JOURNAL OF PHYSICAL ORGANIC CHEMISTRY, VOL. 11, 485–488 (1998)

Improved method of calculating group polarization effects: constants σ_α

Petr Čársky,[1] Petr Nauš[1]† and Otto Exner[2]*

[1]Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, 18223 Praha 8, Czech Republic
[2]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 16610 Praha 6, Czech Republic

Received 12 August 1997; revised 27 November 1997; accepted 28 November 1997

ABSTRACT: An improved model is presented for *ab initio* calculations of the polarizability constants σ_α using Sadlej's basis set. A new set of constants was calculated for common groups and these were compared with previous 3–21G and 3–21G* calculations and with some experimental equilibrium constants in the gas phase. For the time being, the calculations seem to be the most general approach for obtaining these constants whereas the experimental possibilities are limited in scope. © 1998 John Wiley & Sons, Ltd.

KEYWORDS: group polarization effects; polarizability countants; *ab initio*; Sadlej's basis set

INTRODUCTION

The quantitative evaluation of substituent effects in terms of constants σ is one approach to the general problem of elucidating how the individual parts of a molecule influence each other.[1] The constants σ of different kinds also represent an attempt to classify the substituent effects under various labels possibly describing the mechanism of interaction. In the latest stage of development, four kinds of constants are distinguished,[2,3] representing the field-inductive mechanism (σ_F), resonance (σ_R), polarizability (σ_α) and electronegativity (σ_χ). An arbitrary experimental quantity y (e.g. an equilibrium constant[2]) may be then approximated by Eqn (1), where the ρs are proportionality constants and ε is the fitting error:

$$y = y^0 + \rho_F\sigma_F + \rho_R\sigma_R + \rho_\alpha\sigma_\alpha + \rho_\chi\sigma_\chi + \varepsilon \qquad (1)$$

All terms of equation (1) need not be statistically significant in every case. Even when the equation is satisfactory for predicting y, the individual terms need not have a physical meaning. The constants σ are obtainable from various conventional model reactions or properties,[1–4] commonly from the equilibrium constants.[1,2] Their theoretical foundation is not rigorous: the essential factor is the ability of Eqn (1) to predict the experimental facts. For constants σ not easily accessible by conventional experiments, quantum chemical models were proposed by Marriott, Topsom and co-workers.[5–8] These models are also essentially empirical and justified mainly by the agreement with the experimental values. We recently recalculated a set of σ constants by standard modern computational methods and included also some special substituents needed in QSAR.[9]. Comparison with standard sets revealed that the calculated σ are of approximately the same reliability as experimental values but with various types of constants the position differed.

Whereas calculations of σ_F, σ_R and σ_χ proceeded without problems, some questions arose with σ_α. The model is based on calculating the polarization potentials for the molecule XCH_3 and the parent CH_4, σ_α being proportional to the difference. In the original paper,[7] the 3–21G basis set was used for hydrogen and first-row elements; if the substituent X contained atoms from other than the first row, the 3–21G* basis set was recommended. In our opinion, this basis set is inappropriate for this purpose: calculations of molecular polarizability require the use of larger and more flexible basis sets. Therefore, we considered it expedient to recalculate the whole set of constants σ_α by the basis set developed by Sadlej.[10,11] This set was especially devised for SCF calculations of polarizabilities and other electric properties of molecules. However, we did not consider it expedient to improve the model by including electron correlation because we wanted to preserve its original simplicity. The set of substituents selected previously for special use in QSAR was only slightly extended here to

**Correspondence to:* O. Exner, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, 16610 Praha 6, Czech Republic.
Contract/grant sponsor: Grant Agency of the Czech Republic; *contract grant number:* 203/96/0262; *contract grant number:* 203/96/1658.
†Undergraduate at the Faculty of Science, Charles University, Praha.

© 1998 John Wiley & Sons, Ltd. CCC 0894–3230/98/070485–04 $17.50

include also some substituents important for comparison with experimental quantities.

