Journal of Chromatography, 389 (1987) 33–45
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 19 196

STRUCTURAL EFFECTS IN THE LIPOPHILICITY OF DI- AND POLYSUBSTITUTED BENZENES AS MEASURED BY REVERSED-PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

ANNA TSANTILI-KAKOULIDOU
Department of Pharmacy, University of Athens, GR-106 80 Athens (Greece)
and
NABIL EL TAYAR, HAN VAN DE WATERBEEMD and BERNARD TESTA*
School of Pharmacy, University of Lausanne, Place du Château 3, CH-1005 Lausanne (Switzerland)
(Received October 30th, 1986)

SUMMARY

A series of disubstituted benzene derivatives were examined for their lipophilicity as measured by reversed-phase high-performance liquid chromatography (RP-HPLC) and the results were combined with those of a previous study to form a set of 75 compounds. The interactions between substituents were quantified by the difference (T^*) between experimental and calculated (by summing separately determined substituent constants) log k_w values. For *meta*- and *para*-substituted derivatives, the T^* term could be correlated with electronic parameters ($\rho\sigma$ model), while for *ortho* derivatives the *ortho*-effect was expressed by a single indicator I_0 accounting for both steric effects and hydrogen bonds. A single equation (eqn. 16) was derived which described the log k_w values of all 75 compounds with a correlation coefficient of 0.988. This equation was tested with a small series of polysubstituted benzenes and proved successful in predicting their log k_w values.

INTRODUCTION

Hydrophobicity is one of the most important parameters in the relationships between biological activity and chemical structure of drugs and other xenobiotics. A number of indices are frequently used in the literature to express the lipophilic properties of drugs, *e.g.* partition coefficients measured in the *n*-octanol–water system[1] (log P), and extrapolated capacity factors[2] (log k_w) determined using reversed-phase high-performance liquid chromatography (RP-HPLC). Fujita and Hansch[3] have found that partition coefficients show an additive-constitutive character. They introduced the hydrophobic substituent constant π to express the contribution of each substituent to the overall partition coefficient of a compound. For a substituent X, π_x is calculated with the well-known equation

$$\pi_x = \log P_{RX} - \log P_{RH} \quad (1)$$

0021-9673/87/$03.50

where RH is the non-substituted parent compound. However, the π value of a given substituent is often found to vary from one solute structure to another when certain steric and/or electronic interactions of substituents are involved. Different approaches have been used to explain and express such variations in terms of electronic, steric and intramolecular hydrogen bonding effects. In recent years, many studies[4–12] have been performed in order to explain and predict the effects of substituent interactions on partitioning processes of drugs. In a previous paper[12], we have investigated the degree of additivity in log k_w in analogy with log P values. The RP-HPLC technique was used as an alternative means of measuring lipophilicity since it gives more accurate and precise data than the traditional shake-flask method. A series of 43 disubstituted benzenes were examined, and the substituent interactions were expressed in terms of Hammett[13] σ constant for *meta* and *para* isomers ($n = 29$). The importance of hydration–dehydration phenomena involved in the partitioning process prompted us to express the substituents interactions in terms of a hydration factor omega[14] (ω), the general equation being:

$$\log k_{exp.} = \log k_{add.} + m_w \cdot \omega \tag{2}$$

where m_w is a multiplicity factor having an integer value, and where the subscripts exp. and add. refer to measured and calculated (by summation) log k values, respectively. The physical nature of the term $m_w \cdot \omega$ is suggested to be associated with the number of water molecules involved in the partitioning process, reflecting redistribution of partial charges in the molecule. Eqn. 2 appears to be valid for *para-*, *meta-* and *ortho*-disubstituted benzenes. Unfortunately no clearcut rules exist at present to predict the number of occurrences (m_w) of ω. In the present work, we have investigated a series of 75 disubstituted benzenes representing 7 different substituents in all *ortho*, *meta* and *para* combinations, except CH_3–OCH_3 and CH_3–COOH. This series differs from the previous one[12] in that it includes two additional substituents, OCH_3 and COOH. In addition, a series of polysubstituted benzenes were also examined and used to challenge the predictive value of a general equation accounting for substituent interactions.

MATERIALS AND METHODS

Materials and chromatography

Mono-, di- and polysubstituted benzene derivatives were purchased from Fluka (Buchs, Switzerland) and Merck (Darmstadt, F.R.G.) or kindly donated by Delagrange (Paris, France). They all were of analytical grade and used without further purification. Methanol (analyzed reagent grade) was from Merck. The equipment and procedures were as described in previous papers[15,16].

