CORRELATION OF SUBSTITUENT - INDUCED CHEMICAL SHIFTS OF AROMATIC PROTONS: SUBSTITUTED BENZONITRILES AND METHYL BENZOATES

Otto EXNER and Miloš BUDĚŠÍNSKÝ*
Institute of Organic Chemistry and Biochemistry,
Czechoslovak Academy of Sciences, 166 10 Prague 6

Received August 22, 1990
Accepted September 27, 1990

Dedicated to Dr Miroslav Protiva on the occassion of his 70th birthday.

Proton NMR spectra in deuteriochloroform are reported for 53 *meta*- and *para*-substituted benzonitriles and 61 methyl benzoates. The substituent-induced chemical shifts (SCS) were correlated with dual substituent parameters (DSP), with ^{13}C SCS of the adjoining carbon, and with the other ^{1}H SCS using the principal component analysis (PCA). They are controlled by different factors in each position to the variable substituent. In the position 4 the long-range polar effects, as expressed by DSP, are decisive, the conjugative component being more important. Remarkable is also very close correlation of ^{13}C and ^{1}H SCS in this position which has no analogy in the other positions. In the position 3 there is an additional specific effect of heavier halogens. In the position 2 the short-range effect can be quantified by special constants, only for a subgroup of simple donor substituents it is proportional to DSP. In all positions SCS are practically independent of the second substituent, so that the total chemical shifts can be calculated by an additive scheme. The coupling constants $J(2, 3)$, $J(5, 6)$, and $J(2, 6)$ depend on substitution in the same manner, probably are controlled mainly by substituent electronegativity.

In the discussions of substituent chemical shifts (SCS) their correlation to reactivity parameters (substituent constant σ) is of standing interest[1-4]. Most popular of the latter are the dual substituent parameters[5] (DSP), σ_I and σ_R, expressing the inductive and resonance effect, respectively, see Eq. (*1*). Such correlations may serve to several purposes: better understanding of SCS and their relation to electron density, prediction of unknown SCS values, or determining new substituent parameters. On the example of ^{13}C SCS in *meta*- and *para*-substituted benzene derivatives we have recently shown[1,2] that the DSP treatment is not the most efficient procedure: it was both imprecise and overparametrized in the case of *meta* derivatives, at the same time not sufficiently precise in the case of *para* derivatives. Better correlations were obtained with principal component analysis (PCA) which does not refer to any predetermined parameters but derives new parameters from the SCS themselves, this means the components $b, c, \ldots$, loadings $\beta, \gamma, \ldots$ and the location parameter α are optimized in order to fit Eq. (2). Equations (*1*) and (*2*) can be compared with the simplest expression of structural dependence, the additional relationship, Eq. (*3*) which is in use as a chemical tool for assignment of signals in polysubstituted benzene derivatives[6].

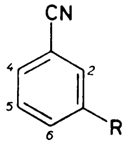
$$\delta = \delta^0 + \varrho_I\sigma_I + \varrho_R\sigma_R \tag{1}$$

$$\delta = \alpha_i + \beta_i b_j + \gamma_i c_j + \ldots \tag{2}$$

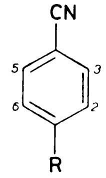
$$\delta = \delta^0 + \sum x_j \tag{3}$$

Equation (*2*) would reduce to Eq. (*3*) if the loadings β_i and γ_i were all equal to unity. Both Eq. (*2*) and (*3*) have in common that the substituent parameters b_j, c_j or x_j are not known a priori but determined only to fit the experimental data: they lack any immediate physical meaning.

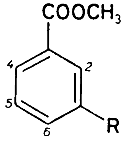
In the mentioned studies[1-3] of ^{13}C SCS we had to use larger and more systematic series of compounds than usual, including substituents of different kind. With these series of compounds in hand we are now able to reinvestigate even the ^1H shifts of aromatic protons and to check their correlations on a broader and more homogeneous experimental material. In the present article we report the ^1H shifts of compounds *I*–*IV*. (The position of a given hydrogen atom is denoted by numerals 1 to 6, related to the variable substituent X; the notations *meta* and *para* are reserved for the mutual position of substituents in a series).



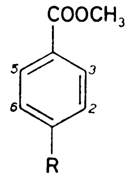
I



II



III



IV

The chemical shifts of aromatic protons in benzene derivatives have been correlated with σ constants since the early development of the NMR spectroscopy, often using unresolved spectra and little precise data. A review from 1978 summarizes the results as follows[4]. Only in the position 4 the SCS are controlled by the electron density and correlate well with DSP, the prevailing contribution comes from the resonance effect. In the position 3 the correlation is worse and there is no agreement[4] whether it can be improved by corrections calculated[7] for magnetic anisotropy and ring current. In the position 2 the third empirical term appears necessary, the pertinent constants Q have been defined just for this particular purpose[8,9]. If SCS are partitioned into four terms according to Eq. (*4*), it is the diamagnetic term δ_d which is connected with the electron density[4]. The long-range term δ_{lr}, associated with electron circulation around the other nuclei, is the main source of deviations. When δ_{lr} is divided into a localized and a delocalized component, the former may be expressed in terms of magnetic anisotropy while the latter is in the case of aromatic compounds due mainly to the ring current. The term with Q constants may account partly for the paramagnetic term δ_p (which is generally of little importance for ^1H) but is certainly more complex in character[4]. The medium term δ_m can be hardly predicted, it can be only reduced by working in the same solvent under identical condition.

$$\delta = \delta_d + \delta_p + \delta_{lr} + \delta_m \qquad (4)$$

The $J(H, H)$ coupling constants in aromatic derivatives have received far less attention in the correlation analysis[10]. Most successful was probably simple additive treatment[11], while correlations with σ constants were rejected as physically impossible[4]. Correlations with substituent electronegativity[10] exhibit large scatter.

EXPERIMENTAL

Substituted benzonitriles and methyl benzoates were characterized previously[1,2].

