Hydrophobicity Parameters Determined by Reversed-Phase Liquid Chromatography. XVI: A New Hydrogen-Accepting Parameter for Monosubstituted Thiophenes and Furans for Correlating Retention Factors and Octanol–Water Partition Coefficients¹⁾

Chisako YAMAGAMI* and Hiromi TACHIKAWA

Kobe Pharmaceutical University; Motoyamakita-machi, Higashinada-ku, Kobe 658–8558, Japan. Received June 3, 2003; accepted July 29, 2003

> We recently proposed a new hydrogen-accepting parameter, $S_{H\text{A}}$, for monosubstituted (di)azines on the basis of the heat of formation calculated by the conductor-like screening model (COSMO) method. In this work, S_{HA} **values for monosubstituted thiophenes and furans were calculated and the results were applied to the analysis of relationships between log** *P* **(***P***: 1-octanol/water partition coefficient) and log** *k* **(***k***: retention factor obtained by re**versed phase HPLC). The S_{HA} parameter was found to work effectively as a hydrogen-bonding parameter in a **range of heteroaromatic compounds.**

> **Key words** partition coefficient; retention factor; hydrogen-accepting parameter; thiophene; furan; conductor-like screening model (COSMO)

In connection with recently developed methodologies for rational drug design, rapid prediction of the hydrophobicity of candidate-compounds before synthesis has increasingly become important.^{2,3)} In particular, much attention has been devoted to predicting the hydrophobicity for heterocyclic compounds. Hydrophobicity, which is generally expressed by the logarithm of 1-octanol/water partition coefficient, log*P*, is often predicted by calculations or estimated from the retention factor, $\log k$, obtained experimentally by reversedphase high performance liquid chromatography (RP-HPLC), and extensive studies have been reported on this subject.^{3—12)} Despite this, no universal method has been established yet to predict reliable values of log *P* for heterocyclic compounds. This is largely due to the fact that parameters like log *P* and log *k* for heterocyclic compounds involve H-bonding effects, of which contributions are difficult to evaluate.

With the RP-HPLC approach, such H-bonding effects often give non-linearity between log *P* and log *k*, leading us to erroneous estimations. Accordingly, a major problem for providing accurate predictions lies in estimating the contributions of H-bonding effects. One widely used procedure for expressing H-bonding effects is to adopt indicator variables, HB_A and HB_D , which take the value of 1 for H-acceptors and H-donors, respectively, and 0 for others.¹³⁾ Although the physical meaning of using such discrete-type parameters is well discussed, 13) their applications are limited to a group of compounds whose H-bonding abilities are constrained within a defined range. We have used the HB_A parameter to analyze the relationship between $\log k$ and $\log P$ for monosubstituted heteroaromatic compounds, $(4^{14}-19)$ such as (di)azines, furans and thiophenes, and also to correlate $log P$ values for (di)azines obtained in different partitioning systems, 20) but in some cases, somewhat arbitrary classification of the substituents into highly and weakly H-accepting ones was needed.18,20) Therefore, in order to deal with a variety of substituents covering a wide range of H-bonding abilities, we have recently proposed²¹⁾ a new H-accepting parameter S_{HA} for (di)azines (pyridines, pyrazines and pyrimidines). Values of S_{H_A} are accessible by semi-empirical MO calculations using the conductor-like screening model (COSMO) method.²²⁾ By using this pre-determined parameter, we formulated Eq. 1 as a general equation for $\log k$ values for monosubstituted (di)azines, AR-X, where X-substituents are non-H-bonding or H-accepting, eluted by various compositions of aqueous methanol solutions. 21)

$$
\log k = a \log P + s S_{HA} + \rho \sigma_I + const. \tag{1}
$$

In Eq. (1), S_{HA} represents the H-accepting parameter of the X-substituent; and σ_1 the inductive electronic constant,²³⁾ expresses the electronic effect of X on H-bonding of the ring *N* atom(s). By examining the coefficients of these correction terms, we concluded that H-bonding effects are minimal in eluents of about 50% aqueous methanol.

