Prediction of Taft's σ^* parameter for alkyl groups and alkyl groups containing polar substituents

Arne Brändström †

Medical Chemistry, AstraZeneca, S-431 83 Mölndal, Sweden

Received (in Lund, Sweden) 22nd February 1999, Accepted and transferred from Acta Chem. Scand. 16th June 1999

X^1, X^2, and X^3 are polar groups, hydrogens or alkyl groups bound to a carbon atom in a group $CX^1X^2X^3$. The value of $\sigma^*(CX^1X^2X^3)$ has been found to be equal to the sum of the contributions of all groups X^1, X^2, and X^3. The contribution of a polar group X to the value of $\sigma^*(CX^1X^2X^3)$ is $b \times \Sigma\sigma^*(X)$ and that of an alkyl group R is $a \times \Sigma[\sigma^*(R) - \sigma^*(H)]$ where $a = 0.202$ and $b = 0.326$. The additivity of the contributions of the groups is good, giving σ^* values usually as good as those obtained by experiments.

Introduction

In 1952 Taft[1] introduced the parameter $\sigma^*(R)$ in eqn. (1), where

$$\sigma^*(R) = [\log(k_R/k_H)_B - \log(k_R/k_H)_A]/2.48 \qquad (1)$$

k_R is the rate constant for the hydrolysis of an ester RCH_2COOR' and k_H is the rate constant for the hydrolysis of the parent compound CH_3COOR'. Due to experimental problems $\sigma^*(R)$ values are not available from this method for all possible groups R. They have therefore often been obtained by a variety of other methods. Many such methods are given by Taft.[1] The most frequently used of these methods to obtain $\sigma^*(R)$ is from pK_a values of substituted carboxylic acids. A high $\sigma^*(R)$ obtained by this method corresponds to a low pK_a value. The difficulties in the measurements of low pK_a values are therefore reflected in the uncertainty observed in the experimental $\sigma^*(R)$ values, which are usually higher the higher the $\sigma^*(R)$ value.

$\sigma^*(R)$ values have been used as a parameter in many correlations of chemical structure against effects. Some examples of this are given in the book by Hansch and Leo[2] who are rather negative about its use. This is more than balanced by the positive approach of Perrin, Dempsey and Serjeant.[3] Lists of experimental $\sigma^*(R)$ values can be found in these two books. The values used in this work are, unless otherwise indicated, those selected by Hansch, Leo and Hoekman[4] and the SD values are those of the experimental values reported by them (a few very deviating values are omitted).

Results and discussion

Predictions of σ^* parameter for alkyl groups

A large number of $\sigma^*(R)$ values, which were not available, were needed to test whether a relationship of Taft's type can explain aqueous solubility values. Since $\sigma^*(R)$ values are reported to be additive[3] it should be possible to calculate them from known values for small groups. Let us now suppose that eqn. (2) is valid. For $R^1 = R^2 = R^3 = H$ we obtain $\sigma^*(CH_3) = 0$

$$\sigma^*(CR^1R^2R^3) = a \times \sum_{i=1}^{3} [\sigma^*(R_i) - \sigma^*(H)] \qquad (2)$$

in agreement with the definition. If eqn. (2) is valid and the parameters a and $\sigma^*(H)$ are known, the sigma value of any alkyl group can be readily calculated. This can be seen by the following examples: $\sigma^*(CH_3CH_2) = a \times [\sigma^*(CH_3) - \sigma^*(H)] = a \times [0 - \sigma^*(H)] = -a \times \sigma^*(H)$, $\sigma^*((C_2H_5)_2CH) = a \times \{2[\sigma^*(C_2H_5) - \sigma^*(H)]\} = a \times \{2[-a \times \sigma^*(H) - \sigma^*(H)]\} = -2a \times \sigma^*(H) - 2a^2 \times \sigma^*(H)$ and so on. We thus obtain $\sigma^*(R) = -\sigma^*(H) \times F(a)$ where $F(a)$ is the potential function $\sum_i n_i \times a^i$.

Table 1 Found and calculated $\sigma^*(R)$ values

R	n_1	n_2	n_3	n_4	σ^*(found)	σ^*(calc.)
CH_3	0	0	0	0	0	0
C_2H_5	1	0	0	0	−0.10	−0.099
n-C_3H_7	1	1	0	0	−0.12	−0.119
$CH(CH_3)_2$	2	0	0	0	−0.19	−0.198
n-C_4H_9	1	1	1	0	−0.13	−0.123
$CH_2CH(CH_3)_2$	1	2	0	0	−0.13	−0.139
$CH(CH_3)C_2H_5$	2	1	0	0	−0.21	−0.218
$C(CH_3)_3$	3	0	0	0	−0.30	−0.297
n-C_5H_{11}	1	1	1	1	−0.15	−0.124
$CH(CH_3)C_3H_7$	2	1	1	0	−0.22	−0.222
$CH(C_2H_5)_2$	2	2	0	0	−0.23	−0.238

Using the value $\sigma^*(H) = 0.49$ and $F(a)$ obtained from formulas of the type above as the independent variable and σ^* values for the alkyl groups listed in Table 1, for which reasonably reliable values are available as dependent variables, we can calculate $a = 0.202$ and $SD = 0.003$ by nonlinear regression which explains 98.5% of the variance. The satisfactory agreement between found and calculated values is readily seen from Table 1.