Proton NMR spectra (at 200 MHz) were measured on a Varian XL-200 FT NMR spectrometer in deuterochloroform with tetramethylsilane as internal reference. To minimize the possible association and temperature effects we measured dilute solutions ($\leq$ 5 mg of substance in 0·4 ml of solution) at the controlled temperature of 30°C. The following typical experimental parameters were used: spectral width 3 kHz, pulse width 4 μs (flip angle 45°), acquisition time 5 s, zero filling to 32 K data points, 64—256 transients. For resolution enhancement a gaussian apodization (with parameters $RE = 0·3$ and $AF = 0·9$) was applied. Standard automatic line listing procedure was used for obtaining the frequencies and intensities of lines. Possible long-range couplings of aromatic protons with substituent protons were eliminated by selective decoupling before the analyses of spectra. The spectra were analyzed using simulation–iteration procedure (SPIN program based on LAOCOON II — standard part of Varian's software version H1Z) to reach the best agreement between experimental and calculated spectrum. The parent monosubstituted compounds — benzonitrile and methyl benzoate — were analyzed as five-spin AA'BB'C systems, the disubstituted derivatives as four-spin ABCD systems (*meta*-disubstituted) or AA'BB' system (*para*-disubstituted). Exception were compounds with two equal substituents giving AB_2C (*meta*) or A_4 (*para*) spectra, and compounds with fluorinated substituents which form more complex spin systems. Chemical shifts and coupling constants obtained are given in Tables I—IV, SCS (relative with respect to the unsubstituted compound are in Tables V and VI). In some cases the simulation–iteration procedure could not be applied and certain parameters could not be determined due to the complications arising from *a*): strongly coupled spin system (suitable input data not available), *b*) overlap with aromatic protons of the substituent, *c*) long-range couplings with fluorine atoms of the substituent, leading to additional splitting (often not sufficiently resolved), *d*) symmetry of the molecule preventing to observe the coupling constants between magnetically equivalent nuclei. In such cases the parameters were either estimated directly from the spectra with a lower accuracy (within the first order approximation) or they could not be determined at all. These cases are indicated in Tables I—IV. The resulting, less precise values were also included into the correlation calculations since the fit was not impaired. We estimate the precision of the latter SCS to 0·005 ppm while the other SCS should be reliable to 0·002 ppm. The *J*(H, H) coupling constants could be reliable to 0·02 Hz.

TABLE I

Proton NMR parameters of meta-substituted benzonitriles in deuteriochloroform

Substituent	H(2)	H(4)	H(5)	H(6)	*J*(2, 4)	*J*(2, 5)	*J*(2, 6)	*J*(4, 5)	*J*(4, 6)	*J*(5, 6)
H	7·644	7·644	7·472	7·603	1·86	0·64	1·26	7·82	1·26	7·65
CH_2Cl	7·688	7·610	7·488	7·635	1·66	0·58	1·76	7·74	1·19	7·86
CH_2Br	7·687	7·588	7·468	7 631	1 62	0 56	1·87	7·66	1·15	7·95
C_6H_5	7·851	7·615	7·532	7·799	1·59	0·62	1·89	7·70	1·14	7·97
$COCH_3$	8·238	7·860	7·648	8·199	1·59	0·62	1·79	7·70	1·26	7·94
COC_6H_5	8·064	7·865	7·631	8·036	1·63	0·63	1·76	7·69	1·22	8·00
COOH	8·386	7·889	7·625	8·318	1·70	0·62	1·71	7·81	1·22	7·92
$COOCH_3$	8·330	7·841	7·591	8·270	1·57	0·65	1·67	7·78	1·35	7·93
$COOCH_2CH_3$	8·334	7·832	7·583	8·273	1·68	0·64	1·70	7·79	1·23	7·93
COCl	8·410	7·967	7·694	8·342	1·64	0·58	1·87	7·79	1·17	8·03
CN	7·981	7·928	7·683	7·928	1·63	0·65	1·63	7·92	[a]	7·92
Cl	7·630	7·559	7·436	7·590	1·53	0·47	2·14	7·76	1·01	8·26
Br	7·795	7·605	7·365	7·747	1·53	0·49	2·01	7·77	1·06	8·19
I	7·981	7·633	7·220	7·943	1·54	0·48	1·77	7·79	1·09	8·06
OH	7·158	7·217	7·334	7·127	1·36	0·48	2·60	7·66	1·03	8·34
OCH_3	7·118	7·215	7·361	7·126	1·39	0·56	2·63	7·60	0·96	8·56
$OCOCH_3$	7·416	7·517	7·482	7·353	1·54	0·48	2·39	7·77	0·88	8·40
NH_2	6·894	7·002	7·209	6·861	1·55	0·49	2·47	7·62	1·00	8·23
$N(CH_3)_2$	6·861	6·925	7·258	6·870	1·39	0·44	2·71	7·44	0·93	8·61
$NHCOCH_3$	8·000	7·319	7·374	7·797	1·10	0·51	1·76	7·23	1·87	8·18
NO_2	8·541	8·012	7·756	8·487	1·56	0·53	2·26	7·74	1·07	8·38
SO_2CH_3	8·262	7·960	7·758	8·198	1·60	0·61	1·86	7·79	1·14	8·00
SO_2NH_2	8·237	7·826	7·656	8·165	1·52	0·62	1·91	7·77	1·17	7·97
SO_2Cl	8·331	8·027	7·810	8·278	1·55	0·55	1·99	7·78	1·10	8·17

[a] The value of parameter could not be determined.

All linear regressions were carried by standard procedures with a freely fitted intercept. The intercept values are omitted, the remaining characteristics are given in Table VII. The characteristic ψ, was suggested for estimating generally the fit of a theory with experiment[12], is in the case of linear regression it is simply related to the correlation coefficient. It was used here to

TABLE II

Proton NMR parameters of *para*-substituted benzonitriles in deuteriochloroform

Substituent	H(2) H(6)	H(3) H(5)	J(2, 3) J(5, 6)	J(2, 5) J(3, 6)	J(2, 6)	J(3, 5)
H	7·472	7·644	7·82	0·64	1·21	1·86
CH_3	7·249	7·491	8·01	0·64	1·84	1·83
$CH_2C_6H_5$	7·265	7·545	7·99	0·70	1·71	1·68
$CH_2SO_2C_6H_5$	7·224	7·562	8·02	0·58	1·82	1·92
CF_3	7·81[a]	7·77[a]	[b]	[b]	[b]	[b]
CH_2Cl	7·507	7·656	8·01	0·57	1·87	1·80
CH_2Br	7·500	7·633	8·03	0·61	1·87	1·84
C_6H_5	7·67[a]	7·71[a]	[b]	[b]	[b]	[b]
$COCH_3$	8·049	7·781	8·17	0·67	1·79	1·70
COC_6H_5	7·868	7·784	8·04	0·25	1·81	2·62
COOH	8·197	7·778	8·13	0·66	1·74	1·70
$COOCH_3$	8·141	7·750	8·13	0·65	1·73	1·71
$COOCH_2CH_3$	8·146	7·750	8·13	0·65	1·75	1·70
COCl	8·232	7·843	8·20	0·61	1·84	1·79
CN	7·800	7·800	[b]	[b]	[b]	[b]
F	7·19[a]	7·69[a]	[b]	[b]	[b]	[b]
Cl	7·469	7·602	8·35	0·61	2·14	2·07
Br	7·636	7·524	8·40	0·57	2·05	2·07
I	7·848	7·366	8·20	0·43	1·96	1·91
OH	6·896	7·466	8·51	0·38	2·82	1·87
OCH_3	6·939	7·542	8·65	0·42	2·73	2·12
$OCOCH_3$	7·225	7·665	8·32	0·35	2·12	2·34
NH_2	6·644	7·401	8·43	0·44	2·49	1·98
$N(CH_3)_2$	6·632	7·442	8·79	0·36	2·82	2·09
$NHCOCH_3$	7·749	7·550	8·53	0·29	2·13	2·13
NO_2	8·366	7·903	8·53	0·53	2·48	1·70
SCH_3	7·259	7·522	8·32	0·57	1·95	2·00
SO_2CH_3	8·088	7·893	8·15	0·62	1·86	1·74
SO_2NH_2	8·058	7·801	8·18	0·68	1·80	1·76
SO_2Cl	8·174	7·938	8·32	0·61	1·99	1·76

