

# Revision of the dual substituent parameter treatment using the DFT-calculated reaction energies

Otto Exner<sup>1</sup> and Stanislav Böhm<sup>2\*</sup>

<sup>1</sup>Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 166 10 Praha 6, Czech Republic

<sup>2</sup>Department of Organic Chemistry, Institute of Chemical Technology, 166 28 Praha 6, Czech Republic

Received 14 December 2006; revised 14 February 2007; accepted 20 February 2007



**ABSTRACT:** The dual substituent parameter (DSP) data treatment is a broadly used procedure correlating the reaction energies or other physical quantities with two sets of substituent constants, inductive ( $\sigma_I$ ) and resonance ( $\sigma_R$ ). It was here revised using the most extensive sets of experimental reactivities available in the literature and two sets of reaction energies calculated at the level B3LYP/6-311+G(d,p): acidities of 4-substituted benzoic acids and 5-(*E*)-substituted penta-2,4-diene-(*E*)-acids with 19 or 15 common substituents. The latter two series enabled us to investigate the substituent effects more systematically than it was ever possible with the experimental data; this means in particular separate treatment of the undissociated acid molecules and of their anions, further separation of donor and acceptor substituents. In addition, the standard statistical treatment was improved when testing the significance of the resonance term.

The DSP treatment is not valid generally, this applies both to the standard reference series and to the series commonly investigated. At best, DSP may be considered to hold for donor substituents but the effects of acceptors are much less variable and do not depend on the constants  $\sigma_R$  nor on any other measure of resonance. The small efficiency of acceptor substituents is due by the fact that the constant functional group (COOH in the standard series) is itself an acceptor. A correct treatment would be to investigate the donor and acceptor substituents separately, donors with an acceptor functional groups, and vice versa; substituents with weak resonance effect should be not included. The popularity and apparent success of the DSP treatment can be attributed to several grounds, most important has been the unbalanced choice of substituents. Copyright © 2007 John Wiley & Sons, Ltd.

Supplementary electronic material for this paper is available in Wiley InterScience at <http://www.interscience.wiley.com/jpages/0894-3230/suppmat>

**KEYWORDS:** benzoic acids; density functional theory; inductive effect; isodesmic reaction; pentadienoic acids; resonance; substituent effect

## INTRODUCTION

The most popular theory of substituent effects, the dual substituent parameter (DSP) treatment,<sup>1–5</sup> assumes that each substituent can be characterized by two independent parameters, the inductive constant  $\sigma_I$  and resonance constant  $\sigma_R$ . A physical property  $Y$  of a substituted compound can then be related to the property  $Y^\circ$  of the unsubstituted compound by means of Eqn (1). The values of  $\sigma_I$  and  $\sigma_R$  have been estimated by rather sophisticated procedures from model reactions;<sup>1–3</sup> once determined they are treated as constants. The proportionality constants  $\rho_I$  and  $\rho_R$  are obtained from multiple linear regression.

$$Y - Y^\circ = \rho_I \sigma_I + \rho_R \sigma_R \quad (1)$$

\*Correspondence to: S. Böhm, Department of Organic Chemistry, Institute of Chemical Technology, 166 28 Praha 6, Czech Republic.  
E-mail: bohm@vscht.cz

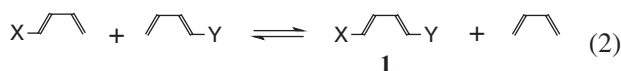
In practice,  $Y$  was most often the reaction Gibbs energy, usually the values of  $pK$  or  $\log k$  were used directly. Of the physical properties, most attention was given to the NMR shifts of  $^{13}\text{C}$ <sup>6</sup> and  $^{19}\text{F}$ .<sup>2</sup> The popularity of the approach may be seen from the number of citations of the fundamental papers,<sup>1,3</sup> 840 and 620, respectively, but the number of actual applications is certainly several times greater.

The fundamental assumption in Eqn (1) is the proportionality of substituent effects:<sup>7</sup> When one substituent acts in a reaction more strongly than another, this must be valid for any other reaction, separately in the term with  $\sigma_I$  and with  $\sigma_R$ . However, it has been known for a long time that for many reactions the resonance of some substituents (either donors or acceptors) was not proportional. The problem was formally solved<sup>1–4</sup> by means of alternative sets of constants  $\sigma_R$  (dual resonance constants), differentiated by a superscript:  $\sigma_R^\circ$ ,  $\sigma_R^{\text{Bz}}$ ,  $\sigma_R^+$ , or  $\sigma_R^-$ . These sets were chosen in the individual reactions

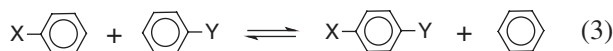
according to the strength of resonance as the case might be. This arbitrariness was seen as an extension of the DSP principle<sup>1–4</sup> but it is merely its breakdown;<sup>8</sup> a possible alternative is extending Eqn (1) with an additional term.<sup>9</sup> In this paper, we shall deal with the simple Eqn (1) with customary constants  $\sigma_R$ .

Further objections against the DSP theory were made from the point of view of general statistics:<sup>10</sup> Significance of the second term was rarely statistically proven and the regression coefficients  $\rho_I$  and  $\rho_R$  were strongly mutually dependent. More specifically, the  $\sigma_R$  values of acceptor substituents were always small and possibly should be equal to zero.<sup>11</sup> This problem was discussed heavily but an agreement was not reached.<sup>12–13</sup> Important contributions were obtained recently from quantum-chemically calculated reaction energies using the principle of isodesmic<sup>14</sup> (and also homodesmic<sup>15</sup>) reactions. In this way interaction energy of both charged and uncharged groups can be determined; in the case of acidobasic reactions one can investigate separately the substituent effect in the ion and in the uncharged molecule of the acid or base.<sup>14</sup> In addition, more systematic data can be obtained than available from the experiments in solution. Numerous results<sup>16–19</sup> confirmed that the inductive effect is proportional in various reactions and even with various physical properties;<sup>2,20</sup> it is expressed simply by the first term of Eqn (1) in all cases when at least one of the interacting groups is charged or strongly polar.