Procedures

In order to combine data published in 1985[12] with the results of the present study, eqn. 3 was established using ten compounds (Table I) measured under the same conditions (pH 7.4, *n*-decylamine as a masking agent):

$$\log k_w(1987) = 1.050(\pm 0.022) \log k_w(1985) + 0.022(\pm 0.050)$$
$$n = 10; \quad r = 0.999; \quad s = 0.036; \quad F = 5919 \tag{3a}$$

TABLE I

CALIBRATION AND INFLUENCE OF EXPERIMENTAL CONDITIONS

No.	*Compound*	*log k_w (pH = 7.4)*		*log k_w (pH = 2.5)*
		(1985)	*(1987)*	
1	C_6H_6	1.987	2.120	2.204
2	C_6H_5–OH	1.274	1.443	1.374
3	–NH_2	0.977	1.037	
4	–NO_2	1.814	1.917	2.064
5	–CH_3	2.603	2.722	2.825
6	–Cl	2.702	2.849	2.959
7	–OCH_3		2.094	2.254
8	–$CONH_2$		0.938	1.273
9	–$COOCH_3$		2.194	2.405
10	–$COOC_2H_5$		2.687	
11	–SO_2CH_3		0.943	1.615
12	–SO_2NH_2		0.829	1.107
13	HO–C_6H_4–3-OH	0.868	0.948	
14	HO–C_6H_4–4-OH	0.395	0.405	
15	NH_2–C_6H_4–2-NH_2	0.632	0.685	
16	NH_2–C_6H_4–3-NH_2	0.273	0.282	

The constant term being non-significant eqn. 3a was forced through the origin (eqn. 3b):

$$\log k_w(1987) = 1.061(\pm 0.021) \log k_w(1985) \quad (3b)$$

All equations in this paper report 95% confidence intervals. The excellent statistical quality of eqn. 3 indeed permits the incorporation of previously studied compounds into the present study.

Benzoic acid and its derivatives could not be measured under the same conditions since the carboxylate anion forms ion pairs with the masking agent *n*-decylamine at pH = 7.4. Therefore, benzoic acid and its derivatives were eluted at pH 2.5 in the absence of *n*-decylamine. The aqueous solution consisted of 0.01 *N* hydrochloric acid, the pH being adjusted with sodium hydroxide. At pH 2.5, retention data for acidic compounds (Table II) must be corrected for solute ionization using eqn. 4:

$$\log k_w(\text{neutral}) = \log k_w(\text{pH } 2.5) + \log (1 + 10^{\text{pH} - \text{p}K_a}) \quad (4)$$

where pK_a is the ionization constant of the compound. Using a number of non-ionizable monosubstituted benzene derivatives under both sets of conditions *i.e.*, (a) pH 7.4 in the presence of *n*-decylamine, (b) pH 2.5 in the absence of *n*-decylamine, it was established that the retention times experience very little variation under these two sets of conditions, as also indicated by the following equation obtained after forcing through the origin:

$$\log k_w(a) = 0.948(\pm 0.054) \log k_w(b)$$
$$n = 7; \quad r = 0.984; \quad s = 0.129; \quad F = 155.6 \quad (5)$$

TABLE II

CAPACITY FACTORS OF BENZOIC ACID DERIVATIVES

HOOC–C_6H_4–R R	*log k_w (pH 2.5)*	*pK_a**	*log k_w^0***	*log k_w^0 (calc)****
2-COOH	1.348 ± 0.068	2.9	1.482	1.405
3-COOH	2.044 ± 0.065	3.7	2.077	1.969
4-COOH	1.880 ± 0.057	3.5	1.918	1.818
2-OCH_3	1.678 ± 0.101	4.09	1.689	1.601
3-OCH_3	2.254 ± 0.080	4.1	2.265	2.147
4-OCH_3	2.216 ± 0.091	4.49	2.220	2.104
2-OH	2.033 ± 0.049	2.98	2.157	2.044
3-OH	1.552 ± 0.059	4.08	1.563	1.481
4-OH	1.368 ± 0.052	4.57	1.371	1.299
2-Cl	1.947 ± 0.080	2.90	2.095	1.986
3-Cl	2.717 ± 0.073	3.84	2.736	2.593
4-Cl	2.726 ± 0.071	4.00	2.740	2.597
2-NO_2	1.144 + 0.068	2.17	1.634	1.549
3-NO_2	1.955 ± 0.058	3.46	2.000	1.896
4-NO_2	2.012 ± 0.056	3.43	2.061	1.953

* pK_a values taken from ref. 21.
** Log k_w (neutral) corrected for solute ionization according to eqn. 4.
*** Log k_w^0 (calc) corresponds to the data transformed from condition b to condition a, using eqn. 5.

TABLE III

SUBSTITUENT CONSTANTS OF MONOSUBSTITUTED BENZENES

Substituent	π**	σ_m**	σ_p**	ρ***
H	0.000	0.00	0.00	0.00
OH	−0.677	0.04	−0.13	1.01
NH_2	−1.082	−0.14	−0.38	0.93
NO_2	−0.203	0.70	0.82	0.00
CH_3	0.602	−0.07	−0.12	0.00
Cl	0.730	0.37	0.27	0.00
OCH_3	−0.025	0.06	−0.16	0.27
COOH	−0.241	0.37	0.46	0.37
$CONH_2$	−1.182	0.28	0.36	
$COOCH_3$	0.074	0.37	0.45	
$COOC_2H_5$	0.567	0.37	0.45	
$CH(CH_3)_2$	1.638	−0.07	−0.15	
C_6H_5	1.992	0.06	−0.01	
SO_2NH_2	−1.291	0.46	0.57	
SO_2CH_3	−1.177	0.60	0.72	
$SO_2C_2H_5$	−0.869	0.60	0.72	
SO_2NHCH_3	−0.802	0.46	0.57	

* The π* values are derived from log k_w data (RP-HPLC conditions).
** The electronic substituent constants are taken from ref. 17.
*** Fujita's ρ values are taken from ref. 5.

