

J-CAMD 296

## Calculation of hydrophobic parameters directly from three-dimensional structures using comparative molecular field analysis

Ki Hwan Kim

*Computer Assisted Molecular Design, Pharmaceutical Products Division, Abbott Laboratories, 100 Abbott Park Road, Abbott Park, IL 60064, U.S.A.*

Received 25 January 1995

Accepted 9 May 1995

**Keywords:** 3D QSAR; CoMFA; GRID; Hydrogen-bond potential; Partial least squares; Hydrophobicity; Reversed-phase liquid chromatography; Capacity ratio; Triazine; Furans

---

### Summary

Capacity ratio ( $\log k'$ ) values, which are a measure of hydrophobicity, were calculated directly from the three-dimensional structures of 17 furans and 54 triazines using the comparative molecular field analysis approach. The  $H_2O$  probe and the GRID force field, including hydrogen-bond potentials, yielded excellent correlations with the  $\log k'$  values. Moreover, the predicted values of  $\log k'$  from 14 additional triazine analogs showed excellent agreement with  $\log k'$  values reported in the literature. Similar results were obtained for the octanol–water partition coefficients ( $\log P$ ) of 17 furans.

---

### Introduction

Partition coefficients, usually expressed as  $\log P$  values, have been used as measures of hydrophobicity [1]. Traditionally, these constants were obtained by the shake-flask method between an aqueous and an immiscible organic solvent – usually 1-octanol [1]. Measurement of  $\log P$  is always recommended. However, this method is time consuming and often encounters problems such as association, adsorption, micelle and/or emulsion formation, impurity effects, as well as the limitation of  $\log P$  to values between about  $-4.5$  and  $+4.5$  by the shake-flask technique [2], although the upper limit of the measurement can be extended to about  $8.0$  by the slow-stir technique [3]. Additionally, compounds that have yet to be synthesized cannot be measured. Such difficulties have resulted in the proposal of many calculation methods [4].

An alternative method for calculating  $\log P$ , the logarithm of the capacity factor ( $\log k'$ ) obtained from reversed-phase high-performance liquid chromatography (RP-HPLC), has been frequently used. In this method, the HPLC retention time is transformed into  $\log k'$  by Eq. 1, where  $t_R$  is the retention time of the compound and  $t_o$  is the column dead time.

$$\log k' = \log (t_R / t_o - 1) \quad (1)$$

Many investigators have shown that  $\log k'$  is directly proportional to  $\log P$  according to the relationship in Eq. 2 [5–14] and have suggested that  $\log k'$  can be used to estimate  $\log P$ . However, the relationship between  $\log P$  and  $\log k'$  is often complicated by the hydrogen-bonding ability of compounds under the chromatographic conditions employed. Specifically, the hydrogen-bonding capability of a compound is often influenced by the type of stationary phase, as well as the composition of the mobile phase in RP-HPLC. Thus, Eq. 3 describes a generalized relationship between  $\log P$  and  $\log k'$ , which includes hydrogen-bonding descriptors. HA and HD are the indicator variables for hydrogen-bonding acceptors and hydrogen-bonding donors, respectively. In heteroaromatic systems, this situation is further complicated by the addition of the electronic interaction term, as described in Eq. 4 [6,7].

$$\log P = a \log k' + b \quad (2)$$

$$\log P = a \log k' + b HA + c HD + d \quad (3)$$

$$\log P = a \log k' + b HA + c HD + d \sigma + e \quad (4)$$

In our previous study [15], we proposed that the comparative molecular field analysis (CoMFA) [16] method,

using the H<sub>2</sub>O probe and the hydrogen-bond potential in GRID [17], can be used to describe the hydrophobic properties of organic molecules in quantitative structure–activity relationship (QSAR) studies. Since then, a number of examples have been presented where such molecular fields reproduced the traditional QSAR hydrophobic parameter log P or  $\pi$  in both linear and nonlinear relationships [18–24]. However, the previous studies have two limitations. First, these CoMFA applications simply illustrated that the CoMFA approach can reproduce the traditional QSAR and failed to demonstrate its applicability in calculating hydrophobic parameters such as log P. Second, most of the examples, particularly those for the nonlinear relationships, included only simple or linear alkyl derivatives. Therefore, it was imperative to address this subject. In our previous studies, we noticed several cases involving high correlations between the CoMFA descriptors and log P as given in Eqs. 5–8, despite the fact that the CoMFA descriptors were not derived for  $\pi$  or log P [23].

$$\pi = 0.112 Z1_{\text{H}_2\text{O}} + 0.049 Z2_{\text{H}_2\text{O}} + 0.423 \quad (5)$$

$$n = 11, s = 0.254, r = 0.958, s_{\text{cv}} = 0.427$$

$$\pi = 0.149 Z1_{\text{H}_2\text{O}} + 0.149 \quad (6)$$

$$n = 8, s = 0.261, r = 0.944, s_{\text{cv}} = 0.327$$

$$\pi = 0.137 Z1_{\text{H}_2\text{O}} + 0.101 Z2_{\text{H}_2\text{O}} + 0.047 Z3_{\text{H}_2\text{O}} + 0.078 Z4_{\text{H}_2\text{O}} + 0.535 \quad (7)$$

$$n = 22, s = 0.431, r = 0.952, s_{\text{cv}} = 0.524$$

$$\log P = 0.141 Z1_{\text{H}_2\text{O}} + 1.174 \quad (8)$$

$$n = 8, s = 0.098, r = 0.998, s_{\text{cv}} = 0.114$$

In order to extend the applicability of CoMFA in calculating the hydrophobic parameter for QSAR studies, we began our investigation using 25 simple benzene derivatives [25]. Equation 9 shows the result. This two-com-

ponent CoMFA model was able to explain 91% of the variance in log P of benzene derivatives. The standard errors of estimation from the fitted model and its leave-one-out cross-validation test were somewhat larger than expected. Nevertheless, Eq. 9 provided an indication for the possible applicability of CoMFA in calculating log P values in other systems.

$$\log P = 0.182(\pm 0.016) Z1_{\text{H}_2\text{O}} + 0.075(\pm 0.015) Z2_{\text{H}_2\text{O}} + 2.401(\pm 0.074) \quad (9)$$

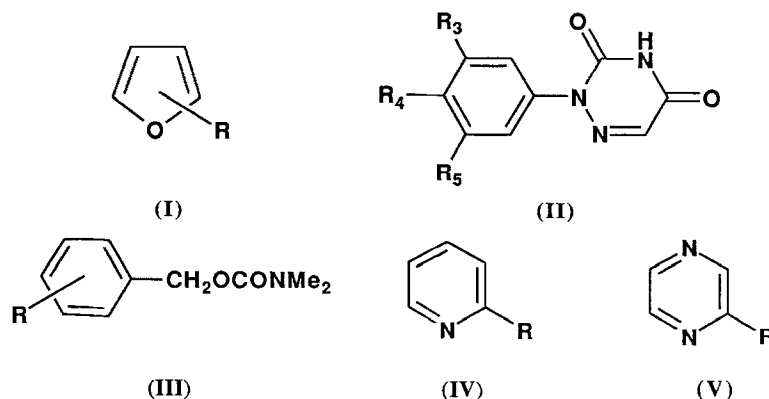
$$n = 25, s = 0.369, r = 0.952, s_{\text{cv}} = 0.504$$

Since an estimation of hydrophobicity in heterocyclic systems is often more complicated [26–30], the investigation was extended toward estimating the hydrophobicity in various systems (I–V). In light of our experience with the hydrogen-bond potential function in describing hydrophobic interactions in biological QSAR, we examined the possibility of calculating the capacity factors (log  $k'$ ) directly from the three-dimensional structures of compounds, using the CoMFA approach. In particular, we are interested in describing the hydrophobicity of molecules in those situations where certain conformations are chosen to represent the bioactive conformations of the ligands that interact with the receptor of interest. In this report, an account of the results for systems I and II is described in more detail. The partial least squares (PLS) method was used in the data analysis [31].