On the other hand, proportionality of the resonance effect has never been proven and there are grounds to believe that it is only a very rough approximation if it exists at all. A model recently suggested referred to 1,4-disubstituted butadienes **1** with the *sp* conformation on the central bond. The reaction energy  $\Delta_2E$  of the isodesmic reaction, Eqn (2) was negative and proportional when X was a donor group and Y an acceptor.<sup>21</sup>

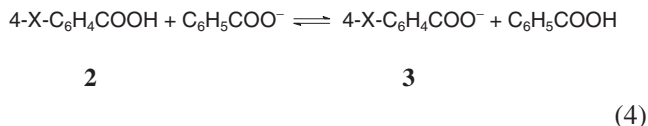


However, when X and Y were two donors or two acceptors, their interaction was destabilizing and not proportional for various groups. Therefore, it was possible to create a scale of the donor ability and another scale of the acceptor ability, but not a uniform scale for all groups. Similar results were obtained<sup>21</sup> also for the reaction energy  $\Delta_3E$  of Eqn (3) but this equation was considered to be less suitable due to crossed conjugation. Note that evaluation of the resonance ability made on a different principle leads also to incoherent scales.<sup>22</sup>

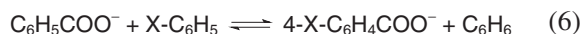
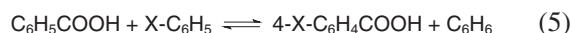


At present DSP treatment seems to be generally accepted since the latest reviews of the correlation analysis<sup>23</sup> do not give particular attention to it. In our opinion, however, it is necessary to enlighten the evident discrepancy between our new results<sup>21,22</sup> and commonly

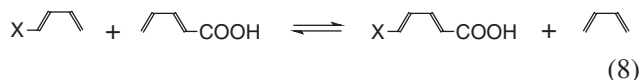
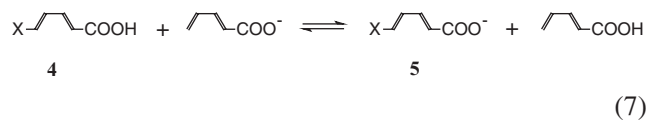
acknowledged DSP. The main problem might be in the charges. Previous investigation of butadiene derivatives<sup>21</sup> **1** was restricted to uncharged groups X and Y,<sup>21</sup> while the DSP theory has dealt mostly with the acidobasic properties; very important was dissociation of 4-substituted benzoic acids **2**, Eqn (4), used even for the definition<sup>1–3</sup> of constants  $\sigma_R$ .



To bridge over the inconsistency between the two concepts, we investigated sets of charged molecules containing the standard group COOH. Reaction energies of the reaction of Eqn (4), most important in the DSP analysis, were obtained within the framework of the density functional theory<sup>24</sup> (DFT) at the level well-tried in the previous work.<sup>18,19,25</sup> Calculations enabled us to include more substituents than in the experimental work,<sup>3,4</sup> particular attention was given to some uncommon substituents. For this purpose it was essential that this theoretical model yielded good agreement with the gas-phase experiments in the case of benzoic acids and common substituents.<sup>25</sup> We were not engaged in the question whether another model (MP2) is not more effective for compounds of this type; this was partly dealt with on another place.<sup>26,27</sup> Further progress was expected from dividing the energy  $\Delta_4E$  of Eqn (4) into two parts: energy  $\Delta_5E$  expressing the substituent effect in the undissociated acid, Eqn (5), and  $\Delta_6E$  expressing the effect in the anion, Eqn (6). This separation was used several times<sup>25</sup> and yielded important information about the nature of substituent effects.<sup>18,25–28</sup>



Last but not least, the common statistical treatment as defined by Eqn (1),<sup>1,3</sup> was extended with particular attention to the second, resonance term in Eqn (1) and its statistical significance. In addition to the benzene derivatives, Eqns (4)–(6), we investigated in the same way butadiene derivatives **4**; that is, their acidities, defined by Eqn (7) and separation of the substituent effects, Eqns (8) and (9).



All the results were consistent. The most important improvement appeared to be the extended number of substituents with sufficiently represented acceptor groups.

## RESULTS AND DISCUSSION

### DSP treatment of the acidity of benzoic acids

Of the two systems investigated here, the acidities of 4-substituted benzoic acids **2**, Eqn (4), will be analyzed in all details since they represent the classic reaction, many times used as the basis reference. The second system, 5-substituted pentadienoic acids **4**, Eqn (7), yielded very similar results and will be treated in briefly subsequently.

The DFT-calculated relative acidities  $\Delta_4 E^\circ$  of **2** are listed in table 1, column 2. They will be treated by the traditional DSP correlation as if they would be results of any experiment. This approach is justified since close correlation of calculated and experimental gas-phase acidities<sup>29</sup> was proved.<sup>4,25,26</sup> From many reported scales of the constants  $\sigma_I$  and  $\sigma_R$  those of Charton<sup>3</sup> were used, probably the most reliable of the available values. The DSP correlation so obtained is quite satisfactorily according to simple statistics (Table 2, entry 1). It is very similar to the reported<sup>4</sup> correlation of experimental  $\Delta_4 G^\circ(298)_g$  of the same reaction, Eqn (4), reproduced in Table 2, entry 2. Nevertheless, there is a difference in the regression coefficients  $\rho_R$  suggesting that this value may be sensitive to the selection of substituents.

According to the common view, these two DSP correlations are successful; in the standard scale<sup>30</sup> the fit would be evaluated as 'satisfactory' or 'excellent',

respectively. Of course, this result was expected since the constants  $\sigma_R$  were derived<sup>3</sup> just from the reaction Eqn (4) (in water solution); proportionality of the calculated acidities and gas-phase experiments was proven<sup>25</sup> and proportionality of gas-phase and water acidities is well known.<sup>29</sup> In the following, more detailed analysis leads to different conclusions, it is because it exceeded the customary treatment in three points:

1. The substituents effects were divided into the effects  $\Delta_5 E$  operating in the undissociated acids, Eqn (5), and  $\Delta_6 E$  operating in the anions, Eqn (6).
2. The selection of substituents was broader and more systematic than in common experimental studies; the number of acceptor substituents was equal to the number of donors.
3. The commonly used statistics were completed by testing the significance of the second, resonance term: standard *F*-test and special graphic test shown later.