The H-bond acidity and basicity scales of Abraham and coworkers^{24—26)} provide one of the most comprehensive Hbonding parameters, applicable to quantitative structure–activity studies, currently available. Unfortunately, the data for heterocyclic compounds are not extensive, probably due to a lack of the experimental data necessary for their calculation. Accordingly, we wished to examine whether our S_{HA} parameter, readily accessible by calculation, would be applicable to estimation of H-bonding effects for other heteroaromatic series. To this end, we determined S_{HA} values for monosubstituted thiophenes (TH-X) and furans (FR-X) for which log *k* values had previously been analyzed by using the HBA parameter,^{16,19)} and reanalyzed the $\log k$ values by Eq. 1 in order to verify the validity of our parameter. Applications were further extended to analyses of log *k* measured in eluents containing acetonitrile as an organic modifier.

Experimental

Chemicals Compounds, TH-X, and FR-X, with the non-H-bonding and H-accepting substituents shown in Table 1 were used. These substituents are similar to those used in the previous studies, $16,19)$ but some strongly H-accepting substituents, including disubstituted derivatives, **32**—**34**, were added in this work to examine more clearly the applicability of the S_{HA} parameter.

Table 1. Physicochemical Parameters for Substituted Thiophenes (TH-X) and Furans (FR-X)

a) Taken from our previous work (ref. 19) unless otherwise noted. *b*) Taken from ref. 27 unless otherwise noted. *c*) Taken from our previous work (ref. 16). *d*) This work.

a) The figures after A represent volume % of acetonitrile in eluents. *b*) Not measured due to long retention times.

Data for log k Retention factors, $\log k$, obtained in eluents containing different concentrations of methanol were taken from our previous work.^{16,19)} Those for additional compounds were measured under the same conditions as those used previously. For thiophenes, log *k* values were also obtained in eluents with acetonitrile as an organic modifier according to the method described for methanolic solutions.¹⁹⁾ Retention times were measured at 25 °C on a Capcell pak C18 SG120 column (Shiseido, 150×4.6 mm I.D.) in CH₃CN–phosphate buffer (0.01 M, pH 7.4) solutions containing 10–60% (v/v) CH₃CN (Table 2). Retention factors, *k*, were calculated by $k = (t_R - t_0)/t_0$ where t_0 and t_R express elution times of methanol and sample, respectively.

Data for log *P* Octanol–water partition coefficients, *P*, were taken from our previous work.^{16,19)} As shown in ref. 16, those for most of the furan derivatives are literature values. The log *P* values for compounds newly added in this study were measured by the conventional shake-flask method.²⁸⁾ The data are given in Table 1.

 S_{HA} **Parameters** The procedure for calculating the S_{HA} parameter was previously described in detail.²¹⁾ First, the minimum energy conformation of each compound (Ar-X) in the gaseous state was established using the AM1 method²⁹⁾ in the MOPAC 93 program package incorporated in an ANCHOR II modeling system.30) By using this conformation as the initial geometry, heats of formation, ΔH _p in various solvents were calculated by the COSMO method²²⁾ which approximates the effects of solvent molecules surrounding the molecule in question with the eps (ε) : dielectric constant) command. In practice, ΔH_f values in the following five dielectric media were calculated: ε =1 (gas), ε =4.8 (chloroform), ε =10.3 (octanol), ε =32.7 (methanol), and ε =78.4 (water). The ΔH_f values calculated for the five dielectric environments for an Ar-X molecule were plotted against the corresponding ΔH_e values for the unsubstituted compound (Ar-H), providing a linear relationship with a positive slope; the higher the dielectric constant of the medium, the lower the ΔH_f value indicating the greater stabilization induced by solvation with the more polar solvent molecules. The slope of this straight line is defined as S_{HA} for the Ar-X compound. As the X-substituents do not contain an H-donating site and the solvents examined are amphiprotic, the S_{HA} parameter is expected to reflect the H-accepting ability of the X-substituent. The