## Method

### Chromatographic capacity factor ( $k'$ )

In the original determination of the retention time of triazines with RP-HPLC, methanol solutions of the triazines were used in a C18 Nova Pak RP-HPLC column. The mobile phase was acetonitrile:1% aqueous acetic acid (1:1) at a flow rate of 1.0 ml/min. The measured retention times were transformed into the logarithm of their capacity ratio (log  $k'$ ) using Eq. 1. Detailed experimental conditions can be found in the original paper by



Scheme 1. Structures of compounds I–V.

McFarland et al. [32]. The reported  $\log k'$  values of furans [33] were originally measured using a commercial Capcellpack C18 column; 50% (v/v) methanol in water was used as the eluent.

#### Partition coefficients

*n*-Octanol–water  $\log P$  values were taken from the literature [33].

#### Molecular modeling

The coordinates of the triazines reported in the previous study were used without any modification. The geometries of the molecules were created and minimized using the SYBYL force field [34]. The charges for all atoms were generated using the Gasteiger–Hückel method. The conformation of each compound was matched to that of diclazuril as closely as possible. The 2-phenyl-1,2,4-triazine-3,5-dione group was overlapped for the CoMFA analysis. Detailed accounts of the molecular modeling and superposition of the molecules have been given in the original publication [35].

The starting coordinates of substituted furans were generated using the graphics modeling package for small molecules at Abbott Laboratories. All geometric variables were optimized with the AM1 method of the AMPAC package [36] on a VAX 9000 computer.

#### CoMFA interaction energy calculation

The hydrophobic, steric and electrostatic potential energy fields of each molecule were calculated at various lattice points surrounding the molecule using  $H_2O$ ,  $CH_3$  and  $H^+$  probes with the program GRID [17]. The  $H_2O$  probe can accept two hydrogen bonds and is capable of donating two hydrogen bonds.

For each substituted triazine or furan molecule, the energies at a total of 2340 or 2100 grid points were calculated, respectively. The CoMFA grid spacing was 2.0 Å in the *x*-, *y*- and *z*-directions. The grid box spanned from *x* = −15.5 to 8.5, *y* = −18.5 to 9.5 and *z* = −14.5 to 7.5 for triazines and from *x* = −10 to 16, *y* = −13 to 15 and *z* = −9 to 9 for furans, respectively. All the lattice points where repulsive steric and electrostatic interaction energies exceeded 4 kcal/mol were truncated to 4 kcal/mol. Any lattice point for which the standard deviation of the energies was less than 0.05 was discarded. The procedures reduced the number of lattice points from the  $H_2O$ ,  $CH_3$  and  $H^+$  probes to 300, 287 and 1932 for the triazines and 149, 163 and 1419 for the furans, respectively.

#### Partial least squares calculations

First, 10 orthogonal latent variables were extracted by the standard PLS algorithm [31]. The 'best' correlation model was chosen on the basis of which model significantly minimized the sum of squares of the difference in activities between the predicted and observed values

(PRESS), using predictions made from a leave-one-out cross-validation test. The final model was further validated by the overall and the stepwise F-statistics. The variables Z1–Z5 in the correlation equations correspond to the first through the fifth latent variables. The probe used in the calculation is indicated as a subscript after each of the latent variable symbols Z. For example, Z1<sub>H<sub>2</sub>O</sub>–Z5<sub>H<sub>2</sub>O</sub> are the first through the fifth latent variables from the  $H_2O$  probe in the PLS analysis.

## Results and Discussion

#### Capacity factors ( $\log k'$ ) of furan analogs

From the 17 monosubstituted 2-furan compounds (I) in Table 1, Eq. 10 was obtained. The experimentally determined  $\log k'$  values of the furan analogs used in the correlation are given in Table 1. Only monosubstituted analogs were used in the analysis; one disubstituted compound (2-Br-5-CO<sub>2</sub>Me) reported in the original paper [33] was not included. In Eq. 10 and the other correlation equations, L is the optimum number of components in the CoMFA model, n is the number of compounds used in the correlation,  $R^2$  and  $R_{cv}^2$  are the squared correlation coefficients of the fitted model and the leave-one-out cross-validated test, and s and  $s_{cv}$  are the standard errors of estimation of the fitted model and the cross-validated test, respectively. F and p are the F-statistics and significance probability of the model. The percentages following the equation present the amount of variance given by each of the latent variables in the model.

$$\begin{aligned} \log k' (\text{furan}) = & 0.103(\pm 0.006) Z1_{H_2O} + 0.027(\pm 0.003) Z2_{H_2O} \\ & + 0.058(\pm 0.007) Z3_{H_2O} + 0.068(\pm 0.010) Z4_{H_2O} \\ & + 0.040(\pm 0.011) Z5_{H_2O} - 0.030(\pm 0.023) \end{aligned} \quad (10)$$

$$\begin{aligned} L = 5, n = 17, s = 0.093, R^2 = 0.978, s_{cv} = 0.434, \\ R_{cv}^2 = 0.541, F = 99.6, p = 0.0001 \\ (54\%, 18\%, 14\%, 9\%, 3\%) \end{aligned}$$

The five-component model (Eq. 10) explains 98% of the variance in  $\log k'$  values with a cross-validated  $R^2$  of 0.541. The correlation is good and the  $R_{cv}^2$  and  $s_{cv}$  are reasonable. The first component accounts for 54% of the variance, whereas the second through the fifth components explain an additional 18%, 14%, 9%, and 3%, respectively. The corresponding four-component model explains 95% of the variance in the data. The excellent correlation in the absence of the hydrogen-bonding indicator variables, as indicated in Eq. 4, shows that the  $H_2O$  probe with the GRID hydrogen-bond potential function adequately describes the hydrogen-bonding effects of various substituents in the  $\log k'$  values.

Since the electronic interactions between the ring oxygen atom and the ring substituent can influence the chromatographic behavior of these compounds in the RP-

TABLE 1  
OBSERVED AND CALCULATED LOG  $k'$  VALUES OF SUBSTITUTED FURANS USING EQ. 10

No.	Substituent	Log $k'$			PLS latent variables				
		Obs.	Calc.	Dev.	Z1 <sub>H<sub>2</sub>O</sub>	Z2 <sub>H<sub>2</sub>O</sub>	Z3 <sub>H<sub>2</sub>O</sub>	Z4 <sub>H<sub>2</sub>O</sub>	Z5 <sub>H<sub>2</sub>O</sub>
1	2-CO <sub>2</sub> Et	0.395	0.381	0.014	0.591	-7.565	6.017	2.460	1.013
2	2-CO <sub>2</sub> Me	0.087	-0.045	0.132	-0.242	-8.865	1.877	0.732	2.309
3	2-COMe	-0.237	-0.203	-0.034	-0.128	-8.548	-0.629	2.869	-2.155
4	2-CONHEt	-0.324	-0.231	-0.093	0.236	-8.524	3.569	-2.846	-0.185
5	2-CONH <sub>2</sub>	-0.806	-0.884	0.078	-1.015	-8.979	-7.645	0.056	-1.674
6	2-CONMe <sub>2</sub>	-0.280	-0.408	0.128	1.008	-8.066	1.471	-4.592	-0.897
7	2-CONHMe	-0.576	-0.421	-0.155	-0.968	-10.315	0.496	-1.439	1.452
8	2-Et	1.002	1.030	-0.028	10.530	9.738	-2.084	-3.008	0.915
9	2-Me	0.648	0.582	0.066	4.471	2.976	-1.950	1.618	1.852
10	2-OMe	0.374	0.419	-0.045	3.379	-1.040	1.048	3.141	-3.621
11	3-CO <sub>2</sub> Et	0.573	0.508	0.065	-1.557	12.074	3.912	1.854	0.446
12	3-CO <sub>2</sub> Me	0.260	0.242	0.018	-2.285	9.822	-1.926	1.204	1.167
13	3-CONHEt	-0.324	-0.343	0.019	-3.198	9.498	0.978	-1.863	-4.313
14	3-CONH <sub>2</sub>	-0.717	-0.698	-0.019	-4.484	0.040	-6.522	1.064	2.458
15	3-CONMe <sub>2</sub>	-0.351	-0.301	-0.050	-4.266	6.425	0.347	-1.961	2.667
16	3-CONHMe	-0.525	-0.509	-0.016	-4.201	7.107	-0.976	-1.810	-1.530
17	H	0.291	0.372	-0.081	2.130	4.223	-1.833	2.518	0.096

HPLC system [33], potential electronic influences were investigated using the H<sup>+</sup> probe in the interaction energy calculations. In previous CoMFA studies [37–39], the H<sup>+</sup> probe was used in describing the electronic effects of substituents on pK<sub>a</sub> values.

