CHEMOMETRICAL ANALYSIS OF SUBSTITUENT EFFECTS. IV. ADDITIVITY OF SUBSTITUENT EFFECTS IN DISSOCIATION OF 3,5-DISUBSTITUTED BENZOIC ACIDS IN ORGANIC SOLVENTS

Oldrich PYTELA, Jiri KULHANEK and Miroslav LUDWIG
Department of Organic Chemistry,
Faculty of Chemical Technology, University of Pardubice, 532 10 Pardubice, The Czech Republic

Received December 14, 1993
Accepted February 20, 1994

Ten 3,5-disubstituted benzoic acids have been synthesized containing all possible combinations of the following substituents: CH_3O, CH_3, Cl/Br, NO_2. The dissociation constants of these acids have been measured in seven organic solvents (methanol, acetone, dimethyl sulfoxide, dimethylformamide, acetonitrile, pyridine, 1,2-dichloroethane). It has been found that the effect of disubstitution is smaller than that due to interaction of substituents or their solvation and represents only about 0.2% of the effect caused by the individual substituents. The additivity in 3,5-disubstitution is about 2 – 3 times as good as that in 3,4-disubstitution. The quantitative description of substituent effects at the 3 and 5 positions is additive within the range of validity of the Hammett equation irrespective of the type of the substituent constants adopted, the addition of the multiplicative term being statistically insignificant. The solvent effect on 3,4- and 3,5-disubstituted derivatives is somewhat different at the same value of the reaction constant, due predominantly to the specific solvation of the 4-CH_3O and 4-NO_2 groups.

The previous communication[1] of this series dealt with evaluation of the problem of disubstitution effect, involving both the literature search with theoretical analysis of the problem and verification of validity of the disubstitution model suggested, applying it to the dissociation of 3,4-disubstituted benzoic acids in organic solvents. The disubstitution effect was found to be additive within the validity range of Hammett equation using various types of substituent constants. The analysis of differences in the Hammett reaction constants for mono- and disubstituted benzoic acids showed that solvation of substituents makes itself felt to various extents in organic solvents. The present communication represents a continuation to the previous one, and its aim is to evaluate the effect of disubstitution in 3,5-disubstituted benzoic acids using the same methods as those used for the 3,4-disubstituted derivatives and adopting the comparison with them.

EXPERIMENTAL

The disubstitution effects were studied with the use of 3,5-disubstituted benzoic acids containing all combinations of four substituents: CH_3O, CH_3, Cl/Br, NO_2. The respective acids were synthesized in known ways or obtained from the sample collection of the Department or were commercial chemicals. The purification methods involved reprecipitation by acidification of the respective salt solutions and further procedures given in Table I. The solvents were purified and dried by the procedures described elsewhere[1], and the same is also true of the methods of measurements of dissociation constants. The experimental results were treated with the help of a personal computer using own programs.

RESULTS AND DISCUSSION

The average values of $\overline{\mathrm{p}K}$ of 3,5-disubstituted benzoic acids in methanol, acetone, dimethyl sulfoxide, dimethylformamide, acetonitrile, pyridine, and 1,2-dichloroethane are given in Table II. The results given can be analyzed in various ways; in order to be able to compare the disubstitution effects in the 3,4- and 3,5-disubstituted benzoic acids we adopted the same method as that in ref.[1]. Also all the matrices analyzed were created in the same structure, and therefore the results of Table II for disubstituted derivatives with different substituents were used twice.

TABLE I

Final methods of purification and physical properties of 3,5-disubstituted benzoic acids

Position		Method of purification[a]	M.p., °C	M.p.[b], °C
3	5			
H	H	e	121 – 122	122
CH_3	CH_3	e	170 – 171	170
CH_3	OCH_3	a	129 – 130	134[c]
CH_3	Cl	e	177 – 179	178
CH_3	NO_2	w	173 – 174	174
OCH_3	OCH_3	a, s	185 – 186	185 – 186
OCH_3	Br	e, a	188 – 189	190 – 191
OCH_3	NO_2	a	193 – 194[d]	–
Br	Br	a, t	217 – 219	219
Cl	NO_2	a	145 – 147	147
NO_2	NO_2	a	206 – 207	206 – 207

[a] Crystallization (e ethanol, a acetic acid, t toluene, w water), s sublimation; [b] ref.[2] if not otherwise stated; [c] ref.[3]; [d] elemental analysis, calculated: 48.73% C, 3.55% H, 7.11% N; found: 48.86% C, 3.80% H, 7.08% N.

