Journal of Chromatography, 295 (1984) 341–353
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 16,775

A CRITICAL APPRAISAL OF LOG *P* FRAGMENTAL PROCEDURES AND CONNECTIVITY INDEXING FOR REVERSED-PHASE THIN-LAYER CHROMATOGRAPHIC AND HIGH-PERFORMANCE LIQUID CHROMATOGRAPHIC DATA OBTAINED FOR A SERIES OF BENZOPHENONES

ANNA KAKOULIDOU* and ROELOF F. REKKER*

Department of Pharmacochemistry, Free University, De Boelelaan 1083, 1081 HV Amsterdam (The Netherlands)

(First received January 20th, 1984; revised manuscript received March 22nd, 1984)

SUMMARY

High-performance liquid chromatography and reversed-phase thin-layer chromatography were used to study the relationships between retention and hydrophobicity in a series of substituted benzophenones. The principal aim of the investigation was the critical appraisal of the existing log *P* fragmental procedures (those of Rekker and of Leo and Hansch) and Kier's connectivity indices. The investigation is to be considered as an extension of a previously published examination of the chromatographic behaviour of simple alkyl benzenes. Substituted benzophenones represent a class of structures where cross-conjugation and steric decoupling of resonance demand precautionary measures in order to keep the quality of the relationships between retention and hydrophobicity parameters at the same high level as attained for alkylbenzenes.

INTRODUCTION

Partition chromatographic data (R_M and log k' values) are increasingly used as substitutes for the log *P* values of organic compounds. The equations used for the transfer of the obtained data:

$$\log P = a' R_M + b' \quad (1)$$

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

are actually extensions of the Collander equation:

$$\log P_1 = a \log P_2 + b \quad (3)$$

* Visiting scientist from the Laboratory of Pharmaceutical Chemistry, Department of Pharmacy, University of Athens, Athens, Greece.

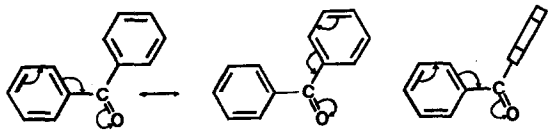
0021-9673/84/$03.00

where P_1 and P_2 represent partition coefficients in solvent systems 1 and 2, respectively.

In a previous paper Koopmans and Rekker[1] reported on the high-performance liquid chromatography (HPLC) of alkylbenzenes. That investigation concerned a simply structured series of congeneric compounds and the capacity factors could be coupled with high precision to hydrophobicity values (correlation coefficients higher than 0.99).

Hydrophobicity was expressed in terms of solvent partition values either from experiments or by calculation (Rekker method and Leo and Hansch method) and by means of Kier's connectivity values. The best results were obtained by application of the Rekker procedure.

This study was performed on a series of benzophenones, including benzophenone, 2-methyl-, 4-methyl-, 2,6-dimethyl-, 2,2'-dimethyl-, 4,4'-dimethyl-, 2,6,2'-trimethyl-, 2,3,5,6-tetramethyl-, 2,6,2',6'-tetramethyl-, 2,3,5,6,2',6'-hexamethyl-, 2,4,6,2',4',6'-hexamethyl-, 2,2'-diethyl-, 2,6,2',6'-tetraethyl-, 2,2'-diisopropyl;, 2-*tert.*-butyl-, 3-chloro-4'-*tert.*-butyl-, 4-chloro-, 4,4'-dichloro-, 3,5,3',5'-tetrachloro- and 4,4'-difluorobenzophenone. They constitute a class of compounds in which cross-conjugation effects operate which will disappear on the introduction of sufficient *ortho* bulk (see formula of benzophenone; A = cross-conjugated and B = decoupled cross-conjugation), so that it stands to reason that a high-quality straightforward correlation as was obtained for the alkylbenzenes is not so evident.



A B

EXPERIMENTAL

The benzophenones investigated were of several origins from laboratory stock. They were all of sufficient purity, no interfering spots or peaks being present in the reversed-phase thin-layer chromatographic (RPTLC) and HPLC recordings. The RPTLC experiments were performed on pre-coated silica gel 60 F_{254} TLC plates (Merck) impregnated with paraffin oil and eluted with two different acetone–water mixtures (65:35 and 60:40). The plates were developed in a closed chromatographic tank, dried at *ca.* 75°C and the spots were located under UV light. R_F values were averaged from at least ten determinations and converted into R_M values via the relationship

$$R_M = \log (1/R_F - 1) \tag{4}$$

The two sets of R_M values could be related with a correlation coefficient of $r = 0.997$. Fig. 1 shows an example of one of the RPTLC experiments.