[a] The value of parameter could not be determined by simulation–iteration procedure; [b] the value of parameter could not be determined.

compare the precision of linear regression and PCA on a uniform scale. PCA was carried out as previously[1], the characteristics measuring the fit are given also in Table VII, the components and loadings are given in Table VIII.

TABLE III

Proton NMR parameters of meta-substituted methyl benzoates in deuteriochloroform

Substituent	H(2)	H(4)	H(5)	H(6)	J(2, 4)	J(2, 5)	J(2, 6)	J(4, 5)	J(4, 6)	J(5, 6)
H	8·033	8·033	7·427	7·542	1·84	0·58	1·31	7·86	1·31	7·49
CH_3	7·853	7·827	7·311	7·356	1·76	0·63	1·61	6·97	1·30	7·56
$CH_2SO_2C_6H_5$	7·714	7·993	7·360	7·330	1·77	0·65	1·97	7·82	1·24	7·72
CF_3	8·30[a]	8·22[a]	7·58[a]	7·81[a]	[b]	[b]	[b]	[b]	[b]	[b]
CH_2Cl	8·060	7·992	7·441	7·586	1·77	0·58	1·88	7·82	1·15	7·67
CH_2Br	8·058	7·964	7·419	7·579	1·72	0·57	1·92	7·82	1·18	7·68
C_6H_5	8·276	8·011	7·504	7·777	1·78	0·59	1·89	7·84	1·16	7·73
$COCH_3$	8·598	8·237	7·561	8·160	1·71	0·60	1·81	7·76	1·24	7·79
COC_6H_5	8·432	8·255	7·578	7·997	1·74	0·63	1·78	7·73	1·24	7·72
COOH	8·676	8·288	7·577	8·296	1·78	0·59	1·77	7·70	1·26	7·88
$COOCH_3$	8·681	8·221	7·526	8·221	1·76	0·58	1·76	7·78	[b]	7·78
CN	8·330	8·270	7·591	7·841	1·67	0·65	1·57	7·93	1·35	7·78
F	7·71[a]	7·83[a]	7·41[a]	7·25[a]	[b]	[b]	[b]	[b]	[b]	[b]
Cl	8·027	7·929	7·386	7·534	1·61	0·47	2·19	7·83	1·08	8·04
Br	8·187	7·976	7·325	7·689	1·59	0·47	2·06	7·82	1·08	8·01
I	8·373	7·994	7·175	7·876	1·62	0·30	1·87	7·87	1·08	7·84
OH	7·614	7·600	7·307	7·090	1·53	0·44	2·65	7·71	1·05	8·15
OCH_3	7·568	7·639	7·342	7·103	1·54	0·45	2·73	7·68	1·01	8·28
OCH_2CH_3	7·557	7·622	7·329	7·090	1·59	0·42	2·66	7·68	1·02	8·26
$OCOCH_3$	7·756	7·907	7·442	7·286	1·57	≈0	2·42	7·82	1·06	8·10
NH_2	7·353	7·421	7·202	6·849	1·57	0·41	2·49	7·70	1·02	7·99
$N(CH_3)_2$	7·40[a]	7·37[a]	7·27[a]	6·90[a]	[b]	[b]	[b]	[b]	[b]	[b]
NO_2	8·855	8·362	7·655	8·409	1·62	0·51	2·34	7·74	1·12	8·23
NCS	7·879	7·931	7·428	7·387	1·63	0·53	2·15	7·85	1·13	7·98
$N=NC_6H_5$	8·577	8·156	7·597	8·109	1·69	0·47	2·00	7·68	1·16	7·96
SCH_3	7·919	7·798	7·348	7·436	1·58	0·51	2·07	7·74	1·11	7·90
SO_2CH_3	8·597	8·317	7·677	8·137	1·63	0·60	1·93	7·81	1·17	7·84
SO_2NH_2	8·577	8·175	7·570	8·109	1·62	0·57	1·95	7·83	1·17	7·85
SO_2F	8·67[a]	8·43[a]	7·74[a]	8·19[a]	[b]	[b]	[b]	[b]	[b]	[b]
SO_2Cl	8·689	8·400	7·730	8·219	1·58	0·55	2·06	7·82	1·16	8·02

[a] The value of parameter could not be determined by simulation–iteration procedure; [b] the value of parameter could not be determined.

DISCUSSION

The following discussion is based on all available values: deviations from the correlations are mentioned in every particular case. As previously[1] the substituents

TABLE IV

Proton NMR parameters of *para*-substituted methyl benzoates in deuteriochloroform