This equation was used to convert the log k_w values of benzoic acid derivatives to the conditions (presence of *n*-decylamine) used for non-acidic compounds.

Data analysis

A number of monosubstituted benzene derivatives were used as standards in order to calculate the π^* value of each substituent (see Table III) in analogy with eqn. 1, the * symbol meaning chromatographic (log k_w) data. The additive values (log $k_{add.}$) of the disubstituted benzene derivatives of general structure X–C_6H_4–Y were calculated by adding the hydrophobic substituent values π^* of X and Y to the log k_w value of benzene. The differences between calculated and experimental log k values are equated with a substituent interaction term T^* in analogy with previous studies based on log P values[4–9].

$$\log k_{exp.} = \log k_{add.} + T^* \quad (6)$$

In order to predict lipophilicity (log P or log k_w values), the interaction term T^* should be expressed in terms of physicochemical parameters. Brändström[4] and then Fujita[5–7] have proposed a useful approach by expressing substituent interactions (as derived from log P data and symbolized by T_I) in terms of electronic parameters (σ_m and σ_p), assuming the interaction between the two aromatic substituents X and Y to be bidirectional:

$$T_I = \rho_Y\sigma_X + \rho_X\sigma_Y \quad (7)$$

where ρ_X and ρ_Y are susceptibility constants to the respective solubility-modifying effects of Y and X. In analogy with eqn. 6, it follows that:

$$\log P_{exp} = \log P_{add} + \rho_X\sigma_Y + \rho_Y\sigma_X \quad (8)$$

Leo[8] has presented an alternative approach which also handles *ortho*-isomers by taking into account the negative *ortho*-effect (symbolized by F_O) and intramolecular hydrogen bonding (F_{HB}). At present, three sets of values for the ρ constants have been calculated by Fujita[5–7], Brändström[4] and Leo[8] (for a comparison of the values, see El Tayar *et al.*[12]). In the present study, we have used Fujita's ρ values together with σ_m and σ_p values taken from Exner[17] (Table III). We consider that the average σ values proposed by Leo are not fully appropriate, the differences between log k values of *meta*- and *para*-isomers often being quite significant, particularly when substituents with distinct electronic effects are present. For the *ortho*-isomers, the corresponding σ_p values were used.

RESULTS AND DISCUSSION

Relationships between HPLC capacity factors and n-octanol–water partition coefficients

In a previous paper[12], we have shown that log P values measured in *n*-octanol–water are well correlated with extrapolated capacity factors (log k_w values), better in fact than with isocratic capacity factors (log k_x). Using the log k_w obtained in the present study (Table IV), we confirm these previous results as shown in eqn. 9.

$$\log k_w = 0.964(\pm 0.065) \log P + 0.180(\pm 0.134)$$
$$n = 72; \quad r = 0.970; \quad s = 0.206; \quad F = 878 \qquad (9)$$

However, three compounds with strongly interacting *ortho* groups (*o*-nitrophenol, *o*-dinitrobenzene and *o*-methoxyphenol) are outliers and are not included in eqn. 9. This observation, which requires further investigations, suggests different solute-solvent interactions in the RP-HPLC and the *n*-octanol–water systems when intramolecular H-bonding plays a role.