The HPLC experiments were performed with a Waters Assoc. HPLC system using UV detection at 254 nm. The column used was an RP type (μBondapack C_{18},

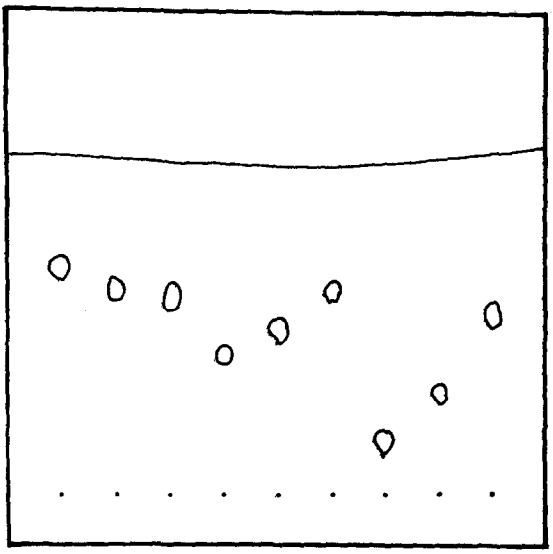


Fig. 1. Example of RPTLC experiments. Spots from left to right: benzophenone, 2-methyl-, 4-methyl-, 2,6,2′,6′-tetramethyl-, 2,6,2′-trimethyl-, 4,4′-difluoro-, 3,5,3′,5′-tetrachloro-, 2,4,6,2′,4′,6′-hexamethyl- and 4,4′-dimethylbenzophenone. Eluent: acetone–water (65:35).

Waters Assoc.). The eluent was de-gassed methanol–water (70:30 or 75:25) and the flow-rate was 2.3 ml/min.

Retention times were expressed as log (capacity factors, k') by:

$$\log k' = \log[(t_r - t_0)/t_0] \tag{5}$$

where t_r represents the retention time of the compound and t_0 denotes the retention time of an unretained peak generated by formamide. The two sets of log k' values could be related with $r = 0.9997$. A representative HPLC trace is shown in Fig. 2.

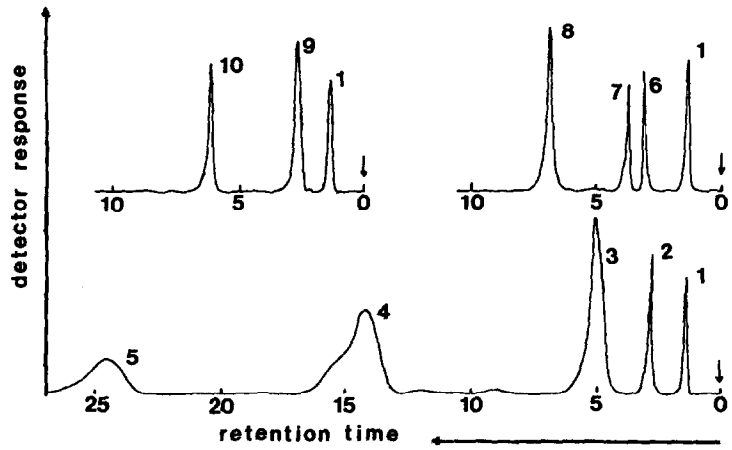


Fig. 2. Example of HPLC trace of benzophenones. Peaks: 1 = formamide (reference); 2 = benzophenone; 3 = 2,6,2′-trimethyl-; 4 = 2,4,6,2′,4′,6′-hexamethyl-; 5 = 3,5,3′,5′-tetrachloro-; 6 = 4,4′-difluoro-; 7 = 4-methyl-; 8 = 2,6,2′,6′-tetramethyl-; 9 = 2-methyl-; 10 = 4,4′-dimethylbenzophenone.

The calculation of the log P of benzophenone in Rekker's system requires one c_M factor of 0.289 to account for cross-conjugation: $2 \cdot 1.840(C_6H_5) - 0.776(CO_{ar}) + 0.289\ (c_M) = 3.193$. In Leo and Hansch's system the cross-conjugation effect is not accounted for but is actually incorporated in their $>CO^{\varphi\varphi}$ fragment value advised for calculation, and taking into account one bond factor the final calculation is as follows: $2 \cdot 1.90(C_6H_5) - 0.50(>CO^{\varphi\varphi}) - 0.12(F_b) = 3.18$*. It should be noted that $>CO^{\varphi}$ equals -1.09 and hence the difference from $>CO^{\varphi\varphi}$ is 0.59. The latter value actually represents the extra contribution of the second phenyl ring to lipophilicity and is, at least partially, the equivalent of Rekker's cross-conjugation factor.