Addition of electronic descriptors to Eq. 10 yielded Eq. 11. The result shows that the electronic descriptors do not improve the correlation obtained earlier. The negligible improvement of Eq. 11 over Eq. 10 indicates that the electronic influences of the substituents that had not been accounted for by the H<sub>2</sub>O probe and the hydrogen-bonding potentials are minimal.

$$\log k' (\text{furan}) = 0.076(\pm 0.004) Z1_{\text{H}_2\text{O},\text{H}^+} + 0.020(\pm 0.003) Z2_{\text{H}_2\text{O},\text{H}^+} + 0.043(\pm 0.005) Z3_{\text{H}_2\text{O},\text{H}^+} + 0.059(\pm 0.009) Z4_{\text{H}_2\text{O},\text{H}^+} + 0.038(\pm 0.009) Z5_{\text{H}_2\text{O},\text{H}^+} - 0.030(\pm 0.020) \quad (11)$$

$$L = 5, n = 17, s = 0.083, R^2 = 0.983, s_{\text{cv}} = 0.323, R_{\text{cv}}^2 = 0.693, F = 127.8, p = 0.0001 (67\%, 8\%, 14\%, 6\%, 3\%)$$

Table 1 lists the calculated log  $k'$  values using Eq. 10, along with the latent variables used and the corresponding experimentally measured log  $k'$  values.

Molecular mechanics potentials, also employed in this study, may be divided into two groups. One group contains a hydrogen-bonding term in addition to the common steric and electrostatic terms. Then, an exponential or 12–10 function is often used to describe the hydrogen-bonding interactions. The other group does not contain a specific hydrogen-bonding term. Such potentials generally rely on the explicit representation of hydrogen atoms and lone-pair electrons to account for the hydro-

gen-bonding directionality. The GRID force field, which was used in this study, includes a hydrogen-bonding term. When the hydrogen-bonding term is not included, it is often assumed that a suitable sampling of the steric and electrostatic fields might provide all the information necessary for understanding their observed biological properties [16]. In order to determine whether the steric and electrostatic fields would provide correlations equivalent to those obtained with the H<sub>2</sub>O probe in the present situation, CH<sub>3</sub> and H<sup>+</sup> probes were also used in the analysis. Equations 12 and 13 summarize the results, demonstrating that the CH<sub>3</sub> and H<sup>+</sup> probes failed to provide correlations comparable to those obtained using the H<sub>2</sub>O probe.

$$\log k' (\text{furan}) = 0.084(\pm 0.019) Z1_{\text{CH}_3} + 0.032(\pm 0.011) Z2_{\text{CH}_3} + 0.050(\pm 0.020) Z3_{\text{CH}_3} - 0.030(\pm 0.074) \quad (12)$$

$$L = 3, n = 17, s = 0.306, R^2 = 0.726, s_{\text{cv}} = 0.504, R_{\text{cv}}^2 = 0.138, F = 11.5, p = 0.0006 (40\%, 20\%, 13\%)$$

$$\log k' (\text{furan}) = 0.071(\pm 0.010) Z1_{\text{CH}_3,\text{H}^+} + 0.023(\pm 0.008) Z2_{\text{CH}_3,\text{H}^+} + 0.040(\pm 0.013) Z3_{\text{CH}_3,\text{H}^+} - 0.030(\pm 0.057) \quad (13)$$

$$L = 3, n = 17, s = 0.235, R^2 = 0.837, s_{\text{cv}} = 0.369, R_{\text{cv}}^2 = 0.538, F = 22.3, p = 0.0001 (61\%, 13\%, 10\%)$$

The results described in this study demonstrate the applicability of CoMFA in calculating log  $k'$  values of a heterocyclic system directly from three-dimensional structures. Similar results (Eqs. 14–16) were obtained for sub-

stituted benzyl *N,N*-dimethylcarbamates (**III**), substituted pyridines (**IV**) and pyrazines (**V**). Details will be described elsewhere.

$$\begin{aligned} \log k' (\text{carbamate}) = & \\ & 0.083(\pm 0.008) Z1_{\text{H}_2\text{O}} + 0.042(\pm 0.006) Z2_{\text{H}_2\text{O}} \\ & + 0.072(\pm 0.011) Z3_{\text{H}_2\text{O}} + 0.032(\pm 0.012) Z4_{\text{H}_2\text{O}} \\ & + 0.830(\pm 0.029) \end{aligned} \quad (14)$$

$L = 4$ ,  $n = 19$ ,  $s = 0.125$ ,  $R^2 = 0.939$ ,  $s_{\text{cv}} = 0.368$ ,  
 $R_{\text{cv}}^2 = 0.355$ ,  $F = 54.0$ ,  $p = 0.0001$   
 (53%, 20%, 18%, 3%)

$$\begin{aligned} \log k' (\text{pyridine}) = & \\ & 0.066(\pm 0.004) Z1_{\text{H}_2\text{O}} + 0.032(\pm 0.004) Z2_{\text{H}_2\text{O}} \\ & + 0.044(\pm 0.006) Z3_{\text{H}_2\text{O}} + 0.013(\pm 0.005) Z4_{\text{H}_2\text{O}} \\ & + 0.002(\pm 0.018) \end{aligned} \quad (15)$$

$L = 4$ ,  $n = 19$ ,  $s = 0.080$ ,  $R^2 = 0.970$ ,  $s_{\text{cv}} = 0.261$ ,  
 $R_{\text{cv}}^2 = 0.609$ ,  $F = 113.1$ ,  $p = 0.0001$   
 (71%, 12%, 12%, 3%)

$$\begin{aligned} \log k' (\text{pyrazine}) = & \\ & 0.070(\pm 0.007) Z1_{\text{H}_2\text{O}} + 0.058(\pm 0.009) Z2_{\text{H}_2\text{O}} \\ & + 0.029(\pm 0.010) Z3_{\text{H}_2\text{O}} + 0.028(\pm 0.014) Z4_{\text{H}_2\text{O}} \\ & - 0.371(\pm 0.035) \end{aligned} \quad (16)$$

$L = 4$ ,  $n = 19$ ,  $s = 0.154$ ,  $R^2 = 0.910$ ,  $s_{\text{cv}} = 0.322$ ,  
 $R_{\text{cv}}^2 = 0.518$ ,  $F = 35.3$ ,  $p = 0.0001$   
 (58%, 25%, 5%, 3%)

In each case, the four-component CoMFA model accounts for 91–97% of the variance in  $\log k'$  values. These results, along with those for the furans (Eq. 10), demonstrate the applicability of CoMFA in calculating  $\log k'$  values, both in carbocyclic systems as well as in various heterocyclic systems.

#### Capacity factors ( $\log k'$ ) of triazine analogs

McFarland et al. [32] used the  $\log k'$  values of 54 compounds as a measure of hydrophobicity in their study of linear discriminants and multiple regression analyses of anticoccidial triazines (**II**). These compounds are listed in Table 2. In the subsequent 3D QSAR study [35], McFarland reported the  $\log k'$  values of 14 additional triazines, calculated from a correlation equation between  $\log k'$  and  $\log P$ . This set is especially interesting, not only because it provides additional evidence for the approach, but also because a larger number of compounds and a wider array of structures are involved. Moreover, the calculated  $\log k'$  values of 14 additional compounds from the derived CoMFA model can be compared with the previously reported values based on a QSAR correlation. Furthermore, the conformation of each compound has already been chosen and its superposition has already been decided for the 3D QSAR study. Thus, the hydrophobicity of the compound set for a 3D QSAR environment can be examined with this data set.

