MULTIVARIATE REGRESSION WITH SUBSTITUENT SHIFT INCREMENTS. IV. 2-(4-X-PHENYL)-1,3-DIHYDRO-2H-ISOINDOLE-1,3-DIONES AND 3-(4-X-PHENYL)-3,4-DIHYDRO-2H-1,3-BENZOXAZINE-2,4-DIONES[+]

Miroslav HOLÍK[a],*, Zdeněk FRIEDL[b], Karel WAISSER[c1] and Jiří GREGOR[c2]

[a] *Department of Theoretical and Physical Chemistry, Faculty of Science, Masaryk University, CZ-611 37 Brno, Czech Republic; e-mail: holik@chemi.muni.cz*

[b] *Faculty of Chemistry, Technical University Brno, CZ-612 00 Brno, Czech Republic; e-mail: friedl@fch.vutbr.cz*

[c] *Department of Inorganic and Organic Chemistry, Pharmaceutical Faculty, Charles University, CZ-501 65 Hradec Králové, Czech Republic; e-mail: [1] waisser@faf.cuni.cz., [2] gregor@faf.cuni.cz*

Dedicated to Professor Otto Exner on the occasion of his 75th birthday.

Two series of *para* disubstituted benzenes were studied: 2-(4-X-phenyl)-1,3-dihydro-2H-isoindole-1,3-diones (**1**) and 3-(4-X-phenyl)-3,4-dihydro-2H-1,3-benzoxazine-2,4-diones (**2**). Their ^1H and ^{13}C chemical shifts were correlated with substituent shift increments (SSI) a_j and z_j, respectively. For ^{13}C chemical shifts, all four z_j values, z_i, z_o, z_m, and z_p, were used to check the assignment and to find out possible variables for improvement of regression equations. Significant deviations from plain additivity were observed in the case of δ_{H3} and δ_{C3} chemical shifts. This can be explained by changes in diamagnetic anisotropy contribution induced by different twist of 4-substituent from the benzene plane caused by variable substituent in position 1.

Key words: Substituent shift increments; NMR spectroscopy, ^1H and ^{13}C; Substituent effects; Linear regression; Proportionality relationship; Phthalimides; Benzoxazines.

It is believed that in 1,4-X,Y-disubstituted benzenes with a variable substituent X and a fixed substituent Y, possibly containing multiple bonds, the twist angle between the benzene ring and the plane of Y substituent is reflected by a slope ρ in the Hammett equation and is therefore constant[1] . However, we have shown[2–6] that the most probable explanation of chemical shifts in the neighbourhood of substituent Y is the

---

+ For Part III see ref.[20]

change of its diamagnetic anisotropy or van der Waals contribution due to variation of the twist angle with the substituent X. The change of the twist in molecules with the change of the substituent can be observed also in the magnitude of the ^4J coupling constants of the benzylidene type[7].

An effect of diamagnetic anisotropy of multiple bond or aromatic ring on the shielding of nuclei is more pronounced in ^1H than ^{13}C NMR spectroscopy. Thus, in 1,4-disubstituted benzenes the most affected NMR signal should be that of H atom in *meta* position to a variable X substituent. For its chemical shift we can write according to the additivity rule Eq. (*1*).

$$\delta_m = 7.27 + a_o(Y) + a_m(X) \, , \tag{1}$$

where a_m and a_o are substituent shift increments (SSI) obtained from spectra of monosubstituted benzenes[8].

$$\delta_m = b_0(Y) + b_1 \, a_m(X) + \varepsilon \tag{2}$$

Equation (*2*) can be used for analysis of chemical shifts of 1,4-disubstituted benzenes with a fixed substituent Y and a variable substituent X; it has been used for checking correctness of ^1H NMR signal assignment[9] and for estimating[2–6] values of slopes b_1. According to the magnitude of b_1 value the investigated cases can be roughly divided into three categories:

1. $b_1 \approx 1$

No essential change in the conformation of group Y, from the viewpoint of diamagnetic anisotropy, is caused by exchanging one X substituent for another.

$$\delta_m = 7.621 + 0.978 \, a_m + \varepsilon \qquad n = 23 \tag{3a}$$

The value 0.978 statistically (95%) insignificantly different from one was obtained for 4-substituted benzonitriles[10]. Also the slope 1.009 found for 5-(*p*-X-phenyl)tetrazoles[11] was not significantly different from one.

$$\delta_m = 8.179 + 1.009 \, a_m + \varepsilon \qquad n = 10 \tag{3b}$$

2. b < 1

In the substances where the dihedral angle between the benzene ring and the plane of substituent Y decreases with the increasing donor character (DID) of the substituent X, *i.e.*, nitro derivative is more twisted than amino one, slope *b* significantly differs from one. Thus, the downfield shift of H_m NMR signal caused by a decrease in the electron density on H_m atom is partially compensated by a decrease in the diamagnetic anisotropy contribution of substituent Y which shifts the signal upfields.