TABLE IV

DATA OF DISUBSTITUTED BENZENES

Compound	*log* $k_{w,exp.}$	*log* $k_{w,add.}$*	*T**	*HS***	*HB***	I_0**
CH_3–C_6H_4–2-CH_3	3.276	3.324	−0.048	0	0	0
CH_3–C_6H_4–3-CH_3	3.380	3.324	0.056			
CH_3–C_6H_4–4-CH_3	3.392	3.324	0.068			
CH_3–C_6H_4–2-Cl	3.568	3.452	0.116	0	0	0
CH_3–C_6H_4–3-Cl	3.564	3.452	0.112			
CH_3–C_6H_4–4-Cl	3.539	3.452	0.087			
Cl–C_6H_4–2-Cl	3.520	3.580	−0.060	0	0	0
Cl–C_6H_4–3-Cl	3.618	3.580	0.038			
Cl–C_6H_4–4-Cl	3.517	3.580	−0.063			
Cl–C_6H_4–2-NO_2	2.517	2.647	−0.130	0	0	0
Cl–C_5H_4–3-NO_2	2.616	2.647	−0.031			
Cl–C_6H_4–4-NO_2	2.448	2.647	−0.199			
OH–C_6H_4–2-Cl	2.325	2.173	0.152	0	0	0
OH–C_6H_4–3-Cl	2.547	2.173	0.374			
OH–C_6H_4–4-Cl	2.471	2.173	0.298			
NH_2–C_6H_4–2-Cl	1.817	1.768	0.049	0	0	0
NH_2–C_6H_4–3-Cl	1.891	1.768	0.123			
NH_2–C_6H_4–4-Cl	1.860	1.768	0.092			
CH_3–C_6H_4–2-NO_2	2.460	2.519	−0.059	0	0	0
CH_3–C_6H_4–3-NO_2	2.547	2.519	0.028			
CH_3–C_6H_4–4-NO_2	2.488	2.519	−0.031			
OH–C_6H_4–2-CH_3	1.854	2.045	−0.191	0	0	0
OH–C_6H_4–3-CH_3	1.908	2.045	−0.137			
OH–C_6H_4–4-CH_3	1.911	2.045	−0.134			
NH_2–C_6H_4–2-CH_3	1.482	1.640	−0.158	0	0	0
NH_2–C_6H_4–3-CH_3	1.567	1.640	−0.073			
NH_2–C_6H_4–4-CH_3	1.598	1.640	−0.042			
CH_3O–C_6H_4–2-Cl	2.608	2.825	−0.217	−1	0	−1
CH_3O–C_6H_4–3-Cl	3.052	2.825	0.227			
CH_3O–C_6H_4–4-Cl	2.979	2.825	0.154			
COOH–C_6H_4–2-Cl	1.986	2.609	−0.623	−2	0	−2
COOH–C_6H_4–3-Cl	2.593	2.609	−0.016			
COOH–C_6H_4–4-Cl	2.597	2.609	−0.012			
CH_3O–C_6H_4–2-OCH_3	1.741	2.070	−0.329	−1	0	−1
CH_3O–C_6H_4–3-OCH_3	2.347	2.070	0.277			
CH_3O–C_6H_4–4-OCH_3	2.144	2.070	0.074			
CH_3O–C_6H_4–2-OH	1.510	1.418	0.092	0	0	0
CH_3O–C_6H_4–3-OH	1.548	1.418	0.132			
CH_3O–C_6H_4–4-OH	1.308	1.418	−0.110			

TABLE IV (*continued*)

Compound	*log* $k_{w,exp.}$	*log* $k_{w,add.}$*	*T**	*HS***	*HB***	I_0**
CH_3O-C_6H_4-2-NH_2	1.338	1.031	0.325	0	1	1
CH_3O–C_6H_4–3-NH_2	1.144	1.031	0.131			
CH_3O–C_6H_4–4-NH_2	0.933	1.031	−0.076			
CH_3O–C_6H_4–2-NO_2	1.896	1.892	0.004	−1	0	−1
CH_3O–C_6H_4–3-NO_2	2.431	1.892	0.539			
CH_3O–C_6H_4–4-NO_2	2.312	1.892	0.420			
COOH–C_6H_4–2-OCH_3	1.601	1.854	−0.253	−1	0	−1
COOH–C_6H_4–3-OCH_3	2.147	1.854	0.293			
COOH–C_6H_4–4-OCH_3	2.104	1.854	0.250			
COOH–C_6H_4–2-COOH	1.405	1.638	−0.233	−1	0	−1
COOH–C_6H_4–3-COOH	1.969	1.638	0.331			
COOH–C_6H_4–4-COOH	1.818	1.638	0.180			
COOH–C_6H_4–2-NO_2	1.549	1.676	−0.127	−1	0	−1
COOH–C_6H_4–3-NO_2	1.896	1.676	0.220			
COOH–C_6H_4–4-NO_2	1.953	1.676	0.277			
COOH–C_6H_4–2-OH	2.044	1.202	0.842	0	2	2
COOH–C_6H_4–3-OH	1.481	1.202	0.279			
COOH–C_6H_4–4-OH	1.299	1.202	0.097			
OH–C_6H_4–2-OH	1.314	0.766	0.552	0	2	2
OH–C_6H_4–3-OH	0.945	0.766	0.187			
OH–C_6H_4–4-OH	0.407	0.766	−0.359			
OH–C_6H_4–2-NH_2	0.972	0.361	0.611	0	2	2
OH–C_6H_4–3-NH_2	0.515	0.361	0.154			
OH–C_6H_4–4-NH_2	0.131	0.361	−0.230			
NH_2–C_6H_4–2-NH_2	0.685	−0.044	0.729	0	2	2
NH_2–C_6H_4–3-NH_2	0.282	−0.044	0.328			
NH_2–C_6H_4–4-NH_2	−0.118	−0.044	−0.040			
NO_2–C_6H_4–2-NO_2	2.080	1.714	0.366	0	1	1
NO_2–C_6H_4–3-NO_2	1.764	1.714	0.050			
NO_2–C_6H_4–4-NO_2	1.571	1.714	−0.142			
NH_2–C_6H_4–2-NO_2	1.860	0.835	1.025	0	3	3
NH_2–C_6H_4–3-NO_2	1.510	0.835	0.675			
NH_2–C_6H_4–4-NO_2	1.476	0.835	0.641			
OH–C_6H_4–2-NO_2	2.517	1.240	1.277	0	3	3
OH–C_6H_4–3-NO_2	2.022	1.240	0.782			
OH–C_6H_4–4-NO_2	2.118	1.240	0.878			

* The additive log k_w values are calculated according to eqn. 1.