Taking the 54 triazines and utilizing the GRID  $\text{H}_2\text{O}$  probe, Eq. 17 was obtained. The five-component model explains 96% of the variance in  $\log k'$  values, with a cross-validated  $R^2$  of 0.865. The first component accounts for 56%, and the second through fifth components explain an additional 18%, 17%, 3% and 2%, respectively. The low  $s_{\text{cv}}$  and high  $R_{\text{cv}}^2$  values indicate the model's successful predictability within the set of compounds included in the analysis. The corresponding three- and four-component models explain 91% and 94% of the variance in the data, respectively.

$$\begin{aligned} \log k' = & \\ & 0.036(\pm 0.001) Z1_{\text{H}_2\text{O}} + 0.043(\pm 0.003) Z2_{\text{H}_2\text{O}} \\ & + 0.038(\pm 0.003) Z3_{\text{H}_2\text{O}} + 0.019(\pm 0.003) Z4_{\text{H}_2\text{O}} \\ & + 0.024(\pm 0.005) Z5_{\text{H}_2\text{O}} + 0.464(\pm 0.014) \end{aligned} \quad (17)$$

$L = 5$ ,  $n = 54$ ,  $R^2 = 0.958$ ,  $s = 0.100$ ,  
 $R_{\text{cv}}^2 = 0.865$ ,  $s_{\text{cv}} = 0.172$ ,  $F = 216.3$ ,  $p = 0.0001$   
 (56%, 18%, 17%, 3%, 2%)

Addition of electrostatic effects using an  $\text{H}^+$  probe (Eq. 18) did not improve the correlation obtained before [40]. The results indicate that there are very little electronic influences from substituents on  $\log k'$  values that have not been dealt with by the  $\text{H}_2\text{O}$  probe and the hydrogen-bonding potential.

$$\begin{aligned} \log k' = & \\ & 0.037(\pm 0.001) Z1_{\text{H}_2\text{O},\text{H}^+} + 0.049(\pm 0.003) Z2_{\text{H}_2\text{O},\text{H}^+} \\ & + 0.026(\pm 0.002) Z3_{\text{H}_2\text{O},\text{H}^+} + 0.020(\pm 0.003) Z4_{\text{H}_2\text{O},\text{H}^+} \\ & + 0.019(\pm 0.004) Z5_{\text{H}_2\text{O},\text{H}^+} + 0.464(\pm 0.012) \end{aligned} \quad (18)$$

$L = 5$ ,  $n = 54$ ,  $R^2 = 0.965$ ,  $s = 0.091$ ,  
 $R_{\text{cv}}^2 = 0.875$ ,  $s_{\text{cv}} = 0.166$ ,  $F = 265.6$ ,  $p = 0.0001$   
 (56%, 24%, 10%, 5%, 2%)

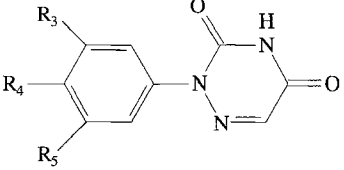
Table 2 lists the experimentally determined  $\log k'$  values, the latent variables used and the corresponding calculated  $\log k'$  values from Eq. 17. Because of the minor contribution of the electronic effects (Eq. 18), Eq. 17 was used in the calculation instead of Eq. 18. Figure 1 shows a plot between the observed and calculated  $\log k'$  values from Eq. 17.

In order to examine whether the steric and electrostatic fields can provide correlations equivalent to those obtained with the  $\text{H}_2\text{O}$  probe in the present situation, the  $\text{CH}_3$  and  $\text{H}^+$  probes were also used in the analysis. The results are presented in Eqs. 19 and 20.

$$\begin{aligned} \log k' = & \\ & 0.038(\pm 0.002) Z1_{\text{CH}_3} + 0.065(\pm 0.006) Z2_{\text{CH}_3} \\ & + 0.023(\pm 0.004) Z3_{\text{CH}_3} + 0.030(\pm 0.007) Z4_{\text{CH}_3} \\ & + 0.464(\pm 0.021) \end{aligned} \quad (19)$$

$L = 4$ ,  $n = 54$ ,  $R^2 = 0.895$ ,  $s = 0.156$ ,  
 $R_{\text{cv}}^2 = 0.745$ ,  $s_{\text{cv}} = 0.236$ ,  $F = 216.3$ ,  $p = 0.0001$   
 (56%, 23%, 6%, 4%)

