−78.
17. Abraham, M. H. Scales of Solute Hydrogen-bonding: Their Construction and Application to Physicochemical and Biochemical Processes. *Chem. Soc. Rev*. **1993**, *22*, 73−83.
18. Abraham, M. H.; Chadha, H. S.; Leitao, R. A. E.; Mitchell, R. C.; Lambert, W. J.; Kaliszan, R.; Nasal, A.; Haber, P. Determination of Solute Lipophilicity, as log P(Octanol) and log P(alkane) Using Poly(Stylene-Divinylbenzene) and Immobilised Artificial Membrane Stationary Phases in Reversed-phase High-Performance Liquid Chromatography. *J. Chromatogr. A* **1997,** *766*, 35−47.
19. Yang, S.; Khaledi, M. G. Chemical Selectivity in Micellar Electrokinetic Chromatography: Characterization of Solute-Micelle Interactions for Classification of Surfactants. *Anal. Chem*. **1995**, *67*, 499−510.
20. Yang, S.; Bumgarner, J. G.; Khaledi, M. G. Chemical Selectivity in Micellar Electrokinetic Chromatography II: Rationalization of Elution Patterns in Different Surfactant Systems. *J. Chromatogr. A* **1996**, *738*, 265−274.
21. Muijselaar, P. G.; Claessens, H. A.; Cramers, C. A. Characterization of Pseudostationary Phases in Micellar Electrokinetic Chromatography by Applying Linear Solvation Energy Relationships and Retention Indexes. *Anal. Chem*. **1997**, *69*, 1184−1191.
22. Poole, C. F.; Poole, S. K.; Abraham, M. H. Recommendations for the Determination of Selectivity in Micellar Electrokinetic Chromatography. *J. Chromatogr. A* **1998**, *798*, 207−222.
23. Valko, K.; Plass, M.; Bevan C.; Reynolds, D.; Abraham, M. H. Relationships between the Chromatographic Hydrophobicity Indices and Solute Descriptors Obtained by Using Several Reversed-phase, Diol, Nitrile, Cyclodextrin and Immobilised Artificial Membrane-bonded High-Performance Liquid Chromatography Columns. *J. Chromatogr. A* **1998**, *797*, 41−55.
24. Seals, T. H., Davis, J. M., Murphy, M. R., Smith, K. W., Stevens, W. C. *Anal. Chem*. **1998**, *70*, 4549−4562.
25. Wu, Y. S.; Lee, H. K.; Li, S. F. Y. Rapid Estimation of Octanol−Water Partition Coefficients of Pesticides by Micellar Electrokinetic Chromatography. *Electrophoresis* **1998**, *19*, 1719−1727.
26. Abraham, M. H.; Chadha, H. S. Applications of a Solvation Equation to Drug Transport Properties. In *Lipophilicity in Drug Action and Toxicology*; Mannhold, R., Kubinyi, H., Timmerman, H., Eds.; VCH: Weinheim (Germany), 1996; Chapt. 18, pp 311−337.
27. Abraham, M. H.; Chadha, H. S., Mitchell, R. C., Hydrogen Bonding. 33. Factors That Influence The Distribution of Solutes between Blood and Brain. *J. Pharm. Sci*. **1994**, *83*, 1257−1268.
28. Young, R. C.; Mitchell, R. C.; Brown, T. H.; Ganellin, C. R.; Griffiths, R.; Jones, M.; Rana, K. K.; Saunders: D.; Smith, I. R.; Sore, N. E.; Wilks, T. J. *J. Med. Chem*. **1988**, *31*, 656−671.

JS990119+