By this procedure $\sigma^*(R)$ for alkyl groups can be converted from an experimental quantity into a parameter dependent only upon the structure and the two parameters a and $\sigma^*(H)$.

We should observe that the formula can only be applied when R is an alkyl group.

Predictions of σ^* parameter for alkyl groups substituted by polar groups

In light of the success in calculating $\sigma^*(R)$ for alkyl groups, a formula was searched for that would allow the calculation of σ^* values when R is an alkyl group substituted by halogens or other polar groups. Suppose that we have a group $CX^1X^2X^3$ where at least one of the X groups is a polar group and the remaining groups are hydrogens or alkyl groups. If the contribution of each polar group to the value $\sigma^*(CX^1X^2X^3)$ is $b \times \Sigma\sigma^*(X)$ and that of each R group, as before, is $a \times \Sigma[\sigma^*(R) - \sigma^*(H)]$, eqn. (3) is thus obtained. In this equa-

$$\sigma^*(CX^1X^2X^3) = b \times \Sigma\sigma^*(X) + a \times \Sigma[\sigma^*(R) - \sigma^*(H)] \qquad (3)$$

---

† Present address: Karlsborg 193, S-271 94 Ystad, Sweden.

Table 2 $\sigma^*(\mathrm{CX}_n\mathrm{H}_{(3-n)}) = b \times n\,\sigma^*(\mathrm{X})$

X	n	σ^*(X)	SD	$\sigma^*(\mathrm{CX}_n\mathrm{H}_{(3-n)})$ found	SD	$\sigma^*(\mathrm{CX}_n\mathrm{H}_{(3-n)})$ calc.
Br	3	2.80	0.10	2.47	0.03	2.74
Br	2	2.80	0.10	1.96	0.01	1.83
Br	1	2.80	0.10	1.00	0.03	0.91
Cl	3	2.94	0.14	2.67	0.20	2.88
Cl	2	2.94	0.14	2.01	0.06	1.92
Cl	1	2.94	0.14	1.01	0.04	0.96
F	3	3.19	0.07	2.60	0.38	3.12
F	2	3.19	0.07	2.05	0.02	2.08
F	1	3.19	0.07	1.10	0.07	1.04
I	2	2.22	0.10	1.62	—	1.45
I	1	2.22	0.10	0.85	0.06	0.72
NO_2	3	4.66	0.29	4.60	0.94	4.56
NO_2	1	4.66	0.29	1.69	0.17	1.52
OH	1	1.37	0.19	0.56	0.04	0.45
SH	1	1.60	0.11	0.56	0.08	0.52
SCF_3	3	2.73	0.03	3.06	—	2.67
SCF_3	1	2.73	0.03	0.75	—	0.89
$SOCF_3$	1	4.30	—	1.62	0.11	1.40
SO_2CF_3	1	4.41	0.09	1.75	0.13	1.44
CN	1	3.64	0.22	1.30	0.17	1.19
NCO	1	2.25	—	0.81	—	0.73
NCS	1	2.62	—	0.94	—	0.85
SCN	1	3.43	0.15	1.18	—	1.12
OCH_3	1	1.77	0.06	0.52	0.07	0.58
OC_2H_5	1	1.68	0.02	0.58	0.04	0.55
SCH_3	1	1.56	0.10	0.42	0.12	0.51
$SOCH_3$	1	2.88[a]	—	1.33	—	0.94
$SO2CH_3$	1	3.68	—	1.32	—	1.20
C≡CH	1	2.15	0.62	0.76	0.12	0.70
$HC{=}CH_2$	1	0.52	0.10	0.12	0.09	0.17
$OCOCH_3$	1	2.51	0.08	0.70	0.11	0.82
$COCH_3$	1	1.65	0.08	0.60	0.02	0.54
$COOCH_3$	1	2.00	0.12	1.06[a]	—	0.65
$COOC_2H_5$	1	2.26	—	0.82	—	0.74
CCl_3	1	2.65	0.20	0.75	—	0.86
CH_2CCl_3	1	0.75	—	0.25	—	0.24
CF_3	2	2.60	0.38	1.32[a]	—	1.69
CF_3	1	2.60	0.38	0.90	0.03	0.85
CH_2CF_3	1	0.90	0.03	0.32	—	0.29
$CF_2CF_2CF_3$	1	2.83	—	0.87	—	0.92
CH_2Br	1	1.00	0.03	0.49	0.10	0.33
CH_2Cl	1	1.01	0.04	0.39	0.00	0.33
CH_2I	1	0.85	0.06	0.41	—	0.28
CH_2OH	1	0.56	0.04	0.21	—	0.18
CH_2CH_2Cl	1	0.39	0.00	0.14	0.01	0.13
CH_2CHCl_2	1	0.65[b]	—	0.25	—	0.21
$CHBrCH_3$	1	1.00	0.08	0.44	—	0.33
$CHClCH_3$	1	1.00	0.05	0.35	—	0.33
C_6H_5	1	0.63	0.06	0.23	0.02	0.21