Molecular connectivity indices have been calculated from zero to sixth order with computer program CFUNC[2].

RESULTS AND DISCUSSION

Straightforward correlations between chromatographic data and calculated log P values or molecular connectivity indices are presented in Table I. Their quality is not as high as that achieved in the above-mentioned investigation on alkylbenzenes[1] and 3,5,3',5'-tetrachlorobenzophenone is a distinct outlier. Although the exclusion of the latter compound gives significantly improved results, they remain unsatisfac-

TABLE I

CORRELATIONS OF RPTLC AND HPLC DATA WITH CALCULATED LOG P VALUES AND MOLECULAR CONNECTIVITY INDICES

Σf_R = summations of Rekker fragmental constants; Σf_{LH} = summations of Leo and Hansch fragmental constants; ${}^0\chi^v$ = zero-order valence connectivities. 95% confidence levels in parentheses. Although perhaps not very significant, it is common practice to give regressor values and confidence levels to three decimal places in the tabulated type of equations.

System	*Regression equations*	*n*	*r*	*s*	*F*	*Eqn.*
RPTLC: acetone–water (65:35)	$\Sigma f_R = 3.896\ (\pm 0.597)\ R_M + 4.532\ (\pm 0.176)$	17	0.947	0.395	131	6
	$\Sigma f_{LH} = 4.300\ (\pm 1.006)\ R_M + 4.743\ (\pm 0.297)$	17	0.888	0.666	56	7
	$R_M = 0.128\ (\pm 0.267)\ {}^0\chi^v - 1.241\ (\pm 0.284)$	17	0.908	0.126	70	8
	3,5,3',5'-Tetrachlorobenzophenone excluded:					
	$\Sigma f_R = 4.640\ (\pm 0.356)\ R_M + 4.562\ (\pm 0.089)$	16	0.987	0.042	528	9
	$\Sigma f_{LH} = 5.477\ (\pm 0.707)\ R_M + 4.791\ (\pm 0.176)$	16	0.964	0.392	186	10
	$R_M = 0.117\ (\pm 0.015)\ {}^0\chi^v - 1.152\ (\pm 0.161)$	16	0.963	0.070	180	11
HPLC: methanol–water (70:30)	$\Sigma f_R = 3.088\ (\pm 0.373)\ \log k' + 3.136\ (\pm 0.251)$	19	0.961	0.321	207	12
	$\Sigma f_{LH} = 3.399\ (\pm 0.602)\ \log k' + 3.197\ (\pm 0.408)$	19	0.921	0.522	95	13
	$\log k' = 0.172\ (\pm 0.026)\ {}^0\chi^v - 1.226\ (\pm 0.280)$	19	0.940	0.123	129	14
	3,5,3',5'-Tetrachlorobenzophenone excluded:					
	$\Sigma f_R = 3.428\ (\pm 0.294)\ \log k' + 2.999\ (\pm 0.184)$	18	0.981	0.224	411	15
	$\Sigma f_{LH} = 3.978\ (\pm 0.452)\ \log k' + 2.964\ (\pm 0.282)$	18	0.968	0.344	235	16
	$\log k' = 0.162\ (\pm 0.169)\ {}^0\chi^v - 1.136\ (\pm 0.191)$	18	0.972	0.078	275	17

* It is logical that this value is in full accordance with the experimental result, because the $>CO^{\varphi\varphi}$ value has been derived exclusively from this observed log P value by subtraction of $2f(C_6H_5)$ and F_b.

tory*. In a further evaluation three possible factors were considered: (1) substitution effects in the benzophenone molecule: (a) halogen-effect; (b) *ortho*-substitution effect; and (c) *para*-substitution effect; (2) systemic differences between RPTLC and HPLC; and (3) limiting factors inherent in the three different modes of parametrization: Rekker's fragmental constant[3,4], Leo and Hansch's fragmental constant[5,6] or Kier's connectivity indices[2,7,8].

Separate treatment of halogen- and non-halogen-substituted ketones gave significantly different slopes for alkyl- and halogen-substituted compounds with the 3-chloro-4-*tert.*-butyl-derivative fitting surprisingly well in the equation for alkyl derivatives, especially when Rekker's system or Kier's indices were applied. The results of these preliminary investigations made us decide to perform individual correlation studies on alkyl- and halogen-substituted benzophenones.