$$\delta_m = 7.654 + 0.303\ a_m + \varepsilon \qquad n = 7 \qquad (3c)$$

The slope value (0.303) in 2-(4'-substituted benzylidene)[3]ferrocenophane-1,3-diones[12] considerably differs from unity and even larger decrease was observed in the case of 4-substituted *N*-benzylideneanilines[2,13]; H_m chemical shifts being almost unaffected by substitution.

$$\delta_m = 7.075 + 0.053\ a_m + \varepsilon \qquad n = 7 \qquad (3d)$$

3. b > 1

When the twist angle between the benzene ring and the substituent Y decreases with decreasing donor character (DDD) of the substituent X, *i.e.*, nitro derivative is less twisted than amino one, then an increase in diamagnetic anisotropy contribution supports the electronic effect and the slope of regression equation (*2*) is larger than one. This effect was observed[5] for 1-(4-X-phenyl)-2,6-dimethyl-4-pyridones

$$\delta_m = 7.235 + 1.275\ a_m + \varepsilon \qquad n = 7 \qquad (3e)$$

and 2-(4-X-phenyl)-1,3-dihydro-2*H*-isoindole-1,3-diones (**1**); both slopes being significantly different from one.

$$\delta_m = 7.452 + 1.252\ a_m + \varepsilon \qquad n = 7 \qquad (3f)$$

A significance of fitted parameters in multivariate regression was tested by the Student t-test by comparing tested value divided by corresponding standard deviation with critical t-value for respective number of freedoms and confidence level (normally 95%), Eq. ($3g$).

$$t = \frac{b - \beta}{s_b} \geq t_{(0.05, n-p)} \tag{3g}$$

With $\beta = 0$ it tests statistical importance of corresponding variable, with $\beta = 1$ it tests a significance of deviation of slope from one.

In ^{13}C NMR spectra the effect of diamagnetic anisotropy is less pronounced than in ^1H NMR and electronic effects may prevail. With multivariate regression[14] of ^{13}C chemical shifts on four ^{13}C SSI, Eq. (4), a preliminary insight into the electronic effects acting in 1,4-disubstituted benzenes has been obtained in the following way[14]:

$$\delta_C = b_0\ 128.5 + b_1\ z_i + b_2\ z_o + b_3\ z_m + b_4\ z_p + \varepsilon\ . \tag{4}$$

For Y group acting as electron donor:

δ_{C1} correlates with z_i and z_m ($-b_3$)

δ_{C3} correlates with z_m and z_p ($-b_4$) .

If Y acts as electron acceptor:

δ_{C1} correlates in addition to z_i also with z_o ($+b_2$) and z_p ($-b_4$)

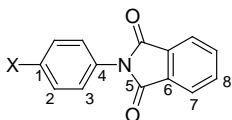δ_{C3} correlates in addition to z_m also with z_o ($-b_2$) and z_p ($\pm b_4$) .

In the case of more complicated substituent Y, like in the mentioned 4-pyridones and N-phenylphthalimides (**1**), δ_{C1} correlated significantly with all SSI, *i.e.*, z_i, z_o, z_m, and z_p; the δ_{C3} gave significant slopes with z_m and z_p its sign depending probably on a mixture of several effects.

Now we have a larger series of 2-(4-X-phenyl)-1,3-dihydro-2H-isoindole-1,3-diones (**1**) and a similar series of 3-(4-X-phenyl)-3,4-dihydro-2H-1,3-benzoxazine-2,4-diones (**2**) and we can repeat some calculations with more confidence and make other to support previous findings[14].
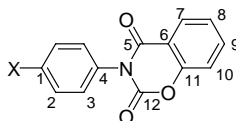
EXPERIMENTAL

Preparations

2-(4-Substituted phenyl)-1,3-dihydro-2H-isoindole-1,3-diones (**1**)[5,18] were prepared by refluxing equimolar amounts (50 mmol) of phthalic anhydride and respective 4-substituted aniline in toluene (100 ml) containing triethylamine (0.8 ml). The reactions were carried out while removing water for 3 h, crude reaction mixtures were filtered hot through a silica gel column and recrystallized from toluene. All compounds gave satisfactory elemental analyses and sharp melting points.

1				**2**			
a	NMe$_2$	**h**	H	**a**	NMe$_2$	**g**	H
b	OMe	**i**	CF$_3$	**b**	OMe	**h**	CF$_3$
c	F	**j**	CN	**c**	F	**i**	CN
d	Cl	**k**	COOMe	**d**	Cl	**j**	COOEt
e	Br	**l**	COMe	**e**	Br	**k**	NO$_2$
f	Me	**m**	NO$_2$	**f**	Me		
g	t-Bu						

3-(4-Substituted phenyl)-3,4-dihydro-2H-1,3-benzoxazine-2,4-diones were prepared[15–17] by adding ethyl chloroformate (5.2 g, 48 mmol) dropwise to a stirred solution of the respective 4-substituted salicylanilide (40 mmol) in dry pyridine (20 ml) cooled with ice. The prepared mixture was heated on a steam bath for 1 h and then poured into 140 ml of 5% hydrochloric acid. Next day the product was filtered off, suspended in 5% potassium hydroxide solution, thoroughly stirred, filtered off and recrystallized from ethanol in yields of 55–65%.