CALCULATIONS

We used the basis set as recommended by Sadlej:[10,11] (6s2p)/[3sls] for hydrogen, (10s6p2d)/[5s3pld] for first-row atoms and (13s10p2d)/[7s5p1d] for second-row atoms. These sets were used both for geometry optimization and calculation of the polarization potentials of the molecules XCH_3 and CH_4. The polarization potential PP is defined[12] as

$$PP = \sum_{a \neq 0} \frac{\langle \Psi_a | H' | \Psi_a \rangle \langle \Psi_a | H' | \Psi_a \rangle}{E_0 - E_a} \tag{2}$$

In this second-order term of the Møller–Plesset theory, the summation extends over all singly excited configuration state functions and the Hamiltonian H' represents coulombic interactions between a positive charge and a XCH_3 or CH_4 molecule. The positive charge is located[12] at a distance of 3 Å from the central carbon atom in the direction along one of the three C—H bonds. Using the Hartree–Fock molecular orbitals φ and orbital energies ε at the optimum geometry, Eqn (2) may be rewritten as

$$PP = 2 \sum_i \sum_j \frac{\langle \varphi_i | H' | \varphi_j \rangle \langle \varphi_j | H' | \varphi_i \rangle}{\varepsilon_i - \varepsilon_j} \tag{3}$$

The pair of indices i and j represents a single electron excitation $i \rightarrow j$ and the H' matrix elements are evaluated for the models $\oplus \cdots H_3CX$ and $\oplus \cdots HCH_3$ as noted above. The constant σ_α is then expressed as the difference:

$$\sigma_\alpha = PP_{CH_3X} - PP_{CH_4} \tag{4}$$

RESULTS AND DISCUSSION

The calculated values of σ_α are given in Table 1. Compared with the values calculated previously,[9] there is general agreement but also some individual differences. In particular, the values for all polar acceptor groups are now smaller in absolute value, for *tert*-butyl the new value is greater and in the case of SO_2CH_3 these two effects compensate. The small polarizability of fluorine, smaller than that of hydrogen, is still more stressed. Statistical comparison (Table 2, line 1) reveals a relatively poor correlation with certain irregular features, in particular a non-zero intercept. We are unable to decide simply which scale is 'better.' Comparison with theoretically based values of polarizability[13] is possible only for isolated atoms. A decision on the basis of experimental data would be feasible in principle but for the time being there is no simple experimental quantity which could be directly correlated with σ_α. Most important applications of σ_α are multiparameter equations such as Eqn (1) correlating the gas-phase basicity or acidity[2], but just the term with σ_α is sometimes statistically insignificant or doubtful. Most significant in this respect was the correlation[2] of the basicities of substituted amines, $XCH_2N(CH_3)_2$. This was later extended by Headley[14] but the coefficients were changed very little to give Eqn (5). For alkyl derivatives $RN(CH_3)_2$, i.e. without an interposed methylene group, Eqn (6) is valid.[14] The reference compound is trimethy-