** The dummy variables account for the negative *ortho*-effect (*HS*), the intramolecular H-bond (*HB*) and the overall *ortho*-effect (I_0).

Substituent interactions

The experimental log k_w values of the 75 disubstituted benzenes were compared with the calculated additive values according to eqn. 6 (Table IV). The resulting interaction terms *T** (Table IV) are presented in Fig. 1. While some groups show weak or negligible interactions, others interact strongly. In several but not all cases, there are marked differences between the positional isomers. There are however few exceptions to the rule that lipophilicity increases when going from a *para* to a *meta* to an *ortho* substitution pattern.

To attempt accounting for these interactions, a stepwise approach is followed. In a first step, two groups are considered, namely *meta*- plus *para*-isomers in one

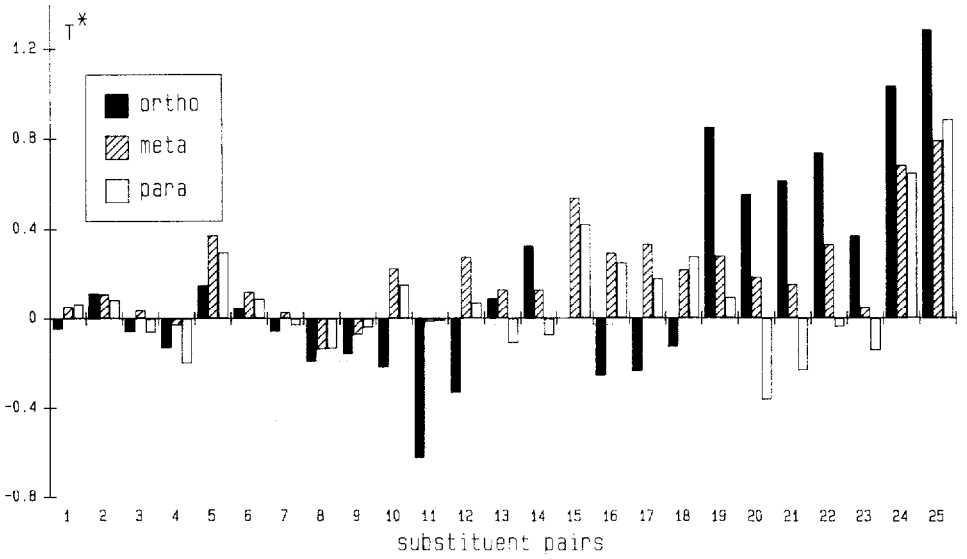

Fig. 1. Substituent interactions (log $k_{exp.}$ − log $k_{add.}$; eqn. 6) for the 25 triplets (*ortho*, *meta* and *para*) of disubstituted benzene derivatives. The numbers in the bottom line refer to the following substituents: (1) CH_3/CH_3; (2) CH_3/Cl; (3) Cl/Cl; (4) Cl/NO_2; (5) Cl/OH; (6) Cl/NH_2; (7) CH_3/NO_2; (8) CH_3/OH; (9) CH_3/NH_2; (10) Cl/OCH_3; (11) Cl/COOH; (12) OCH_3/OCH_3; (13) OCH_3/OH; (14) OCH_3/NH_2; (15) OCH_3/NO_2; (16) $OCH_3/COOH$; (17) COOH/COOH; (18) $COOH/NO_2$; (19) COOH/OH; (20) OH/OH; (21) OH/NH_2; (22) NH_2/NH_2; (23) NO_2/NO_2; (24) NO_2/NH_2; (25) NO_2/OH.

group, and *ortho*-isomers separately. Indeed, treating the *ortho*-isomers separately in a first step is justified by the so-called *ortho*-effect (*i.e.* a combination of electronic and steric interactions)[5]. In a second step, all isomers are merged in a global equation which is then challenged using polysubstituted benzene derivatives.