The *F*-test itself did not reveal any shortcoming: the  $\sigma_R$  term was significant at the very high significance level as it is seldom encountered in chemical applications (Table 2, last column, entry 1 and further). As one of the main problems of multiple regression it was pointed out to the multicollinearity<sup>31</sup> (or intercorrelation<sup>9a</sup>), that is mutual statistical dependence of the explanatory variables. This is no problem here since there is no correlation of  $\sigma_R$  with  $\sigma_I$  (Table 2, entry 3).

Division of substituent effects is presented in the entries 6 and 9 of Table 2. As expected the effects are opposite: the acids **2** are destabilized by electron-attracting substituents while their anions **3** are stabilized. The latter effect is several times greater and both reinforce each other in controlling the acidity. There

**Table 1.** DFT-calculated reaction energies of the isodesmic reactions, Eqns (4)–(9) (in  $\text{kJ mol}^{-1}$ )

Substituent	Equation (4)	Equation (5)	Equation (6)	Equation (7)	Equation (8)	Equation (9)
CH <sub>3</sub>	5.51	-2.47	3.04	7.45	-3.42	4.02
CH <sub>2</sub> Cl	-18.3	1.82	-16.48	-19.54	1.72	-17.82
NH <sub>2</sub>	18.52	-10.29	8.23	28.45	-13.58	14.87
N(CH <sub>3</sub> ) <sub>2</sub>	22.63	-12.62	10.01	32.92	-16.34	16.58
OH	4.40	-5.43	-1.03	9.48	-6.52	2.96
OCH <sub>3</sub>	8.57	-6.15	2.42			
SH	-7.35	-2.75	-10.10			
F	-13.60	0.02	-13.58	-11.47	0.61	-10.86
Cl	-19.73	1.35	-18.38	-17.46	1.55	-15.91
H	0	0	0	0	0	0
CN	-46.72	6.58	-40.14	-50.62	7.36	-43.26
CHO	-36.17	6.17	-30.00	-47.16	6.77	-40.39
COOH	-28.71	5.21	-23.50	-37.99	5.83	-32.16
COOCH <sub>3</sub>	-22.75	4.33	-18.42			
CF <sub>3</sub>	-35.95	6.01	-29.94	-39.80	7.91	-31.89
CCl <sub>3</sub>	-34.48	4.48	-30.02	-36.67	4.87	-31.80
NO <sub>2</sub>	-54.49	8.61	-45.88	-68.20	10.1	-58.1
SO <sub>2</sub> CH <sub>3</sub>	-45.94	7.43	-38.51			
SO <sub>2</sub> CF <sub>3</sub>	-65.05	10.21	-54.84	-76.17	11.40	-64.77

**Table 2.** Statistics of the correlations of reaction energies with substituent constants

Response function <sup>a</sup>	Explanatory variables		Regression coefficients <sup>b</sup>		$R^c$	$s^c$	$N^c$	$\alpha^d$
	$\sigma_I$	$\sigma_R$						
1 $\Delta_4 E$ acidity	$\sigma_I$	$\sigma_R$	-66.5 (52)	-45.5 (31)	0.987	4.2	19	<<0.001
2 $\Delta_4 G^\circ$ exp. <sup>e</sup>	$\sigma_I$	$\sigma_R$	-62.8 (25)	-61.1 (29)	0.9972	1.7	11	
3 $\sigma_R$	$\sigma_I$		0.66 (38)		0.392	0.33	19	
4 $\Delta_4 E$ acidity <b>D</b>	$\sigma_I$	$\sigma_R$	-57.8 (90)	-41.8 (63)	0.960	4.6	10	<<0.001
5 $\Delta_4 E$ acidity <b>A</b>	$\sigma_I$	$\sigma_R$	-70.3 (73)	-71 (29)	0.984	3.6	10	0.05
6 $\Delta_5 E$ acids	$\sigma_I$	$\sigma_R$	10.2 (10)	14.9 (6)	0.9934	0.78	19	<<0.001
7 $\Delta_5 E$ acids <b>D</b>	$\sigma_I$	$\sigma_R$	12.8 (16)	16.3 (9)	0.989	0.82	10	<<0.001
8 $\Delta_5 E$ acids <b>A</b>	$\sigma_I$	$\sigma_R$	9.3 (10)	17.7 (38)	0.953	0.47	10	0.001
9 $\Delta_6 E$ anions	$\sigma_I$	$\sigma_R$	-56.3 (51)	-30.6 (30)	0.979	4.1	19	<<0.001
10 $\Delta_6 E$ anions <b>D</b>	$\sigma_I$	$\sigma_R$	-45.0 (82)	-25.5 (48)	0.932	4.2	10	<0.001
11 $\Delta_6 E$ anions <b>A</b>	$\sigma_I$	$\sigma_R$	-61.0 (69)	-53 (28)	0.981	3.4	10	n.s.
12 $\Delta_5 E$ acids	$\Delta_6$	$E$	-0.56 (7)		0.878	2.2	19	
	anions							
13 $\Delta_4 G^\circ$ sol exp. <sup>f</sup>	$\sigma_I$	$\sigma_R$	-10.1 (6)	-7.25 (36)	0.9934	0.43	18	<<0.001
14 $\Delta_4 G^\circ$ sol exp. <b>D</b> <sup>f</sup>	$\sigma_I$	$\sigma_R$	-10.8 (14)	-7.81 (67)	0.973	0.48	13	<<0.001
15 $\Delta_4 G^\circ$ sol exp. <b>A</b> <sup>f</sup>	$\sigma_I$	$\sigma_R$	-10.4 (3)	-2.5 (11)	0.9991	0.14	6	n.s.
16 $\Delta_7 E$ acidity	$\sigma_I$	$\sigma_R$	-76.0 (69)	-61.9 (43)	0.989	5.4	15	<0.001
17 $\Delta_8 E$ acidity <b>D</b>	$\sigma_I$	$\sigma_R$	-63 (10)	-53.4 (60)	0.976	5.2	8	0.001
18 $\Delta_9 E$ acidity <b>A</b>	$\sigma_I$	$\sigma_R$	-74.3 (90)	-122 (34)	0.989	4.1	8	0.025
19 $\Delta_4 E$ acidity <b>D</b>	$\sigma_I$	$\sigma_D$	-66.4 (74)	-36.7 (39)	0.983	3.0	10	<<0.001
20 $\Delta_4 E$ acidity <b>A</b>	$\sigma_I$	$\sigma_A$	-59.5 (48)	-23.8 (63)	0.9957	1.9	10	0.01
21 $\Delta_5 E$ acids <b>D</b>	$\sigma_I$	$\sigma_D$	11.3 (16)	16.7 (8)	0.993	0.66	10	<<0.001
22 $\Delta_5 E$ acids <b>A</b>	$\sigma_I$	$\sigma_A$	5.9 (13)	7.8 (18)	0.985	0.53	10	0.005
23 $\Delta_6 E$ anions <b>D</b>	$\sigma_I$	$\sigma_D$	-55.1 (64)	-20.0 (34)	0.974	2.6	10	0.001
24 $\Delta_6 E$ anions <b>A</b>	$\sigma_I$	$\sigma_A$	-53.6 (40)	-16.0 (53)	0.9958	1.6	10	0.025
25 $\log k^g$	$\sigma_I$	$\sigma_R$	0.88 (3)	0.80 (3)	0.9979	0.027	11	<<0.001
26 $\log k$ <b>D</b> <sup>g</sup>	$\sigma_I$	$\sigma_R$	0.91 (2)	0.79 (2)	0.9995	0.012	7	<<0.001
27 $\log k$ <b>A</b> <sup>g</sup>	$\sigma_I$	$\sigma_R$	0.74 (9)	1.60 (52)	0.9978	0.032	5	0.1
28 $\Delta G^\circ$ <sup>h</sup>	$\sigma_I$	$\sigma_R$	-49 (14)	-63 (11)	0.940	9.3	14	<<0.001
29 $\Delta G^\circ$ <b>D</b> <sup>h</sup>	$\sigma_I$	$\sigma_R$	-52.0 (63)	-53.6 (50)	0.989	2.9	7	<<0.001
30 $\Delta G^\circ$ <b>A</b> <sup>h</sup>	$\sigma_I$	$\sigma_R$	-43 (22)	-135 (78)	0.786	12.3	8	0.05