Table 3. Analyses of log *k* for Thiophenes (TH-X) and Furans (FR-X) by Eq. 1

	Eluent ^a	Coefficient		const	$n^{b)}$	r^{c}	s^{d}	F^{e}
		log P	$S_{\rm HA}$					
				Methanol-buffer (pH 7.4)				
TH-X	M15	0.737	0.198	-0.286	18^{f}	0.984	0.082	221
		$(0.085)^{g}$	(0.064)	(0.283)				
	M30	0.709	0.074	-0.377	20	0.995	0.053	853
		(0.040)	(0.031)	(0.134)				
	M50	0.579	-0.043^{h}	-0.464	20	0.995	0.048	873
		(0.036)	(0.028)	(0.121)				
	M70	0.437	-0.091	-0.666	20	0.992	0.051	537
		(0.036)	(0.028)	(0.121)				
FR-X	M15	0.761	0.284	-0.304	14	0.974	0.101	103
		(0.118)	(0.095)	(0.304)				
	M30	0.743	0.152	-0.442	14	0.994	0.049	457
		(0.057)	(0.046)	(0.148)				
	M50	0.623		-0.523	14	0.995	0.041	1199
		(0.039)		(0.054)				
	M70	0.541		-0.949	14	0.991	0.049	627
		(0.049)		(0.065)				
				Acetonitrile-buffer (pH 7.4)				
TH-X	A10	0.739	0.133	-0.217	17^{i}	0.992	0.059	436
		(0.063)	(0.046)	(0.205)				
	A20	0.678		-0.176	20	0.998	0.035	3962
		(0.023)		(0.051)				
	A30	0.593	-0.041	-0.265	20	0.997	0.036	1616
		(0.027)	(0.021)	(0.091)				
	A40	0.492	-0.059	-0.302	20	0.995	0.041	904
		(0.031)	(0.024)	(0.104)				
	A50	0.403	-0.063	-0.350	20	0.994	0.038	745
		(0.029)	(0.023)	(0.097)				
	A60	0.386	-0.064	-0.621	20	0.988	0.053	355
		(0.040)	(0.032)	(0.135)				

a) The figures after M and A represent volume % of MeOH and CH3CN in eluents. *b*) Number of compounds used for correlations. *c*) Correlation coefficients. *d*) Standard deviations. *e*) Values of *F*-ratio between regression and residual variances. *f*) Compounds 15 and 18 were excluded because their log *k* values were too long to measure. *g*) Figures in parentheses are 95% confidence intervals. *h*) Justified at the 99.4% level. The log *P* and *S*_{HA} terms of all other equations are justified above the 99.9% level. *i*) Compounds **4**, **15** and **18** were excluded because their log *k* values were too long to measure.

*S*HA values for disubstituted furans, **32**—**34**, were calculated in the same manner: the slope of the straight line in the plot of ΔH_f for a disubstituted furan against ΔH_f for furan being defined as S_{HA} .

Results and Discussion

Comparison of S_{HA} in Different Heteroaromatic Series According to its definition, the S_{HA} value should be close to 1.0 for non-H-bonders and increase with increasing the H-accepting ability. This parameter should constitute a substituent constant specific to each skeletal system. In accord with this, the S_{HA} values in Table 1 demonstrate the following characteristics. (1) S_{HA} values for all alkyl substituents (non-H bonders) are close to unity. (2) S_{HA} values for 2-substituted derivatives are smaller than those for the corresponding 3-substitued derivatives, indicating that the electron-withdrawing property of the ring hetero atom (*S* or *O*) reduces H-accepting abilities of 2-substituents to a greater extent than the 3 substituents. (3) Except for alkyls, S_{HA} values for FR-X are smaller than the corresponding values for TH-X, reflecting the fact that the ring *O* atom is more electron-withdrawing than the ring *S* atom. In fact the S_{HA} values for TH-X are close to those for $Ph-X^{21}$ (monosubstituted benzenes, data not shown) as shown by $S_{HA(Th-X)} = 0.92 S_{HA(Ph-X)} + 0.10$ (*r*= 0.98). This finding conforms to the expectation that heteroatom effects of the ring *S* atom would be small.

Relationship between Retention Factors and Octanol/Water Partition Coefficients By using the parameters given in Table 1, we performed regression analyses of log *k* for thiophenes with methanolic eluents by using Eq. 1 and obtained excellent correlations. Table 3 summarizes the most statistically significant regression coefficients and intercepts at each eluent composition. Although a direct plot of $\log k$ against $\log P$ (Fig. 1A) showed that, in water-rich eluents, strong H-accepting substituents, such as $CO₂R$ and $COMMe₂$, deviate significantly from the linear relationship yielded by non-H-bonders (H and alkyls), addition of the S_{HA} term improved the fit to the extent of providing precise correlations (Fig. 1B, Table 3). The σ_{I} term was statistically insignificant for eluents of all methanol concentrations, conforming to the results previously derived from analyses by using HB_A in place of S_{HA} in Eq. 1.¹⁹⁾

Similar tendencies were observed with analyses of log *k* for the furan series. Even though this series contains a highly H-accepting disubstituted compound like 34, the S_{HA} parameter worked effectively in improving the correlations in water-rich eluents as shown in Table 3 and Fig. 2. The σ_{I} term was again statistically insignificant in all eluents. We believe that the reliability of some log *P* values taken from the literature may be somewhat lower because accurate measurement would be rendered difficult by the high volatility. If

Fig. 1. Relationship between $\log k_{\text{obsd}}$ and $\log k_{\text{calcd}}$ for Thiophenes (TH-X) for M15 Eluent

(A) Plot of log k_{obsd} against log *P*. The numbers represent compound numbers in Table 1. The straight line is drawn through the points for H and alkyls shown by open circles. (B) Plot of log k_{calo} calculated by the correlation in Table 3, against log k_{obsd} . The straight line represents the regression line.