NMR Measurements

Spectral measurements were done on a Bruker Avance DRX-500 spectrometer at 500 (^{1}H) and 125 (^{13}C) MHz, respectively. Samples were measured at concentration 0.1 mol dm^{-3} in CDCl$_3$ (**1**) or DMSO-d_6 (**2**) and chemical shifts were standardized on TMS (^{1}H) or on the central peak of the solvent used[19] (^{13}C): CDCl$_3$ δ 76.91 and DMSO-d_6 δ 39.56. Digital resolution was 0.0002 and 0.007 ppm in ^{1}H and ^{13}C NMR spectra, respectively. Chemical shifts are collected in Table I (**1**) and Table II (**2**).

Calculations

All calculations were carried out with double precision on a PC with Pentium-S/150 MHz processor with 32 MB RAM. Standard subroutines were used for linear regression and significance testing; for assignment a multivariate regression program ASSIGN was used[20]. The semiempirical calculations at the AM1 level[21] were performed under full geometry optimization.

TABLE I
^{1}H (standard TMS) and ^{13}C (standard solvent[19] $\delta(CDCl_3)$ 76.91) NMR chemical shifts of **1** in deuteriochloroform

Compound	X	C1	C2	C3	C4	C5	C6	C7	C8	H2	H3
1a	NMe_2	149.977	112.925	127.533	120.787	167.811	132.070	123.529	134.240	6.820	7.252
1b	OMe	159.260	114.450	127.900	124.340	167.320	131.840	123.590	134.220	7.019	7.337
1c	F	161.988	116.092	128.381	127.763	167.156	131.779	123.796	134.472	7.200	7.429
1d	Cl	133.750	129.250	127.620	130.280	166.900	131.640	123.800	134.500	7.475	7.420
1e	Br	121.758	132.264	127.921	130.905	166.840	131.706	123.820	134.521	7.626	7.356
1f	Me	138.060	129.700	126.400	129.030	167.320	131.840	123.590	134.220	7.292	7.325
1g	t-Bu	151.117	126.053	126.126	129.085	167.374	131.949	123.650	134.278	7.520	7.356
1h	H	128.020	129.050	126.540	131.780	167.190	131.780	123.680	134.320	7.470	7.470
1i	CF_3	129.922	126.222	126.441	135.127	166.719	131.657	123.990	134.715	7.778	7.649
1j	CN	111.421	132.992	126.538	136.098	166.501	131.512	124.160	134.982	7.798	7.695
1k	COOMe	129.280	130.340	125.860	135.650	166.690	131.580	123.840	134.580	8.183	7.603
1l	COMe	136.140	129.150	126.100	136.140	166.770	131.650	123.960	134.690	8.103	7.633
1m	NO_2	146.720	124.420	126.330	136.080	166.390	131.420	124.180	134.980	8.380	7.777

TABLE II
1H (standard TMS) and ^{13}C (standard solvent[19] δ(DMSO-d$_6$) 39.56) NMR chemical shifts of **2** in DMSO-d$_6$

Compound	X	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	H2	H3
2a	NMe$_2$	150.57	112.712	128.243	122.553	160.965	114.492	128.391	125.310	136.086	116.419	152.686	148.417	6.777	7.170
2b	OMe	159.23	114.211	129.573	127.674	160.775	114.780	127.441	125.226	136.227	116.321	152.524	147.762	7.044	7.329
2c	F	161.79	115.800	130.712	131.296	160.613	114.731	127.385	125.247	136.283	116.321	152.496	147.516	7.344	7.493
2d	Cl	134.05	129.038	130.515	133.350	160.522	114.710	127.392	125.296	136.346	116.356	152.217	147.383	7.587	7.478
2e	Br	121.88	132.020	130.853	134.510	160.487	114.703	127.406	125.317	136.367	116.363	152.524	147.347	7.720	7.410
2f	Me	138.140	128.236	129.446	132.548	160.655	114.766	127.427	125.247	136.255	16.335	152.545	147.615	7.305	7.289
2g	H	128.67	128.961	128.553	135.172	160.620	114.780	127.427	125.275	136.297	116.356	152.574	160.620	7.460	7.530
2h	CF$_3$	129.27	126.098	129.784	138.836	160.473	114.682	127.392	125.345	136.431	116.384	152.545	147.256	7.911	7.706
2i	CN	111.73	133.118	130.023	139.364	160.353	114.640	127.385	125.380	136.473	116.384	152.510	147.108	8.010	7.687
2j	COOEt	130.29	129.763	129.115	139.279	160.430	114.682	127.371	125.303	136.374	116.356	152.545	147.221	8.095	7.605
2k	NO$_2$	147.07	124.206	130.325	140.848	160.367	114.640	127.378	125.380	136.480	116.391	152.524	147.495	8.387	7.774