Analysis of substituent interactions in meta- and para-disubstituted benzenes

Meta- and *para*-isomers were treated together since preliminary studies (not shown) indicate that this approach yields the same results as a separate treatment. Using a converted and simplified form of eqn. 8, the highly significant eqn. 10 was obtained:

$$\log k_{exp.} = 0.939(\pm 0.052) \log k_{add.} + 0.775(\pm 0.178)\ \rho\sigma + 0.192(\pm 0.107)$$
$$n = 50; \quad R = 0.988; \quad s = 0.144; \quad F = 722 \qquad (10)$$

where $\rho\sigma$ stands for $\rho_X\sigma_Y + \rho_Y\sigma_X$. In this equation, the coefficient of $\rho\sigma$ deviates from unity; this was expected since the ρ values used were derived from partition coefficient data (log P). An attempt was made to establish ρ^* values derived from log k_w data (Table V) but they are preliminary and need to be confirmed using a larger data base. Therefore, ρ values based on the large set of available log P values are still preferable at present. To evaluate better the contribution of each variable to the explained variance, eqn. 10 was normalized with standardized variables according to Mager and Barth[18]:

$$\log k_{exp.} = 0.952(\pm 0.048) \log k_{add.} + 0.229(\pm 0.047)\ \rho\sigma \qquad (11)$$

TABLE V

ρ^* VALUES FROM HPLC DATA

Substituent	ρ^* *(HPLC)**	n	R	s
OH	0.93(±0.30)	14	0.979	0.16
NH_2	0.75(±0.26)	12	0.986	0.13
OCH_3	0.46(±0.20)	12	0.992	0.09
COOH	0.06(±0.26)	10	0.974	0.10
CH_3	0.01(±0.11)	10	0.998	0.05
Cl	−0.28(±0.21)	14	0.983	0.12
NO_2	−0.43(±0.31)	14	0.914	0.17

* The ρ^* values are obtained from $\pi^*_{X,Y} = \pi^*_{X,H} + \rho^*_Y \sigma_X$.

It should be noted that if the electronic effect is introduced separately for substituents X and Y, the $\rho_X\sigma_Y$ term contributes very little to the regression as shown when comparing eqns. 12 and 13 among themselves and with eqn. 11:

$$\log k_{exp.} = 0.956(\pm 0.044) \log k_{add.} + 0.207(\pm 0.051)\ \rho_Y\sigma_X + \\ + 0.051(\pm 0.045)\ \rho_X\sigma_Y$$
$$n = 50; \quad R = 0.988; \quad s = 0.145; \quad F = 457 \qquad (12)$$

$$\log k_{exp.} = 0.971(\pm 0.044) \log k_{add.} + 0.224(\pm 0.052)\ \rho_Y\sigma_X$$
$$n = 50; \quad R = 0.987; \quad s = 0.149; \quad F = 649 \qquad (13)$$

These equations suggest that the electronic effect of substituent X on the hydrogen bonding ability of Y is dominating. However, since position X and Y are chosen arbitrarily, utilization of the combined term $\rho\sigma$ is certainly preferable. The introduction into eqn. 10 of other electronic parameters (σ_I, σ_R, F, R) instead of $\rho\sigma$ has been examined but does not improve the regression (equations not shown).

Analysis of substituent interactions in ortho-disubstituted benzenes

The structural analysis of *ortho*-isomers is consistently found to be rather complex due to the combination of various proximity effects globally referred to as the "*ortho*-effect"[5]. The *ortho*-effect may be negative (when primarily due to steric hindrance causing a decoupling of resonance between the substituents and the aromatic nucleus) or positive (when primarily due to the formation of intramolecular hydrogen bonds). Thus, the analysis of *ortho*-effects is difficult to perform using a unique and universally applicable set of parameters.

In the series examined here seven compounds are capable of forming intramolecular hydrogen bonds, showing a higher lipophilicity than their *meta*- and *para*-

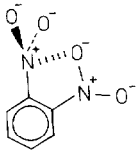

Fig. 2. Postulated intramolecular dipole–dipole interaction in *ortho*-dinitrobenzene.

isomers. In addition, it is puzzling to note that *ortho*-dinitrobenzene behaves in an anomalous way, also showing a higher lipophilicity than its *meta*- and *para*-isomer. We postulate that this increase in lipophilicity is caused by the existence of an intramolecular electrostatic bond between the two nitro groups, as shown in Fig. 2. As a consequence, *ortho*-dinitrobenzene is treated together with the compounds which form internal hydrogen bonds. An indicator *HB* taking the value +1 or +2 is used to distinguish between weaker or stronger hydrogen bonds[19] (Table IV). However, when a nitro group (a strong acceptor of H-bonds) is present in the molecule a higher increase in lipophilicity is observed which deserves a *HB* value of +3.

For the remaining compounds, the introduction of various steric parameters to express the negative *ortho*-effect[5] failed to give meaningful results (equations not shown). Therefore, as done by Leo[8], an indicator (here designated *HS*) was used to express different levels of steric effects; Leo recognized five levels, which we have simplified here to two, *i.e.* *HS* takes the values 0, −1 or −2 (Table IV). Performing the same regression as for the *meta*- and *para*-compounds, and introducing the indicators *HB* and *HS* to account for the H-bond and steric components of the *ortho*-effect, the following equation is obtained:

$$\log k_{\text{exp.}} = 0.946(\pm 0.103) \log k_{\text{add.}} + 0.288(\pm 0.229)\ \rho\sigma + 0.284(\pm 0.129)\ HS + 0.412(\pm 0.134)\ HB + 0.127(\pm 0.258)$$
$$n = 25; \quad R = 0.988; \quad s = 0.129; \quad F = 147.4 \qquad (14)$$