Analysis of Effect of Disubstitution by Means of Analysis of Variance

The data of Table II represent a set of numbers obtained as average values of three measurements. With respect to the standard state, the result of each measurement has the variability composed of the solvent effect, 3-substituent effect, 5-substituent effect, effect of possible interaction between the factors given, and experimental error. Obviously, the model at hand is one of analysis of variance with interactions which gave

TABLE II

Average dissociation constants $\overline{\mathrm{p}K}$ and their standard deviations s for 3,5-disubstituted benzoic acids in methanol (MeOH), acetone (Ac), dimethyl sulfoxide (DMSO), dimethylformamide (DMF), acetonitrile (AN), pyridine (Py), and 1,2-dichloroethane (DCE)

Position		$\overline{\mathrm{p}K}/s$						
3	5	MeOH	Ac	DMSO	DMF	AN	Py	DCE
H	H	9.41	18.20	11.00	12.27	20.70	9.80	20.00
CH_3	CH_3	9.59	18.54	11.29	12.57	20.95	10.57	20.38
		0.01	0.02	0.03	0.04	0.07	0.03	0.04
CH_3	OCH_3	9.44	18.26	10.93	12.32	20.66	10.26	20.11
		0.02	0.07	0.03	0.05	0.03	0.09	0.10
CH_3	Cl	8.94	17.40	10.04	11.35	19.85	9.55	19.58
		0.01	0.08	0.08	0.08	0.05	0.06	0.03
CH_3	NO_2	8.55	16.62	9.36	10.92	19.18	8.86	18.96
		0.02	0.04	0.07	0.09	0.01	0.02	0.02
OCH_3	OCH_3	9.27	18.02	10.71	12.04	20.39	10.03	19.83
		0.01	0.05	0.04	0.03	0.07	0.04	0.09
OCH_3	Br	8.79	17.03	9.79	11.28	19.73	8.97	19.28
		0.01	0.02	0.05	0.02	0.04	0.03	0.04
OCH_3	NO_2	8.40	16.31	9.01	10.37	18.72	8.54	18.74
		0.02	0.06	0.09	0.05	0.04	0.03	0.09
Br	Br	8.29	16.23	8.88	10.11	18.86	8.59	18.96
		0.01	0.09	0.03	0.01	0.04	0.03	0.09
Cl	NO_2	8.03	15.57	8.37	9.60	17.98	8.08	18.16
		0.02	0.02	0.07	0.04	0.09	0.10	0.05
NO_2	NO_2	7.64	14.87	7.67	8.76	17.36	7.30	17.51
		0.00	0.09	0.07	0.01	0.03	0.09	0.10

the variability decomposition given in Table III. As expected, statistically significant are both the basic factors and the interaction terms. This result means that the effects followed, i.e. the effect of disubstitution too, are demonstrable in comparison with the experimental accuracy, but it says nothing about the significance of disubstitution with respect to the precision of the Hammett equation. Nevertheless, the information obtained is important because (inter alia) it allows to form an idea about the share of the respective factor or interaction of factors in the result of experiment. In comparison with the results of analysis of variance for the 3,4-disubstituted derivatives it can be stated that the share of basic factors in the overall variability is roughly comparable ($S_{3,4}$: solvent 6 922.7, substitution from *meta*-position 114.0, from *para*-position 175.8, overall variability 7 225.1, ref.[1]). The mutual interactions between substituents contribute to the total variability less than those between solvent and substituent by a factor of about 20. From the comparison with the sevenfold lower contribution for the 3,4-disubstitution it follows that, with roughly the same level of interaction between solvent and substituent in the substitutions of both types, the additivity is more conclusive for the 3,5-disubstitution. Similarly, also the interaction between 3- and 5-substituents represents only about 0.2% (cf. 0.5% with the 3,4-disubstituted derivatives). Irrespective of the disubstitution type, however, it is obvious that the error caused by neglecting the interaction is lower than or at most comparable with that following from the approximate validity and parametrization of the correlation relationships used. The accuracy of the experiment carried out by the method described[1] is relatively constant, being equal