<sup>a</sup> **D** denotes only donor substituents, **A** acceptor substituents.

<sup>b</sup> Standard deviation in parentheses.

<sup>c</sup>  $R$ , Correlation coefficient;  $s$ , standard deviation from the regression;  $N$ , number of items.

<sup>d</sup> Significance level of the second term,  $F$ -test comparing with a single-parameter equation containing only the  $\sigma_I$  term.

<sup>e</sup> Both the experiments and calculation taken from Ref. 4.

<sup>f</sup> Experimental  $pK$  in 50% (vol) ethanol taken from Ref. 3, the values of  $\sigma_I$  and  $\sigma_R$  from the same source.

<sup>g</sup> Rates of reaction of 4-substituted benzoic acids with diphenyldiazomethane in acetone were taken from Ref. 1, the values of  $\sigma_I$  and  $\sigma_R$  from the same source.

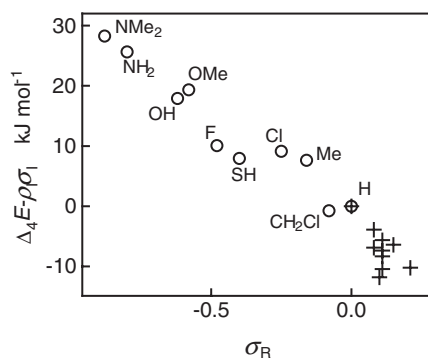
<sup>h</sup> Experimental gas-phase acidities of 4-substituted  $N,N$ -dimethylanilines, Ref. 4; the values of  $\sigma_I$  and  $\sigma_R$  from the same source.

is no proportionality of  $\Delta_5 E$  and  $\Delta_6 E$  (entry 12) as observed<sup>25</sup> for the Hammett set of *meta* and *para*-substituted benzoic acids; in that case it was caused essentially by *meta* substituents.

Significant progress was now achieved when the substituents were divided into donors and acceptors: the latter appeared to be not controlled by  $\sigma_R$ . The acidities  $\Delta_4 E$  of acceptors may be still considered as dependent on  $\sigma_R$  although with a much lower significance than the donors. (Compare the entries 5 and 4 of Table 2.) This dependence is due to the substituent effect in the undissociated acid (entry 8) while the much stronger effects in the anions are not related to the  $\sigma_R$  constants (entry 11).

The statistical tests do not allow easy survey but can be made very objective by the following graphic representation, which was used recently<sup>19</sup> for testing the

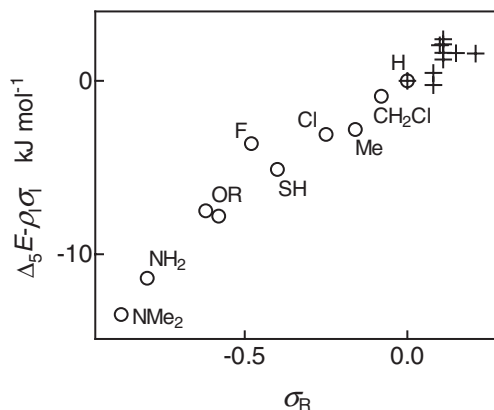
multiple regressions. When the regression has been carried out according to Eqn (1), the first term on the right side is transferred to the left side and the values of  $\Delta_4 E - \rho_I \sigma_I$  are plotted *versus*  $\sigma_R$  to show whether there is an actual dependence on the second explanatory variable. In other words,  $\Delta_4 E$  was correlated step-wise, first with  $\sigma_I$ , then with  $\sigma_R$ . This test was carried out together for all substituents but separately for the acidities  $\Delta_4 E$  (Fig. 1), for the energies of undissociated acids  $\Delta_5 E$  (Fig. 2), and of the anions  $\Delta_6 E$  (Fig. 3). The fundamental difference between donors and acceptors is seen in all graphs. The donors form a sequence with conjugation varying from strong to very weak. The acceptors form a cluster and their substituent effects only have nothing in common with  $\sigma_R$ . Most evident it is with the anions (Fig. 3); with the acidity this effect is weakened (Fig. 1).