The regressor of the $\rho\sigma$ term shows that electronic interactions play a role, but that their intensity has dropped to 1/3 in comparison with the *meta*- and *para*-isomers (eqn. 10). The σ values used in this equation are the σ_p values from Table III according to Fujita's approach, which assumes that σ_{ortho} may be expressed by a combination of σ_p, σ_I and E_S. At first, all these parameters were used, but σ_I and E_S had negligible effect and were statistically insignificant. The confidence limits of *HS* and *HB* in eqn. 14 suggest the combination of the two indicators in a single one I_0 ($I_0 = HB + HS$) which expresses the global *ortho*-effect and takes integer values from −2 to +3 (Table IV). Using the I_0 indicator the following equation is obtained:

$$\log k_{\text{exp.}} = 0.948(\pm 0.080) \log k_{\text{add.}} + 0.260(\pm 0.187)\ \rho\sigma + 0.303(\pm 0.054)\ I_0 + 0.135(\pm 0.167)$$
$$n = 25; \quad R = 0.990; \quad s = 0.112; \quad F = 246.6 \qquad (15)$$

To evaluate the contribution of each independent variable to the explained variance, eqn. 15 was normalized with standardized variables[18]:

$$\log k_{\text{exp.}} = 1.202(\pm 0.101) \log k_{\text{add.}} + 0.111(\pm 0.079)\ \rho\sigma + 0.561(\pm 0.099)\ I_0 \qquad (16)$$

Combined analysis for ortho-, meta- and para-isomers

By combining the meaningful equations obtained above for the separate *ortho*- and *meta*- plus *para*-isomers, a general equation is derived which accounts for the deviation from additivity measured in the present study:

$$\log k_{exp.} = 0.943(\pm 0.043) \log k_{add.} + 0.790(\pm 0.168)\ \rho\sigma + \\ + 0.255(\pm 0.208)\ \rho\sigma_0 + 0.297(\pm 0.050)\ I_0 + 0.170(\pm 0.080)$$
$$n = 75; \quad R = 0.988; \quad s = 0.136; \quad F = 520.2 \qquad (17)$$

In this equation, $\rho\sigma_0$ stands for $\rho_X\sigma_Y + \rho_Y\sigma_X$ and accounts for the *ortho* electronic effect using σ_p values of *para*-isomers. Interestingly, the regressor of the *ortho*-effect indicator I_0 is very close indeed to Rekker's[20] magic constant $C_M = 0.289$. This is certainly not fortuitous and may be related to hydration factors[14]. Eqn. 17, when normalized with standardized variables, becomes:

$$\log k_{exp.} = 1.016(\pm 0.046) \log k_{add.} + 0.206(\pm 0.044)\ \rho\sigma + \\ + 0.054(\pm 0.044)\ \rho\sigma_0 + 0.275(\pm 0.047)\ I_0 \qquad (18)$$

Eqn. 17 is, *stricto sensu*, valid only for the 75 tested solutes. Its potential interest in calculating the lipophilicity of other disubstituted benzene derivatives must be established. In the following section, its applicability to polysubstituted benzene derivatives is examined.

Lipophilicity prediction of polysubstituted benzenes

In this preliminary study, 11 polysubstituted benzene derivatives have been examined (Table VI). Substituent interactions, considered in all cases to be bidirectional, are expressed by a $\rho\sigma$ term for the *meta*- and *para*-position and by a $\rho\sigma_0$ term for the *ortho*-position. For the steric effect and the intramolecular hydrogen bond formation the same indicator (I_0) was used for each pair of *ortho*-substituents as for the disubstituted benzenes. One compound, pyrrogalol, contains three –OH groups in *ortho* position to each other. In this case only one –OH, –OH pair was considered for the intramolecular hydrogen bond formation. To illustrate the procedure, an Appendix details the calculation of one $\log k_{pred.}$ value according to eqn. 17. In Table VI, experimental values ($\log k_{exp.}$), additive values ($\log k_{add.}$) and predicted values ($\log k_{pred.}$) using eqn. 17 are presented. As expected, the correlation between additive and experimental values is poor ($r = 0.802$). In contrast, predicted and experimental

TABLE VI

LIPOPHILICITY DATA OF POLYSUBSTITUTED BENZENES

Compound	*log* $k_{w,exp.}$	*log* $k_{w,add.}$	*log* $k_{w,pred.}$*	*Residuals*
OH–C_6H_3–2-NH_2–4-Cl	1.928	1.091	2.094	−0.166
CH_3–C_6H_3–2-OH–4-Cl	2.720	2.775	2.735	−0.015
Cl–C_6H_3–2-NO_2–4-NO_2	2.270	2.444	2.280	−0.010
CH_3–C_6H_3–2-NO_2–4-NO_2	2.340	2.316	2.217	0.123
CH_3–C_6H_3–3-NO_2–4-NO_2	2.506	2.316	2.622	−0.116
NH_2–C_6H_3–2-NO_2–4-OCH_3	1.977	0.810	1.961	0.016
NH_2–C_6H_3–2-CH_3–5-NO_2	1.992	1.437	1.941	0.051
OH–C_6H_3–2-OH–4–NO_2	2.611	0.563	2.373	0.238
CH_3–C_6H_3–3-OH–6-NO_2	2.634	1.842	2.483	0.151
OH–C_6H_3–2-OH–3-OH	0.832	0.089	0.783	0.049
OH–C_6H_2–2-Cl–4-Cl–5-Cl	4.148	3.633	4.136	0.012