RESULTS AND DISCUSSION

^{1}H NMR Spectra of **1** *and* **2**

For the regression analysis of ^{1}H chemical shifts of signals assigned to hydrogen atoms *ortho*, δ_{H_o}, and *meta*, δ_{H_m}, to variable substituent X, the following model equation, Eq. (*5*), was used.

$$\delta_{H_j} = b_0\ 7.27 + b_1\ a_j + \varepsilon_j\ ,\qquad(5)$$

where a_j is corresponding SSI variable[8] and ε_j is statistical error. Chemical shifts δ_{H_o} in both series **1** and **2** show regular additive behaviour (see Eqs (*6a*) and (*6b*)). The subscripts in parentheses are corresponding standard deviations; if smaller than 0.01 the value is substituted by 0.01).

$$\delta_{H_o}\,(\mathbf{1}) = 1.028_{(0.01)} + 0.975_{(0.01)}\ a_o + \varepsilon_o\qquad(6a)$$

$$s_y = 0.457 \qquad r = 0.9990 \qquad n = 13$$

$$\delta_{H_o}\,(\mathbf{2}) = 1.034_{(0.03)} + 0.978_{(0.05)}\ a_o + \varepsilon_o\qquad(6b)$$

$$s_y = 0.482 \qquad r = 0.9863 \qquad n = 11$$

On the other hand, regression analysis of chemical shifts δ_{H_m} in both series **1** and **2** afforded slope value, b_1, significantly larger than one (see Eqs (*7a*) and (*7b*)).

$$\delta_{H_m}\,(\mathbf{1}) = 1.026_{(0.01)} + 1.246_{(0.03)}\ a_m + \varepsilon_m\qquad(7a)$$

$$s_y = 0.167 \qquad r = 0.9965 \qquad n = 13$$

$$\delta_{H_m}\,(\mathbf{2}) = 1.028_{(0.01)} + 1.330_{(0.10)}\ a_m + \varepsilon_m\qquad(7b)$$

$$s_y = 0.188 \qquad r = 0.9758 \qquad n = 11$$

The slope b_1 (**1**) for the larger set (Eq. (*7a*)), 1.246, does not differ from the value obtained formerly[5], 1.252 (Eq. (*3f*)). This means, that it does not depend on the number of measurements (compounds). The residuals from Eq. (*6a*), ε_m, were transformed to "jackknife" residuals, which can help to detect outliers in data. In order to enhance a numerical value of residual which should belong to a probable outlier, the plain residuals, *i.e.*, differences between experimental and calculated values, ε, are transformed to "jackknife" residuals, ε_{jk}, using Eq. (*7c*).

$$\varepsilon_{jk} = \varepsilon_n \sqrt{\frac{n - p - 1}{n - p - \varepsilon_n^2}} \ , \quad \text{where} \quad \varepsilon_n = \frac{\varepsilon}{s_e \sqrt{1 - h_i}} \qquad (7c)$$

where s_e is standard error of estimate and h_i is the corresponding diagonal element of a transformation matrix $\mathbf{H} = X(X'X)^{-1}X'$. If $\varepsilon_j > 0.2\ (n - p)$ then the corresponding y_i value is considered as suspicious and the regression should be repeated without it. The "jackknife" residual for F derivative showed value 2.74 and the corresponding δ_{H_m} (*F*) datum can be suspected as influential. Omitting this measurement the regression equation (Eq. (*7d*)) was obtained.

$$\delta_{H_m} (\mathbf{1}) = 1.026_{(0.01)} + 1.243_{(0.02)}\ a_m + \varepsilon_m \qquad (7d)$$

$$s_y = 0.174 \qquad r = 0.9980 \qquad n = 12$$

It is clear that neither the number of measurements nor the presence of outlier is responsible for the slope b_1(**1**) larger than one.

In order to support our idea that this increase is due to the change in diamagnetic contribution of the Y group (see above), the geometries of three substances **1** (X = NH_2, H, NO_2) were calculated by semiempirical AM1 method which was developed for the calculation of molecular conformations and electronic properties[21]. The results of calculations (length of C_{aryl}–N bond in Å, dihedral angle in degrees between the aniline benzene ring and the plane of N=C–H atoms are given), X = NO_2, 1.406, 25.36; X = H, 1.412, 29.26; X = NH_2, 1.413, 30.76, are in agreement with our interpretation of the conformational changes induced by the change in the substituent X. When comparing b_1(**1**) = 1.246 with b_1(**2**) = 1.330 it was tempting to attribute the increase in the latter case to the different geometry of the six- and five-membered rings. However, the inspection of "jack-

knife" residuals of Eq. (*7b*) revealed that $\varepsilon_{jk}(NMe_2) = 2.09$ and **1a** derivative can be suspected as outlier. After leaving it out (see Eq. (*7e*)), the lower value of b_1 (1.232) was obtained.