TABLE III
The monitored factors (solvent Sol, *meta*-substitution P3, *para*-substitution P5), sums of squares S, degrees of freedom ν, values of F criterion, and critical values of Fisher–Snedecor distribution F_{crit} at significance level $\alpha = 0.05$ in model of analysis of variance with interactions

Factors	S	ν	F	F_{crit}
Sol	6 821.0	6	$3.80 . 10^5$	2.14
P3	135.4	3	$1.51 . 10^4$	2.64
P5	135.4	3	$1.51 . 10^4$	2.64
Sol + P3	5.4	18	$9.96 . 10^1$	1.65
Sol + P5	5.4	18	$9.96 . 10^1$	1.65
P3 + P5	0.3	9	$1.12 . 10^1$	1.92
Sol + P3 + P5	1.5	54	9.49	1.39
Residual	0.670	224	–	–
Total	7 105.0	335	–	–

to 5.55 . 10^{-2} pK unit for the set analyzed (cf. 5.39 . 10^{-2} pK units for 3,4-disubstituted derivatives).

Analysis of Disubstitution Effect by Means of Additive–Multiplicative Relationships

Our previous communication[1] presented, for evaluation of effect of disubstitution, a relation based on interaction of both substituents which was further simplified approximately to give Eqs (*1*) and (*2*),

$$\Delta G = \Delta G^0 + \alpha_A\sigma_A + \alpha_B\sigma_B + \alpha_{AB}\sigma_A\sigma_B \qquad (1)$$

$$\Delta G = \Delta G^0 + \alpha_A\sigma_A + \alpha_B\sigma_B \; , \qquad (2)$$

where ΔG means the change in Gibbs energy of a process due to substitution, ΔG^0 is the same quantity for the standard substituent (most frequently hydrogen), σ_A and σ_B are parameters describing the substituent effects, and α_A and α_B represent the measure of sensitivity of the ΔG quantity to substitution. Obviously, with regard to the 3- and 5-positions being nondistinguishable, it is $\alpha_A = \alpha_B$, and Eqs (*1*) and (*2*) will simplify to Eqs (*3*) and (*4*) (Hammett equation), respectively:

$$\Delta G = \Delta G^0 + \rho(\sigma_A + \sigma_B) + \alpha_{AB}\sigma_A\sigma_B \qquad (3)$$

$$\Delta G = \Delta G^0 + \rho(\sigma_A + \sigma_B) \; . \qquad (4)$$

Like in the previous communication, the data of Table II were submitted to a treatment by the method of conjugated deviations[1,4]. In the first case we looked for the latent variables obeying Eq. (*3*) and in the second case those obeying Eq. (*4*), simultaneously optimizing the inner values of σ from *meta*-position for the four substituents used. Equation (*3*) could explain 99.10% of the variability of data, and the summary residual deviation[4] expressed as the square root of the quotient of residual sum of squares of non-standardized data and the respective number of degrees of freedom was $s = 1.22 \,.\, 10^{-1}$. Equation (*4*) could explain 98.52% of the variability of data, which corresponds to the summary residual deviation $s = 1.33 \,.\, 10^{-1}$. The test of the hypothesis of equality of residual variances of both correlation relationships gave the value of criterion

$F(84.60) = 1.21$. From the comparison with the critical value of $F_{0.975} = 1.62$ it can be stated that the hypothesis is not rejected at the significance level of $\alpha = 0.05$. Besides the test given, also the test of significance of increment (the addition of multiplicative term in the present case) in the regression can be adopted. The criterion of this test has the value of $F(24.60) = 1.65$, the critical value $F_{0.95} = 1.69$, hence the addition of an additional explaining variable is at the significance limit. In conclusion of this paragraph it can be stated (like in ref.[1] for 3,4-disubstitution) that mutual interactions of substituents cannot statistically be demonstrated, and the substituents effects are additive.