Fig. 2. Relationship between $\log k_{\text{obsd}}$ and $\log k_{\text{calcd}}$ for Furans (FR-X) for M15 Eluent For explanations of symbols see Fig. 1.

it were possible to obtain more accurate $log P$ values, this could even improve the correlations.

By comparing the resultant correlations, we can estimate the contributions of H-bonding effects and so obtain specific information about optimal HPLC conditions for linearity between log *P* and log *k*. Inspection of the correlations given in Table 3 and those so far formulated by Eq. 1 for various (di)azine series²¹⁾ shows that the contributions of the S_{H_A} and σ_{I} (if required) terms, are minimal at 50% MeOH concentration but increase with decreasing methanol concentration in all the series. These additional examples extend our earlier conclusion that the use of eluents containing around 50% MeOH provides the most practical method for predicting $log P.^{14-21)}$

In contrast, many investigations have used $\log k_{\rm W}$ (log *k* at 0% MeOH) as a normalized hydrophobicity parameter.^{8,10)} To resolve the question of which methanol content provides more precise results, we also analyzed $\log k_{\rm W}$ values^{16,19} by Eq. 1. The results are shown as Eqs. 2 and 3.

TH
$$
\log k_{\rm w} = 0.903 \log P + 0.179 S_{\rm HA} - 0.106
$$
 (2)
 $n = 18, r = 0.991, s = 0.079, F = 404$

FR
$$
\log k_{\rm w} = 0.991 \log P + 0.283 S_{\rm HA} - 0.115
$$
 (3)
 $n = 14, r = 0.985, s = 0.093, F = 182$

In both series, the *S*_{HA} term was found to make a greater contribution to $\log k_{\text{W}}$ than to $\log k_{\text{M50}}$, confirming the superiority of the $\log k_{\text{M50}}$ parameter over $\log k_{\text{W}}$.

It is of interest to note that, in the case of thiophenes and furans, not only was the σ _I term insignificant but also the contribution of the S_{HA} term was much smaller than in (di)azines (*e.g.*, $s=0.809$ and $\rho=-0.494$ with M15 for pyrazine series). 21 This can be rationalized in terms of the stronger electronic interactions between the substituent X and the ring hetero atom(s) that should produce more significant changes in H-bonding abilities on both the sites in (di)azines than in furans and thiophenes.

Attempts were also made to apply Eq. 1 to analyzing log *k* for thiophenes measured in acetonitrile solutions. As shown in Table 3, use of the S_{HA} parameter improved the correlations. In both solvent systems, the coefficient of S_{HA} varied from more positive to more negative as the content of organic solvent increased. This means that elution of derivatives with substituents of higher S_{HA} tend to be accelerated relative to the unsubstituted compound (Ar-H) with highly water-rich eluents but to be retarded with eluents of high organic solvent content, suggesting the possibility of reversion of the elution order with the mobile phase composition. Accordingly, for accurate predictions of log *P*, it is very important to select an optimal eluent composition at which the coefficients of S_{HA} and σ_{I} terms are minimal. Since acetonitrile has no Hdonating site, it is only the water in aqueous acetonitrile eluents that bonds to H-accepting sites of the solutes; this would be expected to shift the optimal concentration for aqueous acetonitriles to one more water-rich than that which is optimal for aqueous methanols (50% water). In the present case, such a situation was achieved at around 20% CH₃CN (80%) water). It would be possible to establish the most suitable acetonitrile content in eluents if similar analyses of log *k* were performed for a range of solute systems.

The present work has demonstrated that the S_{HA} parameter effectively works to correlate log *P* with log *k* for various heteroaromatic systems and for different mobile phase systems. By quantifying the H-accepting effects, we have been able systematically to establish optimal HPLC conditions for predicting log *P* with high reliability. In methanolic eluents, we found that intervention of H-accepting effects was minimum at around 50% MeOH regardless of heteroaromatic nucleus. The S_{HA} parameter is also expected to be utilized for describing the characteristics (particularly silanol effects) of different stationary phases, so that an appropriate column free from H-accepting effects can be selected. The definition of a similar H-donor parameter is still to be established. Analyses for compounds containing H-donors are now underway.