[a] This value seems to be out of line and is not used in the regression. [b] Calculated value, no experimental value is available.

tion a is 0.202 as before but b has to be calculated. We may observe that when the polar group X is absent, the equation is reduced to that for the calculation of σ^* values for alkyl groups, eqn. (2).

The different treatment of the contributions of the substituents X and R as σ^*(X) and $[\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]$ leads to a significantly better fit than that obtained when they are treated as $[\sigma^*(\mathrm{X}) - \sigma^*(\mathrm{H})]$ and $[\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]$. This may be due to a real difference, but may also be a consequence of the fact that σ^*(X) values for polar groups are usually not obtained directly from eqn. (1) but from regressions of other types of properties, usually pK_a values, against σ^*(R).[1] Which is the correct explanation can only be settled by a thorough investigation of all primary data from which the σ^*(X) values are obtained. This is not performed here since the results obtained by eqn. (3) are satisfactory for the present purpose.

By iterations in eqn. (3) one might in principle be able to calculate σ^*(X) values for all polar groups. Let us consider the following example: if $\sigma^*[\mathrm{CR'Cl(CRCl_2)}]$ has to be calculated one may start by calculating $\sigma^*(\mathrm{CRCl_2}) = 2b \times \sigma^*(\mathrm{Cl}) + a \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]$. When this expression is introduced into eqn. (3) one obtains eqn. (4). Observe that the group $CRCl_2$ is

$$\sigma^*[\mathrm{CR'Cl(CRCl_2)}] = b \times \{2b \times \sigma^*(\mathrm{Cl}) + a \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]\} + a \times [\sigma^*(\mathrm{R'}) - \sigma^*(\mathrm{H})] + b \times \sigma^*(\mathrm{Cl}) \quad (4)$$

polar and its σ^* value is thus multiplied by b. This equation can be simplified to eqn. (5). We can thus express all σ^* values of

$$\sigma^*(\mathrm{CR'ClCRCl_2}) = a \times [\sigma^*(\mathrm{R'}) - \sigma^*(\mathrm{H})] + (b + 2b^2) \times \sigma^*(\mathrm{Cl}) + a \times b \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})] \quad (5)$$

composed groups as a function of the known parameter a, σ^* values of small fundamental groups which are experimentally known, and the parameter b.

Examples of formulae of this type are given in the captions of Tables 2–4. These can all be summarized in the formula (6),

Table 3 $\sigma^*(\mathrm{CHXY}) = b \times \sigma^*(\mathrm{X}) + b \times \sigma^*(\mathrm{Y})$ or $b \times \sigma^*(\mathrm{X}) + a \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]$

X	Y	σ^*(X)	SD	σ^*(Y)	SD	σ^*(CHXY) found	SD	σ^*(CHXY) calc.
Br	CH_3	2.80	0.10	0	0	1.00	0.08	0.81
Br	C_2H_5	2.80	0.10	−0.10	—	1.02	—	0.79
Cl	CH_3	2.94	0.14	0	0	1.00	0.06	0.86
Cl	C_2H_5	2.94	0.14	−0.10	—	1.05	—	0.84
Cl	CH_2CHCl_2	2.94	0.14	0.65[a]	—	1.26	—	1.17
Cl	CH_2CCl_3	2.94	0.14	0.75	—	1.33	—	1.20
NO_2	CF_3	4.66	0.29	2.60	0.38	2.67	—	2.37

[a] Calculated value, no experimental value is available.