Analysis of Disubstitution Effect in Hammett Equation

The evaluation of additivity or nonadditivity of substituent effects in the sense of Eq. (*4*) depends on the quantitative description of effects of the individual substituents. The initial step consists in expressing the term ($\sigma_A + \sigma_B$) by the first latent variable of the matrix given in Table II. This latent variable involves implicitly the effect of disubstitution in terms of the analyzed set of solvents, and after applying the method of conjugated deviation to the standardized data it described 99.07% of their variability, and the summary residual deviation was $s_{3,5} = 0.10$. The result is comparable[1] with both the analogous calculation for 3,4-disubstituted benzoic acids ($s_{3,4} = 0.09$) and monosubstituted benzoic acids ($s_3 = 0.10$, $s_4 = 0.11$). The value of summary residual deviation obtained by adopting the “internal” substituent constants was given above and it is – according to expectation – somewhat greater ($s_{3,5} = 0.13$). Another step of approximation is represented by adopting the “external” parameters for description of substituent effects. For this purpose we chose the model of the Hammett equation (*4*) with the parametrizations σ^i (ref.[5]) and σ_{Ex} (ref.[6]). The calculation by Eq. (*4*) using σ^i gave the residual standard deviation $s^i_{3,5} = 0.13$ (98.65% explained variability, $s^i_{3,4} = 0.10$, ref.[1]), that using σ_{Ex} gave $s_{3,5Ex} = 0.13$ (98.49% explained variability, $s_{3,4Ex} = 0.13$, ref.[1]). The introduction of general external substituent constants results in the expected somewhat less close fit of correlation, but the differences are not much significant, not even between the individual scales of substituent constants and between the individual substitution models. Hence, again it can be stated that the substituent effects are additive according to Eq. (*4*). An application of more complex models such as (*3*) is of no use in this chemical context.

Analysis of Solvent Effects

Table IV presents the reaction constants and the respective statistical characteristics in the Hammett equation for dissociation constants of 16 disubstituted benzoic acids in 7 organic solvents. When comparing these results with the reaction constants for monosubstituted benzoic acids ($\rho_{3,4}$: methanol 1.47, acetone 2.29, dimethylformamide 2.27, acetonitrile 2.05, ref.[7]) and 3,4-disubstituted benzoic acids ($\rho_{3,4}$: methanol 1.25, acetone 2.20, dimethyl sulfoxide 2.15, dimethylformamide 2.09, acetonitrile 2.10, pyridine 1.59, 1,2-dichloroethane 1.70, ref.[1]) we can observe certain differences which, however, could be connected with the parametrization adopted. This effect, on the other hand, does not influence the evaluation based on the regression of experimental values such as $pK_{3,4}$ and $pK_{3,5}$ in the same solvent. The regression coefficients given as $\rho_{3,4}/\rho_{3,5}$ in Table IV for the individual solvents show distinct deviations from the slope one, the deviation being downward for the measurements in markedly basic solvents and upward for those in acid solvents and also in basic pyridine solvent. Therefrom it is obvious that the reason lies in different specific solvation of the conjugated base since the proton is solvated in the same way in both the cases. The differences in slopes are certainly caused by the extreme points of the dependence, i.e. by substituents with extreme properties located either in 4 or in 5 position. In this respect, special attention should be paid to the methoxy group, which is electron-releasing at the 4 position and electron-withdrawing at the 3(5) position. However, we cannot explain the way in which these differences are projected into the solvent effects. The second alternative is the explanation by an artefact due to a lower accuracy of the measurements in pyridine, but the standard deviations given in Table II do not support this explanation.