Acknowledgments This work was supported in part by a Grant-in-Aid for Scientific Research (No. 15590105) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References and Notes

- 1) A part of this work was based on a lecture presented at LogP2000-The Second Lipophilicity Symposium, University of Lausanne, Lausanne, on 8 March 2000.
- 2) Hansch C., Hoekman D., Gao H., *Chem. Rev.*, **96**, 1045—1076 (1996).
- 3) Testa B., van de Waterbeemd H., Folkers G., Guy R., "Pharmacokinetic Optimization in Drug Research," Wiley-VCH, Weinheim, 2001.
- 4) Fujita T., *Prog. Phys. Org. Chem.*, **14**, 75—113 (1983).
- 5) Leo A., *Chem. Rev.*, **93**, 1281—1306 (1993).
- 6) Chuman H., Mori A., Tanaka H., *Anal. Sci.*, **18**, 1015—1020 (2002).
- 7) Zou J.-W., Zhao W.-N., Shang Z.-C., Huang M.-L., Guo M., Yu Q.-S., *J. Phys. Chem.*, **106**, 11550—11557 (2002).
- 8) Braumann T., *J. Chromatogr.*, **373**, 191—225 (1986).
- 9) Terada H., *Quant. Struct.-Act. Relat.*, **5**, 81—88 (1986).
- 10) Minick D. J., Frenz J. H., Patrick M. A., Brent D. A., *J. Med. Chem.*, **31**, 1923—1933 (1988).
- 11) Lambert W. J., *J. Chromatogr. A.*, **656**, 469—484 (1993).
- 12) Herbert B. J., Dorsey J. G., *Anal. Chem.*, **67**, 744—749 (1995).
- 13) Fujita T., Nishioka T., Nakajima M., *J. Med. Chem.*, **20**, 1071—1081 (1977).
- 14) Yamagami C., "Pharmacokinetic Optimization in Drug Research," Wiley-VCH, Weinheim, 2001, pp. 383—400.
- 15) Yamagami C., Ogura T., Takao N., *J. Chromatogr.*, **514**, 123—136 (1990).
- 16) Yamagami C., Takao N., *Chem. Pharm. Bull.*, **40**, 925—929 (1992).
- 17) Yamagami C., Yokota M., Takao N., *J. Chromatogr. A*, **662**, 49—60 (1994)
- 18) Yamagami C., Yokota M., Takao N., *Chem Pharm Bull.*, **42**, 907—912 (1994).
- 19) Yamagami C., Masaki Y., *Chem. Pharm. Bull.*, **43**, 2238—2242 (1995).
- 20) Yamagami C., Takao N., Fujita T., *J. Pharm. Sci.*, **82**, 155—161 (1993).
- 21) Yamagami C., Kawase K., Fujita T., *Quant. Struct.-Act. Relat.*, **18**, 26—34 (1999).
- 22) Klamt A., Shüürmann G., *J. Chem. Soc.*, *Perkin Trans. 2*, **1993**, 799— 805 (1993).
- 23) Charton M., *Prog. Phys. Org. Chem.*, **13**, 119—251 (1981).
- 24) Abraham M. H., Grellier P. L., Prior D. V., Duce P. P., Morris J. J., Taylor P. J., *J. Chem. Soc.*, *Perkin Trans. 2*, **1989**, 699—711 (1989).
- 25) Abraham M. H., *J. Phys. Org. Chem.*, **6**, 660—684 (1993).
- 26) Abraham M. H., Chadha H. S., Leo A. J., *J. Chromatogr. A*, **685**, 203—211 (1994).
- 27) Lewis S. J., Mirrlees M. S., Taylor P. J., *Quant. Struct.-Act. Relat.*, **2**, 100—111 (1983).
- 28) Yamagami C., Takao N., *Chem. Pharm. Bull.*, **41**, 694—698 (1993).
- 29) Dewar M. J. S., Zoebisch E. G., Healy E. F., Stewart J. J. P., *J. Am. Chem. Soc.*, **107**, 3902—3909 (1985).
- 30) Stewart J. J. P., Fujitsu Ltd., 1993, Tokyo.