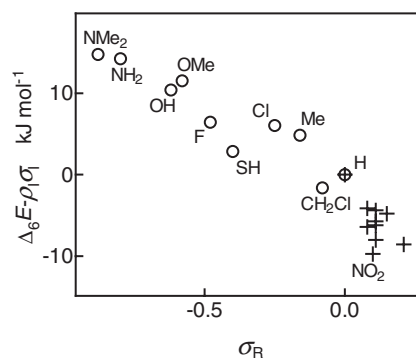
**Figure 1.** DSP treatment of the relative acidities of 4-substituted benzoic acids, Eqn (4), specific test of the resonance term: the reaction energies  $\Delta_4 E$  with the inductive term subtracted have been plotted *versus*  $\sigma_R$ ; ○ donor substituents, + acceptor substituents

Let us stress that these results are not artifacts of the quantum chemical approach nor are they restricted to the gas-phase acidities since essentially the same results were observed for the same reaction, Eqn (4), in solution (50% vol. ethanol), only the selection of substituents was poorer and less systematic.<sup>3</sup> Failure of DSP is remarkable since the values  $\sigma_R$  were determined just from this reaction<sup>3</sup> (in water). The statistics are given in Table 2, entries 13–15. Note particularly the high correlation coefficient for donors (entry 15), which is due exclusively to the inductive component since the resonance is insignificant. In light of such results it appears as a good approximation to use the values  $\sigma_R = 0$  for all acceptors,<sup>11a</sup> or at least in certain reactions.<sup>4</sup> The graphical test is very similar to Fig. 1 and is reproduced in the Supplementary Information (Fig. S1).

The above test seems to be rather sensitive and objective but it is only a graphical picture, not a rigorous statistical procedure. We considered necessary to check it on a couple of examples, partly purely synthetic, partly



**Figure 2.** DSP treatment of the substituent effects in 4-substituted benzoic acids, Eqn (5), specific test of the resonance term: the reaction energies  $\Delta_5 E$  with the inductive term subtracted have been plotted *versus*  $\sigma_R$ ; ○ donor substituents, + acceptor substituents

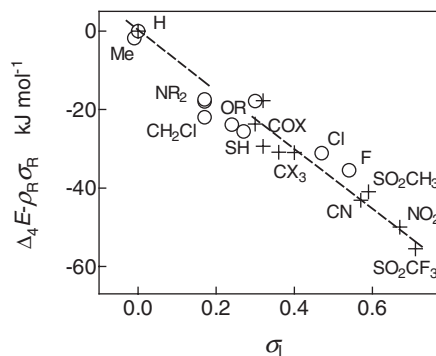


**Figure 3.** DSP treatment of the substituent effects in the anions of 4-substituted benzoic acids, Eqn (6), specific test of the resonance term: the reaction energies  $\Delta_6 E$  with the inductive term subtracted have been plotted *versus*  $\sigma_R$ ; ○ donor substituents, + acceptor substituents

resembling the actual DSP correlations as to the values of  $\sigma_1$  and  $\sigma_R$ . It is good for testing the less important of the two explanatory variables but can yield false results when the dependence on the first explanatory variable is itself very weak. We show as an example the reverse test of our data set of  $\Delta_4 E$ : In this case the dependence on  $\sigma_R$  is still sufficient and when one assumes that it holds, one can test the dependence on  $\sigma_1$ . One gets Fig. 4. There is no disparity of acceptors and donors and the importance of  $\sigma_1$  is evident. The striking difference compared to Fig. 1 is fully convincing: while  $\sigma_1$  is significant both for donors and acceptors,  $\sigma_R$  only for donors.

### Acidity of pentadienoic acids

This second set of compounds **4** was investigated on a smaller number of substituents (Table 1) and in less detail since the results were essentially the same. The statistics given in Table 2, entries 16–18, are similar to those



**Figure 4.** DSP treatment of the substituent effects on the acidity of 4-substituted benzoic acids, Eqn (4), a reverse test of the significance of the inductive term: the reaction energies  $\Delta_4 E$  with the resonance term subtracted have been plotted *versus*  $\sigma_1$ ; ○ donor substituents, + acceptor substituents; the dotted line is only a reference

obtained for benzoic acids **2** (entries 1, 4, and 5). The main difference is that all proportionality factors  $\rho$  are greater (perhaps due to the crossed resonance in **2**). The most important result is the same from both series: The resonance term is insignificant when the correlation is restricted to acceptors (entries 18 and 5). This is seen also in the graphical test; the pertinent graph is very similar to Fig. 1 and is shown only in the Supporting Information (Fig. S2). The fundamental difference between acceptors and donors is confirmed.

### Better scales of resonance

The above results showed that the success of DSP is limited because the effect of acceptor substituents is not properly expressed by the constants  $\sigma_R$ . We tried to improve this procedure by introducing a better scale of resonance, Eqn (10).

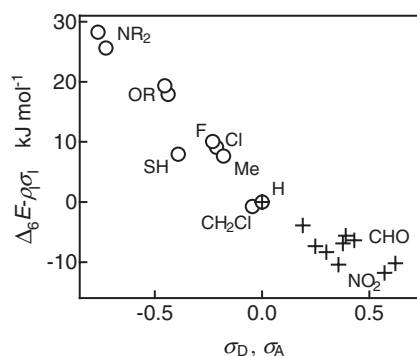
$$\Delta_4 E = \rho_I \sigma_I + \rho_{A,D} \sigma_{A,D} \quad (10)$$

The constants  $\sigma_I$  were taken here from the gas-phase values<sup>19</sup> but the difference compared to standard values<sup>3</sup> is not significant. The symbol  $\sigma_{A,D}$  denotes either the constant  $\sigma_D$  for donors or  $\sigma_A$  for acceptors as the case may be, similarly  $\rho_{A,D}$ . The values of  $\sigma_D$  and  $\sigma_A$ , unbiased measure of conjugation, were derived on the basis of Eqn (2) when X is a donor substituent and Y an acceptor, both uncharged.<sup>21</sup> They are not valid universally but rather in limiting situations:  $\sigma_D$  of a donor when it is conjugated with an acceptor,  $\sigma_A$  of an acceptor when it is conjugated with a (strong) donor. Interaction of two donors or of two acceptors, or interactions of weak substituents cannot be described in this way; acceptors and donors must be correlated separately and the scales of  $\sigma_D$  and of  $\sigma_A$  cannot be merged.<sup>21</sup>