Substituent	H(2) H(6)	H(3) H(5)	J(2, 3) J(5, 6)	J(2, 5) J(3, 6)	J(2, 6)	J(3, 5)
H	7·427	8·033	7·86	0·58	1·33	1·84
CH_3	7·223	7·919	7·99	0·59	1·88	1·93
$CH_2C_6H_5$	7·240	7·944	8·03	0·59	1·89	1·92
$CH_2SO_2C_6H_5$	7·155	7·920	8·07	0·59	1·86	1·95
CF_3	7·70[a]	8·15[a]	[b]	[b]	[b]	[b]
CH_2Cl	7·451	8·022	8·04	0·55	1·76	2·09
CH_2Br	7·450	8·003	8·07	0·58	1·89	1·94
C_6H_5	7·653	8·099	8·19	0·59	1·92	1·98
$COCH_3$	7·997	8·116	8·11	0·55	1·81	1·82
COC_6H_5	7·833	8·142	8·13	0·61	1·73	1·89
COOH	8·167	8·133	8·04	0·55	1·79	1·90
$COOCH_3$	8·084	8·084	[b]	[b]	[b]	[b]
CN	7·750	8·141	8·13	0·65	1·71	1·73
F	7·10[a]	8·05[a]	[b]	[b]	[b]	[b]
Cl	7·406	7·966	8·43	0·48	2·13	2·21
Br	7·573	7·891	8·46	0·31	2·09	2·09
I	7·739	7·797	8·24	0·48	2·01	2·01
OH	6·857	7·951	8·54	0·37	2·20	2·69
OCH_3	6·909	7·986	8·69	0·34	2·20	2·74
OCH_2CH_3	6·890	7·970	8·70	0·32	2·15	2·77
$OCOCH_3$	7·164	8·061	8·36	0·22	2·33	2·33
NH_2	6·632	7·844	8·44	0·39	2·06	2·49
$N(CH_3)_2$	6·641	7·898	8·84	0·33	2·12	2·88
$NHCOCH_3$	7·59[a]	7·99[a]	[b]	[b]	[b]	[b]
NO_2	8·282	8·208	8·63	0·59	2·11	2·11
NCS	7·261	8·024	8·39	0·51	2·04	2·15
$N{=}NC_6H_5$	7·94[a]	8·19[a]	[b]	[b]	[b]	[b]
SCH_3	7·245	7·928	8·36	0·51	2·00	2·05
SO_2CH_3	8·022	8·223	8·26	0·63	1·83	1·86
SO_2NH_2	7·995	8·139	8·28	0·64	1·88	1·88
SO_2F	8·09[a]	8·28[a]	[b]	[b]	[b]	[b]
SO_2Cl	8·117	8·267	8·44	0·59	1·79	2·06

[a] The value of parameter could not be determined by simulation-iteration procedure; [b] the value of parameter could not be determined.

$OCOCH_3$ and COCl, in addition the substituent NCS were excluded from the correlations with DSP since the values of the latter are uncertain. Nevertheless, these substituents were included into correlations not dependent on DSP and were plotted in the graphs. Previously, also the substituent 3-OH was excluded in the case of

TABLE V

Substituent effects in proton NMR spectra of *meta*- and *para*-substituted benzonitriles

Substituent	*Meta*-substituted			*Para*-substituted	
	H(2); H(6)	H(5)	H(4)	H(2); H(6)	H(3); H(5)
CH_3	[a]; [a]	[a]	[a]	−0·223	−0·153
$CH_2C_6H_5$	[a]; [a]	[a]	[a]	−0·207	−0·099
$CH_2SO_2C_5H_6$	[a]; [a]	[a]	[a]	−0·248	−0·082
CF_3	[a]; [a]	[a]	[a]	0·34[b]	0·13[b]
CH_2Cl	0·044; 0·032	0·016	−0·034	0·035	0·012
CH_2Br	0·043; 0·028	−0·003	−0·056	0·028	−0·011
C_6H_5	0·207; 0·196	0·060	−0·029	0·20[b]	0·07[b]
$COCH_3$	0·594; 0·596	0·176	0·216	0·577	0·137
COC_6H_5	0·420; 0·433	0·159	0·221	0·396	0·140
COOH	0·742; 0·715	0·153	0·245	0·725	0·106
$COOCH_3$	0·686; 0·667	0·119	0·197	0·669	0·106
$COOCH_2CH_3$	0·690; 0·670	0·111	0·188	0·674	0·106
COCl	0·766; 0·739	0·222	0·323	0·760	0·199
CN	0·337; 0·325	0·211	0·284	0·328	0·156
F	[a]; [a]	[a]	[a]	−0·29[b]	0·05[b]
Cl	−0·014; −0·013	−0·036	−0·085	−0·003	−0·042
Br	0·151; 0·144	−0·107	−0·039	0·164	−0·120
I	0·337; 0·340	−0·252	−0·011	0·376	−0·278
OH	−0·486; −0·476	−0·138	−0·427	−0·576	−0·178
OCH_3	−0·526; −0·477	−0·111	−0·429	−0·533	−0·102
$OCOCH_3$	−0·228; −0·250	0·010	−0·127	−0·247	0·021
NH_2	−0·750; −0·742	−0·263	−0·642	−0·828	−0·243
$N(CH_3)_2$	−0·783; −0·733	−0·214	−0·719	−0·840	−0·202
$NHCOCH_3$	0·356; 0·194	−0·098	−0·325	0·277	−0·094
NO_2	0·897; 0·084	0·284	0·368	0·894	0·259
SCH_3	[a]; [a]	[a]	[a]	−0·213	−0·122
SO_2CH_3	0·618; 0·595	0·286	0·316	0·616	0·249
SO_2NH_2	0·593; 0·562	0·184	0·182	0·586	0·157
SO_2Cl	0·687; 0·675	0·338	0·383	0·702	0·294

[a] The value of parameter could not be determined; [b] the value (rounded off to 0·01 ppm) was determined using chemical shift data obtained without simulation–iteration procedure (with lower accuracy).

methyl benzoates due to dimer formation[2], but no deviations of this substituent was observed in the present case. Note that inclusion or exclusion of the named

TABLE VI

Substituent effects in proton NMR spectra of *meta*- and *para*-substituted methyl benzoates

Substituent	*Meta*-substituted			*Para*-substituted	
	H(2); H(6)	H(5)	H(4)	H(2); H(6)	H(3); H(5)
CH_3	−0·180; −0·186	−0·116	−0·206	−0·204	−0·114
$CH_2C_6H_5$	^{a} ; ^{a}	^{a}	^{a}	−0·187	−0·089
$CH_2SO_2C_6H_5$	−0·319; −0·212	−0·067	−0·040	−0·272	−0·113
CF_3	0·27^b; 0·27^b	0·15^b	0·19^b	0·27^b	0·12^b
CH_2Cl	0·027; 0·044	0·014	−0·041	0·024	−0·011
CH_2Br	0·025; 0·037	−0·008	−0·069	0·023	−0·030
C_6H_5	0·243; 0·235	0·077	−0·022	0·226	0·066
$COCH_3$	0·565; 0·618	0·134	0·204	0·570	0·083
COC_6H_5	0·399; 0·455	0·151	0·222	0·406	0·109
$COOH$	0·734; 0·754	0·150	0·255	0·740	0·100
$COOCH_3$	0·648; 0·679	0·099	0·188	0·657	0·051
CN	0·297; 0·299	0·164	0·237	0·323	0·108
F	−0·32^b; −0·30^b	−0·02^b	−0·21^b	−0·33^b	0·02^b
Cl	−0·006; −0·008	−0·041	−0·104	−0·021	−0·067
Br	0·154; 0·147	−0·102	−0·057	0·146	−0·142
I	0·340; 0·334	−0·252	−0·039	0·312	−0·236
OH	−0·419; −0·452	−0·120	−0·433	−0·570	−0·082
OCH_3	−0·465; −0·439	−0·085	−0·394	−0·518	−0·047
OCH_2CH_3	−0·467; −0·452	−0·098	−0·411	−0·537	−0·063
$OCOCH_3$	−0·277; −0·256	0·015	−0·126	−0·263	0·028
NH_2	−0·680; −0·693	−0·225	−0·612	−0·795	−0·189
$N(CH_3)_2$	−0·63^b; −0·64^b	−0·16^b	−0·66^b	−0·786	−0·135
$NHCOCH_3$	^{a} ^{a}	^{a}	^{a}	0·16^b	−0·04^b
NO_2	0·822; 0·867	0·228	0·329	0·855	0·175
NCS	−0·154; −0·155	0·001	−0·102	−0·166	−0·009
$N{=}NC_6H_5$	0·544; 0·567	0·170	0·123	0·51^b	0·16^b
SCH_3	−0·114; −0·106	−0·079	−0·235	−0·182	−0·105
SO_2CH_3	0·564; 0·595	0·250	0·284	0·595	0·190
SO_2NH_2	0·544; 0·567	0·143	0·142	0·568	0·106
SO_2F	0·63^b; 0·65^b	0·31^b	0·40^b	0·66^b	0·25^b
SO_2Cl	0·656; 0·677	0·303	0·367	0·690	0·234