TABLE 2  
OBSERVED AND CALCULATED LOG  $k'$  VALUES OF TRIAZINES USING EQ. 17

												
No.	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Log $k'$			Z <sub>1H<sub>2</sub>O</sub>	Z <sub>2H<sub>2</sub>O</sub>	Z <sub>3H<sub>2</sub>O</sub>	Z <sub>4H<sub>2</sub>O</sub>	Z <sub>5H<sub>2</sub>O</sub>	
				Obs.	Calc.	Dev.						
1	CP88817	Cl	CH(CN)C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.78	0.917	-0.137	7.900	3.153	2.551	-3.914	0.264
2	CP25810	Cl	SC <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	1.25	1.183	0.067	7.458	4.730	5.924	-2.438	2.744
3	CP26541	Me	SC <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	1.12	1.161	-0.041	7.593	4.538	5.925	-2.897	2.285
4	CP35759	Me	SC <sub>6</sub> H <sub>4</sub> -4-Ac	Cl	0.58	0.561	0.019	6.292	2.075	-1.537	-5.903	-2.216
5	CP93646	Cl	CH <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	1.21	0.993	0.217	8.688	2.477	2.202	-2.007	2.533
6	CP26245	Cl	SO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.60	0.595	0.005	-9.748	10.650	-0.153	0.058	1.614
7	CP89866	Cl	COC <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.84	0.762	0.078	7.574	-0.072	-1.965	-1.469	5.391
8	CP25673	Me	SC <sub>6</sub> H <sub>4</sub> -4-Cl	Me	1.11	1.158	-0.048	7.270	4.697	5.755	-2.600	2.400
9	CP25722	Me	SO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	Me	0.49	0.521	-0.031	-10.236	10.263	-1.073	0.138	1.360
10	CP25720	Me	S(O)C <sub>6</sub> H <sub>4</sub> -4-Cl	Me	0.24	0.386	-0.146	-3.799	7.543	-3.402	-3.832	-2.474
11	CP29493	Me	OC <sub>6</sub> H <sub>4</sub> -4-CH(OH)Me	Cl	0.99	0.987	0.003	8.872	0.902	1.702	3.019	1.700
12	CP34094	Me	OC <sub>6</sub> H <sub>4</sub> -4-Ac	Cl	0.34	0.456	-0.116	9.624	-4.256	-6.063	2.657	0.206
13	CP35092	Me	OC <sub>6</sub> H <sub>3</sub> -3-Me-4-SMe	Me	0.12	0.295	-0.175	8.380	-4.741	-7.043	-0.226	0.025
14	CP25431	H	CH(OH)C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	1.02	0.953	0.067	7.414	2.492	3.835	0.737	-2.031
15	CP25641	H	SC <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.27	0.338	-0.068	5.539	0.268	-4.133	-3.734	-4.724
16	CP29510	Me	OC <sub>6</sub> H <sub>4</sub> -4-SMe	Me	0.81	0.742	0.068	8.270	-1.187	-1.546	1.994	2.076
17	CP29360	H	OC <sub>6</sub> H <sub>4</sub> -4-SMe	Me	0.69	0.635	0.055	7.325	-1.998	-2.035	5.219	-1.279
18	CP30542	Me	OC <sub>6</sub> H <sub>4</sub> -2-Cl-4-SO <sub>2</sub> -NMeEt	Cl	0.77	0.733	0.037	13.564	-0.779	1.330	-10.986	-1.658
19	CP25636	H	OC <sub>6</sub> H <sub>4</sub> -4-SMe	Cl	0.92	0.966	-0.046	9.996	-0.971	0.430	3.907	3.777
20	CP28412	Me	OC <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.72	0.657	0.063	7.824	-1.868	-1.714	5.025	-1.729
21	CP29401	H	OC <sub>6</sub> H <sub>4</sub> -4-SO <sub>2</sub> Me	Me	-0.06	-0.125	0.065	7.141	-7.544	-10.821	0.074	-4.865
22	CP24115	H	SO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	H	0.26	0.229	0.031	-13.359	7.822	-3.383	6.077	-2.673
23	CP19199	Cl	H	Cl	0.33	0.151	0.179	-9.298	-4.429	5.863	-2.675	1.690
24	CP26263	H	O(naphth-2-yl-6-Br)	Cl	0.93	0.784	0.146	9.832	-2.269	-1.200	6.376	-0.608
25	CP28664	H	OC <sub>6</sub> H <sub>4</sub> -4-I	Cl	1.26	1.350	-0.090	10.528	4.451	9.199	5.051	-5.622
26	CP23539	Me	SO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Br	Cl	0.30	0.266	0.034	-13.023	8.112	-3.071	6.226	-2.808
27	CP25501	H	OC <sub>6</sub> H <sub>3</sub> -2,4-Cl <sub>2</sub>	Cl	0.96	0.986	-0.026	10.264	0.065	1.297	4.684	0.350
28	CP26859	H	SO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	H	0.54	0.593	-0.053	-9.651	10.638	-0.088	-0.021	1.377
29	CP26746	H	OC <sub>6</sub> H <sub>4</sub> -4-Br	Me	0.81	0.834	-0.024	9.077	-1.990	-0.747	7.543	0.511
30	CP25498	H	OC <sub>6</sub> H <sub>3</sub> -2,4-Cl <sub>2</sub>	Me	0.96	0.974	-0.014	9.824	0.097	1.016	5.040	0.658
31	CP91349	H	CH <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.95	0.865	0.085	7.761	1.574	1.695	0.643	-1.065
32	CP33320	Me	OC <sub>6</sub> H <sub>3</sub> -2-Cl-4-SO <sub>2</sub> NH-c-C <sub>3</sub> H <sub>5</sub>	Cl	0.54	0.435	0.105	13.473	-1.778	-2.859	-12.079	-4.653
33	CP23857	Me	SO <sub>2</sub> N(CH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> O	Me	-0.11	-0.030	-0.080	-6.682	2.472	-9.661	-4.883	4.459
34	CP25842	Me	COC <sub>6</sub> H <sub>4</sub> -4-Cl	Me	0.63	0.729	-0.099	6.984	-0.248	-2.279	-1.316	5.611
35	CP19940	H	Br	CF <sub>3</sub>	0.37	0.217	0.153	-10.202	-2.068	5.756	-0.235	-0.096
36	CP18826	H	F	CF <sub>3</sub>	0.18	0.182	-0.002	-9.961	-3.552	5.779	00.849	-0.154
37	CP21745	Me	H	Cl	0.19	0.155	0.035	-9.195	-4.374	5.910	-2.693	1.559
38	CP27207	Cl	SO <sub>2</sub> N(CH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> O	Cl	0.01	-0.043	0.053	-6.427	1.816	-9.724	-4.901	4.781
39	CP23316	H	SO <sub>2</sub> N(CH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> O	Cl	-0.15	-0.243	0.093	-7.730	0.531	-10.650	-1.257	-0.709
40	CP24431	H	SO <sub>2</sub> Ph	H	0.00	-0.048	0.048	-13.898	4.438	-6.456	4.320	-1.151
41	CP25356	H	CH(OH)Ph	Me	-0.08	-0.016	-0.064	0.464	-1.942	-6.809	-3.976	-3.369
42	CP18737	H	H	CF <sub>3</sub>	0.12	0.096	0.024	-10.705	-4.289	6.033	0.137	-1.198
43	CP20726	Me	H	Me	0.07	0.173	-0.103	-9.348	-4.122	5.523	-2.164	2.299
44	CP21262	H	H	CN	-0.25	-0.026	-0.224	-9.981	-5.692	5.406	-1.591	-2.537
45	CP25137	H	OC <sub>6</sub> H <sub>4</sub> -4-Cl	Me	0.75	0.879	-0.129	8.725	-1.548	-0.024	7.773	0.811
46	CP28689	H	S(naphth-2-yl)	Cl	1.19	1.197	-0.007	7.084	5.877	9.088	-1.836	-3.766
47	CP18827	H	H	Cl	0.01	0.017	-0.007	-10.074	-5.148	5.298	-0.806	-2.036
48	CP20807	Cl	OMe	Cl	0.24	0.260	-0.020	-6.833	-3.896	3.460	-0.732	3.923
49	CP22782	H	NO <sub>2</sub>	H	-0.11	-0.231	0.121	-15.380	-2.918	1.660	-0.735	-2.482
50	CP21428	H	OPh	H	0.33	0.198	0.132	-0.015	-5.928	-3.676	4.358	1.977



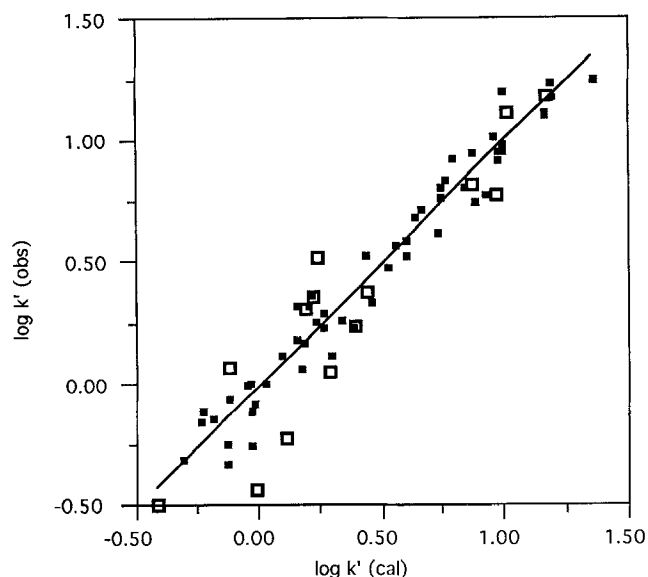


Fig. 1. Observed and predicted  $\log k'$  values of 68 triazines using Eq. 17. The 14 forecasted compounds are indicated with open squares.

lation is similar to the result obtained from the pyrazine series [42], except for the electronic parameter; an electronic parameter  $\sigma_1$  improved the correlation in the pyridine series, whereas it was not required in the pyrazine series.

$$\log k'(\text{furan}) = 0.635 \log P - 0.195 \text{HB}_{\text{AM}} - 0.544 \quad (22)$$

$$n = 18, R^2 = 0.999, s = 0.022$$

The experimentally determined  $\log P$  values of the furan analogs used in the correlations are listed in Table 5 [33]. From the 17 monosubstituted 2-furan compounds, Eq. 23 was derived.

$$\begin{aligned} \log P(\text{furan}) = & 0.173(\pm 0.009) Z1_{\text{H}_2\text{O}} + 0.026(\pm 0.004) Z2_{\text{H}_2\text{O}} \\ & + 0.065(\pm 0.007) Z3_{\text{H}_2\text{O}} + 0.088(\pm 0.012) Z4_{\text{H}_2\text{O}} \\ & + 0.047(\pm 0.013) Z5_{\text{H}_2\text{O}} + 0.922(\pm 0.028) \end{aligned} \quad (23)$$

$$\begin{aligned} L = 5, n = 17, s = 0.116, R^2 = 0.982, s_{\text{cv}} = 0.500, \\ R_{\text{cv}}^2 = 0.548, F = 122.1, p = 0.0001 \\ (66\%, 8\%, 14\%, 8\%, 2\%) \end{aligned}$$

The five-component model explains 98% of the variance in  $\log P$  values with a cross-validated  $R^2$  of 0.548. The correlation is good, and the  $R_{\text{cv}}^2$  and  $s_{\text{cv}}$  are reasonable. The first component accounts for 66% of the variance, and the second through the fifth component explain an additional 8%, 14%, 8% and 2%, respectively. Unlike Eq. 22, no separate parameter is needed in Eq. 23 in order to account for the hydrogen-bonding effects of the substituents. The excellent correlation without the hydro-