Table 1. Calculated polarizability constants σ_α

Substituent	Ref. 7	Ref. 9	This work
CH_3	−0.35	−0.35	−0.32
C_2H_5	−0.49	−0.49	−0.49
C_3H_7	−0.54	−0.54[a]	−0.56
$CH(CH_3)_2$	−0.62	−0.62	−0.66
$CH_2CH(CH_3)_2$		−0.67[a]	−0.72
$CH(CH_3)C_2H_5$		−0.75[a]	−0.82
$C(CH_3)_3$	−0.75	−0.75	−0.84
$CH_2C(CH_3)_3$		−0.82[a]	−0.87
CH_2F		−0.33	−0.25
CHF_2		−0.30[a]	−0.18
CF_3	−0.25	−0.25	−0.12
CH_2OCH_3		−0.45[a]	−0.42
$CH{=}CH_2$	−0.50	−0.54[a]	−0.50
$C{\equiv}CH$	−0.60	−0.66[a]	−0.55
CHO	−0.46	−0.46	−0.32
$COCH_3$	−0.55	−0.55	−0.47
$COOCH_3$		−0.48	−0.37
$CONH_2$		−0.46	−0.37
$CONHCH_3$[b]		−0.52 (*Z*), −0.54[c] (*E*)	−0.46 (*Z*), −0.56 (*E*)
$CON(CH_3)_2$		−0.69	−0.67
COF		−0.36[a]	−0.19
$CSNH_2$		−0.65	−0.66
$CSNHCH_3$[b]		−0.71 (*Z*), −0.80 (*E*)	−0.76 (*Z*), −0.85 (*E*)
$CSN(CH_3)_2$		−0.88	−0.96
CN	−0.46	−0.52	−0.32
NH_2	−0.16	−0.17	−0.16
$N(CH_3)_2$	−0.44	−0.44	−0.52
$NHCOCH_3$[b]		−0.28 (*Z*), −0.47 (*E*)	−0.29 (*Z*), −0.44 (*E*)
$NHCONH_2$[b]		−0.25 (*Z*), −0.38[a] (*E*)	−0.26 (*Z*), −0.39 (*E*)
$NHCSCH_3$[b]		−0.48 (*Z*), −0.68 (*E*)	−0.56 (*Z*), −0.66 (*E*)
$NHCSNH_2$[b]		−0.39[a] (*Z*), −0.55 (*E*)	−0.52 (*Z*), −0.56 (*E*)
NO_2	−0.26	−0.26	−0.09
OH	−0.03	−0.03	+0.06
OCH_3	−0.17	−0.17	−0.13
F	+0.13	+0.13	+0.28
SH	−0.55	−0.53	−0.60
SCH_3	−0.68	−0.68	−0.79
SO_2CH_3	−0.62	−0.63	−0.63
Cl	−0.43	−0.44	−0.37

[a] Calculations carried out now according to the method of Ref. 9.
[b] Conformation on the partially double C—N bond *Z* or *E* as indicated
[c] This is a correction of a misprint in Ref. 9.

© 1998 John Wiley & Sons, Ltd.

Table 2. Correlations of constants σ_α with some experimental quantities (energy in kJ mol^{-1})

Line	Explanatory variables	Response function	Regression coefficients[a]	r^b	SD^b	N^b
1	σ_α (Ref. 9)	σ_α (this work)	1.17(5)[c]	0.963	0.073	46
2	$-\Delta G_{base}$, σ_F[d]	σ_α (this work)	0.038(4), −1.02(12)	0.968	0.071	10
3	$-\Delta G_{base}$, σ_F[d]	σ_α (Ref. 9)	0.034(3), −0.99(10)	0.968	0.060	10
4	$-\Delta G_{base}$, σ_F[d]	σ_α (this work)	0.035(2), −0.92(8)	0.981	0.067	14[e]
5	$-\Delta G_{base}$, σ_F[d]	σ_α (Ref. 9)	0.037(2), −1.05(6)	0.986	0.055	14[e]
6	$-\Delta G_{base}$[d]	σ_α (this work)	0.0349(19)	0.9907	0.050	9[f]
7	$-\Delta G_{base}$[d]	σ_α (Ref. 9)	0.0368(12)	0.9965	0.032	9[f]
8	$-\Delta G_{acid}$, $-\Delta G_{base}$[g]	σ_α (this work)	−0.0092(16), −0.0131(13)	0.972	0.074	9
9	$-\Delta G_{acid}$, $-\Delta G_{base}$[g]	σ_α (Ref. 9)	−0.0096(12), −0.0108(10)	0.976	0.056	9
10	$-\Delta G_{acid}$, $-\Delta G_{base}$[g]	σ_F	0.0044(3), −0.0034(2)	0.9981	0.012	9

[a] Standard deviation in parentheses.
[b] r = Correlation coefficient, SD = standard deviation of the fit and N = number of data.
[c] The intercept of 0.103 is not negligible.
[d] Basicities of substituted dimethylamines $XCH_2N(CH_3)_2$ (Ref. 14).
[e] Including also compounds with the substituents $CH(CH_3)_2$, $CH(CH_3)C_2H_5$, $C(CH_3)_3$ and $C(CH_3)_2C_2H_5$ in place of CH_2X (as it was done in Ref. 14).
[f] Only alkyl substituents including those (in footnote e).
[g] Acidities and basicities of substituted methanols XCH_2OH (Ref. 15).

lamine in both cases.