[a] The value of parameter could not be determined; [b] the value (rounded off to 0·01 ppm) was determined using the chemical shift obtained without simulation–iteration procedure (with lower accuracy).

substituents is immaterial with respect to the global fit. The data matrices for PCA were somewhat more restricted (usually to 23 substituents) in order to get a good filling (see footnotes to Table VII).

Easiest to understand are the ^{1}H SCS in the position 4: according to their correlation with σ_I and σ_R (Table VII) they are controlled essentially by polar effects or say by the electron density. Similar conclusions were drawn from correlations on monosubstituted benzenes[13,15], the fit was similar, the number of compounds somewhat smaller[13] or not given[15]. The correlation with a more sophisticated non-linear variant[16] of DSP, called DSP-NLR, represents only a slight improvement and is not worth of the more complex procedure. The correlations with σ_R^0 instead of σ_R^+ were always worse, confirming tht conjugation is of importance. On the other hand, a significantly better correlation was obtained with PCA, using a small data matrix based on our and literature[13,14] data (Table VII). The loadings in all four series of this data matrix are very close (Table VIII), hence the additive relationship[13,14], Eq. (*3*), can be applied, considering the constant functional group (CN, $COOCH_3$, CH_3) as the second substituent. According to Table VII the additive relationship is only slightly less precise than PCA and could serve well for assignment of signals. The values of the increments x_j are not given, they would be very close to b_4 in Table VIII.

It follows that a small part of substituent effects remains unexplained by DSP. We have searched for such effects, both specific (by analyzing the residuals) and more general (by plotting the components b_4 from Table VIII against various σ constants). We have not found any significant and more common effect. Still most telling is the plot (not shown) of the components b_4 vs the blend of sigma, $\sigma_I + 1{\cdot}35\sigma_R^+$. The largest deviations are caused by strongest donors NH_2 and NMe_2, which are not so strong as predicted, on the other hand the halogens are somewhat stronger. (The latter phenomenon is much more significant in the positions 3 and 2, see below). Some other deviations might appear as fortuitous (Ph, Me, SO_2NH_2) but they are not observed in the correlation of SCS of ^{1}H(4) and ^{13}C(4) in the same molecule (Fig. 1, Table VII). This correlation is very significant and suggests that both ^{1}H and ^{13}C SCS in the position 4 are controlled mainly by polar substituent effects. Nevertheleses, there are also further effects operating as follows from the fact that PCA (components from Table VIII) describe the SCS better than does DSP. Of course, this does not allow decision which constants (DSP or SCS) are better measure of the electron density[15].

The SCS in the position 3 behave also quite regularly: they are very similar in the two series of compounds, in addition also in *meta* and *para* derivatives, see the precise correlation by PCA (Table VII) and close values of loadings β (Table VIII). The preference against DSP seems to be quite dramatic but in this case the cause of deviations can be detected. It is the extraordinary behaviour of the halogen substituents. This effect, not uncovered in previous correlations of monosubstituted

TABLE VII

Correlations of ^{1}H SCS of disubstituted benzene derivatives *I*–*IV*

Method	Compounds	Hydrogen position	Regression coefficients		s^a	R^b	ψ^c
			ϱ_I	ϱ_R			
DSP $(\sigma_R^0)^d$	*I*	2	0.606^e	1.655	0.191 (19)	0.929	0.37
	I	6	0.593^f	1.624	0.170 (19)	0.940	0.34
	II	2, 6	0.753^f	1.789	0.169 (25)	0.948	0.32
	III	2	0.495^e	1.487	0.194 (25)	0.901	0.43
	III	6	0.482^e	1.532	0.195 (25)	0.905	0.42
	IV	2, 6	0.572^f	1.624	0.195 (27)	0.914	0.41
Additivity	^{g}	2 or 6	—	—	0.031 (150)	99.72%	0.058
PCA (1 term)	^{g}	2 or 6	—	—	0.024 (141)	99.83%	0.046
(2 terms)	^{g}	2 or 6	—	—	0.017 (112)	99.93%	0.034
DSP $(\sigma_R^0)^d$	*I*	5^h	0.263^f	0.543	0.076 (19)	0.908	0.42
		5^i	0.364	0.460	0.038 (16)	0.977	0.21
	II	3, 5^h	0.296^e	0.482^f	0.099 (25)	0.826	0.56
		3, 5^i	0.331^f	0.457^f	0.085 (21)	0.879	0.48
	III	5^h	0.264^f	0.443	0.077 (25)	0.880	0.47
		5^i	0.340	0.376	0.048 (21)	0.953	0.30
	IV	3, 5^j	0.229^e	0.324^f	0.082 (27)	0.784	0.62
Additivity	^{k}	3 or 5	—	—	0.022 (98)	98.44%	0.14
PCA (1 term)	^{k}	3 or 5	—	—	0.014 (92)	99.38%	0.092
DSP (δ_R^+)	*I*	4^h	0.355^f	0.462	0.050 (19)	0.989	0.15
		4^i	0.405	0.454	0.041 (16)	0.994	0.11
	III	4	0.352^f	0.435	0.062 (25)	0.979	0.20
		4^i	0.424	0.419	0.049 (21)	0.989	0.15
DSP (σ_R^0)	*I*	4	0.425^f	1.088	0.064 (19)	0.982	0.19
	III	4	0.381	1.024	0.053 (25)	0.984	0.18
DSP-NLR (σ_R^+)	*I*	4	0.334	0.639	0.035 (18)	0.995	0.10
			$\varepsilon = 0.003$				
	III	4^i	0.372	0.565	0.043 (20)	0.992	0.13
			$\varepsilon = 0.261$				
$\delta(^{13}C) - \delta(^1H)$	*I*, *III*	4	0.0576	—	0.036 (48)	0.993	0.12
Additivity	^{l}	4	—	—	0.018 (46)	99.80%	0.056
PCA (1 term)	^{l}	4	—	—	0.014 (43)	99.88%	0.044
PCA (1 term)	*I*–*IV*^{m}	$J_{2,3}$, $J_{5,6}$, $J_{2,6}$	—	—	0.084^m (131)	94.90%	0.25