TABLE 4  
REPORTED AND PREDICTED  $\log k'$  VALUES OF TRIAZINES USING EQ. 17

No.		R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Log k'		
					Rep. <sup>a</sup>	Calc. <sup>b</sup>	Dev.
1R	ENANTIO	Cl	CH(CN)C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.78	0.975	-0.195
59	MEDICLAZ	Cl	C(Me)(CN)C <sub>6</sub> H <sub>4</sub> -4-Cl	Cl	0.82	0.865	-0.045
60	SYNTEX <sup>c</sup>				0.38	0.440	-0.060
61	CP17018	Cl	Cl	H	0.31	0.193	0.117
62	CP18774	CF <sub>3</sub>	Cl	H	0.36	0.220	0.140
63	CP23560	I	H	I	0.53	0.245	0.285
64	CP23240	Cl	SO <sub>2</sub> N(CH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> CH <sub>2</sub>	H	0.07	-0.119	0.189
65	CP23971	Cl	SO <sub>2</sub> N(CH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> CHOH	H	-0.49	-0.409	-0.081
66	CP24343	Cl	SO <sub>2</sub> N(CH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> S	Me	0.05	0.291	-0.241
67	CP24427	Cl	SO <sub>2</sub> NMe <sub>2</sub>	Me	-0.22	0.107	-0.327
68	CP23414	Cl	SO <sub>2</sub> Me	Cl	-0.43	-0.010	-0.420
69	CP25597	Cl	SO <sub>2</sub> C <sub>6</sub> H <sub>4</sub> -4-Cl	H	0.24	0.390	-0.150
70	CP27739	Et	SC <sub>6</sub> H <sub>4</sub> -4-Cl	Me	1.19	1.170	0.020
71	CP28743	Cl	SC <sub>6</sub> H <sub>4</sub> -4- <i>t</i> -Bu	H	1.12	1.015	0.105

<sup>a</sup> From Ref. 18.

<sup>b</sup> Calculated using Eq. 2.

<sup>c</sup>

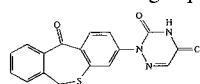




TABLE 5  
OBSERVED AND CALCULATED LOG P VALUES OF SUBSTITUTED FURANS USING EQ. 23

No.	Substituent	Log P			PLS latent variables				
		Obs.	Calc.	Dev.	Z1 <sub>H<sub>2</sub>O</sub>	Z2 <sub>H<sub>2</sub>O</sub>	Z3 <sub>H<sub>2</sub>O</sub>	Z4 <sub>H<sub>2</sub>O</sub>	Z5 <sub>H<sub>2</sub>O</sub>
1	2-CO <sub>2</sub> Et	1.520	1.491	0.029	-0.469	-5.036	7.724	2.493	1.191
2	2-CO <sub>2</sub> Me	1.010	0.853	0.157	-1.367	-7.242	2.714	0.755	2.315
3	2-COMe	0.520	0.589	-0.069	-1.400	-8.222	-0.505	3.034	-2.462
4	2-CONHEt	0.610	0.749	-0.139	-0.572	-6.060	5.245	-2.855	-0.188
5	2-CONH <sub>2</sub>	-0.110	-0.227	0.117	-1.512	-9.769	-8.095	0.447	-3.200
6	2-CONMe <sub>2</sub>	0.410	0.270	0.140	-0.566	-7.746	1.799	-5.049	-0.589
7	2-CONHMe	0.230	0.374	-0.144	-2.053	-8.430	1.231	-1.408	1.418
8	2-Et	2.400	2.429	-0.029	10.369	2.051	-1.692	-2.936	0.700
9	2-Me	1.850	1.753	0.097	4.387	-0.596	-1.987	1.554	1.712
10	2-OMe	1.440	1.502	-0.062	2.421	-3.085	1.212	3.289	-2.722
11	3-CO <sub>2</sub> Et	1.780	1.674	0.106	0.045	13.171	3.783	1.801	0.062
12	3-CO <sub>2</sub> Me	1.280	1.260	0.020	-0.958	10.705	1.294	1.130	0.981
13	3-CONHEt	0.720	0.723	-0.003	-1.111	11.490	0.930	-1.710	-4.477
14	3-CONH <sub>2</sub>	0.090	0.129	-0.039	-2.799	0.975	-8.124	0.254	3.652
15	3-CONMe <sub>2</sub>	0.240	0.279	-0.039	-4.305	7.253	-1.209	-1.263	2.243
16	3-CONHMe	0.340	0.365	-0.025	-2.601	8.790	-1.801	-1.694	-1.405
17	H	1.340	1.459	-0.119	2.491	1.750	-2.518	2.156	0.769

gen-bonding indicator variable shows that the H<sub>2</sub>O probe with the GRID hydrogen-bond potential function effectively describes the hydrogen-bonding properties of various furan substituents.

Although Eq. 22 indicates that no significant additional electronic influences on the log P of furan derivatives are present, it was decided to examine any possible electronic influences on the partitioning behavior in the CoMFA model using the H<sup>+</sup> probe. The result from the combination of hydrophobic and electronic effects provided in Eq. 24 confirms that electronic properties do not play a significant part in the log P values of furan analogs included in this study. The negligible improvement in Eq. 24 over Eq. 23 indicates that the electronic influences of the substituents that have not been described by the H<sub>2</sub>O probe and the hydrogen-bonding potential are extremely small.

$$\begin{aligned} \log P (\text{furan}) = & 0.108(\pm 0.005) Z1_{\text{H}_2\text{O},\text{H}^+} + 0.034(\pm 0.004) Z2_{\text{H}_2\text{O},\text{H}^+} \\ & + 0.044(\pm 0.004) Z3_{\text{H}_2\text{O},\text{H}^+} + 0.069(\pm 0.010) Z4_{\text{H}_2\text{O},\text{H}^+} \\ & + 0.047(\pm 0.011) Z5_{\text{H}_2\text{O},\text{H}^+} + 0.922(\pm 0.025) \end{aligned} \quad (24)$$

$$\begin{aligned} L = 5, n = 17, s = 0.102, R^2 = 0.986, s_{\text{cv}} = 0.392, \\ R_{\text{cv}}^2 = 0.722, F = 158.0, p = 0.0001 \\ (70\%, 9\%, 12\%, 5\%, 3\%) \end{aligned}$$

Table 5 lists the calculated log P values using Eq. 23, along with the latent variables used and the corresponding experimentally determined log P values.

In order to examine whether the steric and electrostatic fields can provide correlations equivalent to those obtained with the H<sub>2</sub>O probe in the present situation, the CH<sub>3</sub> and H<sup>+</sup> probes were again used in the analysis. The results are presented in Eqs. 25 and 26. As in the previous

examples, the CH<sub>3</sub> and H<sup>+</sup> probes failed to yield correlations comparable to those found using the H<sub>2</sub>O probe.

$$\begin{aligned} \log P (\text{furan}) = & 0.148(\pm 0.032) Z1_{\text{CH}_3} + 0.034(\pm 0.015) Z2_{\text{CH}_3} \\ & + 0.922(\pm 0.110) \end{aligned} \quad (25)$$

$$\begin{aligned} L = 2, n = 17, s = 0.453, R^2 = 0.654, s_{\text{cv}} = 0.671, \\ R_{\text{cv}}^2 = 0.185, F = 13.3, p = 0.0006 (52\%, 13\%) \end{aligned}$$

$$\begin{aligned} \log P (\text{furan}) = & 0.103(\pm 0.012) Z1_{\text{CH}_3,\text{H}^+} + 0.035(\pm 0.011) Z2_{\text{CH}_3,\text{H}^+} \\ & + 0.047(\pm 0.013) Z3_{\text{CH}_3,\text{H}^+} + 0.922(\pm 0.068) \end{aligned} \quad (26)$$

$$\begin{aligned} L = 3, n = 17, s = 0.281, R^2 = 0.877, s_{\text{cv}} = 0.470, \\ R_{\text{cv}}^2 = 0.601, F = 30.8, p = 0.0001 (65\%, 12\%, 11\%) \end{aligned}$$