Table 4 $\sigma^*(\mathrm{CX_2Y}) = 2b \times \sigma^*(\mathrm{X}) + b \times \sigma^*(\mathrm{Y})$ or $2b^*(\mathrm{X}) + a \times [\sigma^*(\mathrm{R}) = \sigma^*(\mathrm{H})]$

X	Y	σ^*(X)	SD	σ^*(Y)	SD	$\sigma^*(CX_2Y)$ found	$\sigma^*(CX_2Y)$ calc.
Cl	$COCH_3$	2.94	0.14	1.65	0.08	2.33	2.46
Cl	CH_3	2.94	0.14	0	0	1.53	1.82
Cl	C_2H_5	2.94	0.14	−0.10	—	1.91	1.80
Cl	CH_2CH_2Cl	2.94	0.14	0.39	0.00	2.02	2.04
Cl	CH_2CHCl_2	2.94	0.14	0.65[b]	—	2.10	2.13
Cl	CH_2CCl_3	2.94	0.14	0.75	—	2.16	2.16
F	CF_3	3.19	0.07	2.60	0.38	2.83	2.93
F	CF_2CF_3	3.19	0.07	2.83	—	2.83	3.06
F	$CF_2CF_2CF_3$	3.19	0.07	2.83	—	2.44[a]	3.00
CF_3	F	2.60	0.38	3.19	0.07	3.00	2.74

[a] This value seems to be out of line and is not used in the regression. [b] Calculated value, no experimental value is available.

$$\sigma^*(\mathrm{CXYR}) - a \times n_\mathrm{R} \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})] = b \times [n_\mathrm{X} \times \sigma^*(\mathrm{X}) + n_\mathrm{Y} \times \sigma^*(\mathrm{Y})] \quad (6)$$

where n_X, n_Y, and n_R are the numbers of the groups indicated. A linear regression of $\sigma^*(\mathrm{CXYR}) - a \times n_\mathrm{R} \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]$ against $[n_\mathrm{X} \times \sigma^*(\mathrm{X}) + n_\mathrm{Y} \times \sigma^*(\mathrm{Y})]$ with the intercept 0 should thus give the parameter b. By this procedure one obtains $b = 0.326$ and SD = 0.004. The regression explains 98.2% of the variance. The use of this b value in the calculation of σ^* values for composed groups is presented in Tables 2 to 4 using the formula at the top of each table.

The standard deviation of σ^* values calculated by the formula is 0.13 which is about the same as the standard deviations of the experimental values. The formula converts all σ^*(CXYZ) values from an experimental value to the parameters a and b and the σ^* values of the simple groups X, Y, and Z.

In the calculation of a σ^* value for a group CR′XY it is often helpful to use the σ^* value for the related group CRXY. The contribution of an alkyl group R′ to the total σ^* value is $a \times [\sigma^*(\mathrm{R}') - \sigma^*(\mathrm{H})]$ and that of the alkyl group R is $a \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]$. We thus have $\sigma^*(\mathrm{CR'XY}) = \sigma^*(\mathrm{CRXY}) + a \times [\sigma^*(\mathrm{R}') - \sigma^*(\mathrm{R})]$. If the alkyl group R is replaced by a polar group Z we obtain in a corresponding way eqn. (7). If

$$\sigma^*(\mathrm{CZXY}) = \sigma^*(\mathrm{CRXY}) + b \times \sigma^*(\mathrm{Z}) - a \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})] \quad (7)$$

the change of R to R′ occurs one carbon atom away from the center, we obtain in a similar way: $\sigma^*(\mathrm{CUVCR'XY}) = \sigma^*(\mathrm{CUVCRXY}) + b \times a \times [\sigma^*(\mathrm{R}') - \sigma^*(\mathrm{R})]$ and $\sigma^*(\mathrm{CUVCZXY}) = \sigma^*(\mathrm{CUVCRXY}) + b[b \times \sigma^*(\mathrm{Z}) - a \times [\sigma^*(\mathrm{R}) - \sigma^*(\mathrm{H})]\}$. The parameter b is used since all the groups CR′XY, CRXY, and CZXY are polar. The same rules can also be used in the calculations of σ^* values for groups of the type -CX=CXY, -C(X)=Y.

Acknowledgements

This work was initiated by the project “Lipophilicity”, project manager Docent Robert Lundkvist, Astra, Hässle A.B. Mölndal, Sweden.

References

1 R. W. Taft, in *Steric Effects in Organic Chemistry*, M. S. Newmann, ed., New York, 1956, Chapter 13.

2 C. Hansch and A. Leo, *Exploring QSAR, Fundamentals and Applications in Chemistry and Biology*, p. 70, ACS Professional Reference Book, American Chemical Society, Washington, DC, 1995.

3 D. D. Perrin, B. Dempsey and E. P. Serjeant, *pK_a Prediction for Organic Acids and Bases*, Chapman and Hall, London, 1981.

4 C. Hansch, A. Leo and D. Hoekman, *Exploring QSAR, Hydrophobic, Electronic, and Steric Constants*, ACS Professional Reference Book, American Chemical Society, Washington, DC, 1995.

Paper 9/04819G