Some necessary values of  $\sigma_D$  and  $\sigma_A$  were not available in the original work<sup>21</sup> and were now calculated on the basis of butadiene derivatives, Eqn (2): When Y = NO<sub>2</sub> one can estimate  $\sigma_D$  of various donors X, with Y = NH<sub>2</sub> then  $\sigma_A$  of various acceptors. They are defined<sup>21</sup> by Eqn (11) that includes still a correction for the inductive component,  $\Delta_{\text{ind}}E$ , determined from 1,4-disubstituted bicyclo[2.2.2]octanes. The DFT-calculated energies  $\Delta_2 E$  and resulting new values of  $\sigma_D$  and  $\sigma_A$  are given in Table S2 (Supporting Information), derivation of  $\Delta_{\text{ind}}E$  is described in Table S3.

$$\Delta_2 E - 1.3 \Delta_{\text{ind}}E = 57.71 \sigma_D \sigma_A + 0.028 \quad (11)$$

The results of correlation with Eqn (10) are given in Table 2, entries 19 and 20, separate correlations of the acid molecules and anions in the entries 21–24. At the first sight, there is no spectacular difference compared to the correlations with  $\sigma_R$ . Significant is the smaller dependence on resonance ( $\rho_A$ ) for the acceptors because the scale of  $\sigma_A$  is extended compared to  $\sigma_R$ . In spite of this,



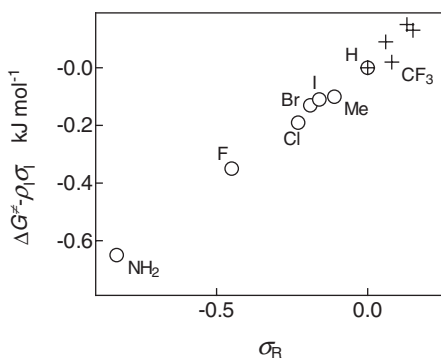
**Figure 5.** Modified DSP treatment of the substituent effects on the acidity of 4-substituted benzoic acids, Eqn (4), specific test of the resonance term: the reaction energies  $\Delta_4 E$  with the inductive term subtracted have been plotted versus the actual resonance constants  $\sigma_D$  or  $\sigma_A$ ;  $\circ$  donor substituents,  $+$  acceptor substituents

the significance level of the resonance term has only little improved, in particular the energy of the anions depends little on  $\sigma_R$  of acceptors (entry 24). The graphical test presented in Fig. 5 looks at the first sight slightly better than Fig. 1 but the great difference between donors and acceptors persists. The tests for the acid molecules and anions are given in Figs. S3 and S4 (Supporting Information), particularly in the former case the dependence on  $\sigma_A$  is poor.

We conclude that replacing the constants  $\sigma_R$  by  $\sigma_D$  and  $\sigma_A$  produces comparable scales for donors and acceptors but the correlations are not much improved. The main problem is that the functional group COOH is itself an acceptor. Wepster<sup>13b</sup> expressed the difference between donors and acceptors by the famous words: ‘Nature has endowed us with a variation of donors whereas the common acceptors form a cluster-like group with less discriminating abilities’. This is true only partly. The variability of acceptors is certainly smaller but it has been made still smaller when evaluated on the basis of ionization of carboxylic acids, that is, on interaction with another acceptor. The unbiased correlation of resonance would need separate correlations of donor substituents with an acceptor functional group and of acceptors with a donor group, using preferably only uncharged functional groups.

### Re-examination of the reported DSP correlations

This section will bring relatively little new information since a systematic re-examination of the published data along the lines of the preceding paragraphs is not possible: There are very few data available with a sufficient number of acceptor substituents. Of 36 reactions correlated by Taft<sup>1</sup> there are 15 acidities of 4-substituted benzoic acids at various conditions; of the

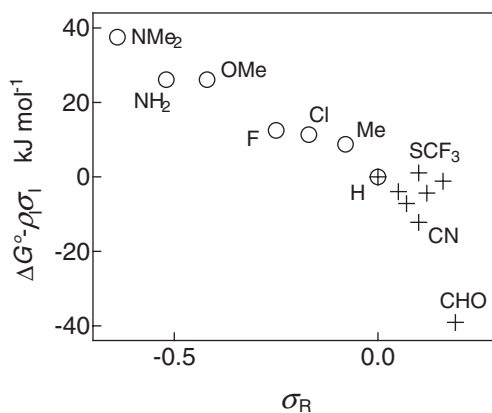


**Figure 6.** DSP treatment of the substituent effects in the kinetics of the reaction of 4-substituted benzoic acids with diphenyldiazomethane in acetone, specific test of the resonance term: the activation Gibbs energies with the inductive term subtracted have been plotted *versus* the resonance constants  $\sigma_R$ ; ○ donor substituents, + acceptor substituents

total number of 326 substituents there are only 72 acceptors. The *meta* series were correlated separately with the constants  $\sigma_R^\ominus$ . The best example one could re-investigate is from kinetics: reaction of 4-substituted benzoic acids with diphenyldiazomethane in acetone. With the values of  $\log k$  and of constants  $\sigma$  as they have been given<sup>1</sup> we got the statistics in Table 2, entries 25–27, and Fig. 6. The dependence on  $\sigma_R$  is insignificant for the acceptor substituents as in the preceding examples.

Of 38 *para* reaction series correlated by Charton,<sup>3</sup> 26 dealt with the acidity of 4-substituted benzoic acids; of all 325 substituents 66 were acceptors, of them the nitro group appeared 36 times. The 23 *meta* series dealt exclusively with the acidity of 3-substituted benzoic acids. The most complete series was the acidity of 4-substituted benzoic acids in 50% (vol) ethanol, dealt with in Fig. S1 (Supporting Information).