Similar results were also obtained from the log P (Eqs. 27–29) values of benzyl *N,N*-dimethylcarbamates (**III**), substituted pyridines (**IV**), and pyrazines (**V**). Further details will be described elsewhere. In each case, the four-component CoMFA model accounted for 94–97% of the variance in log P values. These results, along with those for the benzene derivatives (Eq. 9) and furan derivatives (Eq. 23) described above, demonstrate the applicability of CoMFA in calculating log P values, both in the carbocyclic systems as well as in various heterocyclic systems.

$$\begin{aligned} \log P (\text{carbamate}) = & 0.145(\pm 0.013) Z1_{\text{H}_2\text{O}} + 0.062(\pm 0.010) Z2_{\text{H}_2\text{O}} \\ & + 0.117(\pm 0.018) Z3_{\text{H}_2\text{O}} + 0.051(\pm 0.019) Z4_{\text{H}_2\text{O}} \\ & + 2.321(\pm 0.046) \end{aligned} \quad (27)$$

$$\begin{aligned} L = 4, n = 19, s = 0.201, R^2 = 0.941, s_{\text{cv}} = 0.590, \\ R_{\text{cv}}^2 = 0.384, F = 55.8, p = 0.0001 \\ (56\%, 18\%, 17\%, 3\%) \end{aligned}$$

$$\begin{aligned} \log P (\text{pyridine}) = & 0.141(\pm 0.008) Z1_{\text{H}_2\text{O}} + 0.067(\pm 0.009) Z2_{\text{H}_2\text{O}} \\ & + 0.035(\pm 0.007) Z3_{\text{H}_2\text{O}} + 0.044(\pm 0.012) Z4_{\text{H}_2\text{O}} \\ & + 1.009(\pm 0.031) \end{aligned} \quad (28)$$

$$\begin{aligned} L = 4, n = 20, s = 0.137, R^2 = 0.966, \\ s_{\text{cv}} = 0.420, R_{\text{cv}}^2 = 0.615, F = 106.8, \\ p = 0.0001 (75\%, 13\%, 6\%, 3\%) \end{aligned}$$

$$\begin{aligned} \log P (\text{pyrazine}) = & 0.121(\pm 0.010) Z1_{\text{H}_2\text{O}} + 0.068(\pm 0.010) Z2_{\text{H}_2\text{O}} \\ & + 0.045(\pm 0.012) Z3_{\text{H}_2\text{O}} + 0.038(\pm 0.016) Z4_{\text{H}_2\text{O}} \\ & - 0.396(\pm 0.042) \end{aligned} \quad (29)$$

$$\begin{aligned} L = 4, n = 20, s = 0.186, R^2 = 0.935, \\ s_{\text{cv}} = 0.385, R_{\text{cv}}^2 = 0.668, F = 54.4, \\ p = 0.0001 (65\%, 20\%, 6\%, 3\%) \end{aligned}$$

The results presented here are in line with our experience with the hydrogen-bonding potential function in describing hydrophobic interactions in various biological QSARs [18–24]. These results render further support to the use of the H<sub>2</sub>O probe in describing hydrophobic interactions in 3D QSAR. Even though quite often claims are made stating that capacity factors produce better correlations than log P in QSAR studies, a method has yet to be discovered that allows the chemist to estimate log k' in lieu of synthesis. The availability of various calculation methods of log P values is one of the prime advantages of using log P values in QSAR studies. The present study opens up the possibility of calculating log k', at least in a series where log k' values of some compounds have been measured.

Considering the excellent correlations obtained for log k' or log P values of various systems, one may ask why the CoMFA descriptors derived from the interaction energies using the H<sub>2</sub>O probe and the hydrogen-bonding potentials give the good correlations described above. Various investigators have demonstrated the existence of close correlations between the log P values of organic compounds and their molecular surface area or volume, describing their 'steric' or bulk properties plus the hydrogen-bond donating and accepting properties [43–51]. The H<sub>2</sub>O probe in GRID is an electrically neutral group that has a van der Waals radius of 1.70 Å. It can both donate and accept up to two hydrogen bonds. Therefore, the energy calculated using the H<sub>2</sub>O probe and hydrogen-bonding potential in GRID results from the steric and hydrogen-bonding interactions. Thus, it is the steric and the hydrogen-bond donating and accepting properties of the H<sub>2</sub>O probe that may be responsible for the superb role of the probe mimicking log P. Since there is a close relationship between log P and log k', as indicated in Eqs. 2–4, it would not be surprising to see similar results for log k'.

Numerous methods are available for calculating hydrophobic parameters [4,8,12,14,52–68], but many of these do

not utilize three-dimensional structures. However, some methods do employ three-dimensional structures in the calculation [69–71]. As might be expected, when three-dimensional structures are utilized in the calculations, the computational and human efforts to perform the calculations are usually greater than in the conventional methods of calculation.

## Conclusions

The method described in this study calculates log k' and log P values from the three-dimensional structures of the molecules. The results obtained show much promise in calculating hydrophobic parameters directly from three-dimensional structures. Since an estimation of hydrophobicity in heterocyclic systems is often more complicated, the results are especially encouraging. They indicate that it is not necessary to include log k' along with the CoMFA descriptors from the H<sub>2</sub>O probe in the correlation to describe hydrophobic effects.

## Acknowledgements

The author would like to express sincere gratitude to Dr. James W. McFarland for providing the coordinates of the 68 triazines used in this study.

## References

- 1 Leo, A., Hansch, C. and Elkins, D., *Chem. Rev.*, 71 (1971) 525.
- 2 Deardon, J.C. and Bresnen, G.M., *Quant. Struct.–Act. Relatsh.*, 7 (1988) 133.
- 3 De Bruijn, J., Busser, F., Seinen, W. and Hermens, J., *Environ. Toxicol. Chem.*, 8 (1989) 499.
- 4 Leo, A.J., *Chem. Rev.*, 93 (1993) 1281.
- 5 Terada, H., *Quant. Struct.–Act. Relatsh.*, 5 (1986) 81.
- 6 Miyake, K., Mizuno, N. and Terada, H., *J. Chromatogr.*, 439 (1988) 227.
- 7 Yamagami, C., Takami, H., Yamamoto, K., Miyoshi, K. and Takano, N., *Chem. Pharm. Bull.*, 32 (1984) 4994.
- 8 Minick, D.J., Frenz, J.H., Patrick, M.A. and Brent, D.A., *J. Med. Chem.*, 31 (1988) 1923.
- 9 Yamagami, C., Ogura, T. and Takao, N., *Chem. Pharm. Bull.*, 39 (1991) 1217.
- 10 Thus, J.L.G. and Kraak, J.C., *J. Chromatogr.*, 320 (1985) 271.
- 11 Tayar, N.E., Tsantili-Kakoulidou, A., Roethlisberger, T., Testa, B. and Gal, J., *J. Chromatogr.*, 439 (1988) 237.
- 12 Taylor, P.J., In Ramsden, C.A. (Ed.) *Quantitative Drug Design, Comprehensive Medicinal Chemistry*, Vol. 4, Pergamon, Oxford, 1990, p. 241.
- 13 Hammers, W.E., Meurs, G.J. and De Ligny, C.L., *J. Chromatogr.*, 247 (1982) 1.
- 14 Braumann, T., *J. Chromatogr.*, 373 (1986) 191.
- 15 Kim, K.H., *Med. Chem. Res.*, 1 (1991) 259.
- 16 Cramer III, R.D., Patterson, D.E. and Bunce, J.D., *J. Am. Chem. Soc.*, 110 (1988) 5959.
- 17 GRID, v. 7, Molecular Discovery Ltd., Oxford, 1990.
- 18 Kim, K.H., *Med. Chem. Res.*, 2 (1992) 22.