Alkyl-substituted benzophenones: Rekker's fragmental method

RPTLC data. The straightforward correlation can be expressed by

$$\Sigma f_R = 4.715\ (\pm 0.300)\ R_M + 4.626\ (\pm 0.078)^{**} \tag{18}$$

$$n = 12;\ r = 0.9939;\ s = 0.145;\ F = 813$$

This equation, although satisfactory, is significantly improved when the effect of *ortho* substitution is taken into account. As soon as sufficient *ortho* bulk is introduced into the molecule, cross-conjugation will no longer operate and consequently one factor c_M should be subtracted in the calculation. This is preferably done by applying *kn* as a second parameter (-1 for absence of cross-conjugation and 0 for presence).

$$\Sigma f_R = 4.292\ (\pm 0.221)\ R_M - 0.317\ (\pm 0.113)\ kn + 4.463\ (\pm 0.071) \tag{19}$$

$$n = 12;\ r = 0.9983;\ s = 0.077;\ F = 1449$$

It should be noted that the regressor of *kn* agrees well with the value $c_M = 0.289$ (ref. 4) and comparable statistics are obtained when incorporating the correction for loss of cross-conjugation directly in the calculation: $r = 0.9981$, $s = 0.074$ and $F = 2629$. Details on the experimental data as used in eqns. 18 and 19 are given in Table II.

HPLC data. The equation initially obtained is

$$\Sigma f_R = 3.613\ (\pm 0.297)\ \log k' + 2.936\ (\pm 0.188) \tag{20}$$

$$n = 14;\ r = 0.9875;\ s = 0.191;\ F = 469$$

---

* We are of opinion that satisfactory correlation levels are those with $r \geqslant 0.99$, although in common practice investigators frequently seem satisfied with $r \geqslant 0.95$; the lower value leaves as much as 10% of the variance unaccounted for and is not consistent with the accuracy of RPTLC and HPLC experiments.

** 95% confidence levels in parentheses.

TABLE II

RPTLC DATA (ACETONE–WATER, 65:35) FOR BENZOPHENONES, INCLUDING THE APPLIED PARAMETERS FOR CORRELATION

kn = key number: incidence of constant 0.289 in the regression equation; nm = not measured on account of non-availability at the time of the experiments; see also notes in Table I.

Substituent	R_F	R_M	Σf_R	*kn*	Σf_{LH}	$^0\chi^v$	$^1\chi^v$
None	0.657 ± 0.035	−0.282 ± 0.046	3.193	0	3.18	7.682	4.525
2-Methyl	0.598 ± 0.021	−0.172 ± 0.040	3.712	0	3.84	8.604	4.942
4-Methyl	0.602 ± 0.013	−0.179 ± 0.027	3.782	0	3.84	8.604	4.936
2,6-Dimethyl	0.565 ± 0.016	−0.114 ± 0.028	4.231	−1	4.50	9.527	5.359
2,2′-Dimethyl	0.529 ± 0.008	−0.051 ± 0.013	4.231	0	4.50	9.527	5.359
4,4′-Dimethyl	0.540 ± 0.013	−0.070 ± 0.018	4.231	0	4.50	9.527	5.347
2,6,2′-Trimethyl	0.497 ± 0.004	0.006 ± 0.005	4.751	−1	5.16	10.450	5.775
2,3,5,6-Tetramethyl	nm						
2,6,2′,6′-Tetramethyl	0.432 ± 0.003	0.119 ± 0.005	5.269	−1	5.82	11.372	6.192
2,3,5,6,2′,6′-Hexamethyl	0.315 ± 0.006	0.338 ± 0.013	6.307	−1	7.14	13.218	7.025
2,4,6,2′,4′,6′-Hexamethyl	0.314 ± 0.010	0.340 ± 0.020	6.307	−1	7.14	13.218	7.014
2,2′-Diethyl	nm						
2,6,2′,6′-Tetraethyl	0.192 ± 0.008	0.625 ± 0.023	7.345	−1	7.98	14.201	8.435
2,2′-Diisopropyl	nm						
2-*tert.*-Butyl	0.450 ± 0.005	0.087 ± 0.009	5.269	−1	5.44	11.104	6.192
3-Cloro-4′-*tert.*-butyl	0.304 ± 0.010	0.359 ± 0.021	6.011	0	6.15	12.222	6.694
4-Chloro	0.534 ± 0.019	−0.059 ± 0.033	3.935	+1	3.89	8.800	5.034
4,4′-Dichloro	0.447 ± 0.008	0.092 ± 0.012	4.677	+1	4.60	9.917	5.542
3,5,3′,5′-Tetrachloro	0.167 ± 0.010	0.699 ± 0.031	6.161	+4	6.02	12.153	6.559
4,4′-Difluoro	0.624 ± 0.019	−0.220 ± 0.035	3.611	0	3.46	7.080	4.123