Some good series of data are available among the gas-phase acidities and basicities.<sup>4</sup> The correlations



**Figure 7.** DSP treatment of the experimental basicity of 4-substituted *N,N*-dimethylanilines in the gas-phase, specific test of the resonance term: the reaction Gibbs energies with the inductive term subtracted have been plotted *versus* the resonance constants  $\sigma_R$ ; ○ donor substituents, + acceptor substituents

were extended by inserting a third term into Eqn (1) but this appeared in many cases to be insignificant. We re-investigated the basicity of 4-substituted *N,N*-dimethylanilines using the special constants  $\sigma_I$  and  $\sigma_R$  designed for the gas-phase (Fig. 7).<sup>4</sup> Their use eliminates the application of constants  $\sigma_R^-$ , which would be usually recommended for substituted anilines. The statistics in Table 2, entries 28 to 30, gave the expected result. The reactivities of the acceptor groups are in this case relatively variable but they are not well expressed by  $\sigma_R$ .

## CONCLUSIONS

The DSP relationship is not generally valid for diverse reactions, it is not valid even for the reaction on which it has been defined. The main defect is in the interaction of acceptor substituents with the ionic functional group, which is not dependent on  $\sigma_R$  or on any measure of resonance. When there are few acceptor substituents present, this defect may become less evident. The apparent success of DSP reported so many times in the literature may be thus attributed to several reasons: (1) biased selection of substituents with few acceptors if any, (2) oversimplified statistical treatment of small samples, (3) monotonous selection of reaction series.

The idea that reactivity can be expressed by a combination of inductive effect and resonance is not principally wrong. Reaction series controlled by inductive effect is predictable and proportional in different reactions while resonance may be variable or even absent. It is certainly possible to find some reactions with regularly variable resonance; they should be searched particularly among non-ionic reactions and treated separately for donor and acceptor substituents. The dissociation reactions of carboxylic acids, both in solution and in the gas-phase, are not suitable as standard reference reactions; they were used only because they were experimentally accessible.

Other two-term equations may be fallacious when applied to small samples with biased choice of items. In particular certain equations<sup>32</sup> similar to DSP can be certainly subjected to similar criticism as above.

## CALCULATIONS

The DFT calculations were performed according to the original proposal<sup>24</sup> at the level B3LYP/6-311+G(d,p)//B3LYP/6-311+G(d,p) exploiting the GAUSSIAN 03 program.<sup>33</sup> Full energy optimization and vibrational analyses were carried out in all cases. The minimum-energy conformations were searched for starting from different initial structures but two stable conformations were found only in rare cases (Table S2, footnotes) and

their energies differed only negligibly. No correction for the zero-point energy was introduced.

The energies of the acids **2** and **4** and of their anions **3** and **5** are listed in Table S1 (Supporting Information), energies of some auxiliary compounds in the Tables S2 and S3.

## Acknowledgements

This work was carried within the framework of the research project Z4 055 0506 of the Academy of Sciences of the Czech Republic.

## APPENDIX

### General remarks from the point of view of statistics

The DSP analysis and similar treatments are based on the statistical disciplines called theory of estimation and testing of hypotheses. A series of subjects is defined (population) of which a random sample is taken. From the properties of the sample and the relationships observed in it, the methods of statistical induction give estimates of the properties and relationships existing in the population with a certain probability since there is always a possibility (significance level) that the results have been obtained by chance. In the application in chemistry, already the definition of the population is often not clear.<sup>34</sup> Such definitions as ‘all molecules, stable or unstable’, or ‘all stable molecules with less than  $n$  atoms’, or ‘all molecules described in the literature’ have no clear meaning; they do not decide exactly, which molecule should be included and which not. Evidently the population must be defined as a fuzzy set,<sup>35</sup> a set to which all its elements belong each with a given weight. This weight determines the probability that the element is selected into a sample; the results obtained are valid for the population but can be applied for an individual element only with the proper weight.

In the fuzzy set of all substituents, its elements have obviously very different weights. Small and stable substituents are strongly preferred but other properties may be also of importance. Thus compounds important in practice, commercially available, or easy to prepare will be always favored, similarly compounds with some favorable properties, for instance solubility. Other conceptions can prefer compounds several times treated in the literature (tradition)<sup>36</sup>, or on the contrary those with interesting properties not known from the literature. In practice, selection of a sample is never random but the non-random procedure can be replaced by a fictive random selection from a fuzzy set with different weights.

In the case of the DSP treatment the results are strongly dependent on the definition of the population as a fuzzy

set. If it is defined to cover many known structures with little respect to the importance of the compounds and to the literature tradition, the DSP is not valid as shown conclusively in this paper. If the well-known, often treated conventional compounds are preferred (the sets of recommended substituents<sup>1</sup>), DSP may hold in this very restricted range of validity.