---

[a] Standard deviation in ppm (degrees of freedom in parentheses); [b] multiple correlation coefficient or % of explained sum of squares; [c] statistic measuring generally the fit[12]; [d] with σ_R^+ the fit was worse in all cases but the difference was mostly not significant statistically; [e] the pertinent partial correlation coefficient less than 0·5; [f] the pertinent partial correlation coefficient less than 0·8; [g] the data matrix consisted of ten columns (the first six lines of this table, in addition

monosubstituted benzenes[13], *meta*-substituted toluenes in the positions 2 and 6, and *para*-substituted toluenes[14]), and 23 substituents (as given in Table VIII) with 182 items it was filled to 79%. As outlier 4-aminobenzonitrile was eliminated but the effect on the overall fit was negligible; [h] when $|\sigma_R^0|$ is added as third explanatory variable, this term is insignificant; [i] without halogens; [j] including halogens, their elimination is of little effect; [k] the data matrix consisted of seven columns (similar as in the footnote *g* and 23 lines (substituents of Table VIII); with 127 items it was filled to 79%, no outliers; [l] the data matrix consisted of four columns (similar as in the footnote *g*) and the same 23 substituents; with 72 items it was filled to 78%, no outliers; [m] the data matrix consisted of eight columns and 23 lines (substituents), with 168 items it was filled to 91%, standard deviation in Hz.

TABLE VIII

Components and loadings of PCA for ^1H SCS in benzene derivatives

Substituent	Component[a,b]				Q^c	Compound (Position)	Loadings[a]	
	b_2(1st)	c_2(2nd)	b_3	b_4			β(1st)	γ(2nd)
H	0	0	0	0	2·58	*I* (2)	0·449	−0·513
CH_3	−0·452	0·029	−0·267	−0·503	1·73	*I* (6)	0·435	−0·394
CH_2Cl	0·090	0·019	0·026	−0·078	2·75	*II* (2, 6)	0·463	−0·424
CH_2Br	0·082	0·010	−0·009	−0·134	2·73	*III* (2)	0·412	−0·356
C_6H_5	0·440	0·008	0·146	−0·073	3·41	*III* (6)	0·425	−0·397
$COCH_3$	1·317	−0·033	0·293	0·448	5·06	*IV* (2, 6)	0·455	−0·139
COC_6H_5	0·885	−0·012	0·292	0·447	4·24	Bz (2)	0·462	0·462
COOH	1·732	0·061	0·326	0·585	5·84	*m*-Tol (2)	0·452	0·331
$COOCH_3$	1·518	−0·011	0·225	0·445	5·43	*m*-Tol (6)	0·452	0·437
COCl	1·726	0·030	0·453	0·787	5·82	*p*-Tol (2, 6)	0·446	0·540
CN	0·720	−0·003	0·386	0·610	3·93			
Cl	−0·015	0·012	−0·089	−0·213	2·55	*I* (5)	0·503	—
Br	0·337	−0·002	−0·222	−0·109	3·21	*II* (3, 5)	0·462	—
I	0·775	0·010	−0·499	−0·038	4·04	*III* (5)	0·430	—
OH	−1·171	−0·069	−0·305	−0·971	0·38	*IV* (3, 5)	0·336	—
OCH_3	−1·118	0·006	−0·219	−0·947	0·48	Bz (3)	0·462	—
$OCOCH_3$	−0·597	−0·009	0·032	−0·293	1·46	*m*-Tol(5)	0·458	—
NH_2	−1·700	−0·006	−0·570	−1·419	−0·62	*p*-Tol(3, 5)	0·422	—
$N(CH_3)_2$	−1·601	0·133	−0·403	−1·498	−0·43			
NO_2	1·983	−0·005	0·541	0·810	6·31	*I* (4)	0·457	—
SO_2CH_3	1·368	0·009	0·569	0·688	5·15	*III* (4)	0·457	—
SO_2NH_2	1·307	0·009	0·349	0·375	5·04	Bz (4)	0·462	—
SO_2Cl	1·564	0·019	0·691	0·877	5·52	*m*-Tol (4)	0·445	—

[a] Components and loadings were normalized: b_4 to the value 0·81 for NO_2, b_2 and b_3 to equal loadings in the benzene series (0·462); the data matrices were described in notes *g*, *k*, *l* in Table VII; [b] subscripts 2—4 correspond to the position of H; [c] the values of b_2 scaled with reference to the *Q* values of ref.[19].

benzenes,[13,15] is seen very clearly from a plot of the components vs the blend $\sigma_I + 1{\cdot}5\sigma_R^0$ (Fig. 2), also from the DSP treatment when halogens are excluded (Table VII). The anomaly seems to be pertinent to heavier atoms with large valence shell, it is most conspicuous for iodine. We cannot accept the explanation[7] of these deviations by magnetic anisotropy of some substituents (F, Cl, Br, I, CN) which were referenced to quite a small number of isotropic substituents (H, Me, OMe, NO_2). Unfortunately we were unable to resolve exactly the comlex spectra of fluoro derivatives but we have definitely not observed any deviation of CN (Fig. 2). It is also difficult to understand why the substituents like COX, SO_2X, NO_2 should behave as isotropic. Also the second suggested[7] correction for the position 3, accounting for the ring current changes, is not confirmed from our data. It should lead to correlation with the absolute value of the resonance constants, $|\sigma_R^0|$, but the respective term was not significant (Table VII). We thus assume that these two corrections are artifacts of the theory which was supported by an insufficient number of data[7]. Our conclusion is that ^1H SCS in the position 3 are controlled by polar effect with an additional phenomenon characteristic for heavier atoms or only for heavier halogens. Remarkable is the large contribution from the mesomeric conjugative effect (σ_R), unusual for the position 3 and unexplained for the present. A plot of $^1H(3)$ vs $^{13}C(3)$ SCS in the same molecule (not shown) is complex but its complexity is mainly due to the behaviour of ^{13}C shifts[3] and will not be discussed here.