* The predicted log k_w values are calculated according to eqn. 17.

values are in very good agreement, as shown by eqn. 19 in which the regressor of log $k_{exp.}$ is practically one and the intercept practically zero:

$$\log k_{pred.} = 0.989(\pm 0.109) \log k_{exp.} - 0.004(\pm 0.269)$$
$$n = 11; \quad r = 0.989; \quad s = 0.120; \quad F = 425 \tag{19}$$

At this stage, we conclude that eqn. 17 appears suitable for calculating lipophilic indices (log k_w) of polysubstituted benzenes.

ACKNOWLEDGEMENTS

A.T.-K. is indebted to the University of Athens for a sabbatical leave. B.T. and H.V.d.W. acknowledge the help of the Swiss National Science Foundation (research grant 3.539-0.83).

APPENDIX

When applying eqn. 17, all possible interactions between *meta* and *para* substituents are taken into account for the calculation of the $\rho\sigma$ term, while the $\rho\sigma_0$ term is based on all interactions between *ortho* substituents. For I_0 the values of the corresponding disubstituted benzene are used (Table IV). Taking 2-amino-4-chlorophenol (the first compound in Table VI) as an example:

since for Cl $\rho = 0$, the $\rho\sigma$ term includes: $\rho(NH_2) \cdot \sigma(Cl) + \rho(OH) \cdot \sigma(Cl)$;
$\rho\sigma_0$ includes $\rho(OH) \cdot \sigma(NH_2) + \rho(NH_2) \cdot \sigma(OH)$;
$I_0 = 2$, the value for 2-aminophenol;
applying eqn. 17:

$$\log k_{pred.} = 0.943 \cdot 1.091 + 0.790 \cdot [0.74 \cdot 0.37 + 0.94 \cdot 0.27] + 0.255 \cdot [0.94 \cdot (-0.38) + 0.74 \cdot (-0.13)] + 0.297 \cdot 2 + 0.170 = 2.094$$

REFERENCES

1 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979.
2 T. Braumann, *J. Chromatogr.*, 373 (1986) 191.
3 T. Fujita, J. Iwasa and C. Hansch, *J. Am. Chem. Soc.*, 86 (1964) 5175.
4 A. Brändström, *Acta Pharm. Suec.*, 19 (1982) 175.
5 T. Fujita and T. Nishioka, *Progr. Phys. Org. Chem.*, 12 (1976) 49.
6 T. Fujita, *Progr. Phys. Org. Chem.*, 14 (1983) 75.
7 T. Fujita, *J. Pharm. Sci.*, 72 (1983) 285.
8 A. Leo, *J. Chem. Soc. Perkin Trans. II*, (1983) 825.
9 S. J. Lewis, M. S. Mirrlees and P. J. Taylor, *Quant. Struct.-Act. Relat.*, 2 (1983) 100.
10 R. Franke, S. Dove and R. Kühne, *Eur. J. Med. Chem.*, 14 (1979) 363.
11 W. J. Dunn III, E. Johansson and S. Wold, *Quant. Struct.-Act. Relat.*, 2 (1983) 156.
12 N. El Tayar, H. Van de Waterbeemd and B. Testa, *Quant. Struct.-Act. Relat.*, 4 (1985) 77.
13 L. Hammett, *Physical Organic Chemistry: Reaction Rates, Equilibria and Mechanism*, McGraw-Hill, New York, 2nd ed., 1970.
14 H. Van de Waterbeemd and B. Testa, *Int. J. Pharm.*, 14 (1983) 29.
15 N. El Tayar, H. Van de Waterbeemd and B. Testa, *J. Chromatogr.*, 320 (1985) 293.

16 N. El Tayar, H. Van de Waterbeemd and B. Testa, *J. Chromatogr.*, 320 (1985) 305.
17 O. Exner, in B. Chapman and J. Shorter (Editors), *Advances in Linear Free Energy Relationship*, Plenum, London, 1972, p. 1.
18 H. Mager and A. Barth, *Pharmazie*, 34 (1979) 557.
19 B. Testa, *Principles of Organic Stereochemistry*, Dekker, New York, 1979, p. 28.
20 R. F. Rekker and H. M. De Kort, *Eur. J. Med. Chem.*, 14 (1979) 479.
21 A. Albert and E.P. Serjeant, *The Determination of Ionization Constants*, Chapman and Hall, London, 3rd ed., 1984.