$$-\delta\Delta G\ (\text{kcal}) = 28.7\ \sigma_F + 6.7\ \sigma_\alpha - 0.1 \qquad (5)$$

$$-\delta\Delta G\ (\text{kcal}) = 16.5\ \sigma_\alpha + 5.8 \qquad (6)$$

In the case of Eqn (5), we can consider σ_F as known since it has been obtained by both experimental and theoretical methods with concordant results;[9] then σ_α is the response function. With the experimental data available,[14] we obtained the results given in Table 2, lines 2 and 3. The proportionality constants agree with those in Eqn (5) (in kcal mol^{-1}) within 15%. The precision of the equation is satisfactory (see particularly the correlation coefficient). On the other hand, it is evident that any preference for new values of σ_α over the previous ones[9] is not possible on the basis of these data. We can still include four alkyl substituents not containing a CH_2 group as was done by Headley.[14] The correlation is slightly improved (Table 2, lines 4 and 5), but a decision between old and new values is still impossible: it seems merely that the original values were better.

Using Eqn (6), we obtained the results shown in Table 2, lines 6 and 7. An essential drawback is the restriction to alkyl groups. The difference between correlations with new and previous[9] values of σ_α now seems to be greater. Nevertheless, the difference between lines 6 and 7 is again insignificant statistically, as were also the differences between line 2 and 3 and between lines 4 and 5 (F-test, confidence level $\alpha = 0.10$). The correlations with Eqn (6) are better than those with Eqn (5); it seems that for alkyl groups, the constants σ_α are better defined with a clearer physical meaning than for more polar groups.

Another ingenious idea[15] compares the gas-phase acidities and basicities of the same compounds: substituted alcohols, RCH_2OH. Whereas polarizability effects strengthen both quantities, polar effects strengthen the acidity and weaken the basicity. This was expressed[15] by the equations (7) and (8).

$$RCH_2OH + CH_3O^- = RCH_2O^- + CH_3OH$$
$$-\delta\Delta G_{acid} \approx I + P \qquad (7)$$

$$RCH_2OH + CH_3O^+H_2 = RCH_2O^+H_2 + CH_3OH$$
$$-\delta\Delta G_{base} \approx -I + P \qquad (8)$$

By substracting $\delta\Delta G_{acid}$ from $\delta\Delta G_{base}$, a measure of the inductive effect was obtained.[15] Conversely, we can add Eqns (7) and (8) to obtain a measure of polarizability. The model can be improved. There is no reason why the effects should have exactly the same intensity[16] for both $\delta\Delta G_{acid}$ and $\delta\Delta G_{base}$. We can thus write in the second approximation

$$-\delta\Delta G_{acid} = I + P \qquad (9)$$

$$-\delta\Delta G_{base} = -aI + bP \qquad (10)$$

The model can be tested in the following way. From Eqns (9) and (10), either P or I is eliminated and I is identified with $\rho_F\ \sigma_F$ and P with $\rho_\alpha\ \sigma_\alpha$. We obtain

$$-b\ \delta\Delta G_{acid} + \delta\Delta G_{base} = \rho_F\ (a + b)\ \sigma_F \qquad (11)$$

$$a\ \delta\Delta G_{acid} + \delta\Delta G_{base} = \rho_\alpha\ (a + b)\ \sigma_\alpha \qquad (12)$$

Equation (11) can serve to test how the model works since the constants σ_F are known with reliability. Equation (12) can then indicate whether the calculated constants σ_α have some physical meaning. The results are given in Table 2, lines 8 to 10. The correlation with σ_F seems to be of fairly high precision but the variation of their values is very small: for all alkyl substituents $\sigma_F = 0$. Correlation with σ_α is of similar precision as in the previous model: the difference between old and new σ_α values is again insignificant at $\alpha = 0.10$. This model

© 1998 John Wiley & Sons, Ltd.