Concerning the SCS in the position 2, the main task is to uncover possible ortho effects[17,18] either purely steric or merely short-distance polar effects. An indication

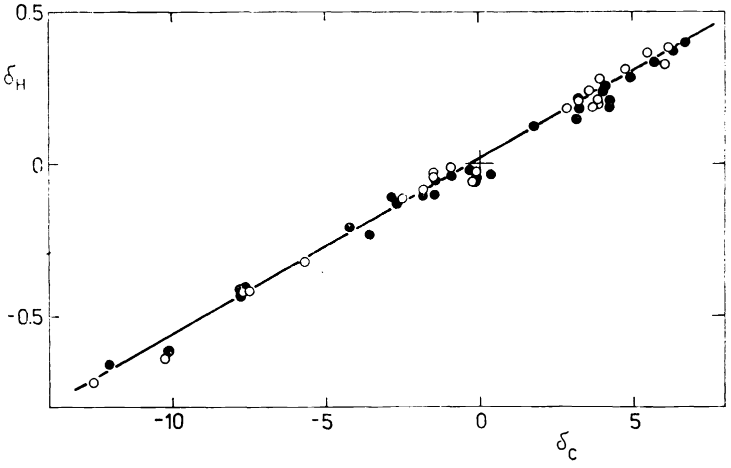


FIG. 1

Comparison of ^1H and ^{13}C SCS on the adjoining atoms of the same molecule; position 4: ○ in *meta*-substituted benzonitriles (*I*) and ● in methyl benzoates (*III*) of the steric effect could be its strenghtening when the hydrogen is placed between two substituents, the so-called sandwich effect[17]. According to our results the latter is quite negligible, see particularly the small differences between the atoms H(2) and H(6) in the derivatives *I* and *III*, further the small differences between the loadings β_2 and β_6 for these two hydrogens (Table VIII). Absence of purely steric effects was deduced formerly from a correlation on monosubstitued benzenes with σ_I and σ_R^{Bz} when an added steric term was insignificant[18]. However, the fit was as bad as in our Table VII and the very large intercept indicated some constant proximity effect of all substituents except hydrogen. Our more detailed analysis revealed first that SCS are practically constants in all series examined, see the excellent fit of PCA and almost constant values of the loadings β_2 and β_6 (Table VIII). The fit itself suggested a possible second term of some significancy but the pertinent components c_2 are very irregular: this term is in fact necessary just to correct some deviations of the substituents NMe_2, OH and COOH. One term is thus sufficient in our opinion, even with this term only the fit is as good as in the case of hydrogen atoms H(3) or H(4). The specific effects were searched for by plotting the components b_2 vs different blends of the DSP: most significant was the combination $\sigma_I + 1\cdot35\sigma_R^+$ (Fig. 3), the same as used for H(4). Proximity effects are evident for the bulky and strongly polar groups COX, SO_2X, and NO_2, further for Ph, heavy halogens and possibly NMe_2, they are absent for OH, OMe, NH_2, CN and substituted methyls. A plot of $^1H(2)$ and $^{13}C(2)$ shifts in the same molecule (Fig. 4) is more complex: the deviations of COX, SO_2X, NO_2 and Ph persist, Br and I are on the line and CN now deviates, indicating some short-range effects even on C(2). Deviations of some substituted methyls, not of all, to the opposite sides are difficult to explain. (Maybe

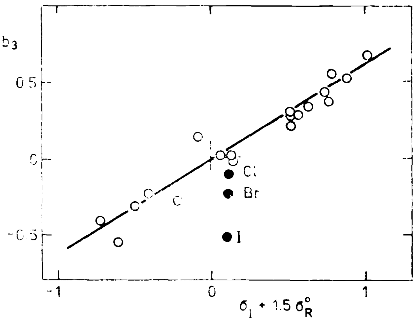


FIG. 2

Plot of the components b_3, expressing the average ^1H SCS in position 3, vs a blend of DSP ($\sigma_I + 1\cdot5\sigma_R^0$); full points belong to heavier halogens

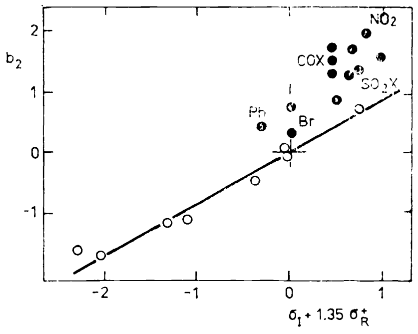


FIG. 3

A similar plot as in Fig. 2 for the position 2 (the component b_2 vs $\sigma_I + 1\cdot35\sigma_R^+$); full points deviate we see here, for the first time, some effect of the ring current in the case of substituents CH_2Ph and CH_2SO_2Ph: these complex substituents were included by us[1-3] just to search for such effects.) The correlation with the constants Q was impaired by the small number of Q values available and by some discrepancies between them[19,20]. Nevertheless, there is a good correlation of our components b_2 with the original[19] Q which enabled us to calculate further values of Q (Table VIII). These values of course can serve only to the same purpose as our b_2, viz. prediction of SCS in position 2 by simple regression. There is no reason whatever to add σ_I and σ_R as further explanatory variables[19,20]: the respective terms would be superfluous and statistically insignificant. In conclusion, SCS in the position 2 depend strongly on short-range effects, are very similar in different series, but can be related to polar effects or electron densities only in a very restricted range (practically substituents NH_2, OH, OR, Me, CH_2Hal, CN).