$$\delta_{H_m}(\mathbf{2}) = 1.030_{(0.01)} + 1.232_{(0.10)}\ a_m + \varepsilon_m \tag{7e}$$

$$s_y = 0.161 \qquad r = 0.9768 \qquad n = 10$$

^{13}C NMR Spectra of **1** and **2**

As the first step in analyzing ^{13}C NMR spectra of both series, the program ASSIGN[20] was used to check the signal assignment. This program uses Eq. (*4*) for fitting all chemical shifts of aromatic carbon atoms with all corresponding SSI , *i.e.* z_i, z_o, z_m, and z_p. The b_j values obtained were tested by the Student *t*-test for the statistical significance on the 95% confidence level. The significant values are in Table III.

While the regression with all SSI is advantageous from the viewpoint of the best fit for signal assignment, for the analysis of data only those variables are relevant which possess statistically significant slope values. Therefore, each data set (*i.e. ipso, ortho, etc.*) was analyzed independently.

Chemical Shifts of C1 (*ipso* to X) Atoms

In the series **1**, the regression of δ_{C1} with corresponding SSI, *i.e.*, z_i, is given in Eq. (*8a*).

$$\delta_{C1}(1) = 0.993_{(0.01)} + 1.005_{(0.01)}\ z_i + \varepsilon \tag{8a}$$

$$s_y = 14.86 \qquad r = 0.9992 \qquad n = 13$$

However, "jackknife" residuals suggest that the value for methyl derivative could be an outlier: $\varepsilon_{jk}(Me) = 2.44 > 0.2 * 11$.

After leaving this measurement out almost same equation (Eq. (*8b*)) was obtained.

$$\delta_{C1}(\mathbf{1}) = 0.992_{(0.01)} + 1.006_{(0.01)} \, z_i + \varepsilon \qquad (8b)$$

$$s_y = 15.52 \qquad r = 0.9995 \qquad n = 12$$

Neither parameters b were changed nor the standard deviation s_y decreased; we can conclude that the value for the Me derivative is not an outlier. Then the remaining SSI values, each individually or in pairs, were tested as possible other explanatory variables but all tests were negative. Statistical significance of addition of another explanatory variable into a regression can be tested by a partial F-test Eq. (*8c*)

$$F = \frac{SR_p - SR_{p+1}}{SR_{p+1}} (n - p) \ge F_{0.05, 1, n-p} \; , \qquad (8c)$$

where SR_p is the sum of squares of residuals for regression with p parameters and SR_{p+1} the same quantity for regression with $p + 1$ parameters.

Only when all of them were taken (like in ASSIGN program) then the result of F-test was positive.

TABLE III
Regression parameters, correlation coefficients, and standard deviation of data calculated by program ASSIGN using four SSI as explanatory variables

C_{atom}	Com-pound	b_0	b_1	b_2	b_3	b_4	r	s_y
1	**1**	0.995	1.024	0.101	−0.465	−0.117	0.9999	14.86
	2	0.998	1.030	0.128	−0.615	−0.124	0.9999	15.61
2	**1**	1.005		1.012		0.041	0.9996	6.79
	2	1.002		1.003		0.081	0.9993	7.45
3	**1**	0.983			0.910	−0.071	0.9693	0.84
	2	1.003			1.211	0.100	0.9646	0.86
4	**1**	1.027				0.867	0.9973	4.87
	2	1.053				0.961	0.9983	5.36

Parameters b were calculated according to Eq. (*4*); only those significant by the Student *t*-test at the 95% level are given.

$$F\text{-test} = 20.79 > F_{(0.05,\,3,\,8)} = 4.07 \text{ and also} > F_{(0.01,\,3,\,8)} = 7.59$$

In the series **2**, the regression of δ_{C1} with corresponding SSI, *i.e.*, z_i, is given in Eq. (*8d*).

$$\delta_{C1}(\mathbf{2}) = 0.995_{(0.01)} + 0.993_{(0.01)}\, z_i + \varepsilon \qquad (8d)$$

$$s_y = 15.61 \qquad r = 0.9991 \qquad n = 11$$

Also in this series, the "jackknife" residuals suggest that the value for methyl derivative could be outlier: $\varepsilon_{jk}(\text{Me}) = 2.06 > 0.2 * 9$.