When considering the effect of *ortho* substitution in the above-described way we obtain

$$\Sigma f_R = 3.289\ (\pm 0.245) \log k' - 0.341\ (\pm 0.151)\ kn + 2.920\ (\pm 0.126) \quad (21)$$

$$n = 14;\ r = 0.9946;\ s = 0.132;\ F = 543$$

This equation, although highly significant in itself, is of a slightly lower quality than eqn. 19; the regressor of *kn* is fairly high (0.341 instead of the expected value of 0.289) and two obviously too high estimates (not real outliers, however) are present: for 4,4'-dimethylbenzophenone, calculated 4.231, estimated 4.433, difference −0.20; for 2,4,6,2',4',6'-hexamethylbenzophenone, calculated 6.307, estimated 6.541, difference −0.23. Because both are *para*-substituted, we decided to omit all three *para*-substituted alkylderivatives from the regression. A re-calculation led to a clear improvement, especially with regard to the regressor of *kn* (0.258; expected value, 0.289):

$$\Sigma f_R = 3.418\ (\pm 0.131) \log k' - 0.258\ (\pm 0.083)\ kn + 2.951\ (\pm 0.072) \quad (22)$$

$$n = 11;\ r = 0.9988;\ s = 0.061;\ F = 1840$$

Some resonance effect from the alkyl substituent in the *para*-position could be a factor in the observed lipophilicity increase, although it apparently escapes observation in TLC experiments; this would indicate that the HPLC method is distinctly more sensitive than RPTLC in detailing lipophilic behaviour, so that, connected with the quantifiability of at least part of the lipophilicity factor[3], an increase of 0.289 is just not feasible in our TLC experiments.

In our revised conceptional approach around the factor 0.289 (ref. 4) we no longer maintain any differentiation, *i.e.*, irrespective of the origin (proximity effect, conjugation increase, ring condensation, etc.) the unique vale of 0.289 is advised, but should the occasion arise a closer check is recommended. With this in mind, we performed a double dummy parametrization in eqn. 21: D_1 for cross-conjugation (present, $D_1 = 0$; absent, $D_1 = 1$) and D_2 for enforcement of resonance effects (if so, $D_2 = 1$; if not, $D_2 = 0$). The result is following high-grade equation:

$$\Sigma f_R = 3.351\ (\pm 0.114) \log k' + 0.272\ (\pm 0.075)\ D_1 - 0.264\ (\pm 0.078)\ D_2 + 2.979\ (\pm 0.065) \quad (23)$$

$$n = 14;\ r = 0.9986;\ s = 0.070;\ F = 1482$$

It should be noted that the regressors of the two dummy parameters actually have identical values not significantly different from 0.289, with opposite algebraic signs. This signifies $c_M = 0.289$ as a suitable constant to express both resonance effect changes from *para*-substituents and cross-conjugation losses. Indeed, similar results as those visualized by eqn. 23 are obtained with a *kn* approach consisting of the introduction of −1 for cross-conjugation loss, +1 for an extra resonance effect and 0 when (a) no extra resonance is present and cross-conjugation is supposedly unaf-

TABLE III

REVERSED-PHASE HPLC DATA (ELUENT METHANOL–WATER, 70:30) FOR BENZOPHENONES, INCLUDING THE APPLIED PARAMETERS FOR CORRELATION

$^4\chi^v$ = fourth-order valence connectivities; $^4\chi^v_{pc}$ = fourth-order valence (patch/cluster) connectivities; elution rate = 2.3 ml/min; t_0 = 1.30 min; see also notes in Tables I and II.