## REFERENCES

- Ehrenson S, Brownlee RTC, Taft RW. *Prog. Phys. Org. Chem.* 1973; **10**: 1–80.
- Hansch C, Leo A, Taft RW. *Chem. Rev.* 1991; **91**: 165–195.
- Charton M. *Prog. Phys. Org. Chem.* 1981; **13**: 119–251.
- Taft RW, Topsom RD. *Prog. Phys. Org. Chem.* 1987; **16**: 1–83.
- Exner O. *Correlation Analysis of Chemical Data* (Chapter 6.1) Plenum Press: New York, 1988.
- For instance: (a) Craik DJ, Brownlee RTC, *Prog. Phys. Org. Chem.* 1983; **14**: 1–73; (b) Reynolds WF, Gomes A, Maron A, MacIntyre DW, Tanin A, Hamer GK, Peat IR. *Can. J. Chem.* 1983; **61**: 2376–2384.
- Wold S, Sjöström M. In *Correlation Analysis in Chemistry, Recent Advances*, Chapman NB, Shorter J (eds). Plenum Press: New York, 1978.
- Ludwig M, Wold S, Exner O. *Acta Chem. Scand.* 1992; **46**: 549–554.
- (a) Charton M. *Prog. Phys. Org. Chem.* 1987; **16**: 287–315; (b) Weeks G, Horák V. *J. Org. Chem.* 1981; **45**: 2068–2071; (c) Nieuwdorp GHE, de Ligny CL, van Houwelingen HC. *J. Chem. Soc., Perkin Trans. 2*: 1979; 537–544.
- (a) Bottino FA, Musumarra G, Rappoport Z. *Magn. Reson. Chem.* 1986; **24**: 31–34; (b) Exner O, Buděšínský M. *Magn. Reson. Chem.* 1989; **27**: 27–36; (c) Clementi S, Fringuelli F, Linda P, Savelli G. *Gazz. Chim. Ital.* 1975; **105**: 281–292.
- (a) Exner O. *Collect. Czech. Chem. Commun.* 1966; **31**: 65–89; (b) Decouzon M, Exner O, Gal J-F, Maria P-C. *J. Phys. Org. Chem.* 1994; **7**: 511–517.
- (a) Palm VA. *Osnovy Kolichestvennoi Teorii Organicheskikh Reaktsii* (Chapter IX-2). Khimiya: Leningrad, 1967; (b) Tsuno Y, Sawada M, Fujii T, Yukawa Y. *Bull. Chem. Soc. Jpn.* 1979; **52**: 3033–3042; (c) Adcock W, Dewar MJS, Gupta BD. *J. Am. Chem. Soc.* 1973; **95**: 7353–7358; (d) Vorpágel ER, Streitwieser A, Jr, Alexandratos SD. *J. Am. Chem. Soc.* 1981; **103**: 3777–3781; (e) Segurado MAP, Reis JCR, de Oliveira JDG. *J. Chem. Soc., Perkin Trans. 2*: 2002; 323–328.
- (a) Taagepera M, Summerhays KD, Hehre WJ, Topsom RD, Pross A, Radom L, Taft RW. *J. Org. Chem.* 1981; **46**: 891–903; (b) Hoefnagel AJ, Oosterbeek W, Wepster BM. *J. Org. Chem.* 1984; **49**: 1993–1997.
- (a) Pross A, Radom L, Taft RW. *J. Org. Chem.* 1980; **45**: 818–826; (b) Exner O. *J. Org. Chem.* 1988; **53**: 1810–1812; (c) Exner O. *Org. Reactivity (Tartu)* 1995; **29**: 1–6.
- George P, Trachtman M, Bock CW, Brett AM. *J. Chem. Soc., Perkin Trans. 2*: 1976; 1222–1227.
- (a) Koppel LA, Mishima M, Stock LM, Taft RW, Topsom RD. *J. Phys. Org. Chem.* 1993; **6**: 685–689; (b) Adcock W, Anvia F, Butt G, Cook A, Duggan P, Grob CA, Marriott S, Rowe J, Taagepera M, Taft RW, Topsom RD. *J. Phys. Org. Chem.* 1991; **4**: 353–360; (c) Sung K. *J. Chem. Soc., Perkin Trans. 2002*; **2**: 1658–1661; (d) Barbour JB, Karty JM. *J. Phys. Org. Chem.* 2005; **18**: 210–216; (e) Wiberg KB. *J. Org. Chem.* 2002; **67**: 1613–1617; (f) Diezy Riega H, Rincón L, Almeida R. *J. Phys. Org. Chem.* 2003; **16**: 107–113; (g) Liu L, Fu Y, Liu R, Li R-Q, Guo Q-X. *J. Chem. Inf. Comp. Sci.* 2004; **44**: 652–657.
- Exner O, Böhm S. *Chem. Eur. J.* 2002; **8**: 5147–5152.
- Exner O, Böhm S. *J. Phys. Org. Chem.* 2004; **17**: 124–130.
- Exner O, Böhm S. *J. Phys. Org. Chem.* 2006; **19**: 393–401.
- Exner O. *Org. Reactivity (Tartu)* 1984; **21**: 3–26.
- Exner O, Böhm S. *J. Phys. Org. Chem.* 2006; **19**: 1–9.
- (a) Greenberg A, Stevenson TA. *J. Am. Chem. Soc.* 1985; **107**: 3488–3494; (b) Helal MR. *J. Mol. Struct. (Theochem)* 2000; **528**: 454–462.



- 255–261; (c) Böhm S, Exner O. *J. Mol. Struct. (Theochem)* 2005; **722**: 125–131.
23. (a) Krygowski TM, Stepień BT. *Chem. Rev.* 2005; **105**: 3482–3512; (b) Exner O, Böhm S. *Curr. Org. Chem.* 2006; **10**: 763–778.
24. (a) Becke AD. *J. Chem. Phys.* 1993; **98**: 5648–5652; (b) Lee C, Yang W, Parr RG. *Phys. Rev. B: Condens. Matter* 1988; **37**: 785–789.
25. (a) Exner O, Böhm S. *J. Org. Chem.* 2002; **67**: 6320–6327; (b) Böhm S, Pařík P, Exner O. *New J. Chem.* 2006; **30**: 384–391.
26. Wiberg KB. *J. Org. Chem.* 2002; **67**: 4787–4794.
27. Böhm S, Exner O. *J. Comput. Chem.* 2006; **27**: 571–577.
28. Exner O, Böhm S. *Phys. Chem. Chem. Phys.* 2004; **6**: 3864–3871.
29. McMahon TB, Kebarle P. *J. Am. Chem. Soc.* 1977; **99**: 2222–2230.
30. Jaffé HH. *Chem. Rev.* 1953; **53**: 191–261.
31. Mager H. *Pharmazie* 1983; **38**: 120–128.
32. (a) Swain CG, Unger SH, Rosenquist NR, Swain MS. *J. Am. Chem. Soc.* 1983; **105**: 492–502; (b) Drago RS, Dadmun AP. *J. Am. Chem. Soc.* 1994; **116**: 1792–1799.
33. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. *Gaussian 03, Revision B.03*. Gaussian, Inc.: Pittsburgh PA, 2003.
34. Exner O. *J. Chim. Phys.* 1992; **89**: 1573–1584.
35. Dubois D, Prade H. *Fuzzy Sets and Systems: Theory and Application*. Academic Press: New York, 1980.
36. Exner O, Böhm S. *Collect. Czech. Chem. Commun.* 2001; **66**: 1623–1637.