And again, no improvement was achieved after leaving this measurement out (Eq. (*8e*)).

$$\delta_{C1}(\mathbf{2}) = 0.994_{(0.01)} + 0.993_{(0.01)}\, z_i + \varepsilon \qquad (8e)$$

$$s_y = 16.45 \qquad r = 0.9994 \qquad n = 10$$

Small improvement in regression can be obtained again only with all the other SSI values as explanatory variables.

$$F\text{-test} = 16.16 > F_{(0.05,\,3,\,6)} = 4.76 \text{ and also} > F_{(0.01,\,3,\,6)} = 9.78\,.$$

Chemical Shifts of C2 (*ortho* to X) Atoms

The results from the ASSIGN program (Table III) suggest that besides the z_o also the z_p variable can be used in regression with δ_{C2}. In the series **1**, the regression of δ_{C2} with z_o is given in Eq. (*9a*).

$$\delta_{C2}(\mathbf{1}) = 1.005_{(0.01)} + 1.023_{(0.01)}\, z_o + \varepsilon \qquad (9a)$$

$$s_y = 6.79 \qquad r = 0.9993 \qquad n = 13$$

From "jackknife" residuals the value for the dimethylamino derivative came out as a possible outlier: $\varepsilon_{jk}(NMe_2) = 2.80 > 0.2 * 11$.

After leaving this measurement out a slight improvement was achieved (Eq. (*9b*)):

$$\delta_{C2}(\mathbf{1}) = 1.005_{(0.01)} + 1.007_{(0.01)} \, z_o + \varepsilon \qquad (9b)$$

$$s_y = 5.87 \qquad r = 0.9995 \qquad n = 12$$

This reduced series was tested by the *F*-test for the addition of z_p variable to the regression equation; the test was negative.

$$F\text{-test} = 4.12 < F_{(0.05, \, 1, \, 9)} = 5.12$$

When this test was performed on the original data it was positive at the 95% confidence level but negative at the 99% one. The necessity to improve the regression including z_p variable can be evoked by that slightly outlying value for the NMe_2 derivative. An analogous situation was observed in series **2**.

$$\delta_{C2}(\mathbf{2}) = 1.003_{(0.01)} + 1.039_{(0.02)} \, z_o + \varepsilon \qquad (9c)$$

$$s_y = 7.45 \qquad r = 0.9984 \qquad n = 11$$

The "jackknife" residuals pointed out that again the value for the dimethylamino derivative could be an outlier: $\varepsilon_{jk}(NMe_2) = 2.81 > 0.2 * 9$.

$$\delta_{C2}(\mathbf{2}) = 1.003_{(0.01)} + 1.013_{(0.02)} \, z_o + \varepsilon \qquad (9d)$$

$$s_y = 6.45 \qquad r = 0.9989 \qquad n = 10$$

The conclusion made on the series **1** data is supported by those of the series **2**. Addition of z_p variable is not necessary.

$$F\text{-test} = 4.89 < F_{(0.05, \, 1, \, 7)} = 5.59$$

Chemical Shifts of C3 (*meta* to X) Atoms

The results from the ASSIGN program (Table III) suggest that besides the z_m also the z_p variable can be used in regression with δ_{C3}. In the series **1**, the regression of δ_{C3} with z_m, is given in Eq. (*10a*).

$$\delta_{C3}(\mathbf{1}) = 0.983_{(0.01)} + 1.005_{(0.17)}\, z_m + \varepsilon \qquad (10a)$$

$$s_y = 0.84 \qquad r = 0.8790 \qquad n = 13$$

Due to a very small susceptibility of the experimental data ($s_y = 0.84$) to the change in substituent X it was necessary to analyze this dependence most carefully. As a suspected outlier we selected the NO_2 derivative: $\varepsilon_{jk}(NO_2) = 2.80 > 0.2 * 11$. After leaving corresponding data out Eq. (*10b*) was obtained.

$$\delta_{C3}(\mathbf{1}) = 0.983_{(0.01)} + 1.10_{(0.13)}\, z_m + \varepsilon \qquad (10b)$$

$$s_y = 0.86 \qquad r = 0.9329 \qquad n = 12$$

Regardless of whether the NO_2 derivative was retained or left out the *F*-test for addition of z_p variable to the regression was positive.

F-test $= 26.18 > F_{(0.05,\ 1,\ 10)} = 4.96$ and also $> F_{(0.01,\ 1,\ 10)} = 10.04$ for $n = 13$

F-test $= 18.62 > F_{(0.05,\ 1,\ 9)} = 5.12$ and also $> F_{(0.01,\ 1,\ 9)} = 10.56$ for $n = 12$

Inclusion of z_p variable led to the better fit, Eq. (*10c*) or (*10d*).

$$\delta_{C3}(\mathbf{1}) = 0.983_{(0.01)} + 0.893_{(0.10)}\, z_m - 0.069_{(0.01)}\, z_p + \varepsilon \qquad (10c)$$

$$s_y = 0.84 \qquad r = 0.9681 \qquad n = 13$$

$$\delta_{C3}(\mathbf{1}) = 0.983_{(0.01)} + 0.951_{(0.09)}\, z_m - 0.055_{(0.01)}\, z_p + \varepsilon \qquad (10d)$$

$$s_y = 0.86 \qquad r = 0.9787 \qquad n = 12$$

Similar behaviour was observed in series **2**. However, in this case the NMe_2 derivative was a suspected outlier. The data including the NMe_2 derivative gave Eq. (*11a*) and those without the NMe_2 compound gave Eq. (*11b*).