Substituent	t_r	*log k'*	Σf_R	*kn*	Σf_{LH}	$^0\chi^v$	$^1\chi^v$	$^4\chi^v_{pc}$	$^4\chi^v$
None	2.83	0.071	3.193	0	3.18	7.682	4.525	0.462	1.807
2-Methyl	3.41	0.210	3.712	0	3.84	8.604	4.942	0.746	2.318
4-Methyl	3.72	0.269	3.712	+1	3.84	8.604	4.936	0.654	2.115
2,6-Dimethyl	3.82	0.288	4.231	−1	4.50	9.527	5.359	0.991	2.897
2,2'-Dimethyl	4.43	0.382	4.231	0	4.50	9.527	5.359	1.030	2.830
4,4'-Dimethyl	5.05	0.460	4.231	+1	4.50	9.527	5.347	0.847	2.423
2,6,2'-Trimethyl	5.19	0.476	4.750	−1	5.16	10.450	5.775	1.275	3.410
2,3,5,6-Tetramethyl	6.54	0.605	5.269	−1	5.82	11.372	6.192	1.748	4.039
2,6,2',6'-Tetramethyl	6.82	0.628	5.269	−1	5.82	11.372	6.192	1.520	3.991
2,3,5,6,2',6'-Hexamethyl	nm								
2,4,6,2',4',6'-Hexamethyl	14.21	0.997	6.307	0	7.14	13.218	7.014	1.809	5.010
2,2'-Diethyl	7.57	0.683	5.269	0	5.58	10.941	6.480	1.052	3.297
2,6,2',6'-Tetraethyl	22.34	1.209	7.345	−1	7.98	14.201	8.435	1.589	4.782
2,2'-Diisopropyl	11.41	0.891	6.307	−1	6.64	12.682	7.246	1.722	4.282
2-*tert.*-Butyl	6.21	0.577	5.269	−1	5.44	11.104	6.192	1.615	3.945
3-Chloro-4'-*tert.*-butyl	14.27	0.999	6.011	+1	6.15	12.222	6.694		4.256
4-Chloro	4.09	0.332	3.935	+1	3.89	8.800	5.034		2.185
4,4'-Dichloro	5.81	0.540	4.677	+1	4.60	9.917	5.542		2.563
3,5,3',5'-Tetrachloro	24.47	1.251	6.161	+4	6.02	12.153	6.559		3.921
4,4'-Difluoro	2.97	0.109	3.611	−1	3.46	7.080	4.123		1.544

fected or (b) when both dummy corrections counterbalance each other. The following equation is obtained:

$$\Sigma f_R = 3.385\ (\pm 0.104) \log k' - 0.254\ (\pm 0.044)\ kn + 2.972\ (\pm 0.061) \quad (24)$$

$$n = 14;\ r = 0.9987;\ s = 0.061;\ F = 2333$$

Comparable statistics are obtained when the correction both for cross-conjugation loss and resonance enhancement are directly incorporated in the calculation: $r = 0.9983$, $s = 0.064$ and $F = 3603$.

Detailed information on log k' and parametrization as applied in the above correlations is given in Table III.

Alkyl-substituted benzophenones: Leo and Hansch's fragmental method

RPTLC data. The straightforward correlation initially obtained can be expressed by

$$\Sigma f_{LH} = 5.623\ (\pm 0.506)\ R_M + 4.950\ (\pm 0.131) \quad (25)$$

$$n = 12;\ r = 0.9879;\ s = 0.245;\ F = 406$$

Its quality is clearly lower than that of eqn. 18. The introduction of a dummy parameter for loss of cross-conjugation leads to a more significant equation with all statistical evidence improved ($D = 1$ for loss of cross-conjugation and $D = 0$ when cross-conjugation is maintained):

$$\Sigma f_{LH} = 5.066\ (\pm 0.545)\ R_M + 0.419\ (\pm 0.280)\ D + 4.736\ (\pm 0.176) \quad (26)$$

$$n = 12;\ r = 0.9928;\ s = 0.200;\ F = 339$$

When coplanarity between the two phenyl nuclei in the benzophenone system is lost owing to sufficient *ortho* bulk, there is no reason to use the fragment $>CO^{\varphi\varphi}$ any longer, the fragment $>CO^{\varphi}$ becoming more suitable in the calculations. This change of "CO" fragment leads to

$$\Sigma f_{LH} = 4.838\ (\pm 0.398)\ R_M + 4.698\ (\pm 0.103) \quad (27)$$

$$n = 12;\ r = 0.9898;\ s = 0.193;\ F = 484$$

The quality of eqn. 27 remains inferior to its predecessor with the dummy parameter; this is easy to understand, however, as the regressor of the dummy parameter is significantly lower than the difference between the fragment values of $>CO^{\varphi\varphi}$ and $>CO^{\varphi}$. This discrepancy will diminish, however, when we assume that *ortho* bulk decreases the flexibility in the molecule, consequently making any bond correction factor (F_b) immaterial. This results in a slight, though significant, improvement of the regression equation:

$$\Sigma f_{LH} = 4.998\ (\pm 0.376)\ R_M + 4.710\ (\pm 0.097) \quad (28)$$

$$n = 12;\ r = 0.9915;\ s = 0.182;\ F = 582$$

HPLC data. The straightforward equation initially obtained can be expressed as follows:

$$\Sigma f_{LH} = 4.195\ (\pm 0.348) \log k' + 2.960\ (\pm 0.220) \quad (29)$$

$$n = 14;\ r = 0.9873;\ s = 0.224;\ F = 463$$

The introduction of two dummy parameters for adequately describing the two different effects observed in HPLC (see above) leads to the following equation:

$$\Sigma f_{LH} = 3.814\ (\pm 0.192) \log k' + 0.409\ (\pm 0.122)\ D_1 - 0.138\ (\pm 0.120)\ D_2 + 2.967\ (\pm 0.099) \quad (30)$$

$$n = 14;\ r = 0.9977;\ s = 0.104;\ F = 866$$

The dummy parameter D_1 is connected with cross-conjugation loss and D_2 with increasing resonance originating from *para* substitution. Eqn. 30 is of good quality; when performing a calculation with $\geq CO^{\varphi\varphi}$ changed for $\geq CO^{\varphi}$, however, D_2, appearing with a low regressor and a high deviation in eqn. 30, is no longer significant and the best equation under these changed conditions of parametrization is now

$$\Sigma f_{LH} = 3.636\ (\pm 0.185) \log k' + 2.933\ (\pm 0.117) \quad (31)$$

$$n = 14;\ r = 0.9951;\ s = 0.119;\ F = 1222$$

and the omission of the bond factor correction (F_b) results in the following equation:

$$\Sigma f_{LH} = 3.749\ (\pm 0.162) \log k' + 2.938\ (\pm 0.103) \quad (32)$$

$$n = 14;\ r = 0.9965;\ s = 0.104;\ F = 1693$$

In conclusion, the Leo and Hansch fragmental system does not lend itself for a proper treatment of extra resonance effects as originating from *para* substitution and although correlations are acceptably high in statistical merits, they remain poorer than those obtained with Rekker's fragmental system.

Alkyl-substituted benzophenones: Kier's molecular connectivities

RPTLC data. The situation with the alkylbenzophenone data can most easily be expressed by applying the ${}^1\chi^v$ index instead of ${}^0\chi^v$ as applied in eqn. 11:

$$R_M = 0.236\ (\pm 0.111)\ {}^1\chi^v - 1.343\ (\pm 0.067) \quad (33)$$

$$n = 12;\ r = 0.9966;\ s = 0.023;\ F = 1475$$

HPLC data. In the correlation of log k' of alkylbenzophenones by means of molecular connectivity indexing, the ${}^0\chi^v$ index gives the best results but only when combined with the higher order ${}^4\chi^v_{pc}$ index:

$$\log k' = 0.167\ (\pm 0.015)\ {}^0\chi^v - 1.220\ (\pm 0.166) \quad (34)$$

$$n = 14;\ r = 0.9844;\ s = 0.037;\ F = 375$$

$$\log k' = 0.215\ (\pm 0.022)\ {}^{0}\chi^{v} - 0.223\ (\pm 0.091)\ {}^{4}\chi^{v}_{pc} - 1.458\ (\pm 0.143) \quad (35)$$
$$n = 14;\ r = 0.9939;\ s = 0.037;\ F = 483$$

The improvement in eqn. 35 with respect to eqn. 34 is evident. The omission of the *para*-alkyl-substituted compounds from the series results in an equation in which the ${}^{1}\chi^{v}$ index gives the best results:

$$\log k' = 0.292\ (\pm 0.017)\ {}^{1}\chi^{v} - 1.221\ (\pm 0.103) \quad (36)$$
$$n = 11;\ r = 0.9956;\ s = 0.032;\ F = 1019$$

The above results give, in addition to those in eqn. 24, a further indication of the exclusive behaviour of *para*-alkyl-substituted benzophenones in HPLC experiments and it can perhaps be expressed most correctly by stating that *para*-alkyl substituents introduce the element of non-congenericity in the original series of compounds. Congenericity can be restored by applying either the parameter *kn* (eqn. 24) or a connectivity index of higher order, ${}^{4}\chi^{v}_{pc}$ (eqn. 35).

Halogen-substituted benzophenones

The regression equations obtained with six halogen-substituted benzophenones, including the unsubstituted compound, are of an unacceptable quality owing to the unsatisfactory fit of the unsubstituted ketone and its 3-chloro-4'.*tert.*-butyl derivative. Exclusion of these two compounds from calculation results in sets of equations that are of good quality for both the RPTLC and the HPLC experiments, the only drawback being that the number of data points (not more than four) is low.