The $J(H, H)$ coupling constants offer less possibility to be correlated with some property of the substituent than SCS. The main reason is the less favourable signal-to-noise ratio, particularly when the substituent is not directly adjoining. The salient feature of the structural dependence is that all substituents produce changes of J in the same direction: hence correlations were attempted with various scales of substituent electronegativity[10,11] rather than with constants σ. However, only for constants $J(2, 3)$ and $J(2, 6)$ an evident dependence was observed, in the other positions the sensitivity to substitution is too small. For this reason we restricted

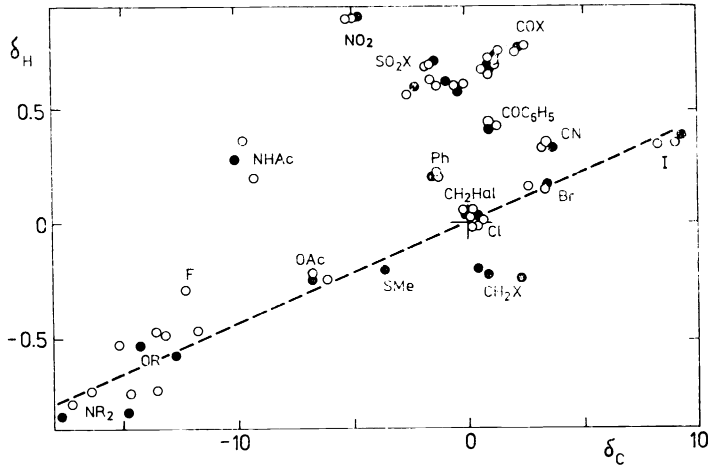


FIG. 4

The same plot as in Fig. 1 for the positions 2 and 6: ○ in *meta*-substituted benzonitriles (*I*), ● in *para*-substituted benzonitriles (*II*) our calculations to coupling constants $J(2,3)$, $J(5,6)$, and $J(2,6)$. The result of PCA (Table VII) is much less satisfactory than for SCS as far as the portion of explained variance is concerned but it is not so bad with respect to the experimental error. The components of the substituents (not listed) were plotted against electronegativity in Fig. 5. Four substituents deviate very markedly which can hardly have any common property (CN,OAc, NMe_2,SMe). We have used a relatively recent scale of electronegativity, called σ_α, the values of which are available for a large number of substituents[21]. Remarkably, more recent values of the same authors[22], which should be considered as improved, give much worsened fit, almost a complete scatter (deviations of the substituents OH, OMe, Cl, Ac, COOMe). Another explanation of the $J(H, H)$ coupling constants could be based on the substituent conjugative ability since the plot vs σ_R^+ and σ_R^- (not shown) yields two slightly non-linear dependences in the two halves (deviations of SCH_3 and NH_2). A decision between these two possibilities is prevented by the lack of reliable electronegativity data.

CONCLUSIONS

In general, the ^1H SCS in benzene derivatives seem to be less complex than ^{13}C SCS (refs[1-3]). They can be described – for each position separately – either by PCA with one term and nearly constant loadings, or – less precisely – by an additive scheme. Their prediction is thus relatively reliable: the components $b_2 - b_4$ (Table VII) can be used together with an empirically determined or estimated loading. The traditional DSP treatment is less effective and the difference is more dramatic than in the case of ^{13}C SCS, refs[1,2]. It follows that there are other factors controlling ^1H SCS, not involved in DSP. They are operating more in the position 3 than in 4, while in the position 2 the short-range forces are prevailing and classical inductive and mesomeric effects can be revealed only in a smaller group of substituents. In their sensitivity to these effects the ^1H SCS are very different from ^{13}C SCS in the

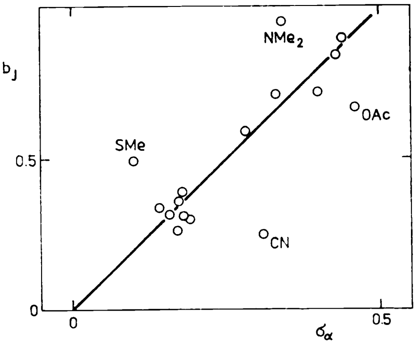


FIG. 5

Plot of the components b_J (for coupling constant $J(2, 3)$, $J(5, 6)$, and $J(2, 6)$) vs substituent electronegativities[21] σ_α

α-position[1,2] and similar merely to SCS of more distant carbon atoms[1]. However, a characteristic and puzzling feature of ^1H SCS is the strong mesomeric effect in the position 3. Returning to Eq. (*4*) we may state that the paramagnetic term δ_p is deciding in the positions 4 and 3, the long-range term δ_{lr} in the position 2 and partly operating in 3. It may be partitioned into anisotropy and ring-current effects but the suggested estimation[7] of these effects was not confirmed in our work.

REFERENCES

1. Exner O., Buděšínský M.: Magn. Reson. Chem. *27*, 27 (1989).
2. Buděšínský M., Exner O.: Magn. Reson. Chem. *27*, 585 (1989).
3. Buděšínský M., Johnels D., Edlund U., Exner O.: Collect. Czech. Chem. Commun. *55*, 368 (1990).
4. Ewing D. F. in: *Correlation Analysis in Chemistry: Recent Advances* (N. B. Chapman and J. Shorter, Eds), p. 357. Plenum Press, New York 1978.
5. Ehrenson S., Brownlee R. T. C., Taft R. W.: Prog. Phys. Org. Chem. *10*, 1 (1973).
6. Günther H.: *NMR Spectroscopy — An Introduction,* Chapter 4. Wiley, Chichester 1980.
7. Yamada H., Tsuno Y., Yukawa Y.: Bull. Chem. Soc. Jpn. *43*, 1459 (1970).
8. Hruska F. ,Hutton H. M., Schaefer T.: Can. J. Chem. *43*, 2392 (1966).
9. Smith W. B., Ihrig A. M., Roark J. L.: J. Phys. Chem. *74*, 812 (1970).
10. Tribble M. T., Traynham J. G. in: *Advances in Linear Free Energy Relationships* (N. B. Chapman and J. Shorter, Eds), p. 143. Plenum Press, London 1972.
11. Hayamizu K., Yamamoto O.: J. Mol. Spectrosc. *25*, 420 (1968).
12. Exner O.: Collect. Czech. Chem. Commun. *31*, 3222 (1966).
13. Hayamizu K., Yamamoto O.: J. Mol. Spectrosc. *28*, 89 (1968); *29*, 183 (1969).
14. Yamamoto O., Hayamizu K., Sekine K., Funahira S.: Anal. Chem. *44*, 1794 (1972).
15. Hehre W. J., Taft R. W., Topsom R. D., Prog. Phys. Org. Chem. *12*, 159 (1976).
16. Bromilow J., Brownlee R. T. C., Craik D. J., Sadek M., Taft R. W.: J. Org. Chem. *45*, 2429 (1980).
17. Reed J. J. R.: Anal. Chem. *39*, 1586 (1967).
18. Charton M.: J. Org. Chem. *36*, 266 (1971).
19. Smith W. B., Proulx T. W.: Org. Magn. Reson. *8*, 567 (1976).
20. Reynolds W. F., Gomes A., Maron A., McIntyre D. W., Tanin A., Hamer G. K., Peat I. R.: Can. J. Chem. *61*, 2376 (1983).
21. Marriott S., Reynolds W. F., Taft R. W., Topsom R. D.: J. Org. Chem. *49*, 959 (1984).
22. Taft R. W., Topsom R. D.: Prog. Phys. Org. Chem. *16*, 1 (1987).

Translated by the author (O.E.).