$$\delta_{C3}(\textbf{2}) = 1.004_{(0.01)} + 0.949_{(0.26)}\ z_m + \varepsilon \qquad (11a)$$

$$s_y = 0.86 \qquad r = 0.7778 \qquad n = 11$$

$$\varepsilon_{jk}(NMe_2) = 2.67 > 0.2 * 9$$

$$\delta_{C3}(\textbf{2}) = 1.005_{(0.01)} + 0.997_{(0.19)}\ z_m + \varepsilon \qquad (11b)$$

$$s_y = 0.85 \qquad r = 0.8761 \qquad n = 10$$

Since in both cases the *F*-test allowed to add explanatory variable z_p to the regression the Eqs (*11c*) and (*11d*) were calculated.

F-test $= 37.08 > F_{(0.05,\ 1,\ 8)} = 5.32$ and also $> F_{(0.01,\ 1,\ 8)} = 11.26$ for $n = 11$

F-test $= 15.96 > F_{(0.05,\ 1,\ 7)} = 5.59$ and also $> F_{(0.01,\ 1,\ 7)} = 12.25$ for $n = 10$

$$\delta_{C3}(\textbf{2}) = 1.004_{(0.01)} + 1.216_{(0.12)}\ z_m + 0.097_{(0.02)}\ z_p + \varepsilon \qquad (11c)$$

$$s_y = 0.84 \qquad r = 0.9643 \qquad n = 11$$

$$\delta_{C3}(\textbf{2}) = 1.004_{(0.01)} + 1.192_{(0.12)}\ z_m + 0.084_{(0.02)}\ z_p + \varepsilon \qquad (11d)$$

$$s_y = 0.85 \qquad r = 0.9639 \qquad n = 10$$

Chemical Shifts of C4 (*para* to X) Atoms

The ASSIGN program (Table III) shows that only the z_p variables correlate with δ_{C4}. In the series **1**, the regression of δ_{C4} with z_p is given in Eq. (*12a*).

$$\delta_{C4}(\mathbf{1}) = 1.025_{(0.01)} + 0.922_{(0.03)}\, z_p + \varepsilon \qquad (12a)$$

$$s_y = 4.87 \qquad r = 0.9951 \qquad n = 13$$

After excluding the NO_2 derivative as possible outlier: $\varepsilon_{jk}(NO_2) = 5.57 > 0.2 * 11$, we obtained Eq. (*12b*).

$$\delta_{C4}(\mathbf{1}) = 1.027_{(0.01)} + 0.958_{(0.01)}\, z_p + \varepsilon \qquad (12b)$$

$$s_y = 4.83 \qquad r = 0.9989 \qquad n = 12$$

In both cases the *F*-test showed non-importance of other explanatory variables.

It was shown that in 1,4-disubstituted benzenes the second order regression[22] of δ_{C4} with σ_I and σ_R constants[23] can give electron demand of the Y substituent.

$$\delta_{C4} = 131.912_{(0.25)} + 3.859_{(0.61)}\sigma_I + 18.602_{(1.01)}\sigma_R - 5.200_{(2.65)}\sigma_R^2 + \varepsilon \quad (12c)$$

$$s_y = 4.87 \qquad r = 0.9967 \qquad n = 13$$

The slope at the σ_R^2 variable, b_3, is significant only at the 90% level. Therefore, the electron demand value $\eta = b_3/b_2 = -0.279$ is significant also at this level. If the suspected outlier NO_2 is left out, Eq. (*12d*) gives b_3 parameter at the 95% level an the calculated η is more probable: $b_3/b_2 = -0.158$.

$$\delta_{C4} = 131.773_{(0.12)} + 4.647_{(0.31)}\sigma_I + 19.754_{(1.01)}\sigma_R - 3.125_{(1.26)}\sigma_R^2 + \varepsilon \quad (12d)$$

$$s_y = 4.83 \qquad r = 0.9993 \qquad n = 12$$

From the comparison of this η datum with those for common substituents[22] it follows that the phthalimido group in **1** can be regarded as a weak electron acceptor.

For series **2** the equation corresponding to (*12a*) reads (*13a*).

$$\delta_{C4}(\mathbf{2}) = 1.053_{(0.01)} + 0.986_{(0.02)}\ z_p + \varepsilon \qquad (13a)$$

$$s_y = 5.36 \qquad r = 0.9977 \qquad n = 11$$

Since no outlier was suspected the regression with σ constants[23] had the form of Eq. (*13b*).

$$\delta_{C4} = 135.202_{(0.36)} + 4.159_{(0.82)}\sigma_I + 17.538_{(1.31)}\ \sigma_R - 7.604_{(3.36)}\sigma_R^2 + \varepsilon \quad (13b)$$

$$s_y = 5.36 \qquad r = 0.9959 \qquad n = 11$$

The b_3 parameter was again significant only at the 90% level. Nevertheless, the negative η value, $b_3/b_2 = -0.433$, allowed to classify the substituent Y in series **2** also as an electron acceptor.