RPTLC data.

$$\Sigma f_{R} = 2.807\ (\pm 0.685)\ R_M + 4.237\ (\pm 0.254) \quad (37)$$
$$n = 4;\ r = 0.9931;\ s = 0.163;\ F = 143$$

$$\Sigma f_{LH} = 2.773\ (\pm 0.713)\ R_M + 4.138\ (\pm 0.264) \quad (38)$$
$$n = 4;\ r = 0.9923;\ s = 0.170;\ F = 129$$

$$R_M = 0.397\ (\pm 0.098)\ {}^{4}\chi^{v} - 0.885\ (\pm 0.264) \quad (39)$$
$$n = 4;\ r = 0.9929;\ s = 0.058;\ F = 140$$

HPLC data.

$$\Sigma f_{R} = 2.284\ (\pm 0.469) \log k' + 3.322\ (\pm 0.330) \quad (40)$$
$$n = 4;\ r = 0.9957;\ s = 0.127;\ F = 202$$

$$\Sigma f_{LH} = 2.259\ (\pm 0.433) \log k' + 3.232\ (\pm 0.304) \quad (41)$$
$$n = 4;\ r = 0.9957;\ s = 0.127;\ F = 233$$

$$\log k' = 0.491\ (\pm 0.085)\ {}^{4}\chi^{v} - 0.695\ (\pm 0.231) \quad (42)$$
$$n = 4;\ r = 0.9965;\ s = 0.051;\ F = 281$$

The necessity for applying a higher order connectivity index, $^4\chi^v$, is once more indicative of the topological complexity of the investigated benzophenone type. With $^0\chi^v$ the correlation coefficients are not higher than 0.9659 and 0.9755, respectively.

It is clear that the small number of data points incorporated in eqns. 37–42 does not allow for any further differentiation by means of an extra parameter, either *kn* or a dummy.

However, when due attention is paid to the mutual fit of the four benzophenone structures in eqns. 37 and 40 and the information obtained is then combined with the difference in slope of eqns. 37 and 40 with eqns. 19 and 21, respectively, it is not difficult to incorporate the complete sets of structures in correlations of surprisingly high quality. The necessary key numbers for eqns. 43 and 44 are given in Tables II and III.

$$\Sigma f_R = 4.247\ (\pm 0.116)\ R_M - 0.312\ (\pm 0.027)\ kn + 4.482\ (\pm 0.034) \qquad (43)$$

$$n = 17;\ r = 0.9982;\ s = 0.077;\ F = 2052$$

$$\Sigma f_R = 3.340\ (\pm 0.081)\ \log k' - 0.251\ (\pm 0.022)\ kn + 3.004\ (\pm 0.053) \qquad (44)$$

$$n = 19;\ r = 0.9984;\ s = 0.069;\ F = 2572$$

Admittedly, it is far from easy to give a satisfactory answer to the question of how to account for the key number sequences of the five halogen-substituted benzophenones just incorporated in the regression equations 43 and 44. We aimed at a highly correlated equation with a quality expressed by an *r* value equal to or even better than 0.99, drawing support from the results obtained in the correlation studies on HPLC data for alkylbenzenes with a parametrization essentially not differing from that applied in this investigation.

The positive *kn* accompanying chlorine substitution in benzophenone points towards a reinforcement of the resonance effect, but the appearance of $kn = -1$ for 4,4'-difluorobenzophenone in the HPLC regression equation is incomprehensible at present, and the same is true of the absence of any effect of the second *para*-chloro substitution.

A study of the effect of chloro substitution in a less complicated set of aromatics is clearly desirable and we intend to extend our efforts in this direction in order to be able to progress to some additional aromatic ketones, including halogen-substituted derivatives.

CONCLUSIONS

The behaviour of benzophenones in RPTLC and HPLC experiments is much more complicated than that of alkylbenzenes. In order to maintain for benzophenones the high statistical evidence attained in a previous study on alkylbenzenes ($r > 0.99$); it was necessary to make due provisions for resonance reinforcement and resonance decoupling, depending on the substitution pattern in the ketone structure.

Rekker's fragmental constants, Kier's connectivity indices and Leo and

Hansch' fragmental constants were applied for parametrization of the experimentally obtained R_M and log k' values. The importance of these parametrizations was evaluated; the results seem to be in favour of Rekker's fragmental system.

REFERENCES

1 R. E. Koopmans and R. F. Rekker, *J. Chromatogr.*, 285 (1984) 267.
2 L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
3 R. F. Rekker, *The Hydrophobic Fragmental Constant; its Derivation and Application; a Means of Characterizing Membrane Systems*, Elsevier, Amsterdam, 1977.
4 R. F. Rekker and H. M. de Kort, *Eur. J. Med. Chem.*, 14 (1979) 479.
5 A. Leo, P. Y. C. Jow, C. Silipo and C. Hansch, *J. Med. Chem.*, 18 (1975) 865.
6 C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York, 1979.
7 L. B. Kier and L. H. Hall, *J. Pharm. Sci.*, 70 (1981) 583.
8 T. W. Schultz, L. B. Kier and L. H. Hall, *Bull. Environ. Contam. Toxicol.*, 28 (1982) 373.