- 19 Kim, K.H., *Quant. Struct.-Act. Relatsh.*, 11 (1992) 309.
- 20 Kim, K.H., Greco, G., Novellino, E., Silipo, C. and Vittoria, A., *J. Comput.-Aided Mol. Design*, 7 (1993) 263.
- 21 Kim, K.H., In *Wermuth, C.G. (Ed.) Trends in QSAR and Molecular Modelling 92 (Proceedings of the 9th European Symposium on Structure-Activity Relationships: QSAR and Molecular Modelling)*, ESCOM, Leiden, 1993, pp. 245-251.
- 22 Kim, K.H., *J. Comput.-Aided Mol. Design*, 7 (1993) 71.
- 23 Kim, K.H., *Quant. Struct.-Act. Relatsh.*, 12 (1993) 232.
- 24 Debnath, A.K., Hansch, C., Kim, K.H. and Martin, Y.C., *J. Med. Chem.*, 36 (1993) 1007.
- 25 Kim, K.H., In *Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications*, ESCOM, Leiden, 1993, pp. 619-642.
- 26 Lewis, S.J., Mirrlees, M.S. and Taylor, P.J., *Quant. Struct.-Act. Relatsh.*, 2 (1983) 1.
- 27 Lewis, S.J., Mirrlees, M.S. and Taylor, P.J., *Quant. Struct.-Act. Relatsh.*, 2 (1983) 100.
- 28 Bradshaw, J. and Taylor, P.J., *Quant. Struct.-Act. Relatsh.*, 8 (1989) 279.
- 29 Fujita, T., *Prog. Phys. Org. Chem.*, 14 (1983) 75.
- 30 Yamagami, C., Takao, N. and Fujita, T., *Quant. Struct.-Act. Relatsh.*, 9 (1990) 313.
- 31 Wold, S., *Technometrics*, 20 (1978) 397.
- 32 McFarland, J.W., Cooper, C.B. and Newcomb, D.M., *J. Med. Chem.*, 34 (1991) 1908.
- 33 Yamagami, C. and Takao, N., *Chem. Pharm. Bull.*, 40 (1992) 925.
- 34 SYBYL, v. 5.4, Tripos Associates, St. Louis, MO, 1991.
- 35 McFarland, J.W., *J. Med. Chem.*, 35 (1992) 2543.
- 36 Dewar, M.J.S., Zoebish, E.G., Healy, E.F. and Stewart, J.J.P., *J. Am. Chem. Soc.*, 107 (1985) 3902.
- 37 Kim, K.H. and Martin, Y.C., *J. Med. Chem.*, 34 (1991) 2056.
- 38 Kim, K.H. and Martin, Y.C., *J. Org. Chem.*, 56 (1991) 2723.
- 39 Kim, K.H., *Quant. Struct.-Act. Relatsh.*, 11 (1992) 127.
- 40 The  $H^+$  probe alone explains 37% of the variance in the data:  

$$\log k' = 0.076(\pm 0.019) Z_{1H^+} + 0.052(\pm 0.014) Z_{2H^+} + 0.464(\pm 0.051)$$

$$L = 2, n = 54, R^2 = 0.374, RMSE = 0.374,$$

$$R_{cv}^2 = 0.222, RMSE_{cv} = 0.413, F = 15.2,$$

$$p = 0.0001 (21\%, 16\%)$$
- 41 CLOGP was calculated by the MEDCHEM program (Daylight Chemical Information System, Irvine, CA).
- 42 Yamagami, C., Ogura, T. and Takao, N., *J. Chromatogr.*, 514 (1990) 123.
- 43 Pearlman, R.S., In *Dunn, W.J.I., Block, J.H. and Pearlman, R.S. (Eds.) Partition Coefficient*, Pergamon Press, New York, NY, 1986, p. 3.
- 44 Leo, A. and Hansch, C., *J. Med. Chem.*, 19 (1976) 611.
- 45 Camilleri, P., Watts, S.A. and Boraston, J.A., *J. Chem. Soc., Perkin Trans. II*, (1988) 1699.
- 46 Dunn, W.J.I., Koehler, M.G. and Grigoras, S., *J. Med. Chem.*, 30 (1987) 1121.
- 47 Leahy, D., *J. Pharm. Sci.*, 75 (1986) 629.
- 48 Tunon, I., Silla, E. and Pascual-Ahuir, J.L., *Protein Eng.*, 5 (1992) 715.
- 49 Kamlet, M.J., Doherty, R.M., Abraham, M.H., Marcus, Y. and Taft, R.W., *J. Phys. Chem.*, 92 (1988) 5244.
- 50 Lien, E.J., Banerjee, S., Khan, A., Gao, H. and Wang, F., In *Wermuth, C.G. (Ed.) Trends in QSAR and Molecular Modelling 92 (Proceedings of the 9th European Symposium on Structure-Activity Relationships: QSAR and Molecular Modelling)*, ESCOM, Leiden, 1993, pp. 303-304.
- 51 El Tayar, N. and Testa, B., In *Wermuth, C.G. (Ed.) Trends in QSAR and Molecular Modelling 92 (Proceedings of the 9th European Symposium on Structure-Activity Relationships: QSAR and Molecular Modelling)*, ESCOM, Leiden, 1993, pp. 101-108.
- 52 Fujita, T., Iwasa, J. and Hansch, C., *J. Am. Chem. Soc.*, 86 (1964) 5175.
- 53 Rekker, R. and Mannhold, R., *Calculation of Drug Lipophilicity*, VCH, Weinheim, 1992.
- 54 Bronto, P., Moreau, G. and Vandycke, C., *Eur. J. Med. Chem.*, 19 (1971) 71.
- 55 Ghose, A. and Crippen, G., *J. Comput. Chem.*, 7 (1986) 565.
- 56 Iwase, K., Komatsu, K., Hirono, S., Nakagawa, S. and Moriguchi, I., *Chem. Pharm. Bull.*, 33 (1985) 2114.
- 57 Viswanadhan, V., Ghose, A., Revankar, G. and Robins, R., *J. Chem. Inf. Comput. Sci.*, 29 (1989) 163.
- 58 Moriguchi, I., Hirono, S., Liu, Q., Nakagome, I. and Matsushita, Y., *Chem. Pharm. Bull.*, 33 (1992) 127.
- 59 Klopman, G. and Wang, S., *J. Comput. Chem.*, 12 (1991) 1025.
- 60 Bodor, N., Gabanyi, Z. and Wong, C.-K., *J. Am. Chem. Soc.*, 111 (1989) 3783.
- 61 Sasaki, Y., Kubodera, H., Matuszaki, T. and Umeyama, H., *J. Pharmacobio-Dyn.*, 14 (1991) 207.
- 62 Kamlet, M., Abboud, J.-L. and Taft, R., *J. Am. Chem. Soc.*, 99 (1977) 6027.
- 63 Tanaka, A., Nakamura, K., Nakanishi, I. and Fujiwara, H., *J. Med. Chem.*, 37 (1994) 4653.
- 64 Abraham, D.J. and Kellogg, G.E., In *Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications*, ESCOM, Leiden, 1993, pp. 506-522.
- 65 Akamatsu, M. and Fujita, T., *J. Pharm. Sci.*, 81 (1992) 164.
- 66 Dorsey, J.G. and Khaledi, M.G., *J. Chromatogr.*, A65 (1993) 485.
- 67 Gaillard, P., Carrupt, P.-A. and Testa, B., *J. Mol. Graphics*, 12 (1994) 73.
- 68 Dubost, J.P., In *Wermuth, C.G. (Ed.) Trends in QSAR and Molecular Modelling 92 (Proceedings of the 9th European Symposium on Structure-Activity Relationships: QSAR and Molecular Modelling)*, ESCOM, Leiden, 1993, pp. 93-100.
- 69 Kellogg, G.E., Semus, S.F. and Abraham, D.J., *J. Comput.-Aided Mol. Design*, 5 (1991) 545.
- 70 Koehler, M., Grigoras, S. and Dunn III, W., *Quant. Struct.-Act. Relatsh.*, 7 (1988) 150.
- 71 Hopfinger, A. and Battershell, R., *J. Med. Chem.*, 19 (1976) 569.