The solvent effects on 3,4- and 3,5-disubstituted benzoic acids can be compared summarily by the PLS method[8]. This method is a generalized parallel to linear regression and represents a description of relation of two (or more) matrices by means of

TABLE IV

Reaction constants ρ, their standard deviations s_ρ, residual standard deviations s, and correlation coefficients r in Hammett equation with parametrizations $\Sigma\sigma^i$ (ref.[5]) and $\Sigma\sigma_{Ex}$ (ref.[6]) for dissociation constants of 3,5-disubstituted benzoic acids (repeated measurements, n = 48) in organic solvents (for symbols see Table II)

Solvent	Parametrization $\Sigma\sigma^i$				Parametrization $\Sigma\sigma_{Ex}$				$\frac{\rho_{3,4}}{\rho_{3,5}}$
	ρ	s_ρ	$s \cdot 10^2$	r	ρ	s_ρ	$s \cdot 10^2$	r	
MeOH	1.27	0.02	6.39	0.993	1.31	0.02	6.89	0.992	1.09
Ac	2.42	0.04	10.8	0.995	2.50	0.04	11.5	0.994	1.03
DMSO	2.35	0.04	12.5	0.993	2.42	0.05	13.2	0.992	1.03
DMF	2.46	0.06	16.4	0.989	2.53	0.06	17.0	0.988	0.96
AN	2.38	0.04	10.9	0.995	2.46	0.04	10.6	0.995	0.99
Py	2.05	0.05	13.6	0.989	2.11	0.05	14.4	0.987	0.84
DCE	1.79	0.03	7.94	0.995	1.85	0.03	7.44	0.996	1.07

their internal characteristics, the so-called latent variables. Like in a linear regression, one matrix represents the dependent variable and its variability is interpreted by the second matrix of explaining variables. In our case we can apply the PLS method to the problem of description of one type of disubstitution by means of another type of disubstitution. In a more logical variant, the matrix formed by the data of Table II is the matrix of explaining variables, and an analogous matrix of data of the 3,4-disubstituted derivatives[1] represents the matrix of the variables being explained. In the calculation by PLS method in this case the first latent variable will explain 99.07% of variability in the matrix of 3,5-disubstituted derivatives and 93.45% of variability in that of 3,4-disubstituted derivatives. An addition of further latent variables does not lead to additional explanation of further valid variability (the 2nd latent variable 93.91%, the 3rd latent variable 94.16%). The dependence of the latent variable u_1 (the matrix being explained) upon the latent variable t_1 (the explaining matrix) has – in accordance with the above-mentioned results – a slope of 0.95 ± 0.06, the most deviating combinations from the regression straight line being 3-CH_3–4(5)-OCH_3, 3-halogen–4(5)-OCH_3, 3-NO_2–4(5)-OCH_3, and 3-OCH_3–4(5)-NO_2, i.e. the combinations with the greatest differences in solvation of the two substitution models compared. The slope value given is not statistically significantly different from one, which means that the reaction constant in model (*4*) is the same for both the substitution types.

REFERENCES

1. Pytela O., Kulhanek J., Ludwig M., Riha V.: Collect. Czech. Chem. Commun. *59*, 627 (1994).
2. Buckingham J. et al.: *Dictionary of Organic Compounds*. Chapman and Hall, New York 1982.
3. Meldrum A. N., Perkin W. H.: J. Chem. Soc. *95*, 1889 (1909).
4. Pytela O.: Collect. Czech. Chem. Commun. *55*, 42 (1990).
5. Pytela O.: Collect. Czech. Chem. Commun. *59*, 381 (1994).
6. Exner O.: *Correlation Analysis of Chemical Data*. Plenum, New York 1988.
7. Ludwig M., Baron V., Kalfus K., Pytela O., Vecera M.: Collect. Czech. Chem. Commun. *51*, 2135 (1986).
8. Wold S., Geladi P., Esbensen K., Ohman J.: J. Chemometrics *1*, 41 (1987).

Translated by J. Panchartek.