evidently offers the best experimental approach to σ_α, not dependent on any other quantity than two gas-phase ionization equilibria. The regression coefficients in Table 2, lines 8–10, agree with the expectations according to Eqns (11) and (12). This confirms that the model is meaningful. However, only a restricted number of substituents can be investigated in this way since protonation must not occur on the substituent. Attempted extension to further derivatives was technically difficult owing to the decomposition of many compounds after ionization.[17] This difficulty was circumvented in a recent approach[18] where $\delta\Delta G_{acid}$ and $\delta\Delta G_{base}$ for a series of alcohols were calculated at the HF/6–31G** level and converted into I and P according to the original uncorrected Eqns (7) and (8). Another difference compared with our treatment was the definition of the substituent in the molecule RCH_2OH (the whole group RCH_2 instead of R as here). For this reason, a direct comparison with our results is possible only with a few values but evidently this combined theoretical–empirical approach[18] yields fairly close relative values to ours. In conclusion, even the above set of model compounds does not allow us to decide between the sets of calculated σ_α constants; evidently a more efficient model system should be sought.

CONCLUSIONS

The procedure outlined here gives in any case the possibility of calculating the constants σ_α in a reproducible, standard way. In our opinion, these calculations are the best possibility of obtaining these constants to date. Experimentally, the constants have been documented only on a few models. This is the reason why we cannot say that using of a theoretically more substantiated model has actually changed the results for the better. Also, the exact physical meaning of σ_α is unclear, particularly in comparison with the inductive effect represented by constants σ_F. Referring to the observable quantities presented here, we can say only that σ_α represens the ability to stabilize a charge, either positive or negative, and σ_F refers to stabilizing the negative and destabilizing the positive charge. For a better understanding, further experimental models would be necessary.

Acknowledgment

Thanks are due to the Grant Agency of the Czech Republic for supporting this work by grants Nos 203/96/0262 (to O. E.) and 203/96/1658 (to P. Č. and P. N.).

REFERENCES

1. O. Exner. *Correlation Analysis of Chemical Data*, Chapt. 5. Plenum Press, New York (1988).
2. R. W. Taft and R. D. Topsom. *Prog. Phys. Org. Chem.* **16,** 1 (1987).
3. R. D. Topsom. *Prog. Phys. Org. Chem.* **16,** 125 (187).
4. C. Hansch, A. Leo and R. W. Taft. *Chem. Rev.* **91,** 165 (1991).
5. S. Marriott and R. D. Topsom. *J. Am. Chem. Soc.* **106,** 7 (1984).
6. S. Marriott, A. Silvestro and R. D. Topsom. *J. Chem. Soc., Perkin Trans. 2* 457 (1988).
7. W.-J. Hehre, C.-F. Pau, A. D. Headley, R. W. Taft and R. D. Topsom. *J. Am. Chem. Soc.* **108,** 1711 (1986).
8. S. Marriott, W. F. Reynolds, R. W. Taft and R. D. Topsom. *J. Org. Chem.* **49,** 959 (1984).
9. O. Exner, M. Ingr and P. Čársky. *J. Mol. Struct. (Theochem)* **397,** 231 (1997).
10. A. J. Sadlej. *Collect. Czech. Chem. Commun.* **53,** 1995 (1988).
11. A. J. Sadlej. *Theor. Chim. Acta* **79,** 123 (1991).
12. S. D. Kahn, C. F. Pau, L. E. Overman and W. J. Hehre. *J. Am. Chem. Soc.* **108,** 7381 (1986).
13. T. K. Ghanty and S. K. Ghosh. *J. Am. Chem. Soc.* **116,** (1994) 8801.
14. A. D. Headley. *J. Am. Chem. Soc.* **109,** 2347 (1987).
15. R. W. Taft, M. Taagepera, J. L. M. Abboud, J. F. Wolf, D. J. DeFrees, W. J. Hehre, J. E. Bartmess and R. T. McIver. *J. Am. Chem. Soc.* **100,** 7765 (1978).
16. M. Charton. in *Similarity Models in Organic Chemistry, Biochemistry and Related Fields*, edited by R. I. Zalewski, T. M. Krygowski and J. Shorter, p. 629. Elsevier, Amsterdam (1991).
17. M. Decouzon, J.-F. Gal and P.-C. Maria. unpublished work.
18. J. Catalán. *J. Phys. Org. Chem.* **9** 652 (1996).

© 1998 John Wiley & Sons, Ltd.