CONCLUSIONS

Thorough testing of regression parameters and residuals obtained from the correlation of ^{1}H and ^{13}C NMR chemical shifts in 1,4-disubstituted benzenes **1** and **2** with substituent shift increments a_j and z_j, respectively, showed that chemical shifts in the *meta* position to variable X substituent, *i.e.* in the *ortho* position to fixed substituent Y, are sensitive to conformational changes concerning the benzene ring and the plane of Y substituent. This effect is most pronounced in ^{1}H chemical shifts causing a significant increase in corresponding slope values over one. In ^{13}C NMR shifts the variable which reflects this conformational change can be substituted by the SSI value for the *para* position since the electron density on C4 is decisive for the bond order and twist angle of the substituent Y. However, this effect is not so clear as in the case of ^{1}H chemical shifts and the reason for the sign of the corresponding slope may consist in a different space in-

terval (and therefore in diamagnetic anisotropy field) around the C=O bond in which the carbon C_m can move.

REFERENCES

1. Craik D. J., Brownlee R. T. C., Sadek M.: *J. Org. Chem.* **1982**, *47*, 657.
2. Holík M., Běluša J., Břicháček J.: *Collect. Czech. Chem. Commun.* **1978**, *43*, 610.
3. Holík M., Potáček M., Švaříček J.: *Collect. Czech. Chem. Commun.* **1978**, *43*, 734.
4. Holík M., Pavéska P., Mlynárik V.: *J. Mol. Struct.* **1984**, *114*, 15.
5. Holík M., Matějková B.: *Collect. Czech. Chem. Commun.* **1990**, *55*, 261.
6. Holík M.: *Chemom. Intell. Lab. Syst.* **1993**, *19*, 225.
7. Haslinger E., Wolschann P.: *Org. Magn. Reson.* **1977**, *9*, 1.
8. Beeby J., Sternhell S., Hoffmann-Ostenhof T., Pretsch E., Simon W.: *Anal. Chem.* **1973**, *45*, 1571.
9. Holík M.: *Org. Magn. Reson.* **1977**, *9*, 491.
10. Exner O., Buděšínský M.: *Collect. Czech. Chem. Commun.* **1991**, *56*, 2234.
11. Ciarkowski J., Kaczmarek J., Grzouka Z.: *Org. Magn. Reson.* **1979**, *12*, 631.
12. Solčániová E., Toma Š., Sališová M., Liptaj T.: *Org. Magn. Reson.* **1983**, *21*, 429.
13. Perez C., Schleinitz K.-D., Gründemann E.: *Z. Chem.* **1982**, *22*, 260.
14. a) Holík M.: *Collect. Czech. Chem. Commun.* **1996**, *61*, 713; b) Holík M.: *J. Mol. Struct.* **1999**, *482–3*, 345.
15. Waisser K., Kubicová L., Klimešová V., Odlerová Ž.: *Collect. Czech. Chem. Commun.* **1993**, *58*, 2977.
16. Wagner G., Singer D., Weuffen W.: *Pharmazie* **1966**, *21*, 161.
17. Einhorn A., von Bagh A.: *Chem. Ber.* **1910**, *43*, 322.
18. a) Grandmougin E., Lang A.: *Chem. Ber.* **1909**, *42*, 4014; b) Forbes E. J., Stacey M., Tatlow J. C., Wragg R. T.: *Tetrahedron* **1960**, *8*, 67; c) Bogert T. M., Wise L. E.: *J. Am. Chem. Soc.* **1912**, *34*, 693; d) Yamamoto T., Kurata Y.: *Can. J. Chem.* **1983**, *61*, 86; e) Gabriel S.: *Chem. Ber.* **1878**, *11*, 2260.
19. Levy G. C., Cargioli J. D.: *J. Magn. Reson.* **1972**, *6*, 143; Günther H.: *NMR Spectroscopy, Basic Principles, Concepts, and Applications in Chemistry*, 2nd ed., p. 532. Wiley, New York 1995.
20. Holík M., Friedl Z., Marchalín Š.: *Collect. Czech. Chem. Commun.* **1999**, *64*, 1696.
21. Dewar M. J. S., Zoebisch E. G., Healy E. F., Steward J. J. P.: *J. Am. Chem. Soc.* **1985**, *107*, 3902.
22. Holík M.: *Magn. Reson. Chem.* **1992**, *30*, 189.
23. Hansch C., Leo A., Taft R. W.: *Chem. Rev.* **1991**, *91*, 165.